2011

Is there a core set of skills for visual analysis across different imaging technologies?

Craig HM Donald
Edith Cowan University

DOI: 10.4225/75/57a016a1ac5c5
Originally published in the Proceedings of the 4th Australian Security and Intelligence Conference, Edith Cowan University, Perth Western Australia, 5th -7th December, 2011

IS THERE A CORE SET OF SKILLS FOR VISUAL ANALYSIS ACROSS DIFFERENT IMAGING TECHNOLOGIES?

Craig H.M. Donald
Edith Cowan University, and Leaderware
craig.donald@leaderware.com

Abstract

This research examines the technological challenges posed by security imaging technologies for human visual analysis of images. Imaging technologies are increasing becoming part of an overall security strategy that incorporates a range of camera technologies, x-ray technologies, and other electromagnetic imaging such as millimetre wave and terahertz based systems. Still and video image types are increasingly becoming presented to viewers or screeners in forms that are only representative in nature and highly abstract, and the use of filters is increasing the complexity of interpretation. Despite a range of factors that are being looked at to enhance visual analysis, the contribution of individualised image processing skills is poorly understood and recognised. The paper explores examples of how an assessment exercise which examines visual analysis, ScanX, correlates against performance in four major studies set in different environments and using both x-ray and CCTV technologies. Correlations show strong relationships to performance despite the difference in image technology and environmental settings, as well as detection targets and criteria. Support for a set of core image analysis skills that can be used across a range of technologies by a common operator group is shown by the research. These skills appear to relate more to the nature of processing applicable to various forms of image rather than the image content itself.

Keywords

Imaging, imaging technology, visual analysis, image analysis, closed circuit television (CCTV), x-rays, operator performance, selection

THE IMAGING TECHNOLOGY CHALLANGE

Modern security systems are increasingly composed of a number of visual imaging technologies at different flow points of the processes being checked and monitored. Whether one is looking at airports, diamond mines, national key points, or even more conventional sites such as shopping centres, a combination of electronic imaging technologies are being used to review people, goods, materials, areas and even processes and movement (Donald 2010). Traditional CCTV cameras, infrared cameras, x-rays of goods, luggage and carry on bags, and penetrative and backscatter full body x-ray technologies are already widespread in operations internationally. This trend is likely to continue as emerging electromagnetic technologies provide increasing capabilities to view electronically at different wavelengths to the human eye. These can include thermal, millimetre wave and terahertz based technologies that are already commercially available (Binstock && Minukas, 2010).

The development of these technologies has seen an increasing divergence from traditional images on screen. For example, a CCTV camera and a thermal camera present very different visual images on screen of the same physical presence especially under different lighting conditions. Moving to a terahertz based system provides a general body shape reflecting blobs of various wavelengths of electromagnetic radiation. X-ray images of bags, other goods, or even people to evaluate illegal or threat conditions or to analyse content, can be viewed with filters that can have a substantial impact on the nature of the display to the viewer. The content moves from the real to the representative where graphic processing creates artificial colouring, changes image composition and removes or enhances only certain content which meets specific wavelength properties. In addition to these, there is an increasing implementation of imaging being used or trialled at other electromagnetic spectrums where the display is entirely synthetic. In all these cases, real items are being displayed in a representative form which may depart radically from the actual outline, shape, texture, surface composition, and form of the real object. The point of these imaging technologies is to identify various items from the way they impact on the wavelength which the measuring technique is using. The objective of the viewer is to analyse the information available and identify threats or problems in the images. This is calling for increasing sophistication in the visual analysis of such images on behalf of the viewer or operators.

VISUAL ANALYSIS DEMANDS

Substantial development cost has been spent on these established and emerging imaging technologies. While extensive efforts are being made to provide automated analytical tools to identify issues within images, these have had varying success. Undoubtedly there are successes in visual analytics or automated recognition systems. Further, we can expect increased presence and sophistication of image or video analytics in security technology. However, at present in some cases they can complicate the decision making processes rather than simplifying it. For example, in aviation x-rays highlighting a potential threat in a bag being scanned for weapons may cause a false alarm, but even more critically, it may distract a viewer or screener from other threats that have not been electronically identified in the bag. In other cases they are limited to basic conditions such as video analytics identifying movement, excessive grouping or contact of people, and basic postures. Identification of more subtle behaviours is beyond most "intelligent systems" at this stage, however. In most cases, human visual analysis of either original, processed, or alert activated images is still necessary.

Human factors research to address the human visual analysis process has addressed some of these visual analysis factors in a number of ways. Research on eye movement (Wickens & McCarley, 2008), concentration spans, spatial and other properties of targets, signal detection characteristics including timing (Huey & Wickens, 1993) are all contributors to detection performance. Aspects such as situational awareness (Sandom, 2001), goal driven attention, and task engagement (Cheyne, Carriére, & Smilek, 2006) may sensitise viewers to the context and task requirements of the detection process. The individual's level of arousal is seen to impact on the motivational intensity and the individual's level of alertness (Davies & Parasuraman, 1982), but there is not always a simultaneous change in detection efficiency during reduced arousal (Koelega, 1996). The role of attention sets in particular has been highlighted by Donald (2011) as critical in determining the frame of reference that operators use to guide observation. However, despite the extensive research into factors that affect an operator's approach to tasks, an understanding of the cognitive visual analysis process that they use as individuals is still poorly understood. Indeed, often the visual analysis term is used as a catch all to reflect matching, memory recall, or simplistic recognition rather than analysis.

Matching or simple recognition activities contrast sharply with the interpretation requirements expected from and implicit in many of the security imaging technologies. Indeed, the more the representative image diverges from the original item in shape, form and composition and the more abstract it is, the more complex the interpretation of such an image becomes. For human visual analysis, the context in which the indicators occur provides important information with which to identify the threat. Whether doing visual analysis of x-rays or CCTV, an understanding of context is critical. For instance, a behaviour in one setting may be perfectly normal, whereas the same behaviour in another setting may be seen as highly suspect. In reviewing the success of CCTV schemes, the Australian Research Council (2006) notes that "The effectiveness of CCTV may be very much dependent on a whole range of issues but in particular the monitoring strategies adopted by camera operators." (pg.ii).

Research on visual analysis also typically adopts a system orientation, in that it is assumed that if the controllable factors that influence how the person relates to the target (e.g, attention sets, frame of reference etc) can be standardised at an optimal level, and the technology itself is standardised, the performance of personnel will be similar. However, we are continually faced with issues such as why one doctor or radiographer spots a problem in an x-ray image and another with similar training and conditions can not. Similarly, there are variations in performance in a range of security settings between people from similar backgrounds, using the same equipment, and who received standardised training. We still get wide variance in aviation x-ray performance and in CCTV detection rates. In a study of X-ray baggage screeners sixty percent of detection errors were issues relating to decision making – screeners could see the threats but made erroneous decisions about them (Security General, 2002). Differences in detection effectiveness between people can occur even over short periods. The inability to make correct decisions is also therefore independent of time induced issues such as fatigue and vigilance. The advent of new imaging technology is not going to solve this, it may in fact make it even more critical as the degree of abstract conceptualisation required becomes greater as the image moves further from the original. Further, we find that analytics often relieve personnel of routine tasks, but increase the expertise required to analyse and decide on cases that get referred by analytics due to ambiguity, lack of resolution, or unrecognised conditions. One is therefore faced with the issue of individual differences in either aptitude, cognitive approach or functioning, or natural skills in which a person can apply the analysis process to the job context.

THE INDIVIDUALISED NATURE OF VISUAL ANALYSIS

If some individuals are better performers than others in such a task, the question then arises as to whether there is a set of core inherent competencies or aptitudes that assist individuals engaged in visual analysis tasks for image intensive operations. Further to this, can these core competencies be applied equally well to technologies producing very different images. Thirdly, can they make a substantial difference to the performance outcome. Fourthly, are they inherent to individuals, or can anyone acquire these skills through training or exposure. From an information technology perspective an additional question arises in asking whether they can be replicated through computerised processing of images.

Research by QinetiQ into cognitive factors in a standardised job profile applied to x-ray screeners highlighted six major factors (Department for Transport, 2003). These included perceptual speed involving the ability to make quick and accurate comparisons between objects, perceptual closure which facilitates identifying specific objects that are hidden or concealed in other objects, deducing involving the ability to apply existing rules and procedures to find solutions to novel problems, visualising which relates the to the ability to imagine how something would look once changes have been made to it, and decision making which involves the ability to make a decision where there is no clear answer or information is complete. Reaction time as reflected in the ability to respond quickly is the sixth factor identified in this respect. While these are standardised factors used for job classification purposes, they provide only a partial understanding of individual differences. However, if we take an instrument that appears to have predictive validity for performance in a variety of image analysis tasks and technologies, we can examine how the cognitive processes required for completion of the task in a technologically neutral way.

METHOD

This paper takes four research studies where such a widely used assessment instrument examining participant competencies was applied as part of the research process. These studies all relate to the effectives or enhancement of performance of x-ray screeners and CCTV operations in a number of international settings. Purposes include assessing predictive validity, or as part of the research design, assigning participants into control and treatment groups for research design purposes. The instrument is designated as ScanX in the case of the x-ray environment, or in the case of use in CCTV operations, the Scanning Exercise where it forms part of a more comprehensive assessment called SAMAE. The term of ScanX will be used for the remainder of the paper. Developed for the evaluation of personnel doing observation and detection tasks in high technology operations, ScanX is one of the few instruments for evaluating individual skills in visual analysis and detection that is used across x-ray, full body x-ray, and CCTV environments.

ScanX test items are a set of graphic objects positioned together in a containing shape which provides the standard background colour. It contains elements of colour, positioning of objects relative to one-another, grouping characteristics, rotation in space, the potential for subtle or obvious loss of information, a degree of pattern information, and the need to encompass the overall graphic as a whole, as the nature of the background shape can alter as well items contained within the shape. The graphic is generic in nature, relates to no standard industry image, and requires not situational awareness from previous experience. However, the process of visual analysis is critical to effective detection in the exercise. Participants have to review objects that appear respond when the target image is different to a reference image provided.

A number of characteristics are required by individuals during the continual processing of the graphic as it appears for each item over the 30 minutes of the exercise. Measurement of performance on the ScanX Exercise is based on two accuracy measures, a reaction time measured in seconds, and a false alarm rate. The accuracy measures include one identifying the percentage of target items within four seconds of them being displayed, and another identifying the percentage of target items detected as long as they are on screen (up to 15 seconds). The 4 second accuracy shows a high speed of processing, whereas the 15 second accuracy score shows sustained attention even after initial processing. The false alarm rate is critical in showing the ability to make accurate decisions about what is seen. A high false alarm rate shows poor judgement or guessing without an appropriate level of insight. The combination of accuracy and false alarms on the same measure produces an efficiency score based on how well or badly the participant does relative to a general distribution of scores on these two variables. A high Efficiency 4 second score for example will show a high accuracy of detection in conjunction with very few false alarms.

The research studies incorporating the ScanX exercise for which validation is available are detailed below.

Study 1. Conducted on 83 currently employed x-ray operators in two major UK airports as part of a UK Department for Transport study (Department for Transport, 2003). Performance was measured by threat image projection (TIP) detection rates on the job. TIP involves the projection into the x-ray image of life like threat objects such as guns, knives and explosive devices. It provides a true detection rate in that it is know how many TIP images were used and the precise response rate to these.

Study 2. An operating diamond mine and involving 16 CCTV surveillance operators who were all full time employees (Donald, 1998). Job performance was measured by levels of theft incident detection identified by the operator as part of the standard work duties, and a separate detection of safety and other violations of personnel on the job. The collection of incident data was based on one month of performance where operators were exposed to the same shift schedules. It is possible that CCTV operators were exposed to different levels of theft rates within their experience of work shifts although a month working period would be expected to even out this kind of exposure.

Study 3. An industrial research project in which a sample of 12 CCTV surveillance operators were drawn from five diamond mines in Southern Africa (Andrew, Landgrebe, & Donald, 2003). They were exposed to a incident detection task of identifying diamond theft behaviours over 3 hours of structured video material. All participants in the sample had experience in the plant process and the detection of theft behaviour was part of their standard job requirement. The footage included a mixture of incident and non-incident footage, either real or acted under realistic conditions. Personnel were given 5 minute breaks at the end of each hour. Job performance was measured by the detection rate of theft behaviour over the full duration of the video footage. As the simulation was designed to incorporate specified theft incidents, the detection requirement rate was known and provides an absolute measure of detection.

Study 4. A PhD research study involving 73 CCTV operators drawn from 5 different companies in the gaming, mining, chemical and town centre environments (Donald, 2011). They were exposed to 90 minutes of continual video material of a plant environment where they had to identify designated physical actions rather than theft behaviour. No participants had any previous exposure to the plant environment in which the video materials were drawn. The video comprised a mixture of video clips, some with no target behaviours and others containing one or more target actions. Actions such as picking up an object, flicking with a broom, kicking and swivelling feet were specified as the target behaviours that had to be detected.

The four studies involved have very different task performance measures and these are gathered in very different settings. In study one, it is detection of aviation security threats in carry-on baggage x-rays, in study two it involves detection of various behaviours in full time working conditions in a general mining area, in study three recognition of theft behaviour occurring in a processing plant that is gathered in a structured research environment, and in study four, identification of simple behaviours that also occur within a plant environment.

Table 1. Studies and associated correlations between ScanX measures and performance outcomes

	Study 1	**Study 2**	**Study 3**	**Study 4**
Accuracy 4 seconds		.559 (.01)	.595 (.041)	.489 (.003)
Accuracy 15 seconds		.213 (.21)	.696 (.01)	.542 (.000)
Efficiency 4 sec	.59 (.000)	.534 (.01)		
Efficiency 15sec	.47 (.000)	.447 (.04)		
Sample size	83	16	12	42

DISCUSSION

The sites in which the different research studies are situationally diverse, the visual environments being processed on screen are very different, the indicators or signals of an event that are used for detection performance are different (although studies two and three would have similarities in this respect), and the data collection conditions very strongly in terms of time demands and the underlying process. Yet, the correlation between the ScanX results and performance in each one of these cases is fairly substantial and consistently between about .44 to .69.

While we expect factors such as salience and small semantic distances to assist in detection, these factors are unrelated when comparing the test exercise to real world conditions. Further, issues such as attention sets may well vary in how well the person does on the tasks, but this no prior attention set that would equip participants to work with the ScanX content. A number of the theoretical issues relating to detection performance therefore do not come into play when looking at the relationship between ScanX and on the job or research study CCTV task performance.

The existence of strong relationships between ScanX scores and the different performance measures in the four studies appears to indicate that there is an underlying core of visual analysis skills or attributes that are contributing to interpretative analysis in different environments. The ScanX task requires rapid review of the image and search behaviour for relevant points. The speed of information processing is therefore a critical attribute. Quick recognition of essential details and the ability to home in on relevant points or characteristics is essential. The speed of information processing comes through when reviewing the impact on threat detection with x-rays. Typically, aviation screeners have approximately 5 seconds per bag in which to screen for threats. The Efficiency 4 second score correlates appreciably higher with treat detection than the Efficiency 15 second score, although both remain high. Colour can be used as a way of indicating inconsistency and pattern definition. Attention to detail is also important, but also the ability to view the broader view of the image and to examine context of individual elements to each other and peripheral information. Use of working memory provides useful reference information in detecting anomalies to the norm and changes in pattern structure and appears to work equally well for both the ScanX exercise and actual x-ray or CCTV scenarios. Sustained concentration is probably also indicated as important over time as a performance requirement for both of these. Poor vigilance would affect both ScanX exercise and work task related outcomes.

CONCLUSION

The nature of the relationship between ScanX performance and the research studies discussed here appear to emphasise the importance of process as opposed to content. Given the appropriate analytical thinking and the conceptual skills required in visual image processing, it can be applied to any different kind of scenario. This is not to say that there are important aspects from the research that do not apply in the actual performance of the job or work task. It is likely that performance can be improved by appropriate attentional sets, frames of reference, and motivational contexts. However, without the core skills required for analysis, even substantial input on these factors may not lead to the required performance level on the job. Personnel with an appropriate generic set of core visual analysis skills would be a major contribution to successful visual image processing with in a range of technologies. Such visual analysis specialists adept at multidisciplinary viewing allow for flexible staffing as different types of imaging technologies are used as part of comprehensive security strategies.

REFERENCES

Andrew, C., Landgrebe, T., & Donald, C. (2003). Enhanced surveillance platform validation testing: Methodology report. Technical report, de Beers, South Africa.

Binstock, J., & Minukas, M. (2010). Developing an operational and tactical methodology for incorporating existing technologies to produce the highest probability of an individual wearing an IED. Unpublished Master's thesis, Naval Postgraduate School, Monterey, California, US.

Cheyne, J. A., Carriere, J. S. A., & Smilek, D. (2006). Absent-mindedness: Lapses of conscious awareness and everyday cognitive failures. Consciousness and Cognition, 15(3), 578-592.

Davies, D.R., & Parasuraman, R. (1982). The psychology of vigilance. New York: Academic Press.

Department for Transport (2003). The selection of X-ray screeners. Transport Security Division, Technical Information Notice 1/03, United Kingdom.

Donald, C. (1998a). Technical report on the validity of the surveillance and monitoring assessment exercise (SAMAE). Technical Report No. 2/1998, Leaderware, SA.

Donald, C. (2001). Vigilance. In J. Noyes, & M. Bransby (Eds), People in control: Human factors in control room design. London: The Institution of Electrical Engineers.

Donald, F.M. (2011). Optimal Characteristics of Inserted Graphic Objects in Stimulating CCTV Operator Vigilance and Performance. Unpublished PhD thesis. University of the Witwatersrand. Johannesburg.

Huey, B. M., & Wickens, C. D. (1993). Workload transition: Implications for individual and team performance. Washington, DC: National Academy Press.

Koelega, H. S. (1996). Sustained attention. In O. Neumann, & A. F. Sanders, (Eds.). Handbook of perception and action, Vol. 3: Attention (pp. 277- 331). London: Academic Press.

Sandtom, C. (2001) Situational awareness.. In J. Noyes, & M. Bransby (Eds), People in control: Human factors in control room design. London: The Institution of Electrical Engineers.

Security General. (2002). Human factors in civil aviation security operations. Technical Report No. 9808 AN/765, International Civil Aviation Organization, Canada.

Wickens, C. D., & McCarley, J. D. (2008). Applied attention theory. London: CRC Press.