UNCLASSIFIED

Defense Technical Information Center Compilation Part Notice

ADP013324

TITLE: Performance Measurement for Visualisation

DISTRIBUTION: Approved for public release, distribution unlimited
Availability: Hard copy only.

This paper is part of the following report:

TITLE: Multimedia Visualization of Massive Military Datasets [Atelier OTAN sur la visualisation multimedia d'ensembles massifs de donnees militaires]

To order the complete compilation report, use: ADA408812

The component part is provided here to allow users access to individually authored sections of proceedings, annals, symposia, etc. However, the component should be considered within the context of the overall compilation report and not as a stand-alone technical report.

The following component part numbers comprise the compilation report:
ADP013309 thru ADP013341

UNCLASSIFIED

Click here to view PowerPoint presentation; Press Esc to exit

Performance Measurement for Visualisation

J.G. Hollands
Defence and Civil Institute Environmental Medicine
1133 Sheppard Ave. W., P.O. Box 2000
Toronto, Ontario, Canada M3M 3B9

Summary:
Key problems with extant visualisation systems for command and control are outlined. A systematic research program investigating five key issues relevant to C2 information visualisation is proposed. The outcome of this work should help to improve our understanding of those factors that make C2 visualisation more effective.

Military command and control (C2) is a complex process: many variables need to be monitored by many people; decisions must be made quickly; stress levels are high given time pressure and life or death consequences. The aim of command or battlefield visualisation software is to display pertinent information in comprehensible form to the commander or command team, so that they can make accurate and timely decisions, ultimately making our forces more effective than enemy forces.

However, despite the widespread development and implementation of command visualisation technology , it is unclear whether such technology actually improves the effectiveness of military forces, or even the command team itself. Command visualisation algorithms, engines, and techniques are being developed at a rapid rate, but the assessment of the approaches is sadly lacking. This is also the case for software more generally (Landauer, 1995, 1997). Although usability methods have increasingly been used to detect and fix more serious software problems (e.g., Nielsen, 1993), the study that compares performance with a new system to an old system (which may be an old computer system, or a pre-existing method not relying on computers) is rare. Does our new technological development really improve the situation or complicate it? The apparent benefit of the new system can be overshadowed by occasional problems or errors that overwhelm the benefits (Landauer, 1997).

It is in some ways not surprising that measurement methods have not been applied to C2 visualisation. Valid measurement involving human behavior in a real-world context is always problematic. In the similarly complex nuclear engineering domain for example, there is little agreement on how human performance should be measured (Voss, 1997). Voss notes that the IEEE Std 845 document (Evaluation of Man-Machine Performance, IEEE, 1988) neglects to specify those types of human performance that are important and necessary to measure in nuclear engineering. Similar problems in specifying appropriate performance measures are likely in C2 visualisation.

What is the system? In addition, it is important that when assessing human performance with a computer, both human and computer are considered as parts of the system. Traditional information-processing approaches have emphasized the human in isolation from computer, or have viewed the situation in static form, ignoring the impact of dynamic control on the human-computer system. In contrast, system designers tend to think of the system as the box on the desktop—forgetting for a moment that for the "system" to do anything useful a human must issue a command and inspect the result, and therefore a complete account of the system must include the human.

All these maxims are especially true in the visualisation domain, where the emphasis has traditionally been on the machine (particularly display software), not on

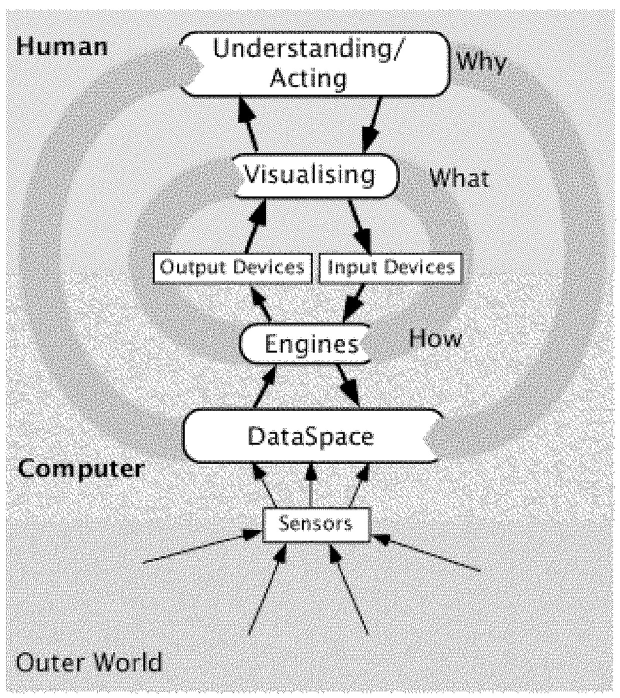

Figure 1. NATO IST-05 Reference Model for Visualisation.

Paper presented at the RTO IST Workshop on "Multimedia Visualization of Massive Military Datasets", held in Quebec, Canada, 6-9 June 2000, and published in RTO-MP-050.

the person. As noted earlier, algorithms and engines are being developed at a rapid pace, but evaluation is lacking. The entire system—including the human—must be considered. To reflect this, the control loop approach represented in Figure 1 is espoused (The IST-05 Reference Model; NATO IST-05, 1999). The Reference Model makes clear that "visualisation" does not refer to displays on a computer screen, but rather to a human activity augmented by such displays. Displaying complex data in a task-relevant way shifts the processing burden to the computer and away from the human, but ultimately, the visualisation must take place in the user's mind, or the display software has not been successful.

When one considers the military C2 context additional concerns become evident. Meister (1989) describes the concept of indeterminacy, or more formally, a determinacy-indeterminacy continuum. In a highly deterministic system inputs (to the user) are usually unambiguous and require little analysis. In contrast, indeterminate systems reflect considerable stimulus ambiguity and uncertainty. Military systems in wartime represent an indeterminate system (Meister, 1989). Any command visualisation situation will therefore reflect this ambiguity. Meister also notes that adversaries are a source of uncertainty because they strive to conceal their actions. This type of uncertainty is not present in supervisory control situations, in contrast.

Thus, it is clear there is a need for a systematic research program investigating factors affecting the effectiveness of C2 visualisation systems.

Proposed research program. The purpose of the proposed work is to develop a command visualisation testbed based on empirical principles, and to develop test protocols by conducting experiments based on relevant military tasks. This testbed and the planned experiments will provide a capability to investigate whether future proposed visualisation algorithms, constructs, and display concepts are consistent with human perception and cognition and whether they improve command decision making.

1. Frame of reference and visual momentum. Various visual momentum (Woods, 1984) techniques are available to allow commanders to transition between or "drill down" and then to "roll up" data with other data of similar types or at different levels (Roth et al. 1997), and their utility will be tested for individual and group displays. A related problem is disorientation or becoming lost when transitioning from one format to another. Use of landmarks in strategic locations and other techniques in development by Vinson (1999) will be tested experimentally.

2. Perceptual bias and reference points. Human judgments of the geometric volumes and areas that are commonly used to depict quantitative values in 3D data representations in statistical graphs and maps are biased (Hollands & Dyre, in press). The use of perspective rendering in 3D displays can also lead to bias. The selection of physical continua to code specific variables and the perceptual biases that result will be examined. In addition, the potential utility of reference points to reduce judgement error in command visualisation systems will be investigated.

3. Modeling mental operations. Follette and Hollands (2000) propose that two factors affect quantitative judgments with graphs: (1) the number of operations necessary; (2) the effectiveness of the perceptual features used as input for the operations. This model requires validation with more complex, dynamic displays as used in command systems. The model also needs to be cross-validated by measurement of eye movements using an eyetracker. A set of experiments is planned to test the predictions of the mental operations framework in command visualisation and cross-validate it using eye movement data.

4. Preattentive processing. When searching a field of symbols on a visual display certain symbols tend to "pop out" or be more salient. This research (e.g., Treisman & Gelade, 1980; Healey et al., 1995) has suggested that target detection and symbol grouping can be made more efficient and reduce attentional demand when pop out occurs. Experiments are planned that will investigate the relations among the perceptual dimensions used in C2 displays with respect to pop out. A better understanding of this relationship should enable the development of display mappings appropriate for different contexts. The eyetracker will provide an understanding of how displays are scanned and how much information in a display can be perceived "at a glance".

5. Mapping data to perceptual continua. Display designers often assign conceptual variables (e.g., vessel co-ordinates, number of torpedoes in task group, radar propagation characteristics) to perceptual dimensions (position, size, shape, colour saturation, colour hue, brightness, etc.) in an arbitrary way. However, perceptual dimensions have specific properties. Some perceptual dimensions are effective for depicting quantitative relationships, others only for order information, still others display only nominal (i.e., categorical) information well (Bertin, 1983; Cleveland, 1985; Wickens & Hollands, 2000). A systematic research program will investigate the effectiveness of various data-display mappings and determine the most effective mapping(s) from the conceptual to the perceptual for command visualisation.

Conclusions. Key problems with extant visualisation systems for command and control were outlined. A systematic research program investigating five key issues relevant to C2 information visualisation was proposed. The outcome of this work should help to improve our understanding of those factors that make C2 visualisation more effective.

References

Bertin, J. (1983). *Semiology of graphics.* Madison: University of Wisconsin Press.

Cleveland, W. S. (1985). *The elements of graphing data.* Pacific Grove, CA: Wadsworth.

Follette, L. J., & Hollands, J. G. (2000). *A theory of mental operations for graphical perception.* Manuscript in preparation.

Healey, C. G., Booth, K. S., & Enns, J. T. (1995). Visualizing real-time multivariate data using preattentive processing. *ACM Transactions on Modeling and Computer Simulation, 5,* 190-221.

Hollands, J. G., & Dyre, B. P. (2000). Bias in proportion judgments: The cyclical power model. *Psychological Review, 107,* 500-524.

Landauer, T. K. (1995). *The trouble with computers: Usefulness, usability, and productivity.* Cambridge, MA: Bradford/MIT Press.

Landauer, T. K. (1997). Behavioral research methods in human-computer interaction. In M. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction* (2nd Ed.) (pp. 203-227). Amsterdam: Elsevier.

IEEE (1988). *IEEE guide to the evaluation of man-machine performance in nuclear power generating station control rooms and other peripheries*. Institute of Electrical and Electronic Engineers Std. 845-1988.

Meister, D. (1989). Indeterminacy and system performance measurements. In *Proceedings of the Human Factors Society—33rd Annual Meeting* (pp.†1182-1186). Santa Monica, CA: Human Factors Society.

NATO IST-05 (1999). *Visualisation of massive military datasets: Human factors, applications, and technologies.* Interim Report (March, 1999), RTO IST-05/TG-002.

Nielsen, J. (1993). *Usability engineering.* Boston: Academic Press Professional.

Roth, S. F., Chuah, M. C., Kerpedjiev, S., Kolojejchick, J., & Lucas, P. (1997). Toward an information visualization workspace: Combining multiple means of expression. *Human-Computer Interaction, 12,* 131-186.

Treisman, A. M. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97-136.

Vinson, N. G. (1999). Design guidelines for landmarks to support navigation in virtual environments. In *CHI 99 Proceedings*. New York: Association for Computing Machinery.

Voss, T. J. (1997). Current human factors standards development efforts within IEEE. In *IEEE Sixth Annual Human Factors Meeting* (pp.†3-1-3-6). Orlando, FL: IEEE.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd Ed.). Upper Saddle River, NJ: Prentice-Hall.

Woods, D. D. (1984). Visual momentum: A concept to improve the cognitive coupling of person and computer. *International Journal of Man-Machine Studies, 21,* 229-244.

Footnotes

Copyright ©2000 Her Majesty the Queen in Right of Canada

Discussion – Paper 14

Command Visualisation

- Problems
  - Users cannot use engines to extract info from data
  - Need to arrange data in the right way for particular tasks
  - Measurement – what do we measure and how do we interpret these measures
  - Need for multiple views

Justin Hollands (DCIEM HCI Group):--

- Measurement important – cannot measure everything
- Example given from Challenger crash showing visualisation of O ring data
  - Temperature vs O ring damage (source E Tufte(1997. Visual Explanations) chart
- DARPA not doing sand table any more
- Multiple view and task dependency
  - Multiple formats
  - Must ease transitions between views
- Info vis is human's capacity to utilize effectively the output from the computer to understand the data.
  - Relies on human capacity
  - Why just computer – could be paper, sound
  - Artifact, process, or result?
- Command visualization testbed
  - To create effective command visualization platform based on empirical performance data captured
  - CTA used to determine type of info should be presented when
- What are we measuring?
  - Error magnitude, signal detection measures (sensitivity and bias)
  - Response time
  - Subjective measures (preferences, situation awareness, workload
- Visual momentum techniques (continuously available global views, gradual transitions, brushing)
- Perceptual measures