

# Segmentation conjointe d'images et copules

## Joint image segmentation and copulas

Stéphane Derrode<sup>1</sup> et Wojciech Pieczynski<sup>2</sup>

<sup>1</sup> École Centrale Marseille & Institut Fresnel (CNRS UMR 6133),  
38, rue F. Joliot-Curie, F-13451 Marseille cedex 20, France.

Tél : int+ 33 4 91 05 47 30, [sderrode@centrale-marseille.fr](mailto:sderrode@centrale-marseille.fr)

<sup>2</sup> Télécom SudParis, CITI department (CNRS UMR 5157),  
9, rue Charles Fourier, F-91011 Evry cedex, France.

Tél : int+ 33 1 60 76 44 25, [wojciech.pieczynski@it-sudparis.eu](mailto:wojciech.pieczynski@it-sudparis.eu)

**Résumé** L'idée à l'origine du Modèle de Mélange Couplé (MMC) est de classer simultanément deux ensembles d'observations en introduisant un *a priori* conjoint entre les deux classifications et un lien statistique entre les deux observations. Nous étudions en particulier le cas de mélanges paramétriques non gaussiens construits à partir de modèles à base de copules et de marginales non gaussiennes. Nous proposons également un algorithme de type ECI pour l'estimation des paramètres, rendant la méthode de classification non supervisée. Des résultats de segmentation à partir d'un couple T1/T2 d'images IRM illustrent le potentiel du MMC par rapport à la classification individuelle des images par deux mélanges indépendants.

**Mots clés** Classification bayésienne, modèle de mélange, segmentation conjointe, copules.

**Abstract** The idea behind the Pairwise Mixture Model (PMM) is to classify simultaneously two sets of observations by introducing a joint prior between the two corresponding classifications and some inter-dependance relations between the two observations. We address the non-Gaussian case using copula-based parametric models and non-Gaussian margins. We also propose an ICE-based parameters estimation procedure which makes the classification algorithm unsupervised. Results on a T1/T2 couple of MRI images illustrate the interest for the PMM when compared to individual classifications obtained with two independent mixture models.

**Key words** Bayesian classification, mixture model, joint segmentation, copulas.

## 1 Introduction

Un problème important en traitement du signal consiste à restaurer un processus inobservable  $\mathbf{x} = \{x_n\}_{n \in \mathbb{N}}$  à partir d'un processus observé  $\mathbf{y} = \{y_n\}_{n \in \mathbb{N}}$ . Dans le « Modèle de Mélange » (MM) classique, les observations sont supposées mutuellement indépendantes et le problème de restauration consiste à classer les données en un nombre fini de  $K$  classes. Lorsque l'on souhaite analyser conjointement deux séries de données, *ie*  $\mathbf{y}^1$  et  $\mathbf{y}^2$ , on peut soit construire un modèle de mélange vectoriel qui fournit une unique classification  $\mathbf{x}$  pour les deux séries d'observations, soit

construire une modèle couplé qui fournit deux classifications  $\mathbf{x}^1$  et  $\mathbf{x}^2$  statistiquement liées [11,1].

Dans le domaine de la segmentation d'image, ces derniers modèles peuvent être d'un grand intérêt dans plusieurs situations. On peut penser à la segmentation conjointe de deux images recalées mais de modalités différentes comme une paire d'images T1 et T2 en Imagerie par Résonance Magnétique (IRM), ou une paire d'images radar et optique en imagerie satellitaire (détection de changements). Le nombre de classes dans les deux images peut être différent. Chaque image apporte une information pour la classification de l'autre. Ces modèles s'étendent sans difficulté à un nombre fini de séries d'observations, permettant par exemple de réaliser la segmentation d'images multi-spectrales.

Récemment [5], nous avons étudié un modèle appelé « Modèle de Mélange Couplé » (MMC) qui étend le modèle proposé dans [11] en prenant en compte, d'une part, un *a priori* commun entre les deux composantes du mélange et, d'autre part, en ajoutant un lien statistique entre les observations par le biais de lois d'attache aux données 2D. Nous proposons dans ce travail d'étendre le travail réalisé dans un cadre gaussien au contexte plus général de modèles paramétriques construits à l'aide de copules [9,2]. Pour l'estimation des paramètres, nous proposons un algorithme itératif du type ECI [10], permettant de ré-estimer les paramètres des copules par le biais de simulations (stochastique) et les paramètres intrinsèques au modèle avec des expressions analytiques (déterministe). La suite de ce papier décrit l'algorithme et propose des résultats comparatifs portant sur un couple d'images IRM.

## 2 Modèle de mélange couplé

Soit  $\mathbf{y} = \{y_1, \dots, y_N\}$  un ensemble de  $N$  données observées ( $y_n \in \mathbb{R}$ ). Soit également  $\mathbf{x} = \{x_1, \dots, x_N\}$  un ensemble de  $N$  données cachées dont chaque élément  $x_n$  appartient à  $\Omega = \{1, \dots, K\}$ . Dans le modèle classique de mélange probabiliste [12], les données  $y_n$  sont considérées comme des réalisations de V.A. réelles  $Y_n$  mutuellement indépendantes et de même distribution définie par le mélange

$$f_{MM}(y_n) = \sum_{k=1}^K \pi_k f_k(y_n), \quad (1)$$

où  $f_k(\cdot) = p(y_n | x_n = k)$  représentent des densités et  $\pi_k = p(x_n = k)$  les probabilités *a priori*.

### 2.1 Le modèle MM

On considère maintenant deux couples de données observées et cachées  $(\mathbf{y}^1, \mathbf{x}^1)$  et  $(\mathbf{y}^2, \mathbf{x}^2)$  avec  $\forall n \in [1, N]$ ,  $x_n^1 \in \Omega^1 = \{1, \dots, K\}$  et  $x_n^2 \in \Omega^2 = \{1, \dots, L\}$ . Nous

noterons  $\mathbf{y}_n = (y_n^1, y_n^2)'$  et  $\mathbf{x}_n = (x_n^1, x_n^2)'$ . Dans ce modèle, les couples de données  $\mathbf{y}_n$  sont des réalisations de vecteurs aléatoires  $\mathbf{Y}_n$  mutuellement indépendants avec la même distribution issue du mélange fini suivant

$$f(y_n^1, y_n^2) = \sum_{k=1}^K \sum_{l=1}^L \gamma_{k,l} f_{k,l}(y_n^1, y_n^2), \quad (2)$$

où  $\gamma_{k,l} = p(x_n^1 = k, x_n^2 = l)$  est l'*a priori* conjoint entre les composantes. Le modèle fait intervenir  $KL$  ddp 2D permettant de modéliser les couples  $(y_n^1, y_n^2)$  conditionnellement aux composantes du mélange  $(k, l)$ . Pour étendre le cas gaussien, nous proposons dans la section 2.3 de modéliser ces distributions à l'aide de copules. Voyons auparavant comment appliquer la décision bayésienne dans ce contexte.

## 2.2 Classification bayésienne

L'estimation de  $\mathbf{x}_n$  à partir de  $\mathbf{y}_n$  par décision bayésienne est réalisée à l'aide d'une fonction de coût  $L$  qui mesure l'erreur entre  $\mathbf{x}_n$  et son estimée  $\hat{\mathbf{x}}_n$  :

$$\hat{\mathbf{x}}_n = \arg \min_{\tilde{\mathbf{x}}_n \in \Omega^1 \times \Omega^2} \sum_{\mathbf{x}_n \in \Omega^1 \times \Omega^2} L(\mathbf{x}_n, \tilde{\mathbf{x}}_n) p(\mathbf{x}_n | \mathbf{y}_n)$$

Nous pouvons considérer les deux fonctions de coût suivantes :  $L_1(\mathbf{x}_n, \tilde{\mathbf{x}}_n) = \mathbb{1}_{x_n^1 \neq \tilde{x}_n^1} + \mathbb{1}_{x_n^2 \neq \tilde{x}_n^2}$  (Maximum A Posteriori - MAP) et  $L_2(\mathbf{x}_n, \tilde{\mathbf{x}}_n) = \mathbb{1}_{\mathbf{x}_n \neq \tilde{\mathbf{x}}_n} = \mathbb{1}_{x_n^1 \neq \tilde{x}_n^1 \text{ ou } x_n^2 \neq \tilde{x}_n^2}$  (Maximum des marginales a posteriori - MPM) où  $\mathbb{1}_c = 1$  si  $c$  est vrai, 0 sinon. En notant  $p(x_n^1 = k | \mathbf{y}_n) = \sum_{l=1}^L p(\mathbf{x}_n = (k, l) | \mathbf{y}_n)$ , nous obtenons

$$\begin{cases} MPM : & \hat{x}_n^1 = \arg \max_{k \in \Omega^1} p(x_n^1 = k | \mathbf{y}_n), \quad \hat{x}_n^2 = \arg \max_{l \in \Omega^2} p(x_n^2 = l | \mathbf{y}_n), \\ MAP : & \hat{\mathbf{x}}_n = \arg \max_{(k,l) \in \Omega^1 \times \Omega^2} p(\mathbf{x}_n = (k, l) | \mathbf{y}_n). \end{cases}$$

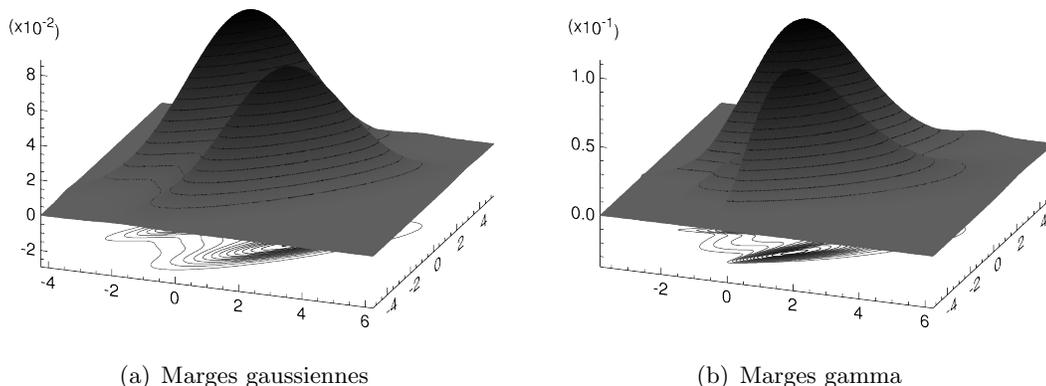
## 2.3 Copules dans les MMC

La théorie des copules [9] permet d'enrichir les capacités de modélisation multivariée (paramétrique) en séparant les problèmes de la recherche d'une forme adéquate pour les marginales et d'une structure de dépendance entre les composantes du vecteur aléatoire. Dans le cas de dimensions 2, les copules se définissent ainsi. Soient  $Y_1$  et  $Y_2$  deux V.A. réelles de Fonctions de Répartition (FdR)  $F^{(1)}$  et  $F^{(2)}$ . Soit  $F$ , la FdR du couple  $(Y_1, Y_2)$  de V.A. Si  $F^{(1)}$  et  $F^{(2)}$  sont continues, alors, d'après le théorème de Sklar, il existe une unique fonction  $C : [0, 1] \times [0, 1]$  de paramètre  $\theta^1$ , appelée « copule », telle que

$$F(y_1, y_2) = C(F^{(1)}(y_1), F^{(2)}(y_2); \theta).$$

---

1.  $\theta$  peut éventuellement représenter un vecteur de paramètres



**Figure 1.** Exemples de mélanges à base de copules gaussiennes ( $K = L = 2$ ) avec des marginales (a) gaussiennes et (b) gamma. Tau de Kendall :  $\tau_{1,1} = 0.8$ ,  $\tau_{1,2} = 0.3$ ,  $\tau_{2,1} = 0.3$  et  $\tau_{2,2} = 0.7$ . Probabilités *a priori* :  $\gamma_{1,1} = 0.30$ ,  $\gamma_{1,2} = 0.20$ ,  $\gamma_{2,1} = 0.15$  et  $\gamma_{2,2} = 0.35$ . Les paramètres exactes des marges utilisées pour construire ces 2 mélanges ne sont pas donnés pour des raisons d’encombrement !

Une copule peut être vue comme une FdR sur  $[0, 1] \times [0, 1]$  avec des distributions marginales uniformes. Il est souvent plus pratique de manipuler les densités et l’équation précédente s’écrit sous la forme

$$f(y_1, y_2) = f^{(1)}(y_1)f^{(2)}(y_2) c(F^{(1)}(y_1), F^{(2)}(y_2); \theta), \quad (3)$$

où  $c$  est la ddp de la copule  $C$ . De très nombreuses copules ont été étudiées, parmi lesquelles les copules à un unique paramètre  $\theta = \tau_{k,l}^2$  suivantes : gaussienne, Student, Clayton et Gumbel-Hougaard<sup>3</sup>.

Il est donc possible de combiner des distributions marginales « classiques » (gaussienne, gamma, béta de 1<sup>re</sup> espèce, . . . , lois du système de Pearson [3]) avec différents modèles de dépendance. Par exemple, la copule gaussienne peut être combinée avec des marges gaussiennes, aboutissant à une gaussienne 2D. Il est plus original d’adjoindre à la copule gaussienne des marges gamma, de manière à construire un nouveau modèle paramétrique (*cf* Fig. 1). En combinant ce potentiel dans le cadre du mélange de l’eq. (3), nous aboutissons à une modélisation d’une grande richesse, capable de prendre en compte des situations très variées sans augmenter significativement la complexité.

2. Le tau de Kendall  $\tau$  est une mesure de dépendance.

3. On trouvera dans [4] les expressions des FdR, des ddp et du tau de Kendall de ces copules.

### 3 Estimation des paramètres

Nous abordons maintenant la question de l'estimation des paramètres intrinsèques au mélange et des paramètres des lois d'attache aux données d'un MMC :

$$\Theta = \{\Theta_{k,l}\}_{(k,l) \in \Omega^1 \times \Omega^2}, \quad \Theta_{k,l} = \left( \gamma_{k,l}, \tau_{k,l}, \theta_{k,l}^{(1)}, \theta_{k,l}^{(2)} \right),$$

où  $\theta_{k,l}^{(1)}$  et  $\theta_{k,l}^{(2)}$  désignent les paramètres des deux marges  $f_{k,l}^{(1)}$  et  $f_{k,l}^{(2)}$ .

La méthode ECI (« Estimation Conditionnelle Itérative ») [10] est une méthode générale d'estimation de paramètres dont le principe ne repose pas sur la maximisation de la vraisemblance (comme EM) mais sur le calcul de l'espérance d'un estimateur  $\hat{\phi}(X, Y)$  du paramètre  $\phi$ , défini sur les données complètes conditionnellement aux données observées  $\phi^{(\ell+1)} = \mathbb{E} \left[ \hat{\phi}(X, Y) \mid Y = y, \phi^{(\ell)} \right]$ . Dans le modèle MMC, cet algorithme conduit à deux situations différentes.

En notant  $\xi_n^{(\ell-1)}(k, l) = p(\mathbf{x}_n = (k, l) \mid \mathbf{y}_n; \Theta^{(\ell-1)})$ , les probabilités conjointes *a posteriori* calculées à partir des paramètres estimés à l'itération précédente, les paramètres intrinsèques s'estiment de manière déterministe selon

$$\gamma_{k,l}^{(\ell)} = \frac{1}{N} \sum_{n=1}^N \xi_n^{(\ell-1)}(k, l) \quad \text{avec} \quad \xi_n^{(\ell-1)}(k, l) = \frac{\gamma_{k,l}^{(\ell-1)} f_{k,l}^{(\ell-1)}(y_n^1, y_n^2)}{\sum_{i=1}^K \sum_{j=1}^L \gamma_{i,j}^{(\ell-1)} f_{i,j}^{(\ell-1)}(y_n^1, y_n^2)}. \quad (4)$$

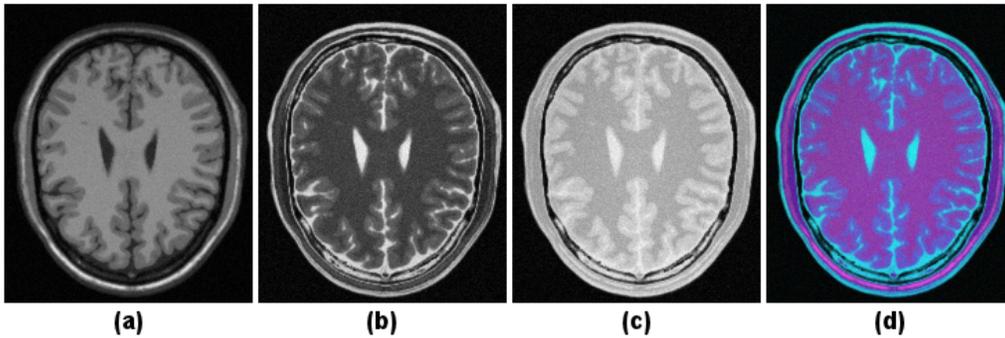
Les paramètres des lois d'attache aux données se calculent à partir d'estimateurs empiriques sur les données complétées  $(\hat{\mathbf{x}}^1, \hat{\mathbf{x}}^2, \mathbf{y}^1, \mathbf{y}^2)$ , où  $\hat{\mathbf{x}}^1 = \{\hat{x}_1^1, \dots, \hat{x}_N^1\}$  et  $\hat{\mathbf{x}}^2 = \{\hat{x}_1^2, \dots, \hat{x}_N^2\}$  sont obtenus, à l'itération  $\ell$ , par simulation, en effectuant tirage selon les distributions

$$p(x_n^1 = k \mid \mathbf{y}_n; \Theta^{(\ell-1)}) = \sum_{l=1}^L \xi_n^{(\ell-1)}(k, l), \quad (5)$$

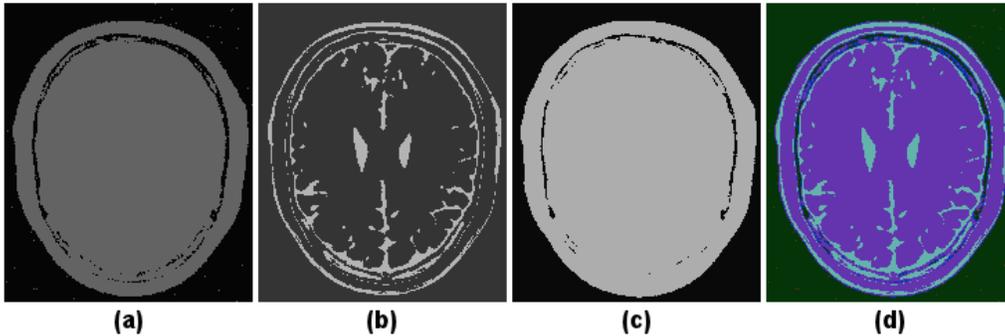
$$p(x_n^2 = l \mid x_n^1 = k, \mathbf{y}_n; \Theta^{(\ell-1)}) = \frac{\xi_n^{(\ell-1)}(k, l)}{p(x_n^1 = k \mid \mathbf{y}_n; \Theta^{(\ell-1)})}. \quad (6)$$

### 4 Segmentation non supervisée d'images

Pour illustrer l'algorithme de classification dans le cadre d'un problème de segmentation d'images, nous avons considéré trois images de modalités différentes d'une coupe de cerveau présentées dans la figure 2. Ces images ont été générées par le simulateur « BrainWeb » ([www.bic.mni.mcgill.ca/brainweb/](http://www.bic.mni.mcgill.ca/brainweb/)) [8]. L'image colorée



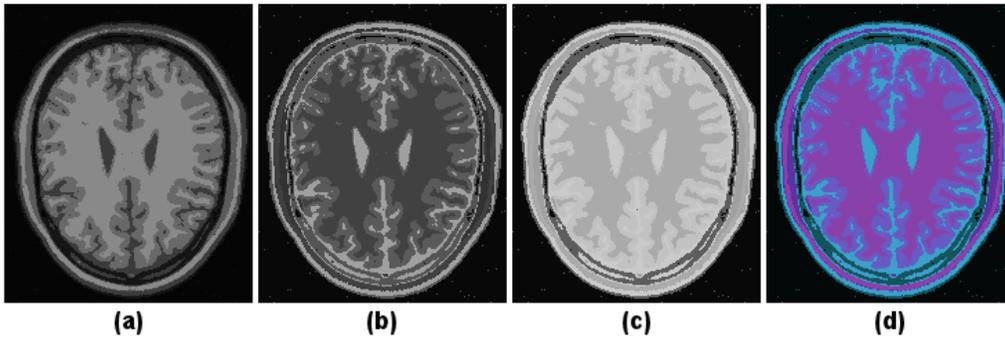
**Figure 2.** Une coupe du cerveau en modalités (a) IRM/T1, (b) IRM/T2, (c) IRM/DP (Densité de Protons) et (d) rendu coloré des 3 images. Dimensions :  $180 \times 216$ .



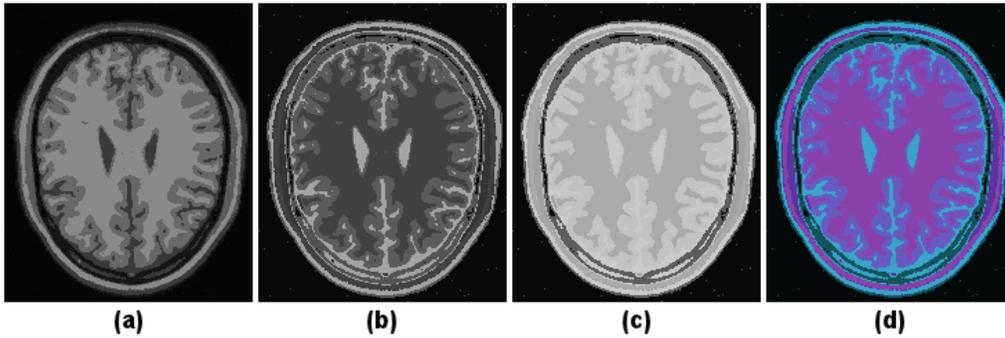
**Figure 3.** Résultats de segmentation (MPM) des 3 modalités par des modèles de mélange gaussien indépendants, avec 2 classes pour chaque coupe.

représente le rendu colorimétrique des trois coupes (couleurs rouge pour T1, verte pour T2 et bleue pour 'DP'). Les résultats de segmentation par des modèles de mélange gaussien indépendants sont reportés dans la figure 3 ; l'image couleur est le composé colorimétrique des 3 classifications. Les résultats obtenus en segmentant conjointement les 3 coupes selon le modèle MMC sont reportés dans la figure 4 lorsque l'on considère des copules gaussiennes et des marges gaussiennes, et dans la figure 5 lorsque l'on considère des copules gaussiennes et des marges Gamma. La différence entre les segmentations indépendantes et les segmentations conjointes est particulièrement notable. Le modèle conjoint est capable de prendre en considération les corrélations entre les informations fournies par chaque bande pour améliorer la classification individuelle de chaque coupe.

Par contre, on ne constate que très peu de différences lorsque l'on compare les résultats de segmentation du modèle MMC où la seule différence réside dans la forme des marginales (gaussiennes *vs* gamma). Nous avons pu constater que les lois



**Figure 4.** Résultats de segmentation (MPM) des 3 modalités par le modèle de mélange couplé MMC, avec des copules gaussiennes et des marges gaussiennes, avec 2 classes pour chaque coupe.



**Figure 5.** Résultats de segmentation (MPM) des 3 modalités par le modèle de mélange couplé MMC, avec des copules gaussiennes et des marges gamma, avec 2 classes pour chaque coupe.

gamma estimées par le MMC étaient très proches en forme des lois gaussiennes, ce qui est cohérent avec la nature gaussienne des bruits utilisés par le simulateur d'images « BrainWeb ». On peut cependant légitimement se poser la question du choix, pour une application donnée, du choix de la meilleure forme pour les densités conditionnelles. Cette question, ouverte, pourra peut être trouver une réponse numérique par le biais du calcul de distances entre les FdR empiriques et les FdR de modèles paramétriques (distance de Kolmogorov) –méthode utilisée avec succès pour la sélection de marges [6]–, ou bien encore en considérant la sélection bayésienne de copules [7].

## 5 Conclusion

Le modèle présenté dans ce travail permet de classer conjointement deux séries de  $N$  données avec des nombres de classes différents. Nous avons en particulier

étendu le cas gaussien étudié dans [5] au cas de mélanges à base de copules. Nous aboutissons ainsi à des mélanges très riches qu'il est cependant toujours possible d'estimer à l'aide d'un algorithme itératif de type ECI. Les résultats comparatifs avec des segmentations ne tenant pas compte de la dépendance montrent l'intérêt de considérer à la fois le lien entre les données cachées et le lien (non nécessairement gaussien) entre les observations.

Comme perspective à ce travail, nous envisageons, d'une part, d'évaluer une méthode de sélection automatique des meilleures copules pour modéliser le mélange parmi un ensemble fini de copules [7] et, d'autre part, d'introduire une structure de chaîne de Markov sur les classifications de manière à définir un modèle de « Chaîne de Markov Cachée Couplée ».

## Références

1. M. Brand. Coupled hidden Markov models for complex action recognition. In *Technical Report 405*, MIT Media Lab Perceptual Computing, June 1997.
2. N. Brunel, W. Pieczynski, and S. Derrode. Copulas in vectorial hidden Markov chains for multicomponent image segmentation. In *Proc. of the ICASSP'05*, Philadelphia, USA, March 19-23 2005.
3. S. Derrode. Description du système de Pearson pour son implémentation en langage C à l'aide de la librairie GSL. Research report, [www.fresnel.fr/perso/derrode/HMC\\_Ext/Doc/Pearson.pdf](http://www.fresnel.fr/perso/derrode/HMC_Ext/Doc/Pearson.pdf), École Centrale Marseille & Institut Fresnel, 2009.
4. S. Derrode. Some details about a few copulas. Research report, [www.fresnel.fr/perso/derrode/HMC\\_Ext/Doc/ListCopulas.pdf](http://www.fresnel.fr/perso/derrode/HMC_Ext/Doc/ListCopulas.pdf), École Centrale Marseille & Institut Fresnel, 2011.
5. S. Derrode and W. Pieczynski. Unsupervised restoration in Gaussian pairwise mixture model. In *19th European Signal Processing conf. (EUSIPCO'11)*, Barcelona, Spain, August 29, September 2 2011.
6. N. Giordana and W. Pieczynski. Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(5) :465–475, 1997.
7. D. Huard, G. Évin, and A.-C. Favre. Bayesian copula selection. *Computational Statistics & Data Analysis*, 51 :809–822, 2006.
8. R.K.-S. Kwan, A.C. Evancs, and G.P. Pike. MRI-simulation based evaluation of image processing and classification methods. *IEEE trans. on Medical Imaging*, 18(11) :1085–1097, 1999.
9. R. B. Nelsen. *An introduction to copulas*. Springer. Lecture Notes in Statistics, New York, USA, 1999.
10. W. Pieczynski. Statistical image segmentation. *Machine Graphics and Vision*, 1(1/2) :261–268, 1992.
11. S. Rogers *et al.* Investigating the correspondence between transcriptomic and proteomic expression profiles using coupled cluster models. *Bioinformatics*, 24(24) :2894–2900, 2008.
12. D. M. Titterton, A. F. M. Smith, and U. E. Makov. *Statistical analysis of finite mixture distributions*. Wiley Series in Probability and Statistics. John Wiley, 1985.