Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007

A New Nonlinear Similarity Measure for Multichannel Biological Signals

Jian-Wu Xu, Hovagim Bakardjian, Andrzej Cichocki, Jose C. Principe

***Abstract*— We propose a novel similarity measure, called the *correntropy coefficient*, sensitive to higher order moments of the signal statistics based on a similarity function called crosscorrentopy. Crossorrentropy nonlinearly maps the original time series into a high-dimensional reproducing kernel Hilbert space (RKHS). The correntropy coefficient computes the cosine of the angle between the transformed vectors. Preliminary experiments with simulated data and multichannel electroencephalogram (EEG) signals during behavior studies elucidate the performance of the new measure versus the well established correlation coefficient.**

I. Introduction

Quantification of dynamical interdependence in multi-dimensional complex systems with spatial extent provides a very useful insight into their spatio-temporal organization. For instance, understanding how functional interaction among different brain regions is important in the context of brain information processing [1]. In practice, the underlying system dynamics are not accessible directly. Only the observed time series can help decide whether two time series collected from the system are statistically independent or not, and further elucidate any hidden relationship between them. Extracting such information becomes more difficult if the underlying dynamical system is nonlinear or the couplings among the subsystems are nonlinear and non-stationary.

There has been extensive research aimed at detecting the underlying relationships in multi-dimensional dynamical systems. The classical methodology employs a linear approach, in particular, the cross correlation and coherence analysis [2], [3]. Cross correlation measures the linear correlation between two signals in the time domain, while the coherence function specifies the linear associations in the frequency domain by the ratio of squares of cross spectral densities divided by the products of two auto-spectra. There have been several extensions of correlation to more than two pairs of time series such as directed coherence, directed transfer functions and partial directed coherence [4], [5]. Unfortunately, linear methods only capture linear relationships between the time series, and might fail to detect nonlinear interdependencies between the underlying dynamical subsystems.

Nonlinear measures include mutual information and state-space methods. The generalized mutual information function appears as a widely used technique for quantifying nonlinear correlation [6]. However, a large quantity of noise-free stationary data is required to estimate these measures based on information theory, which restricts their applications in practice. Rosenblum *et al*. proposed phase synchronization to quantify interdependencies between dynamical systems [7]. In this approach, the instantaneous phase using Hilbert transforms is computed and interdependence is specified in terms of time-dependent phase locking. Another line of research is based on nonlinear dynamical system theory. The similarity-index technique and its modifications have been proposed to measure the nonlinear asymmetric interdependencies between time series by computing the ratio of average distances between index points, their nearest neighbors and their mutual nearest ones [8], [9], [10]. Stam *et al.* proposed the synchronization likelihood to offer a straightforward normalized estimate of the dynamical coupling between interacting systems [11]. There are several drawbacks associated with these techniques based on state space embedding. Estimating the embedding dimension of times series corrupted by measurement noise for a valid reconstruction, searching a suitable neighborhood size and finding a constant number of nearest neighbors are a few of many constraints that severely affect the estimation accuracy [10], [12].

In this paper, we introduce a novel functional measure, called the *correntropy coefficient*, to characterize dynamical interdependencies between interacting systems. Correntropy is a new concept introduced by our research group to quantify similarity using the higher order statistics in random processes based on a reproducing kernel Hilbert space method [13]. By nonlinearly transforming the random processes into a high dimensional RKHS and computing the "conventional" correlation on the transformed signals, correntropy is sensitive to both the higher order statistical distribution information and temporal structure of the original random process. Correntropy can be applied both to one time series, called the *autocorrentropy*, or a pair of multidimensional random variables, called the *crosscorrentropy*. In the context of characterization of interdependence between coupled dynamical systems, the crosscorrentropy is used. In this paper, we work with the *centered crosscorrentropy*, which subtracts the mean of the transformed data in the same spirit of the conventional covariance function. Based on the centered crosscorrentropy, the *correntropy coefficient* is defined to characterize the interdependence between the two transformed time series (or the original two random variables). If two random variables or two time series are

Jian-Wu Xu and Jose C. Principe are with the Computational NeuroEngineering Laboratory, Department of Electrical and Computer Engineering, University of Florida, Gainsville, FL 32611, USA (phone: 1-352-392-2682; fax: 1-352-392-0044; email: {jianwu, principe}@cnel.ufl.edu).

Hovagim Bakardjian and Andrzej Cichocki are with the Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute, Wako-shi, Saitama, 351-0198, JAPAN (phone: 81-48-467-9668; fax: 81-48-467-9686; email: {hova, cia}@brain.riken.jp).

independent, then the correntropy coefficient becomes zero; if two are the same, then it attains maximum value 1; the correntropy coefficient achieves -1 when the two random variables are in the opposite directions.

The paper is organized as follows. In Sec. II, we briefly introduce the newly proposed correntropy concept and present the method of the correntropy coefficient in details. We also explores the correntropy coefficient from geometrical perspective based on the reproducing kernel Hilbert space. Experiments of the correntropy coefficient on simulated data and real EEG signals are presented in Sec. III. We conclude the work in Sec. IV.

II. METHOD

Given two random variables x and y, the "generalized" crosscorrelation function, called *crosscorrentropy* [14] is defined as

$$V(x, y) = E[\kappa(x, y)], \tag{1}$$

where E denotes the statistical expectation operator and $\kappa(\cdot, \cdot)$ is a symmetric positive definite kernel function. The most widely used kernel function is the Gaussian kernel which is given by

$$\kappa(x, y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{(x-y)^2}{2\sigma^2}\}, \tag{2}$$

where σ is the kernel width. When no ambiguity exists we will call crosscorrentropy simply correntropy. Correntropy is also a positive definite function and since the RKHS structure is defined by the kernel, the RKHS induced by correntropy has unique properties. If we apply the Taylor series expansion for the Gaussian kernel in the correntropy definition, it can be easily seen that it compactly contains all even moments of the random variables $(x - y)$ [13]. Hence, correntropy includes higher order statistical information about the random variables.

From the traditional parameter estimation point of view, the extraction of the mean value is the most important preprocessing step before computing the estimate of interest. However, with the presence of the nonlinear transformation in the definition of correntropy, we are unable to explicitly remove the mean and have to rely on a different approach. Therefore we define the "generalized" crosscovariance function, called the *centered crosscorrentropy*, as

$$\begin{aligned} U(x, y) &= E[\kappa(x, y)] - E_x E_y[\kappa(x, y)] \qquad (3) \\ &= \iint \kappa(x, y)(f_{X,Y}(x, y) - f_X(x) f_Y(y)) dx dy \end{aligned}$$

Notice that the crosscorrentropy is the joint expectation of $\kappa(x, y)$, while the centered crosscorrentropy is the difference between the joint and product of marginal expectations of $\kappa(x, y)$.

An important observation here is that when two random variables x and y are independent, in other words, the joint probability density function (PDF) $f_{X,Y}(x, y)$ equals the product of two marginal PDFs $f_X(x) f_Y(y)$, then $E[\kappa(x, y)] = E_x E_y[\kappa(x, y)]$. Therefore the centered crosscorrentropy reduces to zero only if the two random variables are independent. This is a much stronger condition than uncorrelatedness, as required by the conventional cross-covariance function in order to achieve zero. It clearly shows the incorporation of higher order information in centered crosscorrentropy.

By normalizing the centered crosscorrentropy, we can define the "generalized" correlation coefficient, called the *correntropy coefficient*, as

$$\eta = \frac{U(x, y)}{\sqrt{U(x, x) U(y, y)}}, \tag{4}$$

where $U(x, y)$ is the centered crosscorrentropy function for x and y, and $U(x, x)$ and $U(y, y)$ are the centered auto-correntropy functions for variables x and y respectively. The absolute value of the correntropy coefficient is bounded by 1. When two random variables are independent, the correntropy coefficient becomes 0. It attains 1 if two random variables are the same and -1 when they are in opposite direction with a large kernel size. Hence, the correntropy coefficient is a suitable interdependence measure for interacting dynamical systems. If two random variables are uncorrelated but not independent, the correntropy coefficient shall produce a non-zero value (which depends on the parameter, kernel width used in Gaussian kernel). However, the conventional correlation coefficient will be zero. In the context of generalized synchronization, the correntropy coefficient shall characterize both higher order relationship and nonlinearity between interacting systems.

According to the Moore-Aronszajn Theorem [15] of Hilbert space analysis, the positive definite kernel $\kappa(\cdot, \cdot)$ in the definition of the centered crosscorrentropy uniquely determines a data independent reproducing kernel Hilbert space, denoted as $\mathcal{H}_\kappa$. It turns out that the input data is mapped onto the surface of a sphere with the Gaussian kernel [13]. Moreover, it can be easily proved that the centered crosscorrentropy itself is also non-negative definite. Consequently, the centered crosscorrentropy uniquely determines another, data dependent, reproducing kernel Hilbert space, denoted by $\mathcal{H}_U$, by the Moore-Aronszajn Theorem [15]. This enables us to get insights into the geometry of the correntropy coefficient. According to the Mercer's theorem [16], any continuous non-negative symmetric kernel function possesses an eigen-decomposition as follows

$$\begin{aligned} U(x, y) &= \sum_{n=0}^{\infty} \gamma_n \psi_n(x) \psi_n(y) = \langle \Psi(x), \Psi(y) \rangle \qquad (5) \\ &\quad \Psi : x \mapsto \sqrt{\gamma_n} \psi_n(x), \quad n = 1, 2, \ldots, \end{aligned}$$

where γ_n and ψ_n are eigenvalues and eigen-functions for the centered crosscorrentropy respectively, and $\langle \, , \, \rangle$ denotes the inner product between two infinite dimensional vectors. Notice that the nonlinear map Ψ has implicitly embedded the expectation operator so that every vector in $\mathcal{H}_U$ becomes deterministic and contains statistical information of signals, hence it is data dependent. Substituting Eq. (5) into the definition of the correntropy coefficient Eq. (4), we obtain

$$\eta = \frac{\langle \Psi(x), \Psi(y) \rangle}{\|\Psi(x)\| \|\Psi(y)\|} = \cos\theta, \tag{6}$$

where $\|\Psi(x)\|$ and $\|\Psi(y)\|$ are the length of two vectors $\Psi(x)$ and $\Psi(y)$ in $\mathcal{H}_U$ respectively, and θ is the angle between these two vectors. With this geometrical interpretation, the correntropy coefficient essentially computes the cosine of the angle between two nonlinear transformed vectors in RKHS $\mathcal{H}_U$ induced by the centered crosscorrentropy. In particular, if two vectors are orthogonal, then θ is 90° and η equals 0; if two vectors are in the same direction, then θ is 0° and η equals 1, while two vectors are in the opposite direction, θ becomes 180° and η equals -1. Orthogonality between vectors $\Psi(x)$ and $\Psi(y)$ in $\mathcal{H}_U$ corresponds to independence between random variables x and y. When two vectors are in the same or opposite directions, this suggests a strong dependence between random variables x and y.

The RKHS approach to analyze the conventional correlation function was originally proposed by Parzen [17] because the correlation function is also non-negative definite, thus it determines a unique reproducing kernel Hilbert space, denoted as $\mathcal{H}_R$. Grenander analyzed the standard correlation coefficient from RKHS perspective in [18]. Both $\mathcal{H}_R$ and $\mathcal{H}_U$ are data dependent reproducing kernel Hilbert spaces, however $\mathcal{H}_U$ implicitly embeds $\mathcal{H}_\kappa$ which incorporates higher order statistics intrinsic in the data. Therefore the correntropy coefficient requires independence of two signals to make two corresponding vectors in $\mathcal{H}_U$ orthogonal, while the standard correlation coefficient only needs uncorrelatedness. This also reinforces the claim that correntropy coefficient is able to quantify the nonlinearity and higher order relationship between interacting dynamical systems in terms of synchronization detection.

In practice, we only have a finite number of data points or time series samples available from the dynamical system. So we have to work with the estimate of the correntropy coefficient. Substituting the definition of centered crosscorrentropy (3) into the correntropy coefficient (4) and approximating the ensemble average by the sample mean, we can obtain an estimate of the correntropy coefficient directly from data,

$$\hat{\eta} = \frac{\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\kappa(x_i, y_i) - \dfrac{1}{N^2}\displaystyle\sum_{i,j=1}^{N}\kappa(x_i, y_j)}{\sqrt{\kappa(0) - \dfrac{\sum_{i,j=1}^{N}\kappa(x_i, x_j)}{N^2}}\sqrt{\kappa(0) - \dfrac{\sum_{i,j=1}^{N}\kappa(y_i, y_j)}{N^2}}},$$

where N is the total number of samples, $\frac{1}{N^2}\sum_{i,j=1}^{N}\kappa(x_i, y_j)$ is called the *cross information potential* between x and y, $\frac{1}{N^2}\sum_{i,j=1}^{N}\kappa(x_i, x_j)$ and $\frac{1}{N^2}\sum_{i,j=1}^{N}\kappa(y_i, y_j)$ are again the information potential for x and y respectively [19], and $\kappa(0)$ is the value of Gaussian kernel (2) when the argument $(x - y) = 0$.

III. Experiments

In this section, we apply the correntropy coefficient to simulated data and real EEG signals collected from a behavioral experiment paradigm.

A. Two Unidirectionally Coupled Hénon maps

The first experiment compared the correntropy coefficient, the conventional correlation coefficient and the similarity index, a nonlinear interdependence measure proposed in [9], in detecting nonlinear interdependence of two unidirectionally coupled Hénon maps $\mathbf{X}$ and $\mathbf{Y}$ defined as

$$\mathbf{X}: \begin{cases} x_{n+1} = 1.4 - x_n^2 + b_x u_n \\ u_{n+1} = x_n, \end{cases}$$

$$\mathbf{Y}: \begin{cases} y_{n+1} = 1.4 - [Cx_n + (1-C)y_n]y_n + b_y v_n \\ v_{n+1} = y_n. \end{cases}$$

Notice that system $\mathbf{X}$ drives system $\mathbf{Y}$ with a nonlinear coupling strength C. C ranges from 0 to 1 with 0 being no coupling and 1 being complete coupling. Parameters b_x and b_y are both set to 0.3 as canonical values for the Hénon map when analyzing identical systems and $b_x = 0.3$ and $b_y = 0.1$ for non-identical systems. For each coupling strength, we discard the first 10000 time series samples as a transient and keep the next 500 data points for our experiment. The kernel width σ used in Gaussian kernel (2) is chosen to be 0.001 using the Silverman's rule [20]:

$$\sigma = 0.9AN^{-1/5},$$

where A is the smaller value between standard deviation of data samples and data interquartile range scaled by 1.34, and N is the number of data samples. The three measures are calculated between x and y time series.

Fig.1 shows the correlation coefficient, the correntropy coefficient and the similarity index as functions of coupling strength C. It can be seen that the correntropy coefficient generates exactly the same result as the similarity index [9], where both measures attain the value 0 for $C < 0.7$ suggesting no synchronization between two systems and reach the value 1 for $C \geq 0.7$ indicating perfect synchronization. The critical threshold $C = 0.7$ corresponds to the point when the maximum Lyapunov exponent of the response system becomes negative and identical synchronization between the systems takes place [9]. On the other hand the conventional correlation coefficient performs erratically in the unsynchronized region $C < 0.7$. This clearly demonstrates that the correntropy coefficient outperforms the correlation coefficient in characterization of nonlinear coupling between two dynamical systems. Compared to the similarity index, the correntropy coefficient has the advantage of avoiding estimating embedding dimension, choosing nearest neighborhood and other problems associated with state space embedding method [9], [12]. The computational complexity of the correntropy coefficient is still manageable, and the kernel size is easy to estimate.

Next we test how sensitive the correntropy coefficient is to time dependent sudden change in the dynamics of interacting systems due to coupling strength. In experiment, change in coupling strength can cause sudden change in the dynamics of interacting systems, which basically generates nonstationarity in time series. To study such transient dynamical

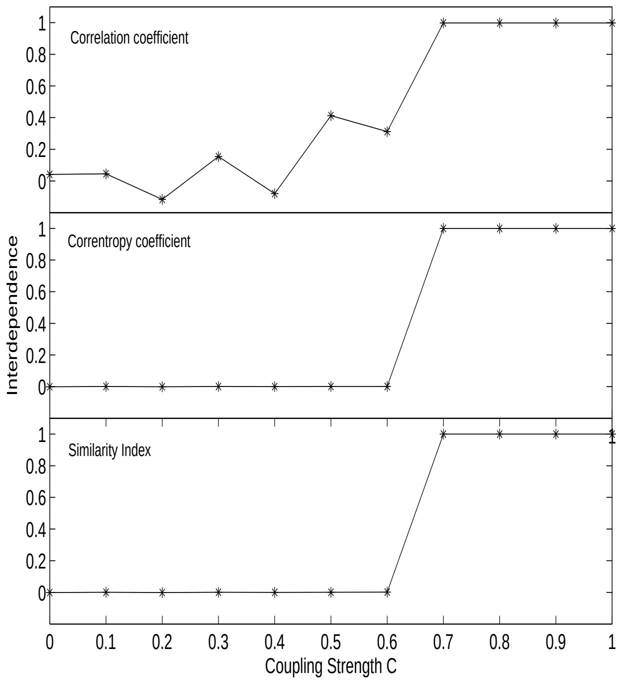

Fig. 1. Comparison of the correlation coefficient, the correntropy coefficient and the similarity index in characterization of synchronization between two unidirectionally coupled Henon map.

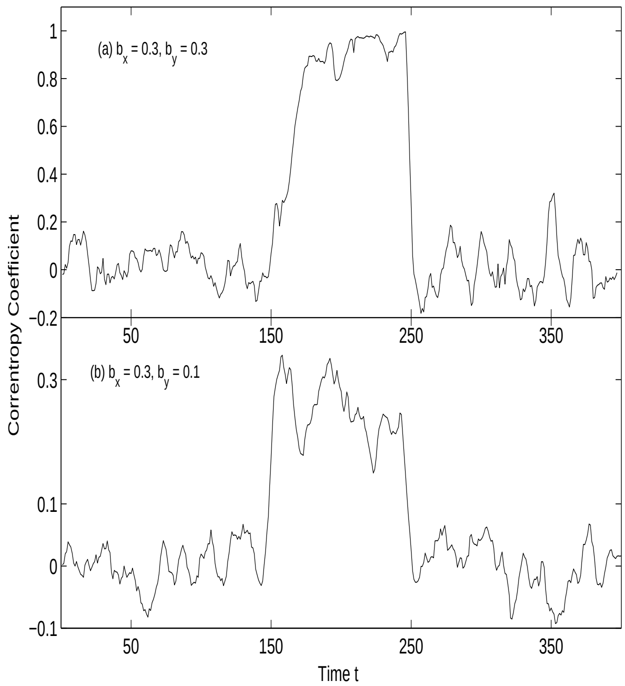

Fig. 2. Time dependent of the correntropy coefficient for unidirectionally coupled Hénon maps (a) identical ($b_x = b_y = 0.3$) and (b) non-identical ($b_x = 0.3$ and $b_y = 0.1$). The coupling strength C = 0.8 for the interval between 150th and 250th time samples and C = 0 for the rest of time.

phenomenon, both identical ($b_x = b_y = 0.3$) and non-identical ($b_x = 0.3$ and $b_y = 0.1$) Hénon maps are considered here. Dynamical systems are coupled only during a single epoch and otherwise uncoupled for both cases [10], [11]. We set the coupling strength $C = 0$ for $n \leq 10150$ and $n \geq 10250$ and $C = 0.8$ for $10150 < n < 10250$. Only 400 data samples are plotted after the first 10000 data are discarded as transient. The sliding window used to compute the correntropy coefficient is chosen to contain 8 data. Kernel size is set to 0.2 for identical map and 0.3 for non-identical map. The results are averaged over 20 independent realizations of different initial conditions ranging 0 to 1. Fig. 2 plots the correntropy coefficient curves for identical and non-identical maps. In uncoupled regions, η fluctuates around 0.01 baseline for identical map and 0.001 for non-identical map. A sharp and clear increase occurs at t = 150 when 0.8 coupling strength between systems $\mathbf{X}$ and $\mathbf{Y}$ is introduced, and there is a sharp and clear decrease in η falling off back to the baseline level when coupling strength between two systems reduces to zero at t = 250. The interval where η is noticeably higher than the baseline level matches nicely to the coupling interval. This phenomenon is observed both in identical and non-identical Hénon maps. Therefore, the correntropy coefficient is potentially able to detect sudden change in the coupling between two interacting dynamical systems with a high temporal resolution, which makes this measure suitable for non-stationary data sets.

B. EEG signals

In the second experiment, we applied the correntropy coefficient to real EEG signals. The electrical potentials on the surface of the scalp of a human subject were measured and recorded with the NeuroScan EEG system (NeuroScan Inc., Compumedics, Abbotsford, Australia). A 64-channel cap was used with electrode locations according to the extended international 10/20 system and with a linked-earlobe reference. Horizontal and vertical electrooculogram (HEOG and VEOG) signals were also recorded for artifact rejection using two sets of bipolar electrodes. The data sampling rate was fixed at 1000Hz and the online band-pass filter range was set to be maximally wide between 0.05Hz and 200Hz. Subjects were presented repeatedly (200 times) with uni-modal auditory and visual stimuli delivered in the central visual and auditory spaces simultaneously and with the same strength to the left/right eyes and ears, as well as with simultaneous cross-modal combinations. For the purpose of this study, only the uni-modal data was used. The visual stimuli consisted of 5x5 black and white checkerboards presented for 10ms, while the auditory stimuli were 2000Hz tones with durations of 30ms. The time interval between the stimuli in any of the experimental conditions was random between 1500ms and 2000ms. Following standard eye-movement artifact rejection procedures and segmentation into single epochs with alignment at the onset of the stimuli, all artifact-free epochs were averaged and normalized to zero mean and unit variance and low-pass filtered at 0-40Hz for further analysis. We then applied the correntropy coefficient

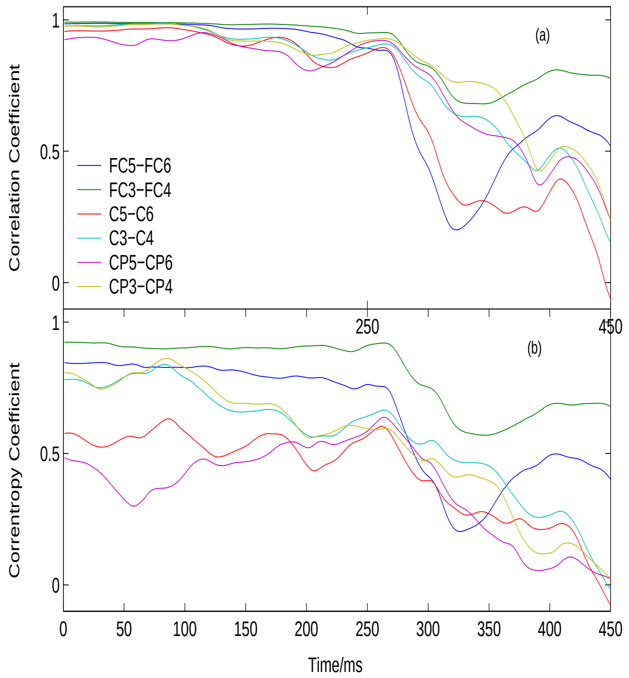

Fig. 3. Comparison of the correlation coefficient and the correntropy coefficient in characterization of synchronization among auditory cortex for audio stimuli EEG signal.

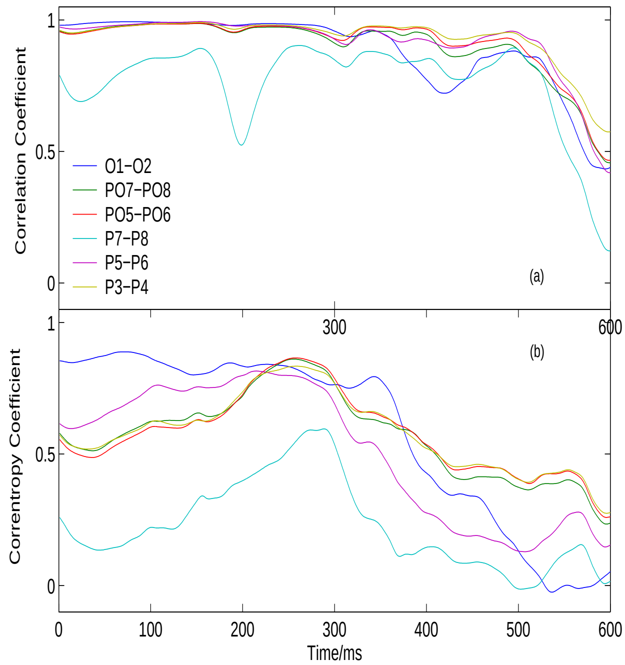

Fig. 4. Comparison of correlation coefficient and correntropy coefficient in characterization of synchronization among occipital cortex for visual stimulus EEG signal.

to the averaged data to quantify the bilateral synchronization or couplings among the corresponding sensory areas of the brain. In order to test whether the correntropy coefficient was able to detect any nonlinear couplings in the EEG signals, the results were compared to the conventional correlation coefficient. A window size of 20ms data is used to calculate both measures corresponding to the duration of a single dipole activation in the cortex [21]. The kernel width σ in Gaussian kernel (2) used in correntropy coefficient was chosen to be 0.4.

Fig.3 (a) and (b) show plots of the correlation and correntropy coefficients for the auditory areas of the brain as a function of time after the subject was exposed only to the audio stimuli. Several bilaterally-symmetrical pairs of electrodes were selected in the vicinity of the auditory cortex, so that both measures were computed for pairs FC5-FC6, FC3-FC4, C5-C6, C3-C4, CP5-CP6, CP3-CP4. As shown in Fig.3 (a) and (b), there are two distinct time intervals 0-270ms and 270-450ms in the auditory response. Both correlation and correntropy coefficients drop at 270ms. This suggests that both measures are able to detect the changes in inter-hemispheric synchronization of the auditory regions. However, as the electrodes are chosen in different locations away from the auditory cortex, it is expected that during the synchronization phase (0-270ms) the synchronization measures for different pairs should be different. Fig.3 (a) shows that the correlation coefficients for all 6 pairs are grouped together and are unable to detect the difference in activation, while Fig.3 (b) suggests that the correntropy coefficient can differentiate successfully the synchronization strength among different areas of the cortex above the left and right auditory regions. Notably, as expected from previous studies, pairs FC5-FC6 and FC3-FC4 exhibit stronger synchronization strength than the others, while most posterior pairs CP5-CP6 and C5-C6 have weaker synchronization strength. Also the synchronization patterns reveal lateral similarity in time for the pairs FC5-FC6 and FC3-FC4, for CP5-CP6 and C5-C6, and for CP3-CP4 and C3-C4. Furthermore the correntropy coefficients for pairs C5-C6, C3-C4 and CP3-CP4 peak simultaneously at 90ms which corresponds to the first mean global field power (MGFP) peak of the EEG signal. These differences indicate that the correntropy coefficient is more sensitive and is able to extract more information as a synchronization measure than the conventional correlation coefficient.

We also compared both measures when applied to the visual cortical areas. The measures are presented in Fig.4 as a function of time when the subject is exposed only to visual stimuli. Again, a window size of 20ms data is used to compute both the correlation and the correntropy coefficients, and the kernel width σ is again set to 0.4 as in the previous case. We also chose bilaterally symmetrical pairs of electrodes O1-O2, PO7-PO8, PO5-PO6, P7-P8, P5-P6 and P3-P4. In Fig.4 (b) the correntropy coefficients for all pairs except for O1-O2 show similar synchronization patterns. The correntropy coefficient increases at first, then reaches a peak around 275ms, after which it drops to lower levels. The maximum values of the correntropy coefficients around 275ms correspond to the peak P1 in the visual evoked potential [22]. As expected the synchronization between occipital channels O1 and O2 has the maximum strength and stays high until it decreases around 350ms. Thus the correntropy coefficient shows that the extra-striate visual networks become increasingly recruited and synchronized until about 275ms after the stimulus onset, while the primary visual cortex is highly synchronous for a longer period of time, until about 350ms after onset. The channels pair P7 and P8 exhibits the weakest synchronization strength since it is located the farthest away from the primary visual cortex compared to other electrode pairs. On the other hand, the correlation coefficients for most channel pairs group together and display the same level of synchronization until its sharp decrease at around 500ms (except for P7-P8). The synchronization between P7 and P8 has irregular patterns with a local minimum around 200ms. This comparison clearly demonstrates that also in this case the correntropy coefficient measure outperforms the correlation coefficient in the quantification of the EEG signal coupling between the bilateral occipital regions of the brain in response to visual stimuli.

IV. Conclusion

In this paper, we propose the correntropy coefficient as a novel nonlinear interdependence measure. Due to a positive definite kernel function, the correntropy coefficient implicitly maps the original random variables or time series into an infinite dimensional reproducing kernel Hilbert space which is uniquely induced by the centered crosscorrentropy function and essentially computes the cosine of the angle between the two transformed vectors. Orthogonality in RKHS $\mathcal{H}_U$ corresponds to independence between original random variables. Comparisons between the correntropy coefficient and the conventional correlation coefficient on simulated two unidirectionally coupled Hénon maps time series and EEG signals collected from sensory tasks clearly illustrate that the correntropy coefficient is able to extract more information than the correlation coefficient in quantification of synchronization between interacting dynamical systems. These preliminary findings warrant further investigation to characterize better and validate this new nonlinear measure of similarity, which is much simpler to apply than the ones available in the literature. The sensitivity to the kernel size and the dependence to the amplitude of the signals must also be fully studied theoretically and practically.

Acknowledgment

This work was partially supported by NSF grant ECS-0601271, Graduate Alumni Fellowship from University of Florida and research scholarship from RIKEN Brain Science Institute.

References

[1] C. J. Stam, T. van Woerkom, and W. S. Pritchard, “Use of nonlinear eeg measure to characterize eeg changes during mental activity,” *Electroenceph. Clin. Neurophysiol.*, vol. 99, pp. 214–224, 1996.

[2] J. C. Shaw, “An introduction to the coherence function and its use in EEG signal analysis,” *Journal of medical engineering & technology*, vol. 5, no. 6, pp. 279–288, 1981.

[3] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*. New York: Wiley, 2000.

[4] L. A. Baccala and K. Sameshima, “Overcoming the limitations of correlation analysis for many simultaneously processed neural structures,” *Progress in Brain Research*, vol. 130, pp. 33–47, 2001.

[5] ——, “Partial directed coherence: A new concept in neural structure determination,” *Biol. Cybern.*, vol. 84, pp. 463–474, 2001.

[6] B. Pompe, “Measuring statistical dependencies in a time series,” *Journal of Statistical Physics*, vol. 73, pp. 587–610, 1993.

[7] M. G. Rosenblum, A. S. Pikovsky, and J. Kurths, “Phase synchronization of chaotic oscillators,” *Phys. Rev. Lett.*, vol. 76, no. 11, pp. 1804–1807, 1996.

[8] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, “A robust method for detecting interdependencies: Application to intracranially recorded eeg,” *Physica D*, vol. 134, pp. 419–430, 1999.

[9] R. Q. Quiroga, J. Arnhold, and P. Grassberger, “Learning driver-response relationships from synchronization patterns,” *Phys. Rev. E*, vol. 61, no. 5, pp. 5142–5148, May 2000.

[10] J. Bhattacharya, E. Pereda, and H. Petsche, “Effective detection of coupling in short and noisy bivariate data,” *IEEE Trans. Syst., Man and Cybern. B*, no. 1, pp. 85–95, February 2003.

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

[12] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, UK: Cambridge University Press, 1997.

[13] I. Santamaria, P. Pokharel, and J. C. Principe, “Generalized correlation function: Definition, properties, and application to blind equalization,” *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2187–2197, June 2006.

[14] J. Xu, P. P. Pokharel, A. R. C. Paiva, and J. C. Principe, “Nonlinear component analysis based on correntropy,” in *Proc. Intl. Joint Conference on Neural Networks (IJCNN'06)*, Vancouver, BC, Canada, July 2006, pp. 3517–3521.

[15] N. Aronszajn, “Theory of reproducing kernels,” *Trans. Amer. Math. Soc.*, vol. 68, no. 3, pp. 337–404, May 1950.

[16] J. Mercer, “Functions of positive and negative type, and their connection with the theory of integral equations,” *Philosophical Trans. of the Royal Society of London*, vol. 209, pp. 415–446, 1909.

[17] E. Parzen, “Statistical inference on time series by hilbert space methods,” *Technical Report 23, Statistics Department, Stanford University*, 1959.

[18] U. Grenander, *Abstract Inference*. New York: John Wiley & Sons, 1981.

[19] J. C. Principe, D. Xu, and J. W. Fisher, “Information theoretic learning,” in *Unsupervised Adaptive Filtering*, S. Haykin, Ed. John Wiley & Sons, 2000, pp. 265–319.

[20] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.

[21] K. Kotani, Y. Kinomoto, M. Yamada, J. Deguchi, M. Tonoike, K. Horii, S. Miyatake, T. Kuroiwa, and T. Noguichi, “Spatiotemporal patterns of movement-related fields in stroke patients,” *Neurol. Clin. Neurophysiol.*, vol. 63, pp. 1–4, 2004.

[22] F. D. Russo, A. Martinez, M. I. Serono, S. Pitzalis, and S. A. Hillyard, “Cortical sources of the eearly components of the visual evoked potential,” *Human Brain Mapping*, vol. 15, pp. 95–111, 2001.