

Neural Decoding Technique to Reconstruct Stimulus from the Evoked fMRI Voxel Responses

Mupparathy Babu Liby¹, Jini Cheriyan²

¹M.Tech Scholar (Signal Processing), Department of Electronics and Communication, TKM Institute of Technology, Karuvelil Hills, Kollam, Kerala, India

²Assistant Professor, Department of Electronics and Communication, TKM Institute of Technology, Karuvelil Hills, Kollam, Kerala, India

Abstract: *In cognitive neuroscience, neural decoding is a technique to reconstruct the stimulus that evoked the neural activity in the brain. The stimulus includes graphical characters, face images, handwritten characters and natural images. The functional Magnetic Resonance Imaging (fMRI) detects the neural activity caused by the varying Blood-Oxygenated-Level-Dependent (BOLD) signal. The proposed technique aims to develop a novel framework for visual image reconstruction of natural images. The exact reconstruction of natural images is challenging due to its complexity. To achieve this accuracy the voxels describing the structural information and semantic information are considered from the lower and higher brain areas respectively. The voxels having the significant information of pixels are then selected. The reconstruction is achieved using both encoding and decoding models. The encoding model is designed using the image prior (dataset consisting of natural images). The model describes how the stimulus features are encoded in the brain activity. The image prior is clustered using multiclass Support Vector Machine classifier by extracting the Gabor features of the images followed by the k-means clustering. Using the Euclidean distance calculation, the maximum probable category is chosen using which the encoding model is designed for further reconstruction. At the decoding end, a hybrid Bayesian framework is formed by gating the clustered image priors and the selected voxels to reconstruct the stimulus from the evoked fMRI voxel responses. The proposed work focuses on developing an accurate, less complex and automatic software technique for visual image reconstruction of natural images.*

Keywords: Visual Image Reconstruction, fMRI voxel responses (BOLD), Multiclass Support Vector Machine classifier, k-means clustering, Hybrid Bayesian framework.

1. Introduction

In cognitive neuroscience, the developments of non-invasive functional neuroimaging methods like Electroencephalography (EEG), Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI), etc. have opened a pathway to peek into one's brain to retrieve the information on the electromagnetic activity happening in response to the individual's perceptual tasks. Among them, the techniques using the fMRI voxel responses are in its developing stage. The fMRI records highly meaningful neural responses caused by the oxygenated hemoglobin supply to the brain due to the person's behavioral action. The brain area consists of numerous voxels where the brain activity will be mapped. A volumetric pixel is abbreviated as voxel which is a three dimensional version of pixel. Voxels may contain vector values such as density, opacity, color and volumetric flow rate which are termed as voxel BOLD responses. A single voxel consists of numerous neurons. The neural activity caused by the neurons will reflect as the responses in the respective voxels. Hence voxel data gives better clarity on the brain activity than pixels and thus used for medical analysis [8]. The involvement of the machine learning algorithms have made it very comfortable in analyzing the fMRI BOLD responses caused inside the brain in response to various stimuli. The stimuli can be handwritten characters, listened audio, perceived alphabets, natural images, video, etc.

The very promising mind reading, brain mapping, lie-detection tasks have been achieved using fMRI in the neuroscience field. The mind reading tasks have used the machine learning algorithms/methods to associate the

distributed brain responses with the subject's thoughts or perception. The Brain Computer Interface application was developed to learn the specific brain activity patterns and then interface with the mental commands. One of the applications in this interface is to learn the mental thoughts of paralyzed patients. The research are undergoing on fMRI data analysis to achieve the reconstruction of person's dream in real time.

Prediction or assessment of the perceptual sensation is a major challenge in the field of neuroscience engineering. The studies related to the reconstruction methods based on fMRI are now small but in its growing stage. The visual image reconstruction allows translating neural brain activity pattern into the visual image that has caused the corresponding brain activity. The proposed work aims in developing novel software that deals with the reconstruction of the stimulus (natural images) from the fMRI BOLD responses. The natural image prior is classified using Support Vector Machine (SVM) classifier followed by the k-means clustering technique. Using test fMRI voxel response (BOLD) data, the maximum probable category of images and corresponding voxels is selected which is used to design the encoding model that determines the model parameters which identifies the significant voxels. The hybrid Bayesian framework formed using the selected voxels and the image cluster determines voxel to pixel mapping thereby reconstructing the stimulus that evoked the test fMRI voxel response data.

2. Literature Survey

In the initial experiments using fMRI, only the location and size of the voxel responses corresponding to stimulus could

be tracked using conventional retinotopic mapping[10]. Hence the technology was limited to stimulus identification only. The accuracy of stimulus identification was improved when images were represented as Gabor wavelets[7]. Later, several experiments were tried on reconstructing stimuli like image bases with different contrasts[9] [6], graphical characters[6], perceived alphabets[3] [2] [1] and natural images[5] [4]. A better understanding of the neural representation was achieved when multi voxel pattern decoders [6] were introduced which predicted the contrast information automatically from the voxel responses. Even though the linear Gaussian framework[5] and hybrid Bayesian framework[1] could reconstruct the graphical characters and alphabets, the accurate reconstruction of natural images could not be achieved till date due to their complexity. The proposed work aims in reconstructing the accurate stimulus (natural images) from the fMRI BOLD responses.

3. Methods / Approach

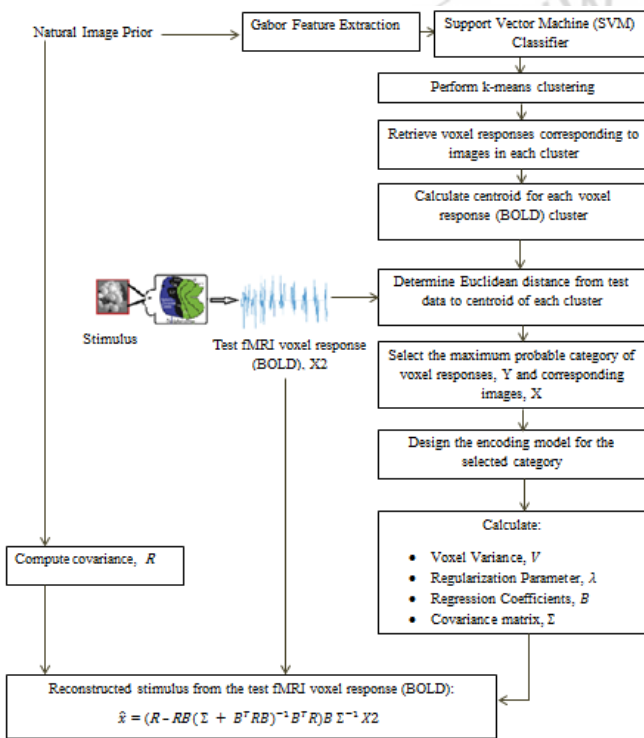


Figure 1: Proposed Block Diagram

3.1 Natural Image Prior Classification

In the proposed method, the dataset of natural image prior consists of various natural images taken from various photographic collections and also the corresponding fMRI recordings. The image prior is assumed to be consisting of images from multiple categories. Hence a fine categorization of the images has to be carried out. The Gabor feature extraction of images gives information on different frequencies and orientation of the same. The Gabor features of images are extracted using the Gabor filter and these features are used in classifying the images using the Support Vector Machine classifier. Since the proposed method deals with natural image consisting of multiple categories, a multiple class SVM framework is used. The SVM classifier

is trained with the training set consisting of Gabor features and its respective labels. Using this training set, the SVM classifier categorizes the images depending on the feature values in each class/labels. The resulting clusters of images are again classified using the k-means clustering technique. The k-means clustering technique groups the images based on their low-level features thereby decreasing the feature description complexity. Hence the natural image priors are divided into cluster, c which consists of images having similar low-level features. The cluster's mean, m_c and covariance, R_c is calculated for each cluster, c .

The voxels corresponding to the images in the clusters are retrieved from the datasets and the centroid of each cluster is calculated. The Euclidean distance from the input test voxel data to each of the cluster centroid is computed. Consider $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$ as two points in the Euclidean n -space, then the Euclidean distance, d from p to q is computed using:

$$d(p, q) = d(q, p) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \quad (1)$$

This is performed to identify the cluster that most resembles the test voxel data. The cluster with minimum distance from test voxel data is taken as the maximum probable category of images. This image category and their associated voxel category are chosen for designing the encoding model.

3.2 Encoding

A forward encoding model is designed with image $x = (x_1, x_2, \dots, x_p) \in R_p$ and the associated measured fMRI brain response $y = (y_1, y_2, \dots, y_q) \in R_q$:

$$y = B'x + \epsilon, \quad \epsilon \in N(0; \Sigma) \quad (2)$$

where ϵ is the zero mean normally distributed noise, B is the Regression coefficients and Σ is the covariance matrix. Let $X = (x^1, x^2, \dots, x^N)' \in R_{xp}^N$ denote the image matrix with x^j denoting the stimulus presented before the subject at the j -th trial. Let $y_i = (y_i^1, y_i^2, \dots, y_i^N)$ denote the associated fMRI voxel BOLD responses in the i -th voxel. For each voxel location (i.e; i -th voxel), the voxel variance, V_i of the responses in the respective voxel location is calculated. The regularization parameter, λ_i is obtained from K-fold cross validation between voxels and pixels using:

$$\hat{\lambda}_i = \operatorname{argmin}_{\lambda \in \Lambda} \operatorname{var}(\hat{\epsilon}_i^K(\lambda)', \dots, \hat{\epsilon}_i^K(\lambda)') \quad (3)$$

where $\hat{\epsilon}_i^K(\lambda)' = y_i^K - X^K \hat{\beta}_i$ are the estimated residuals. The voxels having highest voxel variance compared to the regularization parameter are considered to be carrying most significant information on pixels and hence used for reconstruction. For those voxels, the regression coefficients, $B = (\beta_1, \dots, \beta_N)$ are estimated using the linear regression or classification of multivariate brain analysis:

$$\hat{\beta}_i = (X'X + \lambda_i I_p)^{-1} X'y_i \quad (4)$$

where the amount of regularization is controlled by $\lambda_i \geq 0$. The covariance, $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_1^2)$ is computed using:

$$\widehat{\sigma}_i^2 = \text{var} (\widehat{\epsilon}_i^1(\widehat{\lambda}_i)', \dots, \widehat{\epsilon}_i^K(\widehat{\lambda}_i)') \quad (5)$$

On the other hand, the voxels with variance lesser than the regularization parameter are considered to be less informative and hence discarded from the decoding procedures.

3.3 Decoding

The decoding model is designed such that, given the voxel response for testing; the stimulus (natural image) that has caused the corresponding response is reconstructed from the test voxel responses (BOLD) data, X_2 . The multivariate Gaussian representation of probable image's covariance and mean in canonical form is defined as:

$$\text{Covariance, } Q = (R^{-1} + B \Sigma^{-1} B^T)^{-1} \quad (6)$$

and

$$\text{Mean, } m = QB \Sigma^{-1} X_2 \quad (7)$$

A hybrid Bayesian framework is formed using the selected voxels having most significant information of pixels and selected cluster of natural image priors. The stimulus that has caused the corresponding fMRI voxel responses is reconstructed using the hybrid Bayesian framework which consists of mixture models of natural image priors. The reconstruction of the stimulus from the centralized test fMRI voxel data is obtained using:

$$\widehat{x} = (R - RB (\Sigma + B^T R B)^{-1} B^T R) B \Sigma^{-1} X_2 \quad (8)$$

The reconstructed output will be an average of all the images in the selected maximum probable category. Thus the smoothness of the reconstructed output is proportional to the number of images in the selected category. More the number of images, the more smoothed will be the reconstructed output and vice versa.

4. Results / Discussion

The input to the neural decoding technique is the fMRI voxel responses (BOLD). The voxel responses in the primary visual cortex of the brain are caused by the metabolic activity of the neurons when a person views a visual image. For better clarity of the responses, the voxel responses from the anterior occipital cortex are also taken into account. The dataset consisting of the natural images and their associated fMRI responses taken from the lower and higher brain areas are read into the MATLAB R2015b platform. The dataset was downloaded from <http://crcns.org/data-sets/vc/vim-1>. The stimulus (natural image) that has caused the neural activity in the brain was decoded/ reconstructed from the fMRI voxel responses.

4.1 Natural Image Prior

For an accurate visual image reconstruction, the prior knowledge of the natural image is necessary. The image

prior is the dataset consisting of gray scale natural images taken from various photographs.

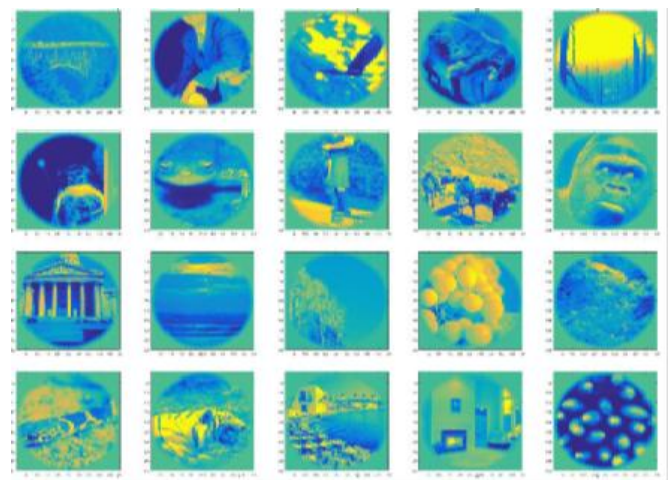


Figure 2: Example of Natural Image Prior

The Figure 2 above shows an example of the dataset consisting of twenty natural images which is used for this work. The image prior is considered to be multimodal which contains natural images of different categories. The image prior categories may be known or unknown. For known image categories, labeling of the natural images using the supervised technique can be performed. When the image prior categories are considered to be unknown, the supervised labeling is not possible. Therefore one of the unsupervised features learning technique, k-means clustering algorithm is suited for clustering. The k-means clustering groups the images into clusters having similar low level features.

In the proposed method, the image prior is divided into group of images. Each group of image is considered as set of images from the same category. Initially k-means clustering was applied on the image prior. Since the k-means clustering technique groups the images based on their low-level features, if the images of animals/humans has mostly scenery as background, they will be grouped with scenery images thereby causing unfine classification results. Out of twenty images, the misclassification happened for 15 images. Thus the classification accuracy could not be achieved here. For better results, the SVM classifier was tried on image pixels directly which resulted in misclassification of 10 images out of 20 images. Accuracy of the classification was very much improved using the Gabor filters on image priors.

The Gabor features of the images are extracted due to their ability of finding high responses in the edges and boundaries of image. The extracted features are given to the SVM classifier. The categorized images are then grouped into clusters within that category using k-means clustering technique which will cluster the images within each category based on their low level feature information. From the final image clusters obtained, the erroneous classification happened only for very few images. The corresponding voxel responses for images in each cluster are retrieved from the database so that the final clusters of voxels are formed.

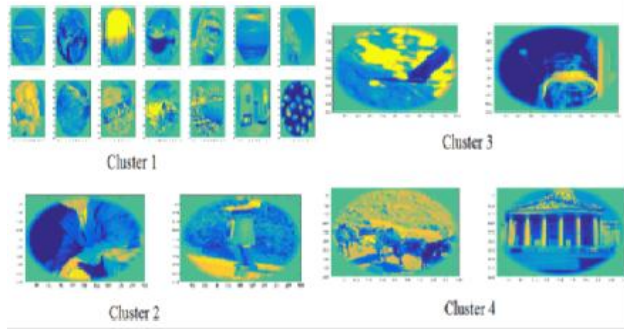


Figure 3: SVM Classifier Output

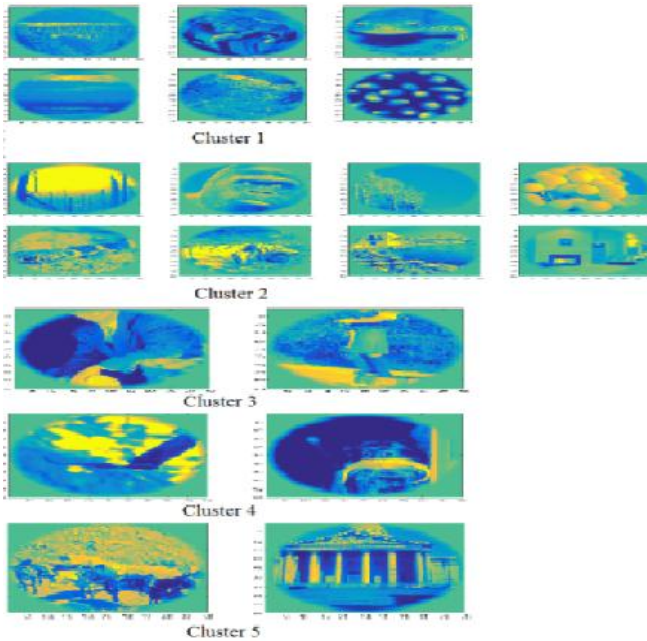


Figure 4: k-means clustering Output

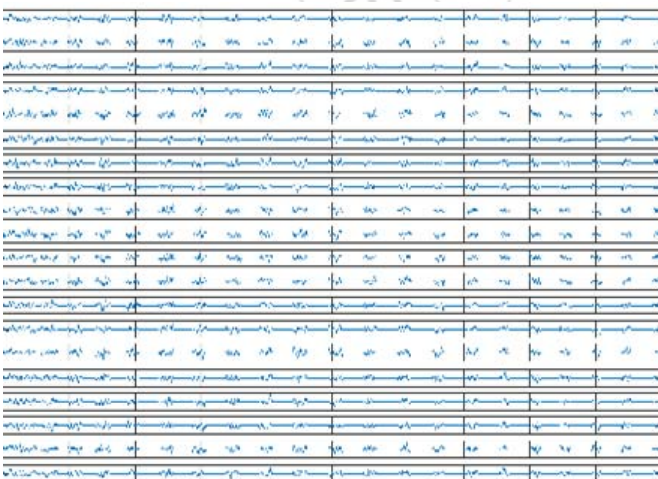


Figure 5: Finally retrieved clusters of voxel responses

The Figure 3 above shows the clustering results using SVM classifier. The Figure 4 above shows the result of k-means clustering applied on the SVM classified output. The Figure 5 above shows the final clusters of voxel responses. These voxel responses corresponding to the images in each cluster are retrieved from the database.

4.2 Stimulus and the test fMRI

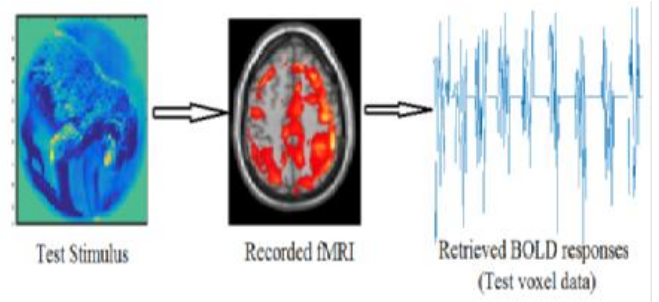


Figure 6: An example of stimulus and test fMRI voxel data

The stimulus is chosen from the given natural image prior set. The corresponding fMRI BOLD responses are pre-processed and also provided in the database. An example of both is shown in the above Figure 6. For the reconstruction of the stimulus, the evoked fMRI BOLD response corresponding to that stimulus is provided to the software as test voxel data.

4.3 Selection of the maximum probable cluster of images and voxel responses

The centroid (mean) of each cluster of voxels is determined. The Euclidean distance from the test voxel data to centroid of each cluster is calculated. The voxel response category that has the minimum Euclidean distance from the test data is chosen to be the maximum probable category which is having high resemblance to the test voxel data. This category of voxel responses and its corresponding images are determined to design the encoding model.

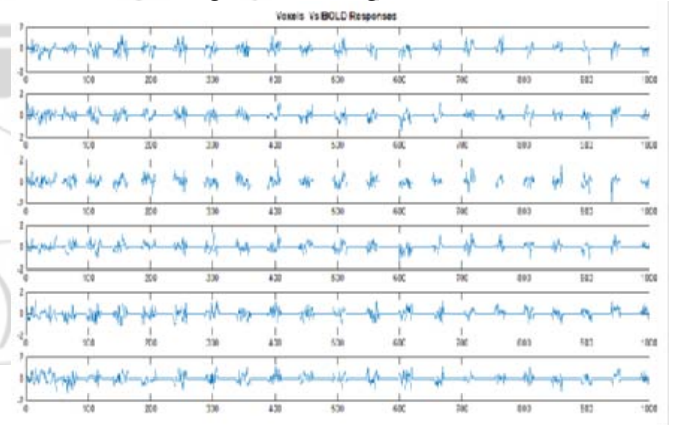


Figure 7: Selected probable cluster of voxel responses

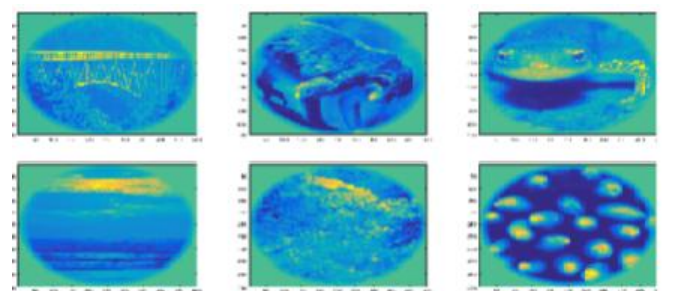


Figure 8: Selected probable cluster of images

The Figure 7 and Figure 8 above shows an example of the

selected category of fMRI BOLD responses and its corresponding natural images. The encoding model is designed for model parameter estimation that contributes to the reconstruction.

4.4 Model Parameter Estimation

The encoding model is designed for the selected category of natural images and the associated fMRI voxel responses of the brain. The model describes how the stimulus features are encoded in the brain responses. The model parameters are computed only for the voxels having significant pixel information and the remaining irrelevant voxels are discarded from the reconstruction procedures. Selection of relevant voxels are done by comparing the voxel variance (V_i) and regularization parameter (λ_i) at each voxel locations. The voxels having higher variance compared to the regularization parameter are considered to be having the significant pixel information. The model parameters computed for selected voxels include Regression coefficients (β_i) and the Covariance matrix (Σ). This automated method of calculating the model parameters in a way reduces the amount of irrelevant voxels thereby decreasing the complexity of the algorithm.

4.5 Decoding

A Hybrid Bayesian framework is formed for reconstructing the natural image from the test fMRI BOLD response. The framework is formed with selected voxels having highest variance and unimodal image cluster. Finally the reconstruction of the stimulus that has caused the test fMRI response is achieved.

The actual dataset of natural image prior is having resolution of 128 x 128 pixels. Also, the dataset of fMRI voxels consists of 25000 voxels. But, due to the limitations in the memory of MATLAB platform used, the resolution size of the natural images had to be reduced to 50 X 50 pixels. Similarly the size of the voxels considered for the simulation was reduced to 1000 voxels. Hence the clarity in the reconstruction result is very poor.

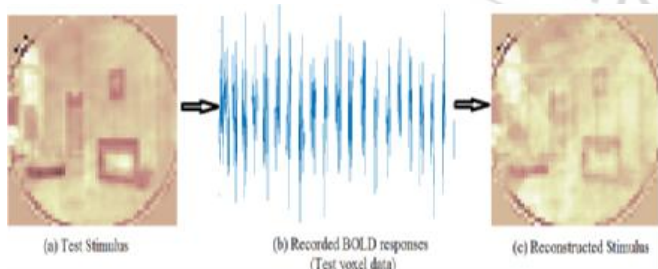


Figure 9: Reconstruction of Stimulus (Trial 1)

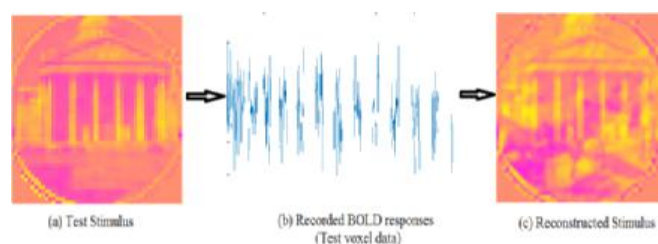


Figure 10: Reconstruction of Stimulus (Trial 2)

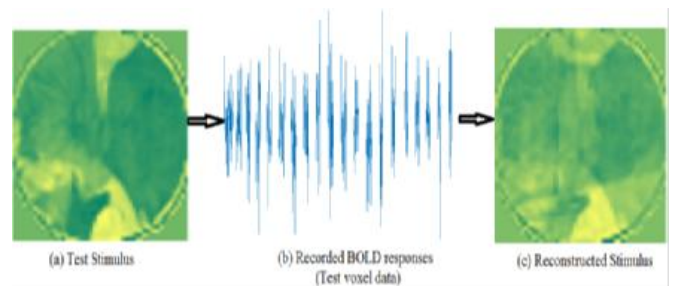


Figure 11: Reconstruction of Stimulus (Trial 3)

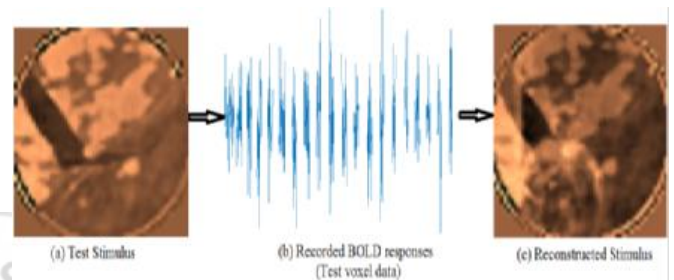


Figure 12: Reconstruction of Stimulus (Trial 4)

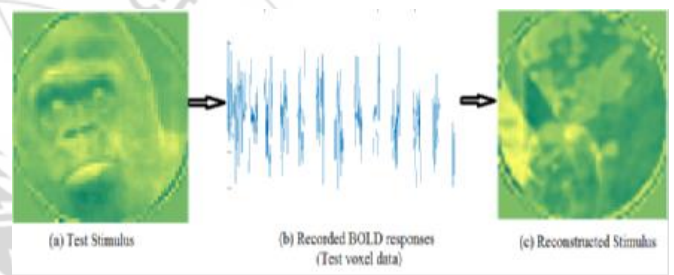


Figure 13: Reconstruction of Stimulus (Trial 5)

The above figures show the results obtained in reconstructing the stimulus image from the test voxel data (BOLD response). In this work, 20 natural images were considered as the image prior. These images were categorized into 5 clusters. Once the probable cluster of voxels and corresponding images are determined, the reconstructed output will be an average of the images in the selected cluster. Since the selected cluster using the test voxel data in Figure 9 contains more images, hence the output of reconstruction seems to be more smoothed. On the other hand, the selected cluster using the test voxel data in Figure 10, Figure 11 and Figure 12 consists of only two images, the reconstructed output seems to be less smoothed. Out of the 20 natural image prior used in this work, four images were wrongly reconstructed. An example of such wrong reconstruction is shown in the below Figure 13.

5. Conclusion

The functional Magnetic Resonance Imaging (fMRI) data can be decoded to retrieve a person's thought or behavior. The applications using fMRI data include mind reading, lie-detection tasks, etc. In earlier works, only the accurate reconstruction could be achieved on perceived alphabets and graphical characters. In the proposed neural decoding method, the fMRI data of a person has been fed as input for decoding. The natural image viewed by the person is successfully reconstructed from the fMRI BOLD responses generated due to the neural activity in the primary visual

cortex of the brain. The image prior classification was tried using k-means clustering which led to high misclassification of images. This was solved to a major extent when the images were classified using SVM classifier by extracting the Gabor features of images followed by the k-means clustering of the output categories obtained from SVM classification. An 80% accuracy in the reconstruction of stimulus from the evoked fMRI voxel responses (BOLD) could be achieved in this work. The complexity could be reduced by classifying the image prior consisting of multiple categories and selecting the most probable category for further reconstruction. An automatic and novel reconstruction framework could be developed for accurate reconstruction of the stimulus from the fMRI responses. Only the clarity of the reconstruction was poorly achieved due to the limitations in the memory of MATLAB version used.

6. Future Scope

The reconstruction of the stimulus has to be achieved for the fMRI voxel responses (BOLD) taken from real time.

References

- [1] Sanne Schoenmakers, UmutGüçlü, Marcelvan Gerven and Tom Heskes. "Gaussian mixture models and semantic gating improve reconstructions from human brain activity". *Frontiers in Computational Neuroscience*, January 2015, Volume8, Article173.
- [2] UmutGüçlü, Marcel A. J. van Gerven. "Unsupervised Feature Learning Improves Prediction of Human Brain Activity in Response to Natural Images". *PLoSComputBiol* 10(8): e1003724. doi:10.1371/journal.pcbi.1003724, August 2014, Volume 10, Issue 8, e1003724.
- [3] Sanne Schoenmakers, Tom Heskes, Markus Barth, Marcel van Gerven. "Linear Reconstruction of Perceived Images from Human Brain Activity". Elsevier Inc., *NeuroImage*xxx(2013) xxx - xxx Impact Factor: 6.36 DOI: 10.1016/j.neuroimage.2013.07.043, July 2013.
- [4] Vincent Q. Vu, Ravikumar, P., Naselaris, T., Kay, K.N. "Encoding and decoding V1 fMRI responses to natural images with sparse nonparametric models". *Ann Appl Stat.* 2011 June; 5(2B): 1159–1182. doi:10.1214/11-AOAS476.
- [5] Thomas Naselaris, Ryan J. Prenger, Kendrick N. Kay, Michael Oliver, and Jack L. Gallant. "Bayesian Reconstruction of Natural Images from Human Brain Activity". Elsevier Inc, *Neuron* 63, 902–915, September 24, 2009.
- [6] Yoichi Miyawaki, Hajime Uchida, Okito Yamashita, Masa-aki Sato, Yusuke Morito, Hiroki C. Tanabe, Norihiro Sadato, and Yukiyasu Kamitani. "Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders". Elsevier Inc, *Neuron* 60, 915–929, December 11, 2008.
- [7] Kendrick N. Kay, Thomas Naselaris, Ryan J. Prenger, and Jack L. Gallant. "Identifying natural images from human brain activity". *Nature* 2008 March 20; 452(7185): 352–355. doi:10.1038/nature06713.
- [8] Martin A. Lindquist. "The Statistical Analysis of fMRI Data". *Statistical Science* 2008, Vol.23, No. 4, 439 - 464 DOI: 10.1214/09 - STS282.
- [9] Yusuke Fujiwara, Yoichi Miyawaki, Yukiyasu Kamitani, Dubois, Jean-Baptiste Poline, Denis LeBihan, and Stanislas Dehaene. "Estimating image bases for visual image reconstruction from human brain activity". *NeuroImage* (2006).
- [10] Bertrand Thirion, Edouard Duchesnay, Edward Hubbard, Jessica, Thirion et al. "Inverse retinotopy: Inferring the visual content of images from brain activation patterns". *NeuroImage* 33 (2006) 1104–1116.

Author Profile



M. B. Liby received her B.E degree in Electronics and Communication Engineering from Visvesvaraya Technological University in 2008. She is currently pursuing second year M.Tech in Signal Processing at TKM Institute of Technology.



Jini Cheriyan received her B.Tech degree in Electronics and Communication Engineering from University of Kerala in 2009 followed by M.Tech in Computer Vision and Image Processing from Amrita University in 2011. She is currently working as Assistant Professor in Electronics and Communication Engineering Department at TKM Institute of Technology.