Denoising Phonocardiogram signals with Non-negative Matrix Factorization informed by synchronous Electrocardiogram

Nafissa DIA[1,2], Julie Fontecave-Jallon[1], Pierre-Yves Gumery[1], Bertrand Rivet[2]
[1]Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP*, TIMC-IMAG, F-38000 Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, F-38000 Grenoble, France
*Institute of Engineering Univ. Grenoble Alpes
[1]surname.name@univ-grenoble-alpes.fr [2]bertrand.rivet@gipsa-lab.grenoble-inp.fr

***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 (gastric noises, breathing noises, etc.). We aim to denoise PCG signals using another cardiac modality, the electrocardiographic (ECG) signal. In this paper, we investigate an informed non-negative matrix factorization to extract signal components out of the noisy PCG signal, considering synchronous ECG information. Our approach is applied and evaluated on a database consisting of real and artificially noisy PCG signals.**

I. Introduction

Cardiac sounds are non-stationary signals recorded and known as phonocardiogram (PCG) signals. Among these sounds, two are particularly audible (Fig. 1): S_1 (which corresponds to the closure of the atrial-ventricular valves and marks the beginning of the ventricular systole) and S_2 (which corresponds to the closure of aortic and pulmonary valves and marks the onset of the ventricular diastole). The analysis of these types of sounds brings information on mechanical function of the heart. However, the raw signals are roughly exploitable because of ambient interference (voice, cough, gastric noise, etc.). It is therefore necessary to remove noise components before interpretation.

Several methods of PCG denoising have been proposed based on principles of adaptive filtering [1], Kalman filtering [2], wavelets [3], or Empirical Modal Decomposition (EMD) [4]. In [1], authors have proposed an application of modified Adaptive Line Enhancement filtering to noisy PCG. This filtering is based on a recursive least squares algorithm. Authors of [2] have looked for removing noise from a PCG signal by considering respiratory sounds as interferences. A Reduced-Order Kalman Filter based on a second order auto-regressive process is used. A decomposition in wavelets has been implemented in [3]: a threshold on transform coefficients allows to separate the useful signal from noise. A new non-linear Empirical Denoising Algorithm approach based on EMD has been proposed in [4] and allows denoising PCG signals without changing the sounds position. Although these methods have shown their interest, they do not exploit, or a little, the important propriety of quasi-periodicity of PCG signals.

In this paper, in order to denoise PCG, we aim to exploit the non-negative matrix factorization (NMF) [5][6] and some derivatives [7][8] in a multimodal context coupling PCG and the electrocardiogram (ECG) sensor (Fig. 1), giving access to the electrical cardiac activity.

It is now common to record physiological signals with different types of sensors leading to multimodal recordings [9][10][11]. Information redundancy and complementarity between modalities can be used for processing and/or interpretation of physiological behaviors. In cardiac analysis, ECG and PCG signals are quasi-periodic and quasi-synchronous. There is always a delay named RS_1, which varies from one heartbeat to another one, between the occurrence of the R peak, corresponding to the positive wave of the ventricular depolarization complex (QRS) of the ECG signal, and the first sound S_1 (Fig. 1). This delay and the associate jitter have to be considered for joint processing of ECG and PCG.

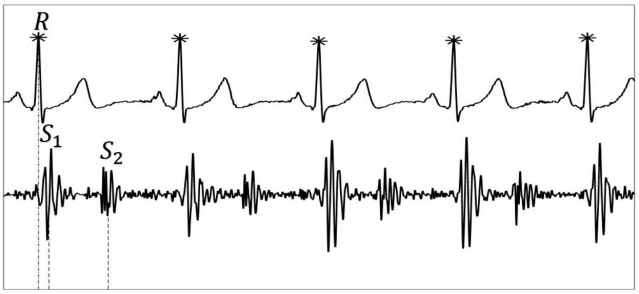


Fig. 1. Synchronous ECG and PCG signals. PCG is a quasi-periodic succession of two bumps S_1 and S_2, following the R peak of ECG signal.

The NMF method decomposes a signal in components. It approximates an observation matrix $X \in \mathbb{R}^{m\times n}$ of positive or zero coefficients by a product of two matrices with non-negative values $W \in \mathbb{R}^{m\times k}$ and $H \in \mathbb{R}^{k\times n}$ with k the number of estimated components: $X = V + N$ where $V = WH$ and N is the residual error from the approximation. With no more constraint, the NMF is not necessarily unique, but under some conditions, it may become unique (for example

ISBN 978-90-827970-1-5

with a sparsity constraint) [12]. For physiological signals, the use of NMF is of interest in time-frequency representations [13]. Indeed, the NMF algorithm applied to the spectrogram is well-suited to identify events with particular spectrum and temporal regularity (quasi-periodicity). In this case, W is the matrix of frequency patterns and H, the matrix of time activations.

Preliminary tests of NMF for PCG denoising have been carried out considering an ECG reference [14]. Satisfactory results were obtained but the multimodal aspect was just exploited as a post-processing for the decomposition achieved by the NMF. Indeed, the ECG modality is only used to select a posteriori the components related to the cardiac signal based on a rythm analysis. It may be of interest to take into account multimodality all along the NMF processing. Hence, we propose now to adapt some existing derivative methodologies [7][8] to explore joint NMF. For one speaker identification during a discussion, [7] proposed to take advantage of both modalities, video and audio. The proposed NMF approach aims to minimize the distance between the activation signals of both modalities. In [8], a general deformation model has been proposed for audio sources separation. This model allows to constraint data according to the processing or interpretation. We aim to investigate these two existing methods and adapt the algorithms in order to propose a NMF solution for PCG decomposition, informed by ECG reference. The proposed method consists then of denoising PCG, using Wiener filtering as in [14]. Evaluation is carried out on a real signals database previously proposed to *Signal Separation Evaluation Campaign* in 2016 (SiSEC 2016) [15].

II. Signals database

Acquisitions were performed on three healthy volunteers at TIMC-IMAG laboratory as part of $MAPO-RCVQ$ protocol (CHU of Grenoble promoter). PCG were recorded with a cardiac microphone ($MLT201$) put on the skin in front of the heart and maintained by a thoracic belt. $D2$ lead ECG signals were simultaneously conditioned with a BioAmp amplifier. All signals were synchronously acquired with a PowerLab data acquisition system (ADInstruments) and sampled at $1kHz$. PCG signals were band-pass filtered between 15 and $300Hz$. Sixteen samples with a duration varying from 10 to 70 $seconds$ have been artificially created from the filtered PCG and different real interference signals (radio, cough, pseudo-periodic noises of breath type, etc.), recorded separately. The database is composed of sixteen samples, each composed of the filtered PCG $s(t)$, the interference signal $n(t)$, the noisy PCG $x(t) = s(t) + n(t)$ and the synchronous ECG $ecg(t)$.

III. Proposed methodology

In this section, before presenting our work related to our application in Section III-B, the basics of NMF are recalled (Section III-A). Finally, the proposed informed NMF is introduced in Section III-C.

A. NMF based on beta divergences

The derivative NMF [7][8] are based on computation of cost functions better known as β-divergences [16]. Such functions quantify the dissimilarity or divergence between the observation (the spectrogram in our case) X and its approximated decomposition $V = WH$ under the constraints of positivity, leading to the following cost function

$$C(W, H) = D_\beta(X \mid V), \tag{1}$$

D_β is the $\beta-$divergence.

To identify the components W and H, multiplicative updates are considered. To ensure the decrease of the β-divergence, an alternative algorithm based on Minimization-Maximization (MM) is applied [16]. In what follows, β is chosen equal to 2, *i.e.* the β-divergence is simply the Euclidian distance. This leads to estimate W knowing H as :

$$W \leftarrow W \otimes \frac{[(WH) \otimes V]H^T}{[(WH)]H^T}, \tag{2}$$

and H knowing W as :

$$H \leftarrow H \otimes \frac{W^T[(WH) \otimes V]}{W^T[(WH)]}, \tag{3}$$

where $\otimes$ and the division are the element-wise multiplication and division, respectively and ^{T} is the transpose operator.

The spectrogram of the noisy PCG $x(t)$ is modeled as $X = WH + N = W_1H_1 + W_2H_2 + N$, where W_1H_1 (*resp.* W_2H_2) corresponds to the signal part, *i.e.* essentially clean PCG, (*resp.* the noise part, *i.e.* composed of interferences).

Compared to a classical NMF for which W and H are first estimated by (2) and (3), respectively and then sorted to identify W_1, W_2 and H_1, H_2 based on prior knowledge, the proposed methodology aims at using another modality (here the ECG) as reference to inform the NMF so that W_1, W_2 and H_1, H_2 are directly identified. To this end, some constraints on H_1 based on a NMF of the ECG signal are used since the PCG signal $s(t)$ and the ECG one share the same origin: the heart beats.

In our approach, the ECG components out of NMF will not be modified but used as reference for updating the components estimated from the noisy PCG. We will thus speak of "informed" NMF instead of joint-NMF.

B. Related work: penalised NMF

The spectrogram of the ECG reference is modeled as : $X_{ref} = W_{ref}H_{ref} + N_{ref}$. The idea in this section, as in [7], is to constraint the activation signals H_1 to be like the H_{ref} ones. Indeed, the ECG and PCG share the same origin so that activation patterns H_{ref} and H_1 should also share the same behavior.

Following [7], one can optimize the following cost function

$$C_1(W_1, H_1, W_2, H_2) = D_2(X \mid W_1H_1 + W_2H_2) + \delta\, P(W_{ref}, H_{ref}, W_1, H_1), \tag{4}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

where P is the penalization function defined by

$$P = \| \Lambda_{ref} H_{ref} - S\Lambda_1 H_1 \|_F^2, \quad (5)$$

with Λ_{ref} and Λ_1 two diagonal matrices whose diagonal elements are the sum of each column of W_{ref} and W_1. S is a diagonal matrix for scaling H_1 in order to compare it with H_{ref}. W_{ref}, H_{ref} and initial values of W_1, H_1, W_2 and H_2 are given by the classic NMF [5] on the ECG and the noisy PCG spectrograms. The number of templates for ECG and PCG are mandatorily the same ($k_1 = k_{ref}$). The penalization term P forces the activation signals H_1 to be superimposed on the H_{ref} ones (up to some scaling factors).

However, this approach has several drawbacks for the considered application. First, the physiological delay between events in both modalities ECG and PCG is not considered in this method. Nevertheless, this may be corrected by integrating a mean delay to readjust positions. In addition, an ECG beat is mainly modeled by the R-peak while a PCG beat is mainly modeled by S_1 and S_2 sounds. Consequently, a direct comparison between H_{ref} and H_1 will surely fail. Therefore the shapes of PCG time activations H_1 are not accurate. Finally, there is a lack of flexibility as the approach imposes to work with the same number of templates for both modalities. This constraint is not suitable with the physiological properties of ECG and PCG signals.

C. Proposed Informed NMF with ECG reference and transformation matrix

To bypass these limits, we investigate the proposition described in [8] in our context. Thus, the considered model is the joint equations:

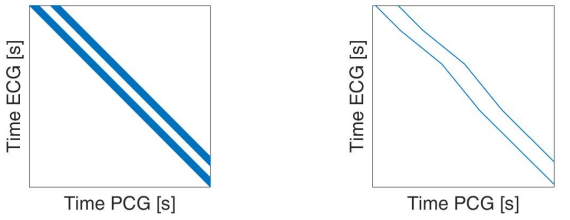
$$X_{ref} = W_{ref} H_{ref} + N_{ref} \quad (6)$$

$$X = W_1 H_{ref} T_1 + W_2 H_2 + N, \quad (7)$$

where T_1 is a transformation matrix that aims at transforming the activation patterns H_{ref} of the ECG into the activation ones $H_1 \simeq H_{ref} T_1$ of the PCG by constraining these latter ones. Indeed, in particular, we intend to consider the succession of two bumps S_1 , S_2 of the PCG in a quasi-periodic manner, as well as the delays between R peaks and S_1 and S_2 sounds (named RS_1 and RS_2 intervals) Fig. 1. To this end, T_1 is defined as a "bidiagonal" and sparse matrix where the first diagonal is centered at RS_1 value and the second diagonal at RS_2 value (Fig. 2). It is worth noting that the "diagonals" should have some widths to take into account the fact that the delays between R peaks and sounds S_1 and S_2 are not constant. W_{ref}, H_{ref} and initial values of W and H are computed by NMF considering k_{ref} and $k = k_1 + k_2$ numbers of templates, with $k_1 = k_{ref}$. The estimation of the components is thus achieved by the minimization of:

$$C_2(W_1, T_1, W_2, H_2) = \| X - W_1 H_{ref} T_1 - W_2 H_2) \|_F^2 \quad (8)$$

Finally, based on MM algorithm, components are iteratively computed by multiplicative updates. We propose two types of structure for the transformation matrix T_1, illustrated in Fig. 2 and described in the following.



(a) T_1 based on literature (b) T_1 based on signals

Fig. 2. Illustration of T_1 structures

1) T_1 based on the literature ($T_{1,lit}$): With this choice, the "diagonals" defining T_1 are chosen based on prior knowledge of average RS_1 and S_1S_2 intervals. According to [17], RS_1 is well modeled by a Gaussian distribution

$$RS_1 \sim \mathcal{N}(\mu_1, \sigma) = \mathcal{N}(70, 20)[ms]. \quad (9)$$

Moreover, in [18], the S_1S_2 interval is estimated to about $300ms$. The combination of these results leads to the following distribution of RS_2

$$RS_2 \sim \mathcal{N}(\mu_2, \sigma) = \mathcal{N}(370, 20)[ms]. \quad (10)$$

The bandwidths $RS_1 \pm \sigma$ and $RS_2 \pm \sigma$ are chosen large enough to take into account the intra- and inter-subject variability; this introduces a too large number of degrees of freedom leading thus to a lack of constraints between H_{ref} and $H_1 \simeq H_{ref} T_1$.

2) T_1 based on the signals themselves ($T_{1,sig}$): With this choice, to overcome the lack of constraints of the prior choice described in the previous paragraph, the two "diagonals" of T_1 will be defined from the signal themselves.

RS_1 and RS_2 intervals are estimated beat by beat for each sample based on two standard NMF [5] with 12 components on noisy PCG and one component on ECG noted H_{ecg}. Let H^* denotes the PCG component that is the most correlated to H_{ecg}. R peaks are first detected on the ECG time activation H_{ecg}, then S_1 and S_2 sounds are detected as local maxima on H^*. The RS_1 and RS_2 values are used to initialize the transformation matrix putting thus more constraints on signal components structure since the width of each diagonal is smaller than in the method described in the previous paragraph.

Finally, the proposed algorithm is summarized in Algorithm 1.

Algorithm 1 Proposed informed NMF

Inputs: H_{ref}, structure of T_1
while stopping criterion not reach **do**
  Update W_1 and W_2 assuming T_1 and H_2 based on (2)
  Update T_1 and H_2 assuming W_1 and W_2 based on (3)
Outputs: W_1, W_2, T_1 and H_2

ISBN 978-90-827970-1-5 © EURASIP 2018

D. PCG Reconstruction with Wiener filtering

Considering the signal and noise components obtained from the proposed algorithm, an estimated PCG reconstruction $\hat{s}(t)$, is provided by applying a Wiener filtering. This classic filter is based on the spectral densities $W_1 H_{ref} T_1$, denoted $P_s(t,f)$, and $W_2 H_2$, denoted $P_n(t,f)$. The impulse response of the filter is :

$$H_{wiener}(t,f) = \frac{P_s(t,f)}{P_s(t,f) + P_n(t,f)}. \quad (11)$$

The estimated short term Fourier transform (STFT) of the estimated PCG $\hat{S}(t,f)$ is thus expressed as

$$\hat{S}(t,f) = H_{wiener}(t,f)\, X(t,f), \quad (12)$$

where $X(t,f)$ is the STFT of the noisy PCG $x(t)$. The denoised PCG $\hat{s}(t)$ is computed as the inverse STFT of $\hat{S}(t,f)$.

IV. Results

In this section, we analyze the performances of our proposed solution by comparison with a standard NMF approach. Evaluation was carried out on the PCG database proposed to SiSEC 2016 [15]. ECG and PCG spectrograms are computed using Hamming window of length 64ms and a shift of one ms, with zero-padding of 512 samples.

A. Evaluation settings

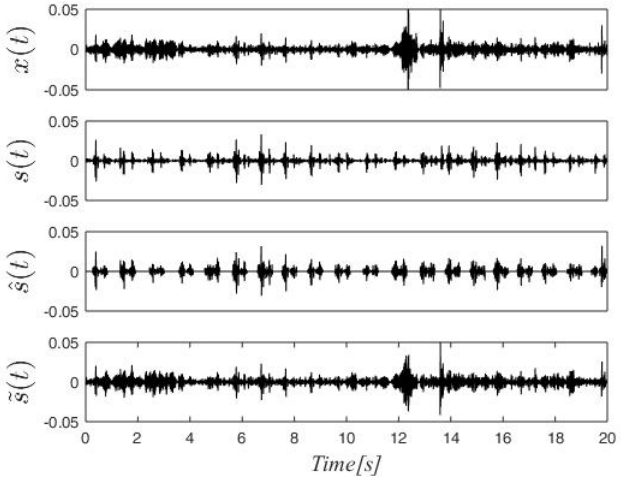
As we focus on two representative physiological events (S_1 and S_2 sounds) in noisy PCG signals, we consider $k_1 = 2$ for the signal components. Moreover, we choose a limited number of noise components ($k_2 = 2$) to force the constraints on the signal time activations. Therefore $k_{ref} = 2$ and $k = 4$.

The results of our informed NMF are compared with a standard NMF [5], using the same number of components. The standard NMF is computed on ECG signal ecg(t) with k_{ref} components and on noisy PCG $x(t)$ with k components. We consider as signal the k_{ref} PCG components whose time activations are the most correlated to the time activation of the first ECG component. The remaining PCG components are considered as noise. The Wiener filtering is applied, as in Section III-D, to reconstruct a PCG denoised estimation, noted $\tilde{s}(t)$.

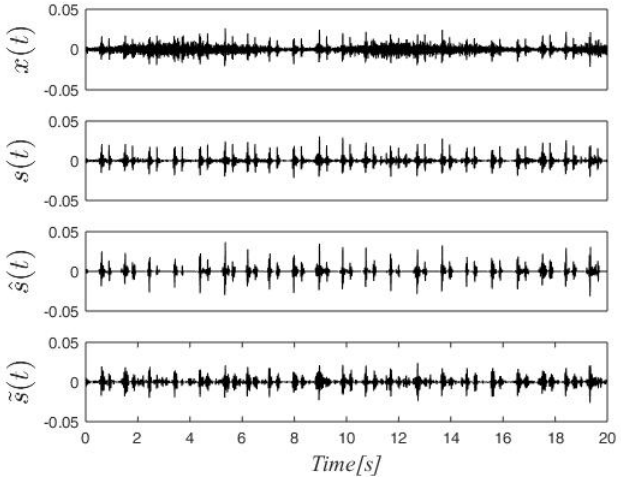
B. Observation of denoised signals

Fig. 3(a) and 3(b) present examples of PCG denoising after NMF and Wiener filtering and compare results of standard NMF and informed NMF for 2 different samples of the database. The T_1 structure is based on the signals themselves ($T_{1,sig}$).

In these examples, informed NMF is able to remove specific noises better than standard NMF. Indeed, in Fig. 3 (a), the impulse noises around $12s$ and $14s$ do not appear anymore in the denoised PCG $\hat{s}(t)$ with informed NMF, which is not the case for $\tilde{s}(t)$ obtained after standard NMF. In Fig. 3 (b), the main noise is stochastic with a quasi-periodic power variation and representative of respiratory interference. Most of these interference are still present in $\tilde{s}(t)$ whereas they have been quite well removed in $\hat{s}(t)$. These two examples show that it is of interest to link the activation profiles H_1 with the ones of the ECG to better estimate the PCG components from the noisy signal.



(a) Sample 1



(b) Sample 6

Fig. 3. Signals examples for samples 1 and 6. $x(t)$: noisy PCG, $s(t)$: original PCG, $\hat{s}(t)$: denoised PCG by informed NMF with $T_{1,sig}$, $\tilde{s}(t)$: denoised PCG by standard NMF

C. Global performance of the proposed methodology

Quality of PCG estimation was quantified by using *BSS Eval Toolbox* [19]. We consider the Signal to Distortion Ratio (SDR) that globally estimates the Signal-to-Noise Ratio (SNR). The difference, expressed in dB, between the SDR computed on the estimation and on the noisy signal represents the performance gain. Boxplots represent the parameters distribution: the line inside the box corresponds to the median value, the boxlimits represent the first and the third quartiles and whiskers highlight extreme values.

The SDR gain is computed on all sixteen samples of the

ISBN 978-90-827970-1-5

SiSEC database for standard and informed NMF, for both T1 structures $T_{1,lit}$ and $T_{1,sig}$.

One can see on Fig. 4 that the overall performances are comparable between the standard NMF and informed NMF with transformation matrix $T_{1,lit}$, and slightly better using the informed NMF with $T_{1,sig}$ than using the standard NMF, as also shown in the previous paragraph. Indeed, the median values of SDR gain are 3.2 dB for standard NMF, 4.1 dB for informed NMF $T_{1,lit}$ and 6.8 dB for informed NMF $T_{1,sig}$.

Moreover, as expected, the best performances are obtained with the T_1 structure driven by the data. Limiting the 'diagonals' bandwidth of T_1 improves the signal and noise components separation and therefore the PCG denoising.

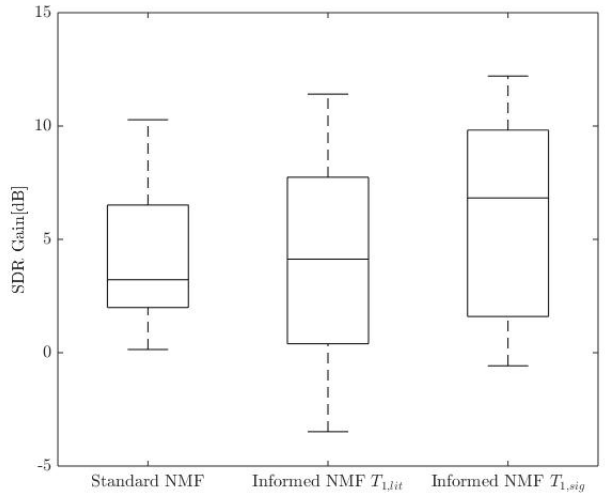


Fig. 4. SDR gain in dB for standard NMF, informed NMF with T_1 based on literature and informed NMF with T_1 based on the signals themselves.

V. Conclusion

In this paper, we have proposed a method of PCG denoising based on NMF informed by constraining the activation profiles of the PCG to be similar to those of the ECG up to a limited transformation. This informed NMF exploits the multimodality better than in our previous work [14]. The results obtained from actual recordings are promising since they allow to remove some specific interference occurring in daily situations.

In perspective, an in-depth analysis of the parameter choices will be investigated. Some issues are noted. In particular it should be possible to constraint each signal component by a different transformation matrix considering thus the specific structure of the different components. Moreover, the quasi-periodic aspect of physiological signals is not taken into account in the informed NMF yet.

Acknowledgment

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) funded by the French program Investissements d'avenir. This work is also partially supported by the French National Research Agency, as part of the SurFAO project (ANR-17-CE19-0012).

References

[1] M. Queensland University of Technology. Signal Processing Research Centre., A. Institute of Electrical and Electronics Engineers., and J. Mazumdar, *Fourth International Symposium on Signal Processing and its Applications : ISSPA 96 : proceedings, main symposium : August 25-30 1996, Royal Pines Resort, Gold Coast, Queensland, Australia*. Queensland University of Technology, January 1996, vol. 1–2.

[2] S. Charleston and M. Azimi-Sadjadi, "Reduced order Kalman filtering for the enhancement of respiratory sounds," *IEEE Trans. on Biomedical Engineering*, vol. 43, no. 4, pp. 421–424, April 1996.

[3] S. R. Messer, J. Agzarian, and D. Abbott, "Optimal wavelet denoising for phonocardiograms," *Microelectronics Journal*, vol. 32, no. 12, pp. 931–941, December 2001.

[4] O. Beya, E. Fauvet, and O. Laligant, "EDA, approche non linéaire de débruitage des signaux cardiaques," in *Proc. COmpression et REpresentation des Signaux Audiovisuels*, Le Creusot, France, November 2013, pp. 166–171.

[5] D. D. Lee and H. S. Seung, "Algorithms for Non-negative Matrix Factorization," in *Proc. Neural Information Processing Systems*, 2001.

[6] F. Sedighin, M. Babaie-Zadeh, B. Rivet, and C. Jutten, "Multimodal Soft Nonnegative Matrix Co-Factorization for Convolutive Source Separation," *IEEE Trans. on Signal Processing*, vol. 65, no. 12, pp. 3179–3190, June 2017.

[7] N. Seichepine, S. Essid, C. Fevotte, and O. Cappe, "Soft nonnegative matrix co-factorizationwith application to multimodal speaker diarization," in *IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, May 2013, pp. 3537–3541.

[8] N. Souviraa-Labastie, A. Olivero, E. Vincent, and F. Bimbot, "Multi-Channel Audio Source Separation Using Multiple Deformed References," *IEEE/ACM Trans. on Audio, Speech, and Language Processing*, vol. 23, no. 11, pp. 1775–1787, November 2015.

[9] D. Lahat, T. Adal, and C. Jutten, "Multimodal Data Fusion: An Overview of Methods, Challenges and Prospects," *in Proc. IEEE*, vol. 103, no. 9, September 2015.

[10] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "Biometric authentication based on PCG and ECG signals: present status and future directions," *Signal, Image and Video Processing*, vol. 8, no. 4, pp. 739–751, May 2014.

[11] S. Noorzadeh, B. Rivet, P.-y. Guméry, and C. Umr, "a Multi-Modal Approach Using a Non-Parametric Model To Extract Fetal Ecg," *in Proc. International Conference on Acoustics, Speech, and Signal Processing*, pp. 832–836, 2015.

[12] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. I. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. John Wiley, 2009, no. June.

[13] M. Niegowski and M. Zivanovic, "ECG-EMG separation by using enhanced non-negative matrix factorization," in *Proc. Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, August 2014, pp. 4212–4215.

[14] N. Dia, J. Fontecave-jallon, B. Rivet, and P.-y. Guméry, "Application De La Factorisation Non-Négative Des Matrices (Nmf) Pour Le Débruitage Des Signaux Phonocardiographiques (Pcg)," in *Proc. Groupe d'Etudes du Traitement du Signal et des Images*, September 2017.

[15] A. Liutkus, F.-R. Stöter, Z. Rafii, D. Kitamura, B. Rivet, N. Ito, N. Ono, and J. Fontecave, "The 2016 Signal Separation Evaluation Campaign," in *Proc. LVA/ICA*. Springer, Cham, February 2017, pp. 323–332.

[16] C. Févotte and J. Idier, "Algorithms for Nonnegative Matrix Factorization with the β-Divergence," *Neural Computation*, vol. 23, no. 9, pp. 2421–2456, September 2011.

[17] S. Noorzadeh, "Extraction de l'ECG du foetus et de ses caractéristiques grâce à la multi-modalité," *http://www.theses.fr*, November 2015.

[18] E. Holt, "Gallop rhythm," *American Heart Journal*, vol. 2, no. 4, pp. 453–469, April 1927.

[19] E. Vincent, R. Gribonval, and C. Fevotte, "Performance measurement in blind audio source separation," *IEEE Trans. on Audio, Speech and Language Processing*, vol. 14, no. 4, pp. 1462–1469, July 2006.