Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Time Delay and Causality in Biological Systems Using Whitened Cross-Correlation Analysis

Mahmoud El-Gohary[1], James McNames[1], Tim Ellis[1], Brahm Goldstein[2]
[1]Biomedical Signal Processing Laboratory, Portland State University, OR, USA
[2]Complex Systems Laboratory, Doernbecher Children's Hospital, Oregon Health and Science University (OHSU), OR, USA

Abstract— **In the study of biological systems, it is often desirable to study the relationship between two simultaneously recorded signals and investigate whether one signal is causing the other. Correlation between signals can be revealed by spectral analysis techniques such as coherence. While coherence reveals the interaction strength between two signals, it does not provide directional information about the direction of causality of the signals, if any. Cross-correlation can be reliably used to test whether a linear association exists between two processes. It can also be used to test whether a time lag exists between the signals by identifying the mean value of their cross-correlation sequence. In this paper, we propose applying a whitening filter to signals prior to estimating the cross-correlation. This whitening removes correlation of the signals with themselves, which generally blurs the cross-correlation over a broad range of lags and limits cross-correlation as a tool for causality analysis. In this application, a Kalman filter is used adaptively to whiten the signals. An example of the increased sensitivity of whitened cross-correlation analysis is given by studying the relationship between the mean intracranial pressure (ICP) and the heart rate (HR) of pediatric patients with traumatic brain injury. Results show that in five recordings from five patients, the heart rate process lags the mean intracranial pressure.**

I. Introduction

In many applications, researchers are interested in the relationship between two different signals. Correlation between signals can be revealed by time-domain techniques such as cross-correlation or frequency-domain techniques such as coherence [2]. Once a correlation has been found, one often wishes to determine whether one process is causing the other. While coherence reveals the interaction strength between two processes, it does not provide directional information about the causal relationship between two interacting signals.

Cross-correlation between two processes can shed some light on the casual relationship between the two processes that generated the two signals. If they are strongly correlated, there are essentially three possibilities: the first process may be causing the second, the second process may be causing the first, or both processes may be caused by a third, unmeasured process. A significant cross-correlation at a lag of zero could be an indication that both processes are caused by a third process, or that the causal relationship has a delay that is too short to be estimated accurately. A statistically significant correlation between two processes with a positive lag could be interpreted as the variation in the reference signal causing the variation in the delayed one, or at least that the delayed signal has no direct causal influence on the reference signal.

A variety of algorithms were developed over the past few decades to estimate time delay. Group delay, which is the first derivative of the phase with respect to frequency, has often been used for time delay estimation. However, it has been shown that a simple causal system such as a bandpass filter can have a negative group delay in most spectral regions [9]. A negative group delay means that there is a group advance; the output waveform emerges before the input. Although this behavior is not in conflict with causality, it shows that the group delay of the system is not a reliable indicator of causality. To estimate the delay, Lindemann presented a procedure based on the Hilbert transform relation between the phase of a linear system and its log gain [6]. This method is limited by the fact that the direction of the relationship between the signals has to be known in advance. Knapp and Carter introduced the time delay estimation of two continuous signals through cross-correlation technique [4]. For discrete signals, Jacovitti and Scarano analyzed the cross-correlation method and used it to detect time delay between two signals by identifying the maximum value of their correlation function [3]. The cross-correlation function has peaks at $\pm D \pm k f_s / f_c$, where D is the time delay, k is an integer, f_s is the sampling frequency, and f_c is the center frequency. Due to the deteriorating effect of the noise on time delay detection, a false peak may be present and cause a false time delay estimate [5].

For a better time lag detection, we propose using whitening filters to whiten the signals before estimating their cross-correlation. This whitening orthogonalizes the data samples that reduces the cross-correlation between signals due to self-correlation of the signals. The practical consequence is that the cross-correlation of the whitened signals contains peaks that are less broad and more indicative of the actual delay between the signals. The process and its whitened version are linearly equivalent and can be obtained from each other through linear causal transformation. If a minimum-phase[1] system $H(z)$ generates a signal $x(n)$ by introducing dependence in the white noise input $w(n)$, its inverse system $H_I(z)$ can be used to recover the input $w(n)$ and is known as whitening filter. Because the random variables of the whitened signal are uncorrelated, each of them add new information or innovation to the process that is not linearly correlated with the preceding observations.

[1]A system satisfies the minimum-phase condition if it has a causal, stable inverse. In the z domain, minimum-phase system has no poles or zeros outside the unit circle.

A clinical scenario where this methodology could be applied is that of the directional relationship between intracranial pressure (ICP) and heart rate (HR) in patients with traumatic brain injury (TBI). It has been demonstrated that the severity of TBI affects efferent autonomic neural pathways to the sinoatrial (SA) Node in the heart [1]. To determine the direction of this effect and whether there is a causal association between intracranial pressure and HR, we apply the whitened cross-correlation proposed in this paper to investigate the causal relationship between the mean ICP and HR in five pediatric patients with TBI. We hypothesize that changes in ICP precede those of HR.

II. Methods

A. Whitening and Kalman Filter

Applying a whitening filter to an observed signal is equivalent to calculating its innovation at every time index. Likewise, the Kalman filter recursions are equivalent to finding an innovation model for the observation. The Kalman filter creates one-step predicted estimates of the state $\hat{x}_{n|n-1}$ which is used to predict the value of the observation y_n [10]. The process y_n is modeled as an autoregressive (AR) process:

$$y_{n+1} = v_n + \sum_{k=0}^{l-1} a_k y_{n-k} \tag{1}$$

where the parameters a_k are the model coefficients, l is the model order, and v_n is the noise or the prediction error when y_{n+1} is predicted from its own past. The state space model of y_n for the Kalman filter

$$y_n = H_n x_n + v_n, \quad n \geq 0, \tag{2}$$

where the state x_n obeys the recursion

$$x_{n+1} = x_n + u_n, \quad n \geq 0, \tag{3}$$

which is a random walk model for the state x_n that represents the model parameters,

$$x_n = [a_1 \; a_2 \; \ldots \; a_l]^T, \tag{4}$$

$$H_n = [y_{n-1} \; y_{n-2} \; \ldots \; y_{n-l}] \tag{5}$$

The measurement noise v_n and the process noise u_n are uncorrelated and are both assumed to be zero-mean white noise processes. It is also assumed that the initial state x_0 has zero mean, covariance matrix Π_0, and is uncorrelated with the two white noise processes that drive the state space model along with x_0. These assumptions are compactly stated as

$$\left\langle \begin{bmatrix} u_n \\ v_n \\ x_0 \end{bmatrix}, \begin{bmatrix} u_n \\ v_n \\ x_0 \\ 1 \end{bmatrix} \right\rangle = \begin{bmatrix} Q_n\delta_{nk} & 0 & 0 & 0 \\ 0 & R_n\delta_{nk} & 0 & 0 \\ 0 & 0 & \Pi_0 & 0 \end{bmatrix}$$

The user must specify Q_n, R_n, and Π_0. The innovations are calculated as the difference between the observed signal and the predicted value, based on a linear combination of the previous observed values,

$$e_n \quad = \quad y_n - \hat{y}_{n|n-1} \tag{6}$$

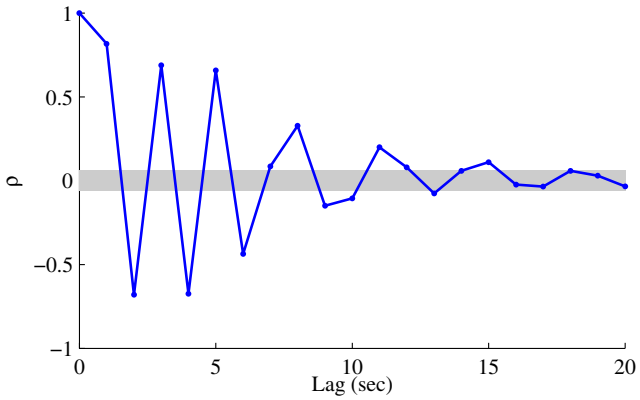

Fig. 1. The partial autocorrelation of the mean ICP with $f_s = 1$ Hz. The gray region represents the confidence bounds which enclose only those lags for which the samples are not significantly correlated. The samples are not significantly correlated after a lag of 15 seconds.

where $\hat{y}_{n|n-1}$ is the linear minimum mean square error (MMSE) estimator of y_n given $(y_0, y_1, \ldots, y_{n-1})$, and $\hat{x}_{n|n-1}$ is the one-step predictions of the state. Therefore, finding the innovations e_n reduces to one-step prediction of the state and can be recursively estimated as follows:

$$e_n = y_n - H_n \hat{x}_{n|n-1} \tag{7}$$

$$R_{e,n} = H_n P_{n|n-1} H_n^* + R_n \tag{8}$$

$$K_{p,n} = (P_{n|n-1} H_n^*) R_{e,n}^{-1} \tag{9}$$

$$\hat{x}_{n+1|n} = \hat{x}_{n|n-1} + K_{p,n} e_n \tag{10}$$

$$P_{n+1|n} = P_{n|n-1} + Q_n - K_{p,n} R_{e,n} K_{p,n}^* \tag{11}$$

$R_{e,n}$ is the innovations covariance, $K_{p,n}$ is the predicted Kalman gain, and P_n is the state error covariance. The recursion is initialized with $\hat{x}_{0|-1} = 0$ and $P_{0|-1} = \Pi_0$.

B. Determining the order of the AR process

One aid to determining the order of an AR process is the partial autocorrelation (PAC) function which is defined as follows. When fitting an AR(p) model, the last coefficient measures the excess correlation at lag p which is not accounted for by an AR($p-1$) model. The PAC function is estimated by fitting AR process of successive higher order. The PAC coefficients are plotted against p, along with the 95% estimated confidence intervals for a white Gaussian noise process with the same variance as the sample variance of the signal. PAC coefficients values that are outside the confidence intervals are significantly different from zero. Fig. 1 shows that PAC of the mean intracranial pressure process ($f_s = 1$ Hz) cuts off after a lag of 15 seconds. A good value for the model order is the smallest lag such that the value of $\rho(l)$ beyond which the coefficient values are not significantly different from zero.

C. Autocorrelation and Cross-Correlation

Autocorrelation coefficients provide an important guide to the properties of a random process. For a white noise, the autocorrelation function (ACF) is given by

$$\rho(l) = \begin{cases} 1 & l = 0 \\ 0 & l = \pm 1,\ \pm 2,\ \ldots \end{cases}$$

Let $x(n)$ and $y(n)$ be two simultaneously recorded data sets of length N. For the process $x(n)$, the biased estimator of the autocorrelation $r_x(l)$ is given by the autocorrelation sequence

$$\hat{r}_x(l) = \frac{1}{N} \sum_{n=0}^{N-l-1} x(n+l)x^*(n) \quad 0 \le l \le N-1 \tag{12}$$

and the ACF of $x(n)$ is given by

$$\hat{\rho}(l) = \frac{\hat{r}_x(l)}{\hat{r}_x(0)} \tag{13}$$

The biased estimate of the cross-correlation between $x(n)$ and $y(n)$ is:

$$\hat{r}_{xy}(l) = \begin{cases} \frac{1}{N} \sum_{n=0}^{N-l-1} x(n+l)y^*(n) & 0 \le l \le N-1 \\ \frac{1}{N} \sum_{n=0}^{N+l-1} x(n)y^*(n-l) & -(N-1) \le l \le -1 \end{cases} \tag{14}$$

D. Mean Delay Definition

Many researchers use the lag at which the cross-correlation between two signals has a maximum value as an estimate of the putative pure time delay. However, in many situations, the cross-correlation sequence has more than one maximum that are not indicative of the actual delay and the relationship between two signals does not consist of a pure delay. In many practical applications, however, the relationship between two stochastic signals is not accurately modeled as a pure delay and the CCF may be significant over a range of lags. For these cases we define the mean delay as the mean of the CCF's energy density function,

$$\hat{\mu}_d = \frac{\sum_{n=-l}^{l} n\hat{\rho}_{xy}^2(n)}{\sum_{n=-l}^{l} \hat{\rho}_{xy}^2(n)} \tag{15}$$

This can be regarded as the center of gravity of the power of the correlation function. Similar definitions are used to determine the mean time of a signal in discussions of the time-frequency uncertainty principle [7].

E. Application to Mean Intracranial Pressure and Heart rate

The technique described above was applied to investigate the relationship between the mean intracranial pressure (ICP) and the heart rate (HR). We obtained the mean ICP and HR signals used in this analysis from two-hour records of the ICP (f_s = 125 Hz) and ECG (f_s = 500 Hz) of five sedated pediatric trauma brain injury (TBI) patients in the ICU at Oregon Health and Science University (OHSU). We computed the mean ICP using a zero-phase (noncausal) elliptic lowpass filter with cutoff frequency $f_c = 0.3$ Hz as shown in Fig. 2. An automatic detection algorithm was used to detect the R wave of the ECG signals. The instantaneous heart rate was computed as the inverse of the interbeat interval. This yields a non-uniformly sampled series, however, which was interpolated it to uniform sampling rate of f_s = 125 Hz using a Gaussian kernel smoother [8]. In addition, applying kernel smoothing reduces quantization error caused by the sample rate of the ECG, as shown in Fig. 3. The signals were lowpass filtered and decimated to 1 Hz before applying the Kalman filter.

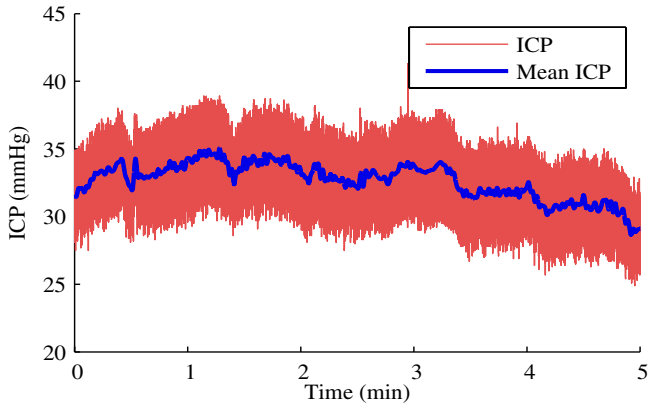

Fig. 2. Five minute segment of ICP and lowpass filtered mean ICP.

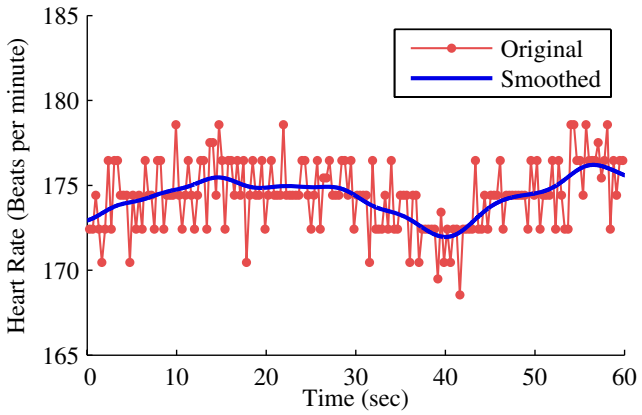

Fig. 3. Sixty second segment of instantaneous heart rate and smoothed estimate. The heart rate, though high, is typical for pediatric TBI patients.

The state space model of the Kalman filter has several parameters that need to be specified before applying the filter. Using the results from partial autocorrelation estimation, we used a Kalman filter order $l = 20$. The initial state variance Π_0, which controls how quickly the model converges from the initial parameter values, was set equal to the identity matrix, $\Pi_0 = I$. For the mean ICP signals, we used $Q = 0.02I$ and $R = 2I$. For the HR signals, $Q = 0.0001I$ and $R = 0.001I$.

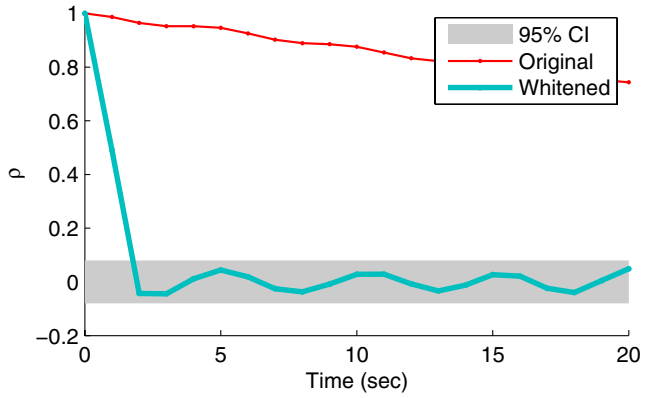

Fig. 4. The autocorrelation of the mean ICP. Samples beyond a lag of 1 second are not significantly correlated for the whitened mean ICP.

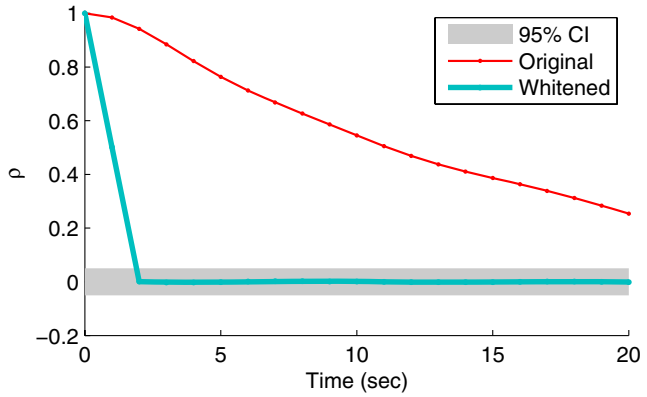

Fig. 5. The autocorrelation of the HR. Samples beyond a lag of 1 second are not significantly correlated for the whitened HR (innovations).

III. Results and Discussion

The autocorrelation estimate of the mean ICP, represented as a dotted line in Fig. 4, shows that the signal has a long memory; the signal has a high degree of correlation among its samples. The autocorrelation estimate of the whitened mean ICP, on the other hand shows a maximum lag at zero then decreases rapidly to zero or to non-significant values which are a characteristic of a white noise. The same result was obtained for the HR as shown in Fig. 5. There was a statistically significant cross-correlation, $\rho > 0.3$ ($p < 0.05$), which indicates that there exists a linear relationship between the mean ICP and HR. The cross-correlation estimate between the signals before whitening showed a negative mean delay for two patients, and positive mean delay for three patients. In contrast, the whitened cross-correlation had a positive mean delay for all five patients. The average time delay across the five patients was 0.625 seconds. These results and Fig. 6 show that the general cross-correlation method could be deceptive while whitening provides a better characterization of the time delay. These results also suggest that the heart rate fluctuations lag, and are possibly caused by, fluctuations in the mean ICP.

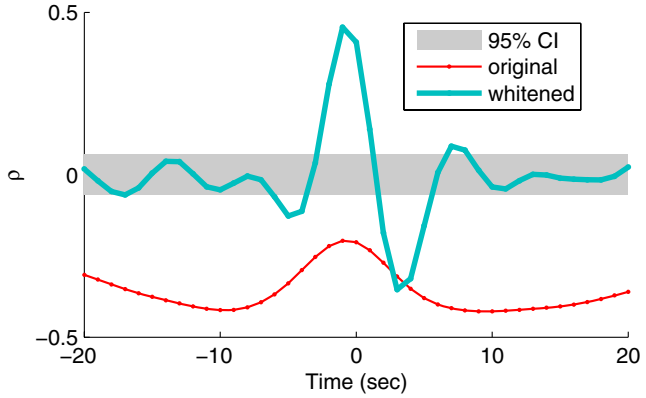

Fig. 6. The cross-correlation between the original observations (mean ICP-HR), and between the innovations with a mean delay of 0.63 seconds. The heart rate lags the mean intracranial pressure.

IV. Conclusion

In this paper, we introduced a new methodology; a whitened cross-correlation analysis, to study the relationship between two simultaneously recorded signals. This whitening provides a better characterization of the time delay by removing the redundant self-correlation among the components of the processes. This technique was used to study the relationship between the mean intracranial pressure and the heart rate. Results suggest that the heart rate fluctuations lag, and are possibly caused by, the fluctuations in the mean ICP. Further studies are necessary to determine the physiologic mechanism by which the heart rate changes depending on the level of the intracranial pressure.

References

[1] Goldstein B, Toweill D, Lai S, Sonnenthal K, and Kimberly B. Uncoupling of the autonomic and cardiovascular systems in acute brain injury. *Am. J. Physiol.*, 275 (4 Pt 2):R1287–R1292, 1998.

[2] G.C. Carter. Coherence and time delay estimation. *Proceedings of the IEEE*, 75:236–255, 1987.

[3] Giovanni Jacovitti and Gaetano Scarano. Discrete time techniques for time delay estimation. *IEEE Transactions on Signal Processing*, 41(2):525–533, February 1993.

[4] C. H. Knapp and G.C. Carter. The generalized correlation method for estimation of time delay. *IEEE Trans. Accoust., Speech, Signal Processing*, 24:320–327, 1976.

[5] Xiaoming Lai and Hans Torp. Interpolation methods for time-delay estimation using cross-correlation method for blood velocity measurement. *IEEE Transactions on Ultrasonic, And Frequency Control*, 46:277–290, 1999.

[6] M. Lindemann, J. Raethjen, J. Timmer, G. Deushl, and G. Pfister. Delay estimation for cortico-peripheral relations. *Journal of Neuroscience Methods*, 111:127–139, 2001.

[7] Patrick J. Loughlin, James W. Pitton, and Les E. Atlas. Proper time-frequency energy distributions and the heisenberg uncertainty principle. *Proceedings of the IEEE-SP International Symposium*, pages 151–154, 1992.

[8] J. McNames, T. Thong, and M. Aboy. Impulse rejection filter for artifact removal in spectral analysis of biomedical signals. *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pages 145–148, 1-5 September 2004.

[9] Morgan W. Mitchell and Raymond Y. Chiao. Causality and negative group delays in a simple bandpass amplifier, March 1999.

[10] A. H. Sayed T. Kailath and B. Hassibi. *Linear Estimation*. 2000.