Comparing Behavioral Receiver Operating Characteristic Curves to Multidimensional Matched Filters

[1#]Krebs, W.K., [2]Scribner, D.A., and [3]McCarley, J.S.

[1]Naval Postgraduate School, Department of Operations Research, 1411 Cunningham Road, Monterey, CA 93943

[2]Naval Research Laboratory, 4555 Overlook Avenue, Optical Sciences Division, Washington, D.C. 20375-5000. Phone number: (202) 767-2225, fax number: (202) 767-9203, e-mail: scribner@axp56r.nrl.navy.mil

[3]Beckman Institute, University of Illinois, 405 North Mathews, Urbana, Illinois 61801. Phone number: (217) 244-3096, fax number: (217) 244-8371, e-mail: mccarley@beckman.uiuc.edu

[#]Current address is Federal Aviation Administration, 800 Independence Avenue, S.W., AAR-100 (Room 907), Washington D.C. 20591. Phone number: (202) 267-8758, fax number: (202) 267-5968, e-mail: william.krebs@faa.gov

Abstract

Human factors experiments can be used to test whether a sensor can improve operator performance for detecting or recognizing a target[1]. Although human factors experiments are of tremendous value, these tests are time consuming and resource intensive. In order to reduce costs associated with collecting behavioral data, a two-dimensional matched filter[2] is proposed. The objective of this paper is to compare and contrast behavioral and matched filter[2] ROC plots to determine whether the matched filter technique is a good predictor of human performance. Five different background images (three infrared band images, chromatic-fused image, and monochromatic-fused image) were used with, and without, a target (airplane) present. False alarm and target detection probabilities were computed and results were plotted on a Receiver Operating Characteristic (ROC) curve. The matched filter ROC curves were then compared to behavioral ROC curves. Results showed that the matched filter ROC curves were similar to behavioral ROC curves with color fusion and long-wave infrared showing the highest sensitivity and mid-wave and short-wave infrared scenes were significantly less sensitive (near chance). These results indicate that the matched filter analysis may be used to model human behavior.

Keywords: Signal Detection Theory, Matched Filter Analysis, Receiver Operating Characteristic, Human Performance Modeling, and Target Detection

<u>Introduction</u>

Military applications require the use of various sensors to determine operational threats and opportunities. The combination of such sensors promise to provide an account of the opposition that is superior to those of individual sensors that operate at particular wavebands. It is desirable to choose the optimal types and combinations of sensor information that are maximally responsive to target types likely to be encountered in the field. This assessment must be done under realistic physical and psychophysical circumstances. Furthermore, it is desirable that the information obtained be modeled productively, i.e. so that experimental results can be interpolated and extrapolated near the conditions under which they are obtained.

This study will modify an existing matched filter model[2] to fit meaningful human performance metrics that can be revised and extended where necessary to represent the data obtained during field tests. This model will be used to evaluate fused imagery systems requirements and performance. Furthermore, this model will indicate what type of data will be needed to validate the type of sensor fusion data to be collected in the future.

In order to assess an operator's ability to detect a target while viewing sensor imagery, different disciplines have developed methodologies to measure operator performance. The Night Vision and Electronic Sensors Directorate (NVESD) developed analytical models to predict target detection ranges for sensors that operate in the visible and infrared bands[3,4.] These electro-optical models provide an adequate prediction of a user's ability to detect a target at any given range. In order to improve the validity of the models, atmospheric conditions, sensor characteristics, target characteristics, clutter, estimated time that an operator searches for the target, and an assortment of other parameters are used to model human performance. Currently, these models are limited to single-band sensors; however, the next generation models may incorporate multi-spectral sensor performance.

Recent technological advances in the design and manufacturing of multi-spectral sensors now allows spatially registered imagery to be mapped to a high speed processor where it can be fused and displayed to an end user[5]. Within the last several years, numerous groups have developed sensor fusion algorithms[2,6-10] that may improve operator performance. These techniques may differ on the algorithm approach, but they all have the same objective: improving the image quality for the observer. Several behavioral studies[1, 11-16] and image quality studies[2,10] have tried to quantify the benefits of sensor fusion, but the results were inconsistent. This is not surprising considering that in many cases different spectral bands were used and a number of other parameters varied as well, such as camera sensitivity, and target and background characteristics.

Tanner and Swets[17] proposed that statistical decision theory might be used to predict operators' decision behavior. Signal Detection Theory is a common technique used by vision scientists to measure subjects' sensitivity and response bias to a set of stimuli[18]. Whether target detection is accomplished through the human visual system or by means of a matched filter, the theory of signal detection requires recognizing a signal plus noise from a steady state noise background. Vision scientists use signal detection theory to measure operator performance to an assortment of stimuli. Similarly, an image-processing algorithm may use a matched filter technique that is based on signal detection theory to predict operators' performance through a sensor. Ideally, the Receiver

Operating Characteristic (ROC) plots derived from both methodologies should yield similar results. The advantage of the matched filter technique allows the system engineer to conduct multiple simulations for a wide variety of backgrounds and target types. These simulations require minimal resources compared to costly human performance field tests.

A matched filter is a two-dimensional (2-D) array, which has been optimized to maximize the signal-to-noise ratio and provide a measure of the spatial correlation between the input image and the reference image[19]. The resulting filters are “tuned” to negate the effects of the background clutter and other noise sources in the image. A matched filter is the optimum linear filter for the detection of the target. Scribner et al.[2] used a matched filter on long-wave infrared (9.0 to 11.6 μm), medium-wave infrared (4.5 to 5.5 μm), and short-wave infrared (2.0 to 2.6 μm) sensors, as well as on a fused single image of these bands. In this approach, spatial-only and spatial-spectral matched filters were derived for the three infrared images and the fused composite image respectively. The intent of these matched filters was to simulate the detection ability and sensitivity of the human visual system. Although the matched filter is commonly used within the physics and engineering communities to quantify sensor performance, it has been used to some extent by the medical field to detect tumors in an x-ray image[20].

The objective of this paper is to compare and contrast behavioral and matched filter ROC plots to determine whether the matched filter technique is a good predictor of human performance. The advantages of the matched filter model are threefold. First, it provides a sensor image fusion metric that can be used to evaluate different sensors. Second, it quantifies the degree of “enhancement” achieved by a fusion process, thus allowing for direct comparisons of the various sensor fusion algorithms. Third, it may have the ability to predict human visual performance across a variety of background and target conditions.

A standard visual search paradigm will be conducted for three different multi-spectral natural scenes. In experiment 1, behavioral ROC plots will be compared to the matched filter ROC plots to determine whether the matched filter technique accurately predicts observers' sensitivity. In experiment 2, eye movement data will be recorded to determine whether observers' scan pattern correlates with the ROC analysis. It is hypothesized that observers viewing a low contrast stimulus will exhibit longer saccade lengths and shorter fixations as well as show a low sensitivity for detecting the target (i.e., less correct responses and more errors).

Experiment 1:

Behavioral Test

Subjects: Fourteen male military officers (mean age 31.7 years old) participated in this visual search study. All subjects had normal (20/20), or corrected to normal, acuity and color vision. Subjects were naive to the purpose of the experiment and none had participated in previous visual search experiments. All subjects signed an informed consent and were briefed on the ethical conduct for subject participation in the Protection of Human Subjects[21].

Apparatus: Stimuli were presented by a VisionWorks computer graphics system[22] on an IDEK MF-8521 high-resolution color monitor (21" X 20" of viewable area, .28mm dot pitch) equipped with a non-glare, anti-reflect, P-22 phosphor. The monitor's resolution was 800 by 600 pixels (x=75.02 and y=74.92 pixels/degree), 98.9 Hz frame-rate, mean chromatically of *Y*= 50.2, *x* = 0.334, *y* = 0.336 (1931 CIE), and a maximum luminance of 100 cd/m^2. Luminance of the monitor was linearized by means of an 8-bit look-up table for each of the red, green, and blue guns. Subjects viewed the monitor from 1.5 meters and were positioned by an adjustable chinrest. Subjects viewed the stimuli under mesopic conditions.

Stimuli: Three single-band stimuli (short-, mid-, and long-wave infrared) and two composite stimuli (fused-color and fused-gray) were selected from a multi-spectral natural scene database. The selection criteria consisted of scenes that contained heterogeneous terrain characteristics and no man-made targets (figure 1).

Insert Figure 1

about here

Each background scene was 320 by 400 pixels (subtended 8.54^0 by 7.24^0 visual angle) with 50 percent of the stimuli containing a randomly placed airplane (subtended 0.1^0 by 0.1^0 visual angle). For each scene, the airplane target spectral characteristics were based on a measured target within the multi-spectral database. The long-wave target pixel spectral values were 255, mid-wave target pixel values were 73, and short-wave target pixel values were 114. The fused color scene spectral values were red=255, green=73, and blue=114. The achromatic fused images were spatially identical to the chromatic fused images; however the achromatic condition was employed to control for luminance effects. For each background scene, the target was present in 50 trials. The target location was generated by a random number generator and then inserted at that particular location. The target placement for each of the 50 locations was identical across the different background types.

The composite stimuli approach is to assign each pixel a color vector defined by the detected power in the registered three- band imagery[2]. Scatterplots (figure 2) of the image ensemble of colors frequently reveal pronounced anti-correlation between short and long wavelengths, consistent with Kirchoffs law (reflective objects which appear bright in the short-wave infrared typically have low emissivity and appear dark in the long-wave infrared).

Insert Figure 2

about here

For a given registered short- and long-wave infrared image pair, the principal component corresponds to the major axis−luminance channel. The orthogonal axis corresponds to the minor axis−color channel. The assignment of luminous intensities to the correlated component is straightforward, but the assignment of color to the uncorrelated features is not immediately obvious. The assignment of a pixel color is based on color opponent process. By a-priori assigning one color to the image intensified

(i^2) and it's color opponent to the infrared *(ir)*, the resulting display shows two and only two opponent colors of various saturation. This makes an immediately intuitive representation as to which spectral bands dominant and by how much. It must be strongly emphasized that this system is mathematically incomplete to allow the perception of actual visible colors in the estimated reflectivity sense. Distinction between various vegetation, soil types, structures, water, and sky is based on coincident phenomenology in each spectral region, not by estimating a physical property such as emissivity.

Procedure: Each subject participated in only one display format. Subjects were instructed to manually respond on a keyboard whether a target was present or absent within the scene. The four response categories were “1” = definitely no target to “4” = definitely a target. At the beginning of each trial, the subject fixated on a cross hair located in the center of the screen. The fixation cross was presented for 200 millisecond, immediately followed by a 1000 millisecond presentation of the experimental stimulus. The stimulus extinguished after the initial presentation or after the subject made a response, whichever came first. The next trial began approximately 1 second after the subject’s preceding response. Accuracy was measured for each trial and no feedback was given for incorrect responses.

Matched Filter Analysis

The matched filter technique is described briefly here and explained in detail along with MATLAB code elsewhere[24]. In general the matched filter analysis paralleled behavioral testing described above. That is the same images and targets were used to generate numerical results. One additional step in the matched filter processing was to blur the image and target very slightly to take into account the modulation transfer function of the display and the human visual system. This was done using a narrow gaussian point spread function with a radius of one pixel. The actual computations were done using MATLAB™ software, which manipulates the image data in a matrix format. Single-band filters were derived using the smoothed 2-D power spectrum of the background and the target template with the target intensity identical to that used in the behavioral testing. The 3-D spatio-spectral (color) filter was derived by considering the target and the background as a 3-D space, with the third dimension being the spectral values of each pixel. In either case a multidimensional matched filter can be derived in the frequency domain using the expression,

$$H\left(\vec{k}\right) = c_0 \frac{S^*\left(\vec{k}\right)}{W\left(\vec{k}\right)} \exp(-i\vec{k}\vec{r}_0)$$

where $\boldsymbol{S}$ is the multidimensional signal representation, $\boldsymbol{W}$ is the multidimensional power spectral density. The spatial frequencies $\boldsymbol{k_x}$ and $\boldsymbol{k_y}$ are image coordinates indices in the frequency domain, and $\boldsymbol{k_f}$ is the spectral band index. The real space filter can be found by computing the inverse 3-D Fourier transform of $\boldsymbol{H}$,

$$h(x, y, l) = F_{3-D}^{-1}\{ H\left(k_x, k_y, k_f\right)\}$$

Processing the image with each respective filter is then done by convolving the filter with several hundred locations in the image. This is accomplished by multiplying filter values times corresponding pixel values aligned at each location. The summed values are stored for each position, giving an indication of false alarms (clutter leakage noise). These values are then compared to a second set of calculated totals produced by the same procedure, but with the target inserted at corresponding locations giving an indication of target detection. By comparing the signal-plus-noise values to the noise values for a given threshold value, false alarm and target detection probabilities can be calculated and displayed in the form of an empirical ROC plot.

Results

Signal detection theory distinguishes operator performance into two categories - sensitivity and response criterion or β[17]. Sensitivity is defined as the difference between the means of the signal plus noise and noise distributions. An observer's response criterion is independent of sensitivity. To calculate an observer's response criterion, β is equal to the ordinate of the signal plus noise distribution at criterion divided by the ordinate of noise distribution at criterion.

Both sensitivity and response criterion is derived from the probability of hits and probability of false alarms for each experimental condition. A ROC plot is a useful illustration of the relationship between sensitivity and response bias. The ROC curve plots on a single graph the joint value of probability of hits and probability of false alarms for each tested condition[25].

For this analysis, behavioral and matched filter ROC plots were compared across the five formats. Figure 3 illustrates the behavioral ROC plots across format types (blue = short-wave infrared; green = mid-wave infrared; red = long-wave infrared; fused gray = monochrome fused; and fused color = color fusion). The fused color and long-wave infrared formats had the highest sensitivity, while the short-wave infrared and fused gray sensor formats were near chance. The mid-wave sensor sensitivity was between the short- and long-wave sensor formats.

Insert Figures 3a, 3b, 3c, 3d, 3e

about here

Figure 4 illustrates the matched filter ROC plot across format types. Again, the sensitivities show similar trends across format types. Moreover, the sensitivities between the matched filter and behavioral ROC plots are very similar. Therefore, the matched filter may be a viable alternative to human performance testing to assess operator detection performance.

Insert Figures 4a, 4b, 4c, 4d, 4e

about here

Experiment 2:

Cognitive scientists record eye movements to understand cognitive processes that occur when an observer is searching for a target[26]. Eye movement data illustrates where and when the eye fixates within the scene; however, the data does not indicate what was processed. Rayner[27] found that our eyes move three to four times per second while searching a scene. These saccadic eye movements enable the observer to extract important high spatial detail from each foveal fixation. Although there has been numerous eye movement studies investigating visual cognition, it is unclear what mechanisms control where and what the eye will fixate on next.

Biederman, Mezzanotte, and Rabinowitz[28] found subjects extract information outside the fovea during scene perception. The parafovea and peripheral vision may extract certain features within a fixation; however, this information may or may not be identified. In order to integrate these parafovea cues into an identifiable object, a fixation is required[29]. To facilitate object identification, the more informative the scene the more likely the observer will fixate on those recognizable regions[29].

On the initial fixation, the observer will obtain a global snapshot of the scene. Next, low-level visual cues such as color, brightness, and contours will guide the observer's eye movements. Therefore, the level of informativeness within the picture will influence subjects' scene comprehension and object identification. It is hypothesized that a color-fused scene contains more informative features about the signal-to-noise ratio then an achromatic scene. Subjects initial eye movements will be guided to a color-fused target due to the target attributes. The achromatic target attributes will not contain enough informative information to capture the observers' visual attention. Furthermore, the eye movement results will correlate with the ROC plots. Targets embedded within the short- and mid-wave infrared scenes will require more saccades to identify the object, while the long-wave and fused conditions will be identified within the first couple fixations.

Methods

Subjects: Ten male military officers participated in this eye movement study. All subjects had normal (20/20), or corrected to normal, acuity and color vision. Subjects were naive to the purpose of the experiment and none had participated in previous visual search experiments. All subjects signed an informed consent and were briefed on the ethical conduct for subject participation in the Protection of Human Subjects[21].

Apparatus: Eye movements were recorded using an ISCAN, Inc. remote eye imaging system[30]. The eye tracker is a video-based system that uses an infrared camera to illuminate the eye and another camera to record the pupil to corneal reflection. The eye tracker then calculates the difference between the pupil and corneal reflection to indicate where the observer is fixating on the stimulus screen. The system operates at a sample rate of 60 Hz and the subject's visual point-of-regard may be determined with an accuracy of better than one degree over a +/- 25 degree horizontal to a +/- 20 degree vertical range.

Stimuli: Same as experiment 1.

Procedure: Each subject participated in only one display format. At the beginning of the experimental session, the subject's head was placed in a chinrest positioned 1 meter from the stimulus monitor. The subject's right eye was then calibrated using a five-point calibration grid displayed on the stimulus monitor. To maintain an accurate calibration between the eye tracker and the stimulus monitor, periodic five-point calibration checks were conducted throughout the experimental session.

At the beginning of each trial, the subject fixated in the center of the screen. Once the subject's eye was in the desired location, the experimenter initiated the trial. A fixation cross was presented for 200 milliseconds, immediately followed by a 5-second presentation of the experimental stimulus. Subjects were instructed to search for the target. Once the target was identified, the subject was to maintain fixation on the target until the stimulus extinguishes. There were 32 target trials and 16 noise trials presented for each format. Subjects' point-of-regard was recorded at 60Hz. No feedback on target identification accuracy was given.

Results

In each trial, the eye movement recording apparatus recorded the observer's fixation point at the rate of 60Hz, and hence a total of 300 data points were obtained per 5-second trial. An analysis software tool was subsequently used to analyze the data with the criterion for minimum fixation time at 40msec and the maximum horizontal and vertical deviation of the eyes at ±5 and ±3 pixels respectively. Thus, the number of fixations, the duration of each fixation, and the distance between fixations could be determined. These data were then tabulated to calculate the mean and the standard error mean values of the fixation duration, number of fixations, and scan path length.

Subjects within the short- and mid-wave infrared and gray fused conditions showed more fixations and longer scan-path lengths compared to the long-wave infrared and color-fused conditions. The long-wave and color-fused targets contained enough informative attributes to guide the subjects' eye movements to the desired location. Figure 5 illustrates a subject's search for a short-wave infrared and another subject's search for a color fused target.

Insert Figures 5a and 5b

about here

The subject immediately identified the color-fused target, while the other subject was not able to find the short-wave infrared target. The subject within the short-wave infrared scene obtained enough global information within the initial fixation to search higher probability areas as to where the target may be located. However, the target's poor contrast inhibited the subject from identifying the location. Alternatively, the fused-color condition provided enough informative information within the first fixation to guide the subject to the target's location. The target's good spatial characteristics and large color contrast easily guided the subject to the appropriate critical region.

These results parallel the ROC results. The short- and mid-wave infrared and gray fused conditions low sensitivities match the eye scan data. Subjects within these conditions were not able to find the target within the first couple fixations which would indicate that their sensitivity should be low. Subjects within the color-fused and long-wave infrared conditions easily identified the target, which would indicate that their sensitivity should be high.

Conclusion

The purpose of this experiment was to compare matched filter analysis with human behavioral signal detection. The matched filter results illustrate that the different sensor format sensitivities are similar to the behavioral sensitivities. Although the ROC sensitivities between each format are not quantitatively identical, the matched filter technique gives excellent qualitative agreement with the human performance tests. Additional refinement of the matched filter should result in even better agreement. Ogawa[24] found that the matched filter ROC was consistently superior to the behavioral ROC. His matched filter did not account for the human visual system inequalities. The gaussian blur was added to the filter to more accurately represent the human visual system resolution limit. The addition of gaussian blur to the filter caused our results to behave more similar to behavioral ROC as compared to Ogawa's results[24]. Additional refinement of the exact amount of gaussian blur to the matched filter should improve the correlation between the two ROC plots. The eye movement results illustrate that the eye was not able to identify the short- and mid-wave infrared and gray fused conditions as well as the color and long-wave infrared conditions. The color and long-wave infrared targets possessed important visual attributes that enabled the subject to identify the target with little to no effort. A surprising finding was the poor performance of the gray fused condition for both the signal detection and eye scan experiments. Subjects guided search for the target was not solely dependent upon spatial content; rather, visual search was mediated by both spatial and color target attributes. This finding indicates that color fusion is more appropriate for targeting applications than monochrome fusion. The color-fused target "pops-out" at the subject, which allows increased signal-to-noise sensitivity.

In summary, the matched filter technique may be a useful technique to predict human visual sensitivity for different sensor types by target characteristics. The matched filter technique will assist system engineers with a rough approximation of a human sensitivity to a target. This information could then be used for rapid prototyping of a system, enhance the predictability of existing electro-optical models, and provide a metric to test multi-spectral sensors. Additional tests will need to be conducted to test the robustness of the matched filter across different signal-to-noise ratios, terrain and target types, and various other atmospheric and illumination conditions. Finally, this matched filter will assist human factors testing by reducing the number of parameters needed to achieve the desired goal. Human factors testing will always be required, but at least the matched filter technique may provide the human factors group a better understanding of how the human will respond in the field.

Acknowledgements

Sponsored by DARPA's Integrated Imaging Sensors Program, Mr. Ray Balcerak is the program manager. A special thanks to James Lowell for his assistance in the eye movement data collection and analysis.

The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, nor the United States Government.

References

1. W.K Krebs, D.A. Scribner, G.M. Miller, J.S. Ogawa, and J. Schuler, "Beyond third generation: A sensor fusion targeting FLIR pod for the F/A-18," In B.V. Dasarathy (Ed.), *Proceedings of the SPIE-Sensor Fusion: Architectures, Algorithms, and Applications II*, 3376, 129–140 (1998).
2. D.A. Scribner, M.S. Longmire, and M.R. Kruer, "Analytic Modeling of Staring Infrared Systems with Multidimensional Matched Filters," *Proceedings of the SPIE*, 890, pp.81-91 (1988).
3. J.A. Ratches, W.R. Lawson, L.P. Obert, R.J. Bergemann, T.W. Cassidy, and J.M. Swenson, "NVL Static Performance Model for Thermal Viewing Systems," US Electronics Command Report ECOM#7403 AD-A011212 Fort Monmouth, N.J. (1973).
4. J.D. Howe, "Electro-Optical Imaging System Performance Prediction," In *Electro-Optical Systems Design, Analysis, and Testing,* (Dudzik Ed.), *The IR & EO Systems Handbook, Vol 4*, Ch. 2, ERIM IIAC, Ann Arbor, MI and SPIE, Bellingham, WA, (1993).
5. R. McDaniel, D.A. Scribner, W.K. Krebs, P. Warren, N. Ockman, and J.S. McCarley, "Image fusion for tactical applications," *Proceedings of the SPIE - Infrared Technology and Applications XXIV*, 3436, 685–695 (1998).
6. A. Toet, L.J. van Ruyven, and J.M. Valeton, "Merging thermal and visual images by a contrast pyramid," *Opt Eng.*, 28, 789–792 (1989).
7. D. Ryan, and R. Tinkler, "Night pilotage assessment of image fusion," In R.J. Lewandowski, W. Stephens & L.A. Haworth (Ed.), SPIE Proceedings on Helmet and

Head Mounted Displays and Symbology Design Requirements II (pp. 50-67). Bellingham, WA, USA: International Society for Optical Engineering (1997).

8. A. Toet, and J. Walraven, "New false color mapping for image fusion," *Opt Eng.*, 35, 650–658 (1996).

9. C.W. Therrien, J. Scrofani, and W.K. Krebs, "An adaptive technique for the enhanced fusion of low-light visible with uncooled thermal infrared imagery," *Proceedings of the IEEE: International Conference on Imaging Processing*, 405-408 (1997).

10. A.M. Waxman, A.N. Gove, D.A. Fay, J.P. Racamato, J.E. Carrick, M.C. Seibert, and E.D. Savoye, "Color night vision: Opponent processing in the fusion of visible and IR imagery," *Neural Networks*, 10, 1–6 (1997).

11. E.A. Essock, M.J. Sinai, J.S. McCarley, W.K. Krebs, and J.K. DeFord, "Perceptual Ability with Real-World Nighttime Scenes: Image-Intensified, Infrared and Fused-Color Imagery," *Human Factors*, 41, 438-452 (1999).

12. M.T. Sampson, W.K. Krebs, D.A. Scribner, and E.A. Essock, "Visual search in natural (visible, infrared, and fused visible and infrared) stimuli," *Investigative Ophthalmology and Visual Science*, (SUPPL) 36, 1362, FT Lauderdale, FL (1996).

13. P.M. Steele, and P. Perconti, "Part task investigation of multispectral fusion using gray scale and synthetic color night vision sensor imagery for helicopter pilotage," *Proceedings of the SPIE Conference on Aerospace/ Defense Sensing, Simulation, and Controls,* 3062, 88–100 (1997).

14. J.S. McCarley and W.K. Krebs, "Visibility of road hazards in thermal, visible, and sensor-fused nighttime imagery," *Applied Ergonomics*, 31, 523-530 (2000).

15. W.K. Krebs, J.S. McCarley, and M.S. Sinai, "Psychophysical assessments of image-sensor fused imagery," (to be submitted).

16. M.J. Sinai, J.S. McCarley, W.K. Krebs, and E.A. Essock, "Psychophysical comparisons of single- and dual-band fused imagery", *Proceedings of the SPIE-Synthetic Advanced Vision*, 3691, 1–8 (1999).

17. W.P. Tanner, and J.A. Swets, "A decision-making theory of visual detection," *Psychological Review*, 401–409 (1954).

18. D.M. Green, and J.A. Swets, *Signal detection theory and psychophysics*, Wiley, New York, USA (1966).

19. W.K. Pratt, *Digital Image Processing*, John Wiley & Sons, Inc, New York (1991).

20. M.P. Eckstein, and J.S. Whiting, " Visual signal detection in structured backgrounds. I. Effect of number of possible spatial locations and signal contrast," *Journal of the Optical Society of America (A)*, 13(9), 1777-87 (1996).

21. Department of the Navy, *Protection of human subjects*, (SECNAV Instruction 3900.39B). Washington, D.C.: Chief of Naval Operations OP-098 (1984).

22. D.J. Swift, S. Panish, and B. Hippensteel, "The use of VisionWorks in visual psychophysics research," *Spatial Vision*, 10, 471–477 (1997).

23. D.A. Scribner, P. Warren, J. Schuler, M. Satyshur, and M. Kruer, "Infrared color vision: an approach to sensor fusion," *Optics and Photonics News*, 8, 27–32 (1998).

24. J.S. Ogawa, *Evaluating color fused image performance estimators*, Master of Science in Operations Research, Naval Postgraduate School (1997).

25. G.A. Gescheider, *Psychophysics method, theory, and publication, 2nd edition.* Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey (1985).

26. K. Rayner, “Eye movements and visual cognition: Introduction”, In K. Rayner (Ed.), *Eye Movements and Visual Cognition Scene Perception and Reading*, Springer-Verlag, New York (1992).

27. K. Rayner, “Eye movements in reading and information processing,” *Psychological Bulletin*, 85, 618–660 (1978).

28. I. Biederman, R.J. Mezzanotte, and J.C. Rabinowitz, “Scene perception: Detecting and judging objects undergoing violation,” *Cognitive Psychology*, 14, 143–177 (1982).

29. K. Rayner, and A. Pollatsek, “Eye movements and scene perception,” *Canadian Journal of Psychology*, 46(3), 342–376 (1992).

30. R. Razdan, and A. Kielar, "Eye tracking for man/machine interfaces,” *Sensors*, September (1988).

Figure Legends

Figure 1. Single-band (upper left) short-wave infrared, (upper right) mid-wave infrared, (lower left) long-wave infrared, and (lower right) fused color was created by taking principle component direction of correlated thermal and visible pixel values as the luminance direction in a transformed space[23]. The airplane target is located in upper right quadrant.

Figure 2. Color fusion algorithm technique. Two highly correlated bands will have a cigar-shaped distribution. The principal component direction (L_1') is the luminance channel and the orthogonal axis (L_2') is the chromatic channel. Increasing color contrast (dotted line) while retaining the luminance characteristics is achieved by re-scaling L_1'. In an actual sensor system, the principal component direction is based on the statistics of the scene (determined adaptively).

Figure 3. Human performance receiver operating characteristic (ROC) plots for fourteen subjects. Each format condition (a) short-wave infrared, (b) mid-wave infrared, (c) long-wave infrared, (d) monochrome fusion, and (e) color fusion had three subjects except for the gray fused condition had two subjects. Subjects within the fused color and long-wave infrared conditions had the highest sensitivity for detecting the target, while short-wave infrared and fused color near chance (d'=0).

Figure 4. Illustrates the matched filter ROC analysis compared to the mean behavioral ROC for five different format types. Each format condition (a) short-wave infrared, (b)

mid-wave infrared, (c) long-wave infrared, (d) monochrome fusion, and (e) color fusion illustrates the matched filter analysis (°) and the mean ROC for each format type (•). Although the ROC sensitivities between each format are not quantitatively identical, the matched filter technique gives excellent qualitative agreement with the human performance tests.

Figure 5a. Subject CH was not able to identify the short-wave infrared target. The subject's initial fixation provides enough global information about where to search within the scene, but the low target contrast does not have enough information to attract the visual system. This result correlates with the low sensitivity for the visual search task. Subjects' sensitivity was near chance (d'=0).

Figure 5b. Subject JL found the target after the first fixation. The color-fused target contained enough visual information to automatically guide the subject to the target location. Thus, the color-fused target had a high level of informativeness, which enabled the subject to identify the target with little effort. Again, this result correlates with the high sensitivity measure within the visual search experiment.

Figure 1.

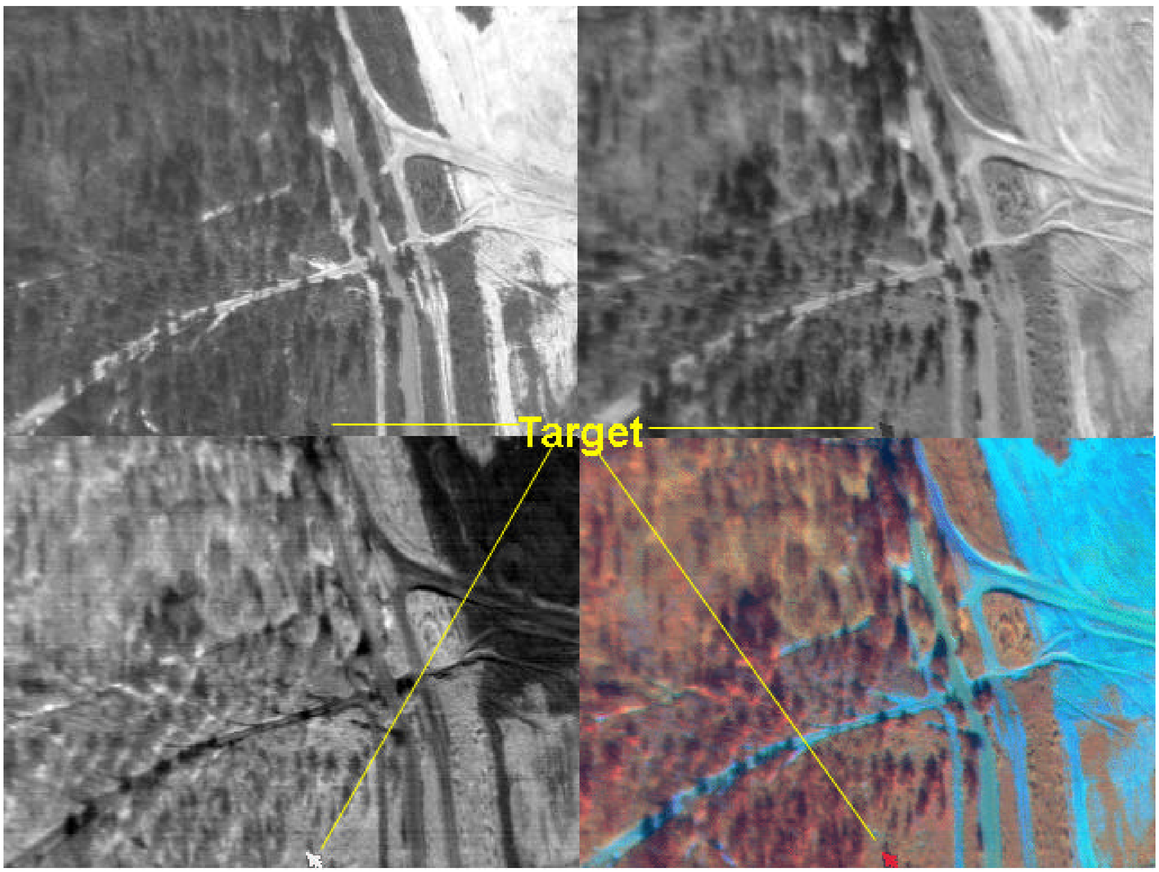
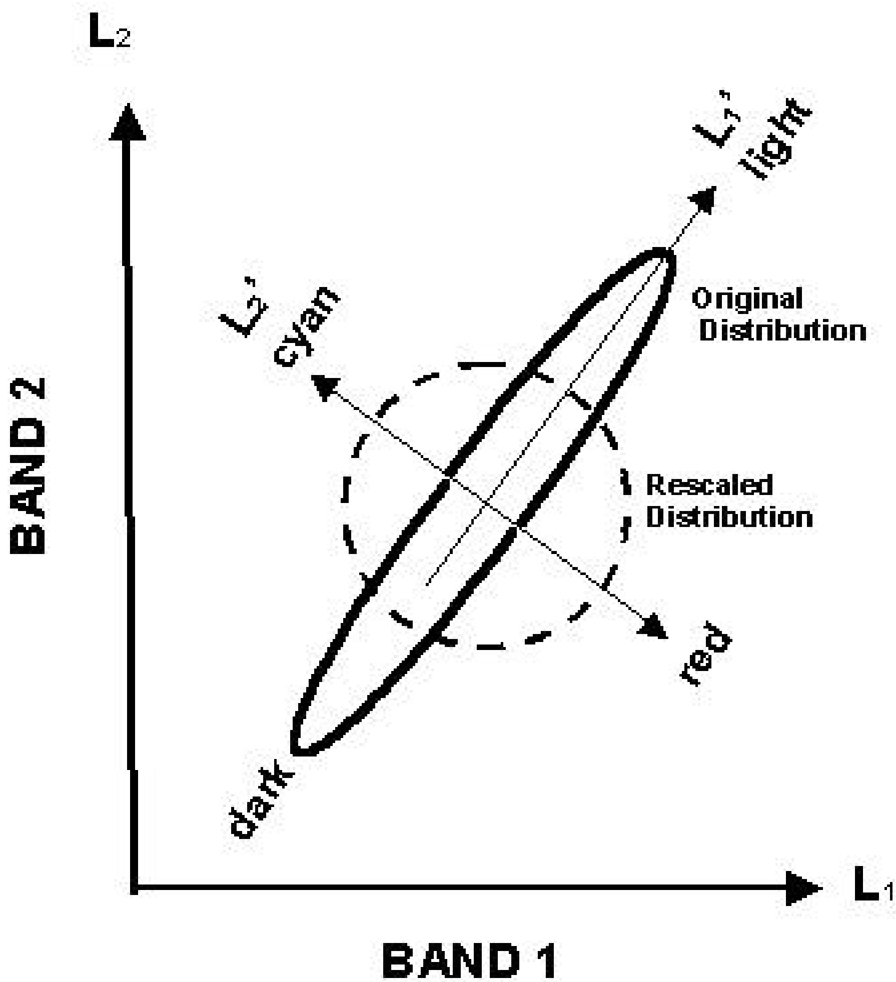


Figure 2



L2
L1'
light
L2'
cyan
Original
Distribution
Rescaled
Distribution
red
dark
BAND 2
L1
BAND 1

Figure 3

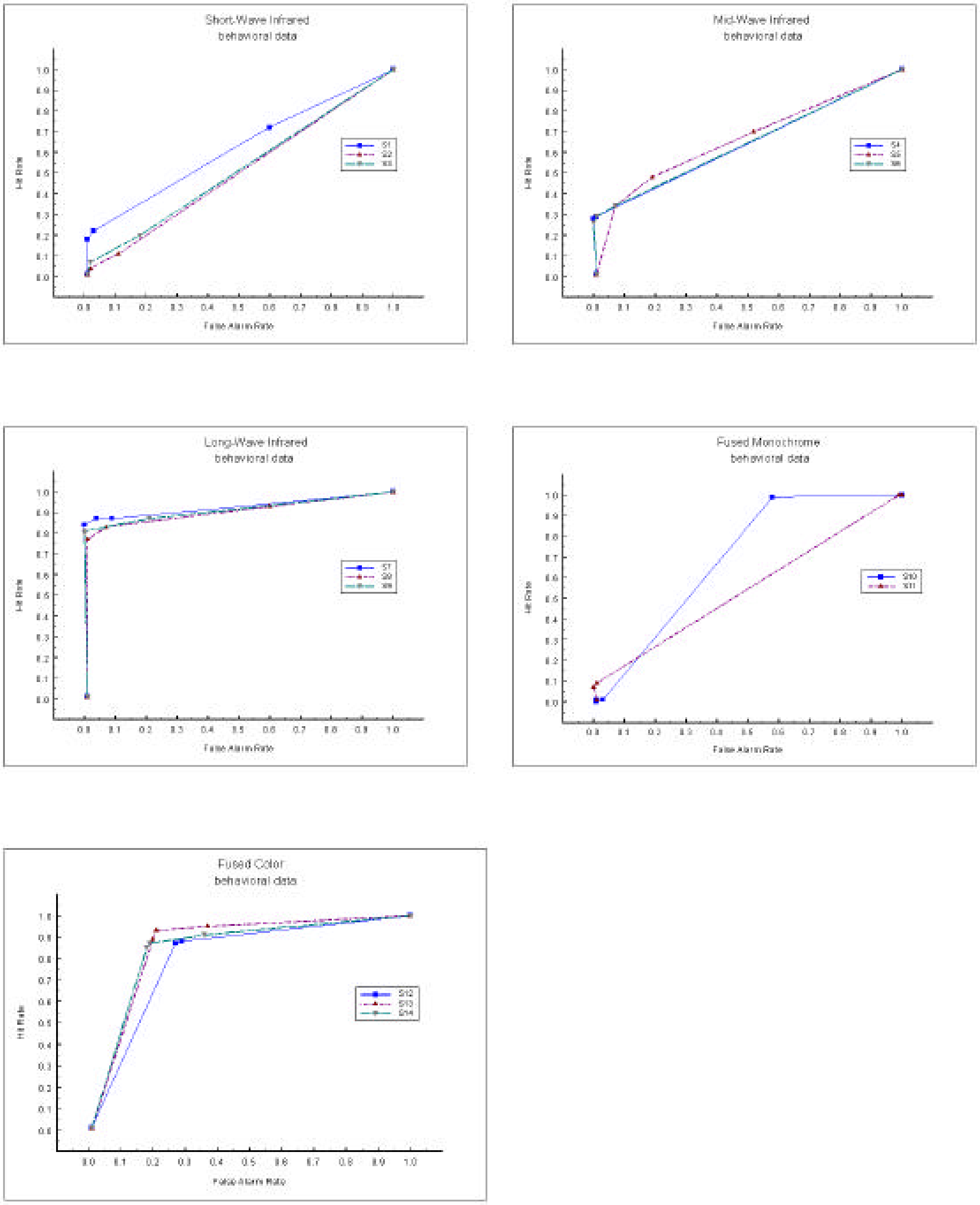

Short-Wave Infrared
behavioral data
S1
S2
S3
Mid-Wave Infrared
behavioral data
S4
S5
S6
Long-Wave Infrared
behavioral data
S7
S8
S9
Fused Monochrome
behavioral data
S10
S11
Fused Color
behavioral data
S12
S13
S14
Hit Rate
False Alarm Rate

Figure 4

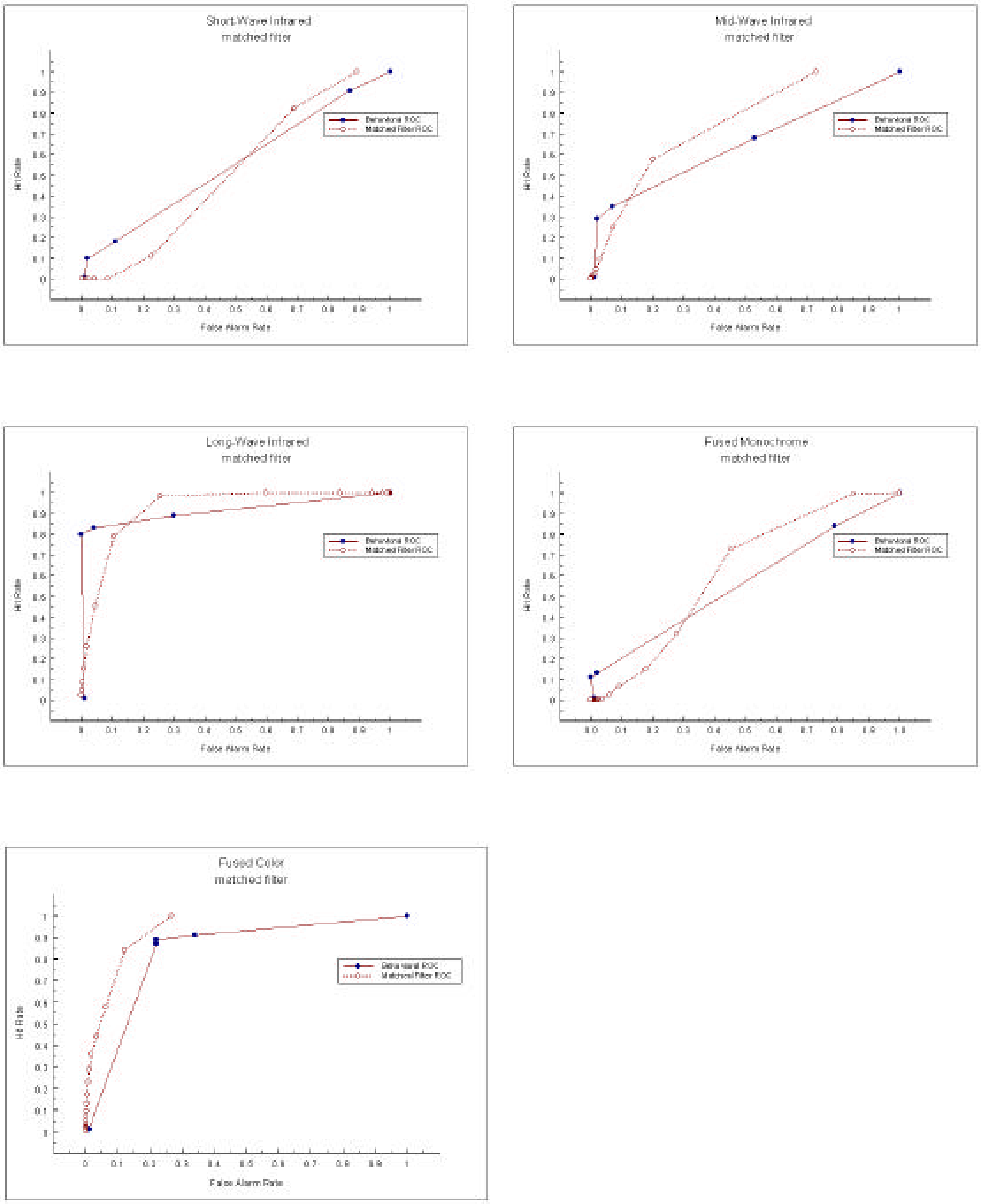

Short-Wave Infrared
matched filter
Hit Rate
False Alarm Rate
Mid-Wave Infrared
matched filter
Hit Rate
False Alarm Rate
Long-Wave Infrared
matched filter
Hit Rate
False Alarm Rate
Fused Monochrome
matched filter
Hit Rate
False Alarm Rate
Fused Color
matched filter
Behavioral ROC
Matched Filter ROC
Hit Rate
False Alarm Rate

Figure 5