An Objective Evaluation of Pathological Voice Based on Complexity Method

Yingji GONG 1, Weiping HU *2

1Department of Mechanical Engineering, Lushan College, Guangxi University of Science and Technology, Liuzhou Guangxi 545616, P.R. China
2 Electronic Engineering College, Guangxi Normal University, Guilin, Guangxi 541004, P.R. China

Abstract — In order to evaluate the Pathological Voice’s signal by parameters objectively, in this paper we adopt HNR, Lempel-Zvi Complexity (LZC) and C0 complexity to analyze Pathological Voice’s signal. Pathological voice and normal voice that classify as 4 degrees by the GRBAS system Total hoarseness degree (G). The result shows that, the existing acoustic parameter HNR is directly proportional to subjective hearing characteristics of Total hoarseness degree, HNR is more effective to distinguish the pathological voice from normal voice; but in the same time, not only the C0 is on good terms with subjective point of view, but also the C0 complexity shows the voice’s hoarse degree more effectively, better able to evaluate Pathological Voice objectively than HNR. It belongs to the voice painless and contactless identification method, which will use as the effect of voice medical process objective evaluation method of sound quality.

Keywords - Pathological voice, harmonics to noise ratio, Lempel-Zvi complexity, C0 Complexity, voice’s hoarse degree

I. INTRODUCTION

Quantitative clinical voice analysis is useful to laryngeal function examination[1]. Various kinds of laryngeal diseases are caused When laryngeal function is abnormal, then format pathological voice. Hoarseness is the cardinal symptom of laryngeal disease, in recent years, objective assessment of speech pathology, making automated speech pathology detection is an active field of research[2-4]. But less attention was paid to Complexity measure for pathological voice signal. The humans’ pronunciation system is a complicated system, thus the voice signal from the pronunciation system must have complexity. As the complexity measure of pathological voice signal, In this paper, compare to the traditional acoustic parameters harmonics-to-noise ratio (HNR)[7], we selected the Lempel-Zvi complexity [8] which measure the system randomness complexity and C0 Complexity which measure non regular components of signal [9].

II. METHODS

A Frequency Domain Relative Harmonics-To-Noise

HNR is the effective parameter reflects the hoarse degree and measure the acoustic parameters of voice quality in traditional, Hu Weiping etc had validated the frequency domain relative HNR the frequency-domain relative signal-to-noise ratio as the acoustic parameters on the effectiveness of normal, pathological voice analysis. Frequency domain relative HNR is computed as follow:

1). Compute the signal’s FFT;

2). Calculations of the resonance peak signal-to-noise ratio $P(k)$, relative value, mapped to each 400hz: band space;

$$\text{HNR}_e = 10\log_{10}(P(k)/P_e(k)) = 10\log_{10}(P(k) - 10\log_{10}P_e(k))$$

B Lempel-Zvi complexity

The Lempel–Ziv complexity analysis of a time series $\{x_i\}$, $i = 1, 2, 3, \cdots, N$, can be carried out as follows.

First encode the time series by constructing a sequence $S$ of the characters 0 and 1 written as: $s(i)$, $i = 1, 2, 3, \cdots, n$, according to the rule,

$$s(i) = \begin{cases} 0 & \text{if } x_i < x_* \\ 1 & \text{if } x_i > x_* \end{cases}$$

Here $x_*$ is a chosen threshold. We shall choose the absolute mean value of the time series to be the threshold.

Here the symbol $p$ is used to denote the operation of deleting the last character in a sequence.

1). Let $P$ and $Q$ denote two subsequences and $PQ$ be the concatenation of $P$ and $Q$. In addition, let $PQp$ be a sequence derived from $PQ$ after its last character is deleted. Here the symbol $p$ is used to denote the operation of deleting the last character in a sequence.

2). Let $v(PQp)$ denote the vocabulary of all different subsequences of $PQp$.

3). At the beginning of the computation, the complexity counter is set to be one. i.e., $c(n) = 1$, with $P = s(1), Q = s(2)$, $PQ = s(1), s(2)$, and $PQp = s(1)$. The Lempel–Ziv complexity measure (LZC) algorithm is listed as following steps,(see Gomez et al. 2006; Zhang et al. 2001).

1). Let $P$ and $Q$ denote two subsequences and $PQ$ be the concatenation of $P$ and $Q$. In addition, let $PQp$ be a sequence derived from $PQ$ after its last character is deleted. Here the symbol $p$ is used to denote the operation of deleting the last character in a sequence.

2). Let $v(PQp)$ denote the vocabulary of all different subsequences of $PQp$.
In general, \( P = (s(1), s(2), \ldots, s(r)) \) and \( Q = (s(r + 1), s(r + 2), \ldots, s(t)) \). Then \( P \cap Q \) is the subsequence of \( P \cap Q \), not a new sequence.

5). Renew \( Q \) to be \( s(r + 1), s(r + 2) \) and see if \( Q \) belongs to \( v(PQp) \).

6). Repeat the previous steps until \( Q \) does not belong to \( v(PQp) \). In this case, \( Q = s(r + 1), s(r + 2), \ldots, s(r + i) \) is not a subsequence of \( PQp \). If \( Q \) belongs to \( v(PQp) \) then \( Q \) is a subsequence of \( PQp \) for a stochastic system, the value of \( C_0 \) is close to 1.

7). Thereafter, \( P \) is renewed to be the sequence: \( P = (s(1), s(2), \ldots, s(r), s(r + i - 1)) \), with \( Q = s(r + i - 1) \). 8). Repeat the previous steps until \( Q \) is the last character. At this point, the complexity counter \( c(n) \) gives the number of different subsequences in \( P \), which is equal to the Lempel–Ziv complexity (LZC). Note that the above procedure uses only two operations: comparison and accumulation, and this makes the computation of \( c(n) \) easy to implement.

A Lempel–Ziv complexity measure which is independent of the length of the sequence can be obtained by normalizing \( c(n) \) with respect to the complexity of a random binary sequence: \( b_n = n/\log_2 n \). The normalized LZC measure is given by

\[
c(n) = \frac{c(n)}{n} \log_2 n \quad (3)
\]

This normalized measure is used for estimating the LZC values in the present study. For a deterministic system that is composed of well defined cycles such as a periodic motion, \( C(n) \) has a value much smaller than 1, whereas for a stochastic system, the value of \( C(n) \) is close to 1.

**C0 Complexity**

C0 Complexity was carried by in a previous paper (Chen et al 1998), F.J Gu and J.H Xu provide the concept of C0 Complexity (Gu et al 2008). C0 Complexity which is aimed to divide original signal into two divergent components, irregular component and regular component are defined as the ratio of irregular component to the original signal. More robust estimation of biomedical signals which are characterized as intense nonlinear and instability can be gotten rapidly by C0 Complexity, such as EEG signal, so we adopt it to analyze voice signal which is nonlinear and instability. And this process obviously avoids over coarse grainning. It can be computed using the following steps:

1. First step: Compute Fast Fourier Transforms (FFT) for original time sequence \( x(t) \) and get it amplitude spectrum \( X(k) = \text{FFT}(x(t)) \).

2. Second step: calculate mean value of amplitude spectrum \( X(k) \), set the part which is larger than mean value as regular component of the sequence where as the part which is equal to or less than mean value as irregular component. \( X(k) \) is given by:

\[
X'(k) = \begin{cases} 
X(k), & |X(k)| < x \\
0, & |X(k)| \geq x 
\end{cases}
\]

Third step: Construct Inverse Fast Fourier Transforms (IFFT) \( x(t) = \text{IFFT}(X'(k)) \) for amplitude spectrum of regular component and get \( x(t) \). Consider \( x(t) \) as regular time sequence.

\[
F_{x(t)} \sim \frac{\text{Area } C_0 \text{ Com}^{-1} \text{ fft} - y}{\text{Area } x(t) - x'(t)}
\]

C0 has a value much smaller than 1, whereas for a stochastic system, the value of C0 is close to 1.

**D Data**

1. Three kinds of typical signal sequence, include:
   a) Discrete periodic sequence: \( X(t) = \sin(2\pi f t) \)
   b) Chaotic sequence generated by Henon map (Abandon the first 2000 points):
   \[
   x(t + 1) = 1.4 - x^2(t) + 0.3y(t + 1)
   \]
   \[
   y(t + 1) = x(t)
   \]
   Initial conditions: \( X(1) = 0.5, Y(1) = 0.5 \)
   c) The mean of randomly generated 0, 1 for the Gaussian noise variance.

2. Voice data used in this study have been obtained from clinical cases in The people's Hospital of Yizhou city they are 20-40 years old, including 6 females and 6 males participated in the experiments. We combine the subjective diagnosis patient self-reported and clinical doctor Subjectively diagnosis. According to the total hoarse GRBAS system is proposed in the Japanese Language Association Degree (G) of perceptual evaluation, including grade 0: normal; 1 mild hoarseness; 2 level moderate; 3 level severe hoarseness; every grade has 3 cases in 12 patients. In a quiet room, they read Chinese 'a' tone, duration is 3 seconds, the sampling frequency is 16KHz, Experiment 24576 data points in the data after 2000 relatively stable.

All data Sequences length 24576 point, the length of the rectangular window 1024 slides, divided into 24 frames, then compute HNR, C0 parameter to every frame.

**III. RESULT**

**A Figures**

Fig.1 and Fig.2 show dynamic process of LZC complexity to three kinds of typical signal sequence:
periodic sequence, Chaotic sequence and random sequence listed in the second part.

![Fig.1. three kinds of typical signal sequence: periodic sequence, Chaotic sequence and random sequence](image1)

![Fig.2. The C0 complexity of periodic sequence, Chaotic sequence and random sequence](image2)

![Fig.3. The HNR comparison to pathological / normal voices hoarseness](image3)

Fig.2, Fig.3, Fig.4 and Fig.5 show dynamic process of voice signal in different hoarse degree which is computed for HNR, LZC Complexity and C0 Complexity, The three groups were similar so, we only show one group between them.

**B Table**

According to the calculation method, every different hoarse degree has three samples, each sample is divided into 24 frames by rectangular window 1024 slides, total frames of every hoarse degree is 72. Compute the mean of 72 frames of every hoarse degree HNR, LZC and C0 parameter, the results list in table 1:

<table>
<thead>
<tr>
<th>Voice signal</th>
<th>HNR</th>
<th>LZC</th>
<th>C0</th>
</tr>
</thead>
<tbody>
<tr>
<td>The normal voice</td>
<td>0.4792</td>
<td>0.2325</td>
<td>0.0208</td>
</tr>
<tr>
<td>Mild hoarseness</td>
<td>0.4237</td>
<td>0.2396</td>
<td>0.0237</td>
</tr>
<tr>
<td>Moderate hoarseness</td>
<td>0.3778</td>
<td>0.2509</td>
<td>0.0423</td>
</tr>
<tr>
<td>Severe hoarseness</td>
<td>0.3689</td>
<td>0.4403</td>
<td>0.0585</td>
</tr>
</tbody>
</table>

![Fig.4. The LZC complexity comparison to pathological / normal voices hoarseness](image4)

![Fig.5. The C0 complexity comparison to pathological / normal voices hoarseness](image5)
IV. DISCUSSION

Using the calculation procedure outlined in the second part, it is well known that a process that is least complex has a LZC and C0 value close to zero, whereas a process with highest complexity will have LZC and C0 value close to one. Lempel–Ziv complexity can also be viewed as a measure of randomness, in the figure (a)(b), we can clearly see that periodic sequence sin(t) is least complex, its LZC and C0 is nearest tends to 0. The C0 is closer to zero than LZC, which is generated by the calculation error caused by coarse graining. The LZC complexity of Chaotic sequence is between 0.6 and 0.7. The LZC complexity of random sequence is close to one. The C0 complexity of Chaotic sequence and random sequence are also between 0.2 and 0.3. All of these verify that we write the calculation method of the 2 complexity. To provide a reference for the complexity of the next step analysis of pathological voice.

Voice disease mainly cause hoarse voice, the hoarse degree performance in the hearing as Increasing with disease worse. Hoarse is the sum of the voice variation, Both vocal irregular vibration component hoarse sound, but also the existence of friction noise components. often use HNR parameter to evaluate signals’ regularity[7], in table 1. Mean HNR is proportional to the patient readme and clinical doctor's subjective auditory, Visible, the subjective classification and according to the objective evaluation parameters is consistent. In comparison with HNR, LZC complexity and C0 complexity. HNR can Distinguish pathological voice and normal voice, normal voice’ HNR is bigger than pathological voice, but at the same time LZC complexity and C0 complexity are better to distinguish different hoarse degree, these cohere with the dynamic change process in figure (e) and (f). So we will discuss further into pathological voice’s LZC complexity and C0 complexity.

In figure (e)(f), we can see that the complexity of normal voice is relatively small, Complexity increases with the hoarse degree. LZC complexity and C0 Complexity can distinguish pathological voice and normal voice partly, but there is little difference between normal voice and mild hoarseness. The main reason for this is that in the time domain, mild hoarseness is minimal Irregular, very close to the normal voice. With the hoarse degree increases, the irregular component of voice signal to increase the degree of randomness, are also on the increase, so LZC complexity value and C0 complexity value are increase with the hoarse degree significantly. C0 complexity are distinguish better in every different hoarse degree than LZC complexity, because when compute LZC complexity, we should encode the time

The comparison of figure (a),(b) (d) (e), the voice signal’s complexity is bigger than periodic signal, but lower than chaotic signal and random signal; The voice signal is more complex than the periodic signal, with a certain degree of chaos and randomness. This is consistent with the subjective understanding, when the aggravation of the vocal organs lesions, the complexity of sounding system rise, hoarseness degree will also increase.

V CONCLUSIONS

Overall, HNR Parameters can do better work in distinguish pathological voice and normal voice, but C0 complexity is more suitable for the objective evaluation of pathological voice hoarseness classification. Can be used as parameters for the objective evaluation of voice. So that, we find a new vision for the objective evaluation of pathological voice use Complexity Method.

ACKNOWLEDGEMENTS

We thank the doctors who help us got the data from the clinical cases, at the same time, we are very touched that the patients who participate in recording pathological voice. This study was supported by the National Natural Science Foundation of China (61362003).

REFERENCES