

Quantitative Study of Geological Target Spotting with the use of Eye Tracking

Tristan Chadwick, Tele Tan, Geoff West and
David A. McMeekin
Curtin University of Technology
Perth, Western Australia
t.tan@curtin.edu.au

Eun-Jung Holden, Mike Dentith and
T. Campbell McCuaig
Centre for Exploration Targeting
University of Western Australia
eunjung@cyllene.uwa.edu.au

ABSTRACT

In this paper we describe the use of eye tracking to quantitatively evaluate and analyse the variations in data interpretation performed by various geoscientists, measuring this against their ability to spot geological targets. We also describe an approach to evaluate the impact data preprocessing (i.e. enhancement) has on one's ability to perform the interpretation task. We adapted a mobile eye tracker to enable it to accurately map the point of gaze to the actual image coordinate instead of the forward facing eye tracker camera allowing the user to move their head as they view. Several visual interpretation tasks were performed by six geoscientists and the results are described in this paper.

Author Keywords

Eye Tracker, Target Spotting, Geological Features.

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

1. INTRODUCTION

Image data is a key component in many fields. Images are a method of gathering data about a subject for posterity, reference, intelligence gathering or discovery. In many cases, images provide vast amounts of information about the observed area whereby trained image interpreters are used to detect and identify occurrences of important target features in the image.

In prospectivity analysis for mineral exploration, geoscientists frequently analyse multiple, noisy and incomplete datasets comprising geological, geochemical, geophysical and remote sensing data. Such analysis is a largely subjective process as their interpretation of data is significantly governed by prior knowledge and experiences as demonstrated clearly from human analysis outputs of

seismic data (Bond et al., 2007). The decision regarding selection of a target within an area undergoing mineral exploration is a task that has to consider and reconcile varying individual opinions from a group of geoscientists. This study aims to quantify and analyse the variations in data interpretation that exist amongst individual geoscientists using an eye tracking system. The specific aims are twofold: (1) examine the variation in their ability to spot targets that are specific geological features within aeromagnetic data; and (2) determine the effectiveness of a data enhancement method that is commonly used for the interpretation. Both goals required the evaluation of eye tracker technology and how it can be most effectively used for studying the analysis of image data. In a similar research by Vilaroni et al (2007), an eye tracker was used to help speed up the cancer labelling process of video data. It also required the adaptation of the eye tracker to suit the situation, construction of multiple pieces of data acquisition, visualisation and analysis software and human interpretation trials to collect these data. The first aim entails the study of the affect of experience and area of expertise on geological data analysis. The second aim entails the study of the impact of data pre-processing or enhancement on the outcome of human visual analytical operations. The overall objective is to utilise the findings of this work to devise future methods to improve the target spotting accuracy involving geological datasets. The cognitive effect on eye movement with relation to various visual analytical tasks is discussed in Section 2. Section 3 details the type of eye tracking system used and how it is adapted to suit the needs of the research. The experimental procedures and settings of the research are described in Section 4, followed by the analysis of the results in Section 5.

2. BACKGROUND

We review the theories concerning the cognitive affects of eye movement, and their implications with regards to this research.

2.1 Bottom-Up

The bottom-up theory of eye movement control is defined as a subconscious process. It proposes that eye movements

are governed by the features of the visual scene (Deangelus and Pelz 2009). Specifically, the fixation density and scan-path between fixations is determined by low-level features such as contrast, brightness, edges and shapes. Early work in understanding eye-movements from this perspective brought about a proposal for a network calculation for an image that determined the visual appeal of each location (Koch and Ullman 1985). Such techniques are continuing to be applied and show high correspondence with tested user views.

2.2 Top-Down

The top-down theory proposes that eye-movements are directly influenced by the conscious task or intention of a person. The foundational work for this theory was performed by Yarbus (1967). While Yarbus was not the first to examine the movements of the eyes in detail (see, Buswell 1935), his work was important in exploring the causal relationship between high-level brain function and eye-movements (Yarbus 1967). His experimentation was performed via the viewing of an image with different intended tasks. Test subjects were initially shown a picture for three minutes and given no specific instruction other than to observe the image. Following this, six other viewings were performed on the same image, but with differing instructions. The results of his research showed notable differences in the patterns produced between tasks. Yarbus therefore concluded that:

“[the eyes] *fixate on those elements of an object which carry or may carry essential or useful information*”
(Yarbus, 1967, p. 211).

Therefore, based on changing tasks, the information provided by certain regions to a person will alter. This will cause them to dynamically adjust their scan and fixation patterns. As a result of this it becomes possible to make deductions about a person’s reasoning or the information of value to them based upon their eye movements.

Given the importance of Yarbus’s original work and the amount of research built upon it, it was recently re-investigated. The investigation was performed with more understanding about the need to consider the bottom-up component. It essentially vindicated the conscious control of major fixations (Deangelus and Pelz 2009). Also, as eye tracking gains popularity in research, the applications continue to show that the top-down theory can be taken advantage of (Einhauser, Rutishauser, and Koch 2008; Rothkopf, Ballard, and Hayhoe 2007; Corsato, Mosconi, and Porta 2008; Adams, Witkowski, and Spence 2008).

2.3 Research Hypothesis

Most likely, the truth lies in a combination of the two theories. Points of interest in a scene are determined via low-level features whose detection is a hard-wired component of our visual system. From this determination of valuable points in a scene, the user consciously decides

which features to look at to obtain the information required. This is most likely supported by subconscious saccades driven by low level features, to gather surrounding information around the consciously chosen target point. For “saliency-map” predictions of eye movement, including consideration of the visual task work, researchers are beginning to evaluate and combine both top-down and bottom-up techniques (Deubel and Schneider 1996; Itti and Koch 2001; Canosa 2005; Peters and Itti 2008).

3. EYE TRACKER SYSTEM

The Mobile Eye Tracker by Applied Science Laboratories was used in this paper. The tracker uses two video cameras and three infra-red light emitting diodes (LED). The two cameras and LEDs are mounted on a pair of standard safety-type glasses, in which, over one of the eyes there is a circular cutout in the perspex lens. This cutout allows for the placement of an adjustable monacle that reflects the infra-red LED light onto the eye surface. The LEDs are arranged in a triangular pattern which is then reflected off the eye and captured by the first video camera. Comparison of the positions of the LED reflections on the eye and the pupil in one of the video cameras determines gaze. The second mounted camera is forward facing and provides an image that represents the subject’s field-of-view (FoV). The technical specification of the eye tracker is shown in Table 1.

Component	Details
Sampling and output rates	30Hz
Measurement principle	Pupil-corneal reflection with custom outdoor enhancements
System accuracy	0.5° visual angle
Resolution	0.1° visual angle
Head movement	Unlimited
Visual range	50° horizontal and 40° vertical

Table 1: Technical specifications of the ASL Mobile Eye Tracker.

Since the Mobile Eye Tracker is designed for mobile applications with no restrictions imposed on the subject’s head movement, it is necessary to adapt the eye tracker to map the eye tracker output to provide the X-Y gaze coordinate relative to the data display monitor as discussed in the next section.

3.1 Eye Tracker and Display Mapping

Due to the nature of operation in a mobile eye tracker, the coordinate points provided are only relative to the video captured from the FoV camera at the time. The data required though for this work, are X-Y coordinates relative to a fixed object (the computer display). This is a particular problem, for example, when the subject has their gaze fixed on the centre of the screen. The eye tracker will give a certain set of X-Y values. If they move their head while maintaining their gaze on the centre of the screen, the eye tracker will give different X-Y coordinates. Based on the X-Y values alone it would appear as though the subject has

moved their point of gaze when in fact they have kept their attention on the same point on the screen and only moved their head. Therefore it was necessary to map these data points provided to be relative to the computer display. The overall steps taken to resolve this issue are as follows:

1. Automatically locate the screen within each video frame.
2. Transform the video frame specific X-Y coordinates to screen relative coordinates.
3. Match the eye tracker data with the image being displayed at the time.

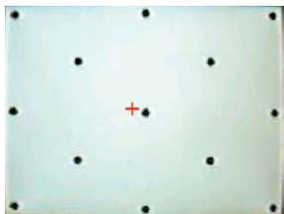
The screen detection algorithm comprises the following frame processing techniques; intensity segmentation, region labelling and line fitting, the details of which can be found in Chadwick (2009).

The mapping of the screen was performed through the use of transformation matrices. A transformation would be obtained that describes the mapping of the screen corners to fill a full 640×480 image such that the distorted screen image captured in the FoV becomes transformed to be the same size and aspect ratio as the true image being displayed at that moment. This transformation can then be used to transform any FoV coordinate into a screen relative coordinate for that frame. A linear transformation of the screen points proved to be insufficient to correctly map the screen contents to the full frame. In order to use a polynomial transformation a minimum of 14 points were required. Intermediary points were therefore interpolated to the screen along the edges in order to support the use of a third-order polynomial mapping. Using this polynomial transformation, the eye tracker's coordinates in the FoV could then be mapped to on-screen coordinates. Figure 1 shows the mapping results.



(a)

A single frame from the FoV video with the eye tracker position overlaid as a red cross-hair.



(b)

The identified and transformed screen content, with the inferred eye tracker position overlaid as a red cross-hair.

Figure 1: Image illustrating the accuracy of the screen detection and mapping of the FoV relative point.

4. EXPERIMENT PROCEDURES

In this experiment, subjects were required to wear the eye tracker and observe the displayed dataset at a distance of about 0.6m from the screen. The eye tracker tracked each subject's eye movement in real time in order to identify the locations of focus with respect to the FoV camera. The mapping operations described in Section 3 were used to transform the tracker FoV coordinate to the pixel location of the displayed image.

Data

In our experiment, an aeromagnetic dataset, specifically a total magnetic intensity (TMI) grid from the granitoid greenstone terrain in the Kirkland Lake area in Ontario, Canada (Ontario Geological Survey, 2003), and an enhanced image of the corresponding data were used.

Aeromagnetic data (Blakely, 1996) is regularly used by the mineral exploration industry for exploration targeting. It provides information on the distribution and relative abundance of magnetic minerals in the Earth's upper crust, allowing mapping of geology in regions that are poorly exposed. Within aeromagnetic data, certain geological features can be mapped based on characteristic affects on the Earth's magnetic field. For example, faults appear as linear discontinuities within the data, and kimberlite pipes (a form of volcanic rocks) appear as circular shaped anomalies.

When analysing aeromagnetic data, various forms of edge/discontinuity enhancement techniques are often applied to aid manual interpretation. One of the most commonly used enhancement techniques is called the first order vertical derivative (1VD) filter (Blakely, 1996). This estimates the magnetic amplitude change with respect to the changes in the distance to the source of the magnetic anomaly. This method takes into account that deep sources cause lower frequency signals within the data whereas the shallow sources cause higher frequency signals.

Through examining the eye movement of the interpreters in their data interpretation process, we analysed the variation in different geoscientists' abilities to spot features of significance and the effectiveness of the 1VD enhancement method. The experiments were conducted on six volunteer geoscientists, each with different skills and expertise.

The experiment consisted of two separate processes. The first process was to examine how the subjects observed the aeromagnetic map when freely viewing the data. The instruction given to the subjects was to make a general assessment of the area. Both the TMI data and the vertical derivative data (1VD) of the same region were displayed for a duration of three minutes. The TMI and 1VD data are shown in Figure 2 and 3 respectively.

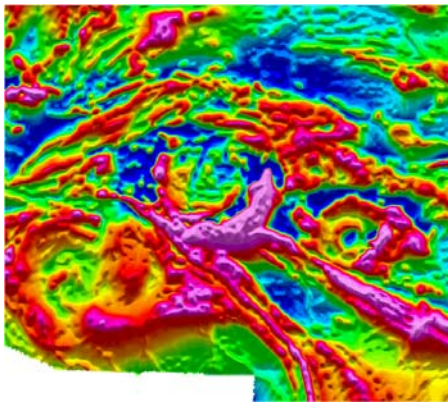


Figure 2: Total Magnetic Intensity (TMI) of the aeromagnetic dataset from granitoid greenstone terrain in Canada (Ontario Geological Survey, 2003).

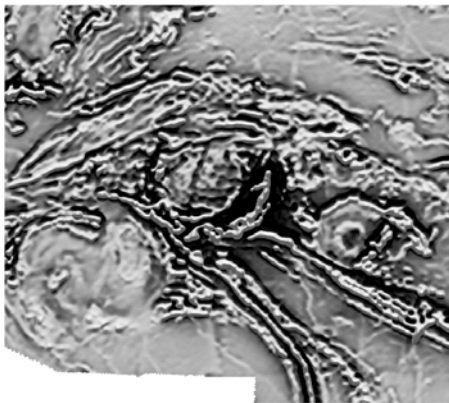


Figure 3: The Vertical Derived (1VD) data (Ontario Geological Survey, 2003).

The second set of experiments focused on task driven viewings. In these tests the subjects were given a specific feature to search for within the image. They were not required to give any physical or verbal response indicating success at locating the features (it was purely a mental exercise). These are the tasks the subjects were instructed to perform:

- Determine the location of any granitoid intrusions (these appeared as large circular feature with different magnetic responses from the background).
- Determine the location of any faults (these appeared as linear discontinuities).
- Determine the location of any kimberlite pipes (these appeared as small circular elevation or depression features).

As the subjects had already performed the free view experiment and became familiar with the data, each of the

tasks was allocated only two minutes. Each of the tasks was performed for the TMI followed by the 1VD image.

5. RESULTS

The following information was used to analyse the eye tracking data:

- Heat maps - using all the viewpoints cumulated over time, including non continuous viewing of areas, a heat map was generated for each test and for each person.
- Total location visits - this was calculated by dividing the image into a 20×20 block grid. The data for that image was processed to determine how frequently an element of the image was visually visited, irrespective of continuity.
- Number of fixations - using the 20×20 grid, the number of times an element was fixated on was obtained. This was simply determined by cumulatively logging the times an element was viewed for an extended period of time and then fixation changed.
- Time spent - using the 20×20 grid, the time actually spent on each element was obtained. This measurement was to remove the time spent when the eyes were travelling between points. This was calculated by determining the fixations and only storing the time relating to them.
- Change in pupil dilation - using the 20×20 grid, the standard deviation in the pupil dilation for each element was calculated. The pupil dilations were baselined to the mean pupil dilation. This measurement gave an indication of how much pupil adjustment occurred when viewed a particular feature.

The findings in relation to the expertise of the different subjects are summarised here. In terms of the performance of test subjects in the free-view experiment, experience influenced the subject's visual search technique. Specifically it was apparent that those with greater experience were more methodical in their search. They tended to examine a wider range of areas within the image, including the areas with subdued magnetic responses. In contrast, those with less experience were more inclined to focus on obvious features, rather than searching for more subtle features.

The ability of test subjects to detect faults regardless of the visualisation used was clearly strong. All subjects identified key features, although it was noted that different people tended to focus on different faults. There was, however, some evidence to suggest that the different kind of features

observed or focused on may be related to experience. Also less experienced subjects were more likely to miss less linear faults. This is most likely a direct result of a less methodical search technique.

In comparing the 1VD image enhancement filter's effects on the features identified or viewed, most participants viewed the two images (1VD and TMI) in a similar fashion. This indicates that little improvement was provided by the image filter. Despite this, two test subjects both detected faults in the 1VD image not seen previously by them in the TMI image. Otherwise no specific effects were seen regarding the identification of granitoid intrusions or kimberlite pipes.

Some examples of the heat maps can be seen in Figure 4 and 5. In these images the different rows in the figures are the different tasks a subject performed and the columns relate to the image being viewed (Magnetic - TMI or Filtered - 1VD). The heat map colours represent weakly viewed areas as light blue, through green, yellow, red and then white for strong viewing. The vertical differences between images relate to the task, whereas horizontal differences between images relate to the image type.

CONCLUSION

We reported on a preliminary study to investigate the use of an eye tracker system to quantitatively analyse the behavioural patterns associated with target spotting involving aeromagnetic dataset. Six geoscientists with various skills and expertise were involved in the visual analytic tasks. We found that those with greater experience were more methodical in their search for targets, usually scanning larger areas compared to the less experienced subjects. The ability of test subjects to detect faults, with or without the enhancement method was clearly strong. All subjects identified significant features, although it was noted that different subjects tended to focus on different features. Future work will involve: the study of the influence of other pre-processing and enhancement techniques on the accuracy of feature recognition; more detailed analysis of the target spotting ability and process by improving the experimental design and possibly incorporating interviews of the participants in the analysis. We would also like to further quantify and evaluate the impact specialised training has on a test subject's ability to accurately recognise features of significance.

ACKNOWLEDGMENTS

We would like to thank the volunteer geoscientists at CET at UWA, who participated in this experiment and the Ontario Geological Survey for the permission to use their data in this research. We also like to acknowledge and thank iVEC (www.ivec.org) for the loan of the eye tracker.

REFERENCES

- Adams, N., Witkowski, M. and Spence, R. (2008). The inspection of very large images by eye-gaze control. In AVI '08: Proceedings of the working conference on Advanced visual interfaces, pp. 111–118.
- Blakely, R.J. (1995). *Potential Theory in Gravity and Magnetic Applications*. Cambridge University Press, New York.
- Bond, C. Gibbs, A., Shipton, Z. and Jones, S. (2007). What do you think this is? "Conceptual uncertainty" in Geoscience Interpretation, in GSA Today, 17:11, pp. 4–10.
- Buswell, G. T. (1935). *How people look at pictures: A study of the psychology of perception in art*. Chicago: University of Chicago Press.
- Chadwick, T. (2009). *Improving and Expanding Rapid Image Sorting Through the Use of Eye Tracking*. B.Sc Honours Thesis, Department of Computing, Curtin University of Technology.
- Corsato, S., Mosconi, M. and Porta, M. (2008). An eye tracking approach to image search activities using rsvp display techniques. In AVI '08: Proceedings of the working conference on Advanced visual interfaces, pp. 416–420.
- Deangelus, M. and Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition* 9999 (1), pp. 1–22.
- Einhauser, W., Rutishauser, U., and Koch C. (2008). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision* 8 (2), pp. 1–19.
- Koch, C. and Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology* 4, pp. 219–227.
- Ontario Geological Survey (2003). *Ontario airborne geophysical surveys, magnetic and electromagnetic data, Kirkland Lake area; Ontario Geological Survey, Geophysical Data Set 1102 - Revised*.
- Rothkopf, C. A., Ballard, D. H. and Hayhoe, M.M. (2007). Task and context determine where you look. *Journal of Vision* 7 (14), pp. 1–20.
- Vilarino, F., Lacey, G., Zhou, J., Mulcahy, H., Patchett, S., 2007, *Automatic Labeling of Colonoscopy video for cancer detection, Pattern Recognition and Image Analysis, Lecture Notes in Computer Science, Springer Berlin/Heidelberg, Vol 4477, pp. 290-297*.
- Yarbus, A. (1967). *Eye movements and vision* (B. Haigh, Trans.). New York: Plenum Pres.

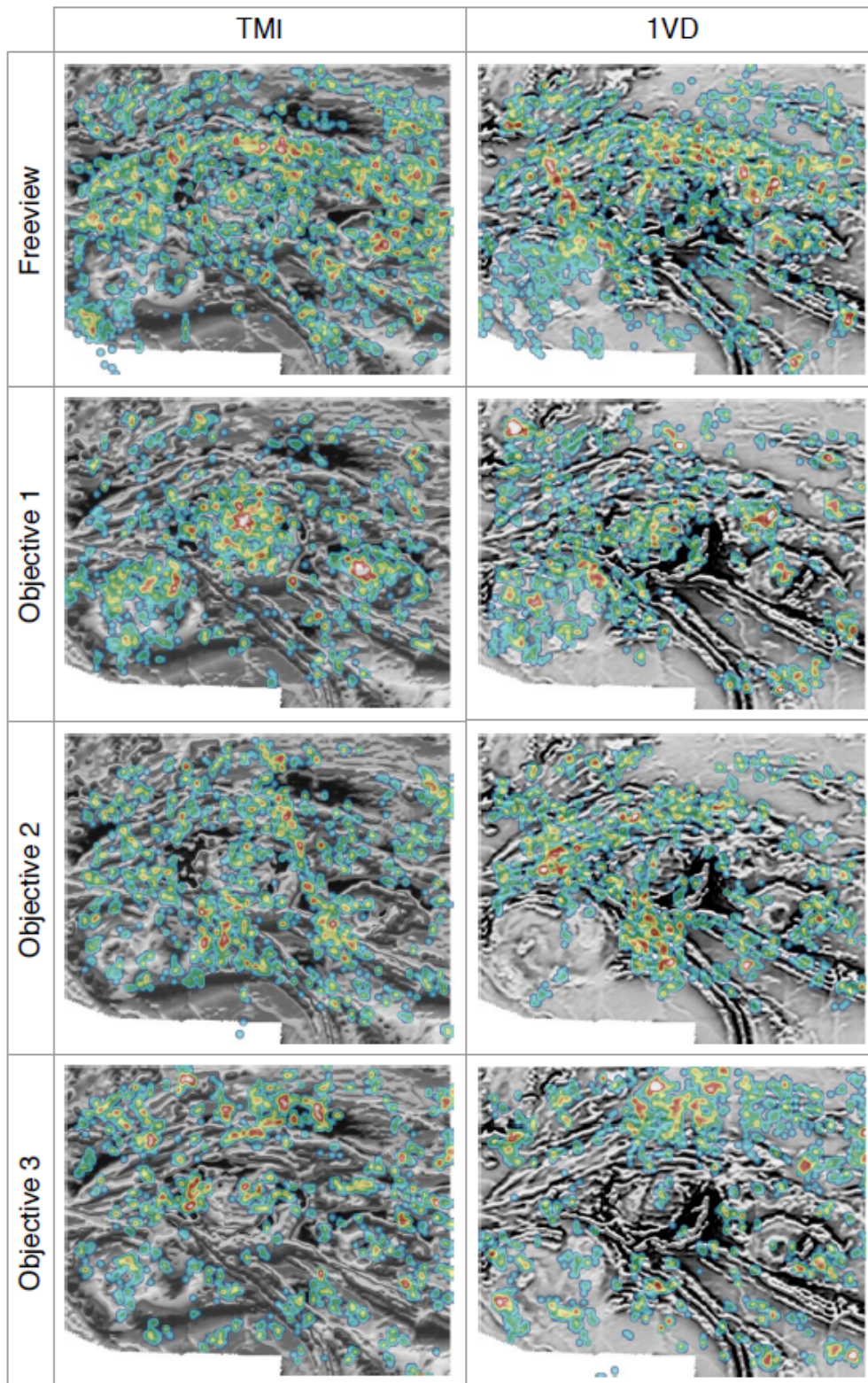


Figure 4: Heat map of the various tasks from an expert subject, overlaid on the TMI and 1VD data. Note that the TMI and 1VD data were shown in colour and in grayscale, respectively, in the experiment, but here, both images are represented as grayscale images as background for the heat maps.

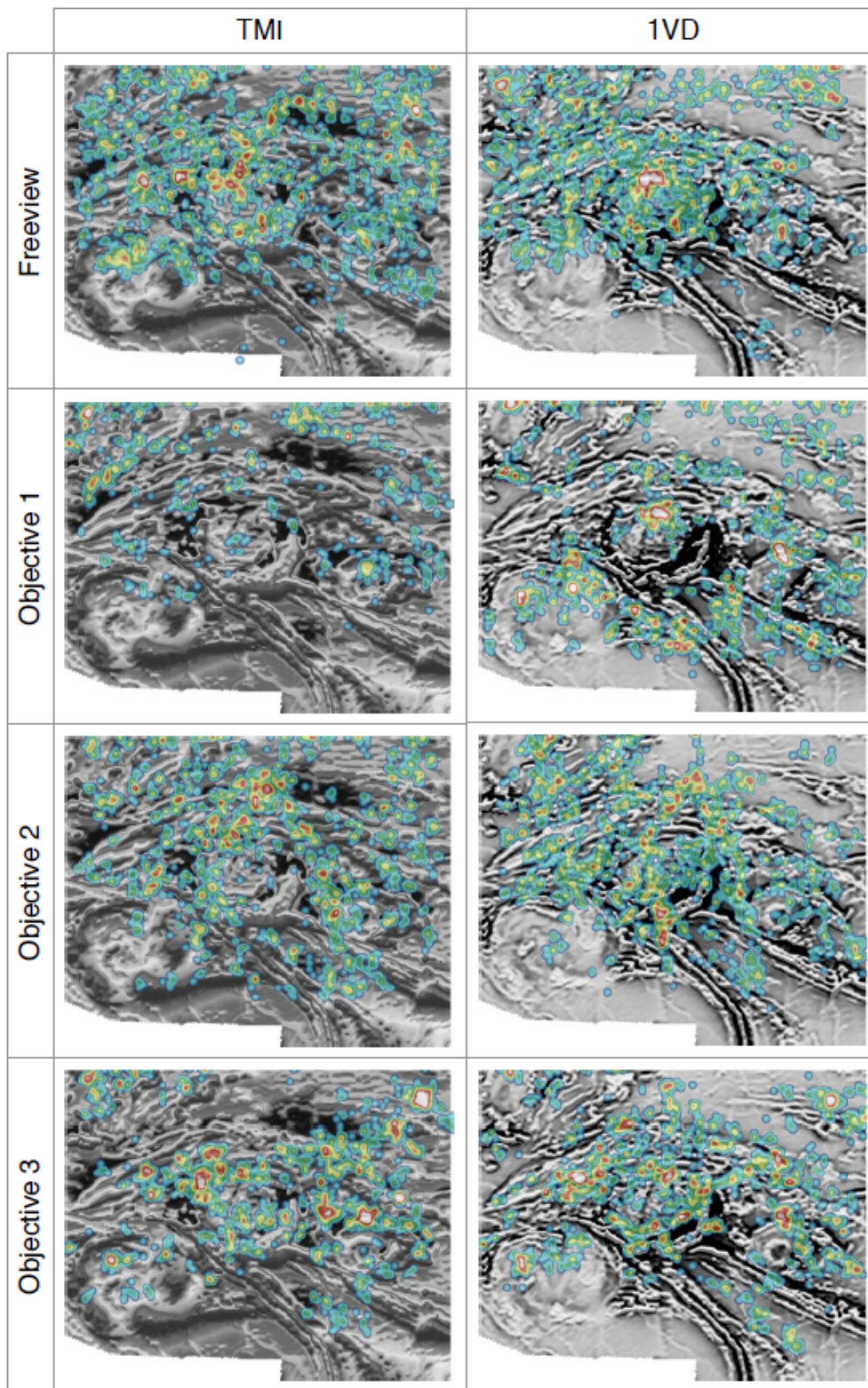


Figure 5: Heat map of the various tasks from a less experience subject, overlaid on the TMI and 1VD data. Note that the TMI and 1VD data were shown in colour and in grayscale, respectively in the experiment, but here, both images are represented as grayscale images as background for the heat maps