

Identification de distributions dans des séries temporelles multivariées présentant des corrélations à délais

Steven LE CAM¹

¹Centre de Recherche en Automatique de Nancy (CRAN), CNRS UMR 7039 - Université de Lorraine
1, av. de la forêt de Haye, 54500 Vandoeuvre-les-Nancy, France
steven.le-cam@univ.lorraine.fr

Résumé – Les modèles par chaînes de Markov cachées ont prouvé leur efficacité pour la segmentation et la restauration de signaux multivariés. En particulier, ces modèles et les extensions associées sont appréciés pour leur robustesse au bruit et pour les outils d'estimation non supervisée qui leurs sont associés. Parmi ces extensions, les Chaînes de Markov Triplet amènent une généralisation permettant de prendre en compte des distributions de bruit complexes ou encore l'introduction d'*a priori* élaborés. Dans cet article, cette modélisation est adoptée avec l'objectif d'identifier les distributions multidimensionnelles à corrélation retardée dans le temps dans des séries temporelles multivariées. Un processus auxiliaire est ainsi introduit pour modéliser les états à retard existant entre les canaux observés. Les paramètres sont estimés à partir des données à l'aide de l'algorithme d'estimation conditionnelle itérative. La méthode est validée et illustrée sur des données simulées.

Abstract – Hidden Markov Chains have proved to provide efficient methods for multivariate signals segmentation and restoration. In particular, these models and their extensions are appreciated for their robustness to noise and for the unsupervised estimation tools they provide. Among them, the triplet Markov chain stands as a fruitful generalization which can handle complex noise distributions or elaborate *prior* schemes. In this paper, this modeling is designed with the objective to identify multidimensional distributions with time-delayed correlation in multivariate time-series. An auxiliary process is introduced to model the existing delayed states between the channels. The parameters are estimated from the data based on the iterative conditional estimation algorithm. The method is validated and illustrated through experiments conducted on synthetic data.

1 Introduction

L'analyse de relations dans les séries temporelles multivariées permettent de révéler des liens de causalités entre phénomènes inter-dépendants, et amène à mieux comprendre les mécanismes qui régissent le système dont sont issues les données. L'identification de telles relations sont d'un intérêt particulier pour la compréhension de phénomènes météorologiques [10], pour l'analyse de séries temporelles en économie [4], ou encore en neuroscience [8] afin d'identifier les interactions entre structures cérébrales. Dans ce dernier cas, les relations à identifier présentent des délais dû au temps de propagation de l'information dans les canaux physiologiques. L'étude de ces relations cérébrales est basée principalement sur des méthodes de quantification statistique linéaire ou non linéaire [8] (corrélation, cohérence temps/fréquence, causalité de Granger, *etc*) à partir desquelles l'identification de liens significatifs entre voies reste une tâche de décision complexe, à laquelle s'ajoute la nécessité d'identifier les délais physiologiques de transfert de l'information. Dans cet article, une méthode d'identification de changement dans les distributions statistiques multidimensionnelles prises par les voies d'observation est proposée, amenant à la caractérisation des interactions et délais entre les voies. Les relations à retard sont ici quantifiées à l'aide d'une mesure d'inter-corrélation linéaire.

Sous hypothèse de stationnarité statistique sur le processus caché à restaurer, les outils d'analyse par chaîne de Markov cachée associés au formalisme bayésien se montrent très efficace pour identifier des séquences d'intérêt dans des contextes très perturbés. L'utilisation de lois multidimensionnelles bien adaptées aux données, notamment via l'utilisation de la théorie des copules, est à même de fournir une estimation des relations linéaires entre les canaux lorsqu'ils sont en phase [2, 6]. En revanche, lorsque des délais sont introduits, il est nécessaire d'étendre le modèle pour pouvoir prendre en compte ce paramètre temporel. Le modèle par chaîne de Markov triplet est de ce fait introduit. En généralisant les modèles par chaînes de Markov cachées [9], il offre un pouvoir de modélisation plus riche permettant notamment de modéliser des aspects non stationnaires du processus caché à restaurer [5], ou encore de prendre en compte des changements stochastiques dans les distributions des données observées [1]. Ce modèle est ici adapté afin d'identifier les changements dans les distributions statistiques des signaux multi-voies, en tenant compte de la présence de corrélations à délai. Dans le cadre de ce travail, une modélisation par lois multidimensionnelles gaussiennes est proposée. L'estimation de l'ensemble des paramètres du modèle est menée à l'aide de l'algorithme d'estimation conditionnelle itérative (ECI), et les performances de la méthode sont illustrées sur un cas en simulation.

2 Modèle de Markov triplet

Soit $X = (X_t)_{1 \leq t \leq T}$ et $\mathbf{Y} = (\mathbf{Y}_t)_{1 \leq t \leq T}$ deux processus stochastiques. X est un processus dit *caché* à valeur dans un ensemble fini $\Omega = \{\omega_1, \dots, \omega_N\}$, et $(\{\mathbf{Y}_t\}_{1 \leq t \leq T} \in \mathbb{R}^M)$ est le processus d'observation, où M voies sont observées sur T échantillons. Les méthodes de segmentation par chaînes de Markov cachées visent à restaurer les états cachés de X à partir des observations \mathbf{Y} . Une loi d'attache aux données modélisant la vraisemblance de \mathbf{Y} conditionnellement à X est définie. Il est possible de proposer des lois bien adaptées aux distributions des données observées, notamment grâce à l'introduction de copules [2, 6, 5]. En revanche, de tels modèles sont incapables d'intégrer une notion de relation à délai entre les canaux, seules les corrélations *instantanées* inter-voies sont considérées. Ce paramètre supplémentaire sera modélisé par un processus auxiliaire noté U , menant au formalisme par chaîne de Markov triplet, et permettant de relâcher la contrainte Markovienne sur le processus caché X .

Soit U un processus stochastique, prenant ses valeurs dans l'ensemble $\Lambda = \{\lambda_1, \dots, \lambda_K\}$. Chaque état de λ_k modélise une configuration de délais particulière entre les canaux. Les délais associés à l'état λ_k sont notés $\tau^{\lambda_k} \in \mathbb{Z}^{M_s}$ (où $M_s = M * (M - 1) / 2$ corrélations bivariées sont à considérer). Une observation des corrélations à délai pour chaque état $\lambda_k \in \Lambda$ est dérivée par calcul des corrélations linéaires entre les canaux de \mathbf{Y} pour les délais $\tau^{\lambda_k} = \{\tau_1^{\lambda_k}, \dots, \tau_{M_s}^{\lambda_k}\}$. Soit y_n et y_m deux canaux de \mathbf{Y} . La corrélation linéaire à délai est donnée par :

$$s_{m_s, t}^k = |\rho_{m_s, t}^{\tau^{\lambda_k}}| \quad (1)$$

$$\text{avec } \rho_{m_s, t}^{\tau} = \frac{\text{cov}(y_{m_p, t}, y_{m_q, [t - \tau]})}{\sqrt{\text{var}(y_{m_p, t}) \text{var}(y_{m_q, [t - \tau]})}} \quad (2)$$

où $\text{cov}(\cdot)$ and $\text{var}(\cdot)$ sont des estimateurs empiriques de la covariance et de la variance calculé sur une fenêtre autour de l'instant t , et $|x|$ désigne la valeur absolue de la valeur réelle x . $\mathbf{S}^k = (\{s_{m_s, t}^k\}_{1 \leq t \leq T, 1 \leq m_s \leq M_s})$ quantifie ainsi les M_s corrélations bivariées à délais pour chaque état λ_k . On note $T = (X, U, \mathbf{O})$, où le couple de variables cachés $V = (X, U)$ est à restaurer à partir du couple d'observations $\mathbf{O} = (\mathbf{Y}, \mathbf{S})$. V est supposé à partir du couple d'observations $\mathbf{O} = (\mathbf{Y}, \mathbf{S})$. V est supposé de Markov, et T est alors une chaîne de Markov triplet, aboutissant à la propriété suivante [5] :

(i) $\mathbf{O}_1, \dots, \mathbf{O}_T$ sont indépendants conditionnellement à $V = (X, U)$.

L'hypothèse supplémentaire suivante est également émise :

(ii) $s_{1, t}^k, \dots, s_{M_s, t}^k$ sont indépendants conditionnellement à u_t

L'évidente dépendance entre \mathbf{S} et \mathbf{Y} étant bien entendu conservée à travers le processus U . Ce modèle de dépendance est illustré sur le graphe fig. 1.

On a alors accès à deux lois de vraisemblance séparables pour \mathbf{Y} et pour \mathbf{S} : $f_{\mathbf{y}}^{\omega_n, \lambda_k}(\mathbf{y}_t^{\lambda_k}) = p(\mathbf{y}_t^{\lambda_k} | v_t = (\omega_n, \lambda_k))$, et $f_{\mathbf{s}}^{\lambda_k}(s_{m_s, t}^k) = p(s_{m_s, t}^k | u_t = \lambda_k)$, où $\mathbf{y}_t^{\lambda_k}$ est le vecteur d'observations \mathbf{y}_t recalé en considérant le vecteur de délais τ^{λ_k} dans l'état λ_k . Dans cette étude, $f_{\mathbf{y}}^{\omega_n, \lambda_k}$ est modélisée par une

loi gaussienne multivariée. Quant à $f_{\mathbf{s}}^{\lambda_k}$, une loi de vraisemblance à support fini entre 0 et 1 doit être choisie. On opte ici pour la famille de lois Beta, offrant une large possibilité de distributions sur ce support borné. Un couple de paramètres $\{\alpha_{k, m_s}, \beta_{k, m_s}\}$ pour chaque état et chaque couple de voies sont de ce fait à estimer.

Dans ce contexte, la vraisemblance des observations conditionnellement aux états cachés s'écrit :

$$p(\mathbf{o}_t | v_t = (\omega_n, \lambda_k)) = f_{\mathbf{y}}^{\omega_n, \lambda_k}(\mathbf{y}_t^{\lambda_k}) f_{\mathbf{s}}^{\lambda_k}(\mathbf{s}_t^k) \quad (3)$$

$$\text{avec } f_{\mathbf{s}}^{\lambda_k}(\mathbf{s}_t^k) = \prod_{m_s=1}^{M_s} f_{s_{m_s, t}}^{\lambda_k}(s_{m_s, t}^k) \quad (4)$$

et T admet la factorisation suivante :

$$p(T) = p(t_1) \prod_{t=2}^T p(t_t | t_{t-1}), \text{ avec } p(t_1) = p(\mathbf{o}_1 | v_1) p(v_1) \quad (5)$$

$$p(t_t | t_{t-1}) = p(\mathbf{o}_t | v_t) p(v_t | v_{t-1}) \quad (6)$$

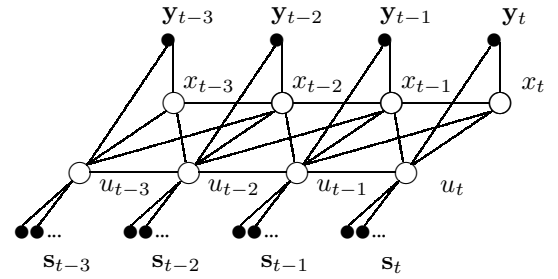


FIGURE 1 – Représentation graphique du modèle

3 Décision bayésienne

Le calcul des marginales *a posteriori* $p(v_t | \mathbf{O})$ est possible à partir du calcul des quantités forward $\alpha_t(v_t) = p(\mathbf{o}_1, \dots, \mathbf{o}_t, v_t)$ et backward $\beta_t(v_t) = p(\mathbf{o}_{t+1}, \dots, \mathbf{o}_T | v_t)$:

$$\alpha_1(v_1) = f_{\mathbf{y}}^{v_1}(\mathbf{y}_1^{u_1}) f_{\mathbf{s}}^{u_1}(\mathbf{s}_1) p(v_1)$$

$$\alpha_t(v_t) = \sum_{v_{t-1} \in \Omega \times \Lambda} \alpha_{t-1}(v_{t-1}) p(v_t | v_{t-1}) \cdot \quad (7)$$

$$f_{\mathbf{y}}^{v_t}(\mathbf{y}_t^{u_t}) f_{\mathbf{s}}^{u_t}(\mathbf{s}_t)$$

$$\beta_T(v_T) = 1$$

$$\beta_t(v_t) = \sum_{v_{t+1} \in \Omega \times \Lambda} \beta_{t+1}(v_{t+1}) p(v_{t+1} | v_t) \cdot \quad (8)$$

$$f_{\mathbf{y}}^{v_{t+1}}(\mathbf{y}_{t+1}^{u_{t+1}}) f_{\mathbf{s}}^{u_{t+1}}(\mathbf{s}_{t+1})$$

La distribution *a posteriori* $p(v_t | \mathbf{O})$ peut alors être obtenue :

$$p(v_t | \mathbf{O}) \propto \beta_t(x_t, u_t) \alpha_t(x_t, u_t) \quad (9)$$

Les marginales *a posteriori* pour les processus x_t et u_t sont immédiatement déduites :

$$\begin{aligned} p(x_t|\mathbf{O}) &= \sum_{u_t \in \Lambda} p(v_t|\mathbf{O}) \\ p(u_t|\mathbf{O}) &= \sum_{x_t \in \Omega} p(v_t|\mathbf{O}) \end{aligned} \quad (10)$$

La restauration de ces deux processus est obtenue à l'aide de l'estimateur bayésien du Mode des Marginales *a posteriori* : $\hat{x}_t = \arg \max_{x_t \in \Omega} p(x_t|\mathbf{O})$ et $\hat{u}_t = \arg \max_{u_t \in \Lambda} p(u_t|\mathbf{O})$. [7] :

$$\begin{aligned} \hat{x}_t &= \arg \max_{x_t \in \Omega} p(x_t|\mathbf{O}) \\ \hat{u}_t &= \arg \max_{u_t \in \Lambda} p(u_t|\mathbf{O}) \end{aligned} \quad (11)$$

4 Estimation des paramètres

Nous nous plaçons dans le cas d'une modélisation non supervisée, les paramètres sont inconnus et doivent être estimés à partir des observations \mathbf{O} . Nous optons ici pour la procédure ECI [6, 5] qui fournit une méthode d'estimation stochastique des paramètres à partir d'une réalisation \hat{v} de la chaîne cachée couple $V = (X, U)$ simulée à l'aide des quantités *a posteriori* $p(v_t|\mathbf{O})$. Les paramètres à estimer sont (i) les paramètres de la chaîne V (probabilités initiales $\pi_{nk} = p(v_1 = (\omega_n, \lambda_k))$ et matrice de transition $a_{nk,ml} = p(v_t = (\omega_m, \lambda_l) | v_{t-1} = (\omega_n, \lambda_k))$), (ii) les paramètres d'attache aux données gaussiens $\{\Sigma_{nk}, \mu_{nk}\}$, (iii) les deux couples de paramètres de lois Beta associés aux valeurs de corrélations. On proposera également une estimation empirique des paramètres de délai τ^{λ_k} à partir de la réalisation \hat{v} .

4.1 Initialisation

Les paramètres *a priori* sur X sont initialisés de façon déterministes : $\pi_{nk} = \frac{1}{N * K}$, $a_{nk,ml} = \frac{99}{100}$ si $(n, k) = (m, l)$, $a_{nk,ml} = \frac{1}{100 * (N * K - 1)}$ sinon. L'initialisation des paramètres de loi normal sont calculées empiriquement à partir des observations \mathbf{Y} . L'initialisation des classes de délai ainsi que des valeurs de délais associés reste une question délicate et conditionne le bon déroulement de l'algorithme. Dans le cadre d'une application neurophysiologique par exemple, on pourra disposer de connaissances cliniques *a priori* sur les délais, ou tout du moins sur certaine borne admissible. Une initialisation empirique à partir des valeurs de corrélations calculées est ici proposée. Le nombre de classe de délai K est déterminé à partir des plages de corrélation significative (supérieur à un seuil donné) évalué dans les corrélations linéaires, les vecteurs de délai associés τ^{λ_k} sont alors immédiatement déduits. Les paramètres d'attache aux données de corrélation $\{\alpha_i, \beta_i\}$ sont initialisés à partir des observations de corrélation correspondant à ces plages de corrélation significatives.

4.2 réestimation ECI

- Calcul des probabilités *a posteriori* :

$\xi_t^{[q]}(n, k) = P(v_t = (\omega_n, \lambda_k) | \mathbf{O})$ and $\Psi_t^{[q]}(n, k, m, l) = P(v_t = (\omega_n, \lambda_k), v_{t+1} = (\omega_m, \lambda_l) | \mathbf{O})$ sont obtenus à l'aide de

l'algorithme forward-backward. Une réalisation $\hat{v}^{[q]} = (\hat{x}^{[q]}, \hat{u}^{[q]})$ est calculée à partir de ces quantités. $\hat{u}^{[q]}$ sera utilisé pour estimer de façon stochastique les paramètres de loi beta ainsi que les valeurs de délai τ^{λ_k} .

- La réestimation des paramètres *a priori* π_{nk} , $a_{nk,ml}$, et gaussien Σ_{nk} , μ_{nk} est donné par :

$$\pi_{nk}^{[q]} = \xi_1^{[q]}(n, k), \quad a_{nk,ml}^{[q]} = \frac{\sum_{t=1}^{T-1} \Psi_t^{[q]}(n, k, m, l)}{\sum_{t=1}^{T-1} \xi_t^{[q]}(n, k)} \quad (12)$$

$$\mu_{nk}^{[q]} = \frac{\sum_{t=1}^T \xi_t^{[q]}(n, k) \mathbf{y}_t^{\lambda_k}}{\sum_{t=1}^T \xi_t^{[q]}(n, k)} \quad (13)$$

$$\Sigma_{nk}^{[q]} = \frac{\sum_{t=1}^T \xi_t^{[q]}(n, k) (\mathbf{y}_t^{\lambda_k} - \mu_{nk}^{[q]}) (\mathbf{y}_t^{\lambda_k} - \mu_{nk}^{[q]})^T}{\sum_{t=1}^T \xi_t^{[q]}(n, k)} \quad (14)$$

- L'ECI amène a une réestimation stochastique des moyennes et variances des populations de corrélation significative $s_{i>0}^{[q]}$ et non significative $s_0^{[q]}$ associées à l'estimation courante $\hat{u}^{[q]}$:

$$\mu_i^{[q]} = \frac{\sum_{s \in s_i^{[q]}} s}{N_i}, \quad \sigma_i^{[q]} = \frac{\sum_{s \in s_i^{[q]}} (s - \mu_i^{[q]})^2}{N_i} \quad (15)$$

où N_i correspond au cardinal de l'ensemble $s_i^{[q]}$. On obtient alors une estimation des paramètres de loi beta à l'itération q [3] :

$$\alpha_i^{[q]} = \mu_i^{[q]} \left(\frac{\mu_i^{[q]} (1 - \mu_i^{[q]})}{\sigma_i^{[q]}} - 1 \right) \quad (16)$$

$$\beta_i^{[q]} = (1 - \mu_i^{[q]}) \left(\frac{\mu_i^{[q]} (1 - \mu_i^{[q]})}{\sigma_i^{[q]}} - 1 \right) \quad (17)$$

- De façon similaire, et dans le cas où on cherche à maximiser les valeurs de corrélation, on peut proposer une réestimation stochastique des paramètres de délais $\tau^{\lambda_k, [q]}$ à partir de la réalisation $\hat{u}^{[q]}$:

$$\tau_{m_s}^{\lambda_k, [q]} = \frac{\sum_{t \in J_k} \tau_{m_s, t}^{\max}}{M_k} \quad (18)$$

où $\tau_{m_s, t}^{\max}$ représente le délai maximisant la corrélation pour le couple de voies m_s à l'instant t , et $J_k = \{t | \hat{u}_t^{[q]} = \lambda_k\}$, avec M_k le cardinal de cet ensemble. Le vecteur d'observation \mathbf{S} est alors à nouveau calculé pour chaque classe λ_k à partir de l'équation (1) avec les délais mis à jour $\tau^{\lambda_k, [q]}$.

5 Résultats

La méthode est illustrées sur un jeu de données simulées. Soit z_1, z_2, z_3 et z_4 trois processus gaussiens centrés, réduits, et non corrélés entre eux. On construit alors les trois canaux d'observation suivants :

$$y_1(t) = d_i (\sqrt{1 - a_k} z_1(t) + \sqrt{a_k} z_4(t)) \quad (19)$$

$$y_2(t) = d_i (\sqrt{1 - b_k} z_2(t) + \sqrt{b_k} z_4(t - \tau_k^1)) \quad (20)$$

$$y_3(t) = d_i (\sqrt{1 - c_k} z_3(t) + \sqrt{c_k} z_4(t - \tau_k^2)) \quad (21)$$

où l'ensemble de paramètres $\{a_k, b_k, c_k, \tau_k^1, \tau_k^2\}$ est associé à un état de corrélation à retard λ_k . Les valeurs de corrélations inter-signaux (a_k, b_k et c_k) sont fixées à 0.3, tandis que les délais sont simulés uniformément entre -8 et 8 . Quatre classes de délais λ_k et deux classes de distributions gaussiennes ω_i sont simulées. Le coefficient d_i permet de décrire les différents états ω_i , pour lesquels on fait ainsi évoluer la variance des distributions observées entre 1 ($d_1 = 1$) et 1.4 ($d_2 = 1.2$). Les différents paramètres gaussiens (variances, moyennes et corrélations) sont fournis dans le tableau 1, avec leur estimée par la méthode proposée pour une réalisation. Le résultat de segmentation est donné sur la figure 2. Tous les états sont bien identifiés. On note quelques erreurs dû aux sauts entre deux classes de délais successifs dans le signal. ce contexte de simulations où les corrélations introduites sont bien identifiables dans \mathbf{S} , les valeurs des délais sont parfaitement retrouvées. Il est à noter que les variances des marginales des distributions gaussiennes étant identiques pour chacune des quatre classes dans Λ , seules les paramètres de corrélations et de délais permettent de distinguer ces classes. La méthode est donc ici évaluée à l'aune de sa capacité à discriminer les différentes configuration de corrélations à retard observées dans le signal.

TABLE 1 – Paramètres des distributions gaussiennes simulées avec leur version estimée, pour chaque couple $v = \{\lambda_i, \omega_j\}$

	d_1	\hat{d}_1	d_2	\hat{d}_2	d_3	\hat{d}_3	a	\hat{a}	b	\hat{b}	c	\hat{c}
λ_1, ω_1	1	1.00	1	1.02	1	0.99	0	0.05	0	0.05	0	0.03
λ_1, ω_2	1.2	1.19	1.2	1.17	1.2	1.19	0	0.02	0	0.04	0	0.01
λ_2, ω_1	1	1.01	1	1.00	1	1.00	0.3	0.31	0	0.01	0	0.04
λ_2, ω_2	1.2	1.19	1.2	1.20	1.2	1.22	0.3	0.30	0	0.00	0	0.04
λ_3, ω_1	1	0.97	1	0.99	1	1.00	0	0.00	0.3	0.32	0	0.01
λ_3, ω_2	1.2	1.21	1.2	1.18	1.2	1.15	0	0.03	0.3	0.3	0	0.01
λ_4, ω_1	1	1.00	1	0.98	1	1.03	0.3	0.32	0.3	0.31	0.3	0.31
λ_4, ω_2	1.2	1.18	1.2	1.20	1.2	1.21	0.3	0.29	0.3	0.27	0.3	0.29

Sur cet ensemble de données simulées où aucune perturbation sur les relations n'est introduite, une approche naïve par seuillage est suffisante afin d'identifier les états à retard éligible λ_k et les délais associés. Cette initialisation devient problématique dans des applications réelles, où les mesures d'observation de la corrélation \mathbf{S} peuvent être perturbées. Dans le cadre de l'analyse de relations entre structures cérébrales, des *a priori* sur les valeurs de retard sont disponibles à partir de l'expertise des neurologues ou de données tractographiques. L'introduction de ces connaissances amènerait à une amélioration de la méthode afin de contrôler l'initialisation et l'estimation de ce paramètre de retard inter-voies. Dans ce contexte applicatif, l'extension de lois gaussiennes vers des distributions plus évoluées (notamment par l'introduction de copules) mieux adaptés aux activités électriques enregistrées est susceptible d'améliorer le pouvoir modélisant de la méthode.

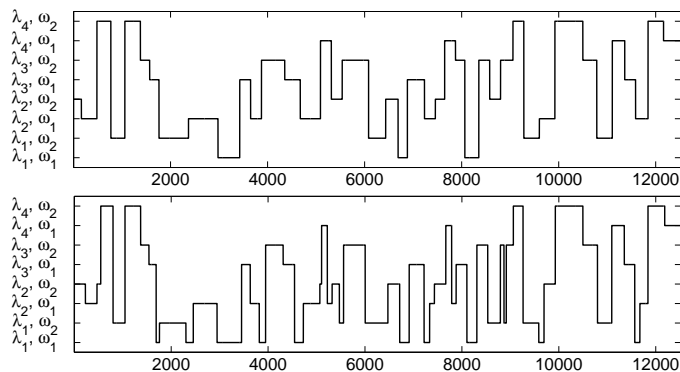


FIGURE 2 – Résultats de segmentation pour les 3 canaux y_1, y_2, y_3 . Vérité-terrain (haut), et résultat de segmentation (bas).

Références

- [1] M.E.Y. Boudaren, E. Monfrini, and W. Pieczynski. Unsupervised segmentation of random discrete data hidden with switching noise distributions. *Signal Processing Letters, IEEE*, 19(10) :619–622, 2012.
- [2] N. Brunel and W. Pieczynski. Unsupervised signal restoration using hidden Markov chains with copulas. *Signal Processing*, 85(12) :2304–2315, 2005.
- [3] M. Evans, N. Hastings, and B. Peacock. *Statistical distributions*, chapter 5 : "Beta distribution". Wiley-Interscience, 2000.
- [4] C. Hiemstra and J.D. Jones. Testing for linear and nonlinear granger causality in the stock price-volume relation. *The Journal of Finance*, 49(5) :1639–1664, 2012.
- [5] P. Lanchantin, J. Lapuyade-Lahorgue, and W. Pieczynski. Unsupervised segmentation of randomly switching data hidden with non-gaussian correlated noise. *Signal Processing*, 91(2) :163–175, 2011.
- [6] S. Le Cam, A. Belghith, Ch. Collet, and F. Salzenstein. Wheezing Sounds Detection Using Multivariate Generalized Gaussian Distributions. *Acoustics, Speech and Signal Processing, IEEE International Conference on*, 2009.
- [7] J. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computer vision. *Journal of the American Statistical Association*, 82(397) :76–89, 1987.
- [8] E. Pereda, R.Q. Quiroga, and J. Bhattacharya. Nonlinear multivariate analysis of neurophysiological signals. *Progress in neurobiology*, 77(1-2) :1–37, 2005.
- [9] W. Pieczynski. Chaînes de markov triplet. *Comptes Rendus de l'Académie des Sciences - Mathématique, Série I*, 335(3) :275–278, 2002.
- [10] MS Santhanam and P.K. Patra. Statistics of atmospheric correlations. *Physical Review E*, 64(1) :016102, 2001.