AN INNOVATIVE METHOD FOR PRESENTATION OF TARGET IMAGERY TO HUMAN OBSERVERS IN A SIMULATED OPERATIONAL ENVIRONMENT

November 1998

Rebekah A. Walrath, Hien T. Nguyen and Glenn A. Rogers
US Army Aberdeen Center for Sensing Technology
Aberdeen Proving Ground, MD 21005-5059

ABSTRACT

Presenting observers with single frames of electronically gathered images of a scene denies the natural variability of signal and noise as they change across time. We discuss a methodology that preserves the temporal fluctuations of signal and noise: thus faithfully representing the images produced by fielded hardware for use as laboratory perception study stimuli. Camouflaged military targets imaged by a 1st Generation Forward Looking Infrared (FLIR) system, were presented to observers in a simulated operational environment. Analog FLIR imagery from a Tube-launched Optically tracked Wire-guided (TOW) sight was digitized and looped to create a dynamic presentation. A test bed was designed to present the images and collect human performance data on a single desktop computer. The performance measures were time to detection/identification and indication of the Visible Center of Mass of the targets. These data were scored using the Hit and Kill criteria from the appropriate military field manuals.

1.0 DISCUSSION

Conducting tests in the field with human observers is expensive and it is increasingly difficult to obtain observers as the military reduces its forces. Laboratory studies that simulate field environments and operational conditions are becoming an economic necessity. However, if the results of laboratory studies are to be generalized to the field environment, maintaining fidelity with the real world is essential.

Perception experiments have traditionally used static imagery for presentation to human observers. It can be argued that this approach is reasonable with high-resolution images that hardware such as 2nd Generation thermal systems provide. This method is *not* acceptable for tactical systems which are much noisier. First generation thermal hardware and image intensifiers are intrinsically noisy systems.

Capturing a single image (from a 1st Generation system) and presenting it to an observer as representative of the real world is erroneous. Such a methodology destroys the fidelity of the

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-11-1998	Conference Proceedings	xx-xx-1998 to xx-xx-1998

4. TITLE AND SUBTITLE
An Innovative Method for Presentation of Target Imagery to Human Observers in a Simulated Operational Environment
Unclassified

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Walrath, Rebekah A. ;
Nguyen, Hien T. ;
Rogers, Glenn A. ;

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
US Army Aberdeen Center for Sensing Technology
Aberdeen Proving Ground, MD21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
Director, CECOM RDEC
Night Vision and Electronic Sensors Directorate
10221 Burbeck Road
Ft. Belvoir, VA22060-5806

10. SPONSOR/MONITOR'S ACRONYM(S)
11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
See Also ADM201041, 1998 IRIS Proceedings on CD-ROM.

14. ABSTRACT
Presenting observers with single frames of electronically gathered images of a scene denies the natural variability of signal and noise as they change across time. We discuss a methodology that preserves the temporal fluctuations of signal and noise: thus faithfully representing the images produced by fielded hardware for use as laboratory perception study stimuli. Camouflaged military targets imaged by a 1st Generation Forward Looking Infrared (FLIR) system, were presented to observers in a simulated operational environment. Analog FLIR imagery from a Tube-launched Optically tracked Wire-guided (TOW) sight was digitized and looped to create a dynamic presentation. A test bed was designed to present the images and collect human performance data on a single desktop computer. The performance measures were time to detection/identification and indication of the Visible Center of Mass of the targets. These data were scored using the Hit and Kill criteria from the appropriate military field manuals.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
			Public Release	9	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703767-9007 DSN 427-9007

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39.18

image produced by the original hardware. In a noisy system, the signal available to the operator contains both signal and noise and these fluctuate over time. It is crucial is that the signal—as it occurs naturally—is preserved and transmitted, unchanged, to the observer. When a single image is selected and presented as a static display, the levels of signal and noise, and their relationship to each other in that image, become fixed values. This artificial modification of that relationship destroys the fidelity of the image and it is no longer representative of reality. When using single images, the integration of visual information across time—which normally occurs within the human visual system—provides no new information, because the point values at any location in the image never change. Only when the imagery presented accurately represents a tactical system can a reliable assessment of the vulnerability of the items under test be obtained.

Our method creates a dynamic display of consecutive frames and preserves the fluctuations of signal and noise that existed in the original hardware. This presentation of the imagery allows the visual system to integrate information temporally, just as would an operator using the system in the field. Why is this so important?

The basic task of the human sensory systems is to detect the presence of energy changes in the environment.[1] Identifying a specific stimulus is an even more difficult task. Since the 1950's, psychologists have used information theory, a concept from communications engineering, as a quantitative method of describing the characteristics of input messages. "Information theory says in part, that the degree to which (transmitted information) the final decoded message reflects the original message depends, in part, on the ability of the system to transmit information without distortion (this is what is meant by the *fidelity* of the system) and on the complexity of the input."[2] Perception or comprehension of sensory input requires isolation of the signal from the noise.

There is a natural survival value in possessing senses designed to provide us with a stable percept of the world. If we brought into consciousness all the inherent noise presented to our sensory systems the world would be a considerably more complex and confusing place than it already is. In particular, our visual system has evolved to a level where much of the noise in the physical stimuli arriving at the retina is removed prior to being presented to our consciousness. The human visual system is constructed to integrate information. There are approximately 125 million photoreceptor cells, 120 million rods and 5 million cones, in the retina. However, the optic nerve; through which information from the photoreceptors leaves the eye, is composed of only 1 million fibers—a compression ratio of 125 to 1![3] Further, the photoreceptors are organized into receptive fields. This complex allocation provides both spatial and temporal

---

1. Coren, S., & Ward, L. M. (1989), Sensation and Perception, (3rd Ed.), New York: Harcourt Brace Jovanovich, p. 16.
2. Ibid., p.29.
3. Coren, S., Porac, C., & Ward, L. M. (1984), Sensation and Perception, (2nd Ed.), New York: Academic Press, Inc., p. 69.

summation. These varied functions of the eye and its structures result in an effective filter that eliminates much of the noise inherent in visual information.

All information within the human body is transmitted through a combination of electrical and chemical events. We all know that our eyes detect the presence of light. Through the process of transduction, the light energy of photons becomes an electrical signal—the product of a chemical reaction of the photopigments in the retina. These electrical signals are then transmitted by neurons. Neurons exist in two states—they are either 'on,' or 'off.'[4]

How does such a simple, two bit, system handle complex visual information? Through coding—both spatial and temporal. Spatial coding within the visual system is complex, elegant, and is preserved from the retina to the visual cortex in the brain. Machine vision systems make use of this same positive correlation between adjacent regions in time through a technique known as pixel averaging. However, the impact of temporal coding is what is applicable to the discussion here.

When a static image from a noisy thermal, or image intensified tactical system is presented to a human observer, sophisticated filtering within the visual system is by-passed. If we are presented with a dynamic view of the world, our visual system will, essentially, sum energy values across time and thus filter out random fluctuations. If however, we view a still image, this temporal variation does not exist. The noise and signal values have been 'frozen' at a particular instant in time. If at the instant the image was captured, noise components significantly mask signal components, the signal can be very difficult to detect. Obviously, this is not representative of what actually happens when the human operator uses these systems. Therefore, data collected using only static images is not valid. The effect of this artificiality becomes very apparent when examining some of the imagery we used. As the amount of noise in the signal increases, it is extremely difficult to see the target when looking at a single frame of imagery. As one steps through the available frames, there is an amazing amount of variability even between consecutive frames. For targets that are difficult to see, in many frames the target may not be visible at all. Yet when these same consecutive frames are looped, creating a dynamic display, the target is readily apparent. We argue that this methodology allows the human visual system to process imagery in the laboratory the same way it would on the battlefield.

---

4. Carlson, N. R. (1986), Physiology of Behavior, (3rd Ed.), Boston: Allyn & Bacon, Inc., p. 198.

The illustrations presented below are a series of frames from one of the movie stimuli. Consecutive frames are presented in Figures 1 through 5, respectively. It is apparent from these pictures that there is considerable variation in the shape of the Bradley Fighting Vehicle (BFVS) in the upper right corner. (The other vehicle in the scene is an M60 tank.)

Figure 1. First frame of movie, B2L334.

Figure 2. Second frame of movie, B2L334.

Figure 3. Third frame of movie, B2L334.

Figure 4. Fourth frame of movie, B2L334.

Figure 5. Fifth frame of movie, B2L334.

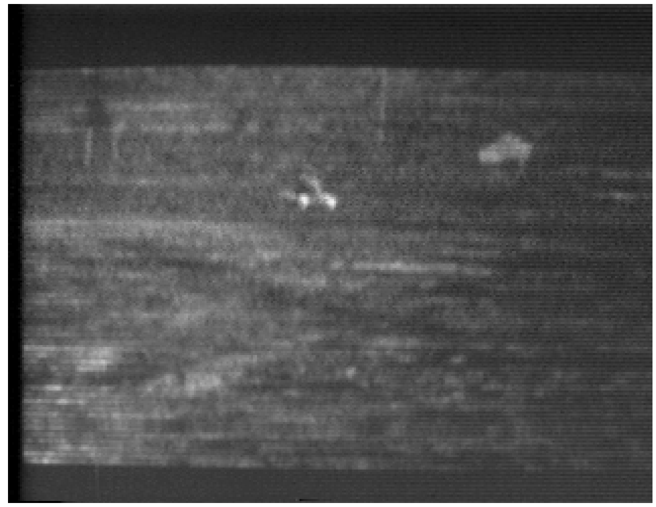

Figure 6. Summation of all frames.

Figure 6 is a representation of the effect of temporal summation within the visual system—information from all of the individual frames is combined. A comparison among these pictures individually, and each with the final version perceived by the human observer, graphically illustrates the importance of dynamic imagery in maintaining the fidelity between the battlefield and the laboratory.

2.0 METHODOLOGY

1.1 Stimulus Production Methodology

The following methodology was employed to obtain the desired dynamic stimuli. Imagery was collected using a standard TOW sight (ANTAS-4a) which had been internally modified using a splitter to provide dual output. One video output went to the eyepiece. The other went to a camera positioned to maintain an eye relief distance identical to the eyepiece. The camera output allowed direct video recording of the analog signal to super VHS tape. This dynamic imagery was then fed directly into a computer system and digitized. The digital movies were then reviewed and specific sections were selected. From these sections, the choice was further narrowed to consecutive frames. The test bed developed for presenting the imagery to the observers created a loop from the selected frames that was viewed by the observer for the desired interval—more about this later.

1.2 Test Bed and Data Acquisition System

Software development created a test bed that accommodates many different forms of imagery, both static and dynamic, and allows rapid prototyping and customization to specific test requirements. For example, simply typing in the appropriate number of seconds can modify the desired image presentation interval—changes to the programming code are not required. Software routines provide both image transformation and efficient image storage and retrieval. This test bed combines both presentation of the desired imagery and collection of the human performance data in a single desktop computer system.

Through a combination of hardware mockups and realistic establishment of situational factors, the experiment was constructed to model tactical sights and operational environments. The experimental apparatus for the TOW sight and that of 2nd Generation thermal mock-ups can be seen below in Figures 7 and 8, respectively. The system was used to collect the following measures of human performance: time to detection/identification, and indication of the Visible Center of Mass (VCM) of camouflaged military targets, imaged with both 1st and 2nd Generation Forward Looking Infrared (FLIR) systems. The VCM is the desired aim point used in gunnery. The data recorded were the x and y coordinates on the image where the observers indicated their aim point. These data were scored using the Hit and Kill criteria from the appropriate military field manuals.

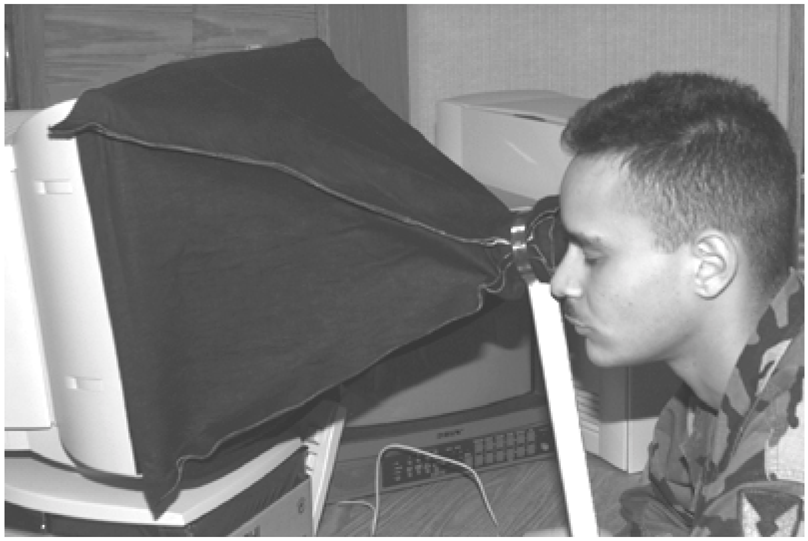

Figure 7. Laboratory apparatus used to simulate a TOW sight.

Figure 8. Laboratory apparatus used to simulate a 2^{nd} Generation display.

1.3 Database

A database was created using a standard office software suite on a PC platform. There were nearly 1000 1st Generation movies and 4000 2nd Generation images produced during this project. Efficient organization and management of this amount of information would have been impossible without a database. The data were stored on a JAZ disc. Again, a custom program allowed batch processing of the imagery into the database along with concurrent decoding of the image filenames. As an end product, the database had three important attributes. First, the movies and images were coupled with the relevant ground truth information such as the range to the target, aspect presented and the time that each image was captured. Second, the data from the observers was included on the same platform, so that for each movie or image the number of correct detections/identifications and the accuracy of the aim point were readily apparent. Third, because the movies and images in the database were the same physical size as the those presented to the observers, a one to one comparison was possible, e.g., the aim-points were obvious, the database proved an invaluable tool for the analysts. Figure 9 is an example of the database created for the 1st Generation movies and the associated human performance data.

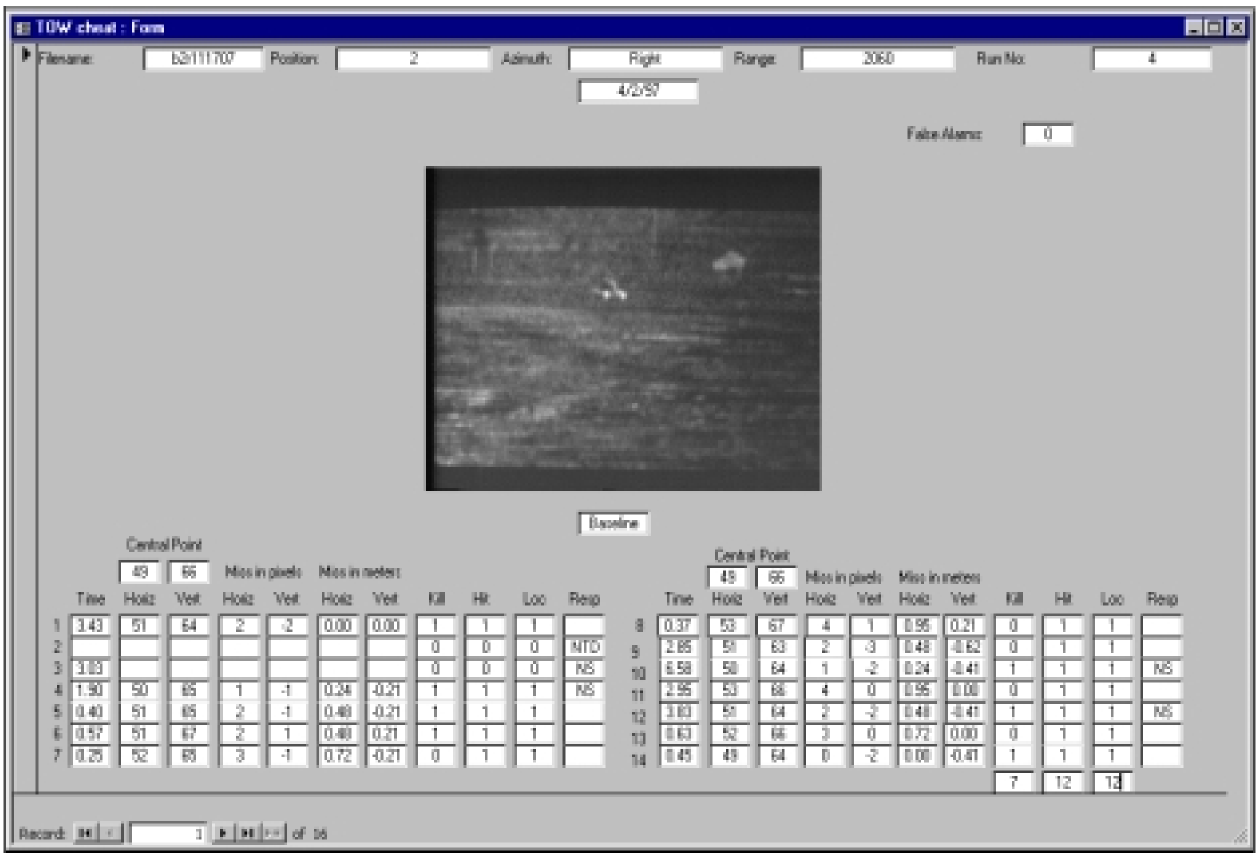

Figure 9. Example of the database which includes both 1st Generation movies, 2nd Generation still imagery, and associated observer data.

3.0 CONCLUSION

It is possible to perform visual perception studies in the laboratory using realistic stimuli that preserve the true nature of the imagery seen by the soldier on the battlefield. Combining commercially available software and program code developed at ATC, a test bed was developed to store and present this imagery and simultaneously record the observer's response in a single desktop system. The database, again a combination of software and in-house programming, archived the imagery and the associated human performance data. Using standard office software, the database operates an a PC platform.