35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Fuzzy central tendency measure for time series variability analysis with application to fatigue electromyography signals

Hong-Bo Xie., and Socrates Dokos, *Member, IEEE*

***Abstract*—A new method, namely fuzzy central tendency measure (*f*CTM) analysis, that could enable measurement of the variability of a time series, is presented in this study. Tests on simulated data sets show that *f*CTM is superior to the conventional central tendency measure (CTM) in several respects, including improved relative consistency and robustness to noise. The proposed *f*CTM method was applied to electromyograph (EMG) signals recorded during sustained isometric contraction for tracking local muscle fatigue. The results showed that the *f*CTM increased significantly during the development of muscle fatigue, and it was more sensitive to the fatigue phenomenon than mean frequency (MNF), the most commonly-used muscle fatigue indicator.**

I. Introduction

Variability analysis can be defined as the comprehensive assessment of the degree and characteristics of patterns of variation in a time series over time [1]. The central tendency measure (CTM) is an important variability analysis measure derived from the Poincaré plot. It is defined as the percentage of data points which fall within a certain radius from the centre of the first difference of Poincaré plot of the original time series. CTM has found applications in many biomedical engineering areas since it was developed. The CTM of blood oxygen saturation (SaO2) signals has been analyzed with Limp-Ziv (L-Z) complexity and approximate entropy methods as a diagnostic test for obstructive sleep [2, 3]. Both the sensitivity and specificity of the CTM method were found to be higher than L-Z complexity and approximate entropy, which are respectively based on symbol dynamics and the regularity of the attractor trajectory. Ramdani et al. examined human postural sway velocity time series using CTM to quantify the smoothness of the underlying dynamics [4]. The results suggested that the CTM of the velocity time series significantly reduced with aging, related to increased postural muscle response and reflex times, or reduced proprioception. Thuraisingham et al. used CTM to depict the degree of variability in electroencephalograph (EEG) signals during eye open and eye closed states, for the purpose of developing a "hands free" brain computer interface [5]. They demonstrated increased EEG variability during eye closed compared to eye open states. An improved component CTM (CCTM) method was also proposed by Thuraisingham to analyze cardiac RR intervals [6]. It counts the number of points present in the four quadrants of the difference plot separately. He employed the technique to differentiate congestive heart failure patients from normal patients, with a classification success rate of 100%. The CTM method was also applied to discriminate between normal heart rate variability and random rhythm generation (RRG) in the ECG of schizophrenic patients [7, 8].

Given a time series with N data points, the calculation of CTM requires *a priori* determination of an unknown parameter, i.e., the radius r of the circular region in the first difference Poincaré plot. Unfortunately, there is no established rule for the selection of the optimal r value. Moreover, incorrect choice of r can result in missing information due to excluded data. In this paper, we present a fuzzy central tendency measure (*f*CTM) for quantifying time series variability. Utilizing the concept of fuzzy sets, similarity between first difference plots is fuzzily defined on the basis of a fuzzy membership function. The proposed *f*CTM was then applied to EMG analysis for monitoring the local muscle fatigue.

II. Methods

A. Fuzzy Central Tendency Measure

Let the scalar time series $x_1, x_2, \cdots, x_N$ be generated by a dynamical system, where n is the number of samples. We can draw scatter plots of first differences of the data which graph $x(i+2)-x(i+1)$ against $x(i+1)-x(i)$, where each $x(i)$ represets the time series' value at time i. These plots, centered around the origin, provide a graphical representation of the degree of variability in the time series. To quantify this variability, the CTM can be obtained from the first differences of the data by choosing a circular region of radius r around the origin, counting the number of those points which fall within the radius, and dividing by the value of total points [9]. A low CTM value demonstrates a large amount of dispersion but a large value concentration near the centre. Assume a time series is with N data points, the CTM can be computed as,

$$CTM(r) = \frac{\sum_{i=1}^{N-2} \delta(d_i)}{N-2}, \tag{1}$$

$$d_i = \left([x(i+2)-x(i+1)]^2 + [x(i+1)-x(i)]^2 \right)^{1/2} \tag{2}$$

$$\delta(d_i) = \begin{cases} 1 & if\ d_i < r \\ 0 & otherwise \end{cases} \tag{3}$$

A Heaviside function $\delta(d_i)$ is used here to count the points in the region. The rigid boundary of the Heaviside function can result in some of information loss. For example, when the data points are just outside the boundary, i.e. the Euclidian distance between $x(i+2)-x(i+1)$ and $x(i+1)-x(i)$ is just larger than the radius r, these points are ignored, while the

Hong-Bo Xie is with the Graduate School of Biomedical Engineering, University of New South Wales, Sydney, 2052, Australia (Tel: 61-4-52582099; e-mail: xiehb@ sjtu.org).

Socrates Dokos is with the Graduate School of Biomedical Engineering, University of New South Wales, Sydney, 2052, Australia (e-mail: s.dokos@unsw.edu.au).

contributions of all data points inside the boundary are handled equally. These problems make the calculation of CTM brittle and the choice of radius r difficult. The CTM value varies dramatically with a slight change in the radius value. However, the problem is able to be avoided if the Heaviside function in Eq. (3) is replaced by a fuzzy membership function. $u(d_i, r)$ to obtain a fuzzy measurement of the "distance" or similarity between $x(i+2)-x(i+1)$ and $x(i+1)-x(i)$ based on their shapes. The fuzzy central tendency measure can thus defined as follow:

$$fCTM(r) = \frac{\sum_{i=1}^{N-2} u(d_i, r)}{N-2}. \quad (4)$$

In practice, a bell-shape function, sigmoid function, Gaussian function, or any other fuzzy membership function that is convex and continuous could be selected to substitute for the Heaviside function. In this study, we used the following Gaussian function for the *f*CTM calculation,

$$u(d_i, r) = \exp(-d_i^2 / r), \quad (5)$$

B. Simulated data

The ability of *f*CTM to discriminate among different degrees of variability was evaluated on several benchmark data sets. These time series included independent, identically distributed (i.i.d.) Gaussian noise, a chirp signal, MIX processes [10], and Logistic map. The simulation of N points MIX(P) process, where P is between 0 and 1, is a sine wave, where $N \times P$ randomly selected points have been substituted for random noise [10]. The Logistic map is defined by [11]

$$x_{i+1} = Rx_i(1-x_i). \quad (6)$$

where R is a control parameter. Data were generated for R=3.5, 3.6, and 3.9. The characteristics of time series for R=3.5 is periodic (period four) dynamics, R=3.6 and 3.9 produce chaotic dynamics with increasing variability.

C. EMG data

The EMG has been often applied to monitor muscular changes [12]. EMG signals used in this paper were detected from human biceps muscle during static voluntary isometric contractions in twelve healthy subjects (mean age $\pm$ std: 30.2 ± 4.9 years). No subjects had any neuro-musculoskeletal disorders and all signed informed consent prior to the experimental testing. A pair of surface EMG electrodes (Axon Systems, Inc., New York, USA) with 25 mm centres distance was put longitudinally on abrased, clean skin, immediately under the thickest point of the biceps. The reference electrode was put on the proximal head of the ulna. In the testing, the subject first executed an elbow flexion against the lever arm to 80% of his/her maximal voluntary contraction (MVC) and maintained this value through visual feedback of the torque showing on the screen. The experiment was stopped when the torque was decreased to approximately 70% of the MVC, indicating the muscle transfers form normal to fatigue state. The amplified rate was 1000 with a 10–400 Hz band-pass filter. Signals from the EMG electrodes were sampled at 1 KHz and digitally stored for further analysis.

III. Results

A. Relative consistency and monotonicity

An important feature of variability measures is the relative consistency, which means if one attractor is smoother than another, then it should process a larger *f*CTM value for all evaluated conditions. Graphically, curves of *f*CTM versus r for various series should not cross over each another. This expectancy was first assessed by using the realizations of the MIX(P) process, where the degree of variability could be specified. The scheme was to compare MIX(0.4), MIX(0.6), and MIX(0.8) series. The anticipant result was that *f*CTM(MIX(0.6)) should be less than *f*CTM(MIX(0.4)), whilst *f*CTM(MIX(0.8)) should be less than *f*CTM (MIX(0.6)). Fig. 1 shows the simulted results with 100-point realizations of the MIX(P) process for *f*CTM and conventional CTM. Apparently, for each tolerance value r, the *f*CTM of MIX(0.6) was notably and strictly lower than that of MIX(0.4), whilst MIX(0.8) was lower than MIX(0.6). These results illustrated the relative consistency of *f*CTM in the simulation. However for CTM, the plots of MIX(0.4) and MIX(0.6) cross over, which demonstrates the lack of relative consistency of CTM in the evaluation.

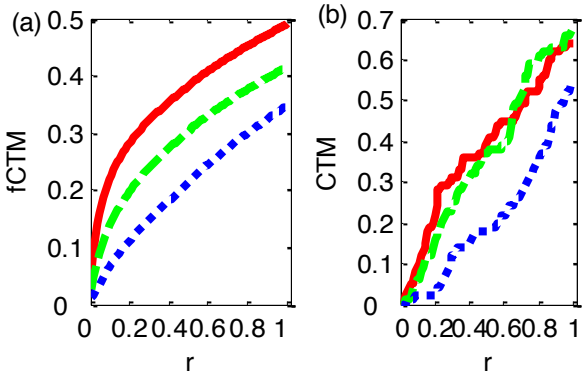

Figure 1. The performances of *f*CTM (a), and CTM (b) statistics that quantify the variability of MIX(0.4) (solid), MIX(0.6) (long dash) and MIX(0.8) (short dash) for N=100.

Gaussian noise and chirp signals are often used to compare the performance of different nonlinear measures on relative consistency [13]. We also assessed the performance of *f*CTM using these signals. Fig. 2 shows the results of *f*CTM and CTM with 200-point realizations of each. The poor results of CTM shown in Fig. 2(b) were similar to approximate entropy in Xie el al. [13], i.e., the variability value of CTM for the chirp signal was less than the white noise signal when r is greater than about 1.5. However, the *f*CTM was able to differentiate the variability between the series pair correctly and with relative consistency. We also found that the values of *f*CTM monotonically increased as r increased for the different signals in Figs. 1 and 2. However, the CTM of chirp signal often abruptly increased or oscillated as r increases. The loss of monotonicity in the CTM caused the difficulty of interpreting signal variability, reducing its differentiation capability.

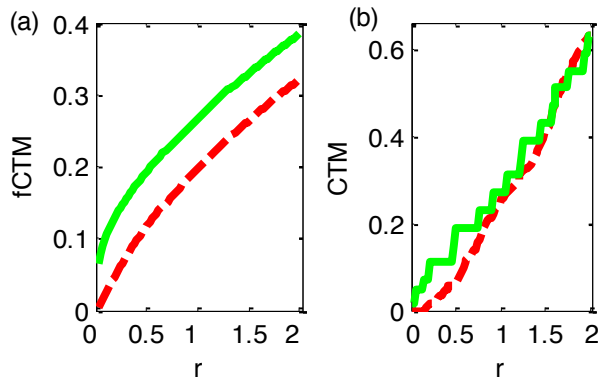

Figure 2. The performances of *f*CTM (a), and CTM (b) statistics that quantify the variability of a chirp signal (solid) and Gaussian noise (long dash) for *N*=200.

B. Robustness to noise

We also assessed the performance of *f*CTM against noise corruption, which is an important feature for real-life applications to noisy signals. This property of both *f*CTM and conventional CTM was evaluated by using Logistic maps with three different *R* values. New time series were obtained by superimposing i.i.d. Gaussian white noise with different noise levels (NL) into clean Logistic series.

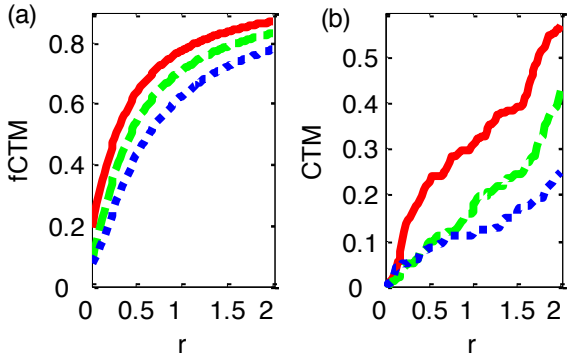

Figure 3. The performances of *f*CTM (a), and CTM (b) in distinguishing Logistic maps with *R*=3.5 (solid), 3.6 (long dash), and 3.9 (short dash) at a noise level of 0.1 and *N*= 500.

For each map, the statistics of both *f*CTM and CTM were obtained for time series with various lengths as *r* ranging from 0.01 to 2. The smoother map produces larger variability values, which are true for both *f*CTM and conventional CTM if *N* and *r* are large enough. However, when *N* and *r* were decreased, the variability measured by CTM became skewed, with results even worse when the series were contaminated by noise. Fig. 3 shows variability values for Logistic maps contaminated with noise (NL = 0.1). The series length was *N* =500 in the calculations. It was found that *f*CTM distinguished Logistic maps with three different *R* values over the whole range of *r* changed from 0.01 to 1 in steps of 0.01. However, CTM could only distinguish the maps with *R*=3.6 and 3.9 for *r* ranging from 0.1 to about 0.5. Fig. 4 shows the performances of the two variability statistics in distinguishing Logistic maps at various noise levels (*r*=0.2). *f*CTM could accurately differentiate between the Logistic series with various control parameters *R* even the NL increased to 0.5. Unfortunately, it was hard for CTM to differentiate between Logistic maps with *R*=3.6 and 3.9 corrupted by noise for *r* = 0.2 and *N*=500. Thus, *f*CTM exhibited better robustness to noise in differentiating various dynamic behaviors.

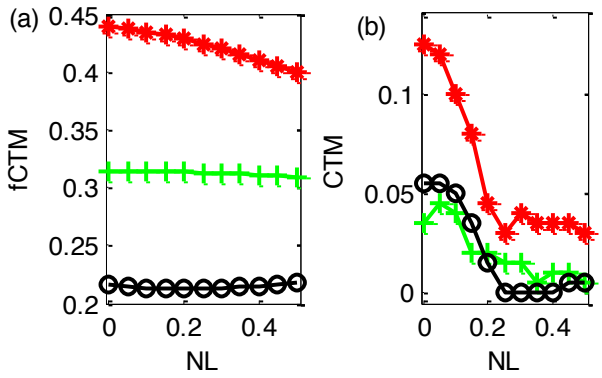

Figure 4. The noise performances of *f*CTM (a), and CTM (b) statistics in distinguishing Logistic maps with *R*=3.5 (star), 3.6 (cross), and 3.9 (circle) at *r*=0.2 and *N*= 500.

C. Results on EMG signal

To indicate the effectiveness of the proposed technique, *f*CTM and CTM of EMG signals were evaluated as indicators of human muscle fatigue during sustained isometric contraction. The EMG is the electrical manifestation of motor unit activities associated with muscle contraction. The reduced variation between slow and fast twitching motor units and the synchronized firing imply that the variability of the EMG signal decreases as the muscle transitions from the normal to the fatigue state. It is thus deduced that values for both *f*CTM and CTM increase with the development of muscle fatigue.

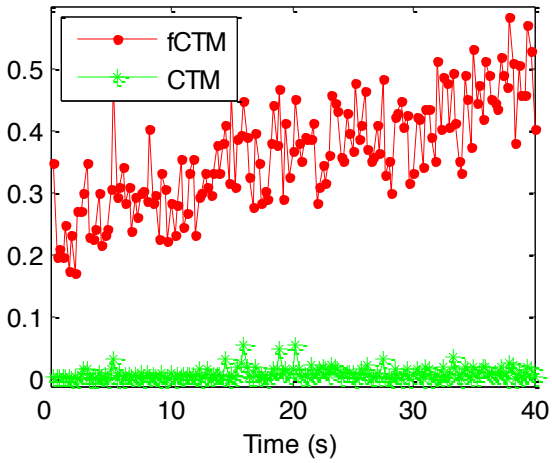

Figure. 5 Time courses of *f*CTM (circle) and CTM (star) for the EMG of subject 3. The window width was 500 ms.

In order to track the changes in *f*CTM and CTM of EMG over time, we segmented the EMG signal into consecutive 500 ms epochs with 50% overlapped rate. For every epoch, the EMG signal was normalized and a value for *f*CTM and CTM was obtained based on Equations (1) and (4), respectively. Fig. 5 shows the time courses of the two measures for subject 3. It can be observed that the *f*CTM increased significantly during the development of muscle fatigue. Unfortunately,

most values of CTM for different EMG epochs were near zero. CTM was not able to detect the small pattern changes in EMG signals recorded from normal to fatigue. The *f*CTM analysis from all the other subjects was similar to the results shown in Fig. 5, demonstrating that the EMG *f*CTM could be used as a new indicator of muscle fatigue.

Table 1. Linear regression slope of *f*CTM and MNF for each subject. All slope values are in units of s^{-1}

subject	*f*CTM	MNF
1	0.0126	-0.0094
2	0.0068	-0.0055
3	0.0066	-0.0056
4	0.0070	-0.0052
5	0.0067	-0.0048
6	0.0065	-0.0037
7	0.0060	-0.0042
8	0.0050	-0.0038
9	0.0059	-0.0042
10	0.0073	-0.0050
11	0.0049	-0.0038
12	0.0086	-0.0066
Mean ± *std*	0.0070 ± 0.0021	–0.0052 ± 0.0015

To date, mean frequency (MNF) has been hailed as the gold standard for muscle fatigue assessment by using EMG under 'static' conditions. The slope of the linear regression has used as a primary quantitative fatigue index [12, 14]. In order to further evaluate the applicability of the *f*CTM statistic for muscle fatigue monitoring, the MNF of our EMG signals was also obtained for comparison, where we observed a time-decrease trend of the MNF. To facilitate comparison, the *f*CTM and MNF were normalized by their respective first epoch values, and a least-square error linear regression was then fitted to each over the period of muscle contraction to obtain the slope. Table. 1 gives the time-regression slopes of *f*CTM and MNF for each subject. One-way ANOVA was performed to test the statistical significance of the results. The slope of *f*CTM was significantly higher than the absolute value of MNF ($p = 0.0213$). These results suggest that *f*CTM is better than MNF in detecting muscle fatigue, since the higher slope indicates that the index is more sensitive to the fatigue effect.

IV. Discussion and Conclusions

We have presented a new variability analysis method, the fuzzy central tendency measure. *f*CTM implements a new rule for determining first order differences similarity in CTM. As opposed to the discontinuous and hard boundary of a Heaviside function in CTM, the continuous and soft boundary of fuzzy membership functions make the *f*CTM statistic change smoothly when the radius *r* is slightly increased. In addition, the *f*CTM also shows better relative consistency and is more robust to noise. When *f*CTM was applied to characterize the variability in EMG signals, it increased significantly and was more sensitive to muscle fatigue than MNF. It can be applied to sports medicine, rehabilitation engineering, and other biomedical-related areas as an alternative measure of local muscle fatigue.

References

[1] A. Bravi, A. Longtin and A. JE. Seely, "Review and classification of variability analysis techniques with clinical applications," *Biomed. Eng. Online*, vol. 10, pp. 1–27, Oct. 201.

[2] D. Alvarez, R. Hornero, D. Abasolo, F. del Campo, and C. Zamarron, "Nonlinear characteristics of blood oxygen saturation from nocturnal oximetry for obstructive sleep apnoea detection," *Physiol. Meas*., vol. 27, no. 4, pp. 399–412, Apr. 2006.

[3] D. Alvarez, R. Hornero, M. Garcia, F. del Campo, and C. Zamarron, "Improving diagnostic ability of blood oxygen saturation from overnight pulse oximetry in obstructive sleep apnea detection by means of central tendency measure," *Artif. Intell. Med.*, vol. 41, no. 1, pp. 13–24, Sep. 2007.

[4] S. Ramdani, B. Seigle, Benoit, D. Varoqui, F. Bouchara, H. Blain, and P. L. Bernard, "Characterizing the dynamics of postural sway in humans using smoothness and regularity measures," *Anns. Biomed. Eng.*, vol. 39, no. 1, pp. 161–171, Jan. 2011.

[5] R. A. Thuraisingham, Y. Tran, P. Boord, and A. Craig, "Analysis of eyes open, eye closed EEG signals using second-order difference plot," *Med. Bio. Eng. Comput.*, vol. 45, no. 12, pp. 1243–1249, Dec. 2007.

[6] R. A. Thuraisingham, "A classification system to detect congestive heart failure using second-order difference plot of RR intervals," *Cardiol. Res. Pract.*, vol. 2009, pp. 807379, 2009.

[7] R. Hornero, D. Abasolo, N. Jimeno, C. I. Sanchez, J. Poza, and M. Aboy, "Variability, regularity, and complexity of time series generated by schizophrenic patients and control subjects," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 2, pp. 210–218, Feb. 2006.

[8] A. Jovic and N. Bogunovic, "Electrocardiogram analysis using a combination of statistical, geometric, and nonlinear heart rate variability features," *Artif. Intell. Med.*, vol. 51, no. 3, pp.175–186, Mar. 2011.

[9] M. E. Cohen, D. L. Hudson, and P. C. Deedwania, "Applying continuous chaotic modeling to cardiac signal analysis," *IEEE Eng. Med. Biol. Mag.*, vol. 15, no. 5, pp. 97–102, May. 1996.

[10] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Nat. Acad. Sci.*, vol. 88, no. 6, pp. 2297–2301, Mar, 1991.

[11] H. B. Xie, Z. M. Gao, and H. Liu, "Classification of ventricular tachycardia and fibrillation using fuzzy similarity-based approximate entropy," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3973–3981, Apr. 2011.

[12] H. B. Xie, and Z. Z. Wang, "Mean frequency derived via Hilbert-Huang transform with application to fatigue EMG signal analysis," *Comput. Meth. Prog. Bio.*, vol. 82, no. 2, pp. 114–120, May. 2006.

[13] H. B. Xie, J. Y. Guo, and Y. P. Zheng, "Fuzzy approximate entropy analysis of chaotic and natural complex systems Detecting muscle fatigue using electromyography signals,", *Ann. Biomed. Eng.*, vol. 38, no. 4, pp. 1483–1496, Apr. 2010.

[14] A. Georgakis, L. K. Stergioulas, and G. Giakas, "Fatigue analysis of the surface EMG signal in isometric constant force contractions using the averaged instantaneous frequency," *IEEE Trans. Biomed. Eng*., vol. 50, no. 2, pp. 262–265, Feb. 2003.