Communications

Adaptive Modeling and Spectral Estimation of Nonstationary Biomedical Signals Based on Kalman Filtering

Mateo Aboy*, Oscar W. Márquez, James McNames, Roberto Hornero, Tran Trong, and Brahm Goldstein

***Abstract*—We describe an algorithm to estimate the instantaneous power spectral density (PSD) of nonstationary signals. The algorithm is based on a dual Kalman filter that adaptively generates an estimate of the autoregressive model parameters at each time instant. The algorithm exhibits superior PSD tracking performance in nonstationary signals than classical nonparametric methodologies, and does not assume local stationarity of the data. Furthermore, it provides better time-frequency resolution, and is robust to model mismatches. We demonstrate its usefulness by a sample application involving PSD estimation of intracranial pressure signals (ICP) from patients with traumatic brain injury (TBI).**

***Index Terms*—Intracranial pressure, Kalman filter, linear models, spectral estimation, traumatic brain injury.**

I. Introduction

Currently, power spectral density (PSD) estimation of physiologic signals is performed predominantly using classical techniques based on the fast Fourier transform (FFT). Nonparametric methods such as the periodogram and its improvements (i.e., Barlett's, Welch's, and Blackman-Tukey's methodologies [1]–[4]) are based on the idea of estimating the autocorrelation sequence of a random process from measured data, and then taking the FFT to obtain an estimate of the power spectrum. The main two advantages of these techniques are their computational efficiency, due to the numerical efficiency of the FFT algorithm, and that they do not make any assumptions about the process except for its stationarity. This makes them the methodology of choice, particularly in situations where long data records need to be analyzed and there is no clear model for the process. Furthermore, the availability of long data records enables one to improve their statistical properties by averaging or smoothing. However, these techniques have some limitations. They require stationarity of the segments studied, do not work well for short data records, and have limited frequency resolution. Since physiologic signals are nonstationary in nature, these techniques are applied following the methodology of the short-time Fourier transform (STFT), where nonparametric methods are applied to short overlapping segments which are assumed to be stationary. This approach has also its limitations. It imposes a piecewise stationary model on the data and, since local stationarity requires the analysis segments to be short in duration, they have limited time-frequency resolution.

Time-frequency resolution can be improved by using parametric methods of PSD estimation. The parametric approach is based on modeling the signal under analysis as a realization of a particular stochastic process and estimating the model parameters from its samples. In the absence of *a priori* knowledge about how the process is generated, parametric PSD is generally performed assuming an autoregressive (AR) model [4]. This is a popular assumption for several reasons: 1) many natural signals such as speech, music or seismic signals have an underlying autoregressive structure; 2) in general, any signal—not necessarily AR in nature—can be modeled as an AR process if a sufficiently large model order is selected; 3) the all-pole structure of AR enables for good spectral peak matching, which makes it a good model candidate for situations where we are more interested in the spectral peaks than valleys; and 4) estimation of the model parameters involves the solution of a linear system of equations, which can be solved efficiently. Even though parametric PSD can improve the frequency resolution, the current techniques for PSD estimation based on AR models (i.e., autocorrelation, covariance, modified convariance, and Burg's methods [5], [6]) assume stationarity. To analyze nonstationary signals they must also assume the signal is locally piecewise stationary.

We describe a methodology to estimate the time-varying AR model parameters of nonstationary signals using an adaptive Kalman filter. This methodology produces instantaneous estimates of PSD, improved time-frequency resolution, and enables for nonstationary spectral analysis in situations where data records are too short and the local stationary model does not work well. The reliability of the algorithm was tested with synthetic data generated from different models (AR, MA, ARMA, and harmonic), and with real data from physiologic pressure signals. Following the description of this methodology, we demonstrate its usefulness by a sample application involving PSD estimation of intracranial pressure signals (ICP) from patients with traumatic brain injury (TBI).

II. Methods

The adaptive Kalman filter algorithm we propose for instantaneous PSD estimation assumes an underlying autoregressive structure of the data. We chose an underlying AR model structure because of its intrinsic generality and peak matching capabilities. These are important properties for the analysis of physiologic signals, since we are usually more interested in estimating the frequency at which the formant frequencies (peaks) occur than the valleys. Starting from this assumption, we modeled a given physiologic signal with a recursion of the form

$$x(n) = \sum_{k=1}^{P} a_k x(n-k) + w(n) \tag{1}$$

where $x(n)$ is the physiologic signal under analysis at instant n, $\{a_k\}_{k=1}^{p}$ are the model parameters, $\{x(n-k)\}_{k=1}^{p}$ are delayed

Manuscript received on March 1, 2004; revised January 2, 2005. A previous version of this paper was presented at the IEEE-EMBS 2004 conference. This work was supported in part by the Northwest Health Foundation, in part by the Doernbecher Children's Hospital Foundation, and in part by the Thrasher Research Fund. *Asterisk indicates corresponding author.*

*M. Aboy is with the Department of Electronics Engineering Technology, Oregon Institute of Technology, Portland, OR 97201 USA. He is also with the Biomedical Signal Processing Laboratory, Department of Electrical and Computer Engineering, Portland State University, Portland, OR 97201 USA (e-mail: mateoaboy@ieee.org).

O. W. Márquez is with the Signal Theory and Communications Department, ETSI-Telecomunicación, University of Vigo, 36310 Vigo, Spain, EU.

J. McNames is with the Biomedical Signal Processing Laboratory, Department of Electrical and Computer Engineering, Portland State University, Portland, OR 97201 USA.

R. Hornero is with the Department of Signal Theory and Communications, ETSI-Telecomunicación, University of Valladolid, 47011Valladolid, Spain, EU.

T. Trong is with the Department of Biomedical Engineering, OGI School of Science and Engineering, Oregon Health and Science University, Portland, OR 97206 USA.

B. Goldstein is with the Complex Systems Laboratory, Department of Pediatrics, Oregon Health and Science University, Portland, OR 97201 USA.

Digital Object Identifier 10.1109/TBME.2005.851465

samples of the signal, and $w(n)$ is assumed to be a random sequence independent and normally distributed with zero mean. Equation (1) can be generalized by allowing the model coefficients $\{a_k\}_{k=1}^{p}$ to be time-variant. The estimation problem within the context of nonstationary processes yields naturally to the discrete Kalman filter (DKF) [7]–[9].

In order to use the DKF, we must have a signal model in state-space form, and the state of the system evolves as a first-order difference equation, and must be estimated from noisy observations. The general form of the state-space model for the linear DKF is given by [8], [10]

$$\begin{aligned}\mathbf{x}(n) &= \mathbf{A}(n-1)\mathbf{x}(n-1)+\mathbf{w}(n)\\ \mathbf{y}(n) &= \mathbf{H}(n)\mathbf{x}(n)+\mathbf{v}(n)\end{aligned} \tag{2}$$

where $\mathbf{x}(n)$ is the state of the system, $\mathbf{A}(n-1)$ is the transition or system matrix, $\mathbf{H}(n)$ is the observation matrix, $\mathbf{y}(n)$ is the vector of observations, and $\mathbf{w}(n)$ and $\mathbf{v}(n)$ are zero-mean white Gaussian noise processes representing system and observation noise, respectively. The system noise and the process noise are assumed to be independent. If the problem can be formulated in state-space according to (2), and if we know $\mathbf{A}(n-1)$, $\mathbf{H}(n)$, and the covariance matrix of $\mathbf{w}(n)$ and $\mathbf{v}(n)$, then we can use the DKF to estimate the state of the system optimally according to the Kalman recursion.

A. Signal Model in State-Space

Since our signal model (1) is a pth-order difference equation, we can transform it to a system of difference equations by defining the state of the system as a p-dimensional vector

$$\mathbf{x}(n)=\begin{pmatrix} x(n)\\ x(n-1)\\ \vdots \\ x(n-p+1)\end{pmatrix}. \tag{3}$$

This enables us to rewrite (1) as a first-order difference equation with time-varying model parameters, and enables us to create a state-space model for the DKF

$$\mathbf{A}(n)=\begin{pmatrix} a_1(n) & a_2(n) & \dots & a_p(n)\\ 1 & 0 & \dots & 0\\ 0 & 1 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \dots & 0\end{pmatrix} \tag{4}$$

$$\begin{aligned}\mathbf{x}(n) &= \mathbf{A}(n-1)\mathbf{x}(n-1)+\mathbf{w}(n)\\ y(n) &= \mathbf{H}\mathbf{x}(n)+\mathbf{v}(n).\end{aligned} \tag{5}$$

The measurement matrix is $\mathbf{H}=(1\ 0\ \dots\ 0)$. However, in order for this state-space model to be useful we need a way to estimate the vector of time-varying coefficients $\mathbf{a}(n)=(a_1(n)\ a_2(n)\ \dots\ a_p(n))^T$ corresponding to the first row of the transition matrix at time n, $\mathbf{A}(n)$.

B. Dual Kalman Filter

The vector of time-varying coefficients that made the first row of the transition matrix $\mathbf{a}(n)$ can also be estimated using a DKF. This is referred as a dual Kalman filter, that is, two DKFs working in parallel to estimate the model parameters and the state of the system [11]–[13].

The estimation of the model parameters $\mathbf{a}(n)$ can be formulated in state space as follows:

$$\begin{aligned}\mathbf{a}(n) &= \mathbf{\Phi}\mathbf{a}(n-1)+\mathbf{e}(n)\\ x(n+1) &= \mathbf{x}^T(n)\mathbf{a}(n)+\mathbf{q}(n).\end{aligned} \tag{6}$$

where $\mathbf{\Phi}$ is a user specified diagonal matrix with entries $(\rho_{ij})_{i=j}$ corresponding to the correlation between $\mathbf{a}(n)$ and $\mathbf{a}(n-1)$, which control the adaptation speed and frequency tracking capabilities of the algorithm. For biomedical signals, where the model parameters change slowly, values close to 1 work well (e.g., 0.995). In the case of $(\rho_{ij})_{i=j}=1$, the system equation becomes $\mathbf{a}(n)=\mathbf{a}(n-1)+\mathbf{e}(n)$. This is a simple Markov process where the vector coefficients evolve following a random walk. The adaptation speed is controlled by the covariance of $\mathbf{e}(n)$. There is a tradeoff between high adaptation speed (fast tracking) and variance of the estimates. The measurement equation of the model, $x(n+1)=\mathbf{x}^T(n)\mathbf{a}(n)+\mathbf{q}(n)$ implements a linear predictor, where the signal at time $n+1$ is estimated from previous values of n according to

$$\begin{aligned}x(n+1) &= \mathbf{x}^T(n)\mathbf{a}(n)+\mathbf{q}(n)\\ &= \sum_{k=1}^{p} a_k(n)x(n-k+1)+\mathbf{q}(n).\end{aligned} \tag{7}$$

In the state-space formulation given by (6), the model parameters $\mathbf{a}(n)$ become state variables. The optimum linear estimate of state of the system $\mathbf{a}(n)$ can be estimated recursively according to

$$\hat{\mathbf{a}}(n|n-1) = \mathbf{\Phi}\hat{\mathbf{a}}(n-1|n-1) \tag{8}$$

$$z(n) = x(n+1) \tag{9}$$

$$\hat{z}(n) = \mathbf{x}^T(n)\hat{\mathbf{a}}(n|n-1) \tag{10}$$

$$\hat{\mathbf{a}}(n|n) = \hat{\mathbf{a}}(n|n-1)+\mathbf{K}(n)\left[z(n)-\hat{z}(n)\right] \tag{11}$$

$$\mathbf{P}(n|n-1) = \mathbf{\Phi}\mathbf{P}(n-1|n-1)\mathbf{\Phi}^T+\mathbf{Q}_e(n) \tag{12}$$

$$\mathbf{\Xi}(n) = \mathbf{x}(n)^T\mathbf{P}(n|n-1)\mathbf{x}(n)^T \tag{13}$$

$$\mathbf{K}(n) = \mathbf{P}(n|n-1)\mathbf{x}(n)\left[\mathbf{\Xi}(n)+\mathbf{Q}_l(n)\right]^{-1} \tag{14}$$

$$\mathbf{P}(n|n) = \left[\mathbf{I}-\mathbf{K}(n)\mathbf{x}^T(n)\right]\mathbf{P}(n|n-1)\Big] \tag{15}$$

where $\hat{\mathbf{a}}(n|n-1)=\mathbf{\Phi}\hat{\mathbf{a}}(n-1|n-1)$ is the best estimate of the state (i.e., AR model parameters) without incorporating the observation at time n, just based on the model structure we imposed on the evolution of $\mathbf{a}(n)$ (prediction), and $\hat{z}(n)=\mathbf{x}^T(n)\hat{\mathbf{a}}(n|n-1)$ is the best estimate of the the measurement $z(n)=x(n+1)$ based on the model. The optimum estimate of the state at time n incorporating the measurement at time n is given by $\hat{\mathbf{a}}(n|n)=\hat{\mathbf{a}}(n|n-1)+\mathbf{K}(n)[z(n)-\hat{z}(n)]$, which is composed of two terms: the best estimate of the state without the measurement at time n, and a weighted difference of the observation at time n and the best estimate of this observation (correction). The weighting factor $\mathbf{K}(n)$ is calculated optimally following the Kalman recursion, and is referred to as the Kalman gain [7], [8], [10].

C. Instantanous PSD Estimation

The theoretical power spectrum of pth-order stationary autoregressive process is given by

$$P_x(e^{jw})=\frac{|b(0)|^2}{\left|1+\sum_{k=1}^{p}a_k e^{-jkw}\right|^2}. \tag{16}$$

If $b(0)$ and $\{a\}_{k=1}^{p}$ can be estimated from data, then we can form an estimate of the power spectrum of a stationary process as

$$\hat{P}_x(e^{jw})=\frac{\left|\hat{b}(0)\right|^2}{\left|1+\sum_{k=1}^{p}\hat{a}_k e^{-jkw}\right|^2}. \tag{17}$$

Since the dual Kalman filter we propose provides estimates of $\{a\}_{k=1}^{p}$ at each time instant, the nonstationary power spectrum is given by

$$\hat{P}_x(e^{jw},n)=\frac{\left|\hat{b}(0,n)\right|^2}{\left|1+\sum_{k=1}^{p}\hat{a}_k(n) e^{-jkw}\right|^2}. \tag{18}$$

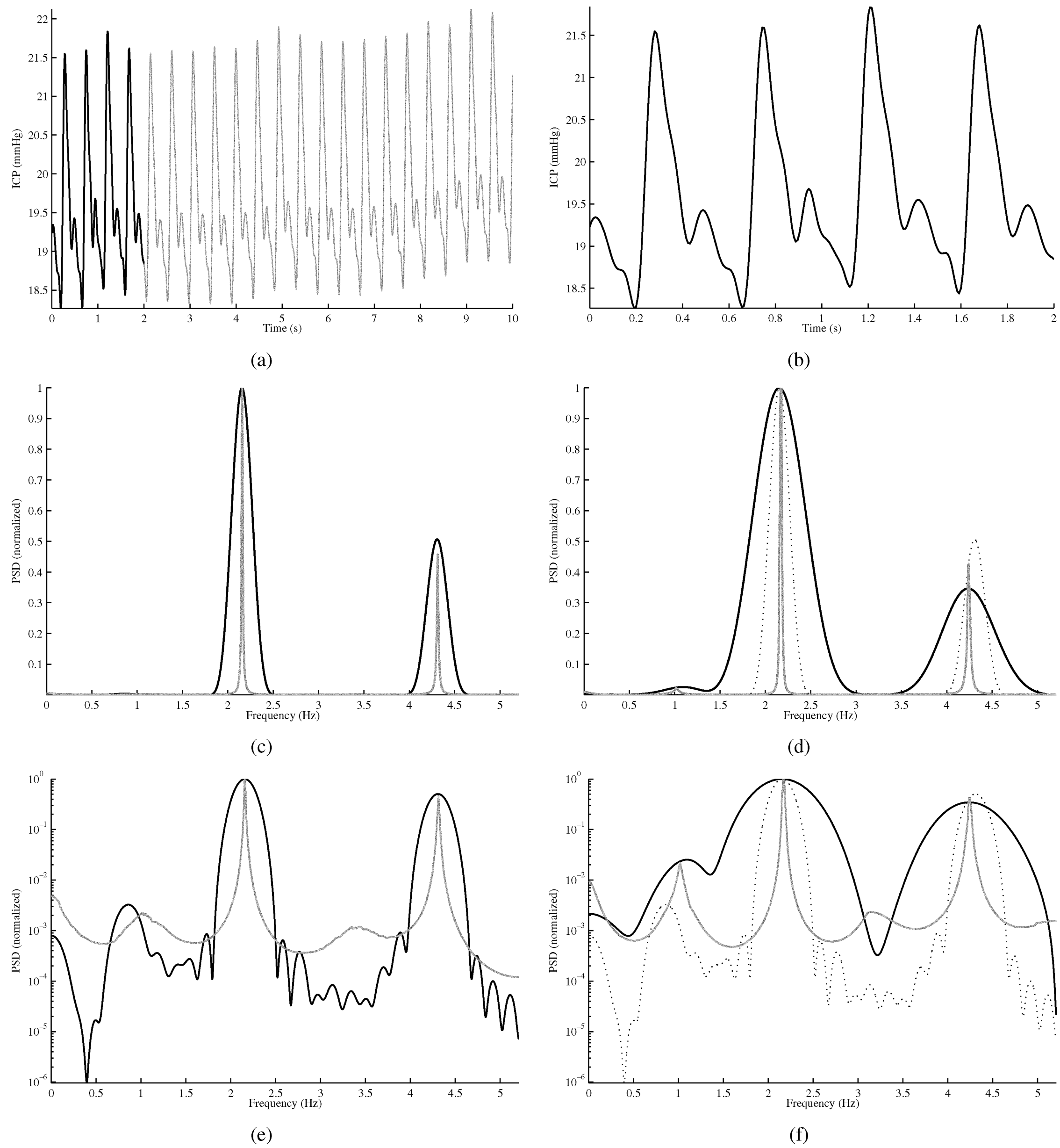

Fig. 1. Representative results of the first comparative study between a nonparametric methodology (Welch's) and the proposed DKF PSD estimator. (a) Plot shows an example of the 10-s ICP segment (light grey) and the 2-s subsegment used for this simulation. (b) Plot of the 2-s subsegment highlighted in (a). (c) Welch's PSD (dark) and DKF PSD (light) estimates corresponding to the 10-s segment. (d) Welch's PSD (dark) and DKF PSD (light) estimates on the 2-s subsegment. The thin line corresponds to Welch's estimate based on the 10-s segment. (e) Welch's PSD (dark) and DKF PSD (light) estimates in the 10-s segment with y-axis in dB scale. (f) Welch's PSD (dark) and DKF PSD (light) estimates in the 2-s subsegment with y-axis in dB scale. The dotted line is the Welch estimate based on the 10 s.

Therefore, the nonstationary PSD given the instantaneous estimates of model parameters $\mathbf{a}(n)$ can be computed according to

$$\hat{P}_x(e^{jw}, n)_{KM} = \frac{1}{|\mathrm{FFT}\,[\bar{\mathbf{a}}(\mathbf{n})]|^2} \tag{19}$$

$$\bar{\mathbf{a}}(\mathbf{n}) = \left(1 - \mathbf{a}(n)^T\right). \tag{20}$$

III. Results

We tested the reliability of the instantanous PSD estimation algorithm with synthetic data generated from different models (AR, MA, ARMA, and harmonic), and with real data from physiologic pressure signals. In the following, we demonstrate its usefulness by a sample application involving PSD estimation of ICPs from patients with traumatic brain injury (TBI).

A. Subjects and Material

This study included ICP signals from patients with significant head injuries who were admitted to the pediatric intensive care unit at Doernbecher Children's Hospital. ICP was monitored continuously using a ventricular catheter or parenchymal fiber-optic pressure transducer (Integra NeuroCare, Integra LifeSciences, Plainsboro, NJ). The ICP monitor was connected to a Philips Merlin CMS patient monitor (Philips, Best, Netherlands) that sampled the ICP signals at 125 Hz. An HPUX workstation automatically acquired these signals through a serial data network and they were stored in files on CD-ROM. Patients were managed according to the standards of care in pediatric intensive care unit at Doernbecher Children's Hospital. The data acquisition protocol was reviewed and approved by the Institutional Review Board at Oregon Health and Science University, and the requirement on informed consent was waived.

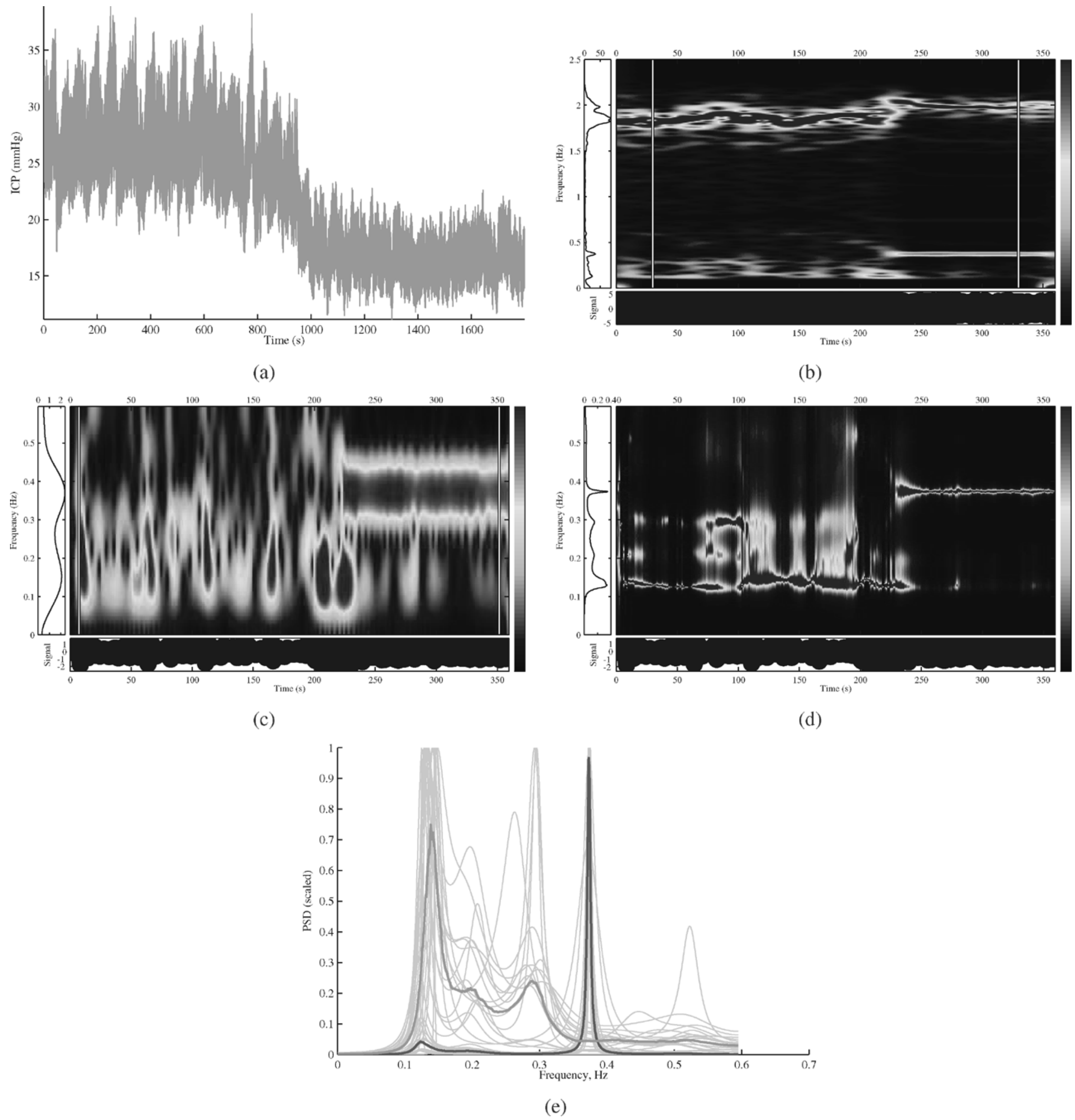

Fig. 2. Representative results of the second comparative study between a nonparametric methodology (Welch's) and the proposed PSD estimator based on the DKF. (a) ICP segment during a period of hypertension $(\mathrm{ICP} > 25\ \mathrm{mmHg})$ and the reduction in mean ICP after mechanical hyperventilation (approximately 800 s). (b) Spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation). In the spectrogram, we can clearly see the cardiac components around 2 Hz and the respiratory component (0.1–0.55) Hz. In the respiratory component, we can note a period of spontaneous breathing (approximately 0–225 s), and the period of mechanical hyperventilation (approximately >225 s). (c) Spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation) generated using the nonparametric PSD estimator with a window of 15 s. (d) Spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation) generated using the instantaneous PSD estimate based on the DKF. The time resolution is of 1 sample. We can appreciate a much better frequency resolution in this case. (e) PSD plot showing the instantaneous PSD estimates (thin light lines) before and after the intervention and their average (thick lines). The average before the change is shown in grey and the average after the change is the black thick line.

B. Comparative Studies

We compared PSD estimates obtained with the proposed Kalman PSD estimation algorithm with those generated by classical nonparametric estimation techniques. For the purposes of this paper, Welch's method of nonparametric PSD estimation was used as the methodology representing the nonparametric methods.

The first comparison was aimed at determining the quality of the PSD estimates of a nearly stationary ICP signal. The PSD of the signal was estimated with Welch's PSD estimator and with the proposed Kalman PSD estimator. PSDs of locally stationary 10-s segments obtained from ICP signals were estimated using both methodologies. Then, 2-s subsegments from these 10-s segments were selected and both methodologies were applied to estimate the PSD corresponding to the 2-s subsegments. The objective of this study was to compare the accuracy of the PSD estimates produced by both methodologies in the 2-s subsegments as estimates of the PSDs estimated in the 10-s segments.

The second study was aimed at comparing the time-frequency resolution of the spectrograms generated using both methodologies. For this purpose, a set of nonstationary ICP signals were used. We selected ICP segments from patients who were undergoing a period of intracranial hypertension (nonstationary conditions), and to whom a therapy of mechanical hyperventilation was applied as a first therapy to reduce elevated ICP. We compared both methodologies at the task of determining the time instant at which the hyperventilation intervention started and determining the change in respiratory frequency applied based exclusively on the ICP spectrogram.

In Fig. 1, we show representative results of the first comparative study between a nonparametric methodology (Welch's) and the proposed PSD estimator based on the DKF. For the purposes of this study, Welch's method was always used with the maximum possible window length, since this yields to the best frequency resolution. Fig. 1(a) shows the 10-s ICP segment (light grey) and the 2-s subsegment subsegment used in this simulation. Fig. 1(b) show a plot of the 2-s subsegment highlighted in Fig. 1(a). In Fig. 1(c), we show the Welch's PSD (dark) and DKF PSD (light) estimates in the 10-s segment, and in Fig. 1(d) we show the results of both methodologies on the 2-s segment. The dotted line corresponds the Welch estimate based on the 10-s segment. Fig. 1(e) and Fig. 1(f) show Welch's PSD (dark) and DKF PSD (light) estimates in dB scale. In Fig. 1(f), the dotted line corresponds to the Welch estimate based on the 10 s. From this results, we can conclude that the PSD estimate generated by our proposed algorithm has better frequency resolution. We also note that the estimate based on the 2-s segment using the Kalman PSD algorithm has even better frequency resolution than Welch's PSD estimate based on 10 s. Observing the plots in dB scale, we can also see how the Kalman PSD estimate does not have sidelobes caused by windowing effects and is smoother.

Results from the second simulation are shown in Fig. 2. In this figure, we show representative results of the second comparative study between a nonparametric PSD estimator and the proposed PSD estimator based on the DKF, which consisted in comparing the two methods time-frequency resolution for nonstationary signals. Fig. 2(a) shows the ICP segment during a period of hypertension $(\mathrm{ICP} > 25\ \mathrm{mmHg})$ and the reduction in mean ICP after mechanical hyperventilation (approximately 800 s). In Fig. 2(b), we show the spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation) using a 40-s window. Examining this spectrogram we can see the cardiac component around 2 Hz and the respiratory component in the range of 0.1–0.55 Hz. In the evolution of the respiratory component, we can note a period of spontaneous breathing (approximately 0–225 s) and a period of mechanical hyperventilation (>225 s). Fig. 2(c) shows the spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation) generated using nonparametric PSD estimation with a window of 15 s with 50% overlap, which enables us to know the time instant of the therapeutic intervention with a time resolution of 15 s. In Fig. 2(d), we show spectrogram of the ICP signal centered around the time of therapeutic intervention (hyperventilation) generated using the instantaneous PSD estimate based on the DKF. In this case, the time resolution is of 1 sample $1/f_s$ s, where f_s is the sampling frequency. Note the higher time and frequency resolution of the Kalman PSD estimate. This enables us to know the precise instant at which the therapeutic intervention occurred and to determine how much the respiratory rate was changed. The change in respiratory rate can be calculated from the plot in Fig. 2(e), which shows the instantaneous PSD estimates (thin light lines) before and after the intervention. The average before the change (grey thick line), and the average after the change (black thick line).

Based only on this simulation study, we cannot claim the DKF is universally better on all nonstationary signals than all other existing PSD estimation methods. Further studies are needed to determine in which situations the DKF PSD estimator performs better than the nonparametric techniques, and what are its limitations. As it was pointed out in the introduction, due to the computational efficiency of the nonparametric PSD estimators, these techniques are well suited for situations where we need to analyze long data records. In this case, especially if we do not need very precise time-resolution, we can increase the window length to improve the frequency resolution of the nonparametric PSD estimates, as long as we do not violate the stationarity assumption. However, in situations where the signals are nonstationary, short, and we need good time-frequency resolution, the instantaneous DKF technique we proposed here may be useful.

IV. Conclusion

The authors described an algorithm to perform instantaneous AR modeling and spectral estimation in nonstationary signals using dual Kalman filters, and demonstrated its potential applicability and usefulness by means of two comparative studies, one on simulated signals and another involving PSD estimation in ICP signals from patients with TBI. The proposed algorithm was compared with Welch's method of PSD estimation. Similar results were obtained when the DKF PSD estimator was compared against other standard nonparametric methods such as the Blackman-Tukey or the modified periodogram. Based on this preliminary study, we conclude that the DKF estimator is able to track changes in the PSD better than a moving window technique, and exhibits good time-frequency resolution when compared with benchmark nonparametric PSD techniques at the task of estimating the PSD of very short data records which are nonstationary. Furthermore, the proposed method does not assume a piecewise stationary model on the data.

References

[1] M. Barlett, "Smoothing periodograms from time series with continuous spectra," *Nature (London)*, vol. 161, no. 8, pp. 686–687, 1948.

[2] P. Welch, "The use of fast fourier transform for estimation of power spectra: a method based on time averaging over short modified periodograms," *IEEE Trans. Audio Electoacust.*, vol. 12, pp. 70–73, 1967.

[3] R. Blackman and J. Tukey, *The Measurement of Power Spectra*. New York: Dover, 1958.

[4] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing. Spectral Estimation, Signal Modeling, Adaptive Filtering and Array Processing*. New York: McGraw-Hill, 2000.

[5] J. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, 1975.

[6] S. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.

[7] R. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME—J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[8] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*. New York: Wiley, 2001.

[9] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

[10] A. Gelb, J. Kasper, R. Nash, C. Price, and A. Sutherland, *Applied Optimal Estimation*. Cambridge, MA: MIT Press, 1974.

[11] E. A. Wan and A. T. Nelson, "Neural dual extended Kalman filtering: applications in speech enhancement and monaural blind signal separation," in *Proc. 1997 Neural Networks for Signal Processing VII*, 1997, pp. 466–475.

[12] ——, "Removal of noise from speech using the dual EKF algorithm," in *Proc. 1998 IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 1998*, vol. 1, 1998, pp. 381–384.

[13] A. T. Nelson and E. A. Wan, "A two-observation Kalman framework for maximum-likelihood modeling of noisy time series," in *Proc. IEEE Int. Joint Conf. Neural Networks*, vol. 3, 1998, pp. 2489–2494.