A Probe Into Decoding Brain Activity Using fMRI and MEG

Pulkit Agrawal
Electrical Engineering and Computer Science
University of California, Berkeley
pulkitag@eecs.berkeley.edu

Abstract

The possibility of deciphering thoughts by measuring brain activity using Functional Magnetic Resonance Imaging(fMRI) has recently been demonstrated. Potentially, this can lead the development of machine interfaces controlled by human thought. However, the cost and the size of the fMRI machine makes it an infeasible imaging modality for this purpose. Electroencephalography (EEG) and Near Infra-red Spectroscopy (NIRS) constitute the class of portable machines which can measure neural activity. However, due to poor signal quality from these devices no substantial success has been reported so far. Magnetoencephalography (MEG) provides a measure of same neural signatures as EEG but with superior signal quality. Thus, success in decoding brain activity measured using MEG can guide the same using EEG. In this work we probe the possibility of informing the MEG signal with fMRI in order to enable successful decoding of neural activations. Initially, using ridge regression we show that it is possible to learn fMRI response to natural vision movies. Next, we explore the performance with MEG data using regression, principal component analysis (PCA) and canonical correlation analysis(CCA) as our tools.

1 Introduction

Recent studies indicate that given a stimuli it is possible to predict the human brain activity as measured by Functional Magnetic Resonance Imaging (fMRI). Mitchell et al [1] built models capable of predicting neural activity associated with thinking about concrete nouns. Once trained, the model is also capable of predicting activations associated with previously unseen words. Kay et al [2] build models capable of predicting neural response to natural vision movies. A more interesting piece of work from Nishimito et al [3] was successful in predicting the visual stimuli subjects were experiencing with good accuracies. All this leads us to believe that it should be possible to engineer models capable of deciphering human thoughts just by measuring neural activity. To be practically useful such a machine should be portable and cheap and fMRI is neither. EEG and NIRS offer portable solutions but with their signal quality no substantial results for predicting brain activity have been reported. MEG (Magnetoencephalography) measures neural activity similar to EEG but provides a cleaner signal. Chan et al [4] have very recently attempted to decode word and category specific spatio-temporal representations using EEG and MEG data. Models which can predict the stimuli given the brain activity would henceforth be referred as decoding models and the ones doing vice-versa as encoding models. A detailed study on encoding and decoding fMRI can be found in [2]. Ideally, we would like to build decoding models using EEG/NIRS. As a first step, we wish to attempt such a model using MEG data.

fMRI measures the change in the blood flow related to neural activity in brain. Basically, when neurons become active they consume oxygen, which in turn increases the blood flow in that region to compensate for the loss in oxygen. The consequent change in blood oxygen level constitutes the BOLD(Blood-oxygen-level dependence) signal recorded by the fMRI machine.A BOLD signal is only observed after a delay of few seconds after the onset of stimuli which leads to a poor temporal resolution. However, fMRI provides an excellent spatial resolution of the order of $1mm^2$. Just like the spatial resolution of image is characterized in terms of pixels, fMRI has voxels. A typical scan has 30,000 to 40,000 voxels. MEG, on the other hand measures the magnetic field associated with the flow of current during neural activation. In contrast to MRI, MEG has an excellent of temporal resolution of the order of ms, but poor spatial resolution ($\sim$ 200-300 channels only).

In this body of work we explore to utilizing the work of [3] to build a decoding model using MEG. [3] is able to predict the visual stimuli given the fMRI image. Now a relatively small set of voxels $\sim$ 1000, accounts for most of the prediction accuracy. Thus, it is a worthwhile effort to see if it would be possible to predict the values of these voxels with the MEG data. If successful we will have a decoding model using MEG data. Initially we build encoding models for fMRI data. With a prior on natural movies, it should be possible to invert the encoding model into a decoding model using (see [3]) the Bayes rule. Next, we present various methods we tried in order to predict MRI from the MEG data. Rest of the paper is organized as following: In section 2 we describe the experimental setup, followed by sec 3 which details the mathematical details of the methods used. Section 4 depicts the results followed by the future directions in section 6.

2 Experimental Setup

MEG and fMRI brain scans were collected from 1 subject. The subject was made to watch 70 minutes of video. The video consisted of random clips from natural movies. 10 sample frames from the video are shown in Fig 1. The video was played at 15 frames/sec and the subject was asked to fixate on a dot at the center of the screen for the entire duration. This visual stimuli is same as used by Nishimito et al. for their experiments and more details can be found in [3]
Out of 70 minutes, 40 minutes of data [corresponding to the intervals: 0-10, 20-30, 40-50, 60-70 mins] was used for training. A particular frame appeared only once in the training sequence. 3 sequences, each of 1 minute were randomly repeated 10 times (3x1x10 = 30 mins) each during the remaining time intervals. These constituted the validation sequence.

Figure 1: Sample Frames from Movies shown to the subject

2.1 fMRI Scan

A fMRI image is obtained every 2 seconds. Each such image consists of 30662 voxels. (A voxel in fMRI scan is analogous to pixel in an image.)The technical specifications of the fMRI scanner and an account of pre-processing of this data can be found in [3].

2.2 MEG Scan

The MEG scan consisted of simultaneous recording of neural activity from 271 channels.The sampling rate was 1200 Hz. The MEG data we collected differs in 2 respects from the conventional way of collecting data:

- We have no repeats for the training set.
- We collect MEG or natural movies in contrast to checkerboard like artificial stimuli.

One of the issue with the recordings that sampling rate of the machine fluctuated between 1150-1250 Hz. The data was collected using the machine at UCSF. A detailed account on cleaning and correction of MEG data can be found in [5]

2.3 Input Stimulus

A motion-energy encoding model employing spatio-temporal gabor filters [3] was used to process visual stimuli. The output of the filter is a 2139-D vector at the rate of 0.5 Hz, i.e. the input stimuli is transformed to the same temporal resolution as the fMRI recordings.

3 Techniques Used

3.1 Ridge Regression

Suppose we are presented with a sequence of data points (total of N) (x_n, y_n), where $x_n \in \Re^P$ is the input to the system and $y_n \in \Re^Q$ is the output. We wish to model y_n as a function of x_n. We can formulate a linear mode, where each component of y_n^k (k=1,2,..Q) can be expressed as (i.e we form an independent model for each component of y_n) :

$$y_n^k = \theta_k^T x_n + \epsilon_n^k$$

where $\epsilon_n^k \sim N(0, \sigma_k)$ and θ_k is the parameter vector for the k^{th} component. Let Y^k be a Nx1 vector whose n^{th} value is y_n. Let X be a NxP matrix with n^{th} row as x_n^T.

$$Y^k = X\theta_k + \epsilon^k \tag{1}$$

The solution to this problem can be obtained by solving $\min_\theta \|Y^k - X\theta_k\|_2^2$. However, this problem is not always well posed. The common cases are when the number of observations is less than the number of parameters which need to be estimated or when the solution is not unique. To account for this often a regularization term is added and we seek to solve,

$$\min_\theta(\|Y^k - X\theta_k\|_2^2 + \|\lambda\theta\|_2^2) \tag{2}$$

where λ is the regularization parameter. From a bayesian perspective this characterizes prior information over the distribution of θ. The estimate of $\theta = \hat{\theta}$ is given by,

$$\hat{\theta} \quad = \quad (X^T X + \lambda^2 I_P)^{-1} X^T Y \tag{3}$$

The value of λ is determined by 10 fold cross validation over the training set. This method of regression is known as ridge regression.

3.2 Principal Component Analysis(PCA)

PCA was discussed in class, hence a description of the same is being skipped here. We use PCA for dimensionality reduction as explained later. Intuitively, PCA provides directions capturing successive maximal variances of the given data.

3.3 Canonical Correlation Analysis(CCA)

Suppose there are 2 devices which can measure the activity of one source. In particular let x $\in \Re^M$ be observation by source 1 and y $\in \Re^N$ be the observation by source 2. Also, suppose we have L observations from each source. Now, if our goal is to predict the recording from one source based on the other source then finding directions in which projections of these sources are highly correlated is a useful pursuit. In fact CCA estimates [6] [7], 2 normalized linear filters $w_x \in \Re^N$ and $w_y \in \Re^M$ such that the correlation between $w_x x^T$ and $w_y^T y$ is maximized. Thus the problem can be formulated as:

$$\begin{aligned} arg \max_{w_x, w_y} w_x^T C_{xy} w_y \\ s.t. w_x^T C_{xx} w_x &= 1 \\ w_y^T C_{yy} w_y &= 1 \end{aligned}$$

where, C_{xx}, C_{yy}, C_{xy} are the covariance matrices. The solution to this problem is:
w_x is an eigen vector of $C_{xx}^{-1} C_{xy} C_{yy}^{-1} C_{yx}$
w_y is an eigen vector of $C_{yy}^{-1} C_{yx} C_{xx}^{-1} C_{xy}$

For higher dimensional settings and ill posed problems using Kernels forms a more appropriate method [7]. Let X = $[x_1, x_2, ...x_L]$ and Y =$[y_1, y_2, ...y_L]$, then the kernels and filters are given by,

$$\begin{aligned} K_X &= X^T X \\ K_Y &= Y^T Y \\ w_x &= X\alpha \\ w_y &= Y\beta \end{aligned}$$

α and β are the solutions of the following generalized eigen value problem:

$$\begin{bmatrix} 0 & K_X K_Y \\ K_Y K_X & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \rho \begin{bmatrix} K_X^2 + \kappa_x K_X & 0 \\ 0 & K_Y^2 + \kappa_Y K_Y \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

where κ_x and κ_y are the regularization parameters.
Since the fMRI and MEG are time delayed with respect with each other and we have little idea about the temporal dynamics it would be interesting to see correlation as a function of delays. [7] have proposed a method called temporal canonical correlation analysis for this purpose, which essentially appends one og the modalities either X or Y with delayed copies of itself and doing a regular KCCA from thereof gives the evolution of correlation with delay.

4 Results

4.1 Encoding Model using fMRI Data

A ridge regression was run between the MRI data and the visual stimulus filtered through motion energy filter as described in sec.2. An independent model is built for each voxel. To account for temporal dependence the response of each voxel at time T is modelled as a linear combination of the stimulus at t = T, T-1 and T-2 time instances. The regularization parameter is varied from 0.0312 to 4096 and selected using a 10 fold cross validation on the training set. The optimum value of $\lambda = 256$. The correlation for each voxel over the validation set is plotted in fig2, whereas the distribution of the correlations is shown in fig3.

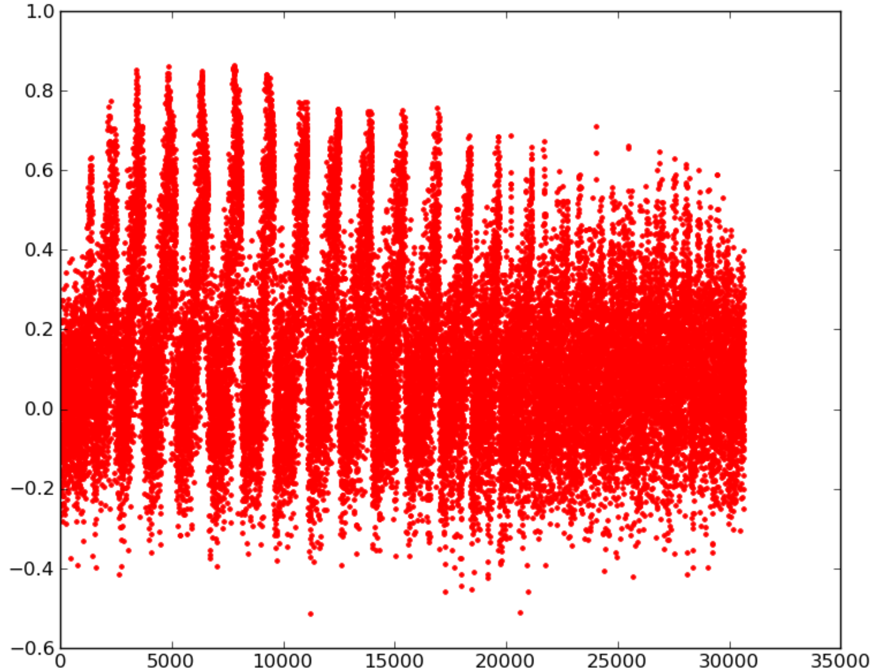

Figure 2: The X axis is the Voxel Number and the Y axis represents the correlation between actual and predicted response for that particular voxel.

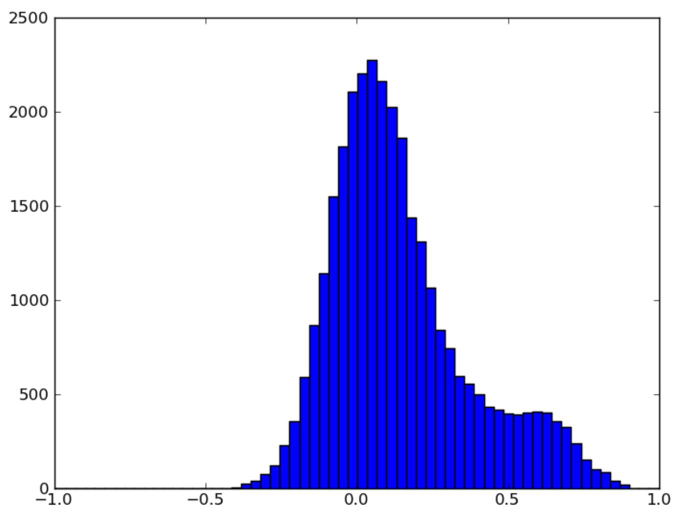

Figure 3: Distribution of Correlation values for all Voxels for ridge parameter $\lambda = 256$.

It should be noted that since the stimulus is visual only, we can only hope to predict the voxel corresponding to the visual regions of the brain. We get a large number of voxels with correlation greater than 0.5 and a fair number over 0.75. They in fact correspond to the visual areas of the brain. The results are very similar as obtained in [3].

4.2 Predicting fMRI using MEG

4.2.1 Regression

MEG data was pre-processed (unless otherwise stated)using median trending, demeaning and subsequently normalizing w.r.t to the standard deviation independently for each channel. It is observed that 99% of the variance in MEG can be captured along the top 10 PCA directions. Consequently the data is reduced 10-D. Also, Due to the difference in the sampling frequencies of MEG and MRI we get 2400 MEG samples in the duration we get 1 MRI reading. Thus we concatenate these 2400 instances to form one big feature vector. Following regressions were tried out:

- The regression between features described above.
- Regression between down sampled versions of MEG at $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{32} .. \frac{1}{128}$ of the sample frequency and the voxel responses.
- PCA channels of MEG and varying number of PCA components in voxel space.

None of the regression models was capable of making any predictions. The typical correlation distribution is shown in fig 4.

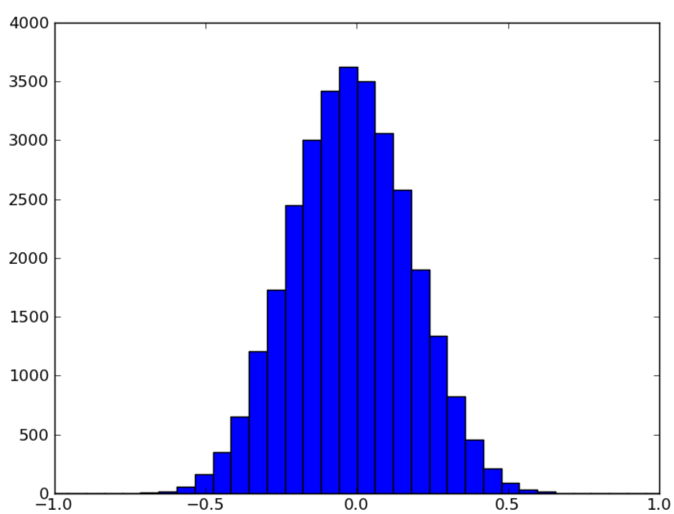

Figure 4: Typical Distribution of Correlation values while regressing with MEG

4.2.2 Frequency Features

Inspired by the presence of information related to various human activities in designated frequency bands of EEG such as Alpha,Theta,Delta [8] etc. and the recent work using EEG and MEG data in [4] it was hypothesized that MEG signal in frequency domain might have predictive information.Thus, a feature described as following was developed:

- Each channel is divided into non-overlapping windows of 6s.
- The raw data in each window is zero meaned and normalized with the standard deviation of the signal value in the particular window.
- A low pass butter-worth filter fig 5 was used (Passband upto 40 hz, stop band attenuation of 30 db beyond 80 Hz) was used to filter each window independently.

- It is found that after this pre-processing the top 10 eigen vectors of 271 channel covariance matrix capture more than 99% of the variance. Thus, the 271 Channels are projected onto these 10 eigen channels. (PCA)

- Each 6s window is further subdivided into 400ms windows overlapping by 200ms each. Thus we have a total of 30 sub-windows.

- Each sub-window is multiplied with a hamming window, zero padded and then the discrete Fourier transform is calculated.

- We retain the DFT coefficients, so that frequencies upto 50 HZ is captured. This gives us 25 coefficients per window. Thus we have 25*30 = 750 features per window per channel. Now we have 10 eigen channels, thus the feature length = 750*10=7500.

- Since, the BOLD response is observable a couple of seconds after the onset of stimuli, the information in the MEG channels in 6s window prioir to the observation of the fMRI scan was hypothesized to be predictive.

Now, a ridge regression was run with λ varying from 0.0312 to 4096 between the obtained feature vector and the observed fMRI response. An independent model is built for each voxel. Ridge regression was found to learn nothing useful. A typical plot showing distribution of correlations is shown in fig 4. Our predictions donot seem to be better than chance predictions.
Now, we only have 1200 training points, but the feature vector length is 7500, thus in-spite of using regularization the regression might have overfit. Thus, regressions were carried out between the feature vector for each channel (750-D) with each voxel, but they failed to provide any promising results.

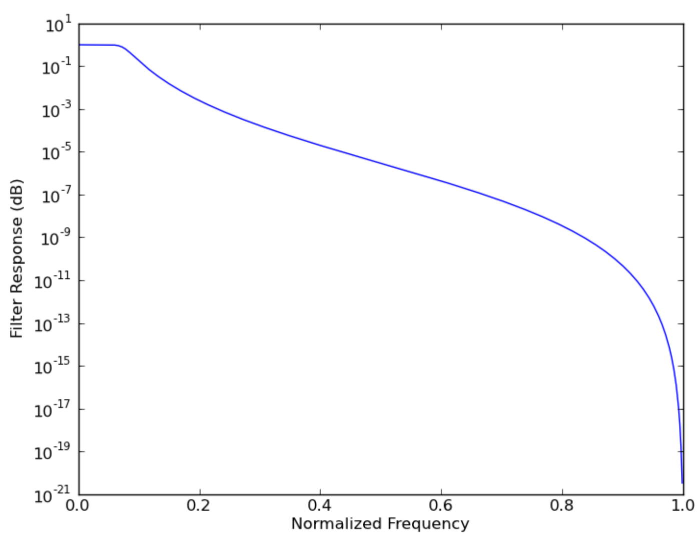

Figure 5: Frequency response of the low pass filter.

4.2.3 Canonical Correlation Analysis

To give an idea of the correlations, the covariance matrix between the eigen MEG and MRI channels is shown in 6.

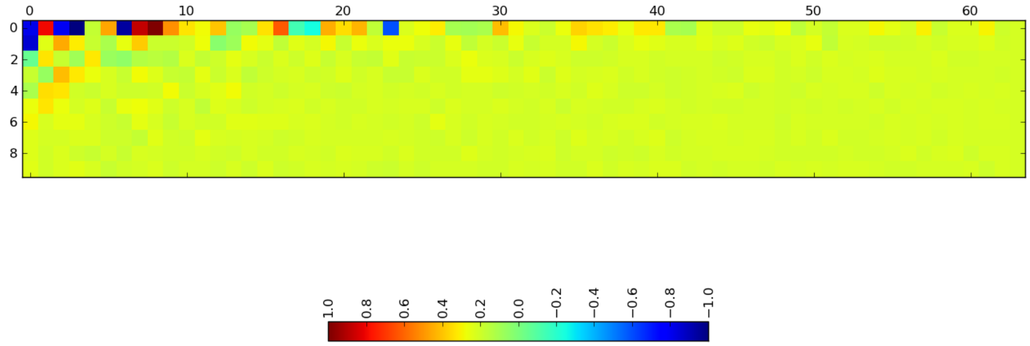

Figure 6: The covariance matrix between top 64 PCA components of MRI and 10 PCA components of MEG

Normal CCA was implemented. The learned directions on the training set gave chance correlations when the data from the validation sets was projected along them. Thus, again we got no success in predicting any voxel from the MEG data.

5 Conclusion

We are able to successfully build an encoding model using the fMRI data. As noted in the introduction it can be easily be extended to build the decoding model. On the other hand we got no success in predicting MRI based on MEG using the techniques described above. Temporal KCCA seems to be an attractive thing to try next.

6 Future Work

Only a very simple version of CCA has been implemented and tested so far. Kernel CCA and Temporal KCCA sound promising. On a different note, one can also try doing source localization for MEG and then learn a model between the sources and voxels corresponding only to visual areas. This would reduce noise in the data, but is severely limited by the accuracy of the source localization models. Moreover, looking at the coherence instead of covariance may help us design a better filter to for extracting predictive MEG frequency components.

Acknowledgments

I thank James Gao for being the subject for both the MRI and MEG data collection. I thank Alex Huth, James Gao and Prof. Jack Gallant for the fruitful discussions and helpful comments. I thank Dr. Srikatan Nagarajan, UCSF for his valuable inputs and for allowing us to collect data from the MEG machine.

References

[1] Tom M Mitchell, Svetlana V Shinkareva, Andrew Carlson, Kai-Min Chang, Vicente L Malave, Robert a Mason, and Marcel Adam Just, "Predicting human brain activity associated with the meanings of nouns.," *Science (New York, N.Y.)*, vol. 320, no. 5880, pp. 1191–5, May 2008.

[2] Thomas Naselaris, Kendrick N Kay, Shinji Nishimoto, and Jack L Gallant, "Encoding and decoding in fMRI.," *NeuroImage*, vol. 56, no. 2, pp. 400–10, May 2011.

[3] Shinji Nishimoto, An T. Vu, Thomas Naselaris, Yuval Benjamini, Bin Yu, and Jack L. Gallant, "Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies," *Current Biology*, pp. 1641–1646, Sept. 2011.

[4] Alexander M Chan, Eric Halgren, Ksenija Marinkovic, and Sydney S Cash, "Decoding word and category-specific spatiotemporal representations from MEG and EEG.," *NeuroImage*, vol. 54, no. 4, pp. 3028–39, Feb. 2011.

[5] J Vrba and S E Robinson, "Signal processing in magnetoencephalography.," *Methods (San Diego, Calif.)*, vol. 25, no. 2, pp. 249–71, Oct. 2001.

[6] Matthew B. Blaschko, Jacquelyn a. Shelton, Andreas Bartels, Christoph H. Lampert, and Arthur Gretton, "Semi-supervised kernel canonical correlation analysis with application to human fMRI," *Pattern Recognition Letters*, vol. 32, no. 11, pp. 1572–1583, Aug. 2011.

[7] Felix Bieß mann, Frank C Meinecke, Arthur Gretton, Alexander Rauch, Gregor Rainer, Nikos K Logothetis, and Klaus-Robert Müller, "Temporal kernel CCA and its application in multimodal neuronal data analysis," *Machine Learning*, vol. 79, no. 1-2, pp. 5–27, 2009.

[8] "Wikipedia article on EEG," .