ASEG-PESA 2013
'THE EUREKA MOMENT'
11-14 AUGUST 2013 MELBOURNE, AUSTRALIA

Towards Understanding and Improving Geoscientific Data Interpretation

Yathunanthan Sivarajah
Centre for Exploration Targeting, University of Western Australia, 35 Stirling Hwy, Nedlands, WA 6009, Australia.
20724504@student.uwa.edu.au

Eun-Jung Holden
Centre for Exploration Targeting, University of Western Australia, 35 Stirling Hwy, Nedlands, WA 6009, Australia.
eun-jung.holden@uwa.edu.au

Roberto Togneri
Electrical, Electronic and Computer Engineering, University of Western Australia, 35 Stirling Hwy, Nedlands, WA 6009, Australia.
roberto.togneri@uwa.edu.au

Mike Dentith
Centre for Exploration Targeting, University of Western Australia, 35 Stirling Hwy, Nedlands, WA 6009, Australia.
michael.dentith@uwa.edu.au

SUMMARY

Geoscientific data interpretation is a highly subjective and complex task as human intuition and biases play a significant role. Based on these interpretations, however, mining and petroleum industries make decisions with paramount of financial implications. As a first step towards understanding and improving the interpretation process, we carried out two experiments to monitor the human-data interactions during the process of identifying 'targets' (porphyry-style intrusive systems) within the aeromagnetic imagery. This is achieved by capturing the eye gaze position using an eye tracker system and the brain responses using electroencephalography (EEG).

The first experiment was intended to analyse the target spotting performance and the data observation patterns. For this experiment participants performed exercises, where the same magnetic image was presented in different orientations. Some key findings include: inconsistencies in target spotting performance within and between the interpreters; an improvement performance when the data were viewed in multiple orientations; and a strong correlation between the target spotting performance and efficient (systematic) data observation pattern. There was no correlation between success in identifying targets and the participants' perception of their expertise.

The second experiment was designed to identify the characteristics of the targets that are easier to detect using EEG. For this experiment images with targets and without targets were presented in a rapid visual display. The analysis on the image characteristics based on the human visual attention model show a strong correlation between target spotting difficulty and dispersion of the visual attention.

Key words: geoscientific data interpretation, interpretation variability, eye tracker, data observation pattern, brain responses.

INTRODUCTION

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. Rankey and Mitchell 2003 carried out a study on six interpreters, designed to analyse the impact of interpretation and the uncertainties associated with it on seismic attribute analysis for the prediction of reservoir properties. 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 21% of the participants successfully identified the three major faults present in the image and 23% 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 a study specifically focusing on mineral exploration, Wastell et al. 2011 studied decision uncertainties. 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 interact with the data whilst interpreting, and its impact on the effectiveness and efficiency of the interpretation (Sivarajah et al. 2012a; Sivarajah et al. 2012b). We monitor the human data interactions during interpretation using an eye tracker (ETS) that traces the eye gaze of the interpreter and using the electroencephalography (EEG) that captures the brain responses. Participants with varying degrees of experience and expertise participated in these experiments (first experiment-fourteen; second experiment-eight). We designed two experiments, where the task was to identify porphyry-style intrusive systems in magnetic data.

For the first experiment participants performed two exercises, where the same magnetic image with multiple porphyry targets was presented in different orientations. The data observation patterns are tracked in real-time using an eye tracker system (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. The results show inconsistencies in target spotting performance within and between interpreters and an improvement when data observed from different orientations.

The interpreter data observation patterns show a strong correlation between the target spotting performance and efficient (systematic) data observation. In addition, there was no correlation between target spotting ability and the participants' perceived expertise.

For the second experiment the images with porphyry and without porphyry were presented in a rapid visual display. The neurological responses were obtained by capturing the EEG during this target spotting exercise. The specific neurological response that corresponds to the target selection (P300 response) is used as a measure to identify the interpreter's responsiveness or lack of responsiveness to the presented visual pattern. We compare these responses to the spatial distribution of salient features within the same images. This technique allows us to quantitatively assess the relationship between interpreter's target detection responses and the spatial spread of salient features within image. The result shows the targets with lower dispersion of visual attention are easier to detect.

METHOD AND RESULTS

In this study participants' eye gaze were captured using a mobile eye tracker available from Applied Science Laboratories and the EEG signals were amplified and recorded using the NuAmp amplifier and the SCAN Express software. At the beginning of each experiment the participants were seated in front of a display monitor at a convenient distance (from 60 – 100cm) and were then fitted with the ETS glasses and EEG head cap.

Experiment 1

14 participants with varying levels of experience and expertise participated in this experiment. The experiment consisted of two exercises, both of which used the same magnetic image, but displayed 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 at the top of the screen (Figure 1a), which will be referred to as the original image. For the second exercise the image was rotated by 180º (Figure 1b), which will be referred as rotated image. Both images were illuminated with a false sun located at the top of the screen, i.e. from actual north in the original image and actual south in the rotated image. Participants were given three minutes for each exercise and had a 30 minute break, during which they were distracted with other tasks. Subjects 10 to 13 saw the rotated image first and all the other subjects saw the original image first.

During the exercise, the participants were asked to press a keyboard button when a target was identified, while fixing their eye gaze at the target 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.

Quantitative measurement of the target spotting performance requires knowledge of the 'ground truth' targets. The survey area is very well explored area and has a number of known deposits, which could be used as the target set, but the number is limited and it does not include many of the magnetic anomalies which have the desired qualities. Thus, for the set of 'ground truth' or 'true' targets for our analysis, we decided to also use the targets generated by a pattern recognition algorithm designed to identifying the magnetic responses of porphyry systems, which is known as the CET Porphyry Detection for Oasis montaj 2012. The set of 'true' targets used to judge the accuracy of the subject's interpretation is a combination of those derived by both methods (42 'ground truth' targets).

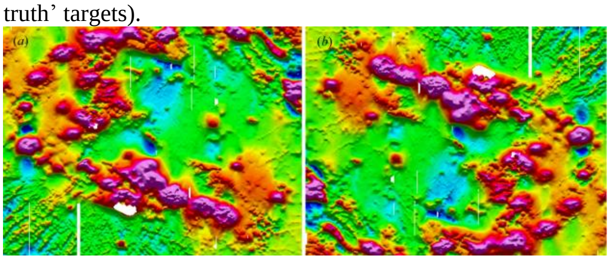

Figure 1. Magnetic images used for the original *(a)* and rotated image exercise *(b)*.

Data was analysed based on the participant's analysis of the original image, the rotated image, and the combination of the two. The repeated identifications of the same targets were removed from the analysis. Based on this information we quantified target identification performance by calculating Recall and Precision, defined as follows,

$$\text{Recall (R)} = N_{ik} / N_k$$

$$\text{Precision (P)} = N_{ik} / N_s$$

Where, N_k is the number of targets in image ($N_k = 42$).
N_{ik} is the number of targets correctly identified by the participant. N_s is the total number of targets 'identified' (without duplication) by the participant.

We used two different eye gaze-based measures to quantify the data observation patterns. These are the scan path length and scan path duration. Scan path length is the total distance traced by the eye during the exercise (measured in pixels). We also computed the mean scan path length between target identifications. An equivalent temporal measure is scan path duration between target identifications. These parameters are indications of the efficiency of the data analysis.

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

Variability in target spotting performance: Figure 2 shows the target spotting performance of all 14 participants for the original and rotated image exercises and also both exercises together (combined exercise). Recall ranges from 0.17 to 0.67 with an average of 0.36 for the original image exercise and for the rotated image exercise it ranges from 0.10 to 0.43 with an average of 0.26 (Figure 2). There is significant variation between the results from the different exercises as completed by each individual and between the performances of individual interpreters. Precision calculations indicate all participants had relatively high precision (greater than 80%) except subject 1, 7 and 14 in original and rotated image exercises (Figure 3). Interestingly, the level of self-assessed expertise did not correlate with participants' performance in terms of target detection accuracy. Here we use the outcomes of the target identification exercise as a measure of expertise, referring to them as high achievers and low achievers.

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, regardless of the order in which the exercises were performed (subject numbers 10 to 13 performed the rotated image exercise first). Four subjects obtained better performance in the rotated image exercise. This shows the variability in target spotting performance based on the orientation of the data within each subject. The important outcome from this component of the experiment is that orientation and/or direction of illumination does make a difference.

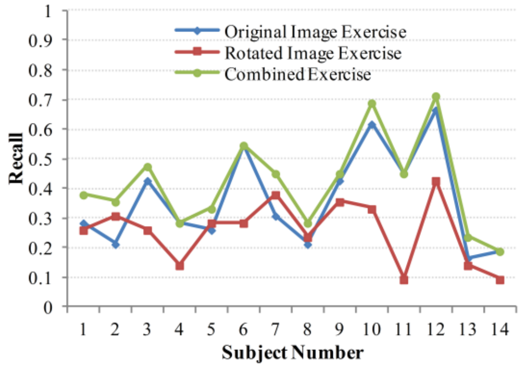

Figure 2. Recall versus subject number for three different exercises.

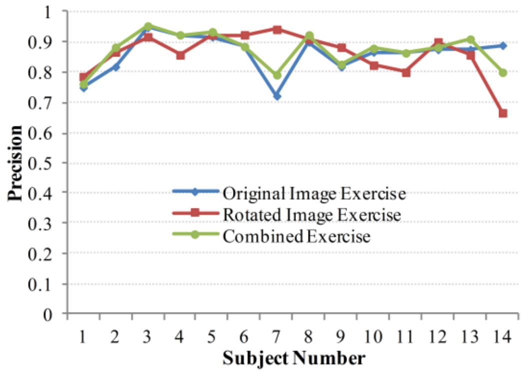

Figure 3. Precision versus subject number for three different exercises.

Even though most of the subjects performed better with the original image there are some times missed targets, which were identified in the rotated image. This is represented by the performance improvement obtained in the combined analysis (Figure 2). Eight participants obtained higher recall in the combined case than the original image exercise, while others had no improvements (subject number 4, 6, 11 and 14), but 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%. Comparable precision rates (Figure 3) and the overall improvement in the recall rates for the combined exercise quantitatively shows that the geoscientific data interpretation performance can be improved by observing the data from multiple orientations during interpretation.

Variability in data observation patterns: Data observation patterns differ significantly among subjects. 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 scan path length and mean scan path duration. 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 1).

	Original image exercise (r)	Rotated image exercise (r)
Mean scan path length vs recall	-.675*	-.800*
Mean scan path duration vs recall	-.692*	-.748*

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

Table 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.

Strong negative correlation between the mean scan path length and the recall indicates, as the performance increases the mean scan path length decreases, which means that high achievers observed the data more efficiently than the low achievers. The negative correlation between the mean scan path duration and recall indicates the decision making difficulty increases (longer scan path duration) with the decrease in target spotting performance. That is high achievers made the decisions quickly when compared to the low achievers. Therefore these results quantitatively show a strong correlation between the target spotting performance and efficient data observation patterns.

Experiment 2

Eight participants with varying levels of experience and expertise participated in this study. 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 sequence was identical for all the participants. Each image was shown for 1000ms with an inter-stimulus interval of 1000ms. User responses were also collected by requesting the participants to respond to the target images by pressing a key on the key board, as soon as they detect a porphyry (the target) in an image.

A well-known image saliency analysis method proposed by Itti and Koch 2000 is used for this study. In this model, the retinal input of the human vision system is first processed in parallel to generate a set of feature maps based on colour, intensity or orientation. These maps are then combined at each location to generate a topographic saliency map which is then used to prioritise and select the salient region locations. Only the features from this location are selected for further analysis.

The expected dispersion of visual attention for each target image is measured using the saliency map output. Two separate measures are used. The first is the major axis length of the largest object within each saliency map. The other is the dispersion of circular peaks within the saliency map.

Interpreter target detection is evaluated by detecting the EEG signal that is associated with target detection; this is known as the P300 response. The P300 response is indicated by the positive deflection in the EEG signal around 300ms after the presentation of a target visual stimulus. In this study we used the detectability of the P300 responses as a measure to quantify the saliency of the features within the images. The ranking of the target images using the P300 responses is compared with the ranking of the target images based on the image saliency analysis.

Captured EEG signals were pre-processed to minimise the artefacts present in the signals. Ocular artefacts were minimised by applying blind source separation (BSS) and the other artefacts were minimised by filtering. Brain responses corresponding to the presented images were obtained from the pre-processed signals by extracting an epoch (0 – 700ms) after the presentation of each image. Based on these epochs the images were classified into images with porphyries and without porphyries using a support vector machine (SVM) (Sivarajah et al. 2012b). Each target image was shown ten times, the target identification of each image for each subject was calculated. These results were used to calculate the correct identification of each target image by the eight subjects. The percentage of correct identification of each target image out of 80 presentations (each target image repeated 10 times and presented to 8 subjects = 80 total presentations).

The target images were ranked based on the percentage of correct identification of the targets (highest to the lowest) by all subjects using the P300 responses (Figure 4). The target images also ranked based on the dispersion of saliency using major axis length of the saliency map (lowest to the highest) and the standard deviation of the positive peaks (lowest to the highest) within the salience image. Statistically significant Pearson's correlation coefficients were obtained between the ranking based on the P300 responses and the ranking based on the dispersion of saliency map (ranking based on the major axis length: 0.74; ranking based on the standard deviation of the positive peaks within saliency maps: 0.81). This shows a strong positive correlation between the interpreter target detection and the low dispersion of saliency images. This is a significant finding for two reasons. This finding has potentially established a link between interpreter P300 responses and dispersion of saliency. The other significance is that it demonstrates a potential in using human attention model-based saliency measure to assess and optimize data display and enhancement methods.

CONCLUSIONS

In this study we presented the findings from two experiments to understand the geoscientific interpretation process. Data observation patterns show a strong correlation between the target spotting performance and efficient (systematic) data observation pattern. The target spotting performance results show inconsistencies within and between the interpreters and an improvement when the data were viewed in multiple orientations. There was no correlation between success in identifying targets and participants' perceived level of their expertise. The analysis on the image characteristics based on the human visual attention model show a strong correlation between target spotting difficulty and dispersion of the visual attention. Our findings are useful in identifying effective data display methods, which can provide a roadmap for the training of geoscientific data interpreters.

ACKNOWLEDGMENTS

We would like to acknowledge IVEC for the loan of the eye tracker. We would also like to thank the volunteer geoscientists, who participated in the experiments; and Barrick Gold of Australia Ltd for the permission to use their data in this research. We also acknowledge The University of Western Australia for providing a Scholarship for International Research Fees (SIRF) for this study.

REFERENCES

Bond, C.E., A.D. Gibbs, Z.K. Shipton, and S. Jones. 2007. "What Do You Think This Is? 'Conceptual Uncertainty' in Geoscience Interpretation." *GSA Today* 17 (11): 4. doi:10.1130/GSAT01711A.1.

CET Porphyry Detection for Oasis montaj. 2012. "CET Porphyry Detection for Oasis Montaj." Accessed November 15. http://www.geosoft.com/products/software-extensions/cet-porphyry-detection.

Frodeman, Robert. 1995. "Geological Reasoning: Geology as an Interpretive and Historical Science." *Geological Society of America Bulletin* 107 (8) (August): 960 –968. doi:10.1130/0016-7606(1995)107<0960:GRGAAI>2.3.CO;2.

Itti, L., and C. Koch. 2000. "A Saliency-based Search Mechanism for Overt and Covert Shifts of Visual Attention." *Vision Research* 40 (10-12): 1489–1506.

Rankey, E. C., and J. C. Mitchell. 2003. "Interpreter's Corner—That's Why It's Called Interpretation: Impact of Horizon Uncertainty on Seismic Attribute Analysis." *The Leading Edge* 22 (9): 820. doi:10.1190/1.1614152.

Sivarajah, Y., Holden, E.-J., Togneri, R., Dentith, M., and Shragge, J., 2012a. "Analysing Variability in Geophysical Data Interpretation by Monitoring Eye Gaze Movement." *ASEG Extended Abstracts* 2012 (1) (January 1): 1–4.

Sivarajah, Y, Holden, E.-J., Togneri, R., Tan, T., Price, G., and Kovesi, P., 2012b. "Studying the Impact of Feature Saliency for Geoscientific Target Detection Using Electroencephalography (EEG)." In *2012 International Conference on Digital Image Computing Techniques and Applications (DICTA)*, 1 –6. doi:10.1109/DICTA.2012.6411707.

Wastell, Colin A, Mike Etheridge, Maureen McMahon, Gillian Lucas, and Lisa Hartley. 2011. "The Impact of Cognitive Predispositions on Exploration Decisions in the Minerals Industry." *Applied Cognitive Psychology* 25 (3) (May 1): 469–479. doi:10.1002/acp.1715.

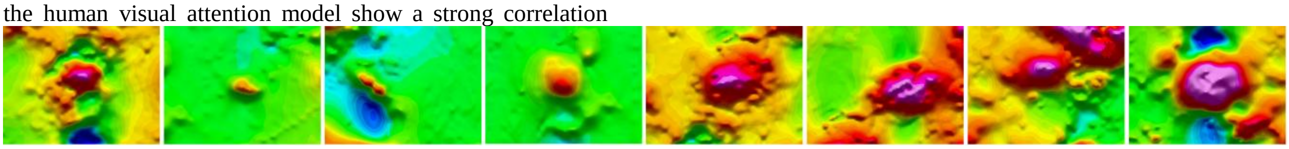

Figure 4. Easy to detect targets to difficult targets (from left to right) based on the P300 responses.