Proceedings of the 2005 IEEE
Engineering in Medicine and Biology 27th Annual Conference
Shanghai, China, September 1-4, 2005

Information Capacity of Brain Machine Interfaces

Gregory J Gage
Department of Biomedical Engineering
University of Michigan
Ann Arbor, MI 48109, USA
Email: gagegreg@umich.edu

Edward L Ionides
Department of Statistics
University of Michigan
Ann Arbor, MI 48109, USA
Email: ionides@umich.edu

Daryl R Kipke
Department of Biomedical
and Electrical Engineering
University of Michigan
Ann Arbor, MI 48109, USA
Email: dkipke@umich.edu

***Abstract*— Brain Machine Interfaces (BMIs) are emerging as an important research area in clinical therapy. A large range of potential BMI control signals can be found in the brain. In increasing order of volume of brain tissue being sampled, these signal includes recordings of electric discharges from multi unit activity (MUA), summed population activity of thousands of neurons via local field potentials (LFPs), and electrical activity recorded from either the surface of the brain via electrocorticograms (ECoGs) or the surface of the scalp via electroencephalograms (EEGs). While each of these signals have been studied separately, it has been difficult to compare the potential that each signal has for general prosthetic control across studies. Information theory has been proposed as an abstract measurement to bridge this gap, however the maximum information rates of any experiment is limited by the parameters defined by that experiment (*e.g.* inter-trial interval length, number of targets). Here we propose a different measure of information, which we call *information capacity*, which measures the maximum possible information rate that a signal can provide. An advantage of measuring information capacity is that it can readily be compared between different signals and different tasks. We show how to calculate information capacity making linear Gaussian assumptions, and we discuss more general possibilities. We present a case study involving a rat BMI task involving either MUA or LFP signals.**

I. Introduction

A brain machine interface (BMI) is a communication channel between the brain and the external environment. The channel typically interfaces with the brain via electrodes on or near the surface of the brain, or implanted in the brain. The channel interfaces with the external environment via a prosthesis, which may be a physical device or a computer monitor. A motivation for studying BMIs is the possibility of developing prostheses to overcome various physical handicaps. The flow of information may be in either direction along the channel. For an output BMI (appropriate to spinal injury or ALS patients) the aim is to extract intention from the brain to control a prosthesis. The reverse scenario is an input BMI (appropriate for overcoming hearing loss or blindness). Here we consider output BMIs – the input case has similar considerations but is beyond the scope of this work.

The theory of Shannon information [1], [2] provides a way to evaluate the effectiveness of a communication channel. Previous work has used Shannon information to measure ability of a BMI at specific tasks [3] – we call this *task information*. This work develops a task-independent measure which we call *information capacity* (IC). Information capacity is related to the theoretical channel capacity of the BMI. Information capacity is a particularly useful concept for comparing different BMIs, possibly applied to different tasks.

Section II(A) gives an overview of Shannon information and defines our measure of information capacity. Section III presents a case study where information capacity is used to compare different modalities of brain interfaces for a rat BMI. Section IV discusses other situations where information capacity may be a useful quantity to consider. A general calculation of information capacity under Gaussian assumptions is relegated to an appendix.

II. Methods

A. An information-theoretic framework for BMIs

In the classical communication setting [1, Chapter 8], a source message x_t drawn from a known probability distribution is to be sent through a communication channel. The source is encoded by some known function to give the transmitted signal $g(x_t^h)$ where $x_t^h = \{x_s,\ s \le t\}$, the history of x_t. The communication channel adds noise, giving rise to the received signal $y_t = g(x_t^h) + q_t$. y_t is then decoded to give $\hat{x}_t$. Possibly one is more interested in decoding some function of x_t, say $z_t = h(x_t^h)$. This is estimated by $\hat{z}_t$ which might be the plug-in estimator $\hat{z}_t = h(\hat{x}_t^h)$ or some other function of y_t. The *information rate* between x_t and y_t is the supremum of possible rates of error-free transmission when x_t encodes a digital signal. The *channel capacity* is the supremum of the information rate over choices of $g(.)$ in some appropriate class of functions.

In the BMI setting, x_t quantifies the intention of the subject. This is a difficult quantity to study, and cannot usually be measured directly. In animal experiments, the subject's intention is manipulated by giving a reward for successfully completing a task. It is then assumed that the subject is intending to complete the task; caution is required when the subject is not attending to the task. y_t is a measure of neural activity. z_t is taken to be the goal of the task, whereas x_t must describe how the task is carried out. In an item selection task where the subject moves a cursor to select one of several objects, x_t is the intended cursor motion and z_t is the intended object selection. In a tracking task where the subject is asked to move a cursor along a specified route, the intention process and the task coincide ($x_t = z_t$).

0-7803-8740-6/05/$20.00 ©2005 IEEE.

Determining $g(.)$ is the neural encoding problem. Progress on this has been made in certain systems [2], but in practice for BMIs a simple linear model has proved applicable [4], [5]:

$$y_t = Hx_t + q_t, \quad q_t \sim N(0, Q) \tag{1}$$

where $x_t \in R^m$, $y_t \in R^n$ and H is an $n \times m$ matrix. To decode the neural signal, we need to make an assumption about how x_t is generated. x_t has been previously taken to be the solution to a linear, Gaussian difference equation [6], [5]

$$x_{t+1} = Ax_t + w_t, \quad w_t \sim N(0, W). \tag{2}$$

One advantage of this model is that it is analytically convenient – the Kalman filter algorithm provides an optimal Bayesian decoding to find $\hat{x}_t$. Another advantage is that it generalizes to any task requiring the same intention space. However, the experiment cannot usually enforce the subject's intention to be a random draw from this distibution. Typically, the mean and covariance of x_t are set to match the mean and covariance of an ideal response that earns the reward for the subject. A leap of faith or a body of experimental evidence is required to show that the model is relevant and reasonable. The utility of this model is already firmly established as a way of decoding intention in BMIs. Section III gives a demonstration of information capacity calculated using this model. If information capacity provides a quantity that is experimentally stable, and is found to provide a level playing field for comparing different experiments and different BMIs, we deduce that the violations of model assumptions are not great enough to upset the utility of the model.

The *task information* is the information rate between z_t and $\hat{z}_t$. For a selection task with objects labeled $1, \ldots, K$ the Shannon information per trial is

$$\mathcal{I}(z_t, \hat{z}_t) = \sum_{j=1}^{K} \sum_{k=1}^{K} p_{z\hat{z}}(j,k) \log_2 \left(\frac{p_{z\hat{z}}(j,k)}{p_z(j) p_{\hat{z}}(k)} \right)$$

where $p_{z\hat{z}}(j,k) = P\{z_t = j, \hat{z}_t = k\}$. Assuming that the trials are independent, the information rate $\overline{\mathcal{I}}(z_t, \hat{z}_t)$ is $\mathcal{I}(z_t, \hat{z}_t)$ divided by the mean time per trial.

The *information capacity* is the information rate between x_t and y_t. The information rate, in bits per unit time, may be calculated for Gaussian models using a spectral method: for the model given by (1) and (2) in the particular case where x_t is a one-dimensional process,

$$\begin{aligned} \overline{\mathcal{I}}(x_t, y_t) \quad = \quad & -\int_0^{1/2} \log_2 \Big\{ 1 - H' \big[HH' + (1 + A^2 \\ & -2A\cos(2\pi\nu)) Q/W \big]^{-1} H \Big\} \, d\nu. \end{aligned} \tag{3}$$

For certain non-Gaussian models, the calculation of information capacity may still be possible [7]. We view refinements in calculating information capacity as being, for now, secondary to the primary issue of developing information capacity as a tool for investigating BMIs. Differences between BMIs and more conventional communication channels are that $g(.)$ is a property of the BMI (rather than a parameter to be selected by the communications engineer) and x_t cannot be directly observed. If $g(.)$ is thought of as fixed, then the information capacity is the conventional channel capacity. In practice $g(.)$ varies, due to neural plasticity and/or instability of the recording device, and must be estimated in an adaptive way. We have called $\overline{\mathcal{I}}(x_t, y_t)$ *information capacity* of the BMI rather than the *channel capacity* to highlight the differences from the standard communication theory framework.

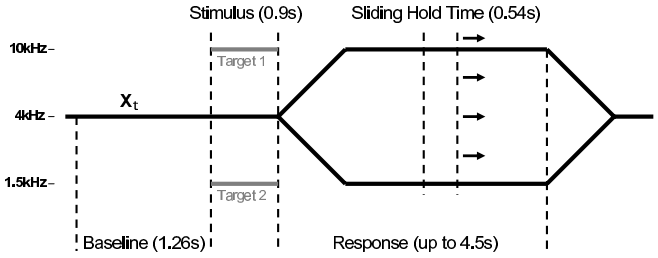

Fig. 1. Behavioral Paradigm. Trials were composed of a fixed baseline period (1.26s) followed by a 0.9s presentation of a target tone. Subjects had up to 4.5s to match the target tone and hold it within a criterion window for the required hold time (0.54s). The auditory tone frequencies that were played back to the subjects were the initial target cue followed by 90ms (or 30ms) feedback pips during the response window. The unobserved ideal response or "intended" response (x_t) is shown here for a two-target task. Trials were separated by a random intertrial window.

B. BMI paradigm used in case studies

The experiments analyzed below used 177 sessions from 7 rats [4]. Each session consisted of 100 trials in which a cortical control system was used to perform a one-dimensional auditory analog of a center out reaching task [8]. In center out reaching experiments, the hand (or cursor, in the case of brain controlled tasks) is held at the center of a circle until a target cue is placed in one of a fixed number of points on the perimeter of the circle. The subject's task is to move the hand (cursor) into the target position and hold it for a fixed amount of time. Target acquisition must be completed within an allowable response period. In our auditory version, an audio cursor is represented by sound pips representing the one-dimensional location within the logarithmically-spaced 250Hz to 16kHz frequency spectrum. Baseline firing rates were mapped to the center of the frequency space, and trials began with the presentation of a target tone at a given frequency. Subjects had a fixed amount of time to match the target frequency using the auditory cursor. The movement of the auditory cursor was dependent on the real-time decoding of the cortical signals as described below. As with the center-out paradigm, trials are marked either as "correct" (held at correct target frequency for the hold period), "wrong" (held at an incorrect target frequency for the hold period), or "late" (no answer within the response period). Timing for a single trial in these experiments is shown in figure 1.

A Kalman filter was implemented to decode, or infer, an estimate of the control signal ($\hat{x}_t$) for the auditory cursor from the observed electrophysiological signal (y_t). This signal was either a train of 90ms spike bins (n=6 rats) or 30ms RMS values of the local field potential (n=1) of the motor cortex.

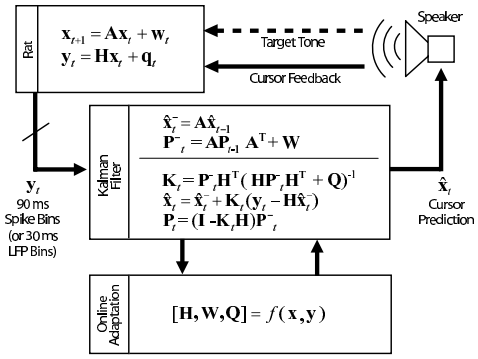

Fig. 2. Closed loop cortical control schematic. Spike bins (or binned RMS of LFPs), $\mathbf{y}_t$, from the motor cortex were decoded using a Kalman filter to predict the cursor frequency ($\hat{x}_t$). The predicted frequency was fed back to subjects via a speaker every 90ms (30ms for LFP) of the response window.

After each trial, an adaptive algorithm iteratively modified the filter parameters to minimize the squared error between the predicted control signal, $\hat{x}_t$, and the "intended" control signal, x_t. A schematic of the closed loop cortical control system is shown in figure 2.

III. Experimental Results

A. Information Capacity of Multi Unit Recordings

Offline analysis of *information capacity* was performed on data obtained from 6 rats that were trained to control the auditory cursor to reach a fixed target over 24 sessions. In these experiments, rats started from a naïve state, and learned to control the cursor in 6.8±1.2 (mean±SD) sessions. Information capacity was calculated for each trial of a session using equation (3) with the parameters H, K, W, A that were adapted and used for each trial. The mean information capacity across all trials in a session was calculated for each subject.

Figure 3 shows the results for the mean performance and the mean information capacity as a function of training session. The average amount of information capacity for all subjects in this task was 0.32 bits per 90ms bin, or 3.6 bits/s. Regression analysis showed that the amount of IC increases significantly as the subject learns to control the auditory cursor over the first 13 sessions ($p<0.01$, one-way ANOVA), while the last 11 sessions did not show evidence of an increasing trend (*ns*, one-way ANOVA). However, the performance (task information) of the subjects continued to show a positive trend during these final 11 sessions ($p<0.01$, one-way ANOVA). This suggests that the control signal may have reached it's information limit, however the subjects were able to better use the information contained in the signal to complete the required task.

We examined the spectrum of equation 3 for one subject (KCD-05). Figure 3 shows the spectrum of 3 trials (04, 27, 52) from session 24. Note that most of the power is contained in frequencies < 0.1 cycles/bin, indicating that the control signal has the largest affect on the order > 1.11s.

The actual *task information* was lower than the information capacity. For example, in the MUA experiments using two response states (maximum of 1 bit of information per trial), subjects eventually were able to perform at a level of 0.7

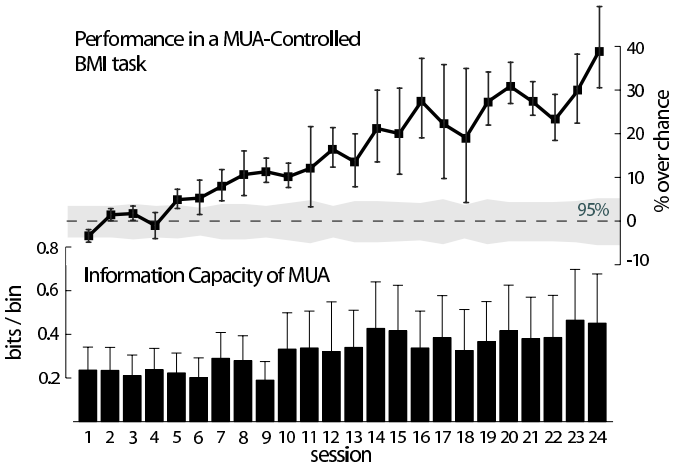

Fig. 3. Information Capacity (IC). The amount of information capacity contained the in the 90ms spike-binned signal as a function of the number of training sessions for rats (n=6) for a single target reaching task. The mean IC for all session was 0.32 bits/bin, which for a 90ms bin is equivalent to 3.6 bits/s. Error bars indicate standard error.

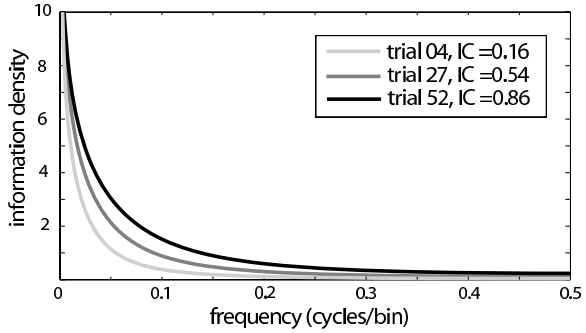

Fig. 4. Information capacity spectrum for for 3 trials (04, 27, 52) from one subject (KCD-05) on session 24 of a single target task. The information capacity spectrum is defined as the integrand in (3).

bits/trial. The average length of a trial was 5.9s, thus providing for an average *task information rate* of 0.11 bits/s.

B. Information Capacity of Local Field Potentials

We also analyzed information capacity from 11 behavioral sessions from one subject who was trained to control an auditory cursor to a fixed target using local field potentials of the motor cortex (figure 5). The mean information capacity across all trials was 0.13 bits/bin, or for a 30ms bin used in this study, 4.3 bits/s. The regression coefficient of the amount of IC contained in the LFP signal as a function of training session was small, positive and significant (0.002 bits per bin/session, $p<0.01$, one-way ANOVA).

IV. Discussion

During MUA controlled experiments, there were changes in the number of units used to control the cursor from session to session (mean = 11.5 units per session, coefficient of variance = 30.1%). We found that the number of units had a significant effect on the amount of IC in the signal (0.01 bits per bin/unit, $p<0.01$, one-way ANOVA). This was expected as there was little correlation between sites and

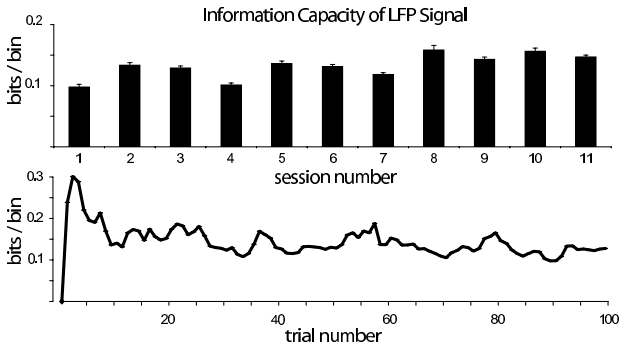

Fig. 5. Top: mean information capacity of all LFP sessions. Bottom: Information capacity of LFP signal across the 100 trials of Session 9.

thus the information capacity should increase approximately linearly with the number of sites. This relationship did not confound our finding that IC improved over training, as the number of units did not have significant trend with respect to session number. The LFP signals, on the other hand, are highly correlated between sites so the information capacity should not increase much with additional sites. Although we find here that LFP has a slightly higher information capacity than MUA, we expect that this relationship would change if larger electrode arrays were used. The LFP experiments analyzed here used 16 channels for control on all sessions.

Since information capacity relies on adaptively estimating the parameters in the model, one must check that parameter estimation does not have too large an effect on the estimated information capacity. When too little data are available to estimate the parameters, our experience suggests that overfitting to noise increases the estimated information capacity (as can be seen in the first 7 trials of Figure 5).

Note that the scale of the unobserved intention process x_t is arbitrary. Changing the scale of x_t causes corresponding scale changes in W, H and Q that cancel out in the calculation of information capacity. Different tasks may have intention processes on different scales, but that in itself will not affect information capacity. The parameters H and Q represents the ability of a subject to modulate the signal in response to a change in intention. This may depend on the task being attempted, but one can argue heuristically that the dependence should be weak: for example, if a cursor is being used to carry out a task, the ability to control the cursor may be transferable across different cursor-based tasks.

A natural question to ask is how close one can come to achieving the information capacity by a careful choice of task. Previous BMI research using rhesus monkeys [9] has shown task information rates as high as 5.6 bits/s, based on rapid sequences of trials.

We do not propose that information capacity should replace task information as a way to quantify BMI performance, but information capacity does complement task information by being only weakly dependent on the task, and allows for comparisons that would not be possible simply using task information.

APPENDIX

Theorem 4 of [10] gives a spectral formula for the rate of Kullback-Leibler divergence between jointly Gaussian stationary processes. Since mutual information is the Kullback-Leibler divergence between the joint distribution and the product distribution (if x_t and y_t had the same marginal distributions, but were independent), this theorem amounts to

$$\begin{aligned}\overline{\mathcal{I}}(x_t, y_t) &= \int_0^{1/2} \Big\{ \log|\Gamma_x(\nu)| + \log|\Gamma_y(\nu)| - \log|\Gamma(\nu)| \Big\}\, d\nu. \\ &= -\int_0^{1/2} \log\left| I - \Gamma_x^{-1}(\nu)\Gamma_{xy}(\nu)\Gamma_y^{-1}(\nu)\Gamma_{yx}(\nu) \right| d\nu.\end{aligned}$$

Here, $\Gamma(\nu)$ is the joint spectral density of x_t and y_t at frequency ν cycles per unit time, and $|.|$ denotes the determinant. To evaluate this for (1) and (2), define $\tilde{x}_t = x_t - Ax_{t-1}$ and $\tilde{y}_t = y_t - Ay_{t-1}$. This ensures that $\tilde{x}_t$ and $\tilde{y}_t$ are stationary, even if x_t and y_t are not. Since information rate is preserved by applying a differencing operator, $\overline{\mathcal{I}}(x_t, y_t) = \overline{\mathcal{I}}(\tilde{x}_t, \tilde{y}_t)$. Now,

$$\begin{aligned}\tilde{x}_t &= w_t \\ \tilde{y}_t &= Hw_t + q_t - Aq_{t-1}\end{aligned}$$

and so $\Gamma_{\tilde{x}}(\nu) = W$, $\Gamma_{\tilde{x}\tilde{y}}(\nu) = WH'$ and $\Gamma_{\tilde{y}}(\nu) = HWH' + (I - Ae^{-2\pi i\nu})Q(I - A'e^{2\pi i\nu})$. This gives

$$\begin{aligned}\overline{\mathcal{I}}(x_t, y_t) &= -\int_0^{1/2} \log\Big| I - H'\big[HWH' + \\ &\qquad (I - Ae^{-2\pi i\nu})Q(I - A'e^{2\pi i\nu})\big]^{-1} HW \Big|\, d\nu,\end{aligned}$$

which simplifies to (3) in our one-dimensional case.

ACKNOWLEDGEMENTS

The authors wish to thank Rio Vetter for providing the Local Field Potential cortical control data.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: Wiley, 1991.

[2] R. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes*. MIT Press, 1997.

[3] D. Taylor, S. Helms Tillery, and A. Schwartz, "Information conveyed through brain-control: Cursor versus robot," *IEEE Trans. Neural Systems and Rehab. Eng.*, vol. 11, pp. 195–199, 2003.

[4] G. Gage, K. Ludwig, K. Otto, E. Ionides, and D. Kipke, "Naive coadaptive cortical control," *J. Neural. Eng. (submitted)*, 2005.

[5] W. Wu, M. J. Black, D. Mumford, Y. Gao, E. Bienenstock, and J. P. Donoghue, "Modeling and decoding motor cortical activity using a switching kalman filter," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 6, pp. 933–942, 2004.

[6] E. Brown, L. Frank, D. Tang, M. Quirk, and M. Wilson, "A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal cells." *J. Neurosci.*, vol. 18, pp. 7411–7425, 1998.

[7] V. Solo, "System identification with analog and counting process observations ii: Mutual information." Submitted to IEEE Conference on Decision and Control, 2005.

[8] D. Taylor, S. Helms Tillery, and A. Schwartz, "Direct cortical control of 3d neuroprosthetic devices," *Science*, vol. 296, no. 5574, pp. 1829–32, 2002.

[9] G. Santhanam, S. Ryu, B. M. Yu, A. Afshar, and K. V. Shenoy, "A high performance neurally-controlled cursor positioning system." IEEE 2nd International Conference on Neural Engineering, 2005.

[10] D. Kazakos and P. Papantoni-Kazakos, "Spectral distance measures between gaussian processes," *IEEE Trans. Aut. Control*, vol. 25, pp. 950–959, 1980.