Brain-Implantable Biomimetic Electronics as the Next Era in Neural Prosthetics

THEODORE W. BERGER, MICHEL BAUDRY, ROBERTA DIAZ BRINTON, JIM-SHIH LIAW,
VASILIS Z. MARMARELIS, FELLOW, IEEE, ALEX YOONDONG PARK, BING J. SHEU, FELLOW, IEEE,
AND ARMAND R. TANGUAY, JR.

Invited Paper

    An interdisciplinary multilaboratory effort to develop an implantable neural prosthetic that can coexist and bidirectionally communicate with living brain tissue is described. Although the final achievement of such a goal is many years in the future, it is proposed that the path to an implantable prosthetic is now definable, allowing the problem to be solved in a rational, incremental manner. Outlined in this report is our collective progress in developing the underlying science and technology that will enable the functions of specific brain damaged regions to be replaced by multichip modules consisting of novel hybrid analog/digital microchips. The component microchips are "neurocomputational" incorporating experimentally based mathematical models of the nonlinear dynamic and adaptive properties of biological neurons and neural networks. The hardware developed to date, although limited in capacity, can perform computations supporting cognitive functions such as pattern recognition, but more generally will support any brain function for which there is sufficient experimental information. To allow the "neurocomputational" multichip module to communicate with existing brain tissue, another novel microcircuitry element has been developed—silicon-based multielectrode arrays that are "neuromorphic," i.e., designed to conform to the region-specific cytoarchitecture of the brain. When the "neurocomputational" and "neuromorphic" components are fully integrated, our vision is that the resulting prosthetic, after intracranial implantation, will receive electrical impulses from targeted subregions of the brain, process the information using the hardware model of that brain region, and communicate back to the functioning brain. The proposed prosthetic microchips also have been designed with parameters that can be optimized after implantation, allowing each prosthetic to adapt to a particular user/patient.

    Keywords—*Biomimetic signal processing, hippocampus, mixed signal, multisite electrode array, neural engineering, neural network, neural prosthetic, neuron-silicon interface, pattern recognition, VLSI.*

I. INTRODUCTION

    One of the true frontiers in the biomedical sciences is repair of the human brain: developing prosthetics for the central nervous system to replace higher thought processes that have been lost due to damage or disease. The type of neural prosthetic that performs or assists a cognitive function is qualitatively different than the cochlear implant or artificial retina, which transduce physical energy from the environment into electrical stimulation of nerve fibers [1], [2], and qualitatively different than functional electrical stimulation (FES), in which preprogrammed electrical stimulation protocols are used to activate muscular movement [3]. Instead, we consider here a neural prosthetic designed to replace damaged neurons in central regions of the brain with silicon neurons that are permanently implanted into the damaged region. The replacement silicon neurons would have functional properties specific to those of the damaged neurons, and would both receive as inputs and send as outputs electrical activity to regions of the brain with which the damaged region previously communicated. Thus, the

prosthetic being proposed is one that would replace the computational function of damaged brain, and restore the transmission of that computational result to other regions of the nervous system. Such a new generation of neural prosthetic would have a profound impact on the quality of life throughout society, as it would offer a biomedical remedy for the cognitive and memory loss that accompanies Alzheimer's disease, the speech and language deficits that result from stroke, and the impaired ability to execute skilled movements following trauma to brain regions responsible for motor control.

Although the barriers to creating intracranial, electronic neural prosthetics have seemed insurmountable in the past, the biological and engineering sciences are on the threshold of a unique opportunity to achieve such a goal. The tremendous growth in the field of neuroscience has allowed a much more detailed understanding of neurons and their physiology, particularly with respect to the dynamic and adaptive cellular and molecular mechanisms that are the basis for information processing in the brain. Likewise, there have been major breakthroughs in the mathematical modeling of nonlinear and nonstationary systems that are allowing quantitative representations of neuron and neural system function to include the very complexity that is the basis of the remarkable computational abilities of the brain. The continuing breakthroughs in electronics and photonics offer opportunities to develop hardware implementations of biologically based models of neural systems that allow simulation of neural dynamics with true parallel processing, a fundamental characteristic of the brain, and real-time computational speed. Fundamental advances in low-power designs have provided the essential technology to minimize semiconductor circuit heat generation, thus increasing compatibility with temperature-sensitive mechanisms of the brain. Finally, the complementary achievements in materials science and molecular biology offer the possibility of designing compatible neuron-silicon interfaces to facilitate communication between silicon computational devices and the living brain.

II. Essential Requirements for an Implantable Neural Prosthetic

In general terms, there are six essential requirements for an implantable microchip to serve as a neural prosthetic. First, if the microchip is to replace the function of a given brain tissue, it must be truly biomimetic, i.e., the neuron models incorporated in the prosthetic must have properties of real biological neurons. This demands a fundamental understanding of the information processing capabilities of neurons that is experimentally based. Second, a neural prosthetic is desired only when a physiological or cognitive function is detectably impaired (according to neurological or psychiatric criteria). Physiological/cognitive functions are the expression, not of single nerve cells, but of populations of neurons interacting in the context of a network of interconnections. Thus, biologically realistic neuron models must be capable of being concatenated into network models that can simulate physiological/psychological phenomena. Third, the neuron and

neural network models in question must be miniaturized sufficiently to be implantable, which demands their implementation in at least microchip circuitry. Given the known signaling characteristics of neurons, such an implementation will most likely involve hybrid analog/digital device designs. Fourth, the resulting microchip or multichip module must communicate with existing, living neural tissue in a bidirectional manner. Given that both electronic and neural systems generate and respond to electrical signals, this is feasible, though the region-specific, nonuniform distribution of neurons within the brain places substantial constraints on the architecture of neuron-silicon interfaces. Fifth, the variability in phenotypic and developmental expression of both structural and functional characteristics of the brain will necessitate adaptation of each prosthetic device to the individual patient. Some provision for "personalizing" an implantable prosthetic must be anticipated and included in the neuron/network model and the device design. Finally, there is the critical issue of power required for the prosthetic device. Not only will supplying power be difficult given implantation of a set of microchips into the depths of the brain (versus the periphery as with a cochlear implant), but cellular and molecular mechanisms found in the brain are highly temperature sensitive, so that any solution must minimize heat generation to remain biocompatible.

We will describe here an interdisciplinary multilaboratory effort to develop such an implantable, computational prosthetic that can coexist and bidirectionally communicate with living neural tissue. We will deal with five of the above requirements; only the issue of power will not be addressed here. Although the final achievement of an implantable prosthetic remains years in the future, it is nonetheless our position that the path to such a goal is now definable, allowing a solution path to be defined and followed in an incremental manner. Outlined in this report is our collective progress to date in developing the underlying science and technology that will enable the functions of specific brain regions to be replaced by multichip modules consisting of novel hybrid analog/digital microchips. The component microchips are "neurocomputational," incorporating experimentally based mathematical models of the nonlinear dynamic and adaptive properties of real brain neurons and neural networks. The resulting hardware can perform computations supporting cognitive functions such as pattern recognition, but more generally will support any brain function for which there is sufficient experimental information. To allow the "neurocomputational" multichip module to communicate with existing brain tissue, another novel microcircuitry element has been developed —silicon-based multielectrode arrays that are "neuromorphic," i.e., designed to conform to the region-specific cytoarchitecture of the brain. When the "neurocomputational" and "neuromorphic" components are fully integrated, our vision is that the resulting prosthetic, after intracranial implantation, will receive electrical impulses from targeted subregions of the brain, process the information using the hardware model of that brain region, and communicate back to the functioning brain. The proposed prosthetic microchips also have been
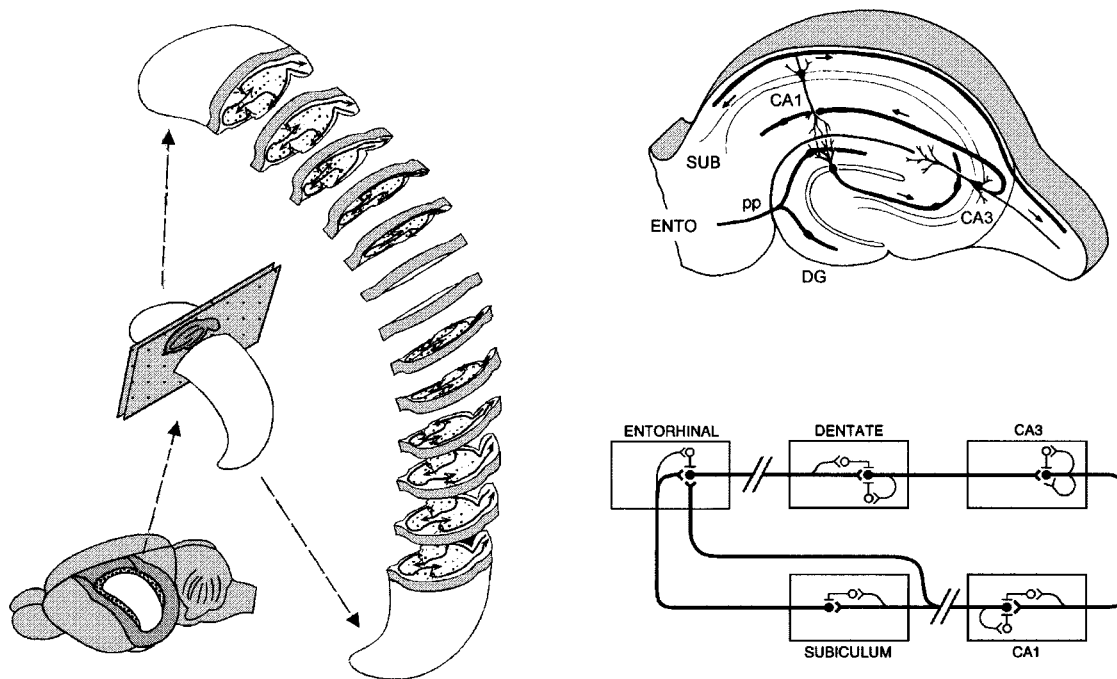
Fig. 1. Left panel: Diagrammatic representation of the rat brain (lower left), showing the relative location of the hippocampal formation on the left side of the brain (white); diagrammatic representation of the left hippocampus after isolation from the brain (center), and slices of the hippocampus for sections transverse to the longitudinal axis. Right panel: Diagrammatic representation of one transverse slice of hippocampus, illustrating its intrinsic organization: fibers from the entorhinal cortex (ENTO) project through perforant path (pp) to the dentate gyrus (DG); granule cells of the dentate gyrus project to the CA3 region, which in turn projects to the CA1 region; CA1 cells project to the subiculum (SUB), which in the intact brain then projects back to the entorhinal cortex. In a slice preparation, return connections from CA1 and the subiculum are transected, creating an open-loop condition for experimental study of hippocampal neurons.

designed with parameters that can be optimized after implantation, allowing each prosthetic to adapt to a particular user/patient.

III. THE SYSTEM: HIPPOCAMPUS

The computational properties of the prosthetic being developed are based on the hippocampus, a cortical region of the brain involved in the formation of new long-term memories. The hippocampus lies beneath the phylogenetically more recent neocortex, and is composed of several different subsystems that form a closed feedback loop (see Fig. 1), with input from the neocortex entering via the entorhinal cortex, propagating through the intrinsic subregions of hippocampus, and then returning to neocortex. The intrinsic pathways consist of a cascade of excitatory connections organized roughly transverse to the longitudinal axis of the hippocampus. As such, the hippocampus can be conceived of as a set of interconnected, parallel circuits [4], [5]. The significance of this organizational feature is that, after removing the hippocampus from the brain, transverse "slices" (approximately 500 μm thick) of the structure may be maintained *in vitro* that preserve a substantial portion of the intrinsic circuitry, and thus allow detailed experimental study of its principal neurons in their open-loop condition [6], [7].

The hippocampus is responsible for what have been termed long-term "declarative" or "recognition" memories

[8]–[11]: the formation of mnemonic labels that identify a unifying collection of features (e.g., those comprising a person's face), and to form relations between multiple collections of features (e.g., associating the visual features of a face with the auditory features of the name for that face). In lower species not having verbal capacity, an analogous hippocampal function is evidenced by an ability, for example, to learn and remember spatial relations among multiple, complex environmental cues to navigate and forage for food [12]. Major inputs to the hippocampus arise from virtually all other cortical brain regions, and transmit to hippocampus high-level features extracted by each of the sensory systems subserved by these cortical areas. Thus, the hippocampus processes both unimodal and multimodal features for virtually all classes of sensory input, and modifies these neural representations so that they can be associated (as in the case of forming a link between a face and a name) and stored in long-term memory in a manner that allows appropriate additional associations with previously learned information (the same face may have context-dependent names, e.g., first-name basis in an informal social setting versus position title in a formal business setting), and that minimizes interference (the same name may be associated with several faces). After processing by the hippocampal system, new representations for important patterns are transmitted back to other cortical regions for long-term storage; thus, long-term memories are not stored in hippocampus, but

propagation of neural representations through its circuitry is required for a re-encoding essential for the effective transfer of short-term memory into long-term memory.

Although developing a neural model for long-term memory formation (or any other cognitive function) may initially appear somewhat daunting, there is a rational approach to the problem. Information in the hippocampus and all other parts of the brain is coded in terms of variation in the sequence of all-or-none, point-process (spike) events, or temporal pattern (for multiple neurons, variation in the spatio-temporal pattern). The essential signal processing capability of a neuron is derived from its capacity to change an input sequence of interspike intervals into a different output sequence of interspike intervals. The resulting input–output (I/O) transformations in all brain regions are strongly nonlinear, due to the nonlinear dynamics inherent in the molecular mechanisms comprising neurons and their synaptic connections [13]. As a consequence, the output of virtually all neurons in the brain is highly dependent on temporal properties of the input. The I/O transformations of neurons in hippocampus and neocortex—the regions of the brain subserving pattern recognition—are the only "features" that the nervous system has to work with in constructing representations at the cortical level. Identifying the nonlinear I/O properties of neurons involved in pattern recognition is equivalent to identifying the feature models that endow the brain with its superior feature extraction capability. I/O properties of synapses and neurons are not static, but are altered by biological learning mechanisms to achieve an optimal feature set during memory formation for a new pattern. Identifying activity-dependent forms of synaptic plasticity of neurons involved in pattern recognition is equivalent to identifying the biological "learning rules" used in optimizing feature sets.

IV. BIOMIMETIC MODELS OF HIPPOCAMPAL NEURON PROPERTIES

A. Quantifying I/O Nonlinearities of Hippocampal Neurons

In order to incorporate the nonlinear dynamics of biological neurons into neuron models for the purposes of developing a prosthetic, it is first necessary to measure them accurately. We have developed and applied methods for quantifying the nonlinear dynamics of hippocampal neurons [6], [7], [14]–[18] using principles of nonlinear systems theory [19]–[23]. In this approach, properties of neurons are assessed experimentally by applying a random interval train of electrical impulses as an input and electrophysiologically recording evoked output of the target neuron during stimulation (Fig. 2, upper left). The input train consists of a series of impulses (as many as 4064), with interimpulse intervals varying according to a Poisson process having a mean of 500 ms and a range of 0.2–5,000 ms. Thus, the input is "broad-band" and stimulates the neuron over the majority of its operating range, i.e., the statistical properties of the random train are highly consistent with the known physiological properties of hippocampal neurons. Nonlinear response

properties are expressed in terms of the relation between progressively higher order temporal properties of the sequence of input events and the probability of neuronal output, and are modeled as the kernels of a functional power series. In the case of a third-order estimation

$$y(t) = G_0 + G_1[h_1, x(t)] + G_2[h_2, x(t)] \\ + G_3[h_3, x(t)] + \cdots$$

where
$y(t)$ output;
(G_i) set of functionals;
(h_i) set of kernels which characterize the relationship between the input and output

$$G_0(t) = h_0$$
$$G_1(t) = \int h_1(\tau)x(t - \tau)d\tau$$
$$G_2(t) = 2 \iint h_2(\tau, \tau + \Delta)x(t - \tau)x(t - \Delta - \tau)d\Delta d\tau$$
$$G_3(t) = 6 \iiint h_3(\tau, \tau + \Delta_1, \tau + \Delta_1 + \Delta_2) \\ \times x(t - \tau)x(t - \tau - \Delta_1) \\ \times x(t - \tau - \Delta_1 - \Delta_2)d\Delta_1 d\Delta_2 d\tau.$$

The train of discrete input events defined by $x(t)$ is a set of δ-functions. The first-, second-, and third-order kernels of the series are obtained using a variety of estimation procedures [22]–[24].

To clarify the interpretation of the kernels in the context of results for a typical granule cell of the hippocampus, the first-order kernel, $h_1(\tau)$, is the average probability of an action potential output occurring [with a latency of (τ)] to any input event in the train. The intensity of stimulation was chosen so that the first-order kernel had a probability value of 0.4–0.5 (Fig. 2, bottom left). The second-order kernel, $h_2(\tau, \Delta)$, represents the modulatory effect of any preceding input occurring (Δ) ms earlier on the most current impulse in the train (Fig. 2, top right). Second-order nonlinearities are strong: intervals in the range of 10–30 ms result in facilitation as great as 0.3–0.4 (summing the first and second-order values, the probability of an output event is 0.8–1.0 for this range of intervals). The magnitude of second-order facilitation decreases as interstimulus interval lengthens, with values of (Δ) greater than 100 ms leading to suppression, e.g., interstimulus intervals in the range of 200–300 ms decrease the average probability of an output event by approximately 0.2. The third-order kernel, $h_3(\tau, \Delta_1, \Delta_2)$, represents the modulatory effects of any two preceding input events occurring Δ_1 ms and Δ_2 ms earlier on the most current impulse that are not accounted for by the first- and second-order kernels (Fig. 2, bottom right). The example third-order kernel shown is typical for hippocampal granule cells, and reveals that combinations of intervals less than approximately 150 ms leads to additional suppression of granule cell output by as much as 0.5. This third-order nonlinearity represents in part saturation of second-order facilitative effects.
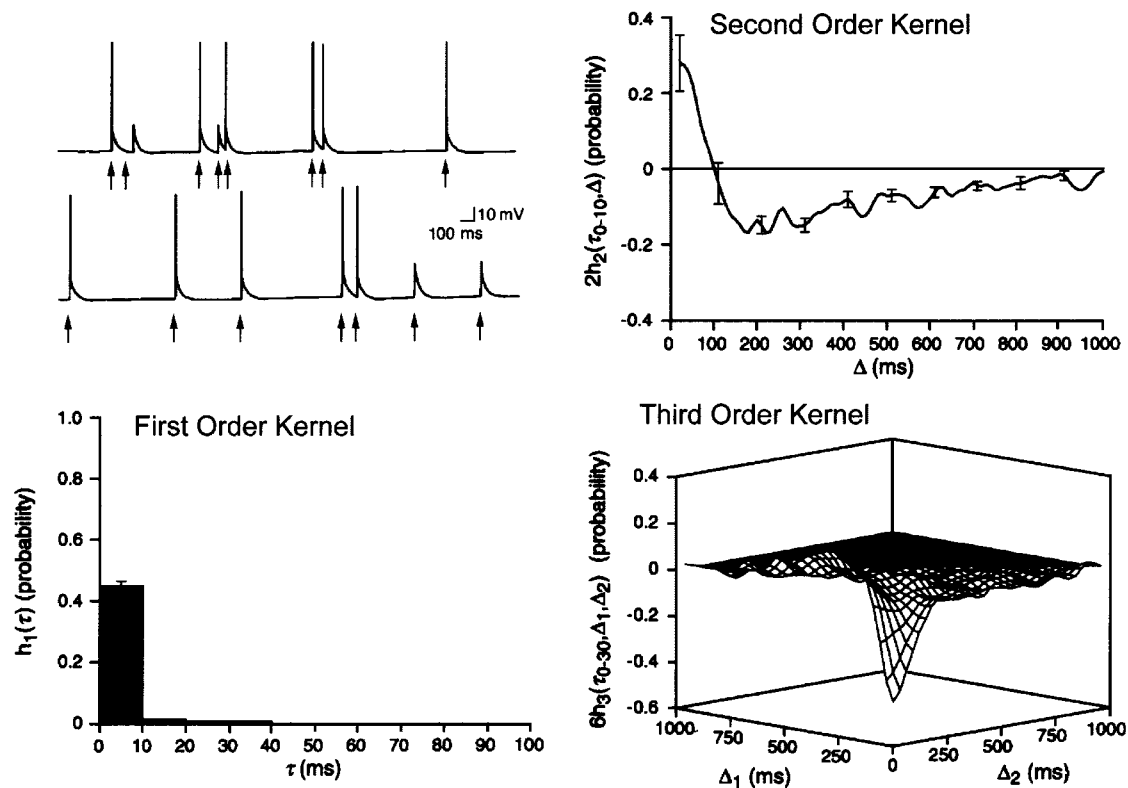
Fig. 2. Left upper panel: Sample electrophysiological recording from a hippocampal granule cell during random impulse train stimulation. Each arrow indicates when an electrical impulse is applied to perforant path inputs (see Fig. 1). Large, positive-going, unitary (action potential) events indicate when an input generated an output response from the granule cell; smaller, positive-going events (e.g., to first impulse and last two impulses) indicate when an input generated only a subthreshold response (no output). The time delay (latency) from the input event (arrow) to the granule cell response is equivalent to the parameter τ in the equations in the text (all latencies are less than 10 ms); the intervals between input events is equivalent to the parameter Δ in the equations in the text. Left lower panel: First-order kernel, $h_1(\tau)$, which represents the average probability of an action potential output occurring [with a latency of (τ)] to any input event in the train. Right upper panel: Second-order kernel, $h_2(\tau, \Delta)$, which represents the modulatory effect of any preceding input occurring (Δ) ms earlier on the most current impulse in the train. Right lower panel: Third-order kernel, $h_3(\tau, \Delta_1, \Delta_2)$, which represents the modulatory effects of any two preceding input events occurring Δ_1 and Δ_2 ms earlier on the most current impulse that are not accounted for by the first and second-order kernels.

B. Improved Kernel Estimation Methods

The output of hippocampal and other cortical neurons exhibits a dependence on input temporal pattern that is among the greatest of any class of neuron in the brain, because of a wide variety of voltage-dependent conductances found throughout their dendritic and somatic membrane. Despite this, I/O models of the type described here provide excellent *predictive* models of cortical neuron behavior. Depending on the circumstances, kernels to the third order, and sometimes even to the second order alone, can account for 80%–90% of the variance of hippocampal neuron output. Until recently, high-order nonlinearities have been difficult to estimate accurately: traditional kernel estimation methods (e.g., cross-correlation) are highly sensitivity to noise and, thus, require long data sequences. To circumvent this problem, we have developed several novel methods for estimating nonlinearities that are significantly more efficient and result in substantially improved kernel estimates [24]–[31]. Several of the new methods involve the use of feedforward artificial neural networks (ANN). We have compared the Volterra–Wiener (cross-correlation) and ANN models in terms of their prediction ability on test data. Results showed two major advantages of the new-generation methodologies: 1) a significant reduction in the required data length (by a factor of at least ten) to achieve similar or better levels of prediction accuracy and 2) an ability to model higher order nonlinearities that could not be detected using traditional kernel estimation methods. In addition, we have recently developed methods capable of estimating nonstationary processes, and demonstrated their efficacy with long-term forms of hippocampal cellular plasticity [32]–[35]. The ability to accurately characterize nonstationarities provides the opportunity for extending the applicability of this approach to modeling adaptive properties of hippocampal and other cortical neural systems as well.

In total, the kernel functions represent an experimentally based model that is highly accurate in describing the functional dynamics of the neuron in terms of the probability of neuron output as a function of the recent history of the input.

As such, the kernels provide a mathematically "compact" representation of the resulting composite dynamics, as each of the many contributing biological processes need not be represented individually, or for that matter, even be known. In addition, because of the broad-band nature of the test stimulus, the model generalizes to a wide range of input conditions, even to input patterns that are not explicitly included in the random impulse train. As such, the kernels not only provide the basis for a biologically realistic neural network model, but also perhaps an ideal basis for an implantable neural prosthetic: an I/O model can be substituted for a neuron on which the model is experimentally based, without regard to the variability in neural representations that must exist from individual to individual, or the nearly infinite range of environmental stimuli that would give rise to those representations.

V. NEURAL NETWORK MODELS WITH BIOLOGICALLY REALISTIC DYNAMICS

A. Conventional Artificial Neural Networks

Brain-like processing is often modeled mathematically as "artificial neural networks," or networks of processing elements that interact through "connections." In artificial neural network models, a connection between processing elements—despite the complexity of the synaptic nonlinear dynamics described above—is represented as a single number to scale the amplitude of the output signal of a processing element. The parameters of an artificial neural network can be optimized to perform a desired task by changing the strengths of connections according to what are termed "learning rules," i.e., algorithms for when and by how much the connection strengths are changed during optimization. This simplification of a synapse as a number results in two fundamental limitations. First, although a processing element can be connected to a large number of other processing elements, it can transmit only one identical signal to all other elements. Second, only the connection strength can be changed during the optimization process, which amounts to merely changing the gain of the output signal of a processing element.

B. The "Dynamic Synapse" Neural Network Architecture

In an effort to develop more biologically realistic neural network models that include some of the temporal nonlinear signal processing properties of neurons, we have developed the "Dynamic Synapse" neural network architecture [36]. In this scheme, processing elements are assumed to transmit information by variation in a series of point-process (i.e., all-or-none) events, and connections between processing elements are modeled as a set of linear and nonlinear processes such that the output becomes a function of the time since past input events [Fig. 3(a)]. By including these dynamic processes, each network connection transforms a sequence of input events into another sequence of output events. In the brain, it has been demonstrated that the functional properties of multiple synaptic outputs that arise from a given neuron are not identical. This characteristic of the brain also has been incorporated as a second fundamental property of

Dynamic Synapse neural networks: although the same essential dynamics are included in each synapse originating from a given processing unit, the precise values of time constants governing those dynamics are varied. The consequence arising from this second property is that each processing element transmits a spatio-temporal output signal, which, in principle, gives rise to an exponential growth in coding capacity. Furthermore, we have developed a "dynamic learning algorithm" to train each dynamic synapse to perform an optimized transformation function such that the neural network can achieve highly complex tasks. Like the nonlinear dynamics described above and included in the Dynamic Synapse network models, this learning algorithm also is based on experimentally determined, adaptive properties of hippocampal cortical neurons (which cannot be reviewed here; see [32]–[35]), and is unique with respect to neural network modeling in that the transformation function extracts invariant features embedded in the input signal of each dynamic synapse. The combination of nonlinear dynamics and dynamic learning algorithm provides a high degree of robustness against noise, which is a major issue in processing real biological signals in the brain, as well as real-world signals, as demonstrated in our case studies of speaker-independent speech recognition described below.

C. Application to Speech Recognition

Current state-of-the-art speech recognition technology is based on complex multistage processing that is not biologically based. Although commercial systems can demonstrate impressive performance, they are still far from performing at the level of human listeners. To test the computational capability of the dynamic synapse neural network, two strong constraints were imposed: the network must be simple and small, and it must accomplish speech recognition in a single step, i.e., with no preprocessing stages. Our system not only achieved this goal, but as will be described below, also performed better than human listeners when tested with speech signals corrupted by noise, marking the first time ever that a physical device has outperformed humans in a speech recognition task [37]–[39].

1) Invariant Feature Extraction: Two characteristics of speech signals, variability and noise, make its recognition a difficult task. Variability refers to the fact that the same word is spoken in different ways by different speakers. Yet, there exist invariant features in the speech signal, allowing the constant perception of a given word, regardless of the speaker or the manner of speaking. Our first application of the Dynamic Synapse neural network model to speech recognition aimed at extracting those invariant features for a word set with very difficult discriminability, e.g., "hat" versus "hut" versus "hit" (14 words in total), spoken by eight different speakers. The variability of two signals can be measured by how well they correlate with each other. As seen in Fig. 3(b) (lower left), the speech waveforms of the same word spoken by two speakers typically show a low degree of correlation, i.e., they are quite different from each other. However, the dynamic synapse neural network can be trained to produce highly correlated signals for a given word [Fig. 3(b), lower
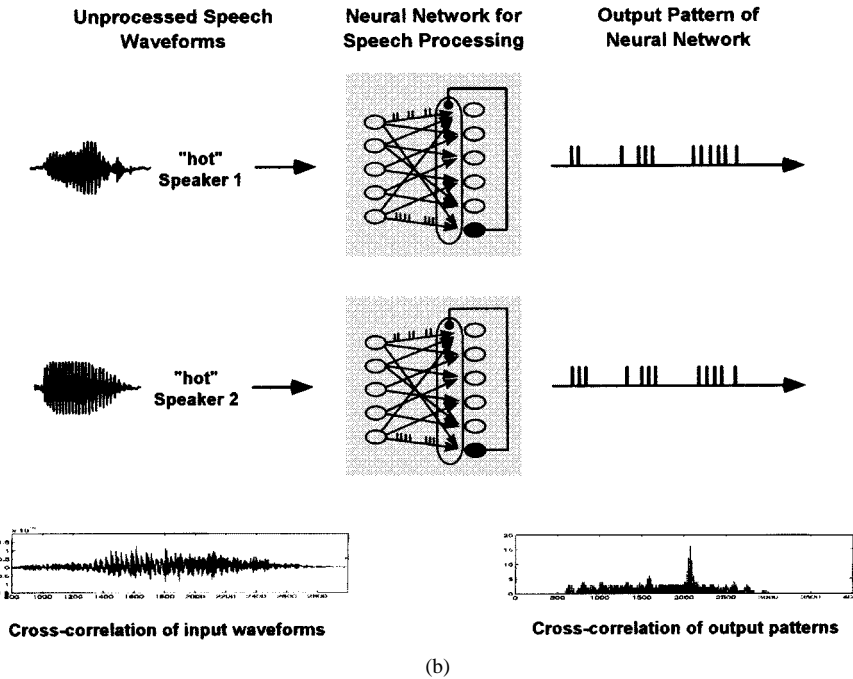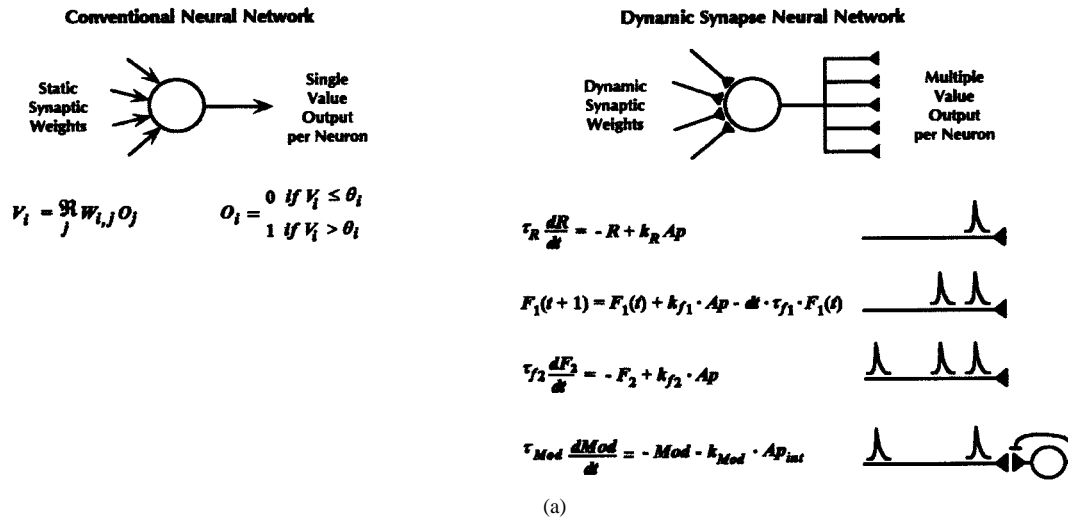
Conventional Neural Network

Static Synaptic Weights

Single Value Output per Neuron

$$V_i = \Re_j W_{i,j} O_j \qquad O_i = \begin{cases} 0 & \text{if } V_i \le \theta_i \\ 1 & \text{if } V_i > \theta_i \end{cases}$$

Dynamic Synapse Neural Network

Dynamic Synaptic Weights

Multiple Value Output per Neuron

$$\tau_R \frac{dR}{dt} = -R + k_R Ap$$

$$F_1(t+1) = F_1(t) + k_{f_1} \cdot Ap - dt \cdot \tau_{f_1} \cdot F_1(t)$$

$$\tau_{f_2} \frac{dF_2}{dt} = -F_2 + k_{f_2} \cdot Ap$$

$$\tau_{Mod} \frac{dMod}{dt} = -Mod - k_{Mod} \cdot Ap_{int}$$

(a)

Unprocessed Speech Waveforms

Neural Network for Speech Processing

Output Pattern of Neural Network

"hot" Speaker 1

"hot" Speaker 2

Cross-correlation of input waveforms

Cross-correlation of output patterns

(b)

Fig. 3. (a) Properties of a processing element of a traditional artificial neural network versus properties of a processing element of a biologically realistic Dynamic Synapse neural network. (b) Conceptual representation of speaker-independent word recognition identification by a Dynamic Synapse neural network. Inputs to the network are digitized speech waveforms from different speakers for the same word, which have little similarity (low cross-correlation) because of differences in speaker vocalization. The two networks shown are intended to represent the same network on two different training or testing trials; in a real case, one network is trained with both (or more) speech waveforms. On any given trial, each speech waveform constitutes the input for all five of the input units shown in the first layer. Each unit in the first layer of the network generates a different pulse-train encoding of the speech waveform ("integrate and fire neurons" with different parameter values). The output of each synapse (arrows) to the second layer of the network is governed by four dynamic processes [see (a)], with two of those processes representing second-order nonlinearities; thus, the output to the second layer neurons depends on the time since prior input events. A "dynamic learning rule" modifies the relative contribution of each dynamic process until the output neurons converge on a common temporal pattern in response to different input speech signals (i.e., high cross-correlation between the output patterns).

right]. Thus, the dynamic synapse neural network can extract invariant features embedded in speech signals that are inherently very difficult to discriminate, and can do so with no preprocessing of the data (only the output from a microphone was used) using a core signal processing system that is extremely small and compact.

2) Robustness with Respect to White Noise: To test the robustness of the invariant features extracted by the dynamic synapse neural network, the network was first trained to recognize the words [yes, no] randomly drawn from a database containing utterances by some 7000 speakers with no added noise. We then evaluated the performance of the model
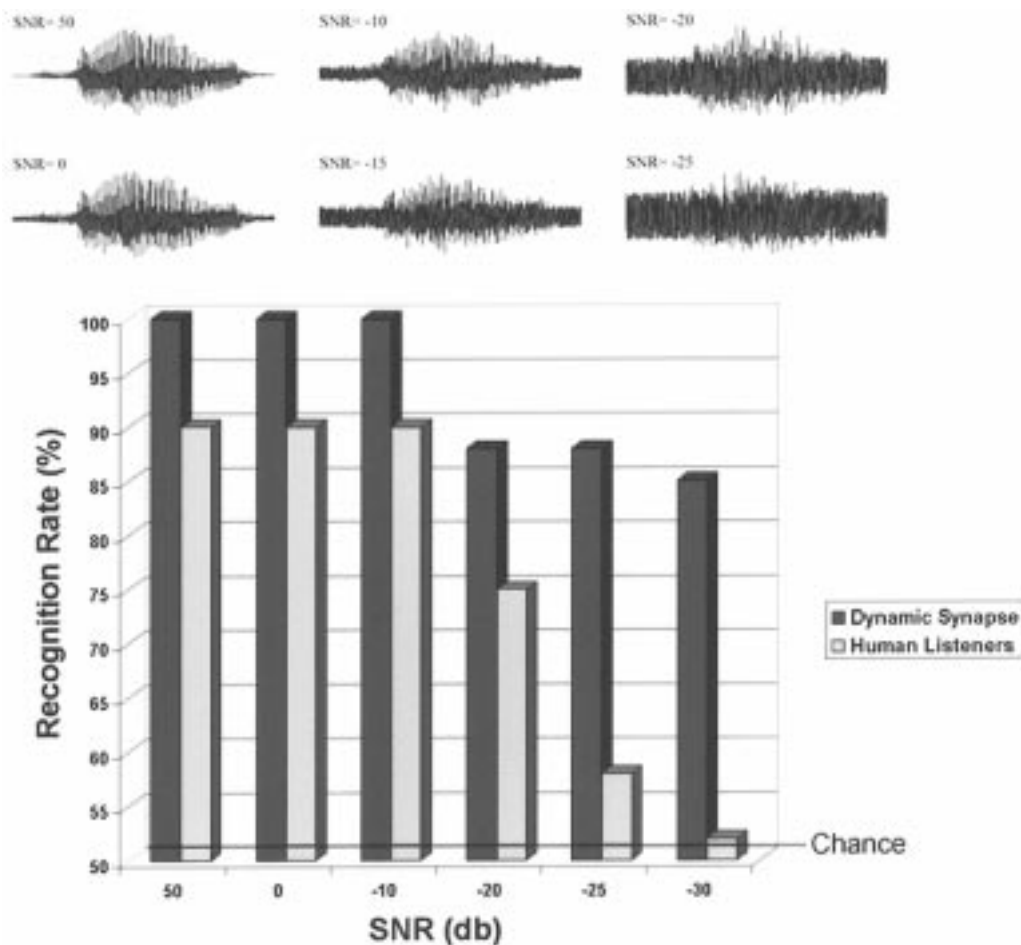
Fig. 4. Comparison of recognition rates by the Dynamic Synapse neural network system (dark bars) and human listeners (gray bars) for speaker-independent identification of the words "yes" and "no" when increasing amounts of white noise are added to the speech waveforms. Note that a 50% recognition rate is equivalent to chance.

when the speech signals not used during training were corrupted with progressively increasing amount of white noise [measured by the signal-to-noise- ratio (SNR) in decibels]. Results showed that our model is extremely robust against noise, performing better than human listeners tested with the same speech data set (Fig. 4). This is first time ever that a speech recognition system has outperformed human listeners, and the dynamic synapse system did so by a considerable margin.

3) Comparison with a State-of-the-Art Commercial Product and Robustness with Respect to Conversational Noise: The objective of this study was to compare the performance of the dynamic synapse neural network and one of the best state-of-the-art commercially available systems, namely, the Dragon Naturally Speaking speech recognition system. Since the Dragon system operates in a speaker-specific mode (the system is trained specifically for the one user), the speech signals consist of the words ["yes," "no," "fire," "stop"] spoken by a single speaker. The training of the Dragon system was done in two stages. In the first stage, the system was fully trained using the material provided by the manufacturer. In the second stage, it was further trained using the four target words.

Once the training was complete, both the dynamic synapse neural network and the Dragon system were tested with noise-added speech signals at various SNRs. A realistic "conversational" type of noise was used: a recording of one female speaker and one male speaker voice-reading newspapers simultaneously, along with the broadcast of a news program on the radio. The same noise-added speech signals were used to test human performance (average of five subjects). The results (Fig. 5) show that the Dragon system is extremely sensitive to noise and performs poorly under noisy conditions. Its performance degrades to 50% correct when the SNR was $+$ 20 dB, whereas both the dynamic synapse neural network and human listeners retained 100% recognition rate. The Dragon system failed to recognize any word when the SNR was $+$ 10 dB while the Dynamic Synapse neural network and human listeners performed at 100% and 90% recognition rates, respectively. Furthermore, the Dynamic Synapse neural network was highly robust and performed significantly better than human listeners when the SNR dropped below $+$ 2.5 dB. For example, for SNR ranging from 0 to -5 dB, human performance varied from 30% to 15% correct rate, while the Dynamic Synapse neural network retained a 75% correct rate. These findings show
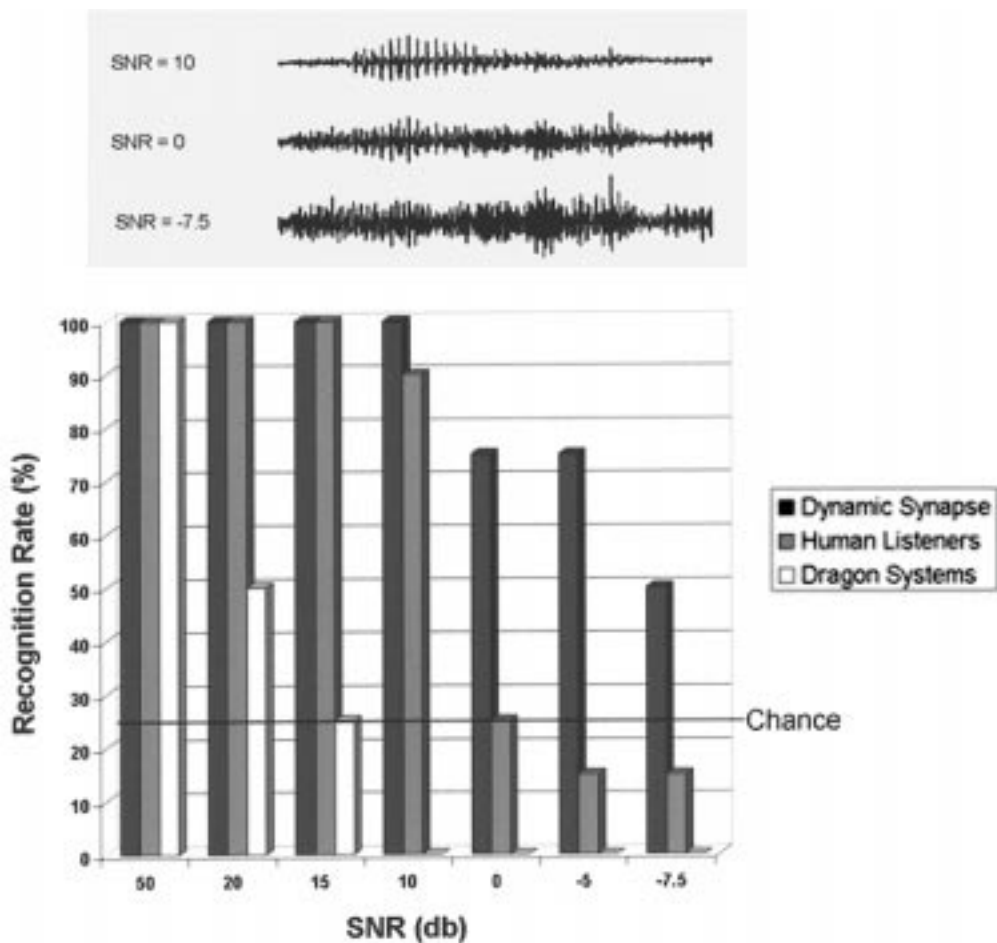
Fig. 5. Comparison of recognition rates by the Dynamic Synapse neural network system (dark bars), human listeners (gray bars), and the Dragon Naturally Speaking System for speaker-specific identification of the words "yes" and "no" when increasing amounts of conversational noise (see text) are added to the speech waveforms. Note that a 25% recognition rate is equivalent to chance.

that human listeners perform far better than the Dragon system in terms of robustness against noise. Performance degradation under noisy conditions is well-documented for all speech recognition systems based on conventional technology, like that used in the Dragon system. In contrast, the Dynamic Synapse neural network significantly outperformed the Dragon system, demonstrating a robustness superior to human listeners under highly noisy conditions.

The significance of these findings with respect to developing a neural prosthetic for replacing cognitive functions is several-fold. First, the Dynamic Synapse neural network used in the above studies is remarkably small: only 11 processing units and 30 synapses. The computational power of such a small network suggests that extremely large neural networks will not be required for developing replacement silicon-based circuitry for the brain. Second, the speaker-independent applications of the Dynamic Synapse technology were performed using an unsupervised learning algorithm, meaning that the features of the variable speech signals upon which successful word recognition were based were not identified *a priori*; the network was allowed to find an optimized feature set independently. In the context of an implantable prosthetic, this is obviously a desirable advantage in the sense that it may be reasonable to consider devices

that adapt to the host brain by optimizing a set of initial parameters. Given that we know so little about the features used in pattern recognition for many parts of the brain, depending on their *a priori* identification would represent a substantial impediment to progress. Third, the robustness of the trained Dynamic Synapse system clearly suggests that combining biologically based nonlinear dynamics with biologically based learning rules may provide a new paradigm for identifying algorithms of the brain for feature extraction and pattern recognition, and opens the possibility for studying radically novel feature sets not predictable on the basis of current theoretical frameworks.

VI. Analog VLSI Implementations of Biologically Realistic Neural Network Models

To this point, we have addressed issues concerning the first two essential requirements for an implantable neural prosthetic. We have shown that it is possible to obtain experimentally based, biologically realistic models that accurately predict hippocampal neuron behavior for a wide range of input conditions, including those known to be physiologically relevant. In addition, we have shown that the fundamental nonlinear dynamic properties of hippocampal neu-

rons can provide the basis for a neural network model which can be trained, according to biologically realistic learning rules, to respond selectively to temporal and spatio-temporal patterns coded in the form of point-process spike trains as are found in the brain. Moreover, pattern recognition by the network model can be accomplished even when input signals are embedded in substantial amounts of noise, a characteristic both of real-world conditions and of signaling in the brain. Below we will address the third essential requirement, namely, the need to implement neuron and neural network models in silicon, so that miniaturization will allow intracranial implantation.

A. Design and Fabrication of Programmable Second-Order Nonlinear Neuron Models

We have designed and fabricated several generations of hardware implementations of our biologically realistic models of hippocampal neural network nonlinear dynamics using analog VLSI technology [40]–[42]. The model expressions of the first- and second-order kernel functions describing those dynamics are computed in analog current-mode instead of digital format in order to fully exploit massively parallel processing capability. The particular objective of the design described here was to incorporate programmable second-order nonlinear model-based parameters so that a flexible, generally applicable hardware model of hippocampal nonlinearities could be developed. A fabricated and tested 3×3 neural network chip is shown in Fig. 6.

The information transmitted among neurons is encoded in the interpulse intervals of pulse trains. Different synaptic weights can be applied to the input pulse trains. Each neuron executes the convolution of a model-based second-order kernel function as

$$h_2(\Delta) = a \cdot 10^{-b \cdot \Delta} - c \cdot 10^{-d \cdot \Delta}.$$

The parameters a, b, c, d, and an $h_1(\tau)$ offset are programmable not only so that the same design can accommodate nonlinearities characteristic of different subpopulations of hippocampal neurons, but also so that training-induced modification of nonlinearities can be accommodated.

The programmable pulse-coded neural processor for hippocampal region was fabricated by a double-polysilicon triple-metal process with linear capacitor option through the MOSIS service. Each neuron contains two input stages connected to two outputs of other neurons in the network. The exponential decay in the above expression is implemented by a modified wide-range Gilbert multiplier and a capacitor. During initialization of the chip, the initial state potentials are loaded to the state capacitors. The parameter values are stored on capacitors. These analog values are refreshed regularly by off-chip circuitry and can be changed by controlling software. Bias voltages to set the multipliers and variable resistors in the correct operational modes also are required.

When operating with a 3.3-V power supply, simulation results show a 60-dB dynamic range. Depending on the com-
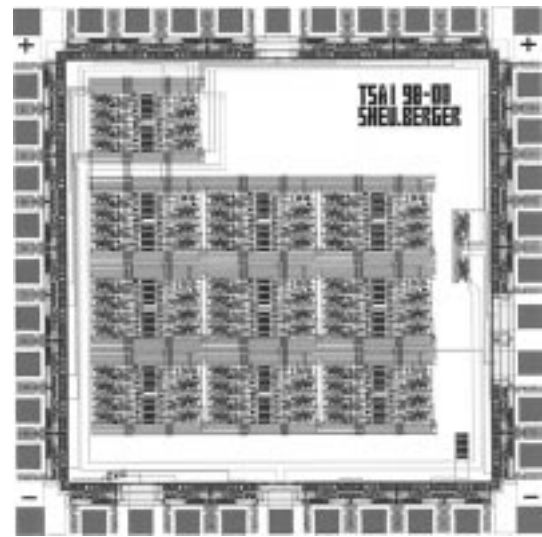
(a)
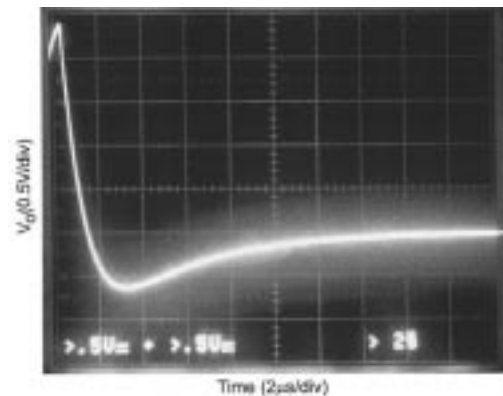
(b)

Fig. 6. Hybrid analog/digital VLSI implementation of a 3×3 network of hippocampal neuron models with second-order nonlinear properties. (b) shows a second-order kernel function generated by on-chip circuitry (compare with the second-order kernel shown in Fig. 2). The first-order kernel value and the second-order nonlinear function are programmable from off-chip circuitry.

plexity of the multiplier design, the resistance can vary from 300 Ω to 300 kΩ. If the state potential is larger than the threshold when an input pulse arrives, an output pulse is generated. Testing of fabricated chips shows reproducibility of experimentally determined I/O behavior of hippocampal neurons with a mean square error of less than 3%.

B. Design of A High-Density Hippocampal Neuron Network Processor

Although it is not yet known how many silicon neurons will be needed for an effective prosthetic, the number is likely to be in the hundreds or thousands. This demands a capability to scale up the type of fundamental design described above. To accomplish this goal, we have utilized concepts of neuron sharing and asynchronized processing to complete design of a high-density neuroprocessor array consisting of 128×128 second-order nonlinear processing elements on a single microchip [43]. Each single processor is composed of four data buffers, four indium bump flip-chip
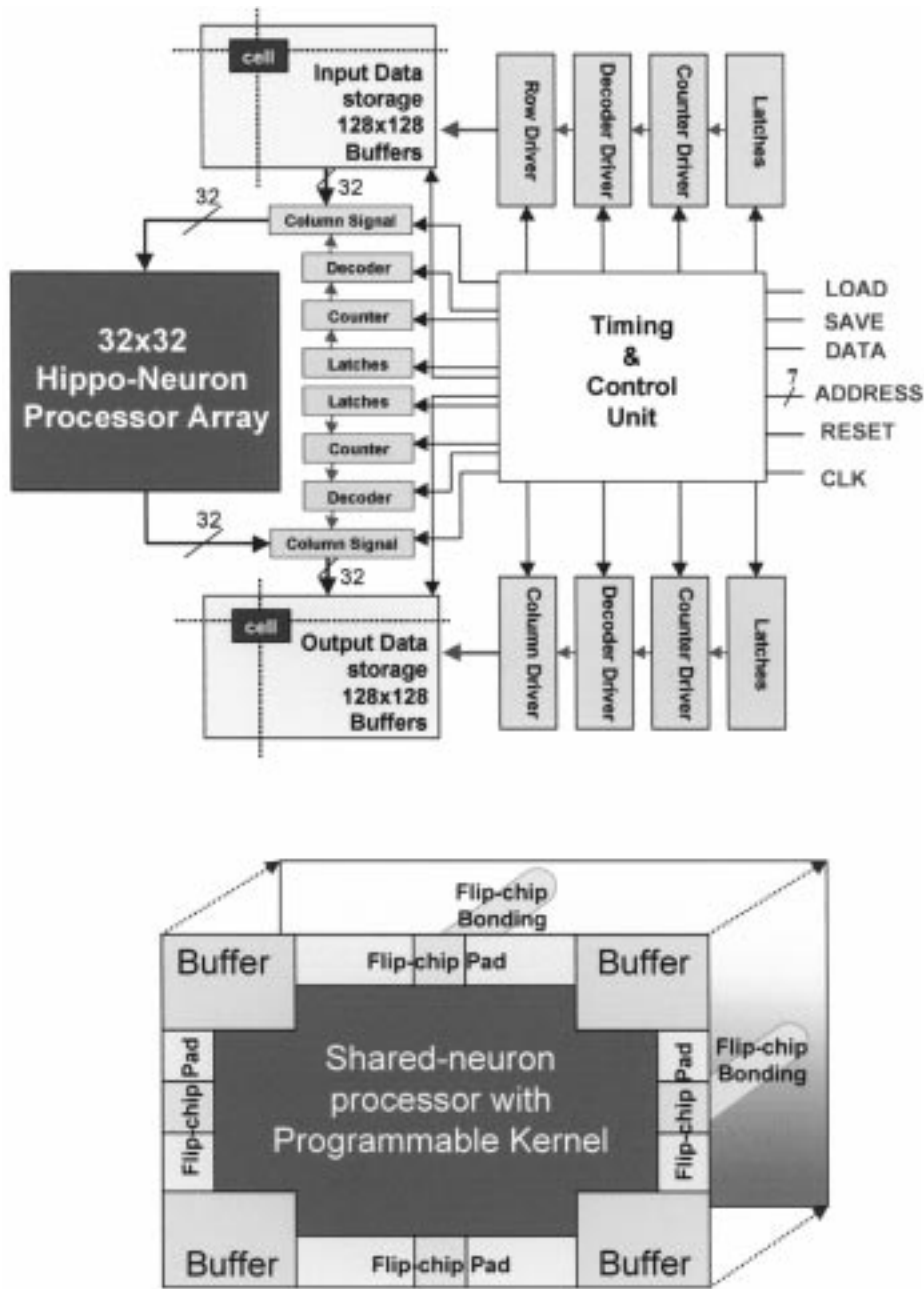
Fig. 7. Schematic diagram for a scalable version of the programmable second-order nonlinear neuron neural processor shown in Fig. 6. This layout is scalable to a 128×128 neuron model network.

bonding pads (see below), and one shared-neuron model with second-order nonlinear properties (Fig. 7). The processing procedure is as follows: 1) the input data are held in an input memory as the data arrive; 2) the input array is divided in 16 parts, with each part a 32×32 array; 3) each part of the input data is sent to the processor array, with each neuron processing four buffered data, one at a time; 4) all parameters of the kernel function are updated; and 5) after all 16 data parts have been processed, results are stored in an output buffer array.

This design provides not only for programmable kernel parameters, but also incorporates indium bumps (four per processor) for flip-chip bonding to a second connectivity matrix chip. This design (see Fig. 8) allows for considerable

connection flexibility by separating circuitry dedicated to processor dynamics from circuitry dedicated to connection architecture. With the additional technology for flip-chip bonding, the combined multichip module (not yet fabricated) will function much like a multilayer cellular neural network (CNN) structure [44].

C. VLSI Implementation of a Dynamic Synapse Neural Network

The VLSI implementation of a limited capacity dynamic synapse neural network has been designed and fabricated using TSMC 0.35 μm technology, as shown in Fig. 9 [45]. The dynamic synapse neural network chip includes six input
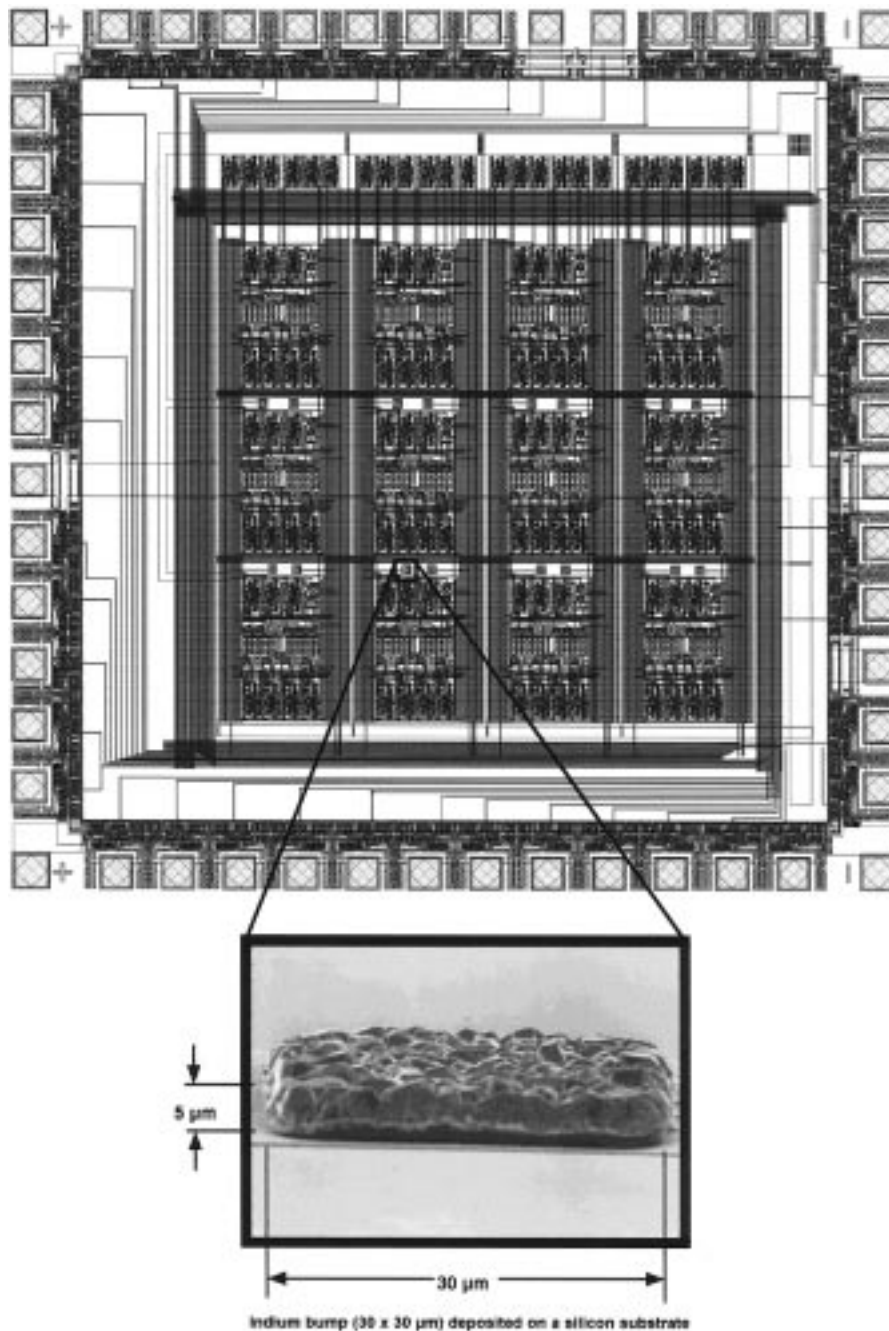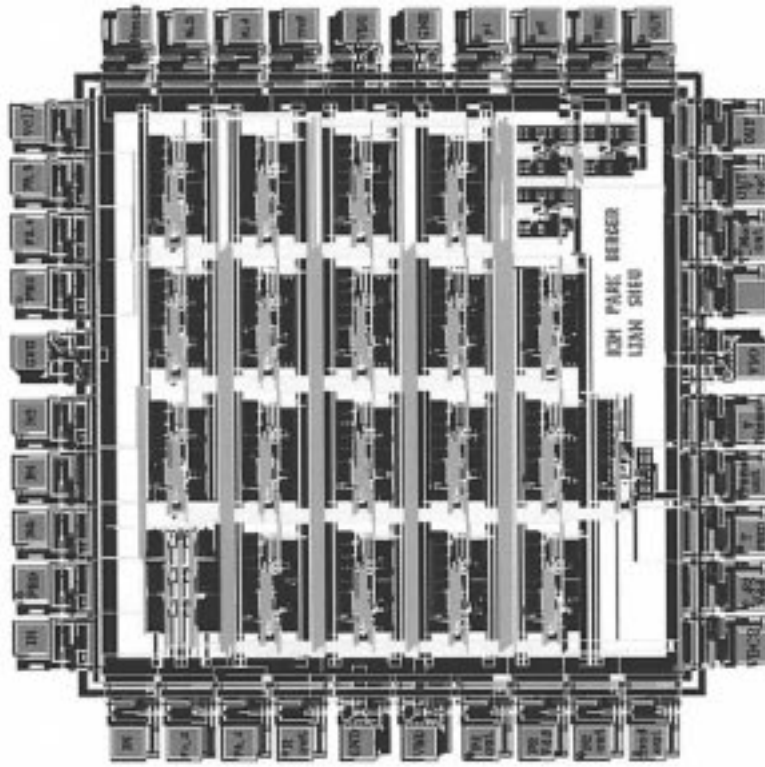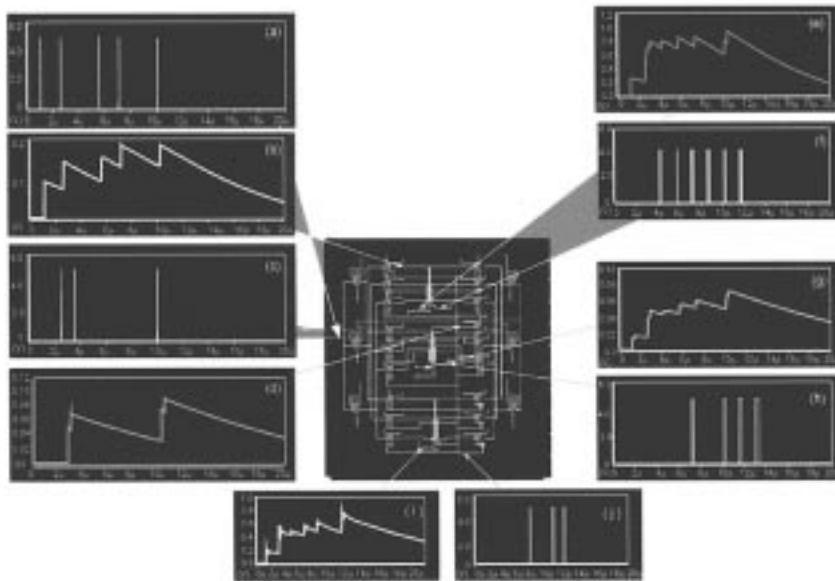
Fig. 8. Hybrid analog/digital VLSI implementation of a 4 × 4 network of hippocampal neuron models with second-order nonlinear properties designed using the layout scheme shown in Fig. 7. Also shown in the inset is an indium bump (two are included for each neuron model; one for input, one for output) that allows flip-chip bonding of this neuron processing microchip to a second connectivity microchip (not shown) so that nonlinear processor properties and network connectivity properties are incorporated in different microchips of a multichip module.

neurons, two output neurons, one inhibitory neuron, 18 dynamic synapses, and 24 I/O pads. Each synapse consists of seven differential processing blocks, two hysteresis comparators, one AND gate, two transmission gates, and biasing circuitry. As described in the previous section, the functional properties of each synapse are determined by four dynamic processes, each having different time courses. Three of the processes are excitatory, one is inhibitory; two of the processes represent different second-order nonlinearities

The resistor-capacitor exponential decay circuit for the dynamic processes was implemented using poly (poly1/poly2) capacitance and NMOS active registers to save chip area. The voltage-controlled active NMOS channel resistance and current source were used to achieve the programmability of parameter values of the dynamic synaptic neural network by controlling biases. Each differential equation processing block was implemented with fully programmable voltage controlled active resistors, poly capacitors and a current

(a)

(b)

Fig. 9. (a) Hybrid analog/digital VLSI implementation of a six-input, two-output unit Dynamic Synapse neural network. The circuit design also includes one additional processing unit as part of the output layer that functions to provide feedback to the dynamic synapses. In total, there are 18 dynamic synapses. Network connectivity is fixed. (b) Results of a circuit simulation showing input and output pulse events, and analog potentials equivalent to excitatory and inhibitory synaptic events generated in the network connections.

source. Each differential processing circuit consists of two MOSFETs for active resistors, one poly capacitor, three control MOSFETs, two transmission gates, and one inverter. A novel, efficient low-power analog summation circuit was developed without using op-amps, which require significant silicon area and higher power consumption.

The capacity of this prototype Dynamic Synapse microchip is limited (because of the small number of output

neurons), and not yet fully determined because the upper capacity depends in large part on the decoding scheme used for distinguishing different temporal patterns, an issue which is currently still under investigation. Nonetheless, the successful implementation of this neural network model demonstrates that biologically realistic nonlinear dynamics that perform a high-level pattern recognition function can be realized in hardware. We are currently working on an expanded design that will provide for 400 dynamic synapses and on-chip implementation of the dynamic learning rule used to optimize feature extraction by the network.

What we have attempted to clarify in this section are several points relevant to a hardware implementation of biologically realistic neural network models. First, nonlinear dynamics (at least to the second order) characteristic of hippocampal and other cortical neurons can be efficiently implemented in mixed analog/digital VLSI. The designs not only can be programmable, to accommodate adaptive alterations in the dynamics of the microchip neuron models, but also are scalable to substantial numbers of processing elements. Considerable flexibility can be realized by separating the circuitry implementing processing element nonlinearities from the circuitry implementing the connectivity among the elements. Processing element and connectivity microchips then can be integrated as a multichip module. Finally, a prototype of a dynamic synapse neural network capable of limited speech recognition function has been designed, fabricated, and tested, demonstrating that a biomimetic neural network performing a cognitive function of neurological interest is feasible. Although the capacity of dynamic synapse neural network microchips fabricated to date is admittedly not large, it is critical to distinguish between functionality that significantly alleviates clinical symptomatology and a functionality that reproduces capabilities of an intact brain. A stroke patient that has lost all capability for speech need not be provided with a 5000-word vocabulary to substantially improve their quality of life; a vocabulary of even 20 words would constitute a marked recovery of function. Even the next-generation microchip neural networks will have a capacity that warrants considering their future clinical use, provided other technical barriers, such as interfacing with the living brain, can be overcome.

VII. NEURON-SILICON INTERFACE

The major issues with regard to an effective neuron-silicon interface that supports bidirectional communication between the brain and an implantable neural prosthetic include: 1) density of interconnections; 2) specificity of interconnections; and 3) biocompatibility and long-term viability. The issue of density of interconnections refers to the fact that virtually all brain functions are mediated to a degree by a mass action of neural elements, i.e., changing the activity of one neuron in a system is unlikely to have any substantial influence on the system function, and, thus, on the cognitive process that depends on that function. The neuron-silicon interface must be designed so that a
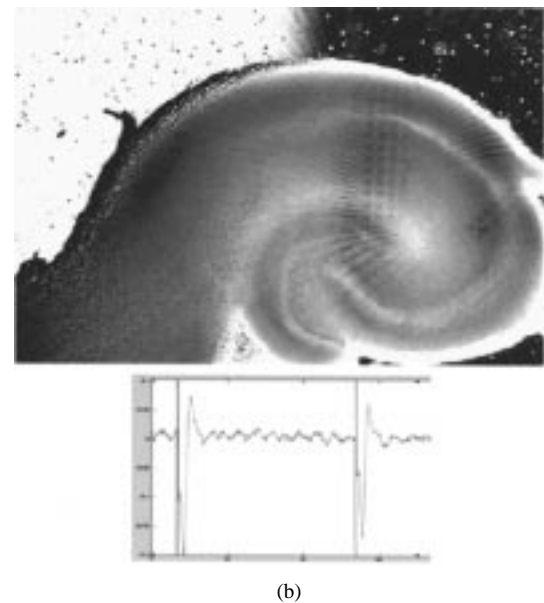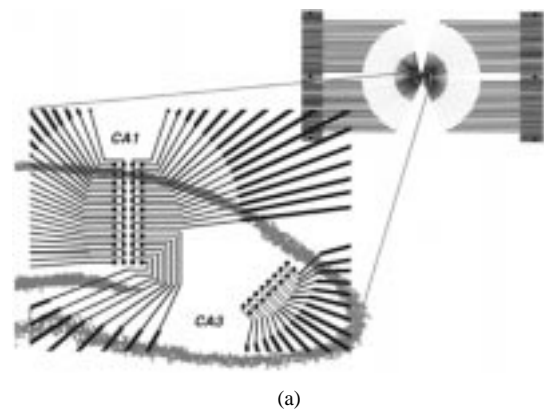

(a)

(b)

Fig. 10. (a) Schematic layout of a conformal multisite electrode array designed for electrical stimulation of CA3 inputs to the CA1 region of hippocampus. (b) Photomicrograph of a hippocampal slice positioned on a conformal array fabricated on the basis of the layout shown in (a). Bottom panel in (b): Two extracellular field potential responses recorded from one of the electrode sites in the rectangular array located in CA1 following two stimulation impulses administered to two of the electrode sites (bipolar stimulation) in the rectangular array located in CA3.

large number of neurons are affected by the implanted microchip. The issue of specificity of interconnections refers to the fact that neurons comprising a given brain region are not randomly distributed throughout the structure: the majority of brain systems have clear and definable "cytoarchitecture." For the hippocampus, the major features of this cytoarchitecture are a dense grouping of cell bodies into cell layers, with dendritic elements oriented perpendicular to those layers (see Fig. 1). The issue of specificity also extends to the organization of intrinsic circuitry: in the case of the hippocampus, the entorhinal-to-dentate-to-CA3-to-CA1-subiculum pathway is composed of different cell populations that are spatially segregated from one another. Any neuron-silicon interface must be designed to be consistent with the cytoarchitectural constraints of the target tissue. Finally, the issue of long-term viability refers
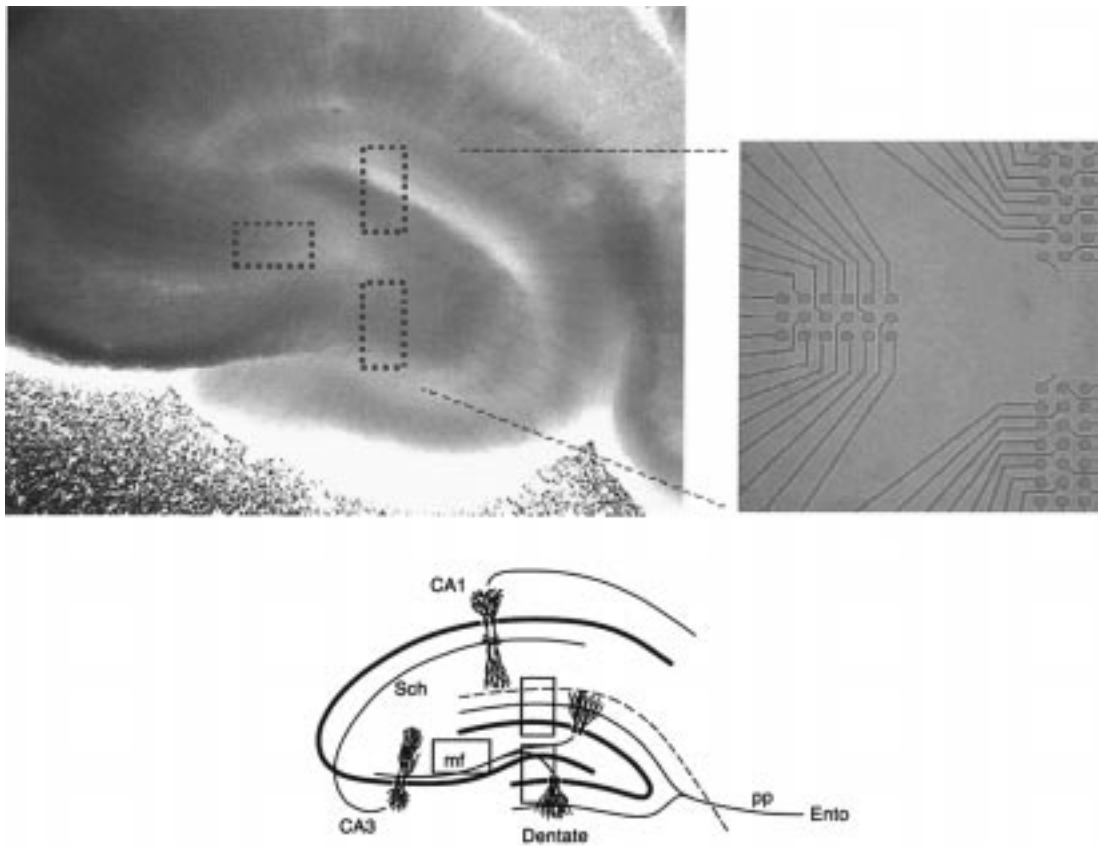
Fig. 11. Left upper panel: Photomicrograph of a hippocampal slice placed over a conformal multisite electrode array designed and fabricated for stimulation and recording of activity from the dentate gyrus and CA3 regions. Right upper panel: Detailed visualization of the three sets of electrodes included in the dentate-CA3 array: each consists of a 3 × 6 electrode site rectangular array, with the two vertically oriented arrays designed for stimulation/recording from the dentate gyrus, and the horizontally oriented array designed for stimulation/recording from the CA3 region. Bottom panel: Schematic representation of a transverse section through the hippocampus illustrating the relative locations of its subfields.

to the obvious problems of maintaining effective functional interactions between a microchip and brain tissue on a time-scale of years, as periodic replacement of an implant is not likely to be a realistic option.

A. Density and Specificity

With regard to the first two issues of density and specificity, one can either attempt to integrate these design considerations into that of the computational component of the prosthetic, or separate the computational and interface functions into different domains of the device and thus deal with the design constraints of each domain independently. We have opted for the latter strategy, developing silicon-based multisite electrode arrays with the capability to both electrophysiologically record and stimulate living neural tissue. The fundamental technologies required for multichannel bidirectional communication with brain tissue already exist commercially and are being developed further at a rapid rate [46], [47]. Silicon-based 64- and 128-electrode site recording/stimulating arrays having spatial scales consistent with the hippocampus of mammalian animal brain (significantly smaller than even that of the human) are now routinely used in ours and several other laboratories [46]–[56]. In the very near future, electrode densities sufficient to influence the majority of neurons in a two-dimensional plane of a targeted brain region will be operational.

The majority of commercially available multisite electrode arrays have a uniform geometry, however, leaving the issue of specificity unresolved. For this reason, we have focused the majority of our research with respect to neural-silicon interfaces on designing multisite arrays for which the spatial distribution of electrode sites conforms to the cytoarchitecture of the target brain region, i.e., array geometries specific to the hippocampus [57]. For example, one multisite electrode array that we have fabricated and tested was designed for CA3 inputs to the CA1 region of rat hippocampus. Two rectangular arrays were constructed using silicon nitride and indium-tin-oxide (ITO): one 2 × 8 array of electrodes oriented for stimulation of CA3 axons that course through the dendritic region of CA1, and a second 4 × 12 array positioned and oriented for recording CA1 dendritic and cell body responses evoked as a consequence of stimulation through the first array (Fig. 10). This particular conformal probe had sixty-four $40 \times 40\ \mu$m stimulating-recording pads,

a 60-μm center-to-center interelectrode distance within each array. The silicon nitride layer was deposited over the ITO electrodes, providing insulation both between the various electrodes and between each electrode and the hippocampal tissue. Silicon nitride layers were patterned to provide apertures only at the electrode tips. Silicon nitride films approximately 1500 Å thick were deposited using the plasma enhanced chemical vapor deposition (PECVD) technique. Electrical characterization using a VLSI electronic probing station showed excellent insulation capability and electrical isolation, with less than 1.8% crosstalk level on adjacent recording pads on the SiNx-insulated probes, measured over a frequency range from 100 Hz to 20 kHz with a sinusoidal waveform and 50–1000 mV-rms signal amplitudes. Experimental testing with acute rat hippocampal slices consistently demonstrated evoked extracellular field potentials with signal-to-noise ratios greater than 10 : 1.

Additional mask designs have been completed and fabricated successfully that incorporate several key modifications. First, the recording-stimulating pads have been resized to 30 μm diameters, a size approaching the diameter of a single neuron cell body. Combined with smaller center-to-center distances between pads, the smaller pad feature size will enable higher density arrays for greater spatial resolution for interfacing with a given brain region and, thus, better monitoring and control of that region. Second, several new layouts have included different distributions of stimulation-recording pads that geometrically map several subregions of the hippocampus (Fig. 11). This represents the beginnings of a library of interface devices that will offer monitoring/control capabilities with respect to different subregions of hippocampus, and ultimately other brain structures as well. In addition, more recent designs have utilized gold as the stimulation-recording electrode material to allow for higher injection current densities during stimulation. Electrical characterization of the most recent generation conformal neural probe arrays indicate, despite the higher density of electrodes, less than a 4.1% crosstalk level on adjacent recording pads.

B. Biocompatibility and Long-Term Viability

Many of the problems with respect to biocompatibility and long-term viability cannot be fully identified until the working prototypes of multielectrode arrays described above have been developed to the point that they can be tested through chronic implantation in animals. Nonetheless, we have begun to consider these issues and to develop research strategies to address them. One of the key obstacles will be maintaining close contact between the electrode sites of the interface device and the target neurons over time. We have begun investigating organic compounds that potentially can be used to coat the surface of the interface device to increase its biocompatibility, and thus, promote outgrowth of neuronal processes from the host tissue and increase their adhesion to the interface materials. Poly-d-lysine and laminin are known to be particularly effective in promoting adhesion of
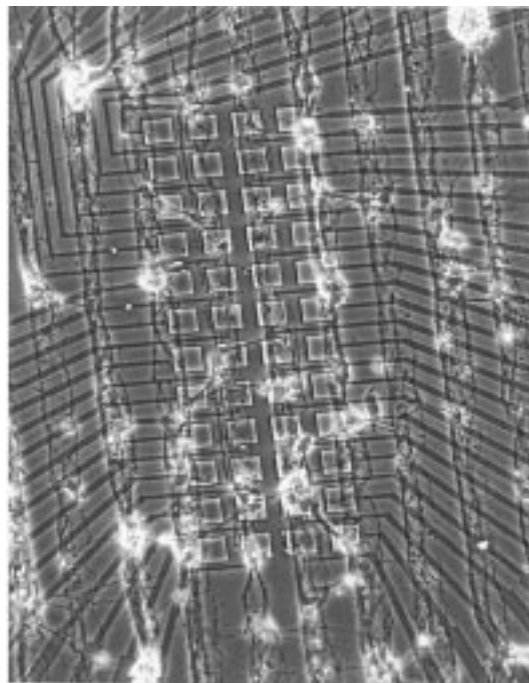

Fig. 12. Photomicrograph of the larger, 4 × 12 set of electrodes of the conformal multisite array shown in Fig. 10, i.e., the set designed for stimulation/recording from the CA1 region. This figure shows the results of culturing dissociated hippocampal neurons on the conformal array after its surface was coated with poly-d-lysine and laminin. The poly-d-lysine and laminin were applied in 40-μm-wide lines (same width as the electrode pads) oriented parallel to the long axis of the rectangular array. Note that cultured neuron cell bodies (phase-contrast bright) and their dendritic and axonal processes adhered and grew primarily along these linear paths, and for the region of the 4 × 12 array, almost directly over the electrode pads.

dissociated neuron cultures (cultures prepared from neonatal brain; neurons are prepared as a suspension and then allowed to adhere, redevelop processes, and reconnect into a network) onto inorganic materials [55], [56], and we have investigated their efficacy with regard to our hippocampal conformal multisite electrode arrays [57], [58]. Poly-d-lysine and laminin were applied to the surface of the conformal arrays shown in Fig. 10, but application was limited to linear tracks aligned with the long axis of each column of electrode sites in the rectangular array. When dissociated hippocampal neurons were prepared on the surface of the array, the adhesion of cells and the extension of their processes were restricted to the treated regions, i.e., hippocampal neurons were attracted, attached, and proliferated synaptic connections almost exclusively in parallel, linear tracks over the columns of electrodes (Fig. 12). Though this represents only an initial thrust into the issues of biocompatibility, it is through approaches such as these that we anticipate finding solutions to biocompatibility problems.

Although much of the electrophysiological testing of interfaces to date has been completed using acutely prepared hippocampal slices (which remain physiologically viable for 12–18 h), we also have begun using hippocampal slice *cultures* for testing the long-term viability of the neuron-silicon interface [49]. The latter preparations involve slices of
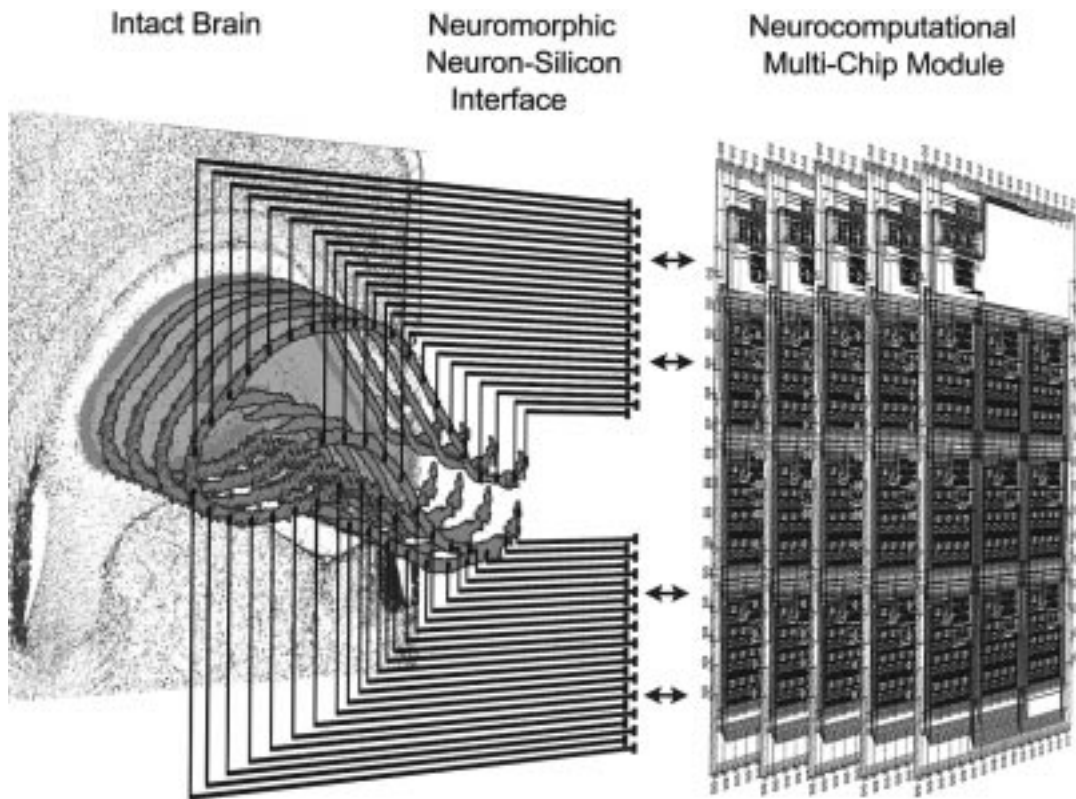
Fig. 13. Conceptual representation of an implantable neural prosthetic for replacing lost cognitive function of higher cortical brain regions. The concept is illustrated here in the context of the specific case of a prosthetic substituting for a portion of the hippocampus. The two essential components of the prosthetic system are a "neurocomputational" multichip module that performs the computational functions of the dysfunctional or lost region of hippocampus, and a "neuromorphic" multisite electrode array that acts as a neuron-silicon interface to allow the neurocomputational microchips to both receive input from, and send output to, the intact brain.

hippocampus placed onto a semipermeable membrane in contact with tissue culture media, and maintained long-term in a culture incubator [53]. Slice cultures can be prepared directly onto multisite electrode arrays, which then can be taken out of the incubator and tested periodically to examine the robustness of the electrophysiological interaction with the hippocampal tissue. Preliminary findings have revealed that bidirectional communication remains viable for at least several weeks, though we have yet to systematically test long-term functionality. The main point to be made here is that novel preparations like the slice culture will provide highly useful platforms for identifying and resolving viability issues.

VIII. CONCLUSION

The goal of this paper was to bring into focus what we believe will be one of the premier thrusts of the emerging field of neural engineering: to develop implantable, neural prosthetics that can coexist and bidirectionally communicate with living brain tissue, and thus, substitute for the lost cognitive function due to damage and/or disease (Fig. 13). Because of a convergence of progress in the fields of neuroscience, molecular biology, biomedical engineering, computer science, electrical engineering, and materials science,

it is now reasonable to begin defining the combined theoretical and experimental pathways required to achieve this end. We have shown here major progress on four of the essential requirements for an implantable neural prosthetic, in the context of a series of experimental and modeling studies concentrating on the hippocampus: biologically realistic neuron models that can effectively replace the functional properties of hippocampal cells, the concatenation of the neuron model dynamics into neural networks that can perform a pattern recognition problem of cognitive and neurological relevance, the implementation of biologically realistic neural network models in VLSI for miniaturization, and the development of silicon-based multisite electrode arrays that provide for bidirectional communication with living neural tissue.

This progress does not constitute a set of final solutions to these four requirements. Additional work is needed with respect to nonlinear models of neuron dynamics, both with respect to characterization of higher order nonlinearities and particularly with respect to cross-input nonlinearities. All neurons receive inputs from more than one other source, and interactions between separate inputs most likely results in nonlinearities specific to those interactions which cannot be characterized by our present experiments or modeling. Likewise, the Dynamic Synapse neural network models must be

expanded both in terms of number of processing elements and numbers of network layers to begin approaching the complexity and mass action of brain subsystems for which a neural prosthetic will substitute. VLSI implementations of neural network models must be scaled up as well, and better incorporate efficient and novel interchip transmission technologies to achieve the high densities required for intracranial implantation. Critically, future generations of biomimetic devices will require low-power designs to be compatible with the many temperature-sensitive biological mechanisms of the brain, an issue that our program has yet to address. Finally, there remains much concerning organic-inorganic interactions that need to be investigated for long-term compatibility between silicon-based technology and neural tissue. Although these problems are formidable, the rapid advances now occurring in the biological and engineering sciences promise equally rapid progress on the various elements of the global problem of intracranial implantable neural prosthetics, particularly given the synergy that should emerge from cooperative efforts between the two sets of disciplines.

REFERENCES

[1] G. E. Loeb, "Cochlear prosthetics," *Ann. Rev. Neurosci.*, vol. 13, pp. 357–71, 1990.
[2] M. S. Humayun, E. de Juan, Jr., J. D. Weiland, G. Dagnelie, S. Katona, R. J. Greenberg, and S. Suzuki, "Pattern electrical stimulation of the human retina," Vision Res., submitted for publication.
[3] K. H. Mauritz and H. P. Peckham, "Restoration of grasping functions in quadriplegic patients by functional electrical stimulation (FES)," *Int. J. Rehab. Res.*, vol. 10, pp. 57–61, 1987.
[4] P. Andersen, T. V. P. Bliss, and K. K. Skrede, "Lamellar organization of hippocampal excitatory pathways," *Exp. Brain Res.*, vol. 13, pp. 222–238, 1971.
[5] D. G. Amaral and M. P. Witter, "The three-dimensional organization of the hippocampal formation: A review of anatomical data," *Neurosci.*, vol. 31, pp. 571–591, 1989.
[6] T. W. Berger, T. P. Harty, X. Xie, G. Barrionuevo, and R. J. Sclabassi, "Modeling of neuronal networks through experimental decomposition," in *Proc. IEEE 34th Mid. Symp. Cir. Sys.*, 1992, pp. 91–97.
[7] T. W. Berger, G. Chauvet, and R. J. Sclabassi, "A biologically based model of functional properties of the hippocampus," *Neural Netw.*, vol. 7, pp. 1031–1064, 1994.
[8] T. W. Berger and J. L. Bassett, "System properties of the hippocampus," in *Learning and Memory: The Biological Substrates*, I. Gormezano and E. A. Wasserman, Eds. Hillsdale, NJ: Lawrence Erlbaum, 1992, pp. 275–320.
[9] H. Eichenbaum, "The hippocampus and mechanisms of declarative memory," *Behav. Brain Res.*, vol. 3, pp. 123–33, 1999.
[10] M. L. Shapiro and H. Eichenbaum, "Hippocampus as a memory map: Synaptic plasticity and memory encoding by hippocampal neurons," *Hippocampus*, vol. 9, pp. 365–84, 1999.
[11] L. R. Squire and S. M. Zola, "Episodic memory, semantic memory, and amnesia," *Hippocampus*, vol. 8, pp. 205–11, 1998.
[12] J. O'Keefe and L. Nadel, *The Hippocampus as a Cognitive Map*. Oxford, U.K.: Oxford Univ. Press, 1978.
[13] J. Magee, D. Hoffman, C. Colbert, and D. Johnston, "Electrical and calcium signaling in dendrites of hippocampal pyramidal neurons," *Ann. Rev. Phys.*, vol. 60, pp. 327–46, 1998.
[14] T. W. Berger, J. L. Eriksson, D. A. Ciarolla, and R. J. Sclabassi, "Nonlinear systems analysis of the hippocampal perforant path-dentate projection. II. Effects of random train stimulation," *J. Neurophys.*, vol. 60, pp. 1077–1094, 1988.
[15] ——, "Nonlinear systems analysis of the hippocampal perforant path-dentate projection. III. Comparison of random train and paired impulse analyses," *J. Neurophysl.*, vol. 60, pp. 1095–1109, 1988.
[16] T. W. Berger, G. Barrionuevo, S. P. Levitan, D. N. Krieger, and R. J. Sclabassi, "Nonlinear systems analysis of network properties of the hippocampal formation," in *Neurocomputation and Learning: Foundations of Adaptive Networks*, J. W. Moore and M. Gabriel, Eds. Cambridge, MA: MIT Press, 1991, pp. 283–352.
[17] T. W. Berger, G. Chauvet, and R. J. Sclabassi, "A biologically based model of functional properties of the hippocampus," *Neural Netw.*, vol. 7, pp. 1031–1064, 1994.
[18] S. S. Dalal, V. Z. Marmarelis, and T. W. Berger, "A nonlinear positive feedback model of glutamatergic synaptic transmission in dentate gyrus," in *Proc. 4th Joint Symp. Neural Comp.*, vol. 7, 1997, pp. 68–75.
[19] P. Z. Marmarelis and V. Z. Marmarelis, *Analysis of Physiological Systems: The White-Noise Approach*. New York: Plenum, 1978.
[20] W. J. Rugh, *Nonlinear Systems Theory: The Volterra/Wiener Approach*. Baltimore, MD: John Hopkins Univ. Press, 1981.
[21] R. J. Sclabassi, D. N. Krieger, and T. W. Berger, "A systems theoretic approach to the study of CNS function," *Ann. Biomed. Eng.*, vol. 16, pp. 17–34, 1988.
[22] H. Krausz, "Identification of nonlinear systems using random impulse train inputs," *Bio. Cyb.*, vol. 19, pp. 217–230, 1975.
[23] Y. W. Lee and M. Schetzen, "Measurement of the kernels of a nonlinear system by cross-correlation," *Int. J. Control*, vol. 2, pp. 237–254, 1965.
[24] V. Z. Marmarelis, "Identification of nonlinear biological systems using Laguerre expansions of kernels," *Ann. Biomed. Eng.*, vol. 21, pp. 573–589, 1993.
[25] V. Z. Marmarelis and M. E. Orme, "Modeling of neural systems by use of neuronal modes," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 1149–1158, 1993.
[26] V. Z. Marmarelis and X. Zhao, "Volterra models and three-layer perceptrons," *IEEE Trans. Neural Networks*, vol. 8, pp. 1421–1433, 1997.
[27] D. N. Krieger, T. W. Berger, and R. J. Sclabassi, "Instantaneous characterization of time-varying nonlinear systems," *IEEE Trans. Biomed. Eng.*, vol. 39, pp. 420–424, 1992.
[28] M. Iatrou, T. W. Berger, and V. Z. Marmarelis, "Modeling of nonlinear nonstationary dynamic systems with a novel class of artificial neural networks," *IEEE Trans. Neural Networks*, vol. 10, pp. 327–339, 1999.
[29] ——, "Application of a novel modeling method to the nonstationary properties of potentiation in the rabbit hippocampus," *Ann. Biomed. Eng.*, vol. 27, pp. 581–591, 1999.
[30] K. Alataris, T. W. Berger, and V. Z. Marmarelis, "A novel network for nonlinear modeling of neural systems with arbitrary point-process inputs," *Neural Networks*, vol. 13, pp. 255–266, 2000.
[31] M. Saglam, V. Marmarelis, and T. Berger, "Identification of brain systems with feedforward artificial neural networks," *Proc. World Congr. Neural Networks*, pp. 478–481, 1996.
[32] X. Xie, T. W. Berger, and G. Barrionuevo, "Isolated NMDA receptor-mediated synaptic responses express both LTP and LTD," *J. Neurophys.*, vol. 67, pp. 1009–1013, 1992.
[33] E. Thiels, G. Barrionuevo, and T. W. Berger, "Induction of long-term depression in hippocampus **in vivo** requires postsynaptic inhibition," *J. Neurophys.*, vol. 72, pp. 3009–3016, 1994.

[34] X. Xie, J. S. Liaw, M. Baudry, and T. W. Berger, "Novel expression mechanism for synaptic potentiation: Alignment of presynaptic release site and postsynaptic receptor," *Proc. Nat. Acad. Sci.*, vol. 94, pp. 6983–6988, 1997.

[35] M. Baudry, "Synaptic plasticity and learning and memory: 15 years of progress," *Neurobiol. Learning Memory*, vol. 70, pp. 113–118, 1998.

[36] J. S. Liaw and T. W. Berger, "Dynamic synapse: A new concept of neural representation and computation," *Hippocampus*, vol. 6, pp. 591–600, 1996.

[37] ——, "Computing with dynamic synapses: A case study of speech recognition," in *Proc. IEEE Int. Conf. Neural Networks*, 1997, pp. 350–355.

[38] ——, "Robust speech recognition with dynamic synapses," in *Proc. IEEE Int. Conf. Neural Networks*, 1998, pp. 2175–2179.

[39] ——, "Dynamic synapses: Harnessing the computing power of synaptic dynamics," *Neurocomputing*, vol. 26–27, pp. 199–206, 1999.

[40] R. H. Tsai, B. J. Sheu, and T. W. Berger, "VLSI design for real-time signal processing based on biologically realistic neural models," in *Proc. IEEE Int. Conf. Neural Networks*, vol. 2, 1996, pp. 676–681.

[41] ——, "Design of a programmable pulse-coded neural processor for the hippocampus," in *Proc. IEEE Int. Conf. Neural Networks*, vol. 1, 1998, pp. 784–789.

[42] ——, "A VLSI neural network processor based on a model of the hippocampus," *Analog Integr. Circuits Signal Process.*, vol. 15, pp. 201–213, 1998.

[43] R. H. Tsai, J. C. Tai, B. J. Sheu, A. R. Tanguay Jr, and T. W. Berger, "Design of a scalable and programmable hippocampal neural network multi-chip module," *Soc. Neurosci. Abstr.*, vol. 25, p. 902, 1999.

[44] L. O. Chua and L. Yang, "Cellular neural networks: Applications," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1273–1290, Oct. 1988.

[45] Y. Park, J.-S. Liaw, B. J. Sheu, and T. W. Berger, "Compact VLSI neural network circuit with high-capacity dynamic synapses," in *Proc. IEEE Int. Conf. Neural Networks*, vol. 4, 2000, pp. 214–218.

[46] U. Egert, B. Schlosshauer, S. Fennrich, W. Nisch, M. Fejtl, T. Knott, T. Müller, and H. Hämmerle, "A novel organotypic long-term culture of the rat hippocampus on substrate-integrated multielectrode arrays," *Brain Res. Protoc.*, vol. 2, pp. 229–242, 1998.

[47] O. Hiroaki, K. Shimono, R. Ogawa, H. Sugihara, and M. Taketani, "A new planar multielectrode array for extracellular recording: Application to hippocampal acute slice," *J. Neurosci. Meth.*, vol. 93, pp. 61–67, 1999.

[48] T. W. Berger, W. Soussou, G. Gholmieh, R. Brinton, and M. Baudry, "Multielectrode array recordings from the hippocampus *in vitro*," *Soc. Neurosci. Abstr.*, vol. 25, p. 902, 1999.

[49] G. Gholmieh, W. Soussou, R. Brinton, A. F. Nordholm, M. Baudry, and T. W. Berger, "Monitoring of trimethylopropane phosphate's (TMPP) neurotoxic effect on hippocampal slices and cells using multielectrode arrays," *Soc. Neurosci. Abstr.*, vol. 25, p. 902, 1999.

[50] K. Shimono, F. Brucher, R. Granger, G. Lynch, and M. Taketani, "Origins and distribution of cholinergically induced β rhythms in hippocampal slices," *J. Neurosci.*, vol. 20, pp. 8462–8473, 2000.

[51] G. W. Gross, A. N. William, and J. H. Lucas, "Recording of spontaneous activity with photoetched microelectrode surfaces from mouse spinal neurons in culture," *J. Neurosci. Meth.*, vol. 5, pp. 13–22, 1982.

[52] B. C. Wheeler and J. L. Novak, "Current source density estimation using microelectrode array data from the hippocampal slice preparation," *IEEE Trans. Biomed. Eng.*, vol. 33, pp. 1204–1212, 1986.

[53] L. Stoppini, S. Duport, and P. Correges, "New extracellular multirecording system for electrophysiological studies: Application to hippocampal organotypic cultures," *J. Neurosci. Meth.*, vol. 72, pp. 23–33, 1997.

[54] M. Han, P. Nasiatka, G. Gholmieh, W. Soussou, M. Baudry, T. W. Berger, and A. R. Tanguay, "Conformally mapped neural probe arrays for multisite stimulation and recording," *Soc. Neurosci. Abstr.*, vol. 26, p. 184, 2000.

[55] D. A. Stenger, J. J. Hickman, K. E. Bateman, M. S. Ravenscroft, W. Ma, J. J. Pancrazio, K. Shaffer, A. E. Schaffner, D. H. Cribbs, and C. W. Cotman, "Microlithographic determination of axonal/dendritic polarity in cultured hippocampal neurons," *J. Neurosci. Meth.*, vol. 82, pp. 167–73, 1998.

[56] C. D. James, R. Davis, M. Meyer, A. Turner, S. Turner, G. Withers, L. Kam, G. Banker, H. Craighead, M. Isaacson, J. Turner, and W. Shain, "Aligned microcontact printing of micrometer-scale poly-L-lysine structures for controlled growth of cultured neurons on planar microelectrode arrays," *IEEE Trans. Biomed Eng.*, vol. 47, pp. 17–21, 2000.

[57] W. Soussou, G. Yoon, G. Gholmieh, R. Brinton, and T. W. Berger, "Characterization of dissociated hippocampal neurons cultured on various biochemical substrates and multielectrode arrays for the creation of neuronal networks," *Soc. Neurosci. Abstr.*, vol. 25, p. 903, 1999.

[58] ——, "Network responses of dissociated hippocampal neurons cultured onto multi-electrode arrays," *Soc. Neurosci. Abstr.*, vol. 26, p. 1699, 2000.

Theodore W. Berger received the Ph.D. degree in physiological psychology from Harvard University, Cambridge, MA, in 1976, receiving the McKeen Cattell Award from the New York Academy of Sciences for his thesis research. He conducted postdoctoral research at the University of California, Irvine from 1977, 1978, and was an Alfred P. Sloan Foundation Fellow at The Salk Institute from 1978 to 1979.

Since 1992, he has been Professor of Biomedical Engineering and Biological Sciences at the University of Southern California, and is a member of the Executive Committee of the Program in Neuroscience. He joined the Departments of Neuroscience and Psychiatry at the University of Pittsburgh in 1979, being promoted through to the level of Full Professor in 1987. During that time, he received a McKnight Foundation Scholar Award, twice received an NIMH Research Scientist Development Award, and was elected a Fellow of the American Association for the Advancement of Science in 1986. He became Director of the Center for Neural Engineering in 1997, an organization that helps to unite the numerous USC faculty with cross-disciplinary interests in neuroscience, engineering, and the medicine. The development and application of biomimetic signal processing systems and implantable neural prosthetics is the main focus of his research.

Dr. Berger has received an NIMH Senior Scientist Award, was awarded the Lockheed Senior Research Award in 1997, was elected a Fellow of the American Institute for Medical and Biological Engineering in 1998, and was recently appointed to the Division of Space Life Sciences Science Council of the Universities Space Research Association.

Michel Baudry received the engineering degree from Ecole Polytechnique, Paris, France, in 1971 and the Ph.D. degree in biochemistry from the University of Paris VII in 1977.

He was an Associate Researcher at the National Center for Scientific Research in Paris from 1976 to 1980 before joining the Department of Psychobiology at the University of California at Irvine. He moved to the University of Southern California, Los Angeles, in 1989 and promoted to Full Professor in 1993. He has published more than 250 papers in technical journals and conference proceedings, and co-edited several technical books. He is the co-founder of Synaptics, Inc., a privately-owned company in San Jose, CA, which develops technology that makes it possible for people to communicate with the digital devices of today and tomorrow. He is also the co-founder of Eukarion, Inc., a privately-owned company in Bedford, MA, which develops new therapeutics for age-related diseases. He is on the Editorial Boards of *Neurobiology of Learning and Memory* and *e-Biomed*.

Roberta Diaz Brinton received the Ph.D. degree in psychobiology and neuropharmacology from the University of Arizona in 1984 as an NIH Predoctoral Fellow. She continued her postdoctoral training in Neuroendocrinology at Rockefeller University as an NIH Postdoctoral Fellow from 1984 to 1988. In 1987 and 1988, she was an invited scientist in The Netherlands at the Brain Research Institute, Amsterdam, and the Rudolf Magnus Institute of Pharmacology, of Utrecht University and at The Weizmann Institute, Rehovot, Israel.

She is Associate Professor of Molecular Pharmacology and Toxicology and of Biological Sciences (Neurobiology), and also serves as the Director of the USC Science, Technology and Research Program (STAR). She co-founded the biotechnology company CoCensys in 1988. Dr. Brinton joined the faculty of University of Southern California, School of Pharmacy and the Neuroscience Program in 1988 and was promoted to Associate Professor of Molecular Pharmacology and Toxicology and Neuroscience in 1994. Her laboratory was named the Kenneth T. and Eileen L. Norris Foundation laboratory for Neuroscience Research in 1999 (http://pharmweb.usc.edu/brinton-lab). Research in the Brinton Laboratory has two levels of expression, elucidation of fundamental cellular mechanisms of cognitive function and the application of those principles to the discovery and design of therapeutic molecules and devices for the treatment of disorders of the nervous system. In 1989, Dr. Brinton became director of The STAR Program, which provides junior and senior high school students in East and South Central Los Angeles the opportunity to learn science by joining a USC basic science research team as part of their high school science curriculum (http://www.usc.edu/hsc/USCSTAR/). She serves on the scientific advisory boards for the Institute for the Study of Aging and the Diabetes Insipidus Foundation. She has served on scientific review boards for the National Institutes of Health and serves on the Executive Committee of the USC Program in Neuroscience.

Jim-Shih Liaw received the Ph.D. degree in computer science from the University of Southern California in 1993.

He joined the faculty in the Biomedical Engineering Department at the University of Southern California (Los Angeles) in 1996, where he established the Laboratory for Neural Dynamics in 1998. His research involves employing mathematical models and computer simulations to study neural systems at different scales of organization and applying the knowledge of neural principles to developing new technology. He has published more than 45 articles in technical journals and conference proceedings and book chapters. His latest research focuses on the study of the role of synaptic dynamics in neural processing and the application of the dynamic synapse neural networks to pattern recognition tasks such as speech recognition.

Vasilis Z. Marmarelis was born in Mytiline, Greece, on November 16, 1949. He received the Dipl. degree in electrical engineering and mechanical engineering from the National Technical University of Athens in 1972, and the M.S. and Ph.D. degrees in engineering science (information science and bioinformation systems) from the California Institute of Technology, Pasadena, in 1973 and 1976, respectively.

After two years of postdoctoral work at the California Institute of Technology, he joined the faculty of Biomedical and Electrical Engineering at the University of Southern California, Los Angeles, where he is currently Professor and Director of the Biomedical Simulations Resource, a research center funded by the National Institutes of Health since 1985 and dedicated to modeling/simulation studies of biomedical systems. He served as Chairman of the Biomedical Engineering Department from 1990 to 1996. His main research interests are in the areas of nonlinear and nonstationary system identification and modeling, with applications to biology, medicine, and engineering systems. Other interests include spatio-temporal and nonlinear/nonstationary signal processing, high-resolution 3-D ultrasonic imaging, and analysis of neural systems and networks with regard to information processing. He is co-author of the book *Analysis of Physiological System: The White Noise Approach*(New York: Plenum, 1978; Russian translation: Moscow, Mir Press, 1981; Chinese translation: Academy of Sciences Press, Beijing, 1990) and editor of three volumes on Advanced Methods of Physiological System Modeling (1987, 1989, 1994). He has published more than 100 papers and book chapters in the area of system and signal analysis.

Alex Yoondong Park, Photograph and biography not available at the time of publication.

Bing J. Sheu (Fellow, IEEE) received the Ph.D. degree in electrical engineering from University of California, Berkeley, in 1985.

He joined the faculty at Electrical Engineering Department at University of Southern California, Los Angeles, during 1985–1998, and was promoted to Full Professor in 1997. He moved to microelectronic design automation industry in early 1999 and worked at Avant! Corporation, Fremont, CA, during January 1999–April 2000, and has been with Nassda Corporation, Santa Clara, CA, since May 2000. He served as Architecture Track Chair of IEEE ICCD Conference, and then as Technical Program Chair in 1997, General Chair in 1998. He was Technical Program Chair of 1996 IEEE International Conference on Neural Networks. He has published more than 195 papers in technical journals and conference proceedings, and co-authored/co-edited several technical books.

Dr. Sheu was a recipient of Engineering Initiation Award from the National Science Foundation in 1987; the 1991 and 1992 Best Presenter Awards at IEEE International Conference on Computer Design, co-recipient of IEEE TRANSACTIONS ON VLSI SYSTEMS Best Paper Award in 1995, IEEE Guillemin-Cauer Award in 1997, and IEEE CAS Society Golden Jubilee Award in 2000. He served as VP-Conferences in IEEE CAS Society in 1998. He was President-Elect, President, and Past President for IEEE Circuits and Systems Society in 1999, 2000, and 2001, respectively. He served as CAS Editor of *IEEE Circuits and Devices Magazine* during 1995 and 1996, Editor-in-Chief of IEEE TRANSACTIONS ON VLSI SYSTEMS during 1997 and 1998, and founding Editor-in-Chief of IEEE TRANSACTIONS ON MULTIMEDIA in 1999.

Armand R. Tanguay, Jr., received the B.S. degree in physics (*cum laude*) from the California Institute of Technology in 1971, and the M.S., M.Phil., and Ph.D. degrees in engineering and applied science from Yale University in 1972, 1975, and 1977, respectively.

He is Associate Professor of Electrical Engineering, Materials Science, and Biomedical Engineering at the University of Southern California. He is a founding member of the Center for Photonic Technology, serving as both Deputy Director and Director. He has further served as Director of the Center for Neural Engineering, and is a member of the Neuroscience Graduate Program at USC. His research interests and experience include the crystal growth and characterization of optical and optoelectronic materials; dielectric and optical thin film physics; thin film deposition technology and characterization; device processing by ion beam milling and etching techniques; electronic/photonic packaging including multichip module integration by flip-chip bonding; physical optics; the physics and technology of electrooptic, optoelectronic, and integrated optical devices (including spatial light modulators, photorefractive volume holographic optical elements, diffractive optical elements, and advanced integrated optical signal processors); photonic implementations of neural networks; smart cameras; immersive panoramic cameras; chaos in neural networks; VLSI-based multielectrode neural probes and neural unit array prostheses for the brain; hybrid biological/electronic/photonic computational modules; and the fundamental and technological limitations of optical information processing and computing.

Dr. Tanguay is a Fellow of the Optical Society of America, and has received the Yale University Harding Bliss Prize, the USC Faculty Service Award, and the Rudolph Kingslake Medal and Prize of the Society for Photo-Optical Instrumentation Engineers. He has recently been elected a Fellow of the American Association for the Advancement of Science (AAAS).