31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Fuzzy Logic Based Classification and Assessment of Pathological Voice Signals

B. Seyed Aghazadeh, H. Khadivi Heris

Abstract—**In this paper an efficient fuzzy wavelet packet (WP) based feature extraction method and fuzzy logic based disorder assessment technique were used to investigate voice signals of patients suffering from unilateral vocal fold paralysis (UVFP). Mother wavelet function of tenth order Daubechies (d10) was employed to decompose signals in 5 levels. Next, WP coefficients were used to measure energy and Shannon entropy features at different spectral sub-bands. Consequently, using fuzzy c-means method, signals were clustered into 2 classes. The amount of fuzzy membership of pathological and normal signals in their corresponding clusters was considered as a measure to quantify the discrimination ability of features. A classification accuracy of 100 percent was achieved using an artificial neural network classifier. Finally, fuzzy c-means clustering method was used as a way of voice pathology assessment. Accordingly, fuzzy membership function based health index is proposed.**

I. Introduction

UNILATERAL vocal fold paralysis (UVFP) occurs from a dysfunction of the recurrent or vagus nerve innervating the larynx and causes a characteristic breathy voice. UVFP most commonly occurs following a surgical iatrogenic injury to the vagus or recurrent laryngeal nerve resulting in glottal incompetence, either partial or complete, because of the poor or reduced vocal fold closure.

Physiological alterations of vocal cords cause unhealthy patterns of cord vibration and decrease in patient speech signal quality known as voice pathologies. Consequently, the detection of incipient damages to the cords is useful for improving the prognosis, treatment and care of such pathologies. Physicians often use invasive techniques, such as Endoscopy, to diagnose symptoms of voice disorders. It is, however, possible to identify disorders using certain features of speech signal in a non-invasive way [1]. Schuck *et al* [2] used Shannon entropy and energy features of wavelet packet decomposition, and employed the best basis algorithm for normal/pathological speech signal classification. Fonseca *et al* [3] employed mean square values of reconstructed signals in discrete wavelet transform sub-bands and least square support vector machine (LS-SVM) classifier for identification of signals from patients with vocal fold nodules and normal signals. Guido *et al* [4] tried different wavelets in the search for voice disorders. Mother wavelet of Daubechies with support length of 20 (db10) was found to be the best wavelet for speech signal analysis among commonly used wavelets. Behroozmand *et al* [5] used genetic algorithm for optimal selection of wavelet packet based energy and Shannon entropy features for identification of patient speech signal with unilateral vocal fold paralysis (UVFP). The results showed that the decomposition level of five is the best level to analyze pathological speech signals. Umapathy *et al* [6] used Local discriminant bases (LDB) and wavelet packet decomposition to demonstrate the significance of identifying discriminant WP subspaces.

Fuzzy wavelet packet based feature extraction method has been proposed by Li *et al* and has been applied to biological signal classification [7]. In contrast to the standard methods of feature extraction used in WPs, this method of discriminatory feature extraction from wavelet packet coefficients is based on the fuzzy set criterion. Yang *et al* [8] applied fuzzy wavelet packet method to feature extraction from electroencephalogram (EEG) signals. The results show that this method is promising for the extraction of EEG signals in brain-computer interfaces (BCIs). Seyed Aghazadeh [9] *et al* introduced a fuzzy membership-based criterion to quantify vocal fold disorders. Having clustered the training data into normal/pathologic classes, they showed that the membership of a sample in the normal class can be considered as a health index in order to assess the voice pathology.

This work aims to identify patients with UVFP by extracting an effective feature vector containing fewer features, higher discrimination accuracy, and lower order of computational complexity in comparison with the results obtained by genetic algorithm based optimal feature [5]. Additionally, fuzzy c-means method was used to quantify the amount of disorder.

The rest of this paper is organized as follows: methods and materials are reviewed in next section, section 2; the results of this study are discussed in section 3; section 4 presents the conclusions.

Babak Seyed Aghazadeh (email: seyedaghazab@vcu.edu) is a graduate student of department of mechanical engineering, Virginia Commonwealth University, Virginia, USA

Hossein Khadivi Heris (e-mail: Hossein.khadiviheris@mail.mcgill.ca) is a graduate student of department of mechanical engineering, McGill University, Montreal, Canada

II. MATERIALS AND METHODS

A. Wavelet packet transform

Recently, wavelet packets (WPs) have been widely used by many researchers to analyze voice and speech signals. There are many outstanding properties of wavelet packets that encourage researchers to employ them in many widespread fields. It has been shown that sparsity of coefficients' matrix, computational efficiency, and time-frequency analysis can be useful in dealing with many engineering problems. The most important, multiresolution property of WPs is helpful in voice signal synthesis.

The hierarchical WP transform uses a family of wavelet functions and their associated scaling functions to decompose the original signal into subsequent sub-bands. The decomposition process is recursively applied to the both low and high frequency sub-bands to generate the next level of the hierarchy. WPs can be described by the following collection of basis functions:

$$W_{2n}(2^{p-1}x-l)=\sqrt{2^{l-p}}\sum_{m}h(m-2l)\sqrt{2^{l-p}}W_n(2^p x-m) \quad (1)$$

$$W_{2n+1}(2^{p-1}x-l)=\sqrt{2^{l-p}}\sum_{m}g(m-2l)\sqrt{2^{l-p}}W_n(2^p x-m) \quad (2)$$

where p is scale index, l the translation index, h the low-pass filter and g the high-pass filter with

$$g(k)=(-1)^k h(l-k) \quad (3)$$

The WP coefficients at different scales and positions of a discrete signal (*f(m)*) can be computed as follows:

$$C^{p}_{n,k}=\sqrt{2^p}\sum_{m=-\infty}^{\infty}f(m).W_n(2^p m-k) \quad (4)$$

$$C^{p-1}_{2n,l}=\sum_{m}h(m-2l).C^{p}_{n,m} \quad (5)$$

$$C^{p-1}_{2n,l}=\sum_{m}g(m-2l).C^{p}_{n,m} \quad (6)$$

For a particular sequence of wavelet packet coefficients, energy in its corresponding sub-band can be computed as:

$$Energy_n=\frac{1}{N^2}\sum_{k=1}^{n}\left|C^{p}_{n,k}\right|^2 \quad (7)$$

The Shannon entropy as another extracted feature for classification of signals can be computed through the following formula:

$$Entropy_n=-\sum_{k=1}^{n}\left|C^{p}_{n,k}\right|^2\log\left|C^{p}_{n,k}\right|^2 \quad (8)$$

Due to the noise-like effect of irregularities in the vibration pattern of damaged vocal folds, the distribution manner of such variations within the whole frequency range of pathological speech signals is not clearly known. Therefore, it seems reasonable to use WP rather than discrete wavelets transform (DWT) to obtain more detailed sub-bands.

B. Fuzzy Set-Based Feature Selection Criterion

With fuzzy sets we allow any pattern x_k to belong to several classes to varying degrees. Assuming u_{ik} a membership grade of pattern x_k to class i we have:

$$u_{ik}=\left[\sum_{j=1}^{c}\left(\left\|x_k-v_i\right\|^2/\left\|x_k-v_j\right\|^2\right)^{1/(b-1)}\right]^{-1} \quad (9)$$

where c is the number of clusters, $v_i=\sum_{k\in A_i}x_k/N_i$ is the mean of class i, A_i is the set of indexes of the training patterns belonging to class i, N_i is the number of class i training patterns, $\|.\|$ is the Euclidean distance and $b>1$ is the fuzzification factor that modifies the shape of membership grades. For the labeled training patterns in feature space, X, we define a membership function based on the criterion $F(X)\in(0,N]$ to evaluate the classification ability of X as follows [7]:

$$F(X)\in\sum_{i=1}^{c}\sum_{k\in A_i}u_{ik} \quad (10)$$

The larger the values of *F(X)*, the higher the classification (discrimination) abilities of the feature space X. In fuzzy set based optimal WP decomposition for each labeled original signal a full WP decomposition to maximum level of five has been performed. The mother wavelet function is chosen to be the tenth order Daubechies (db10). Consequently, features (i.e. energy and entropy) of all signals in each node have been clustered using Fuzzy Clustering Method (FCM). Most discriminant nodes have been identified according to the parameter *F(X)*, and the signal energy and entropy in those nodes have been used to construct the feature vector applied to artificial neural network (ANN) classifier.

C. Fuzzy Set-Based Health Index

Fuzzy c-means (FCM) is used to cluster data from patients with severe UVFP and Normal signals. Having clustered the data set each point's membership in patient cluster (UVFP) can be interpreted as the amount of disorder in the signal. Also, the fuzzy membership of a given test sample in Normal cluster is an appropriate measure of similarity between the given sample and normal signals. Thus, the adopted similarity measure is proposed as a Health Index.

D. Database

Sustained vowel phonation samples from subjects from the Disordered Voice Database [10], model 4337, version 1.03 (Kay Elemetrics Corporation, Lincoln park, NJ) were used in this study. Subjects were asked to sustain the vowel /a/; voice recordings were made in a sound proof booth on a DAT recorder at a sampling frequency of 44.1 kHz.

III. Results and Discussion

Signals were decomposed by mother wavelet of tenth order Daubechies (db10) to 5 levels of decomposition. Next, energy and Shannon entropy at each decomposition sub-band were calculated. Then, fuzzy logic based feature extraction method was applied to construct an optimal feature vector of length 8 according to the nodes' discrimination ability which can separate normal and pathological (UVFP) voice signals. Table 1 shows the most discriminant nodes in terms of energy or entropy feature, with their discrimination abilities ($F(X)$/ number of data$\times$100).

In order to make a comparison, two feature vectors of lengths 8 were extracted: 1) with equal portion of discriminant energy and entropy nodes, 2) containing the most discriminant features. Approximately 65 percent of the data was used to train a feedforward backpropagation multilayer classifier neural network with 3 hidden layers, and the remaining 35 percent was set aside as the test and validation data. The feature vector obtained by considering nodes' discrimination ability resulted in classification accuracy of 100 percent, and showed a better performance in comparison with the feature vector including equal portion of energy and entropy nodes with the approximate classification accuracy of 96 percent.

Figure 1 shows the wavelet packet tree and the discriminant nodes listed in Table 1. As can be seen, selected sub-bands are distributed over the whole available frequency ranges. This means that the frequency range of such decisive features outgrows the spectral range of the speech signals in some of the conventional applications. In other words, this shows that pathological factors do not influence specific frequencies, which accentuates the role of WP decomposition with equal decomposition of both high and low frequencies. Also, due to the selection of discriminant

TABLE 1
PARTICIPATING NODES AND THEIR DISCRIMINATION ABILITY

Node	Energy	Entropy	Discrimination ability (%)
31	*		76.22
34	*		73.32
25	*		67.17
29	*		66.95
28	*		66.61
37		*	66.82
38		*	65.85
17		*	62.71
32		*	61.50
5		*	61.27

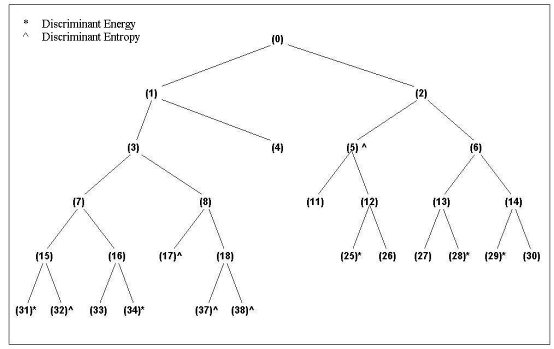

Fig. 1. The most discriminant nodes in terms of signal energy or entropy

features from WP decomposition the effect of noise was eliminated automatically and a lower computational cost was achieved. As a case in point, energy of each voice signal at most discriminant node (31) of WP decomposition was calculated. Figure 2 shows the efficiency and discrimination ability of the selected node.

In the following, fuzzy set-based evaluation of disorder degree was implemented to voice signals. In this sense, obtained signal features of normal individuals and patients suffering from severe unilateral vocal fold paralysis (UVFP) were considered to cluster the data into two classes of normal and UVFP. According to the effective selection of discriminant features using fuzzy wavelet packet based feature extraction method, the fuzzy memberships of all the UVFP training data in the UVFP class were higher than 0.5 and the fuzzy memberships of the normal voices in the normal class were higher than 0.5. Next, signals of patients with mild UVFP were chosen according to the diagnostic information on the database CD and passed on to the obtained FCM to calculate their memberships in the UVFP class. A two-dimensional projection of feature space is given in figure 3, namely, onto the energy at nodes 31 and 34.

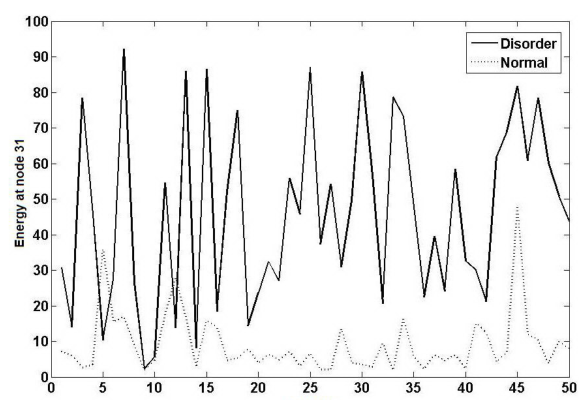

Fig. 2. Discrimination ability of the node (31)

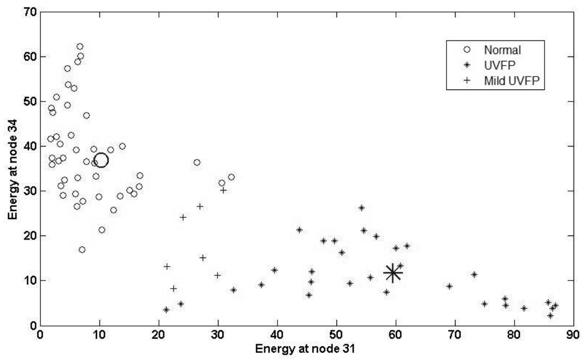

Fig. 3. A two-dimensional projection of feature space onto energy at nodes 31 and 34

As can be seen, the evaluating data of mild UVFP lie between two regions of normal and UVFP classes. The fuzzy membership grades of these data are presented in Table 2. According to this table, fuzzy memberships of mild UVFP samples in UVFP cluster are lower in comparison with severe UVFP samples. Considering a cutting value of 0.5, all mild UVFP data are classified in UVFP class. The obtained results not only present a satisfying usage of fuzzy membership grade as a health index in the assessment of voice pathologies, but also demonstrate the effectiveness and the discriminant nature of selected features using fuzzy wavelet packet based feature extraction method.

The simulation results show that fuzzy wavelet packet based feature extraction method, and neural network classifiers are effective tools in voice signal analysis. Also, the feature vector obtained considering nodes' discriminant ability outperforms the one including equal portion of nodes for the features of energy and entropy in terms of classification accuracy. Moreover, fuzzy membership-based health index was proven to be effective in the quantification of voice pathologies of UVFP.

IV. CONCLUSIONS

In this study, classification of voice signals into two groups of normal and patients with unilateral vocal fold paralysis (UVFP) has been presented in the context of fuzzy logic. Fuzzy wavelet packet based feature extraction method was utilized to find the optimal feature vector of length 8 from energy and Shannon entropy features in WP decomposition sub-bands. In the following, the obtained feature vector was passed on to an artificial neural network (ANN) classifier. The simulation results show that by using the fuzzy wavelet packet based optimal feature vector of length 8 applied to an ANN classifier, the classification accuracy of 100 percent can be achieved, which despite its relatively short length, outperforms feature vectors obtained by other methods. Furthermore, fuzzy membership-based health index was proposed to quantify the amount of disorder. It has been shown that fuzzy membership grade of a given voice sample in the patient cluster can be used as a measurement of disorder quantity.

TABLE 2
HEALTH INDEXES OF SAMPLE DATA

Membership in UVFP cluster	Health index	Description
0.23	77%	Normal
0.31	69%	Normal
0.42	58%	Normal
0.56	44%	Mild UVFP
0.57	43%	Mild UVFP
0.59	41%	Mild UVFP
0.60	40%	Mild UVFP
0.63	37%	Mild UVFP
0.78	22%	UVFP
0.84	16%	UVFP
0.91	9%	UVFP
0.96	4%	UVFP

ACKNOWLEDGMENT

This research was supported by grants number 2R01DC005788-05A2 from National Institute on Deafness and other Communication Disorders. This work was also supported by Canadian Institute of Health Research Grant No. MOP 84495.

REFERENCES

[1] M. N. Viera, F. R. Mclnnes, M. A. Jack, "Robust FO and Jitter estimation in the Pathological voices," in Proc. of ICSLP96, Philadelphia, 1996, pp.745 -748.

[2] A. Schuck Jr., L. V. Guimaraes, J. O. Wisbech, "Dysphonic voice classification using wavelet packet transform and artificial neural network," in: Proc. of the 25th IEEE Annual EMBS International Conference, Cancum,Mexico, September 2003, pp. 2958-2961.

[3] E. S. Fonseca, R. C. Guido, A. C. Silvestre, J. C. Pereira, "Discrete wavelet transform and support vector machine applied to pathological voice signals identification," in: Proc. Of the seventh IEEE International Symposium on Multimedia, ISM'05.

[4] R. C. Guido, J. C. Pereira, E. Fonseca, F. L. Sanchez, L. S. Vierira, "Trying different wavelets on the search for voice disorders sorting," in: Proc. of the 37th IEEE International Southeastern Symposium on System Theory, 2005, pp. 495-499.

[5] R. Behroozmand, F. Almasganj, "Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis," Computers in Biology and Medicine, Vol. 37, pp. 474-485, 2007.

[6] K. Umapathy, S. Krishnan, "Feature analysis of pathological speech signals using local discriminant bases technique," Med. Bio. Eng. Comput., Vol. 43, pp. 457-464, 2005.

[7] D. Li, W. Pedrycz and N. Pizzi, "Fuzzy wavelet packet based feature extraction method and its application to biomedical signal classification," IEEE Transactions on biomedical engineering, Vol. 52 no.6 pp. 1132-1139, 2005.

[8] B. Yang, G. Yan, T. Wu and R. Yan, "Subject-based feature extraction using fuzzy wavelet packet in brain-computer interfaces," Signal processing, 87 (2007) 1569–1574.

[9] B. S.Aghazadeh, H. Khadivi, M. Nikkhah-Bahrami, "Nonlinear analysis and classification of vocal disorders," in: Proc. of the 29th IEEE Annual EMBS International Conference, Lyon, France, August 2007, pp. 6199-6202.

[10] Disordered voice database (CD-ROM), Version 1.03, Massachusetts Eye and Ear Infirmary, Kay Elemetrics Corporation, Boston, MA, Voice and Speech Lab., October 1994.