COMPENSATION DES RÉPONSES EN FRÉQUENCES DES CAPTEURS EN SÉPARATION AVEUGLE DE SOURCES

Massoud BABAIE-ZADEH[1,2], Christian JUTTEN[1] et Kambiz NAYEBI[2]

[1] Institut National Polytéchnique de Grenoble (INPG), Laboratoire des images et des signaux (LIS)
[2] Université technologie de Sharif, Téhéran, Iran
mbzadeh@yahoo.com Christian.Jutten@inpg.fr knayebi@sina.sharif.ac.ir

RÉSUMÉ

La séparation aveugle de source est un problème fondamental en traitement des signaux. Dans le cas général de mélanges convolutifs, le critère d'indépendance est complexe car il doit s'appliquer à un processus aléatoire. Dans cet article, nous traitons un cas particulier, dans lequel le canal introduit un mélange linéaire instantané et le filtrage est lié à la réponse en fréquence des capteurs. Dans ce cas, on montre le critère d'indépendance des variables aléatoires est suffisant pour réaliser la séparation.

1. INTRODUCTION

La séparation aveugle de source est un problème important en traitement des signaux et a été largement étudié dans la dernière décennie. Dans le cas instantané linéaire, le mélange est de la forme :

$$\mathbf{x} = \mathbf{A}\mathbf{s} \qquad (1)$$

où $\mathbf{s}$ est le vecteur des sources, $\mathbf{x}$ est le vecteur des observations, et $\mathbf{A}$ est la matrice mélangeante. Pour séparer les sources, on estime une matrice de séparation, $\mathbf{B}$, telle que :

$$\mathbf{y} = \mathbf{B}\mathbf{x} \qquad (2)$$

On montre [2] que l'indépendance des composantes de $\mathbf{y}$, est suffisante pour obtenir la séparation (avec des indéterminations sur les ordres et le gain des sources).

Dans le cas convolutif, les matrices de mélanges et de séparation sont des matrices de filtres, fréquemment des filtres à réponse impulsionnelle finie [5, 4, 1]. Pour ce type de mélanges, on montre que l'indépendance des sorties est suffisante pour obtenir la séparation (à un filtre près) [6]. Mais, le critère d'indépendance de y_1 et y_2 ne se restreint pas à l'indépendance de $y_1(n)$ et $y_2(n)$. Il doit inclure l'indépendance des paires $y_1(n)$ et $y_1(n-m)$ pour toutes les paires de m et n.

Dans ce papier, nous considérons un cas particulier de mélanges convolutifs dans lequel le canal induit un mélange linéaire instantané, mais les réponses fréquentielles des capteurs sont modélisées par des filtres. Dans ce cas spécial, on montre que l'indépendance de

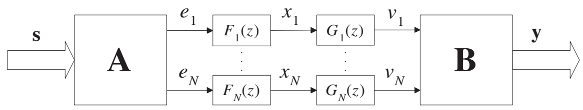

FIG. 1 – Le système mélangeant-séparateur

$y_1(n)$ et $y_2(n)$ à l'instant n est suffisante pour la séparation. Ce résultat permet de concevoir un algorithme très simple pour séparer ce type de mélanges.

2. PRÉLIMINAIRES

Considérons le système de figure 1, où les sources sont d'abord mixées par un mélange linéaire et instantané, puis filtrées, composante par composante, par les filtres $F_i(z)$, associés à la réponse fréquentielle des capteurs ou des amplificateurs qui suivent ces capteurs. Pour séparer le mélange, on doit d'abord compenser l'effet du filtrage (en estimant des filtres G_i), puis estimer une matrice de séparation. Dans la suite, le filtre total, sur chaque voie i sera noté :

$$H_i(z) = F_i(z)G_i(z) \qquad (3)$$

Il est clair que si :

$$H_i(z) = c_{ij}H_j(z) \qquad \forall i, j \qquad (4)$$

et BA est la produit d'une matrice de permutation et d'une matrice diagonale, alors les sorties sont indépendantes. Donc, les sources ne peuvent être restituées qu'à un filtre près.

Avec ce mélange convolutif particulier, nous pouvons montrer les résultats suivants :

Théorème 1 *Soient 2 sources s_i centrées iid, dont au plus une est gaussienne, soit la matrice $\mathbf{A} = \begin{pmatrix} 1 & a_1 \\ a_2 & 1 \end{pmatrix}$ régulière telle que $a_1 a_2 \neq -1$, et soient les filtres vérifiant la condition : $\exists m : h_1(m)h_2(m) \neq 0$, si $y_1(n)$ et $y_2(n)$ sont indépendants, alors $y_1(n)$ et $y_2(n-m)$ sont indépendants $\forall m$.*

La preuve est donnée dans l'annexe.

Ce théorème montre que, pour les mélanges particuliers considérés, l'indépendance de $y_1(n)$ et $y_2(n)$ est une condition nécessaire et suffisante d'indpendence des processus alatoires y_1 et y_2. Ce théorème s'étend facilement au cas $N \times N$ en considérant l'indépendance par paire.

On peut aussi montrer le rsultat suivant :

Théorème 2 *Sous les hypothèses du théorème 1 et si la matrice A satisfait la condition : $a_1 \neq 0$ ou $a_2 \neq 0$, l'indépendance de $y_1(n)$ et $y_2(n)$ implique que $H_1(z) = cH_2(z)$, où c est une constante arbitraire.*

La preuve est donnée dans l'annexe.

3. EQUATIONS D'ESTIMATION

Les filtres $F_i(z)$ associés aux capteurs ou aux amplificateurs sont typiquement des filtres causaux que l'on peut modéliser par des filtres à réponse impulsionnelle finie (RIF) d'ordre réduit (1 à 3) qui vérifient $f_i(0) \neq 0, \forall i$.

Les filtres compensateurs $G_i(z)$ sont choisis de type RIF, causaux, d'ordre M, de la forme :

$$G_i(z) = 1 + \sum_{j=1}^{M} g_{ij} z^{-j} \tag{5}$$

Compte tenu de l'hypothèse $f_i(0) \neq 0$ et de la forme des filtres G_i, l'hypothèse $h_i(0) \neq 0$ du théorème 1 est satisfaite, et le théorème peut s'appliquer.

Soit $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ le vecteur aléatoire des sorties, on utilise l'information mutuelle, $I(\mathbf{y})$, comme critère d'indépendance :

$$I(\mathbf{y}) = \sum_{k=1}^{N} H(y_k) - H(\mathbf{y}) \tag{6}$$

où $H(\cdot) = -E \ln p_{(\cdot)}(\cdot)$ est l'entropie de Shannon.

3.1. Séparation de la partie instantanée

En utilisant les relations de changements de variables dans les densités, on peut exprimer l'entropie de $\mathbf{y}$ en fonction de l'entropie de $\mathbf{v} = (v_1, v_2, \ldots, v_N)^T$:

$$H(\mathbf{y}) = H(\mathbf{v}) + \ln|\det(\mathbf{B})|. \tag{7}$$

Le gradient de $H(\mathbf{y})$ par rapport à $\mathbf{B}$ s'écrit :

$$\frac{\partial H(\mathbf{y})}{\partial \mathbf{B}} = \frac{\partial}{\partial \mathbf{B}} |\det(\mathbf{B})| = \mathbf{B}^{-T} \tag{8}$$

De plus, à partir de la définition de l'entropie et de $y_j = \sum_{k=1}^{N} b_{jk} v_k$, on obtient :

$$\frac{\partial}{\partial b_{ij}} \left(\sum_{k=1}^{N} H(y_k) \right) = -E\left\{\psi_i(y_i) v_j\right\} \tag{9}$$

où $\psi_i = \frac{p'_{y_i}(\cdot)}{p_{y_i}(\cdot)}$ est la fonction score de y_i.

En combinant (8) et (9), nous avons :

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{B}} = -E\left\{\boldsymbol{\psi}_{\mathbf{y}}(\mathbf{y})\mathbf{v}^T\right\} - \mathbf{B}^{-T} \tag{10}$$

où l'on note :

$$\boldsymbol{\psi}_{\mathbf{y}}(\mathbf{y}) = \left(\psi_1(y_1), \psi_2(y_2), \ldots, \psi_N(y_N)\right)^T. \tag{11}$$

3.2. Compensation des filtres

De l'équation (7), on calcule :

$$\begin{aligned} \frac{\partial H(\mathbf{y})}{\partial g_{ij}} = \frac{\partial H(\mathbf{v})}{\partial g_{ij}} &= -E\left\{\frac{\partial}{\partial g_{ij}} \ln p_{\mathbf{v}}(\mathbf{v})\right\} \\ &= -E\left\{\frac{\partial \ln p_{\mathbf{v}}(\mathbf{v})}{\partial v_i} . \frac{\partial v_i}{\partial g_{ij}}\right\} \end{aligned} \tag{12}$$

Donc :

$$\frac{\partial H\left(\mathbf{y}(n)\right)}{\partial g_{ij}} = -E\left\{\varphi_{\mathbf{v},i}\left(\mathbf{v}(n)\right) x_i(n-j)\right\} \tag{13}$$

où $\varphi_{\mathbf{v},i}(\mathbf{v})$ est la i-ème composante de la fonction score conjointe [1] de $\mathbf{v}$, définie par :

$$\begin{aligned} \boldsymbol{\varphi}_{\mathbf{v}}(\mathbf{x}) &= \nabla_{\mathbf{x}} \ln p_{\mathbf{v}}(\mathbf{x}) \\ &= \frac{1}{p_{\mathbf{v}}(\mathbf{x})} \left(\frac{\partial p_{\mathbf{v}}(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial p_{\mathbf{v}}(\mathbf{x})}{\partial x_N} \right)^T \end{aligned} \tag{14}$$

Le calcul de la dérivée de $H(y_k)$ par rapport à g_{ij} donne :

$$\frac{\partial H(y_k)}{\partial g_{ij}} = -E\left\{b_{ki}\psi_k(y_k)\frac{\partial v_i}{\partial g_{ij}}\right\} \tag{15}$$

Donc :

$$\frac{\partial}{\partial g_{ij}} \sum_{k=1}^{N} H(y_k) = -E\left\{ \left(\sum_{k=1}^{N} b_{ki}\psi_k\left(y_k(n)\right) \right) x_i(n-j) \right\} \tag{16}$$

En notant la matrice $N \times M$, $\mathbf{G} = (g_{ij})$, et les vecteurs $\mathbf{X}(n) = [\mathbf{x}(n-1), \mathbf{x}(n-2), \cdots, \mathbf{x}(n-M)]$ et $\psi_{diff}(n) = \varphi_{\mathbf{v}}(\mathbf{v}(n)) - \mathbf{B}^T\psi_{\mathbf{y}}(\mathbf{y}(n))$, et en appliquant (13) et (16), on a :

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{G}} = E\left\{diag\left(\boldsymbol{\psi}_{diff}(n)\right)\mathbf{X}(n)\right\} \tag{17}$$

4. ALGORITHME

La mise en oeuvre des équations d'estimations requiert l'estimation préalable des fonctions score marginales et multivariables [3]. Nous proposons un algorithme par bloc pour la séparation, mais une version adaptative est bien entendu envisageable.

En ce qui concerne la partie instantanée, l'équation d'estimation est identique au cas instantané (10). Donc, l'algorithme d'estimation de $\mathbf{B}$ s'écrit :

$$\mathbf{B}_{n+1} = (I + \mu_1 \mathbf{D})\mathbf{B}_n \quad (18)$$

où :

$$d_{ij} = \begin{cases} 1 - y_i^2 & \text{if } i = j \\ \hat{E}\left\{\psi_i(y_i)y_j\right\} & \text{if } i \neq j \end{cases} \quad (19)$$

Les filtres G_i sont ajustés selon (17), à partir d'un algorithme de descente de gradient :

$$\mathbf{G}_{n+1} = \mathbf{G}_n - \mu_2 \hat{E}\left\{diag\left(\boldsymbol{\psi}_{diff}(n)\right)\mathbf{X}(n)\right\} \quad (20)$$

où $\hat{E}$ est l'espérance mathématique empirique, et μ_2 est le pas d'adaptation.

5. RÉSULTATS EXPÉRIMENTAUX

Pour valider notre méthode, nous avons choisi des filtres $F_i(z)$ passe-bas du premier ordre :

$$F_i(z) = \frac{1}{1 - \alpha_i z^{-1}} \quad (21)$$

et des filtre compensateurs de la forme :

$$G_i(z) = 1 + g_i z^{-1}. \quad (22)$$

Ce choix règle le problème d'indétermination (la constante c) qui apparaît dans le théorème 2.

Dans la première expérience, 100 échantillons de deux signaux iid, à distribution uniforme, sont mélangés par la matrice :

$$A = \begin{pmatrix} 1 & 0.5 \\ 0.3 & 1 \end{pmatrix}$$

Les coefficients des filtres sont $\alpha_1 = 0.6$ et $\alpha_2 = 0.8$. Nous avons utilisé $\mu_1 = 0.1$ et $\mu_2 = 0.05$ (constante). Les résultats de séparation, obtenus avec l'algorithme ci-desus, sont moyennés sur 50 expériences. On trouve, en moyenne, $g_1 = -0.5954$ et $g_2 = -0.7942$, avec des variances de 0.0023 et 0.0008, respectivement.

Dans une deuxième expérience, on a choisi une matrice de mélange mal conditionnée :

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 5 \end{pmatrix}$$

Nous avons utilisé $\mu_1 = 0.1$ et $\mu_2 = 0.0005$. En moyenne sur 50 expériences, on obtient $g_1 = 0.6004$ et $g_2 = 0.8001$, avec des variances de 7.1×10^{-5} et 3.7×10^{-5}, respectivement.

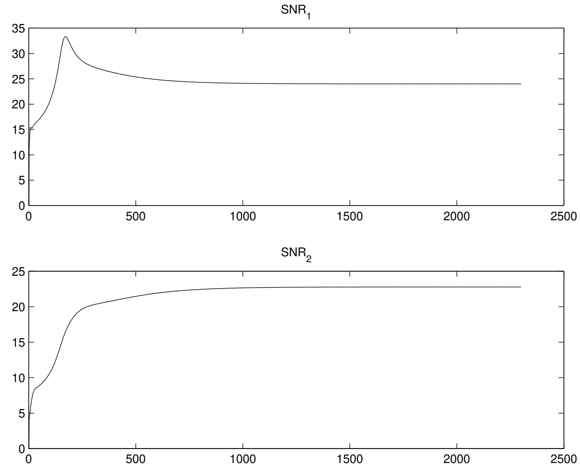

FIG. 2 – Les RSBs versus le nombre d'itération, pour séparer deux sources non-iid.

On voit que l'estimation de filtre est meilleure dans le cas des mélanges mal conditionnés, ce qui peut s'expliquer parce que l'algorithme (18) est alors plus sensible à la déformation introduite par les filtres. Cette hypothèse doit cependant être vérifiée par un calcul de performances.

Le théorème 1 a été énoncé et démontré sous l'hypothèse de sources iid. La robustese de l'algorithme vis-à-vis de cette hypothèse est intéressante à tester. Dans la troisième expérience, nous avons mélangé 100 échantillons de deux sources non-iid : un sinusoïde, et un signal triangulaire, par la matrice :

$$A = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

suivie des filtres de la première expérience.

Pour évaluer la qualité de la séparation, on définie le rapport du signal au bruit (RSB), par l'/'quation (en supposant qu'il n'y a pas de permutation) :

$$\text{RSB}_i = 10 \log_{10} \frac{E\left\{y_i^2(n)\right\}}{E\left\{\left.y_i^2(n)\right|_{s_i(n)=0}\right\}} \quad (23)$$

où $\left.y_i(n)\right|_{s_i(n)=0}$ s'indique la i-ème sortie quand la i-ème source s'annule.

On a choisi les pas d'adaptation $\mu_1 = 0.1$ et $\mu_2 = 0.01$. Figure 2 s'indique les courbe des RSBs versus le nombre d'itération. Cette expérience montre que l'algorithme reste robuste pour les signaux non-iid.

6. CONCLUSION

Nous avons présenté un algorithme de séparation dans le cas de mélanges convolutifs particuliers. Pour ces mélanges, l'indépendance instantané des sorties permet de restituer les sources un filtre près. Actuellement, nous étendons cette étude aux mélanges post-non linéaires à mémoire, obtenus en insérant une non-linéarité (modélisant la distortion de chaque capteur) entre la matrice de mélange et chaque filtre.

A. ANNEXE

Preuve du théorème 1 : En notant la matrice de séparation $B = \begin{pmatrix} 1 & b_1 \\ b_2 & 1 \end{pmatrix}$, on a :

$$\begin{aligned} y_1(n) = & \sum_{k=-\infty}^{\infty} \{h_1(k) + a_2 b_1 h_2(k)\} \, s_1(n-k) \\ & + \sum_{k=-\infty}^{\infty} \{a_1 h_1(k) + b_1 h_2(k)\} \, s_2(n-k) \end{aligned} \tag{24}$$

$$\begin{aligned} y_2(n) = & \sum_{k=-\infty}^{\infty} \{b_2 h_1(k) + a_2 h_2(k)\} \, s_1(n-k) \\ & + \sum_{k=-\infty}^{\infty} \{a_1 b_2 h_1(k) + h_2(k)\} \, s_2(n-k) \end{aligned} \tag{25}$$

Supposons que $y_1(n)$ et $y_2(n)$ sont indépendants. En appliquant le théorème de Darmois [2] aux variables aléatoires $s_1(n-k)$ et $s_2(n-k)$, on doit avoir, $\forall k$:

$$\left\{ \begin{array}{rcl} b_2 h_1(k) + a_2 h_2(k) & = & 0 \\ a_1 h_1(k) + b_1 h_2(k) & = & 0 \end{array} \right. \tag{26}$$

ou :

$$\left\{ \begin{array}{rcl} h_1(k) + a_2 b_1 h_2(k) & = & 0 \\ a_1 b_2 h_1(k) + h_2(k) & = & 0 \end{array} \right. \tag{27}$$

Supposons maintenant $\exists m$ tel que $h_1(m)h_2(m) \neq 0$, en écrivant (26) pour $k = m$, on déduit :

$$b_1 = -a_1 \frac{h_1(m)}{h_2(m)} \tag{28}$$

$$b_2 = -a_2 \frac{h_2(m)}{h_1(m)} \tag{29}$$

Posons $c = \frac{h_1(m)}{h_2(m)}$. Pour $k \neq m$, on trouve $h_1(k) = ch_2(k)$ en reportant dans (26), ou le système suivant en reportant dans l'équation (27) :

$$\begin{pmatrix} h_2(m) & -a_1 a_2 h_1(m) \\ -a_1 a_2 h_2(m) & h_1(m) \end{pmatrix} \begin{pmatrix} h_1(k) \\ h_2(k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{30}$$

Le déterminant de ce système est :

$$\left(1 - a_1^2 a_2^2\right) h_1(m) h_2(m) \tag{31}$$

est ne peut pas s'annuler selon les hypothèses. Les solutions uniques de (30) sont donc $h_1(k) = h_2(k) = 0$, et sont aussi solutions de (26). Ainsi, (26) est satisfaite pour tout k et on a $h_1(k) = ch_2(k)$. Les sorties $y_1(n)$ et $y_2(n)$ sont alors :

$$y_1(n) = \sum_{k=-\infty}^{\infty} \{h_1(k) + a_2 b_1 h_2(k)\} \, s_1(n-k)$$

et :

$$y_2(n) = \sum_{k=-\infty}^{\infty} \{a_1 b_2 h_1(k) + h_2(k)\} \, s_2(n-k)$$

et l'on voit aisément que $y_1(n)$ and $y_2(n-m)$ sont indépendants, $\forall m$. De la même manière, ce raisonnement peut s'appliquer à l'équation (27). □

D'après ce résultat, on peut déduire le théorème 2. En effet, si $a_1 \neq 0$ ou $a_2 \neq 0$, alors $\forall k$, $h_1(k) = ch_2(k)$, où c est une constante réelle. Dans le cas contraire ($a_1 = 0$ et $a_2 = 0$), c'est-à-dire pour une matrice A identité, l'indépendance peut être préservée avec des filtres quelconques. □

RÉFÉRENCES

[1] M. Babaie-Zadeh, C. Jutten, and K. Nayebi. Separating convolutive mixtures by mutual information minimization. In *Proceedings of IWANN'2001*, Granada, Spain, Juin 2001.

[2] P. Comon. Independent component analysis, a new concept ? *Signal Processing*, 36(3) :287–314, 1994.

[3] W. Härdle. *Smoothing techniques with implementation in S.* Springer-Verlag, 1991.

[4] C. Simon. *Séparation aveugle des sources en mélange convolutif.* PhD thesis, l'université de Marne la Vallée, Novembre 1999. (In French).

[5] H.L. Nguyen Thi and C. Jutten. Blind sources separation for convolutive mixtures. *Signal Processing*, 45 :209–229, 1995.

[6] D. Yellin and E. Weinstein. Criteria for multichannel signal separation. *IEEE Trans. Signal Processing*, pages 2158–2168, August 1994.