

Classification of Simple Stimuli Based on Detected Nerve Activity

THOMAS D. COATES JR.,
LINDA J. LARSON-PRIOR,
SETH WOLPERT, AND FRED PRIOR

*An Interface and
a Signal Processing
Methodology That
Use Measured Neural
Signals to Image
Overall Axonal
Activity in Intact
Peripheral Nerve*

The development of a direct, bidirectional interface between the human nervous system and a machine would radically alter the way we interact with computers and each other. The peripheral nervous system offers a number of advantages over the brain as the location for certain types of neural interfaces. These advantages include ease of accessibility and direct correspondence between external sensory stimuli and nerve activity.

A number of interfacing strategies have been proposed for connecting peripheral nerves to computers. One such strategy utilizes the limited ability of peripheral nerves to regenerate following transection and anastomosis. These “regeneration” interfaces [1] consist of a perforated substrate resembling a sieve with electrodes placed around some or all of the holes. The interface is placed between the ends of a cut peripheral nerve and the nerve is allowed to regenerate. Once nerve fibers have regrown through the holes of the interface the electrodes can be used to record and stimulate the fibers. Unfortunately, regenerative neural interfaces require that a nerve be transected to accomplish the interface. Regenerative interfaces may provide a good solution in cases where limb amputation has occurred. However, since one of the goals of this research is to create an interface that can eventually be used in both healthy and injured persons, an interfacing methodology that does not rely on nerve transection had to be developed. Other methods such as suction electrodes or patch clamp recording techniques can obtain single-fiber recordings with excellent signal-to-noise ratio (SNR) but are too bulky to be used on a large number of fibers simultaneously.

A compact and less-invasive alternative to these methods are nerve cuff electrodes [2]-[4]. The “cuff” consists of a short, flexible tube that has been slit longitudinally to allow it to be opened and placed around a nerve. Arranged around the inner circumference of the cuff are electrodes that detect the electric field produced by sources within the lumen. Cuff electrodes have a long history of use in the peripheral nervous system for both recording and stimulation [3], [5]-[10]. A self-sizing spiral cuff electrode has recently been used in humans for implementing the stimulation interface of a prototype visual prosthesis [4]. When fabricated from medical-grade silicone, elastomer cuff electrodes can provide a mechanically flexible and biologically well-tolerated platform for deploying a recording/stimulation electrode array.

Multiple electrodes can be used to obtain different versions of the evoked, bulk electrical activity. That is, each electrode would see the signals originating from sources closest to it as being larger in amplitude than those further away. Hence, the signal detected by one electrode would consist of the same sources mixed together in different amounts than the signal detected by another electrode in a different position. If differential pairs of electrodes are used, then signals closest to

the negative electrode would appear inverted while those near the positive electrode would not be inverted. To be useful as a neural interface, complex electrical activity recorded from a nerve or other neural source must be interpreted. In the work reported here, a cuff electrode containing multiple pairs of differential detectors was used to explore the feasibility of using measured neural signals to image overall axonal activity in intact peripheral nerve.

MINIS: Minimally Invasive Neural Interfacing System

MINIS consists of four parts: an in vivo multielectrode nerve cuff placed around an intact ensheathed whole nerve, wavelet based signal processing, information-theoretic data summarization, and a cascade correlation neural network. The system was validated using the visual system of *Limulus polyphemus* (common horseshoe crab). In our application the implantation of the cuff electrode requires surgery to expose the nerve but does not require removal of the sheath and surrounding connective tissue, hence the term “minimally invasive.”

Limulus

Limulus is a good experimental system to demonstrate the efficacy of the MINIS hardware/software as it has been the subject of intense study for over half a century. The body of work pertaining to *Limulus*, and specifically the *Limulus* visual system, is vast. Additionally, the lateral eye and lateral eye optic nerve are easily accessible for surgeries and specimens are hardy and easy to maintain prior to use. *Limulus* has a highly developed compound eye which, as in mammals, uses lateral inhibition to perform edge enhancement [11]. One visual unit (ommatidium) consists of one lens, several photoreceptors, and a neuron whose axon connects the visual unit to the brain and to neighboring visual units (Figure 1).

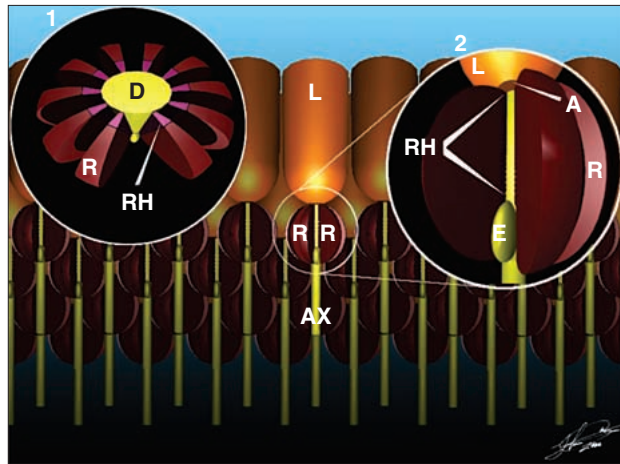


Fig. 1. Diagram of the receptor surface of the lateral eye. L = cylindrical lens, A = aperture, R = reticular cells, RH = rhabdomeres, E = neuron cell body, D = dendrite, AX = axon. Light enters from the top of the illustration, travels down the lens, and strikes the photosensitive cells surrounding the dendrite. Inset 1 shows the radial symmetry of the photoreceptor complex for one visual unit (lens removed). Inset 2 is a closeup of the region of the eccentric cell. One reticular cell on each photoreceptor complex has been removed to allow visualization of the eccentric cell body. Pigment-producing cells and lateral connections are not shown.

A pattern of light incident on the surface of the eye is represented by the rate/timing of firing of the visual units illuminated by the given portion of the light/dark pattern. The coding of the spatial image into the neural representation is defined by a series of equations originally elucidated by Hartline [12] and greatly expanded upon by others [13]–[15]. Thus, a visual scene is coded by the parallel activity of all the visual units in the lateral eye. The mapping of the visual units to the fibers of the optic nerve is stereotypical. Axons from horizontal strips of units contribute to the bundles that form the optic nerve. The packaging of the bundles in the optic nerve replicates the spatial relationship between the strips (five to seven on average) of visual units on the eye (Figure 2). The placement of the fibers in the bundles also preserves the position of the unit in the horizontal strip [16].

Thus, the *Limulus* visual system is well characterized both in terms of the anatomy (mapping of visual units to fibers) and its coding properties, making it an ideal testbed for MINIS.

Stimuli

Stimuli (Table 1) consisted of 11 light patterns and one control pattern (no illumination) delivered to the surface of the lateral eye by eight fiber-optic cables. One end of each fiber was connected to an LED that was driven by one line of an 8-bit digital port. The other end of the fiber was held in place perpendicular to the surface of the eye by a hood constructed of opaqued beeswax. The peak emission wavelength of the LEDs was 520–525 nm and was matched to the peak spectral response of the photoreceptors in the lateral eye [17]. Illumination at the surface of the eye by one optic fiber produces a vigorous response in a small group of visual units that in turn produce focal activity in the optic nerve.

Stimulus patterns were presented in a set that began with pattern 1 and ended with pattern 12 (Table 1). Each set took approximately 4 minutes to complete and was presented 100 times (6.7 hours per experiment). For data processing purposes, all the presentations of a given stimulus pattern are referred to as a trial. Thus, a trial consists of 100 presentations of a given pattern and there are 12 trials per experiment, one for each pattern. Individual stimulus patterns were presented for

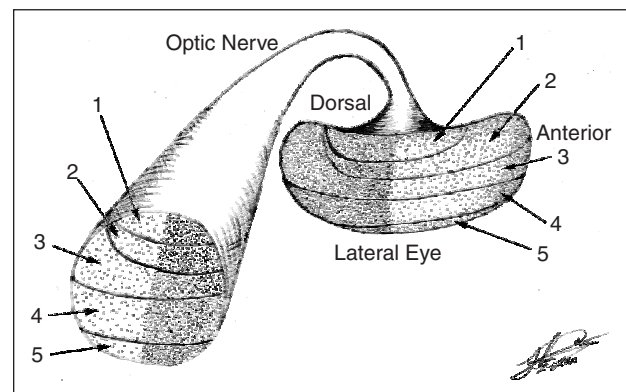
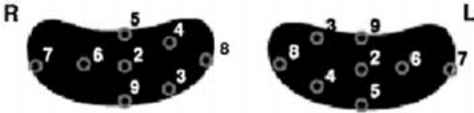
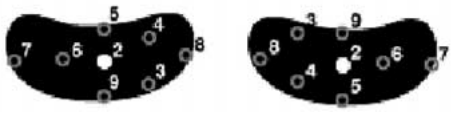
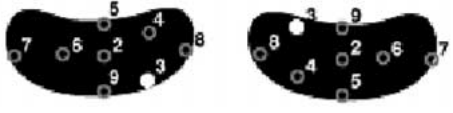
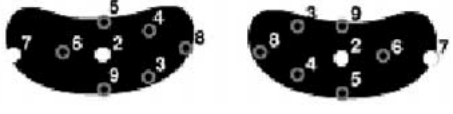
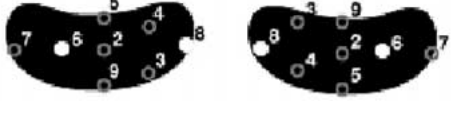
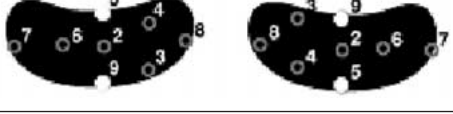


Fig. 2. Illustration showing the general mapping scheme of horizontal strips of visual units to their corresponding fibers located in the optic nerve. The shading present on the left half of the lateral eye and right half of the optic nerve is to indicate that the anterior/posterior mapping of units is also preserved. The visual unit density (units/mm²) is greatest along the center horizontal strips (3 and 4) and decreases at the top and bottom (1 and 5) of the eye.

The trained network for a given specimen was very specific to the specimen-interface-nerve configuration on which the data used to build the training/testing sets originated.

1.3 s followed by an interstimulus interval of 18.7 s. Each presentation generated four data records each comprising 32,768 sample points. That is, there was one record for each of the four differential electrode pairs in the cuff.

TABLE 1. Stimulus patterns. The outline of the lateral eye is shown in black with the stimulus areas superimposed. The areas are numbered from 2 to 9 to correspond to the eight single location stimulus patterns. Pattern 1 was the control (no stimulus). In patterns 10, 11, and 12 two locations were illuminated simultaneously. The 12 patterns are numbered top to bottom, left to right beginning with 1. Patterns for both the left and right eyes are shown in each cell of the table; however, only one eye was used in any given experiment. Only the first and last three patterns of the 12 are shown for space considerations.

1	
2	
3	
10	
11	
12	

Signal Acquisition

Our cuff electrode design (Figure 3) was based on an electrode layout consisting of four differential electrodes symmetrically arranged around the inner circumference of a short piece of silicone tubing. The active areas of the positive and negative poles of the differential pairs were located approximately opposite to one another to insure that they were recording through a large cross-section of the nerve. There was enough deviation in positioning of the electrodes to insure that the zero point in the centers of the differential pairs did not coincide. Had this been the case it would have created a recording "dead spot" in the center of the nerve.

The signals detected by the cuff electrodes were buffered by a unity gain headstage (Neuralynx EIB-27, impedance $10^{13} \Omega$) then passed to a differential amplifier (Neuralynx Lynx-8) with a gain of 50,000. Analog bandpass filtering was applied from 10 Hz to 9 kHz to eliminate unwanted dc components and to prevent aliasing when the signals were digitized. The four channels of data were acquired into the computer using a 12-bit A/D card (National Instruments AT-MIO-16F5) at a sampling rate of 25 kHz per channel. Data acquisition began approximately 150 ms prior to the presentation of the stimulus and continued for 1.31 s yielding 131K ($32K \times 4$) points of data per presentation.

Signal Processing

Signals were processed offline using the discrete wavelet transform (DWT) [18]. The DWT has the ability to utilize a number of different kernels (basis functions) that enable better matching of the filter with the features of the signal to be extracted. Several possible choices for the basis functions resemble action potentials (Figure 4) and produce better results than other methods for denoising neural data.

Processing the 32K data records with the DWT returns 15 levels of coefficients. The lowest levels contain coefficients that describe the dc components, 60-Hz noise, etc. The highest levels contain coefficients for the high-frequency noise portion of the signal. By throwing out (zeroing) coefficients and then reconstructing the signal, various types of filtering operations can be performed. We chose to reconstruct our signal using a Daubechies 4 wavelet basis and the coefficients from DWT level 10 resulting in a signal that is band limited to approximately 300-900 Hz. This band encompasses the frequency of the primary lobe of the action potential, which is approximately 500-600 Hz based on spectral analysis of the unprocessed data.

The SNR of MINIS was determined using the method illustrated in Figure 5. The beginning of every record contains approximately 150 ms of noise. This period is the latency between when the stimulus is presented to the Limulus lateral eye and

when the resulting activity is observed in the recording of the optic nerve activity. The data contained within this latency period (Figure 5) is the baseline noise measurement used in the SNR calculation for the given record. Following the latency period there is approximately 250 ms of large amplitude transient response (Figure 5) before sensory adaptation reduces the amplitude of the activity to the steady-state level seen in the remainder of the record. The activity during the transient response period is used for the “signal” part of the SNR measurement since it can be easily identified in both filtered and unfiltered data. The peak-to-peak voltage for both the latency and the transient response period of the record being considered is found using (1):

$$V_{pp} = |\max| + |\min| \quad (1)$$

where V_{pp} is the peak-to-peak voltage within the region of interest (ROI) of the signal, $|\max|$ is the absolute value of the maximum voltage measured in the ROI, and $|\min|$ is the absolute value of the minimum voltage measured in the ROI. V_{pp} was converted to root mean square (RMS) prior to computing the SNR (2) for the given data:

$$\text{SNR} = \frac{T_{\text{vrms}}}{L_{\text{vrms}}} \quad (2)$$

where T_{vrms} is the RMS voltage measured during the transient response period and L_{vrms} is the RMS voltage measured during the latency period preceding the transient response.

The improvement in the SNR following wavelet-based filtering is found using (3):

$$I_{\text{SNR}} = 20 * \log \frac{B}{A} \quad (3)$$

where I_{SNR} is the improvement in decibels of the SNR following signal processing, B is the SNR before processing, and A is the SNR after processing.

Information-Theoretic Data Summarization

Initial data analysis utilized amplitude histograms built from the signal reconstructed from wavelet level 10. Amplitude histograms were converted to probability densities (PDs) using (4):

$$P_i = \frac{n_i}{N} \quad (4)$$

where P_i is the approximate probability density for the given voltage bin in the amplitude histogram, n_i is the number of counts in the given voltage bin, and N is the total number of counts in the histogram.

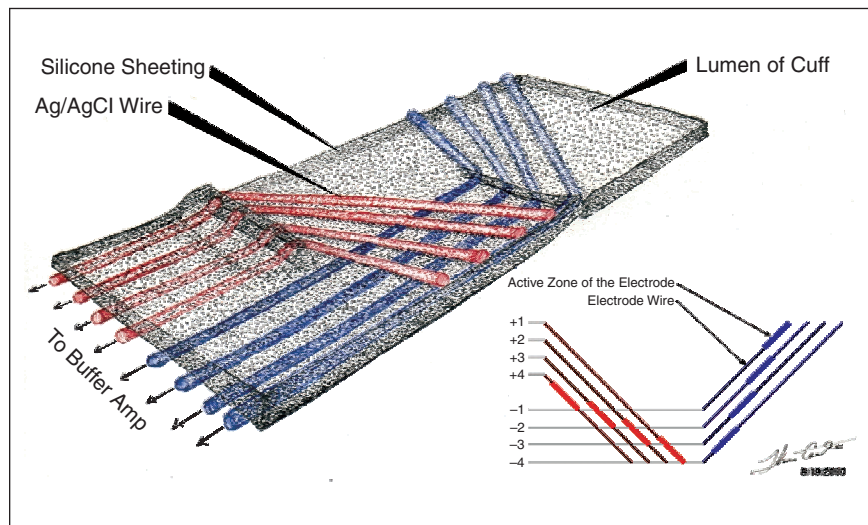


Fig. 3. Our four-pair differential cuff electrode was formed from 0.076-mm diameter silver wire and 0.356-mm thick silicone sheeting. The wires that formed the electrodes passed through slits in the silicone sheeting and traversed the top of the sheet at approximately a 45° angle. The electrodes were adhered to the sheet and insulated from the extracellular environment by a thin layer of silicone elastomer (not shown). The active area (zone) of an electrode was the portion that was in direct electrical contact with the nerve sheath. To form the active zones indicated in the inset above, the elastomer was selectively removed using microdissecting scissors and forceps. The sheet was then rolled to form a tube (electrodes on the inside) that was then placed around the intact nerve.

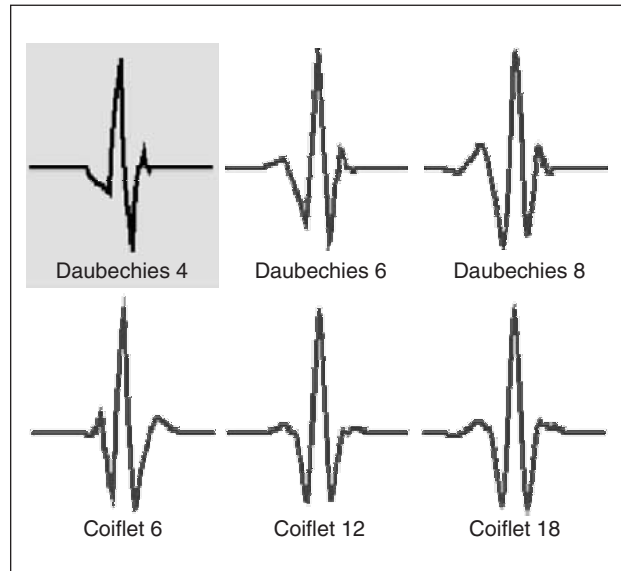


Fig. 4. Members from both the Daubechies and Coiflet wavelet families. All of these wavelets have characteristics in common with an action potential: sharp voltage transitions and a large positive peak in the center followed by a smaller negative peak. However, of these six choices the Daubechies 4 most resembles an action potential because of the relative sizes of its major peaks. The D4 wavelet basis was the one used in this research for denoising recorded neural data.

The PD of a discretely sampled distribution is an approximation of the area under the probability density function. By converting the amplitude distribution to the PD, information-theoretic measures can be used to further analyze the data.

Entropy (5) is a measure of the complexity of a given system [19]. In information theory, entropy relates to the minimum number of bits required to code a data set of a given complexity. The more complex the data the more information it contains and the more bits that are needed for encoding. We computed the entropy for the PD of our data set using (5):

$$H(X) = -\sum_{x=1}^N p(x) \log_2 p(x) \quad (5)$$

where $H(X)$ is the entropy of the random variable x expressed in bits, N is the number of intervals (bins), and $p(x)$ is the probability density for the given interval.

The Kullback-Leibler distance [(6) and (7)] is used to determine how easily two distributions can be distinguished from one another. The formula is similar to that of (5) except the \log_2 of the ratio of the distributions being compared is used:

$$D(p||q) = \sum_{x=1}^N p(x) \log_2 \frac{p(x)}{q(x)} \quad (6)$$

$$D(q||p) = \sum_{x=1}^N q(x) \log_2 \frac{q(x)}{p(x)} \quad (7)$$

where $D(p||q)$ and $D(q||p)$ are the Kullback-Leibler distances between probability densities p and q , N is the number of intervals (bins), and $p(x)$ and $q(x)$ are the probability densities on identical intervals for the two distributions being compared.

It is important to note that the Kullback-Leibler (K-L) distance is not symmetrical. That is, the distance from p to q is not equal to the distance from q to p . To compensate for the asymmetry the J-Divergence is computed (8) and used for comparison. The J-Divergence is computed by taking the mean of the K-L distances from p to q and from q to p for the two distributions:

$$J_{(p,q)} = \frac{D(p||q) + D(q||p)}{2} \quad (8)$$

where $J(p,q)$ is the J-Divergence for the distributions p and q , $D(p||q)$ is the K-L distance from p to q , and $D(q||p)$ is the K-L distance from q to p .

$H(X)$ provides a compact descriptor of the quantity of activity detected by a given electrode pair. Since the wavelet-based processing step band-limited the signal, $H(X)$ described the quantity of activity occurring at a given time and within the given range of frequencies. Little or no activity resulted in a narrow distribution with a low $H(X)$ (Figure 6, left). Maximal activity resulted in a broad distribution with a high $H(X)$ (Figure 6, right). $J(p,q)$ is useful as an early indicator of how well the neural

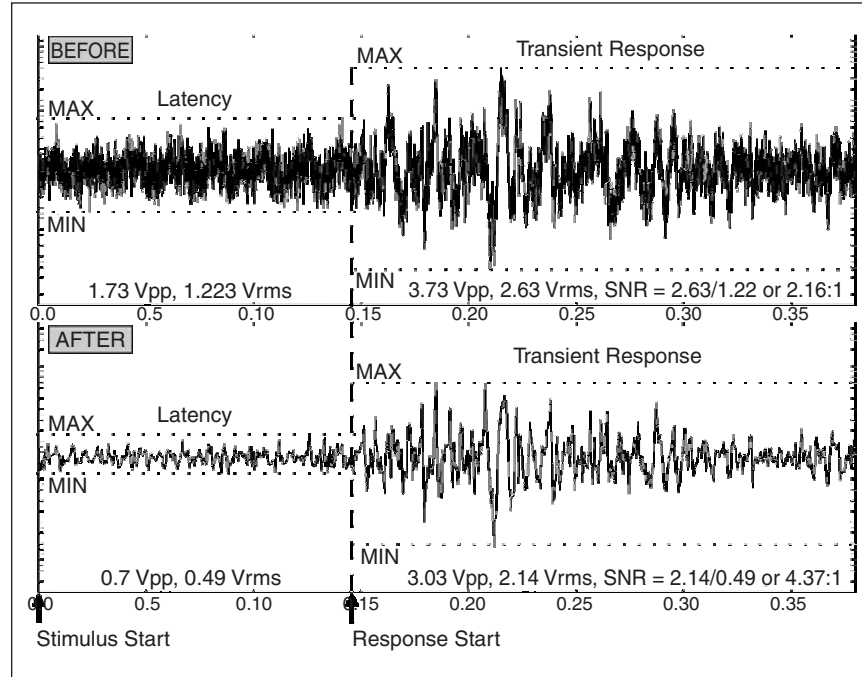


Fig. 5. Measuring the signal-to-noise ratio before and after signal processing. The top plot shows measurements on the signal before filtering, and the bottom plot shows measurements on the same waveform following wavelet-based filtering. The stimulus is presented to the lateral eye and recording begins at time 0 (stimulus start). Approximately 150 ms later the response is observed in the recording (response start). The maximum and minimum values used in the SNR calculations for this record are denoted by labeled dotted lines. The measurements for the latency and transient response periods are shown beneath the given segment.

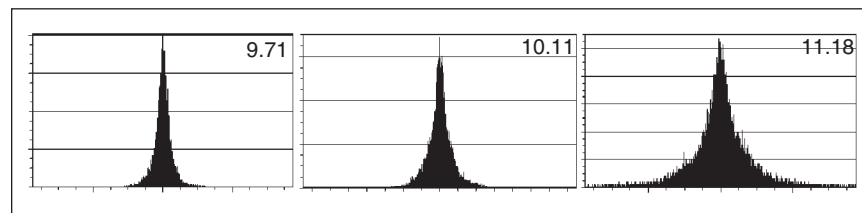


Fig. 6. Example demonstrating the relationship between the probability density $H(X)$ and $J(p,q)$ for band-limited neural data from a control recording (left) and two different stimulation patterns (center and right). The data shown is for a single electrode pair. The $H(X)$ measure for the given distribution is shown in the top right corner of each plot. The $J(p,q)$ for the center and left densities (compared to the control) were 0.36 bits and 2.03 bits, respectively.

When the network becomes overfitted it performs increasingly well at identifying the activity that corresponds to the data on which it was trained while becoming worse with novel data.

network would perform on its task. A low value for $J(p, q)$ for a given combination of patterns often indicated difficulty learning to distinguish those patterns from one another. Higher values indicated greater probability of success at learning to distinguish individual patterns (Figure 6).

$J(p, q)$ was also used to evaluate the overall quality of the recordings. Low values for $J(p, q)$, when comparing recordings for stimulated activity to controls, indicate data that are increasingly difficult to distinguish from background noise. In such cases the neural net will produce less accurate results or completely fail at the recognition task.

The difference in the mean J-Divergence (DMJD) provided an accurate predictor of the performance of the neural network during pattern classification. The DMJD was computed by first finding the $J(p, q)$ between the probability density of the data $p(x)$ produced by each stimulus pattern and that of the control $q(x)$. Thus, one cycle of 12 patterns produced 44 values for $J(p, q)$ (11 patterns compared to the control * 4 channels of data). Then the absolute value of the difference between all the nonredundant combinations of the 44 values was found, yielding four tables.

The mean of corresponding cells in the four tables was computed, yielding a single table (see Table 4). The rows and columns of the DMJD matrix correspond to those of the score matrices (see Table 3).

TABLE 2. Results of the SNR calculations for all specimens. The results are based on the analysis of record 4, presentation 1, trial 8 of the first experiment for each specimen.

Specimen	Before	After	Δ dB
1	1.84	2.75	3.48
2	2.23	2.5	1.00
3	1.99	3.28	4.31
4	2.16	4.37	6.05
5	1.30	2.92	7.06
Mean	1.90	3.15	4.38

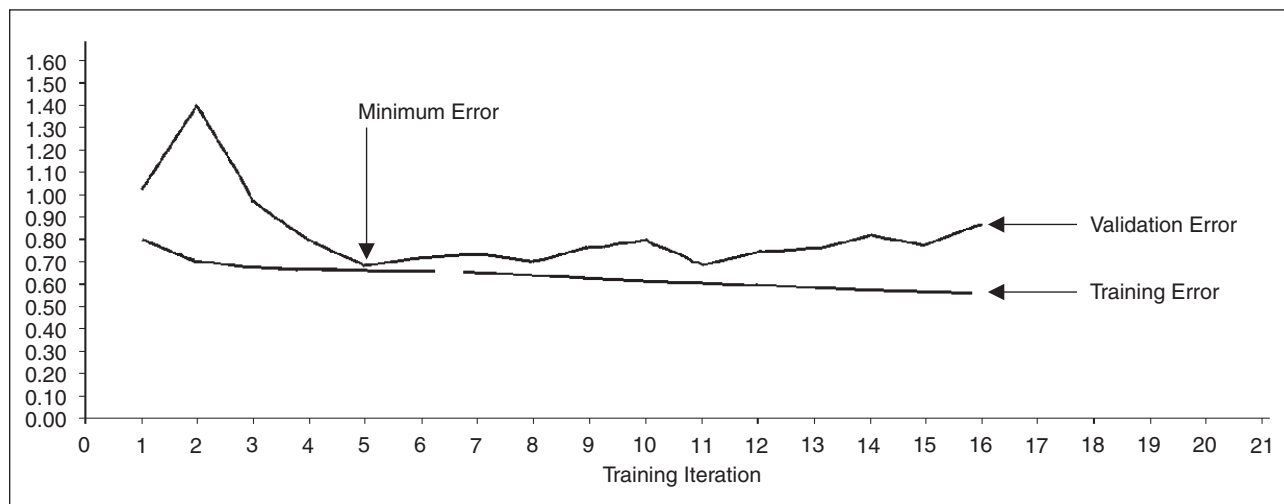


Fig. 7. Training the cascade correlation neural network. The top line on the plot is the mean square error (MSE) for the network during forward operation with the validation set. The bottom line is the MSE for the network during training using the training set. Since the network learns only the training set, the validation set provides an independent measure of the network's ability to generalize what it is learning to other data. The point labeled "minimum error" represents the optimum training point where the network will best be able to abstract what it has learned to the test data. Continuing training past this point results in an overfitted network that will perform better on the training data but progressively worse on the validation or testing data as training proceeds. This can be seen on the plot as divergence in the training and validation error curves as the network error for the training set falls and the validation set error rises.

Cascade Correlation Neural Network

The Stuttgart Neural Network Simulator (SNNS) [21] was used to evaluate a number of network types including time delay [20][21] and radial basis function neural networks [21], [22]. Ultimately, a back-propagation neural network [23] was constructed using the cascade correlation algorithm [24] and trained using the quickprop [25] learning rule (9). The primary selection criteria were accuracy of the results, speed, and ease of use:

$$\Delta(t+1)w_{ij} = \frac{S(t+1)}{S(t) - S(t+1)} \Delta(t)w_{ij} \quad (9)$$

where w_{ij} is the weight between units i and j , $\Delta(t+1)$ is the actual weight change, $S(t+1)$ is the partial derivative of the error function w_{ij} , and $S(t)$ is the previous partial derivative.

Cascade correlation (CC) is used to build and train optimal multilayer networks. CC begins with a minimal network consisting of an input and output layer. By creating a hidden layer and adding new units, it minimizes the overall error of the network until it falls below a predetermined value. CC is often referred to as a meta-algorithm rather than a learning rule since it directs the addition of hidden units and monitors the learning process but calls a preselected learning algorithm to actually train the net. (For a complete discussion of the mathematical background the reader is directed to the seminal paper by Scott

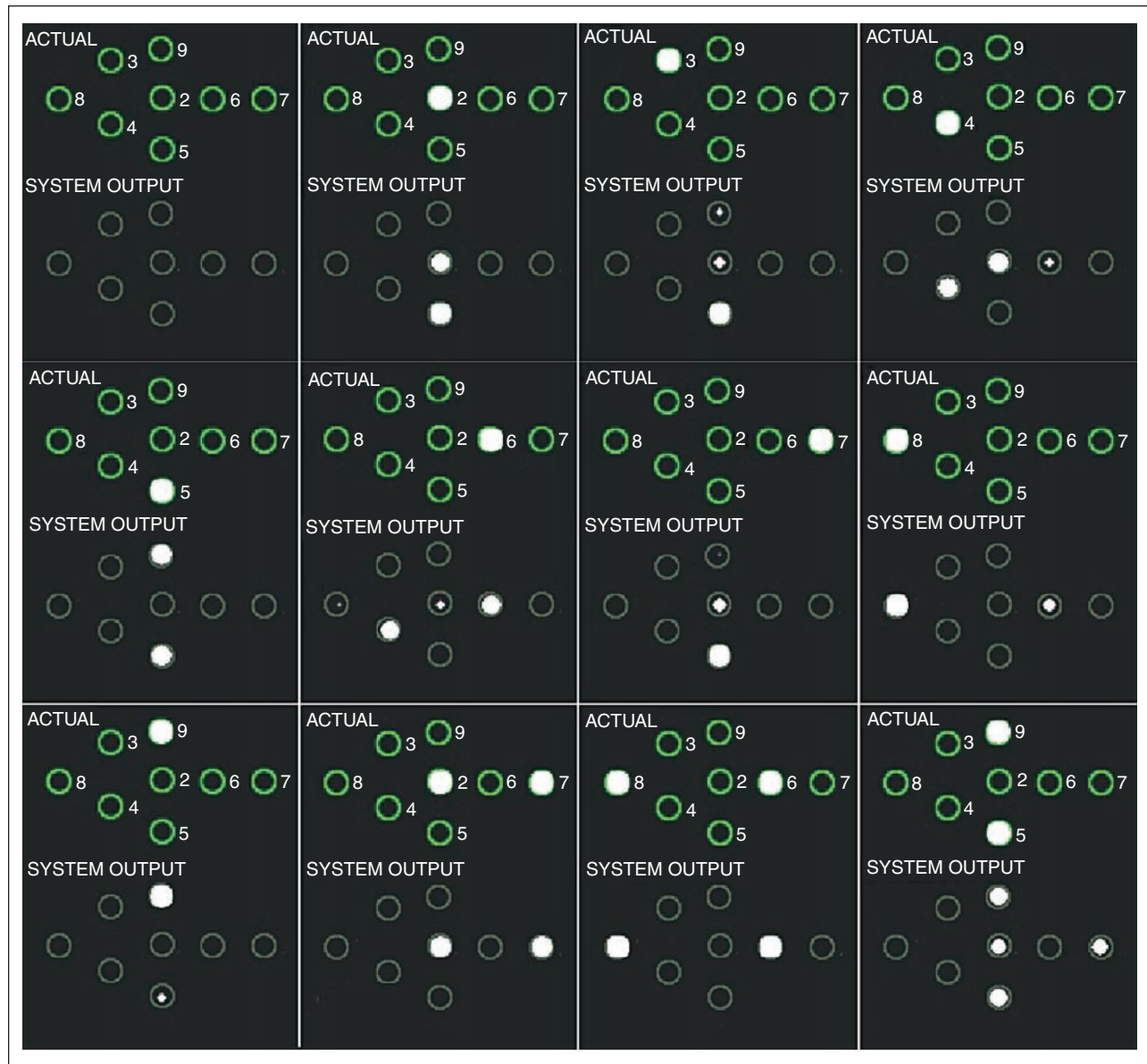


Fig. 8. Twelve frames of animation showing one sequence of patterns presented to the lateral eye. The patterns/frames are numbered left to right, top to bottom. Thus, pattern 1 (no stimulus) appears in the top left corner of the picture while pattern 12 appears in the bottom right corner. The portion of each frame marked "ACTUAL" indicates the stimulation pattern presented to the lateral eye. The "SYSTEM OUTPUT" section indicates what MINIS said the stimulation pattern was. The intensity of the point in the system output section indicates the relative certainty of the system that the given point was illuminated by the stimulus.

Fahlman [24]. The implementation and operation of the cascade correlation algorithm in SNNS is discussed in detail in the freely available SNNS users guide [21].)

The final network topology for this application had four inputs and eight outputs. One input pattern for the ANN consisted of four $H(X)$ values, one for each channel of data acquired from each of four electrode pairs in the cuff. The size of the hidden layer was solely determined by the CC algorithm and ranged from five to 18 units. Each of the eight outputs corresponded to one of the eight stimulus locations (Table 1). All the layers of the network were fully connected.

To minimize the chance of overfitting the network, each cycle of training was followed by a cycle of testing with an independent data set (termed the validation set). The network was saved prior to each new cycle of training. During initial training the error between the training and validation sets decreased. As training progressed the error between the set would begin to increase again (Figure 7). The network configuration prior to the point of increase is the configuration that was used.

Scoring

There was one output from the neural network for each of the 11 stimulated conditions. The output of the neural network for

each stimulus presentation was scored for accuracy using the following method. First, a threshold value was selected below which a given output was considered zero. This was necessary since the double precision floating point output of the network was rarely 0, but zeros on the outputs were needed to indicate pattern 1 (no stimulus). If there were outputs that were nonzero following thresholding, the outputs were evaluated by matching the highest valued 1 (for single location stimuli) or 2 (dual location stimuli) outputs to the actual stimulus location(s). The scores reported in Table 3 (top and bottom) are in percentages of the total presentations evaluated for the given trial.

Results

Five adult specimens of *Limulus polyphemus* were used in this study. An average of three studies per animal were performed. Data exhibiting low SNR were excluded from analysis. The results discussed in this article are based on analysis of data originating from seven experiments conducted with specimens 3, 4, and 5.

The output of the neural network simulator maps to an image of the stimulus pattern. Cinematic images (Figure 8) could be generated for each experiment to provide a qualitative measure of performance. However, analysis of

TABLE 3. (Top) Cumulative scores from seven experiments (329 presentations of each pattern) using neural data. (Bottom) Cumulative scores from experiments using Poisson noise (500 presentations per pattern).

	1	2	3	4	5	6	7	8	9	10	11	12	DK
1	100	0	0	0	0	0	0	0	0	—	—	—	0
2	0	25	12	0	33	0	13	1	17	—	—	—	0
3	0	13	0	0	33	1	15	2	36	—	—	—	0
4	0	27	4	32	3	11	18	3	4	—	—	—	0
5	0	2	0	0	56	0	2	0	38	—	—	—	2
6	0	23	0	7	1	38	7	24	0	—	—	—	0
7	0	12	1	0	31	2	23	0	32	—	—	—	0
8	0	12	0	1	0	24	1	62	0	—	—	—	0
9	0	0	0	0	9	0	0	0	91	—	—	—	0
10	—	—	—	—	—	—	—	—	—	47	5	1	46
11	—	—	—	—	—	—	—	—	—	1	97	0	3
12	—	—	—	—	—	—	—	—	—	8	0	55	36
	1	2	3	4	5	6	7	8	9	10	11	12	DK
1	18	6	8	8	10	13	17	8	12	—	—	—	0
2	16	12	5	8	15	8	14	10	11	—	—	—	0
3	17	9	6	6	11	12	18	9	12	—	—	—	0
4	18	12	12	4	13	10	15	10	6	—	—	—	0
5	18	10	10	6	11	9	17	8	10	—	—	—	0
6	26	10	11	4	9	11	15	6	8	—	—	—	0
7	22	8	10	6	11	11	15	9	8	—	—	—	0
8	22	10	9	4	6	12	20	9	8	—	—	—	0
9	24	8	11	5	11	9	15	10	8	—	—	—	0
10	—	—	—	—	—	—	—	—	—	8	10	10	72
11	—	—	—	—	—	—	—	—	—	7	7	17	69
12	—	—	—	—	—	—	—	—	—	9	7	15	69

the cumulative scores (Table 3, top) proved to be more useful in determining the overall performance of the system at its task.

Each experiment generated a score matrix organized as shown in Table 3 (top and bottom). The leftmost column in each table contains the number of the desired output pattern

(i.e., what the system's answer should have been). The top-most row in each table contains the pattern number for the actual output from the system (i.e., the system's answer). For example, looking at row 2 in the top table we would read "When stimulus pattern 2 was presented, the system said pattern one for 0% of the total presentations, pattern two for 25% of the total, etc." The highlighted diagonal delineates the position in which correct classifications are located, and numbers on either side of the diagonal are misclassifications. The last column labeled "DK" indicates the percentage of the data that could not be classified as belonging to one of the 12 patterns. It is important to note that the results are arranged in rows, the values in a given row sum to 100%, and the table is not row-column invertible. The score matrices from the seven experiments were averaged to produce the values shown in Table 3 (top).

To determine the portion of the system's responses that might be attributable to random processes, data sets were constructed from Poisson distributed noise and subjected to identical processing as the data collected from the in vivo preparations. The result of these experiments (Table 3, bottom) was a numerical value that served as the decision point for significance (termed the criterion point) of the neural data scores (Table 3, top).

Composite scores from four experiments (3,360 total presentations) shown in Table 3 (top) indicate high accuracy at identifying stimulus patterns 1, 8, 9, 11, and 12. When compared to the criterion levels given in Table 3 (bottom) for the corresponding table locations, patterns 10, 5, 6, 4, and 2 (listed in decreasing order) were also significant but not as remarkable as the previously mentioned patterns. Pattern 3 was never correctly identified during any of the seven experiments, a state which is itself unlikely. Referring to Table 3 (bottom), for pattern 3 the number of correct responses produced by chance was six, giving a probability score of 0.012 (6/500). Thus for 329 presentations of pattern 3, we would expect a score of 4 by random chance. The results from Table 3 (top) that were greater than the corresponding locations in Table 3 (bottom) were plotted as a stacked bar chart [Figure 9(a)] to permit comparison of both the correct and incorrect responses from the system. The results given in Table 3 (bottom) are shown as a stacked bar chart below [Figure 9(b)] the experiment

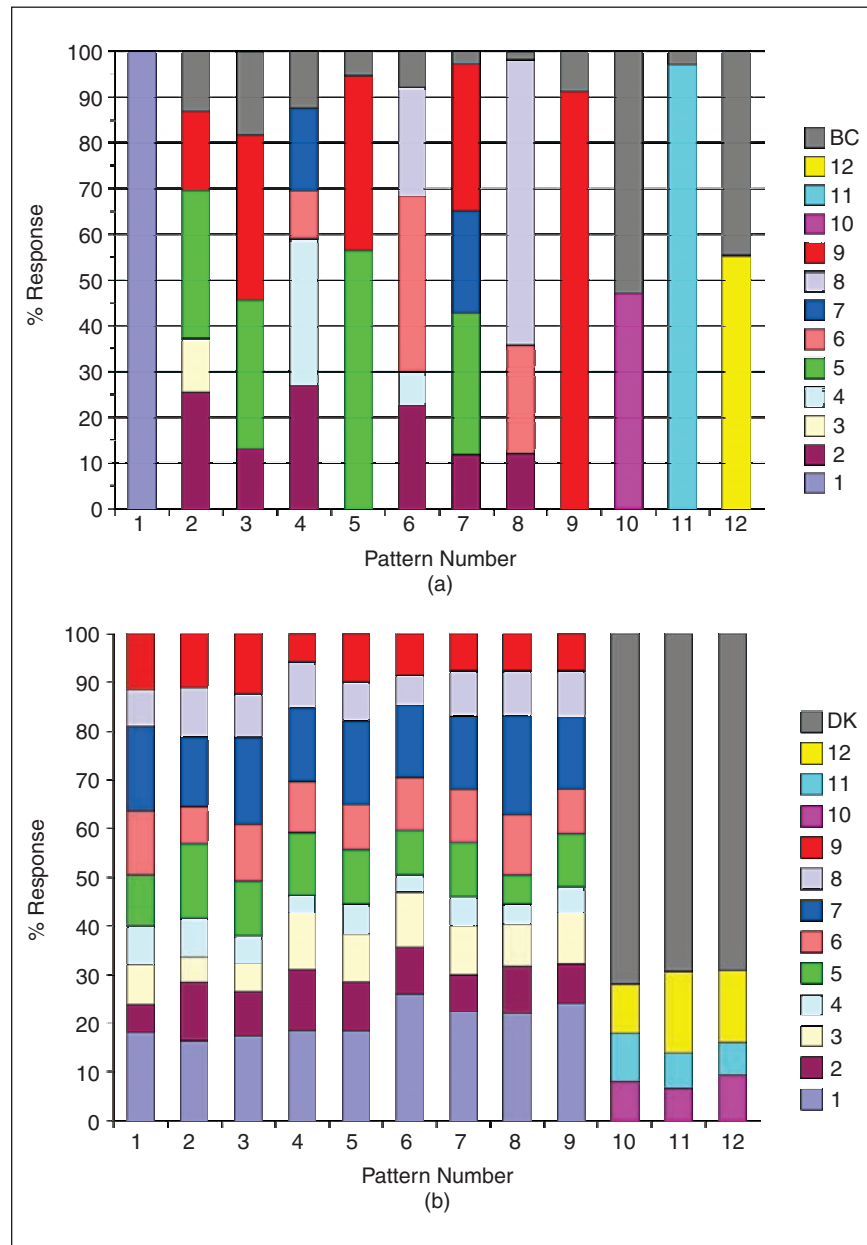


Fig. 9. (a) Bar chart representation of the scores from Table 3 (top), which were greater than the corresponding values in Table 3 (bottom); that is, the values above the criterion point. Each numbered bar corresponds to the desired output from the system for the given pattern. The colors on the bar represent the relative percentages of the system responses for the desired output that were classified/misclassified as the patterns denoted by the color code at the right of the chart. The grey portion of the bar represents the percentage of the responses for the given pattern that were either unclassifiable or below the criterion point (BC on the legend). (b) Stacked bar chart representation of the scores from Table 3 (bottom) using Poisson distributed noise to train/test the system. The grey portion of the bar (DK on the legend) represents the percentage of responses that were unclassifiable.

Though it's doubtful a given source could ever be at the exact centers of all four pairs in a hand-made cuff, being near the centers impacts the SNR and thus the accuracy for that pattern.

scores to permit direct comparison of the system performance on both random and neural data.

Training the neural network with data from one experiment then testing it with data from the same specimen but a different experiment results in high accuracy at the stimulus/no-stimulus recognition task (Pattern 1) and greatly decreased accuracy with the other patterns. Training the network with data from one specimen and testing it with data from a different specimen was also attempted with unsatisfactory results. Therefore, the trained network for a given specimen was very specific to the specimen-interface-nerve configuration on which the data used to build the training/testing sets originated.

The DMJD was used to flag combinations of neural data whose probability distributions (and thus their entropy values) may not differ enough to allow successful classification by the trained neural network. The mean value of all the cells in the table (generally around 0.32 bits) for a given experiment was used as the a priori decision value. A value in a given cell that was below the decision value was considered an indicator of a potential misclassification. After the network was trained and the score matrices (Table 3, top) constructed, the predictions made using the DMJD results (Table 4) were compared to the system scores for accuracy.

The accuracy of the DMJD as a predictor of neural data misclassification at various decision values was assessed (Figure 10). Values for the DMJD that were below the decision value indicated a potential problem with the neural net partitioning the data for the two patterns. Decision values from 0.1 to 0.9 bits were evaluated along with the median (0.28 bits) and the mean (0.32 bits). It was determined from this analysis that a value of approximately 0.4 bits would have produced the highest accuracy (approximately 87%) while the efficacy of the mean and median were approximately 82%.

Discussion

$H(X)$ provides a single measure on which the computer can reliably distinguish between the stimulus/no-stimulus conditions because the $H(X)$ values for each condition cluster in distinct domains (Figure 11) in the input space of the neural network. The two domains are well separated because the experiments are conducted in the dark and there is a striking difference in the nerve activity between the stimulated and unstimulated conditions. While useful as a basic test of system function, in a real-world application it is unlikely that there will be a strictly unstimulated condition. Thus, how the system performs with the patterns containing stimulation (patterns 2-12) is of particular importance.

TABLE 4. Composite DMJD result matrix from one experiment. The rows and columns are organized as in Table 3 with two exceptions. First, there are no numbers on the diagonal since the $J(p,q)$ of a pattern compared to itself is 0. Second, since the $J(p,q)$ is the mean of $D(p||q)$ and $D(q||p)$ the DMJD values are symmetric about the diagonal; thus, the table is row-column invertible.

	1	2	3	4	5	6	7	8	9	10	11	12
1		0.40	0.35	0.53	0.32	0.59	0.37	0.70	0.27			
2	0.40		0.04	0.13	0.08	0.19	0.03	0.31	0.13			
3	0.35	0.04		0.17	0.03	0.24	0.02	0.35	0.08			
4	0.53	0.13	0.17		0.21	0.06	0.16	0.18	0.26			
5	0.32	0.08	0.03	0.21		0.27	0.05	0.38	0.05			
6	0.59	0.19	0.24	0.06	0.27		0.22	0.11	0.32			
7	0.37	0.03	0.02	0.16	0.05	0.22		0.33	0.10			
8	0.70	0.31	0.35	0.18	0.38	0.11	0.33		0.43			
9	0.27	0.13	0.08	0.26	0.05	0.32	0.10	0.43				
10											0.33	0.13
11										0.33		0.46
12										0.13	0.46	

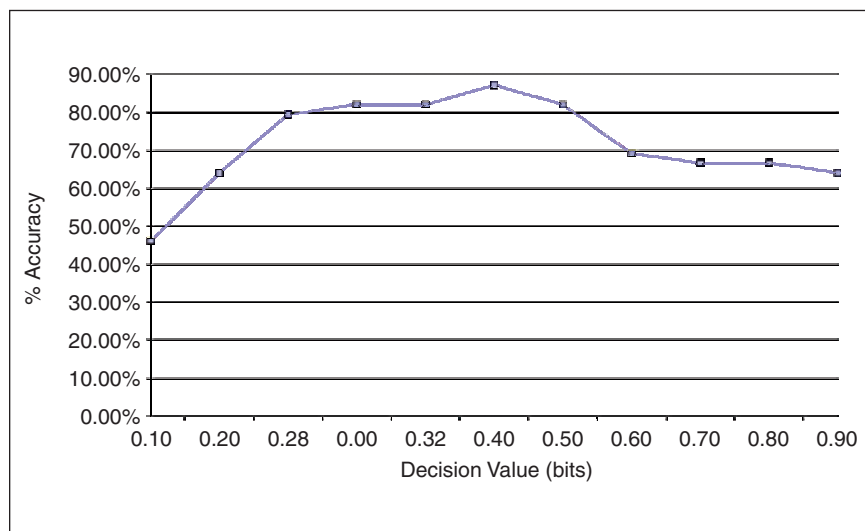


Fig. 10. Accuracy of the DMJD as a predictor of system performance at various decision values. In addition to increments of 0.1 bits (X axis) the accuracy of the median (0.28) and the mean (0.32) as decision values are also shown. The optimal decision value for this data was 0.4 bits.

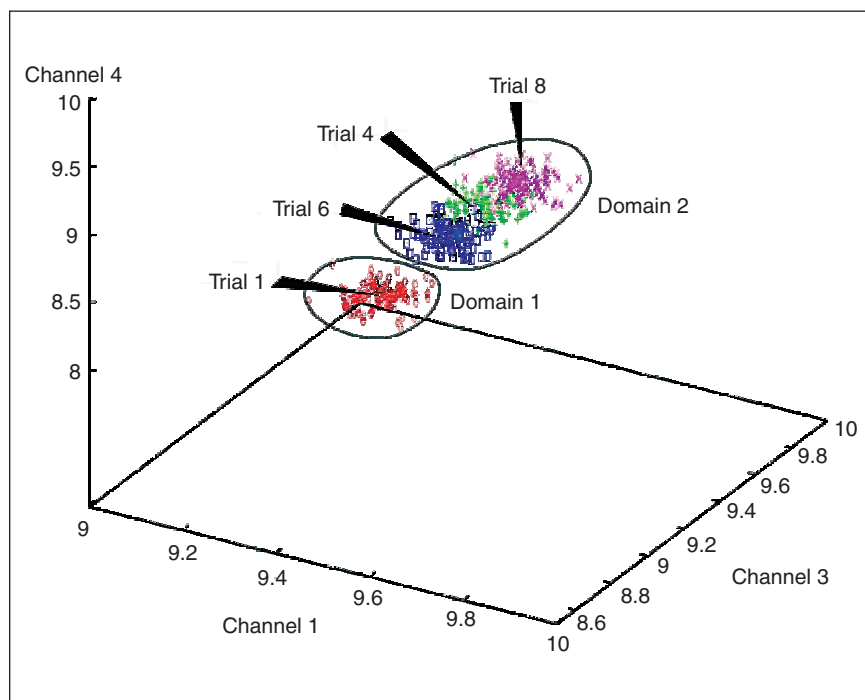


Fig. 11. State-space plot for three out of four dimensions of a partial data set for a single specimen. The X, Y, and Z axes are $H(X)$ values in bits for the designated data acquisition channel. Domain 1 consists of the $H(X)$ data for all the presentations for the control. Domain 2 consists of clusters formed by the $H(X)$ data from the presentations of the remaining patterns. Here the term trial is used to refer to all the presentations of the same pattern. Thus, trial 1 refers to all the presentations of pattern 1, trial 2 is pattern 2, etc. Only clusters of $H(X)$ data from trials 4, 6, and 8 are shown here for clarity. There are 400 presentations (100 per trial) shown on this plot.

The input space for the neural net has four dimensions, one for each differential electrode pair. Within the cluster formed by the $H(X)$ data for the stimulated condition (Figure 11, domain 2) there are subdomains that contain the data for identical patterns (Figure 11, trials 4,6,8). One of the tasks of the neural net is to learn to partition this multidimensional space in such a way as to identify the different domains/subdomains and map them to desired output states.

However, partitioning is not the only task the network must be able to perform. In order to correctly classify novel data the network must be able to generalize about its inputs. Thus, it is critical to halt the training process before the network becomes overfitted. The cascade correlation algorithm trains the network and constructs the hidden layer; however, it is not simply a matter of the network weights becoming suboptimal with overtraining but also the size of the hidden layer growing too large. When the network becomes overfitted it performs increasingly well at identifying the activity that corresponds to the data on which it was trained (i.e., the training set) while becoming worse with novel data (Figure 7).

Several correctable problems have been identified that degrade system performance both when training/testing using data from the same specimen, same experiment (Table 3 and Figure 9) and training with data from one experiment and testing with data from a different experiment (same specimen).

First, it is difficult to reliably stabilize the cuff electrode in an acute recording situation and specimen movement produces artifacts that decrease the system's accuracy. Research indicates that when used in vivo in a chronic recording situation, connective tissue will encapsulate the cuff and mechanically stabilize it [8], thus preventing shifting of the electrode position due to movement of the subject. The impedance of an implanted cuff will sharply increase during the first few days following implantation due to trapped air and edema between the nerve and electrodes. This results in a decrease in the SNR and a subsequent decrease in system accuracy. In a chronic implantation situation, trapped air is resorbed and the edema resolves

as the damage heals. Studies conducted in cats by Loeb and Peck [8] demonstrate that around 20 days postoperative the impedance has decreased and stabilized at a value approximately double its measured preimplantation value. The doubling of the impedance is primarily due to the ingrowth of connective tissue between the nerve and the electrodes but should pose little problem with recordings.

Second, the $H(X)$ values used in this implementation are derived from the activity occurring in a narrow frequency band (wavelet level 10). This provides a coarse measure of activity that permits demonstration of this signal processing methodology using very simple stimuli. However, more levels must be utilized to improve accuracy on the simple stimuli and to allow the system to distinguish between the activity arising due to more complex stimulation patterns. Additionally, the information contained in each level might be better utilized by thresholding/windowing the coefficients into separate classes and then computing $H(X)$ for each class. Each group of $H(X)$ values is equivalent to a set of coordinates that locates a point in the neural net's multidimensional input space. With more dimensions in the input space of the neural net there would be a greater chance that the data generated by complex stimulus patterns would cluster.

Third, the electrode cuffs fabricated for this research were limited to a maximum electrode density of four pairs. This was the largest number of electrodes that could reliably be placed on a hand-made cuff. A cuff electrode constructed using microfabrication techniques on a flexible thin-film substrate could achieve much higher electrode density. This means more views of the nerve activity could be generated resulting in improved clustering of the data in the neural network's input space and better accuracy. Additionally, there is the possibility of integrating the buffer amplifier stage and the electrodes on the thin film to further improve the SNR.

Finally, it is not necessary to reconstruct the signal following wavelet decomposition. The entropy can be computed on the wavelet coefficients directly, thus saving several moderately expensive computing steps. This performance tweak would improve system efficiency allowing more wavelet levels to be processed in real time, thus improving accuracy.

The system's misclassifications on patterns 2-12 proved to be instructive. Referring to Figure 9 (top) we see that patterns 2 and 3 are frequently misclassified as patterns 5 and 9. Referring back to Table 1, we see that with the exception of pattern 2, the regions stimulated lie within the least receptor-dense portion of the eye. Thus, much of the activity produced by stimulating location 3 yields data that cluster in the same region of the neural net's input space as patterns 5 and 9. Furthermore, some of the activity produced by stimulating location 5 clusters in the region for pattern 9.

Pattern 2 stimulates an area of high receptor density that presumably produces a maximal response. However, the nerve fibers from the center of the eye are located in the center of the optic nerve and thus in the approximate center of the differential electrodes. A source in the exact center of a differential pair of electrodes would not be detectable since it would be seen equally by both the positive and negative electrodes. Though it's doubtful a given source could ever be at the exact centers of all four pairs in a hand-made cuff, being near the centers impacts the SNR and thus the accuracy for that pattern. Additionally, we speculate that there may be

shielding effects that reduce the received signal strength from fibers lying in the center of a nerve. More analysis is needed to determine if the misclassification results from some deficiency of the hardware/software, the anatomy of lateral eye/optic nerve, or both.

In the present study the DMJD was a useful measure of the quality of the recorded data and an accurate predictor of system performance. On-the-fly calculation of the DMJD can easily be incorporated into future MINIS software to provide an estimate of system performance when given real-time streaming data. The mean or median of the computed $J(p,q)$ provides a dynamic and reasonably accurate method for computing a decision value for the DMJD (Figure 10). Further investigation is needed to determine if the decision value computed using this method should be weighted to improve its accuracy.

The current system represents a first step toward the realization of the detector portion of a minimally invasive, direct peripheral nervous system interface. While the system is targeted at producing an interface to the peripheral nervous system, it would be relatively easy to adapt it to the central nervous system. The basic premise of the technology is that multiple electrodes placed around a volume conductor can detect the fluctuations in the field produced by electrically active sources in that volume. Theoretically, it should not matter if that volume is the space within a cuff electrode or the space between an array of electrodes in the brain or spinal cord.

Thus far the results are encouraging; however, more work is needed before this system could be used to reliably drive a prosthesis or interact with a virtual environment. As this article goes to press work is underway to greatly enhance the performance of the system both in terms of accuracy and speed. The focus of the ongoing research is to first improve the accuracy of the system to near 100% on the simple stimulus set. Following that, the next experiments will utilize more complex monochrome/greyscale patterns to enable further refinement of the imaging capabilities of the system. The goal of these experiments is to produce a near real-time stream of images from the lateral eye while the animal is in its natural environment.

Acknowledgments

The work discussed in this article was conducted at the Pennsylvania State University as part of Dr. Coates' doctoral research. In addition to the coauthors, Dr. Coates would like to thank Robert Milner, Ph.D.; Steven Dear, Ph.D.; and Steve Levison, Ph.D. A special thanks goes to Robert Barlow, Ph.D. and Fred Dodge Ph.D., for their helpful commentary and suggestions. For more information on this research, neural interfacing, and related fields visit <http://www.neuropunk.org>.

Thomas D. Coates Jr. received his undergraduate degree in biology from Virginia Commonwealth University in 1995 and his Ph.D. in neuroscience from the Pennsylvania State University in 2001. Dr. Coates is presently involved in neural interfacing research at the University of Kentucky and is the founder/ Webmaster of Neuropunk.Org. His research interests include direct brain-computer interfacing, virtual reality, and artificial intelligence.

Linda J. Larson-Prior received the B.A. degree in anthropology and Arabic in 1972 from Ohio State University and the M.A. in anthropology in 1978 from Case Western Reserve University. She received the Ph.D. in neuroscience from Kent State University (Northeast Ohio Universities College of Medicine) in 1986 and took postdoctoral training in neuroscience at Northwestern University College of Medicine and at the University of Chicago. Dr. Larson-Prior is currently an associate professor in neuroscience at Touro University College of Osteopathic Medicine (TUCOM). She moved to TUCOM from Penn State University College of Medicine in 1998. Her research interests include serotonergic modulation of cerebellar cortical networks and the neural processing of high-rate sensory stimulation in the human auditory system.

Seth Wolpert received his B.A. in biology and his B.S.E.E. with biomedical and computer options from Rutgers University in 1979. Upon graduation, he took a position as a design engineer with the MOS Memory Group of Texas Instruments Inc. In January of 1984, he returned to Rutgers University, where he pursued a Ph.D. in biomedical engineering. Upon graduation, he took a position with the Department of Electrical and Computer Engineering at the University of Maine. In June of 1995, he joined the faculty of electrical engineering at Penn State University-Harrisburg, where he continues to teach courses in electronics, semiconductor devices, analog and digital IC design, and artificial neural networks. Dr. Wolpert has published numerous articles in journals, conference proceedings, and books. He has conducted and mentored research with funding from the NSF and many other corporate and academic sponsors. He has twice chaired the IEEE Northeast Bioengineering Conference.

Fred Prior received the B.A. degree in chemistry in 1974, the M.A. in anthropology in 1976, and the M.S. in biomedical engineering in 1984 all from Case Western Reserve University. He received the Ph.D. in computer science from Illinois Institute of Technology in 1992. Dr. Prior is currently the director of PACS Product Line Management for the Health Imaging Division of Eastman Kodak Company developing medical image management systems for radiology applications. He was assistant professor of radiology, bioengineering, and neuroscience at Penn State University College of Medicine. His research interests include medical workflow management systems and modeling and simulation of neural information processing.

Address for Correspondence: Thomas D. Coates, Jr. PhD, Neuropunk.org, PO Box 910385, Lexington, KY, 40591-0385. E-mail: tcoates@neuropunk.org

References

- [1] G.T.A. Kovacs, C.W. Stormont, and J.M. Rosen, "Regeneration microelectrode array for peripheral nerve recording and stimulation," *IEEE Trans. Biomed. Eng.*, vol. 39, no. 9, pp. 893-902, 1992.
- [2] J.D. Sweeney and J.T. Mortimer, "An asymmetric two electrode cuff for generation of unidirectionally propagated action potentials," *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 541-549, June 1986.
- [3] C. Veraart, W.M. Grill, and J.T. Mortimer, "Selective control of muscle activation with a multipolar nerve cuff electrode," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 640-653, July 1993.
- [4] C. Veraart, C. Raftopoulos, J.T. Mortimer, J. Delbeke, D. Pins, G. Michaux, A. Vanlierde, S. Parrini, and M.C. Wanet-Delfalque, "Visual sensations produced by optic nerve stimulation using a self-sizing spiral electrode cuff," *Brain Res.*, vol. 813, no. 1, pp. 181-186, Nov. 30, 1998.
- [5] R. Baratta, M. Ichie, S.K. Hwang, and M. Solomonow, "Orderly stimulation of skeletal muscle motor units with tripolar cuff electrode," *IEEE Trans. Biomed. Eng.*, vol. 36, pp. 836-843, Aug. 1989.
- [6] M. Brunner and U.T. Koch, "A new cuff electrode for reversible conduction blocking," *J. Neurosci. Meth.*, vol. 38, no. 2-3, pp. 259-265, July 1991.
- [7] W.M. Grill and T. Mortimer, "Stability of the input-output properties of chronically implanted multiple contact nerve cuff stimulating electrodes," *IEEE Trans. Rehab. Eng.*, vol. 6, pp. 364-373, Dec. 1998.
- [8] G.E. Loeb and R.A. Peck, "Cuff electrodes for chronic stimulation and recording of peripheral nerve activity," *J. Neurosci. Meth.*, vol. 64, no. 1, pp. 95-103, Jan. 1996.
- [9] P.J. Slot, P. Selmar, A. Rasmussen, and T. Sinkjær, "Effect of long-term implanted nerve cuff electrodes on the electrophysiological properties of human sensory nerves," *Artificial Organs*, vol. 21, no. 3, pp. 207-209, Mar. 1997.
- [10] J.J. Struijk, M. Thomsen, J.O. Larsen, and T. Sinkjær, "Cuff electrodes for long-term recording of natural sensory information," *IEEE Eng. Med. Biol. Mag.*, vol. 18, pp. 91-98, May/June 1999.
- [11] R.B. Barlow Jr., "Inhibitory fields in the Limulus lateral eye," *J. Gen. Physiol.*, vol. 54, no. 3, pp. 383-396, Sept. 1969.
- [12] H.K. Hartline, "The receptive fields of optic nerve fibers," *Amer. J. Physiol.*, vol. 130, p. 690, 1940.
- [13] S.E. Brodie, B.W. Knight, and F. Ratliff, "The spatiotemporal transfer function of the limulus lateral eye," *J. Gen. Physiol.*, vol. 2, pp. 167-202, 1978.
- [14] E. Kaplan and R.B. Barlow Jr., "Properties of visual cells in the limulus lateral eye in situ," *J. Gen. Physiol.*, vol. 66, pp. 303-326, 1975.
- [15] C. Passaglia, F. Dodge, E. Herzog, S. Jackson, and R. Barlow, "Deciphering a neural code for vision," *Proc. Nat. Acad. Sci.*, vol. 94, num. 23, pp. 12649-12654, 1997.
- [16] S.C. Chamberlain and R.B. Barlow Jr., "Neuroanatomy of the visual afferents in the horseshoe crab (Limulus polyphemus)," *J. Comparative Neurol.*, vol. 192, pp. 387-400, 1980.
- [17] W.H. Farenbach, "The visual system of the horseshoe crab Limulus polyphemus," *Int. Rev. Cytology*, vol. 41, pp. 285-349, 1975.
- [18] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992.
- [19] T.M. Cover and J.A. Thomas, *Elements of Information Theory*. New York, NY: Wiley, 1991, pp 1-42.
- [20] A. Waibel, "Modular construction of time delay neural networks for speech recognition," *Neural Computation*, vol. 1, no. 1, pp. 39-46, 1989.
- [21] A. Zell, G. Mamier, M. Vogt, N. Mache, R. Hübner, S. Döring, K. Herrmann, T. Soyöz, M. Schmalzl, T. Sommer, A. Hatzigeorgiou, D. Posselt, T. Schreiner, B. Kett, G. Clemente, J. Wieland, and J. Gatter. (1999). Stuttgart Neural Network Simulator User Manual (version 4.2), University of Stuttgart, Institute for Parallel and Distributed High Performance Systems. Stuttgart, Germany, [Online]. Available: <http://www.informatik.uni-stuttgart.de/ipvr/bv/projekte/snns/>.
- [22] M. Vogt, "Generalized radial basis functions" in *Einem simulator neuronaler netze*, Universität Stuttgart, Institute for Parallel and Distributed High Performance Systems (IPVR), Stuttgart, Germany, Diplomarbeit 875, 1992.
- [23] D.E. Rummelhart, G.E. Hinton, and R.J. Williams, "Learning internal representations by error propagation," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition; Vol. 1: Foundations*, D.E. Rummhart and J.L. McClelland, Eds. Cambridge, MA: MIT Press, 1986.
- [24] S.E. Fahlman, "The recurrent cascade correlation architecture," School of Computer Science, Carnegie Mellon University, Pittsburg, Pa, Tech. Rep. CMU-CS-91-100, 1991.
- [25] S.E. Fahlman, "Faster learning variations on back-propagation: An empirical study," In *Connectionist Models Summer School*, T.J. Sejnowski, G.E. Hinton, and D.S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1988.