SIGNAL DETECTION THEORY ANALYSIS OF CATEGORY-BASED VISUAL SEARCH IN NATURAL MOVIES

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

ELECTRICAL - ELECTRONICAL ENGINEERING

By

Osman Tutaysalgır

August 2016

Signal Detection Theory Analysis of Category-Based Visual Search in Natural Movies
By Osman Tutaysalgır
August 2016

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Tolga Çukur(Advisor)

Emine Ülkü Sarıtaş

Didem Gökçay

Approved for the Graduate School of Engineering and Science:

Levent Onural
Director of the Graduate School

ABSTRACT

SIGNAL DETECTION THEORY ANALYSIS OF CATEGORY-BASED VISUAL SEARCH IN NATURAL MOVIES

Osman Tutaysalgır
M.S. in Electrical - Electronical Engineering
Advisor: Tolga Çukur
August 2016

The human brain changes its inner hierarchy and connection strength between the neurons in order to apprehend the real world. In visual search, It is thought that the human brain changes the sensitivity of neurons in the favor of the attended object. Here, we investigate these tuning shifts of the voxels in signal detection theory perspective. Brain activities of human subjects were recorded while they were watching a natural movie. To assess the attentional effect on the human brain, the decoding procedure was employed on the BOLD responses and the natural movie stimuli. Decoding procedure tries to predict the stimuli that form the BOLD responses. In order to bridge the gap between the stimuli and BOLD responses, logistic regression which is a classification algorithm is applied to form models of the subjects' brain. The model performances were assessed with d-prime, ROC and AUC parameters.

Our results suggest that category-selective regions in the human brain boost their detection performances further for the objects that they are not inherently selective.

Keywords: Visual Search, Logistic Regression, D-prime, Receiver Operating characteristic (ROC), Area Under Curve (AUC) .

ÖZET

KATEGORİSEL GÖRSEL DİKKATİN SİNYAL ALGILAMA TEORİSİ ILE İNCELENMESI

Osman Tutaysalgır
Elektrik - Elektronik Mühendisliği, Yüksek Lisans
Tez Danışmanı: Tolga Çukur
Ağustos 2016

İnsan beyni, gerçek dünyayı anlamlandırabilme için kendi iç hiyerarşisini ve nöronlar arasındaki bağlantı gücünü değiştiren bir yapıya sahiptir. Bu doğrultuda, insan beyninin dikkat edilen nesneye göre nöronların hassasiyetini değiştirdiği düşünülmektedir. Bu çalışmada voxellerde meydana gelen bu duyarlılık değişimleri sinyal algılama teorisi çerçevesinde ele alınacaktır. Deneklerin beyin aktiviteleri, onlara bir video izletilirken kayıt altına alınmıştır. Dikkatin beyin üzerindeki etkisini anlayabilmek için hem beyin hücrelerinin aktivite sinyalleri (BOLD) hem de videodaki objelere çözümleme (decoding) prosedürü uygulanmıştır. Bahsedilen çözümleme (decoding) prosedürü beyin aktivitelerine neden olan uyarıcıyı (stimuli) tahmin etmeye çalışmaktadır. Bu bağlamda, etki ve tepkiler arasındaki bağıntının kurulması amacıyla bir sınıflandırma algoritması olan "logistic regression" ile deneklerin beyin modelleri oluşturulmaya çalışılmıştır. Bu modellerin performansları d-prime, ROC ve AUC parametreleriyle hesaplanmıştır ve değerlendirilmiştir.

Bu araştırma sonucu ortaya çıkan sonuçlara göre, insan beyninin kategori-seçici bölgeleri görsel dikkatle birlikte seçici olmadıkları nesneleri algılayabilmek için o nesnelerin seçiciliğini seçici oldukları nesnelere göre daha fazla arttırmaktadır.

Anahtar sözcükler: Görsel Dikkat, Logistic Regression, D-prime, ROC, AUC.

Acknowledgement

Firstly, I would like to express my kindest gratitude to my advisor Dr. Tolga Cukur for his immense support, patience, motivation and sharing his knowledge during my MS study. His guidance helped me in all the time of research and writing of this thesis. Besides my advisor, I would like to thank Dr. Emine Ulku Saritas and Dr. Didem Gokcay for their insightful comments and encouragement to widen my research from various perspectives. I also want to mention about my friends: Salman, Yonus, Gokhan, Mahmut and Mehmet. I would like to thank them one by one for their encouragement and support during the years we are together. I also would like to thank TUBITAK for TUBITAK BIDEB and TUBITAK 3501 114E546 supports throughout my MS study.

Last but not the least, I would like to thank my family especially my granddad, Osman, for supporting me spiritually throughout writing this thesis and my my life in general.

Contents

1 Introduction **1**

2 Functional Magnetic Resonance Imaging **4**

2.1 Statistical Analysis of FMRI Data 6

2.2 Data Acquisition 8

2.3 FMRI Noise Sources and Characteristics 9

2.3.1 Thermal and System Noise 9

2.3.2 Physiological Noises 10

2.4 Components of FMRI Preprocessing 10

2.4.1 Quality Assurance 10

2.4.2 Slice Time Correction 11

2.4.3 Head Motion Correction 11

2.4.4 Spatial Normalization 12

2.4.5 Temporal and spatial smoothing 12

3 Human Brain and Vision Related Parts 14

3.1 Brain and Its Structure 14

3.2 Cerebrum and Its Subparts 15

3.2.1 Vision Related Regions 17

4 Methods 20

4.1 Subject 20

4.2 Stimuli 20

4.3 Experiment 21

4.4 MRI Data Acquisition 22

4.5 Data Pre-Processing 22

4.6 Category Model 23

4.7 Model Fitting 23

4.7.1 fMRI Models 24

4.7.2 Signal Detection Theory and D-prime 27

4.7.3 Logistic Regression 32

4.7.4 Searchlight Analysis 34

4.7.5 Procedures 34

5 Results 37

5.1 Conclusion 57

A Appendix **63**

A.1 MRI: Magnetic Resonance Imaging 63

A.2 MRI Physics 64

A.3 Image Formation 68

A.3.1 Bloch Equations 71

List of Figures

2.1 A typical hemodynamic response. Measured BOLD activity starts to increase two seconds after from beginning of the neural activity. It reaches its maximum about 5s after the onset. After the neural stimulus ends, hemodynamic response amplitude falls below a baseline and recovers its initial state. The figure is reinterpreted from [1]. 5

4.1 Signal and noise trial. The threshold value is selected to minimize false alarm rate and maximize true positive rate for a given experiment. D-prime value represents how far means of these two curves are located between each other. This figure is adapted from [2]. . 30

4.2 ROC curve with different d-prime values. Discrimination between the trials increases with the d-prime value. Different d-prime values yield various ROC curves. This figure is reinterpreted from [3]. . . 31

4.3 Sigmoid function. Our modeling function is changed to estimate the binary classification [4]. 33

5.1 Decoding performance difference of the 'Human' object between attended and unattended stimuli. This illustration was obtained by subtracting the d-prime values of the 'Human' when subjects were attending and not attending to the object of interest. All of the regions of the brain, whether they are inherently selective to the 'Human' object or not, enhance their detection performances with visual search. Error bars show standard deviation that was calculated by bootstrapping d-prime values across subjects. 38

5.2 Decoding performance difference of 'Vehicle' object between attended and unattended stimuli. This graph is calculated by subtracting obtained d-prime values from the subjects' two conditions (attending or not attending the category of interest). All of the regions benefit from the visual search and increases their detection performances. 39

5.3 Decoding differences between two object categories. This graph illustrates the decoding sensitivity changes over attention in different ROIs. This plot is acquired by subtracting the Fig:5.1 from Fig:5.2. 'Human' and 'Scene' selective areas shows different characteristics. Human selective areas take advantage of the visual search of 'Vehicle' more than the visual search of 'Human'. Scene selective areas tend to show the opposite relation. 41

5.4 Area under the curve graph in group analysis, horizontal axis represents the decoding improvement in 'Human' object with attention, vertical axis represents the increase in 'Vehicle' object decoding with attention. Scene selective areas enhance their areas under the ROC curves to a greater extent on visual search of 'Human' objects. Human selective regions tends to show the opposite relation. 44

5.5 ROC curve for subject 'TC' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate. 45

5.6 ROC curve for subject 'SN' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate. 46

5.7 ROC curve for subject 'JG' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate. 47

5.8 ROC curve for subject 'AH' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate. 48

5.9 ROC curve for subject 'AV' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate. 49

5.10 Flatmap representation of subject 'TC', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention. 52

5.11 Flatmap representation of subject 'SN', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention. 53

5.12 Flatmap representation of subject 'JG', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention. 54

5.13 Flatmap representation of subject 'AH', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention. 55

5.14 Flatmap representation of subject 'AV', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention. 56

A.1 Slice selection: Hydrogen atoms are tipped with slice selection gradient. Hydrogen atoms within the selected slice are tipped right after the RF pulse. In the left pane, the magnetization vector points toward a discrete set of angles. In the real world, however, particles are exposed to a slightly different magnetic field and therefore form continuous bands of tipped angles. This figure is adapted from [1]. 69

A.2 A sample MR sequence. The RF and the slice selection gradient excite a 2-D plane in a 3-D volume. After the excitation phase, phase and frequency encoding gradients are applied. Frequency and phase encoding gradients constitute a sufficient k-space coverage. Sampling occurs at the same time as the frequency encoding gradient. This illustration is adapted from [5]. 71

A.3 An MR sequence and its k-space coverage. If we use number 1 of the phase encoding gradient, we will cover the line 1 on the k-space. By changing the G_x gradient, we shift from the middle to the left and from the left to the right of the k-space and sample as we use the G_x gradient. The figure is adapted from [6]. 74

Chapter 1

Introduction

The brain is a highly flexible organ that evaluates, transforms and modifies the data from the sensors of the human body. It also governs many different tasks related to cognition and perception. It is a network interconnection system that consists of over 100 billion nerve cells. Although an individual nerve cell has limited capacity, the interconnection of this system controls not only motor muscle movements such as walking but also fulfills feature specific, computationally expensive tasks such as categorization of seen objects. The human brain's process of understanding starts with the signaling pathways of several sensory organs, the communication between the neural cells and their formation of the interconnection system.

Vision is one of the most advanced features of a living organism. Most of our perception of the real world is obtained from our eyes. Vision starts in the retina. The retina converts analog real world signals into electrical ones [7]. Signals from the retina are transmitted to the cerebral cortex (that is the outermost layer of the brain) via neurons. This data first arrives at the primary visual cortex where the basic shape and orientation of the object are identified. In the higher levels of the processing, the brain combines information from several subregions to build the perception. It also selects information about specific attributes to construct short and long term memory [7].

The human visual system carries out various tasks in order to apprehend the surrounding world. Recognition is one of the most computationally expensive and challenging task. To recognize an object, the brain needs to compare the perceived object with countless possibilities and find the right match. Palmer et al. [8] described the perception process such that the brain converts an object into a reference plane. Size and orientation of the object seen in the reference plane are compared with the objects stored in the memory. This model creates an advantage on computational intensity in that only a few models need to be stored in the memory and processed in order to recognize a particular object.

Olshausen et al [9] presented a mechanism to explain how objects are represented in visual areas. They tried to explain how attention and pattern recognition occur in the brain. According to their model, input and output relationship between different brain regions are dynamical, changed with neural connections. Also, the strength of these neural connections are modified without loss of spatial and temporal resolution.

In earlier studies, researchers tried to evaluate the relationship between isolated individual voxel responses and the the stimuli represented. These analyses are not sensitive enough to decode information in certain aspects. These classical methods select the voxels that show significant response to the experiment stimuli and take the average of their responses. Despite spatial averaging smoothing the data, it blurs the dataset and might eliminate unique information. Furthermore non-significant voxels might convey stimuli-specific information, and this information can't be taken into account for processing [35, 36].

In another recent study [10], researchers investigated the attentional effect on extracting categorical information. They proposed that visual search of a particular object biases the processing in the favor of that object so that only the attended object is represented in high-level cortical areas. In another study [11], researchers tested whether task-relevant features of the attended category are selected. According to their results, when subjects paid attention motion, they found out that the acquired signal is bigger as subjects see moving faces instead of moving houses in face-selective areas. Reddy et al [12] also showed that attention

removes the clutter effects and increases visual search performances under natural viewing conditions.

In this thesis, we tested how attention changes tuning shifts of the category-selective regions of the human brain. We used a signal detection theory perspective and modeled different areas of the brain as detectors of the target category. Our results suggest that the detection performances of the category-selective regions are the highest for objects for which they are not inherently selective.

Outline of the thesis is as follows: In Chapter 2, building blocks of the MRI (magnetic resonance imaging), fMRI (functional magnetic resonance imaging) and fMRI signal processing are briefly explained. In Chapter 3, we talk about the human brain and its category-selective regions. In Chapter 4, methods used in this study are elaborated. In Chapter 5, obtained results are presented and explained.

Chapter 2

Functional Magnetic Resonance Imaging

As the information processing takes places in the brain, the energy required is provided by the vascular system in the form of glucose and oxygen. Oxygen is transmitted to the brain via blood. It appends to the hemoglobin molecules in the blood. It changes the magnetic property of the hemoglobin. Functional magnetic resonance imaging (fMRI) uses these alterations in the magnetic field to form images of neural activity correlated with physiological changes. Contrast based on neural activities is also known as the blood-oxygenation-level-dependent (BOLD) contrast.

When oxygen molecules attach themselves to the hemoglobin, hemoglobin molecules show the diamagnetic property. Diamagnetic materials have no magnetic moments when they are placed inside a magnetic field, thus they do not effect the MR signals. Deoxygenated hemoglobin, on the other hand, shows paramagnetic property, and it can be used as a contrast material for MR. This phenomenon was first discovered by Seiji Ogawa, who was a research scientist at Bell Laboratories [1].

Deoxygenated hemoglobin has a small T_2 value that leads MR components

on the transverse plane to decay faster and reduce the MR signal (underlying physical phenomena are described in the appendix part). Although oxygenated hemoglobin has no effects on the MR signal, neural activity related MR images are collected with the help of the oxygen molecules. Whenever a neural activation occurs in a brain region, the concentration of oxygenated hemoglobin molecules increase more than that of the deoxygenated hemoglobin molecules. Therefore, signal decrease due to the T_2 effect is relatively reduced.

The MR signal triggered by neural activity is known as the hemodynamic response. Neurons usually fire within milliseconds after the stimulus. But hemodynamic responses start 1-2 seconds following the neural event and usually reach their maximum within the following 5 seconds. After a few seconds, hemodynamic responses fall below their baseline and refine to their initial state within 10 seconds.

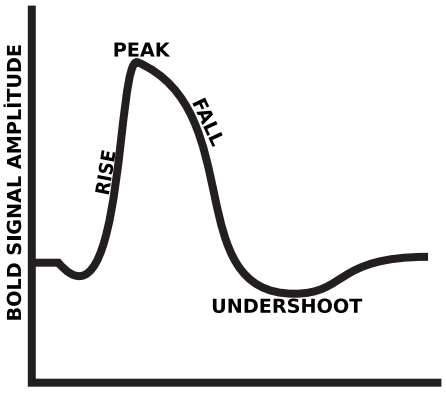

Figure 2.1: A typical hemodynamic response. Measured BOLD activity starts to increase two seconds after from beginning of the neural activity. It reaches its maximum about 5s after the onset. After the neural stimulus ends, hemodynamic response amplitude falls below a baseline and recovers its initial state. The figure is reinterpreted from [1].

The spatial resolution of an MR image is defined as the separability of a voxel from the nearby spatial location. It is dependent on several different factors relating to the MR sequence and the machine. On the MR device, the field of view parameter describes the total imaged size in 2 dimensions, and it is usually expressed in centimeter. On the MR sequence, the total number of imaged voxels is defined as sampling period on each frequency and phase encoding gradient. So,

for example, a field of view of 25 by 25 cm and 125 samples on each encoding gradient results in a 2mm x 2mm in 2-D. The third dimension that is needed to construct the voxel is provided by the thickness of the slice that is formed with the slice selection gradient [13].

Increasing the voxel size is beneficial for fMRI studies. SNR of fMRI signals are dependent on the BOLD and if we decrease the spatial resolution by a factor of 2, BOLD signals reduce by half so does the SNR value. The second effect of increasing the voxel size is that time of acquisition decreases as we increase the voxel size. This increase in the acquisition time causes images to suffer from T_2^* effects and reduces the BOLD signal. On the contrary, increasing the voxel size too much causes a decrease in discrimination performances [13].

Temporal resolution can be described as the ability to discriminate different consecutive stimuli from the BOLD signals. fMRI can ensure a temporal resolution that is on the order of a few seconds. Like spatial resolution, temporal resolution is dependent on both the MR sequence and BOLD signal characteristic. Repetition time of the MR sequence is the first parameter in constructing the temporal resolution. Repetition time might change from 500ms to 3000 ms in a typical MR experiment [1]. The second factor that effects the temporal resolution is the hemodynamic response characteristic. As mentioned in previous sections, BOLD responses last around 10 seconds to recover their original state. It is better to sample the BOLD signals in a small interval because the smaller repetition intervals help us better identify the BOLD response of a neural activity.

2.1 Statistical Analysis of FMRI Data

In a typical fMRI study, researchers need to find out how to deal with the inherently noisy data coming from a complex spatiotemporal structure. To elucidate the underlying reasons behind the fMRI data, scientists utilize statistics as a valuable tool.

fMRI is a noninvasive imaging modality that is used to unveil brain functions. In a typical fMRI experiment, the researchers acquire a set of images (directly measuring the blood oxygenation level in different parts of the brain) that are related to the current task of the brain. This novel technique uncloaks the brain actions and inner hierarchies which previously were very difficult to gather information about. Stages of the systematics of neural connections and task-related responses of the brain can now be understood with a simple noninvasive brain scan.

In an fMRI study, researchers try to achieve information about several features and actions of the brain. These comprise of task-related activation, hierarchies, the cooperation of different parts of the brain and psychological state of the subject [13]. However, the images that are acquired by a standard MRI machine are inherently noisy and contain complex spatiotemporal structure. Scientists try to clear data from unwanted noise and simplify the complicated data set to achieve more meaningful results with the use of statistical properties.

Several components make an analysis of an fMRI study prone to errors: These include head movement during the scan and inconsistency in data, including the variability between or within the subjects during the time course of fMRI. The analysis of the fMRI data is to deal with all these inherent problems, and remedies for these problems can be grouped into several components. Correction of the fMRI data set may be accomplished by [14]

- Fixing the spatial distortion on the data set
- Alignment of the images over the time course to get rid of the relative head movement
- Alignment of the time sequence of slices and the subject to create a framework so that the data can be used in group analysis
- Smoothing the data temporally and spatially to reduce the noise over the relatively weak signals

After these steps, the data might be ready for the further analysis.

2.2 Data Acquisition

In this section, an overview of data collection is presented in a concise manner. The data collected in an fMRI experiment is an aggregation of several MRI images. These images are obtained while the subject is carrying out an individual task. As mentioned earlier, a subject is located inside an MRI machine. The hydrogen atoms (H) in the body align with the magnetic field. A radio frequency pulse is used to tip over the aligned H atoms from their initial position. Immediately after the RF pulse, atoms that have been tipped precess from their aligned point in a way that they induce a current in the receiver coils of the MRI machine. The fully aligned slices in the brain can be altered via gradient coils throughout the experiment to complete the full brain images as they change the magnetic field over the brain.

The raw MRI images represent the spatial frequency of the proton densities of the brain tissue. Different tissues in the body create distinct spatial frequencies. Thus, they form a different contrasts on the reconstructed image. To acquire meaningful images, k-space representation of the imaged tissue needs to be sampled adequately. Different sampling schemes can be applied (uniform and non-uniform) over the k-space. In the uniform sampling, uniformly distributed points on the k-space are sampled, while in the non-uniform sampling, the k-space is sampled with different trajectories. Each sampling scheme has different advantages and disadvantages regarding resolution, SNR, and speed. Once raw data are collected, the Fourier transform is applied to it. Using the Fourier transform on the raw data creates actual images of the imaged tissue, since the raw data only represents the spatial frequencies of the tissues.

Researchers scan subject's brain while the subjects are performing certain tasks. Extracting the BOLD (blood-oxygen-level dependent) responses from the MRI images is quite challenging and the results are often indeterminable due to

spatial and temporal changes in the imaged brain tissue, the variability of the MR machine during the scan and physiological effects like breathing, head motion and heart beats [1]. Preprocessing algorithms are applied to readily acquired images in order to overcome these issues.

2.3 FMRI Noise Sources and Characteristics

2.3.1 Thermal and System Noise

MR imaging studies suffer from thermal noise related to the free motion of the electrons inside the subject and the machine. During the slice selection and excitation stages, the temperature inside the MR hardware increases, which in turn causes the electrons to collide with the atoms more frequently and results in unwanted current distortion inside the imaging machine. The same incident happens to the receiver hardware of the device. As induced currents pass through the receiver device, temperature increases, which results in more collision inside the hardware and causes more distortion on the image. Besides the thermal noise, another major type of noise is the MR system noise that is caused by the unsteadiness and variability in the imaging hardware. Known causes for this type of noises are [1]

- Static magnetic field inhomogeneities which in turn deteriorate the quality of image,which can also disrupt the contrast and geometry of the image because of the varying resonant frequency of the hydrogen atoms
- Instability of the gradient coils and their characteristics similar to the magnetic field inhomogeneities, which changes the shape and location of parts of the image
- Off-resonance effects in the RF transmitter and receiver circuitries, and drift in the magnetic field of the primary magnets, which result in inefficient excitation and causes enormous decrease in image quality.

2.3.2 Physiological Noises

The human body is not a stationary, inert object. Muscles interact with each other, the respiratory system changes the position of the body and heartbeats make small variations in the status of the body or subject may swallow or move his/her head unintentionally. These minor changes create motion related artifacts in MR signals. Even if we don't take into consideration the intrinsic noise sources of MRI hardware, these subject related movements create excessive degradation in the fMRI signals. BOLD related signals in MR are relatively small compared to the non-BOLD (anatomically) related signals. Since these variations are digitized, artifacts due to problems mentioned above can be modeled and removed from the digitized data if sampling is sufficient according to the Nyquist sampling rate.

2.4 Components of FMRI Preprocessing

The success and reliability of statistical processing algorithms depend on the consistency of the raw fMRI data sets. The primary objective of the preprocessing stage of the fMRI study is to ensure that the data is free from the artifacts causing the problems as mentioned above. Preprocessing has several steps.

2.4.1 Quality Assurance

Quality assurance is the starting point of the data analysis. The first, most important and easiest method in quality assurance is eye examination of the raw data because many artifacts in the raw data can be seen with the naked eye. There are several programs that can show time series images as an animation and this way any scanner dependent distortion can be detected quickly. Furthermore, researchers perform tests to ensure the statistical consistency of the raw data. No further processing is done on the data without passing this stage in order to save researchers from unnecessary work [13].

2.4.2 Slice Time Correction

In most standard MR scanners, data acquisition occurs in an interleaved manner, which means that the collected data belongs to the different time instants. The BOLD responses of adjacent slices would not be identical in consecutive slices even if they are related to the same event. These gaps between acquisition times form problems in data analysis. One remedy for this artifact is to interpolate time during preprocessing. Through interpolation the MR signals are corrected as if they are collected in the same instant so that the differences are minimized.

2.4.3 Head Motion Correction

Physical and mental status of the subject cannot be completely stable during 1-2 hour long sessions. Tasks that involve muscle movements, swallowing, and respiration, are the underlying causes of head motion during experiments. Several precautions (small breaks during a session to relieve subject, use of head stabilizer) can be done during and before the scan to prevent head motion. Preventing head movement is relatively easy compared to correcting it [1]. Although cautions are made, head movements still occur during the sessions. Some types of head movement only change the location of the voxel in the resulting images and can be corrected with software tools while others transform the data into meaningless throughputs.

Aligning consecutive images to a reference image is called co-registration. Since the shape and the volume of the brain don't change, the rigid body transform can be used to align the images. In this transformation, a cost function is applied to maximize the similarities between images. Six variables (x-axis, y-axis, z-axis, roll, pitch, yaw) are formulated to shift the input images onto the reference images as close as possible [13]. Once the alignment is successfully implemented, co-registered images are spatially interpolated to reduce the noise that comes with the co-registration and this procedure is applied to all acquired images.

Co-registration is also used between functional and structural images. fMRI images are relatively low-resolution images that only show a silhouette of the brain. Structural brain images, on the other hand, show more anatomical details. With structural images, the regions of interests in the brain can easily be located, and activation patterns can easily be seen.

2.4.4 Spatial Normalization

Morphology of the brain across subjects is different from each other. These variations prevent us from an inter-subject analysis. Spatial normalization techniques are types of co-registration that scale images to a normalized framework. Spatial normalization algorithms determine the sizes of individuals brains and try to compress or enlarge them to fit a known space. With this powerful tool, fMRI signals of different subjects can be aggregated for group analysis and let us test our hypothesis across subjects. A well-known and widely used method for normalization framework is the Talairach space. This scheme was created by Jean Talairach and imitated from the brain of an elderly woman [1].

2.4.5 Temporal and spatial smoothing

Temporal and spatial filtering are applied to the collected fMRI signals to increase the functional SNR [1]. Functional SNR is a term that expresses the differences between BOLD activities of ROI in two different conditions. In a nutshell, a functional SNR value enables us to detect and differentiate states of the experimental conditions.

Sampled data might be corrupted by physiological effects such as breathing and heartbeats. These effects can be diagnosed by applying a physiological examination on the subject before the experiment (typical respiration rate for humans

is around 0.2-0.3 Hz, and heart rate is 1-1.5 Hz). If we sample the data sufficiently, we can eliminate these physiological variables through the use of appropriate filters without effecting the task-dependent BOLD signals. Rectification of corruptions is also dependent on the experiment itself. If the experiment is fast-event based this means that conditions change every 1-2 second. In this case, the BOLD signals can be coupled with these unwanted effects, which makes removing them much more challenging. Researcher need to take precautions before the experiment. Another problem with the BOLD signals is the temporal correlation between the time points of the experiment. These irregularities cause data to become non-eligible to be used in statistical analysis and cause results to be more erroneous [13]. The remedy for this is whitening (intentionally adding white noise in) the data so that the temporal correlation between pairs is corrupted.

Spatial smoothing is applied in order to reduce the high-frequency component in the fMRI data sets. There are several reasons to reduce the spatial frequency in the data. Firstly, spatial smoothing blurs the images and increases the functional SNR. Generally, activation in the human brain occurs in multiple voxels. Smoothing data across voxels reduces susceptibility effects and minimizes unwanted noise in the data set. Secondly, it decreases the variability and the mismatch that is not adjusted with the spatial normalization. Most common method for spatial smoothing is convolution with a Gaussian kernel [14]. The degree of smoothing on the data set depends on the variance of the Gaussian kernel in that more voxels are involved in the smoothing with a large width kernel. There are different measures to the required Gaussian kernel but in general starting from a twice-voxel sized kernel and increasing from that start point is recommended because there is a trade-off between the increase in the functional SNR and the attenuation of the meaningful information. A Gaussian kernel with a smaller width has no positive effects on SNR and a wider one diminishes the significant information on the brain.

Chapter 3

Human Brain and Vision Related Parts

3.1 Brain and Its Structure

The brain is the main organ that is responsible for decision-making processes and sensory information. It is located inside the skull. The human brain shares the same structure as other mammals but, it has a more advanced cerebral cortex. It reacts real world stimuli and interacts with the environment. It contains millions of neuron cells, and each cell contributes to the transfer and the processing of specific information gathered from the other organs in the body [7].

The brain is the central organ of the central nervous system and interacts with the rest of the body via the spinal cord. It constitutes a centralized control mechanism over the body. The way to control the body is forming muscles' activities through the spinal cord and releasing feature specific hormones. It also processes sophisticated sensory inputs and integrates them to collect information from the environment. From a neurologist's point of view, the brain is the biological computer that collects and processes information from the world.

Brain consists of three main parts of which each one is responsible for a different type of processing. Basic descriptions of the brain parts are as follows [15]:

- **Cerebrum:** Cerebrum is the most advanced and largest part of the brain. It is located at the top of the brain. It comprises of two hemispheres (right and left) and their cortices. Cerebrum carries out functions such as triggering and coordination of movement, visual and audial processing, learning and interpretation, sense of touch and emotion.
- **Brainstem:** The brainstem is adjacent to the spinal cord and can be thought as the continuation of it. It structurally connects the cerebrum and cerebellum to the spinal cord. The brainstem controls all the automatic life dependent functions such as cardiac and respiratory systems.
- **Cerebellum:** Cerebellum is located under the cerebrum at the back of the head. It plays important roles in the motor movements. The cerebellum also coordinates some intentional actions such as speech, balance, and coordination.

The first aim of this study is to understand the attentional effects of the human visual system. Different subjects' brains were scanned and stimuli related data were collected while they were watching a movie. Evaluation of the attentional effects was made on the category-selective and attention-related regions of interest in the human brain. These category-selective and attention-related regions were localized with the MR scans for different subjects, , the data from which were used in this study. These regions of interest will be further explained in the following sections.

3.2 Cerebrum and Its Subparts

The cerebrum is the largest and the most developed part of the human brain. It includes two hemispheres each controlling opposite sides of the human body. The

hemisphere on the right side is considered to control the creativity and artistic skills. The other hemisphere is thought to relate to the arithmetic, comprehension and writing abilities.

The surface of the cerebrum is called the cerebral cortex. Most of the computational processing takes place in this area of the brain. It has a structure with folds that are called gyri and sulci which enhance its surface, therefore, increasing its computational power. It is the most superior part of the brain and encompasses more than half of it. Nearly all of the interesting calculations (such as thinking, perception, language processing, sensory information processing) occur in this area. Cerebral cortex is directly involved with the sensory informations and perception of the surrounding world. The cerebral cortex is subdivided into four parts [15]:

- **Frontal lobe:** This part is found at the front of the brain. Its functions are correlated with attention, concentration, short term memory, motivation, and judgment. It also controls the daily activities like walking and plays a role in speaking and writing skills.
- **Parietal Lobe:** The parietal lobe is adjacent to the frontal lobe , stretching towards to the end of the brain. The parietal lobe interprets sensory information and integrates information coming from different locations in the body. It is also responsible for the information about spatial and visual perception. It plays important roles in language processing.
- **Occipital Lobe:** Occipital lobe is located at the rear of the brain behind the parietal lobe. It is the primary visual processing center in the human brain. It comprises of most of the visual processing regions.
- **Temporal Lobe:** Temporal Lobe is located beneath the other three lobes towards to the brainstem. It interprets informations to reveal and organize the long term memory. It is also responsible for language and comprehension.

In summary, the cerebral cortex performs most of the processing related to the

sensory information and maintains cognitive functions of the human brain. Regions associated with the visual system and attention are further explained in the next sections.

3.2.1 Vision Related Regions

Discovery of the noninvasive imaging modality, specifically of the fMRI, facilitates the study of the human visual cortex and therefore enhances the understanding of the hierarchies of it. Two principles are employed to explain the visual cortex. In the first model (called the hierarchical processing), an image is first expressed in a local and conventional form. Through this stage of the processing, the image is turned into a more complete and complex representation [16]. In the second model, different properties of the visual scene are handled with parallel hierarchical streams [17]. Both these themes are approved by several studies in the fMRI researches.

In this thesis, not the whole human visual cortex, but only the human and scene selective regions are described and assessed. The recent finding about these areas for these two objects of interest are clarified. Attention-related regions in the brain are also explained. List of the areas of interests and their role in the higher visual system is as follows:

- **FFA (Fusiform Face Area):** This section, as its name indicates, is activated with perception of faces. Kanwisher, et al. found that activation pattern of this area six times greater when subjects passively watch human faces [18].
- **EBA (Extrastriate Body Area):** This region is named to indicate that it shows significantly higher responses for body parts than for inanimate objects [19]. Kanwisher et al. also showed that this area shows relatively higher responses to human body parts except for faces.
- **MT (Middle Temporal Visual Area):** This region responds more significantly when the subject sees non-stationary objects. This area performs

a significant role in the recognition of the motion. MT is also selective for biological motion, meaning the region is activated when a subject views walking people, hand and mouth movements of a person [20].

- **PPA (Parahippocampal Place Area):** Results in [21] indicates that this region is processing the information about the spatial domain. They reported that the PPA is responsible for scene recognition and also gives more marked responses to familiar places.

- **RSC (Retrosplenial Cortex):** RSC is involved with the spatial navigation, memory related actions and scene processing in the brain [22]. Damages that occur on the RSC might result in significant memory loss and navigation deficiency. In another study [23], It is stated that RSC is more selective when the number of stable items in the scene is increased.

- **TOS (Transverse Occipital Sulcus):** This area is located on the dorsal occipital-temporal cortex and is activated for construction and scene related stimuli [24].

- **RET (Retinotopic (Early Visual Areas)):** Early visual areas are represented under the category of RET. In this area, a 2-D projection of the 3-D objects is demonstrated. Retinotopic mapping was introduced by the research of Tootell et al. [25].

- **LOC (Lateral Occipital Complex):** This area is associated with the recognition of an objects. LOC is considered the general purpose system for object analysis and categorization [26].

- **V7** This area responds strongly to the spatial attention. V7 also shows robust activation in response to several different categories of stimuli[27].

- **IPS (Intraparietal Sulcus):** IPS serves as one of the visual attention centers in the brain. This area is activated in directed attention regardless of the stimuli presented [28].

- **FEF (Frontal Eye Field):** This section is responsible for the generation and control of the eye movements and adjusting the visual attention [29].

- **SEF (Supplementary Eye Field):** According to the work of Purcell et all [30], SEF is the main region responsible for auditing the visual search performance. This study also states that it has a minor effect on continuing search behavior.

- **FO (Frontal opperculum):** This region is responsible for controlling the activity of other brain areas to fulfill cognitive tasks such as attention [31].

Chapter 4

Methods

In this work, we investigate the effects of attention on the responses of the human brain. Our experiment, preprocessing and methods to analyze the data are explained in the following parts.

4.1 Subject

Five healthy male subjects took part in the experiment. Age of subjects were between $25-30$. All subjects had normal or corrected to normal vision, and they all signed a consent form. The procedure of the experiment was approved by The Institutional Review Board at the University of California Berkeley.

4.2 Stimuli

To understand the effects of attention, a natural movie clip was created using $10-20s$ long short clips without repetition. For each attention condition (human or vehicle) $1800s$ of a natural film was created. Visual angle for the video was 24° x 24° and its resolution was 512x512. Each movie was created in a similar

manner where only one of the attended objects appeared in the movie for $450s$, they co-appeared for $450s$, and none of them appeared for $450s$. Attended objects appeared in immensely different sizes, shapes, visual angles, and positions. A fixation point for the subjects (0.16° square) was located on the film and its color was switched at 1 Hz to ensure its visibility. The resultant video was shown to the subjects with a mirror and projector system during the scan.

4.3 Experiment

All experimental data were acquired at the University of California Berkeley, as part of a previous study on visual attention [33]. This thesis reanalyzes that dataset to address different neuroscientific questions than the previous study. Each subject was scanned seven times to collect not only functional data but also anatomical, functional localization and retinotopic mapping data in the original experiment. The functional data was collected in one session. During this session, six attention conditions each lasting $600s$ were performed by the subjects (three runs for the human and three runs for the vehicle category objects). Subjects were fixated to a point while they were seeking for a category of interests in the movie. To stay alert during sessions, they pressed a button whenever the attended object group was seen in the film. A cue word was shown before each run to inform the subjects about the attended category. The attended category was switched after each run and same with all the other runs. To create mutually exclusive stimuli clusters, each type of stimulus (human, vehicle, both of them, none of them) was randomly picked and evenly distributed within and between the runs. The hemodynamic offsets between the voxel responses and the movie stimuli were compensated by adding the last $10s$ of the film to the start of the film of that run. The offset data were extracted from further processing. Subjects passively watched $7200s$ of natural video in the other three sessions. These extra sessions were not used in this study.

4.4 MRI Data Acquisition

The MRI data acquisition took place in the University of Berkeley with a $3T$ Siemens scanner with a 32-channel head coil. The MRI sequence for the functional data was a T_2 weighted gradient-echo echo planar imaging that was modified with water-excitation RF pulse. Head motion was prevented by using a foam padding. Sequence parameters was as follows: Slice number = 32, TR = $2s$, echo time = 34 ms, flip angle = $74\,^{\circ}$, voxel size = 2.24 x 2.24 x $3.5mm^3$, FOV = 224 x 224 mm^2. Anatomical data was collected with a T_1 - weighted MP-RAGE sequence. The parameters for anatomical scan were voxel size = 1 x 1 x 1 mm^3 and FOV = 256 x 212 x 256 mm^3. The anatomical and the retinotopic data for two subjects were collected with a $1.5T$ MRI scanner. Retinotopic mapping data were collected in four different sessions. Each session lasted nine minutes with various stimuli (rotating polar wedges and widening and shrinking rings). Motor localizer was a ten-minute scan. Each subject performed six different motor tasks (hand, mouth, foot, speech, rest and saccade blocks). Middle temporal visual area (MT) was localized with using four 90 seconds natural movie as stimuli. Category-selective regions were extracted with six 4.5 minutes scans. Stimuli used in the visual localization consisted of places, faces, animals, objects and human body parts. Auditory cortex was localized with various types of sound stimuli [32, 33].

4.5 Data Pre-Processing

Non-Brain tissue was extracted from the brain tissue using the Brain Extraction Tool (BET). Functional data between and within the runs was aligned to the first functional image acquired from the subject. The alignment was done with the Statistical Parameter Mapping Toolbox (SPM8). The BOLD responses arising from button press task were identified and removed from further processing. The cortical surface of the subjects was constructed with Caret5 software. Voxels of the cortical surface were determined such that the voxels which is located

within the $4mm$ radius of the cortical surface were selected. To remove the low-frequency drift (baseline wander) from the datasets, the Savitzky-Golay filter (which uses the low-degree polynomial to smooth the data without distorting the signals) was applied. No extra spatial and temporal smoothing was applied. The datasets for each subject were normalized to become zero mean and unit variance (z-score) in order to enable us to compare the results from different subjects easily. No subjects' data were transformed into known brain space. Instead, two-dimensional flat-map representations were used. The data from the functional localizer and the retinotopic mapping was used to identify the particular region of interest (ROI) for each subject.

4.6 Category Model

WordNet Lexicon [34] database was used to label the object categories within the natural movies. In this database, English words are not organized based on their spellings but they are arranged with respect to their meanings and semantic features. Conceptually related words are linked to each other within hierarchical orders. The natural movie was labeled by three raters. Since the words are connected in a hierarchical manner, the existence of any category in the movie also denotes that its higher order classes also exists. It means that the presence of a 'lorry' also indicates the existence of following: 'vehicle', 'conveyance'. The raters labeled 604 different object categories in the film. 331 higher order categories of these objects were also included. The stimulus matrix ($category \times time$) for further processing is composed of $1s$ and $0s$ (former indicates the presence, and the latter represents the absence of the category of interest).

4.7 Model Fitting

In this work, we used the BOLD responses and the natural movie as inputs to our system and tried to create a model that reveals the effects of attention on the

human brain. As a starting point, we chose decoding routines as our modeling procedure between the stimuli and the response. In other words, we first tried to predict the stimuli that formed the obtained responses based on constructed models. Secondly, some metrics were computed to evaluate the performance of the model and clarify how attention effects decoding performance. Besides, these procedures were applied to different regions of the brain to gather a complete set of results on the visual attention.

Some alterations were done on both stimuli and the response data sets to apply a decoding procedure successfully. Firstly, stimulus matrix (category $\times$ time course) was down-sampled by 2 to balance time points of the BOLD signals and the stimuli. Additionally, some changes were applied to the response matrix to compensate for the hemodynamic shift. Originally, the response matrix contained the responses of each voxel over the period of the experiment (voxels $\times$ time course). We added shifted versions of the each voxel response as a new column to the response matrix. The new matrix ($4 \times$ voxels $\times$ time course) contains the time slided version of each voxel response from $2s$ to $6s$.

We modeled voxel responses as input and stimuli as output in order to create a transformation model. We divided our datasets into training, cross-validation and evaluation chunks to form and evaluate a model. Logistic regression algorithm was applied on the training datasets to obtain the models of the brain. The performance metrics were calculated on the cross-validation and the evaluation sets. We repeated this procedure 1000 times with different training, cross-validation and evaluation sets to create aggregations of the results and evaluate their statistical consistency.

4.7.1 fMRI Models

Decoding models use voxel responses to estimate information about the real world (stimuli). Their complimentary operation is called encoding models that uses the opposite relationship between stimuli and response. If we formulate the response and the stimuli as probabilistic events, we can express our decoding and encoding

models with these prior probabilities. Let

- **P(r)** be the probability of the voxel responses
- **P(s)** be the probability of stimuli that form the responses.
- **P(r,s)** be joint probability that r is the response, s is the stimulus
- **P(r|s)** be conditional probability of a formed response r given that stimuli is present.
- **P(s|r)** be the conditional probability of stimuli was present given that response is occurred.

variables represent the probabilistic events. Joint probability can be described as the conditional probability of evoked response given stimuli times stimuli or the conditional probability of stimuli given formed response multiplied by the probability of response.

$$P(r, s) = P(r|s)P(s) = P(s|r)P(r) \tag{4.1}$$

A decoding model can be modeled as the likelihood of stimuli given response **P(s|r)** on the other hand an encoding model can be modeled as the likelihood of BOLD response patterns given the stimuli present **P(r|s)**. Using Bayes Theorem these probabilities can be expressed with each others.

$$P(r|s) = \frac{P(s|r)P(r)}{P(s)} \tag{4.2}$$

$$P(s|r) = \frac{P(r|s)P(s)}{P(r)} \tag{4.3}$$

From these equation sets, we formulated our decoding procedure and tried to model human brain with prior stimuli and response probabilities.

In our work, we used decoding models to assess attentional effects on the human brain. We collected BOLD responses from each subject while they were viewing the movie sets. These datasets have low SNR value and massive size.

Because of these inherent characteristics of the fMRI datasets, tremendous efforts needs to be exerted on them to reveal any meaningful data.

Multivariate analysis leverages pattern recognition and, unlike the single voxel analysis, uses a population of voxels as a whole to increase its sensitivity [37]. It also includes the responses of non-significantly sensitive voxels in the calculation and extracts substantial information. Multivariate analysis are used by a majority of the decoding studies in this area [36]. This procedure can disclose the information hidden in a set of data which is its advantage over conventional spatial averaging type analysis [38, 39].

Decoding studies might also be used to decode mental and cognitive states of the subject's brain. Previously studies can only show which stimuli or task activate which region of the brain entirely because of resolution constraints. By increasing sensitivity using the multivariate analysis, decoding studies shows significant improvements and can even detect sudden $(2s-4s)$ changes in the mental states.

The main advantage of multivariate analysis comes from representing the responses and the stimuli as high dimensional spaces [37]. Responses of the brain and the stimuli are represented in these high-dimensional vector spaces, and every component and property of stimuli or activity and voxel of interest are described as new dimensions. If we measure the 1000 voxel over 900 s of time course, we represent the response matrix with 900x1000 dimensional space. Similarly, the stimuli pattern over the time course is formed and described with 1×900 dimensional space. Once these representation spaces are created, they may be manipulated with various mathematical tools to unveil substantial information and differentiate the several mental states required for one task. Multivariate analysis use machine learning algorithms and try to constitute a decision threshold that discriminates differences between the cognitive states and the high dimensional patterns [37].

In terms of decoding and encoding, multivariate analysis may be considered a manipulation and evaluation of the response and stimuli vectors on the hyperspace. Decoding can be evaluated in order to form a transformation matrix that converts response patterns to stimuli patterns. Although cortical topography and anatomical structure are the essential characteristics of brain responses, these features are discarded. They can be evaluated with other analysis methods such as searchlight [40].

In our work, we performed our analysis on different regions of the brain. Voxels on each region were treated as a whole, and logistic regression algorithms were applied to each of them separately. We began our analysis by splitting response and stimuli matrices as training, test and evaluation subsets. These datasets were entirely independent of each other. The decision threshold that determines the presence of stimuli on a given response pattern was computed on the training sets with different regression parameters for each subject. Best regularization parameter was selected on the test set. Validation of the decision threshold and regularization parameter was done on the evaluation set. Test and evaluation data had no effect on the classifier to have a viable testing. We perform this procedure on each subject for both of the attention conditions. The performance of the classifiers was calculated with the d-prime parameter.

4.7.2 Signal Detection Theory and D-prime

Signal detection theory helps us to detect information and elucidate how decision is made under ambiguity. This method provides an invaluable mathematical framework to probe the relationship between stimuli and response in psychological studies. Consider the following example to come to a deeper understanding of detection theory in psychological studies. A person is passively viewing the natural video while their BOLD responses are recorded. BOLD responses for different objects are distinct from each other for a particular ROI. Because of the loss of attention the subject could not recognize the object of interest and this creates noise on the BOLD responses. If we try to determine the object using the data

from the BOLD responses, the signal detection theory accounts those noises into its calculation and try to formulate an optimal decision threshold. The threshold might be used to resolve the presence of the object.

To explain detection framework, some basic terminology needs to be explained. A complete picture can be described with 4 different metrics in detection theory. The first metric for performance of a detector is the hit rate. Hit rate can be explained as the proportion of the correct YES responses when stimuli is present during the basic YES/NO trials.

$$\text{Hit Rate} = \frac{\sum_{correct} \text{Yes}}{\sum \text{target}} \tag{4.4}$$

It is obvious that a good detector performs with a high hit rate. But this metric unfortunately is not enough to completely evaluate the performance of the detector. This is because the hit rate disregards the information when there is no stimuli presented during the trial. The hit rate depends solely on the correct answers and does not account for the mistake. The second metric for the detector is known as false alarm. The false alarm rate is the number of deceptive YES answers when there is no stimuli present on the trial. False alarm rate can be explained as follows:

$$\text{False Alarm Rate} = \frac{\sum_{false} \text{Yes}}{\sum \text{no target}} \tag{4.5}$$

The third metric in detection theory stands for misses in the signal trials. It evaluates the performance of a trial with its wrong answers. It calculates the ratio of the wrong responses on the signal trials (when stimuli are present on the trial).

$$\text{Miss Rate} = \frac{\sum_{false} \text{No}}{\sum \text{target}} \tag{4.6}$$

The last metric of the detection theory calculates the correct differentiation of the absence of the stimuli on the noise trials.

$$\text{Correct Rejection Rate} = \frac{\sum_{correct} \text{No}}{\sum \text{no target}} \tag{4.7}$$

These 4 metrics can be tabulated as follows:

	Yes Response	No Response
Stimuli Present	Hit	Miss
Stimuli Absent	False Alarm	Correct Rejection

In detection theory, the metrics that are most commonly used are hit and false alarm rate. Others might be though of a complement of the first two. Although, these two metrics together might explain the behaviour of the detectors under certain conditions, one single quantity that stands for the sensitivity might be more desirable. Signal detection theory uses both hit and false alarm rates to create a concrete mathematical formulation for the detection process. It also provides a metric for detectability useful for ambiguous decision-making processes. In the signal detection framework, hit and false alarm rate is turned into significant quantities.

Modeling trials as probabilistic events led us to utilize signal detection framework. Simplest form of the model is the Gaussian model in which both distributions are the same but one of them is shifted to one side. We can model noise and signal trials as follows:

$$X_n \sim N(0, 1) \tag{4.8}$$

$$X_s \sim N(d', 1) \tag{4.9}$$

where N represents the Gaussian distribution

D-prime value explains the difference in the means the two identical Gaussian curves. When the d-prime value is large, the two curves are well separated from each other which is an indication of an acute detector of the conditions. Since we model the trials as equal variance Gaussian models as above, the d-prime value can be calculated from the hit and false alarm rates as follows [2].

$$\text{d-prime} = z(\text{Hit Rate}) - z(\text{False Alarm Rate}) \tag{4.10}$$

Where z is inverse cumulative distribution of the normal Gaussian distribution.

Weak signals can be hidden because of the bias effects [41]. Here in our study we directly measure the attention effect on the category selectivity of particular

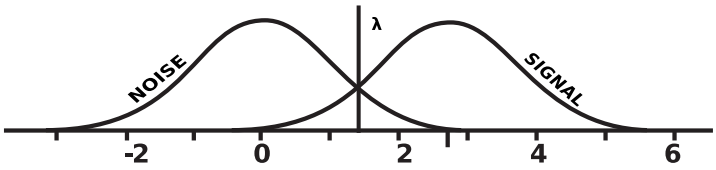

Figure 4.1: Signal and noise trial. The threshold value is selected to minimize false alarm rate and maximize true positive rate for a given experiment. D-prime value represents how far means of these two curves are located between each other. This figure is adapted from [2].

brain regions. The d-prime metric enables us to measure bias-free statistical tests on our datasets.

In our work, the signal detection theory connects response of the subjects to the natural movie stimuli. Responses for both cases (when stimuli are present and absent) can be considered as two separated (with different means) Gaussian curves. We trained our model on the training sets for several regularization parameters using logistic regression. D-prime value is calculated to assess model performances. The best regularization parameter was selected in the trial. The final d-prime value was found on evaluation sets. D-prime value on our experiment shows us how particular ROI responds to a specific category and how its responses change with the directed attention.

Another metric used to evaluate the performance of a classifier is the receiver operating characteristic (ROC) plots. ROC curves illustrate the classifier performance while the decision threshold is varied. ROC curves are plotted with the false alarm rate located on the horizontal axis and the hit rate on the vertical axis and show trade-off between these two metrics [42]. With this two-sided mapping, the performance of the detector may easily be evaluated based on a single curve. If the probability distributions of the signal and the noise trial are known, the ROC curve can be plotted as the cumulative distribution function of both of the distributions that are calculated from various decision thresholds to inf.

$$\text{false alarm} = \int_{T}^{inf} f_0(x)dx \tag{4.11}$$

$$\text{hit rate} = \int_{T}^{inf} f_1(x)dx \tag{4.12}$$

On the ROC curve, the best possible outcome is the one in upper left corner that indicates %100 hit rate and %0 false alarm rate. On a given trial if the two Gaussian-shaped curves are well separated from each other, we might easily choose an appropriate decision criterion to increase the hit alarm rate. A suitable threshold changes ROC to take a shape of an upward bow like curve. On the other hand, an entirely random classification might be located on the diagonal line [42].

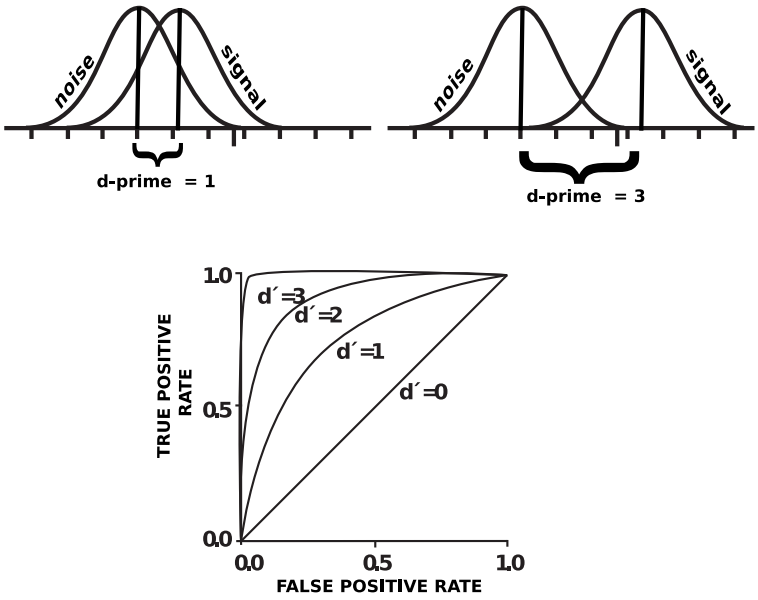

Figure 4.2: ROC curve with different d-prime values. Discrimination between the trials increases with the d-prime value. Different d-prime values yield various ROC curves. This figure is reinterpreted from [3].

ROC curves are two-dimensional plots of the classifier's performance on a given trial. The area under the ROC curve is calculated to evaluate the performance of classifier with a single value. The area under the curve (AUC) always takes a value between 0 and 1 since probability distribution is used to calculate the curve. Better classifiers constitute higher AUC values for a given trial. On the other hand, a random classifier forms the diagonal line on the ROC curves and its area is 0.5, so no practical classifier should have a value below 0.5.

4.7.3 Logistic Regression

In our experiment, we are interested in the relations of input and output between real-world stimuli and BOLD responses of the subjects. We were trying to find a model of the subjects' brains that enables us to classify information. Because we modeled the natural movie as a binary variable, '1' is the indication of the presence of the object category. Classification is the name for this procedure and its one of the cornerstones of statistics and machine learning algorithms.

In the previous parts, we modeled response and stimuli datasets as conditional probabilities of one another. Since we are using decoding procedure we are using **P(s|r)** as a starting point and any unknown parameter in the probability distribution can be estimated with suitable estimators. In case of linear regression we can approximately model stimuli as a function of response patterns.

$$h_\omega(x) = \omega_0 + \omega_1 x_1 + \omega_2 x_2 + ... \tag{4.13}$$

$$h(x) = \sum_{i=0}^{n} \omega_i x_i = \omega^T x \tag{4.14}$$

Above formula ω_i are weights that maps input to output variables, n is the number of inputs, x represents input patterns and y represent the output. We first assume that our stimuli datasets contains values of 1 and 0 so that we can change our modeling function for logistic regression as follows:

$$h_\omega(x) = g(y\omega^T x) = \frac{1}{1 + e^{-y\omega^T x}} \tag{4.15}$$

$g(-y\omega^T x)$ is known as the logistic function and take a value between 0 and 1 [4]. If we formulate our decoding assumption with a logistic function, formulation for the decoding procedure becomes as follows:

$$P(s = 1, 0|r; \omega) = h_\omega(r) = \frac{1}{1 + e^{-s\omega^T r}} \tag{4.16}$$

where ω represent the model parameters, s represents the stimuli, r represents the response patterns

There are several ways to estimate the model parameters on the above equation. A common method is to use the maximum-posteriori (MAP) estimator [43].

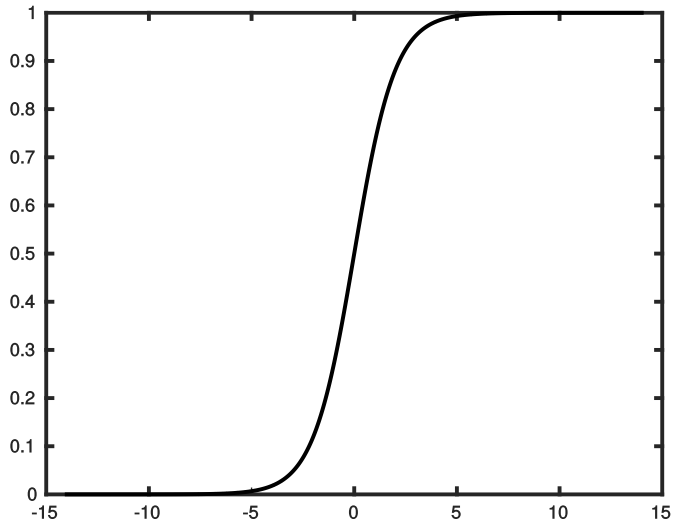

Figure 4.3: Sigmoid function. Our modeling function is changed to estimate the binary classification [4].

The starting point in the MAP estimator is defining a prior probability. Let us define a gaussian probability with zero mean and parametric deviation. Variance parameter in the equation below is known as the regularization parameter and accelerates the algorithm in high dimensional spaces [44].

$$P(\omega) \sim N(0, \lambda^{-1} I) \tag{4.17}$$

On a given response and stimuli dataset, we want to find model parameters ω that maximize the likelihood function. Let's model likelihood of m different training samples

$$l(\omega) = -\sum_{i=1}^{m} log(1 + exp(-s_i \omega^T r_i)) + \frac{\lambda}{2} \omega^T \omega \tag{4.18}$$

For maximizing the above equation, we need to take the partial derivative with respect to ω and find the point that maximizes the likelihood. There are several iterative algorithms to find the parameter spaces but explanation for those algorithms is beyond scope of this thesis.

4.7.4 Searchlight Analysis

Searchlight analysis technique is easily applied to the multivariate analysis in fMRI researches. Its aim is to identify and characterize informative areas on the human brain. The important assumption underlying the searchlight analysis is that adjacent voxels in the brain show similar brain activation patterns over time courses of the trials. Searchlight analysis takes advantage of this assumption and applies the multivariate analysis technique to the brain responses. With the searchlight analysis we might answer the question of where the localization occurs in the brain for specific stimuli and therefore how the spatial structure of the human brain is connected [45].

Searchlight technique creates synthetic spheres, composed of several adjacent voxels, for each voxel of the brain. Those groups of voxels might indicate the pattern of interest far better than the voxelwise analysis [40].

In our procedure, we created circular spheres for each voxel on the cortical surface of the subject's brain. We again divided our data into training, evaluation and test samples. We trained our model for each sphere with the logistic regression and found the best regularization parameter that yielded better d-prime values on the evaluation set. We repeated this procedure for two different attention conditions. In each attention condition, we decoded two separate categories of object and found the differences of d-prime for each attention condition. The difference in the d-prime value is visualized on the flat maps of the subject's brain which are the projection of a human brain on a flat surface.

4.7.5 Procedures

In this work, readily available data of responses from 5 different subjects and stimuli were used. Data was collected at the University of Berkeley and used in several studies [33, 32]. In this study, we applied preprocessing algorithms on the

data to remove the unwanted noisy effects. Two attention conditions were performed by the subjects in the original experiment. During the functional scan, attention conditions were alternated for each category of objects and the subjects were warned about the category of interests before each session. Originally, stimuli datasets consisted of 1705 different category and action labels but in this study, we extracted human and vehicle stimuli from the large datasets to apply our procedures. Category selective areas (FFA, EBA, MT, PPA, RSC, TOS) and attention control regions (IPS, FEF, SEF, FO) were extracted from functional data of each subject for further processing. But some attention-related regions were not analyzed in this study such as temporoparietal junction, precuneus, dorsolateral prefrontal cortex and inferotemporal cortex. Attentional effects on the response of the human brain were studied, and several performance metrics were calculated for each category and subject.

On the modeling side, training data from the stimuli and the responses of each subject were extracted from the original datasets in a way temporal correlation between stimuli and response matrices maintained. Both response and stimuli data were divided into chunks. The chunks were selected randomly to construct the training dataset. Nearly %80 of the response data and its associative stimuli were used as the training dataset. Logistic regression routine was applied to the training datasets with different regularization parameters. Before the logistic regression method was used, areas of the brain that are related to vision and attention were selected for both of the hemispheres for each subject. Responses from the same regions from different hemispheres of each subject were aggregated together and used to form the models. Voxels in each region of interest created a cluster allowing each voxel to contribute to the modeling function of the logistic regression. This way we took advantage of the benefits of the multivariate analysis. The logistic regression routine was applied to each subject and the region of interests separately and thirteen different model parameters for each subject were constructed.

The best regularization parameter and model were selected from the test set and performance metrics were calculated on the evaluation sets for each region of all subjects. Both datasets were chosen randomly like train datasets, and they

were mutually-exclusive. D-prime values for different regularization parameters were calculated for each region of each brain. To calculate the d-prime values, hit and false alarm rates were obtained. Based on the findings from the test set, regularization parameter and regression model that yield higher d-prime values were applied to the evaluation set. Final d-prime values for each region of interests of each subject were calculated based on selected regularization and model parameters.

This procedure was performed over 1000 times with original datasets. In each iteration, subsets were reselected randomly, and they were guaranteed to be mutually exclusive sets. Best regularization parameter and model were reselected in each trial. The final d-prime value was calculated by taking the means of the d-prime values of each case on the evaluation sets. Bootstrapping were applied to the ultimate results of all of the subjects and statistical significance were assessed.

ROC curves and their AUC were calculated for different ROIs of each subject. For each ROI, several regularization parameters were used to obtain best modeling parameters. Optimal regularization parameters were obtained on the test sets for corresponding ROIs of each subject. Based on those regularization parameters optimum models were extracted on the evaluation set. False and true positive rates that resulted from the model parameters were calculated. This algorithm was iterated 100 times to retrieve a concrete result.

For searchlight analysis, voxels on the cortical surface of the brain were selected. Cubes from the surrounding voxels were created for each selected voxel. We modeled the stimuli separately for each of the created cubes, and each voxel within the cube contributed to it. This mentioned procedure was applied to each voxel and d-prime values were obtained. The findings for each voxel were plotted on the flat map representation of each subject.

Chapter 5

Results

As mentioned earlier, the subjects were viewing an aggregation of short movie clips of 1800 s. Subjects were performing two attention conditions for distinct categories of objects. Different brain regions were thought to be activated by these two classes of objects in previous studies [19, 21, 18]. We collected brain responses of these two sets of stimuli for different attention conditions in our case.

In the first graph below, we calculated the d-prime values for the human decoding performance when the subjects were attending to the human and vehicle objects. That means we calculated two d-prime values for human, one for when the subjects are attending to the human objects, and one for when the subjects are attending to the vehicle. We differentiated these d-prime values to evaluate the decoding performance improvement.

Above mentioned procedure was applied to the data each one of the subjects. The aggregation of the d-prime values was obtained. The bootstrap method was used in the collection of the d-prime values of the subjects and re-sampled 10000 times. Blue bars on the below graph show the mean of the bootstrapped d-prime values for the different regions of interests and red lines on each bar show the sample standard deviation of those re-sampled d-prime values.

Although, all of the regions show improvement in d-prime values, regions related with attention (IPS, FEF, SEF, FO) shows relatively high performance improvement on the decoding performances and all of the ROIs show statistically significant results.

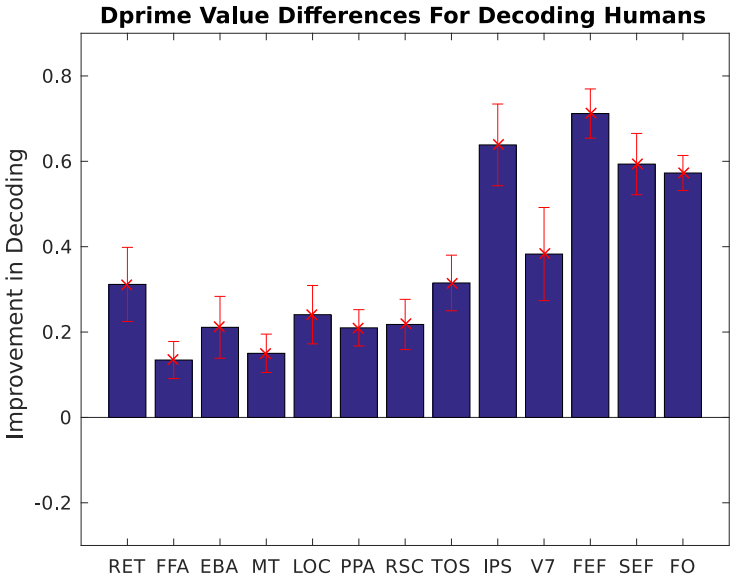

Figure 5.1: Decoding performance difference of the 'Human' object between attended and unattended stimuli. This illustration was obtained by subtracting the d-prime values of the 'Human' when subjects were attending and not attending to the object of interest. All of the regions of the brain, whether they are inherently selective to the 'Human' object or not, enhance their detection performances with visual search. Error bars show standard deviation that was calculated by bootstrapping d-prime values across subjects.

Table 5.1: Selected Regularization parameters for human decoding improvement

	RET	FFA	EBA	MT	LOC	PPA	RSC	TOS	IPS	V7	FEF	SEF	FO
Decod Human 'TC'	400	400	400	6400	400	6400	1600	1600	1600	1600	1600	1600	6400
Decod Human 'SN'	6400	6400	1600	6400	1600	400	400	1600	6400	6400	1600	400	6400
Decod Human 'JG'	1600	1600	6400	6400	1600	1600	6400	400	1600	6400	1600	1600	6400
Decod Human 'AH'	1600	1600	400	1600	1600	400	400	400	1600	1600	6400	400	6400
Decod Human 'AV'	1600	1600	1600	1600	1600	1600	400	400	1600	6400	400	400	400

In the above table, selected regularization parameters for decoding 'human' object is represented. Best regularization parameters were selected on the test sets for each subject and ROI separately.

In the second graph below, d-prime values for vehicle decoding case are calculated. We estimated the increase in the decoding performances by calculating two distinct d-prime values when subjects are attending one of the object categories. These two d-prime values were subtracted to assess the performance increase. Similarly to the first graph, attention specific regions benefit more that the other ROIs from the attentional effect and increases their decoding performances. Also, all of the ROI in the brain boost their performances when they account for the attentional effect.

Here, we have so far analyzed the attentional effect on the decoding performances that were acquired while the subjects were attending the both of the stimuli by turns. If we describe each ROIs as the detector of each category of object, attention changes the tune of the detectors and increases their sensitivity.

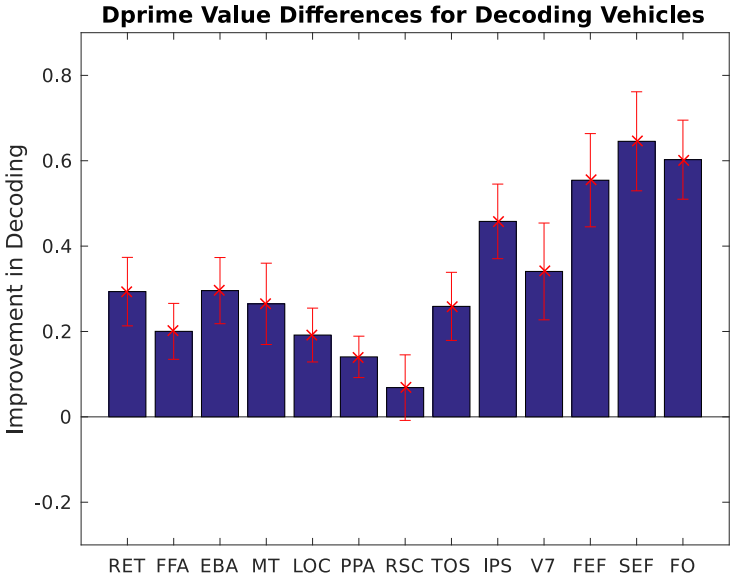

Figure 5.2: Decoding performance difference of 'Vehicle' object between attended and unattended stimuli. This graph is calculated by subtracting obtained d-prime values from the subjects' two conditions (attending or not attending the category of interest). All of the regions benefit from the visual search and increases their detection performances.

Table 5.2: Selected Regularization parameters for vehicle decoding improvement

	RET	FFA	EBA	MT	LOC	PPA	RSC	TOS	IPS	V7	FEF	SEF	FO
Decod Vehicle 'TC'	1600	1600	1600	1600	400	400	400	400	1600	6400	400	400	1600
Decod Vehicle 'SN'	1600	400	400	400	1600	1600	1600	1600	1600	1600	400	100	400
Decod Vehicle 'JG'	6400	400	100	400	1600	400	1600	400	400	6400	400	400	400
Decod Vehicle 'AH'	6400	1600	1600	400	1600	1600	1600	1600	6400	1600	1600	400	1600
Decod Vehicle 'AV'	6400	1600	400	1600	6400	400	1600	400	1600	1600	400	400	100

In the above table, regularization parameters that were for decoding 'vehicle' object is shown. Best regularization parameters were selected on the test sets for each subject and ROI separately.

In the third graph below, we differentiate the two obtained two d-prime values for each ROIs from the previous figures. Obtained results are bootstrapped and their means are shown as blue bars, and their standard deviations are shown as red lines. Finding for the final results can be stated as follows:

- FFA is inherently sensitive to faces. When we analyse decoding performances of this region, it can be clearly seen that its vehicle decoding performance increases more than its human decoding performance.
- EBA is selective for the human body and body part. It shows the same characteristic with FFA.
- MT is responsible for detecting biological motions like human walking. Its vehicle decoding performance benefits more from attention than its human decoding performances.
- Areas innately selective for places and scene processing (PPA, RSC, TOS) show the opposite relation to the face discriminating areas. Their human decoding performances are enhanced by the attentional effect.
- Attention-related areas (IPS, FEF, SEF, FO) are divided into two categories. IPS and FEF increase their human decoding performances with the attention. SEF and FO seem as if they benefit more in case of the vehicle attention, but their results are not significant.

- Object-sensitive ROIs (LOC and V7) increase their human sensitivity more than any other object. However, the results regarding the LOC are not significant; they show the same characteristic.
- Position-selective region (RET) shows insignificant results and can't be evaluated.

Our results suggest that attending to an object boosts the detection performance most significantly in brain regions that are not inherently sensitive to that object.

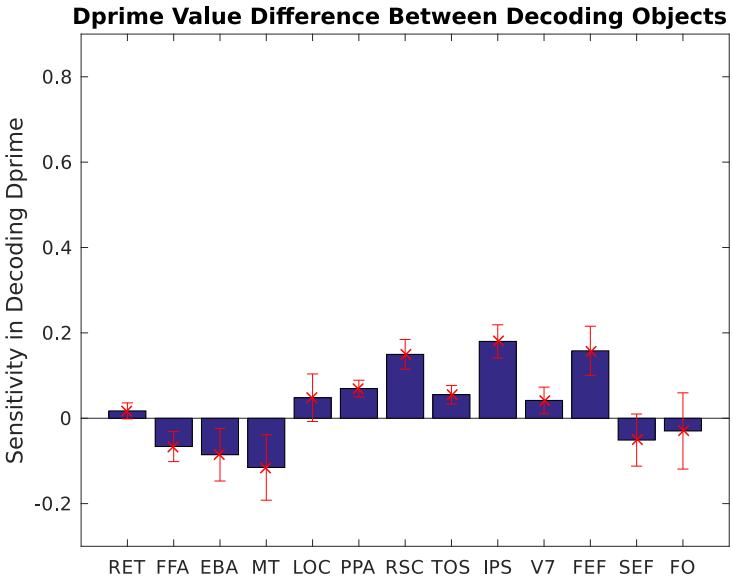

Figure 5.3: Decoding differences between two object categories. This graph illustrates the decoding sensitivity changes over attention in different ROIs. This plot is acquired by subtracting the Fig:5.1 from Fig:5.2. 'Human' and 'Scene' selective areas shows different characteristics. Human selective areas take advantage of the visual search of 'Vehicle' more than the visual search of 'Human'. Scene selective areas tend to show the opposite relation.

In our second analysis, we analyze the decoding performance by calculating the ROC and the area under the ROC. As we did in the d-prime calculation, we divided response and stimuli data of 'vehicle' and 'human' objects for three

mutually exclusive datasets with different sizes. After that, we implemented logistic regression routine with several distinct regression parameters on the training datasets. Models were formed for each RO of the subjects with the constraint of regularization parameters and the best model was selected. This algorithm was iterated for 100 times for selected ROIs. To calculate the ROC curve and the area under it, we aggregated the selected stimuli, the response datasets, and best model parameters and found the true and false positive rates to obtain the ROC for each subject.

ROC curves represent how well the model estimates the outcome variables. Upper bowed shaped curve depicts a successful estimator and the degree of the success can be understood from the form of the figure. In the below graphs, we visualized four different ROC curves for each subject and each ROI. Object and human selective areas were picked from the ROI list, and its outcomes were demonstrated. In each graph, first two letter represent the subject's initials, consecutive two letters indicates the following:

- HH stands for decoding Human when attending Human case
- VH stands for decoding Human when attending Vehicle case
- VV stands for decoding Vehicle when attending Vehicle case
- HV stands for decoding Vehicle when attending Human case

Our previous findings are supported by these results. Attending to an object and decoding that particular object in that datasets increases the performance of the model performances in general. Model performances are improved more in the vehicle attention cases for human selective ROIs, and the exact opposite occurs in the human attention case for scene selective ROIs.

In below graphs, the first column demonstrates the ROC curves of attending and decoding the same object category. The second column in each ROI demonstrates the decoding performances without the attention for that object category.

For PPA, although some discrepancies is found, attention increases the detection performances of the human objects across the subjects. In RSC and TOS, all of the subjects benefit more from attention on human detection and our previous findings are supported with the ROC graphs.

For the human selective ROIs, visual search enhances the detection performances of the vehicle objects more from the BOLD signals. For FFA, all of the subjects gain advantage in the favor of the vehicle object category. In spite of some discrepancies, EBA and MT increase their performances further for the vehicle object detection across the subjects.

Our second analysis of the ROC curves was performed by calculating the area under the curves and trying to estimate the improvement in sensitivities for different object and ROIs. The areas under four different ROC curves for each ROI of each subjects were calculated, and their averages were found. Acquired mean values from various subjects were aggregated and bootstrapped over ten thousand times. Human and vehicle decoding improvements and their standard variations were calculated and plotted.

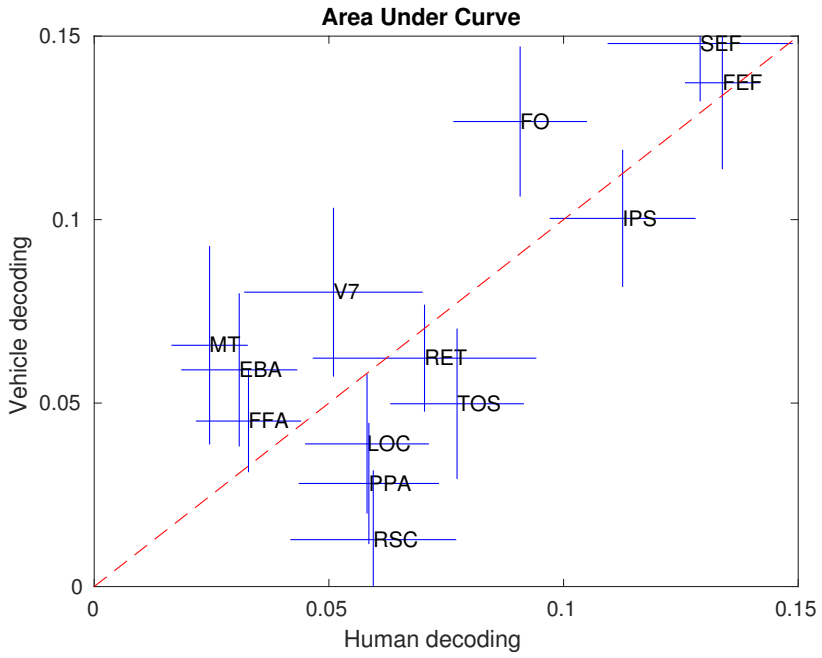

Figure 5.4: Area under the curve graph in group analysis, horizontal axis represents the decoding improvement in 'Human' object with attention, vertical axis represents the increase in 'Vehicle' object decoding with attention. Scene selective areas enhance their areas under the ROC curves to a greater extent on visual search of 'Human' objects. Human selective regions tends to show the opposite relation.

In figure 5.4, the human and vehicle decoding sensitivity increases were plotted on the horizontal and vertical axes respectively. These sensitivity increases are calculated with the area under the previously found ROC curves. The red line on the graph was plotted on the diagonal line meaning human, and vehicle detection is the same. Anything above the red line stands for better vehicle detection. Results can be interpreted as follows:

- Inherently face-selective ROIs such as FFA, MT, and EBA upsurges their vehicle detection performances when the attention is applied while viewing the natural movie.
- PPA, RSC, and TOS are known to be more responsive to the spatial domain

and scene processing. These ROIs benefit more from attention in human detection cases.

- Attention specific ROIs like SEF, IPS, FEF and FO increases their performances on both object categories. These results are consistent with our previous findings.

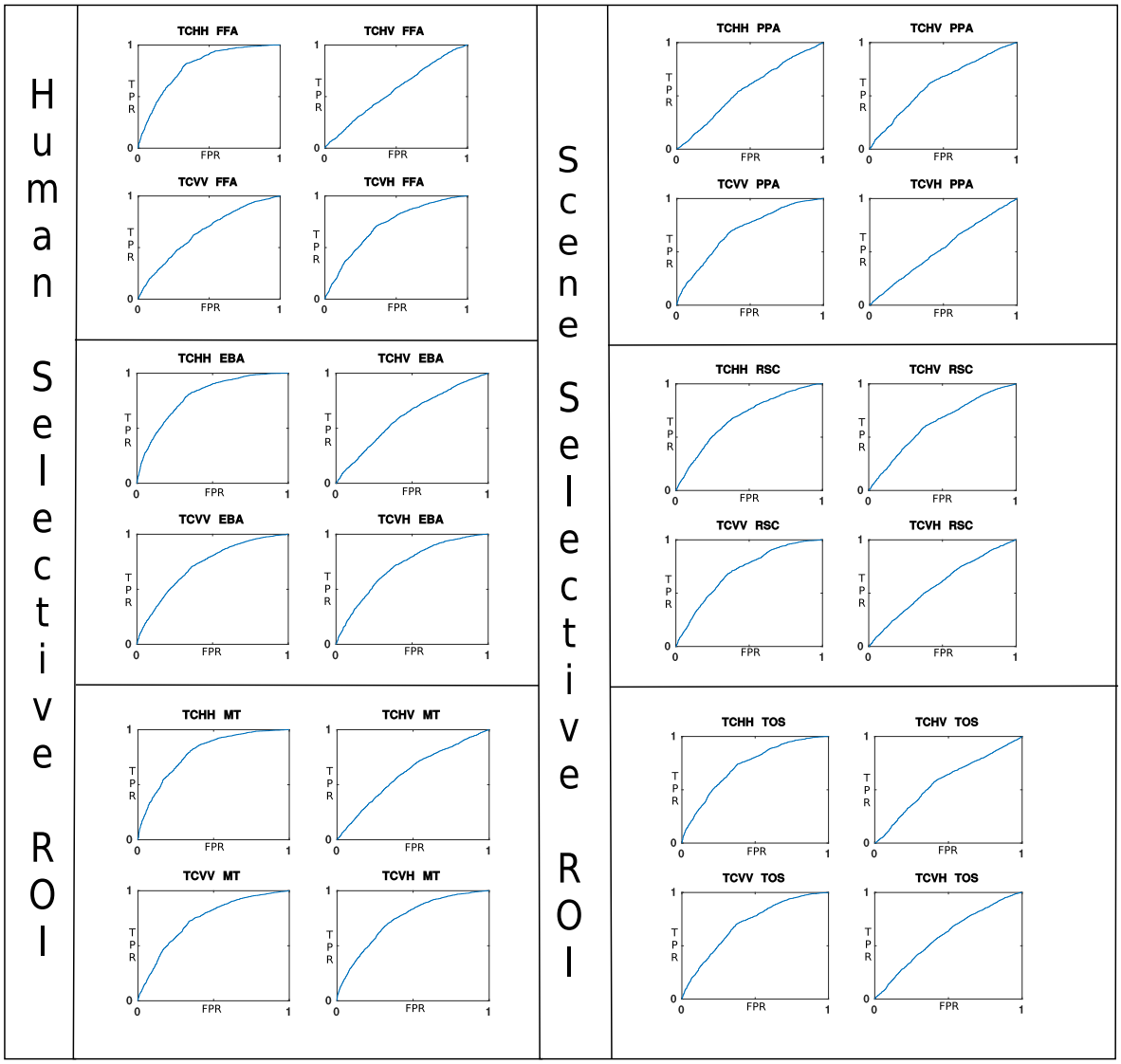

Figure 5.5: ROC curve for subject 'TC' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate.

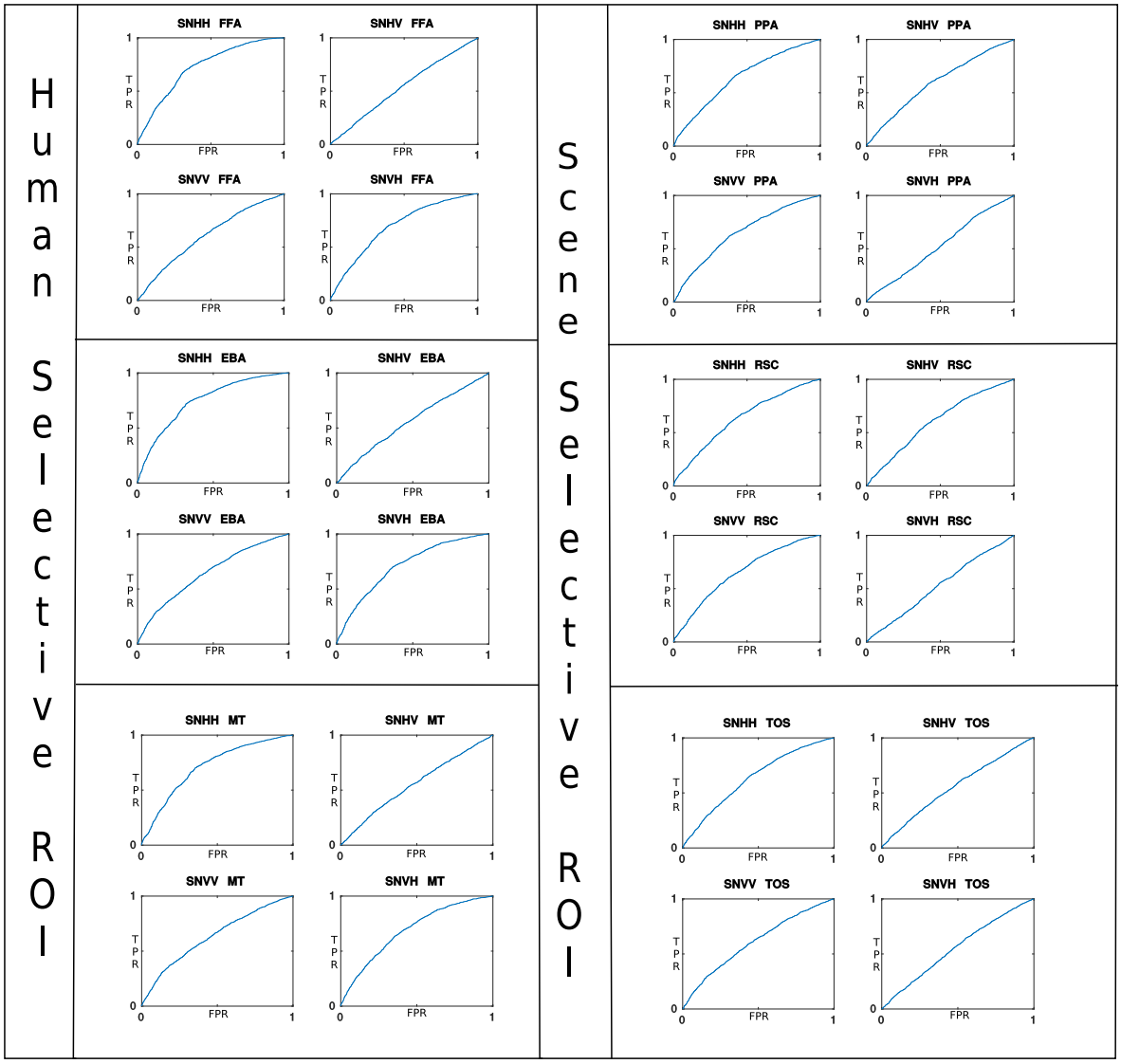

Figure 5.6: ROC curve for subject 'SN' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate.

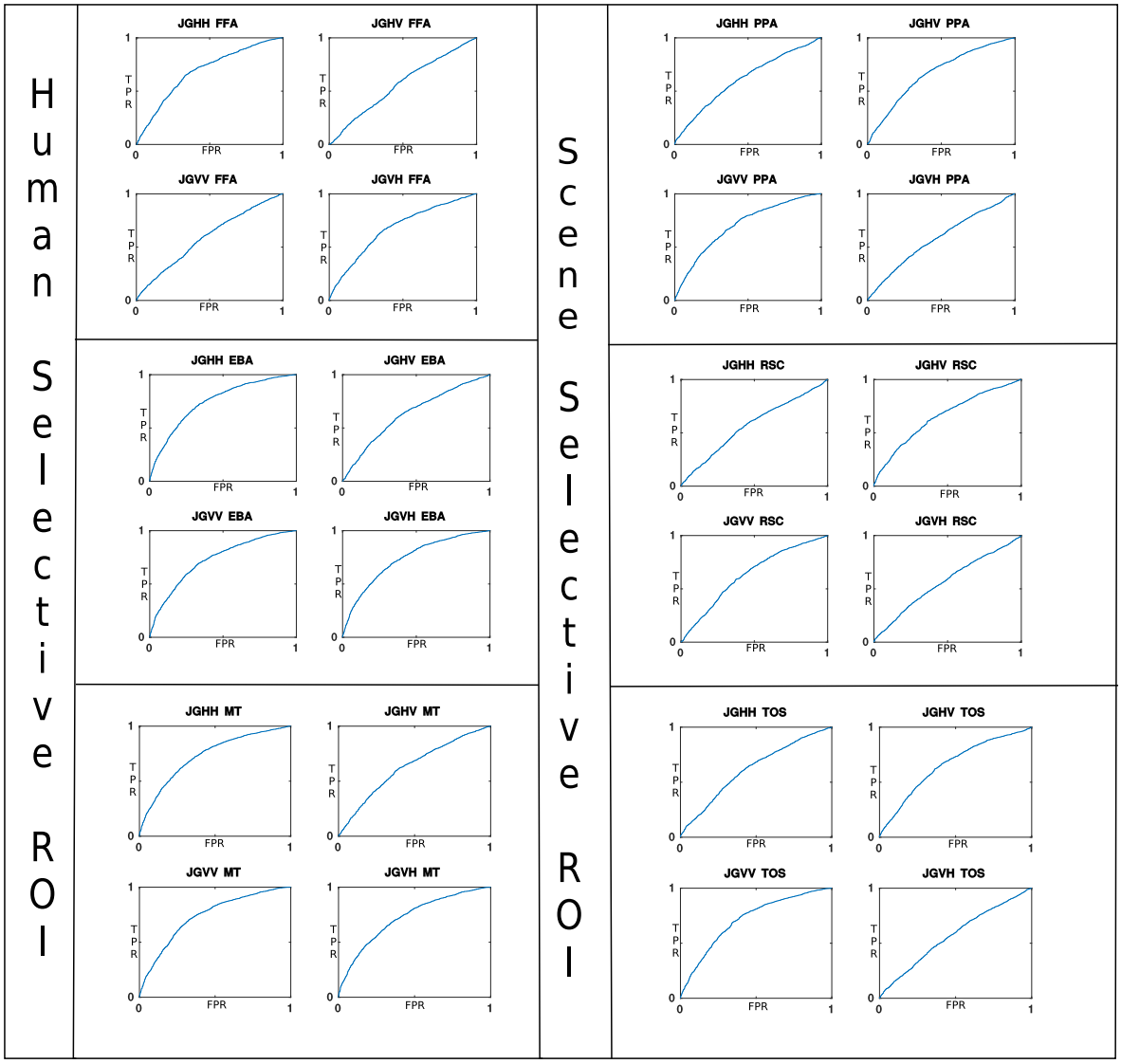

Figure 5.7: ROC curve for subject 'JG' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate.

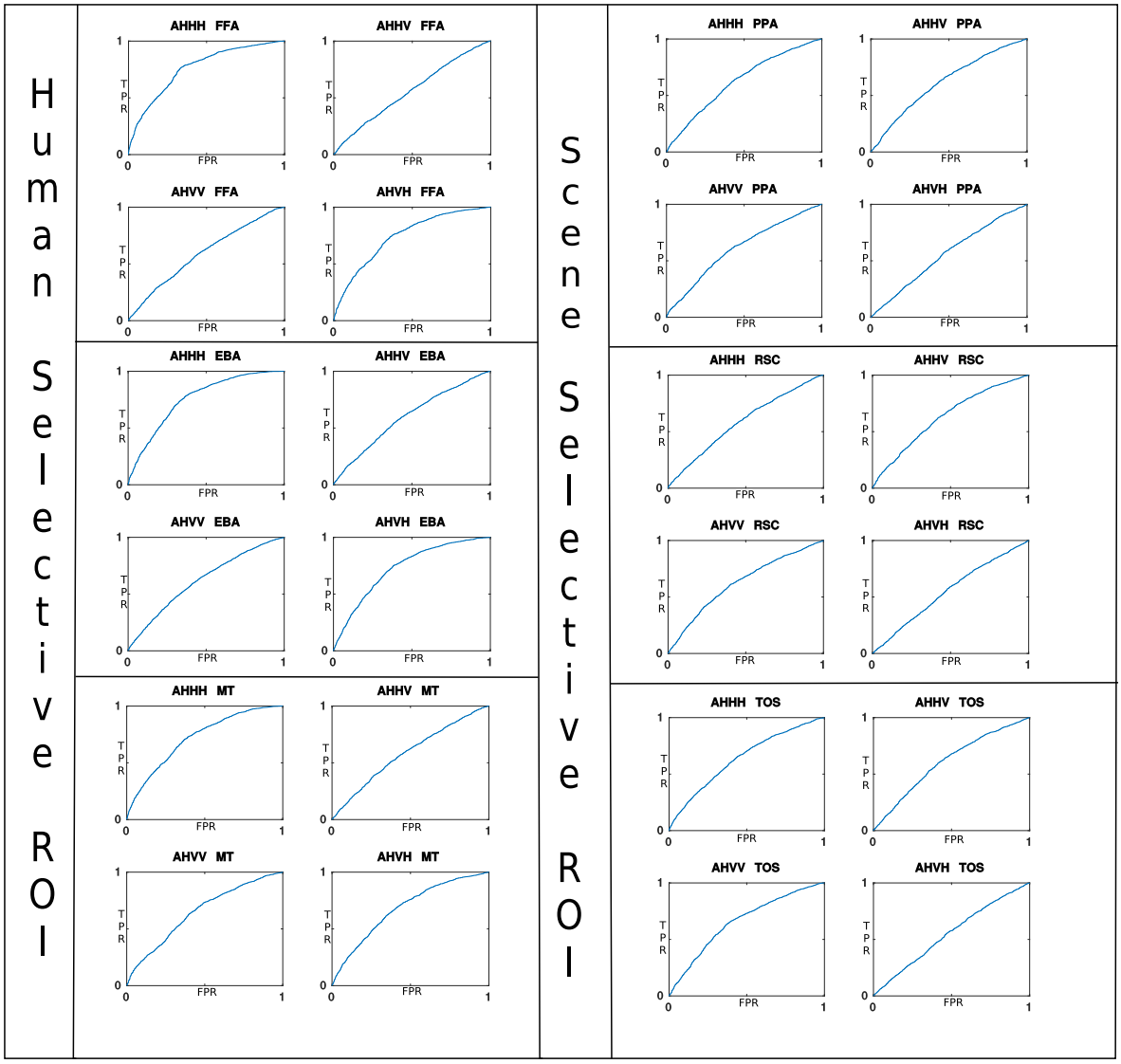

Figure 5.8: ROC curve for subject 'AH' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate.

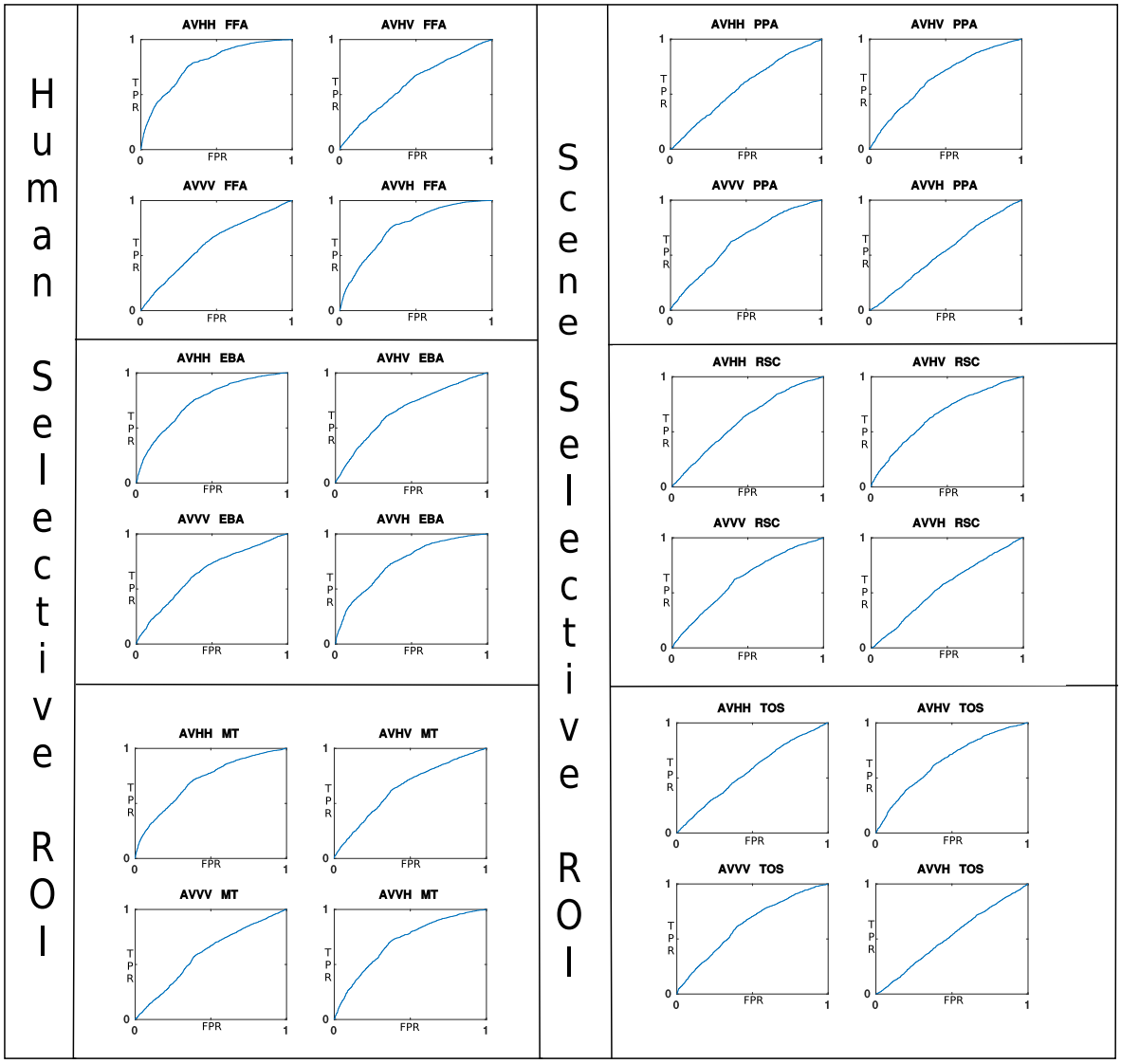

Figure 5.9: ROC curve for subject 'AV' with selected ROIs. FFA, EBA and MT selected as human selective ROIs and PPA, RSC, TOS selected as object selective ROIs. Horizontal axis represents the false positive rate and vertical axis represents the true positive rate.

In our third analysis, we created small cubes around the voxels on the cortical cortex to understand the attentional effect on the whole brain map. Each cube consisted of 27 voxels to use the advantages of the multivariate analysis. We conducted our analysis on each cube separately. Obtained results were attained to the central voxel. Response and stimuli data were separated into training,

test and evaluation subgroups. The training set consisted of nearly 80 percent of response and stimuli data, and was used to obtain model parameters. Logistic regression algorithm was used again in this analysis to acquire parameters with different regularization parameter. Those model parameters were evaluated on the test set. The best model was selected to obtain the final value from the evaluation set. The d-prime metric was calculated for the assessment of the model parameters. Four different d-prime metrics were calculated. Two of these were for the decoding human when the subject was attending to the human object and the vehicle object. Others were computed for the vehicle decoding case when subjects were attending human and vehicle objects.

3-D human cortical cortex is transformed in 2-D in below graphs. Each subject's brain was specially scanned and transformed. Regions of the brain were extracted from functional localization scans. Those regions were represented in the flat map representation. In below graphs, the d-prime value of human and vehicle decoding were subtracted from each other and shown on the maps. Blue regions in the figures show that improvement on vehicle decoding is stronger than human decoding. Red areas indicate a relationship that is the exact opposite. Findings in our third analysis can be summarized as follows:

- FFAs in nearly all subjects increase their vehicle decoding performance to a larger extent with attention. Majority of the voxels of different subject responses are similar and support our findings.
- Vehicle sensitivity of region EBA is enhanced more than the human sensitivity with attentional effect. It is the most robust region across all subjects. This supports our previous findings.
- The decoding performances for the vehicle is increased in area MT with attentional effect except some discrepancies on one subject. Although region MT is inherently selective for human detection and biological motion, it benefits from vehicle attention more than human attention.
- Results from part PPA indicate sensitivity increases more on the human object. Responses of two subjects have some deviation within the region,

but the majority of the voxels from these subjects enhance their decoding performances to a greater extent in the human case. Findings are consistent and tuning of the detectors shift toward human object further.

- Attention effects region RSC in a similar manner like the other scene sensitive regions. The sensitivity of the voxels within this area increases more for the human when the attention is taken into account. One of the subject shows some discrepancy and several of its voxels reveal a contrasting relationship with attention but the other subjects gain more from human attention.

- In region TOS, results are relatively poor with respect to other ROIs. The results from one of the subjects are consistent with our previous findings. Results of the other subjects are changeable within the region and can not be evaluated clearly.

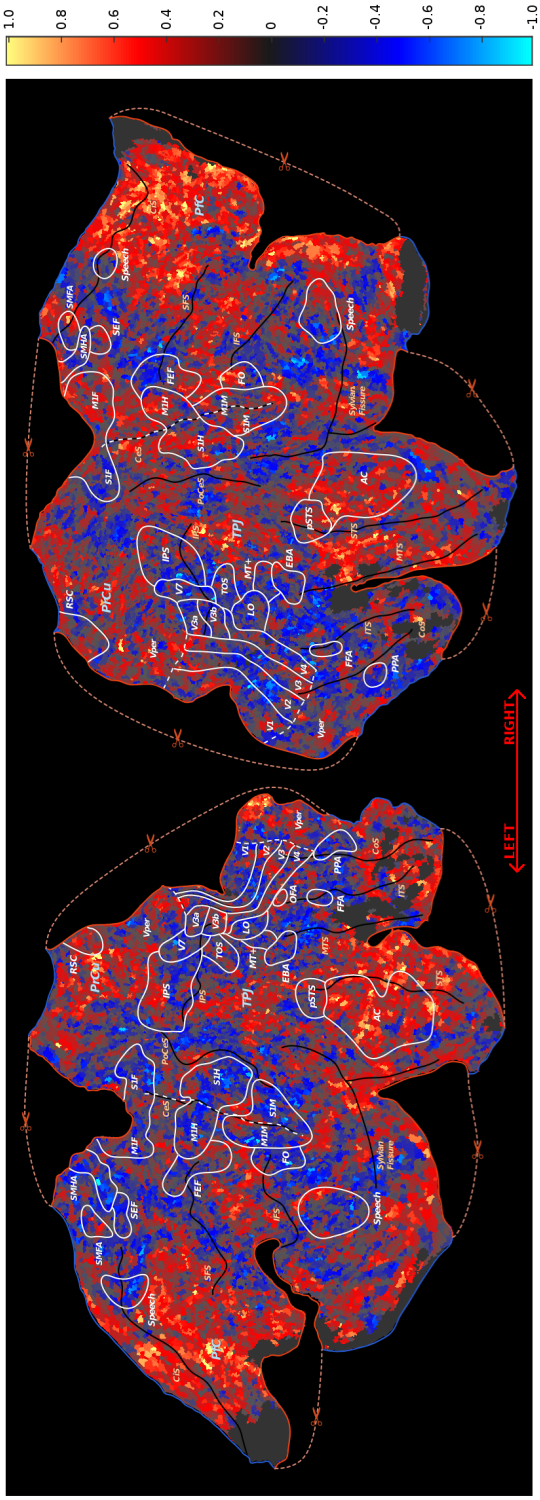

Figure 5.10: Flatmap representation of subject 'TC', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention.

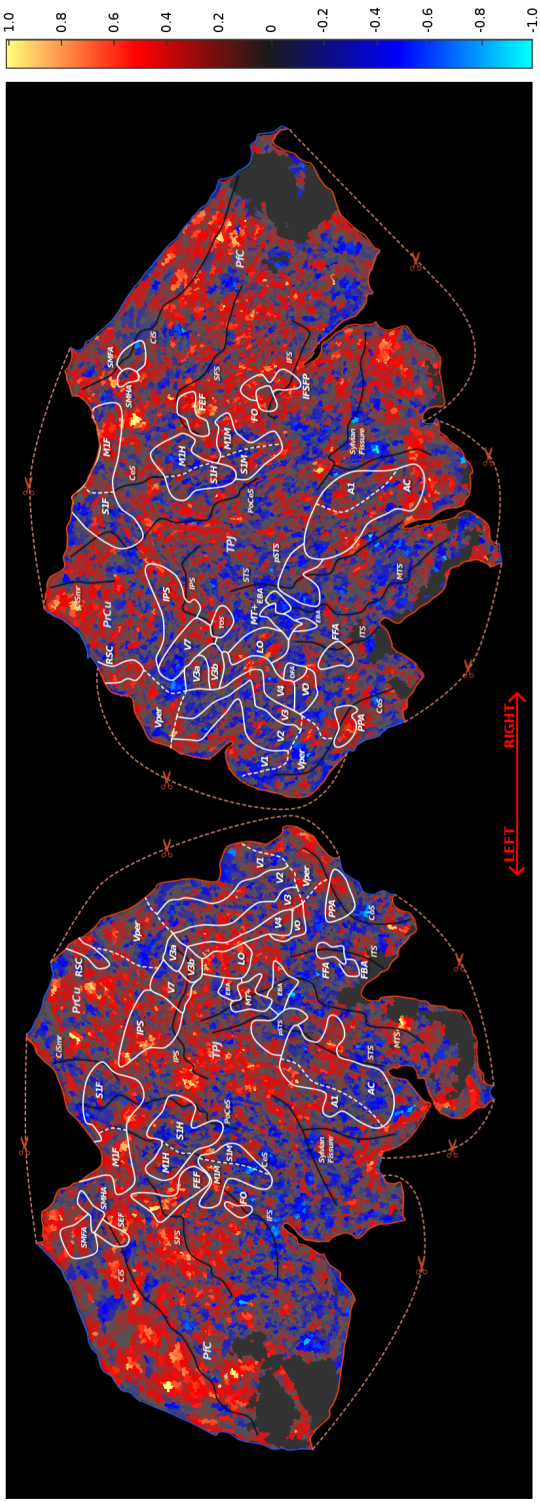

Figure 5.11: Flatmap representation of subject 'SN', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention.

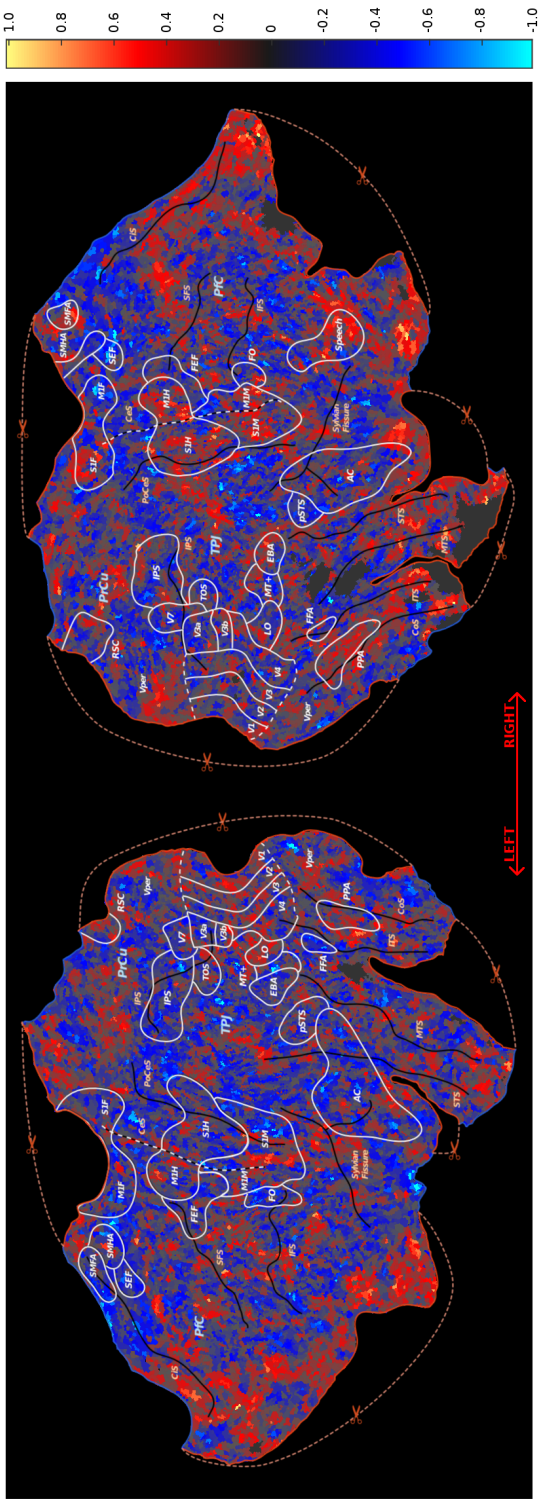

Figure 5.12: Flatmap representation of subject 'JG', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention.

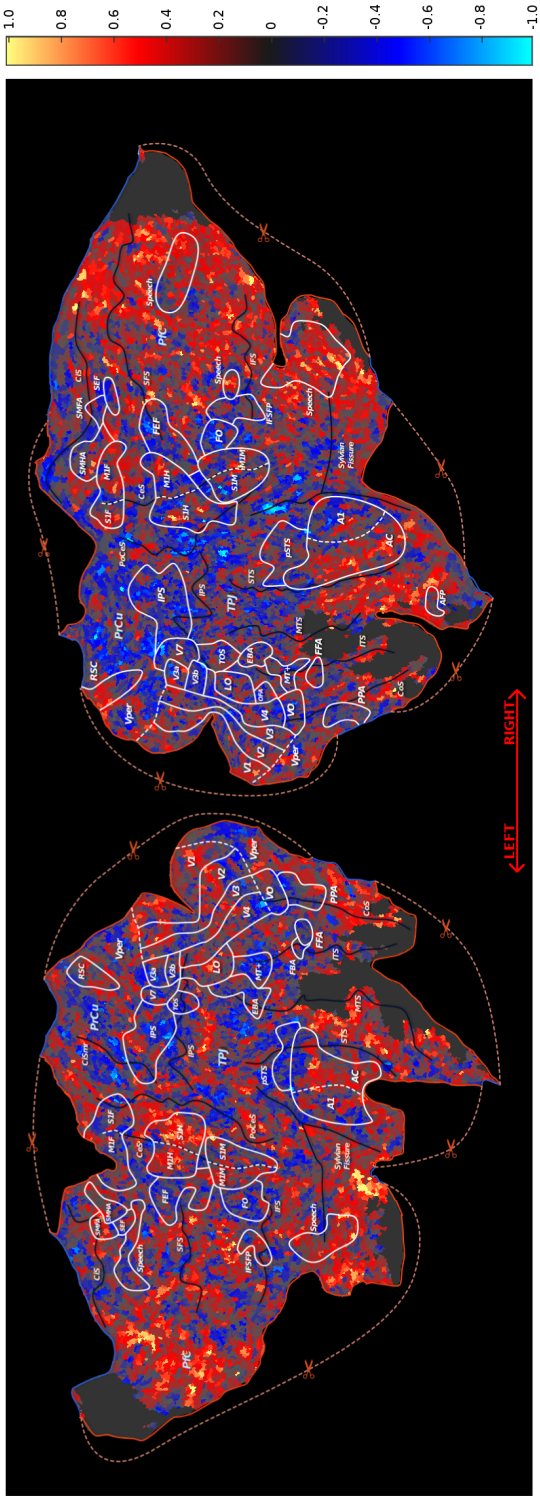

Figure 5.13: Flatmap representation of subject 'AH', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention.

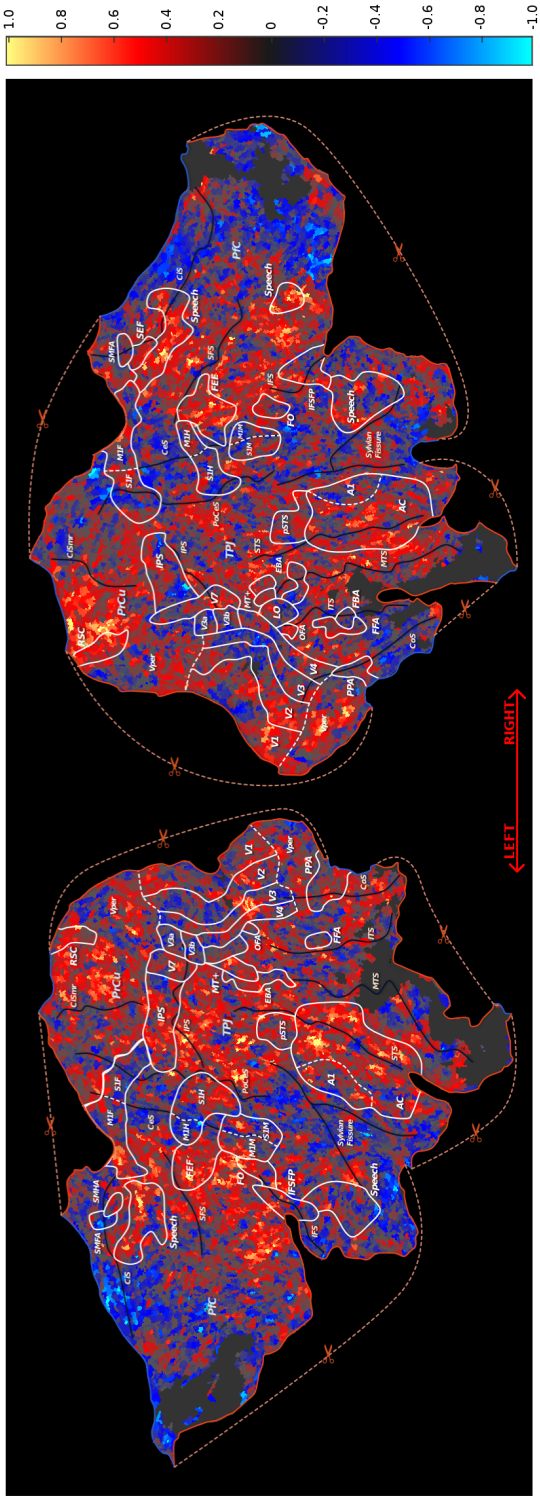

Figure 5.14: Flatmap representation of subject 'AV', red areas represent increase in human detection performance with attention, blue areas represent sensitivity enhancement in vehicle detection with attention.

5.1 Conclusion

We modeled neural populations in the human brain as target detectors during category based visual search. We applied decoding procedure on the neural responses and natural movie stimuli using logistic regression algorithm. We calculated several performance metrics for the decoding procedure. We found that category-selective areas in ventral-temporal show greater attentional improvement in detection of objects, for which they are not inherently selective for. This finding also strongly suggests that the human brain dynamically alters neural tuning to enhance target representations broadly across the brain at the expense of behaviorally-irrelevant items.

Attentional changes in detection of task-irrelevant objects and actions can be examined as a future work. But this result remains open questions whether spatial and feature-based attention also operates through similar neural mechanisms that dynamically redefine neural tuning profiles.

Bibliography

[1] S. Huettel, A. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*. Freeman, 2009.

[2] T. Wickens, *Elementary Signal Detection Theory*. Oxford University Press, 2001.

[3] S. F. Choice, "Signal detection theory," 1997.

[4] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.

[5] Z.-P. . Liang and P. C. Lauterbur, *Principles of magnetic resonance imaging: a signal processing perspective*. The Institute of Electrical and Electronics Engineers Press, 2000.

[6] D. Nishimura, *Principles of magnetic resonance imaging*. Stanford University, 1996.

[7] E. Kandel and J. Schwartz, *Principles of Neural Science, Fifth Edition*. Principles of Neural Science, McGraw-Hill Education, 2013.

[8] S. Palmer and B. C. S. P. University of California, *The Psychology of Perceptual Organization: A Transformational Approach*. Berkeley cognitive science report, Berkeley Cognitive Science Program, Institute of Cognitive Studies, University of California, Berkeley, 1983.

[9] B. Olshausen, C. Anderson, and D. Van Essen, "A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing

of information," *The Journal of Neuroscience*, vol. 13, no. 11, pp. 4700–4719, 1993.

[10] M. V. Peelen, L. Fei-Fei, and S. Kastner, "Neural mechanisms of rapid natural scene categorization in human visual cortex," *Nature*, vol. 460, no. 7251, pp. 94–97, 2009.

[11] K. M. O'Craven, P. E. Downing, and N. Kanwisher, "fmri evidence for objects as the units of attentional selection," *Nature*, vol. 401, no. 6753, pp. 584–587, 1999.

[12] L. Reddy and N. Kanwisher, "Category selectivity in the ventral visual pathway confers robustness to clutter and diverted attention," *Current Biology*, vol. 17, no. 23, pp. 2067 – 2072, 2007.

[13] M. A. Lindquist, "The statistical analysis of fmri data," *Statist. Sci.*, vol. 23, pp. 439–464, 11 2008.

[14] R. A. Poldrack, J. Mumford, and T. Nichols, *Handbook of functional MRI data analysis*. Cambridge: Cambridge University Press, 2011.

[15] R. Carter, *The Human Brain Book*. DK Publishing, 2009.

[16] E. DeYoe and D. V. Essen, "Concurrent processing streams in monkey visual cortex," *Trends in Neurosciences*, vol. 11, no. 5, pp. 219 – 226, 1988.

[17] M. Mishkin, L. G. Ungerleider, and K. A. Macko, "Object vision and spatial vision: two cortical pathways," *Trends in Neurosciences*, vol. 6, pp. 414 – 417, 1983.

[18] N. Kanwisher, J. McDermott, and M. M. Chun, "The fusiform face area: A module in human extrastriate cortex specialized for face perception," *The Journal of Neuroscience*, vol. 17, no. 11, pp. 4302–4311, 1997.

[19] P. E. Downing, Y. Jiang, M. Shuman, and N. Kanwisher, "A cortical area selective for visual processing of the human body," *Science*, vol. 293, no. 5539, pp. 2470–2473, 2001.

[20] K. A. Pelphrey, T. V. Mitchell, M. J. McKeown, J. Goldstein, T. Allison, and G. McCarthy, "Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion," *The Journal of Neuroscience*, vol. 23, no. 17, pp. 6819–6825, 2003.

[21] R. Epstein, A. Harris, D. Stanley, and N. Kanwisher, "The parahippocampal place area: Recognition, navigation, or encoding?," *Neuron*, vol. 23, no. 1, pp. 115 – 125, 1999.

[22] S. D. Vann, J. P. Aggleton, and E. A. Maguire, "What does the retrosplenial cortex do?," *Nature Reviews Neuroscience*, vol. 10, no. 11, pp. 792–802, 2009.

[23] S. D. Auger and E. A. Maguire, "Assessing the mechanism of response in the retrosplenial cortex of good and poor navigators," *Cortex*, vol. 49, no. 10, pp. 2904 – 2913, 2013.

[24] U. Hasson, M. Harel, I. Levy, and R. Malach, "Large-scale mirror-symmetry organization of human occipito-temporal object areas," *Neuron*, vol. 37, no. 6, pp. 1027 – 1041, 2003.

[25] R. Tootell, M. Silverman, E. Switkes, and R. De Valois, "Deoxyglucose analysis of retinotopic organization in primate striate cortex," *Science*, vol. 218, no. 4575, pp. 902–904, 1982.

[26] K. Grill-Spector, Z. Kourtzi, and N. Kanwisher, "The lateral occipital complex and its role in object recognition," *Vision Research*, vol. 41, no. 1011, pp. 1409 – 1422, 2001.

[27] R. B. Tootell, N. Hadjikhani, E. Hall, S. Marrett, W. Vanduffel, J. Vaughan, and A. M. Dale, "The retinotopy of visual spatial attention," *Neuron*, vol. 21, no. 6, pp. 1409 – 1422, 1998.

[28] E. Wojciulik and N. Kanwisher, "The generality of parietal involvement in visual attention," *Neuron*, vol. 23, no. 4, pp. 747 – 764, 1999.

[29] J. D. Schall, "On the role of frontal eye field in guiding attention and saccades," *Vision Research*, vol. 44, no. 12, pp. 1453 – 1467, 2004. Visual Attention.

[30] B. A. Purcell, P. K. Weigand, and J. D. Schall, "Supplementary eye field during visual search: Salience, cognitive control, and performance monitoring," *The Journal of Neuroscience*, vol. 32, no. 30, pp. 10273–10285, 2012.

[31] T. Higo, R. B. Mars, E. D. Boorman, E. R. Buch, and M. F. Rushworth, "Distributed and causal influence of frontal operculum in task control," *Proceedings of the National Academy of Sciences*, vol. 108, no. 10, pp. 4230–4235, 2011.

[32] A. Huth, S. Nishimoto, A. Vu, and J. Gallant, "A continuous semantic space describes the representation of thousands of object and action categories across the human brain," *Neuron*, vol. 76, no. 6, pp. 1210 – 1224, 2012.

[33] T. Çukur, S. Nishimoto, A. G. Huth, and J. L. Gallant, "Attention during natural vision warps semantic representation across the human brain," *Nature neuroscience*, vol. 16, no. 6, pp. 763–770, 2013.

[34] G. Miller, "Wordnet: A lexical database for english," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[35] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby, "Beyond mind-reading: multi-voxel pattern analysis of fmri data," *Trends in cognitive sciences*, vol. 10, no. 9, pp. 424–430, 2006.

[36] T. Naselaris, K. N. Kay, S. Nishimoto, and J. L. Gallant, "Encoding and decoding in fmri," *NeuroImage*, vol. 56, no. 2, pp. 400 – 410, 2011. Multivariate Decoding and Brain Reading.

[37] J. V. Haxby, A. C. Connolly, and J. S. Guntupalli, "Decoding neural representational spaces using multivariate pattern analysis," *Annual Review of Neuroscience*, vol. 37, no. 1, pp. 435–456, 2014. PMID: 25002277.

[38] K. N. and B. P., "Analyzing for information, not activation, to exploit high-resolution fmri.," *Neuroimage*, vol. 38, no. 4, pp. 649–662, 2007.

[39] R. Leila, N. G. Kanwisher, and R. VanRullen, "Attention and biased competition in multi-voxel object representations," *Proceedings of the National*

Academy of Sciences of the United States of America, vol. 106, no. 50, pp. 21447–21452, 2009.

[40] N. Kriegeskorte, R. Goebel, and P. Bandettini, "Information-based functional brain mapping," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 10, pp. 3863–3868, 2006.

[41] D. G. E. Barry H. Kantowitz, III Henry L. Roediger, *Experimental Psychology*. Wadsworth Publishing, 9th ed., 2008.

[42] T. Fawcett, "An introduction to roc analysis," *Pattern Recogn. Lett.*, vol. 27, pp. 861–874, June 2006.

[43] T. P. Minka, "A comparison of numerical optimizers for logistic regression," tech. rep., 2003.

[44] K. Nigam, J. Lafferty, and A. McCallum, "Using maximum entropy for text classification," in *IJCAI-99 workshop on machine learning for information filtering*, vol. 1, pp. 61–67, 1999.

[45] J. A. Etzel, J. M. Zacks, and T. S. Braver, "Searchlight analysis: Promise, pitfalls, and potential.," *NeuroImage*, vol. 78, pp. 261–269, 2013.

Appendix A

Appendix

A.1 MRI: Magnetic Resonance Imaging

Magnetic resonance imaging is a noninvasive imaging modality [1] that enables us to monitor human body. Various body parts including the spinal cord, brain, chest, lungs, ankles, and feet can be visualized, and problems can be detected without any physical invasion of the human body. MR (magnetic resonance) also enables us to obtain metabolic information and gives rise to invaluable possibilities in the physiological studies. It has some advantages over other imaging methodologies such as CT (computed tomography) or PET (positron emission tomography). The most significant benefit of MR over other imaging techniques is that MR does not use ionizing radiation or radioactive traces and thus it does not induce any known health risks on the human body. Although MR requires sophisticated and expensive hardware for reliable images, it has impressively better soft tissue contrast and spatial resolution than other imaging techniques. The goal of this chapter is to explain the fundamental physical and mathematical principles that underly MR.

A.2 MRI Physics

The building block of MR relies on the physical principles of the atomic nuclei that consist of protons, neutrons, and electrons. The proton and the neutron are located in the center of the nucleus. The proton is considered to have a positive charge while the neutron is neutral. An electron that has negative charge exists in the spherical shells around nuclei. Under normal conditions, protons spin around their axis because of thermal energy and this spinning results in spin angular momentum which is expressed as:

$$S = \hbar I \tag{A.1}$$

where $\hbar$ is Plank's constant and I is the spin operator.

Angular momentum creates a current loop on the surface of the nucleus, thereby forming a magnetic dipole moment on the atomic nucleus when it is located inside a magnetic field.

$$\mu = \gamma S = \gamma \hbar I \tag{A.2}$$

where γ is the gyromagnetic ratio that is unique to each different atomic nucleus.

The magnetic dipole moment and the angular momentum of particles are the foundations of the NMR (nuclear magnetic resonance). The Magnetic dipole moment of particles that have an even number of protons and neutrons is nearly zero. Therefore, they cannot be used in the MR studies. A nucleus that has an odd number of protons might be utilized in MR applications such as H and C. Static magnetic field (B_0) and radio frequency field (RF) are used to handle the NMR phenomenon in an MR study [6].

Under normal conditions, particles spin randomly and their magnetic dipole moments might cancel each other. To remove this cancellation effect, atoms are exposed to a static magnetic field (B_0), and they align with the direction of the magnetic field. Thereby magnetic dipole moments can't eliminate each other and elicit net magnetization moment (M) of the whole volume in the B_0 direction. By convention, the z-axis is called the longitudinal axis, and x-y axis is

called transverse axis in a typical MR study [5]. The net magnetization vector is fundamental to the NMR, but we cannot quantify it under the equilibrium state. RF signal is applied to the particles to rotate away from their equilibrium state to the transverse plane.

When the directions of the B_0 and M are different, precession occurs on the particle, which results in a net torque on the particle. Torque also equals the derivative of the angular momentum (S). Equations are as follows:

$$\tau = \mu \times B_0 \tag{A.3}$$

$$\frac{dS}{dt} = \mu \times B_0 \tag{A.4}$$

Above equation also results in the following when we multiply both sides with γ :

$$\frac{d\mu}{dt} = \mu \times \gamma B_0 \tag{A.5}$$

If we try to evaluate above equation on the whole volume, we obtain the net magnetization vector (M).

$$\frac{dM}{dt} = M \times \gamma B_0 \tag{A.6}$$

Net magnetization vector (M) is the sum of the magnetization of the whole particles within a material. The magnetization vector continues its rotation around the z-axis. The frequency of this rotation unique to different atoms and is known as the larmour frequency. The solution for the above equation points the same results that M vector precesses with the larmour frequency [6].

When the RF field is applied, the net magnetization vector is tipped by some angle from its equilibrium state. This procedure is known as excitation. After the excitation, the magnetization vector that has been tipped, continues its rotation while it tries to align with the B_0 field again. This precession creates a small electro-magnetic force (EMF), and EMF can be detected by the gradient coils tuned to the larmour frequency. The amplitude of the EMF depends on the excitation angle. Supposing the particles are completely tipped on the transverse plane (i.e., 90° tip from the longitudinal plane), it generates the maximum EMF on the gradient coils [1].

After the excitation, transverse magnetization vector (M_{xy}) starts to decay. Meanwhile, the longitudinal vector (M_z) eventually starts to align with the B_0. The realignment of the M vector is called relaxation. Spin relaxation does not continue for a long time and usually lasts a few seconds. T_1 and T_2 time constants are associated with the M_z and M_{xy} vectors during the realignment of the net magnetization vector.

The relaxation of the longitudinal component behaves according to the following equation:

$$\frac{dM_z}{dt} = -\frac{M_z - M_0}{T_1} \tag{A.7}$$

Solution for the above differential equation is as follows:

$$M_z = M_0 + (M_z(0) - M_0)\exp^{-t/T_1} \tag{A.8}$$

where

M_z longitudinal magnetization vector

M_0 is the net magnetization vector before excitation

$M_z(0)$ the longitudinal magnetization right after excitation

T_1 spin-lattice time constant, expressing a return to the equilibrium state.

T_1 is the called spin-lattice time constant and expresses a revival of the magnetization vector on the z-axis. It can easily be seen in the equations that M_z reaches M_0 when time proceeds to infinity. T_1 value is dependent on the static magnetic field and increases as the magnetic field increases. For a $90\,^\circ$ excitation pulse, above equation can be rearranged as follows since $M_z(0) = 0$

$$M_z = M_0(1 - \exp^{-t/T_1}) \tag{A.9}$$

the relaxation on the transverse plane can be expressed with the following equation

$$\frac{dM_{xy}}{dt} = -\frac{M_{xy}}{T_2} \tag{A.10}$$

For a $90\,^\circ (M_{xy} = M_0)$ excitation, the solution is as follows:

$$M_{xy} = M_0 \exp^{-t/T_2} \tag{A.11}$$

where

M_z is the transverse magnetization vector

M_0 is the net magnetization vector before excitation

T_2 is called the spin-spin time constant and expresses the decrease in the transverse magnetization vector. Spins in the material are phase coherent right after the excitation. As the time passes the coherence between the spins disappear, and they gradually become out of phases. This phenomenon is called transverse relaxation. The transverse relaxation is caused by spin-spin interactions and static magnetic field inhomogeneity. We can see in the equation that as time passes, the transverse magnetization gradually becomes insignificant.

T_2 decay is faster than the T_1 decay and consequently, time is limited when collecting MRI data on the excited particles. In order to obtain a high resolution and fine-grained anatomical data, many excitation pulses are used in a typical MRI study.

Equation A.7 and A.11 are called the Bloch equation, are combined to describe the NMR. In the Bloch equation, the excitation and relaxation of the magnetization vector, the MR signal formed on the receiver coils and the T_1, T_2 effects are described in a single equation.

$$\frac{dM}{dt} = \gamma M \times B_0 + \frac{1}{T_1}(M_0 - M_z) - \frac{1}{T_2}(M_x + M_y) \tag{A.12}$$

where

γ = gyro-magnetic ratio

T_1 = transverse relaxation time constant

T_1 = longitudinal relaxation time constant

M_0 = initial magnetization before excitation stage

A.3 Image Formation

Particles precess with different frequencies as they are exposed to a distinct magnetic fields. The improvement that makes MR imaging possible is the spatially variable magnetic field that are superimposed on one another. The basis for the image formation in the MR studies lies on that fact and first introduced by the Lauterbur in 1972 [5]. As Lauterbur demonstrated, the MR signal in 3-D volume is divided into subcomponents. Each of them is exposed to a magnetic field that has modulated strength in the longitudinal direction and thereby precesses at different frequencies. Images can be acquired with a fine tuned receiver hardware that is adjusted to the larmour frequency.

Direct three-dimensional imaging is computationally expensive and slower to acquire images of an object. Also, 3-D imaging is not applicable to the functional imaging of the brain [1]. The common first step is to reduce the dimension of the object to the 2-D space. By changing the slices with slice selection gradients, we might form images of a 3-D volume.

A systematic way to acquire MR image formation might be divided into three routine procedures, since we cannot expect to receive images with modulating the magnetic field randomly. First, a slice of a 3-D object is selected with the slice selection gradient field along with the RF pulse. The slice selection gradient field usually changes along with the longitudinal direction, is applied to the object in order to ensure that only the slice of interest is precessing with the same frequency of RF pulse. So, only the particles within the selected slice absorb the RF signal and are tipped on the transverse plane. After the RF pulses, particles within the selected slice contribute to the MR signal. When the designated part of the object is selected, phase and frequency encoding gradients are used to resolve a 2-D image.

In the real world, we cannot create discrete bands of frequencies, and therefore cannot form discrete tipped angles on the magnetization vector as seen on the left of Figure: A.1. Instead, we form a continuous frequency band on the slice

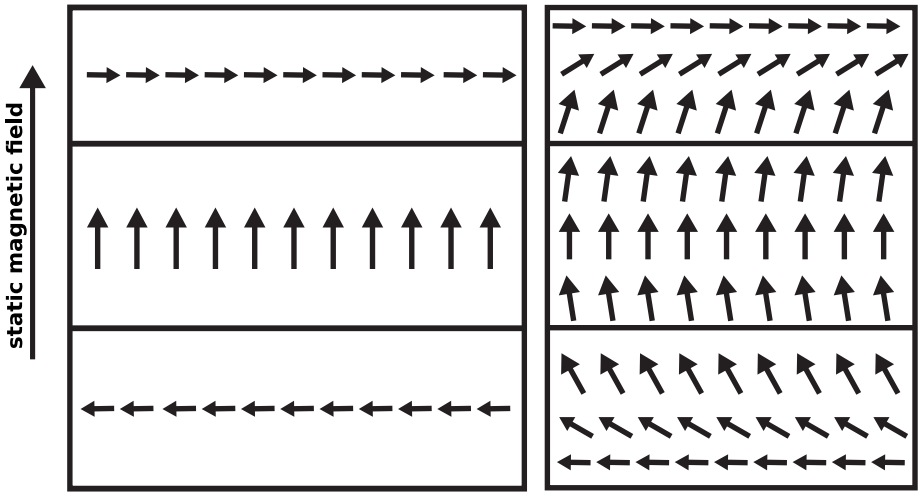

Figure A.1: Slice selection: Hydrogen atoms are tipped with slice selection gradient. Hydrogen atoms within the selected slice are tipped right after the RF pulse. In the left pane, the magnetization vector points toward a discrete set of angles. In the real world, however, particles are exposed to a slightly different magnetic field and therefore form continuous bands of tipped angles. This figure is adapted from [1].

(illustrated on the right Fig: A.1). The excitation requires that the RF pulse is coherent with the larmour frequency range. The remedy for the selection problem is that if we know the static magnetic and the gradient field and the slice location correctly, we might form an RF pulse with desired frequencies. Furthermore, we might alter the thickness of the slice with the bandwidth of the RF pulses.

Instantly after the excitation pulse, magnetization vectors on the transverse plane experience T_1 and T_2 decay. Both of these result in a loss of the MR signal. Because of the instant decay in the MR signal, following the slice selection, another gradient field needs to be applied instantly to the object. Other gradients provide a change in frequency for the particles and thereby information on the 2-D slice.

When we excite a slice, particles in that slices show precession with a known frequency. However, we cannot obtain a fine-spatial resolution from this because the whole slice contributes to the MR signal. Lauterbur tried to solve this resolution problem and constructed an experiment for it [1]. In this experiment, he

used two physically separated vials on which he applied variable magnetic fields. The magnetic fields applied on the vials were different. Since the particles within the vials were exposed to a different magnetic field, they precessed with distinct frequencies. This procedure was later called frequency encoding. MR signals constructed after frequency encoded gradients also have transverse relaxation but the only difference is that slower oscillations (caused by the frequency encoded gradient) are superimposed on the higher oscillations (caused by the static magnetic field). With the help of signal processing algorithms, variable oscillations might provide us with information about the positions and fill rate of the vials. In summary, physical differences of the vials are encoded on the frequency of the collected MR data but only in one dimension.

Frequency encoding gradients provide information only in one dimension. Creating a fine-grained two-dimensional image requires a second spatial gradient field, applied on the other axis, which is known as phase encoding. The fundamental goal of using spatially different gradient fields is to collect a distribution of the particles on the 2-D plane. For example, if we apply a frequency encoding gradient on the x-axis and use two different amplitude phase encoding gradients on the y-axis, we obtain two different MR signals that change vertically due to the variable gradient field on the y-axis. The distribution of the spins across the 2-D plane is acquired by changing the phase encoding gradient, and this distribution might be used to estimate the characteristic of the object. The different amplitude phase encoding gradient is constructed through iterated excitation of the sample.

In figure A.2, After the RF and slice selection gradient, all particles on the slice are precessing with the same frequency, and they are all in-phase within the slice. The phase encoding gradient field is turned on, and the phase of the spins linearly increases as we move along in upper y direction. After switching off the phase encoding gradient, frequency encoding gradient is turned on. This magnetic field modulates the precession frequency of the particles on the x-axis. As we keep exciting the same slice with different phase encoding gradients, we create a distribution densities of the spin across x and y-directions.

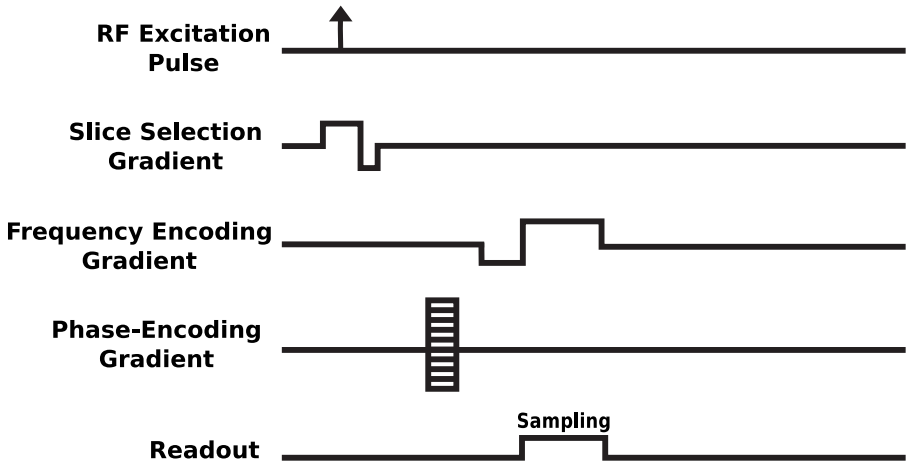

Figure A.2: A sample MR sequence. The RF and the slice selection gradient excite a 2-D plane in a 3-D volume. After the excitation phase, phase and frequency encoding gradients are applied. Frequency and phase encoding gradients constitute a sufficient k-space coverage. Sampling occurs at the same time as the frequency encoding gradient. This illustration is adapted from [5].

A.3.1 Bloch Equations

Effects of the static and gradient magnetic fields can be explained mathematically with Bloch equations. In the previous parts, we explained that the static magnetic field B_0 aligns the particles and creates a net magnetization vector along with its direction. Above we include new magnetic fields to change the magnetization vector and therefore clarify spin densities over the 3-D space. Bloch equations can be solved with the additional phase and frequency encoding gradients. At this point we can write the Bloch equations for the sake of clarity as follows [6]:

$$\frac{dM}{dt} = \gamma M \times B + \frac{1}{T_1}(M_0 - M_z) - \frac{1}{T_2}(M_x + M_y) \tag{A.13}$$

where B contains both static and gradient field vectors.
We can arrange above equations as follows:

$$\frac{dM_x}{dt} = M_y.\gamma B - \frac{1}{T_2}(M_x) \tag{A.14}$$

$$\frac{dM_y}{dt} = M_x.\gamma B - \frac{1}{T_2}(M_y) \tag{A.15}$$

$$\frac{dM_z}{dt} = -\frac{M_z - M_0}{T_1} \tag{A.16}$$

Above equations clarify the net magnetization vector over time. As explained in the previous parts, transverse magnetization undergoes to decay with T_2 time constant, and longitudinal magnetization recovers with the T_1 time constant.

Solutions for the longitudinal magnetization (A.16) and transverse magnetization (A.14-A.15) are as follows:

$$M_z = M_0 + (M_z(0) - M_0)\exp^{-t/T_1} \tag{A.17}$$

$$M_x = (M_x(0)\cos\omega t + M_y(0)\sin\omega t)\exp^{-t/T_2} \tag{A.18}$$

$$M_y = (M_y(0)\cos\omega t - M_x(0)\sin\omega t)\exp^{-t/T_2} \tag{A.19}$$

We can write the transverse magnetization vector in complex form:

$$M_{xy} = M_x + iM_y \tag{A.20}$$

$$\begin{aligned} M_{xy} &= (M_x(0) + iM_y(0))\exp^{-t/T_2}(\cos\omega t - i\sin\omega t) \\ &= M_{xy}(0)\exp^{-t/T_2}\exp^{-i\omega t} \end{aligned} \tag{A.21}$$

Particles exposed to gradient fields that are oriented on the x, y and z axes during the excitation and relaxation period and magnetic field B over temporal and spatial locations can be formulated as follows:

$$B(t) = B_0 + G_x(t)x + G_y(t)y + G_z(t)z \tag{A.22}$$

Precession frequency ($\omega = \gamma B$) is dependent on the gyro-magnetic ratio and magnetic field. If we write the magnetic field (A.22) into the transverse magnetization (A.21) and integrate over time, we have the magnetization vector with respect to the time interval.

$$M_{xy}(x,y,z,t) = M_{xy0}(x,y,z)\exp^{-t/T_2}\exp^{-i\gamma B_0 t}\exp^{-i\gamma\int_0^t (G_x(\tau)x + G_y(\tau)y + G_z(\tau)z)d\tau} \tag{A.23}$$

In above equations, the transverse magnetization vector is explained with the initial magnetization vector M_{xy0}, the T_2 decay effect, the phase value caused by static magnetic field B_0 and the aggregated gradient effects.

In a typical MR experiment, every excited particle contributes to the signal. The MR signal can be explained in the following manner.

$$S(t) = \int_x \int_y \int_z M_{xy}(x, y, z, t) dx dy dz \tag{A.24}$$

$$S(t) = \int_x \int_y \int_z M_{xy0}(x, y, z) \exp^{-t/T_2} \exp^{-i\omega_0 t} \exp^{-i\gamma \int_0^t (G_x(\tau)x + G_y(\tau)y + G_z(\tau)z)d\tau} dx dy dz \tag{A.25}$$

Equation A.25 represents the connection between the measured signal and the magnetization vectors. The $e^{-i\omega_0 t}$ term explains the modulation and The e^{-t/T_2} term indicates the decay of the MR signal. These terms do not comprise spatial location and we may ignore them for the sake of clarity. After removing them, we come up with a more meaningful MR signal equation.

$$S(t) = \int_x \int_y \int_z M_{xy0}(x, y, z) \exp^{-i\gamma \int_0^t (G_x(\tau)x + G_y(\tau)y + G_z(\tau)z)d\tau} dx dy dz \tag{A.26}$$

However, the equation A.26 explains the collected MR signal in 3-D, while 3-D imaging is computationally expensive and susceptible to hardware defects [5]. The slice selection gradient enables us to obtain the signal in the 2-D domain and the equation A.26 changes into a simpler and more meaningful form. The slice selection gradient excites a slice of the 3-D volume and no particles outside that slice generate an MR signal.

$$S(t) = \int_x \int_y M_{xy0}(x, y) \exp^{-i\gamma \int_0^t (G_x(\tau)x + G_y(\tau)y)d\tau} dx dy \tag{A.27}$$

The selected slice is formed by the amplitude of the gradient, bandwidth of the excitation field and the center frequency of the excitation pulse. Slice thickness is defined by the bandwidth of the excitation field and slice location is determined by the amplitude of the gradient and the center frequency of the RF field. The whole object is scanned with successive changes to the slice location, achieved by sliding the center frequency over the z-axis. Magnetization vector after slice selection can be calculated by integrating whole vector over the slice width on the z axis.

$$M(x, y, t) = \int_{z_0 - \delta z}^{z_0 + \delta z} M(x, y, z, t) dz \tag{A.28}$$

Although we change the equation A.26 to the equation A.27 after the slice selection, this equation is still hard to interpret. To form a better understanding of

the MR signal, MR researchers constitute a k-space notation in MRI studies [6]. it shows frequency components of the magnetization vector in a new 2-D axis. The k-space notation represents the gradient fields as the time integral of each gradient field multiplied by an appropriate constant.

$$k_x = \frac{\gamma}{2\pi}\int_0^t G_x(\tau)d\tau \qquad k_y = \frac{\gamma}{2\pi}\int_0^t G_y(\tau)d\tau \tag{A.29}$$

If we put above equation on the A.27, we end up with a known equation that represents the relationship between image and k-spaces. The signal we acquire from the MR machine is basically a 2-D Fourier transform of the magnetization vector. Inverse Fourier transform converts the acquired signal to a spatially meaningful image.

$$S(t) = \int_x \int_y M_{xy0}(x,y)\exp^{-i2\pi k_x(t)x - i2\pi k_y(t)y}\,dxdy \tag{A.30}$$

To construct an informative image from the signals, we used gradient fields to cover the whole k-space. The gradient fields are used to control the sampling path in the k-space. By commuting the strength of the gradient, we might regularly sample whole k-space. The phase encoding gradient is used to select a different k_y row, and the frequency encoding gradient gradually changes the precessing frequency on the k_x line. The MR signal is sampled within sequence to construct a meaningful image.

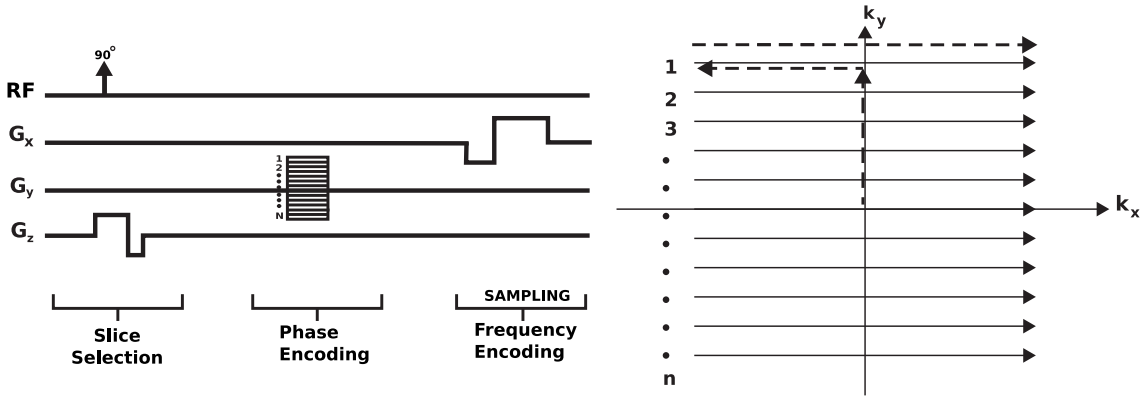

Figure A.3: An MR sequence and its k-space coverage. If we use number 1 of the phase encoding gradient, we will cover the line 1 on the k-space. By changing the G_x gradient, we shift from the middle to the left and from the left to the right of the k-space and sample as we use the G_x gradient. The figure is adapted from [6].