Monitoring of Human-Data Interactions
Towards Understanding the Interpretation Process
of Geoscientific Data

Yathunanthan Sivarajah

School of Earth and Environment

2015

This thesis is presented for the degree of
Doctor of Philosophy
of The University of Western Australia
Abstract

Qualitative data interpretation forms the basis of important decisions with significant social, financial, and environmental implications across diverse sectors of our society including the medical, legal, astronomy, and resource sectors. These interpretations typically involve the recognition of anomalies or specific features of interest within data. In this process, interpreters’ intuition and biases play an important role, resulting in outcomes that are highly subjective and uncertain. This thesis presents a study that aims to understand and address uncertainties in geoscientific data interpretation, specifically focusing on geological target spotting within magnetic geophysical data.

Previous studies on geological interpretation uncertainties have mainly focused on analysing the interpretation outcomes and identifying factors that influenced these outcomes. This thesis on the other hand aims to understand and address uncertainties in human-data interactions, that is, the process by which the interpretation outcomes are reached. This is achieved by monitoring and quantifying physiological and neurological responses of interpreters during target spotting exercises through an eye tracker system (ETS) and electroencephalography (EEG) techniques. These technologies are widely used in various fields, such as human-computer interaction (HCI), brain-computer interface (BCI), clinical studies, and web marketing, but their use in geoscience is in its infancy. Various experiments were carried out to capture ETS and EEG data during target spotting exercises involving the identification of prescribed geological ‘targets’, specifically porphyry-style intrusive systems within magnetic data. Interpreters with varying levels of expertise and experience participated in the experiments, and their data observation patterns were profiled using an ETS and their brain responses were monitored using EEG techniques.

This thesis reports three applications of the ETS and EEG techniques, and their potential to be used to understand biases in geological target spotting. The first application used eye gaze movements, captured through ETS to identify effective data observation practices. Variability in data observation patterns of different interpreters and their target spotting outcomes were analysed. The results show that target spotting performances can be improved by observing the data more
systematically and by observing it from multiple orientations. These findings may be useful in devising a roadmap for the training of magnetic data interpreters.

The second application used ETS data to identify image locations of potential visual attention biases of interpreters. The eye gaze fixation patterns derived from the ETS data were compared with saliency maps. These maps, which represent visual saliency of an image, were generated using various algorithms. The results showed a close correlation of the eye gaze fixation maps with the image saliency maps of a well-known algorithm, the Itti’s saliency model (ITTI) algorithm. This study then proposed the use of this image saliency algorithm to predict and compensate for potential visual attention biases when choosing a set of magnetic data enhancement methods for interpretation. A set of experiments using three different magnetic datasets demonstrated the feasibility of using saliency maps as a tool to select a set of enhanced images for interpretation. This was achieved by analysing the similarities and differences in visual saliency regions between different enhanced images and to ensure the wide coverage of saliency regions across different enhanced images.

Finally, the third application investigated the feasibility of using brain responses to automatically identify geological target selection epiphany, without the need for behavioural responses. Experiments were conducted using target and non-target images and a P300 response associated with target detection was identified from the EEG data using a pattern classification technique, namely support vector machine (SVM). A single trial average classification accuracy of 83% was obtained, which demonstrates the effectiveness of using the P300 response to detect geoscientific target selections.
Acknowledgements

I am grateful to many people whose help and support made this endeavour possible. First and foremost I would like to express my sincere thanks and gratitude to all of my supervisors. My coordinating supervisor, Eun-Jung Holden infused into me the motivation and enthusiasm for research. She gave me the impetus for research work thereby moulding a research worker out of me. She always gladly helped me and guided me anytime I needed help. This continuous guidance and support was always a source of inspiration for my research. I am short of words to express my gratitude to her.

A special word of thanks is reserved for Roberto Togneri. I am grateful for his encouragements, supervision, and sharing his wealth of knowledge during all discussions. My sincere thanks goes to Michael Dentith for introducing me to geophysics and guiding me to communicate my research with others. I would also like to thank Greg Price for his encouragement and insightful feedbacks, and Tele Tan for his advice and guidance.

There are also many others I would like to thank for their contribution to this thesis. I had very useful discussions with Mark Lindsay regarding the analysis of interpreter visual biases, and with Cuntai Guan from the Institute for Infocomm Research (Singapore) who gave me valuable feedback and technical advice on electroencephalography. My sincere thanks goes to Daniel Wedge for his support and friendship throughout the study and the proof reading of this thesis. I would like to thank Joanne Edmondston for her support and helpful advices. I am grateful to Jason Wong for his participation in the initial experimental setups, Jeffrey Shragge for providing the seismic data for initial experiments, and Jelena Markov for her support. My thanks are also due to all the staff members of the School of Earth and Environment and also the Centre for Exploration Targeting at The University of Western Australia for helping me in many ways throughout my study. I specially appreciate the volunteer geoscientists, who participated in the experiments.

I thank iVEC for the loan of the eye tracker system and Barrick Gold of Australia Ltd for permitting me to use their data in this research. I also thank Geological Survey of Ontario, Canada for Kirkland Lake aeromagnetic data. I acknowledge the Scholarship for International Research Fees (SIRF), University International
Stipend and the PhD Completion Scholarship offered by The University of Western Australia.

Lastly and most importantly, I would like to thank my Dad and Mum for teaching me the value of hard work. I am grateful to my parents for their support and prayer for my success. I also wish to thank my wife as she has been the tower of strength and encouragement during times of trials and tribulations. Her encouragements stood in good stead during my studies. This venture would not have been possible without her.
To my parents,
Mr & Mrs Sivarajah,
for their love, endless support,
and encouragement.
Table of Contents

Abstract i
Acknowledgements iii
Table of Contents vii
List of Tables xi
List of Figures xiii
Publications arising from this thesis xv
Contribution of the candidate to the published work xvii
Abbreviations xix

1 Introduction 1
  1.1 Motivation of the research ............................... 1
    1.1.1 Geoscientific data interpretation .................. 2
    1.1.2 Human-data interaction monitoring .......................... 3
  1.2 Thesis focus ........................................ 6
  1.3 Research objectives ................................... 6
  1.4 Research contributions .................................. 7
  1.5 Thesis organisation .................................... 8

2 Review of ETS and EEG technologies and applications 11
  2.1 Introduction ........................................ 11
  2.2 Eye tracker system (ETS) ................................ 11
    2.2.1 Quantifying eye gaze measurements ..................... 14
    2.2.2 Previous applications of ETS .......................... 16
  2.3 Electroencephalography (EEG) ............................ 17
    2.3.1 Electrodes ........................................ 18
    2.3.2 Event related potential (ERP) ......................... 20
List of Tables

2.1 Technical specifications of the mobile ETS .................................. 15
4.1 Pearson product-moment correlation between different eye gaze measures and the target spotting performance ........................................... 50
5.1 Performances of ITTI, GBVS, and HFT algorithms .......................... 66
5.2 Parameters used for the reduction to the pole (RTP) calculations ...... 68
5.3 RMS error values between saliency maps of enhanced magnetic Data 1 72
5.4 RMS error values between saliency maps of enhanced magnetic Data 2 74
5.5 RMS error values between saliency maps of enhanced magnetic Data 3 74
6.1 Pseudo-peak approach ANOVA analysis results ............................. 89
6.2 10-trial average ERP peak approach ANOVA analysis results .......... 90
6.3 Target spotting performances of eight participants. .......................... 91
## List of Figures

2.1 Events occurring during observation ..................................... 12
2.2 Suction cap and a sample position of the eyelids ..................... 12
2.3 The apparatus used to record eye movement .............................. 13
2.4 ETS glasses ................................................................. 13
2.5 Digital video cassette recorder and the analysis computer ............ 14
2.6 Sample image of eye gaze movements .................................... 16
2.7 The 3D representation of the 10-20 electrode settings ................. 18
2.8 Participant wearing the EEG cap ........................................ 19
2.9 EEG amplifier with connections to the electrodes ..................... 19
2.10 Oddball paradigm ......................................................... 21

3.1 Experiment setup ............................................................ 26
3.2 Image used for the calibration of the ETS ............................... 28
3.3 ETS data pre-processing results ......................................... 29
3.4 Hough transform parameters ............................................. 31
3.5 High-pass filtered EEG signals .......................................... 32
3.6 Low-pass filtered EEG signals .......................................... 33
3.7 Scalp distribution of ICA decomposed sources .......................... 34
3.8 ICA decomposed source signals ......................................... 35
3.9 Ocular artefact-minimised EEG signals .................................. 35

4.1 Typical magnetic response of a porphyry-style mineralisation ........ 40
4.2 Experimental setup .......................................................... 41
4.3 ETS data pre-processing results ......................................... 42
4.4 Displayed magnetic image with the true targets ........................ 43
4.5 Magnetic images used for target spotting exercises .................... 44
4.6 Target spotting performances of fourteen participants ................. 46
4.7 Eye-trace plots for four high achievers and four low achievers ....... 47
4.8 Heat map plots obtained for a high achiever and for a low achiever .. 50
4.9 The number of participants who identified each ground truth target . 51
4.10 Fixation map ................................................................. 52
Publications arising from this thesis

Publications directly relevant to this thesis (fully refereed)


2. Yathunanthan Sivarajah, Eun-Jung Holden, Roberto Togneri, Michael Dentith and Mark Lindsay. Visual Saliency and Potential Field Data Enhancements: Where is your attention drawn? *Accepted for publication in Interpretation*. (Chapter 5)


Additional publications


Contribution of the candidate to the published work

The contribution of the candidate in all the published papers was 80%. The candidature was responsible for the original idea. He developed and conducted all the experiments, implemented the algorithms, and did all the analysis. The papers were also written by the candidate. The co-authors reviewed the papers and provided valuable feedbacks for further improvement.
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1VD</td>
<td>first vertical derivative.</td>
</tr>
<tr>
<td>2D</td>
<td>two dimensional.</td>
</tr>
<tr>
<td>3D</td>
<td>three dimensional.</td>
</tr>
<tr>
<td>AGC</td>
<td>automatic gain control.</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance.</td>
</tr>
<tr>
<td>AS</td>
<td>analytic signal.</td>
</tr>
<tr>
<td>AUROC</td>
<td>area under receiver operating characteristic.</td>
</tr>
<tr>
<td>BCI</td>
<td>brain-computer interface.</td>
</tr>
<tr>
<td>BSS</td>
<td>blind source separation.</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid.</td>
</tr>
<tr>
<td>DVCR</td>
<td>digital video cassette recorder.</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalography.</td>
</tr>
<tr>
<td>ERP</td>
<td>event related potential.</td>
</tr>
<tr>
<td>ETS</td>
<td>eye tracker system.</td>
</tr>
<tr>
<td>FoV</td>
<td>field-of-view.</td>
</tr>
<tr>
<td>GBVS</td>
<td>graph based visual saliency.</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>HCI</td>
<td>human-computer interaction.</td>
</tr>
<tr>
<td>HFT</td>
<td>hypercomplex Fourier transform.</td>
</tr>
<tr>
<td>ICA</td>
<td>independent component analysis.</td>
</tr>
<tr>
<td>IR</td>
<td>infrared.</td>
</tr>
<tr>
<td>ITTI</td>
<td>Itti’s saliency model.</td>
</tr>
<tr>
<td>LDA</td>
<td>linear discriminant analysis.</td>
</tr>
<tr>
<td>LED</td>
<td>light emitting diode.</td>
</tr>
<tr>
<td>LPF</td>
<td>low-pass filter.</td>
</tr>
<tr>
<td>MRP</td>
<td>movement-related potential.</td>
</tr>
<tr>
<td>RBF</td>
<td>radial basis function.</td>
</tr>
<tr>
<td>RMS</td>
<td>root mean square.</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristics.</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest.</td>
</tr>
<tr>
<td>RSVP</td>
<td>rapid serial visual presentation.</td>
</tr>
<tr>
<td>RTP</td>
<td>reduction to the pole.</td>
</tr>
<tr>
<td>SVM</td>
<td>support vector machine.</td>
</tr>
<tr>
<td>TDR</td>
<td>tilt derivative.</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Motivation of the research

Recent technological advances and improved accessibility of data acquisition and processing methods have contributed to the collection of vast volumes of data in many sectors. Examples include two dimensional (2D) and three dimensional (3D) ultrasound technologies for medical diagnosis (Benacerraf et al., 2005; Cohen et al., 2001; Lazebnik and Desser, 2007); radio-telescope technology for the imaging of galaxies in astronomy (Kemball et al., 2010); deoxyribonucleic acid (DNA) analysis for forensics (Benecke, 1997; Moody, 1989); and remote sensing (Lucieer et al., 2011) and drilling technologies (Wang et al., 2011a) for the exploration and extraction of resources. The increase in data volume and data types bring new challenges for these sectors as data needs to be processed, interpreted, and transformed into knowledge to maximise the return on investment in data collection.

One of the major challenges for the data to knowledge process is how we understand and address systemic uncertainties in our biased interpretation and decision making. This problem was clearly demonstrated by Dror and Hampikian (2011), who studied the variability in DNA typing interpretation. They asked 17 independent DNA experts from North America, of different age and gender and with a range of experience and different levels of academic training, to examine a mixed DNA sample obtained from a real criminal case where the DNA examiners concluded that the suspect could not be excluded from contributing to the DNA. In their study, even though all of the participants typed the DNA sample in the same laboratory using the same sample, reagents equipment, and interpretation guidelines, only one participant came to the same conclusion as the analyst from the criminal case (‘cannot be excluded’). Four examiners concluded the results were ‘inconclusive’ and the rest (12 examiners) thought the suspect should be ‘excluded’.

Similar human biases and uncertain decisions are also major issues for geoscientific data interpretation. However, uncertainty in geological data interpretation
is further attributed by the complexity of the task which aims to predict geological events occurred in distant past using limited data (Frodeman, 1995). Further, geological data interpretation is a difficult task to improve due to the challenge in accessing the ground truth which is only possible to collect through costly drilling and analysing drill hole data. As a first step towards understanding and addressing uncertainties associated with geoscientific data interpretation, this thesis focuses on understanding how interpreters reach decisions by monitoring quantitative indicators of their interaction with data. This is achieved by harnessing the advances in HCI and BCI technologies to quantitatively profile interpreters’ physiological and neurological responses during interpretation.

1.1.1 Geoscientific data interpretation

Geoscientific data interpretation is a complex task that aims to understand natural phenomenon that occurred in the distant past using a limited set of observations (Frodeman, 1995). Thus, interpreters’ intuition and biases play an important role, which make the interpretation outcomes highly subjective. Yet despite this subjectivity, government agencies, and energy and mineral industries make decisions of significant financial and social importance based on these interpretations such as decision regarding mineral exploration, groundwater supply, natural hazard predictions, and CO$_2$ and nuclear waste storage. While geoscience communities are largely aware of this issue, there have been only a limited number of published reports on the study of interpretation uncertainty and the factors contributing to those uncertainties.

In 2007, Bond et al. (2007) examined the factors affecting the interpretation of seismic data, which are routinely used by the oil and gas industry to determine subsurface structures. They carried out a geological interpretation experiment using a single synthetic seismic dataset, produced by forward modelling. In their experiment, 412 geoscientists from various backgrounds with different education levels and varying experience were asked to interpret the synthetic dataset. Only 21% of the participants interpreted the ‘correct’ tectonic setting and only 23% identified the three major faults present in the data. Analysis of the interpretations showed that the number of years of experience did not influence the ability of the geoscientists to correctly interpret the data as 76% of the participants with more than 15 years of experience were incorrect. Instead, they found that the participants who used multiple interpretation techniques, such as identification of features and horizons, drawing straight lines, annotations, sketches, and writing, had a higher probability of producing the right answer.

In a more recent study, Bond et al. (2011) took the interpretations of 184 struc-
tural geology experts from the previous study (Bond et al., 2007) to further analyse the impact of background on interpretation outcomes. They found that the number of years of experience as an interpreter, technical speciality, or gender had no significant impact on the experts to correctly interpret the data. However, they did find that education level and profession (working in academia or industry) did have a significant impact on interpretation performance. They found that academics were 2.5 times more likely, postgraduate degree holders (Master’s and/or PhD) were 6 times more likely, and academics with postgraduate degrees were 14.5 times more likely to produce the correct interpretation compared to other interpreters. They also found that interpreters who work in academia are more likely to use effective techniques, such as horizon interpretation, evolutionary sketches, and annotation, which may have contributed to the quality of their interpretation. Bond et al. (2011) also found that the interpreters who validated their interpretation by checking geometric and evolutionary feasibility were 3 times more likely to produce the correct results than others who did not use any validation techniques. Overall, they concluded that the use of interpreters with postgraduate training and the use of validation techniques can minimise interpretation uncertainty.

The above mentioned studies mainly focus on ‘right-wrong’ data interpretation outcomes and the factors that influence the outcome. However, to improve the interpretation outcome it is important to better understand the process by which the interpretation is reached. Geoscientific data interpretation is largely a pattern recognition task, where features of geological significance are sought within a single dataset or across multiple datasets through visual inspection. In this process, how data is displayed and observed has a significant impact on interpretation outcomes. This is demonstrated by the development and the use of numerous data enhancement and visualisation techniques to assist visual inspection of geophysical data (Miller and Singh, 1994; Rajagopalan and Milligan, 1994; Roest et al., 1992), which aim to enhance the regions of interest within data to attract interpreters’ visual attention. Thus, to understand and improve geoscientific data interpretation, it is important to analyse and address the data observation biases, and identify effective data inspection practices that can compensate for these biases. To achieve this, it is necessary to monitor how the interpreter interacts with the data.

1.1.2 Human-data interaction monitoring

Monitoring of human-data interactions is an active area of research in which different techniques are used to capture human responses. Some of the monitoring techniques used in this field include eye tracking (Buscher et al., 2009, 2010; Cutrell and Guan, 2007; Vilarino et al., 2007), brain monitoring (Esfahani and Sundararajan, 2012;
1.1. Motivation of the research

Farwell and Donchin, 1988; Nakayama and Abe, 2010), heart rate variability monitoring (Hercegfi, 2011), gesture recognition (Wang et al., 2011b), monitoring of facial expressions (Branco et al., 2006), pupillometry (which is the measurement of the pupil diameter) (Andreassi, 2000; Oliveira et al., 2009; Tullis and Albert, 2010), and skin conductance (Hercegfi et al., 2009).

In this thesis, two specific monitoring techniques, eye gaze tracking using an ETS and brain signal monitoring using EEG techniques, are employed to monitor and profile the interpreter-data interactions (physiological and neurological responses respectively) during interpretation. These techniques have been used widely to quantitatively capture human responses in the fields of HCI and BCI.

Eye tracker systems (ETSs)

The ETSs capture eye gaze positions and movements, and have been used in a wide range of applications. Vilarino et al. (2007) used eye tracking to identify regions of interest (ROIs) when experts were asked to find anomalies within medical images. Several other studies have assessed user behaviours in web searching (Buscher et al., 2009, 2010; Cutrell and Guan, 2007; Granka et al., 2004) and the effectiveness of internet advertising (Drèze and Huss herr, 2003; Hervet et al., 2011).

In the geological area, there have been two recent studies using ETS. Maltese et al. (2013) used ETS to study how students learn through observation during geological field work on their individual mapping and group learning exercises. However, their results were based largely on the captured scene video, rather than the ETS data. The other reported work was a pilot study conducted by Chadwick et al. (2010), who demonstrated the feasibility of capturing eye gaze data to monitor and analyse interpretation of magnetic data. This feasibility study formed the foundation of the ETS application components of this thesis.

Chadwick et al. (2010) used the ETS to understand the differences between data observation patterns from free-viewing vs target spotting exercises, and the impact of a data enhancement method on target detection. A data observation pattern was quantified using visitation frequencies of eye gaze on each location in the data in a form of a heat map. These heat maps were qualitatively assessed, and more experienced geoscientists were found to methodically search a wider area and identify more targets when compared to less experienced geoscientists. Similar to the work of Chadwick et al. (2010), this study used the ETS to understand data observation patterns of interpreters during a geological target detection exercise. However, this study extended the processing and the use of ETS.

In Chapter 3, a robust data observation tracking technique is implemented to track the eye motion over the image based on the ETS data. The eye gaze move-
ments are used to identify effective data observation practices for a specific target detection exercise. To achieve this, data observation patterns of different interpreters were quantitatively compared. Eye gaze tracking measures, such as fixations, saccades, scanpath lengths, and scanpath durations, were used to quantify and analyse variations in data observation patterns. In this study, behavioural responses (i.e. key presses) were also collected to analyse the relationship between different data observation patterns and their associated target spotting outcomes. This is explained in detail in Chapter 4. In addition, this thesis extends the use of ETS to understand and predict the regions where the visual attention of interpreters is likely to be drawn, which is explained in Chapter 5.

Electroencephalography (EEG)

In traditional BCIs, the intentions of users such as key pressing and selections, are identified by capturing EEG signals while the users mentally select simple targets such as shapes, colours, or letters (Esfahani and Sundararajan, 2012; Farwell and Donchin, 1988; Nakayama and Abe, 2010). There have been few studies that report using EEG signals to identify complex targets. Gerson et al. (2006) used event related potential (ERP) P300 responses, which are related to target selection, to classify the natural images into images with and without people. In their target spotting exercise, stimulus complexity was associated with the position, scale, and pose of the human targets. They found that EEG based responses in combination with button clicks are effective for the identification of targets. Huang et al. (2011) developed a BCI to search for surface-to-air missile sites within satellite images by detecting ERP responses within EEG. Their results showed that the ERP-based interface is five times faster than traditional target spotting exercises. More recently, some studies have reported using both EEG and ETS in HCI applications. Lee et al. (2010) developed a 3D BCI using ETS and EEG. They controlled 2D navigation by tracking eye gaze movements and controlled depth directional navigation using EEG. In a recent study, Kamienkowski et al. (2012) proposed a BCI using EEG and ETS to detect targets (i.e. a letter) during a free-viewing target search exercise.

Monitoring of the human-data interactions by combining the ETS and EEG responses will provide a more complete analysis of the interpretation process. As a first step towards this goal, this study explored the feasibility of using EEG signals to identify complex geological target selection. To the best of our knowledge, the use of EEG has not been attempted for complex geoscientific target detection, which provides the motivation for this thesis.
1.2 Thesis focus

This thesis focuses on understanding the geoscientific data interpretation practices that lead to improved interpretation outcomes. This is achieved by quantitatively monitoring and profiling interpreter data interactions by capturing interpreter data observation patterns, brain responses, and target selections during the interpretation process. Eye gaze movements were captured using an ETS, brain responses were obtained from EEG, and the target selections through button clicks.

In this study, the data were captured during target spotting experiments in which the task was to identify responses associated with porphyry-style mineralisation within magnetic data. Magnetic data are one of the essential datasets for mineral exploration, as they are relatively cheap to acquire over a large area, when compared to other geophysical mapping methods such as gravity and electromagnetic surveys. Many national and regional government agencies collect these data and make them available at nominal or no cost, to promote exploration in their jurisdiction. Mining and exploration companies routinely supplement these data with higher resolution datasets acquired within the area.

Experiments were designed to capture the ETS and EEG data while the participants identified prescribed targets within magnetic images. Eye gaze data were used to identify the interpreter data observation patterns and the areas of interest within the magnetic data. These data observation patterns were then used to understand the variability between and within participants and the impacts of data observation patterns on the target spotting performances, which helped to identify the effective interpretation practices. Identification of areas of interest within magnetic data using eye gaze fixations from ETS data helped to select the human attention model that can generate visual attention maps. This was used to analyse the interpreter visual biases which are introduced when magnetic data are viewed after applying different data enhancement and filtering techniques. The EEG responses were used to investigate the feasibility of using EEG techniques as a tool to identify the relationship between the target spotting difficulty and the characteristics of the targets.

1.3 Research objectives

The primary goal of this research is to quantitatively monitor and profile the human-data interactions during geoscientific interpretation processes by capturing the data observation patterns, brain responses, and behavioural responses of the interpreters. The ultimate goal is to use them to better understand the interpretation process and identify effective interpretation practices. More formally, this thesis aims to
achieve the following objectives:

1. Use ETS, image analysis techniques, and statistical analysis to understand the variability in data observation patterns between and within individuals for the identification of effective interpretation practices (Chapter 4).

2. Use ETS and image processing techniques to select a tool that can computationally predict the interpreters’ visual attention for the analysis of interpreter visual biases when magnetic data are visualised after applying different enhancement and filtering methods (Chapter 5).

3. Use EEG recordings and signal processing techniques to identify the brain response patterns that are associated with targeting important geological features within the magnetic data to measure the target spotting difficulty (Chapter 6).

1.4 Research contributions

This thesis makes four major contributions:

1. First ever collection of ETS and EEG data during complex geological target spotting exercises within magnetic data. The appropriate experimental setup and data capture methodologies for the analysis of ETS and EEG data were developed. This was achieved by analysing various experimental setups and selecting the best configuration that enabled the capturing of the eye gaze movements and the brain responses during complex geological target spotting exercises.

2. Quantitative analysis of the interpreter data observations and new insights for effective interpretation of magnetic data. Previous studies showed the effectiveness of ETS and the possibility of using this system for geoscientific data interpretation. To our knowledge this is the first study which quantifies the data observation patterns, and statistically analyses the relationship between the data observation patterns and the target spotting performance.

3. First reported study of the selection of a human attention model for complex magnetic images to predict potential interpreter visual biases. This can be used to generate interpreter visual attention maps computationally which will produce these maps quickly without requiring complicated experimental setup and signal capturing. These attention maps will not only provide guidance for the interpreters but also provide feedback for software and data enhancement algorithm developers.
4. First report of the classification of complex magnetic images into targets and non-targets using ERP P300 responses. The promising results obtained in this thesis provide the basis for further analysis of the variations in brain responses to identify the geoscientific target spotting task difficulty.

1.5 Thesis organisation

In line with the regulations of The University of Western Australia, this thesis is organised as a series of published or accepted manuscripts in fully refereed international journals. Firstly, in Chapter 2, the necessary background information about ETS and EEG techniques, and their applications are provided. Next, in Chapter 3, the experimental design and setup for the ETS and EEG data capture and preprocessing are detailed. Chapters 4 to 6 constitutes three manuscripts which are the applications of ETS and EEG towards understanding geological target spotting.

Chapter 4 presents a study which captures and analyses eye gaze movements using an ETS to understand how individuals interact with data. This is then used to identify effective interpretation practices. Fourteen participants were used in the experiments and they performed two target spotting exercises within aeromagnetic data. During the exercises the data was displayed in different orientations and the participants were requested to identify as many targets as possible within a three-minute interval. The accuracy and efficiency of the geological target detection for individual interpreters were assessed using the captured data observation patterns and the interpreter feedback. This chapter was published in Interpretation (Sivarakjah et al., 2013).

Chapter 5 considers the selection of a suitable human visual attention model for the identification of visual biases in data observation when interpreting enhanced magnetic data using commonly used filtering and enhancement techniques. The effectiveness of various human visual attention models were analysed by comparing the heat maps generated using those models with the interpreter data observation patterns captured using the ETS. This study provides a useful guide for choosing enhancement methods, based on saliency maps that minimise unintended visual biases in magnetic data interpretation and some recommendations for the use of commonly used enhancement methods for different types of magnetic data. This chapter was accepted with minor revision for publication in Interpretation (Sivarakjah et al., 2014a).

Chapter 6 investigates the identification of brain response patterns that are associated with geological target spotting. This is achieved by detecting ERP P300-like responses that are associated with target detection from the EEG of interpreters.
For this analysis the EEG was captured from eight participants during a target spotting task within aeromagnetic images. Classification of the images into targets and non-targets based on the ERP P300-like responses was carried out and correlated with the button click responses to determine whether ERP P300-like responses can be reliably used for the analysis of variations in the target identification difficulty. This research was published in International Journal of Human-Computer Studies (Sivarajah et al., 2014b)

Finally, Chapter 7 provides the conclusions and proposes avenues of future research.
1.5. Thesis organisation
Chapter 2

Review of ETS and EEG technologies and applications

2.1 Introduction

This chapter presents a review of eye tracking and EEG monitoring techniques. The first part of this chapter introduces ETS technology and the different measures which are derived from ETS data to quantify data observation patterns. It then reviews previous studies of eye tracking specifically in the field of geoscience. The second part of this chapter describes EEG technology and of EEG signal capture, and then discusses the use of EEG for spotting complex targets.

2.2 Eye tracker system (ETS)

Eye movements control the ability to observe details of a scene as it is not possible to process information outside of the foveal vision without moving the eyes (Rayner, 1998). While performing basic tasks, such as reading or simply looking at a scene, eyes move on average every 200 - 350 ms. When a scene is viewed, only a small portion of the scene registered in the vision system is processed. Under natural viewing conditions, significant changes in an image remain invisible to observers which is referred to as change blindness (O’Regan et al., 1999). In the visual system, new information is only acquired during fixation and vision is suppressed during saccade (Henderson, 2003). Fixation is defined as maintaining eye gaze on a particular location for at least 100 - 150 ms (Viviani, 1990) and saccade is the quick eye gaze movement from one location to another. Eyes typically fixate on locations that are surprising, salient, or significant (Loftus and Mackworth, 1978). As shown in Figure 2.1, eyes will move from one fixation to another, once the visual information related to that fixation has been processed and a new location has been selected based on what is present in peripheral areas (Goldberg and Kotval, 1999; Henderson, 2007).
Buswell (1935) and Yarbus (1967) pioneered research addressing the impact of a task on eye gaze and fixation. Yarbus (1967) designed miniature optical devices or suction ‘caps’ to capture eye movements. He constructed different caps which varied in shape, size, and the materials used. These caps were fixed to the anesthetised surface of the subject’s eye using suction. The cap used to record eye movements in his experiments is shown in Figure 2.2a. In order to prevent blinking, the eyelids were secured using adhesive plasters (Figure 2.2b). The eye movements were then recorded using an apparatus consisting of a stand, two light sources, a chin rest, and a control panel as shown in Figure 2.3. The chin rest was used to steady the subject’s head during the experiment. Different light sources and accessories were used for the different types of caps tested.

In an experiment carried out by Yarbus (1967), one participant was asked to view a painting under seven different instructions. On first viewing, no specific instruction was given to the participant when observing the image. In the subsequent six viewings, the participant was given different instructions such as “remember the clothes worn by the people”, “remember the position of people and objects in
the room”, and “give the ages of the people” (Yarbus, 1967, p. 174). In another experiment by Yarbus (1967), seven participants were asked to observe the same painting without any instruction. Significant variations were observed between the eye movement patterns if different instructions were given to the participant and these variations were found to be greater than variations on eye movements on free-viewing of an image by different individuals. The study concluded that the eyes “fixate on those elements of an object which carry or may carry essential or useful information” (Yarbus, 1967, p. 211). These findings were later confirmed by several studies which showed that the intention of a person influences their eye movements (DeAngelus and Pelz, 2009; Einhäuser et al., 2008; Rothkopf et al., 2007).

With advances in modern technology, the cumbersome ETS that was devised
2.2. Eye tracker system (ETS)

Figure 2.5: (a) Sony digital video cassette recorder with attached Applied Science Laboratories circuit board for data processing; (b) Analysis computer - used for real time calibration or post-data collection analysis.

by Yarbus (1967) has now been replaced by a pair of glasses that is equipped with cameras and processors to track eye gaze (Figures 2.4 and 2.5). In this study, a mobile ETS built by Applied Science Laboratories\(^1\) (ASL) is used. This ETS uses two video cameras (a scene camera and an eye camera) and three harmless near infrared (IR) light emitting diodes (LEDs), which are mounted on a pair of standard safety glasses (Figure 2.4). Over the right eye there is a circular cut-out in the lens which allows the placement of an adjustable monocle called a ‘combiner’. The combiner is partially reflective in the near IR and IR range, and is attached to the lens using a ball joint to allow adjustment. The IR light beam from the IR LEDs, which are placed in a triangular pattern, are reflected by the combiner onto the eye surface. The forward facing scene camera captures the participant’s field-of-view (FoV), and the eye video camera captures the eye image and corneal reflection sent back by the combiner. By comparing the positions of the reflected IR and the pupil, the eye gaze locations are determined with respect to the FoV by the Applied Science Laboratories circuit board, which is attached to a digital video cassette recorder (DVCR) (Figure 2.5a). The calculated eye gaze locations and the FoV video can be either recorded in the video cassette using the DVCR (Figure 2.5a) or on a laptop (Figure 2.5b). The technical specifications of this ETS are shown in Table 2.1.

2.2.1 Quantifying eye gaze measurements

In the experiments reported in Chapters 4 and 5, an ETS was used to capture interpreters’ eye gaze movements over the data during interpretation. Based on the

\(^1\)Applied Science Laboratories, http://www.asleyetracking.com
Table 2.1: Technical specifications of the Applied Science Laboratories mobile ETS.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sampling and output rates</td>
<td>30 Hz</td>
</tr>
<tr>
<td>Measurement principle</td>
<td>Pupil-corneal reflection with custom outdoor enhancements</td>
</tr>
<tr>
<td>System accuracy</td>
<td>0.5° visual angle</td>
</tr>
<tr>
<td>Resolution</td>
<td>0.1° visual angle</td>
</tr>
<tr>
<td>Head movement</td>
<td>Unlimited</td>
</tr>
<tr>
<td>Visual range</td>
<td>50° horizontal and 40° vertical</td>
</tr>
</tbody>
</table>

captured eye gaze locations, observations can be quantified using different eye gaze measures. These measures include: number of fixations, location of fixation, and scanpath length and duration.

The number of fixations indicates the number of visual components a user processed. This measure is often used in combination with the locations of fixation to identify the areas of interest (DeAngelus and Pelz, 2009; Goldberg and Kotval, 1999). The location of fixation identifies the areas on which observer’s attention is focused in the FoV. In an encoding task, a higher number of fixations on a particular area means that the area is of greater interest (Poole et al., 2005). The scanpath length is the summation of the distances between the gaze points that are sampled at a fixed sampling rate as shown in Figure 2.6. A longer scanpath length indicates less systematic data observation or poor performance (Goldberg and Kotval, 1999). The scanpath duration is the amount of time spent between two eye gaze points, which is the number of sample points between them × sampling duration as shown in Figure 2.6. Note that this measure will incorporate fixations and saccades, i.e. scanpath duration = fixation durations + saccade durations. Longer scanpath durations indicate less efficient search (Goldberg and Kotval, 1999).

These eye gaze measures have been used in various eye tracking applications. DeAngelus and Pelz (2009) used eye gaze fixation locations and the number of fixations to identify the ROI when viewing a painting. Scanpath length and duration were used to compare the quality of a computer interface by Goldberg and Kotval (1999).
2.2. Eye tracker system (ETS)

Figure 2.6: Sample image of eye gaze movements. The points 1 - 10 represent ten consecutive eye gaze sampling points generated by an ETS. The scanpath length between sampling points 1 and 10 is $a + b + c + d + e + f + g + h + i$, while their scanpath duration is $10 \times$ sampling duration.

2.2.2 Previous applications of ETS

The ETSs have been applied widely for various applications. Eye gaze movements have been analysed during reading (Reichle et al., 1998, 2003), web searching (Buscher et al., 2009, 2010; Cutrell and Guan, 2007; Granka et al., 2004), and to study internet advertising (Drèze and Huss herr, 2003; Hervet et al., 2011). They have also been used in HCI, mainly for people with disabilities (Hutchinson et al., 1989; Levine, 1981, 1984) and to help autistic children to learn social skills (Ramloll et al., 2004). Jerald and Daily (2002) used ETS to track the eye gaze focus of people during video conferencing.

In a medical study by Vilarino et al. (2007), an ETS was used to assist the labelling of ROIs in endoscopy images. In their study, experts were fitted with an ETS and asked to focus their visual attention in a ROI when detecting a polyp (target) in an endoscopy video. When they detected an ROI, they were asked to press and hold a key. Based on the eye gaze locations at the time of key press the polyp ROIs were labelled automatically as positive and the non-polyp ROIs as negative. These were then used to train a machine learning classifier in an automated polyp detection system.

Two geological applications using ETS have also been reported. Maltese et al. (2013) observed the behaviours of students in geological field work using a mobile ETS. Eye gaze data were collected from six students (five undergraduate students and one master’s student) during two different geological field mapping exercises. In the first exercise, three students were given a set of materials including a topographic map and a pair of stereo photos, and they were asked to map the area. During the exercise, the students studied hand samples as well as outcrops and overall landscape. This resulted in inaccurate eye gaze measurements due to the wide
range of distances of objects from the viewing points. To get accurate results for different situations, the device should have been re-calibrated. Thus, even though this study used an ETS, the reported results were mainly based on the scene video rather than based on the eye gaze tracking measures. Their experiments showed that different students spent the exercise time differently in various activities such as working at an outcrop, working on mapping, moving and searching for outcrops, taking notes, and taking samples and measurements. In the second exercise, they captured the eye movements of three students during a group learning activity in the field. The results showed that tasks such as locating themselves in space often distracted students’ field learning.

In another geological study, Chadwick et al. (2010) analysed the data observation patterns of geophysical magnetic data interpreters using an ETS during a target spotting task. This was a feasibility study for research reported in this thesis. Six geoscientists with different levels of skills and expertise performed two different experiments. Total magnetic intensity data and the vertical derivative data of the same region were displayed during both experiments. In the first experiment, participants were allowed to freely observe the data for three minutes. For the second experiment, participants were given two minutes to perform three specific target spotting tasks including the identification of granitoid intrusions, faults, and kimberlite pipes. The results showed that the participants with more experience scanned larger areas and had a more methodical search pattern compared to the less experienced participants, and the more experienced participants identified more targets. Their findings also showed that all participants were able to identify key features, in both of the images presented. However, no behavioural responses were captured in their study and the quantitative analyses were performed using heat maps, which represent the accumulated visitation frequencies at different locations.

### 2.3 Electroencephalography (EEG)

Another technique used for monitoring interpreter-data interactions in this thesis was EEG which records the electrical neural activity at different locations across the scalp. An EEG measures the variations in voltage caused by the ionic current flows within the brain (Niedermeyer and Silva, 2005). Electrical activity of the exposed cerebral hemispheres of rabbits and monkeys was first presented by Caton in the British Medical Journal in 1875. Based on this and others’ work on animals, Berger extended the study to involve human subjects and recorded the first electroencephalogram of the human brain in 1929. The EEG was initially used in clinical applications (Hyland et al., 1939), and later it became a valuable tool in BCIs due
to its non-invasive nature and high temporal resolution. The EEG is captured by multiple electrodes that are placed on the scalp.

2.3.1 Electrodes

The locations of the electrodes over the scalp are defined in the ‘10-20 system’ (Jasper, 1958), which is recognised as the standard system for placing EEG electrodes. This system enables the reproducibility of EEG signals and also makes it possible to compare signals between different subjects. The numbers ‘10’ and ‘20’ indicate the distance between electrodes which are 10% or 20% of the distance between nasion and inion or left and right pre-auricular points (Figure 2.7). The American EEG Society recommends using 21 electrodes in the standard electrode placement, but more recently the recommendation was modified to include more electrodes (Reilly, 2005).

Electrodes are labelled using a letter and a number, which indicate the cortical lobes and hemisphere locations respectively (Figure 2.7). Letters F, T, C, P, and O stand for frontal, temporal, central (which refers to the central region instead of being called central lobe), parietal, and occipital lobes respectively. The additional letter ‘z’ indicates the electrodes placed on the midline. In addition to these letters, A, Pg, and Fp are used to refer to earlobes, nasopharyngeal, and front polar sites respectively. Even numbers are used to indicate the electrodes placed on the right hemisphere and odd numbers are used for the electrodes placed on the left hemisphere (Reilly, 2005). A participant wearing the EEG electrode cap and the
locations of the electrodes are shown in Figure 2.8. Typically captured EEG signals using those electrodes are in a microvolt range, thus they need to be amplified before being recorded. In the experiments described in next chapter (Chapter 3), a NuAmps\textsuperscript{2} amplifier (Figure 2.9) was used to amplify the EEG signals.

\textsuperscript{2}NuAmps, http://www.neuroscan.com/nuamps.cfm
2.3.2 Event related potential (ERP)

An ERP is the brain response to an external stimulus (visual, somatosensory, or auditory). Captured EEG signals represent thousands of simultaneously ongoing brain activities, which makes the detection of the brain responses to a single stimulus from a single trial very difficult. In order to overcome this problem, each stimulus is presented to the subject multiple times (multiple trials) and the EEG responses extracted with respect to the stimulus presentation time are averaged across multiple trials to obtain the ERP responses. This process enhances detection of the brain responses caused by the stimulus by averaging out the background EEG activities.

Components of ERP

The ERP response signal contains a sequence of positive and negative voltage peaks, which are caused by ‘higher’ cognitive processes that may involve memory, expectation, attention, and changes in mental state. Most of these different components of the ERP response are referred to by a letter which corresponds to the polarity of the peak (either N - negative or P - positive), followed by a number which is either the latency (time delay) of the peak with respect to the stimulus onset in milliseconds, or the component’s numerical order in the ERP response. For example, the first substantial peak that occurs in the negative side at a latency of 100 ms after the stimulus presentation is referred as N100 or N1. One of the widely researched components of ERP is the P300 response, which is related to decision making.

2.3.3 P300 response

The P300 signal is defined as the positive deflection in the ERP around 300 ms or more after the presentation of the stimulus (Duncan et al., 2009). An example of a P300 response is shown later in Figure 6.2 in Chapter 6. It is also referred as the P3 wave since it is the third major positive peak in the late sensory evoked potential (Ritter et al., 1968).

Elicitation of P300

P300 responses are normally elicited using the ‘oddball paradigm’, where infrequent target stimuli (‘oddball’) and frequent standard stimuli are presented in a random sequence (Figure 2.10). This oddball paradigm was first used by Ritter and Vaughan (1969), and in this paradigm, the task is to identify the target stimuli either by mental acknowledgement or by pressing a button. The process of stimulus evaluation, especially a successful response to the infrequent target stimulus elicits the requi-
site P300 response. The goal directed ‘target’ aspect is thus important in P300 elicitation, as utilised in this study.

![Figure 2.10: Sequence of images presented in a typical oddball paradigm, where images with green boxes and red boxes are shown to the participant in a random sequence and the task is to identify the infrequent green boxes.](image)

2.3.4 Previous studies based on ERP and P300 responses

In the field of HCI, ERP P300 responses are widely used in BCI applications. A well-known application is the BCI speller, where letters and numbers are arranged in a grid, and the rows and columns of the grid are flashed in random sequence (Farwell and Donchin, 1988). The user focuses attention on a particular character and when the corresponding row or column is flashed, the P300 response is elicited. By detecting the P300 response from the captured EEG, the user selected characters are identified and slowly typed.

The P300 responses are also used for the identification of target detection within images. Typically, target detection can be recognised by human behavioural responses such as button click. However, ERP responses are preferred over behavioural responses for a number of reasons. The button click reaction times comprise a number of different cognitive processes and it is difficult to identify the particular process that is attributed to the variation in button click reaction times (Luck, 2005). In contrast, the ERP responses provide a continuous measurement between the stimulus presentation and the response (Luck, 2005). This helps to identify which processes are being impacted by variations in data and experimental conditions. Another advantage of ERP is that it provides responses even when there is no behavioural change.

Gerson et al. (2006) developed a real-time BCI that sorted a set of images by separating images containing a target from images without a target. They carried
out experiments to sort the images containing people in a sequence of images presented using the rapid serial visual presentation (RSVP) paradigm. In the RSVP paradigm, images are presented one after the other in a rapid fashion, usually every few hundred milliseconds. Gerson et al. (2006) used two sets of experiments to capture brain and button click responses. In the first block of experiments, they displayed two target images which contained one or more people in a natural scene. These target images were mixed with 98 non-target images in an image sequence and these images were shown to five participants, one at a time. Each image was displayed for 100 ms and the EEG signal was captured. In the second block, participants were asked to press the left mouse button immediately after identifying a target image in the sequential image display. Participants were asked to double click the button if target images were shown one after the other. These 98 non-target and two target images in a sequence was randomly selected from a pool of 251 non-target and 33 target greyscale images. In each block, the participants were presented with a total of 5000 images in sequences of 100 images. A Fisher linear discriminator was used to classify the ERP P300 responses, and the first 2500 images were used for training and the remaining 2500 images for testing. Their results indicated that there is no significant difference in classification performance between the EEG and the button press based methods. However, the EEG based method was superior for the subjects who identified a lesser number of target images correctly during the button press exercise. They suggested that integrating the EEG and behavioural responses would provide a better classification of images.

In another study by Mathan et al. (2008), ERP responses were used to identify surface-to-air missile sites from greyscale satellite images. In this pilot study, seven military image analysts were asked to identify targets within the images presented in a RSVP paradigm at a rate of 100 ms per image. Subjects were asked to press a key when they identified a target and their EEG signal was also recorded during the experiment. They classified the images into targets and distractors based on the button clicks, and also on the ERP responses using a machine learning based classifier, namely the SVM. Outputs of these two classifications were fused together to produce probability contour maps. Their results showed that the combined use of button clicks and ERP responses provided an increase in speed of the target detection process over six times compared with traditional search without compromising target detection accuracy.

Recently, Huang et al. (2012) analysed the relationship between ERP response characteristics and the level of difficulty of a target detection task. Twelve subjects were asked to identify surface-to-air missile sites from satellite images using a RSVP paradigm, i.e. similar to the work of Mathan et al. (2008). The subjects were asked
to press the space bar when they identified the targets. The difficulty of the target detection task differed in two aspects; visual stimulus complexity (target difficulty) and task difficulty. They varied the task difficulty by presenting the stimulus for different durations (increasing the image duration decreases the task difficulty). Each participant was involved in five sessions of experiments and the image display duration for each session was 25, 50, 100, 150, and 200 ms. Target images were divided into four categories depending on the target difficulty (easy, medium-easy, medium-hard, and hard). Their results showed that easier targets require a relatively short duration to reach the detection rate of 75%, whilst a longer image duration was needed to achieve the same detection rate for more difficult targets. They also found that as the task difficulty increases, ERP latency becomes longer, and both task and target difficulties are correlated with ERP responses.

2.4 Summary

This chapter describes two widely used human-data interaction monitoring techniques; ETS and EEG. The first part of this chapter described the ETS, the different measures used in various eye tracking studies and their applications. In particular the pilot study by Chadwick et al. (2010) showed the effectiveness of ETS in analysing interpreter data observation patterns during geological target spotting. The second part of this chapter provides background on the EEG including the electrodes placement, ERP, and P300 responses. Some key studies using EEG for identification of complex targets are presented. These studies show the usefulness of P300 responses for the identification of target detection. A study by Mathan et al. (2008) demonstrated that combined use of button clicks and ERP responses can speed up the target detection process more than six times when compared with traditional searching, without compromising the target detection accuracy.

Both ETS and EEG techniques were used in our experiments to quantitatively monitor and profile the interpreter-data interactions. The ETS technique was used for identification of interpreter data observation patterns and the areas of interests within magnetic data in Chapters 4 and 5 respectively. The EEG technique was used to analyse the possibility of identifying the target selections by detecting the P300 responses in Chapter 6. The experimental setup, ETS and EEG data capture, and pre-processing steps are described in the next chapter.
2.4. Summary
Experimental data capture and pre-processing

3.1 Introduction
In this study, two geoscientific target spotting experiments were designed and carried out to monitor human-data interactions during the interpretation process. During these experiments the interpreter eye gaze movements were captured using an ETS and the brain responses were obtained from EEG. The first part of this chapter describes the experimental setup and the tasks that were requested from the interpreters. It then provides the details of the methodology for pre-processing ETS data for analysis. This involves the development of an image analysis algorithm to transform eye gaze coordinates (with respect to the FoV video frame) on to the displayed image. Finally, it explains the EEG data capture and pre-processing, mainly focusing on removing artefacts from noisy EEG signals.

3.2 Experiments
Different target spotting experiments were carried out with magnetic data, where the participants were asked to identify geological targets that have characteristics suggestive of gold-copper-rich porphyry systems. The relevant magnetic anomalies have a distinctive ‘Mexican hat’ like character comprising sub-circular magnetic highs with surrounding annular lows (Holden et al., 2011; Hoschke, 2011) (An example image is later shown in Figure 4.1 in Chapter 4). The magnetic survey data used in the experiments are over a very well explored area that has a number of known deposits.

Fourteen participants (six females and eight males) with varying levels of experience and expertise participated in the experiments. All of the participants were trained geophysicists or geologists with at least some experience in magnetic data interpretation. All the participants had normal or corrected to normal vision (i.e. using contact lenses) and had at least a basic understanding about the magnetic
3.2. Experiments

anomaly associated with porphyry-style mineralisation. Prior ethics approval was obtained from The University of Western Australia, Human Research Ethics Office and the participants consented to participate in the experiments (a sample consent form is included in Appendix B).

![Experiment setup](image)

Figure 3.1: Experiment setup: Participant wearing ETS glasses and EEG cap seated in front of the display monitor. A second desktop computer was used to record EEG signals and the laptop was used to record ETS data.

Participants were seated in front of a display monitor (52 cm x 33 cm) at a convenient distance (from 60 to 100 cm) and were then fitted with ETS glasses and the EEG cap to capture their eye gaze movements and brain responses respectively (Figure 3.1). Two desktop computers and a laptop computer were used in the experiments. One desktop computer was used to display the magnetic images and record user responses, and the other desktop computer recorded the EEG signals. The laptop computer was used to perform ETS calibration and data recording. Written instructions were displayed on the monitor screen at the beginning of each exercise. The participants were requested to press a button on the keyboard as soon as they identified targets. Following the experiments all participants were asked to rank themselves (from 1 to 10) for their own perceived level of expertise for this task.

Two types of target spotting experiments were carried out. The first experiment involved a magnetic image with multiple targets, which will be referred to as a large-scale magnetic image throughout this thesis. This image type is later shown in Figure 4.5 in Chapter 4. The second experiment used small-scale images where each image only covered a small area, which may typically contain a single target. Example small-scale images with target and without target are shown later in Figure 5.5 in Chapter 5.
Chapter 3. Experimental data capture and pre-processing

The first experiment focused on capturing data observation patterns of participants using the ETS during two target spotting exercises. The same magnetic image was used for both of those exercises, but the image was displayed in different orientations. For the first exercise, the image was displayed in a ‘normal’ fashion, that is north towards the top of the screen, and for the second exercise the image was rotated by 180°. These two images are later shown in Figure 4.5 in Chapter 4. The participants were given three minutes for each exercise and had a 30-minute break between the two exercises, during which they performed other tasks. The participants were requested to identify as many porphyry targets as possible within the given time. They were asked to press a key on the keyboard while fixing their eye gaze on the target when selecting a target. The captured eye tracking data from these exercises were used for the identification of effective data observation patterns in Chapter 4; and for the identification of a suitable saliency detection algorithm in Chapter 5. The EEG data obtained in this experiment are kept for later analysis.

The second experiment was conducted to monitor target spotting related brain responses of interpreters using EEG technique and to generate interpreter visual attention maps using data observation patterns. This experiment used an ‘oddball’ paradigm (which was previously described in Chapter 2) involving a random display sequence of target and non-target images. The large-scale magnetic image was cropped into small images, some with single porphyry targets, and others without targets. The target images were obtained from the regions of known deposits and the non-target images were randomly selected from regions known not to contain any porphyry-style anomalies. In the visual display, eight target images were repeated 10 times and 50 non-target images were repeated six times. These images were shown in a random sequence, but the same sequence was used for all the participants. The images were displayed on the centre of the monitor for 1000 ms with an inter-image interval of 1000 ms in which a blank screen was shown. The participants were requested to respond to target images by pressing a key on a keyboard, as soon as they spotted the target. The eye tracking data obtained in this experiment were used for the identification of a suitable saliency detection algorithm in Chapter 5 and the captured EEG data were used for the identification of brain response that is associated with target spotting epiphany in Chapter 6.

3.3 ETS data capture and pre-processing

A mobile ETS by Applied Science Laboratories, introduced in Chapter 2, was used to capture eye gaze movements. This ETS needs to be calibrated for each participant since even a small change in the optics placement can reduce data accuracy. The
ETS calibration is performed by requesting the participants to fixate their eye gaze at known locations and marking those points on the FoV video. Thirteen points were used to cover the entire monitor as shown in Figure 3.2. The accuracy of the eye gaze coordinates was verified between different exercises.

Figure 3.2: Image used for the calibration of the ETS, which contains 13 spots covering the entire monitor screen.

The ETS follows the interpreters’ eye gaze movements in real time to calculate the eye gaze locations with respect to the FoV at a fixed interval. A sample FoV video frame with the recorded eye gaze location is shown in Figure 3.3a. The data observation pattern analysis, however, requires the identification of the location of eye gaze within the displayed image. Frame by frame transformation of the eye gaze coordinates is required as the location of the displayed image with respect to the FoV video frame changes between frames due to participants’ head movements. A robust algorithm was developed to transform the FoV coordinates to the image coordinates using image processing techniques. This algorithm firstly identifies the corners of the displayed image and then transforms the eye gaze coordinates by analysing each FoV video frame.

Ideally, the entire monitor screen should lie within the FoV throughout the experiment. However, this is practically not feasible as participants need to be seated without any body (particularly head) movements. Even though the participants were asked to minimise their body movements during the experiment, there were a few frames where the displayed image was partially outside the FoV. These frames were removed from the sequence.

To ensure easier identification of the displayed image, a black background was used behind the monitor. However, some bright spots (due to the reflected light from the monitor screen) were captured in the FoV video. These spots were removed by applying morphological operations (Gonzalez and Woods, 2002, p. 519 - 566).

The next step was to identify the coordinates of the corners of the displayed image in each of the FoV frames. Simple corner detection algorithms were tested but
found to be unsuitable as there were dark features near the corners in the displayed image, as shown in Figure 3.3a. Thus, an alternative strategy was implemented. The four border lines of the displayed image were identified by an edge detection technique, and then the corners were obtained by extrapolating these lines and calculating their intersections. Ideally, the edges of the displayed image should appear as a straight line, but they were deformed due to the lens distortion of the FoV camera (Figure 3.3a). This was rectified by applying a lens distortion transformation (Figure 3.3b).

**Lens distortion removal:** Distortion calibration methods measure how much edges are distorted and then minimise these distortions by varying the distortion parameters. Based on the first order distortion model (Devernay and Faugeras, 1995), the undistorted coordinates \((x_u, y_u)\) are related to the distorted ones \((x_d, y_d)\) as follows,

\[
\begin{align*}
x_u &= x_d + (x_d - c_x)kr_d^2 \\
y_u &= y_d + (y_d - c_y)kr_d^2
\end{align*}
\]  
\[ (3.1) \]
3.3. **ETS data capture and pre-processing**

where \(k\) is the first order distortion parameter, \((c_x, c_y)\) are the coordinates of the centre of distortion, and \(r_d\) is the distorted radius from the centre point of the image. The radius value \(r_d\) is defined as follows,

\[
r_d = \sqrt{\left(\frac{x_d - c_x}{s_x}\right)^2 + (y_d - c_y)^2}
\]

(3.2)

where \(s_x\) is the distortion aspect ratio.

In order to obtain the undistorted image the distortion parameters \(k\), \(c_x\), \(c_y\), and \(s_x\) need to be identified. The algorithm described in Devernay and Faugeras (1995) was used for the identification of the distortion parameters. This algorithm firstly identifies the most dominant segments in the image, which are straight lines in the real world but appear as curves in the images. To calculate how much each segment is distorted, a least squares fitting method was used to find the best fit line from the edge pixels. A distortion error \((E_0)\) is calculated by obtaining the sum of squares of distances from the edge pixels to the least squares fitted line. This algorithm then optimises the distortion parameters \((k, c_x, c_y, \text{ and } s_x)\) to minimise the distortion error by varying the optimisation parameters and computes another distortion error \((E_1)\) using the optimised parameters. These steps are repeated until the relative change of error \(((E_0 - E_1)/E_1)\) becomes lower than a certain threshold. After completing this process the optimum values for the distortion parameters will be identified. Finally, using these parameters the undistorted image is obtained using Equation 3.1.

**Edge Detection:** The lens distortion corrected FoV video frames were thresholded and the edges were detected using the widely used Canny edge detection method (Canny, 1986) (Figure 3.3c). The Canny edge detection algorithm firstly applies a Gaussian filter for smoothing. It then calculates first order derivatives in the vertical and horizontal directions to estimate the edge gradient strength and direction. Next, a non-maximal suppression technique is applied to thin the edges by setting the value of a pixel to zero, if its gradient magnitude is less than the magnitude of its two neighbours in the gradient direction. The Canny algorithm then uses a hysteresis thresholding technique to mark edge pixels. The hysteresis thresholding technique uses two thresholds \(T_1\) and \(T_2\), with \(T_1 > T_2\). The algorithm traces the edge pixels using the directional information and the pixels with intensity greater than \(T_1\) are marked as genuine or strong edges. The low threshold value is used to allow the detection of faint parts of the continuing edges. The pixels with intensity value less than \(T_1\) and greater than \(T_2\) are marked as edges, if they are already connected with edge pixels.
Image border line detection: The image border lines were detected by identifying the lines that fit the edge pixels using the Hough transform method (Duda and Hart, 1972). In the Hough space, lines are represented by a pair of parameters, namely rho ($r$) and theta ($\theta$) (Figure 3.4). All potential lines that may go through a single point ($x, y$) can be represented by the following equation.

$$r = x \cos \theta + y \sin \theta$$  \hspace{1cm} (3.3)

![Figure 3.4: Hough transform parameters.](image)

The Hough transform uses an accumulator array where an array element at ($i, j$) stores the number of edge pixels along a line defined by $r = i$ and $\theta = j$. The dominant lines in the image are represented by the peaks within the accumulator array as a higher number of edge pixels are counted for dominant lines. The peaks were identified by thresholding the accumulator array and line parameters $r$ and $\theta$ corresponding to the peaks were calculated. The results of this process were the four border lines of the displayed image as shown in blue lines in Figure 3.3d. The corners of the displayed image were obtained by calculating the intersection points of border lines (Figure 3.3e).

Coordinate Transform: The final step was to transform the eye gaze coordinates on to the displayed image coordinates. This was achieved by using a 2D homography, which transforms a point’s coordinates from one image plane to another. A 2D homography matrix $H$ is defined as,

$$x = Hx'$$  \hspace{1cm} (3.4)

where $x$ and $x'$ are the homogeneous coordinates of a point in two different image planes. The calculation of the homography matrix ($H$) requires at least four pairs of coordinates in two different image planes (Hartley and Zisserman, 2003). The normalised direct linear transformation algorithm given by Hartley and Zisserman...
(2003, p. 109) was used to calculate a homography matrix for each FoV video frame using four image corner coordinates. Finally, the eye gaze coordinates on the displayed image were calculated using the homography matrix (Figure 3.3f). MATLAB\textsuperscript{1} implementations of the Hough transform and 2D homography calculations were obtained from Kovesi (2000).

### 3.4 EEG data acquisition and pre-processing

The EEG signals were captured from 36 scalp sites (extended international 10 - 20 system). Ag/AgCl electrodes mounted on a cap (Quik cap from Compumedic\textsuperscript{2}) were used in the experiments and electrode impedances were kept below 10 kΩ for all participants. These signals were amplified using the NuAmps\textsuperscript{3} amplifier and recorded in the SCAN Express\textsuperscript{4} software at a sampling rate of 1000 Hz with earlobe electrodes average as the voltage reference \((A1 + A2)/2\). Image presentation onset markers were sent to the amplifier using E-Prime\textsuperscript{5}.

![Figure 3.5: High-pass filtered (at 0.5 Hz) EEG signals of 36 electrodes. Image presentations are represented by the vertical lines.](image)

Recorded signals were high-pass filtered at 0.5 Hz to remove the drift in the EEG (Figure 3.5) and then low-pass filtered at 30 Hz to remove unwanted background EEG activity (Figure 3.6). Next, the ocular artefacts (artefacts which are mainly caused by eye blinks and eye movements) present in the EEG were minimised by applying blind source separation (BSS) using independent component analysis (ICA),

\begin{itemize}
  \item \textsuperscript{1}MATLAB, http://www.mathworks.com.au
  \item \textsuperscript{2}40 Channel Quik-Cap, http://www.neuroscan.com
  \item \textsuperscript{3}NuAmps, http://www.neuroscan.com/nuamps.cfm
  \item \textsuperscript{4}SCAN Express, http://www.neuroscan.com
  \item \textsuperscript{5}E-Prime, http://www.pstnet.com/eprime.cfm
\end{itemize}
which is a standard technique used for the removal of ocular artefacts present in
the EEG signals (Flexer et al., 2005; Joyce et al., 2004; Jung et al., 2000a; Vigário,
1997).

The ICA is a computational method to find a linear representation of compo-
nents by assuming that the components are non-Gaussian signals and statistically
independent (Hyvärinen and Oja, 2000). The ICA is widely used to separate ocular
artefacts from EEG data. Makeig et al. (1996a) first applied ICA to EEG. The
capability of ICA in separating artefactual and neurally generated EEG sources is
demonstrated by Makeig et al. (1999) and Jung et al. (2000a).

The ICA is efficient when: (i) the sources are temporally independent; (ii) the
sensor inputs are a linear combination of the source signals; and (iii) the number
of sources is equal to the number of sensors (Jung et al., 2000a). The EEG signals are
not phase-locked and volume conduction is considered to be linear, thus EEG data
satisfies the first two assumptions (Jung et al., 2000a). Even though the number
of independent sources in EEG is not known, Makeig et al. (1996b) demonstrated
the effectiveness of ICA in separating relatively large and temporally independent
sources from a large number of low-level and temporally independent sources using
synthetic EEG signals.

The basic ICA model is defined as \( x = As \), where \( s \) represents the \( m \) indepen-
dent source signals \( \{s_1, s_2, s_3, \ldots, s_m\} \), \( x \) represents the \( m \) different recorded mixtures
\( \{x_1, x_2, x_3, \ldots, x_m\} \), and \( A \) is the unknown mixing matrix. In this model, neither
the source signals nor the mixing process is known. The solution to this problem
can be expressed as \( s = Wx \) and the task is to find the \( m \times m \) demixing matrix
\( W \). There are different algorithms used to decompose the source signals from the
mixture (Hyvärinen, 1999; Hyvärinen and Oja, 2000). One of these is the infomax
3.4. EEG data acquisition and pre-processing

Figure 3.7: Scalp distribution of the ICA decomposed sources obtained from a representative participant - brain activity increases as colour changes from blue to red as shown by the colour bar. Source 1 indicates high frontal activity.

method (Bell and Sejnowski, 1995), which separates the source signals from mixtures by maximising the joint entropy of the signals. In this study, the infomax ICA algorithm implemented in EEGLAB (Delorme and Makeig, 2004) was used to separate 36 independent components from the filtered 36 channel signals. The scalp distribution of each of the source signals are shown in Figure 3.7. In these data the first component shows high frontal activity (Figure 3.7) which also resembles the characteristics of ocular artefacts, in particular eye blink artefacts (Figure 3.8). This component was removed and the remaining components were back projected onto channel signals to obtain ocular artefact-minimised EEG signals (Figure 3.9). This is achieved by calculating $x'$ using $x' = (W)^{-1}s'$, where $s'$ is the new source signals obtained by making the rows representing ocular artefact in matrix $s$ as zero. This process was repeated for each participant to minimise the ocular artefacts present in the EEG signals.
Chapter 3. Experimental data capture and pre-processing

3.5 Summary

This chapter describes the experimental setup, exercise tasks, and ETS and EEG data capture and pre-processing. The participants’ eye gaze movements were captured using an ETS, which records the eye gaze coordinates with respect to the FoV camera frame. An algorithm was developed to transform these eye gaze locations onto the displayed image. This algorithm firstly corrects the lens distortion introduced by the FoV camera and identifies coordinates of the displayed image corners using Canny edge detection and the Hough transform. It then transforms the eye
gaze coordinates on to the displayed image by applying a 2D homography. These transformed eye gaze data were used to calculate different eye gaze measures and these measures were analysed for the identification of effective data interpretation practices in Chapter 4 and for the identification of the interpreter visual attention biases in Chapter 5. The brain responses were captured using EEG technique. The background brain activity artefacts were minimised by filtering the EEG signals and ocular artefacts were minimised by applying BSS using ICA. Artefact-minimised EEG signals were further analysed to detect the P300 responses associated with geoscientific target spotting in Chapter 6.

Chapter 4

Identifying effective interpretation methods for magnetic data by profiling and analysing human-data interactions

Abstract

Geoscientific data interpretation is a highly subjective and complex task because human intuition and biases play a significant role. Based on these interpretations, however, the mining and petroleum industries make decisions with paramount financial and environmental implications. To improve the accuracy and efficacy of these interpretations, it is important to better understand the interpretation process and the impact of different interpretation techniques, including data processing and display methods. As a first step towards this goal, we aim to quantitatively analyse the variability in geophysical data interpretation between and within individuals. We carried out an experiment to analyse how individuals interact with magnetic data during the process of identifying prescribed targets. Participants performed two target spotting exercises where the same magnetic image was presented at different orientations. The task was to identify the magnetic response from porphyry-style intrusive systems. The experiment involved analysing the data observation pattern during the interpretation process using an ETS that captures the interpreter’s eye gaze motion and the target spotting performance. The time at which targets were identified was also recorded. Fourteen participants with varying degrees of experience and expertise participated in this study. The results show inconsistencies within and between the interpreters in target spotting performance. The results show a correlation between a systematic data observation pattern and target spotting performance. Improved target spotting performance was obtained when the magnetic image was observed from multiple orientations. These findings will help to identify and quantify the effective interpretation practices, which can provide a roadmap for the training of geoscientific data interpreters and contribute towards
the understanding of the uncertainties in the data interpretation process.

4.1 Introduction

An understanding of the most effective and efficient ways to interpret geophysical datasets is important because decisions with significant economic, environmental, and also social implications are made on the basis of the interpretations. Geoscientific data interpretation is based on pattern recognition. The interpreter looks for patterns in the data that are compatible with their expectations of the results of geological processes and how these appear in geophysical data. For example, an anastomosing pattern of lineaments in aeromagnetic imagery or a down-converging pattern of reflector offsets in seismic data may be interpreted as having geometric characteristics consistent with a flower structure and a 'wrench-tectonic' geological environment.

Many types of geophysical data are presented as images; for example, seismic reflection sections and maps of variations in the intensity of potential fields. There is no universally accepted method for interpreting geophysical imagery. The interpretation outcomes are heavily dependent on the skills and experiences of interpreters, who inevitably brings their personal biases to the process. Interpretational skills are often acquired by informal interactions with experienced colleagues, by attending short courses, through studying the literature, or by trial and error.

In this paper, we describe a study that captures, analyses, and seeks to understand how individuals approach the qualitative interpretation of aeromagnetic data. The data used in our experiments is an aeromagnetic dataset from an area of known porphyry-style mineralisation and the exercise, which acts as a vehicle to understand how different individuals approach its analysis, involves identification of magnetic responses indicative of prospective geology. Data observation patterns are tracked in real time using an ETS during the interpretation process. Then, together with interpreter feedback, we assess the accuracy and efficiency of geological target detection within magnetic data for individual interpreters. Although the use of an ETS is new to geoscientific research, it is widely used in studies of HCI (Goldberg and Kotval, 1999; Prendinger et al., 2007) and automated classification of, for example, video data acquired for medical purposes (Vilarino et al., 2007).
4.1.1 Previous analysis of the interpretation of geoscientific data

Understanding of the earth’s subsurface based on the interpretation of geoscientific data is a challenging task. The complex natural environment needs to be predicted based on multiple datasets (geophysical, geological, and geochemical) each with its own characteristics and limitations. The data are often ambiguous, incomplete, inaccurate, and of low resolution (Frodeman, 1995). There are a limited number of published studies on how geoscientists interpret their data. The approaches used and desired outcomes are extremely variable. Variability or human biases in interpretations have been researched in the areas of oil and gas exploration (Bratvold et al., 2002) and seismic data interpretation for petroleum related studies (Bond et al., 2007; Rankey and Mitchell, 2003). Rankey and Mitchell (2003) carried out a study designed to qualitatively and quantitatively analyse the impact of interpretation and the uncertainties associated with it on seismic attribute analysis for the prediction of reservoir properties. They used six different interpreters for this study, and among these, two were experts and the rest were intermediate. Another study by Bond et al. (2007) analysed the results of the interpretation of some seismic data. In their study, a synthetic seismic image was interpreted by 412 participants with varying levels of experience and training. Their results showed that only 23% of the participants successfully identified the three major faults present in the image and 21% identified the correct tectonic setting. Rankey and Mitchell (2003) found that seismic interpretations are influenced by the interpreter’s biases based on previous experience, preconceived notions, types of data available, data quality, and geological understanding, whereas Bond et al. (2007) claimed that prior knowledge had a greater influence.

In an alternative approach, Welland et al. (2006) studied how the type of colour display impacts the interpretation of seismic data. They emphasise the non-linear nature of human colour perception, that is, equal amounts of change in blue and yellow tones are not perceived as equal by the viewer. To reduce the interpreter bias in interpretations, they proposed a psychological colour space to code the seismic data.

In a study specifically focusing on mineral exploration, Wastell et al. (2011) studied decision uncertainties. They carried out an experiment involving 94 individuals from mineral exploration companies. The participants were asked to identify the chances of either finding a deposit that can be mined or the existence of the ore of a particular metal, estimated purchase price, and to propose what additional exploration data was required to increase confidence in their decision. They reported that variability in mineral exploration decision making is due to human predispositions,
such as rational thinking and cognitive closure.

Unlike the studies summarised above, our research focuses on a more fundamental problem. We seek to understand how interpreters observe the data while interpreting and its impact on the effectiveness and efficiency of the interpretation. We monitor and profile the eye gaze of an interpreter while they are interpreting using an ETS. Our preliminary studies (Chadwick et al., 2010; Sivarajah et al., 2012a) demonstrated the feasibility of capturing eye gaze data to effectively monitor and analyse the data observation patterns during geoscientific target spotting exercises on magnetic and seismic datasets.

4.2 Experiment and data processing

We present a study that compares how different individuals interact with the same magnetic dataset and how this is related to their effectiveness in identifying exploration targets. The experiment involves exercises that require participants to recognise targets that have characteristics suggestive of gold-copper-rich porphyry systems. The relevant magnetic anomalies have a distinct ‘Mexican hat’ like character comprising sub-circular magnetic highs with surrounding annular lows (Holden et al., 2011; Hoschke, 2011) (Figure 4.1). Data observation patterns are captured by the ETS along with user feedback on the targets recognised. Based on the results, we quantitatively analyse the impact of such variations as displaying the data in different orientations and how the observation strategies of different interpreters affect individual and overall target identification performance.

Figure 4.1: Typical magnetic response of a porphyry-style mineralisation. The mineralised one appears as an elevated sub-circular feature with surrounding annular lows (Courtesy of Barrick Gold of Australia Ltd).

4.2.1 Capturing eye gaze motion using the ETS

We used a mobile ETS available from Applied Science Laboratories. The tracker uses two video cameras and three IR LEDs, which are mounted on a pair of standard safety glasses (Figure 4.2a). There is a circular cut-out in the right lenses of the
glasses. This cut-out allows for the placement of an adjustable monocle that reflects the IR light beam from the LEDs, which are arranged in a triangular pattern, onto the eye surface. Eye gaze is determined by comparing the reflected IR light from cornea and the pupil, which are captured by the first camera. A forward facing second camera records the interpreter's FoV.

![Experiment participant wearing ETS glasses and (b) the experiment setup.](image)

Figure 4.2: (a) Experiment participant wearing ETS glasses and (b) the experiment setup. Note that the skull cap allows EEG data to be captured during the experiment, but the results are not described in this paper.

Figure 4.2 shows the ETS glasses and the experiment setup. At the beginning of each experiment, the participants were seated in front of a display monitor at a convenient distance (from 60 – 100 cm) and were then fitted with the ETS glasses. The ETS needs to be calibrated for every participant; this process is required to enable accurate calculation of the eye gaze coordinates. This calibration is achieved by requesting the participants to fix their gaze on known locations and marking those points on the FoV video frame. We used 13 points to cover the entire monitor screen. The accuracy of the eye coordinates was revalidated between the exercises.

The ETS follows the interpreter's eye movement in real time to calculate the locations of focus with respect to the FoV camera video frames (Figure 4.3a). These eye gaze locations should be transformed to the corresponding coordinates with respect to the displayed data region. This process requires the analysis of each video frame to identify the location of the displayed data region (that is, the corners of the monitor screen) with respect to the FoV video frame and transforming the eye gaze locations. We developed a robust algorithm to estimate these locations using image processing techniques which involves the following steps. Firstly, it performs camera lens distortion correction to transform the FoV image that has barrel distortion as shown in Figure 4.3a into a perfect perspective projection image.
as shown in Figure 4.3b. Secondly, edge detection is performed to find the boundaries of the data region within the FoV image as shown in Figure 4.3c. Thirdly, it identifies the four corners of the rectangular data region as shown in Figure 4.3d. Finally, the eye gaze coordinates, which were previously with respect to the FoV image are transformed into the coordinates with respect to the displayed data region using a 2D homography as shown in Figure 4.3e.

![Figure 4.3: (a) The FoV frame from the ETS with the red cross indicating the location of the eye gaze; (b) Lens distortion corrected image; (c) Detected edges of the monitor screen on the FoV frame; (d) Calculated corners represented by four red crosses; (e) Computed eye gaze location (red cross) on the displayed image.](image)

4.2.2 Test dataset

The ground magnetic data used in the experiment were collected with a line spacing of 100 m and gridded at a 25 m cell size. The data have been upward continued to 50 m to suppress short wavelength responses, which are mostly noise originating in the near surface. Finally, the data was reduced to the pole to, ideally, give symmetrical responses centered on deposits. The location of the study area has been withheld due to agreed commercial confidentiality.

Quantitative measurement of the target spotting performance requires ground truth data. In practice, this is not fully achievable because one cannot definitely recognise the entire set of true targets in the image without a prohibitive amount of drilling in the survey area. The survey used in this experiment is over a very well explored area that has known deposits that could be used as the target set. However, the number of targets is still limited and they do not include many of the magnetic anomalies that have the desired qualities. Although recognition of what are actual targets is appealing, the experiment is ultimately one of pattern recognition. Thus, for the set of targets for our analysis, we decided to also use the targets generated by a pattern recognition algorithm designed to identify the magnetic responses of porphyry systems (Holden et al., 2011), which is known as the CET Porphyry Analysis Extension for Geosoft Oasis Montaj\textsuperscript{TM}. The set of true targets used for training the algorithm is referred to as the CET Porphyry Data Set.

\footnote{Centre for Exploration Targeting (CET) Porphyry Analysis Extension for Geosoft Oasis Montaj\textsuperscript{TM}, http://www.geosoft.com/products/software-extensions/cet-porphyry-detection}
targets used to judge the accuracy of the participant’s interpretation is a combination of those derived by either of these means. These 42 targets are shown in Figure 4.4.

Figure 4.4: Displayed magnetic image with the true targets shown by black circles.

4.2.3 Interpreter tasks

Fourteen participants with varying levels of experience and expertise participated in this study. All of our participants were trained geophysicists or geologists with experience in magnetic data interpretation. All had normal or corrected to normal vision. The experiment consisted of two exercises, both of which use the same magnetic image, but displayed it in different orientations. The intent was to understand the impact on interpretation of viewing the data in different orientations. For the first exercise, the image was displayed in a normal fashion, i.e. with north towards the top of the page (Figure 4.5a), which will be referred to as the original image. For the second exercise, the image was rotated by 180° (Figure 4.5b), which will be referred to as the rotated image. Both images were illuminated with a false sun located as if higher up the page, i.e. from actual north in the original image and actual south in the rotated image. This is because of the well-known problem that illumination from below the image can cause apparent inversion of the topography in an image which includes false-sun illumination. The disadvantage of this approach is that the two images are slightly different. Participants were given three minutes for each exercise and had a 30-minute break, during which they were distracted with other tasks. Participants 10 to 13 saw the rotated image first and all the other participants saw the original image first.

Written instructions were displayed on the monitor screen at the beginning of each exercise. During the exercise, the participants were asked to press a keyboard button as soon as they identified a target, while fixing their eye gaze at the target
4.2. Experiment and data processing

Figure 4.5: Magnetic image used for (a) the original image exercise and (b) the rotated image exercise.

location. We captured the data eye gaze movements during the target spotting task using the ETS. We also recorded the time of the target identifications by capturing the button click times. Following the experiment, all participants were asked to rank themselves (from 1 to 10) in their level of expertise for this task.

4.2.4 Characterisation of interpreter performance

Two aspects of the interpreter’s interaction with the data were assessed: (i) success rate, which is based on comparison of the user-identified targets and the ground truth targets; and (ii) data observation patterns, which provides information about the individual’s approach to target identification.

Success rate

Data were analysed based on the participants’ analysis of the original image, the rotated image, and the combination of the two. Data recorded include porphyries correctly selected as targets (true positives) and also the incorrect identifications (false positives). Note that, in some cases, participants selected the same target multiple times when viewing an image, but these repeat identifications were not included in the analysis. When the data from the analysis of the original and rotated images were combined, repeat identifications were also not included. Based on this information, we quantified target identification performance by calculating ‘Recall’ $R$ and ‘Precision’ $P$ defined as follows,

\[ R = \frac{N_{ik}}{N_k} \]
\[ P = \frac{N_{ik}}{N_s} \]
where $N_k$ is the number of true targets in the image; $N_k = 42$ (Figure 4.4). Here, $N_{ik}$ is the number of targets correctly identified by the participant, and $N_s$ is the total number of targets identified (without duplication) by the participant. The ratio between the identified false targets and all the targets identified (without duplication) by the participant is represented by $1 - P$.

**Observation patterns**

We used four different eye gaze-based measures to quantify the data observation patterns. These are the scanpath length, scanpath duration, number of saccadic eye movements, and the number of fixations.

Scanpath length is the total distance traced by the eye during the exercise (measured in pixels). We also computed the mean scanpath length between target identifications. An equivalent temporal measure is scanpath duration between target identifications. These parameters are indications of the efficiency of the data analysis (Goldberg and Kotval, 1999).

Fixation is defined as maintaining the eye gaze on a particular location for at least $100 - 150$ ms (Viviani, 1990). The eyes typically fixate on locations that an individual finds to be surprising, salient, or significant (Loftus and Mackworth, 1978). Saccades represent the rapid movement of the gaze point from one location to another. By definition, the number of saccades is equal to the number of fixations minus one. Put simply, eye gaze can be thought of as comprising a series of fixations separated by saccades. There are various ways to define what constitutes a fixation in terms of spatial variation about a specific location. We used the algorithm described in Goldberg and Kotval (1999), which identifies the fixations by detecting continuous observations at a particular location (within 40 pixels radius) for at least 166 ms. During fixation, the human visual system is processing the visual information, observing the peripheral areas of the visual field and planning the next saccade based on what is present in the peripheral areas (Goldberg and Kotval, 1999).

We calculated the number of saccades between the target selections (button clicks) for each participant and obtained the average number of saccades between the target selections. We also computed the number of fixations and their spatial distribution, which represent the magnetic features the participant considered as possible targets. We divided the image into four quadrants and counted the number of fixations per quadrant for each participant and obtained the average number of fixations in each quadrant by all the participants for the original and rotated image exercises.
4.3 Results and discussion

Using the individual target spotting performances and the eye gaze profile data, we analysed the individual variability in three different aspects: target spotting performance, data observation patterns, and impact of data observation from multiple orientations.

4.3.1 Variability in target spotting performance

Figure 4.6 shows the target spotting performance of all 14 participants for the original and rotated image exercises and both exercises together. Recall ranges from 0.17 to 0.67 with an average of 0.36 for the original image exercise and from 0.10 to 0.43 with an average of 0.26 for the rotated image exercise (Figure 4.6a). There is significant variation between the results from the different exercises as completed by each individual and between the performance of individual interpreters. Interestingly, the level of self-assessed expertise did not correlate with participants’ performance in terms of target detection accuracy (correlation coefficients between the ranking based on their self-assessments and the rankings based on the target spotting performances in original, rotated, and combined exercises are 0.14, 0.44, and 0.19, respectively). Analysing the impact of past experience in the accuracy and efficiency of data interpretation is an intriguing area of research, but it requires a large number of participants and is considered to be beyond the scope of this study. Here, we use the outcomes of the target identification exercise as a measure of expertise, referring to them as high achievers and low achievers.

Figure 4.6: Target spotting performances of fourteen participants in original, rotated, and combined image exercises based on their target selection and ground truth data. (a) Recall versus participant number; R = 1 means identification of all true targets. (b) Precision versus participant number; P = 1 means all identified targets by the participant are true targets.
Figure 4.7: Eye-trace plots for four high achievers: (a) Participant 12, (c) Participant 10, (e) Participant 6, and (g) Participant 11 — more systematic observation and for
4.3. Results and discussion

four low achievers: (b) Participant 13, (d) Participant 14, (f) Participant 8, and (h) Participant 2 — more random observation. The numbers indicate the button clicks in their order (zero indicates the initial eye gaze location at the beginning of the experiment). Black numbers on white background represent the true positives, while the white numbers on the black background represents the false positives. Different colour traces indicate the eye gaze path in between consecutive button clicks (the colour changes from blue to red as shown in the vertical colour bar with the button click times). The horizontal colour bars at the bottom of each image indicate the portion of the experiment time spent on target selection. The total length of the bar (green and red) represents the experiment time, and the green region represents the amount of time spent on the target selections (end of green region indicates the final button click time).

4.3.2 Impact of data observation from different orientations

Most of the participants (ten participants) performed better in the original image exercise than the rotated image exercise (with an average recall difference of 92%), regardless of the order in which the exercises were performed (participant numbers 10 to 13 performed the rotated image exercise first). Four participants obtained better performance in the rotated image exercise (with an average recall difference of 22%). This shows the variability in target spotting performance based on the orientation of the data within each participant. Even though the participants saw the same targets in the first and second exercises, the second (rotated image) exercise resulted in lower success rates. Which image was seen first did not affect this result. This demonstrates that working with the data in the first exercise did not help them to perform better in the later exercise. This is probably because the exercise is too short for the gaining of significant advantage from the earlier assessment. It is hard to explain why performance was better with the original image. The most likely explanation is that it is related to the different (relative to the data) illumination direction. This is difficult to correct because many people visualise apparent topography inverted according to whether a dataset is illuminated from the north or from the south. It was for this reason that a consistent (relative to display) illumination direction was used. The important outcome from this component of the experiment is that orientation and/or direction of illumination does make a difference. Determining how and why will be the subject of a further study.

Even though most of the participants performed better with the original image, there are sometimes missed targets, which were identified in the rotated image. This is represented by the performance improvement obtained in the combined analysis (Figure 4.6a). Ten participants obtained higher recall in the combined case than the original image exercise, while others had no improvements (participant number
4, 6, 11, and 14); however, in comparison to the rotated image exercise all of the participants had an improvement in the combined task. Average increase in the recall rate for the combined viewing over the original image exercise is about 20% and over the rotated image exercise is 81%. Precision calculations indicate participants had relatively high precision (greater than 75%) in the original and rotated image exercises (Figure 4.6b). Comparable precision rates and the overall improvement in the recall rates for the combined exercise quantitatively shows that the geoscientific data interpretation performance could be improved by practising to observe the data from multiple orientations during interpretation. Geoscientists traditionally interpret data from different orientations by rotating the maps because printed maps were used for interpretation in the early days. Nowadays, most interpretations are done within some kind of spatial analysis software package. Our results demonstrate the need for a functionality to flexibly view data in different orientations.

4.3.3 Variability in data observation patterns

Data observation patterns differ significantly among participants. Figure 4.7 shows examples of two different observation patterns. In these figures, differently coloured eye gaze traces represent the data observation between any two target selections (button clicks). The data in Figure 4.7 show the eye traces of high-achieving interpreters and low-achieving interpreters. The data observation patterns of low achievers are far more complicated and have more random eye gaze motions. The data in Figure 4.7 show the high achievers have selected the targets more systematically, moving from a target to another nearby. In contrast, the low achievers employed an inefficient method of scanning much larger areas of the image in search of the next target. Figure 4.8 shows the data observation heat maps, which are generated based on the total amount of time spent on each location during the target spotting exercise. Heat maps show that the high achiever has visited the entire data region focusing on different porphyry-style features, whereas the low achiever has spent more time on certain specific features. This difference was true in a general sense for all participants, although not universal.

We computed the Pearson product-moment correlation to determine the relationship between different eye gaze measures and the target spotting performance (recall). The eye gaze measures used for this analysis are the mean scanpath length, mean scanpath duration, and mean number of saccades. Statistically significant strong negative correlation was obtained for the original image exercise and rotated image exercise between these eye gaze measures and the recall (Table 4.1).

Strong negative correlation between the mean scanpath length and the recall indicates that performance increases as the mean scanpath length decreases, show-
ing that high achievers observed the data more efficiently than low achievers. The negative correlation between the target spotting performance and the scanpath duration indicates that the difficulty in decision making increases with the decrease in target spotting performance, with high achievers making the decisions more quickly than low achievers. In addition, the increase in the number of saccades between the target selections as the performance declines shows that low achievers had to search harder to identify the targets than high achievers. Overall, these results quantitatively show a strong correlation between target spotting performance and efficient data observation patterns.

Table 4.1: Pearson product-moment correlation (r) between different eye gaze measures and the target spotting performance (recall) for the original image and rotated image exercises.

<table>
<thead>
<tr>
<th></th>
<th>Original image exercise (r)</th>
<th>Rotated image exercise (r)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean scanpath length</td>
<td>-0.675*</td>
<td>-0.800*</td>
</tr>
<tr>
<td>versus recall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean scanpath duration</td>
<td>-0.692*</td>
<td>-0.748*</td>
</tr>
<tr>
<td>versus recall</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean number of saccades</td>
<td>-0.649**</td>
<td>-0.723*</td>
</tr>
<tr>
<td>versus recall</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significant at 0.01 (*) and 0.05 (**) levels; n = 14.

Figure 4.8: Heat map plots obtained (a) for a high achiever and (b) for a low achiever, where the number of visitations increases from blue to red.
4.3.4 Target-driven analysis

We divided the ground truth porphyries into three categories based on the number of participants that selected them as targets. If a ground truth porphyry is selected by many participants it is classed as an easy target. If it is selected by very few participants it is considered to be a difficult target. Features that were selected by none of the participants were categorised as very difficult. Figure 4.9 shows easy targets (in green circles), difficult targets (in yellow circles), and very difficult targets (in red circles) for the original and rotated image exercises.

![Figure 4.9: The images show the number of participants who identified each ground truth target (a) for original image exercise and (b) for the rotated image exercise. Green circles indicate the easy targets (porphyries selected as targets by more than seven participants), yellow circles indicate the difficult targets (porphyries selected as targets by less than seven participants, but at least by one participant), and the red circles indicate very difficult targets (porphyries selected as targets by none of the participants). The blue triangles indicate the porphyries, where more than 50% of the participants fixated, but failed to select them as targets.](image)

We identified the locations of fixations during the exercises for each participant. Figure 4.10 shows the locations of the fixations by one participant for the original image exercise. Fixation information can explain whether the participant who failed to select the difficult and very difficult targets have actually observed them. Therefore, we calculated the percentage of participants who actually fixated at those locations. Missed targets where more than 50% of the participants fixated are indicated by blue triangles in Figure 4.9. Most of these are in regions where the targets are close together and the magnetic signature is confused by interference. There are some targets, which were not visited (fixated) by at least half of the participants, and this is comparatively more in the rotated image exercise than the original image exercise. Missing a target by not visiting them could be minimised by systematically observing the data. Our ongoing study focuses on finding the relationship between...
geological target detection and an image saliency model based on colour, intensity, and orientation (Sivarajah et al., 2012b).

Figure 4.10: Small red dots indicate different fixations the participant had on the displayed image during the original image exercise.

To quantify the target-driven data observation, we analysed the level of interest each participant had in different parts of the image. We divided the image into four quadrants (Figure 4.11) and calculated the number of fixations each participant had in each quadrant by computing the average number of fixations by all participants in each quadrant for the original and rotated image exercises (Figure 4.12). The number of fixations in each quadrant indicates the level of interest of that specific region in the data. Quadrant 2 in Figure 4.11a is the area with the most number of ground truth targets, and the results from our original image exercise shows a higher number of fixations than other quadrants, as expected. However, in the rotated image exercise, Quadrant 2 in Figure 4.11b did not have a higher number of
fixations than other quadrants. It shows that the same features appearing in different parts of the data seem to attract different levels of attention from interpreters, which affects their interpretation performance. Again, reducing this effect requires viewing the data from different orientations.

![Figure 4.12: Average number of fixations by all the participants on each quadrant for the original image and the rotated image exercises. Error bars indicate the standard deviations.](image)

**4.4 Conclusions**

In this paper, we present the findings from an ongoing study that aims to better understand the geoscientific interpretation process. Fourteen participants with varying levels of experience and expertise participated in the identification of gold-copper-rich porphyry systems within magnetic data presented as images. Target spotting performance shows significant variations between and within individuals based on the orientation of the image, irrespective of the image presentation sequence. The data observation patterns obtained using the ETS show significant variation among interpreters. The results show a correlation between a systematic data observation pattern and target spotting performance. Target spotting performance improves when data are observed from two different orientations. These findings may provide practical guidance on how to train geoscientific interpreters to improve their performance and accuracy by effective interpretation techniques, such as observing data from multiple orientations and observing the data systematically.

The most intriguing aspect of our results is the different performances as a function of data orientation. Explaining this phenomenon requires further research. Our data suggest that the order the images are presented is not the reason. There was no obvious change in observation behaviour between the first and second viewing, and when questioned, most participants did not even recognise that they had
observed the same dataset twice. As noted above, the two images were slightly different because of the preference for sun-illumination from the top of the display, but this makes little apparent difference to the appearance of the targets. Specific experiments are required to understand this intriguing result.

Our ongoing and future research focuses on combining the detection of the brain responses associated with the target spotting with the eye gaze tracking to better understand the interpretation process. Our previous research has successfully shown that target identification can be reliably correlated with brain responses. Such a combination of data might provide insights as to why the participants who fixated at some locations subsequently did not select them as targets. The combination of the brain responses with the eye tracking will not only provide invaluable information for the interpreters, but it will also provide feedback for the software and hardware developers to identify efficient display methods.
Interpretation of gravity and magnetic data for exploration applications is primarily based on pattern recognition, where geophysical signatures of geological features associated with localised characteristics, are sought within data. A crucial control on what comprises noticeable and comparable characteristics in a dataset is how images displaying those data are enhanced. Interpreters are provided with various image enhancement and display tools to assist their interpretation, though the effectiveness of these visualisation tools to improve geological feature detection is difficult to measure. We address this challenge by analysing how alternative visualisation tools impact on the interpreter’s visual attention when interpreting the data, as features which are more salient to the human visual system are more likely to be noticed. This study uses geological target spotting exercises within images generated from magnetic data to assess commonly used magnetic data visualisation methods for their visual saliency. Our aim is achieved in two stages. In the first stage, we identify a suitable saliency detection algorithm that can computationally predict visual attention of magnetic data interpreters. The computer vision community has developed various image saliency detection algorithms, and in this study we assess which algorithm best matches the interpreter’s data observation patterns for magnetic target spotting exercises. In the second stage, we apply this saliency detection algorithm to understand potential visual biases for commonly used magnetic data enhancement methods. This study provides a guide to choosing image enhancement methods, based on saliency maps that minimise unintended visual biases in magnetic data interpretation and some recommendations for identifying exploration targets in different types of magnetic data.
5.1 Introduction

Interpretation of magnetic or any other geoscientific data is primarily based on a pattern recognition process where geological features are sought within data and their spatial associations analysed. It is common practice to process or enhance the data prior to display to bring out characteristics perceived to be useful to the interpreter. The combination of high- and/or low-pass filtering, colour contour mapping, and sun-angle shading is widely used by the potential field data interpreters. Conventionally, interpreters select different data enhancement methods (Blakely, 1995) based on their prior knowledge of these methods or by trial and error to enhance specific features of interest or data characteristics. In practice, it is common that interpreters use multiple enhancement methods, for example using both high- and low-pass filtering to bring out anomalies associated with causative sources at different depths, or using multiple different high-pass filters to find discontinuities within data. In addition, images are also visualised using different colour display and shading methods.

Welland et al. (2006) reported the impact of human visual perception of colours on seismic data interpretation. Even though the findings are not described in detail, their study is based on the nonlinear nature of human colour perception, where the same amount of change in different bands in the visual spectrum, such as yellow and blue in an image, are not perceived as the same change by the interpreter. To address this, they proposed a modified colour bar to compensate for visual bias in the interpretation of seismic data. For potential field data, we previously reported the impact of human-data interactions on geological target spotting (Sivarajah et al., 2013). This study showed that viewing of data in two different orientations and carrying out a systematic target search impact the target spotting performance. Evidently, how we view and interact with data plays a significant role in data interpretation.

In the fields of psychology and computer vision, there has been active research on understanding and emulating human visual attention. In our visual and other sensory systems, a key attention mechanism is saliency: a quality that makes certain items, such as objects, faces, and sounds, ‘stand out’ from their surroundings. Thus, visual saliency is typically associated with contrast from its neighbours, such as a bright object within a dark image background, and is called the ‘bottom-up’ influence. Visual saliency can also be influenced by memory or anticipatory mechanisms through training, for example, identifying your child’s face in a school group photograph or looking at moving cars when crossing the road. This is called the ‘top-down’ influence. In psychology, human attention has been modelled using both the
bottom-up and top-down influences including the learning of attention prioritisation using these influences (van de Laar et al., 1997). The computer vision community, on the other hand, focuses on emulating the bottom-up influence computationally using saliency detection algorithms (Borji and Itti, 2013; Harel et al., 2006; Itti et al., 1998). There are a large number of algorithms developed to identify image saliency. These are based on: (i) a biological model using spatial contrasts in colour, intensity, and orientation (Itti et al., 1998); (ii) purely computational approaches using frequency analysis (Achanta and Susstrunk, 2010; Achanta et al., 2008, 2009); or (iii) a combination of the two (Harel et al., 2006). Visual attention maps computed using these algorithms are called saliency maps.

In previous work, saliency maps have been used for various applications, such as scene classification (Siagian and Itti, 2007), text detection (Sun et al., 2010), object detection (Walther et al., 2002), visual search (Elazary and Itti, 2010), and automatic seam line detection for the merging of optical remote-sensing images (Yu et al., 2012). In another study, Su et al. (2004) investigated the possibility of using the inverted saliency model for display enhancement of natural images. We present a novel study of human attention based on saliency models for the task of analysing interpreter biases. We aimed to determine whether saliency maps can effectively represent interpreters’ visual attention for magnetic data, and then be used to understand potential biases in data observation when interpreting the data using different visualisation methods. This research was conducted in two stages.

In the first stage, we compared interpreters’ visual attention maps with saliency maps generated from three widely known saliency detection algorithms. The interpreters’ visual attention was determined by identifying eye gaze fixation locations. Fixation is defined as maintaining eye gaze at a particular location for at least 100 - 150 ms (Viviani, 1990). As visual attention moves to a new location, the eye gaze will try to follow (Deubel and Schneider, 1996) and typically fixate on locations that an individual finds to be surprising, salient, or significant (Loftus and Mackworth, 1978). In order to capture this information, we carried out a target spotting experiment and the interpreters’ eye gaze movements were acquired using an ETS. In this experiment, the task was to identify responses associated with porphyry-style mineralisation within magnetic data. Our preliminary studies (Chadwick et al., 2010; Sivarajah et al., 2012a) demonstrated the feasibility of capturing ETS data to effectively monitor and analyse the human-data interactions during target spotting exercises on magnetic and seismic datasets. For this study, we used two separate target spotting exercises. In the first exercise, we displayed small-scale images, each containing either a single target or background noise. In the second exercise, we displayed a large-scale image containing multiple targets. A set of saliency maps
were generated from the magnetic images using different saliency algorithms. These saliency maps were then compared with the eye tracking results and the saliency algorithm which generated the saliency maps that had the closest match to the interpreters’ data observation was identified. Previously, researchers have used a similar approach to demonstrate the correlation of saliency maps and interpreter-data-interactions using ETS for natural images (Harel et al., 2006; Li et al., 2013). However, such analysis has not been conducted to date for geoscientific data interpretation.

In the second stage, we applied the selected saliency algorithm to predict how widely used magnetic data enhancement methods will impact human visual attention during interpretation. The regions in the data likely to attract attention were highlighted using the selected saliency detection algorithm from the first stage, revealing potential visual biases in interpretation. This analysis was performed based on the assumption that a target can be more easily identified if it is located within a region that attracts interpreter attention than if it is located in a region that does not draw interpreter attention. We propose that this information can be used to guide the selection of enhancement methods to compensate for these biases by identifying the enhancement methods that produced dissimilar and complementary saliency maps. Potentially, such insight can also assist in the design of new data enhancement and filtering methods.

In this paper, Section 2 reports the experimental details, ETS data capture and processing, and the selection the most suitable saliency detection algorithm. Section 3 presents the analysis of the interpreter biases using the selected saliency detection algorithm to evaluate commonly used enhancement methods, and the limitations and applicability of the findings. Finally in Section 4, our conclusions and on-going research are discussed.

5.2 Interpreter visual attention vs image saliency

Our study analyses the effectiveness of saliency maps in predicting interpreters’ visual attention and then selects the most suitable saliency detection algorithm for magnetic data. The interpreter visual attention maps were captured through an experiment requiring participants to recognise targets that have characteristics suggestive of gold-copper-rich porphyry systems. The relevant magnetic anomalies have a distinctive ‘Mexican hat’ like character comprising sub-circular magnetic highs with surrounding annular lows (Holden et al., 2011; Hoschke, 2011) as shown in Figure 5.5a-1. All the interpreters who participated in this study were trained geophysicists or geologists with experience in magnetic data interpretation and had
normal or corrected-to-normal vision (i.e. using contact lenses).

![Figure 5.1: Magnetic data with multiple porphyry-style mineralisation and the arrow indicating the main porphyry belt. Note the ‘Mexican hat’ geometry of the anomalies. Data courtesy of Barrick Gold of Australia Ltd.](image)

The survey used in this experiment is over a mature exploration area that contains a number of known deposits confirmed by field drill tests. Figure 5.1 shows that the main porphyry belt runs from the top left corner to the bottom right (outlined) and consists of porphyritic intrusions with dacitic, granodioritic, quartz dioritic, and dioritic compositions. Volcaniclastic and pyroclastic breccias are present, along with shale-siltstone, sandstone, minor volcanic rocks, and mafic to intermediate dykes. The strong positive circular to elliptical magnetic responses correspond to porphyry-style deposits. Ground magnetic data were collected with a line spacing of 100 m and gridded with a 25 m cell size. The data have been upward continued to 50 m to suppress noisy short wavelength responses which mostly originate from the near surface. Finally, the data were reduced to the pole in order to give symmetrical responses and centre the anomaly peak over the centre of the porphyritic intrusions. The magnetic image was illuminated with a false sun located in the north side of the region covered by the data at an inclination of 45°.
5.2.1 Experiment setup

Participants were seated in front of a display monitor (52 cm x 33 cm) at a convenient distance (from 60 to 100 cm) and were then fitted with ETS glasses to capture their eye gaze movements (Figure 5.2). In order to maximise the participants' engagement with the target spotting task, they were requested to respond to targets by pressing a key on a keyboard, as soon as they spotted an anomaly likely to indicate a porphyry deposit (which we term ‘a porphyry’). Participants were requested to perform two different exercises during this experiment to capture the data observations patterns: spotting targets within small-scale images and within a large-scale image. Written instructions were displayed on the monitor at the beginning of each exercise.

For Exercise 1, the magnetic image was cropped to small images with porphyries, i.e. ‘target’, and without porphyries, i.e. ‘non-target’ images (Figure 5.5 - top row). These target images were obtained from the regions where known deposits are located and the non-target images were selected from regions where there were no significant porphyry-style anomalies. These target and non-target images were displayed in a rapid fashion to six participants for the target identification. In the visual display, target images and non-target images were displayed on the centre of the monitor (within 23.4 cm x 17.4 cm area). These images were shown in a random sequence, but the sequence was identical for all the participants. We displayed the images for 1000 ms with an inter-image interval of 1000 ms in which a blank screen was shown.

In Exercise 2, the magnetic image with multiple targets (Figure 5.6a) was displayed on the entire monitor for three minutes. Fourteen interpreters participated in this exercise and they were requested to identify as many porphyries as possible within that time.
5.2.2 ETS data acquisition and processing

This study used a mobile ETS available from the Applied Science Laboratories. The ETS uses two video cameras and three IR LEDs, which are mounted on a pair of standard safety glasses. There is a circular cut-out in the right ‘lens’ of the glasses. This cut-out allows for the placement of an adjustable monocle that reflects the IR light beam from the LEDs (which are arranged in a triangular pattern) onto the eye surface. Eye gaze is determined by comparing the reflected IR light from the cornea and the pupil, which are captured by the first camera. A forward facing second camera records the interpreter’s FoV. The ETS needs to be calibrated for every participant; in order to enable accurate calculation of individual eye gaze coordinates. Calibration is achieved by requesting the participants to fix their gaze on known locations and marking those points on the FoV video frame. We used 13 points to cover the entire monitor.

The ETS records the FoV camera video frames together with the locations of the eye gaze with respect to FoV. The eye gaze locations with respect to the displayed data region were calculated by analysing each video frame to identify the location of the displayed data region (the corners of the monitor) with respect to the FoV video frame. An image processing algorithm was developed to calculate eye gaze coordinates on the displayed data region.

The algorithm firstly corrects the known barrel distortion introduced by the camera lens (Figure 5.3a) and transforms the FoV image into a perfect perspective projection image (Figure 5.3b). It then identifies the boundaries of the data region using edge detection and the Hough transform (Figure 5.3c). Based on the boundaries, the algorithm calculates the four corners of the rectangular data region within the FoV (Figure 5.3d). Finally, it transforms the eye gaze coordinates from the FoV frame to the image/data frame (Figure 5.3e) using a 2D homography (Hartley and Zisserman, 2003). Figure 5.4a shows the data observation pattern generated as a track plot using the calculated eye gaze locations.

![Figure 5.3: (a) The FoV frame from the ETS with the red cross indicating the location of the eye gaze; (b) Lens distortion corrected image; (c) Detected edges of the monitor screen on the FoV frame; (d) Calculated corners represented by four red crosses; (e) Computed eye gaze location (red cross) on the displayed image.](image-url)
5.2. Interpreter visual attention vs image saliency

Figure 5.4: (a) Interpreter data observation pattern plotted on top of the displayed data as a track plot; (b) Blue dots on the displayed images indicate the fixation location calculated from the data observation data - RTP magnetic intensity changes from blue to pink as the value increases as shown by the colour bar; (c) Visual attention heat map - saliency increases as colour changes from blue to red as shown by the colour bar.

The interpreter eye gaze fixation locations were identified using the calculated eye gaze locations with respect to the displayed images (Figure 5.4b). There are various methods used to identify the fixation locations from the calculated eye gaze locations. We implemented the algorithm described in Goldberg and Kotval (1999), which identifies eye gaze fixations based on the continuous observation at a particular location (within 40 pixels radius area) for at least 100 ms. Based on the calculated fixation locations for each participant for each of the observed images, we generated the fixation maps by placing a Gaussian smoothed circle on the locations of the fixations (with a radius of 40 pixels). We then obtained the interpreter visual attention 'heat maps' for the magnetic images by pixel wise averaging these fixation maps across participants (Figure 5.4c). The average fixation map represents the accumulated fixations of all the participants.

5.2.3 Image saliency algorithms

The image saliency detection algorithms used in this study were selected to represent three different approaches in modelling human attention which are based on biological, purely computational approaches, or a combination of the two. We selected widely known image saliency algorithms in these categories, namely the visual attention model based algorithm by Itti et al. (1998) referred to here as ITTI, the hypercomplex Fourier transform (HFT) method (Li et al., 2013), and the graph based visual saliency (GBVS) approach (Harel et al., 2006). All of the leading saliency detection models are based on the following three steps (Harel et al., 2006):

1. Extraction of image features / feature maps.
2. Generation of activation maps / conspicuity maps.

3. Obtaining the saliency map through selection / normalisation.

The ITTI method adapts the saliency-based visual attention model of Koch and Ullman (1985). In this model, the visual input of the human vision system is first processed in parallel to generate a set of image features for different channels, such as colour, intensity, and orientation, across multiple spatial scales. The feature maps for these different channels are calculated based on the centre-surround differences, by taking the difference between the smaller and larger scale image features. These feature maps at different scales are then combined and normalised, resulting in three conspicuity maps representing colour, intensity, and orientation separately. Finally, these three conspicuity maps are combined using equal weights to generate the saliency map.

The HFT considers the saliency detection as a frequency domain problem and defines a concept of non-saliency using global information. In this method, feature maps are computed based on colour, intensity, and motion, but for static images they are based on colour and intensity. The amplitude spectrum, phase spectrum, and eigenaxis spectrum are computed from the feature maps. Spikes in the amplitude spectrum correspond to repeated patterns (non-salient regions) in the spatial domain. These repeated patterns are smoothed with Gaussian kernels to suppress the non-salient regions. The saliency map at each scale is derived using the smoothed amplitude spectrum, and original phase and eigenaxis spectrum. The final saliency map is selected by choosing the best scale with minimal saliency map entropy.

In the GBVS method, the image features are calculated using the ITTI method, but the activation and normalisation steps are implemented using a graph based approach. This is achieved by joining all the nodes (pixels) of the feature maps to generate a fully connected directed graph (Bang-Jensen and Gutin, 2008). The directed edges are assigned with a weight, which is proportional to the dissimilarity between the end nodes of the edges and to their closeness. A Markov chain (Norris, 1998) is defined on the graph to estimate the equilibrium distribution and an activation measure is obtained from pairwise contrast. The normalisation step is performed on these activation maps by using another Markovian process on the graph that is constructed from the activation map to generate the saliency map.

5.2.4 Selection of the most suitable saliency detection algorithm

We compare the saliency maps generated using the three algorithms with the participants’ visual attention maps to identify the saliency detection algorithm that can
suitably predict areas that will draw interpreters’ visual attention. From Exercise 1, we used eight target and eight non-target images for this analysis. The non-target images used for this analysis were selected based on the presence of some kind of magnetic anomaly within the displayed region. Note that some of the non-target images displayed during the exercise 1 did not have any anomalies. We generated the interpreter visual attention maps for these 16 images (maps obtained for four target and four non-target images are shown in Figure 5.5b) and for the large-scale magnetic image displayed in Exercise 2 (Figure 5.6b). The interpreter visual attention maps were thresholded to obtain binary images, which were used as the ground truth saliency regions for the identification of the most suitable saliency detection algorithm.

Figure 5.5: (a) 1 to 4 - Images shown with significant porphyry-style mineralisation (target images); (a) 5 to 8 - Images shown without any significant porphyry-style mineralisation (non-target images) - see Figure 5.4b colour bar; (b) Average fixation heat maps; (c) Saliency heat maps based on ITTI algorithm; (d) Saliency heat maps based on GBVS algorithm; (e) Saliency heat maps based on HFT algorithm. See Figure 5.4c saliency colour bar.

We generated saliency maps from the 16 magnetic images used for this analysis from Exercise 1 (Figure 5.5) and the magnetic image used in Exercise 2 (Figure 5.6) using ITTI, GBVS, and HFT saliency detection algorithms. The performance of these saliency detection algorithms in identifying the interpreter visual attention were calculated by analysing how well these saliency maps match with the ground truth saliency regions. This is achieved by obtaining the binary images of the saliency maps generated using the algorithms by varying the threshold (from 0 to 1) and comparing them pixel wise with the ground truth saliency regions. Based on this comparison, the true positive rate (TPR) and the false positive rate (FPR)
Figure 5.6: (a) Magnetic image displayed during Exercise 2 containing multiple targets - see Figure 5.4b colour bar; (b) Heat map generated using average participants’ fixations maps; (c) ITTI saliency heat map; (d) GBVS saliency heat map; (e) HFT saliency heat map. See Figure 5.4c saliency colour bar.

Figure 5.6: (a) Magnetic image displayed during Exercise 2 containing multiple targets - see Figure 5.4b colour bar; (b) Heat map generated using average participants’ fixations maps; (c) ITTI saliency heat map; (d) GBVS saliency heat map; (e) HFT saliency heat map. See Figure 5.4c saliency colour bar.

were calculated (Equation 1 and 2) by assuming the salient regions as ‘target’ and non-salient regions as ‘background’.

\[
\text{TPR} = \frac{TP}{P} \quad (5.1)
\]

\[
\text{FPR} = \frac{FP}{N} \quad (5.2)
\]

where \( TP \) – number of true positive pixels, \( FP \) – number of false positive pixels, \( P \) – number of target pixels, and \( N \) – number of background pixels.

Using the true positive and false positive rates at different thresholds, we plotted the receiver operating characteristics (ROC) curves (Fawcett, 2006). Even though the ROC curve is widely used in the field of machine learning to quantify the performances of different binary classification algorithms, it is also used to analyse the performances of the saliency detection algorithm by treating it as a binary classification problem (Harel et al., 2006). A ROC curve can be summarised by a single value by calculating the area under receiver operating characteristic (AUROC) curve. The AUROC value ranges from 0 to 1 and as the classifier performance increases the AUROC value also increases (Fawcett, 2006).

We calculated the AUROC performances of three saliency detection algorithms using the selected 16 images displayed in Exercise 1 and the results are shown in Figure 5.7. We also calculated overall performance of these algorithms in identifying the interpreter visual attention from small-scale images by obtaining the average AUROC (Table 5.1). These average AUROC values show all three algorithms perform similarly in identifying the interpreter visual attention maps of small-scale magnetic data with minor performance differences. In the study by Harel et al. (2006), the GBVS algorithm performed better than the ITTI algorithm in identifying the salient regions within natural images and Li et al. (2013) showed that the HFT algorithm outperformed both ITTI and GBVS in identifying saliency regions within natural images. In our study ITTI and GBVS obtained comparatively higher AUROC values for all small-scale magnetic images (Figure 5.7) when compared to HFT, and
5.2. Interpreter visual attention vs image saliency

Figure 5.7: The AUROC curve obtained for the eight target images (1 to 8) and eight non-target images (9 to 16) using three saliency detection algorithms.

GBVS obtained the highest average AUROC value (Table 5.1).

Table 5.1: Performances of ITTI, GBVS and HFT algorithms in detecting interpreter visual attention.

<table>
<thead>
<tr>
<th></th>
<th>ITTI</th>
<th>GBVS</th>
<th>HFT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exercise 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(Average AUROC)</td>
<td>0.8957</td>
<td>0.8999</td>
<td>0.8815</td>
</tr>
<tr>
<td><strong>Exercise 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(AUROC)</td>
<td>0.7903</td>
<td>0.6731</td>
<td>0.6741</td>
</tr>
</tbody>
</table>

For Exercise 2, we calculated the AUROC values for the three saliency detection algorithms using the same analysis method used for Exercise 1 and the results are shown in Table 5.1. The ITTI algorithm outperformed in identifying the regions that attract interpreter visual attention within magnetic images with multiple targets when compared to GBVS and HFT. Even though the HFT algorithm identified the saliency region effectively for small-scale images, it did not perform well in the large-scale image. This poor performance of the HFT with the large-scale image could be due to the wrong selection of the saliency map from multiple scale saliency maps using the entropy method as mentioned by the developers (Li et al., 2013). The superior performance of ITTI over the others could be due to the multi-scale strategy used in their algorithm which is based on the human attention model. The ITTI saliency map highlights the bottom left corner of the image whereas the
participants have fixated in that region comparatively less when compared to other regions (Figure 5.6). This discrepancy could be due to the top-down influence caused by the target spotting task which requires the identification of ‘Mexican hat’ like structures.

We selected the ITTI algorithm as the most suitable saliency detection algorithm for the magnetic data as it performed well with both the small-scale and large-scale data.

5.3 Analysis of interpreter visual bias

In this section we analyse the visual biases on magnetic data after commonly used filtering and enhancement techniques, and select the enhancement and filtering methods that will minimise these biases for interpretation. The locations within the data that will most likely attract visual attention during the interpretation process are highlighted by the saliency maps and thus show the potential visual biases in the data observation. We used the selected ITTI saliency detection algorithm as a proxy to identify the distribution of interpreter attention on the observed data. Three different magnetic datasets were used in this analysis:

1. The magnetic data used in the Exercise 2 - Data 1 (Figure 5.1).
2. Kirkland Lake, Ontario, Canada - Data 2 (Figure 5.8).
3. Kimberley Basin, Western Australia, Australia - Data 3 (Figure 5.9).

Details of Data 1 are explained in Section 2. The aeromagnetic data from Kirkland Lake in Ontario was acquired along north–south survey lines spaced at 200 m and with a nominal terrain clearance of 73 m. The gridded aeromagnetic image has a 40 m cell size. The area consists of komatiitic, tholeiitic, and calc-alkaline volcanic rocks overlain by clastic sedimentary and alkali volcanic rocks, which have been intruded by felsic to intermediate alkalic bodies and dykes (Ispolatov et al., 2008). The thick sequence of ultramafic-mafic volcanic rocks generates the strongest positive magnetic responses within this data (Figure 5.8).

The aeromagnetic data in Figure 5.9 is from the Kimberley Basin in Western Australia (WA) and was obtained by merging the 1993 Mt. Elizabeth, WA dataset with the 2012 Charnley, WA dataset. Both surveys were captured along north-south survey lines. The Mt. Elizabeth survey was flown at an altitude of 60 m, line-spacing of 400 m, and tie-spacing of 5000 m. The Charnley survey was flown at an altitude of 50 m, line-spacing of 200 m, and tie-spacing of 2000 m. The high-amplitude magnetic responses in this area are primarily due to the tholeiitic basalt of the
5.3. Analysis of interpreter visual bias

Figure 5.8: Aeromagnetic data from Kirkland Lake in Ontario, Canada. The arrow indicates a strong magnetic response generated by a thick sequence of ultramafic-mafic volcanic rocks. Data courtesy of Ontario Geological Survey, Canada.

Carson Volcanics, and the extensive dolerite sills and dykes of the Hart Dolerite. The south-western region (outlined) exhibits a different geological character reflecting the presence of Paperbark Supersuite rocks which largely comprise felsic granitoid plutons and several later dykes (Sheppard et al., 2012).

Table 5.2: Parameters used for the RTP calculations

<table>
<thead>
<tr>
<th>Data</th>
<th>Inclination (degrees)</th>
<th>Declination (degrees)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The RTP data was provided</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>73.597</td>
<td>-11.902</td>
</tr>
<tr>
<td>3</td>
<td>-47.198</td>
<td>2.895</td>
</tr>
</tbody>
</table>

All datasets were reduced to the pole using the parameters shown in Table 5.2 and enhanced using five commonly used techniques: first vertical derivative – 1VD, automatic gain control – AGC (Rajagopalan, 1987), analytic signal – AS (Roest et al., 1992), low-pass filter – LPF, and tilt derivative – TDR (Miller and Singh, 1994). Details of these methods are provided in Appendix A. These enhancement
methods are commonly used to enhance anomalies with certain characteristics within the magnetic data.

The first vertical derivative (1VD) relatively enhances steep gradients within signals and thus it is often used in interpretation to resolve the effects of adjacent anomalies and sharpen the anomalies over source bodies (Gunn et al., 1995). The automatic gain control (AGC) enhances the data to provide equal emphasis to both low and high amplitude signals, assisting in emphasising low amplitude anomalies. The controlling parameters of AGC are the window size over which the gain is computed, and the gain functions controlling signal and noise amplification. We used a window size of 30 cells for all three datasets and maximum gain corrections of 30, 100, and 30 for Data 1, 2, and 3 respectively.

The analytic signal (AS) generates maxima or peaks directly over small discrete bodies and over the edges of large bodies, thus AS maps provide an indication of the location of source body edges and corners (for example, lithological contacts, faults and shears zones, or fault block boundaries). The AS is particularly effective in regions where magnetic remanence or low latitude responses can adversely affect interpretation. The low-pass filter (LPF) is used to eliminate high frequency anomalies. It is applied to enhance longer wavelength anomalies produced by geological features at a given depth. We used 500 m, 5000 m, and 6000 m as the low-pass cut-offs for Data 1, 2, and 3 respectively. The tilt derivative (TDR) is used to enhance the weak magnetic anomalies, the signals of which are normally overwhelmed by high amplitude anomalies.
We displayed all the images in the same way as the magnetic images displayed during the eye tracking experiment for consistency. That is, the images were displayed in colour and illuminated with a false sun located on the north side of the region covered by the data at an inclination of 45°. In some cases, some of these enhanced images are displayed in greyscale during interpretation, but may produce different saliency maps when compared to colour images.

5.3.1 Results and discussion

Saliency maps for each of these enhanced magnetic images were generated using the ITTI saliency detection algorithm. Figures 5.10, 5.11, and 5.12 show the enhanced magnetic images and their corresponding saliency maps. The distribution of the highlighted areas of the saliency map varies with the specific enhancement method. These variations indicate different enhancement methods attract interpreters' visual attention differently. The analysis of the differences among these maps will help to understand the potential visual biases in interpreter data observations and to minimise the unintended biases during the interpretation. We quantified this difference by computing the root mean square (RMS) error (Equation 3) between two saliency maps:

\[
\text{RMS Error} = \left[ \frac{1}{n(X)} \sum_{x \in X} (f(x) - g(x))^2 \right]^{1/2}
\]

(5.3)

where \( f \) and \( g \) are the two saliency maps and \( n(X) \) is the number of pixels in the pixel raster \( X \).

We calculated the RMS error between all the combinations of the saliency maps and the results are shown in Tables 5.3, 5.4, and 5.5. The RMS error value increases between two images as they become more dissimilar.

The LPF method focuses the interpreter's visual attention to a narrow area compared to all the other methods in all three datasets. In Data 1, most of the known porphyry-style mineralisation is located on the belt running from top-left corner to the bottom-right corner (Figure 5.1). The LPF effectively attracts visual attention towards this belt (Figure 5.10e), whereas other enhancement methods such as RTP, AGC, and TDR are likely to draw attention mainly to the south-western region of the data which represents short wavelength shallow features. In Data 2, LPF highlights the thick sequence of ultramafic-mafic volcanic rocks in the central region (Figure 5.11e). The saliency map of Data 3 LPF covers a wider region compared to the other two datasets probably due to the strong and widespread magnetic response of the Carson Volcanics and Hart Dolerite sills (Figure 5.12e). Conversely, AGC and TDR enhancements attract interpreter visual attention to a wider area of the data, thus obtaining higher RMS error with the LPF data when compared...
to other combinations of methods in all three datasets (Table 5.3, 5.4, and 5.5). In most cases the LPF highlights areas not highlighted by AGC and TDR (Figure 5.10 and 5.11). This result indicates AGC and TDR enhancements complement LPF in attracting the human visual attention to different aspects of the geology. The selection of enhancement methods that produce complementary distribution of visual saliency for interpretation can ensure a more thorough examination of the data.

Some enhancement methods produced very similar distributions of human visual attention. For example, 1VD and AS of Data 1 produced almost identical saliency maps with an RMS error of 0.0668. We calculated residual signals for all three datasets by removing the regional effects from the magnetic data. This approximates the shallow crustal magnetic response, and can emphasise the higher frequency signal. The regional data were obtained by upward continuing the Data 1, 2, and 3 to 3000 m, 12 000 m, and 20 000 m respectively based on trial and error. Upward continuation is the process that transform data as if it is observed from a surface above the actual data observation surface, and results in smoothing high-frequency anomalies compared to low-frequency anomalies. Residual magnetic data had a similar saliency distribution as the RTP data with RMS errors of 0.0232,
5.3. Analysis of interpreter visual bias

Table 5.3: Calculated RMS error values between saliency maps of enhanced magnetic Data 1

<table>
<thead>
<tr>
<th></th>
<th>RTP</th>
<th>1VD</th>
<th>AGC</th>
<th>AS</th>
<th>LPF</th>
<th>TDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>0.0972</td>
<td>0.2267</td>
<td>0.1105</td>
<td>0.1962</td>
<td>0.1820</td>
<td></td>
</tr>
<tr>
<td>1VD</td>
<td>0</td>
<td>0.2306</td>
<td>0.0668</td>
<td>0.2148</td>
<td>0.1767</td>
<td></td>
</tr>
<tr>
<td>AGC</td>
<td>-</td>
<td>-</td>
<td>0.2223</td>
<td>0.3087</td>
<td>0.1266</td>
<td></td>
</tr>
<tr>
<td>AS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.1980</td>
<td>0.1807</td>
<td></td>
</tr>
<tr>
<td>LPF</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.2916</td>
<td></td>
</tr>
<tr>
<td>TDR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

0.0290, and 0.0192 for Data 1, 2, and 3 respectively. These results show that some enhancement methods are likely to attract interpreter visual attention similarly. The identification of this fact will help to minimise unintended biases towards certain regions by limiting the usage of redundant enhancement methods in interpretation.

The above results show how saliency maps can be used to analyse potential visual biases in interpretation when using different enhancement methods. Saliency mapping can assist in the selection of enhancement methods with complementary saliency distributions to analyse the entire data with minimum unintended biases. This will help to achieve minimum redundancy with maximum information gain. The results show that it is more effective to include TDR and/or AGC filtered datasets in the set of standard interpretation enhancements as they attract interpreter visual attention to different types of anomalies and regions in all datasets than is the case in other methods. To complement TDR and AGC, we recommend 1VD and LPF for Data 1; 1VD, LPF, and AS for both Data 2 and 3 based on saliency distribution.

As demonstrated in our previous study (Sivarajah et al., 2013), there are variations in data observation within and between interpreters, thus it is impossible to generate the human visual attention maps computationally. Therefore in this study we used the most suitable saliency detection algorithm (ITTI) that produces saliency maps close to the interpreter visual attention maps for the analysis of interpreter biases. Even though there could be minor variations in the ITTI saliency maps when compared to the actual interpreter visual attention maps, they can be used effectively as a guide in the selection of enhancement methods for interpretation.

The possibilities of using saliency maps for data with complex anomalies and
linear features require further analysis. This can be achieved by capturing the eye gaze movements from interpreters for the selection of the most suitable saliency detection algorithm for specific anomalies. Even though it requires the capturing of eye gaze data, once a most suitable saliency algorithm is selected, it can be used for all future interpretations. The quick generation of saliency maps helps to quantify the impact of different enhancement methods and could help in improving

Figure 5.11: Enhanced magnetic data (Data 2) using commonly used enhancement techniques (left) - see Figure 5.4b colour bar and the corresponding saliency heat maps (right) - see Figure 5.4c colour bar (a) RTP; (b) 1VD; (c) AGC; (d) AS; (e) LPF; (f) TDR.

Figure 5.12: Enhanced magnetic data (Data 3) using commonly used enhancement techniques (left) see Figure 5.4b colour bar and the corresponding saliency heat maps (right) - see Figure 5.4c colour bar (a) RTP; (b) 1VD; (c) AGC; (d) AS; (e) LPF; (f) TDR.
Table 5.4: Calculated RMS error values between saliency maps of enhanced magnetic Data 2

<table>
<thead>
<tr>
<th>Data 2</th>
<th>RTP</th>
<th>1VD</th>
<th>AGC</th>
<th>AS</th>
<th>LPF</th>
<th>TDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>0</td>
<td>0.0964</td>
<td>0.2663</td>
<td>0.1396</td>
<td>0.3274</td>
<td>0.2021</td>
</tr>
<tr>
<td>1VD</td>
<td>-</td>
<td>0</td>
<td>0.2643</td>
<td>0.0984</td>
<td>0.3178</td>
<td>0.1909</td>
</tr>
<tr>
<td>AGC</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.2511</td>
<td>0.5167</td>
<td>0.1386</td>
</tr>
<tr>
<td>AS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.3006</td>
<td>0.1780</td>
</tr>
<tr>
<td>LPF</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.4310</td>
</tr>
<tr>
<td>TDR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

The interpretation outcomes.

Saliency maps can also be used by enhancement algorithm developers to develop new techniques to compensate or augment biases, depending on the desired results. For example, developing algorithms that suppress regions of high saliency (which may be due to noise) to draw interpreter visual attention to other regions in the data, or vice versa.

Table 5.5: Calculated RMS error values between saliency maps of enhanced magnetic Data 3

<table>
<thead>
<tr>
<th>Data 3</th>
<th>RTP</th>
<th>1VD</th>
<th>AGC</th>
<th>AS</th>
<th>LPF</th>
<th>TDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTP</td>
<td>0</td>
<td>0.1055</td>
<td>0.1465</td>
<td>0.1636</td>
<td>0.1234</td>
<td>0.2566</td>
</tr>
<tr>
<td>1VD</td>
<td>-</td>
<td>0</td>
<td>0.0952</td>
<td>0.1049</td>
<td>0.1576</td>
<td>0.1738</td>
</tr>
<tr>
<td>AGC</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.1504</td>
<td>0.1936</td>
<td>0.1513</td>
</tr>
<tr>
<td>AS</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.1882</td>
<td>0.1748</td>
</tr>
<tr>
<td>LPF</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>0.2770</td>
</tr>
<tr>
<td>TDR</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
</tbody>
</table>

The identification of human visual attention using saliency detection algorithms could be improved by incorporating some top-down cues. Top-down information can be obtained based on either the geology of the area or what the interpreter is
looking for in the data, that is specific exploration targets. Another aspect that can be investigated using this method is the analysis of the impact of different colour display methods on the interpreter visual attention. For example greyscale versus colour and different colour schemes. This will help to understand the visual biases introduced by various display methods which can be used for the selection of most suitable display methods by minimising those biases.

5.4 Conclusions

This paper presents a study on the identification of an interpreter visual attention detection method. This is achieved by capturing interpreter data observation patterns while interpreting magnetic images. We used an ETS to capture geoscientists’ eye gaze movements and derived fixation locations, from which the interpreters’ visual attention maps were obtained. Comparison of these visual attention maps with saliency maps shows the effectiveness of the saliency detection algorithms in identifying salient regions within geoscientific magnetic images. The ITTI algorithm outperformed the HFT and GBVS methods in identifying areas attracting interpreter visual attention within magnetic images. Based on these findings, we propose that the ITTI algorithm is the most suitable saliency detection algorithm with which to create human visual attention maps for magnetic data.

The ITTI algorithm was then used for analysing potential interpreter biases on three magnetic datasets which were enhanced using common methods. The ITTI saliency analysis shows variations in the distribution of human visual attention depending on image enhancement method. The residual signal produced very similar saliency maps to the RTP images whereas LPF produced very different saliency maps across the three magnetic datasets. The TDR and/or AGC methods produced a wider dispersion of saliency covering various types of anomalies and regions in all three datasets used in our analysis. The results show that it is better to include images enhanced using TDR and AGC methods in the standard set of enhanced images used for the interpretation. In addition to TDR and AGC, we recommend 1VD and LPF for data with many small and strong magnetic responses, and 1VD, LPF, and AS for data with a dominant, and strong magnetic response or widespread high frequency magnetic response.

Saliency assessments provide invaluable input for both interpreters and software developers to guide the selection of effective methods and processes in capturing, enhancing, and displaying geoscientific images with minimum visual bias. The findings of this study establish a reliable framework to assess interpreter data observation biases for magnetic images based on the image saliency analysis, which can be adopted
to evaluate other types of geoscientific datasets.
Chapter 6

Quantifying target spotting performances with complex geoscientific imagery using ERP P300 responses

Abstract

Geoscientific data interpretation is a challenging task, which requires the detection and synthesis of complex patterns within data. As a first step towards better understanding this interpretation process, our research focuses on quantitative monitoring of interpreters’ brain responses associated with geoscientific target spotting. This paper presents a method that profiles brain responses using EEG technique to detect P300-like responses that are associated with target spotting for complex geoscientific data. In our experiment, eight interpreters with varying levels of expertise and experience were asked to detect features, which are likely to be gold-copper-rich porphyry systems within magnetic geophysical data. The target features occur in a noisy background and often have incomplete shape. Magnetic images with targets and without targets were shown to participants using the ‘oddball’ paradigm. The ERPs were obtained by averaging the EEG epochs across multiple trials and the results show delayed P3 response to the targets, likely due to the complexity of the task. The EEG epochs were classified and the results show reliable single trial classification of EEG responses with an average accuracy of 83%. The result demonstrated the usability of the P300-like responses to quantify geoscientific target spotting performance.

6.1 Introduction

Interpretation of geoscientific data provides the basis for the modelling and understanding of the earth’s sub-surface. These data interpretations directly influence decision making at all levels within resources (minerals, oil, gas, and geothermal)
industries and other research communities, such as ground water research. Indirectly, it will also impact the decisions of government agencies that draw upon these interpretations. Whilst these decisions have significant economic and social implications, geoscientific data interpretations are highly subjective and often inconsistent (Bond et al., 2007; Rankey and Mitchell, 2003; Sivarajah et al., 2013) as they heavily rely on interpreters’ ability to recognise and synthesise patterns within data.

Geophysical data specifically represent the earth’s physical properties. For example, magnetic data collected through airborne and ground surveys represent the combination of the crustal rock magnetism and earth’s magnetic field from the core. These data are usually represented using colour displays as shown in Figure 6.1. The interpretation of such data is based on pattern recognition - identifying geophysical signatures of geological features within data, which are often associated with localised anomalous characteristics and discontinuities, as well as the spatial relationships between features.

Figure 6.1: Magnetic data with multiple porphyry-style mineralisation. Courtesy of Barrick Gold of Australia Ltd.

Quantitative monitoring of the human-data interactions during the interpretation process can provide invaluable information about the impact of different data display methods on the interpretation outcome and may help to identify effective interpretation practices and strategies. While it can be argued that user behavioural responses such as button clicks can be captured easily and reliably, the button click reaction times comprise of a number of different cognitive processes and it is difficult to identify the process that attributed to the variation in reaction times (Luck, 2005). In contrast, neurological responses provide continuous measure during target spotting task (Luck, 2005) and they are known to have lower latency and lower variation in latency than behavioural responses. This has led to the use of neurological
responses for practical applications (Huang et al., 2011).

This is the first study of its kind which focuses on monitoring the neurological responses of magnetic data interpreters to identify their target detection for complex geological features within magnetic geophysical data. This is achieved by identifying ERP P300-like responses that are associated with target detection from EEG of interpreters.

P300 signals are widely researched in the areas of BCI (Birbaumer and Cohen, 2007) and in clinical studies (Duncan et al., 2009; Polich and Herbst, 2000). P300 responses are normally elicited using the ‘oddball paradigm’, where infrequent target stimuli (‘oddball’) and frequent standard stimuli are presented in a random sequence. In these exercises the task is to identify the infrequent target stimuli either by mental acknowledgement or by pressing a button. The process of stimulus evaluation, especially to a successful response to the target stimulus will elicit the requisite P300 response.

![Figure 6.2](image)

**Figure 6.2**: The ERP signal obtained for a simple target detection task by averaging across 100 target trials; 0 ms indicates the onset of the stimulus. In this experiment, red and green boxes were shown and the participant was asked to silently count the number of green boxes.

The P300 signal is mainly characterised by its amplitude, latency, and the scalp distribution. The P300 signal amplitude and latency are mainly affected by biological factors and experiment modality. Some of these factors include gender, seasonal cycles (Deldin et al., 1994), exercise (Yagi et al., 1999), age (Fjell and Walhovd, 2001), arousal level (Polich and Kok, 1995), complexity of the task (Caryl and Harper, 1996; Luo and Sajda, 2009; Ting et al., 2011), inter-stimulus interval (Polich, 1990), target to target interval (Gonsalvez and Polich, 2002), target probability (Polich, 1990), and image presentation sequence. The amplitude of the P300 response is defined as the maximum voltage within the time window, with respect to the pre-stimulus mean baseline voltage. The P300 amplitude is maximal in the midline region and decreases from the parietal region to the frontal region (Johnson, 1993). This scalp distribution changes significantly with age (significant reduction
in P300 response in the parietal region for older subjects) (Smith et al., 1980). It is very difficult to identify the P300 response (which is on the order of few microvolts) from raw EEG signals due to the strong background EEG activity (which is on the order of tens of microvolts). Conventionally, the P300 response is obtained by averaging the EEG traces across multiple trials (Figure 6.2). This process, termed grand averaging, suppresses the background EEG activity and channel noise, and enhances the P300 response. Grand averaging is not preferred in many cases since it is not possible to identify the trial-by-trial variations in ERP (amplitude and latency variations) and it requires a large number of samples.

Analysis of variations in brain responses during target spotting requires the identification of the P300 responses from a single trial or from few trials. However, the variations in the P300 response (amplitude, latency, and scalp distribution) and the difficulty in removing the background EEG activity from the raw EEG makes the detection of P300-like signal from a single trial sample more complicated and challenging. There have been different approaches proposed by the researchers for the classification of the signals with or without P300 responses. The SVM (Thulasidas et al., 2006) and linear discriminant analysis (LDA) (Hoffmann et al., 2008) are widely used techniques in the field of BCI.

Conventionally, detection of P300 responses is applied to simple classification tasks, such as identification of shapes, colours, and letters. However, there have been a few studies that used more complex tasks. For example, P300-like responses were used for the detection of people (Luo and Sajda, 2009) within images, where task complexities are associated with variations in position, scale, and pose of the people. A study by Gerson et al. (2006) used P300 and behavioural responses to prioritise the images (by moving the target images in front of an image stack). In other studies (Huang et al., 2011; Mathan et al., 2008) ERP responses were used to search for visual targets within satellite images.

Figure 6.3: Typical magnetic response of a porphyry-style mineralisation – appear as an elevated sub-circular feature with surrounding annular lows.

In our study, P300 responses were used to identify the detection of magnetic anomaly patterns of likely gold-copper-rich porphyry systems. A typical footprint for these mineral systems appears within magnetic data as an elevated sub-circular
feature with surrounding annular lows (Holden et al., 2011) (see Figure 6.3).

Magnetic porphyry target spotting is considered as a highly complex task since the patterns are contrasted with noisy background and vary significantly in shape and size within data. These are due to complex geological background and data sampling noise. Figure 6.4 shows some of the target (top row) and non-target (bottom row) images used for the experiment.

![Magnetic Images](image)

Figure 6.4: Magnetic images. Top row: some of the images shown with porphyry-style mineralisation (target images). Bottom row: some of the images shown without any significant porphyry-style mineralisation (non-target images).

In this paper, Section 2 explains our methodology including the experiment setup, data acquisition and pre-processing, classification, and the electrode selection. Section 3 presents the results and discussion which includes the ERP analysis and the classification performance. Finally in Section 4, the summary and on-going research are presented.

### 6.2 Methodology

#### 6.2.1 Experiment setup

Eight healthy participants (five females and three males with ages ranging between 22 and 34) with varying levels of experience and expertise participated in this study. The small number of participants was due to the fact that participants require some experience or knowledge about porphyry footprints within magnetic data. This is a specialised field and we only had access to a limited number of potential candidates. All of our participants are trained as geophysicists at least to an undergraduate level, except for one geologist who is experienced in magnetic data interpretation. Note that in this paper, we will not discuss the relationship between the participant...
performance and their level of expertise, as the small number used for this study does not warrant drawing any generalised conclusions. Prior ethics approval was obtained from The University of Western Australia, Human Research Ethics Office, and the participants consented to participate in the experiments. All the participants had normal or corrected to normal vision and had at least a basic understanding about the magnetic anomaly associated with porphyry-style mineralisation.

Participants were seated in front of a display monitor (52 cm x 33 cm) at a convenient distance (from 60 to 100 cm). The images were displayed on the centre of the monitor screen (within 23.4 cm x 17.4 cm area). We used two computers for this experiment, one used to display the images and record user responses and the other one to record the EEG signals.

The magnetic image was cropped down to small images with porphyries (Figure 6.4 - top row), i.e. target images, and without porphyries (Figure 6.4 - bottom row), i.e. non-target images. These ground truth targets were known locations of porphyry deposits, which are confirmed by field drill tests. In the visual display, eight target images were repeated 10 times and 50 non-target images were repeated six times. These images were shown in a random sequence using the ‘oddball’ paradigm, but the sequence was identical for all the participants. In the standard oddball paradigm, images are displayed for 50 - 150 ms (Duncan et al., 2009; Mathan et al., 2008) but this timing was too short for the participants to properly observe the magnetic data. We displayed the images for 1000 ms with an inter-stimulus interval of 1000 ms. This setting was determined through several experiments to reach a compromise between participant comfort and P300 elicitation. This increase in display time resulted in long experiment, causing possible participant fatigue. In order to minimise the experiment time we reduced the image repetitions to an acceptable minimum. The experiment took approximately 20 minutes and the participants were not allowed to take any breaks during this task. Ocular artefacts were minimised by displaying the images in the centre of the display screen and requesting the participants to fixate their eye gaze at a white cross on a black background during the inter-stimulus intervals. The participants were requested to respond to target images by pressing a key on a keyboard, as soon as they spotted the porphyry-style mineralisation to enable them to actively engage with the task.

### 6.2.2 Data acquisition and pre-processing

The EEG was acquired at 1000 Hz sampling frequency from 36 scalp sites (extended international 10-20 system), using Ag/AgCl electrodes mounted on a cap (Quik cap from Compumedic). These signals were amplified and recorded using the NuAmps amplifier and the SCAN Express software with earlobes average as the reference (A1
The EEG signals were time locked to the stimuli by sending the markers to the amplifier using E-Prime. All electrode impedances were kept below 10 kΩ for all the participants. Recorded EEG signals were band-pass filtered between 0.5 Hz and 30 Hz to remove unwanted noise signals present in the EEG.

We adopted BSS using ICA, (Jung et al., 2000b) which is a standard technique used for the removal of ocular artefacts present in the EEG signals. These signals were then low-pass filtered at 8 Hz. The ERP epochs with 100 ms pre-stimulus interval and 1000 ms post-stimulus interval were extracted for each image trigger, and the baseline was corrected using the pre-stimulus mean voltage. Button press responses ranged from 369 ms to 1343 ms (with median of 663 ms). In order to minimise the artefacts caused by the button presses, we extracted the response-epochs from the onset of the stimulus to a 600 ms interval within the ERP epochs (Huang et al., 2011). While the effect of physical artefacts (button press movement and eye blink) has been minimised, the physiological aspects of motor preparation are a potential confound (Salisbury et al., 2001). Even though movement-related potentials (MRPs) are reverse time locked to the actual movement, they may be coincident with the P300 responses. The MRP is a measure of brain activity that leads to voluntary muscle movement. It is not clear that the MRP effects are necessarily coincident with the ERP (Verleger et al., 2006). Given the greater variability seen in reaction time, these potentials are not expected to be time locked to the stimulus. However, advanced techniques to address this limitation, as originally canvassed in previous discussion of this confound (Kok, 1988), would be valuable.

### 6.2.3 Classification

The standard technique used to visualise and classify the P300 responses is by averaging the responses across multi-trials to minimise the background EEG activity. Multi-trial averaged ERP signals will help to identify the characteristics (amplitude, latency, and scalp distribution) of the P300-like responses for this particular target spotting task. This is achieved by averaging the ERP epochs across 80 target trials and 300 non-target trials.

We also performed single trial classification of the images based on P300-like responses using a SVM classifier, as it is not sensitive to overtraining (Jain et al., 2000) and performs well for large dimensional features (Bennett and Campbell, 2000; Burges, 1998). The SVM is a supervised machine learning algorithm, which is based on statistical learning theory. Even though it was created as a linear classifier, it is used as a non-linear classifier by projecting data into a higher dimensional space using kernel techniques. A non-linear SVM with a radial basis function (RBF) kernel implemented in LIBSVM (Chang and Lin, 2011) was used to classify the
Response-epochs from selected electrodes were concatenated to form a feature vector and scaled to the interval $[0, 1]$ (Hsu et al., 2003). Corresponding class labels for these feature vectors were obtained from the ground truth data. Response-epochs were classified into epochs that belong to the images with porphyry-style anomaly or without porphyry-style anomaly, by performing leave-one-out cross-validation. For the better classification of the epochs, the optimum values for the two regularization parameters (kernel width of Gaussian kernels: $\sigma^2$ and the cost parameter: $C$) of the RBF SVM needs to be calculated. These parameters were obtained by performing 10-fold cross-validation on the training data by varying the $\sigma^2$ from 0.00003 to 8 and the $C$ from 0.03 to 32000. For each leave-one-out classification, the parameter combination, which produced the highest balanced classification rate is used to train the SVM classifier. Normally classification accuracy is considered for the selection of the parameters, but in our case it is not an appropriate measure, as our data is highly unbalanced (80 targets and 300 non-targets). Balanced classification rate is defined as the arithmetic mean of the sensitivity and the specificity. Sensitivity and specificity are defined as follows,

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \\
\text{Specificity} = \frac{TN}{TN + FP}
\]

where $TP$ - true positive, $TN$ – true negative, $FN$ – false negative, and $FP$ – false positive.

### 6.2.4 Electrode selection

A combination of four middle line electrodes (Fz, Cz, Pz, and Oz) and four parietal electrodes (such as P3, P4, P7, and P8) is widely used for the detection of P300 responses in BCI type applications for better classification accuracy (Hoffmann et al., 2008; Krusienski et al., 2006). Selection of a minimum number of electrodes is always advantageous, as the classification time depends on the length of the feature vector. We carried out the P300 classification using different electrode combinations, as shown in Figure 6.5, to identify the combination of electrodes that leads to better classification performance for this specific target spotting task.

The SVM P300 classification performances for different electrode configurations were obtained by performing 10-fold cross-validation. Figure 6.6 shows different electrode configuration classification performances obtained from eight participants. Conf. 1 and Conf. 3 produced a similar performance when compared to other configurations, and Conf. 2 produced a slightly lower performance. Conf. 3 obtained a good overall performance across all the participants when compared to Conf. 4.
Figure 6.5: Different electrode combinations used for the classification of the P300 signal; (a) Conf. 1 - all electrodes with 1080 features; (b) Conf. 2 – middle line electrodes with 180 features; (c) Conf. 3 – electrodes selected based on the P300 scalp distribution (see Figure 6.9) with 390 features; (d) Conf. 4 – electrodes based on the BCI type application with 240 features (T5 and T6 electrodes were used instead of P7 and P8 electrodes).

We selected Conf. 3 for our analysis as this configuration performed well for all participants with a limited number of electrodes.

6.3 Results and discussion

6.3.1 Event related potential (ERP) analysis

The multi-trial averaged ERP signal obtained for target and non-target image presentations from eight different participants at the electrode location CPz (by averaging across 80 target trials and 300 non-target trials) is shown in Figure 6.7. Multi-trial averaging is carried out based on the assumption that the different re-
6.3. Results and discussion

Figure 6.6: Box plot shows the image classification accuracies ($\text{accuracy} = \frac{\{\text{truepositive} + \text{truenegative}\}}{\{\text{truepositive} + \text{truenegative} + \text{falsepositive} + \text{falsenegative}\}}$) obtained for eight participants using four different electrode configurations.

Repeated target images elicit similar P300-like response. The average P3 target latency for eight participants was around 527 ms. This increase in latency when compared to the P3 latency of a simple task, which is around 300 to 450 ms (Comerchero and Polich, 1999; Duncan et al., 2009), could be due to higher level of difficulty in the geoscientific target spotting task. This agrees with the findings of Caryl and Harper (1996), Luo and Sajda (2009), and Ting et al. (2011) on the latency of P300 for complex tasks. The P300-like amplitudes range from 10 µV to 26 µV for all the participants except participant 4 (around 4.5 µV).

The ERP images at channel location CPz for target and non-target images were rendered using EEGLAB (Delorme and Makeig, 2004) and are shown in Figure 6.8 (top graphs). This shows the ERP voltage variations throughout the trials. There is a consistent peak activity that could be observed between the 400 ms and 700 ms interval for the target trials and comparatively less activity could be observed within that time window for the non-target trials.

The ERP scalp distributions obtained using EEGLAB are shown in Figure 6.9 for target and non-target images at a 200 ms timing interval. The scalp maps of the ERP signal for the target and non-target images at 600 ms latency indicate that the P300-like peak mainly spans across the central and parietal regions of the scalp (Figure 6.9).
Figure 6.7: Multi-trial ERP responses obtained by averaging ERP responses across 80 target and 300 non-target trails for eight different participants.
6.3. Results and discussion

Figure 6.8: Top graphs show the amplitude variation of the ERP signal for the target (left) and non-target (right) images over time at different trials (smoothed across eight trials); bottom graphs show the amplitude variation of the average ERP across those 80 target trials (left) and 300 non-target trials (right). The $x$-axis represents time and 0 ms indicates the onset of the stimulus. These results are obtained from a representative participant (participant 3) at electrode location CPz.

Figure 6.9: Top row scalp maps show the ERP brain activity for a target image presentation at 200 ms interval. Bottom row scalp maps show the ERP brain activity for a non-target image presentation at 200 ms interval (different colours indicate the voltages in $\mu$V). These results are obtained from a representative participant (participant 3), although no individual distribution showed clear evidence of eye blink or movement confound.

6.3.2 Classification performance

Multi-trial analysis

The multi-trial average represents canonical analysis for a P300-like paradigm, but the single value result is unhelpful for comparative sensitivity analysis. In order to illustrate the continuum of classification for varying trial number, a pseudo-peak approach was adopted to allow group based comparison from a single average value.
For each participant the single trial ERP signal amplitude for the target and non-target, at the latency of the P300-like responses within the 80 trial averaged ERP was obtained. Obviously, each pseudo-peak is unlikely to exactly match the true (unknown) trial peak; however, we assume that P300-like responses occur at a similar time instant for different target detection, as there is a clear evidence of a P300-like response in the corresponding multi-trial average. Based on this assumption, we demonstrate the separability of responses of target images from other images based on a single value within a single sweep. We performed analysis of variance (ANOVA) between the target and non-target amplitudes for every participant and the results are shown in Table 6.1. Higher $F$-values obtained with $p < 0.01$ for all the participants except participant 4 ($p = 0.28$) shows significant difference in P300-like amplitudes with respect to the image type (target or non-target image) as expected in this paradigm, and the canonical average can classify targets.

Table 6.1: The mean amplitudes of P300-like responses and standard deviations in parentheses for the pseudo-peak approach and the results of ANOVA.

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Target</th>
<th>Non-target</th>
<th>$F$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>19.16(14.87)</td>
<td>8.9(16.5)</td>
<td>25.38*</td>
</tr>
<tr>
<td>2</td>
<td>10.32(10.15)</td>
<td>3.42(10.21)</td>
<td>28.88*</td>
</tr>
<tr>
<td>3</td>
<td>18.21(14.23)</td>
<td>10.08(13.01)</td>
<td>23.73*</td>
</tr>
<tr>
<td>4</td>
<td>4.44(9.98)</td>
<td>3.1(9.9)</td>
<td>1.16</td>
</tr>
<tr>
<td>5</td>
<td>19.46(16.47)</td>
<td>7.38(12.31)</td>
<td>52.23*</td>
</tr>
<tr>
<td>6</td>
<td>25.21(13.33)</td>
<td>8.59(13.45)</td>
<td>96.74*</td>
</tr>
<tr>
<td>7</td>
<td>16.81(9.66)</td>
<td>8.55(13.43)</td>
<td>26.59*</td>
</tr>
<tr>
<td>8</td>
<td>16.00(14.09)</td>
<td>8.54(13.55)</td>
<td>18.81*</td>
</tr>
</tbody>
</table>

Significant at 0.01 (*) levels.

We also calculated the multi-trial averaged responses for target (8 images) and non-target (50 images) images for all participants and obtained the peak amplitude of the ERP responses within the 450 ms and 600 ms interval. The ANOVA analysis was performed between the target and non-target peak amplitudes for each participant and the results are shown in Table 6.2. The results show higher $F$-values with $p < 0.01$ for all the participants except participant 4 ($p = 0.69$) and the mean voltages of the targets are higher than the mean voltages of the non-target responses.
Table 6.2: The mean amplitudes of P300-like responses and standard deviations in parentheses for the 10-trial average ERP peak approach and the results of ANOVA.

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Target</th>
<th>Non-target</th>
<th>F-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.05(4.45)</td>
<td>15.38(6.13)</td>
<td>8.70*</td>
</tr>
<tr>
<td>2</td>
<td>10.82(3.36)</td>
<td>6.64(4.16)</td>
<td>7.29*</td>
</tr>
<tr>
<td>3</td>
<td>19.19(7.47)</td>
<td>13.2(5.02)</td>
<td>8.52*</td>
</tr>
<tr>
<td>4</td>
<td>5.38(2.21)</td>
<td>4.79(4.15)</td>
<td>0.15</td>
</tr>
<tr>
<td>5</td>
<td>20.46(4.20)</td>
<td>10.24(5.38)</td>
<td>26.16*</td>
</tr>
<tr>
<td>6</td>
<td>25.44(4.96)</td>
<td>10.51(5.56)</td>
<td>51.03*</td>
</tr>
<tr>
<td>7</td>
<td>18.32(3.16)</td>
<td>11.12(5.59)</td>
<td>12.53*</td>
</tr>
<tr>
<td>8</td>
<td>17.74(5.19)</td>
<td>10.94(4.76)</td>
<td>13.75*</td>
</tr>
</tbody>
</table>

Significant at 0.01 (*) levels.

The results show that there is a significant difference in amplitude of the P300-like responses with respect to the image type, that is, the images could be classified based on the multi-trial average P300-like responses with more than 99% confidence except for participant 4. The low F-value obtained for participant 4 could be due to lower amplitudes of the P300-like responses (Figure 6.7d).

**Single trial analysis**

Further reducing the number of trials used for the multi-trial analysis will make the detection of the peak within the ERP responses difficult due to background EEG noise. Clearly a more sophisticated classification approach, such as SVM, is required for the single trial analysis. Single trial P300 response classification performances for eight participants are shown in Table 6.3. All participants’ single trial classification achieved relatively high accuracy (>78%) with an average of about 83% (accuracy is the proportion of the true results i.e. true positives and true negatives in the classification outcome). We obtained an average sensitivity of 59% and an average specificity of 89% for this classification task. Sensitivity performance could be increased by presenting more images in the experiment.

Results show the reliable classification of images based on P300-like responses. It is difficult to compare the classification performance with other ERP studies due to varying experimental conditions, image type, and classification methods used. In
most of the studies single trial classification accuracy between 65% and 90% was achieved (Bayliss and Ballard, 1999; Kaper et al., 2004; Krusienski et al., 2006). In another BCI study (Daubigney and Pietquin, 2011) sensitivity of more than 50% was obtained. On this basis, comparable performance was obtained for the single trial classification based on P300-like responses for our complex geoscientific target detection task.

Additionally, we conducted a further analysis to validate our EEG based target spotting performance by comparing them with the target spotting performance of the individual participants using the button click responses (Table 6.3). The classification accuracies based on the button clicks range from 74% to 96% across participants, which show the level of difficulty of the task. We calculated the Pearson’s product moment correlation between classification accuracies based on the button clicks and P300-like responses, and it was 0.75. This shows that there is a positive correlation between these two performance measures and the ERP P300-like response based classification can be used for the analysis of the target spotting performances of geoscientific interpreters.

Table 6.3: Target spotting performances of eight participants.

<table>
<thead>
<tr>
<th>Participant no.</th>
<th>Accuracy based on SVM classification(%)</th>
<th>Accuracy based on button clicks(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>81</td>
<td>87</td>
</tr>
<tr>
<td>2</td>
<td>84</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>84</td>
<td>93</td>
</tr>
<tr>
<td>4</td>
<td>78</td>
<td>74</td>
</tr>
<tr>
<td>5</td>
<td>79</td>
<td>90</td>
</tr>
<tr>
<td>6</td>
<td>86</td>
<td>93</td>
</tr>
<tr>
<td>7</td>
<td>86</td>
<td>91</td>
</tr>
<tr>
<td>8</td>
<td>84</td>
<td>96</td>
</tr>
<tr>
<td>Average</td>
<td>83</td>
<td>90</td>
</tr>
</tbody>
</table>
6.4 Summary and on-going research

This paper presents a geoscientific study that monitors human brain responses associated with the detection of complex geological targets, namely gold-copper-rich porphyry systems within magnetic images. Our study monitored interpreters’ brain responses using EEG technique, and these signals were analysed to identify P300-like responses that are associated with target detection. The delayed P3 responses obtained for all the participants and the participant performances based on button clicks indicate the level of difficulty of the task. Multi-trial average analysis shows that the target images could be identified based on the P300-like responses. Comparable classification accuracies obtained from the single trial analysis show that the P300 paradigm-based target identification is possible in complex geoscientific images. While we minimised the artefacts in the pre-processing stage, the possibility of ocular movement and button click confounds impacting the pattern classification is still needed to be further investigated. Strong positive correlation between the target spotting performance based on the button clicks and the classification accuracy based on the P300-like responses shows that the P300-like responses could be used to measure the target spotting performances.

Our on-going and future research includes the following two aspects. Firstly, the brain response analysis will be combined with interpreters’ eye gaze tracking to understand the interpretation process. The combined use of an ETS and EEG will allow the analysis of the interpretation process for a large dataset consisting of multiple targets. Our previous research successfully showed that eye gaze tracking reliably provided data observation patterns of interpreters (Sivarajah et al., 2013). Currently, we are in the process of combining this with EEG P300 detection to determine the relationship between the data viewing pattern and the target detection during the interpretation process. Secondly, we are extending the proposed study to understand perceived image saliency, when applying different magnetic data displays and enhancement methods by analysing the variations in P300-like responses. These results will be compared against the reaction time and interpreter feedback through questionnaires. This will help to understand the impact of different display methods for magnetic data interpretation, which can provide a roadmap for the training of interpreters. In this planned study, other single trial P300 amplitude estimation methods such as spatio-temporal filtering methods (Li et al., 2009) will be investigated.
Chapter 7

Conclusion

7.1 Summary

This thesis has demonstrated that monitoring of interpreter data interactions can be used to understand the geoscientific data interpretation process. The interpreter data interactions were quantitatively monitored by capturing data observation patterns using an ETS and brain responses using EEG. This study is the first of its kind to extend the use of these technologies to geoscience.

Interpreter data interactions were monitored during various target spotting exercises in which the task was to identify gold-copper-rich porphyry systems within magnetic images. Fourteen participants with varying levels of experience and expertise participated in the target spotting experiments. Three studies of ETS and EEG data were used to understand this geological target spotting. The first study used ETS data to identify effective data observation practices for target spotting by comparing the eye gaze traces and target spotting performances of different interpreters (Chapter 4). The second study used ETS data to understand human visual attention on magnetic images and identified a saliency detection algorithm to predict potential locations of interpreters’ visual attention (Chapter 5). The third study used EEG data to realise the potential of automatically identifying interpreters’ target spotting response through P300-like response detection using a machine learning technique (Chapter 6).

A target spotting experiment involving the analysis of data observation patterns for a large-scale image is presented in Chapter 4. In this experiment, the same magnetic image was presented in two different orientations to the participants: one was in the original orientation and the other one was rotated by 180°. Some participants were shown the original image first and the rotated image later, while other participants were shown the images in the reverse order. The target spotting performances show high variability between and within participants, but for most participants the overall target spotting performance was improved when the data was observed from
7.1. Summary

Multiple orientations. The data observation patterns showed a strong correlation between high target spotting performance and a more systematic data observation strategy. Different target spotting performances were obtained based on the orientation of the data, but the order in which the images were presented had no impact on the target spotting performances. The analysis of the interpreters’ visual attention shows that the same features appearing in different parts in the FoV seem to attract different levels of visual attention, again demonstrating the importance of the data observation from multiple orientations and the importance of systematic data observation. These findings may help in training geoscientific interpreters to improve their target spotting performances and accuracy.

Chapter 5 demonstrates the use of ETS data to understand interpreters’ visual attention during data inspection. Saliency maps from three different algorithms (ITTI, GBVS, and HFT) were compared with the eye gaze fixation patterns obtained from the ETS data. The results show that the ITTI algorithm outperforms others in the large-scale magnetic image with multiple porphyry targets. The ITTI algorithm was then used to understand potential visual attention biases when selecting a set of different enhancement and filtering methods for interpretation. The saliency maps were generated from three magnetic datasets after applying five commonly used enhancement and filtering methods. Some enhancement methods have similar distribution of human visual attention whereas others show significantly different distributions within the same dataset. Three important observations were obtained from the datasets: (i) residual signal produces very similar saliency maps as the RTP data; (ii) LPF data generates saliency maps that are significantly different from other enhancement methods; and (iii) the TDR and/or AGC produce a wider dispersion of saliency covering various types of anomalies and regions.

Interpreters can use saliency maps to minimise visual attention biases by selecting a set of enhancement and filtering methods which show a wider or different distribution of human visual attention, and by avoiding methods that produce similar saliency maps. These saliency assessments not only guide the selection of enhancement or filtering methods, but can also provide invaluable input for software developers. The possibility of generating these maps computationally enables a quick visual attention analysis without special experimental setup as opposed to the capturing of eye gaze movements.

Chapter 6 investigates the potential use of P300-like responses obtained from EEG data for the identification of target detection epiphany. The EEG data were captured during an oddball paradigm-based target spotting exercise. Multi-trial average analysis showed a clear difference in P300-like response amplitudes based on the presence or absence of targets within images. The delayed P300 responses
obtained for the target images and relatively lower target spotting performance obtained from button click target selections show the level of difficulty of the geo-scientific target spotting task. The ANOVA analysis performed on the multi-trial averaged ERP signals shows that the images could be classified into images with targets and without targets based on multi-trial average ERP with more than 99% confidence for all the participants except one. Single trial classification of the images based on P300-like responses using a SVM obtained an average accuracy of 83% across eight participants. This finding demonstrates that the identification of complex targets by detecting P300-like responses is possible. The positive correlation between the image classification accuracies based on the button clicks and the P300-like responses demonstrates the potential of using P300-like responses as a measure to quantify target spotting performances.

7.2 Future work

This work can be extended to better understand the interpretation process in a number of ways.

- The interpretation process can be further explored by combining the ETS and EEG responses. This first requires the identification of the regions of interest based on the eye gaze fixations captured using the ETS. Then, each fixation start time can be used as the onset of a stimulus. The EEG signals can then be analysed to identify the brain responses, as reported by Kamienkowski et al. (2012). This will help to capture participant’s responses without interfering with workflow. The outcome of this type of study will provide an insight into why some of the eye gaze fixated features were not selected as targets by participants. This study was beyond the scope of this thesis as it requires ETS with a higher sampling rate to match the EEG capture rate (Kamienkowski et al., 2012).

- The analysis of the variations in ERP P300 responses can be extended to quantify the difficulty of target spotting task as shown in the study of Huang et al. (2012). Such analysis will be useful to understand and assess the effectiveness of different data enhancement and display methods. According to Comerchero and Polich (1999), and Polich (1987), when tasks are easier the P300 response should be quicker (low latency) with a higher amplitude. This means that a better enhancement technique should generate images that evoke P300 signals quickly with more discernible peaks than others. These results can then be compared or combined with target selection button click reaction times and
interpreter feedback obtained from questionnaires to assess the effectiveness of different enhancement and display methods. Another aspect to consider is interpreter’s familiarity with the data being presented, as this will have some impact on the P300 latency. This prior knowledge about the data will need to be collected from questionnaires, and used to analyse the impact of biases on P300 response. To achieve this, EEG processing and pattern recognition methods need to be extended.

• The calculation of human visual attention maps can be improved by fusing some task-specific cues into the identification of saliency regions. These cues can be generated based on either the geology of the region or the characteristics of the features that are being interpreted. This will help to analyse interpreter biases based on the task as well as the data.

• Welland et al. (2006) investigated the impact of human visual perception of colours on seismic data interpretation. This could be quantified by extending the saliency analyses to understand the visual biases introduced by various display methods. This could include the analysis of greyscale and different colour display methods on interpreter visual attention, which may help to select the best display method.

This thesis provides the basis for a potentially ground-breaking approach to understanding and improving geoscientific data interpretation. These findings in combination with the proposed future work, offer the promise of enhanced geoscientific interpretation.
Appendix A

Geophysical data enhancement and filtering methods

A.1 Reduction to the pole (RTP)

As the name suggests, this technique transforms the induced magnetic responses to that which would be measured in a vertical magnetic field, for example at the magnetic poles. RTP transforms the asymmetric responses of vertical components to a simpler symmetric form. This simplifies the anomalies by centering the ‘highs’ over the causative magnetic bodies. The RTP assumes that all bodies are magnetised by induction and it is not a valid technique, where there are appreciable remanent magnetism is presents.

A.2 Vertical and horizontal derivatives

Derivative filters are used to calculate the spatial rate of change of the magnetic field in the vertical and horizontal direction respectively. Derivative filters are important to enhance anomalies originating in the near surface, however, they enhance noises as well, which limits the use of high order derivatives.

A.2.1 First vertical derivative (1VD)

First vertical derivative also known as vertical gradient is define as follows,

\[ 1VD = -\frac{\partial A}{\partial z} \]

where \( A \) is the magnitude of the total magnetic field.

1VD is one of the most commonly used edge/discontinuity enhancement techniques. This enhancement removes the long wavelength component, thus helps to resolve closely spaced anomalies. This filtering can be noisy because it amplifies short wavelength noise.
A.2.2 Total horizontal derivative (THD)

\[ \text{THD} = \sqrt{\left(\frac{\partial A}{\partial x}\right)^2 + \left(\frac{\partial A}{\partial y}\right)^2} \]

where \( A \) is the magnitude of the total magnetic field.

In this enhancement, maxima represent source edges provided contacts are vertical. This filtering method often generates more accurate locations of faults when compared to first vertical derivative. However, it should be used in combination with other transformation techniques, such as RTP, for magnetic data.

A.3 Analytic signal (AS)

\[ \text{AS} = \sqrt{\left(\frac{\partial A}{\partial x}\right)^2 + \left(\frac{\partial A}{\partial y}\right)^2 + \left(\frac{\partial A}{\partial z}\right)^2} \]

where \( A \) is the magnitude of the total magnetic field.

The AS generates maxima or peaks directly over small discrete bodies and over the edges of large bodies, thus analytic signal maps provide an indication of the location of source bodies. It is particularly effective in regions where magnetic remanence or low altitude can adversely affect interpretation, but is susceptible to noise.

A.4 Tilt derivative (TDR)

\[ \text{TDR} = \tan^{-1}\left[\frac{\text{VD}}{\text{THD}}\right] \]

where VD is the vertical derivative and THD is the total horizontal derivative.

The output of this filter always lies between \( +90^\circ \) and \( -90^\circ \) because of the nature of arctan function. This method is used to enhance the weak magnetic anomalies, the signals of which are normally overwhelmed by high amplitude anomalies.

A.5 Spectral filtering

Spectral filtering is a method used to separate a particular band of frequencies in the data. In the magnetic data low frequencies or long wavelengths represent the deep anomalies and the high frequencies or short wavelengths represent the shallow features. Low-pass filters are used to separate regional signals and the high-pass filters are used to separate residual signals in the data.
A.6 Automatic gain control (AGC)

The AGC is a window-based amplitude filtering method which increases low amplitude signals whilst suppressing high amplitude signals (Rajagopalan and Milligan, 1994).
A.6. Automatic gain control (AGC)
Appendix B

Sample consent form

CONSENT FORM
PROFILING OF HUMAN DATA INTERPRETATION

I (given name) ______________________ (surname) ______________________ consent to my data recorded from eye and head trackers and EEG sensor for use in the study that profiles human data interpretation. I have read the information provided and any questions I have asked have been answered to my satisfaction. I have been advised as to what data is being collected, what the purpose is, and what will be done with the data upon completion of the research.

As long as no other identification information is included, I have no objection to the following:

My data may be published in scientific textbooks, journals, conference proceedings, presentations and other appropriate media. [YES / NO]

My data may be distributed for the purpose of research. [YES / NO]

I MAY / MAY NOT be sent information regarding future data acquisition sessions on my email address:

Email: ______________________

( ______________________ )

Signature of participant  date:

The Human Research Ethics Committee at the University of Western Australia requires that all participants are informed that, if they have any complaint regarding the manner, in which a research project is conducted, it may be given to the researcher or, alternatively to the Secretary, Human Research Ethics Committee, Registrar's Office, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009 (telephone number 6488-3703). All study participants will be provided with a copy of the Information Sheet and Consent Form for their personal records.

Eun-Jung Holden
Research Associate Professor,
Centre for Exploration Targeting, School of Earth & Environment, UWA
Tel (08) 6488 5806
eunjung@cyllene.uwa.edu.au
Bibliography


P. Caryl and A. Harper. Event related potentials (ERPs) in elementary cognitive

T. Chadwick, T. Tan, G. West, D. A. McMeekin, E.-J. Holden, M. Dentith, and
T. C. McCuaig. Quantitative study of geological target spotting with the use
of eye tracking. In *Proceedings of the 2010 workshop on Eye gaze in intelligent
ACM.

C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM

L. S. Cohen, P. F. Escobar, C. Scharm, B. Glimco, and D. A. Fishman. Three-
dimensional power doppler ultrasound improves the diagnostic accuracy for ovar-

M. D. Comerchero and J. Polich. P3a and P3b from typical auditory and visual

E. Cutrell and Z. Guan. What are you looking for?: an eye-tracking study of
information usage in web search. In *Proceedings of the SIGCHI conference on

L. Daubigney and O. Pietquin. Single-trial P300 detection with kalman filtering
and SVMs. In *Proceedings of the European Symposium on Artificial Neural Net-
works, Computational Intelligence and Machine Learning*, pages 399–404, Bruges,
Belgique, Apr. 2011.


A. Delorme and S. Makeig. EEGLAB: an open source toolbox for analysis of single-
trial EEG dynamics including independent component analysis. *Journal of Neu-

H. Deubel and W. X. Schneider. Saccade target selection and object recognition:
1837, June 1996.


T. Hoschke. Geophysical Signatures of Copper-gold Porphyry and Epithermal Gold Deposits, and Implications for Exploration. ARC Centre of Excellence in Ore Deposits, University of Tasmania, 2011.


P. Kovesi. MATLAB and Octave functions for computer vision and image processing, 2000.


A. Lucieer, S. Robinson, and D. Turner. Unmanned aerial vehicle (UAV) remote sensing for hyperspatial terrain mapping of antarctic moss beds based on structure from motion (SfM) point clouds. In Proceedings of the 34th International


