Nonlinear analysis of auscultation signals in Traditional Chinese Medicine using Wavelet Packet Transform and Approximate Entropy

Jianjun Yan* and Yong Shen

Center for Mechatronics Engineering,
East China University of Science and Technology,
Shanghai, PR China
E-mail: jjyan@ecust.edu.cn
E-mail: shenyong198601058@163.com

Yiqin Wang* and Fufeng Li

Center for TCM Information Science and Technology,
Shanghai University of TCM,
Shanghai, PR China
E-mail: wangyiqin2380@sina.com
E-mail: fufeng_lee@hotmail.com
*Corresponding authors

Chunming Xia

Center for Mechatronics Engineering,
East China University of Science and Technology,
Shanghai, PR China
E-mail: cmxia@ecust.edu.cn

Rui Guo and Chunfeng Chen

Center for TCM Information Science and Technology,
Shanghai University of TCM,
Shanghai, PR China
E-mail: guoruier@sina.com
E-mail: chenchunfeng79@163.com

Zhongyan Gu and Xiaojing Shen

Center for Mechatronics Engineering,
East China University of Science and Technology,
Shanghai, PR China
E-mail: gu_zhong_yan@126.com
E-mail: handsomesxj@163.com

Abstract: The distinctive characteristic of Wavelet Transform (WT) is that it can well characterise the local information of signals in time-frequency domain, and Wavelet Packet Transform (WPT) has a more subtle decomposition method than WT. The purpose of this paper is to analyse the auscultation signals in Traditional Chinese Medicine (TCM) utilising WPT and Approximate Entropy (ApEn). In this paper, a new scheme was presented for analysing the Auscultation Signals consisted of qi-deficient, yin-deficient and normal people. In the first stage, voice signal were decomposed into approximation and detail coefficients using WPT. Then the ApEn values of these signals were computed based on these coefficients. The differences of the ApEn values and the meaning of which for signals among three kinds of samples were discussed. Finally the conclusion can be drawn that the distributions of ApEn in different frequency ranges for all signals of three kinds of samples have their special characteristics. The ApEn values for three kinds of samples were used as the feature vectors for Support Vector Machine (SVM) classifier and some impressing results of classifications can be obtained.

Biographical notes: Jianjun Yan is an Associate Professor at Center for Mechatronics Engineering, East China University of Science and Technology. He received his PhD from the Huazhong University of Science and Technology in 2003. Dr. Yan's main research areas include image processing, biomedical signal processing, artificial intelligence and the modernisation auscultation of TCM.

Yong Shen is a Graduate Student of East China University of Science and Technology. His research interest is in the area of biomedical signal processing.

Yiqing Wang is a Professor at Center for TCM Information Science and Technology in Shanghai University of TCM. She received her doctor degree from the Shanghai University of TCM in 2002. She engages in the study of objective TCM diagnosis and clinical & laboratory study on treatment of asthma with Chinese herbs.

Fufeng Li is an Associate Professor at Center for TCM Information Science and Technology in Shanghai University of TCM. She received her doctor degree from the Shanghai University of TCM in 2007. Her research interests are in the areas of the objectivation of TCM four diagnostic methods, image processing and biomedical signal processing.

Chunming Xia is an Associate Professor at Center for Mechatronics Engineering, East China University of Science and Technology. He received his PhD from the University of Glasgow in 2003. His main research areas include biomedical signal processing and fault detection and diagnosis.

Rui Guo is Researcher at Center for TCM Information Science and Technology in Shanghai University of TCM. She received her Master Degree from Tongji

University in 2004. Her research interests are in the areas of image processing, signal processing; especially image fusion and biomedical signal processing.

Chunfeng Chen is a Graduate Student of Shanghai University of TCM. Her research interest is in the area of the objective auscultation of TCM.

Zhongyan Gu is a Undergraduate Student of East China University of Science and Technology. His research interest is in the area of biomedical signal processing.

Xiaojing Shen is a Graduate Student of East China University of Science and Technology. His research interest is in the area of biomedical signal processing.

1 Introduction

The auscultation of traditional Chinese medicine (TCM) plays an important role in Auscultation and Olfaction, which has a distinctive analysis way and comprehensive theory. According to the theory of TCM, sound, the outward sign of vital activities, can reflect the functional activities of zang-fu organs, the wax and wane of the qi, blood and body fluids, and the conditions of their physiology and psychology. Some abnormal sound usually appears in their bodies for the invading of exogenous pathogenic factors and dysfunction of their bodies when people get diseases (Wang, 2006). Therefore the sound which includes a large sum of information has influences on the organisation function of variable organs. Consequently it's of vital significance to identify the locations and the properties of the disease. And listening to sounds can not only diagnose the local lesion, but also diagnose the pathological changes of the internal organs.

The traditional auscultation of TCM mainly depend on the auditory of experienced Chinese physician to identify the asthenia and sthenia, and the visceral lesions. Nevertheless, there are a series of factors which may have influence on the results of diagnosis:

- the diagnosis is not objective enough due to the difference of physician' hearing

- it is very difficult to discern the information of diseases effectively for physician who lacks clinical experience

- physician' hearing and memory deterioration may affect diagnostic objectivity as the increasing of their ages

- it is easy for physician to feel fatigue when his diagnosis last too much time, the environment with too much noise will easily cause confused, misdiagnosis, and missed diagnosis.

All the factors mentioned above usually lead to inconsistent even totally different results. In addition, the way of identifying the sound of patients is more qualitative rather than quantitative. Therefore, the results of traditional auscultation of TCM are unpersuasive. In this condition, some objective auscultation using a lot of modern voice processing methods for traditional Chinese medical diagnosis is urgently needed to be developed.

In recent years, the researches on objective auscultation for traditional Chinese medical diagnosis and related works are not so many and the research of auscultation of TCM is still rarely studied. However, the research results achieved by previous researches should still be mentioned. Gao and Li (2007) did some systematical research, exploited, collected, and textually research the theory of Five Zang-organs harmonise Pitch in emperor's internal classic. He published four monographs about the theory of Five Zang-organs harmonise Pitch. On clinical application aspect, Mo et al. (1998) has used the Plotter to record the sonagram analysis of patients with cough. Qi-deficiency, yin-deficiency and TCM excessive syndrome were chosen as samples to do the research. And the result is that the diagnostic specificity of group vowel and sound of cough is higher. Zhang (1985) proposed the physiologic voice diagnosis and the pathological voice diagnosis on the auscultation of TCM, he concluded that the physiologic voice diagnosis is the premise for curing the pathological voice diagnosis. Five modern methods about auscultation: aerodynamics diagnosis, laryngostroboscope diagnosis, sound spectrum analysis diagnosis, X-ray electromyography diagnosis were presented. Yang et al. (1997) proposed voice analysis for qi-deficiency patients. He compared the qi-deficiency and the non-deficiency using the parameters: average number of zero-crossings, the variations in local peaks and valleys, and he found that the parameters used have some differences for the samples: qi-deficiency and non-deficiency. Zhang et al. (1998) studied on the voice analysis of yin-deficiency patients and proposed three parameters: zero-crossing average numbers, variations on peaks and valleys and variations on formant frequencies as index to identify the qi-defiency sydrome. The result showed that the accuracy using this three parameters reached 88.24%. Two kinds of methods: traditional methods and nonlinear methods were illuminated briefly by Wang et al. (2008). Delay Vector Variance method was applied to detect the nonlinearity of the vowel /a/ signals of normal persons and patients with deficiency syndrome by Yan et al. (2008). All the research mentioned above has acquired some achievements. However, a unified scheme for justifying standard has not reached an agreement due to the lack of information database. In addition, linear is the idealisation and approximation of nonlinear, and the characteristic of nonlinear for voice signal is realised by researchers gradually. There are weak points on the traditional methods of Speech Signal Processing. Therefore it is of great importance to analyse the voice signal based on the nonlinear methods. So there is a long way to go from the assisting clinical practice for auscultation before some Satisfying results about objective auscultation were acquired.

The voice signal is a kind of nonstationary random signals. Its frequency changes along with the time changes. Fourier Transform (FT) is the traditional way of dealing with the voice signal, while it does not do well in local analysis. The Short-Time Fourier Transform (STFT) improves the drawbacks to some extent, while it does not have a high time resolution in the high frequency domains together with a high frequency resolution in the low frequency domains. Wavelet analysis is a kind of time-frequency analysis technique. Its non-stationary signal analysis technology has reached a mature stage. The WT has particular advantages for characterising signals in different scales in time as well as frequency domains. It uses a series of oscillating functions with different frequencies as window function to deal with transient signals by modulating its resolution in different time-interval. Higher time resolution was applied at higher frequencies, and higher frequency resolution was applied at lower frequency components in WT. This feature of WT is more useful for analysing and characterising the signals which include the non-stationary components (Chiu et al., 1998). The SVM is a machine

learning technique, which originate form the statistics theory (Vapnik, 1998) and is applied for the classification of the voice samples. The advantage of SVM is that it can model highly non-linear systems and the special properties of the decision surface ensure very good generalisation. Our research is aimed at applying the modern techniques of voice analysis to analyse and identify the samples of the normal, qi-deficiency, and yin-deficiency. The goal of our research is to propose a new scheme using ApEn and WPT for the purpose of making a quantitative analysis of auscultation in TCM diagnosis. We will take the clinical data as basis to extract the feature parameters of voice signal by using the WPT in order to identify the three samples. First of all, voice signals were decomposed into four layers of wavelet decomposition coefficients, then the ApEn values for all these sub-band coefficients were computed. Among these ApEn values, different sub-bands may possess different effects for identifying different groups of samples.

2 Technical background

2.1 DWT and WPT for individual signal processing

Generally, wavelets are well crafted to have specific properties which make them available for signal processing. Wavelets can be combined, using a "shift, multiply and sum" technique named convolution to pick up information from the unknown signal. The continuous wavelet transform (CWT) of a signal, $x(t)$, is the integral of the signal multiplied by scaled and shifted versions of a wavelet function ψ and is defined by:

$$\text{CWT}(a, b) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right) \mathrm{d}t \tag{1}$$

where parameters a, b are so called the scaling (reciprocal of frequency) and time localisation or shifting parameters, respectively. If the scales and shifts are selected based on powers of two, so-called dyadic scales and positions, then the wavelet analysis will be much more efficient. Such analysis is obtained from the DWT which is defined as:

$$\text{DWT}(j, k) = \frac{1}{\sqrt{|2^j|}} \int_{-\infty}^{\infty} x(t) \psi\left(\frac{t - 2^j k}{2^j}\right) \mathrm{d}t \tag{2}$$

where a and b are replaced by 2^j and $2^j k$, respectively. In practical uses, the algorithm of multi-resolution signal decomposition proposed by Mallat (1986, 1989) is generally used. The implementation of the orthogonal wavelet decomposition can be defined as:

$$C^j(n) = C^{j-1}(n) * \tilde{H} = \sum_{k=-\infty}^{+\infty} \tilde{h}(k) C^{j-1}(n - 2^j k) \tag{3}$$

$$D^j(n) = C^{j-1}(n) * \tilde{G} = \sum_{k=-\infty}^{+\infty} \tilde{g}(k) C^{j-1}(n - 2^j k) \tag{4}$$

where $\tilde{H}$ and $\tilde{G}$ are the filters corresponding to the wavelet function. $C^j(n)$ is the low frequency part with the frequency lower than 2^j, and $D^j(n)$ is the frequency between 2^j and 2^{j+1}. The orthogonal decomposition analysis can be carried out using a pair of compact support functions which can be represented by two sequences $\{Pk\}$. The sequences contain all the information about the scaling function $\phi[k]$, and the sequence $q_k = (-1)^k p_{-k+1}$ which characterises its corresponding wavelet function $\Psi[k]$. Based on these sequences, we introduce a family of functions $\{\mu_n\}, n = 2l$ or $2l+1$, then the symbol $l = 0,1,...,N$, called a wavelet packet which is a generalization of the orthogonal wavelet $\Psi[k]$, which is used to improve the performance of wavelets for time–frequency localization (Chui, 1992; Gutitkez et al., 2001)
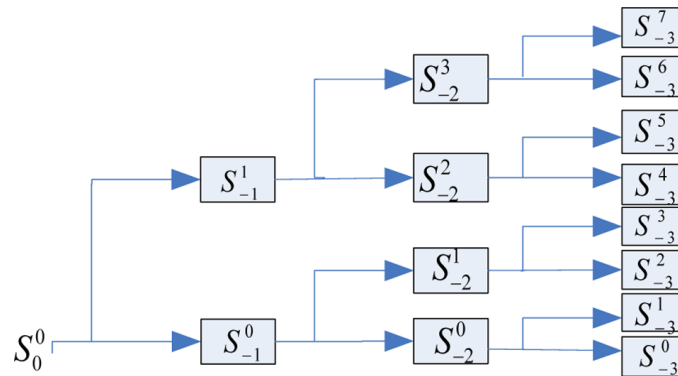
$$\mu_{2l}[k] = \sum_k p_k \phi[2k - n] \tag{5}$$

$$\mu_{2l+1}[k] = \sum_k q_k \phi[2k - n]. \tag{6}$$

Because of the capability of decomposing not only the low frequency of the signal but also the higher-frequency octaves, Wavelet packets are specially used when the better frequency localisation is searched.

2.2   Wavelet Packet Decomposition (WPD)

Wavelet packet decomposition is a wavelet transform where the signal is passed through more filters than DWT. In the DWT, each level is calculated by passing only the previous approximation coefficients through a high and low pass filters. However in the WPD, both the detail and approximation coefficients are decomposed. Given a finite energy signal whose scaling space is assumed as S_0^0, S_0^0 can be decomposed into small subspaces S_j^n by wavelet packet transform in dichotomous way (see Figure 1).

Figure 1   The structure of Wavelet Packet Transform S_j^{n-1} shows the nth subspace the *j*th resolution level (see online version for colours)

The dichotomous way is defined by the following recursive scheme.

$$S_{j+1}^n = S_j^{2n} \oplus S_j^{2n+1} \quad j \in Z; n \in Z_+ \tag{7}$$

where $(j \leq 0)$ is the resolution level and $\oplus$ denotes orthogonal decomposition S_{j+1}^{n}, S_{j}^{2n} and S_{j}^{2n+1} are three close spaces corresponding to $S_n(t)$, $S_{2n}(t)$ and $S_{2n+1}(t)$, $S_n(t)$ satisfies the following equation (Coifman and Wickerhauser, 1992):

$$S_{2n}(t) = \sqrt{2} \sum_{k \in Z} h(k) S_n(2t - k)$$

$$S_{2n+1}(t) = \sqrt{2} \sum_{k \in Z} g(k) S_n(2t - k)$$

(8)

$h(k)$ and $g(k)$ are the coefficients of the low-pass and the high-pass filters, respectively. The sequence of function $\{S_n\}$ ($n = 0, 1, \ldots, \infty$), which is generated from a given function S_0 by equation (11), is called wavelet packet basis function.

2.3 Approximate Entropy (ApEn)

The most valued statistics like the median, mean and standard deviation have been widely used in voicing and speech analysis, though they are seldom considered as simple variability statistics. But for linear systems e.g., periodic signals, mean and standard deviations are sufficient. However, it is suggested that an adaptive method of quantification should be require by more complicated signals (Manickam et al., 2005). ApEn describes the rate of producing new information, Approximate entropy is a measure that regularity or predictability of a time series (Pincus, 1991). A low ApEn indicates a less complicated signal while a large ApEn value indicates more complex signals ApEn, unlike Shannon's entropy, takes into account the temporal order of points in a time sequence and is therefore a preferred measure of randomness or regularity. For an N sample time series $\{u(i): 1 \leq i \leq N\}$, given m, form vector sequences X_1^m through X_{N-m+1}^m as:

$$X_i^m = \{u(i), u(i+1), \ldots, u(i+m-1)\}, \quad i = 1, \ldots, N-m+1$$

(9)

where the symbol m means the length of compared window. For each $i \leq N - m + 1$, let $c_i^m(r)$ be $(N - m + 1)^{-1}$ times the number of vectors X_j^m within r of X_i^m. By defining:

$$\phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} \ln C_i^m(r)$$

(10)

where ln is the natural logarithm, Pincus (1991) defined the parameter:

$$\text{ApEn}(m, r) = \lim_{N \to \infty} [\phi^m(r) - \phi^{m+1}(r)].$$

(11)

2.4 Support Vector Machine (SVM)

The SVM is a useful machine learning technique that has been successfully applied in the classification area. Classifying data is a common task in machine learning. In most cases, the data to be classified is linearly nonseparable but nonlinearly separable. The nonlinear support vector classifier will be used. The main idea is to transform the original data into a high-dimensional features space. Although SVM only treat the two-class problem, the multi-class classification problem can always be converted into the two-class problem. The two-class problem and the multi-class classification problem are both

involved in this paper. So the 'all vs. all' method was used to transfer the multi-class classification problem into two-class problem. Thus, though the classifier is a hyperplane in the high-dimensional feature space, it may be non-linear in the original input space (Ding and Dubchak, 2001).

In order to construct a nonlinear support vector classifier, the product (x, y) is replaced by a kernel function $K(x, y)$. The following are some common used kernel functions:

Polynomial (homogeneous):

$$k(x, x') = (x \cdot x')^d. \tag{12}$$

Polynomial (inhomogeneous):

$$k(x, x') = (x \cdot x' + 1)^d. \tag{13}$$

Radial Basis Function:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2), \quad \text{for } \gamma > 0. \tag{14}$$

Gaussian Radial basis function:

$$k(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right). \tag{15}$$

Hyperbolic tangent:

$$k(x, x') = \tanh(\kappa x \cdot x' + c), \quad \text{for some (not every) } \kappa > 0 \text{ and } c > 0. \tag{16}$$

The goal of SVM is to produce a model which predict target values of data instances in the test set, for which only the attributes are given. The following decision function is applied to determine which class the sample belongs to.

$$f(x) = \text{sgn}\left(\sum_{i=1}^{l} y_i a_i^* k(x_i, x_j) + b^*\right). \tag{17}$$

The parameters a_i^* and b^* is the optimum solution for specific.
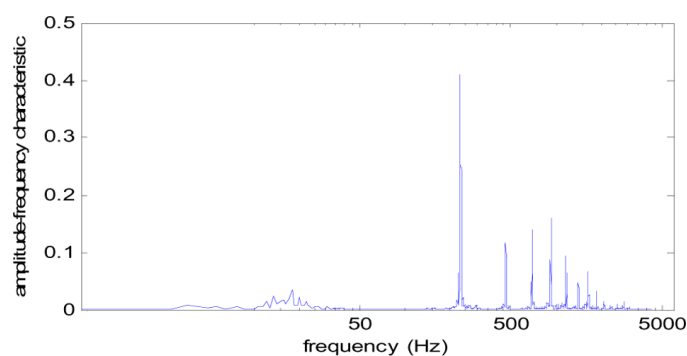
3   Methods

3.1   Clinical data

All the data showed above is provided by the TCM Syndrome laboratory of Shanghai University of TCM who is the partner of our research. WPT and SVM were used to analyse and identify the voice signals collected from the normal people, yin-deficient and Qi-deficient patients. The time-frequency feature parameters were extracted to characterise the asthenia and sthenia of the patient's voice. And the samples which will be analysed are comprised of voice signals from people of different age and sex. The detailed information is listed in Table 1.
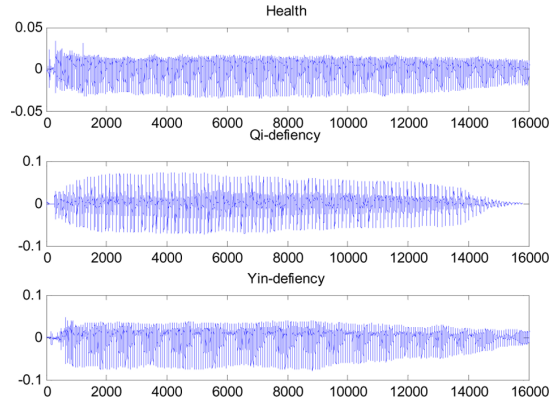
Table 1 Samples information

Normal		Qi-deficiency	Yin-deficiency	Sum
Sample number	27	116	38	181
Man	9	39	11	59
Woman	18	77	27	122
Max. age	54	76	80	–
Min. age	19	6	18	–
Average age	*24.9*	42.4	52.1	–

The voice is recorded using the high-performance microphone (the band is AKG model HSD171) and a 16-bit A/D converter which is connected to a computer. The frequency response range of the microphone is 60 Hz~17 kHz, its sensitivity is 1 mv/Pa (–60 dBV), impedance is 600 ohms. And the sample frequency is 16 kHz. All the samples of voice were collected by the acquisition system developed based on Visual C++ 6.0. The endpoint detection algorithm was applied to remove the non-voice portions of the leading and the trailing of each utterance. The vowel /a/, /i/, /u/, /e/, /o/ were usually chosen as patient's utterance to pronounce. In this research, the vowel /a/ was chosen as the utterance. Each subject produced a stable phonation of a sustained English vowel /a/ lasting about one second. The reason for choosing the vowel /a/ is that it is easy for whether patients or normal people to pronounce. In addition, the voice can be easily sent out by both the normal and the deficient. In addition, the vocal organ is not abuttal and there is no obstacle in cavity when some one is sending out the vowel. The pronunciation flow is unblocked, and periodical waveform can be produced in this process. Therefore, recently the vowel /a/ was mainly chosen to be used as voice parameters (G.H. Ling, 2004). The spectrum of the vowel /a/ is shown in Figure 2.

Figure 2 Spectrum of the vowel /a/ (see online version for colours)

The voice signal was collected in a quiet room with an echo suppression in which the noise index meets with the requirement. Figure 3 is the illustrative examples of the normal group, qi-deficiency group and yin-deficiency group. They are the representative signals from these three groups. Nevertheless, more objective analysis is needed to illustrate the differences among the three samples.

Figure 3   Original signals of normal, qi-deficiency and yin-deficiency (see online version
for colours)

3.2   *Preprocessing of voice signal using WPT*

In the first stage of processing of sample identification, the voice signals including
three kinds of samples were analysed using WPT. Considering the length of data
needed by ApEn (500–1000), four levels wavelet packet decomposition was applied
as the preprocessing step for all the qi-deficiency, yin-deficiency and normal subjects.
Under this level, the minimum length of the signal is 500 if the length of original signal
is 8000 under the sample frequency 16 kHz.

3.3   *The ApEn computation*

In the second stage, ApEn values of approximation and detailed coefficients at each level
of the wavelet decomposition were computed. ApEn of a time series is closely related to
two parameters including embedding dimension m and tolerance r were set to 2, 0.2 times
the standard deviation of the data respectively based on suggestions by Pincus. ApEn
values for S_j^{n-1} (means the nth subspace the jth resolution level) coefficients were
computed for all three kinds of samples. The values of ApEn for all three kinds of
samples were showed in Tables 2–4.

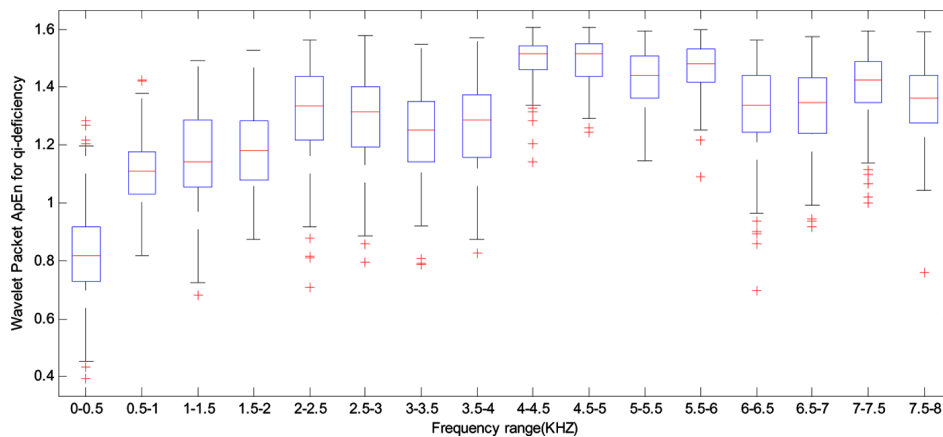**Table 2**   ApEn for the first and the second sub-bands coefficients

Frequency range (kHz)	Qi-deficiency		Yin-deficiency		Normal	
	mean (ApEn)	*SD*	*mean* (ApEn)	*SD*	*mean* (ApEn)	*SD*
S_{-1}^0 (0–4)	6.25E-01	1.36E-01	6.13E-01	1.54E-01	5.79E-01	1.11E-01
S_{-1}^1 (4–8)	1.20E+00	2.20E-01	1.15E+00	2.72E-01	1.26E+00	2.29E-01
S_{-2}^0 (0–2)	8.18E-01	1.73E-01	7.97E-01	1.91E-01	7.19E-01	1.41E-01
S_{-2}^1 (2–4)	1.16E+00	1.93E-01	1.10E+00	2.00E-01	1.24E+00	1.83E-01
S_{-2}^2 (4–6)	1.38E+00	1.94E-01	1.39E+00	1.51E-01	1.43E+00	1.37E-01
S_{-2}^3 (6–8)	1.31E+00	1.63E-01	1.26E+00	2.25E-01	1.36E+00	1.45E-01

4 Experimental results
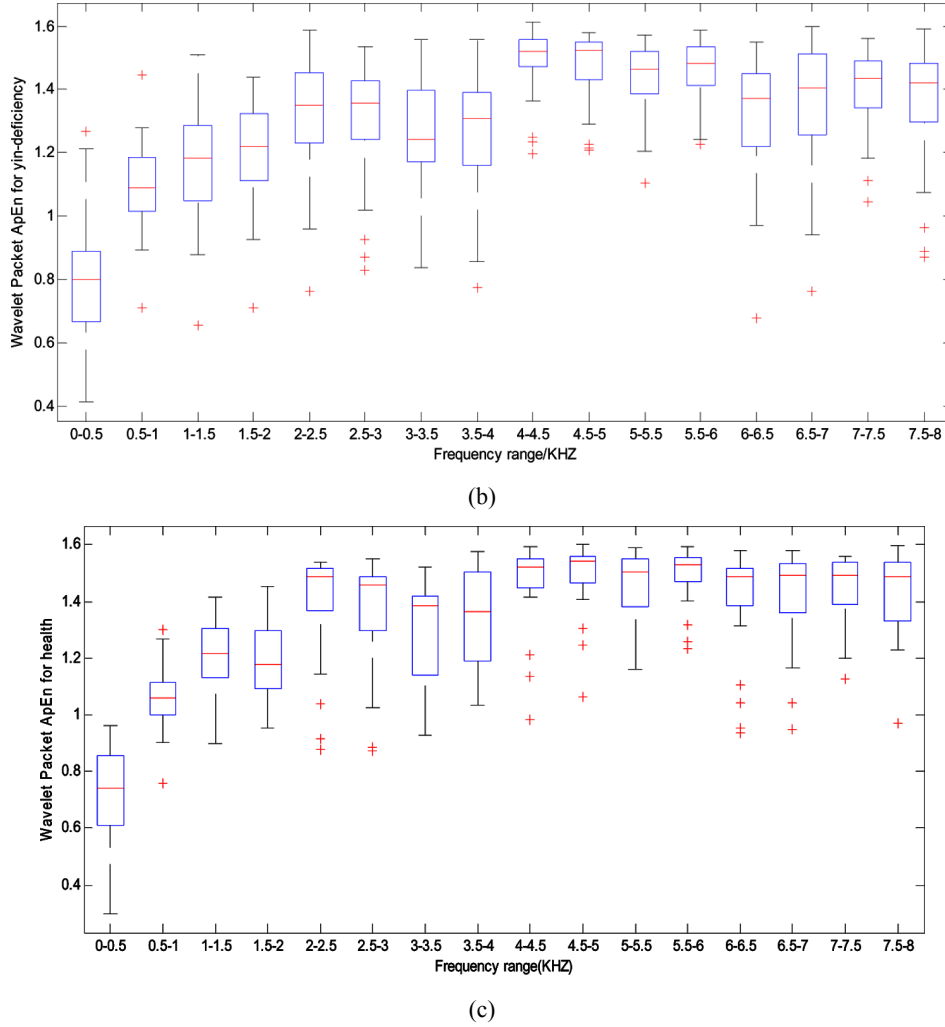
4.1 Results on ApEn values for APT coefficients

Voice signal from qi-deficiency, yin-deficiency and normal samples were decomposed into sub-bands using WPT. If the length of original signal is 8 kHz under the sample frequency 16 kHz, the frequency ranges for these sub-bands were as follows S_{-1}^{n} (0–8 kHz, the frequency interval is 4 kHz, $n = 0,1$), S_{-2}^{n} (0–8 kHz, the frequency interval is 2 kHz, $n = 0, 1, 2, 3$), S_{-3}^{n} (0–8 kHz, the frequency interval is 1 kHz, $n = 0, 1, 2, …, 7$), S_{-4}^{n} (0–8 kHz, the frequency interval is 0.5 kHz, $n = 0, 1, 2, …, 15$). ApEn values of the approximation and detail coefficients at these sub-bands were computed for the all three kinds of samples. Figure 3 shows the box plot of ApEn values of the fourth level of wavelet packet coefficients that is S_{-4}^{n} ($n = 0, 1, 2, …, 15$). The top and bottom of each rectangular box denote the 25th and 75th percentiles, respectively, with the median shown inside the box. From Figure 4, basically, difference among three kinds of sample can be observed easily. The numbers of upper and lower outliner in qi-deficiency people were more than yin-deficiency and normal people in the frequency range 0–0.5 kHz, which means that the ApEn data fluctuates sharply in this frequency range. The median for qi-deficiency people was larger than the other people. From frequency range 0.5–4 kHz, the number of lower outliner in normal was larger than the other people. In the frequency range 6–6.5 kHz, the length between upper and lower hinge of (b) is larger than (a) and (c), which means the variation range of people with yin-deficiency is larger than people with qi-deficiency and normal people. In addition, the ApEn data fluctuates more slightly than people with yin-deficiency and normal people. From frequency range 1–8 kHz, there were no upper outliners in all three kinds of samples.

Figure 4 Boxplot of ApEn in different frequency ranges for people with: (a) qi-deficiency; (b) yin-deficiency and (c) normal persons (see online version for colours)

(a)

Figure 4   Boxplot of ApEn in different frequency ranges for people with (a) qi-deficiency; (b) yin-deficiency and (c) normal persons (see online version for colours) (continued)

(b)

(c)

The mean ApEn values of all coefficients for three kinds of samples of four level sub-bands were listed in Tables 2–4. As the tables show, 2, 4, 8, and 16 ApEns corresponding to these four level sub-bands respectively. After the feature selection, these data can be used as the vectors to classify three kinds of samples. For better displaying the changes of ApEn values from low frequencies to high frequencies and the differences of ApEn among three kinds of sample, Figure 4 was drew. From Figure 5, the following conclusions can be obtained:

- In general, the ApEn value is on the increase form low frequencies to high frequencies. For people with qi-deficiency and yin-deficiency, ApEn values between them are not significant in both the low frequencies and high frequencies.

- In the frequency range 0–1.5 kHz, the ApEn values for normal people are lower than which for people with qi-deficiency and yin-deficiency. In the frequency range 1.5–8 kHz, the ApEn values for normal people become higher than which for the other two kinds of people.

- The differences among three kinds of samples display highly in the frequency ranges 1.5–4 kHz and 6–7.5 kHz.

According to Tables 2–4, good classification results can be speculated if ApEn values for $S_{-1}^0, S_{-1}^1, S_{-2}^3, S_{-3}^1, S_{-3}^2, S_{-3}^3, S_{-3}^6, S_{-3}^7, S_{-4}^3, S_{-4}^4, S_{-4}^5, S_{-4}^6, S_{-4}^7, S_{-4}^{12}, S_{-4}^{13}$ and S_{-4}^{14} were chosen as input vectors for classifier. Considering the meaning of ApEn, in the frequencies mentioned above, the degree of complexity for these frequencies were higher than the other frequencies, and the degree of complexity for group normal in these frequencies were higher than the other two groups. In other words, the probability of a new model appears relatively larger. The possible reason for this is that the spectral distribution for normal patient is more flat than deficient patients, which means it is close to equilibrium state. If a signal is more close to equilibrium state, the degree of complexity of it is higher.

Table 3   ApEn for the third sub-bands coefficients

Frequency range (kHz)	Qi-deficiency		Yin-deficiency		Normal	
	mean (ApEn)	SD	mean (ApEn)	SD	mean (ApEn)	SD
S_{-3}^0 (0–1)	8.47E-01	1.73E-01	8.28E-01	1.75E-01	7.04E-01	1.77E-01
S_{-3}^1 (1–2)	1.05E+00	1.73E-01	1.03E+00	1.83E-01	1.03E+00	2.18E-01
S_{-3}^2 (2–3)	1.28E+00	1.66E-01	1.28E+00	1.74E-01	1.36E+00	1.70E-01
S_{-3}^3 (3–4)	1.25E+00	1.65E-01	1.21E+00	1.92E-01	1.28E+00	1.92E-01
S_{-3}^4 (4–5)	1.49E+00	9.69E-02	1.47E+00	1.01E-01	1.46E+00	1.64E-01
S_{-3}^5 (5–6)	1.44E+00	1.26E-01	1.45E+00	8.94E-02	1.48E+00	1.12E-01
S_{-3}^6 (6–7)	1.34E+00	1.52E-01	1.31E+00	2.20E-01	1.42E+00	1.55E-01
S_{-3}^7 (7–8)	1.37E+00	1.50E-01	1.35E+00	1.91E-01	1.42E+00	1.67E-01

Table 4   ApEn for the fourth level sub-bands coefficients

Frequency range (kHz)	Qi-deficiency		Yin-deficiency		Normal	
	mean (ApEn)	SD	mean (ApEn)	SD	mean (ApEn)	SD
S_{-4}^0 (0–0.5)	8.24E-01	1.74E-01	7.97E-01	1.87E-01	6.99E-01	1.99E-01
S_{-4}^1 (0.5–1)	1.11E+00	1.09E-01	1.10E+00	1.24E-01	1.06E+00	1.20E-01
S_{-4}^2 (1–1.5)	1.16E+00	1.64E-01	1.17E+00	1.68E-01	1.20E+00	1.34E-01
S_{-4}^3 (1.5–2)	1.18E+00	1.35E-01	1.21E+00	1.46E-01	1.19E+00	1.42E-01
S_{-4}^4 (2–2.5)	1.31E+00	1.70E-01	1.33E+00	1.74E-01	1.40E+00	1.87E-01
S_{-4}^5 (2.5–3)	1.29E+00	1.52E-01	1.31E+00	1.70E-01	1.36E+00	1.90E-01

Table 4   ApEn for the fourth level sub-bands coefficients (continued)

Frequency range (kHz)	Qi-deficiency		Yin-deficiency		Normal	
	mean (ApEn)	SD	mean (ApEn)	SD	mean (ApEn)	SD
S_{-4}^6 (3–3.5)	1.24E+00	1.49E-01	1.26E+00	1.81E-01	1.30E+00	1.65E-01
S_{-4}^7 (3.5–4)	1.26E+00	1.53E-01	1.27E+00	1.87E-01	1.34E+00	1.62E-01
S_{-4}^8 (4–4.5)	1.49E+00	7.68E-02	1.49E+00	8.86E-02	1.47E+00	1.41E-01
S_{-4}^9 (4.5–5)	1.49E+00	8.20E-02	1.49E+00	9.39E-02	1.49E+00	1.17E-01
S_{-4}^{10} (5–5.5)	1.43E+00	1.02E-01	1.44E+00	1.07E-01	1.46E+00	1.14E-01
S_{-4}^{11} (5.5–6)	1.46E+00	8.78E-02	1.46E+00	1.01E-01	1.49E+00	9.27E-02
S_{-4}^{12} (6–6.5)	1.32E+00	1.60E-01	1.32E+00	1.82E-01	1.41E+00	1.79E-01
S_{-4}^{13} (6.5–7)	1.33E+00	1.39E-01	1.36E+00	1.95E-01	1.41E+00	1.81E-01
S_{-4}^{14} (7–7.5)	1.40E+00	1.22E-01	1.40E+00	1.16E-01	1.45E+00	1.15E-01
S_{-4}^{15} (7.5–8)	1.35E+00	1.30E-01	1.36E+00	1.72E-01	1.43E+00	1.41E-01

Figure 5   ApEn for the fourth level sub-band (see online version for colours)

4.2   Classification using SVM

In this paper, the libsvm software was used to identify the auscultation signal. The feature parameters with remarkable differences were chosen as the input vectors meeting with the format of the libsvm. The type of the SVM is C-SVC, and the RBF function was chosen as the kernel function for nonlinear training and testing after numerous experiments. The parameters *c* and *r* were obtained using cross-validation and grid-search. In the process of classification, 145 instances in the 181 were chosen as training data and the other 36 instances were chosen for testing data. Table 5 shows the classification results using SVM.

Table 5   Prediction Accuracies using SVM

Group numbers		*Training sample number*	*Predicting sample number*	*Correct number*	*Accuracy (%)*
1	Normal	22	5	3	89.3
	Qi-deficiency	93	23	22	
2	Normal	22	5	4	85.0
	Yin-deficiency	30	8	7	
3	Qi-deficiency	93	23	19	80.6
	Yin-deficiency	30	8	6	
4	Qi-deficiency	93	23	22	
	Yin-deficiency	30	8	5	83.3
	Normal	22	5	3	

5   Conclusions

In recent years, the objective auscultation of TCM has become more and more important as the constant happens of misdiagnosis or missed diagnosis on clinical diagnosis. However the researches on discrimination analysis of TCM for voice signals make slow progress. Due to the non-stationary and complexity of Chinese auscultation signals, voice parameters which can better differentiate non-deficient and deficient subjects have not been found. Nowadays, the research on analysis of TCM Symptoms has long been a hot issue. The vowel /a/ is chosen as the utterance for each patient to pronounce in this study. The parameter ApEn has widely been used in many fields once proposed, especially in the complexity analysis of short physiology signal time series. However, so far it has not been reported that the ApEn has been used in the voice signals of auscultation analysis. In this study, WPT and ApEn were combined together to analyse the voice signals of qi-deficiency, yin-deficiency, normal people. The sub-bands of voice signals through WPD revealed that the ApEn values for voice signals of all three kinds of samples in different frequencies have different classification results. For example, there were larger differences for ApEn values among three kinds of samples in the frequency ranges 1.5–4 kHz and 6–7.5 kHz than in other frequency, and the difference of ApEn values between health people and the other people are distinct. All these conclusions mentioned above are important for the objective auscultation of TCM.

Acknowledgements

References

Chiu, C.C., Chang, H.H. and Yang, C.S. (1998) 'Objective auscultation for traditional Chinese medical diagnosis using novel acoustic parameters', *Computer Methods and Program in Biomedicine*, pp.99–107.

Chui, C.K. (Ed.) (1992) *An Introduction to Wavelets*, Academic Press, London.

Coifman, M.V. and Wickerhauser, M.W. (1992) 'Entropy-based algorithms for best basis selection', *IEEE Trans. Inform. Theory*, Vol. 38, pp.713–718.

Ding, C.H. and Dubchak, I. (2001) 'Multi-class protein fold recognition using support vector machines and neural networks', *Bioinformatics*, Vol. 17, pp.349–358.

Gao, Y.T. and Li, J.W. (2007) *Five Zang-organs Harmonize Pitch-Emperor's Internal Classic-the Theory and Technology that have Lost for 2000 Years*, China Ancient Books Press, Beijing.

Gutitkez, J., Alchntara, R. and Medina, V. (2001) 'Analysis and localization of epileptic events using wavelet packets', *Medical Engineering and Physics*, pp.623–631.

Ling, G.H. (2004) A Study of Automatic Analysis for Listening Examination of Chinese Medicine, Chung Yuan Christian University, pp.12–13.

Mallat, S.G. (1986) 'A theory for multi-resolution signal decomposition: the wavelet representation', *IEEE Trans. Pattern Anal. Machine Intell.*, pp.674–693.

Mallat, S.G. (1989) 'Multi-resolution approximations and wavelet orthonormal bases of L2 (T)', *Trans. Amer. Math. Soc.*, pp.69–87.

Manickam, K., Moore, C.J., Willard, T. and Slevin, N. (2005) 'Quantifying aberrant phonation using approximate entropy in electrolaryngography', *Speech Communication*, pp.312–321.

Mo, X.M., Cai, G.X. and Zhang, J.L. (1998) 'The research of clinical experiment about the objective auscultation', *Chinese Journal of Basic Medicine in Traditional Chinese*, pp.37–39.

Pincus, S.M. (1991) 'Approximate entropy as a measure of system complexity', *Proc. Natl. Acad. Sci., USA*, Vol. 88, pp.2297–2301.

Vapnik, V.N. (1998) *Statistics Learning Theory*, Wiley, New York.

Wang, H.J., Yan, J.J. and Wang, Y.Q. (2008) 'Digital technology for objective auscultation in traditional chinese medical diagnosis', *Proceedings of 2008 International Conference on Audio, Language and Image Processing*, Shanghai, pp.1, 2.

Wang, Y.Q. (2006) *Diagnostics of Traditional Chinese Medicine*, Higher Education Press, Beijing.

Yan, J.J., Wang, Y.Q. and Wang, H.J. (2008) 'Nonlinear analysis in TCM acoustic diagnosis using delay vector variance', *The 2nd International Conference on Bioinformatics and Biomedical Engineering*, Shanghai, pp.1–5.

Yang, J.R., Zhang, H.R. and Qiu, C.G. (1997) *Research of Modernization of Auscultation and Olfaction Traditional Chinese Medicine-Voice Analysis of the Qi-deficiency Patients*, Medical University Hospital, Graduate institute of Chinese medical science, TaiZhong, China.

Zhang, H.R., Qiu, C.G. and Yang, Z.X. (1998) 'The research of the modernization of traditional chinese medicine auscultation and olfaction-the voice analysis for qi-deficiency and yin-deficiency patient', *Engineering Science and Technology and the Applications of Chinese and Western Medicine*, pp.15–17.

Zhang, N.H. (1985) *Concise Voice Medicine*, People's Medical Publishing House, Beijing.