HEART RATE ESTIMATION FROM PHONOCARDIOGRAM SIGNALS USING NON-NEGATIVE MATRIX FACTORIZATION

[1,2]*Nafissa Dia,* [1]*Julie Fontecave-Jallon,* [1]*Pierre-Yves Gumery,* [2]*Bertrand Rivet*

[1]Univ. Grenoble Alpes, CNRS, CHU Grenoble Alpes, Grenoble INP*, TIMC-IMAG, Grenoble, France
[2]Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, Grenoble, France
*Institute of Engineering Univ. Grenoble Alpes

ABSTRACT

Electrocardiogram (ECG) is classically considered for heart rate (HR) estimation. However in certain conditions, its use may be difficult and alternative techniques, such as phonocardiograhpy (PCG), are investigated. For PCG signals, in most studies, the challenge is to detect and annotate the heart sounds S_1 and S_2, which may become quasi-impossible in case of noise. In this paper, we present a novel approach of HR estimation from PCG signals based on non-negative matrix factorization (NMF), applied to the spectrogram of PCG, considered as a source-filter model. Compared to state of the art methods, specific considerations based on the signal properties have been included to ensure the reliability of the decomposition. HR estimations obtained from noise-free and noisy real PCG signals are evaluated by comparison to HR estimation from synchronous ECG.

Index Terms— Phonocardiogram, Heart rate estimation, Non-Negative Matrix Factorization

1. INTRODUCTION

Electrocardiogram (ECG) is the reference technique for heart rate (HR) monitoring in adults; the cardiac frequency is usually estimated according to the detection of prominent R waves on ECG signals [1]. However, besides the classical practical limitations such as movement sensibility, the use of electrodes also appears to be a problem when used on elderly people. Therefore, due to the population aging and the increase of health monitoring needs, there is a current interest in alternative non-invasive techniques to replace ECG for HR monitoring in different situations. Among others, Phonocardiography (PCG) is a diagnostic graphical method of recording echoes that accompany mechanical vibrations originating in the heart and vessels. It is used to register heart sounds and murmurs in the diagnosis of heart diseases. It uses a microphone attached at the surface of the chest wall to the usual auscultatory points of the heart. Among cardiac
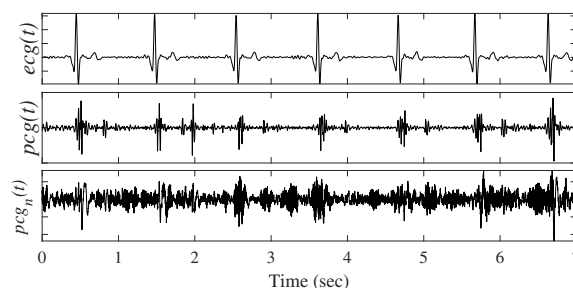
Fig. 1. Synchronous signals. From top to bottom: ECG, PCG and noisy PCG.

sounds, two, noted respectively S_1 and S_2, are particularly audible and correspond to the closure of respectively the atrial-ventricular valves (beginning of the ventricular systole) and the aortic and pulmonary valves (onset of the ventricular diastole). As shown in Fig.1, a PCG signal is therefore a succession of two bumps S_1 and S_2, following the R peak of a synchronous ECG signal.

S_1 and S_2 occur in all healthy individuals and are more or less easy to detect. Therefore, most studies on PCG analysis deal with the detection of these first two heart sounds and can be divided according to whether they use a different signal than PCG (ECG, carotid / jugular pulse etc.) or the only PCG. A wide amount of methods uses the signal envelop detection for analyzing heart sounds. Common used methods are based on homomorphic filtration, signal energy computation (Square energy, Shannon entropy, Shannon energy) or Hilbert transform [2, 3, 4, 5]. Other proposed methods are based on Short Time Fourier transformation and Wavelet transformation in order to investigate the exact features of the heart sound [6, 7]. More sophisticated methods include probabilistic models such as Hidden Markov Models for segmentation of heart sounds [8, 9] or neural networks [10, 11].

The amount of proposed methods highlights the difficulty of detection of heart sounds and this difficulty is further increased in noisy conditions, as illustrated in Figure 1 with a noisy PCG on which S_1 and S_2 cannot be identified at each cardiac cycle. Most of the algorithms will suffer from a lack of robustness for sounds detection in such conditions.

Moreover, for HR monitoring, the classification of S_1 and S_2 sounds is essential to estimate the cardiac beat-to-beat intervals as S_1-S_1, similarly as R-R intervals on ECG signals.

Therefore, we propose a new method for HR monitoring from PCG signals which needs the only source PCG, allows to avoid the distinction between S_1 and S_2 and is robust to noise. The method is based on the Non-negative Matrix Factorization (NMF) approach [12, 13, 14] applied on the spectrogram of the PCG signal.

2. PROPOSED METHODOLOGY

In this section, we will first detail the reasons why PCG signals are modelled as source-filter based on their quasi-harmonic structure. Indeed, this structure carries an important information namely heart rate which will be estimated in a second step using a NMF-based algorithm presented at the end of the section.

2.1. Signal modelling

To model a PCG signal $x(t)$, which is a succession of heart beats, one can write that

$$x(t) = \sum_i s_i(t - \tau_i), \qquad (1)$$

where τ_i are the cardiac beats instants and s_i is the shape of S_1 and S_2 bumps succession at the ith heart beat. This expression can be rewritten as

$$x(t) = \sum_i s_i(t) * \delta(t - \tau_i), \qquad (2)$$

which is nothing else but a time-varying filtering of an input signal $e(t) = \sum_i \delta(t - \tau_i)$ by a time-varying filter whose impulse response $s_i(t)$ depends on the considered time. Indeed, due to the non-stationarity of physiological signals, both the duration, $\Delta_i = \tau_{i+1} - \tau_i$, and the shape, $s_i(t)$, of each beat can be different from one beat to an other one. Consequently, the PCG signal can be modelled as a source-(time-varying) filter model. To study such non-stationary signals, it is common to apply a time-frequency representation like the spectrogram. It is expected that the spectrogram of the signal (2) exhibits an harmonic structure due to the excitation term (i.e. the input signal $e(t)$) whose fundamental frequency is time-varying and whose amplitude is modulated by a time-varying envelop due to the shape of each heart beat. As shown in Fig. 2, the quasi-harmonic structure is clearly shown and the fluctuation of the fundamental frequency represents the evolution of the heart rate that has to be estimated.

2.2. Estimation of heart rate by NMF

In time-frequency domain, the short-time Fourier transform (STFT) $X(f, t)$ of $x(t)$ is expressed as

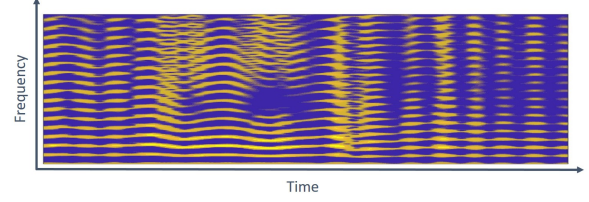$$\forall (f, t), \quad X(f, t) = X^{(e)}(f, t) X^{(\varphi)}(f, t), \qquad (3)$$

Fig. 2. Spectrogram of a noisy-free PCG (F_s= 1kHz, window = $4s$, noverlap = $64ms$, nfft = $8s$)

where $X^{(e)}(f, t)$ (resp. $X^{(\varphi)}(f, t)$) is the STFT of the excitation (resp. the filter). This equation can be recast into matrix form to model the spectrogram as

$$\mathbf{X} = \mathbf{X}^{(e)} \odot \mathbf{X}^{(\varphi)}, \qquad (4)$$

$\mathbf{X} \in \mathbb{R}^{F \times N}$, $\mathbf{X}^{(e)} \in \mathbb{R}^{F \times N}$ and $\mathbf{X}^{(\varphi)} \in \mathbb{R}^{F \times N}$, with F the number of frequency bins and N the number of time windows, whose (f, t)th elements are $|X(f, t)|^2$, $|X^{(e)}(f, t)|^2$ and $|X^{(\varphi)}(f, t)|^2$, respectively. $\odot$ is the Hadamard product (i.e. the element-wise multiplication).

An efficient way to analyze a spectrogram $\mathbf{A}$ is the NMF [12, 13] which allows to factorize it as a product of two non-negative matrices $\mathbf{WH}$ of lower rank than the rank of $\mathbf{A}$. Applied on the source-filter model, this leads to [14]

$$\mathbf{X} \simeq \left(\mathbf{W}^{(e)} \mathbf{H}^{(e)} \right) \odot \left(\mathbf{W}^{(\varphi)} \mathbf{H}^{(\varphi)} \right), \qquad (5)$$

where $\mathbf{W}^{(e)}$ and $\mathbf{W}^{(\varphi)}$ are the spectral templates of the excitation and filter, respectively and $\mathbf{H}^{(e)}$ and $\mathbf{H}^{(\varphi)}$ are their related temporal amplitudes.

However, this NMF decomposition (5) suffers from a problematic ambiguity: without any constraint, the two terms $\mathbf{W}^{(e)} \mathbf{H}^{(e)}$ and $\mathbf{W}^{(\varphi)} \mathbf{H}^{(\varphi)}$ are perfectly interchangeable due to the element-wise multiplication. One way to overcome this identifiability issue is to add some constraints on both excitation and filter parts. The two proposed constraints are firstly to fix $\mathbf{W}^{(e)}$ such that each column is a Dirac comb with different fundamental frequencies. Consequently, the role of $\mathbf{H}^{(e)}$ is to rightly select one consistent cardiac frequency among all these present in $\mathbf{W}^{(e)}$ at each time. And secondly to ensure that $\mathbf{W}^{(\varphi)}$ is used to model the spectral envelop of the PCG beats and $\mathbf{H}^{(\varphi)}$ is their temporal evolution, a smoothness constraint is added on the columns of $\mathbf{W}^{(\varphi)}$ which models the spectral envelop of each PCG beat. Contrary to [14], this smoothness property is not ensured by decomposing $\mathbf{W}^{(\varphi)}$ on a dictionary of smooth functions but by adding a constraint term into the criterion as detailed in the next section. This allows us to reduce the number of parameters to be estimated and thus speed up the algorithm.

2.3. Algorithm based on NMF

To estimate the excitation and filter parts by NMF, the following cost function is minimized[1]

$$C\big(\mathbf{H}^{(e)}, \mathbf{W}^{(\varphi)}, \mathbf{H}^{(\varphi)}\big) = D_2(\mathbf{X}|\mathbf{V}) + \gamma_s S\big(\mathbf{W}^{(\varphi)}\big), \quad (6)$$

where $\mathbf{V} = \big(\mathbf{W}^{(e)}\mathbf{H}^{(e)}\big) \odot \big(\mathbf{W}^{(\varphi)}\mathbf{H}^{(\varphi)}\big)$, $D_2(\cdot|\cdot)$ is the Frobenius norm between $\mathbf{X}$ and $\mathbf{V}$

$$D_2\big(\mathbf{X}|\mathbf{V}\big) = \frac{1}{2} \parallel \mathbf{X} - \mathbf{V} \parallel_{\mathrm{F}}^2 \quad (7)$$

and $S\big(\mathbf{W}^{(\varphi)}\big)$ is a smoothness constraint on the columns of $\mathbf{W}^{(\varphi)}$

$$S(\mathbf{W}^{(\varphi)}) = \sum_{k_\varphi=1}^{K_\varphi} \sum_{f=2}^{F} \Big(W_{f,k_\varphi}^{(\varphi)} - W_{f-1,k_\varphi}^{(\varphi)}\Big)^2, \quad (8)$$

where K_φ is the number of filter components. $S\big(\mathbf{W}^{(\varphi)}\big)$ constrains the spectral templates of the filter to have small variations with respect to the frequency. However some scaling factor ambiguities also affect the decomposition (5): indeed substituting $\mathbf{W}^{(e)}$ and $\mathbf{H}^{(e)}$ by $\mathbf{D}^{(W)}\mathbf{W}^{(e)}\mathbf{D}^{(e)}$ and $\big(\mathbf{D}^{(e)}\big)^{-1}\mathbf{H}^{(e)}\mathbf{D}^{(H)}$, respectively and $\mathbf{W}^{(\varphi)}$ and $\mathbf{H}^{(\varphi)}$ by $\big(\mathbf{D}^{(W)}\big)^{-1}\mathbf{W}^{(\varphi)}\mathbf{D}^{(\varphi)}$ and $\big(\mathbf{D}^{(\varphi)}\big)^{-1}\mathbf{H}^{(e)}\big(\mathbf{D}^{(H)}\big)^{-1}$, respectively, where all $\mathbf{D}$ matrices are diagonal with positive entries, leads to the same approximation. To overcome these scaling ambiguities, the smoothness constraint is changed to

$$S\big(\mathbf{W}^{(\varphi)}, \mathbf{H}^{(e)}, \mathbf{W}^{(\varphi)}\big) =$$
$$\sum_{k_e,k_\varphi,n} \big(H_{k_e,n}^{(e)}\big)^2 \big(H_{k_\varphi,n}^{(\varphi)}\big)^2 \sum_{f=2}^{F} \Big(W_{f,k_\varphi}^{(\varphi)} - W_{(f-2),k_\varphi}^{(\varphi)}\Big)^2, \quad (9)$$

To optimize (6) with the smoothness constraint (9), $\mathbf{H}^{(e)}$, $\mathbf{W}^{(\varphi)}$ and $\mathbf{H}^{(\varphi)}$ components are updated alternatively using a majoration minimization (MM) algorithm. Due to the lack of space, the detail of the auxiliary functions are omitted but are based on the Févotte *et al.* paper [15]. This leads to the following multiplicative updates of $\mathbf{H}^{(e)}, \mathbf{W}^{(\varphi)}, \mathbf{H}^{(\varphi)}$

$$\mathbf{H}^{(e)} \leftarrow \mathbf{H}^{(e)} \odot \Big([\mathbf{W}^{(e)}]^T (\mathbf{V}^{(\varphi)} \odot \mathbf{X})\Big)$$
$$\oslash \Big([\mathbf{W}^{(e)}]^T (\mathbf{V}^{(\varphi)} \odot \mathbf{V}) + \gamma_s \mathbf{H}^{(e)} \odot (\mathbf{1}_{K_e} \mathbf{c}_n)\Big),$$

with $\mathbf{1}_{K_e}$ a vector of ones of length K_e and $\mathbf{c}_n$ a row vector whose l-th entry is $c_n(l) = \sum_f (W_{f,k_\varphi}^{(\varphi)} - W_{(f-2),k_\varphi}^{(\varphi)})^2 \big(\mathbf{h}_l^{(\varphi)}\big)^{.2}$, where $\mathbf{h}_l^{(e)}$ is the l-th column of $\mathbf{H}^{(e)}$ and $(\cdot)^{.2}$ is the element wise square power.

$$\mathbf{H}^{(\varphi)} \leftarrow \mathbf{H}^{(\varphi)} \odot \Big([\mathbf{W}^{(\varphi)}]^T (\mathbf{V}^{(e)} \odot \mathbf{X})\Big)$$
$$\oslash \Big([\mathbf{W}^{(\varphi)}]^T (\mathbf{V}^{(e)} \odot \mathbf{V}) + \gamma_s \mathrm{Diag}(\delta_{\mathbf{W}^{(e)}}) \big(\mathbf{1}_{K_\varphi} \boldsymbol{\beta}_{\mathrm{H}^{(e)}}\big) \odot \mathbf{H}^{(\varphi)}\Big),$$

---

[1]Recall that as discussed into the previous section, $\mathbf{W}^{(e)}$ is kept fixed during the estimation process.

with $\mathrm{Diag}\big(\delta_{\mathbf{W}^{(e)}}\big)$ a diagonal matrix whose k_φ-th entry is $\sum_f (W_{f,k_\varphi}^{(\varphi)} - W_{f-1,k_\varphi}^{(\varphi)})^2$ and $\boldsymbol{\beta}_{H^{(e)}}$ a row vector whose l-th entry is $\sum_{k_e} (H_{k_e,l}^{(e)})^2$.

The equations to update $\mathbf{W}^{(\varphi)}$ depend on the index of the row as follows:

$$\forall f, \quad \mathbf{W}_{f,:}^{(\varphi)} \leftarrow \mathbf{W}_{f,:}^{(\varphi)} \odot \Big(\big(\mathbf{X}\odot\mathbf{V}^{(e)}\big)_{f,:} \big(\mathbf{H}^{(\varphi)}\big)^T + \gamma_s \boldsymbol{\Lambda}_{f,:}^{(\varphi)}\Big)$$
$$\oslash \Big(\big(\mathbf{V}\odot\mathbf{V}^{(e)}\big)_{f,:} \big(\mathbf{H}^{(\varphi)}\big)^T + \gamma_s \boldsymbol{\Gamma}_{f,:}^{(\varphi)}\Big),$$

where $(.)_{f,:}$ corresponds to the f-th row of the related matrix, with if $f = 1$

$$\boldsymbol{\Lambda}_{1,:}^{(\varphi)} = \big(\mathbf{W}_{1,:}^{(\varphi)} + \mathbf{W}_{2,:}^{(\varphi)}\big)\Delta_H \quad \text{and} \quad \boldsymbol{\Gamma}_{1,:}^{(\varphi)} = 2\,\mathbf{W}_{1,:}^{(\varphi)}\Delta_H,$$

if $2 \leq f \leq F - 1$

$$\boldsymbol{\Lambda}_{f,:}^{(\varphi)} = \big(2\mathbf{W}_{f,:}^{(\varphi)} + \mathbf{W}_{(f-1),:}^{(\varphi)} + \mathbf{W}_{(f+1),:}^{(\varphi)}\big)\Delta_H$$
$$\boldsymbol{\Gamma}_{f,:}^{(\varphi)} = 4\,\mathbf{W}_{f,:}^{(\varphi)}\,\Delta_H,$$

and if $f = F$

$$\boldsymbol{\Lambda}_{F,:}^{(\varphi)} = \big(\mathbf{W}_{F,:}^{(\varphi)} + \mathbf{W}_{(F-1),:}^{(\varphi)}\big)\Delta_H \quad \text{and} \quad \boldsymbol{\Gamma}_{F,:}^{(\varphi)} = 2\,\mathbf{W}_{F,:}^{(\varphi)}\Delta_H$$

with $\Delta_H = \sum_n (\boldsymbol{\beta}_{H^{(e)}} \sum_{k_\varphi} \boldsymbol{\beta}_{H^{(\varphi)}})$ where $\boldsymbol{\beta}_{H^{(e)}}$ a row vector whose l-th entry is $\sum_{k_e} (H_{k_e,l}^{(e)})^2$ and $\boldsymbol{\beta}_{H^{(\varphi)}} = (\mathbf{H}^{(\varphi)})^{.2}$.

The initialization of this iterative algorithm is crucial. To this end, some physiological considerations are taken into account to both simplify the choice of the initial matrices $\mathbf{H}^{(e)}$, $\mathbf{W}^{(\varphi)}$ and $\mathbf{H}^{(\varphi)}$ and to speed up the convergence. Each column of $\mathbf{W}^{(e)}$ is a Dirac comb with fundamental frequencies from 30 bpm (*i.e* $0.5Hz$) to 180 bpm (*i.e* $3Hz$), with K_e=100, modulated by an average PCG signal envelop. The choice of a modulated dictionary $\mathbf{W}^{(e)}$ leads to an easier initialization of $\mathbf{W}^{(\varphi)}$. Indeed, the K_φ=2 components of $\mathbf{W}^{(\varphi)}$ model thus the departure from the average PCG envelop instead of itself. Consequently, $\mathbf{W}^{(\varphi)}$ is initialized around 1 and the number of iterations in the NMF can be reduced. Finally, $\mathbf{H}^{(e)}$ is initialized around 1 as the cardiac frequency to be selected is unknown and $\mathbf{H}^{(\varphi)}$ is also initialized around 1 since the temporal evolution of the PCG is unknown.

3. RESULTS: HR ESTIMATION FROM PCG

In this part, synchronous ECG and PCG signals considered for performances evaluation will be presented. Then, HR estimation results from $\mathbf{H}^{(e)}$ will be depicted.

3.1. Signals database

A real signals database previously proposed to Signal Separation Evaluation Campaign in 2016 (SiSEC 2016) [16] is considered for evaluation. Acquisitions at 1kHz were performed on three healthy volunteers at TIMC-IMAG laboratory. PCG
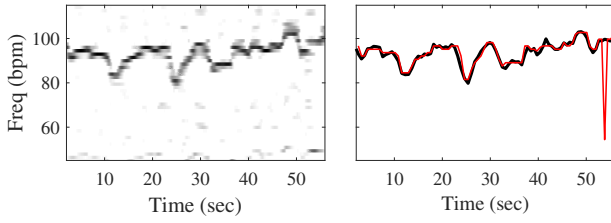
Fig. 3. Estimation of the HR from a clean PCG (sample 9). Left: $\mathbf{H}^{(e)}$ from proposed NMF. Right: HR_{PCG} (red line) is estimated from $\mathbf{H}^{(e)}$ and compared to HR_{ECG} (black line).

signals were recorded with a cardiac microphone (MLT201) put on the skin in front of the heart and band-pass filtered between 15 and 300Hz. Noisy PCG signals were then artificially created from the filtered PCG and different real interference signals (pseudoperiodic noises of breath type, cough, radio, etc.), recorded separately. Synchronous D2 lead ECG signals were also recorded. The database finally consists of 16 samples with a duration from 10 to 70 seconds, each composed of a clean filtered PCG $pcg(t)$, an artificially noisy PCG $pcg_n(t)$ and the synchronous ECG $ecg(t)$.

The classical algorithm of Pan-Tomkins [1] was applied on each ECG to detect R-waves, compute R-R intervals and estimate the heart rate. This HR estimation is noted HR_{ECG}.

3.2. HR estimation from $\mathbf{H}^{(e)}$

The proposed NMF approach is applied on PCG signals. For an easier use of the source-filter modelling and to keep the excitation spectral template $\mathbf{W}^{(e)}$ as simple as described in Section 2, we consider the rectified PCG signals.

The algorithm allows to estimate the temporal amplitudes of the excitation and filter, as well as the spectral template of the filter. For HR estimation, we focus on the excitation temporal amplitude $\mathbf{H}^{(e)}$. One example can be observed on Fig. 3 for sample 9. A time-varying fundamental frequency is emerging and corresponds to the instantaneous cardiac frequency detected on PCG. Thereby, instantaneous heart rate from PCG is estimated from $\mathbf{H}^{(e)}$, as the frequency of maximum power at each column of $\mathbf{H}^{(e)}$. This estimation is then downsampled so as to associate one PCG cardiac frequency to each R-wave of the ECG signal. The resulting HR estimation on clean filtered PCG signals $pcg(t)$ (respectively on artificially noisy PCG signals $pcg_n(t)$) is noted HR_{PCG} (respectively HR_{nPCG}). These estimations are compared with the one obtained from the reference ECG HR_{ECG}.

3.3. Performances of the proposed method

The performance of the proposed method is evaluated throughout the quality of HR estimation from PCG signals. First, considering the clean PCG signals, the relative error between HR_{ECG} and HR_{PCG} is computed for each sample,
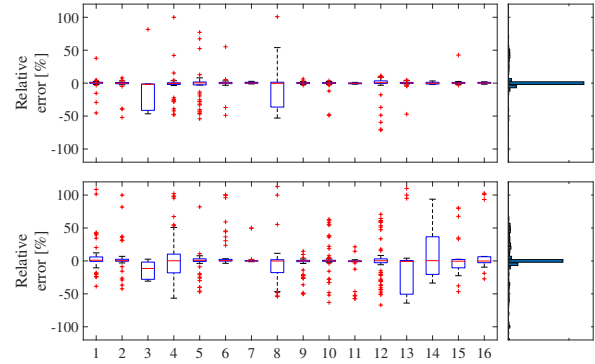

Fig. 4. Performance of HR estimation for noise-free PCG and noisy PCG signals for 16 samples. Relative error between HR_{ECG} and HR_{ECG} (top) or HR_{nPCG} (bottom).

as $\epsilon = (HR_{PCG} - HR_{ECG})/HR_{ECG}$. Results are presented in Fig.4 as boxplots, which represent the parameters distribution: red lines correspond to the median, the boxlimits represent the first and the third quartiles, whiskers highlight extreme values and red x-marks correspond to outliers.

For most of samples, HR estimation from clean filtered PCG signals is very similar to HR estimation from ECG. Indeed, the median of the relative error for all samples is 0.1% and 770 heart cycles over 812 lead to a relative error lower than 5%. This is confirmed by the global histogram which highlights a peaky distribution around 0. Considering the noisy PCG signals, a similarly good behavior is observed. The median of the error is 0.1% for the overall set of noisy samples. The related histogram shows a slight higher spread of the relative errors. However, more than 84% (686 over 812) of heart cycles are detected with an error lower than 5%.

4. CONCLUSION

In this study, an improvement of a NMF decomposition based on a source-filter model has been proposed to estimate the HR from PCG in real conditions from noisy signals. Based on signal properties, constraints have been added to avoid identifiabilities issues. The proposed approach prevents from the difficult detection of S_1 and S_2 cardiac sounds as most often considered in PCG analysis (see Section 1) and allows a very satisfactory HR estimation from noise-free PCG signals. The origins of estimation errors mainly lie on the transition from $\mathbf{H}^{(e)}$ to the HR computation. Some properties of HR may be introduced for a more robust estimation.

In case of noisy signals, for which the sounds detection and the cardiac frequency estimation are quasi-impossible on temporal signals, the HR estimation obtained with our methodology is promising. We intend to add several improvements to the method, considering on one hand the sparse property of the excitation temporal amplitude $\mathbf{H}^{(e)}$, and on the other hand a joint NMF of synchronous noisy ECG and PCG signals.

5. REFERENCES

[1] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, March 1985.

[2] H. Liang, S. Lukkarinen, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelogram," in *Computers in Cardiology 1997*, Sept 1997, pp. 105–108.

[3] Samjin Choi and Zhongwei Jiang, "Comparison of envelope extraction algorithms for cardiac sound signal segmentation," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1056 – 1069, 2008.

[4] Amir A. Sepehri, Arash Gharehbaghi, Thierry Dutoit, Armen Kocharian, and A. Kiani, "A novel method for pediatric heart sound segmentation without using the ecg," *Computer Methods and Programs in Biomedicine*, vol. 99, no. 1, pp. 43 – 48, 2010.

[5] D. Gill, N. Gavrieli, and N. Intrator, "Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model," in *Computers in Cardiology, 2005*, Sept 2005, pp. 957–960.

[6] A. Djebbari and F. Bereksi Reguig, "Short-time fourier transform analysis of the phonocardiogram signal," in *ICECS 2000. 7th IEEE International Conference on Electronics, Circuits and Systems (Cat. No.00EX445)*, Dec 2000, vol. 2, pp. 844–847 vol.2.

[7] Jung Jun Lee, Sang Min Lee, In Young Kim, Hong Ki Min, and Seung Hong Hong, "Comparison between short time fourier and wavelet transform for feature extraction of heart sound," in *Proceedings of IEEE. IEEE Region 10 Conference. TENCON 99. 'Multimedia Technology for Asia-Pacific Information Infrastructure' (Cat. No.99CH37030)*, Sept 1999, vol. 2, pp. 1547–1550 vol.2.

[8] S E Schmidt, C Holst-Hansen, C Graff, E Toft, and J J Struijk, "Segmentation of heart sound recordings by a duration-dependent hidden markov model," *Physiological Measurement*, vol. 31, no. 4, pp. 513, 2010.

[9] L. G. Gamero and R. Watrous, "Detection of the first and second heart sound using probabilistic models," in *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)*, Sept 2003, vol. 3, pp. 2877–2880 Vol.3.

[10] Cota Navin Gupta, Ramaswamy Palaniappan, Sundaram Swaminathan, and Shankar M. Krishnan, "Neural network classification of homomorphic segmented heart sounds," *Applied Soft Computing*, vol. 7, no. 1, pp. 286 – 297, 2007.

[11] Chung-Hsien Wu, Ching-Wen Lo, and Jhing-Fa Wang, "Computer-aided analysis and classification of heart sounds based on neural networks and time analysis," in *1995 International Conference on Acoustics, Speech, and Signal Processing*, May 1995, vol. 5, pp. 3455–3458 vol.5.

[12] Pentti Paatero and Unto Tapper, "Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values," *Environmetrics*, vol. 5, no. 2, pp. 111–126, 1994.

[13] Daniel D Lee and H Sebastian Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788, 1999.

[14] J. Durrieu, A. Ozerov, C. Févotte, G. Richard, and B. David, "Main instrument separation from stereophonic audio signals using a source/filter model," in *EUSIPCO*, Aug 2009, pp. 15–19.

[15] Cédric Févotte and Jérôme Idier, "Algorithms for non-negative matrix factorization with the β-divergence," *Neural computation*, vol. 23, no. 9, pp. 2421–2456, 2011.

[16] Antoine Liutkus, Fabian-Robert Stöter, Zafar Rafii, Daichi Kitamura, Bertrand Rivet, Nobutaka Ito, Nobutaka Ono, and Julie Fontecave, "The 2016 signal separation evaluation campaign," in *International Conference on Latent Variable Analysis and Signal Separation*. Springer, 2017, pp. 323–332.