AN INFORMATION THEORETIC ANALYSIS OF 256-CHANNEL EEG RECORDINGS: MUTUAL INFORMATION AND MEASUREMENT SELECTION PROBLEM

Akira Date

RIKEN Brain Science Institute
Laboratory for Advanced Brain Signal Processing
2-1 Hirosawa, Wako, Saitama 351-0198 JAPAN
date@bsp.brain.riken.go.jp

ABSTRACT

How does one part affect another in the brain? How much information can we extract from the brain data? The multi-channel EEG recording system is now available to study this issue. We recorded 256-channel EEGs while a subject performed a visual discrimination task, and obtained mutual information between a visual stimulus condition and signals from single and multiple EEGs. Here we report preliminary results which show a power of the 256 channel recording system. In particular, we show examples that the best two informative independent measurements are not the two best.

1. INTRODUCTION

Probability, statistics, and information theory are playing an expanding role in interpreting data and in understanding the principles of representation and computation in the nervous system. This paper considers how many bits of information we can extract from EEG data. Pooling of *independent* signals from individual sources can provide more reliable information than that encoded by any single source (see e.g. [1]). In EEG recordings, signals recorded nearby electrodes which modulate in time share a similar wave form, i.e., there exists a strong correlation. Therefore the utility of pooling seems to be greatly reduced. We are interested in how does the benefit of pooling work or not work as the number of observation channels grows in EEG recordings. This is a particular interest in this paper.

2. METHODS

In brief, a subject was trained to report the direction of motion of a random dot display in which all dots moved coherently *Left* or *Right*. This is the simplest version of the task originally designed for monkeys by Newsome and Paré [2]. The subject was required to see a fixation point which appeared at the center of the monitor throughout the experiment which lasted about 5 minutes. Beginning with 4-seconds blank screen, a motion stimulus was presented for 4 seconds, the subject indicated his judgment (Left or Right) by pushing a switch as soon as the motion stimulus was disappeared. Between the motion stimuli, only the fixation point remained on the liquid monitor for 4 seconds. These processes were repeated for 5 minutes, and at the same time we recorded 256 channel EEGs from entire brain with Geodesic Sensor Net (Electrical Geodesics, Inc.) with sampling rate at 500Hz where in analysis each signal from a electrode was normalized as it modulates between 0 and 1. The detail about the sensor, e.g., configuration of the electrodes and the electrode numbers which we use in this paper can be found at `http://www.egi.com` In this paper, we analyzed the first one minute data where the seven (three in Left motion) stimuli were presented to the subject. The task was easy and the subject performed 100% correctly.

Mutual Information: Let $x_i(t), i = 1, \cdots, 256$ denote the signal from ith electrodes at tth sampled time, where $0 \leq x_i(t) \leq 1$ and each x_i was discretized into equally spaced 20 categories or bins ($x_{i,1}, \cdots, x_{i,20}$) to have a empirical distribution $P(X_i)$. Let $y(t)$ denote a stimulus condition which is 1 when the motion stimulus was presented, and 0 when it's not, i.e. $P(Y = 0) = 0.5$ and $P(Y = 1) = 0.5$. That is, we disregard Left and Right condition in this paper. With these $P(X_i)$, we can calculate its entropy

$$H(X_i) = -\sum_{a=1}^{20} P(X_{i,a}) \log P(X_{i,a}),$$

where 20 is the number of bins, and we can also obtain conditional entropy $H(X_i|Y)$, and mutual information

$$I(Y; X_i) = I(X_i; Y) = H(X_i) - H(X_i|Y)$$

by which we know how much information X_i tells about Y. For obtaining $P(X_i)$ and $P(X_i|Y)$, we used only each middle of 2sec of 4 sec data in both stimulus on and off

The work was supported by MEXT Grant-in-Aid 1250240 for Scientific Research on Priority Areas.

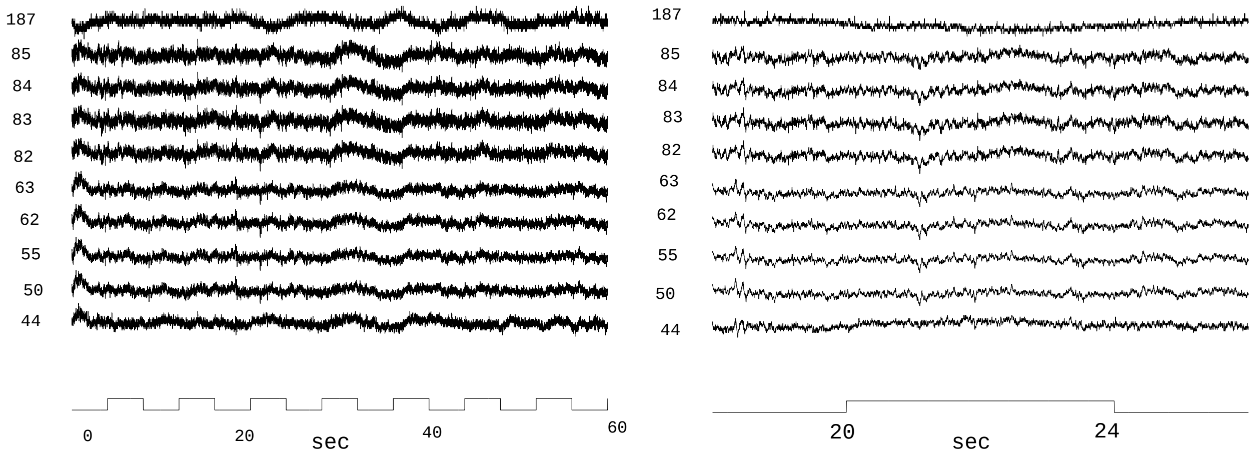

Fig. 1. EEG data obtained with EGI Geodesic Sensor Net 256 channel. Left upper: Observations from 10 different electrodes for the first 1 minute. Left bottom: visual stimulus condition, blank (0), visual stimuli on (1). Right: same to the Left except a time scale (8 sec.) Signals are from electrodes 44 (bottom), 50, 55, 62, 63, 82, 83, 84, 85, and 187 respectively.

periods. In total, 14,000 sample points (data of 14 periods of 2sec) were used to have a empirical distribution, i.e.,

$$P(X_{i,a}) = \frac{\text{number of times we observe} x_i(t) \text{in category } a.}{14000}$$

A conditional distribution was estimated by the same manner based on the samples of 7,000 points. Since Y is binary in this analysis, mutual information $I(Y;X_i)$ is equal or less than 1. Furthermore we can easily consider $I(Y;X_i,X_j)$ in the same manner.

3. RESULTS

3.1. Mutual information on single observations X_i

Figure 1 shows signals from 10 electrodes for the first 1 minute. We obtained distribution $P(X_i)$,$i = 1,\cdots,256$ and also conditional distribution $P(X_i|Y=0)$, and $P(X_i|Y=1)$ to calculate mutual information $I(Y;X_i)$ for all i.

In the upper row of Fig.2, three examples of $P(X_i)$, $P(X_i|Y=0)$, $P(X_i|Y=1)$ are shown where from the left $i = 127, 187$, and 62. Since two conditional distributions of $P(X_{128}|Y=0)$,$P(X_{128}|Y=1)$(upper left in Fig.2) shares similar form, this electrode doesn't tell us much about Y. On the contrary, those of X_{62}(upper right) are clearly divided into two different area. Actually, X_{62} was the second best informative signal. The best informative signal was X_{187} which has distributions shown in upper middle in Fig.2. The shape of two conditional distribution was complemental and multiple peaks.

In the bottom of Fig. 2, we show the histogram of $I(Y;X_i)$, $i = 1,\cdots,256$. The distribution has a peak around 0.01

Table 1.
Best 10 informative channels and some others about stimulus on-and-off.

i	Mutual Information $I(Y;X_i)$
187	0.62509183
62	0.50570727
44	0.46895245
90	0.44303724
55	0.43309851
82	0.43151794
63	0.41826862
50	0.40560261
61	0.39445616
9	0.38991076
101	0.33062429
256	0.29846919
84	0.04314735

which suggests that many of the signals from single electrodes tell us little about stimulus on-off information about Y. Table 1 shows the most 10 informative single channels, and at the bottom right of Fig. 2, $I(Y;X_i)$ is depicted according to the electrode number where the electrode which attached a close number tended to be in nearby location. We can see that electrodes around 50 to 60 has relatively high values of $I(Y;X_i)$ where they may share a similar wave form.

We looked at the locations of electrodes which has high mutual information. We found that channels which satisfies the condition $I(Y;X_i) > 0.3$ concentrated on the left hemi-

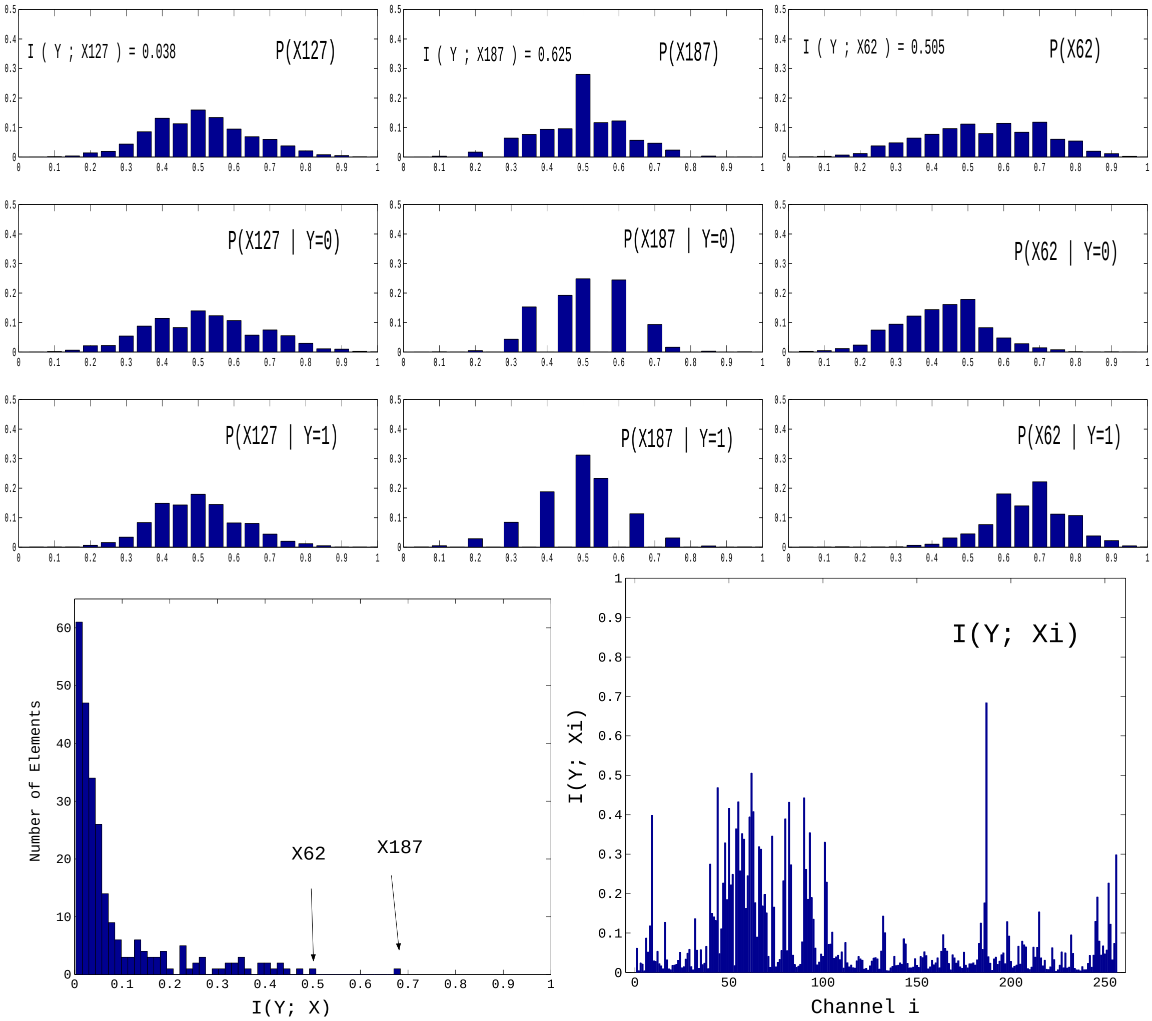

Fig. 2. Distribution of amplitude of EEG signal (upper: Marginal, middle and bottom: conditional), and distribution of mutual information $I(Y;X_i)$ for all i (bottom left: histogram).

sphere from the parietal to the left temporal part without an exception, although occipital area must have been highly activated during the visual stimulus presentation.

3.2. Mutual information on a pair of observation

In the left of Fig. 3, we show the histogram of $I(Y;X_i,X_j)$, $i,j = 1,\cdots,256$, and Table 2 shows the most 10 informative pairs of channels. The best informative pair was channels (82, 84) which gave us 0.930 bits about Y. Although channel 84 itself was less informative, $I(Y;X_{84}) = 0.043$, combined with the channel 82, these pair became most informative. Actually we can see

$$I(Y;X_{82},X_{84}) > I(Y;X_{82}) + I(Y;X_{84})$$

$$0.930 > 0.432 + 0.043.$$

The two best single informative channels were (62,187) which was 6th-best pair (see Table 2). We can represent the following:

$$I(Y;X_{62},X_{187}) < I(Y;X_{62}) + I(Y;X_{187})$$

$$0.831 < 0.625 + 0.506$$

although this expression is not appropriate since $I(Y;X_i,X_j) \leq 1$. In Fig. 1, we depicted the wave forms of these channels.

4. CONCLUSIONS

We looked at mutual information between EEG(s) and a visual on-off condition. We observed that the two individual best channel were not the two best. This kind of observation has been studied theoretically (e.g. see [3]). This

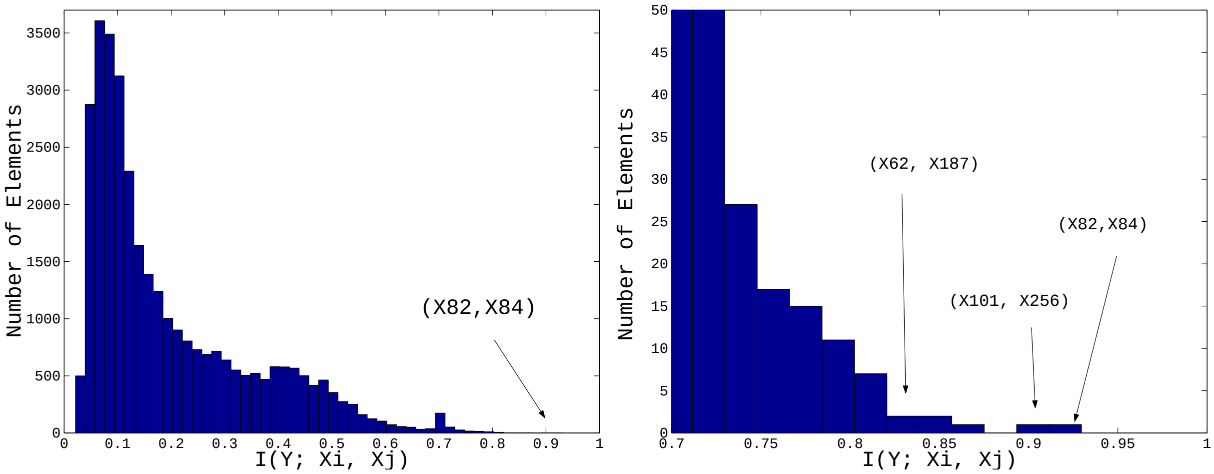

Fig. 3. Distribution of mutual information $I(Y; X_i, X_j)$. Right: The area of high informative components is expanded.

Table 2.
Best 10 informative pairs of channels about stimulus on-and-off.

(i, j)	Mutual Information $I(Y; X_i, X_j)$
(82,84)	0.92963693
(101,256)	0.90659774
(90,256)	0.86410887
(82,85)	0.84229231
(44,187)	0.84024080
(62,187)	0.83097042
(90,187)	0.82867276
(82,187)	0.81985277
(55,187)	0.81475721
(101,187)	0.81373467

analysis became possible by using this kind of 256-channel system. Although we showed mutual information on only single and pairs of EEG signals, we can calculate that of triplets, quadruples of signals.

One thing we will study is to see whether how much information EEGs tell us about directions, Left or Right, even for 4 or 8 directions. And we will conduct an experiment with different subject to know whether it is consistent evidence that informative signals comes from left hemisphere in temporal not occipital region.

In this paper, we defined the problem to choose a single and two best raw signals among 256 channels by maximizing the mutual information Then the pair signals we chose was interpreted as the best pair. To know how this is related to ICA, we will also compare the distributions of mutual information based on the signals before and after the process of ICA.

Acknowledgments

I thank Professor Andrzej Cichocki for giving me an opportunity to do this experiment and the analysis. I also thank colleagues of his laboratory for technical assistance during the course of the experiments reported herein, especially Tomasz M. Rutkowski, Hovagim Bakardjan, and Yo-ichi Miyawaki.

5. REFERENCES

[1] Michael N. Shadlen Ehud Zohary and William T. Newsome, “Correlated neuronal discharge rate and its implications for psychophysical performance,” *Nature*, vol. 370, pp. 140–143, 1994.

[2] William T. Newsome and Edmond B. Paré, “A selective impairment of motion perception following lesions of the middle temporal visual area (MT),” *The Journal of Neuroscience*, vol. 8, pp. 2201–2211, 1988.

[3] Thomas M. Cover and Jan M. Van Campenhout, “On the posible orderings in the measurement selection problem,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-7, pp. 657–661, 1977.