Time-frequency scaling transformation of the phonocardiogram based of the matching pursuit method.

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

Time-Frequency Scaling Transformation of the Phonocardiogram Based of 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*— A time-frequency scaling transformation based on the matching pursuit (MP) method is developed for the phonocardiogram (PCG). The MP method decomposes a signal into a series of time-frequency atoms by using an iterative process. The modification of the time scale of the PCG can be performed without perceptible change in its spectral characteristics. It is also possible to modify the frequency scale without changing the temporal properties. The technique has been tested on 11 PCG's containing heart sounds and different murmurs. A scaling/inverse-scaling procedure was used for quantitative evaluation of the scaling performance. Both the spectrogram and a MP-based Wigner distribution were used for visual comparison in the time-frequency domain. The results showed that the technique is suitable and effective for the time-frequency scale transformation of both the transient property of the heart sounds and the more complex random property of the murmurs. It is also shown that the effectiveness of the method is strongly related to the optimization of the parameters used for the decomposition of the signals.**

***Index Terms*— Heart murmurs, heart sounds, matching pursuit (MP) method, phonocardiogram, signal processing, time-frequency analysis, time-frequency scaling, wavelet transform.**

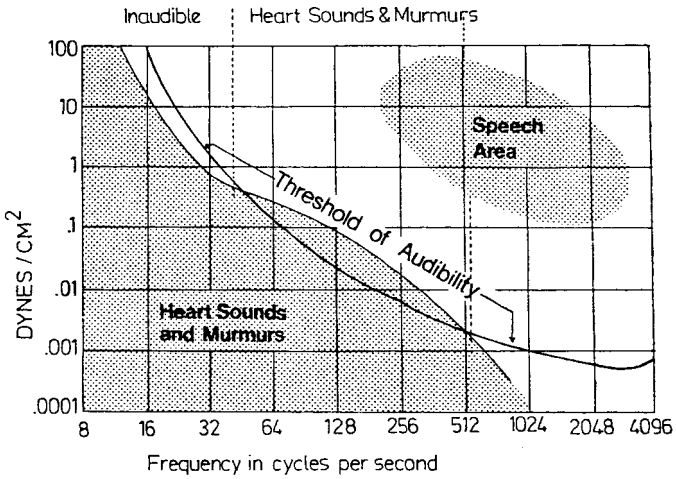

Fig. 1. Intensity of heart sounds and murmurs in correspondence with the threshold of audibility and speech (from [4]). An important part of the intensity and frequency distribution of the heart sounds and murmurs is out of the human hearing range.

I. Introduction

A HUMAN can hear sound vibrations from 16–18 000 Hz [2], [4], but this audible range is highly related to the sound intensity. Fig. 1 shows that the audible range of the heart sounds above the threshold of audibility is approximately from 40–500 Hz, which represents a very small audible area. The speech area is also displayed, indicating the optimal range for auditory acuity. This explains why physicians can sometimes detect atrial (S4) and ventricular filling (S3) sounds by palpation rather than by auscultation [1], [2]. The complex structure of the phonocardiogram (PCG) generates additional difficulties for auscultation. For instance, it is sometimes difficult to distinguish two close components because the time interval between them is too small. This difficulty increases when the heart rate increases. As an example, it is often difficult to recognize an opening snap when it is immediately followed by a long systolic murmur. This paper is a first attempt to remove these limitations of auscultation through time-frequency scaling of the PCG based on the matching pursuit (MP) method. The basic idea of time-scaling of the PCG is to expand the signal in time while keeping the same spectral properties. For frequency scaling, two aspects are interesting: 1) compress or expand the frequency band of the signal and 2) shift a frequency band up or down to a desired frequency range without changing the temporal properties of the signal. Finally, joint time-frequency scaling can be applied to a signal to modify both its time and frequency characteristics. This time-frequency scaling of the PCG could find interesting applications in the diagnosis of heart disease and the teaching of auscultation.

Time-frequency scaling techniques have been widely applied to the speech signal to compress its frequency bandwidth for transmission or to provide an aid for people with impaired hearing (shift the speech signal into an audible frequency range). A review of the literature shows that a variety of methods have been developed for time-frequency scaling of speech signals. These include: the sampling methods for time

X. Zhang is with the IRCM, Universit́ 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.

compression and expansion proposed by Lee [5], the time-domain harmonic scaling algorithms proposed by Malah [6], the scaling based on the short-time Fourier analysis and synthesis developed by Portnoff [8], and a method based on the sinusoidal representation of speech proposed by Quatieri and McAulay [10], [11]. The later method yields modifications of speech in both the time and the frequency domains with high quality. From the short-time spectrum of the signal, the technique extracts the frequency, amplitude, and phase of the sine waves, and the transformations are implemented by modifying these parameters before synthesis. However, this method was found to be ineffective for short complex transient sounds. Recently, a new approach for time-scaling of these sounds was proposed by Quatieri *et al*. [9], which uses subband signal representation and channel phase correction.

In phonocardiography, no time-frequency scaling method has been proposed so far, due to the complex structure of the PCG and the difficulty in finding an analysis-synthesis method suitable to both the heart sounds and the heart murmurs. In a companion paper [12], the MP method is described for the analysis and synthesis of PCG's. This technique was found very suitable to the properties of both the heart sounds and the heart murmurs. In this paper, we first introduce three types of transformation: time-scaling, frequency-scaling, and joint time-frequency scaling. In Section III, the PCG data base and a quantitative evaluation of the scaling transformations is described. Then, the scaling transformations are applied in Section IV to normal and pathological PCG's signals. The results are evaluated by quantitative analysis in the time and the time-frequency domains, visual comparison of time-frequency representations (TFR's), and auditory appreciation. The TFR's of the PCG's are computed with two methods: the spectrogram of the signal and the MP-based Wigner distribution as described in Section III. Finally, a discussion and a conclusion are presented in Section V.

II. Theory

A. Time-Scaling of PCG Signals

The MP method represents a signal $x(t)$ as a combination of an infinite number of time-frequency atoms [7]. It can be written as

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

with

$$h_i(t) = \beta_i \cdot g_i(t) \cdot u_i(t) \tag{2}$$

and

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

where a_i are the expansion coefficients. The parameters s_i (the scale factors) are used to control the width of the waveform envelope, and p_i are used to specify their temporal location. The parameters β_i are normalizing factors to keep the norm of $h_i(t)$ equal to one.

The purpose of time scaling the PCG is to change the rate of presentation while keeping the perceptual quality of the original signal. For a uniform change in the time scale, the time t of the original PCG is mapped into the transformed time scale t' through the mapping

$$t' = \gamma \cdot t. \tag{4}$$

In our application, time scale expansion is more useful than time scale compression. Thus, γ is always larger than one. In the MP method, the temporal properties of a time-frequency atom are related with the time-position p_i and the scale s_i which are modified for time-scaling of the PCG. For an input signal $x(t)$, the reconstructed signal $x'(t)$ by the MP method for m time-frequency atoms is given by

$$x'(t) = \sum_{i=0}^{m-1} a_i g_i\left(\frac{t - p_i}{s_i}\right) \cos(2\pi f_i t + \phi_i). \tag{5}$$

For time scaling, the time-position and the scale factors are modified to give

$$p_i' = \gamma \cdot p_i, \quad s_i' = \gamma \cdot s_i. \tag{6}$$

The module, frequency and phase are not changed, so that the time-scaled version is

$$x'(t') = \sum_{i=0}^{m-1} a_i g_i\left(\frac{t' - \gamma \cdot p_i}{\gamma \cdot s_i}\right) \cos(2\pi f_i t' + \phi_i). \tag{7}$$

B. Frequency-Scaling of PCG Signals

Frequency-scaling by the MP method is performed by scaling the frequency f of each time-frequency atom of the signal by using

$$f' = \xi \cdot f \tag{8}$$

where ξ is a scaling constant. Thus, the frequency-scaled signal can be expressed as

$$x'(t) = \sum_{i=0}^{m-1} a_i g_i\left(\frac{t - p_i}{s_i}\right) \cos(2\pi \cdot \xi \cdot f_i t + \phi_i). \tag{9}$$

C. Joint Time-Frequency Scaling of PCG Signals

Sometimes a joint time-frequency scaling is desired for changing both the time and the frequency properties of a signal. The transformation is the combination of the time-scaling and the frequency-scaling described above. Thus

$$x'(t') = \sum_{i=0}^{m-1} a_i g_i\left(\frac{t' - \gamma \cdot p_i}{\gamma \cdot s_i}\right) \cos(2\pi \cdot \xi \cdot f_i t' + \phi_i). \tag{10}$$

III. Materials and Method

The data base of 11 typical PCG's described in the companion paper [12] for the analysis and synthesis was also used for testing time-frequency scaling transformations.

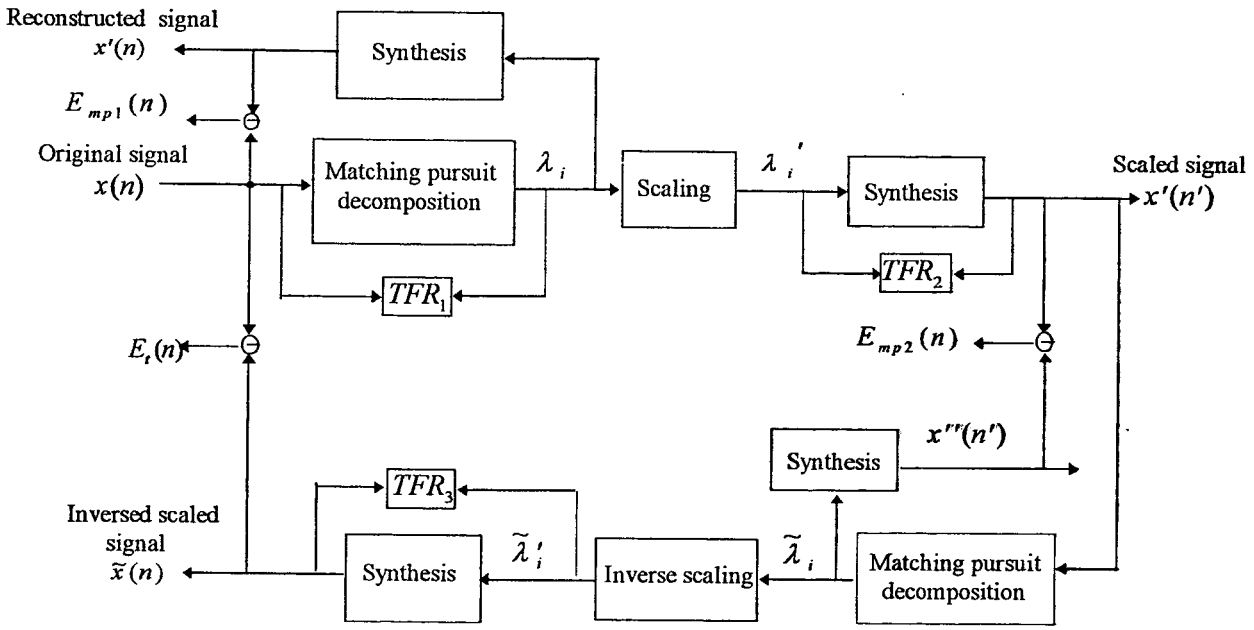

Fig. 2. The evaluation procedure for time-frequency scaling of PCG signals. A scaling/inverse scaling procedure was used to evaluate the performance of the MP method for the time-frequency scaling of the PCG.

A. Evaluation of the Scaling Transformation Process

Fig. 2 illustrates the procedure used for evaluating the scaling transformation process of a PCG signal $x(n)$. The MP method decomposes $x(n)$ into a number of time-frequency atoms represented by a set of parameters λ_i. The summation of these time-frequency atoms provides a reconstructed version $x'(n)$, with an error $E_{mp1}(n)$. The evaluation of this step was discussed in details in [12]. Scaling the parameters of the atoms in time and/or frequency and the synthesis yields the transformed signal $x'(n')$, which could be more useful than the original signal.

After transformation, the temporal and/or the frequency properties of the signal are changed. Therefore, the quantitative evaluation of the scaling process based on the comparison of the scaled signal and the original signal is difficult. To facilitate comparison, an inverse scaling of the scaled signal is introduced to transform it back to a version of the original signal. For instance, if we time scale $x(n)$ by a factor of 2.0, an inverse scaling of the resulting $x'(n')$ by a factor 0.5 will transform it back to a signal $\tilde{x}(n)$. Ideally, this transformed signal should be identical to the original signal, if no distortion is introduced in both the MP decomposition and the time scaling procedure. The TFR can be calculated for the four temporal signals: the original signal $x(n)$, the reconstructed signal $x'(n)$, the time- (or/and frequency-) scaled signal $x'(n')$, and the corresponding inverse scaled signal $\tilde{x}'(n)$. By visually comparing the TFR's of these signals, we can qualitatively appreciate the effects of the scaling transformations. Furthermore, this is compatible with the human peripheral auditory system which also performs a sort of time-frequency analysis of the signal.

As shown in Fig. 2, the original signal goes through twice the MP decomposition, the scaling process, and the synthesis process during the scaling/inverse scaling process. Assuming that the synthesis process generates negligible error compared with decomposition and scaling, the first MP decomposition generates the error $E_{mp1}(n)$, which is the residue energy between the original signal $x(n)$ and the reconstructed signal $x'(n)$. The second MP decomposition generates the residue energy error $E_{mp2}(n)$ between the scaled signal $x'(n')$ and the reconstructed scaled signal $x''(n')$. It is possible to isolate the scaling errors by minimizing $E_{mp1}(n)$ and $E_{mp2}(n)$ using a small value of the energy threshold ε^2 to stop the MP iterative process and by computing the total error $E_t(n)$ between $x(n)$ and $\tilde{x}(n)$. If we also assume that the two scaling processes contribute equal error, then the scaling error can be estimated as

$$E_s = \frac{E_t - E_{mp1} - E_{mp2}}{2}. \tag{11}$$

The normalized root-mean-square error (NRMSE) was used for calculating E_{mp1}, E_{mp2}, and E_t by using

$$\text{NRMSE} = 100 * \sqrt{\frac{\sum_{i=0}^{N-1} e^2(n)}{\sum_{i=0}^{N-1} x^2(n)}} \tag{12}$$

where N is the number of samples of $x(n)$ and $e(n)$ is the difference between the original and the reconstructed signals.

Mallat and Zhang [7] proposed to use the sum of the Wigner distributions of all the individual atoms composing a signal to represent its energy distribution in the time-frequency plane. This approach, which automatically removes the cross-terms between any two Wigner distributions of time-frequency atoms because each atom is a monocomponent, is called the MP-based Wigner distribution in the present paper to distinguish it from the conventional Wigner distribution applied to the signal

itself. The MP-based Wigner distribution of the function $x(n)$ in (1) was, thus, represented as

$$X(n,\omega) = \sum_{i=1}^{m} a_i^2 \cdot Wh_i\left(\frac{n-p_i}{s_i}, s_i(\omega - 2\pi \cdot f_i)\right) \quad (13)$$

where a_i, s_i, p_i, f_i are the parameters of the ith time-frequency atom and

$$Wh(n,\omega) = 2e^{-2\pi\left(n^2+(\omega/2\pi)^2\right)} \quad (14)$$

is the Wigner distribution of the time-frequency atom $h(n)$ if $g_i(t) = 2^{1/4}e^{-\pi\cdot(t-p_i/s_i)^2}$.

The spectrogram of the function $x(n)$ was calculated by

$$X(n,\omega) = \sum_{m=-L/2}^{L/2} x(n+m)w(m)e^{-iwm} \quad (15)$$

where $w(m)$ is a Hanning window. The window length is $L+1$. Both the MP-based Wigner distribution and the spectrogram were used for the visual appreciation of the TFR's of the original signals and that of the inverse scaled signals.

B. Optimization of Parameters

In the companion paper [12], the effect of J on the decomposition of PCG's was studied, and it was found that a good analysis and synthesis performance can be obtained for $J = 7$. In order to determine the sensitivity of E_{mp1}, E_{mp2}, and E_s as a function of the number of atoms used in the scaling process, the energy threshold used to stop the MP iterative process was varied from $\varepsilon^2 = 10^{-6}$ to 10^{-4} for a J value of seven. In this test, a time-scaling factor $\gamma = 2$ was used. In a second test, we optimized the scaling process by selecting the values of parameters M and J minimizing E_t. To force the limiting number of time-frequency atoms M to be the stopping criteria for the decomposition process (thus, $m = M$), the energy threshold ε^2 was set to zero. Parameter M was then varied over a range of 50–250 with steps of 50 for $J = 6, 7$, and 8.

IV. Results

A. Time-Scaling of the PCG's

Table I shows that the number of atoms required to correctly represent a PCG signal vary significantly depending on the type of heart pathology. It also demonstrates that E_s did not change when the threshold criteria ε^2 was increased from 10^{-6} to 10^{-4}, even if the total number of atoms M was highly reduced (from two up to more than ten times). These results clearly demonstrate that E_s can be isolated from the scaling/inverse scaling process, since E_{mp1} and E_{mp2} are negligible compared to E_s. The maximal correlation levels between $E_{mp1}(n)$ and $E_t(n)$ and between $E_{mp1}(n)$ and $x(n)$, for each PCG signal and for $J = 7$, were always <0.081 for $\varepsilon^2 \leq 10^{-5}$ and <0.21 for $\varepsilon^2 = 10^{-4}$. According to this test, it appears that $E_{mp1}(n)$ and $E_t(n)$ were not correlated when $\varepsilon^2 \leq 10^{-5}$. When $\varepsilon^2 = 10^{-4}$, these error signals were weakly correlated for Z1–Z4 and not correlated for the other signals. A similar test was also performed between $E_{mp1}(n)$ and the inverse scaled representation of the $E_{mp2}(n)$ signal for $\varepsilon^2 = 10^{-5}$. The maximal correlation levels were generally low for all 11 signals (between 0.05 and 0.18).

TABLE I
The Scaling Errors E_s (in %) and the Atom Numbers M (in Parenthesis) of the 11 PCG's as a Function of the Energy Threshold Level ε^2 for Time-Scaling by a Factor of $\gamma = 2$. The Corresponding J Value and E_{mp} Error ($E_{mp} \cong E_{mp1} \cong E_{mp2}$) Are Given at the Top of the Table

Signal	$\varepsilon^2 = 10^{-6}$ J = 7 $E_{mp} \cong 0.10\%$	$\varepsilon^2 = 10^{-4}$ J = 7 $E_{mp} \cong 1.0\%$	$\varepsilon^2 = 10^{-4}$ J = 8 $E_{mp} \cong 1.0\%$
Z1	12.8 (468)	12.4 (38)	13.0 (40)
Z2	14.5 (557)	14.0 (51)	15.2 (51)
Z3	9.5 (495)	9.9 (54)	11.6 (46)
Z4	16.9 (712)	16.5 (52)	11.7 (63)
Z5	21.7 (1175)	21.3 (313)	13.4 (317)
Z6	15.6 (1097)	14.9 (250)	19.4 (248)
Z7	18.1 (1018)	17.2 (257)	37.5 (238)
Z8	8.9 (828)	8.6 (134)	25.4 (125)
Z9	15.4 (949)	12.3 (208)	11.2 (196)
Z10	15.8 (1078)	15.2 (114)	7.2 (87)
Z11	25.3 (2000)	25.0 (854)	32.6 (813)
Mean	15.9 (943)	15.2 (211)	18.0 (202)

As shown in Table I, a test was also performed for $J = 8$ and $\varepsilon^2 = 10^{-4}$. The results shows that E_s is very sensitive to the largest octave value J, since it is significantly modified for some signals (Z4–Z8, Z10, and Z11). Consequently, in the second test, the values of J and M were varied to minimize the total error between $x(n)$ and $\tilde{x}(n)$. The results are shown in Fig. 3. Fig. 3(a)–(c) gives the average NRMSE's of the reconstruction errors $(E_{mp1} + E_{mp2})$, the time-scaling error E_s, and the total error E_t of the 11 PCG's for different values of M (50, 100, 150, 200, and 250) when $J = 6$, 7, and 8, respectively, and decrease as M increases for all the cases, and $J = 8$ and 7 give similar errors which are lower than those of $J = 6$. On the contrary, E_s increases with M, and for $J = 7$ it is lower than for $J = 8$ and 6. For a given J, E_t of the whole process does not vary significantly as a function of M. It is clear that $J = 7$ gives the minimum E_t for all the atom numbers. The best time-scaling performance was obtained when $J = 7$ and $M = 150$.

Fig. 4(a) shows an example of the original PCG of aortic regurgitation, and Fig. 4(b) shows its time-scaled version by a factor $\gamma = 2$ in two sections. We can see that both the waveform and the envelope of the signal are preserved after the time-scale transformation. The corresponding MP-based Wigner distributions and spectrograms of the PCG's of Fig. 4(a) and (b) are shown in Fig. 6(a1) and (b1) and Fig. 6(a2) and (b2), respectively. The spectrograms were computed by using a Hanning window of 66.7 ms with a window shift of 1.33 ms (2.76 ms for the scaled PCG). The duration of the discrete Fourier transform was the same as that of the Hanning window. It can be noticed that the time and frequency resolutions of the MP-based Wigner distributions

TABLE II
THE MINIMUM TOTAL ERRORS (IN %) OF THE 11 PCG'S WITH THE CORRESPONDING J VALUE AND NUMBER OF ATOMS M FOR TIME-SCALING BY A FACTOR OF $\gamma = 2$

Signal	NRMSE of E_{mp1} (%)	NRMSE of E_{mp2} (%)	NRMSE of E_s (%)	NRMSE of E_t (%)	J	m
Z1	0.25	0.23	12.47	25.43	7	200
Z2	0.37	0.58	12.07	25.09	8	150
Z3	0.46	0.62	8.89	18.86	7	100
Z4	0.38	0.38	10.69	22.15	8	150
Z5	1.51	2.14	12.10	27.85	8	150
Z6	0.99	1.69	14.05	30.78	7	250
Z7	1.05	2.05	16.17	35.44	7	250
Z8	1.69	1.75	6.76	16.96	7	100
Z9	0.95	1.45	10.52	23.44	8	200
Z10	0.93	0.50	7.03	15.49	8	100
Z11	7.45	6.64	14.94	43.96	7	250
Mean	1.46	1.64	11.43	25.95		

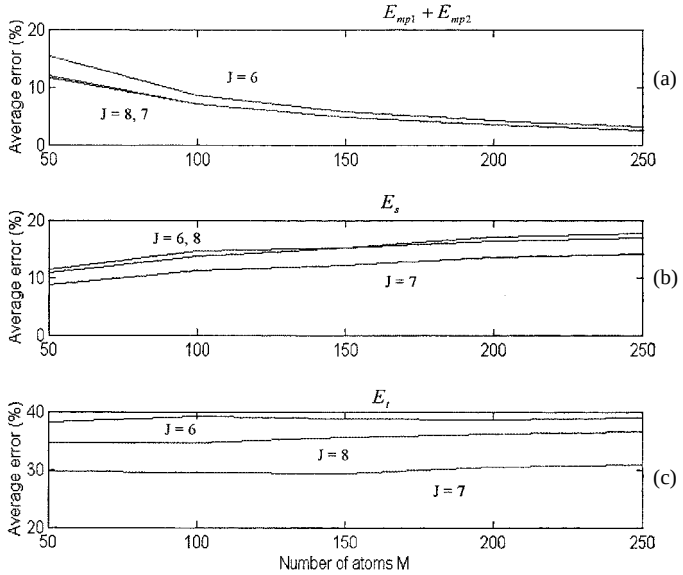

Fig. 3. Average errors of the time-scaling transformation as a function of the number of atoms M and the maximum octave value $J = 6, 7$, and 8. (a) The reconstruction error $(E_{mp1} + E_{mp2})$. (b) The time scaling error (E_s). (c) The total error E_t. The best value of J which is seven is relatively independent of the value of M.

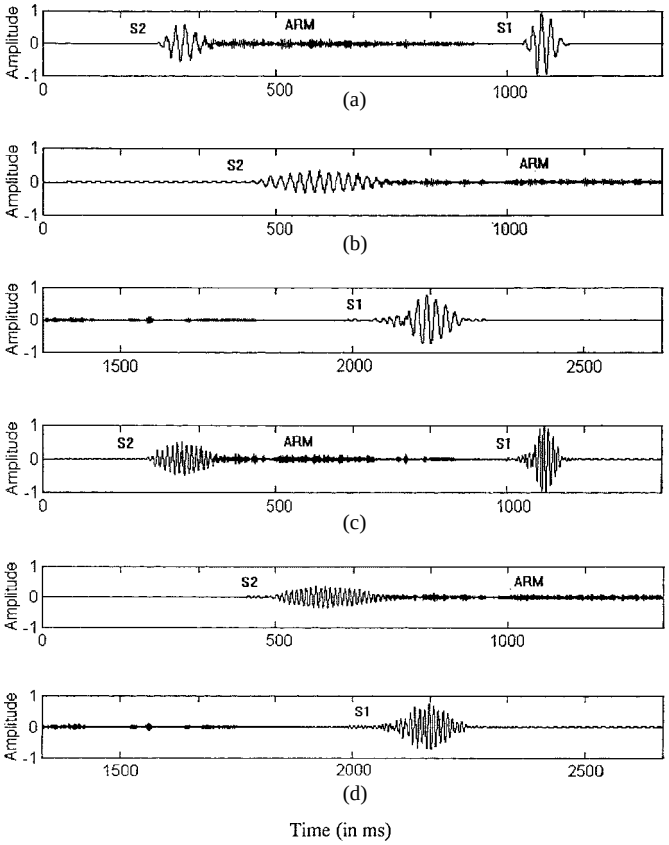

Fig. 4. (a) The original PCG of aortic regurgitation. (b) The time scaled signal of (a) by factor $\gamma = 2$. (c) The frequency scaled signal of (a) by factor $\xi = 2$. (d) The time-frequency scaled signal of (a) by factors of $\gamma = 2$ and $\xi = 2$. This figure clearly shows the effectiveness of the method for time-frequency scaling of the PCG.

are much higher than those of the spectrograms. The MP-based Wigner distributions (which has no cross-term) give a clearer presentation of the PCG while the spectrograms provide a blurred TFR of the corresponding signal. For both TFR techniques, the frequency resolution of the time scaled PCG is better than that of the original PCG. The first and the second heart sounds have a duration of about 100 ms and a frequency bandwidth below 180 Hz. In the spectrogram, each sound appears to have two spectral components while in the MP-based Wigner distribution more components can be distinguished. Both TFR's show that the murmur of aortic regurgitation begins almost with the second heart sound and ends approximately 130 ms before the onset of the first heart sound. Its frequency bandwidth is between 180–400 Hz. The granularity of the murmur is clearly seen on the MP-based Wigner distributions.

Table II gives, for a time scaling transformation by a factor of 2.0, parameters J and M, and the minimum total error for each of the 11 PCG's. These values were selected such that the maximum value of E_{mp1} or E_{mp2} was $<2.2\%$, except for

Z11. The mean value of E_t is 25.95% which is about 3.5% less than the result obtained by using $J = 7$ and $M = 150$ for all PCG's [the minimum point of Fig. 3(c)].

B. Frequency-Scaling

The total errors between the original PCG's and the inverse scaled PCG's after frequency-scaling turned out to be almost the same as those of the time-scaling when the scaling factors were identical. This phenomenon is due to the scaling operation of the MP. For instance, (16) is for time-scaling with $t' = \gamma \cdot t, p_i' = \gamma \cdot p_i, s_i' = \gamma \cdot s_i$, while (17) describes frequency-scaling by a factor ξ

$$x'(\gamma \cdot t) = \sum_{i=0}^{m-1} a_i g_i\left(\frac{\gamma \cdot t - \gamma \cdot p_i}{\gamma \cdot s_i}\right) \cos\left(2\pi \cdot f_i \cdot \gamma \cdot t + \phi_i\right) \tag{16}$$

$$x'(t) = \sum_{i=0}^{m-1} a_i g_i\left(\frac{t - p_i}{s_i}\right) \cos\left(2\pi \cdot \xi \cdot f_i t + \phi_i\right). \tag{17}$$

When $\gamma = \xi$, each time scaled atom has exactly the same number of cycles of oscillation as the frequency scaled atom but with a longer duration (γ times). For instance, if we take one sample for every two in the time scaled signal of Fig. 4(b), essentially, the frequency scaled signal shown in Fig. 4(c) is obtained. Fig. 6(a3) and (b3) shows the MP-based Wigner distributions and spectrograms of Fig. 4(c), respectively. We can see that the corresponding atoms are shifted up in frequency by a factor of $\xi = 2$.

C. Joint Time-Frequency Scaling

The 11 PCG's were also used to evaluate the joint time-frequency scaling technique, with the octave values $J = 6$, 7, and 8, and the number of atoms $M = 50$, 100, 150, 200, and 250. Fig. 7 shows the NRMSE's for the different values of M and J. The results are similar to those of Fig. 3, but the total error E_t is always lower in Fig. 7 than in Fig. 3, and $J = 8$ gives the minimum error. From the waveform of the joint time-frequency scaled signal shown in Fig. 4(d), it can be seen that both the duration and the oscillations of each sound component have been doubled while the envelope of the PCG is similar to that of the other PCG's. Fig. 6(a4) and (b4) shows the TFR's of the joint time-frequency scaled PCG of Fig. 4(d). Because of this double scaling transformation, each time-frequency atom is better separated from the other ones. The improvement of the time and frequency resolutions is more striking for the spectrogram.

V. Discussion and Conclusion

The results of this study show that the performance of the time-frequency scale transformations is sensitive to the variation of the maximum octave value J and the number of atoms M used in the decomposition process. We also found that the values of the parameters giving the best performance for the time scale transformation for each PCG varied significantly, and no common rules could be obtained to define fixed parameters for different PCG's.

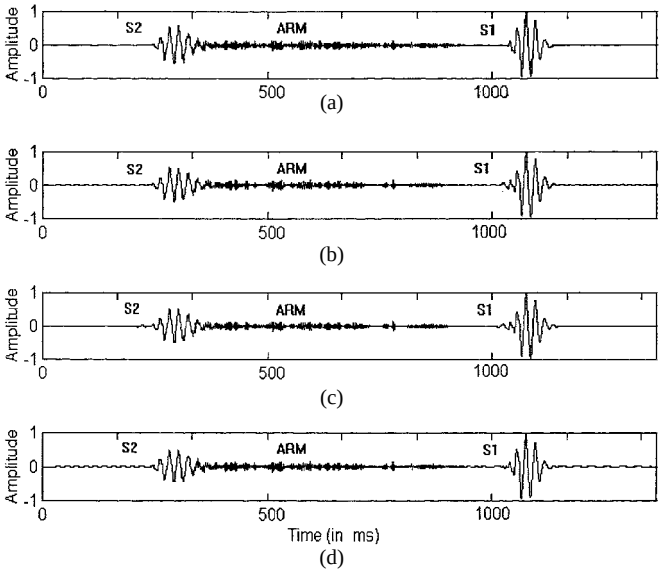

Fig. 5. (a) The original PCG of aortic regurgitation. (b) The inverse time scaled signal of Fig. 4(b) by factor $\gamma = 0.5$. (c) The inverse frequency scaled signal of Fig. 4(c) by factor $\xi = 0.5$. (d) The inverse time-frequency scaled signal of (a) by factors of $\gamma = 0.5$ and $\xi = 0.5$. The inverse-scaled PCG signals appear to be very similar to the original PCG signal.

In Fig. 4, the maximal amplitudes of the time scaled and the joint time-frequency scaled PCG's are slightly lower than those of the original PCG and of the frequency scaled PCG. This is due to the norm of $hi(t)$ which must be equal to one. Consequently, increasing the window scale factor "s" results in a decrease of the window amplitude. The inverse scaling has the opposite effect, as shown in Fig. 5(b)–(d) where the inverse scaled versions of the PCG's of Fig. 4(b)–(d) are presented.

In Table II, the large and errors appearing for Z11 are due to the high complexity and random structure of the pericardial rub signal. When the number of atom was increased above 250 to reduce the E_{mp} errors to less than 2.5%, the scaling error increased rapidly. For $M = 750$, the errors were $E_{mp1} = 1.29\%, E_{mp2} = 1.95\%, E_s = 23.81\%$, and $E_t = 50.86\%$. Comparing the errors of the time-scaling and the joint time-frequency scaling of Figs. 3 and 7, it is observed that the errors caused by the scaling transformation are smaller for the joint time-frequency scaling of all 11 PCG's. This was not expected because the joint time-frequency scaling involves the modification of three parameters instead of two for the time-scaling process. In an attempt to explain this phenomenon, we examined the phase variation between the decomposed atoms of the original PCG's and the corresponding time scaled PCG's or joint time-frequency scaled PCG's. We found that after joint time-frequency scaling, the phases of the atoms were better preserved than those of the time scaled PCG's. The increased time-frequency resolution of the joint time-frequency scaling could also be responsible for this effect. Further studies are required to elucidate this finding.

The time-frequency shifted versions of the 11 PCG signals have been informally evaluated through listening by an experienced cardiologist. He found that the time-scaling of the PCG signals kept the sound quality of the original signals while having a slower tempo. We believe that the time-frequency scaling of the PCG may find important applications

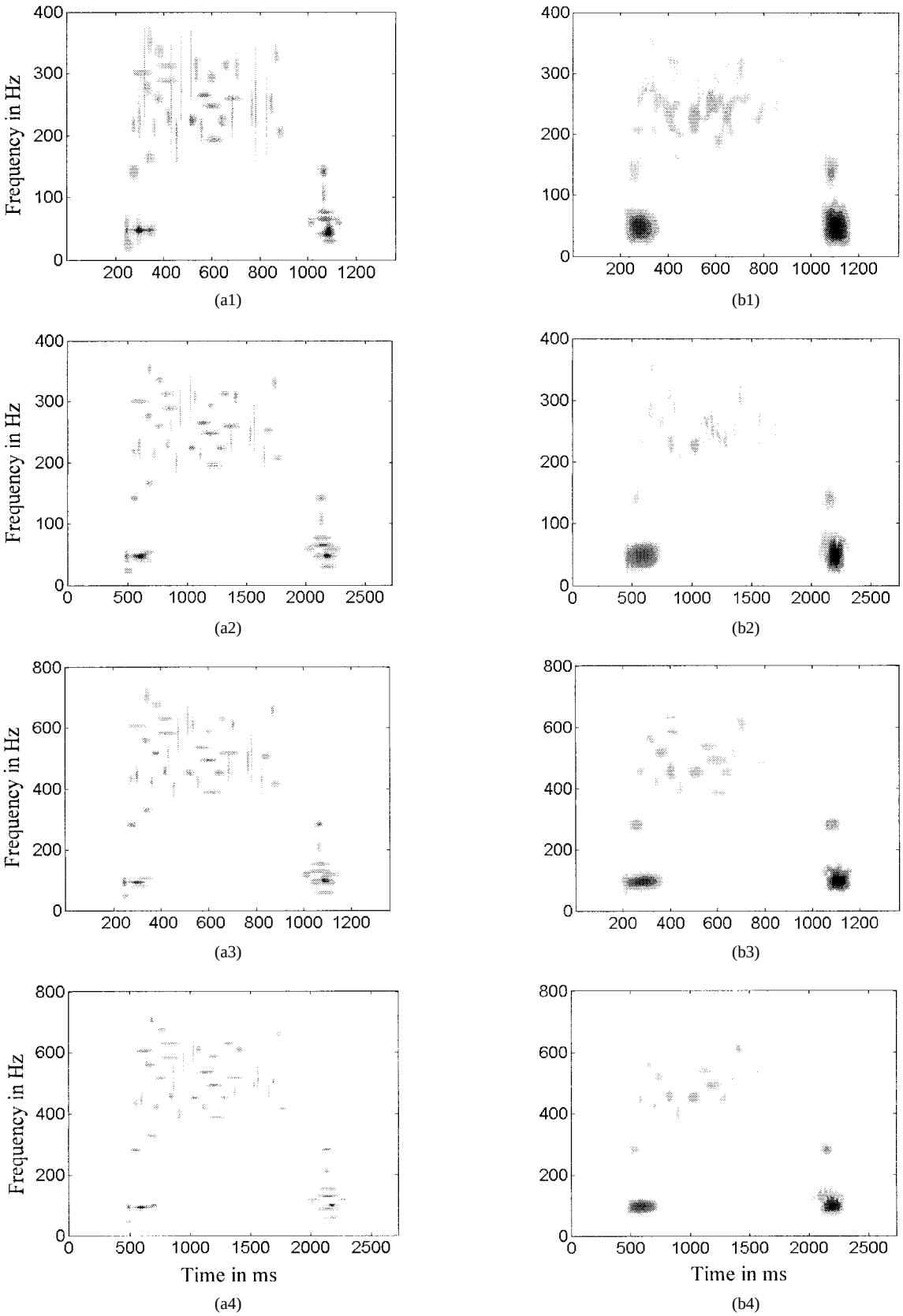

Fig. 6. The MP-based Wigner-distributions (left panels) and spectrograms (right panels) of (a1), (b1) the original signal of the pathological PCG of Fig. 5, (a2), (b2) the time scaled signal by factor of $\gamma = 2$, (a3), (b3) the frequency scaled signal by factor of $\xi = 2$, and (a4), (b4) the time-frequency scaled signal by factors of $\gamma = 2$ and $\xi = 2$. This figure qualitatively shows that the structure of the original PCG signal is preserved after time, frequency, and both time and frequency scaling of the PCG signal.

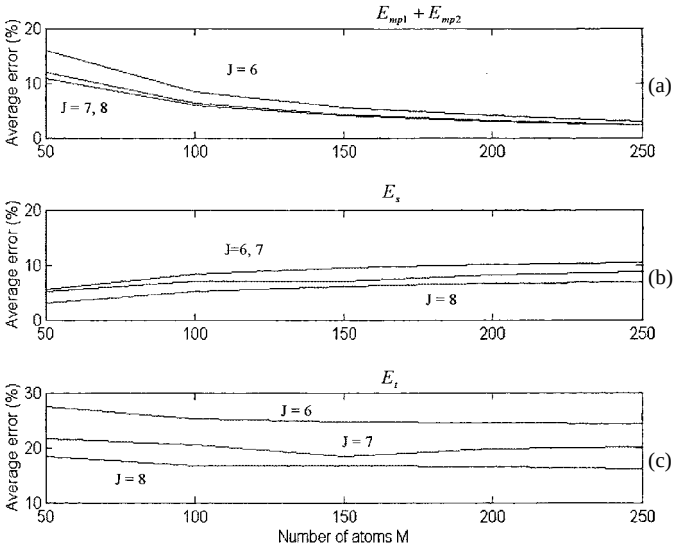

Fig. 7. Average errors of the joint time-frequency scaling transformation as a function of the number of atoms M and the maximum octave value $J = 6$, 7, and 8. (a) The reconstruction error $(E_{mp1} + E_{mp2})$. (b) The time-frequency scaling error (E_s). (c) The total error E_t. The best value of J, which is eight, is relatively independent of the value of M.

in the improvement of 1) the teaching of auscultation and 2) the diagnosis of heart and heart valve disease. An example of the first application is the following: A student is using a computer having a data base of normal and pathological PCG's for learning of cardiac auscultation. The complex cases of auscultation are known and the scaling procedures (time, frequency, or both) improving the detection and identification of sound components difficult to distinguish by auscultation are available as additional files. The student can listen alternatively and repeatedly to these files using an electronic stethoscope or a set of headphones, and he can also make visual analyses of the corresponding PCG signals and TFR's on the computer video monitor. This allows the student to make all the necessary auditory and visual comparisons to better detect and identify the difficult component(s) in the original PCG signal. As an example of the second application, a physician is using a state-of-the-art electronic stethoscope comprising a digital signal processor able to perform the time scaling of the PCG signal with a variable time expansion factor or the real-time frequency compression of mechanical valve clicks to provide a better diagnosis of the status of the heart and the heart valves [3].

In summary, we have developed a time-frequency scaling technique for the PCG based on the MP method. Although some distortion of the signal structure cannot be avoided by the scaling transformations, the results of our study show that this approach is very suitable for auditory perception. The scaling of the heart sound signals is easier than that of the murmur signals, due to their relatively simple structure. The improvement of the scaling transformation for both the heart sound and the murmur signals, however, will require further studies.

ACKNOWLEDGMENT

The authors would like to thank Dr. P. Pibarot for the auditory comparison of the original and scaled PCG's, and S. G. Mallat and Z. Zhang for providing the MP program package. They would also like to thank Dr. H. Sava for his helpful advises during the review process of this paper and D. Chen and Z. Qin for performing the correlation studies between the PCG signals and the error signals.

REFERENCES

[1] J. Abrams, "Current concepts of the genesis of heart sounds II. Third and fourth sounds," *JAMA*, vol. 239, pp. 2029–2030, 1978.
[2] ——, "Cardiac auscultation: The Stethoscope and Cardiac Sound," in *Essentials of Cardiac Physical Diagnosis*, J. Abrams, Ed. Philadelphia, PA: Lea & Febiger, 1987.
[3] 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.
[4] A. Leatham, *Auscultation of the Heart and Phonocardiography*, 2nd ed. London, U.K.: Churchill Livingstone, 1975.
[5] F. F. Lee. "Time-compression and expansion of speech by the sampling method," in *Speech Enhancement*, J. S. Lim, Ed. Cambridge, MA: Massachusetts Inst. Technol., 1983, pp. 286–290.
[6] D. Malah, "Time-domain algorithms for harmonic bandwidth reduction and time scaling of speech signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 121–133, 1979.
[7] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397–3415, 1993.
[8] M. R. Portnoff, "Time-scale modification of speech based on short-time Fourier analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 374–390, 1981.
[9] T. F. Quatieri, R. B. Dunn, and T. E. Hanna, "A subband approach to time-scale expansion of complex acoustic signal," *IEEE Trans. Speech Audio Processing*, vol. 3, pp. 515–519, 1995.
[10] T. F. Quatieri and R. J. McAulay, "Shape invariant time-scale and pitch modification of speech," *IEEE Trans. Signal Processing*, vol. 40, pp. 497–510, 1992.
[11] ——, "Speech transformations based on a sinusoidal representation," *IEEE Trans. Acoustics Speech Signal Processing*, vol. 34, pp. 1449–1464, 1996.
[12] X. Zhang, L. G. Durand, L. Senhadji, H. C. Lee, and J. L. Coatrieux, "Analysis-synthesis of the phonocardiogram based on the matching pursuit method," this issue, pp. 962–971.