

Neural Spike Sorting Under Nearly 0-dB Signal-to-Noise Ratio Using Nonlinear Energy Operator and Artificial Neural-Network Classifier

Kyung Hwan Kim and Sung June Kim*

Abstract—We report a result on neural spike sorting under conditions where the signal-to-noise ratio is very low. The use of nonlinear energy operator enables the detection of an action potential, even when the SNR is so poor that a typical amplitude thresholding method cannot be applied. The superior detection ability facilitates the collection of a training set under lower SNR than that of the methods which employ simple amplitude thresholding. Thus, the statistical characteristics of the input vectors can be better represented in the neural-network classifier. The trained neural-network classifiers yield the correct classification ratio higher than 90% when the SNR is as low as 1.2 (0.8 dB) when applied to data obtained from extracellular recording from *Aplysia* abdominal ganglia using a semiconductor microelectrode array.

Index Terms—Extracellular recording, neural-network classifier, neural spike sorting, nonlinear energy operator, signal-to-noise ratio.

I. INTRODUCTION

For purpose of studying information transmission within the nervous system, the extracellular recording of a neural signal is very useful, since it enables simultaneous monitoring of activities of several nearby neurons. Usually the recorded waveform includes action potentials from several cells which are physically near the electrode site and, thus, it should be classified into spike trains from each individual cell for further analysis in which the spike train is to be considered to be a point process [1], [2].

During the past three decades, various classification methods ranging from simple amplitude discrimination to a neural-network classifier have been applied to the issue of spike sorting [3]–[8], and some have been implemented in commercially available packages such as Spike2 (Cambridge Electronic Design Ltd., Cambridge, U.K.), and MAP system (Plexon Inc., Dallas, TX). It appears that when the signal-to-noise ratio (SNR) is sufficiently high, any of the suggested schemes are capable of yielding satisfactory results. Recently, by using a neural-network classifier, more than 80% correct classification has been reported when the SNR is higher than four [5]. However, in actual extracellular recording, the cases often arise where the SNR is much lower. Recent advances in electrode fabrication technology enables the simultaneous recording of many adjacent neurons, up to hundreds of channels. These electrode arrays would, however, be expected to suffer from higher impedance noise because of the reduced size of the electrode site. Furthermore, if one uses a semiconductor neural probe with on-chip CMOS circuitry, more noise problems arise, as the result of the noise characteristic of CMOS preamplifier. The noise characteristic of CMOS circuit is inferior to that of the typical instrumentation amplifiers that are used for extracellular recording, due to the high $1/f$ (flicker) noise level of the MOSFET [9], [20], [21].

Manuscript received May 3, 1999; revised May 30, 2000. This work was supported by the Ministry of Health and Welfare, Korea, under Grant HMP-98-E-1-0006. Asterisk indicates corresponding author.

K. H. Kim is with the School of Electrical Engineering, Seoul National University, Kwanak-gu, Seoul 151-742, Korea (e-mail: khkim@helios.snu.ac.kr).

*S. J. Kim is with the School of Electrical Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul 151-742, Korea (e-mail: kim@helios.snu.ac.kr).

Publisher Item Identifier S 0018-9294(00)08531-1.

A typical extracellular recording consists of a background noise floor and distinct spikes. However, closer inspection of the recording shows that it also contains spikes whose amplitude is similar to that of the background noise. As shown in Fig. 1, it is not unusual for the SNR to be as low as 0 dB. In previous studies, efforts have been largely exerted on the classification method to enhance classification accuracy, to enable real-time operation, or to resolve overlapped action potentials [4]–[6]. Less attention has been paid to the correct detection of action potential although this is critical for the neural spike sorting system. Bankman and Janssewitz have described procedures to elaborate the threshold determination procedure [11]. However, in cases where SNR is very poor as in Fig. 1, the correct detection of an action potential by thresholding is very difficult, irrespective of how well the threshold level is adjusted. Thus, methods which permit the utilization of information other than amplitude needs to be employed for spike detection.

Time-frequency analysis methods or matched filter detection represent possible solutions. Yang and Shamma have used the Haar transform [7] and Gozani and Miller have employed matched filter detection [8], but, neither of these studies have shown detection and classification performance tests under very low SNR. In addition, the matched filter detection requires *a priori* knowledge concerning the waveform of the target signal and the background noise and, thus, cannot be used in the initial data collection for the training set of the supervised classifier. Chandra and Optican implemented a neural-network-based method [5]. Their system employs the neural network for spike detection as well as for classification and, thus, the detection performance can be much better than that of the simple threshold method. However, for the initial data collection to obtain training data, they also used a thresholding method. Moreover, the output of neural network must be calculated for all data samples, and, as a result, it can be computationally demanding.

In this paper, we employ a nonlinear energy operator (NEO) to utilize the instantaneous frequency and amplitude information simultaneously for the detection of action potential for both training and classification stages. The output of the NEO is proportional to the product of the instantaneous amplitude and frequency of the input signal, thus, highlighting the action potential peak. Its computational burden is slight [12], [13] and, thus, it has another advantage considering real-time and multichannel operation. It can be applied to the training stage because no exact *a priori* knowledge about the waveform of the target signal is necessary.

We show the detection and classification performance of our method, as it is applied to experimental data obtained from abdominal ganglia of *Aplysia*. The detection of the neural spikes is possible when the SNR is very low, and this enables data collection and clustering to obtain the training set required for the supervised neural-network classifiers under low SNR. The training set thus obtained better represents the statistical characteristics of the input vectors than that obtained by simple thresholding. This is because accurate detection by thresholding is possible only for a relatively high SNR, but, in fact, the classifier can operate under lower SNRs. In the classification stage, by separating the spike detection and the classification module it is possible to achieve good performance under conditions of low SNR with lower computational requirements than is demanded by the method employed by Chandra and Optican [5].

For the classification, we used supervised classifiers because unsupervised classification was not effective under low SNR. We employed a multilayer perceptron (MLP) and radial basis function network (RBFN) as trainable classifiers. The performances of these neural-network classifiers are at least theoretically equivalent to that of

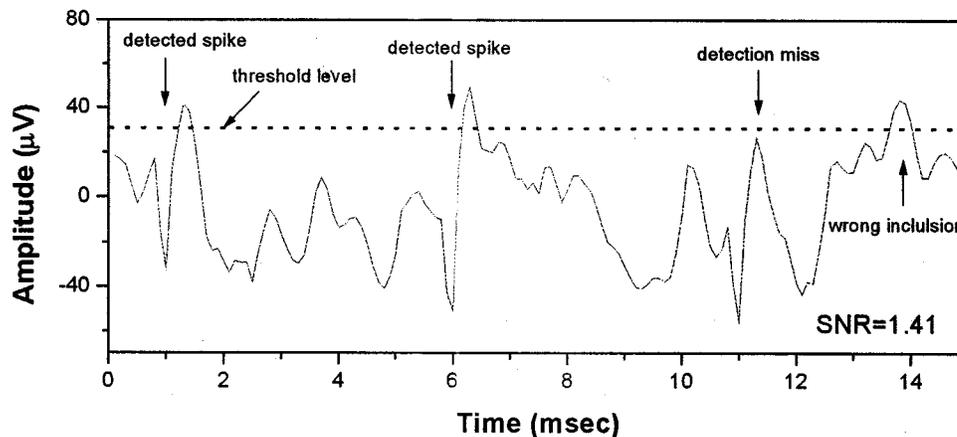


Fig. 1. A typical extracellular recording from an abdominal ganglion of *Aplysia* with a semiconductor microelectrode. The arrows indicate action potential firing. The problem of spike detection using amplitude (or power) thresholding is illustrated. An erroneous inclusion of noise peak or missing a neural spike is very frequent even when the threshold level has been carefully chosen.

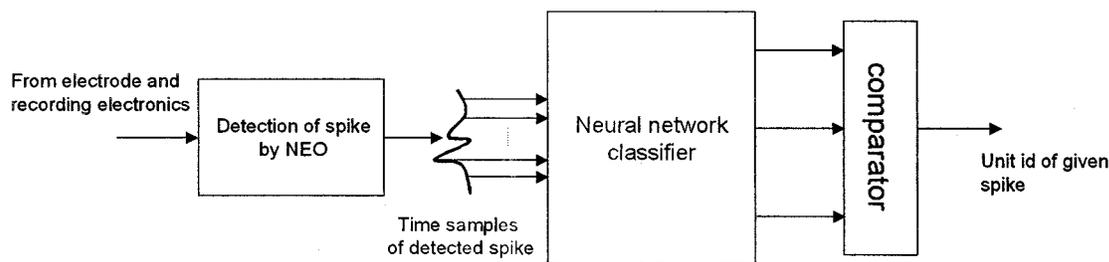


Fig. 2. Overall structure of the spike sorting system used in the present study.

the statistical methods [14]. In practice, it is possible to achieve better performance through their adaptive capabilities of repetitive learning underlying probability distribution of input pattern vectors [15].

II. METHODS

A. Experimental Methods

Aplysia Californica with body weights of 200–350 grams were selected. The animals were anesthetized with an injection of a volume of isotonic $MgCl_2$ equal to approximately one-half of their body weight, after which, they were mounted on a dissection dish which was filled with artificial sea water (ASW) (pH: 7.6, composition: NaCl 460 mM, KCl 10 mM, HEPES 10 mM, $CaCl_2$ 11 mM, and $MgCl_2$ 54 mM).

Recording experiments were performed in a Faraday cage attached to a vibration isolation table to reduce power line interference and artifacts due to vibration.

A semiconductor neural probe with one shank which contains five iridium (Ir) electrode sites were used. These were obtained from the Center for Neural Communication Technology, University of Michigan [10]. The areas of the Ir electrode sites were $75 \mu m^2$, and the impedance levels were 2–3 $M\Omega$ at 1 kHz. The probe was mounted on a custom-made socket and connected to a unity gain JFET buffer (DAM80P, World Precision Instruments, Sarasota, FL) which was attached to a micromanipulator.

After the abdomen was opened, connective tissue was removed, and ganglia were located by visual inspection under a stereo zoom microscope, the electrode was moved with the micromanipulator, so that the probe shank was located into the ganglia. A ground electrode was placed in the saline (i.e., ASW) about 10 cm distant from the target ganglion.

The output from the buffer stage was sent to a main amplifier with a bandpass filter (DAM80, World Precision Instruments, Sarasota, FL). The gain was 1000 and the passband was from 100 Hz to 5 kHz. After bandpass filtering, the signal was monitored with a digital oscilloscope (LeCroy 9304AM) the output of which was sent to an audio speaker, a personal computer via 12-bit 10000 samples/s analog-to-digital converter, and a digital data recorder with VCR (VR-10B, Instrutech Corp., Port Washington, NY). All experiments were performed at room temperature, with no special intentions to control the temperature.

B. Detection and Classification Methods

Fig. 2 shows the overall structure of the spike sorting system used in the studies. The sorting system takes the raw recording waveform as input, and processes it with the NEO, in order to extract the action potentials. The recording waveform is sent to the classifier only when the output of the NEO exceeds the predetermined threshold. Thus, the neural-network output is calculated only for the time samples of the detected action potential. A comparator after the output layer yields the most probable unit number given as input. The NEO, ψ , is defined as

$$\psi(x(t)) = \left(\frac{dx(t)}{dt} \right)^2 - x(t) \left(\frac{d^2x(t)}{dt^2} \right).$$

It can easily be shown that the output of the NEO is proportional to the product of the amplitude and frequency of the input signal [12]. For a discrete-time sequence $x(n)$, the NEO is given as $\psi(x(n)) = x^2(n) - x(n+1)x(n-1)$. The NEO has been used for the amplitude and frequency demodulation, the analysis of speech signals, and

recently for the detection of a spiky waveform in the electroencephalogram (EEG) [12], [13]. All of these studies have made use of the fact that the NEO can simultaneously consider the “instantaneous” amplitude and frequency information of the input signal. By the “instantaneous” amplitude and frequency, we mean the amplitude and frequency of the dominant sinusoidal component at any particular time. More formally, they can be defined by the Hilbert transform pair [22]. When an action potential is fired, it is possible to see an instantaneous increment of signal amplitude and frequency using time-frequency analysis such as short-time Fourier transform or Wigner–Ville distribution [16]. The output of the NEO is convolved with a Bartlett window in order to eliminate the spurious peaks due to the cross terms and background noise [16]. For the data being considered herein, we were able to consistently obtain satisfactory results for window lengths of 6–12.

For the determination of the detection threshold level, the level was manually adjusted based on data segments containing 20 action potentials, aiming at minimizing the erroneous inclusion of noise (or false alarm) and the detection miss. A statistically optimal determination of the threshold level was not attempted because this requires a complete knowledge of the probability distribution function (pdf). Although we do, in fact, have knowledge on the pdf for the test set, we attempted to devise and test a method which is generally applicable to common experimental recordings. For our *Aplysia* data, the threshold selection was not a difficult problem even for the low SNR as we show later (see Fig. 5). For the training stage, a high threshold level is desirable, in order to minimize false alarms (erroneous inclusions of noise).

Two representative neural networks, multilayer perceptron (MLP), and radial basis function network (RBFN) were employed as classifiers. The training procedure of the overall neural spike sorting system was as follows. To obtain initial training data, action potentials were detected by processing the raw extracellular recording, the SNR of which was approximately 2.0 with the NEO. The SNR is defined as shown at the bottom of the page. Each detected spike was aligned according to the peak value. After that, fuzzy c-means clustering [19] and further refinement (for example, removal of the outliers) by a human supervisor were performed, in order to obtain the waveforms of the initial training set. The training set consists of approximately 300 action potentials for each unit.

The classifiers take 25 sample points (2.5-ms segment) of the detected spikes as input, both in the training and operation stages. The number of the output nodes of the classifiers was same as the number of units to be identified. The classifiers were trained so that the output node which corresponds to the given input spike yielded the maximum output value. The MLPs with one hidden layer were trained by a modified backpropagation method with adaptive learning rate [18]. A sigmoidal activation function was employed for all layers. The number of the hidden nodes was determined experimentally while simultaneously considering the sum-squared-error and generalization performance on a test set which was different from the one used for training. We were able to obtain satisfactory results for 10–12 hidden nodes. During the training of the RBFN, the number of hidden layer nodes was increased until an error goal is satisfied. The centers of radial basis function were changed by error-correction learning [18].

Test sets for various SNRs were generated from an autoregressive moving average (ARMA) model of the background noise and the pre-

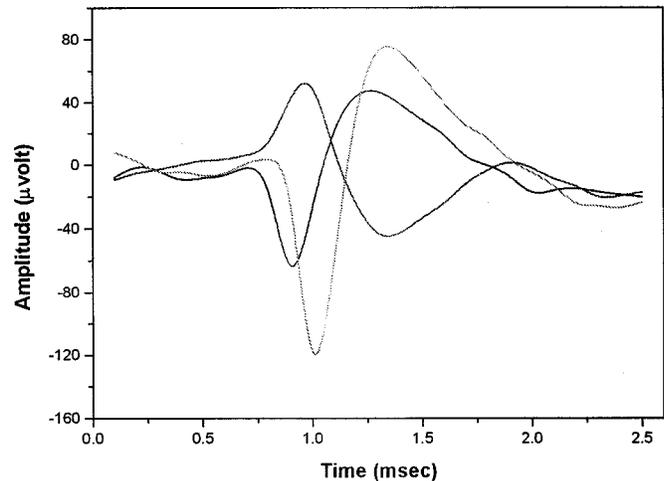


Fig. 3. Average waveforms of the initial training samples extracted from extracellular recording of *Aplysia* abdominal ganglion with semiconductor microelectrode array.

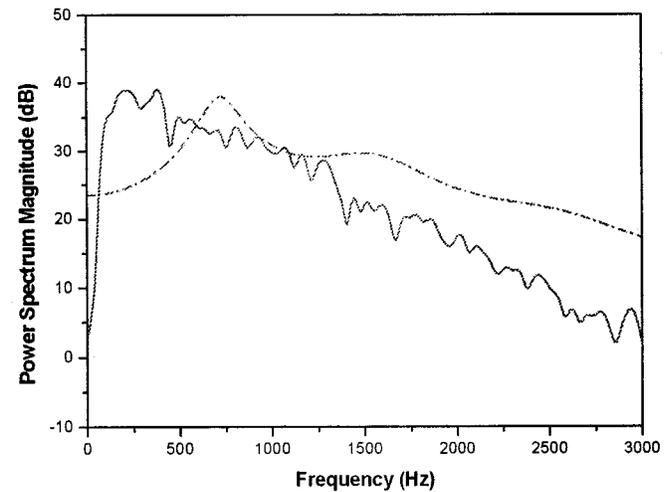


Fig. 4. Power spectral densities of the action potential with minimum amplitude shown in Fig. 3, and the background noise.

viously obtained initial training set. The ARMA(4,4) model could satisfactorily represent the background noise of our recording. The model coefficients were obtained by the extended Yule–Walker method from a noise segment of the data which had not been used to obtain the training set [17].

III. RESULTS

The overall sorting system was tested for three-unit classification. The average waveform of each unit extracted from the *Aplysia* recording is shown in Fig. 3. From the power spectral densities of the action potential and that of the background noise shown in Fig. 4, it is obvious that the neural spike and the background noise share the same spectral range. The signal shows a bandpass type spectrum, and the

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

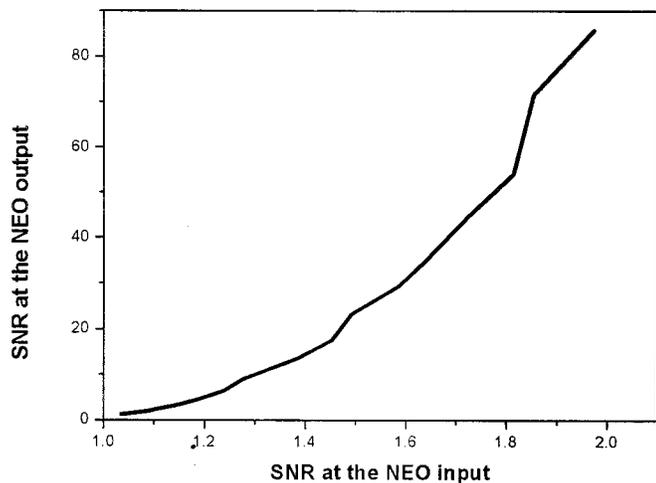
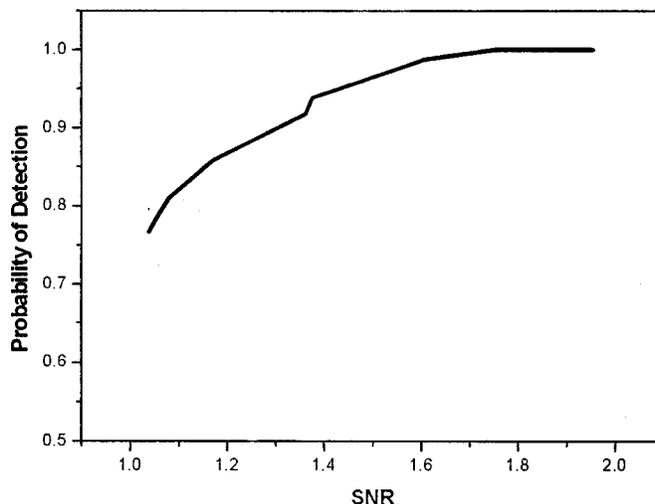
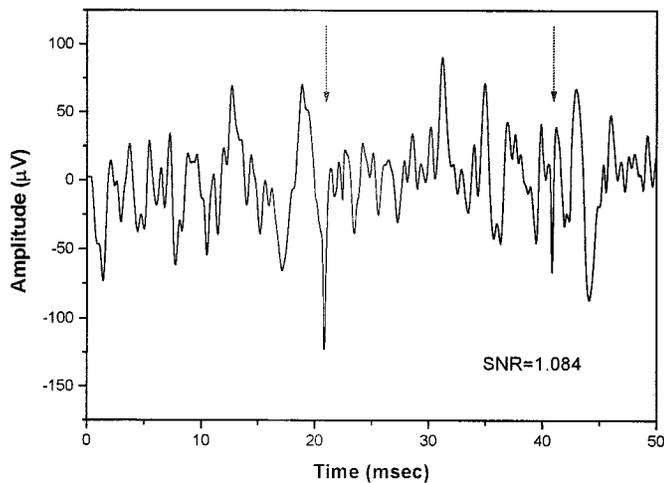


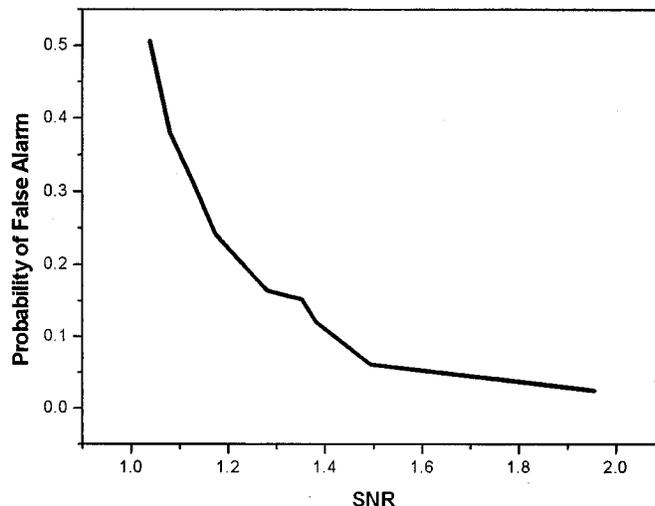
Fig. 5. SNR before and after the processing by the NEO.



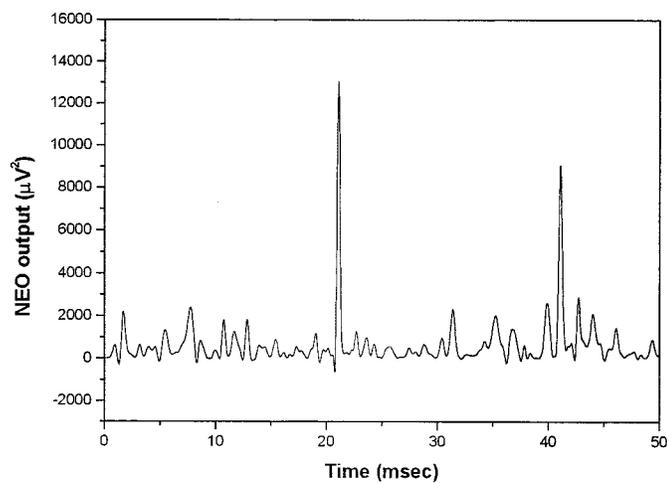
(a)



(a)



(a)



(b)

Fig. 6. (a) Input and (b) output waveform of the NEO spike detector when SNR = 1.084. The determination of the threshold level is now a trivial problem.

Fig. 7. (a) Probability of detection and (b) probability of false alarm for various values of SNR.

background noise shows a low-pass type spectrum having a 1/f-type shape with the effect of bandpass filtering. This spectral overlap of the signal and noise spectrum is common in the extracellular neural

signal recording. The performance testing was done for various levels of background noise having the spectral characteristic of Fig. 4.

Fig. 5 shows the enhancement in the SNR after the raw waveform is processed using the NEO. Fig. 6 shows the input and output waveform of the NEO detector when the SNR is 1.084. From these two figures it is clear that the determination of the threshold level becomes much more straightforward when the signal is processed using the NEO, even for the case of very low SNR. In order to investigate the detection performance of the NEO, the test result for the correct detection ratio and the false alarm (i.e., wrong inclusion of noise) ratio at various SNRs are shown in Fig. 7. The correct detection was identified when the NEO output exceeded the threshold level and the location of the peak value was within 0.5 ms from the true location of the action potential peak. A false alarm was indicated when the NEO output was higher than the threshold level where, in fact, no neural spike existed. These ratios were calculated by dividing the number of correct detections and false alarms by the total number of detected spikes including the false alarm. When the SNR is lowered to a value below 1.2, the false alarm ratio becomes higher than 0.2 and the correct detection ratio becomes lower than 0.9, and, as a result, it is possible that the classification performance deteriorates significantly.

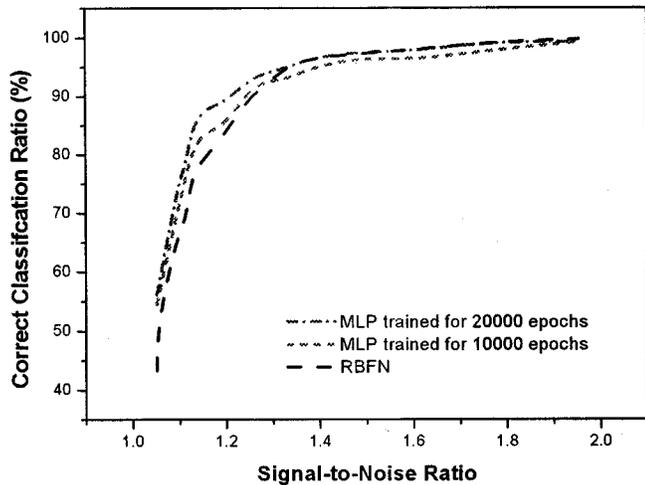


Fig. 8. CCR at various SNRs (See text for the details).

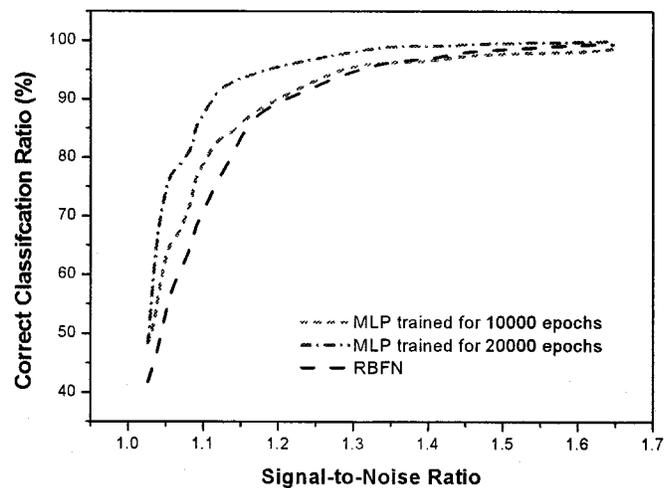


Fig. 10. CCR at various SNRs for the data shown in Fig. 9.

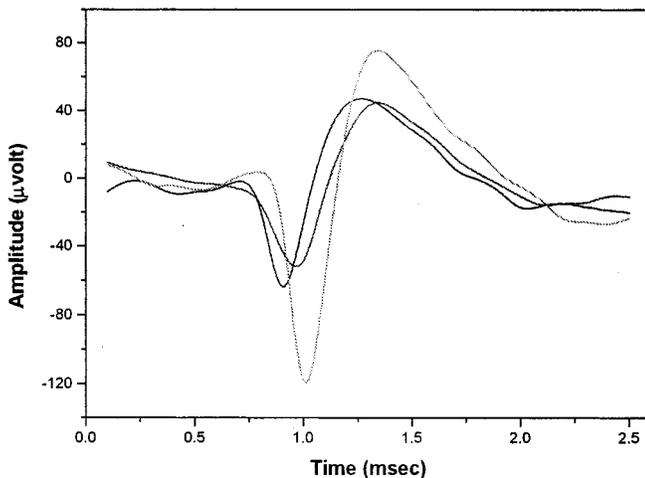


Fig. 9. Average waveforms of another data set in which two units among three have very similar target waveforms.

Overall classification test results for the MLP and RBFN under various SNR values are illustrated in Fig. 8. The result is for the MLP classifier trained for 10 000 and 20 000 epochs, since further training increased the performance only slightly. The RBFN classifier in Fig. 8 was trained so that the obtained sum-squared-error was the same as that of the MLP classifier trained for 30 000 epochs. Our performance measure was the correct classification ratio (CCR) which was defined as follows; $CCR = (\text{total number of correctly detected and classified spikes} - \text{number of the detection miss}) / (\text{total number of correctly detected spikes in the test data}) \times 100 (\%)$.

The CCR rapidly decreased for cases where the SNR was lower than 1.1, and the CCR was saturated to nearly 100% when the SNR was increased to levels higher than 1.4–1.5.

Another data set in which two units among the total three have very similar target waveforms were analyzed, in order to assess the discrimination ability of our method. In our recording, there is no evidence that the two similar waveforms in Fig. 9 arose from two different units, because the recording represented spontaneous activity from *Aplysia* abdominal ganglion. However, when each unit can be selectively stimulated, Mirfakhraei and Horch showed that it is possible to train a supervised neural-network classifier which is capable of discriminating such similar waveforms [6]. Our aim here was to test whether our method

can be applied to such a case. In order to obtain the training set, background noise having spectral characteristic shown in Fig. 4 were added to the mean waveforms shown in Fig. 9. The same method as above was applied for the test on this data. The classification performance for this case is shown in Fig. 10.

IV. DISCUSSION

Numerous studies on the classification of multiunit extracellular neural signal recording have been reported. To date, the results are satisfactory when the SNR is sufficiently high. The main issue of study has been moving toward the resolution of overlapped action potentials [4], [5]. However, in order to use spike sorting more extensively in real neurophysiological experiments, the SNR required for successful operation must be lowered further.

The classification performance of neural networks is so powerful that it can be used to solve pattern recognition problems which appear to be more difficult than spike sorting under low SNR, but this is valid only when an appropriate training set is available. Hence, it is predicted that the prerequisite for spike sorting under low SNR would be on efficient ways for extracting action potentials from high background noise. Spike detection is also important for the operation of the classifier after it is trained.

In general, signal detection is accomplished by making the target signal conspicuous while suppressing noise, and by applying the thresholding method [22]. In previous studies, less attention has been paid to spike detection than to classification. As a result, most of them applied the thresholding directly to the raw recording waveform. Exceptions involved the use of matched filter detection [8] and neural-network detection [5]. The matched filter provides a theoretically optimal linear filter under Gaussian noise, but assumes complete knowledge of the target waveform and the spectral characteristics of the noise. A detector based on a neural network also requires a knowledge on the target signal waveform and the background noise, for the training. Another shortcoming is in that, in the case of neural-network, output must be calculated for every input data samples, while in our method it is calculated only when the action potential is extracted. Our NEO detector is much faster and requires less information on the target signal. It takes advantages of the fact that the firing of action potential provides the temporal increase of the signal amplitude and frequency. In typical situations when the bandpass filtering is done, in order to reduce the noise components outside the bandwidth of the target signal, the NEO detection appears

to be efficient. However, when the spectral characteristics of the noise is so similar to that of the signal, the detection performance may be degraded. This problem may be severe in the case of cortical recordings where the electrical activities from many neurons which are not tightly coupled to the electrode become the dominant noise source. Novel methods for the detection need to be developed to deal with this problem.

Several studies have reported on spike sorting using an unsupervised classifier (for example, [23] and [24]), but none of them have verified its performance under low SNR. Our experiences with the unsupervised classifiers were not satisfactory for the case of low SNR. Thus, the unsupervised method was not considered in this paper. However, at least for the acquisition of initial training data set required for the supervised classifier, unsupervised clustering is necessary. In this regard, a study designed to assess the performances of the unsupervised classification methods under low SNR is necessary. In addition, a study of how to enable the resolution of overlapped action potentials, especially under low SNR condition, is also needed.

In conclusion, we have achieved neural spike sorting with a high success rate under conditions of very low SNR. This was due to the improved detection scheme in collaboration with the ability of supervised neural networks to learn the underlying probability distribution of given training examples.

ACKNOWLEDGMENT

The authors wish to thank Prof. B. G. Kang, the Institute for Molecular Biology and Genetics at Seoul National University for his help with the experimental methods.

REFERENCES

- [1] D. H. Perkel, G. L. Gerstein, and G. P. Moore, "Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains," *Biophys. J.*, vol. 7, pp. 419–440, 1967.
- [2] W. Bialek, R. de Ruyter van Steveninck, and D. Warland, "Reading a neural code," *Science*, vol. 252, pp. 1854–1857, 1991.
- [3] B. C. Wheeler and W. J. Heetderks, "A comparison of techniques for classification of multiple neural signals," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 752–759, Dec. 1982.
- [4] M. S. Lewicki, "A review of methods for spike sorting: The detection and classification of neural action potentials," *Network: Computation Neural Syst.*, vol. 9, pp. R53–R78, 1998.
- [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. Mirfakhraei and K. Horch, "Classification of action potentials in multi-unit intrafascicular recordings using neural network pattern-recognition techniques," *IEEE Trans. Biomed. Eng.*, vol. 41, pp. 89–91, Jan. 1994.
- [7] X. Yang and S. A. Shamma, "A totally automated system for the detection and classification of neural spikes," *IEEE Trans. Biomed. Eng.*, vol. 35, pp. 806–816, Oct. 1988.
- [8] 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.
- [9] J. Ji and K. D. Wise, "An implantable CMOS circuit interface for multiplexed microelectrode recording arrays," *IEEE J. Solid-State Circuits*, vol. 27, pp. 433–443, Mar. 1992.
- [10] S. L. BeMent, K. D. Wise, D. J. Anderson, K. Najafi, and K. L. Drake, "Solid-state electrodes for multichannel multiplexed intracortical neuronal recording," *IEEE Trans. Biomed. Eng.*, vol. BME-33, pp. 230–241, Feb. 1986.
- [11] I. N. Bankman and S. J. Janselowitz, "Neural waveform detector for prosthesis control," in *Proc. 17th Ann. Conf. IEEE EMBS*, 1995, pp. 963–964.
- [12] 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.
- [13] S. Mukhopadhyay and G. C. Ray, "A new interpretation of nonlinear energy operator and its efficacy in spike detection," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 180–187, Feb. 1998.
- [14] D. W. Ruck, S. K. Rogers, M. Kabrisky, M. E. Oxley, and B. W. Suter, "The multilayer perceptron as an approximation to Bayes optimal discriminant function," *IEEE Trans. Neural Network*, vol. 1, pp. 296–298, Dec. 1990.
- [15] L. Holmstrom, P. Koistinen, J. Laaksonen, and E. Oja, "Neural and statistical classifiers-Taxonomy and two case studies," *IEEE Trans. Neural Network*, vol. 8, pp. 5–17, Jan. 1997.
- [16] D. L. Jones and D. L. Parks, "A resolution comparison of several time-frequency representations," *IEEE Trans. Signal Processing*, vol. 40, pp. 413–420, Feb. 1992.
- [17] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*. New York: Wiley, 1996.
- [18] S. Haykin, *Neural Networks*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.
- [19] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.
- [20] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor*. New York: McGraw Hill, 1989.
- [21] K. H. Kim and S. J. Kim, "Noise performance design of CMOS preamplifier for the active semiconductor neural probe," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1097–1104, Aug. 2000.
- [22] C. W. Helstrom, *Elements of Signal Detection and Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
- [23] I. N. Bankman, K. O. Johnson, A. Menkes, S. D. Diamond, and D. O'Shaughnessy, "Automated analyzer for on-line recognition of neural waveforms in extracellular recordings of multiple neurons," in *Proc. Annu. Int. Conf. IEEE EMBS*, vol. 14, 1992, pp. 2852–2853.
- [24] F. Ohberg, H. Johansson, M. Bergenheim, J. Pedersen, and M. Djupsjobacka, "A neural network approach to real-time spike discrimination during simultaneous recording from several multi-unit nerve filaments," *J. Neurosci. Meth.*, vol. 64, pp. 181–187, 1996.