Multifractal Detrended Fluctuation Analysis of Congestive Heart Failure Disease Based on Constructed Heartbeat Sequence

SHAN LI (Student Member, IEEE)
Key Laboratory of Modern Acoustics, School of Electronic Science and Engineering, Ministry of Education, Nanjing University, Nanjing 210023, China
e-mail: lishan@smail.nju.edu.cn
This work was supported by Doctoral Interdisciplinary Research and Innovation Project of Nanjing University under Grant 201704.

ABSTRACT Heart rate variability (HRV) can be used as a common detection method for congestive heart failure (CHF). Existing researches regarding HRV, including both linear indicators and nonlinear characteristics, are mostly based on the RR intervals of the ECG signal. This article proposed a sequence that can reflect the regulation of sympathetic and parasympathetic nerve on heart rate, and on this basis, conducted multifractal detrended fluctuation analysis (MFDFA). We extracted multifractal features to quantitatively compare the complexity of proposed sequence between the healthy and CHF groups. Results showed that abnormal physiological and pathological conditions due to the weakening of autonomic nerve control can reduce the complexity of the heartbeat signal. Estimate the separation performance of all features, the best discrimination is obtained for the area under the mass index spectrum $S_1$ as providing 100% accuracy in separating the Healthy Young and CHF groups, and 90.93% separation accuracy between the Healthy Elderly and CHF groups. This work provide a good basis for the diagnosis of CHF with a novel perspective.

INDEX TERMS Congestive heart failure, heartbeat time series, multifractal detrended fluctuation analysis.

I. INTRODUCTION
Congestive heart failure (CHF) refers to a pathological state due to the decrease in systolic and diastolic functions of the heart, which in turn cannot meet the needs of systemic tissue metabolism [1]. Despite the great progress in diagnosis and treatment, morbidity and mortality are still high, and the prevalence is rising [2]. Around 26 million people worldwide are affected by CHF [3], of which, the prevalence is 1.5%-2.0% in developed countries, and more than 10% in patients >70 years [4]. CHF is a progressive disease, early diagnosis and intervention are extremely important to delay the progress and improve the quality of patients’ life.

Heart rate variability (HRV) can be used as a diagnostic basis for abnormal rhythm or sinus arrhythmias in patients with CHF [5], [6]. HRV conventionally refers to the fluctuation of time intervals between adjacent heartbeats, which is generated by the dual modulation of sinus node by the sympathetic and parasympathetic nervous system (SNS & PNS) [7]–[9]. The increase and decrease of heart rate are respectively related to the excitement of the body’s sympathetic and parasympathetic nerves. The increase in sympathetic tension is often associated with the occurrence of severe arrhythmia, at which time HRV will reduce to a certain extent. Nonetheless, the increase in parasympathetic tone can improve the threshold of arrhythmia, thereby reducing the occurrence of arrhythmia. At this time, HRV appears as a certain increase. Generally, HRV can monitor the tension balance of the sympathetic-parasympathetic nervous system and is a prognostic indicator for evaluating autonomic nervous function and related diseases.

Extensive researches have been carried out regarding HRV, proposing both linear indicators and nonlinear characteristics [10], [11]. Traditional analysis, linear or time-frequency methods, are based on the Fourier transform (FT), which is effective for the analysis of periodic and stationary signals, but can not well reflect the local time-varying characteristics of the signal. Although Wavelet Transform (WT) can focus on any details of the signal, it is not sensitive enough to catch the exact jump time of the frequency components with small amplitude. In distinguishing deterministic “random” signals like heartbeat, nonlinear dynamic parameters are more effective than time-frequency analysis methods. Multifractal analysis proved the chaotic characteristics of biomedical signals.
like EEG, ECG, etc. [12], [13], and has discovered the loss of physiological complexity with CHF [14]. Multifractal is actually a generalization of the fractal system when a single scale exponent is not enough to describe the dynamics of the object, and should be expressed as a continuous exponential spectrum (singular spectrum $f(\alpha) \sim \alpha$). Improved formalisms have been developed around multifractal analysis, such as the wavelet transform modulus maxima (WTMM) method and the multifractal detrended fluctuation analysis (MFDFA) method that does not require the modulus maxima procedure. The comparison between these two methods was concerned by Kantelhardt et al. [15], suggesting that the results of MFDFA are slightly more reliable than those of WTMM. Therefore, hereafter we will extract features on the basis of MFDFA to characterize the complexity of heartbeat sequence.

Previous studies on HRV have been based on the RR intervals of ECG signal. Next, we will propose a sequence that helps reflect the alternation of the increase and decrease of heart rate due to the sympathetic and parasympathetic nerve regulation, which will provide powerful diagnostic basis for predicting the physiological and pathological changes.

In this article, we adopted the MFDFA method to analyze the heartbeat signal of CHF patients, and the specific process is as follows: Based on the regulation of heart rate by the autonomic nerve, we first constructed a sequence, that is, the monotonous increase to monotonous decrease amplitude ratios as derived from RR intervals. Then the complexity of the sequence is quantified by multifractal features, including the width of multifractal spectrum, the area under the mass index spectrum and the multifractal spectrum area. Finally, we compared the separation performance of different features.

II. METHODS AND RESULT

A. DATA

There is no obvious limitation in the age of onset of CHF, so both healthy young people and healthy elderly people were selected as the control group. The ECG recordings used herein were from the PhysioNet [16], which is a repository of freely-available medical research data and managed by the MIT Laboratory for Computational Physiology. Recordings of 14 CHF patients (NYHA class 3-4) were from the BIDMC Congestive Heart Failure Database [17], which is 20 hours in duration, and sampled at 250 Hz. The original analog electrical signals were made with ambulatory ECG recorders (bandwidth 0.1-40Hz). Recordings of 19 healthy young (aged 21-34) and 15 healthy elderly people (aged 68-85) were from the Fantasia database [18]. They underwent rigorous screening and had ECG taken during 2 hours of continuous supine rest. The signals were digitized at 250 Hz. To ensure accuracy, annotation files of both patients and healthy subjects were detected by an automated detector and additional manual checks.

B. PREPROCESSING

Based on the regulation of heart rate by the autonomic nerve, an amplitude-ratio sequence AR was constructed (Fig. 1). Thereinto, Fig. 1(a) is a segment of ECG signal where RR$(n)$ is the time interval between peaks $R_n$ and $R_{n+1}$; Fig. 1(b) is the extracted RR$(n)$ sequence; Fig. 1(c) shows the construction process of AR (Eq. 1), where $m^+_i$ represents the trough position of RR$(n)$, and as the sequence grows, $m^-_i$ and $m^+_{i+1}$ are the nearest peak and trough, respectively. $IA(i)$ represents the amount of monotonous increase in the heartbeat intervals, indicating the activation of parasympathetic nerve. Correspondingly, $DA(i)$ represents the increased sympathetic regulation. Therefore, AR can reflect the balance of cardiac autonomic nerve.

\[
AR(i) = \frac{IA(i)}{DA(i)} = \frac{RR(m^-_i) - RR(m^+_i)}{RR(m^+_i) - RR(m^+_{i+1})},
\]

Fig. 2 shows the probability density function of (a) the RR interval exhibiting a normal distribution, and (b) the AR sequence presenting a power-law distribution. All subjects in the database show similar results.

Once a certain power-law distribution is found, we can deduce whether this is due to the existence of fractals.

C. METHODS

Based on the MFDFA proposed by Kantelhardt et al. [19], we extracted the multifractal features of AR sequence in i) healthy young people, ii) healthy elderly people and iii) CHF...
patients, and compared the differences between them. The following are the important steps involved in this method:

1) Calculate the cumulative sum of the de-averaged time series

\[ Y(i) = \sum_{k=1}^{i} [x(k) - x_{ave}], \quad i = 1, 2, 3, \ldots, N, \]  

where \( x_{ave} \) is the average of non-stationary time series \( x(i) \).

2) Use the box of size \( s \) to divide \( Y(i) \) in order, and discard the parts that cannot be divided, then \( N_s \) non-overlapping boxes can be obtained. In order to make full use of the series, repeat the above segmentation for inverted \( Y(i) \). Ultimately, \( 2N_s \) boxes will be obtained altogether.

3) Fit the series in each box by least squares and calculate the corresponding variance. For \( v, v = 1, 2, 3, \ldots, N_s \),

\[ F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} [Y((v-1)s+i) - y_v(i)]^2, \]  

and for \( v = N_s + 1, N_s + 2, N_s + 3, \ldots, 2N_s \), there is

\[ F^2(s, v) = \frac{1}{s} \sum_{i=1}^{s} [Y[N - (v - N_s)s+i] - y_v(i)]^2, \]  

where \( y_v(i) \) is the fitting polynomial of the series in the \( v \)-th box.

4) Calculate the \( q \)-th order fluctuation function

\[ F_q(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(s, v)]^q \right\}^{\frac{1}{q}}. \]  

Here, order \( q \) can be any real value except 0. For \( q = 0 \), \( F_0(s) \) can be calculated by a logarithmic averaging procedure

\[ F_0(s) = \exp\left[\frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln[F^2(s, v)]\right] \sim s^{h(0)}. \]  

5) Vary box size \( s \) and repeat the step 2)~4) above. Obviously, \( F_q(s) \) will increase as \( s \) goes up. For long-range correlated series, \( F_q(s) \) will be a power function of \( s \)

\[ F_q(s) \propto s^{h(q)}. \]  

In general, the value of exponent \( h(q) \) may depend on \( q \). For stationary time series, \( h(2) \) is identical to the well-known Hurst-exponent \( H \). Thus, \( h(q) \) can be called the generalized dimension \( D_q \). For \( q > 0 \), the box \( v \) with large variance \( F^2(s, v) \) may dominate the fluctuation function \( F_q(s) \). In this case, \( h(q > 0) \) basically describes the scaling behavior of boxes with large fluctuations. Accordingly, \( h(q < 0) \) represents the scaling behavior of boxes with small fluctuations. While for monofractal time series, \( h(q) \) is constant for all values of \( q \). Therefore, \( 0.5 < H < 1 \) indicates a (positively) long-range correlated series; \( H = 0.5 \) corresponds to a random walk series; and \( 0 < H < 0.5 \) demonstrates that the series is long-range anti-correlated, that is, the future increment is negatively correlated with the past increment, and the series has abrupt jump reversal.

6) Mass index \( \tau(q) \) can be calculated by the following relationship with \( h(q) \)

\[ \tau(q) = qh(q) - 1. \]  

Obviously, for a monofractal time series with unique \( h(q) \), \( \tau(q) \) will grow linearly with respect to \( q \), and for a multifractal time series, \( \tau(q) \) will depend nonlinearly on \( q \). The greater the curvature of curve \( \tau(q) \sim q \), the more complicated the series is.

Singularity strength \( \alpha \) and singularity spectrum \( f(\alpha) \) can be obtained by Legendre transformation

\[ \alpha = h(q) + qh'(q), \]  

\[ f(\alpha) = q[\alpha - h(q)] + 1, \]  

where \( \alpha \) represents the fractal dimension of subset series in a box, called the local fractal dimension. Due to the large number of boxes, a sequence of different dimensions can be obtained. And \( f(\alpha) \) is used to describe the frequency of the dimension. The width of spectrum \( \Delta \alpha = \alpha_{\text{max}} - \alpha_{\text{min}} \) can be used to denote the range of fractal dimensions and can provide the complexity information of the fractal body. A larger \( \Delta \alpha \) represents a greater complexity.

D. RESULTS

The multifractal features of \( AR \) sequence in i) healthy young people, ii) healthy elderly people and iii) CHF patients were extracted based on MFDFA method. Thereinto, the sequence length of \( AR \) was uniformly intercepted as \( N = 1249 \) in accordance with the minimum of all subjects. Research on the MFDFA of nonstationary time series [19] shows that,
systematic deviations from the scaling behavior in Eq. 7 will occur for \( s < 10 \). Additionally, fluctuation function \( F_q(s) \) will become statistically unreliable for large sizes \( s > N/4 \). Thus, in this article, it is an adequate range for \( s \) varying from 10 to 200. Moreover, we chose \( q \in [-5, 5] \) to ensure the effectiveness and stability.

In order to select the adequate order of the polynomial fitting function, we applied the MFDFA, varying the order from 1 to 3. Fig. 3 shows the results of generalized dimension \( D_q \sim q \). The curves are just slightly different from each other. In terms of simplicity of calculation, we adopted detrending polynomial fitting functions of order 1 in the MFDFA.

Fig. 4 is the results of MFDFA for the healthy and CHF groups. Fig. 4a shows the change of generalized dimension \( D_q \), singularity strength \( \alpha \) and singularity spectrum \( f(\alpha) \) with order \( q \). \( \alpha \) decrease monotonically with \( q \), indicating that \( AR \) is multifractal characterized in all groups. Fig. 4b gives the mass index curve \( \tau(q) \sim q \). It can be concluded from Eq. 8 that, the linear dependence of \( \tau(q) \) on \( q \) indicates a monofractal characteristic, and that more complex \( AR \) will correspond to greater curvature. Fig. 4c exhibits the relation between singularity spectrum \( f(\alpha) \) and singularity strength \( \alpha \), where \( \Delta \alpha \) is a typical quantitative indicator of multifractal degree. From the three subplots showed in Fig. 4, conclusion about the degree of complexity that healthy young > healthy elderly > CHF patients can be obtained.

Next, we extracted four features from Fig. 4 that can characterize the degree of multifractality, and performed quantitative comparison between different groups:

1) The width of multifractal spectrum \( \Delta \alpha \) (Fig. 4c),

2) Area under the mass index spectrum \( S_{1\tau} \) (Fig. 5a), which is enclosed by the curve \( \tau(q) \sim q \) (blue solid line) and its end-to-end connection (black dotted line). \( S_{1\tau} \) contains the two-dimensional space information of mass index spectrum.

3) The classical multifractal spectrum area \( S(f) = \int f(\alpha)d\alpha \). Integration of \( f(\alpha) \) and \( \alpha \) contains all dimension information of \( AR \) sequence.

4) Area under the multifractal spectrum \( S_{1f} \) (Fig. 5b) proposed by Wang et al. [20]. \( S_{1f} \) is enclosed by curve \( f(\alpha) \sim \alpha \) (blue solid line) and its end-to-end connection (black dotted line). \( S_{1f} \) effectively avoids the deviation caused by the movement of curve position, and can better reflect the non-linear nature of \( AR \) sequence.

The higher the nonlinear complexity of \( AR \), the greater the features above. Table 1 lists the statistical results of four features, suggesting that the complexity of \( AR \) sequence is more in healthy groups than those with CHF groups, and that, data of young people are more complex than the elderly. Significant difference between the healthy and CHF groups (\( p < 0.01 \)) can be obtained in all features (Fig. 6), among which, \( S_{1\tau} \) (Fig. 6c) can effectively distinguish between healthy young people and healthy elderly people (\( p < 0.05 \)).
not simply change with other factors, such as external stimuli, intensity of physical activity, and changes in body posture. In fact, aging and disease can lead to a decrease in the body’s autonomic control, which in turn affect the degree of nonlinearity in the electrocardiographic system. However, there is no definite mathematical equation to describe such a complex nonlinear dynamic system, so it is very necessary to simulate its complexity through experiments.

In this article, we first proposed a sequence that is derived from ECG signals and can reflect the regulation of autonomic nerve to heart rate. Next, for subjects with different physiological and pathological conditions, we used MFDFA to analyze complexity for the proposed sequence. The following conclusions can be obtained:

1) AR sequence, which is derived from RR intervals, can reflect the alternating changes of heart rate due to the regulation of sympathetic and parasympathetic nerves. Different from a typical normal probability distribution showed by RR sequence, AR presents a power-law distribution with long tail. The emergence of power-law distribution allows us to speculate whether this is due to the existence of fractals.

2) The multifractal analysis of AR sequence confirms the conclusion that aging and disease are often accompanied by a decrease in HRV. Additionally, quantitative results of the complexity of cardiac dynamics for each group are shown as, healthy young > healthy elderly > CHF patients. The decrease in the nonlinear complexity of the electrocardiographic system may be caused by changes in the self-similar structure of the heart tissue.

3) Features extracted from MFDFA, \( \Delta \alpha, S_1, S_\tau \) and \( S_1 \) are all sensitive to the difference between the healthy and CHF groups. Thereinto, the best discrimination is obtained for \( S_1 \) as providing 100% accuracy in separating the Young and CHF groups, and 90.93% separation accuracy between the Elderly and CHF groups.

### Table 2. Separation Performance of \( \Delta \alpha, S_1, S_\tau \) and \( S_1 \) Between the Healthy (Young, Y; Elderly, E) and Patient (CHF, C) Groups.

| Feature  | \( \Delta \alpha \) | \( S_1 \) | \( S_\tau \) | \( S_1 \) |
|----------|---------------------|-----------|-------------|---------|
| Y        | 0.66                | 0.84      | 1.57        | 0.60    |
| C        | 1.20                | 1.01      | 2.15        | 0.10    |
| E        | 1.12                | 0.61      | 0.78        | 0.39    |
| Y        | 0.77                | 0.43      | 1.16        | 0.59    |
| CHF      | 0.68                | 0.54      | 0.81        | 0.21    |
| E        | 0.76                | 0.53      | 0.79        | 0.23    |

### III. DISCUSSION

The multifractal phenomenon in heartbeat dynamics is an inherent characteristic of its control mechanism, which does for data with small samples, we adopt the following measures to estimate the separation performance between healthy and CHF groups.

1) The ratio \( \eta \)

\[
\eta = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2},
\]

where \( \mu_1, \mu_2 \) and \( \sigma_1, \sigma_2 \) are respectively the means and standard deviations of the features of the two groups being compared. For a good discrimination, difference between the means is much greater than the standard deviation, which will result in \( \eta \gg 1 \).

2) Statistical distance \( d^2 \), which represents the “distance” from the means of the “boundary” that minimizes incorrect classification for both groups.

\[
d^2 = \left( \frac{\mu_1 - \mu_2 + \sigma_1^2 \times \mu_1}{\sigma_1^2 + \sigma_2^2} \right)^2.
\]

Similarly, the larger the \( d^2 \) the better the discrimination provided by the method.

3) The total separation rate \( TSR \)

\[
TSR = 1 - \frac{1}{2} \times [L/(2 \times \sigma_1) + L/(2 \times \sigma_2)],
\]

which ranges in \([0,1]\). Thereinto, \( L \) denotes the length of the overlap of standard deviations between the healthy and CHF groups. \( TSR = 1 \) means a clear separation and that the feature can separate subjects into two nonoverlapping groups.

Table 2 shows the performance of \( \Delta \alpha, S_1, S_\tau \) and \( S_1 \) in the discrimination between healthy and CHF groups. It can be obtained that \( S_1 \) performs better, and provides 100% accuracy in separating the Young and CHF groups, and 90.93% accuracy in separating the Elderly and CHF groups.
sequence in CHF patients [23], [24], and the results showed only differences in the width of the multifractal spectrum $\Delta \alpha$.

Therefore, we believe that this article has great advantages and contributions in the research of CHF detection based on MFDFA.

In this article, we provide a good basis for the diagnosis of CHF in different ages. Besides, the presentation of AR sequence is noteworthy as it relates the regulation of cardiac autonomic nerve to the multifractal complexity of body’s heartbeat signal, which provides a novel perspective for the detection of CHF disease.

REFERENCES

[1] J. Wang, Z. Li, J. Chen, H. Zhao, L. Luo, C. Chen, X. Xu, W. Zhang, K. Gao, B. Li, and J. Zhang, “Metabolomic identification of diagnostic plasma biomarkers in humans with chronic heart failure,” Mol. BioSyst., vol. 9, no. 11, pp. 2618–2626, 2013.

[2] A. L. Bui, T. B. Horwich, and G. C. Fonarow, “Epidemiology and risk profile of heart failure,” Nature Rev. Cardiology, vol. 8, no. 1, pp. 30–41, Jan. 2011.

[3] P. Poniksowski, S. D. Anker, K. F. AlHabib, M. R. Cowie, T. L. Force, S. Hu, T. Jaarsma, H. Krum, V. Rastogi, L. E. Rohde, C. Samal, H. Shimokawa, B. B. Siswanto, K. Sliva, and G. Filippatos, “Heart failure: Preventing disease and death worldwide,” ESC Heart Failure, vol. 1, no. 1, pp. 1–45, Sep. 2014.

[4] A. Mosterd and A. W. Hoes, “Clinical epidemiology of heart failure,” Heart, vol. 93, no. 9, pp. 1137–1146, Sep. 2007.

[5] J. Olligs, D. Lintz, D. G. Dechering, L. Eckardt, and P. Müller, “Heart rate—A complex prognostic marker in acute heart failure,” JIC Heart Vascularure, vol. 26, Feb. 2020, Art. no. 100456.

[6] Y. Isler, A. Narin, M. Ozer, and M. Perc, “Multi-stage classification of congestive heart failure based on short-term heart rate variability,” Solitons Fractals, vol. 118, pp. 145–151, Jan. 2019.

[7] Z. Z. Taralov, K. V. Terziyski, and S. S. Kostianew, “Heart rate variability as a method for assessment of the autonomic nervous system and the adaptations to different physiological and pathological conditions,” Folia Medica, vol. 57, nos. 3–4, pp. 173–180, Apr. 2016.

[8] N. Herring, M. Kalla, and D. J. Paterson, “The autonomic nervous system and cardiac arrhythmias: Current concepts and emerging therapies,” Nature Rev. Cardiol., vol. 16, no. 12, pp. 707–726, Dec. 2019.

[9] K. E. Speer, S. Semple, N. Naumovski, and A. J. McKune, “Heart rate variability for determining autonomic nervous system effects of lifestyle behaviors in early life: A systematic review,” Physiol. Behav., vol. 217, Apr. 2020, Art. no. 112806.

[10] A. Bagirathan, J. Selvaraj, A. Gurusamy, and H. Das, “Recognition of positive and negative valence states in children with autism spectrum disorder (ASD) using discrete wavelet transform (DWT) analysis of electrocardiogram signals (ECG),” J. Ambient Intell. Hum. Comput., May 2020.

[11] C. Besson, M. Saubade, V. Cremaux, G. P. Millet, and L. Schmitt, “Heart rate variability: Methods, limitations and clinical examples,” Revue Medicale Suisse, vol. 16, no. 701, pp. 1432–1437, 2020.

[12] A. L. Bui, T. B. Horwich, and G. C. Fonarow, “Epidemiology and risk profile of heart failure,” Nature Rev. Cardiology, vol. 8, no. 1, pp. 30–41, Jan. 2011.

[13] D. Jiao, Z. Wang, J. Li, F. Feng, and F. Hou, “The chaotic characteristics detection based on multifractal detrended fluctuation analysis of the elderly 12-lead ECG signals,” Phys. A, Stat. Mech. Appl., vol. 540, Feb. 2020, Art. no. 123234.