micturition process. Here, we have used the new algorithm to examine more clearly the signal frequencies and persistent signal intensities of EUS and bladder at the same time.

The datasets can be obtained from: http://oz.nthu.edu.tw/~d907911/FD.html

IV. Conclusion

The main advantage in our approach is that the signal frequencies and fractal dimensions can be obtained via the SDF simultaneously. The only demerit in this approach is that the computational amount is not of the same order as conventional FFT. However, when accuracy is more important, the proposed method will be more desirable. It is believed that one can benefit from using this proposed method in the evaluation of physiological functions.

Acknowledgment

The authors would like to thank W. C. Lin of Taichung Veterans General Hospital for performing the animal experiments and providing the EMG and CMG data for this paper.

References

[1] B. B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1983.
[2] J. Feder, *Fractals*. New York: Plenum, 1988.
[3] I. Rodriguez-Carreno, A. Malanda-Trigueros, L. Gila-Useros, J. Navallas-Irujo, and J. Rodriguez-Falces, "Filter design for cancellation of baseline-fluctuation in needle EMG recordings," *Comput. Methods Programs Biomed.*, vol. 81, pp. 79–93, 2006.
[4] R. Ferenets, T. Lipping, A. Anier, V. Jantti, S. Melto, and S. Hovilehto, "Comparison of entropy and complexity measures for the assessment of depth of sedation," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 6, pp. 1067–1077, Jun. 2006.
[5] R. Carvajal, N. Wessel, M. Vallverdu, P. Caminal, and A. Voss, "Correlation dimension analysis of heart rate variability in patients with dilated cardiomyopathy," *Comput. Methods Programs Biomed.*, vol. 78, pp. 133–140, 2005.
[6] Y. C. Lai, I. Osorio, M. A. Harrison, and M. G. Frei, "Correlation-dimension and autocorrelation fluctuations in epileptic seizure dynamics," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 65, pp. 031921–031921, 2002.
[7] J. Bhattacharya, H. Petsche, and E. Pereda, "Long-range synchrony in the gamma band: Role in music perception," *J. Neurosci.*, vol. 21, pp. 6329–6337, 2001.
[8] G. Henderson, E. Ifeachor, N. Hudson, C. Goh, N. Outram, S. Wimalaratna, C. D. Percio, and F. Vecchio, "Development and assessment of methods for detecting dementia using the human electroencephalogram," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 8, pp. 1557–1568, Aug. 2006.
[9] S. C. Liu and S. Chang, "Dimension estimation of discrete-time fractional Brownian motion with applications to image texture classification," *IEEE Trans. Image Process.*, vol. 6, no. 8, pp. 1176–1184, Aug. 1997.
[10] O. I. Craciunescu, S. K. Das, J. M. Poulson, and T. V. Samulski, "Three-dimensional tumor perfusion reconstruction using fractal interpolation functions," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 4, pp. 462–473, Apr. 2001.
[11] K. L. Chan, "Quantitative characterization of electron micrograph image using fractal feature," *IEEE Trans. Biomed. Eng.*, vol. 42, no. 10, pp. 1033–1037, Oct. 1995.
[12] T. E. Southard, K. A. Southard, J. R. Jakobsen, S. L. Hillis, and C. A. Najim, "Fractal dimension in radiographic analysis of alveolar process bone," *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, vol. 82, pp. 569–576, 1996.
[13] S. Chang, S. T. Mao, S. J. Hu, W. C. Lin, and C. L. Cheng, "Studies of detrusor-sphincter synergia and dyssynergia during micturition in rats via fractional Brownian motion," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 8, pp. 1066–73, Aug. 2000.
[14] Y. C. Chang and S. Chang, "A fast estimation algorithm on the Hurst parameter of discrete-time fractional Brownian motion," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 554–559, Mar. 2002.
[15] S. Chang, S. J. Hu, and W. C. Lin, "Fractal dynamics and synchronization of rhythms in urodynamics of female Wistar rats," *J Neurosci. Methods*, vol. 139, pp. 271–9, 2004.
[16] J. L. Doob, *Stochastic Processes*. New York: Wiley, 1953.
[17] E. Wong and B. Hajek, *Stochastic Processes in Engineering* . New York: Springer-Verlag, 1985.
[18] S. Chang, S. T. Mao, T. P. Kuo, S. J. Hu, W. C. Lin, and C. L. Cheng, "Fractal geometry in urodynamics of lower urinary tract," *Chin. J. Physiol.*, vol. 42, pp. 25–31, 1999.

The Use of Wavelet Packet Transform and Artificial Neural Networks in Analysis and Classification of Dysphonic Voices

César David Paredes Crovato and Adalberto Schuck*

***Abstract*—This paper presents a dysphonic voice classification system using the wavelet packet transform and the best basis algorithm (BBA) as dimensionality reductor and 06 artificial neural networks (ANN) acting as specialist systems. Each ANN was a 03-layer multilayer perceptron with 64 input nodes, 01 output node and in the intermediary layer the number of neurons depends on the related training pathology group. The dysphonic voice database was separated in five pathology groups and one healthy control group. Each ANN was trained and associated with one of the 06 groups, and fed by the best base tree (BBT) nodes' entropy values, using the multiple cross validation (MCV) method and the *leave-one-out* (LOO) variation technique and success rates obtained were 87.5%, 95.31%, 87.5%, 100%, 96.87% and 89.06% for the groups 01 to 06, respectively.**

***Index Terms*—Acoustical analysis of voices, artificial neural network, dysphonic voice classification, wavelet packet transform.**

I. Introduction

The Videolaringoscopy and the Videostrobolaringoscopy are well established procedures for larynx pathology diagnosis [1]. However, many alternative techniques for automatic classification of voice quality have been proposed, based on linear prediction coding (LPC) and inverse filtering [2]–[4], artificial neural network (ANN) [5]–[7], acoustical indexes [8]–[10] and time-frequency techniques [11]–[15].

Because of insufficient number of subjects in some pathology group, in [10] the voice samples were grouped in six clusters based on acoustic and similarity of sounds. These groups are Chronic Laryngitis, Degenerative, Incorrect Mobility, Organic Alterations, Organic Growths, and Normal.

In [13], a linear discriminator between normal and pathological subjects based on wavelet packet (WPT) and best basis algorithm (BBA) is presented. This classifier is based on the presence or absence of certain best base tree (BBT) nodes for each sample of voice. The base function and cost function were *Symlet* 5 and *Shannon Entropy*, respectively; this choice gave better sensibility for detecting pathologies and higher gender insensibility. In [14], the authors proposed an extension

Manuscript received December 16, 2005; revised November 5, 2006. *Asterisk indicates corresponding author.*

C. D. P. Crovato is with the Departamento de Engenharia Elétrica, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, CEP 90.035-190, Brazil (e-mail: cesarcrovato@yahoo.com.br).

*A. Schuck is with the Departamento de Engenharia Elétrica, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, RS, Av. Osvaldo Aranha 103, CEP 90.035-190, Brazil (e-mail: schuck@eletro.ufrgs.br).

Digital Object Identifier 10.1109/TBME.2006.889780

TABLE I
NUMBER OF HIDDEN NEURONS VERSUS FINAL ERRORS

GN°*	Group	Pathology	Neurons	FP_{error}	FN_{error}	C_{error}	G_{error}
1	Chronic Laryngitis	Laryngitis	20	1.5%	10.9%	12.5%	43.6%
2	Degenerative	Cancer, granulomas, leukoplakia	25	0%	4.6%	4.6%	35.9%
3	Incorrect Mobility	Unilateral paralysis, unilateral and hypercinetic disfonia	10	1.5%	12.5%	12.5%	48.5%
4	Organic Alterations	Reinke's edema, degenerative edemas	5	0%	0%	0%	27.6%
5	Organic Growths	Polyps, nodules and cysts	25	1.5%	3.1%	3.1%	31.9%
6	Normals	Normal subjects	20	3.1%	10.9%	10.9%	43.3%

*Group Number

of [13], using a multilayer perceptron (MLP) to obtain a nonlinear discriminator between normal and dysphonic subjects' voices. The ANN was fed by the BBT nodes entropy values, and *cross validation* (CV) [16] for performance evaluation.

This paper proposes the use of WPT, BBA, and ANN to classify the voices in the database used in [13] and [14], in some of the pathology groups proposed in [10], now using the *multiple cross validation* (MCV) validation method and the *leave-one-out* (LOO) technique [16].

II. SYSTEM PROPOSED

A. Feature Extraction System: WPT, BBA, and BBT

According to [17], the wavelet packet (WP) coefficients can be calculated by

$$\lambda_{sf}(p) = \langle w_{sfp}, x \rangle = \int_{\Re} 2^{-s/2} w_f(2^{-s}t - p)x(t)dt \qquad (1)$$

where s is the scale index, p is the translation index, and ef is the frequency index; $f, s, p \in \Re$.

As described in [18], to perform the WPT, the wavelet function (w_f) used was *Symlet* 5 up to level $s = 5$. The BBA chooses the BBT based on the analysis of some cost function, all the other bases are pruned. In [13], [14] and this paper, the cost function used was *Shannon Entropy*. It is interesting to note that the morphology of the orthogonal BBT obtained, depends on the characteristics of the original signal x, and the wavelet and cost functions chosen to decompose the signal. This feature makes the WPT and BBT a preclassifier system, as discussed in [13].

B. Classification System: A Neural Network System

Classification is performed by a system composed by 6 ANN trained to identify each pathology group proposed in [10]. The use of 6 ANN, instead of one with 6 outputs, is justifiable by the fact that the nonlinearity degree in the input space can be evaluated, according to the number of hidden neurons [16], the number of hidden neurons increase if the input space heave more relevant characteristics for classification.

The entries were the BBT entropy values. These ANN are MLP with 3 layers with *tanh* as activation function and were trained by a BKP algorithm. The output can be evaluated by looking for the *i-esim* output of each i independent ANN. The threshold of acceptance of positive identification in the group i is: $y_i \geq 0, \mathrm{PAC} \in \mathrm{Group}_i$. The statistic tool used to select and test the models was MCV. The metric used to evaluate the evolution of training and validation was the *mean average percentual error* (MAPE). [16].

III. METHODS AND MATERIALS

A. Pathologic Voices Database

The voices in the database (as described in detail in [18]) were recorded at Dr. N. Steffen's office, at PUCRS Hospital. Each subject was diagnosed using videolaryingoscopy and videostrobolaryingoscopy. The sounds recorded were the emissions of the sustained vowel /a/. Each recording was digitized with a rate of 25 ksps [14], and the most stable part of the signal was chosen to be analyzed. A total of 64 voice samples were effectively used. This number of samples is realistic and compatible with the existing works. Besides, a 7th group containing all the pathologies not quoted in the other groups in [10], was added to test the generalization power of ANN.

B. WPT, BBA, and BBT

All the calculations were performed by Matlab v5.3. A full binary tree was generated by means of WPT with 64 nodes. Then the BBA was applied to obtain a BBT. All the nodes' entropy values were arranged in a 64-element row vector. The vector entropy's value used for the pruned nodes is zero. These vectors were arranged in matrix form, where each line of the matrix is related to each subject. The procedure adopted was to normalize each line by its maximum absolute number [18].

C. Training, Validation, and Testing of the Independent Pathologies Groups Neural Networks

In each MLP, the first layer contains 63 input neurons, receiving the normalized entropy node value of the BBT. The number of hidden neurons was chosen by experimentation, using the MCV method. The output layer contains just 01 neuron, which gave an error/hit output. In the MCV method, all the N available (patients) must be grouped in k subsets. Since there are 64 subjects, but 14 of them belonging to the 7th group, so there are 50 subjects $(N = 50)$ and k was 5. Then, subsets M_1 to M_5 were created with 10 subjects from all of the pathology groups. With the combination of these subsets someone will be able to create 5 new subsets, using 4 of them for training and one of them for validation. The number of neurons in the hidden layer was limited to half the number of input neurons [16]. ANN with different numbers of hidden neurons, was tested, choosing the topology which gave the lesser validation MAPE. After chosen the optimum number of neurons in the hidden layer, the ANN was trained again, with the LOO method.

IV. RESULTS, DISCUSSION, AND CONCLUSION

The final results, which specify the final configurations and the average generalization errors, are shown in Table I. A global success rate of 89.07% was obtained and this result is compatible with those found

in the literature: In [2], 69%; in [4], 54.79%; in [5], 95.1%; in [10], 62.33%; in [12], 91%; in [13], 67.2%; in [14], 84%; in [18], 86.89% were 13.11%, 4.92%, 4.92%, 6.56%, 8.20%, and 13.11% for groups 1 to 6, respectively. In assessing the subjects belonging to the 7th group, the system gave some incorrect classifications, patients from postoperatory were identified as belonging to the chronic laryngitis because of the extreme similarity of voice quality. Due to the high average G_{error}, causal by the small number of subjects, it is interesting that future works use bigger databases. Other suggestions are: assessing different cost functions in BBA and/or the use of radial base function ANN. All in all, it is important to remark that this system is able to classify the probability of some subject having a specific disease in physicians' office, using just a simple set of hardware/software. This approach can be very useful to help diagnosis.

REFERENCES

[1] R. J. Baken and R. F. Orlikoff, *Clinical Measurement of Speech and Voice*, 2nd ed. San Diego, CA: Singular Thomsom Learning, 2000.

[2] D. G. Childers and K. S. Bae, "Detection of laryngeal function using speech and electroglottographic data," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 1, pp. 19–25, Jan. 1992.

[3] B. Fritzell, "Inverse filtering," *J. Voice*, vol. 6, no. 2, 1992.

[4] M. O. Rosa, M. J. C. Pereira, and M. Grellet, "Adaptive estimation of residue signal for voice pathology diagnosis," *IEEE Trans. Biomed. Eng.*, vol. 47, no. 1, pp. 96–104, Jan. 2000.

[5] S. Hadjitodorov, B. Boyanov, and B. Teston, "Laryngeal pathology detection by means of class-specific neural maps," *IEEE Trans. Inf. Technol. Biomed.*, vol. 4, no. 1, pp. 68–73, Mar. 2000.

[6] C. E. Martinez and H. L. Rufiner, "Acoustic analysis of speech for detection of laryngeal pathologies," in *Proc. IEEE-EMBS Int. Conf. Inf. Tech. Appl. Biomed.*, 2000, pp. 2369–2372.

[7] R. T. Ritchings, M. Mcguillion, and C. J. Moore, "Pathological voice quality assessment using artificial neural networks," presented at the 2nd Int. Workshop Models and Analysis of Vocal Emissions for Biomedical Application, Firenze, Italy, 2001.

[8] M. Frohlich, D. Michaelis, and S. H. Werner, "Acoustic 'breathiness measures' in the description of pathologic voices," in *Proc. 1998 IEEE Int. Conf. Acoustics Speech and Signal Processing.*, vol. 2, pp. 937–940.

[9] M. O. Rosa, M. Greller, and A. Carvalho, "Signal processing and statistical procedures to identify laryngeal pathologies," in *6th Int. Conf. IEEE Circuits and Systems Society Electronics, Circuits and Systems.*, 1999.

[10] M. O. Rosa, C. J. Pereira, and A. Carvalho, "Evaluation of neural classifiers using statistic methods for identification of laryngeal pathologies," in *Proc. 5th Brazilian Symp. Neural Networks*, Brazil, Dec. 9–10, 1998, pp. 220–225.

[11] G. V. D. Wouwer, P. Scheunders, and D. V. Dyck, "Wavelet-Filvq Classifier for Speech Analysis," in *Proc. 13th Int. Conf. Pattern Recognition (ICPR'96)*, Antwerp, Belgium, 1997, vol. 4, p. 214.

[12] K. Umapathy *et al.*, "Discrimination of pathological voices using a time-frequency approach," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 3, pp. 421–421, Mar. 2005.

[13] A. Parraga, M. A. Zaro, and A. Schuck Jr., "Quantitative assessment of the use of continuous wavelet transform in the analysis of the fundamental frequency disturbance of the synthetic voice," *Med. Eng. Phys. PII: S*, vol. 2, no. 00050-4, pp. 1350–4533.

[14] A. Schuck, Jr., and A. Parraga, "On the use of the wavelet packet transform as a feature extractor for pathological voice assessement," presented at the IFMBE Proceedings EMBEC'02 2nd Eur. Medical and Biological Engineering Conf., Vienna, Austria, 2002.

[15] A. Schuck, Jr, L. V. Guimarães, and O. Wisbeck, "Dysphonic voice classification using wavelets packets transform and artificial neural networks," in *Proc. 25th Annu. Int. Conf. IEEE-EMBS*, Cancún, Mexico, 2003, pp. 2958–2961.

[16] S. Haykin, *Neural Network: A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1998.

[17] M. V. Wickerhauser, *Adapted Wavelet Packet Analysis From Theory to Software*. Wellesley, MA: AK Peters, Ltd., 1994.

[18] C. D. P. Crovato and A. Schuck, Jr, "Classificação de sinais de voz utilizando a transformada wavelet packet e redes neurais artificiais," in *Proc. III Congresso Latinoamericano De Engenharia Biomedica*, João Pessoa, Paraíba, Brazil, 2004, vol. 5, pp. 1027–1030, ISBN: 85-98739-01-04.

Use of Sample Entropy Approach to Study Heart Rate Variability in Obstructive Sleep Apnea Syndrome

Haitham M. Al-Angari* and Alan V. Sahakian

***Abstract*—Sample entropy, a nonlinear signal processing approach, was used as a measure of signal complexity to evaluate the cyclic behavior of heart rate variability (HRV) in obstructive sleep apnea syndrome (OSAS). In a group of 10 normal and 25 OSA subjects, the sample entropy measure showed that normal subjects have significantly more complex HRV pattern than the OSA subjects ($p < 0.005$). When compared with spectral analysis in a minute-by-minute classification, sample entropy had an accuracy of 70.3% (69.5% sensitivity, 70.8% specificity) while the spectral analysis had an accuracy of 70.4% (71.3% sensitivity, 69.9% specificity). The combination of the two methods improved the accuracy to 72.9% (72.2% sensitivity, 73.3% specificity). The sample entropy approach does not show major improvement over the existing methods. In fact, its accuracy in detecting sleep apnea is relatively low in the well classified data of the physionet. Its main achievement however, is the simplicity of computation. Sample entropy and other nonlinear methods might be useful tools to detect apnea episodes during sleep.**

***Index Terms*—Approximate entropy, heart rate variability, nonlinear signal processing, obstructive sleep apnea, power spectral density, sample entropy.**

I. INTRODUCTION

Heart rate variability (HRV) varies from wakefulness to sleep due to normal changes in the autonomic system activities. Sympathetic tone drops from wakefulness over nonrapid eye movement (NREM) sleep stages, while it shows an increase in REM sleep [1]. Parasympathetic activity increases from wakefulness over NREM sleep [2]. Spectral analysis of HRV is used to evaluate the activity of the autonomic nervous system. Low-frequency (LF) components (0.04–0.15 Hz) evaluate the sympathovagal balance while high-frequency (HF) components (0.15–0.4 Hz) estimate the parasympathetic tone related to respiratory rhythm [3]. In sleep disorders, impairment of the autonomic nervous system is observed. Studies of muscle sympathetic nerve activity (MSNA) have shown an increase in sympathetic tone in patients with obstructive sleep apnea syndrome (OSAS) during sleep and wakefulness [4], [5]. These findings were supported by results from spectral analysis. The HF power was significantly diminished and LF/HF ratio was enhanced in awake OSA patients, which indicates a drop in the parasympathetic tone associated with an increase in the sympathetic tone [6]. At the start of the apnea episode however, RR intervals lengthen which indicates an increase in the vagal activation [7]. There is also a noticeable increase in MSNA, peaking immediately prior to apnea cessation. Arousal at the termination of an apnea initiates a burst of sympathetic activity (associated with an increase in blood pressure and heart rate). This is observed as cyclical variation (progressive bradycardia followed by abrupt tachycardia) of the heart rate [8].

Nonlinear analysis of time series provides a parameter set that quantifies the characteristics of the system attractor even when the model

Manuscript received January 2, 2006; revised October 29, 2006. *Asterisk indicates corresponding author.*

*H. M. Al-Angari was with Northwestern University, Evanston, IL 60208 USA. He is now with the Electrical and Computer Engineering Department, King AbdulAziz University, P O Box 80204, Jeddah, 21589 Saudi Arabia (e-mail: hangari@kau.edu.sa).

A. V. Sahakian is with the Electrical Engineering and Computer Science Department, Evanston, IL 60208 USA (e-mail: sahakian@ ece.northwestern.edu).

Digital Object Identifier 10.1109/TBME.2006.889772