31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Using an Open-Loop Inverse Control Strategy to Regulate CA1 Nonlinear Dynamics for an *in vitro* Hippocampal Prosthesis Model

Min-Chi Hsiao*, *Student Member, IEEE*, Dong Song, *Member, IEEE*,
and Theodore W. Berger, *Senior Member, IEEE*

***Abstract*—A modeling-control paradigm to regulate output of the hippocampus (CA1) for a hippocampal neuroprosthesis was developed and validated using an in vitro slice preparation. Our previous study has shown that the VLSI implementation of a CA3 nonlinear dynamic model can functionally replace the CA3 subregion of the hippocampal slice. The propagation of temporal patterns of activity from DG→VLSI→CA1 reproduces the activity observed experimentally in the biological DG→CA3→CA1 circuit. In this project, we incorporate an open-loop controller to optimize the output (CA1) response. Specifically, we seek to optimize the stimulation signal to CA1 using a predictive dentate gyrus (DG)-CA1 nonlinear model (i.e., DG-CA1 trajectory model) and a CA1 input-output model (i.e., CA1 plant model), such that the ultimate CA1 response (i.e., desired output) can be first predicted by the DG-CA1 trajectory model and then transformed to the desired stimulation intensity through the CA1 inverse plant model. Laguerre-Volterra kernel model for random - interval, graded - input, contemporaneous – graded - output system is formulated and applied to build the DG-CA1 trajectory model and the CA1 plant model. The inverse model to transform desired output to input is also derived and validated. We validated the paradigm in hippocampal slices, and results showed the CA1 response evoked by the controlled stimulation signal reinstated the CA1 response evoked by the trisynaptic pathway.**

I. Introduction

A neuroprosthesis is a device that interfaces with the nervous system to improve or restore impaired neuronal function. In the recent past, we have seen many promising biomedical innovations helping individuals to pursue a better quality of life. For example, functional electrical stimulation (FES) has helped people with spinal cord injuries to move their arms and legs; and cochlear implants have helped restore hearing to people who have auditory impairments. A new class of neuroprostheses works through decoding neural signals which are meant to activate another region of the nervous system. Such neuroprosthesis attempts to replace the computational functions of damaged brain regions in order to restore neural communications between intact brain regions. Berger was the first to propose this idea of developing a "cognitive" neuroprosthesis to restore long-term memory formation [1], [2]. The idea was to design a biomimetic model of the nonlinear dynamics of the hippocampus – a model that would capture how hippocampal circuitry re-encodes, or transforms, incoming spatiotemporal patterns of neural activity (i.e., the contents of short-term memory) into outgoing spatiotemporal patterns of neural activity (i.e., the contents of long-term memory). Such a prosthesis requires bidirectional communication between different brain regions in order to replace the impaired neural transmissions. The idea is that once the device records normal neural signals, it transforms them into their corresponding output signals, producing activity in brain regions formerly inaccessible due to impairment. In this way, the damaged neural circuitry can be rebuilt, or at least functionally replaced.

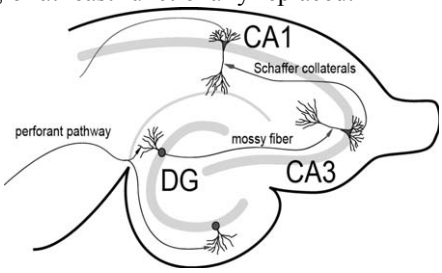

Fig. 1. The intrinsic trisynaptic pathway in a hippocampus slice.

A hippocampal slice is an ideal biological preparation to use for such a proof of concept. The major projection in a transverse hippocampal slice includes a cascade of excitatory synaptic connections involving the dentate gyrus (DG), CA3, and CA1 regions, as seen in Fig. 1. This trisynaptic circuit can be maintained in a transverse slice preparation, and each synaptic transformation is highly nonlinearly modulated. Using a typical engineering perspective, the hippocampal slice preparation can be represented as the composite of the input-output functions of the DG, CA3, and CA1 subsystems (Fig. 2A). The evoked field potentials in each subsystem are measured as input-output signals. For example, the CA3 response (field excitatory postsynaptic potentials amplitude) can be conceived as the input signal to CA1, and CA1 can be conceived as the final output system in the preparation. An analogy to the design of such a hippocampal neuroprosthesis is shown in Fig. 2B, where CA3 is assumed to be damaged, such that neurotransmission cannot be completed. In the replacement model, the intact region DG sends its outputs to

Manuscript received April 7, 2009. This work was supported in part by the NSF (BMES ERC and BITS Program), DARPA (HAND Project), ONR (Adaptive Neural System Program), NIBIB, and the Brain Restoration Foundation.

* Min-Chi Hsiao is with the Department of Biomedical Engineering at University of Southern California (USC), Los Angeles, CA 90089 USA (Phone: 213-740-8061; e-mail: mhsiao@usc.edu).

Dong Song is with the Department of Biomedical Engineering, Center for Neural Engineering at University of Southern California (USC), Los Angeles, CA 90089 USA. (e-mail: dsong@usc.edu).

Theodore W. Berger is with the Department of Biomedical Engineering, Program in Neuroscience, and Center for Neural Engineering at University of Southern California (USC), Los Angeles, CA 90089 USA. (e-mail: berger@bmsr.usc.edu).

the hippocampal prosthesis model. The signals propagate through the model to generate corresponding outputs to stimulate another intact region (CA1). The goal of this study is to regulate CA1 output activities as illustrated in Fig. 2B to correspond to the CA1 activities produced through the original biological pathway as shown in Fig. 2A.

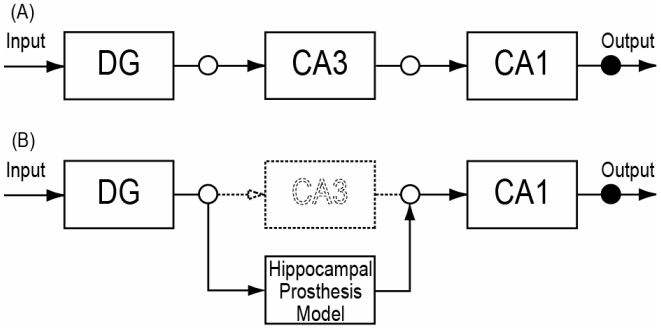

Fig. 2. A schematic diagram of (A) the trisynaptic pathway, and (B) a hippocampal prosthesis model functionally replacing the original pathway.

This paper describes the development of a model of hippocampal prosthesis in a context of a modeling- control framework. This framework includes a nonlinear modeling approach to mimic computation in the hippocampal subsystems, and an open-loop inverse controller to drive system output (Fig. 3). In this study , we applied Volterra kernel model to build DG-CA1 trajectory model and CA1 plant model. An inverse model is then formulated to transform the CA1 output into stimulation amplitudes based on the CA1 plant model. Using the desired CA1 output predicted by the DG-CA1 trajectory model and converting it through the inverse CA1 plant model, makes it possible to derive the optimal stimulation, enabling CA1 to produce activities similar to the original CA1 activities.

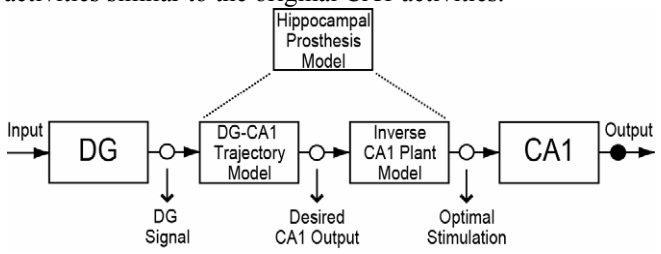

Fig. 3. Schematic diagram of proposed modeling-control frameworks for the development of a hippocampal prosthesis model.

II. Materials and Methods

Hippocampal slices from 8 to 10-week-old male Sprague-Dawley rats (250 to 300 g) were prepared. Electrophysiology data were collected through an extracellular recording technique using an MEA60 system. Details on the slice preparation and experimental setup can be found in [3].

A. Data collection

In this study, biphasic currents with a 100 μs duration in each phase were applied to all stimulation patterns. Different stimulation trains were programmed in MC_Stimulus and used to study the nonlinear properties of different regions. There was a 5 to 7 minute waiting period between each stimulus train.

1. RIT-induced trisynaptic data

An external bipolar electrode of twisted Nichrome wires was used to elicit the trisynaptic response. Paired-pulse or quadruplet-pulse electrical stimulation was applied to the perforant pathway of each slice using the external electrode to generate electrophysiological responses propagating throughout the trisynaptic pathway (evoked field potentials with progressively longer latencies in DG, CA3, and CA1). When the full trisynaptic response was observed, we stimulated the slice with a series of random inter-impulse-interval trains (RITs). Four 300-pulse Poisson distributed RITs of a fixed current intensity (biphasic, 150 to 300 μA) were delivered to the perforant path (1200 impulses; range of intervals: 2 ms to 5 s; mean frequency: 2 Hz). Response amplitudes from selected channels in DG and CA1 regions were analyzed to build the DG-CA1 trajectory model.

2. RARIT-induced monosynaptic data

Paired-pulse stimulation was applied to the *stratum radiatum* from a pair of stimulation electrodes in the conformal array in order to elicit the monosynaptic CA1 response. The pair of stimulation electrodes was selected according to their location and their ability to evoke typical paired-pulse facilitation. In this set of experiments, the amplitudes of the RIT used to evoke trisynaptic responses were modified to a Gaussian distributed amplitude (mean amplitude: 150 μA, which is the mean amplitude observed in the RIT-induced trisynaptic dataset). Once the pair of stimulation electrodes were determined, four 300-pulse random-amplitude, random-interval trains (RARITs) were delivered to the slice. A channel from the CA1 region was selected and fEPSP amplitudes were analyzed for suitability in constructing the CA1 plant model.

3. DARIT-induced monosynaptic data

In this set of experiments, the amplitude of the RARITs were reformed using the optimal stimulation amplitudes calculated from the inverse CA1 plant model (See Section II-D), called desired-amplitude RITs (DARITs). Four 300-pulse DARITs were delivered to the slice through the same pair of stimulation electrodes (as in Section II-A-2). A channel from the CA1 region was selected and the fEPSP amplitudes were analyzed, comparing them to the RIT-induced trisynaptic CA1 response amplitudes.

B. Configuration of the Laguerre-Volterra kernel model

A single-input, single-output discrete model was derived from Volterra series as expressed below:

$$y(n) = k_0 + \sum_{m=0}^{M} k_1(m)x(n-m) + \sum_{m_1=0}^{M}\sum_{m_2=0}^{M} k_2(m_1,m_2)x(n-m_1)x(n-m_2) + \ldots \quad (1)$$

The zero-order kernel k_0 is the value of output y when the input is absent. First-order kernels k_1 describe the relationship between the m^{th} input x_m and y. Second-order kernels k_2 describe the relationship between each unique pair

of input x_{m1}, x_{m2} and y. M is the number of inputs in the series and $m=0$ denotes the present input. The input to the system can be expressed as a series of variable-amplitude, random-interval delta functions:

$$x(t_i) = \sum_{i=1}^{I} A_i \delta(t - t_i) \tag{2}$$

where i is the index number of impulses and I is the total number of impulses. The time of occurrence of the i^{th} impulse is t_i. In the DG-CA1 trajectory model experiment, A_i is the DG population spike amplitude; in the CA1 plant model experiment, A_i is the RARITs stimulation amplitude. Because the input amplitude is varied, in order to isolate influence from present input, we considered the zero lag terms in (1) independently, as follows:

$$\begin{aligned} y(n) &= k_0 + k_1(0)x(n) + \sum_{m=1}^{M} k_1(m)x(n-m) \\ &+ k_2(0,0)x(n)x(n) + \sum_{m_1=1}^{M} k_2(m_1,0)x(n-m_1)x(n) \\ &+ \sum_{m_2=1}^{M} k_2(0,m_2)x(n)x(n-m_2) + \sum_{m_1=1}^{M}\sum_{m_2=1}^{M} k_2(m_1,m_2)x(n-m_1)x(n-m_2) + \ldots \end{aligned}$$

and can be then rearranged as:

$$\begin{aligned} y(n) &= k_0 + k_1(0)x(n) + k_2(0,0)x(n)^2 + \sum_{m=1}^{M} k_1(m)x(n-m) + \\ &\sum_{m_1=1}^{M}\sum_{m_2=1}^{M} k_2(m_1,m_2)x(n-m_1)x(n-m_2) + 2\sum_{m=1}^{M} k_2(m)x(n)x(n-m) + \ldots \end{aligned} \tag{3}$$

In order to reduce the number of open parameters, estimation of the kernels is facilitated by expanding them on the orthonormal Laguerre basis functions L:

$$L_l(m) = \alpha^{(m-l)/2}(1-\alpha)^{1/2}\sum_{k=0}^{l}(-1)^k \binom{m}{k}\binom{l}{k}\alpha^{l-k}(1-\alpha)^k$$

where α is the discrete-time Laguerre parameter ($0 < \alpha < 1$). The convolution of Laguerre basis functions L and inputs x can be represented as

$$v_l(t_i) = \sum_{t_i - \mu < t_j \le t_i} A_j L_l(t_i - t_j)$$

where t_i is the time of occurrence of the current impulse in the input-output sequence and t_j is the time of occurrence of the j^{th} impulse prior to the present impulse within the kernel memory window μ. The adapted Laguerre expansion of Volterra kernels with L basis functions can be rewritten as:

$$\begin{aligned} y(t_i) &= c_0 + A_i c_1(0) + A_i^2 c_2(0,0) + \sum_{l=1}^{L} c_1(l)v_l(t_i) \\ &+ \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} c_2(l_1,l_2)v_{l_1}(t_i)v_{l_2}(t_i) + 2A_i\sum_{l=1}^{L} c_2(l)v_l(t_i) + \ldots \end{aligned} \tag{4}$$

where c_0, c_1, c_2… are the kernel expansion coefficients. Since the number of basis functions (L) can be made much smaller than the memory length, the number of open parameters is greatly reduced by this expansion technique.

C. Model parameter estimation and kernel reconstruction

The kernel expansion coefficients (c_0, c_1, c_2…) can be estimated via the least-squares method, and can be used to reconstruct the Volterra kernels (k_i) using Laguerre basis functions.

D. Implementation of the inverse Laguerre-Volterra kernel model

The inverse model was built to transform the output (i.e., desired output of a CA1 region) to the input (i.e., desired input stimulation to a CA1 region). In order to develop the inverse model such that it would be based on the Laguerre-Volterra (LV) model, the original equation (4) was rearranged to (5):

$$\begin{aligned} &[c_2(0,0)]A_i^2 + [c_1(0) + 2\sum_{l=1}^{L} c_2(l)v_l(t_i)]A_i \\ &+ [c_0 + \sum_{l=1}^{L} c_1(l)v_l(t_i) + \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} c_2(l_1,l_2)v_{l_1}(t_i)v_{l_2}(t_i) - y(t_i)] = 0 \end{aligned} \tag{5}$$

In Eq. (5), the desired output y and the coefficients c_0, c_1, c_2… were obtained during process of model estimation. All the convolution terms could also be determined using the coefficients and previous stimulation amplitudes A_{i-1}, A_{i-2}…. Because all the terms are determined, Eq. (5) becomes a quadratic equation with unknown desired input stimulation (A). It can be simplified as

$$aA^2 + bA + c = 0 \tag{6}$$

where

$$a = c_2(0,0)$$

$$b = c_1(0) + 2\sum_{l=1}^{L} c_2(l)v_l(t_i)$$

$$c = c_0 + \sum_{l=1}^{L} c_1(l)v_l(t_i) + \sum_{l_1=1}^{L}\sum_{l_2=1}^{L} c_2(l_1,l_2)v_{l_1}(t_i)v_{l_2}(t_i) - y(t_i)$$

such that the transformation of the inverse model (output to input) becomes an operation of solving A in Eq. (6). The roots of the quadratic equation can be solved as follows:

$$A_+ = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{; and} \quad A_- = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

In this study, the roots are all calculated from A_+ term, because it provides positive solutions.

E. The modeling-control framework experiment protocol

In this study, six experiments were conducted. The experiment protocol to regulate CA1 nonlinear dynamics for an *in vitro* hippocampal prosthesis involves the following steps:

1. Stimulating the perforant path with RITs, and analyzing the trisynaptic responses in DG and CA1.
2. Building a DG-CA1 trajectory model using an LV kernel modeling approach, and predicting desired CA1 outputs.
3. Stimulating the Schaffer collaterals with RARITs, and analyzing the monosynaptic responses in CA1.
4. Building a CA1 plant model using the LV kernel modeling approach, and formulating an inverse CA1 plant model.
5. Applying the desired CA1 output from Step 1 to the inverse CA1 plant model to calculate the optimal stimulation amplitudes.
6. Stimulating Schaffer collaterals with the optimal stimulation from Step 5, and analyzing the responses in CA1.
7. Comparing the CA1 responses from Step 1 to those in Step 6.

III. Result

A. The DG-CA1 trajectory model and the prediction results

The RIT-induced hippocampal trisynaptic data were analyzed for use in building the DG-CA1 trajectory model. The amplitudes of evoked DG population spikes were used as measures of the input to the system, and the amplitudes of evoked CA1 fEPSPs were used as measures of output of the system. Model estimation was completed using PS amplitudes and the intervals of the input-output sequences. From six datasets, the slice response amplitudes were analyzed and compared with the predicted amplitudes. The mean NMSE was 5.85±1.15%. This shows that this trajectory model captures well the nonlinearity between DG and CA1.

B. The CA1 plant model and the prediction results

The RARIT-induced CA1 monosynaptic data were analyzed for use in building the CA1 plant model. The random amplitudes of the RARITs were used as measures of input to the system, and the amplitudes of evoked CA1 fEPSPs were used as measures of the output of the system. The LV kernel model was applied to study the nonlinearity of the CA1 system. Model estimation was completed using stimulation intensities and intervals of the input-output sequences. The NMSE between slice response and model prediction was 5.61±4.99%, averaged from six datasets. This low NMSE shows that the CA1 plant model can accurately predict CA1 amplitudes based on stimulation amplitudes.

C. The inverse CA1 plant model implementation and validation results

The implementation of the inverse CA1 plant model is accomplished using RARIT-induced monosynaptic data. The output predictions were acquired from the CA1 plant model, and applied as the y in Eq. (5). The input stimulations (A) were calculated as described in the previous section. The operations for solving the root were run through all data points in order to process the transformation from output into input.

D. The modeling-control framework simulation results

Following the protocol in the previous section, CA1 desired output is first predicted through the DG-CA1 trajectory model, and then applied to the inverse CA1 plant model to derive the optimal stimulation amplitudes. The amplitudes were used to formulate DARITs as described in Section II-A-3. The DARITs were sent into the slice and the monosynaptic CA1 responses were recorded. The proposed modeling-control framework was intended to evoke from CA1 activities similar to the original CA1 activities. Thus, DARIT-induced monosynaptic CA1 amplitudes were compared with RIT-induced trisynaptic CA1 response amplitudes. Results from six experiments are reported here, examples from two of those six are shown in Fig. 4. The accuracy of the comparison was evaluated using NMSE of the amplitude and the average NMSE was 15.72±8.17%.

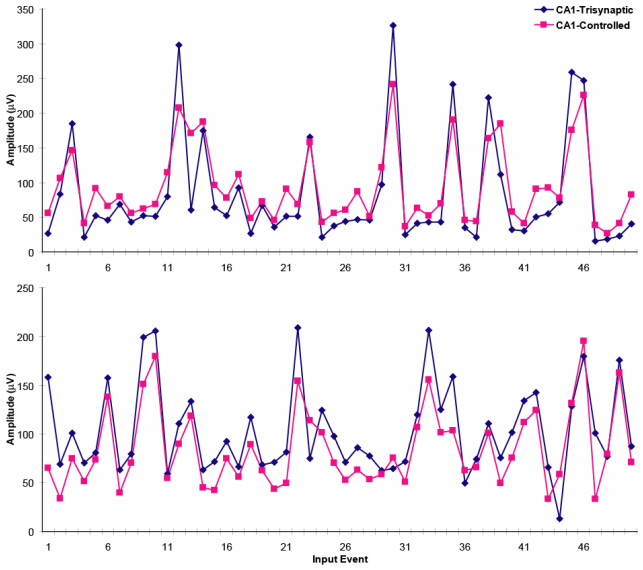

Fig. 4. A comparison of CA1 field EPSP amplitudes in response to RIT-induced trisynaptic response (CA1-trysynaptic in blue diamonds) and DARIT-induced monosynaptic response (CA1-Controlled in red squares)

IV. Discussions and Conclusions

In this study, an LV modeling approach was applied to build the DG-CA1 trajectory model and CA1 plant model. The present input m=0 terms in the discrete form of Volterra series was moved out (Eq. 3), thus when constructing the model using Laguerre expansion technique, the coefficients (Eq. 4) were estimated independently. This method manifests the contribution from the present input amplitude. Both models can accurately predict the system outputs for broadband inputs, and the small NMSE values indicate that the kernels computed from the experiment datasets sufficiently capture the nonlinear dynamics of each model. The implementation of the inverse model was carried out by solving a quadratic equation derived from a LV model. This operation allows us to convert the desired output response amplitudes to input stimulation amplitudes in a dynamic, recursive manner. In conclusion, a control theory-based paradigm was formulated to regulate CA1 dynamics for our *in vitro* hippocampal prosthesis. This new paradigm essentially predicted CA1 desired output through DG-CA1 trajectory model, and then was applied to the inverse CA1 plant model to derive the optimal stimulation amplitudes. Lastly, the optimal stimulations drive CA1 to the optimal responses. The low NMSE shows that the control theory-based paradigm could drive CA1 system to our desired output. In the future, we will incorporate this paradigm into a real time processing platform.

V. References

[1] T.W. Berger *et. al.*, IEEE Eng. Med. Biol. Mag., 2005
[2] D. Song *et. al.*, IEEE Trans. On Biomed. Eng., 2007
[3] M-C Hsiao *et. al.*, Proceedings of the IEEE-EMBC, 2008