Using Cognitive Task Analysis and Eye Tracking to Understand Imagery Analysis

Laura Kurland, Abigail Gertner, Tom Bartee, Michael Chisholm and Scott McQuade

The MITRE Corporation
202 Burlington Rd
Bedford, MA 01730
{lkurland,gertner,bartee,chisholm,smcquade}@mitre.org

1   INTRODUCTION

The National Geospatial-Intelligence Agency (NGA) is the national-level producer of Geospatial Intelligence, serving both policy makers and DoD elements. One core task of Geospatial Intelligence Analysts is to develop intelligence through the exploitation of imagery (including overhead, airborne, and video sources), with geospatial data and additional intelligence sources supporting the analysis process. Currently there is a gap between the exploitation and analysis capabilities of senior analysts and incoming junior analysts. Some expert knowledge is not well documented or "institutionalized," creating a barrier to knowledge transfer either through classroom training or on-the-job learning.

To begin addressing these concerns, we conducted a study of analysts performing a realistic area search task. In a collaborative effort, we have worked with the NGA/Production organization and they have provided us with analysts to support the project. One outcome of this work is a structured knowledge representation of the goals and the search strategies that analysts use in the process of solving typical Geospatial Intelligence Analysis problems. Furthermore, by studying analysts with a range of experience, from novice to expert, we have identified strategies that are typically associated with "expert" performance and that less experienced analysts may need help to develop. We have also recorded the eye movements and tool usage of analysts during imagery search and have used these to study the analysts search behavior in detail.

2   EXPERIMENT

Using a Cognitive Task Analysis (CTA) framework for knowledge elicitation and "think aloud" protocol (Meichenbaum & Asarnow, 1979; Merrill, 1994), our cognitive research focuses on several aspects: strategies and issues that novice, journeyman and expert Imagery Analysts (IAs) think about and the tool operations they use while doing a real-world area search task and the nature of their domain understanding. This effort encompasses several components:

- Instrumented Testbed: Development of an instrumented environment to capture qualitative and quantitative data by recording time-stamped eye tracking, screen capture of an Electronic Light Table application for imagery analysis (Remote View), keyboard and mouse clicks, "think-aloud" audio protocols, and video recordings (over the shoulder) to capture pointing gestures.

- Levels of Knowledge: Assessing the development of domain understanding by interviewing subject matter experts (SMEs) and comparing their understanding to that of novices and journeymen. In this case, the domain chosen was thermal (non-nuclear) power plants because the information was publically available.

- Problem Development: Development of a suitable problem for broad area search. After several pilot experiments using MITRE journeymen image analysts, a problem set was developed (with the SME's assistance) that would provide a good performance spread, containing items that were rated by the SME as easy, fairly difficult and very difficult to find – where difficulty could be determined by a combination of size, manipulations needed to obtain image clarity and functional domain understanding.

From Expert-Novice studies in various domains conducted in the past 20 years (physics – Chi et al, 1988; Feltovich et al, 1997; situational awareness- Endsley & Garland, 2000), it is well known that novices tend to see just what is in front of them and their understanding is shallow, mostly at a descriptive level (e.g., a "pulley problem" as opposed to an acceleration or force problem) and/or it appears to be procedural in nature ("first you do this, then you do..."). For this project, several experimental tasks were designed to characterize novice, journeyman and expert imagery analyst domain knowledge, including functional understanding, tool use (procedural knowledge) and how they went about searching an image. In addition, analysts were interviewed about their background and about their day-to-day work tools and sources, as much as could be discussed in an unclassified context. The experiments are described in the following section.

2.1   Subjects

Subjects were volunteers and came from two specialty areas: Petroleum and Industry (see Table 1). Six NGA analysts participated in the study. Two MITRE analysts also participated in the pilot experiment with the same format and materials. However the pilot search image was not as clear due to a technical problem. Their data are included in the results, except where noted. In our analysis, subjects are categorized according to years of experience, field specialty and agency. Novice is defined as less than 4 years of experience, Journeyman is defined as 4-10 years of experience and Expert is defined as more than 10 years of experience.

Table 1: Experimental Subjects

Expertise	# Years	Field, Specialty	Agency
Expert	25	Petroleum	NGA
Novice	<1	Industries	NGA
Expert	17	Industries	NGA
Novice	<1	Petroleum	NGA
Expert	20	Petroleum	NGA
Journeyman	8	Industries	NGA
Pilot subjects			
Journeyman	4 as an IA; 10 total	Sensors	MITRE
Journeyman	6 as an IA; 11 total	photogrammetry	MITRE

2.2 Area Search Task

The imagery used for the experimental Area Search task was selected to be fairly representative in size, difficulty and would take approximately an hour. A single wide-area, unclassified image (1m CIB®) was of an urban area of Baghdad (from year 2000). The image file was opened in Remote View (RV), version 2.0 in Windows XP. There was a small overview image to the left of the main window. (Note: The original imagery used in this experiment has been substituted with imagery from Google Earth™ and Google Local™ throughout this paper.)

The image contained several thermal (oil-powered) power plants varying in difficulty that would give a performance spread among novices, journeymen and even experts. Difficulty was determined by several parameters: size, type of image enhancements or manipulations to increase visibility of specific telltale indications, and location. There were multiple foils that had characteristics which could be mistaken as a power plant, such as water treatment plants and various light industries. The image was of an urban area, which is more difficult to search than one with fewer buildings and man-made structures.

Subjects were told that their task was simply to find any power plants that they could in the image. They were not given any additional information, such as what country the image was from nor if the image was georeferenced, which it was not. Subjects were instructed to ask the experimenter questions so that we could find out *what* information was important, the *order* in which the information was requested, and *how* they used that information.

2.3 Results

From our discussions with and observations of IAs performing the area search task, we were able to identify a number of factors that contribute to a successful search. These include: knowing what items to look for and what items to rule out; having the ability to follow the search paths with the highest probability of success based on domain knowledge, country knowledge, and features of the overall image; and

skillful use of the available tools to adjust the image's appearance. Search procedures are meticulous and usually involve multiple images of the same area as well as additional intelligence sources in the production of the final report. Expert IAs were found to have much more high-level knowledge they could apply in performing the search task. Additionally, the nature of their expertise is closely aligned to their domain area, and it is clear the strategies and techniques are domain specific (e.g. power plants), and do not necessarily apply to another closely related domain (e.g. petroleum). Finally, experts outside of the power plant domain did not employ the software tool to their best advantage and were not necessarily the most successful in finding the harder power plants in the search task.

2.3.1 Concept Map

One of our goals was to capture the nature of area search: to determine the order and importance of specific information with respect to the image, the country it came from, what are important telltale indications, and how the information affects an analyst's detection and determination of an item of interest, which is in this case a power plant. Furthermore, we also wanted to determine how an analyst's functional domain understanding affects the detection and determination of power plants in imagery.
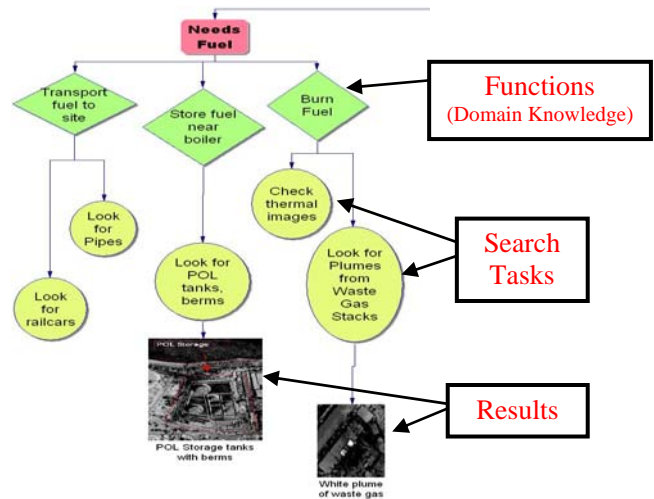

Figure 1: Concept map of functional domain knowledge, search tasks and image results (images are from Google Earth™).

Based on these observations and interviews, we have combined the functional domain knowledge, procedural tool manipulations (Figure 1), and strategic search knowledge (Figure 2) into a concept map representation (Novak, 1998) to assist less experienced analysts in conducting more effective and efficient searches. Moreover, the representation includes links to background information that explains key aspects of the domain, procedures and strategies.
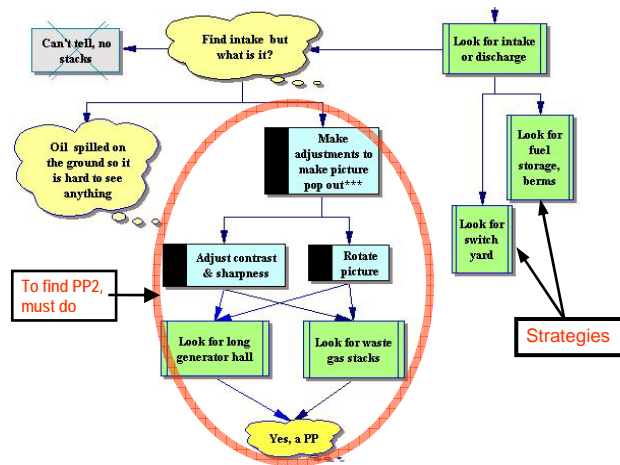
Figure 2: Analyst search strategy attempts

2.3.2 Eye tracking

During the broad area search task, we tracked subjects' eye movements using a remote desktop-mounted eye tracker (SMI's iViewX). The eye tracking served several purposes in the experiment: (a) to record the search paths of analysts during the experiment and determine features of successful vs. unsuccessful search, (b) to observe the proportion of eyes-on-imagery as well as the analysts' use of the toolbars and overview image window, and (c) to observe the differences in low-level eye movement behavior in experts and novices.

2.3.2.1 Search Patterns

The eye tracker gives us information about where on the screen the user is looking. To study the search patterns reflected by analysts' eye movements it was necessary to convert the XY screen coordinates coming from the eye tracker into XY pixel coordinates relative to the image being searched. This was done by getting information from RemoteView about what portion of the image was currently visible on the screen (taking into account zoom level, rotation, and position) and performing the necessary transformation on the gaze position coordinates. This information was stored as a "snail trail" representing the analyst's search path through the entire image.
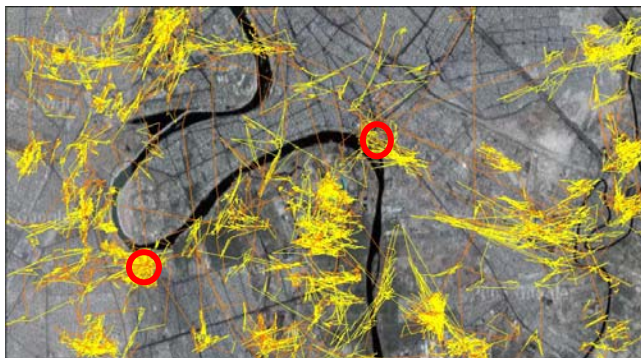

Figure 3: Recorded search path for S1 (circles indicate the locations of power plants). Image is from Google Local™.

Figure 3**Figure 3** shows the scan pattern of one subject in our experiment (S1) over the entire image. When we examine the search paths of the six IAs over the entire image, what we find (not surprisingly) was that the analysts who performed the most strategic search found the power plants the soonest. The best strategy was to start by searching along the river that passes through the middle of the image, since power plants tend to need a source of water for cooling. What is slightly surprising is that the most successful analysts were not always the experts. The novices tended to search more strategically than the experts and in fact both novices and the journeyman found both of the power plants (although one novice failed to identify one of them as a power plant), while two of the expert analysts missed power plant #2 (however, both of these were experts in petroleum, not industries, which is the domain area that would be most relevant for finding power plants). Note that while S1's eyes passed over the location of power plant 2 several times, he never fixated in that area (i.e., his eyes didn't stop moving long enough to see the power plant).

2.3.2.2 Eyes-on-imagery

To evaluate the proportion of eyes-on-imagery, the screen area was divided into several areas of interest: the main image view, the toolbar above, the menu, the overview window to the left (which shows a zoomed-out view of the entire image), the area above the image that was not part of the toolbar, and the area below the image. Table 2 shows the percentage of time spent looking at the different areas of the screen. In general, the proportion of eyes-on-imagery was quite high (average 88.2%), while the time spent viewing the toolbar (2.4%) and the overview (4.6%) was rather low. This is a positive result as it means that in general the interface is not distracting the analysts much from spending time viewing imagery. The number of fixations on the menu was close to zero, but because the area of the menu is so small it is within the error range of the eye tracker.

Table 2: Percentage of time spent on areas of the screen

	image	toolbar	overview	menu	other
S1	0.857	0.003	0.009	0	0.134
S2	0.934	0.014	0.032	0	0.034
S3	0.895	0.042	0.036	0.002	0.069
S4	0.851	0.044	0.036	0	0.111
S5	0.835	0.026	0.133	0	0.033
S6	0.921	0.016	0.031	0	0.048
Avg	0.882	0.024	0.046	0.000	0.038

Another interesting question is whether the subjects' use of the various interface features correlated with their success on the search task. In fact, there is a significant correlation between the percentage of time spent looking at the image and success at locating the power plants (t=-5.22, df = 5, p<.05). The average percentage of time spent looking at the image for unsuccessful searches was 84.7%, while the average for successful searches was 91.7%.

2.3.2.3   Low-level eye movement behavior

To explore the differences between the eye movements of experts and novices, we decided to look at three quantitative features of the analysts' eye gaze: fixation duration, distance between fixations, and frequency of fixations. For the purposes of this analysis, we grouped the journeyman (8 years experience) with the experts. The mean fixation duration for the expert subjects was 189.32 ms while the mean for the novices was 259.85 ms. This difference is significant according to a two-tailed t-test (t=-6.39, df=4,p<0.05). Likewise, there was a consistent difference in the length of saccades, with the average for experts being 52.33 mm and the average for novices being 64.75. This difference was not significant, however, due to the small sample size (t=-1.91, df=4, p>0.1). On the other hand, the average fixation frequency is about the same for the two groups: 188.63 for experts and 186.5 for novices. These results suggest that experts tend to take in what they are looking at more quickly, and that they move in shorter jumps from location to location, which may result in a more efficient search pattern. However, in light of the fact that two of the experts were the only subjects who didn't find all of the power plants, it is important to note that this more efficient low-level eye movement does not necessarily translate into improved performance at the macro level.

2.3.3   ELT Tool Usage

In addition to the subjects' eye movements, we were also interested in how they used the ELT software. We wanted to know how they manipulated the imagery in order to prepare it to be searched. This includes the use of image enhancement tools, such as adjusting brightness and contrast, as well as zooming in and out and rotating the image. To collect this information, we used a tool called Epiplex (http://www.epiance.com), which allowed us to extract a record of widget manipulation events from the ELT software without having to modify or interact with the software itself in any way.

Table 3: Summary of ELT tool usage

	Bright-Ness	Contrast	Haze	Sharp-ness	DRA	Zoom	Rotation
S1		4	3	9	3	202	12
S2	1	4	2	6	30	163	9
S3	2	15	2	5	1	35	2
S4						60	3
S5	3	2	2	3	2	16	4
S6						31	

Table 3 summarizes the number of times each of the subjects used a number of common image manipulation functions: brightness, contrast, haze, sharpness, DRA (dynamic range adjustment) , zoom and rotation. One thing to note here is that there is a categorical split between analysts who used the image processing tools and those who didn't. Those analysts who used the tools tended to use all of them, while two of them did not use the image adjustment tools at all. Several of the analysts seem to have a "favorite" image manipulation

(e.g. contrast for S3 and DRA for S2). Another interesting observation is the difference in the use of the zoom tool. Some analysts zoomed in and out frequently, while others did hardly at all. There is no correlation between the *frequency* of zooming and success on the search task. However, we did find a significant relationship between *zoom level* and target identification: the average zoom level when a target was identified was 1:1.498, while the average zoom level when a target was passed over but *not* identified was 1:2.31 (t = -2.18, df = 13, p < .05). In other words, a contributing factor to whether or not a target is identified seems to be whether or not the analyst was zoomed in enough to be able to see crucial identifying features clearly.

3   CONCLUSION

We have described an experiment that we conducted to *make explicit* the processes and knowledge that expert Image Analysts employ. As a result of our work, we identified successful and unsuccessful strategies for performing imagery search in a particular sub-domain, as well as codifing the functional and procedural knowledge necessary to work as an analyst in that domain. Ultimately, our goal is to use this knowledge to develop tools that can support novice analysts in learning their craft and help to make the job of more experienced analysts more efficient. Thus we are applying the heuristics and strategies learned through this knowledge elicitation process to the development of new training products, as well as a software tool called "Intelligent Image Queuing" which uses expertise encoded from analysts to help prioritize incoming imagery to be searched.

4   ACKNOWLEDGMENTS

5   REFERENCES

Chi, M.T.H., Glaser, R. & Farr, M.J. (1988). The Nature of Expertise. Hillsdale, NJ: Lawrence Erlbaum.

Endsley, M and Garland, D. 2000. Situation Awareness Analysis and Measurement. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers

Feltovich, PJ, Ford, KM, Hoffman, R.R.(Eds.), 1997. Expertise in Context. MIT Press.

Meichenbaum, D. & Asarnow, J. Cognitive-behavioral modification and metacognitive development, 1979.

Merrill, M. David (1994). Instructional Design Theory. Englewood Cliffs, NJ: Educational Technology Publications

Novak, J. D. (1998). Learning, creating, and using knowledge: Concept maps(R) as facilitative tools in schools and corporations. Mahweh, NJ, Lawrence Erlbaum Associates.