



HAL
open science

Analysis-synthesis of the phonocardiogram based on the matching pursuit method.

Xuan Zhang, Louis-Gilles Durand, Lotfi Senhadji, Howard C. Lee, Jean-Louis Coatrieux

► **To cite this version:**

Xuan Zhang, Louis-Gilles Durand, Lotfi Senhadji, Howard C. Lee, Jean-Louis Coatrieux. Analysis-synthesis of the phonocardiogram based on the matching pursuit method.. IEEE Transactions on Biomedical Engineering, 1998, 45 (8), pp.962-71. 10.1109/10.704865 . inserm-00460389

HAL Id: inserm-00460389

<https://inserm.hal.science/inserm-00460389>

Submitted on 28 Feb 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analysis-Synthesis of the Phonocardiogram Based on the Matching Pursuit Method

Xuan Zhang, Louis-Gilles Durand,* *Senior Member, IEEE*, Lotfi Senhadji, *Member, IEEE*, Howard C. Lee, and Jean-Louis Coatrieux, *Fellow, IEEE*

Abstract—The matching pursuit method of Mallat and Zhang is applied to the analysis and synthesis of phonocardiograms (PCG's). The method is based on a classical Gabor wavelet or time-frequency atom which is the product of a sinusoid and a Gaussian window function. It decomposes a signal into a series of time-frequency atoms by an iterative process based on selecting the largest inner product of the signal (and the subsequent residues) with atoms from a redundant dictionary. The Gaussian window controls the envelope duration and time position of each atom; and the sinusoid represents the frequency. The method was applied to two sets of PCG's: one with very low-noise level and the other with 10% noise energy. Each data base includes 11 PCG's representing the normal and the pathological conditions of the heart. The normalized root-mean-square error (NRMSE) was computed between the original and the reconstructed signals. The results show that the matching pursuit method is very suitable to the transient and complex properties of the PCG's, as it yielded excellent NRMSE's around 2.2% for the two sets of 11 PCG's tested.

Index Terms— Analysis-synthesis, heart murmurs, heart sounds, matching pursuit method, phonocardiogram, signal processing, wavelet transform.

I. INTRODUCTION

A phonocardiogram (PCG) is a recording of the acoustic waves produced by the mechanical action of the heart. It generally consists of two kinds of acoustic vibrations: the heart sounds and the heart murmurs [12]. The heart sounds are low-frequency transient signals produced by the vibration of the heart valves after closure and opening, and/or by the vibration of the whole myocardium and the associated structures. The murmurs are noise-like signals having a more complex structure as they are caused by the turbulence of blood flow. They can be heard sometimes in normal hearts, but most generally in abnormal hearts.

X. Zhang is with the IRCM, Université de Montréal, Montréal, P.Q. H2W 1R7 Canada and the Department of Electrical Engineering, McGill University, Montréal, P.Q. H3A 2A7 Canada.

*L.-G. Durand is with the IRCM, Université de Montréal, Montréal, P.Q. H2W 1R7 Canada, and the Department of Electrical Engineering, McGill University, Montréal, P.Q. H3A 2A7 Canada (e-mail: durandlg@ircm.umontral.ca).

L. Senhadji and J.-L. Coatrieux are with the Laboratoire Traitement du Signal et de l'Image, INSERM, Université de Rennes I, Campus de Beaulieu, 35042 Rennes Cedex France.

H. C. Lee is with the Department of Electrical Engineering, McGill University, Montréal, P.Q. H3A 2A7 Canada.

Although heart auscultation has been recognized for a long time as an important tool for the diagnosis of heart disease, its accuracy is questionable due to the limitations of auditory perception of the sounds which are closely spaced in time or have low-frequency components. In addition, as the art of heart auscultation is difficult to learn for medical students, it is desirable to develop new methods for helping them to learn faster and easier [4], [11]. We propose to apply time-frequency scaling transformations to the PCG to solve some of the difficulties found during auscultation and help the physicians to provide a better diagnosis of heart diseases. Some examples of the application of time-frequency scaling transformations of the PCG for this purpose can be summarized as follows.

- 1) The temporal rate of the various PCG components can be reduced by time-scale expansion. This could help the physicians to better detect the different sound components which are difficult to distinguish on the original PCG sequence, especially when the heart rate is increased. Also, it may help medical students to better understand the PCG structure and learn faster the art of auscultation.
- 2) Time-scaling can be useful in veterinary medicine for the auscultation of small animals like cats, which have fast heart rate up to 200 bpm.
- 3) Frequency expansion could be helpful to better distinguish the heart sounds with multiple low-frequency components.
- 4) Heart sounds produced by mechanical heart valves have a much higher frequency bandwidth (1–50 kHz) than biological heart sounds [4]. Frequency compression could be useful to shift the ultrasonic part (above 18 kHz) to the sensitive range of human hearing (2–10 kHz) and help the physician to better appreciate these mechanical sounds.

In summary, the time-frequency scaling transformations could be used to expand the temporal rate of the PCG in time without changing its perceptual qualities, or alternatively shift the frequency components while the temporal characteristics are maintained. In order to achieve this objective, we need high-quality analysis-synthesis of the PCG in our approach. In the present paper, we propose an analysis-synthesis method that is suitable for time-frequency scaling transformations and applicable to both the transient property of the heart sounds and the random property of the heart murmurs. The results of a study of the time-frequency scaling of the PCG are reported in a companion paper [14].

The previous studies performed for the analysis and synthesis of the PCG were focused only on the heart sounds. For instance, a pole-zero model (or autoregressive moving average model) was used by Joo *et al.* [5] for the reconstruction of the second heart sound in patients with a bioprosthetic heart valve implanted in the aortic position. Another method proposed by Köymen is the damped sinusoid model [6] which was used to study the power spectra of the first heart sound produced by mechanical prosthetic heart valves implanted in the mitral position. Yu, Chen, and Durand [10] also evaluated the damped sinusoid model by using the mean filter of forward and backward predictor for the analysis-synthesis of the aortic component of the second heart sound in dogs. The pole-zero and the damped sinusoid models were proven to be suitable models for the analysis-synthesis of heart sounds.

The wavelet transform has been used recently for the analysis of the electrocardiogram and evoked potentials [1], [9]. But, to our knowledge, it has not been applied previously for the analysis-synthesis of the PCG. In our preliminary study, a few different wavelet families were tested for the PCG analysis and synthesis, including the discrete wavelet transform proposed by Bertrand *et al.* [1], the orthonormal wavelet decomposition as used in Senhadji *et al.* [9], and the matching pursuit method developed by Mallat and Zhang [7]. The results showed that all these methods gave small errors for both the heart sounds and the more general type of the PCG which includes both the heart sounds and the heart murmurs.

We also found that the damped sinusoid method and the matching pursuit method had promising potential for time-frequency scaling of heart sounds and murmurs. These two methods basically originate from the same principle: that a complex and transient signal can be described by the parameters of the envelope and frequency of each sinusoidal component. Four parameters are used for each component in the damped sinusoid model: the damping factor, the amplitude, the frequency, and the phase. In addition to these parameters, the matching pursuit method extracts another parameter called the time-transition or the time-position. The damped sinusoid method gave good reconstruction results for short transient signals, but not for the heart murmurs with longer duration. Therefore, a segmentation procedure was needed to separate the PCG into short segments in order to satisfy the requirements of the damped sinusoid method. The steps for segmenting the PCG's are complex and time consuming. Also, the reconstruction results are affected by the discontinuity between consecutive segments. The matching pursuit method has been used for analysis-synthesis of signals such as speech and chirps [7]. Instead of using a damped sinusoid wavelet, a Gaussian wavelet with a varying standard deviation and time-transition is used to model the local waveform of the nonstationary signal. The advantage of this method is its capability of analyzing signals with longer time duration. Due to these characteristics, the reconstruction of PCG's based on the matching pursuit method provided a much better approximation of the original PCG's than the damped sinusoid model [13], and therefore only the matching pursuit method is presented in this paper. In Section II, the matching pursuit method is introduced briefly and a data base of 11 typical

PCG's used for testing is described. This will be followed by the analysis-synthesis results and the conclusion.

II. MATERIALS AND METHODS

A. The Matching Pursuit Method

The analysis or decomposition of a signal by the matching pursuit algorithm is a complex iterative process which was developed independently by Mallat and Zhang [7] and by Qian and Chen [8]. The method is based on a dictionary which contains a family of functions called time-frequency atoms. The decomposition of a signal is performed by projecting the signal over the function dictionary and by selecting the atoms which can best match the local structure or waveform of the signal. A complete redundant dictionary of time-frequency atoms is generated by scaling, translating, and frequency modulating a normalized window function $h(t)$ which is the product of a wavelet envelope function $g(t)$ and a sinusoid $u(t)$. The matching pursuit method represents a signal as

$$x(t) = \sum_{i=0}^{+\infty} a_i \cdot h_i(t) \quad (1)$$

with

$$h_i(t) = \beta_i \cdot g_i(t) \cdot u_i(t) \quad (2)$$

and

$$\begin{aligned} g_i(t) &= g\left(\frac{t-p_i}{s_i}\right) \\ u_i(t) &= \cos(2\pi f_i t + \phi_i) \end{aligned} \quad (3)$$

where a_i are the expansion coefficients. The square of a_i represents the part of the signal energy associated with atom $h_i(t)$. The scale factors s_i are used to control the width of the envelope of $h_i(t)$, and the parameters p_i control the temporal placement. The parameters β_i are normalizing factors to keep the norm of $h_i(t)$ equal to one. The parameters f_i and ϕ_i are the frequency and the phase of the cosine function $u_i(t)$, respectively. Therefore, five parameters need to be extracted for each atom by the matching pursuit method. For simplicity, we let λ_i represents the parameter set $(a_i, s_i, p_i, f_i, \phi_i)$. In our application

$$g_i(t) = 2^{1/4} e^{-\pi \cdot (t-p_i/s_i)^2}. \quad (4)$$

In practice, a large redundant and complete dictionary (also called the Gabor dictionary) of time-frequency atoms is first built. The dictionary also includes two special cases: One is when $s = 1$ and $f = 0$, and for this, the discrete-time window function $h(n)$ is taken as an impulse at position p . The other case is when $s = N$ and $p = 0$, and is a sine wave represented by $(1/N)e^{i(2\pi \cdot k/N)n}$. In order to decompose a signal $x(t)$ into a set of atoms which can best describe the time-frequency structure of the signal, an iterative orthogonal projection of $x(t)$ onto the dictionary is necessary. Let $h_0(t)$ be one of the atoms in the dictionary. The first projection decomposes the signal into two parts as [7]

$$x(t) = \langle x(t), h_0(t) \rangle h_0(t) + r^1 x(t) \quad (5)$$

where $\langle x(t), h_0(t) \rangle = \int_{-\infty}^{+\infty} x(t) \cdot \bar{h}_0(t) \cdot dt$ denotes the inner product of $x(t)$ and $h_0(t)$, and $\bar{h}_0(t)$ is the complex conjugate of $h_0(t)$. The first term in the right-hand side of (5) is the projection of $x(t)$ onto the atom $h_0(t)$, and the second term $R^1x(t)$ is the residual vector after approximating $x(t)$ in the direction of $h_0(t)$. It is clear that these two components are orthogonal to each other. Because $\|h(t)\|$ is always equals to one, the energy of the signal can be expressed by

$$\|x(t)\|^2 = |\langle x(t), h_0(t) \rangle|^2 + \|R^1x(t)\|^2. \quad (6)$$

In this case, the best match with the signal $x(t)$ is obtained when the residue $\|R^1x(t)\|$ is minimum. In other words, $h_0(t)$ is the best choice among all the atoms in the sense that it gives the maximum inner product with $x(t)$. After this first step, the same process is repeated by using the residue $R^1x(t)$ instead of $x(t)$. If this procedure is repeated until the signal is decomposed into m components, $x(t)$ is represented as

$$x(t) = \sum_{i=0}^{m-1} \langle R^i x(t), h_i(t) \rangle h_i(t) + R^m x(t) \quad (7)$$

where $R^i x(t)$ is the signal residue for the i th iterations and $R^0 x(t) = x(t)$. Similar to (6), $\|x(t)\|^2$ can be written as

$$\|x(t)\|^2 = \sum_{i=0}^{m-1} |\langle R^i x(t), h_i(t) \rangle|^2 + \|R^m x(t)\|^2. \quad (8)$$

It can be shown [7] that as $m \rightarrow \infty$ the signal can be represented as an infinite series of time-frequency atoms from the dictionary without any distortion

$$x(t) = \sum_{i=0}^{\infty} \langle R^i x(t), h_i(t) \rangle h_i(t) \quad (9)$$

$$\lim_{m \rightarrow \infty} R^m x(t) = 0$$

and the energy of the signal is

$$\|x(t)\|^2 = \sum_{i=0}^{\infty} |\langle R^i x(t), h_i(t) \rangle|^2. \quad (10)$$

From the above description, it can be seen that the matching pursuit method finds the time-frequency atoms in a decreasing energy order. The higher-energy components of the signal are always extracted first. These higher-energy components are regarded as the coherent part of the signal due to the similarity between their waveforms and the signal. In practice, two thresholds are used for stopping the iterative process: one is a specified limiting number M of time-frequency atoms, and the other is a residual energy level ε^2 . The iteration process can be stopped either by reaching the preset number M , or by satisfying the following relation between the signal residue and the energy threshold:

$$|R^i x(n)|^2 < \varepsilon^2. \quad (11)$$

TABLE I
THE ZENECA DATA BASE OF NORMAL AND PATHOLOGICAL PCG RECORDINGS

SIGNALS	DESCRIPTION
Z1	First and second heart sounds
Z2	Third heart sound
Z3	Fourth heart sound
Z4	Summation gallop
Z5	Aortic regurgitation
Z6	Mitral regurgitation
Z7	Aortic stenosis
Z8	Mitral stenosis
Z9	Pulmonary stenosis
Z10	Atrial fibrillation
Z11	Pericardial rubbing

B. Data Acquisition

The data base used to evaluate the method is from a compact disc called "Heart Sounds" containing a selection of 11 PCG's of normal and pathological heart sounds and heart murmurs.¹ These PCG's are characteristics of important clinical conditions and they were recorded for educational purpose. All the 11 different digitally recorded signals were used in our study. Table I gives a short description of these PCG's. The first column of the table contains the label of each PCG as recorded on the compact disc, and the second column its description. Each PCG was played back with a Sony CD player (model D-125) and digitized with 16-bit resolution at a sampling frequency of 3 kHz by a Sound Blaster 16 card (Creative Technology Ltd.) in a Pentium computer. Before digitizing the data, the PCG's were low-pass filtered at 1 kHz with an eighth-order Butterworth filter to prevent frequency aliasing.

Since the PCG's recorded on the compact disc are for educational purposes, the background noise level is relatively low. In our study, a significant background noise was introduced to better simulate the recording of PCG's in hospitals. In practice, the noise sources contributing to the PCG recorded on the thorax may be composed of noise produced by the recording instrumentation, ambient noise, thoracic muscular noise, peristaltic intestine noise, and respiratory noise. The contribution of each source may vary significantly depending on the technical characteristics of the recording instrumentation, the recording environment and the physiological status of the subject. In previous studies performed by our group in patients and in animals, it was shown that the mean signal-to-noise ratio of the PCG was around 30 dB and the power

¹The compact disk containing a selection of normal and pathological heart sounds characteristics of important clinical conditions is produced by Zeneca Pharmaceuticals, Alderley House, Alderley Park, Macclesfield, Cheshire, England, SK10 4TG (project 1693), February 1993. (Fax: 44 16 25 51 74 36).

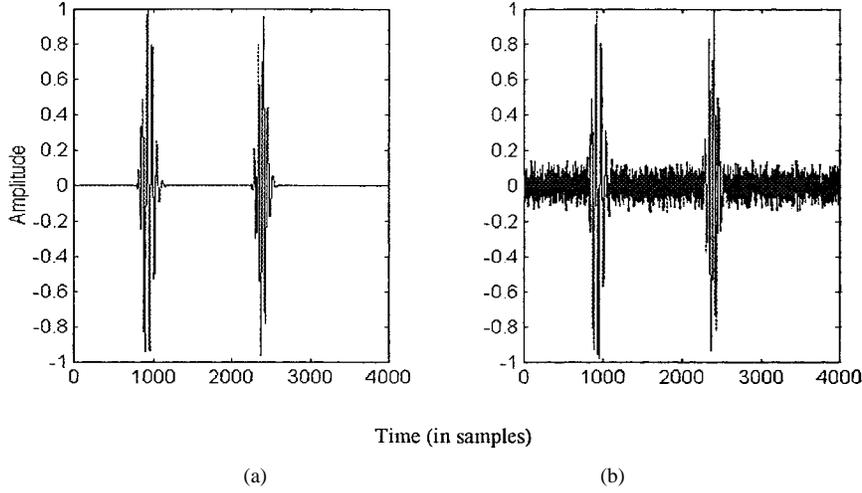


Fig. 1. (a) PCG of the normal first and second heart sounds data (Z1). (b) The PCG in (a) with 10% added Gaussian white noise. The noisy PCG is used to simulate a PCG recorded in relatively severe background noise conditions.

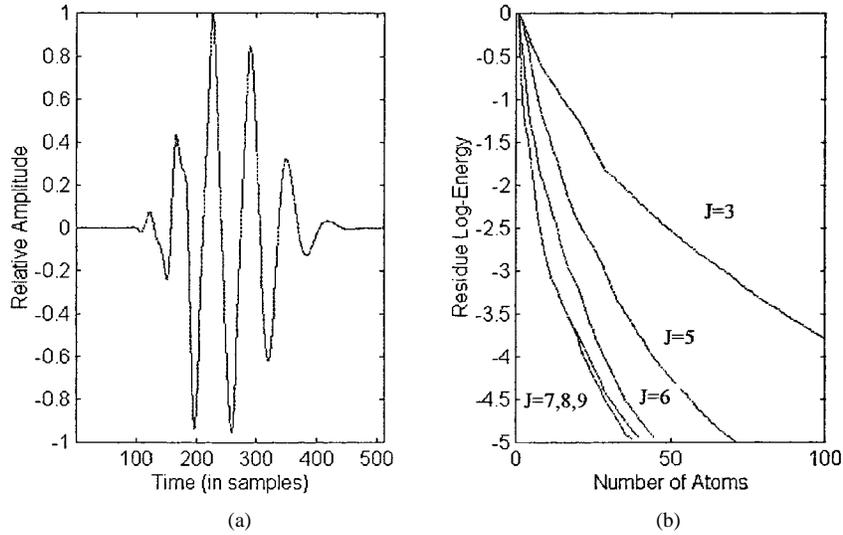


Fig. 2. (a) The first heart sound. (b) The residual log energy curves for different maximum octave values J of the scale s of the time-frequency atoms. ($\log_{10} \varepsilon^2 = 5, M = 100$). A value of $J = 7$ provides the decomposition with the lowest number of atoms.

spectrum of the PCG background noise was relatively uniform between 20 and 50 Hz and decreased at a rate of approximately -20 dB/decade between 50 and 500 Hz [2], [3]. In the present study, we have used a Gaussian white noise with 10% of the energy of each PCG signal to simulate PCG signals recorded in relatively severe background noise conditions. Fig. 1(a) is a recording of the normal heart sound signal Z1, and Fig. 1(b) is the same signal with the added white noise.

C. Data Processing

The matching pursuit software package developed by Mallat and Zhang was run on a UNIX station. The software was written in the C language. The matching pursuit algorithm could be very time consuming if each iteration must cover all the atoms of a large and redundant dictionary. Mallat and Zhang [7] proposed to discretize the dictionary into a subdictionary to simplify the computation. If the signal has N samples, according to (2) and (3) and with $f = 2\pi \cdot k/N$, the

discretized window function of $h(t)$ is

$$h(n) = \beta \cdot g\left(\frac{n-p}{s}\right) \cos\left(\frac{2\pi \cdot k}{N} \cdot n + \phi\right), \quad 0 \leq n < N \quad (12)$$

where β is the normalized factor that sets $\|h(n)\| = 1$. The window scale can only be inside the range of $[1, N]$, while the time-position p and the frequency index k are all integers between zero and $N - 1$. In order to reduce the computation, the scale is also limited to an exponential relation with a dilation factor α such that

$$s = \alpha^j \quad (13)$$

where j is the octave of the scale s which varies between zero and $\log_{\alpha} N$. Hence, only the scales s_i which lie in $[1, N]$ and satisfy (13) are selected from the dictionary. In the software developed by Mallat and Zhang, $\alpha = 2$. Therefore, the signal duration was always zero-padded to a power of two in our

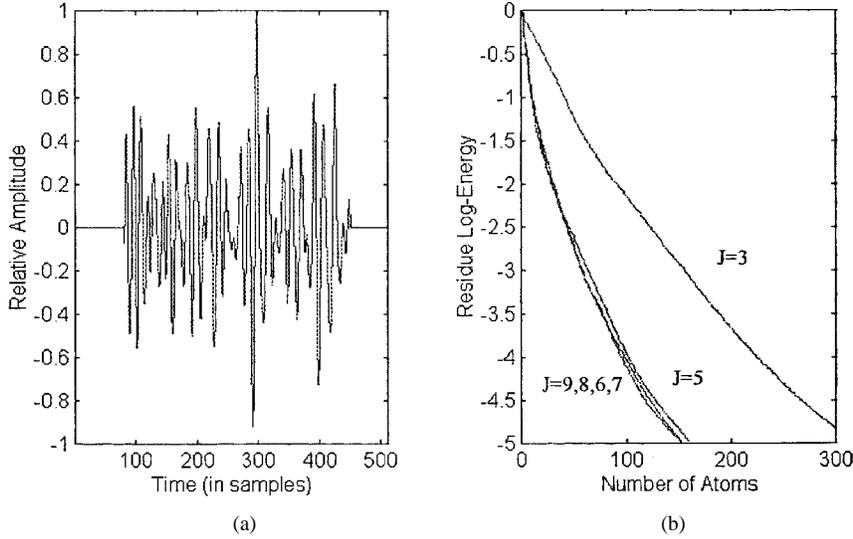


Fig. 3. (a) The heart murmur. (b) The residual log-energy curve of different J values. A value of $J = 9$ or 8 provides the decomposition with the lowest number of atoms.

study. Because of the nonstationarity of the PCG, an upper bound called the maximum octave value $J \leq \log_2 N$ was used for selecting the atoms during decomposition.

In order to get the best way to decompose the signal, i.e., to find the closest projection of the signal in the time-frequency atom dictionary, we studied the residual log-energy as a function of the number of time-frequency atoms for different maximum octave values J of the scale s . The residual energy is calculated by

$$R_m = \log_{10} \frac{E - \sum_{i=1}^m a_i^2}{E} \quad (14)$$

where $E = \sum_{n=0}^{N-1} x^2(n)$ is the total energy of the signal $x(n)$, and a_i is the amplitude of the i th time-frequency atom. That is, R_m is the energy left from the signal after extracting the first m time-frequency atoms by the matching pursuit method. The performance of the method was evaluated on more than one cardiac cycle (about 4000 samples, or 1.3 s) of each digitized signal. The following normalized root-mean-square error (NRMSE) between the original signal and the reconstructed signal was used

$$\text{NRMSE} = 100 \cdot \sqrt{\frac{\sum_{i=0}^{N-1} e^2(n)}{\sum_{i=0}^{N-1} x^2(n)}} \quad (15)$$

where $e(n)$ is the difference between the original and the reconstructed signals. Also, the original signals and the reconstructed signals were played back through the Sound Blaster card for auditory comparison by a cardiologist as well as by the authors of this paper.

III. RESULTS

Fig. 2(b) shows the residual energy curves for the first heart sound of Fig. 2(a) extracted from the data Z1. A fixed energy

TABLE II
THE NUMBER OF m TIME-FREQUENCY ATOMS OBTAINED AT THE END OF THE DECOMPOSITION AND THE CORRESPONDING NRMSE'S IN THE RECONSTRUCTION OF THE FIRST HEART SOUND FOR DIFFERENT MAXIMUM OCTAVE VALUE J ($\varepsilon^2 = 10^{-5}$, $M = 100$)

J	9	8	7	6	5	3
m	40	40	38	45	71	100
NRMSE	0.37	0.37	0.31	0.32	0.31	1.25

TABLE III
THE NUMBER OF m TIME-FREQUENCY ATOMS OBTAINED AT THE END OF THE DECOMPOSITION AND THE CORRESPONDING NRMSE'S IN THE RECONSTRUCTION OF THE MURMUR FOR DIFFERENT MAXIMUM OCTAVE VALUE J ($\varepsilon^2 = 10^{-5}$, $M = 300$)

J	9	8	7	6	5	3
m	168	168	182	175	191	300
NRMSE	1.78	1.78	0.31	0.31	0.31	0.32

threshold of $\varepsilon^2 = 10^{-5}$ and a limiting number $M = 100$ of time-frequency atoms were used for the decomposition process. The signal had 512 samples, so that the maximum octave values of $J = 9, 8, 7, 6, 5$, and 3 were tested. Table II gives the number of m time-frequency atoms obtained from the matching pursuit decomposition, and the NRMSE between the original signal and the reconstructed signal. It can be easily seen that for $J = 9, 8, 7, 6$, and 5 , the iteration was stopped by the same preset energy threshold, with almost the same reconstruction error. For $J = 3$, the iteration was stopped by the limiting number M , indicating that more atoms would be needed to reach the energy threshold. This can also be seen in Fig. 2(b), where the residual log-energy curves for $5 < J < 9$ were stopped with the same preset threshold, at almost the same residual log-energy ($R_m = -5$), while

TABLE IV
THE NRMSE OF THE PCG SIGNALS WITHOUT THE ADDED SIMULATED BACKGROUND NOISE

SIGNAL	Number of time-frequency atoms m ($M=1000$)	NRMSE in %
Z1	39	2.20
Z2	50	2.00
Z3	53	2.11
Z4	67	2.22
Z5	300	2.23
Z6	189	2.28
Z7	204	2.26
Z8	125	2.32
Z9	170	2.33
Z10	44	2.16
Z11	536	2.21
Mean		2.20

for $J = 3$, the decomposition stopped at a higher level of residual energy. The curves for $J = 9, 8$, and 7 are very similar, but when $J = 7$, the decaying slope is little steeper and the reconstruction error is less. Therefore, we can say that it provides the best approximation and the most compact representation of the signal with the matching pursuit method, as confirmed in Table II since the number of atoms m and the NRMSE are both minimum when $J = 7$.

Fig. 3(a) shows a murmur signal with 512 samples extracted from the data Z6. Table III gives the number of m time-frequency atoms obtained from the decomposition, and the NRMSE between the original and the reconstructed signals. The corresponding residual energy curves for different values of J are shown in Fig. 3(b). The energy threshold was $\epsilon^2 = 10^{-5}$ and the maximum number of atoms was limited to $M = 300$. Compared to the first heart sound signal, the waveshape of the murmur changes faster, i.e., it involves more atoms with shorter envelopes. Hence, a smaller octave value like $J = 5$ generated an energy log-residual curve similar to those of $J = 9, 8, 7$, and 6 . Among them, $J = 6$ gives the best performance, as shown in Table III. Fig. 4 shows the comparison of the residual energy curves of the first heart sound and the murmur obtained by using the same energy threshold and a J value of 7 . The murmur signal needs much more time-frequency atoms to reach the same energy threshold.

The matching pursuit method with a maximum octave value of $J = 6$ was applied to the analysis and synthesis of the two sets of the 11 PCG signals, one set with low noise and the other one with an additional white noise as described in Section II. To stop the decomposition process, the energy threshold was set at $\epsilon^2 = 5 * 10^{-4}$ and the limiting number of time-frequency atoms was set at $M = 1000$. For the PCG's

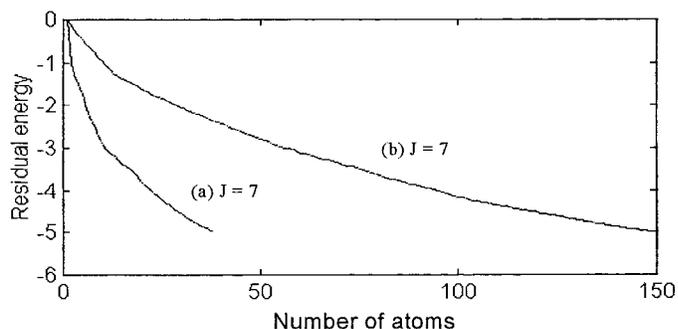


Fig. 4. Residual log-energy curves of (a) the first heart sound and (b) the murmur. Because the murmur is a random signal, it requires much more atoms than the first heart sound to reach the same residual energy level.

with low background noise, the number of m time-frequency atoms at the end of each analysis was always less than 1000 because the energy threshold was reached first. The value of m also varied according to the structure of the signal. For relatively simple PCG's containing only the heart sounds, a small number of atoms varying between 39 and 67 gave very good reconstruction results (NRMSE $\leq 2.2\%$). For signals containing more components, like the murmurs, m varied between 125 and 536. Table IV gives the NRMSE's and the required numbers of m atoms for all the signals.

In the first study on analysis-synthesis of the noisy PCG's, the stopping criterion was set at the same level as for the low-noise PCG's described above. It was found that all the iteration processes were stopped by the limiting number $M = 1000$, before reaching the energy threshold, because of the presence of the added noise. Table V gives the NRMSE's of the reconstructed noisy PCG's. The mean value of the NRMSE's for the 11 noisy PCG's is 6.08%, which is much larger

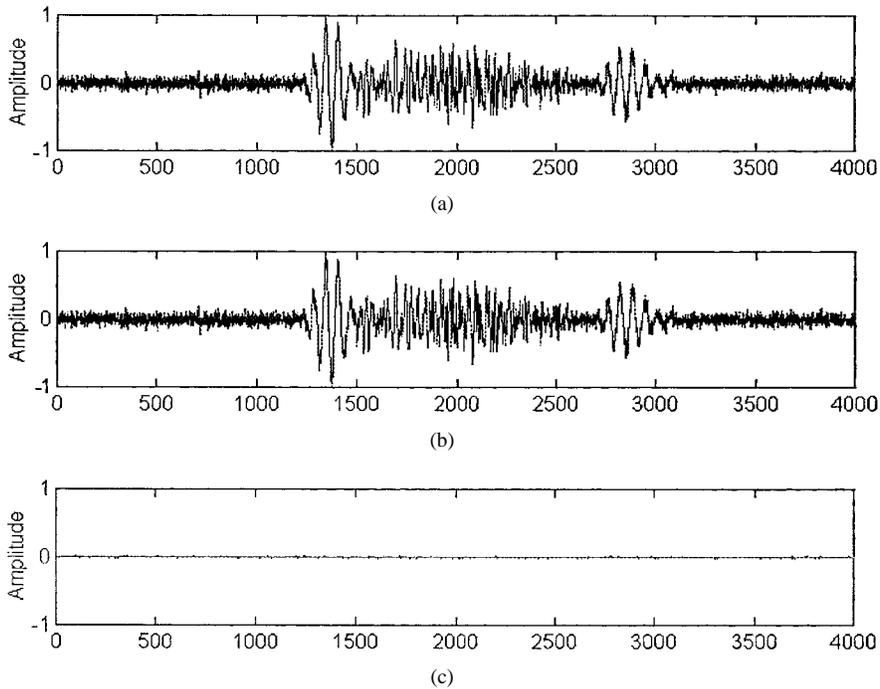


Fig. 5. Some typical results for noisy PCG reconstruction (a) The noisy PCG of data Z7 (aortic stenosis). (b) The reconstructed noisy PCG. (c) The difference between these two PCG's. This figure clearly demonstrates that the matching pursuit method can reconstruct noisy PCG's with very low signal error.

than the error for the low-noise PCG's. In order to further reduce the NRMSE's, the limiting iteration number M was increased to 2000. The second column of Table VI gives the number of atoms required to reach the energy threshold, and the corresponding NRMSE's are shown in the third column. It can be seen that the errors of the reconstructed PCG's with background noise were reduced by representing each of them with more time-frequency atoms. The NRMSE's of the third column in Table VI are similar to those shown in Table IV because the algorithm was stopped by using the same energy threshold. Fig. 5 shows some typical reconstruction results of the noisy PCG of data Z7. A total of 1571 time-frequency atoms were used for reconstructing the PCG. The error between the original and the reconstructed PCG's is negligible.

Unlike the values of m shown in Table IV which varied with the different PCG's, the numbers of atoms for reconstructing the noisy PCG's were almost constant. The matching pursuit method decomposes the signals based on the energy of the components, with all the significant events (the important energy part, or the coherent part) extracted first (within the first 536 iterations as shown in Table IV). Thus, the additional atoms are probably required to represent the noise. For illustration, Fig. 6(c) shows the amplitude of the time-frequency atoms as a function of the atom index for the noisy PCG of data Z1 shown in Fig. 6(b). Only the first 100 time-frequency atoms among all the 1530 are shown. The amplitude curve shows that the required number of atoms for reconstructing the PCG without including the background noise corresponds to the value of a transition point where there is a rapid change of slope. The value of the transition point which separates the 1530 time-frequency atoms into the significant part (corresponding to data Z1 with low noise) and

TABLE V
NRMSE OF THE SIGNALS WITH THE ADDED
SIMULATED BACKGROUND NOISE ($M = 1000$)

SIGNAL	NRMSE in % ($m=M=1000$)
Z1	5.78
Z2	5.63
Z3	5.85
Z4	5.82
Z5	6.12
Z6	6.30
Z7	6.34
Z8	6.06
Z9	6.20
Z10	5.70
Z11	7.13
Mean	6.08

the simulated background noise is 15 or less. In Fig. 6(d), the reconstructed signal by using the first 11 atoms is displayed. It contains all the significant events appearing in the low-noise signal from data Z1, shown in Fig. 6(a), and the NRMSE between it and the signal is 6.4%. Hence, adding more time-frequency atoms for analysis-synthesis of the noisy PCG's actually is mostly required to reduce the error caused by the background noise which always has lower energy, as shown in

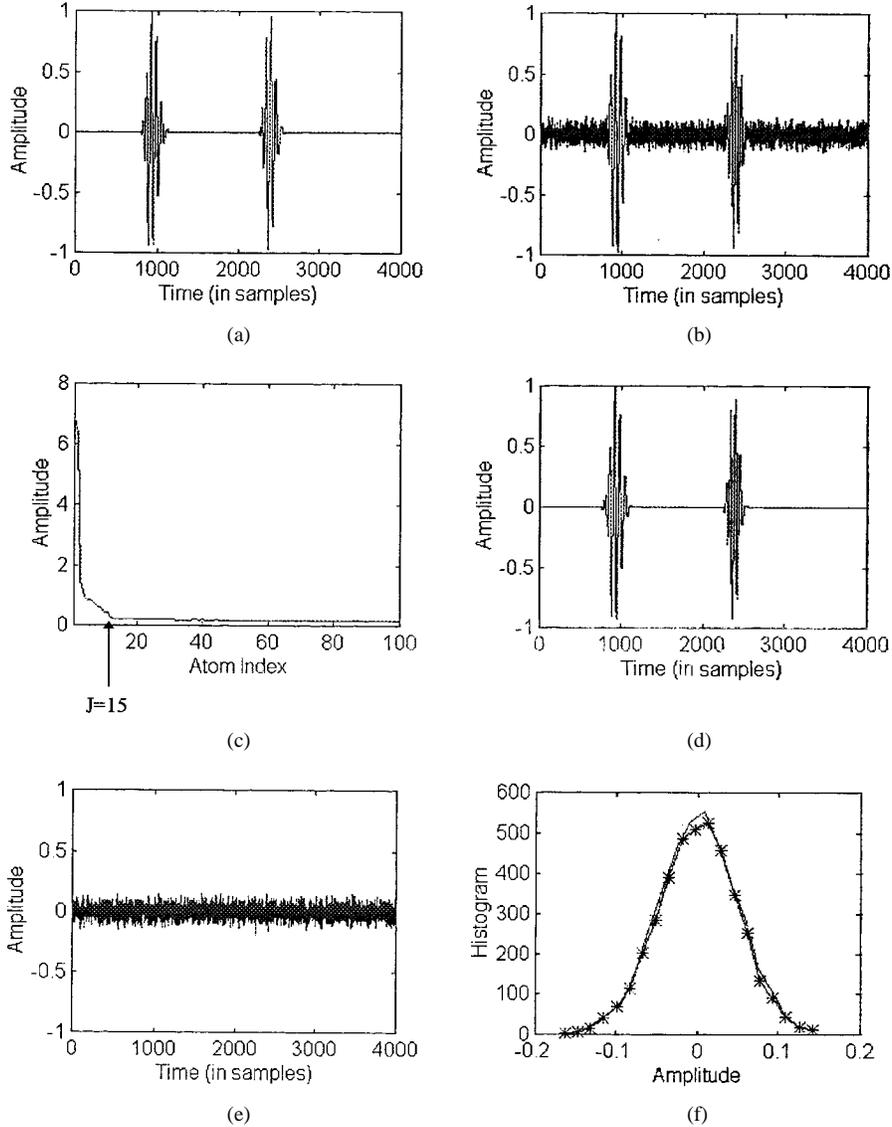


Fig. 6. (a) The low-noise PCG of data Z1 (first and second heart sounds). (b) The same data Z1 with added noise. (c) The amplitude curve of the first 100 time-frequency atoms of the noisy PCG in (b). (d) The reconstructed PCG by using the first 11 time-frequency atoms. (e) The difference between (b) and (d). (f) Histograms of the added noise (*) and the difference signal (e). This figure shows that the matching pursuit method can act as a powerful filter for Gaussian background noise removal.

Fig. 6(e). Therefore, the additional atoms after the transition point do not provide useful information on the significant events.

The error signals between the original signals and the reconstructed signals were very low because the energy threshold used to stop the decomposition process was very low: $\epsilon^2 \leq 5 * 10^{-4}$. The error signal was generally distributed across the complete duration of the signal. Also, it was very difficult to visually found significant differences between the graphics of the original signals and those of the synthesized signals. Thus, the NRMSE was a very good index of difference between the original and the reconstructed signal. According to the results of the present study, the best values of J for analyzing PCG with heart sounds and murmurs is six or seven and the maximal value of M should be no more than 600. The original and the reconstructed versions of the two sets of 11 PCG's using the number of atoms of Table IV and

Table VI, respectively, were compared through listening by a cardiologist and some of the authors of the paper. The reconstructed PCG's were found to sound very natural.

IV. CONCLUSION

The present study shows that the classical Gabor wavelet, as the product of a Gaussian window function with a sinusoid, is a very suitable basis for the analysis and synthesis of the PCG's. With the matching pursuit method, the coherent part or the significant events of the signal is always extracted first by the iterative process, since the decomposition of a signal is based on the selection of the maximum inner product between a time-frequency atoms and the signal (or its residue). This can be clearly seen from the results of the analysis-synthesis of the noisy PCG where the heart sounds are recovered by using the first 11 time-frequency atoms. In Fig. 6(f), histograms of the added noise (*) and the removed noise show that

TABLE VI
NRMSE OF THE SIGNALS WITH THE ADDED
SIMULATED BACKGROUND NOISE ($M = 2000$)

SIGNAL	Iteration number m ($M=2000$)	NRMSE in (%)
Z1	1530	2.22
Z2	1515	2.22
Z3	1532	2.23
Z4	1533	2.23
Z5	1574	2.23
Z6	1571	2.23
Z7	1571	2.23
Z8	1561	2.21
Z9	1567	2.22
Z10	1529	2.22
Z11	1651	2.22
Mean		2.22

almost all the added noise has been removed by recovering the signal from the 11 atoms. This phenomenon indicates that the matching pursuit method can act as a powerful filter for Gaussian noise removal like the one used in the present study. Therefore, this method has the potential of improving the quality of PCG's recorded in the hospital environments. Another important application of the matching pursuit method, which is the object of the companion paper [14], is its potential for the time-frequency scaling transformation of PCG's, since the Gabor wavelet provides good control of both the time and the frequency parameters of each atom. The accuracy of the decomposition of a PCG by the matching pursuit method is determined by the limiting iteration number or the energy threshold, and by the complexity of the signal structure. For the same energy threshold, complex signals need more iterations to reduce the reconstruction errors. A smaller energy threshold always yields better reconstruction results.

ACKNOWLEDGMENT

The authors would like to thank Dr. P. Pibarot for the auditory comparison of the original and reconstructed PCG's, and S. G. Mallat and Z. Zhang for providing the matching pursuit program package.

REFERENCES

[1] O. Bertrand, J. Bohorquez, and J. Pernier, "Time-frequency digital filtering based on an invertible wavelet transform: An application to evoked potentials," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 77–88, 1994.
[2] G. Cloutier, M. C. Grenier, R. Guardo, and L. G. Durand, "Spectral analysis of closing sounds produced by Ionescu-Shiley bioprosthetic aortic heart valves. Part 2: Computer simulation of aortic closing sounds and estimation of their truncation level and signal-to-noise ratio," *Med. Biol. Eng., Comput.*, vol. 25, pp. 492–496, 1987.

[3] L. G. Durand, Y. E. Langlois, T. Lanthier, R. Chiarella, P. Coppens, S. Carioto, and S. Bertrand-Bradley, "Spectral analysis and acoustic transmission of mitral and aortic valve closure sounds in dogs. Part I: Modeling the heart/thorax acoustic system," *Med. Biol. Eng., Comput.*, vol. 28, pp. 269–277, 1990.
[4] L. G. Durand and P. Pibarot, "Digital signal processing of the phonocardiogram: Review of the most recent advancements," *CRC Crit. Rev. Biomed. Eng.*, vol. 23, pp. 163–219, 1995.
[5] T. H. Joo, J. H. McClellan, R. A. Foale, G. S. Myers, and R. S. Lees, "Pole-zero modeling and classification of phonocardiograms," *IEEE Trans. Biomed. Eng.*, vol. BME-30, pp. 110–118, 1983.
[6] H. Köymen, B. K. Altay, and Y. Z. Ider, "A study of prosthetic heart valve sounds," *IEEE Trans. Biomed. Eng.*, vol. BME-34, pp. 853–863, 1987.
[7] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, 1993.
[8] S. Qian and D. Chen, "Signal representation using adaptive normalized Gaussian functions," *Signal Processing*, vol. 36, pp. 1–11, 1994.
[9] L. Senhadji, G. Carrault, J. J. Bellanger, and G. Passariello, "Comparing wavelet transforms for recognizing cardiac patterns," *IEEE Eng. Med., Biol. Mag.*, vol. 14, pp. 167–173, 1995.
[10] Y. Tang, C. Danmin, and L. G. Durand, "The synthesis of the aortic valve closure sound of the dog by the mean filter of forward and backward predictor," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 1–8, 1992.
[11] M. E. Tavel, "Cardiac auscultation. A glorious past-but does it have a future?" *Circ.*, vol. 93, pp. 1250–1253, 1996.
[12] A. G. Tilkian and M. B. Conover, "Understanding heart sounds and murmurs with an introduction to lung sounds." Philadelphia: Saunders, 1984.
[13] X. Zhang, "Analysis-synthesis and time-frequency scaling of phonocardiogram," masters thesis, McGill Univ., Montréal, P.Q., Canada, 1996, pp. 1–98.
[14] X. Zhang, L. G. Durand, L. Senhadji, H. C. Lee, and J. L. Coatrieux, "Time-frequency scale transformation of the phonocardiogram based on the matching pursuit method," this issue, pp. 972–979.