

Application de la factorisation non-négative des matrices (NMF) pour le débruitage des signaux phonocardiographiques

NAFISSA DIA^{1,2}, JULIE FONTECAVE-JALLON¹, PIERRE-YVES GUMERY¹, BERTRAND RIVET²

¹ Laboratoire TIMC-IMAG, Université Grenoble Alpes, Faculté de Médecine, 38706 La Tronche cedex, France

² Laboratoire GIPSA-Lab, UMR 5216, 11 Rue des Mathématiques, 38400 Saint-Martin-d'Hères

¹ prenom.nom@univ-grenoble-alpes.fr

² bertrand.rivet@gipsa-lab.grenoble-inp.fr

Résumé - Les signaux phonocardiographiques (PCG) sont d'intérêt pour l'analyse du fonctionnement mécanique du cœur. Cependant, ils sont difficilement exploitables directement, dû aux interférences et aux bruits ambiants. Dans cet article, la factorisation non-négative de matrices (NMF) est considérée pour la décomposition du signal PCG bruité. Nous proposons alors d'identifier les composantes associées au signal utile grâce à la multimodalité en exploitant le signal ECG synchrone. Cette approche est appliquée à une base de données établie à partir de signaux PCG réels, artificiellement bruités, utilisée pendant la campagne d'évaluation SiSEC 2016. Les performances sont supérieures à ce qui a été obtenu avec la décomposition modale empirique proposée par les participants et sont comparables pour la sélection multimodale considérée et celle effectuée manuellement par un expert.

Abstract - The phonocardiographic signals (PCG) are of interest for the analysis of the cardiac mechanical function. However, they are not always directly exploitable because of ambient interference. In this paper, non-negative matrix factorization (NMF) is considered for the decomposition of the noisy PCG signal. We also propose to identify the components associated to the signal thanks to multimodality by using the synchronous ECG signal. This approach is applied to a database consisting of real and artificially noisy PCG signals, used during the SiSEC 2016 evaluation campaign. Performance is superior to what was obtained with empirical modal decomposition proposed by the participants and is comparable for the multimodal selection considered and the manual selection by an expert.

1 Introduction

Les bruits cardiaques sont des signaux non-stationnaires enregistrés et connus sous le nom de signaux phonocardiogrammes (PCG). Parmi ces bruits, deux sons sont particulièrement audibles : S1 (qui correspond à la fermeture des valves auriculo-ventriculaires et marque le début de la systole ventriculaire) et S2 (qui correspond à la fermeture des valves aortique et pulmonaire et marque le début de la diastole ventriculaire). L'analyse de ces sons permet de donner des informations sur le fonctionnement mécanique du cœur. Cependant, les signaux PCG bruts (Fig.2) sont difficilement exploitables dû aux interférences et aux bruits ambiants (voix, toux, bruit gastrique, etc.). Il est donc nécessaire de débruiter ces signaux pour pouvoir les interpréter.

Diverses méthodes de débruitage du PCG ont été proposées, se basant sur les principes du filtrage adaptatif [1], du filtrage de Kalman [2], des ondelettes [3] ou de la décomposition modale empirique (EMD) [4]. Dans [1], les auteurs proposent une application du filtrage ALE (Adaptive Line Enhancement) modifié aux signaux PCG bruités. Ce filtrage est basé sur un algorithme des moindres carrés récursif. Les auteurs de [2] cherchent à débruiter un signal PCG en considérant les sons respiratoires comme interférences. Un filtre de Kalman à Ordre Réduit (ou ROKF pour "Reduced-Order Kalman Filter") fondé sur un processus autorégressif de second ordre est utilisé. Une décomposition en ondelettes est mise en œuvre dans [3] : un seuil sur les coefficients de la transformée

permet de séparer le signal utile du bruit. Une nouvelle approche empirique non linéaire EDA (Empirical Denoising Algorithm) basée sur l'EMD est proposée en [4] et permet le débruitage des signaux PCG sans altérer la position des sons. Si ces méthodes ont montré leur intérêt dans le débruitage des PCG, il n'en demeure pas moins qu'elles n'exploitent pas, ou seulement implicitement, une propriété importante des PCG : leur quasi-périodicité.

Dans cet article, nous proposons une méthode de débruitage du PCG couplant la factorisation non-négative de matrices (NMF) [5,12] et la multimodalité. La NMF est une méthode de décomposition d'un signal en composantes. Elle approxime une matrice $A(m \times n)$ à coefficients positifs ou nuls par un produit de deux matrices à valeurs non négatives $W(m \times k)$ et $H(k \times n)$ avec k le nombre de composantes estimées: $A \approx WH$. Sans contrainte supplémentaire, la NMF n'est pas unique, mais sous certaines conditions, elle peut le devenir (par exemple avec une contrainte de parcimonie) [6]. Pour des signaux physiologiques, il est d'intérêt d'appliquer la NMF sur des représentations temps-fréquence [7]. L'algorithme NMF appliqué au spectrogramme est en effet particulièrement bien adapté pour identifier des événements ayant un spectre particulier avec une certaine régularité temporelle comme les signaux quasi-périodiques. Dans ce cas la matrice W est constituée de motifs (fréquentiels) et la matrice H de signaux d'activation (temporels). De plus, il est maintenant classique d'enregistrer des signaux physiologiques avec plusieurs types de capteurs

conduisant à des enregistrements multimodaux. La redondance, mais également la complémentarité d'informations entre les modalités, sont mises à profit dans le traitement et/ou l'interprétation des phénomènes [8]. Dans notre étude, la modalité PCG sera couplée à l'électrocardiogramme (ECG).

Dans la suite de cet article, le paragraphe 2 introduit les signaux, la décomposition utilisée et le couplage multimodal proposé ainsi que les critères d'évaluation. Les résultats sont présentés en 3 et sont suivis des conclusions et perspectives en 4.

2 Matériel et méthodes

2.1 Signaux d'étude

Les signaux utilisés ont initialement été proposés à la campagne SiSEC 2016 (*Signal Separation Evaluation Campaign*) [9]. Des signaux PCG ont été enregistrés avec un microphone cardiaque (MLT201) mis sur la peau en face du cœur et maintenu par une ceinture thoracique. Des signaux ECG ont simultanément été acquis (amplificateur BioAmp) de manière synchrone (PowerLab, ADInstruments). Les acquisitions ont été réalisées sur 3 volontaires sains au laboratoire TIMC-IMAG, dans le cadre du protocole MAPO-RCVQ (promoteur CHU de Grenoble). Les signaux PCG, échantillonnés à 1KHz ont été filtrés passe-bande entre 15 et 300 Hz. Seize échantillons, d'une durée variant d'une dizaine de secondes à plus d'une minute, ont ensuite été artificiellement créés à partir de ces PCG et de différents signaux réels d'interférence (radio, toux, bruits pseudo-périodiques type respiration, ...), enregistrés séparément. La base de données est composée au final de 16 échantillons, avec pour chacun (Fig.2) : le signal PCG filtré d'origine $s(t)$, le signal d'interférences $b(t)$, le signal PCG bruité ($x(t)=s(t)+b(t)$) et le signal ECG synchrone $ecg(t)$.

2.2 Algorithmes de débruitage

La méthodologie proposée est schématisée Fig.1 et implémentée sous Matlab®. L'analyse et l'étude des signaux physiologiques non stationnaires se font fréquemment sur la base de représentations temps-fréquence. Soit $X(t, f)$ le spectrogramme de $x(t)$ à l'instant t et la fréquence f . La NMF est alors appliquée sur la matrice $A = |X(t, f)|^2$ de taille $m \times n$, où m est le nombre de fréquences et n le nombre de fenêtres temporelles de la transformée de Fourier à court terme (TFCT). La durée de la fenêtre est fixée à 64 ms, une durée inférieure à la durée moyenne d'un son S1 (100-150 ms) et d'un son S2 (70-140 ms) [10] de façon à avoir une bonne localisation temporelle des

événements. Le nombre de fréquences de la transformée de Fourier discrète est choisi égal à 512. Dans le cas de cette étude, nous avons fixé à $k=12$ le nombre de composantes à identifier et l'algorithme utilisé pour la NMF est l'algorithme des moindres carrés alternés.

Une fois obtenue la décomposition de la matrice A , l'étape suivante consiste à sélectionner, parmi les k composantes, celles correspondant au signal utile PCG, notées $W_s(f)$ et $H_s(t)$. Les autres composantes sont alors associées au bruit et notées $W_b(f)$ et $H_b(t)$. Pour éviter de faire cette sélection manuellement par un expert, nous proposons dans cette étude une sélection multimodale automatique en exploitant l'ECG qui a été enregistré de façon synchrone. Les composantes pour lesquelles l'intercorrélation entre les signaux d'activation du PCG bruité et le signal d'activation $H_{ecg}(t)$, issu de la NMF (à 1 composante) appliquée au spectrogramme de l'ECG, est supérieure au seuil λ sont identifiées comme étant liée au PCG, les autres sont alors considérées comme représentant le bruit. Il est à noter que si le seuil n'est atteint pour aucune des composantes, une seule composante est sélectionnée, celle qui présente la plus grande valeur d'intercorrélation. A partir de cette sélection, les densités spectrales de puissance du signal et du bruit sont calculées, $S_s(t, f) = W_s(f) \cdot H_s(t)$ et $S_b(t, f) = W_b(f) \cdot H_b(t)$.

La reconstruction se fait ensuite en estimant $\hat{S}(t, f)$ la TFCT du signal $s(t)$, à partir du filtrage de $X(t, f)$, pour obtenir $\hat{s}(t)$. Deux filtres, classiquement utilisés dans la littérature, sont ici étudiés, le filtre de Wiener :

$$H_{wiener} = \frac{S_s(t, f)}{S_s(t, f) + S_b(t, f)} \quad (1)$$

ou le filtre noté $H_{dsp}(t, f)$ défini par :

$$H_{dsp} = \frac{S_s(t, f)^{\frac{1}{2}}}{|X(t, f)|} \quad (2)$$

Ce second filtre correspond à estimer $\hat{S}(t, f)$ comme étant $\hat{S}(t, f) = S_s(t, f)^{\frac{1}{2}} \cdot e^{j\varphi_x(t, f)}$ avec $\varphi_x(t, f)$, la phase de $X(t, f)$.

Plusieurs méthodes seront comparées : la méthode M_{manu} basée sur la sélection manuelle des composantes de la NMF et la méthode M_{auto} basée sur leur sélection automatique grâce à l'ECG. Chacune de ces méthodes sera scindée en deux en fonction du filtre de reconstruction mis en œuvre : M^{wiener} utilisant le filtre de Wiener (1) et M^{dsp} utilisant le filtre simple DSP (2).

La figure 2 illustre les signaux à disposition : un signal ECG et un signal PCG enregistré dans des conditions optimales, tous 2 synchrones, ce signal PCG bruité artificiellement et le signal estimé à partir de ce dernier avec la méthode M_{auto}^{wiener} (seuil $\lambda=0,75$).

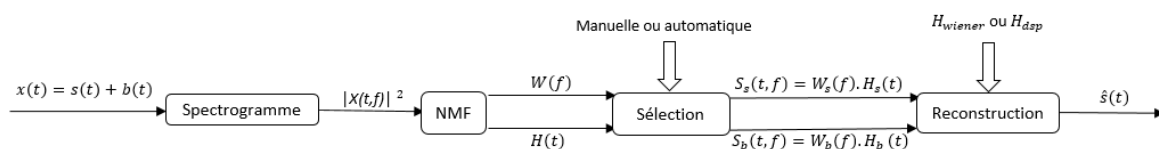


Figure 1 : Schéma-bloc de l'algorithme de débruitage (et variantes possibles)

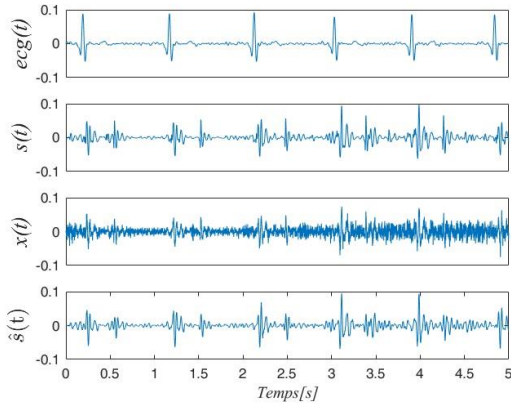


Figure 2 : Signaux ECG $ecg(t)$, PCG brut $s(t)$, PCG bruité $x(t)$ et PCG débruité $\hat{s}(t)$ (méthode M_{auto}^{wiener} , seuil =0,75)

2.3 Méthodes d'évaluation

La qualité de l'estimation du PCG $\hat{s}(t)$ est évaluée à l'aide de la *BSS Eval Toolbox* [11], qui permet la mesure de différents paramètres :

- le SDR (Signal to Distortion Ratio) estime globalement le RSB (Rapport Signal sur Bruit);
- le SIR (Signal to Interference Ratio) quantifie les interférences dues aux autres sources dans celle estimée;
- le SAR (Signal to Artifacts Ratio) quantifie les artéfacts résiduels dans l'estimation du signal et mesure la linéarité de la méthode d'extraction.

De plus, le gain de performance entre le PCG estimé $\hat{s}(t)$ et le signal bruité enregistré $x(t)$ est défini comme la différence entre le critère calculé sur l'estimation et celui calculé sur le signal bruité, tous deux exprimés en dB.

La distribution des paramètres est représentée par des boîtes à moustaches pour lesquelles la ligne rouge correspond à la médiane, les extrémités des boîtes correspondent aux 25^{ème} et 75^{ème} percentiles et les moustaches mettent en évidence les valeurs extrêmes.

3 Résultats

Cette partie présente les résultats obtenus par la méthode proposée pour l'estimation du PCG et sont comparés avec les résultats obtenus lors de la campagne d'évaluation SiSEC 2016 [9]. Dans un premier temps, l'influence du seuil utilisé pour la sélection automatique des composantes est étudiée (3.1), puis l'influence du choix du filtre de reconstruction est étudiée (3.2). Enfin, la méthode proposée utilisant la NMF est comparée aux deux extractions proposées lors de l'évaluation SiSEC 2016 (3.3).

3.1 Choix du seuil

La figure 3 montre la distribution des gains de SDR pour différentes valeurs du seuil appliqué pour la sélection automatique des composantes relatives au PCG. Chaque boîte est obtenue pour 20 tests réalisés pour chacun des 16 échantillons. Pour chaque test, l'initialisation de l'algorithme de NMF est faite de façon aléatoire et différente. La reconstruction est réalisée avec le filtre de Wiener. La figure 3 montre que le choix du seuil influe sur le gain du SDR : le gain médian augmente

de 1,7dB ($\lambda=0,2$) à 9,5dB ($\lambda=0,75$) puis diminue à 7,5dB ($\lambda=0,9$). Il est intéressant de noter que, comme attendu, plus la valeur du seuil est élevée et moins il y a de composantes qui sont détectées comme contenant du PCG. Les résultats (t-test à $\alpha=0,01$) montrent cependant un gain de SDR maximum pour un seuil à 0,75 qui sera donc la valeur choisie dans la suite de cette étude.

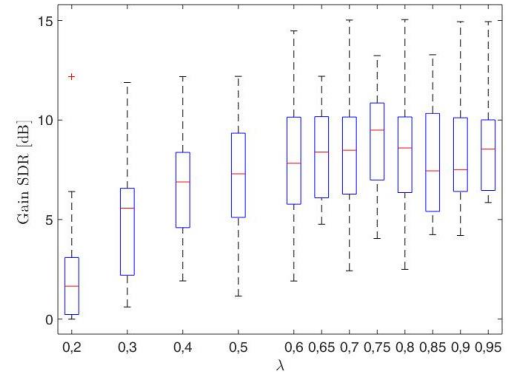


Figure 3 : Gains en SDR pour différentes valeurs de seuil d'intercorrélation λ

3.2 Influence du filtre de reconstruction

L'influence du filtre de reconstruction est maintenant analysée, par le gain de SDR, sur les 2 types de sélection. Les résultats sont présentés sur la figure 4. Comme au 3.1, chaque boîte est obtenue pour chacun des 16 échantillons et plusieurs tests sont réalisés (initialisation de la NMF différente à chaque test), 20 tests pour la sélection automatique et 5 pour la sélection manuelle.

Que ce soit pour la sélection automatique ou pour la sélection manuelle, on observe des performances légèrement meilleures pour la reconstruction effectuée avec le filtre de Wiener. Le gain médian en SDR est de 8dB pour M_{manu}^{wiener} contre 7,8dB pour M_{manu}^{dsp} et de 9,3dB pour M_{auto}^{wiener} contre 8,1dB pour M_{auto}^{dsp} . Par la suite, c'est donc le filtre de Wiener qui sera considéré.

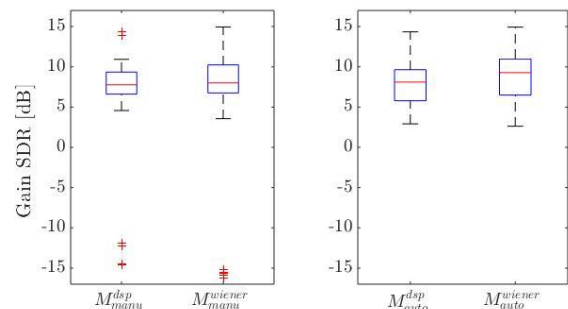


Figure 4 : Gains en SDR selon le filtre de reconstruction et le type de sélection des composantes du signal. La sélection automatique est réalisé avec un seuil $\lambda=0,75$.

3.3 Comparaison de méthodes

Dans ce paragraphe, l'approche multimodale proposée est comparée aux résultats obtenus lors de la campagne SiSEC 2016 [9] où 2 approches de type décomposition modale empirique étaient proposées. La figure 5 compare les performances de débruitage de ces 2 méthodes avec celles obtenues par l'approche proposée fondée sur la NMF avec sélection manuelle ou automatique des composantes du PCG. Il est à noter que les performances

des 2 méthodes EMD n'ont été analysées dans [9] que sur 15 échantillons au lieu des 16 (absence d'un échantillon pour un candidat). D'autre part, pour SiSEC, les signaux PCG avant bruitage et ECG n'étaient pas disponibles.

L'application de la NMF avec sélection manuelle M_{manu}^{wiener} , comparable en terme de signaux exploités à ce qui a pu être fait par les participants, est plus performante que l'utilisation des méthodes EMD. Le gain médian en SDR est de 8dB pour M_{manu}^{wiener} contre 1,5 et 5,7dB pour les 2 propositions EMD. Les gains médians en SIR sont comparables mais on observe moins de dispersion pour M_{manu}^{wiener} . Quant au SAR, il est de 10,4dB pour M_{manu}^{wiener} contre 0,85 et 5,5dB avec l'EMD montrant ainsi une estimation de bien meilleure qualité.

La sélection automatique des composantes du signal, grâce à l'ECG, M_{auto}^{wiener} donne des valeurs de performances équivalentes à la sélection manuelle par un expert et même légèrement meilleure pour le SAR, ce qui confirme l'intérêt de l'approche multimodale proposée.

Ainsi, les méthodes rejettent toutes de façon comparable les sources concurrentes de bruit (SIR identique entre méthodes) mais les méthodes fondées sur la NMF permettent *in fine* une meilleure estimation du PCG par rapport à l'EMD (SDR et SAR plus élevés).

4 Conclusions et perspectives

Dans cette étude, nous avons proposé une méthode de débruitage multimodale du PCG par l'ECG fondée sur la NMF. L'identification des composantes obtenues par NMF utilise l'ECG enregistré simultanément au PCG. Les résultats montrent tout l'intérêt de cette approche. En effet, l'approche multimodale NMF surpasse les approches fondées sur l'EMD et donne au moins les mêmes résultats qu'une sélection manuelle faite par un expert. Pour cela, il est nécessaire d'enregistrer en plus du PCG un ECG synchrone. Si cela peut représenter une contrainte, cela reste d'un grand intérêt physiologique puisque l'ECG fournit des informations complémentaires au PCG (électriques et mécaniques).

En perspective à cette étude multimodale préliminaire, il pourra être intéressant de proposer une NMF multimodale telle que proposée dans [12], mais il faudra pouvoir prendre en compte les variabilités temporelles entre modalités. En effet, à chaque battement cardiaque caractérisé par le complexe QRS sur l'ECG est associée la paire de sons S1 et S2 du PCG décalée d'un temps τ appelé *jitter*, physiologiquement non stationnaire.

5 Remerciements

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) funded by the French program Investissements d'avenir

6 Références

- [1] M.A. Tinati, A. Bouzerdoum, J. Mazumdar, "Modified adaptive line enhancement filter and its application to heart sound noise cancellation", *Proc. Int. Symp. Signal Processing and its Applications*, 2, pp. 815–818, 1996.
- [2] S. Charleston, M.R. Azimi-Sadjadi, "Reduced order Kalman filtering for the enhancement of respiratory sounds", *IEEE Trans. Biomedical Engineering* 44, pp. 1006–1019, Oct 1997.
- [3] S.R. Messer, J. Agzarian, D. Abbott, "Optimal wavelet denoising for phonocardiograms", *Microelectronics Journal* 32, pp. 931–941, Dec 2001.
- [4] O. Beya, E. Fauvet, O. Lalignat « EDA, approche non linéaire de débruitage des signaux cardiaques », *Proc. CORESA*, pp. 166-171, Le Creusot, France, 2013.
- [5] D.D. Lee, H.S. Seung, "Algorithms for Non-negative Matrix Factorization", *Advances in Neural Information Processing Systems* 13, pp. 556-562, 2000.
- [6] A. Cichocki, R. Zdunek, A. Phan, S. Amari, "Nonnegative Matrix and Tensor Factorizations", *Wiley*, pp.24-25
- [7] M. Niegowski and M. Zivanovic, "ECG-EMG Separation by Using Enhanced Non-Negative Matrix Factorization", *Proc. Int. Conf. Eng. Medicine & Biology Soc. (EMBC)*, pp. 4212-4215, Chicago, USA, 2014.
- [8] D. Lahat, T. Adalı, C. Jutten, "Multimodal Data Fusion: An Overview of Methods, Challenges and Prospects," *Proceedings of the IEEE* 103(9), pp.1449-1477, 2015.
- [9] A. Liutkus, F. Stöter, Z. Rafii, D. Kitamura, B. Rivet, N. Ito, N. Ono and J. Fontcave, "The 2016 Signal Separation Evaluation Campaign", *Proc. LVA/ICA*, pp.323-332, Grenoble, France, 2017.
- [10] A. Moukadem, "Segmentation et Classification des signaux non-stationnaires". Thèse de doctorat. Université de Nice Sophia Antipolis, 2012.
- [11] E. Vincent, R. Gribonval, and C. Févotte, "Performance measurement in blind audio source separation," *IEEE Trans. Audio, Speech and Language Proc.*, 14(4), pp. 1462-1469, 2006.
- [12] F. Sedighin, M. Babaie-Zadeh, B. Rivet, and C. Jutten, "Multimodal Soft Nonnegative Matrix Co-Factorization for Convolutional Source Separation," *IEEE Trans. Signal Processing*, 2017 (accepted for publication).

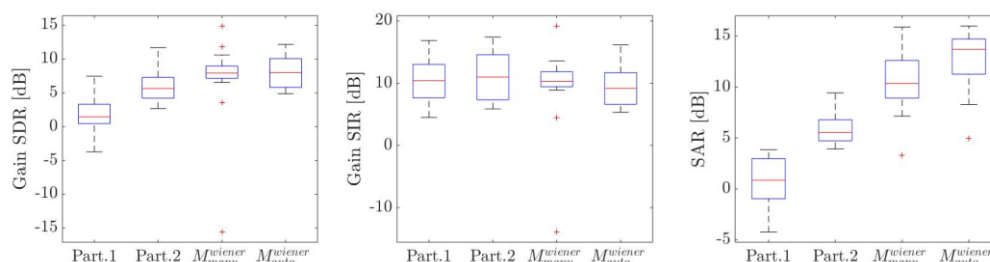


Figure 5 : Gains en SDR et SIR, et SAR pour les 2 méthodes de type EMD proposées par les participants 1 et 2 de [9] et pour nos méthodes basées sur la NMF avec sélection manuelle ou automatique (grâce à l'ECG) des composantes correspondant au signal PCG. Pour ces 2 dernières, l'estimation du signal PCG débruité est faite au travers du filtre de reconstruction de Wiener. Pour la méthode automatique, le seuil d'intercorrélacion est fixé à 0,75.