A Wavelet-Based Method for Action Potential Detection From Extracellular Neural Signal Recording With Low Signal-to-Noise Ratio

Kyung Hwan Kim*, *Member, IEEE,* and Sung June Kim, *Member, IEEE*

Abstract—We present a method for the detection of action potentials, an essential first step in the analysis of extracellular neural signals. The low signal-to-noise ratio (SNR) and similarity of spectral characteristic between the target signal and background noise are obstacles to solving this problem and, thus, in previous studies on experimental neurophysiology, only action potentials with sufficiently large amplitude have been detected and analyzed. In order to lower the level of SNR required for successful detection, we propose an action potential detector based on a prudent combination of wavelet coefficients of multiple scales and demonstrate its performance for neural signal recording with varying degrees of similarity between signal and noise. The experimental data include recordings from the rat somatosensory cortex, the giant medial nerve of crayfish, and the cutaneous nerve of bullfrog. The proposed method was tested for various SNR values and degrees of spectral similarity. The method was superior to the Teager energy operator and even comparable to or better than the optimal linear detector. A detection ratio higher than 80% at a false alarm ratio lower than 10% was achieved, under an SNR of 2.35 for the rat cortex data where the spectral similarity was very high.

Index Terms—Action potential detection, extracellular neural signal recording, signal-to-noise ratio, Teager energy operator, wavelet transform.

I. INTRODUCTION

THE extracellular recording of neural signals consists of action potentials from several neurons near the electrode site, and background noise. Since information of the nervous system is encoded in the form of firing frequency or firing time [1], the first procedure in the interpretation of neuronal signals is the detection of the action potential firing, i.e., the neural spike. In spite of the fundamental importance of this, only a few studies on neural spike detection have appeared in the literature. In most cases, major efforts have been made to optimize experiments so that the recorded waveforms are of sufficient quality to enable reliable detection by simple traditional methods. However, situations are often encountered where the signal-to-noise ratio

(SNR) of the recording is so poor as to prohibit neural spike detection using simple thresholding, and in some cases, such as the recording from a long-term implanted electrode, precise experimental control cannot be achieved. Moreover the statistical characteristics of background noise can be very similar to those of the target signal (action potential). This statistical similarity arises from the fact that the major noise source can be electrical potentials from neurons that are not coupled sufficiently tightly to electrode sites. This "neuronal" or "biological" noise may result in more serious problems in the case of a recording from cortex or ganglia where the density of neurons can be high. This statistical similarity prohibits the satisfactory enhancement of SNR using conventional signal processing techniques such as bandpass filtering, and as a result, the detection problem becomes much more difficult to solve.

Because the computational power required for the use of more sophisticated algorithms is now readily available, efforts directed toward the development of signal processing techniques for action potential detection are within the realm of possibility. The situation is similar to the problem of QRS detection from an electrocardiogram (ECG), in that the goal is to detect short transient waveforms in the presence of background noise, however, this requires less computational power due to the lower sampling rate employed. Numerous studies on the development of the QRS detection algorithm have been reported during past several decades [2], [3].

A considerable number of studies on the classification of multiunit extracellular recordings can be found in the literature [4]–[7]. Most of these are concentrated on improvements in classifier performance. However, the action potential detection is a preliminary step which can dominate the overall performance of a neural spike sorting system. Bankman and Janselewitz [8] described a method for the elaborate determination of threshold level: they determined it by analyzing the probability density function (pdf) of the background noise, i.e., by modeling the initial segment of the recording where no neural spike fires as a white Gaussian noise. However, no matter how well the threshold level is selected, the direct application of thresholding has fundamental limitations in case of low SNR values that are the focus of this paper. Background noise modeling as a white Gaussian noise is also not appropriate, as demonstrated by Fig. 4 of this paper and by Fee *et al.* [9] who showed similar plots of the spectra of action potentials and background noises. Most methods for transient signal detection use appropriate signal processing

techniques before applying thresholding, in order to boost the target signal while shrinking the background noise [10]. In this procedure, *a priori* knowledge of the characteristics of signal and noise are utilized. Another study by Bankman *et al.* [4] reported that a prewhitening filter enhanced SNR by assuming a complete knowledge of the spectral characteristic of the background noise. Chandra and Optican [5] reported on a neural-network-based action potential detection and classification [5]; but this implies the availability of a training set with a known class label, which is unrealistic in cases of frequently observable low SNR recordings. In real situations, especially when the SNR is very low, this quantitative information is not available and can even be time-varying.

Recently wavelet-transform-based methods for neural spike detection have been reported. Oweiss and Anderson [11] proposed an array signal processing technique for neural spike detection that involves signal subspace estimation and threshold denoising in the wavelet packet domain. They showed that this method enables the generalized likelihood ratio test (GLRT) without computing the noise-only covariance matrix. This method is effective when a multichannel electrode is used. Nakatani *et al.* [12] reported on a method for determining the scales to be considered and the threshold levels at each scale, for detection and waveform denoising. Their focus appears to be on the latter, and they compared its performance only with that of conventional wavelet denoising and showed that its detection performance is not as good as that of conventional wavelet denoising.

In this paper, we present a novel wavelet-based nonlinear method for the detection of action potential from single channel extracellular recording. We show its performance on various experimental recordings where the SNR is low, and the similarity between the characteristics of the target signal and background noise is also significant. These include data from the somatosensory cortex of the rat, the crayfish giant medial nerve, and the cutaneous nerve of the bullfrog. Differently from the method reported by Nakatani *et al.* [12] based on the wavelet denoising, our method is derived from nonlinear combination of multiple approximations of matched filters. Our method utilizes the point-wise product of wavelet transform (WT) coefficients over several selected scales. The wavelet basis function has a "spiky" waveform with a short time support, and thus, a single subband of the wavelet decomposition can be regarded as performing the function of the matched filtering, without *a priori* knowledge of the target signal and noise. From classical detection theory, the matched filter (along with the prewhitening filter) is known to provide an optimal linear method for signal enhancement under Gaussian noise [10]. However, this solution requires a precise knowledge of the target signal waveform and the spectral characteristics of the noises. Unfortunately, in many cases, including ours, none of this information is easily obtainable, and the characteristics of the noise are even nonstationary. The performance of our detection method was compared with those of the above inapplicable optimal linear method, and the Teager energy operator (TEO) [6], [13] which is a nonlinear method exploiting the instantaneous rise in amplitude and frequency of the dominant frequency component.

II. EXPERIMENTAL METHODS

A. Recording From Rat Somatosensory Cortex

A semiconductor microelectrode fabricated by the authors' group [14] was used for recording from the somatosensory cortex of Sprague–Dawley rats. Animals with body weights from 200–300 g were anaesthetized with urethane (1 g/kg, i.p.). After mounting the animals in a stereotaxic frame, a craniotomy (2–3 mm diameter) was performed over the primary somatosensory (SI) cortex using the bregma as the initial point of reference. Detailed methods pertaining to this surgery can be found in [15]. The semiconductor microelectrode was inserted into the forepaw area of the SI with a micromanipulator. Stainless-steel wire located in the subcutaneous region on the side of the cerebellum was used as a reference electrode. Electrical stimulation was provided by a bipolar concentric stimulating electrode and consisted of monophasic square pulses (pulse width 0.1 ms, frequency 1 Hz) passed through a stimulator (Model 1830, World Precision Instruments) with an isolation unit to provide constant current ($50 \sim 500 \, \mathrm{mA}$). The stimulating electrode was inserted under the center of the receptive field and was fixed firmly to prevent any movement. The amplified and bandpass filtered (300 Hz to 3 kHz) signal is transferred to a digital storage oscilloscope for visual inspection and then to the data acquisition system for digitization and subsequent storage and analysis using a personal computer. The sampling rate for the digitization was 20 ksamples/s, and the signal was subsequently resampled at 10 ksamples/s by a cascade of an antialiasing finite impulse response (FIR) low-pass filter designed by the Kaiser window method [16] and decimation.

B. Recording From Crayfish Medial Giant Nerve

A planar-type microelectrode array was used for recording from the medial giant nerve of crayfish. The medial giant nerve is a bundle-structured set of motor axons and sensory axons crossing the body center, and can be easily laid down on the surface of the planar microelectrode. Details relating to the dissection can be found in [17]. The electrode site was made of gold, and the impedance of the electrode site was in the range of 200–800 kΩ at 1 kHz. The evoked response to the mechanical stimulus applied to the tail-fin by probing lightly with a glass probe was recorded. The signal from the microelectrode array was preamplified by an ac coupled, fully differential amplifier (ACamp08, AC Instruments Corp.) having an input impedance higher than 1 GΩ, further amplified, and bandpass filtered (300 Hz to 3 kHz).

C. Recording From Bullfrog Cutaneous Nerve

A hook-type stainless-steel electrode was used for recording from the cutaneous nerve of the bullfrog responding to a mechanical stimulus given to the skin. The electrode impedance was about 10 kΩ at 1 kHz. After dissecting the backside skin of the bullfrog, a peripheral nerve bundle is hooked by the electrode. The nerve bundle, containing the axons from the cutaneous mechanoreceptor, was identified by monitoring the response to the mechanical stimulus applied to the skin by probing with a glass micropipette. Low-pass filter with 10-kHz cutoff frequency was used along with 60-Hz notch filter.
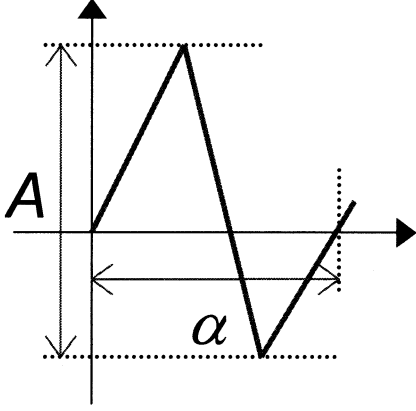
Fig. 1. Transient signal characterized by amplitude A, and scale α.

III. ACTION POTENTIAL DETECTION USING WAVELET COEFFICIENT COMBINATION

A. Detection Algorithm

Consider following equation that defines WT:

$$W(\alpha, \tau) = \int_{-\infty}^{\infty} x(t) \frac{1}{\alpha^{1/2}} \psi\left(\frac{t-\tau}{\alpha}\right) dt \qquad (1)$$

where the real numbers α and τ denote scale and translation, respectively. As shown by (1), the WT essentially carries out a correlation analysis between the input signal ($x(t)$) and the translated and dilated version of a reference signal called the mother wavelet ($\psi(t)$) [18]. Hence, it would be expected that the output would have local maxima where the input signal most closely resembles the analysis template, i.e., the wavelet function. The wavelet basis functions have "spiky" (biphasic or triphasic and so on) waveforms with compact support, and accordingly are similar in shape to the action potential. In addition, the basis function is dilated over many scales. Thus, at some particular scales, it can act as a number of effective approximations of the matched filter, even though the exact waveform of the target signals is not known. This can also be stated more formally as follows. Suppose the goal is to detect a transient signal $s_0(t)$ that can be characterized by amplitude A, arrival time τ, and duration α, as shown in Fig. 1. The detection problem can be posed as a hypothesis testing as follows:

$$H_0 : x(t) = n_0(t)$$
$$H_1 : x(t) = A \cdot s_0\left(\frac{t-\tau}{\alpha}\right) + n_0(t). \qquad (2)$$

Here $n_0(t)$ denotes the background noise and is assumed to be a colored Gaussian random process. Assuming that the spectrum of $n_0(t)$ is known, the problem can be transformed into the one assuming a white Gaussian random process by application of prewhitening filter as follows:

$$H_0 : x(t) = n(t)$$
$$H_1 : x(t) = A \cdot s\left(\frac{t-\tau}{\alpha}\right) + n(t) \qquad (3)$$

where the noise $n(t)$ is now a white Gaussian random process. Here it is assumed that the target signal is known up to

unknown parameters A, τ, and α, which denote amplitude, arrival time, and scale, respectively. By GLRT [10], which uses the maximum-likelihood estimates of unknown parameters, Neyman–Pearson test statistic [10] is obtained as follows:

$$
\begin{aligned}
&\Lambda_1(x) \\
&= \frac{p(x|\hat{A}, \hat{\alpha}, \hat{\tau}, H_1)}{p(x|H_0)} \\
&= \frac{\exp\left(-\int_0^T \left[x(t) - \hat{A} \cdot s\left(\frac{t-\hat{\tau}}{\hat{\alpha}}\right)\right]^2 dt\right)}{\exp\left(-\int_0^T [x(t)]^2 dt\right)} \\
&= \exp\left(\int_0^T \left(2x(t) \cdot \hat{A} \cdot s\left(\frac{t-\hat{\tau}}{\hat{\alpha}}\right) - s^2\left(\frac{t-\hat{\tau}}{\hat{\alpha}}\right)\right) dt\right) \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_1. \quad (4)
\end{aligned}
$$

By taking logarithms and simplifying, we obtain

$$
\begin{aligned}
l_1(x) &= \int_0^T x(t) \cdot \frac{1}{\hat{\alpha}^{1/2}} \cdot s\left(\frac{t-\hat{\tau}}{\hat{\alpha}}\right) dt \\
&= \max_{\alpha, \tau}\left[\int_0^T x(t) \cdot \frac{1}{\alpha^{1/2}} \cdot s\left(\frac{t-\tau}{\alpha}\right) dt\right] \underset{H_0}{\overset{H_1}{\gtrless}} \lambda_2(\hat{A}, \hat{\alpha}, \hat{\tau}).
\end{aligned}
$$
$$(5)$$

Here the right side of (5) indicates the dependency of the threshold level on the amplitude, arrival time, and scale. Note that it is obvious from (5) that the amplitude A does not affect the test statistic, but only changes the threshold level and, thus, a lack of knowledge of A is not relevant. From (1) and (5), we can deduce that the wavelet analysis at the scale that yields the maximum output is equivalent to the GLRT, provided that exact template waveform of the signal is used as the wavelet basis function, i.e., $\psi(t) = s(t)$. In fact, the derivation of (5) is equivalent to that of the matched filter as shown in [10], except for the assumption that the scale α is unknown. It becomes clear that wavelet analysis may provide a valuable practical tool for the approximation of matched filtering, although there are several assumptions that cannot be satisfied in an actual situation of neural signal recording. For example, the wavelet basis $\psi(t)$ cannot be exactly the same as the template waveform $s(t)$, and the background noise cannot always be assumed to be a Gaussian random process. Our purpose here is to show that by performing wavelet decomposition over many scales and selecting some of those that yield faithful representation of the signal, it is possible to obtain a number of useful approximations of the matched filtering.

Therefore, it is possible to implement an effective action potential detection scheme by combining the wavelet coefficients calculated over a series of multiple scales. We use the point-wise product of the wavelet coefficients over some successive scales. This exploitation of the point-wise product is intended to perform the role of logical "AND" operation (although not precisely) over multiple scales. Details of the procedure are as follows. First, the wavelet coefficients and their absolute values are calculated for 5 dyadic scales, i.e., $2^1 \sim 2^5$, and the scale where the absolute value yields a maximum (this scale is called $2^{j\max}$) is selected. Subsequently, as shown below in (6), $P(n)$, the point-wise product of the wavelet coefficients over three
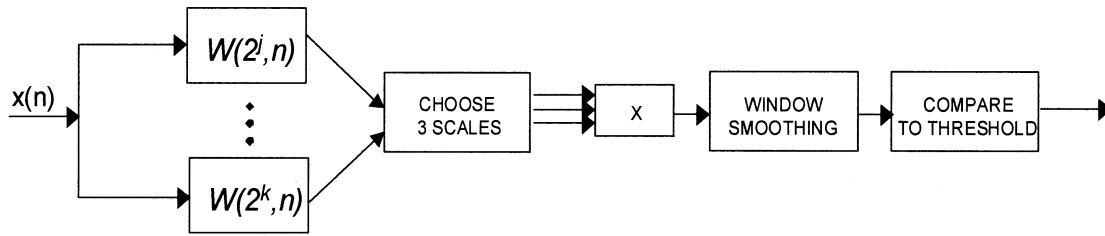
Fig. 2.   Block diagram of the proposed detection method (DWT product detector).
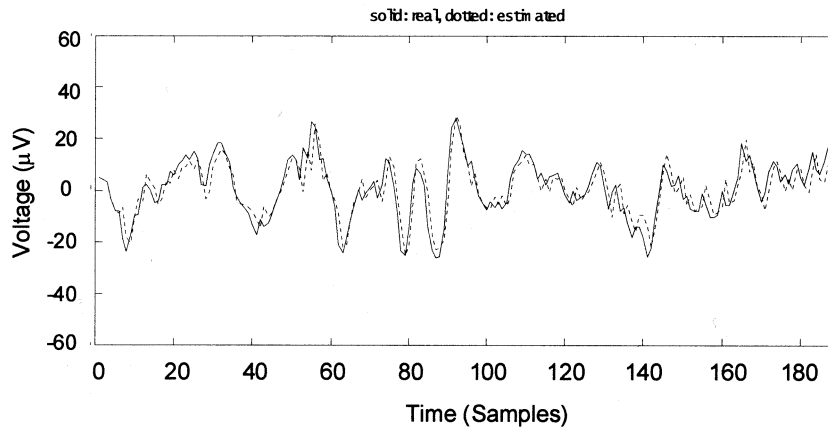

Fig. 3.   Modeling of the background noise of the crayfish medial giant nerve recording. Solid line: actual data. Dotted line: estimation by the fifth-order AR model.

consecutive scales up to $2^{j_{\max}}$ is calculated for all time samples, n, as follows:

$$P(n) = \prod_{j=j_{\max}-2}^{j_{\max}} \left| W(2^j, n) \right|. \qquad (6)$$

This is similar to an efficient edge detection technique used in the field of image processing. Before the advent of wavelet theory, Rosenfeld and Thurson [19] devised an edge detection method employing the calculation of the point-wise product of smoothed gradients, where the degree of smoothing takes dyadic values. More recently, Sadler and Swami [20] reported on a statistical analysis of Rosenfeld's method. However, these studies provide no guidelines for the selection of the scales to be used in product calculation, and it must be chosen carefully so as to be suited for a specific desired application. We found that the choice of three consecutive dyadic scales up to $2^{j_{\max}}$ is appropriate for most cases of action potential detection. The inclusion of excessively large or small scales considerably decreased the signal peaks because the correlation between the analysis template (dilated wavelet basis) and the target signal (action potential) is decreased for those scales. The inclusion of too coarse a scale also decreased the signal peaks in the resulting waveform of $P(n)$ because of the large mismatch in the location of the peaks among different scales.

The wavelet decomposition into dyadic scales can be computed by several means. It can be shown that a rapid computation is possible when the translation is limited to the integer multiples of the scale [18], [21]. This yields a nonredundant, orthogonal representation and is usually called discrete wavelet transform (DWT). The DWT can be performed by successive applications of a bank of quadrature mirror filters and decimation by factor of two between them. A slightly more complicated version, which is called stationary or undecimated DWT (UDWT) and has the advantage of translation-invariance, can be obtained by omitting the decimation [21]. We tested both the decimated DWT and UDWT, and the performances of detection were similar for our data. However, the UDWT was chosen considering its potential benefit in more exact waveform estimation. The coiflet basis function [22] was used for our system.

The DWT product $P(n)$ is smoothed by convolution with the Bartlett window to mitigate malicious effects of spurious peaks due to cross terms, background noise, and slight mismatches in the location of the signal peaks over different scales. The window length was determined empirically, taking all these effects into consideration and found to be suitable when it is about half the duration of the target signal. The final output of the proposed action potential detector, $T(n)$, is expressed as follows:

$$T(n) = w(n)^* P(n) = w(n)^* \left(\prod_{j=j_{\max}-2}^{j_{\max}} \left| W(2^j, n) \right| \right). \quad (7)$$

Here $w(n)$ denotes the Bartlett window. Block diagram describing the DWT product method is shown in Fig. 2.

B. Performance Test

Performance tests were carried out while varying the level of the background noise, and the similarity between the signal and noise. For this comprehensive test, a large amount of data sets representing real characteristics of various experimental neural signal recordings were required. A time-series prediction technique was used for the modeling of experimental background
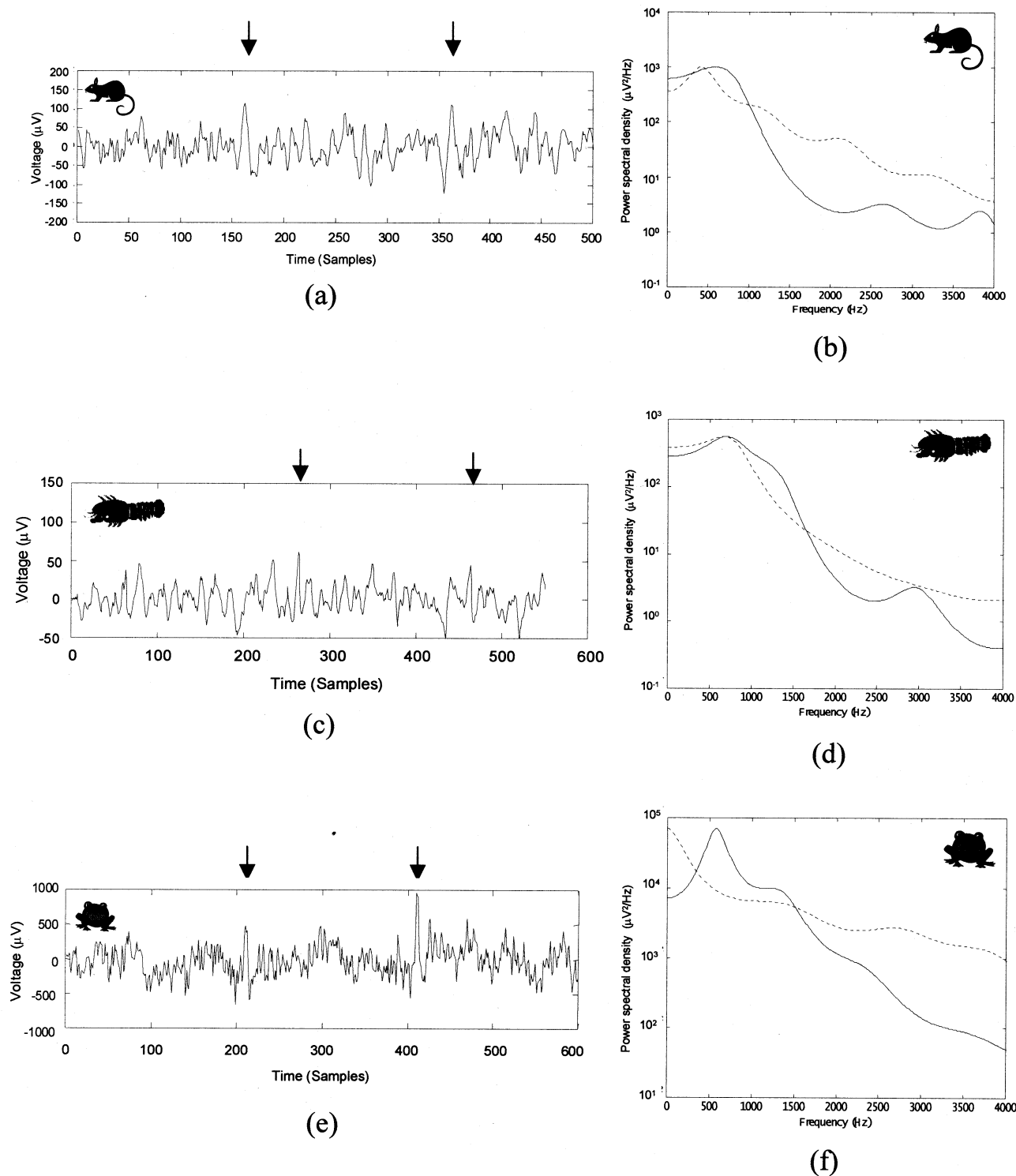
Fig. 4. Waveforms (left columns) and power spectra (right columns) of the three neural signals considered in this paper. (a) and (b): Rat somatosensory cortex recording. (c) and (d): Crayfish recording. (e) and (f): Bullfrog recording. The arrows indicate the position of action potential firing.

noise in order to construct this large data set. From the template waveform and the model of the noise, a test data set of arbitrary SNR could be generated. Template waveforms were obtained by averaging distinct action potential waveforms from the experimental recordings, using our detection method assisted by visual inspection. Several linear and nonlinear time-series modeling techniques such as the autoregressive (AR), autoregres-

sive moving average (ARMA) model [23], and time-delayed feedforward neural network (TDNN) [24] were tried for the modeling of the background noise as colored Gaussian noise. For our rat cortex and crayfish data, we were able to obtain satisfactory results using the AR model as shown in Fig. 3. For the bullfrog data, the ARMA was more appropriate for the background noise modeling, however, the AR model with a
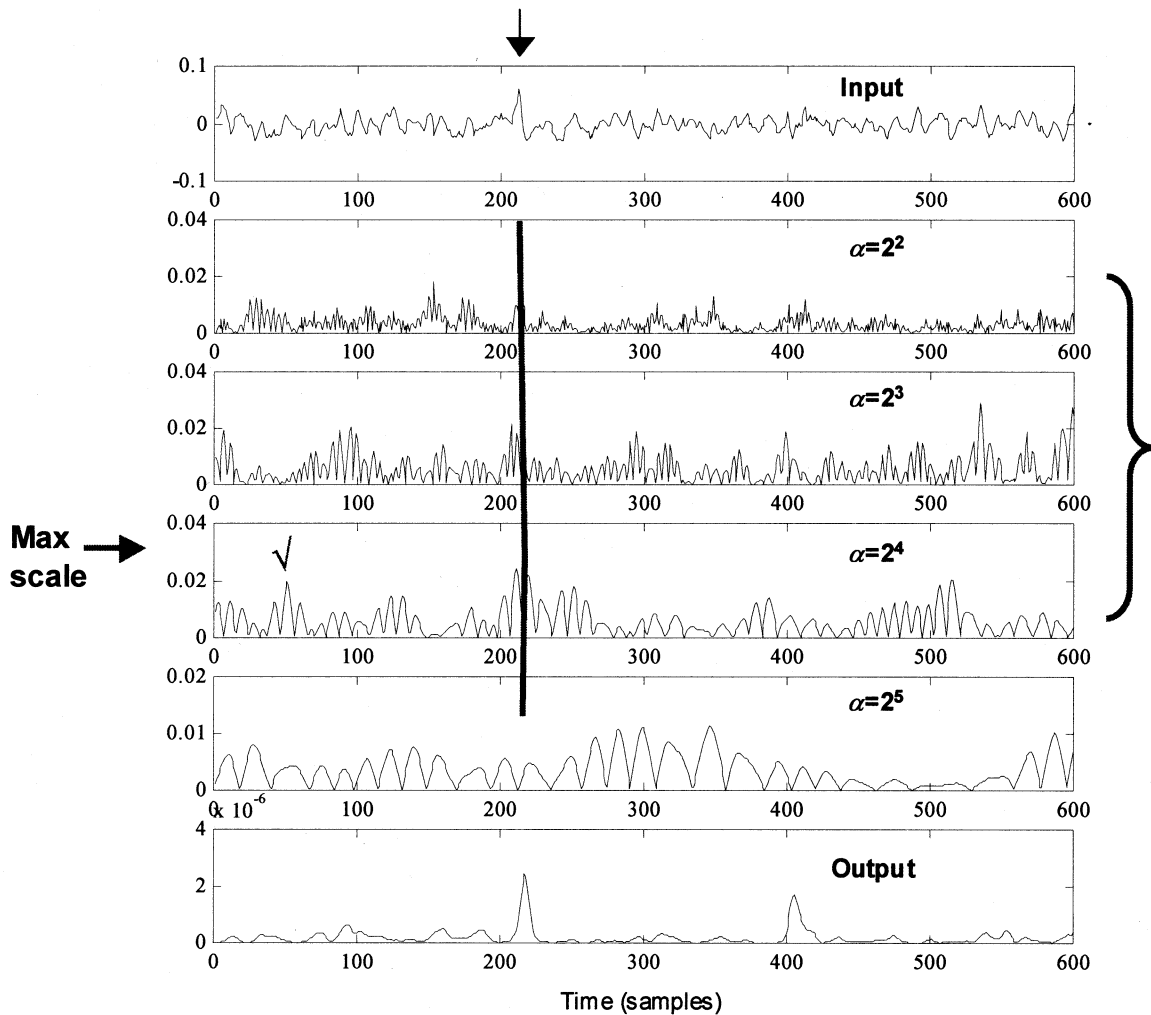
Fig. 5. Illustration of the rationale for the DWT detector. As denoted by the thick line, which is nearly vertical, the peaks in the wavelet transform coefficients due to the transients in input signal appear at nearly same time point over successive scales, while the peak due to the background noise (denoted as "$\sqrt{}$") is prominent only on a single scale. "}" indicates the scales that are included in the detection. The arrow indicates the position of action potential firing.

high order was also acceptable. The order of the AR models was determined using Akaike's information criteria [23], but the order selection was not crucial to the performance of the time-series modeling, since satisfactory prediction was possible for a large range of AR model orders. We used the fifth-order AR model for the rat cortex and crayfish data, and the fifteenth-order model for the bullfrog data. Calculation of the AR model coefficient was performed by solving the Yule–Walker equation [23], or using the third-order spectra-based method [25]. The SNR was defined as ratio of powers of target signal waveform and noise, as follows:

$$\text{SNR} = \left(\frac{\text{peak-to-peak value of action potential waveform}}{\text{root-mean-square value of pure noise segment}} \right)^2 .$$

IV. RESULTS

Fig. 4 shows the waveforms (left columns) and power spectra (right columns) for the three recordings. The arrows in Fig. 4(a), (c), and (e) indicate the position of action potential firing. The

solid and dotted lines in Fig. 4(b), (d), and (f) correspond to the power spectra of the action potentials and those of the background noises, respectively. The SNRs of the data shown in Fig. 4 were about 2.16 for the rat cortex recording [Fig. 4(a) and (b)], 2.15 for the crayfish recording [Fig. 4(c) and (d)], and 2.13 for the bullfrog recording [Fig. 4(e) and (f)], respectively. Significant spectral overlap between the signal and background noise is evident for all the recordings considered in this paper, and this has been also shown by Fee *et al.* [9]. The large amount of spectral overlap makes the action potential detection problem more difficult. Fig. 5 shows the rationale for the DWT product detector. As denoted by the thick line in Fig. 5, the peaks in the wavelet transform coefficients due to the transient in the input signal appear at nearly the same time points over successive scales, while the peak due to the background noise (denoted by "$\sqrt{}$") is prominent only on single scale.

A comprehensive performance test of the proposed detector, and a comparison with the matched filter and the TEO are presented below. The threshold level for detection was determined by manual adjustment, based on data segments containing about
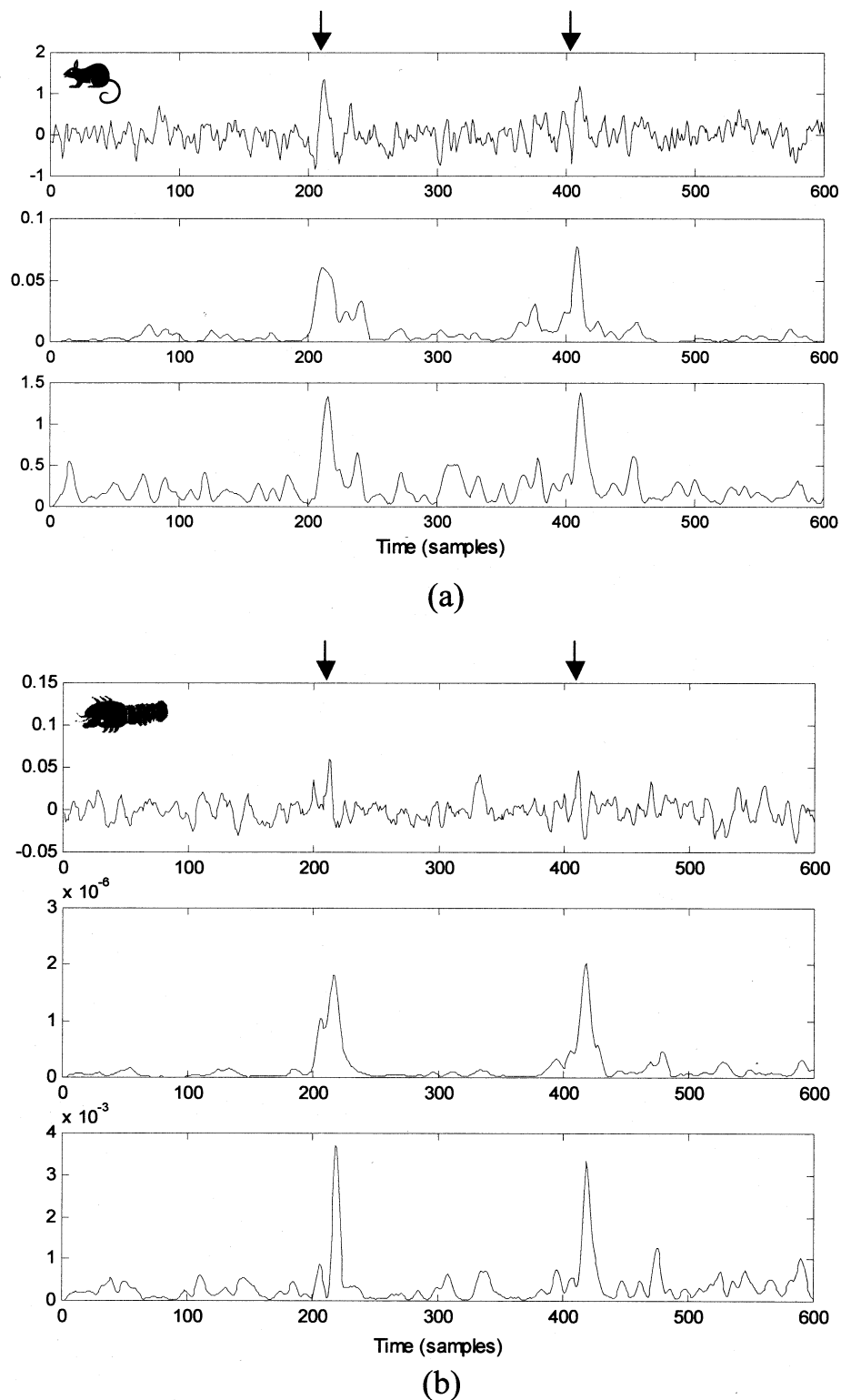
Fig. 6. (a) The input waveform of the recording from rat somatosensory cortex (upper panel), and the output waveforms when the input is applied to the DWT product detector (middle panel) and the TEO (lower panel) detector. (b) The input waveform of the crayfish recording (upper panel), and the output waveforms when the input is applied to the DWT product detector (middle panel) and the TEO (lower panel) detector. The arrows indicate the position of action potential firing.

20 action potentials while attempting to minimize the wrong inclusion of noise (i.e., false alarms) and detection misses. A statistical determination of the threshold level was not attempted here because it requires a complete knowledge of the pdf [10]. Although it was possible to obtain knowledge of the pdf for the test data set by Monte Carlo method [10], our goal was to devise and test a method that is generally applicable to common experimental situations.

Fig. 6(a) shows the input waveform of the recording from the somatosensory cortex of the rat, and the output waveforms
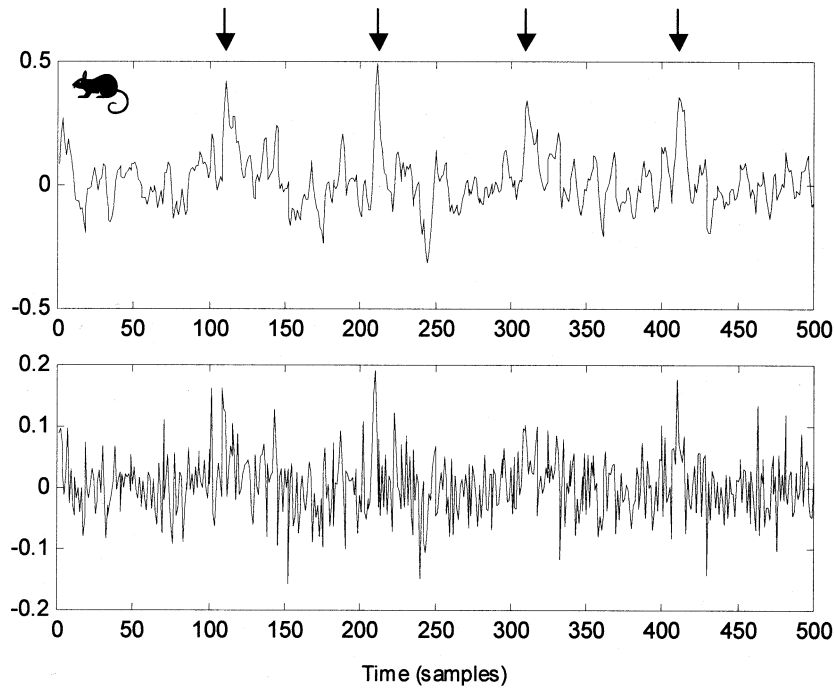
Fig. 7.   The limited efficacy of the prewhitening filter for action potential detection. Upper panel: Rat cortex recording with SNR≈ 2.76. The arrows indicate the position of action potential firing.
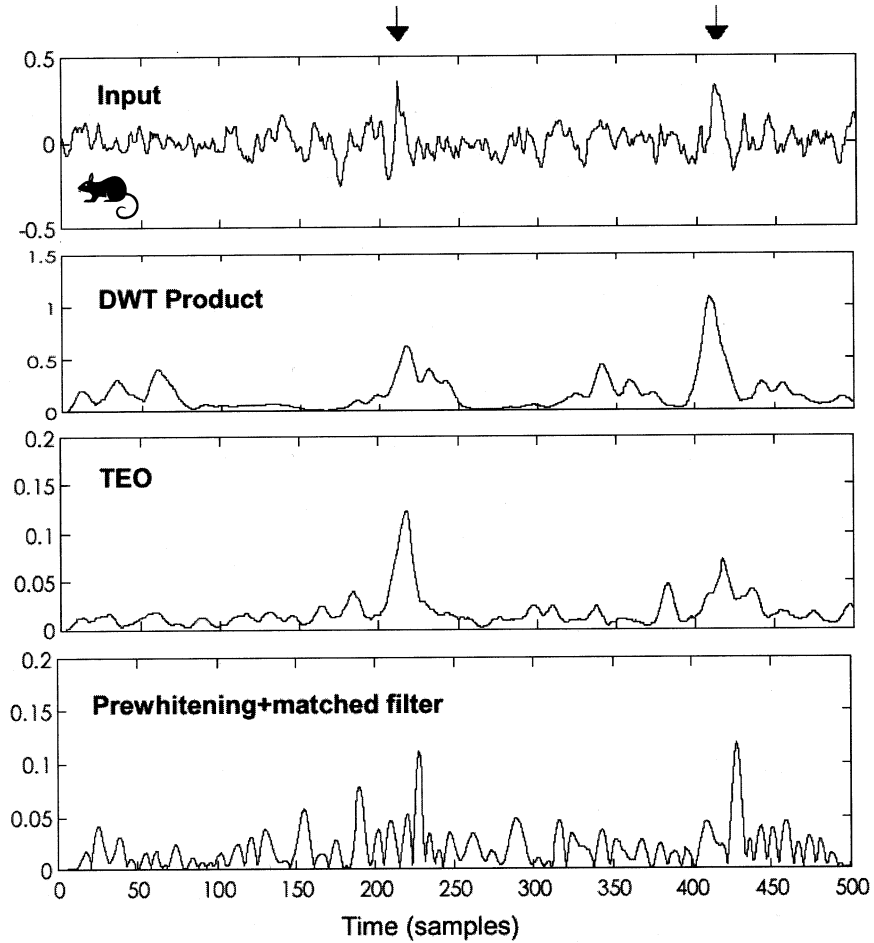

Fig. 8.   Comparison of the signal quality after processing by the matched filter and proposed detectors. Input: Rat cortex recording, SNR≈ 2.33. The arrows indicate the position of action potential firing.
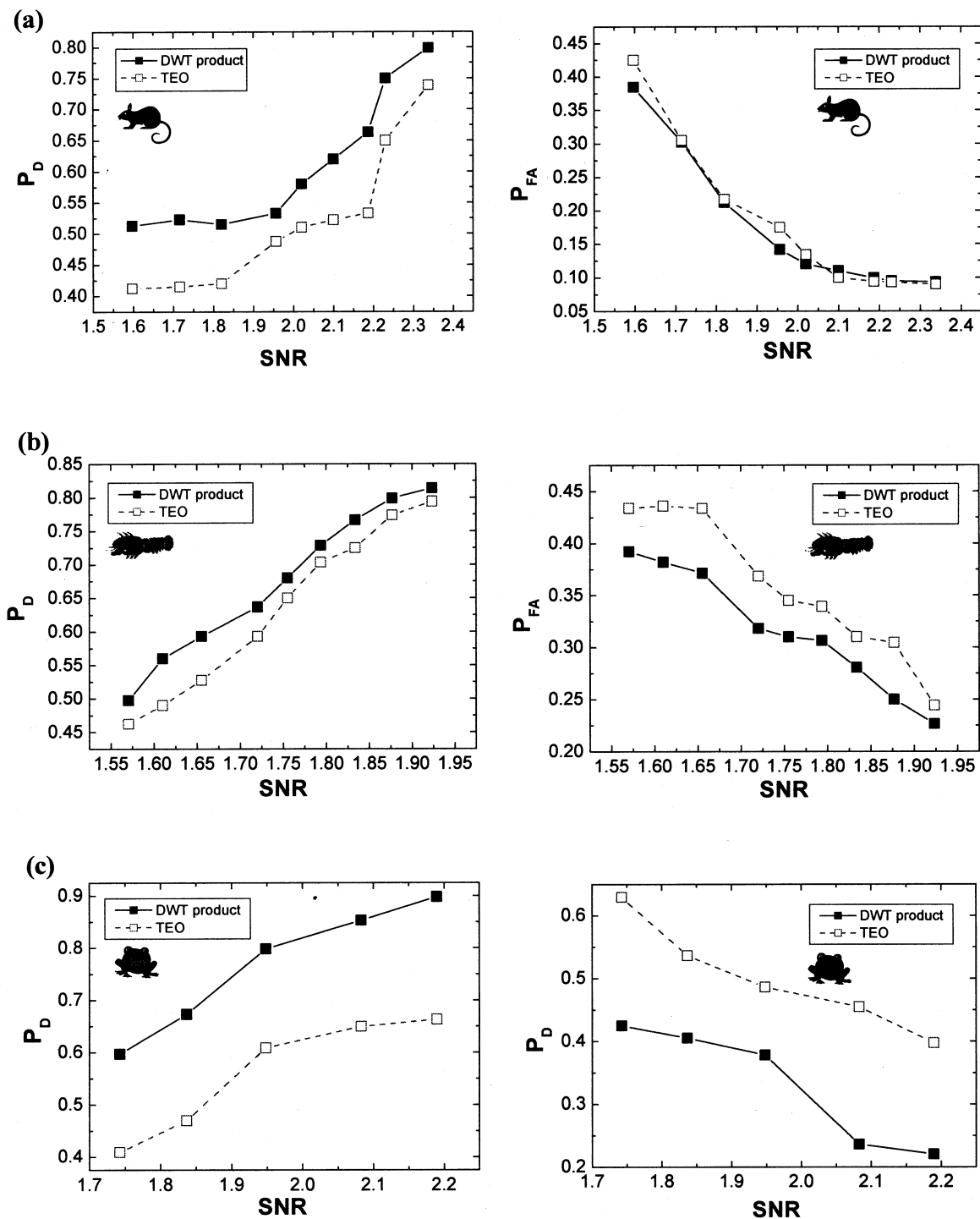
Fig. 9. Left columns: P_D versus SNR. Right columns: P_{FA} versus SNR. Solid lines: DWT detector. Dotted lines: TEO detector. (a) Rat cortex recording. (b) Crayfish recording. (c) Bullfrog recording.

when the input is applied to the DWT product detector and the TEO detector. The upper panel shows the input waveform, and the SNR is about 2.52. The output from the TEO detector (bottom panel) has more distinct peaks than the output from the DWT detector (middle panel) at points where the action potentials fire. However, the TEO detector appears to produce a larger amount of spurious peaks. This phenomenon

also occurred in other situations, when recordings from the crayfish [Fig. 6(b)] and the bullfrog (not shown) were used as inputs.

The performance of the proposed method was compared with those of the TEO, the prewhitening filter, and the matched filter with preceding prewhitening filter [10]. Fig. 7 shows the limited efficacy of the prewhitening filter for the action potential

detection. The prewhitening filter was helpful only for the crayfish data. In the case of the rat cortex data shown in Fig. 7 and the bullfrog data (not shown), signal qualities were somewhat deteriorated. Fig. 8 shows a comparison of the signal quality after processing by the matched filter with prewhitening, with our detectors. It is not possible to judge whether this linearly optimal method is superior or inferior to our detector from Fig. 8 alone. Later we will show a more detailed comparison of detection performance by receiver operating characteristic [10]. In order to investigate detection performances more thoroughly, we plotted the probability of detection (P_D) and the probability of false alarm (P_{FA}) at various levels of SNR. Fig. 9(a) shows the results for the rat recordings. The solid and dotted lines in Fig. 9 correspond to the DWT and the TEO detectors, respectively. The test was performed using data that included 1,000 action potentials. In Fig. 9(a) the level of P_{FA} is intentionally controlled to be similar for the two detection methods, by adjusting the threshold level. For similar P_{FA}, the DWT product detector yielded a higher P_D than the TEO detector, for all SNR levels. Fig. 9(b) shows the result for the crayfish recording, where we attempted to hold the value of P_D to a similar level. Here, lower P_{FA} (i.e., better performance) could be obtained by the DWT product detector. This improved performance of the DWT detector was more profound for the case of the bullfrog recording shown in Fig. 9(c).

A comparison of the detection performance by means of the receiver operating characteristic (ROC) is presented. The ROC curve was obtained by plotting P_D as a function of P_{FA}[10]. Each data point in a curve corresponds to each different threshold value. Fig. 10(a)–(c) shows the results for the rat cortex data, crayfish data, and bullfrog data, respectively. In Fig. 10(a), the TEO and DWT product detector shows similar level of P_D when a high level of P_{FA} was allowed, but the latter shows a superior performance for a low level of P_{FA}. Both were superior to the matched filter with prewhitening. For the crayfish data [Fig. 10(b)], the difference in performance between the two detectors was larger. The performance of the TEO detector was slightly inferior compared to the matched filter for a low P_{FA}. The DWT product detector showed consistently higher performance than that of the matched filter. For the bullfrog data, as shown in Fig. 10(c), both were not as good as the optimal linear method, however, the performance of the DWT product detector was comparable. In summary, the proposed detector showed high performance, superior or comparable to that of the matched filter with prewhitening, which is unrealizable.

Finally the application of the proposed detector to an actual experimental recording from the rat somatosensory cortex is presented. Fig. 11(a) and (b) shows a collection of the waveforms detected by the proposed method and a scatter plot of its principal components, respectively. Two clusters are distinct in the scatter plot in Fig. 11(b). This is also clear from Fig. 11(a) where waveforms from two units are present. It contains 210 detections. However, in Fig. 11(c) where conventional amplitude thresholding was used to detect action potentials, it is not apparent whether two units are present in the detected waveforms. It contains 566 detections. We can deduce from
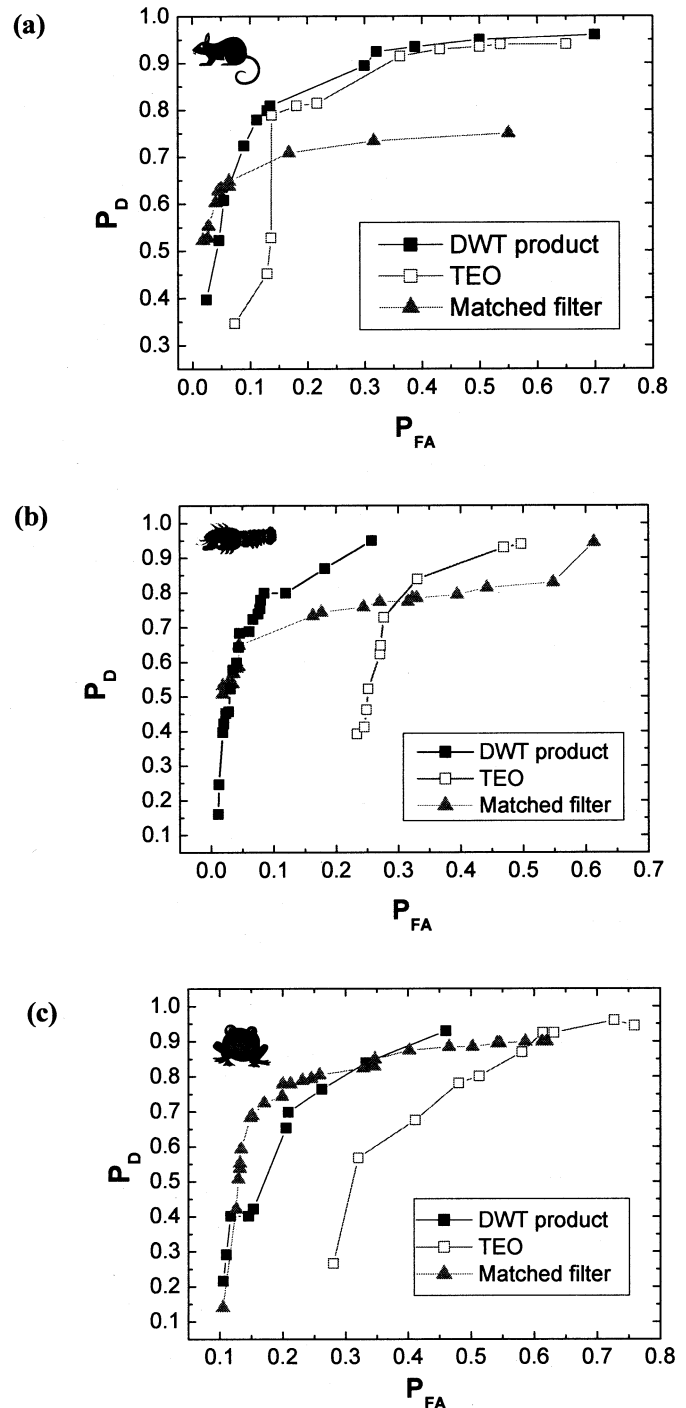

Fig. 10. Comparison of detection performance by the ROC curve. (a) Rat cortex data (SNR$\approx$ 2.33) (b) Crayfish data (SNR$\approx$ 2.38) (c) Bullfrog data (SNR$\approx$ 1.91).

the scatter plot in Fig. 11(d) that the detected waveforms in Fig. 11(c) include numerous false alarms, since there is a noise cluster here in-between the two clusters that are also present in Fig. 11(a). When we increased the threshold level to remove these false alarms, many correct action potentials were also removed and a considerable number of noise segments still remained. This is shown in Fig. 11(e) and (f). Here again, 210 detections were included, but a large portion of these is from the false inclusion of noise.
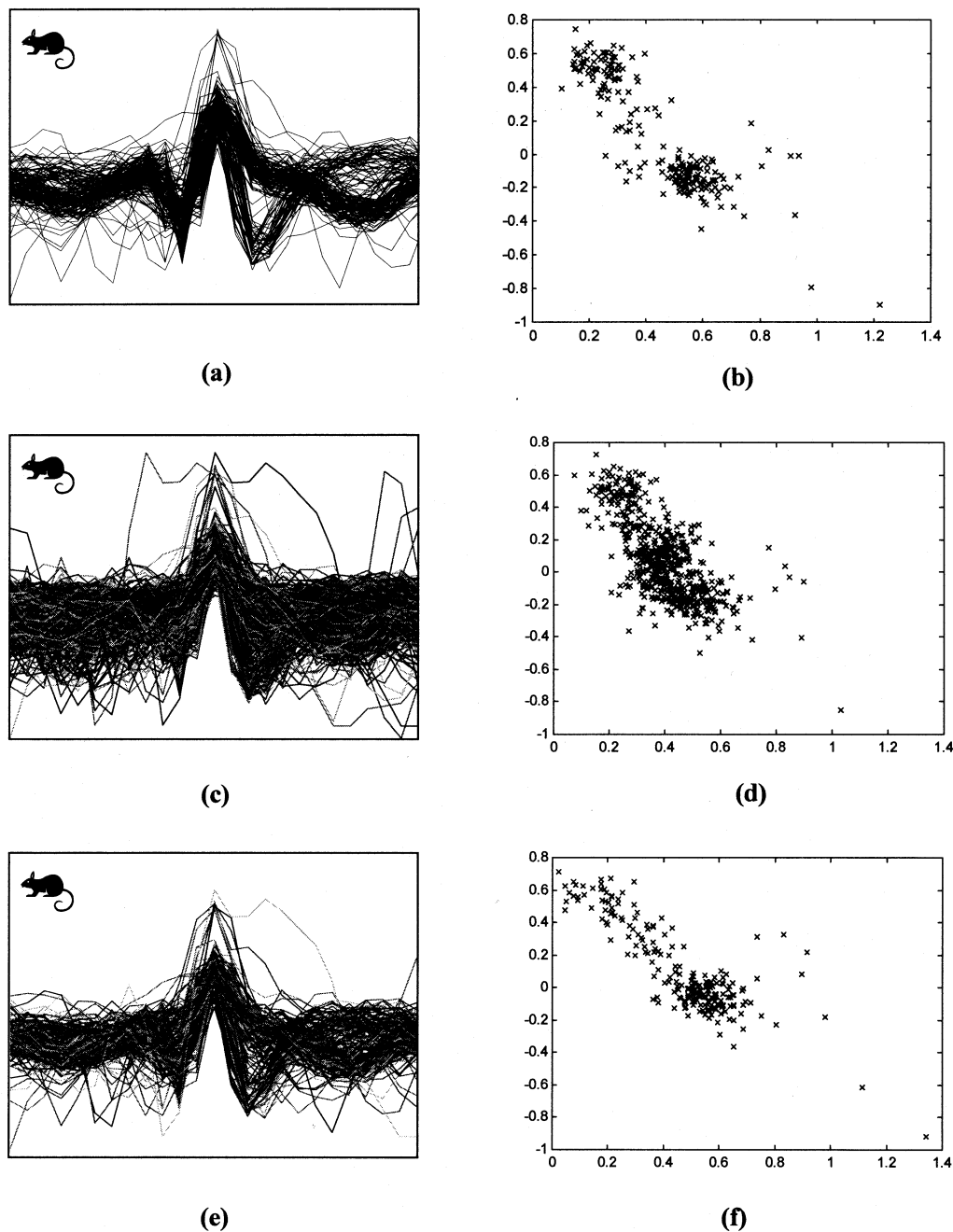
Fig. 11. (a) Collection of the waveforms detected by the proposed detector. (b) Scatter plot of the first two principal components extracted from the data shown in (a). (c) Collection of the waveforms detected by amplitude thresholding. (d) Scatter plot of the first two principal components extracted from the data shown in (c). (e) Collection of the waveforms detected by amplitude thresholding when a higher threshold level was used in order to remove the false inclusion of noise segments. (f) Scatter plot of the first two principal components extracted from the data shown in (e).

V. DISCUSSION AND CONCLUSION

The detection method proposed in this paper does not require "quantitative" *a priori* information on either the target signal and background noise, and only involves "qualitative" information that is common to the neural signal recordings. This is also valid for the previously described TEO detector [6], where we simply made use of the fact that the instantaneous frequency and amplitude of the dominant frequency component increases when the action potential fires. This is a clear advantage over several previously proposed action potential detection methods, such as [7] and [5]. In order to use the method of Bankman *et al.* [4] or that of Gozani and Miller [7], a complete knowledge of the background noise and the action potential waveforms of all the units present in the recording is required. Neural-network-based action potential detection [5] necessitates a training set with known class label, and this requirement is equivalent to a complete quantitative knowledge of the target signal and noise.

Among the two methods proposed by the authors, the current method, the DWT product detector, shows superior performance for all the recordings considered to date. From the viewpoint of computational requirements, the TEO required about

ten times less number of floating point operations and, thus, is better suited to real-time implementation. However, when implemented on a Pentium III PC with 500-MHz clock speed, the difference in execution time was not great (1.02 times faster than the wavlet method for 10^5 data points processing). Some methods that are similar to our wavelet-based action potential detector have been proposed for QRS detection from ECG signals [2], [3], [28]. Studies by Kadambe *et al.* [2] and Li *et al.* [28] involve searching for all the local maxima of the absolute value of the wavelet coefficients within the analysis window over some successive scales and, therefore, can be time-consuming. This becomes much more problematic in case of neural signal processing where the sampling rate is about 100 times higher than the case of ECG processing. Moreover, QRS detection seems to be a much easier problem than the action potential detection from extracellular neural signal recordings, with respect to both the SNR and the similarity between the signal and background noise.

Several wavelet-based methods for neural signal analysis have been reported [11], [12], [26], [29]. Hulata *et al.* [29] proposed an application of wavelet packet decomposition for neural spike detection and sorting, although emphasis was given to the latter. Their method does not appear to be appropriate for a fully automated system, since it requires the initial manual grouping of signals into several neural spikes and noise to find the basis. An action potential detection method utilizing wavelet denoising was proposed by Nakatani *et al.* [12]. This method is based on the removal of the background noise which is effective when it shows white spectrum. Their performance test was demonstrated for cases where the background noise was a Gaussian random process that was lowpass-filtered to the cutoff frequency of 4 kHz, and the action potential was bandlimited to 100 Hz to the 2 kHz. Hence, background noise can be treated as a white Gaussian process if the sampling rate is changed to 4 kHz (which is feasible from the Nyquist theorem) and, thus, this is similar to the problem of band-limited signal detection under white Gaussian noise, where the wavelet denoising is effective. Their method may be less effective in cases where the background noise shows a spectral characteristic similar to that of the action potential, and the benefit of their method is mainly on accurate waveform estimation by noise removal.

Oweiss and Anderson [11], [26] described a series of array signal processing techniques for neural spike detection and sorting. They exploited signal subspace estimation and threshold denoising in the wavelet packet domain. Their methods yield high performance under low SNR, and seem to provide efficient solutions when a multichannel electrode array is employed and the activities of all the units under consideration are recorded by all the channels under investigation. Another multichannel technique, blind source separation using independent componenet analysis, has been applied to the problem of neural spike sorting [27], but its performance under low SNR has not been presented.

The possibility exists for improving the performance of the proposed detector by applying some slight modifications. For example, it might be possible to utilize information provided by incorporating a more elaborate selection of subbands. Wavelet packet decomposition might be applied in order to broaden the extent of subbands to be included for the detection. An appropriate criterion should be devised to guide in the selection of scales. A novel technique other than the point-wise product should be developed to combine the information from the selected subbands to achive an improved detection performance.

In conclusion, an action potential detector using the point-wise product of wavelet coefficients of multiple scales is presented. The method involves multiple approximations of matched filters and nonlinear combination thereof. The proposed method was tested for various SNRs and degrees of spectral similarity between the signal and background noise. The detection performance achived was better than that of the TEO [6] and even comparable or superior to that of the optimal linear detector. Since the proposed action potential detector does not require a quantitative knowledge of the signal and noise of the recording under investigation, it can be utilized for the online or first offline analysis of neural signal recordings, where such quantitative information is not readily available.

REFERENCES

[1] F. Rieke, D. Warland, R. De Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press, 1996.

[2] S. Kadambe, R. Murray, and G. F. Boudreaux-Bartels, "Wavelet transform-based QRS complex detector," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 838–847, July 1999.

[3] B. U. Kohler, C. Hennig, and R. Orglmeister, "The principles of software QRS detection," *IEEE Eng. Med. Biol. Mag.*, vol. 21, pp. 42–57, Jan.-Feb. 2002.

[4] I. N. Bankman, K. O. Johnson, and W. Schneider, "Optimal detection, classification, and superposition resolution in neural waveform recordings," *IEEE Trans. Biomed. Eng.*, pp. 836–841, Aug. 1993.

[5] R. Chandra and L. M. Optican, "Detection, classification, and superposition resolution of action potentials in multiunit single channel recordings by an on-line real-time neural network," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 403–412, May 1997.

[6] K. H. Kim and S. J. Kim, "Neural spike sorting under nearly 0 dB signal-to-noise ratio using nonlinear energy operator and artificial neural network classifier," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1406–1411, Oct. 2000.

[7] S. N. Gozani and J. P. Miller, "Optimal discrimination and classification of neuronal action potential waveforms from multiunit, multichannel recordings using software-based linear filters," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 358–372, Apr. 1994.

[8] I. N. Bankman and S. J. Janselewitz, "Neural waveform detector for prosthesis control," *Proc. 17th Ann. Conf. IEEE EMBS*, pp. 963–964, 1995.

[9] M. S. Fee, P. P. Mitra, and D. Kleinfeld, "Variability of extracellular spike waveforms of cortical neurons," *J. Neurophysiol.*, vol. 76, pp. 3823–3833, 1996.

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[11] K. Oweiss and D. J. Anderson, "A multiresolution generalized maximum-likelihood approach for the detection of unknown transient multichannel signals in colored noise with unknown covariance," in *Proc. ICASSP*, vol. 3, 2002, pp. 2993–2996.

[12] H. Nakatani, T. Watanabe, and N. Hoshiyama, "Detection of nerve action potentials under low signal-to-noise ratio condition," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 845–849, Aug. 2001.

[13] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "On amplitude and frequency demodulation using energy operators," *IEEE Trans. Signal Processing*, vol. 41, pp. 1532–1550, Apr. 1993.

[14] T. H. Yoon, E. J. Hwang, D. Y. Shin, S. I. Park, S. J. Oh, S. C. Jung, H. C. Shin, and S. J. Kim, "A micromachined silicon depth probe for multichannel neural recording," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1082–1087, Aug. 2000.

[15] H. C. Shin, H. J. Park, and J. K. Chapin, "Differential phasic modulation of short and long latency afferent sensory transmission to single neurons in the primary somatosensory cortex in behaving rats," *Neurosci. Res.*, vol. 19, pp. 419–425, 1994.

[16] A. V. Oppenheim and R. W. Schafer, *Discrete-time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[17] S. J. Kim, M. Kim, and W. J. Heetderks, "Laser-induced fabrication of a trans-substrated microelectrode array and its neurophysiological performance," *IEEE Trans. Biomed. Eng.*, vol. BME-32, p. 497, 1985.

[18] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proc. IEEE*, vol. 84, pp. 626–638, Apr. 1996.

[19] A. Rosenfeld and M. Thurson, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, pp. 562–569, 1971.

[20] B. M. Sadler and A. Swami, "Analysis of multiscale products for step detection and estimation," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1043–1051, Apr. 1999.

[21] S. Mallat, *A Wavelet Tour of Signal Processing*. San Diego, CA: Academic, 1998.

[22] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Soc. Ind. Appl. Math., 1992.

[23] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.

[24] S. Haykin, *Neural Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall, 1999.

[25] A. Swami and J. M. Mendel, "ARMA parameter estimation using only output cumulants," *IEEE Trans. Acoust. Speech. Signal Processing*, vol. 38, pp. 1257–1265, July 1990.

[26] K. G. Oweiss and D. J. Anderson, "A unified framework for advancing array signal processing technology of multichannel microprobe neural recording devices," *Proc. 2nd IEEE Conf. Microtechnology in Medicine and Biology*, pp. 245–250, 2002.

[27] G. D. Brown, S. Yamada, and T. Sejnowski, "Independent component analysis at the neural cocktail party," *Trends Neurosci.*, vol. 24, pp. 54–63, 2001.

[28] C. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using wavelet transforms," *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 21–28, Jan. 1995.

[29] E. Hulata, R. Segev, Y. Shapira, M. Benveniste, and E. Ben-Jacob, "Detection and sorting of neural spikes using wavelet packets," *Phys. Rev. Lett.*, vol. 85, pp. 4637–4640, Nov. 2000.
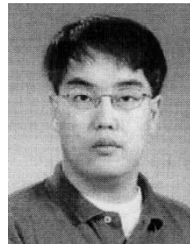
Kyung Hwan Kim (M'02) was born in Seoul, Korea, in 1973. He received the B.S. degree from Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejon, Korea, in 1995. He received the M.S. and Ph.D. degrees from the School of Electrical and Computer Engineering, Seoul National University, Seoul, Korea, in 1997, and 2001, respectively.

Since March 2001, he has been working as a Member of the Research Staff in the Human-Computer Interaction Laboratory, Samsung Advanced Institute of Technology, Yongin, Korea. Since March 2003, he has also been working as a Visiting Scholar at the Functional Magnetic Resonance Imaging (fMRI) Laboratory, Brain Science Research Center, KAIST. His research interests include biomedical signal processing, pattern recognition, and instrumentation with emphasis on their application to neural signals, for neuroscience research, neural prosthesis, and affective human-computer interface. He is also interested in advanced functional neuroimaging technology using a combination of electroencephalography and fMRI and its application to cognitive neuroscience and clinical problems.
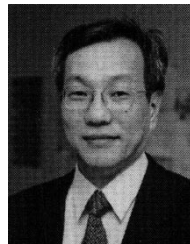
Sung June Kim (S'79–M'84) received the B.S. degree in electronics engineering from Seoul National University, Seoul, Korea, in 1978. He received M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1981, and 1983, respectively. His Ph.D. degree thesis was on the photo-induced fabrication of trans-substrate microelectrode arrays based on silicon substrate and their neurophysiological applications.

From 1983 to 1989, he worked as an MTS in Bell Laboratories, Allentown, PA, and in Murray Hill, NJ. At Bell Labs, he studied design and processing of silicon VLSI, and the process and device development of optoelectronic integrated circuits (OEICs) for long-wavelength optical communications. In 1989, he returned to Korea to join the Department of Electronics Engineering and the Inter-university Semiconductor Research Center (ISRC), Seoul National University, where he is now a Full Professor in the School of Electrical Engineering. Since 2000, he has been the Director of Nano-Bioelectronics and System Research Center (NBS-ERC) which is funded by Korea Science and Engineering Foundation (KOSEF). His research interests are in areas of bioelectronics, bioinstrumentation, neural prosthesis, and optoelectronic semiconductor devices. He has published about 60 papers, and has managed many research projects in the areas mentioned. His web site is at http://helios.snu.ac.kr/.