Time-Frequency Representations for Continuous Adventitious Lung Sounds

B.A. Reyes[1], S. Charleston-Villalobos[1], T. Aljama-Corrales[1], and R. González-Camarena[2]

[1] Univ. Autónoma Metropolitana, Electrical Engineering Department, Mexico City, México,schv@xanum.uam.mx
[2] Universidad Autónoma Metropolitana, Health Science Department, Mexico City, México

Abstract—**Four different time-frequency representation (TFR) techniques were tested on synthetic wheezes to determine the best one in order to analyze the time-frequency course of this kind of continuous adventitious lung sounds (LS). The best TFR may be help to gain a deeper understanding of the genesis of the wheeze and its relation with lung diseases. In this study, the Spectrogram was included as the classical analysis tool in the field but it has disadvantages when working with nonstationary signals. We also include the Reassigned Spectrogram, the TFR obtained via the Time-Varying Autoregressive Modeling (TVAR), and the more recently developed the Hilbert-Huang Spectrum based on Empirical Mode Decomposition (EMD). Since the theoretical TFR of a synthetic wheeze is known beforehand, performance measurements can be obtained and used to select the appropriate TFR. Performance indexes are based on both the TFR image as well as signal approaches. According to the performance indexes, the Hilbert-Huang Spectrum was the best TFR with ρ equals to 0.9247, ρ_{mean} of 0.9521, NRMSE of 0.0601 and res_{TF} of $2.47x10^{-6}$. Finally, the Hilbert-Huang Spectrum (HHS) was applied to real wheezes acquired from diffuse interstitial pneumonia patients. TFR with both HHS and Spectrogram were contrasted to point out the meaningful differences.**

Keywords— **Time-frequency analysis, empirical mode decomposition, respiratory sounds, adventitious lung sounds, wheezes.**

I. Introduction

Continuous adventitious lung sounds considered in this work are the so-called wheezes. According to the Computarized Respiratory Sound Analysis (CORSA) guidelines [1], a wheeze is acoustically characterized by a periodic waveform with a dominant frequency commonly higher than 100 Hz with duration greater than 100 ms and they can be monophonic or polyphonic. Physiopathological mechanisms that explain the genesis of wheezes are not well understood. In line with the fluid dynamic flutter theory [2], whezees are produced by airway walls vibration and gas airflow. Walls oscillations start when gas velocity inside airways reaches a critical value given by dimensions and mechanical properties of airway walls and mechanical properties of the gas.

In time domain, the Time Expanded Waveform Analysis (TEWA) of wheezes has shown that they possesses sinusoidal deflections and continuous undulations [3]. While in frequency domain, description of this sound is classically done by computing their spectrums with a FFT algorithm. Additionally, autoregressive modeling (AR) has also been used.

The first time-frequency representation (TFR) of wheezes was done using the sound spectrograph [4]. Later, sucesive spectrums obtained via the FFT where plotted showing that this method is useful to follow changes in wheezes parameters as frequency and duration during medical treatment [5]. Since then, TFRs have also been used for automated analysis; the spectrogram [7] is included as the classical tool in the field of respiratory sounds, as well as others as AR modeling [6] and the continuous *wavelet* transform [8]. Although the importance of TFRs of wheezes, there have not been questioning about if the classical spectrogram is the best TF technique for analyzing this adventitious lung sound or if another more recent developed technique performs better allowing to extract information more accurately and even, providing better understanding of its genesis. According to the former idea, we propose to compare the performance of four time-frequency representations (TFR) techniques, including the classical tool, with the help of synthetic sounds. Later, the best TFR is applied to real acquire wheezes to contrast the results and point out meaningful differences.

II. Time-frequency representations techniques

The existence in nature of signals with a time-varying spectrum, leads to the requirement of a function that allows us to determine which frequencies of the signal are present at certain time, such signal-dependent function is a time-frequency representation.

A. The Spectrogram

The spectrogram of a signal $s(t)$ is given by the square magnitude of its Short-Time Fourier Transform as [9]

$$SP(s;t,\omega)=\left|\frac{1}{\sqrt{2\pi}}\int s(\tau)h(\tau-t)e^{-j\omega\tau}d\tau\right|^2 \qquad (1)$$

where $h(t)$ is a temporal window used to enhance the signal information around time t. However, $h(t)$ implies a tradeoff between localization in time and frequency, dependency that is supported by the uncertainty principle. Temporal

J. Folgueras Méndez et al. (Eds.): CLAIB 2011, IFMBE Proceedings 33, pp. 1086–1089, 2013.
www.springerlink.com

windows used for the spectrogram computation are the same for harmonic analysis [10].

B. TVAR Modeling and Instantaneous Power Spectrum.

For a discrete stochastic process $\{s(n)\}$ the TVAR model is given by

$$s(n) = -\sum_{k=1}^{M} a_k(n)s(n-k) + v(n) \quad (2)$$

where M is the model order, the set $\{a_k(n)\}_{k=1,\ldots,M}$ are the TVAR coefficients at time index n, $\{s(n\text{-}k)\}_{k=1,\ldots,M}$ represent past samples of $s(n)$, and $\{v(n)\}$ is a white noise process. Among the several procedures to estimate the AR coefficients, the Burg algorithm is one of the most used. This method is based on the minimization of the sum of the forward and backward errors of the lineal prediction filters, subject to the constraint that the AR coefficients satisfy the Levinson-Durbin recursion to ensure a stable system [11]. After estimating the TVAR parameters for all signal extend, the TFR can be computed by

$$P_{Burg}(s;n,\omega) = \frac{\sigma^2(n)}{\left|1+\sum_{k=1}^{M}\hat{a}_k(n)e^{-j\omega k}\right|^2} \quad (3)$$

where σ^2 represents the variance of model error at time n.

C. The Reassigned Spectrogram.

The reassignment method was introduced to increment the energy concentration of signal's components in the spectrogram and later on, it was reintroduced and extended to more distributions [12]. One can express the SP of a signal $s(t)$ as the 2D convolution between its Wigner-Ville distribution (WVD) and the WVD of a temporal window $h(t)$. Here, the WVD of $h(t)$ delimits a time-frequency domain around (t,ω), where a weighted average of the values of the WVD of $s(t)$ is obtained. In the reassignment method, each value obtained for (t,ω) is moved to another point (t',ω'), the center of gravity of this domain, which is more representative of the signal's energy location in the neighborhood of (t,ω) to produce the reassigned spectrogram

$$SPR(s;t',\omega) = \iint SP(s;t,\omega)\delta\left(t'-\hat{t}(s;t,\omega)\right)\delta\left(\omega-\hat{\omega}(s;t,\omega)\right) \quad (4)$$

whose value at any point (t',ω') is the sum of all the spectrogram values reassigned to this point.

D. The Hilbert-Huang Spectrum.

Huang et. al. developed the empirical mode decomposition (EMD), a technique for processing nonstationary and nonlinear signals [13]. This decomposition is based on identify and extract the intrinsic oscillatory modes of the signal $s(t)$, the so-called IMFs, through its characteristics time scales and directly from the original data via a sifting process. After this sifting process, $s(t)$ can be represented as

$$s(t) = \sum_{k=1}^{N} IMF_k(t) + r_N(t) \quad (5)$$

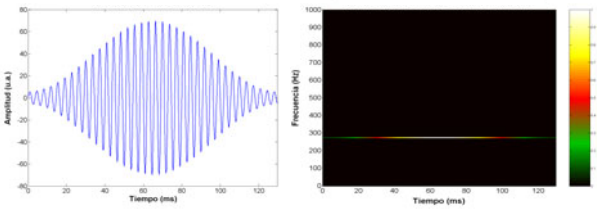

Fig. 1 Synthetic continuous adventitious lung sound. a) Wheeze time waveform. b) Ideal TFR, the number of frequency bins was fixed at 512

where N represents the number of extracted IMFs, and $r_N(t)$ can be interpreted as the residue component. Due to local conditions imposed on each IMF, it is possible to obtain without ambiguity its associated analytical or complex signal $z(t)$, and then their amplitude and phase by

$$z(t) = IMF(t) + jH\{IMF(t)\} = A(t)e^{j\varphi(t)} \quad (6)$$

where $H\{\bullet\}$ corresponds to the Hilbert transform, $A(t)$ is the instantaneous amplitude, and $\varphi(t)$ is the instantaneous phase of each oscillatory mode. The instantaneous frequency $f_i(t)$ can be obtained using Ville's definition, consequently, the signal $s(t)$ can be represented by

$$s(t) = \sum_{k=1}^{N} A_k(t)e^{j2\pi\int f_{i,k}(t)dt}. \quad (7)$$

Equation (7) can be interpreted as a generalization of the Fourier expansion of $s(t)$. Now, amplitude and frequency are time-dependent functions, allowing to obtain a TFR called the Hilbert-Huang Spectrum denoted as HHS.

III. Methodology

A. Synthetic Wheezes

Wheezing sound simulation was done by generating a waveform similar to those found in non-healthy subjects and that satisfy CORSA standards. We simulated a monophonic wheeze with a central frequency of 275 Hz and duration of 130 ms by

$$y(t) = A(t)\sin\left(2\pi f_0 t\right) \quad (8)$$

where $A(t)$ is the amplitude modulation function, and f_0 is the frequency of the pitch. Different authors have used a Hamming function as the amplitude modulation function. The sampling frequency was set to 5 kHz. The resulting wheeze waveform and its ideal or theorical TFR, produced by mapping $A(t)$ and f_0 functions to an image, are shown in figure 1.

B. TFRs Performance Measurements.

The ideal TFR of the synthetic wheeze in Fig. 1(b) offers a reference to compare the performance of the previously introduced TFRs and to select the best TFR to analyze real acquired wheeze sounds.

B.1. Central 2D Correlation

This index is given by

$$\rho = \frac{\sum_l \sum_k TFR_d(l,k) TFR_e(l,k)}{\sqrt{\sum_l \sum_k \left(TFR_d(l,k)\right)^2 \sum_l \sum_k \left(TFR_e(l,k)\right)^2}} \quad (9)$$

where $TFR_e(l,k)$ corresponds to the ideal TFR, $TFR_d(l,k)$ corresponds to the estimated TFR for each technique, and l and k represent the discrete time and frequency indexes, respectively [14]. This index was used in two ways: 1) in a global way, i. e., applying the former equation directly considering the TFR as a complete image, and 2) in a local way dividing the TFR in sub-images and then, computing an average value ρ_{mean} to asses local behavior. A correlation index value closer to one is associated with the best performance.

B.2. NRMSE

It is possible to estimate the instantaneous frequencies of a signal by means of the first moment of its TFR. Then, the relative error index is defined as

$$NRMSE = \sqrt{\frac{\sum_l \left(IF_d(l) - IF_e(l)\right)^2}{\sum_l \left(IF_d(l)\right)^2}} \quad (10)$$

where $IF_d(l)$ and $IF_e(l)$ correspond to the instantaneous frequencies obtained from the ideal and estimated TFR, respectively, and l represents the discrete time index. According to the former definition, the TFR with the minimum value of *NRMSE* is associated with the best performance.

B.3. Time-Frequency Resolution Index

This index measures the energy concentration or resolution of the estimated TFRs, and is given by

$$res_{TF} = \frac{\sum_l \frac{\sum_k TFR_e(l,k)}{\max_k \{TFR_e(l,k)\}}}{N} \cdot \frac{\sum_k \frac{\sum_l TFR_e(l,k)}{\max_l \{TFR_e(l,k)\}}}{N} \quad (11)$$

where $TFR_e(l,k)$ corresponds to the estimated TFRs by each TF technique, N represents the number of time and frequency samples in the TFR, and l and k represent the discrete time and frequency indexes, respectively [14]. The minimum res_{TF} value implies the TFR with best performance.

C. Computation of TFRs

In general, there is a tradeoff in performance for a TFR that depends on different parameters as for the Spectrogram. This tradeoff is *non trivial* and requires a *prior* knowledge of the theoretical signal TF distribution. Fortunately, we possess the ideal TFR of the synthetic wheeze. Consequently, the TFRs were computed and compared for different set of parameters, as shown in table 1, and the best one for each technique was selected according to the performance index ρ_{mean} for a final comparison between different TFRs. Each TFR, including the ideal one, was normalized with respect to their maximum and minimum values to produce an intensity image in the interval [0,1].

D. Real acquired wheezes

Wheezes sounds were selected from a database created by the digital signal and image processing laboratory of the Universidad Autónoma Metropolitana at Mexico City; the multichannel acquisition system is described in [15]. Particularly, wheezes were obtained from diffuse interstitial pneumonia patients from the National Institute of Respiratory Diseases (INER) at Mexico City. Patients were asked to breathe at a maximum airflow of 1.5 L/s, and the breathing maneuver was controlled by a physician, including an initial 5 s period of inspiratory apnea followed by 5 s of breathing and a final 5 s period of expiratory apnea. Afterward FIR digital filtering, temporal windows including wheezes were visually extracted and finally, they were processed by the selected TFR technique.

IV. Results

The parameters of selected TFRs for the synthetic wheeze are shown in table 1 along with their corresponding performance indexes. The corresponding selected TFRs of synthetic wheeze are shown in figure 2.

Table 1. Sets of parameters used to compute the TFRs and performance indexes

TFR	Set of parameters used		Parameters selected	ρ	ρ_{mean}	*NRMSE*	res_{TF}
SP	Window type	Rectangular, Hamming and Blackman-Harris.	Hamming	0.5300	0.8851	0.0074	0.0029
	Window length	9 linearly equidistant values in [13 - 65] ms.	65 ms				
P_{BURG}	Model Order	[2 - 5].	M=4	0.4564	0.8798	0.0199	0.0048
	Window type	Hamming and Blackman-Harris.	Hamming				
	Window length	9 linearly equidistant values in [13 - 65] ms.	65 ms				
SPR	Window type	Rectangular, Hamming and Blackman-Harris.	Hamming	0.8821	0.9163	0.1400	5.83×10^{-5}
	Window length	9 linearly equidistant values in [13 - 65] ms.	13 ms				
HHS	-	-	-	0.9247	0.9521	0.0601	2.47×10^{-6}

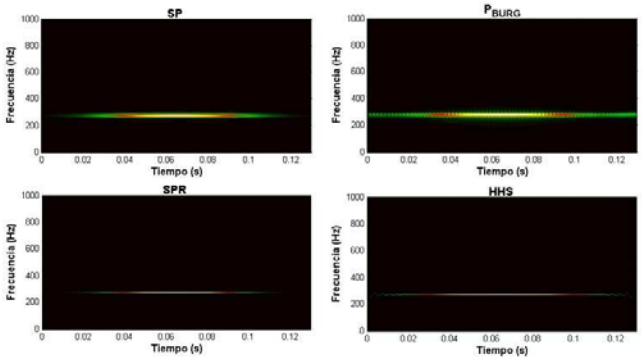

Fig. 2 Selected time-frequency representations for the synthetic wheeze according to criteria in equations (9)-(11)

V. Discusion

According to the performance indexes, and as expected by visual inspection, HHS results the best TFR technique for the continuous adventitious lung sound, having the best performance achieving three of four indexes. Then, HHS technique was applied to real acquired wheezes. An example is shown in figure 3 where the Spectrogram is compared with the HHS. As can be seen, the HHS shows a time-varying spectral behavior that is lost in the Spectrogram due to the HHS improved TF resolution.

VI. Conclusions

HHS, based on empirical mode decomposition, offers a better TF resolution compared with the classical spectrogram to analyze continuous adventitious lung sounds. This improved resolution revealed for real wheeze certain time-varying spectral behavior that could help to understand in a better way the physiological genesis of this kind of lung sounds.

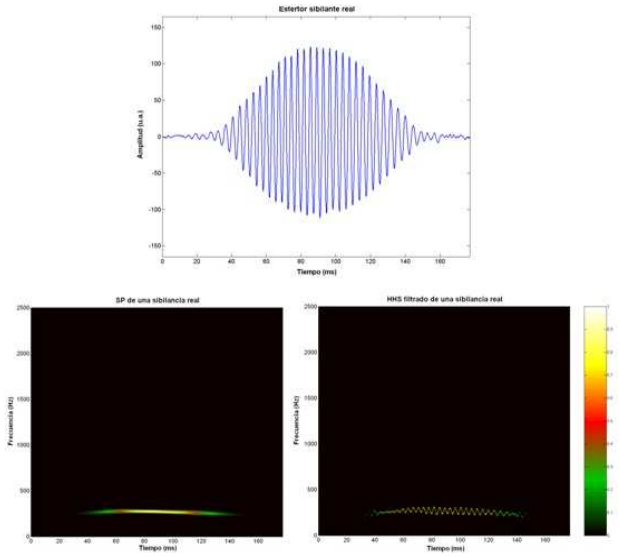

Fig. 3 Real acquired wheeze. a) Time waveform. b) Spectrogram, Hamming window of 65 ms. c) HH Spectrum, smoothed version

References

1. Sovijärvi A, Dalmasso F, Vanderschoot J, Malmberg L, Righini G, Stoneman S (2000) Definitions of terms for applications of respiratory sounds, Eur. Respir. Rev., vol. 10, pp. 597-610
2. Gavriely N, Palti Y, Alroy G, Grotberg J (1984) Measurement and theory of wheezing breath sounds, J. Appl. Physiol., vol. 57, no. 2, pp. 481-492
3. Murphy R, Holford S, Knowler W (1977) Visual lung-sound characteri-zation by time-expanded wave-form analysis, N. Eng. J. Med., vol. 296, pp. 968-971
4. McKusick V, Jenkins J, Webb G (1955) The acoustic basis of the chest examination: studies by means of sound spectrography, Am. Rev. Tuberc., vol. 72, pp. 12-34
5. Baughman R, Loudon R (1984) Quantitation of wheezing in acute asth-ma," Chest, vol. 86, pp. 718-722
6. Rosqvist T, Paajanen E, Kallio K, Rajala H, Katila T, Piirilä P, Malmberg P, Sovijärvi A (1995) Toolkit for lung sound analysis," Med. Biol. Eng. Comput., vol. 33, pp. 190-195
7. Homs-Corbera A, Fiz J, Morera J, Jané R (2004) Time-frequency detec-tion and analysis of wheezes during forced exhalation, IEEE Trans. Biomed. Eng., vol. 51, pp. 182-186
8. Taplidou S, Hadjileontiadis L, Kitsas I, Panoulas K, Penzel K, Gross V, Panas S (2004) On applying Continuous Wavelet Transform in wheeze analysis, Proc. 26th Annual Int. Conf. IEEE-EMBS, pp. 3832-3835
9. Cohen L (1995) Time-Frequency Analysis. Englewood Cliffs, NJ: Prentice-Hall
10. Harris F (1978) On the use of windows for harmonic analysis with the discrete Fourier transform" Proc. IEEE, vol. 66, no. 1, pp. 51-84
11. Marple S (1989) A tutorial overview of modern spectral estimation," Proc. IEEE-ICASSP, pp. 2152-2157
12. Auger F, Flandrin F (1995) Improving the readability of time-frequency and time-scale representations by the reassignment method, IEEE Trans. on Signal Proc., vol. 43, no. 5, pp. 1068-1089
13. Huang N E, Shen Z, Long S R, Wu M C, Shih H H, Zheng Q, Yen N C, Tung C C and Liu H H (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. Roy. Soc. London, Ser. A, vol. 454, no. 1971, pp. 903-995
14. Rankine L, Stevenson N, Mesbah M, Boashash B (2005) A Quantitative Comparison of Non-Parametric Time-Frequency Representations" Proc. 13th Eur Sig. Proc. Conf. (EUSIPCO2005), Turkey
15. Charleston-Villalobos S, González-Camarena R, Chi-Lem G, Aljama-Corrales T (2007) Crackle sounds analysis by empirical mode de-composition, IEEE Eng. Med. Biol. Mag., pp. 40-47