31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

# Detection of synchrony in biosignals using cross fuzzy entropy

Hong-Bo Xie., Yong-Ping Zheng, *Senior Member, IEEE.*, and Jing-Yi, Guo

***Abstract*—A new method, namely Cross fuzzy entropy (C-FuzzyEn) analysis, that could enable the measurement of the synchrony or similarity of patterns between two distinct signals, was presented in this study. Tests on simulated data sets showed that C-FuzzyEn was superior to the conventional cross sample entropy (C-SampEn) in several aspects, including giving entropy definition in case of small parameters, better relative consistency, and less dependence on record length. The proposed C-FuzzyEn was then applied for the analysis of simultaneously recorded electromyography (EMG) and mechanomyography (MMG) signals during sustained isometric contraction for monitoring local muscle fatigue. The results showed that the C-FuzzyEn of EMG-MMG decreased significantly during the development of muscle fatigue. The time-decrease trend of C-FuzzyEn is similar to the mean frequency (MNF) of EMG, the commonly used muscle fatigue indicator.**

## I. Introduction

SYNCHRONY or similarity between bivariate time series are traditionally quantified with linear methods, usually coherency and spectral estimates [1, 2]. Pincus [3] developed approximate entropy (ApEn) measure from theories in the field of nonlinear dynamic analysis and chaos. Cross approximate entropy (C-ApEn) which is the generalization of ApEn was then applied by different authors for measuring synchrony between time series [4]. However, the C-ApEn is a direction dependent statistic where C-ApEn(*m,r,N*)(*v*|*u*) and its direction conjugate C-ApEn(*m,r,N*)(*u*|*v*) are unequal in most cases due to logarithm computation inside the summation [5]. Richmann and Moormann further proposed sample entropy (SampEn) by not counting self matching of template and extended it to cross sample entropy (C-SampEn) [5]. C-SampEn is direction independent and shows better relative consistency than C-ApEn in some situations. If two records show lower C-SampEn than another with one set of parameters of tolerance $r$ and embedded dimension $m$, they also have lower C-SampEn with other parameters of $r$ and $m$ [5]. Due to the above properties of C-SampEn, it has been used as an alternative nonlinear statistic to analyze the functional connectivity of distinct brain regions [6] and assess the relationship between arterial blood pressure and renal sympathetic nerve activity of anaesthetized rats [7].

Recently, a new time series regularity measure, fuzzy entropy (FuzzyEn), has been proposed [8]. In this study, we generalized the FuzzyEn approach for comparing two distinct, yet intertwined time series in a network so as to assess their degree of pattern synchrony. We applied the exponential function to bound two vectors $X_i$ and $Y_j$ from two different time series, and defined this new synchrony measure as cross-fuzzy entropy (C-FuzzyEn). We demonstrated that C-FuzzyEn would provide more freedom of parameter selection and is independent of data length. We then analyzed the C-FuzzyEn between Electromyography (EMG) and Mechanomyography (MMG) signals recorded simultaneously from biceps brachii during an exhausting isometric contraction.

This work was supported in part by the Hong Kong Research Grant Council (PolyU 5331/06E), The Hong Kong Polytechnic University (1-BB69), and Jiangsu University, PRC (07JDG40).

H. B. Xie was with Jiangsu University, Jiangsu, 212013 PR of China. He is now with the Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, P.R.China (e-mail: xiehb@sjtu.org).

Y. P. Zheng (corresponding author) is with the Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, P.R.China (Tel: 852-27667664; fax: 852-23624365; e-mail: ypzheng@ieee.org).

J. Y. Guo is with the Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong SAR, P.R.China (e-mail: 05901780r@inet.polyu.edu.hk).

## II. Methodology

### A. Cross Fuzzy Entropy

In the two proposed related asynchrony measures C-ApEn(*m,r,N*) and C-SampEn(*m,r,N*), similarity of vectors is based on Heaviside function. In a Heaviside function, the boundary is rigid: the contributions of all the data samples included are treated equally, whereas the data samples just outside are not considered. The hard boundary causes discontinuity, which may lead to abrupt changes of entropy values when the tolerance $r$ changes slightly, and even to the failure in entropy definition if no template-match can be found for small tolerance $r$. To deal with this problem, we substituted a Gaussian function for the Heaviside function in calculation of the similarity between two vectors.

Given two normalized time series of $N$ points $\{u(i): 1 \le i \le N\}$ and $\{v(j): 1 \le i \le N\}$ with unit standard deviation and $m$, in both C-SampEn and C-ApEn, vectors are formed directly from the original $m$ consecutive $u$ and $v$ values as

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

$$Y_j^m = \{v(j), v(j+1), \cdots, v(j+m-1)\} \quad (2)$$

and the distance $d_{ij}^m$ between $X_i^m$ and $Y_j^m$ is defined as

$$d_{ij}^m = d[X_i^m, Y_j^m] = \max_{k \in (0,m-1)} |u(i+k) - v(j+k)| \quad (3)$$

Under the definition, $X_i^m$ and $Y_j^m$ are considered to be similar only when $\max_{k \in (0,m-1)} |u(i+k) - v(j+k)| \le r$. So vectors'

similarity is totally determined by their absolute coordinates, which may fail in the application to signals from different sources with different fluctuations of complex nonlinear variability in synchronicity analysis.

In the definition of C-FuzzyEn, the vectors $\{X_i^m, i=1,\cdots,N-m\}$ and $\{Y_j^m, i=1,\cdots,N-m\}$ are formed as follows:

$$X_i^m = \{u(i), u(i+1), \cdots, u(i+m-1)\} - \bar{u}(i) \quad (4)$$

$$Y_j^m = \{v(j), v(j+1), \cdots, v(j+m-1)\} - \bar{v}(j) \quad (5)$$

here $X_i^m$ and $Y_j^m$ represent $m$ consecutive $u$ and $v$ values commencing with the $i$th and $j$th point, respectively, and generalized by removing baseline

$$\bar{u}(i) = \frac{1}{m}\sum_{l=0}^{m-1} u(i+l) \quad (6)$$

$$\bar{v}(j) = \frac{1}{m}\sum_{l=0}^{m-1} v(j+l) \quad (7)$$

The distance between the two vectors $X_i^m$ and $Y_j^m$ is defined as the maximum absolute difference of the corresponding scalar components of them

$$d_{ij}^m = d[X_i^m, Y_j^m] = \max_{k\in(0,m-1)} |u(i+k) - \bar{u}(i) - (v(j+k) - \bar{v}(j))| \quad (8)$$

Given $n$ and $r$, calculate the synchrony or similarity degree $D_{ij}^m$ between $X_i^m$ and $Y_j^m$ through a fuzzy function $\mu(d_{ij}^m, n, r)$

$$D_{ij}^m(n,r) = \mu(d_{ij}^m, n, r) \quad (9)$$

where the fuzzy function $\mu(d_{ij}^m, n, r)$ is the exponential function

$$\mu(d_{ij}^m, n, r) = \exp(-(d_{ij}^m)^n / r)) \quad (10)$$

Define the function $\phi^m$ as

$$\phi^m(n,r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{N-m}\sum_{j=1}^{N-m} D_{ij}^m\right) \quad (11)$$

Similarly, form $\{X_i^{m+1}\}$ and get the function $\phi^{m+1}$

$$\phi^{m+1}(n,r) = \frac{1}{N-m}\sum_{i=1}^{N-m}\left(\frac{1}{N-m}\sum_{j=1}^{N-m} D_{ij}^{m+1}\right) \quad (12)$$

Theoretically, we can define the parameter C-FuzzyEn($m$, $n$, $r$) of the two sequences as the negative natural logarithm of the conditional probability $\phi^{m+1}/\phi^m$

$$\text{C-FuzzyEn}(m,n,r) = -\lim_{N\to\infty}(\ln(\phi^{m+1}/\phi^m)) \quad (13)$$

In fact, the number of data points $N$ is finite and the result obtained through the above steps is an estimate of C-FuzzyEn when the data length is $N$, which can be denoted as

$$\text{C-FuzzyEn}(m,n,r,N) = \ln\phi^m(n,r) - \ln\phi^{m+1}(n,r) \quad (14)$$

There are three parameters that must be determined for each calculation of C-FuzzyEn. The first parameter $m$, as in C-ApEn and C-SampEn, is the length of sequences or the dimension of the vector to be compared. The other two parameters $r$ and $n$ determine the width and the gradient of the boundary of the exponential function respectively. The criterion and methods for the selection of parameters $r$ and $n$ in C-FuzzyEn estimation are similar as those used in the estimation of FuzzyEn and have been previously described [8]. Experimentally, it is convenient to set the tolerance $r$ between 0.1 and 0.3, and choose small integers for the $n$ selection.

### *B. EMG and MMG Data Sets*

The EMG and MMG signals analyzed in this paper were recorded during the voluntary isometric contractions of 12 healthy human subjects (eight males and four females, age: $30.2 \pm 4.9$ ). None of them had history of any neuromuscular disorder and each gave written informed consent prior to the experiment. The experiment was approved by the ethical committee. A pair of surface EMG self adhesive conductive gel electrodes (Axon Systems, Inc., New York, USA), with their centre 25 mm apart from each other were placed on abrased, clean skin longitudinally, immediately under the thickest point of the Biceps. The EMG reference electrode was placed on the proximal head of the ulna. The accelerometer (EGAS-FS-10-VO5, Entran Inc, Fairfield, NJ) was placed as close as possible to the EMG electrodes so as to record the MMG on the axis orthogonal to the muscle from the same motor territory. It was fixed to the skin using double adhesive tape. When the experiment began, the subject was asked to perform an elbow flexion against the lever arm to 80% of his/her maximal voluntary contraction (MVC) and maintain this value through visual feedback of the torque reading on the screen. The test was stopped when the torque dropped to approximately 70% of the MVC, which indicates the muscle is fatigued. The gain of the EMG signal was 2000 with a 10–400 Hz bandwidth and that of the MMG signal was 5000 with a bandwidth 5–400 Hz. Signals from the EMG electrodes and accelerometer were acquired at 1 KHz per channel and stored in computer for further C-FuzzyEn analysis.

## III. RESULTS

Before the pattern synchronization measures were applied to EMG-MMG analysis, they were firstly tested on two simulated datasets. The first coupled series is a MIX(0.5) and an i.i.d. uniform random sequences. The MIX($P$) time series is a sine wave of $N$ points, where $N \times P$ randomly chosen points have been replaced with random noise ($P$ is between 0 and 1) [5]. Fig. 1 shows the performances of C-FuzzyEn(2,2,$r$,$N$) and C-SampEn(2,$r$,$N$) on the two sequences with different lengths of $N$=100 and $N$=50. Another input parameter $r$ varies from 0.001 to 1.0 in steps of 0.001. Due to the direction dependence, the results of C-ApEn were not reported here. C-SampEn gave no value when $r$ is smaller than 0.07 for $N$=100 as shown in Fig. 1(a), and when $r$ is smaller than 0.18 for $N$=50 in Fig. 1(b). So, the calculation of C-SampEn was confronted with the problem of parameter limitation; the shorter the datasets are, the larger the minimum tolerance $r$ is needed. But the problem did not affect the calculation of C-FuzzyEn even for rather small $r$ value (here 0.001).

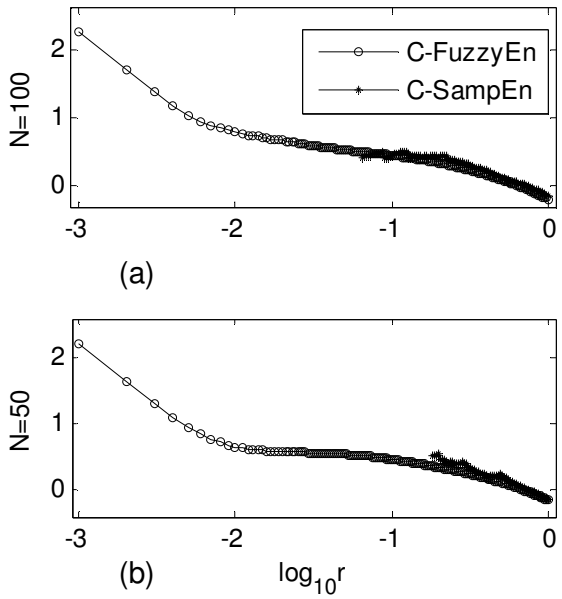


Fig. 1. C-FuzzyEn(2,2,*r*,*N*) and C-SampEn(2,*r*,*N*) as functions of tolerance threshold *r* for MIX(0.5) and a sequence of independent identical distribution uniform numbers for *r* varying from 0.001 to 1.0 in steps of 0.001 for (a) N=100 and (b) N=50. Abscissa and ordinate are all logarithm with base 10

Then, we examined two i.i.d uniform random numbers of different lengths to test whether the C-FuzzyEn values would change as a function of the record length. The data length ranged from 50 to 500 in steps of 10. Fig. 2(a) and (b) shows the effects of increasing data length on the estimation of C-FuzzyEn(*m*,2,0.2,*N*) and C-SampEn(*m*,0.2,*N*) with embedded dimensions $m = 2$ and 3, respectively. The Mean and standard deviation of C-FuzzyEn(*m*,2,0.2,*N*) are $2.237 \pm 0.020$, $1.977 \pm 0.029$ for $m = 2$ and 3, respectively, and C-SampEn(*m*,0.2,*N*) $2.151 \pm 0.161$, $2.088 \pm 0.331$. Both Fig. 2(a) and Fig. 2(b) indicated that the C-FuzzyEn yielded consistent and fairly robust estimates with very low standard deviations for various data lengths, i.e., it is largely independent of record length. This finding is rather valuable for many practical applications where data from long recordings are not available. In comparison, it was revealed that the C-SampEn didn't have the property. When $m = 2$, for shorter data length, the C-SampEn estimate showed more fluctuations. It became worse with significant fluctuations in the extensive data length span when $m = 3$.

Fig. 3 shows the typical raw EMG and MMG signals acquired from subject 3 in the course from rest to fatigue. In order to monitor the change of C-FuzzyEn of EMG-MMG with time, each signal was first segmented into consecutive, nonoverlapping epochs of 500 ms. For each epoch, the EMG and MMG signals were normalized and the procedure as described above was followed for the calculation of C-FuzzyEn. Fig. 4(a) shows the time course of the C-FuzzyEn for subject 3. It could be observed that the C-FuzzyEn decreased significantly during the course of muscle contraction. A linear regression analysis was conducted on the sequential C-FuzzyEn values. The results from all the other subjects with different epoch lengths were similar as shown in Fig. 4(a), indicating that the C-FuzzyEn of EMG-MMG might qualify as a new indicator of muscle fatigue. The decreasing of C-FuzzyEn was confirmed by a statistically significant negative slope for all participants (ANOVA, $\alpha = 0.05$, $p < 0.02$)

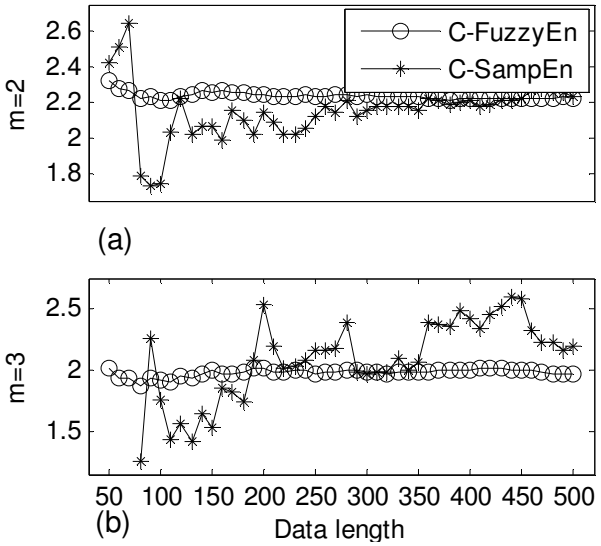


Fig. 2. (a) C-fuzzyEn(*m*,2,0.2,*N*) and (b) C-SampEn (*m*,0.2,*N*) for *N* varying from 50 to 500 in steps of 10.

During sustained contractions, when muscle length and tension are held constant, muscle conduction velocity decreases with fatigue and this phenomenon is reflected in EMG signals through a decrease of its mean frequency (MNF) [9]. MNF so far has been hailed as the gold standard for muscle fatigue assessment by using EMG under 'static' conditions [9]. The slope and intercept of the linear regression for the time course of MNF have also served as an important quantitative fatigue index [9]. In order to further examine the effectiveness of the C-FuzzyEn statistic for muscle fatigue assessment, the MNF analysis was also applied to the EMG with different epoch lengths for comparison purpose. The time course of the MNF of subject 3 with 500ms epoch length is shown in Fig. 4(b). A similar time-decrease trend in the two variables was observed. Then, the C-FuzzyEn and MNF were normalized by their respective first epoch value. A least-square error linear regression was fitted to each normalized C-FuzzyEn and MNF over the period of contraction to obtain the slope and intercept respectively. The Mann–Whiteney U test compared the slope and intercept of normalized C-FuzzyEn and MNF obtained from all the subjects. The test results demonstrated that there is no significant difference for both slope and intercept between the C-FuzzyEn and MNF. The results suggested that C-FuzzyEn may potentially become a new reliable method for muscle fatigue assessment.

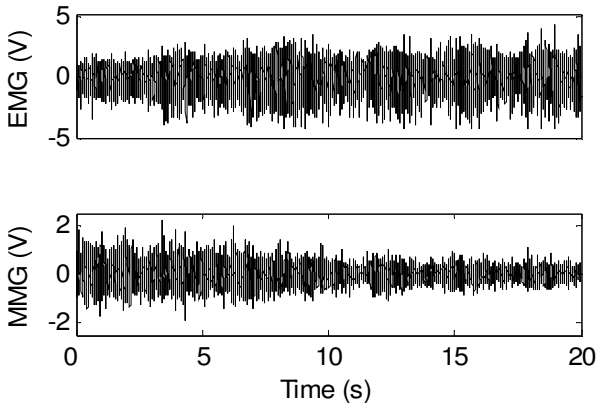


Fig. 3. The typical EMG (a) and MMG (b) signals acquired from subject 3 during sustained isometric contraction.

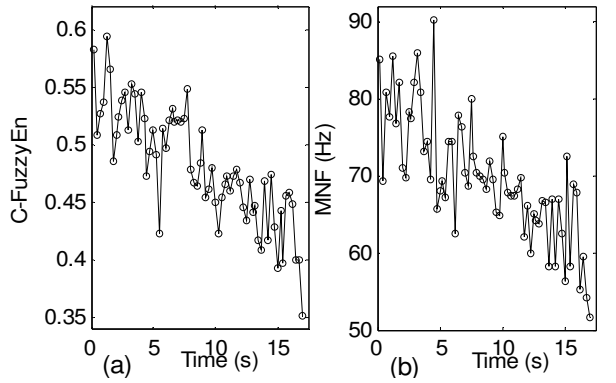


Fig. 4. Time courses of EMG-MMG C-FuzzyEn (a) and EMG MNF (b) of subject 3. The duration of the analysis window is 500 ms.

## IV. Discussion and Conclusions

In this paper, we proposed a new pattern synchrony testing method namely cross fuzzy entropy (C-FuzzyEn), for measuring the degree of synchrony or similarity between two series and applied it for the analysis of muscle fatigue using EMG and MMG signals. C-FuzzyEn describes the synchronicity of patterns rather than the time synchronization of the signals. It assigns a nonnegative number to the synchronicity of patterns in two series, with smaller values corresponding to more common features in the pattern architecture and larger values indicating larger signal differences in their pattern architecture.

Like C-SampEn, C-FuzzyEn is the negative natural logarithm of the conditional probability that two time series of length $N$, having similar patterns for $m$ points within a boundary $r$, will also repeat for $m+1$ points. Similar to C-SampEn, C-FuzzyEn possesses the direction independent property that C-ApEn lacks. However, unlike C-SampEn, where the similarity of two vectors is based on Heaviside function, C-FuzzyEn employs exponential function to bound two vectors' similarity. In addition, the method employed to construct $m$-dimension vectors in C-FuzzyEn is also different from that of C-SampEn and C-ApEn. In C-FuzzyEn, the vector sequences are generalized by removing the baseline using equations (4) and (5). In this way, vectors' similarity, determined by the distance $d_{ij}^m$ between the vectors as expressed in equation (8), depends on their shapes rather than their absolute coordinates, which makes the synchronicity definition fuzzier. So, the C-FuzzyEn outperforms C-SampEn not only because it doesn't suffer limitation of parameters, but also because it is independent to data length. Actually it is especially applicable for short data sets. Nonlinear methods of signal analysis can be more useful when characterizing internal structure of complex dynamics. However, typical nonlinear analysis would demand a large number of data points to reconstruct dynamics of the regulation system. The C-FuzzyEn proposed in this study can be applied to a relatively small amount of serial data for comparing two distinct, yet intertwined time series in a network. Since most physiological data are of limited length, C-FuzzyEn is thus especially attractive to such real data analysis.

As an application, we analyzed the C-FuzzyEn change of EMG-MMG signals during sustained isometric contraction. It was found that the pattern synchronization of EMG-MMG is significantly increasing in the fatigue state. Though the mechanism underlying this phenomenon is not clear, the results suggest that the C-FuzzyEn is another choice to quantify the muscle fatigue. C-FuzzyEn can also be applied to quantify joint pattern synchrony between any other two separate but linked time series with short data length in noisy background.

## Acknowledgment

This work is partially supported by the Hong Kong Research Grant Council (PolyU 5331/06E), The Hong Kong Polytechnic Unversity (1-BB69), and Jiangsu University, PRC (07JDG40).

## References

[1] A. Porta, G. Baselli, F. Lombardi, N. Montano, A. Malliani, and S. Cerutti S, "Conditional entropy approach for the evaluation of coupling strength," *Biol. Cybern*. Vol. 81, pp. 119–129, 1999.

[2] C.J. Stam, and B.W. van Dijk, "Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets," *Physica D.* vol.163, pp. 236–251, 2002.

[3] S.M. Pincus, "Approximate entropy as a measure of system complexity," *Proc. Natl. Acad. Sci*. vol. 88, pp. 2297–2301, 1991.

[4] P.Y. Liu, S.M. Pincus, and D.M. Keenan, "Analysis of bidirectional pattern synchrony of concentration-secretion pairs: implementation in the human testicular and adrenal axes," *Am J. Physiol-Regul. Integr. Comp. Physiol*. Vol. 288, pp. R440–R446, 2005.

[5] J.S. Richmann, and J.R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am J. Physiol. Heart. Circ. Physiol*. vol. 278, pp. H2039–H2049, 2000.

[6] Z.H. Hu; and P.C. Shi, "Interregional Functional Connectivity via Pattern Synchrony," in: 9th Int. *Conf. Control, Automation, Robotics and Vision,* pp. 1–6, 2006.

[7] T. Zhang, Z. Yang, and J.H. Coote, "Cross-sample entropy statistic as a measure of complexity and regularity of renal sympathetic nerve activity in the rat," *Exp. Physiol.* vol. 92, pp. 659–669, 2007.

[8] W.T. Chen, Z.Z. Wang, H.B. Xie, and W.X. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans Neural. Syst. Rehabil. Eng*. Vol. 15, pp. 266–272, 2007.

[9] C.J. De Luca, "Myoelectric manifestations of localized muscular fatigue," *Crit. Rev. Biomed. Eng*. vol. 11, pp. 251-279, 1984.