Research on Objective Auscultation of TCM Using Wavelet Packet and Support Vector Machine

Jianjun Yan, Xiaojing Shen

***Abstract*—The goal of this study is to provide objective analysis and quantitative research for the auscultation in Traditional Chinese Medicine (TCM) using wavelet packet transform (WPT) and support vector machine (SVM). Based on WPT, the voice signals are decomposed into six layers wavelet coefficients. This paper proposed Shannon entropy as feature parameter extracted from wavelet packet coefficients to make analysis of health, qi-vacuity and yin-vacuity subjects. Then the feature values used as vectors were put into SVM to be trained and predicted, and the classification results showed that the overall accuracy of the health group, qi-vacuity group and yin-vacuity group reached to 80.84%. It is proved that our method is effective for auscultation research of TCM.**

***Index Terms*—auscultation signals, wavelet packet transform, Shannon entropy, support vector machine; traditional Chinese medicine.**

I. INTRODUCTION

The auscultation, which is an important part in four diagnosis of Traditional Chinese Medicine (TCM), is the approach to diagnose a disease by listening to patients' speech, respiration, cough. Every voice sent out by human vocal organs includes plenty of information about personality characteristics and physiological, psychological state, and it is able to reflect the functional activities of zang-fu organs and abundance or declining of the qi, blood and body fluids. As a result, when one person falls ill, some abnormal sounds usually appear in his body due to exogenous pathogenic factors or functional disorders. Then physicians of TCM can detect the vocal changes to confirm the property and position of illness, pathogenic or normal vicissitude state, etc.

As computer science and information technology are drawn into the field of TCM, researchers have proposed some methods in the area of objective auscultation. X. H. Zhang proposed several modern methods on the auscultation [1]: aerodynamics diagnosis, throat-dynamic mirror diagnosis, the sound spectrum analysis graphics diagnosis, X-ray diagnosis and EMG diagnosis; X. M. Mo, etc. made voice frequency spectrum analysis and research on the voice of cough patients making use of digital sonagraph [2]; X. L. Wang analyzed the tuberculosis patients' voice using electronic computer, and explored the differences of amplitude perturbation of the vowel /a:/ [3]; J. J. Yan and H. J. Wang made studies on nonlinearity of the vowel /a:/ signals of healthy persons and

patients with deficiency syndrome applying delay vector

Jianjun Yan, Center for Mechatronics Engineering, East China University of Science and Technology, Shanghai, China, 86-21-64252074.

Xiaojing Shen, Center for Mechatronics Engineering, East China University of Science and Technology, Shanghai, China.

variance [4, 5]; Z. X. Yang proposed four novel acoustic parameters, the average number of zero-crossings, the variations in local peaks and valleys, the variations in first and second formant frequencies, and the spectral energy ratio, to analyze and identify the characteristics among non-vacuity, qi-vacuity, and yin-vacuity subjects [6]. These methods provided a good basis for objective auscultation. However, there are no standard indicators of the auscultation used in clinic practices. So the key for the development of auscultation is to extract the effective feature parameters by taking advantage of modern signal processing technology. Considering the voice signal is usually non-stationary and time-varying, wavelet packet transform can be used to analyze the voice signal based on time-frequency domain transform and the feature of the multi-scale and multi-resolution decomposition and reconstruction. It makes decomposition in both the low-frequency and high-frequency at the same time, and self-adaptively determines the signal resolution in different frequency bands [7].

So this paper applied wavelet packet transform for analyzing voice signals of health, qi-vacuity, and yin-vacuity subjects. Then Shannon entropy parameter was proposed to make a quantitative analysis of auscultation based on 6-layers wavelet packet coefficients. Finally, statistical analysis was made to obtain effective feature values which were inputted into SVM classifiers, and classification analysis was performed to realize automatic identification of health, qi-vacuity and yin-vacuity subjects.

II. MATERIAL AND METHODS

A. Material

In a physiological sense [8], qi-vacuity means deficiency of vital energy and weakness of the qi. Yin-vacuity is the pathology manifestation of insufficiency of the yin aspect and depletion of body fluids and blood. In this paper, we mainly analyze and discern the voice signals of health, qi-vacuity and yin-vacuity subjects. There are a total of 181 subjects including 27 cases of health, 38 cases of yin-vacuity and 116 cases of qi-vacuity. In addition, our research cooperation partner, TCM Syndrome Lab in Shanghai University of Traditional Chinese Medicine supplied all samples, which were collected from the outpatient department of the Subsidiary Hospital of Shanghai University of Traditional Chinese Medicine. The high-powered microphone (Brand AKG, Model HSD 171, frequency response range of 60Hz-17kHz, sensitivity of 1mv/Pa (-60dBV), impedance of 600ohms) and a 16-bit A/D converter card (Brand CME Xcorpio, frequency response range of 20Hz-20 kHz, dynamic range of 100dB) are used to record voice signals. An endpoint detection algorithm was adopted to remove the non-voice portions of the leading and the trailing of each utterance. Finally, the information of subjects and dialectic result of TCM auscultation were correctly gathered and high-fidelity voice signals were collected through the combination of the software system developed by Visual C++ 6.0 and hardware devices.

In this research, the vowel /a:/ was chosen as the utterance for each subject to pronounce[9], the main reason is that subjects can easily pronounce /a:/ without any guidance. In addition, when one pronounces the vowel/a:/, there is no obstruction in the cavity, pronouncing flow is smooth, and pronouncing vibration generates periodical waveform. So the vowel /a:/ was selected to make speech analysis, it can decrease the interference and reduce complexity and uncertainty of the analysis. Subjects' pronunciation time was fixed about one second, and the sampling frequency is at 16KHz (16025Hz). Therefore, the most common voice parameters, pitch frequency or pitch variations, the power of the voice, and the speed of speech, are less useful in analyzing the changes of the voice. Entropy is a statistical measure of randomness that can be used to characterize the uncertainty of the signal. So Shannon entropy was presented as characteristic parameters to analyze the changes of the voice of the health, qi-vacuity and yin-vacuity subjects.

B. Wavelet packet

Wavelet transform is a sort of time-frequency analysis, which uses different scales to obtain the best resolution of time-domain and frequency-domain in different parts of the signal. But the resolution analysis only produces further decomposition in the low-frequency part so that the high-frequency part is not subdivided. Wavelet packet transform provide a more precise decomposition for signal analysis. It can produce further decomposition in the high-frequency part, so frequency bands are subdivided synchronously in the low-frequency and high-frequency part. In addition, wavelet packet can select self-adaptively the signal resolution in different frequency bands. Thus wavelet packet improves the time-frequency resolution.

The common wavelet function includes Harr wavelet, Daubechies wavelet, Coiflet wavelet, Morlet wavelet, Mexican Hat wavelet and so on. Through compared with the analysis results of db, coif, sym wavelet functions, db4 wavelet function with high energy concentration is ultimately chosen for the analysis of the voice signals of health, qi-vacuity and yin-vacuity subjects. The voice signals are decomposed into 6-layers wavelet packet coefficients, by which the voice signals are reconstructed. Then the characteristic parameters of every frequency band are extracted [10].

In order to effectively analyze the subjects' voice signals, the analysis programs were developed under the MATLAB environment [11]. Entropy is mainly used to measure the regularity of the information. When the entropy is smaller; the regularity of information is stronger. So Shannon entropy is selected and extracted as auscultation parameter from wavelet packet coefficients. Then Shannon entropy parameters of every frequency band are used to make quantitative analysis of the voice of the health, qi-vacuity and yin-vacuity subjects. The Shannon entropy parameter is defined as follows:

$$E = \sum_i E(s_i) = -\sum_i s_i^2 \log(s_i^2) \qquad (1)$$

where the signal S is a group of orthogonal basis on the coefficient, and Si is the i coefficients of S in an orthogonal basis, while the entropy E is the value stack up by a certain transform of the coefficients of per orthogonal basis. In addition, because of different pronunciation time and sound intensity of different samples, the parameter values can not be used directly. So the ratio of the entropy of every decomposed node and the entropy of the root node is taken as feature parameter for analysis in this paper. After the processing of wavelet packet decomposition and reconstruction for the voice signals of health, qi-vacuity and yin-vacuity subjects, feature parameters are computed for every frequency band which corresponds with decomposed node.

C. Support vector machine

Support vector machine (SVM) is used to classify the features of the three groups. SVM is a novel regression technique and a new class of training algorithm based on the statistical learning theory [12]. The principle of this theory is based on the structural risk minimization (SRM). The SVM carry out the model of recognition by determining the hyperplane of separation with the maximum distance to the narrowest points of the formation positioning. These points are called the vectors of support. If the data are not separable linearly in the input space, one can apply a non-linear transformation. In this section, the model which is linearly separable is applied from a mathematical point of view. In the long run, we select SVM as classifier for the recognition of subjects.

In order to avoid the randomness of one-time training and predicting, the training and predicting are repeated many times. In this process, 70% samples are randomly selected for training, and the remaining 30% of samples are used to predict. After training and predicting, the mean value of classification accuracies is taken as the final accuracy.

III. Results and Discussion

After the analysis of the results above mentioned in section two which are all derived from the contrast of Shannon entropy parameters, the characteristic parameters having significant differences are merged and put into SVM for identification and classification of the voice signals of health, qi-vacuity, and yin-vacuity subjects. The average prediction accuracies of classification results are shown in Table 3.1, and the impressive results were gained as followed:

(1) The average prediction accuracy of health and yin-vacuity group is the highest in the above four groups, and its average accuracy reached 91.38%; but the classification prediction accuracy of qi-vacuity and yin-vacuity group is low, the average prediction accuracy is only 78.728%.

(2) The classification prediction accuracy of health and qi-vacuity and yin-vacuity group is 80.84% in a general level.

(3) The overall average classification accuracy of four groups respectively reached 81.82%, 78.72%, 91.38% and 80.84%, the overall effect shows that the classification analysis using SVM is helpful for identification of the voice signals of health, qi-vacuity, and yin-vacuity subjects.

Table 1 Classification results by using SVM

NO	*Group*	*Training samples*	*Testing samples*	*Average accuracy*
G1	Health	17	10	81.82%
	Qi- vacuity	81	35	
G2	Qi- vacuity	81	35	78.72%
	Yin- vacuity	26	12	
G3	Health	17	10	91.38%
	Yin- vacuity	26	12	
G4	Health	17	10	80.84%
	Qi- vacuity	81	35	
	Yin- vacuity	26	12	

Due to the complexity and non-stationarity of the voice signal, we used a data acquisition card and auscultation system developed for acquiring signals correctly. Consequentially, the hardware device can provide an accurate and high-fidelity voice signals. In our study, the vowel /a:/ was chosen as the utterance for each subject to pronounce, it can decrease the interference and reduce complexity and uncertainty of the auscultation analysis. The results of wavelet packet analysis of all samples can provide useful information for TCM physicians.

Based on the results of classification analysis in Table 1, it seems that the voice parameters extracted guarantee an effective classification. We should note that these classification results must be considered preliminary through using Shannon entropy parameter. The subjects were simply classified qualitatively into four groups in our experiment, but the results of classification analysis were particularly impressive. This research is a beginning in an ongoing effort to modernize the auscultation in TCM, and more attempts will be done in the future.

IV. Conclusion

This research presented a quantitative analysis of auscultation based on wavelet packet analysis and support vector machine. The auscultation signals were decomposed by wavelet packet transform, and the Shannon entropy parameters were extracted from wavelet packet coefficients of every node. Then, these parameters of the voice signals of health, qi-vacuity and yin-vacuity subjects were identified automatically using SVM, and the classification results are satisfied.

In the future, we will collect more clinical voice signals and develop the auscultation system for assisting clinical practices.

Acknowledgment

The authors would like to thank Shanghai University of Traditional Chinese Medicine for supplying all samples.

References

[1] X. H. Zhang, *Concise sound medical*, People's Health Publishing House, Beijing,1985, pp. 33-54.

[2] X. M. Mo, “Preliminary study on making use of sound ring instrument for the diagnosis of deficiency syndrome of the lung with cough,” *Journal of Traditional Chinese Medicine Research*, 1987, (3), pp. 43.

[3] X. L. Wang, “Frist exploration of the smell diagnosis on tuberculosis,” Summary of Graduate Student Thesis, 1992; (3):29.

[4] H. J. Wang, J. J. Yan, Y. Q. Wang, R. Guo, F. F. Li, C. M. Xia, T. C. Ma, “Digital technology for objective auscultation in Traditional Chinese Medical diagnosis,” Proceedings of 2008 International Conference on Audio, Language and Image Processing, Shanghai July 7–8, 2008.

[5] J. J. Yan, Y. Q. Wang, H. J. Wang, F. F. Li, C. M. Xia, R. Guo, T. C. Ma. “Nonlinear analysis in TCM acoustic diagnosis using delay vector variance,” The 2nd International Conference on Bioinformatics and Biomedical Engineering (2008), Shanghai May16-18, 2008.

[6] Z. X Yang, “Objective auscultation for traditional Chinese medical diagnosis using novel acoustic parameters,” *Computer Methods and Programs in Biomedicine*, 2000, vol 62, pp. 99–107 .

[7] A. H.Tewfik, D. Sinha, P. Jorgensen, “On the Optimal Choice of a Wavelet for Signal Representation,” *IEEE Trans on Information Theory*, 1992, 38(2); pp. 747-765。

[8] C. H. Horng, “The principles and methods of diagnostics,” The Illustrations of Chinese Medicine, Lead Press, Taipei, Taiwan, 1993 (Chapter 4).

[9] L. Z. Hou, D. M. Han, “Selection of the vowel sound in the throat sound detection,” *Journal of Audiology and speech Diseases*, 2002:(10)16.

[10] S. H. Liang, I. Hartimo, “A feature extraction algorithm based on wavelet packet decomposition for heart sound signals,” Proc of IEEE-SP intre.Symp, USA, IEEE, 1998:93-96

[11] C. H. Dong, *Theory and application of the wavelet toolbox in Matlab*, National Defense Industry Press, Beijing: 2004

[12] Z. Q. Bian, X. G. Zhang, *Pattern Recognition*, 2nd ed, Beijing: Tsinghua University Press, 2000 (in Chinese).

Jianjun Yan received his PhD from Huazhong University of Science and technology. From 2004 he worked in Huadong University of Science and technology. His research focus is in machine learning and biological signal processing.

Xiaojing Shen received his MD from Huadong University of Science and technology. From 2012 he worked in ZTE Company. His research focus is in signal processing.