JOURNAL OF THE TENNESSEE ACADEMY OF SCIENCE 73(1–2):47–52 JANUARY–APRIL 1998

ENHANCED SEPARATION OF CRACKLES AND SQUAWKS FROM VESICULAR SOUNDS USING NONLINEAR FILTERING WITH THIRD-ORDER STATISTICS

LEONTIOS J. HADJILEONTIADIS AND STAVROS M. PANAS

Department of Electrical & Computer Engineering, Aristotle University of Thessaloniki, GR-540 06 Thessaloniki, Greece

ABSTRACT—The separation of pathological discontinuous adventitious sounds (DAS) from vesicular sounds (VS) is of great importance in the analysis of lung sounds, since they are related to certain pulmonary pathologies. An automated way of revealing the diagnostic character of DAS by isolating them from VS, based on their nonstationarity, is presented in this paper. A combination of nonlinear filtering with third-order statistics is implemented, resulting in a modification of the nonlinear digital STationary-NonSTationary (ST-NST) filter, proposed by Arakawa et al. (1986), for improving its performance in noisy environments. The use of third-order statistics in estimating the AR-model's order and coefficients provides more reliable estimates of the stationary part of the input signal. The implementation of this modified ST-NST filter (mST-NST) in fine crackles, coarse crackles and squawks was examined. Our results indicate that a reliable and accurate separation of these adventitious sounds from vesicular sounds can be achieved, even in the presence of additive Gaussian noise.

Pulmonary diagnosis is often based on the analysis of acoustical pulmonary signals, since the generated acoustical energy, produced by air flow during inspiration and expiration, is highly correlated with pulmonary dysfunction. Pulmonary dysfunction is caused by anatomical or physiological changes in the pulmonary system and is characterized by changes in the acoustical properties of the various parts or organs involved (Cohen, 1990). Thus, when narrowing of a portion of the tracheobronchial tree occurs, turbulent flow may cause the generation of specific acoustical noise, i.e. adventitious sounds.

Adventitious sounds are divided into two major classes: continuous and discontinuous sounds (Loudon et al., 1984), and are only heard in pathological cases, indicating an underlying physiological malfunction. The first class contains wheezes and rhonchi, characterized by a relatively long duration (250 ms), and a sharp peak in the power spectral density function in the range of 400 Hz (wheezes) or in the range of 200 Hz or less (rhonchi). The second class contains crackles and squawks, characterized mainly by their time domain features such as a relatively short duration (20 ms), the initial deflection width (IDW) and the two cycle duration (2CD).

Crackles are discrete, nonmusical sounds, which, when they appear, behave as a nonstationary explosive noise superimposed on breath sounds. Their only useful categorization is between fine and coarse crackles, with IDW = 0.90 ms; 2CD = 6.0 ms and IDW = 1.25 ms; 2CD = 9.50 ms (Cohen, 1990), respectively. Fine crackles (or Velcro™ sounds) are exclusively inspiratory events which tend to occur in mid-to-late inspiration and repeat in similar patterns over subsequent breaths. They have been credibly established to result from the explosive reopening of small airways that had closed during the previous expiration. They are connected either to congestive heart failure or to pulmonary fibrotic diseases such as asbestosis and idiopathic interstitial fibrosis. Coarse crackles are found in early inspiration and occasionally in expiration as well. They are of a "popping" quality (not Velcro™-like) and tend to be less reproducible from breath to breath. Furthermore, they apparently arise from fluid in small airways; can change pattern or clear after coughing, implying a transient character in their production mechanism; and are related with chronic bronchitis. Squawks are a combination of wheezes and crackles; although they appear as short inspiratory wheezes, they are heard in association with fine crackles (in fact, they may be initiated with a crackle). They are related to allergic alveolitis and interstitial fibrosis and are caused by the explosive opening and fluttering of the unstable airway which causes the short wheeze (Kraman, 1993).

From the aforementioned description of the discontinuous adventitious lung sounds it is evident their separation from vesicular sounds could reveal significant information, since the structure of the DAS isolates its diagnostic character. In order to achieve automated separation, their nonstationarity must be taken into account. Consequently, the use of highpass filtering fails to separate the nonstationary sounds, destroying the waveforms. Furthermore, level slicer can not overcome the small amplitude of fine crackles. Application of time-expanded waveform analysis in crackle time domain analysis (Loudon et al., 1984; Murphy et al., 1977) results in separation; however, it is time consuming with large interobserver variability. In the case of fine crackles, nonlinear processing proved to overcome the aforementioned problems, resulting in objective and accurate results (Ono et al., 1989).

The extension of nonlinear filtering by means of higher-order statistics results in an enhanced separation of all discontinuous adventitious sounds from vesicular ones, since AR-modeling based on third-order statistics provides more reliable estimations of the predicted signal. In this paper, the implementation of a nonlinear filter based on third-order statistics, in separation of crackles (fine & coarse) and squawks (nonstationary waves) from vesicular sounds (stationary waves) is presented. Its performance is evaluated through experimental results, which prove the establishment of an efficient and objective method.

METHOD

In this section, the description of the proposed nonlinear filter (mST-NST) is presented. Our work builds on the ST-NST filter proposed by Arakawa et al. (1986). In our case, the prediction filter performs autoregressive prediction based on third-order statistics (AR-TOS). The equation describing the autoregressive model is:

$$y_n + \sum_{i=1}^{p} a_i y_{n-i} = w_n, \qquad a_0 = 1, \tag{1}$$

where y_n represents a p^{th} order AR process of N samples ($n = 0,\ldots,N-1$), a_i are the coefficients of the AR-model, and w_n are i.i.d., non-Gaussian, third-order stationary, zero-mean, with $E\{w_n^3\} = \beta \neq 0$ and y_n independent of w_l for $n < l$. Since w_n is third-order stationary, y_n is also third-order stationary, assuming it is a stable AR model. For the model of equation (1) we then have (Nikias et al., 1993):

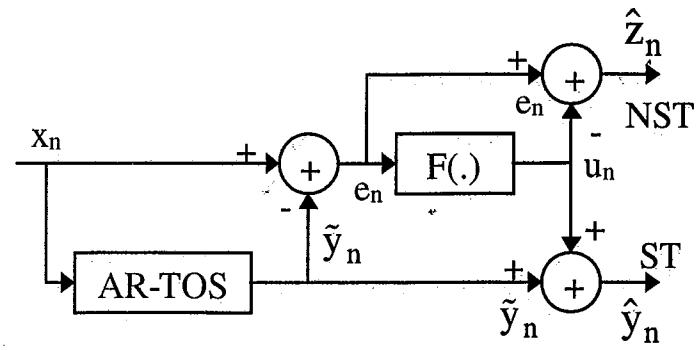
$$R(\tau_1, \tau_2) + \sum_{i=1}^{p} a_i R(\tau_1 - i, \tau_2 - i) = \beta \cdot \delta(\tau_1, \tau_2), \tag{2}$$

where, $R(\tau_1, \tau_2)$ is the third order moment or cumulant sequence of the AR process and $\delta(\tau_1, \tau_2)$ is the 2-d unit impulse function. If τ_1 and τ_2 are in the range of $\tau_1 = 0, 1,\ldots, p$ and $\tau_2 = 0, 1,\ldots, p$, respectively, we find the a_i coefficients with the Optimized AR Method (OARM), proposed by An et al. (1990), using all third-order cumulants of the pxp plane (and not only along the line $\tau_1 = \tau_2$). This method uses third-order statistics to formulate an overdetermined system of equations for a_i and β with a least squares solution. The calculation of AR model's order p is reduced to a rank determination problem. According to Giannakis et al. (1990), the rank of a matrix C_e, defined as

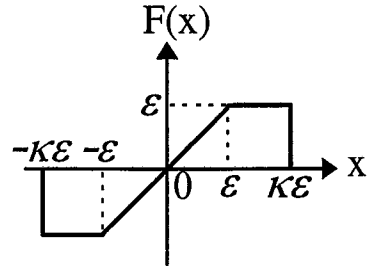
$$C_e = \begin{bmatrix} R(1,-\bar{p}) & \cdots & R(\bar{p},-\bar{p}) \\ \vdots & \vdots & \vdots \\ R(1,0) & \cdots & R(\bar{p},0) \\ & \vdots & \\ R(\bar{p},-\bar{p}) & \cdots & R(2\bar{p}-1,-\bar{p}) \\ \vdots & \vdots & \vdots \\ R(\bar{p},0) & \cdots & R(2\bar{p}-1,0) \end{bmatrix} [\bar{p}(\bar{p}+1)x\bar{p}], \tag{3}$$

formed by exact third-order cumulants, is equal to p, even when only $\bar{p}$, the upper bound of p, is known. The rank of C_e is p, and equals the maximum number of nonzero singular values of C_e. In practice, all the singular values of the matrix $\hat{C}_e$ (formed after replacing the third-order statistics with their estimates) will be nonzero. A subjective rule to select the "effective" AR order p is to find the largest drop among two successive normalized singular values of $\hat{C}_e$. The use of this method is more reliable than applying the SVD approach to a similar matrix containing the sampled autocorrelation estimations, when additive colored Gaussian noise is present (Giannakis et al., 1990).

The most profound motivations behind the use of third-order statistics in the estimation of a_i are (Proakis et al., 1992): i) Suppression of Gaussian noise since third-order statistics of Gaussian signals are identically zero, and ii) preservation of the true phase character of the signal since third-order statistics do not suppress the phase information of the signal, as the second-order statistics (autocorrelation) do. Hence, when the analysis waveform consists of a non-Gaussian signal in additive Gaussian noise, the parameter estimation of the original signal using third-order statistics takes place in a high signal-to-noise ratio (SNR) domain. Furthermore, the parametric presentation of the process is more accurate and reliable (Nikias et al., 1993).



(a)



(b)

FIG. 1. (a) Schematic diagram of the modified ST-NST separating filter (mST-NST). AR-TOS is the AR-model based on Third-Order Statistics. (b) Definition of nonlinear function F.

Let the input signal x_n be a summation of two types of signals: the stationary signal, that can be expressed by an autoregressive model, and the nonstationary signal, composed of random impulsive waves, whose occurrence rate is low (Ono et al., 1989). Under these circumstances we can separate nonstationary from stationary signals using the filter depicted in Fig. 1(a).

According to Fig. 1(a), x_n is the input, $\hat{z}_n$ is the nonstationary output, and $\hat{y}_n$ is the stationary output. The prediction filter based on third-order statistics performs autoregressive prediction of the stationary waves $\hat{y}_n$ of the input signal x_n, as follows:

$$\tilde{y}_n = \sum_{i=1}^{p} a_i \hat{y}_{n-i}. \tag{4}$$

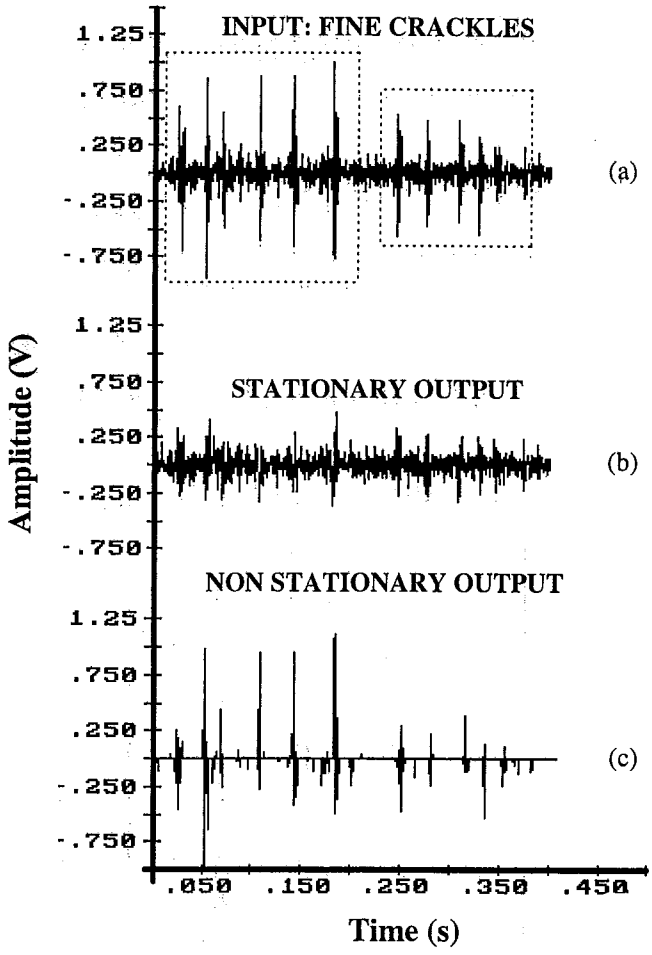


FIG. 2. (a) A time section of 0.41 sec of fine crackles recorded from a patient with pulmonary fibrosis. (b) The stationary output of mST-NST filter (VS). (c) The nonstationary output of mST-NST filter (DAS).

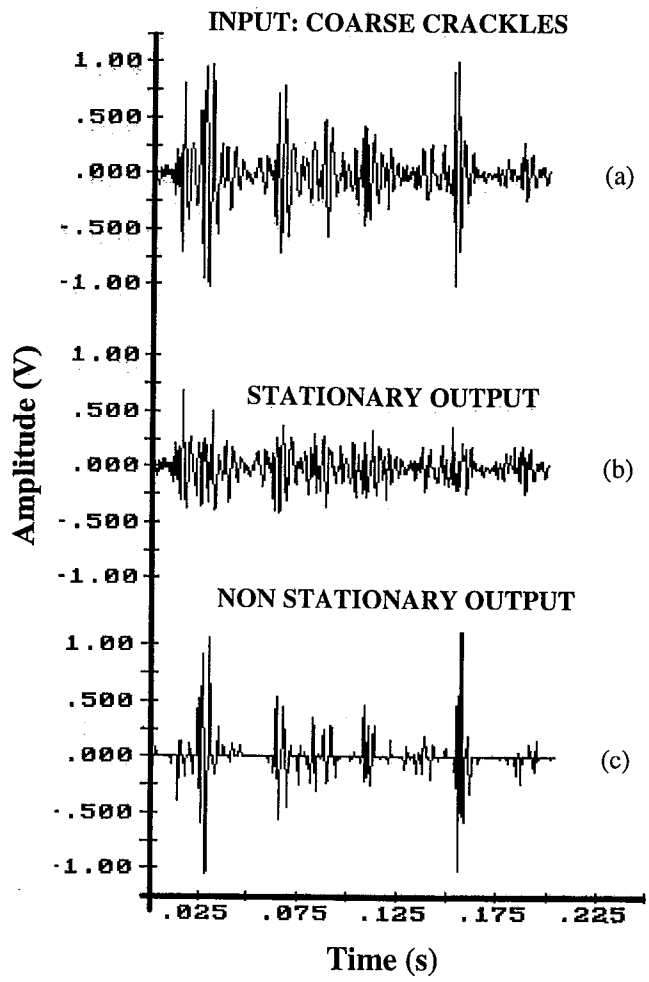


FIG. 3. (a) A time section of 0.25 sec of coarse crackles recorded from a patient with chronic fibrosis. (b) The stationary output of mST-NST filter (VS). (c) The nonstationary output of mST-NST filter (DAS).

The prediction error e_n for the input x_n is obtained by subtracting the predicted output from the input, i.e.:

$$e_n = x_n - \tilde{y}_n. \tag{5}$$

The e_n is then processed by the nonlinear function $F(*)$ for determining u_n, which is the part included in the stationary component, as follows:

$$u_n = F(e_n). \tag{6}$$

According to the ST-NST nonlinear filter by Arakawa et al. (1986), the nonlinear function $F(*)$ is defined as in Fig. 1(b). The value of the parameter ϵ is determined so that the probability of detection of nonstationary waves is given by a certain value γ, i.e.:

$$\int_{-\epsilon}^{\epsilon} p(x)\, dx = 1 - \gamma, \tag{7}$$

where, $p(x)$ is the probability density function of the prediction error of the original signal. In our case, we suppose that the probability density function (*pdf*) of the prediction error of the original signal is the *pdf* of a zero-mean, non-Gaussian distribution. An exponential distribution with mean equal to θ, shifted to left, was used. The delayed exponential *pdf* was selected in order to obtain non zero third-order statistics (since it is non symmetrical) and zero mean. Furthermore, from a physiologic perspective, the exponential distribution is preferable for populations where the observations involve items whose status changes over time. This fact resembles the nonstationary character of DAS.

The nonstationary output $\hat{z}_n$ and the stationary output $\hat{y}_n$ are obtained as follows:

$$\hat{y}_n = \tilde{y}_n + u_n \text{ and } \hat{z}_n = e_n - u_n. \tag{8}$$

If $-\epsilon < e_n < \epsilon$, then the stationary output is the input signal itself and the nonstationary component is zero, i.e., $\hat{y}_n = x_n$ and $\hat{z}_n = 0$. If $|e_n| > \kappa\epsilon$, then the stationary output is the autoregressively estimated value itself and the nonstationary output is the whole prediction error, i.e., $\hat{y}_n = \tilde{y}_n$ and $\hat{z}_n = e_n$. If $-\kappa\epsilon < e_n < -\epsilon$ and $\epsilon < e_n < \kappa\epsilon$, the input signal is processed as quasistationary in the sense that a part of the prediction error is included in the stationary output and the other is regarded as the nonstationary component, i.e., $\hat{y}_n = \tilde{y}_n - \epsilon$ and $\hat{z}_n = e_n + \epsilon$, $\hat{y}_n = \tilde{y}_n + \epsilon$ and $\hat{z}_n = e_n - \epsilon$, respectively. In all cases, the input signal x_n is equal to the sum of $\hat{y}_n$ and $\hat{z}_n$.

IMPLEMENTATION

The method was implemented on a IBM-PC (Pentium/120MHz) using the programming language ASYST 4.1 (Keithley Instruments, Inc, Taunton, MA). Pre-classified signals, corresponding to fine crackles (from pulmonary fibrosis), coarse crackles (from chronic bronchitis), and squawks (from interstitial fibrosis) were drawn from a lung sound data-base (Kraman, 1993). Nine cases were investigated (Table 1). After an antialiasing filter, the signals were digitized with a 12-Bit A/D converter at a sampling rate of 2.5KHz. A section of 15s of every signal was digitized and processed by the nonlinear filter. Successive records of $N = 512$ or $N = 1024$ samples were divided into $M = 8$ sections each, and the third-order statistics of each record were estimated by averaging the cumulants of each section. Using equation 2, the β and a_i coefficients were estimated, in order to predict the stationary part of the input signal (per record). The values of γ, θ and κ were settled to be, in most cases, 0.04, 0.1 and 4, respectively, as a trade-off between not separating vesicular sounds and separating crackles and squawks. The value of order p was found to be equal to 2.

RESULTS AND DISCUSSION

Results obtained with the mST-NST filter on different kinds of DAS (fine/coarse crackles & squawks) are presented in this section. The processed records were selected so the main structure morphologies of the DAS would be clearly encountered. Figures 2, 3 and 4 depict the experimental results of applying the nonlinear filter of Fig. 1 to fine crackles, coarse crackles and squawks, respectively. From Fig. 2 we can see the separation of two repeated similar patterns of fine crackles from vesicular sounds, including six fine crackles each (included in dotted squares). The observed similarity results from the explosive reopening of small airways that had closed during the previous expiration. The abnormal airway closure that precedes the "crackling" reopening is due to increased lung stiffness. From this figure, the explosive nonstationary character of fine crackles, along with their short time duration are evident, since they are superimposed on stationary vesicular sounds with lower amplitude.

In Fig. 3, the separation of seven coarse crackles from vesicular sounds, with "popping" quality and tendency to be less reproducible, is depicted. This observed randomness in crackles' occurrence is due to the presence of fluid in the small airways. Although this sequence of coarse crackles differs from the previously described fine crackles in both structure and shape, the mST-NST filter still performs satisfactorily. This becomes clear from Figs. 3 (b) & (c) where the true locations of the input coarse crackles and their time duration and morphology are accurately identified, while the shape characteristics of VS also are accurately retained. Furthermore the separation of two isolated squawks from vesicular sounds can be observed (Fig. 4). The two main characteristics of a squawk, i.e., an underlying fine crackle, followed by a short wheeze with an almost exponential decay, are easily identified in the first squawk. These characteristics are clearly reproduced in the nonstationary output of the mST-NST filter depicted in Fig. 4(c). The pure vesicular sounds are accurately reconstructed from the mST-NST filter stationary output depicted in Fig. 4(b).

Apart from visual comparisons between the original DAS and those picked up by the mST-NST filter, a quantitative analysis of the performance of the mST-NST filter for each case was performed by means of the rate of detectability, D_{R1}, defined as follows:

$$D_{R1} = \left(1 - \frac{N_R - N_{E1}}{N_R}\right) \cdot 100, \qquad (9)$$

where N_R is the number of visually recognized DAS by a physician (considered as the true number of DAS in the input signal), and N_{E1} is the number of estimated DAS using the mST-NST filter. Analytical results of the above mentioned parameters are presented in Table 1. These results indicate that the mST-NST

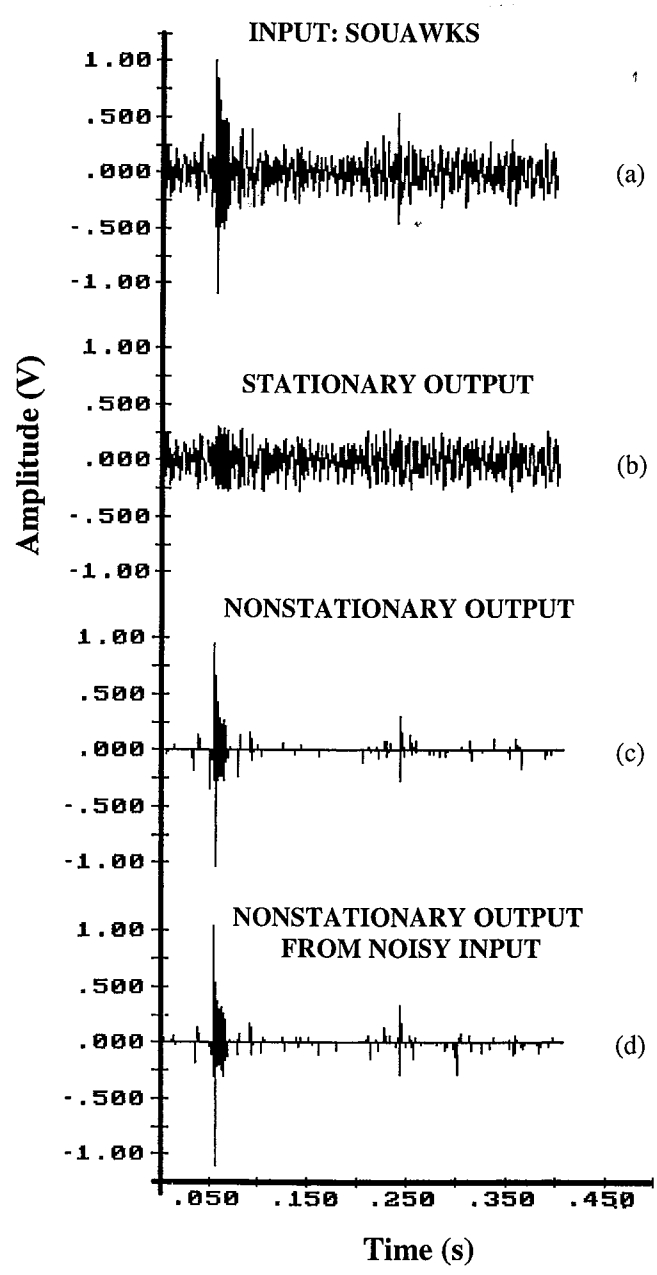


FIG. 4. (a) A time section of 0.41 sec of squawks recorded from a patient with interstitial fibrosis. (b) The stationary output of mST-NST filter (VS). (c) The nonstationary output of mST-NST filter (DAS). (d) The nonstationary output of mST-NST filter (DAS) from noisy input [+GN(mean = 0, variance = 0.03)].

TABLE 1. The performance of mST-NST filter vs ST-NST filter for different cases of DAS.

Cases	DAS type	Diagnosis	N	M	N_R	Noise-free case				Additive Gaussian noise [mean = 0, variance = 0.03]			
						N_{E1}	N_{E2}	$D_{R1}(\%)$	$D_{R2}(\%)$	N'_{E1}	N'_{E2}	$D'_{R1}(\%)$	$D'_{R2}(\%)$
C1	FC	PF	1024	8	12	12	10	100	83.0	12	9	100	75.0
C2	FC	PF	1024	8	7	7	6	100	85.7	7	5	100	71.4
C3	FC	PF	1024	8	7	7	6	100	85.7	7	5	100	71.4
C4	CC	CB	512	8	7	7	4	100	57.1	7	4	100	57.1
C5	CC	CB	1024	8	8	8	6	100	75.0	8	5	100	62.5
C6	CC	CB	1024	8	11	11	6	100	54.5	11	6	100	54.5
C7	SQ	IF	1024	8	2	2	2	100	100	2	1	100	50.0
C8	SQ	IF	1024	8	8	8	6	100	75.0	8	5	100	62.5
C9	SQ	IF	1024	8	2	2	2	100	100	2	2	100	100

FC: Fine Crackles; CC: Coarse Crackles; SQ: Squawks; PF: Pulmonary Fibrosis; CB: Chronic Bronchitis; IF: Interstitial Fibrosis. Used values: p = 2, $\gamma = 0.04$, $\theta = 0.1$, $\kappa = 4.0$ (mST-NST filter); p = 15, $\gamma = 0.04$, $\kappa = 2.0$ (ST-NST filter; Arakawa et al., 1986).

filter performs very well, since all DAS were efficiently detected. To quantitatively compare the mST-NST and ST-NST filters, the ST-NST filter was implemented and applied to our database. Therefore, a measure of the performance of the ST-NST filter was defined as D_{R2}, in a similar manner to the definition of D_{R1} (equation (8)), as follows:

$$D_{R2} = \left(1 - \frac{N_R - N_{E2}}{N_R}\right) \cdot 100, \qquad (10)$$

where N_{E2} is the number of estimated DAS using the ST-NST filter. Results of DAS analysis with the ST-NST filter are presented in Table 1. Comparing these results to those obtained with the mST-NST filter, it is evident that the mST-NST filter performs better than the ST-NST filter, in all cases of each type of DAS.

To examine the contribution of third-order statistics in the accurate separation of the stationary and nonstationary part, we applied the nonlinear filtering in input signals with additive colored Gaussian Noise with zero mean and variance equal to 0.03. The N_{E1}, N_{E2}, D_{R1}, and D_{R2} parameters were re-calculated for the noisy inputs, resulting in the new values of N'_{E1}, N'_{E2}, D'_{R1}, and D'_{R2} parameters, respectively (Table 1). From Table 1 it is evident that in the presence of additive Gaussian noise the performance of the mST-NST filter remained unchanged, while that of the ST-NST filter was deteriorated, since its rate of detectability, in most cases, was reduced, due to the additive Gaussian noise. Additionally it was found that: i) the estimation of the order p was not influenced by the Gaussian noise, and ii) the nonstationary and stationary outputs of the mST-NST filter were almost identical to those derived from the noise-free case, demonstrating the noise-robustness of the mST-NST filter. An example of these results, in the case of squawks, is depicted in Fig. 4(d).

The effect of the mST-NST filter on input breath sounds also was tested by listening to its stationary outputs after Digital-to-Analog (D/A) conversion ('wav' format archives). During this evaluation procedure, the DAS were almost practically not heard, confirming a good separation performance by the mST-NST filter.

Furthermore, from our experiments, it was clear that, when the AR prediction of VS is performed by third-order statistics, the estimated values of model's order p are kept low enough ($p \leq 3$), compared to Ono's method ($p = 15$) (Ono et al., 1989). Finally, in order to maintain the probability of detecting DAS constant, to keep the stability of the filter, and to capture the changes of VS among patients, sites and states of ventilation, the $F(*)$ function was set, and the order p and a_i coefficients of the AR-TOS model were calculated adaptively, at each record of the input signal. In this way, this adaptive procedure updated the mST-NST filter parameters every time a new record of the incoming signal was processed. A step by step methodology of the mST-NST filter is summarized in Appendix 1.

A combination of nonlinear filtering with third-order statistics was presented. This method, is a modification of the ST-NST filter proposed by Arakawa et al. (1986), and differs in the way the stationary part of the signal is modeled by an AR-model. The use of third-order statistics in estimating AR-model's order and its coefficients, lead to more reliable prediction of the stationary part of the signal, improving the overall performance of the ST-NST filter.

Finally, experiments have shown that this modified nonlinear filter can separate fine crackles, coarse crackles and squawks, even when they are contaminated by Gaussian noise or by noise with symmetrical *pdf*.

LITERATURE CITED

AN, G. K., S. B. KIM, AND E. J. POWERS. 1990. Optimized parametric bispectrum estimation. Proc. ICASSP', 88:2392–2395.

ARAKAWA, K., D. H. FENDER, H. HARASHIMA, H. MIYAKAWA, AND Y. SAITOH. 1986. Separation of nonstationary component from the EEG by a nonlinear digital filter. IEEE Trans. Biomed. Eng., 33:724–726.

COHEN, A. 1990. Signal processing methods for upper airway and pulmonary dysfunction diagnosis. IEEE Eng. in Medicine & Biology Mag., 3:72–75.

GIANNAKIS, B. G., AND J. M. MENDEL. 1990. Cumulant-based order determination of non-gaussian ARMA models. IEEE Trans. on Acoustics, Speech, and Signal Processing, 38: 1411–1423.

KRAMAN, S. S. 1993. Lung Sounds: an introduction to the inter-

pretation of the auscultatory finding. American College of Chest Physicians, Northbrook, Illinois.

LOUDON, R., AND R. L. H. MURPHY, JR. 1984. Lung Sounds. Am. Rev. Respir. Dis., 130:663–673.

MURPHY, R. L. H. JR., S. K. HOLFORD, AND W. C. KNOWLER. 1977. Visual lung-sound characterization by time-expanded waveform analysis. New Eng. J. Med., 296:968–971.

NIKIAS, L. CH., AND A. M. PETROPULU. 1993. Higher-order spectra analysis: a nonlinear signal processing framework. PTR Prentice-Hall, New Jersey.

ONO, M., K. ARAKAWA, M. MORI, T. SUGIMOTO, AND H. HARASHIMA. 1989. Separation of fine crackles from vesicular sounds by a nonlinear digital filter. IEEE Trans. Biomed. Eng. 36:286–291.

PROAKIS, G. J., CH. M. RADER, F. LING, AND CH. L. NIKIAS. 1992. Advanced digital signal processing, Macmillan, New York.

APPENDIX 1

In this appendix, a step by step methodology of mST-NST filter is illustrated.

Step 1: Initialize: N (number of samples per record), x_n (input signal, n = 1,. . .,N), L (number of records), M (number of sections per record).

Step 2: For i = 1,2,. . ., L.

- Set the values of γ, θ, κ (parameters of function F(*)), $\tilde{p}$ (estimation of the upper bound of AR-TOS model order).
- Find the value of parameter ϵ using equation (7).
- Construct the function F(*) using Fig. 1(b) and the values of γ, θ, κ and ϵ.
- Find the third-order statistics of each section M of x_n, and average them to get the final estimation of $R(\tau_1,\tau_2)$.
- Estimate the AR-TOS model order p using the method proposed by Giannakis et al. (1990).
- Find the a_i coefficients and β of AR-TOS model using the estimated value of p, and equations (1) & (2) (An et al., 1990).
- Calculate the prediction error e_n using equations (4) & (5).
- Calculate the function u_n using the equation (6).
- Calculate the nonstationary output $\hat{z}_n$ and the stationary output $\hat{y}_n$ using equation (8) and the structure of Fig. 1(a).

Step 3: Next i.