Benefits of photosimulation and sensor fusion for threat detection

E. Bankowski, Ph.D., D. Bednarz, Ph.D., D. Bryk, R. Jozwiak,
K. Lane, T. Meitzler, Ph.D., and E.J. Sohn

U.S. Army TACOM
Survivability Technology Area
Warren, MI 48937

ABSTRACT

Visible, infrared (IR) and sensor-fused imagery of scenes that contain occluded camouflaged threats were compared for hit and miss differences in the probability of detection of objects. Response times were also measured. Image fusion was achieved using a Gaussian Laplacian pyramidal approach with wavelets for edge enhancement. Three types of images were also ranked in terms of better probability of detection of concealed weapons. Detecting potential threats that are camouflaged or difficult to see is important not only for military acquisition problems but, also for crowd surveillance as well as tactical use such as on border patrols. Imaging and display technologies that take advantage of sensor fusion are discussed in this paper.

I. BENEFITS OF SIMULATION IN VISUAL PERCEPTION LAB.

A method is described for using the TACOM photosimulation laboratory environment to detect threats and to evaluate the effectiveness of camouflage for military vehicles. There are distinct advantages to acquiring images at the field site and then bringing them back for observer testing in a laboratory environment. Laboratory testing provides a repeatable, secure, and low-cost way to generate realistic performance data for threat detection and vehicle evaluation for the purposes of signature testing, measurement of the effectiveness of camouflage relative to a baseline vehicle, and calibration and validation of target acquisition models.

Three tests are described by the authors. In the first test a baseline LAV is compared to a treated LAV in the TARDEC Photosimulation Laboratory using imagery collected from the field in the manner prescribed by an experimental design. Two tests were performed for RPG detection for homeland defense. Twenty five civilian subjects took the Rocket Propelled Grenade (RPG) detection test. The first RPG detection test consisted of 116. There were images of soldiers holding the RPG. There were also images of soldiers without RPG. The test was done on a flat panel monitor. There were three types of images: visual, IR and fused. The probability of detection (PD) of the RPG was estimated as a ratio of correct answers to the total number of pictures. The response time was also measured. The objective was to see the relationship between the probability of detection and the entropy metric. The second RPG detection test was ranking three types of images: visual, IR and fused. The test subjects were shown 32 sets of three images: visual, IR and fused (the total number of images was 96). The RPG was present in every picture. They had to rank images which image was better in terms of detecting the RPG. Using high- resolution graphics projectors, the imagery can be presented in the controlled environment of the lab in such a manner as to obtain observer data with confidence levels approaching 99%. It is the authors opinion that, the benefits of the high degree of confidence achieved using the repeatability offered by the lab environment, far outweighs any so-called 'realistic' advantage of having multiple teams of observers present at a field location for threat detection and vehicle testing.

II. PREVIOUS LAV-25 TEST.

1. Introduction. The Light Armored Vehicle (LAV) Family of Vehicles (FOV) was developed to provide the Marine Corps with enhanced mobile warfare capabilities. The LAV FOV includes several variants which utilize light armor protection from small arms, light machine gun fire, artillery projectile fragments, and mine fragments. Each variant was designed for a specific mission function and was mounted on a common chassis.

2. Experiment. The test design implemented was an extension of visual detection requirements provided to our lab from PM LAV. Initially the PM requested only two range points at standard engagement ranges. We suggested having more ranges in between the critical range points to obtain a probability of detection versus range curve, as is more typical for these kinds of tests. A test matrix was developed in full-factorial form and 24-bit color imagery was collected using a Kodak 460 digital camera. The images were prepared for the photosimulation test and then presented to 30 subjects. The experimental factors and levels with their values are shown below in Table 1. The photosimulation test in the lab was arranged so that a pixel IFOV subtended by the display was less than one minute of arc and the displayed image represented a unity magnification or 1X representation to the subject. The first test was meant to emulate naked-eye vision. Prior to the actual test, the subjects were instructed on the purpose of the test as well as required to take a pre-test in which they could become familiar with the imagery and software. None of the pictures used in the pre-test were used in the actual test, however, the images were from the same set. The test procedure was to display an image with a time-out of 30 seconds. The imagery is cropped so that no scrolling is required. The target can appear within one of five possible regions. The soldier must use the mouse to "click-on" what her or she thinks is a target, based on the training. The tests are done in a dark room in which the subjects are 'dark-adapted' to maximize contrast differences in the images.

3. Results. Analysis of the first test results has shown that most subjects obtained a score of only 20 % detection. This is not unreasonable given the difficulty of the imagery. The ranges are not unusual for such a test, however the high degree of clutter and in particular the height of the grass on the terrain makes it difficult for the unaided eye to detect common cure features of the vehicle. A second test was arranged at a power of 3X. The imagery from the field was of sufficient resolution so that there was no noticeable increase in pixilation of the imagery an in increase in magnification. The presentation in the lab was randomized, this is a very good reason to use the lab.

Region	
1	Top-Left
2	Top-Right
3	Lower-Left
4	Lower-Right
5	Center
Vehicle Type	
1	Baseline (old LAV)
2	SLEP + ADCAM
3	SLEP + ADCAM - ADCAM bowplane
Aspect angle	
1	Front
2	30 degree
3	Side
Lighting	
1	Front Lit
2	Back Lit
Weather condition	
1	Clear
2	Overcast
Range (km)	
1	1
2	1.5
3	2
4	2.5
5	3

Table 1: Factor matrices for the visual detection test

When making inferences about differences in a factor in a perception experiment in the laboratory we want to make the experimental error as small possible. This requires that we remove the variability between subjects from the experimental error. The design we use to accomplish this is a factorial experiment run in a randomized complete block. By using this design with the subjects as blocks we form a more homogeneous experimental unit on which to compare different factors. This experimental design improves the accuracy of the comparisons among the different factors by eliminating the variability among the subjects. Within a block, the order in which the treatment combinations are run is randomly determined. It is usually not possible to implement this experimental design in the context of a traditional field test.

The pictures below in Fig. 1 through Fig. 6 were used for training observers as to what kind of vehicles they would be looking for in the test.

Fig. 1: Baseline side

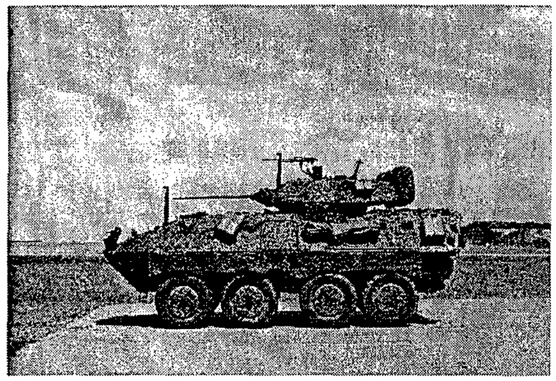

Fig. 2: ADCAM side

The charts in Fig. 3 below show the results of measuring the X and Y chromaticity values of the monitors that were used in the test. The values measured were compared to standard values and found to be virtually identical.

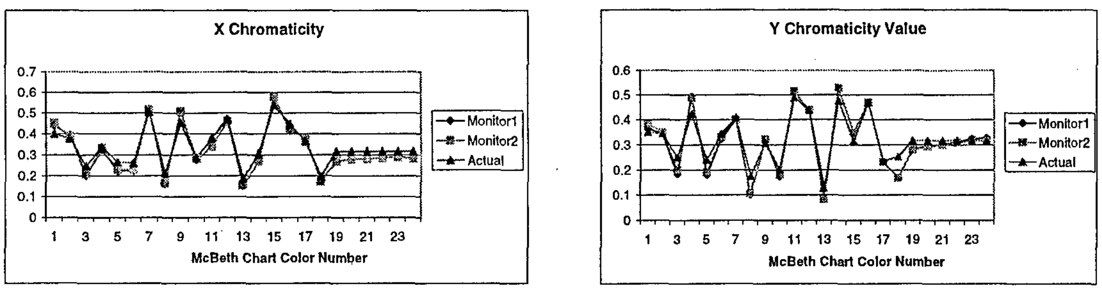

Figure 3: CIE X and Y monitor chromaticity calibration charts

4. Analysis Below in Table 2 is the ANOVA table for the baseline vehicle and the other experimental factors. The treated vehicle has been excluded because of security classification. The power of the experimental design methodology is shown here in that the significance of individual factors and of their interactions are available. Using this kind of a test, one can obtain not only a model curve of the detection probability versus any factor in the test, but, one can also obtain the relative importance of the individual factors.

Table 2 shows that the aspect angle was the least important factor in the experiment. Figure 4 shows the model generated logistic curve of the probability of detection versus the distance from LAV, or range. The model shows a good fit.

Tests of Between-Subjects Effects

Dependent Variable: RANK of RESPONSE

Source	Type IV Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
Corrected Model	185215009[b]	89	2081067.514	26.782	.000	2383.607	1.000
Intercept	1159759015	1	1159759015	14925.408	.000	14925.408	1.000
SKY_COND	2918024.068	2	1459012.034	18.777	.000	37.553	1.000
RANGE	161301308	9	17922367.58	230.650	.000	2075.852	1.000
ASPECT	944347.896	2	472173.948	6.077	.002	12.153	.887
SKY_COND * RANGE	5751990.053	18	319555.003	4.112	.000	74.025	1.000
SKY_COND * ASPECT	2459473.720	4	614868.430	7.913	.000	31.652	.998
RANGE * ASPECT	6204854.010	18	344714.112	4.436	.000	79.853	1.000
SKY_COND * RANGE * ASPECT	5397720.861	36	149936.691	1.930	.001	69.465	1.000
Error	128521871	1654	77703.671				
Total	1641367780	1744					
Corrected Total	313736880	1743					

a. Computed using alpha = .05

b. R Squared = .590 (Adjusted R Squared = .568)

Table 2: ANOVA of test factors

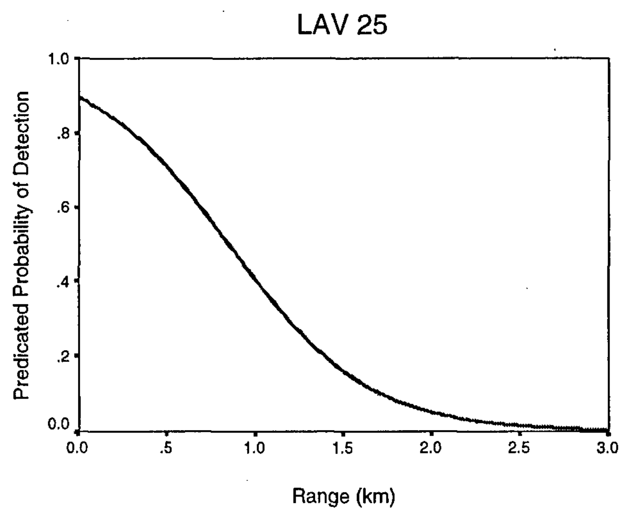

Figure 4. Logistic curve fit to the model from the subject responses

Figure 5 shows the plot of the entropy versus the probability of detection.

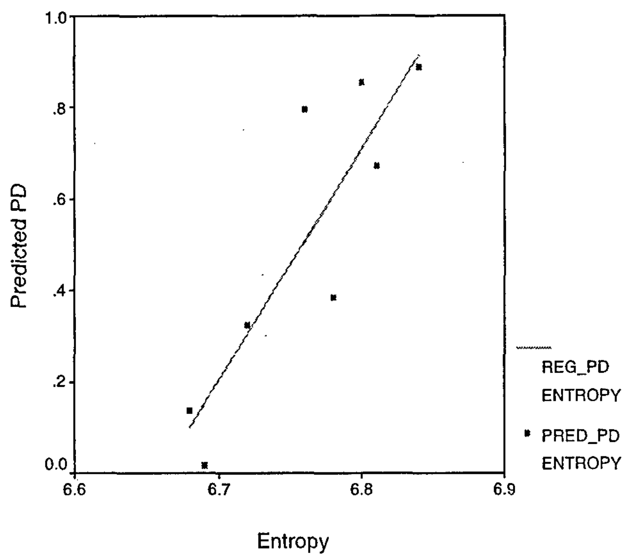

Figure 5. The plot of entropy versus the probability of detection.

There was a correlation between the entropy and the probability of detection, as it was observed in the LAV test.

III. THE RPG DETECTION TEST FOR HOMELAND DEFENCE.

1. Experimental Test Procedure. Twenty five civilian subjects were picked up randomly from the employee population of TARDEC. Each subject took the Rocket Propelled Grenade (RPG) detection test, and ranking the quality of images test.

Test No. 1 was the RPG detection test. It consisted of 116 images in a random order. There were images of soldiers holding the RPG. There were also images of soldiers without RPG. The test was done on a flat panel monitor. There were three types of images: visual, IR and fused. Some images had brightness, contrast and noise adjusted. Each image was shown for one second. After viewing each image for one second, the test subject was asked to click "Yes" or "No", depending whether he could see the RPG in the picture, then go to the next image. The probability of detection (PD) of the RPG was estimated as a ratio of correct answers to the total number of pictures. The response time was also measured. The objective was to see the relationship between the probability of detection and the entropy metric. Figure 6 shows three identical images of a man with the RPG: the image on the left side is the visual image, the image in the middle is the infra-red (IR) image, and the image on the right is the fused image.

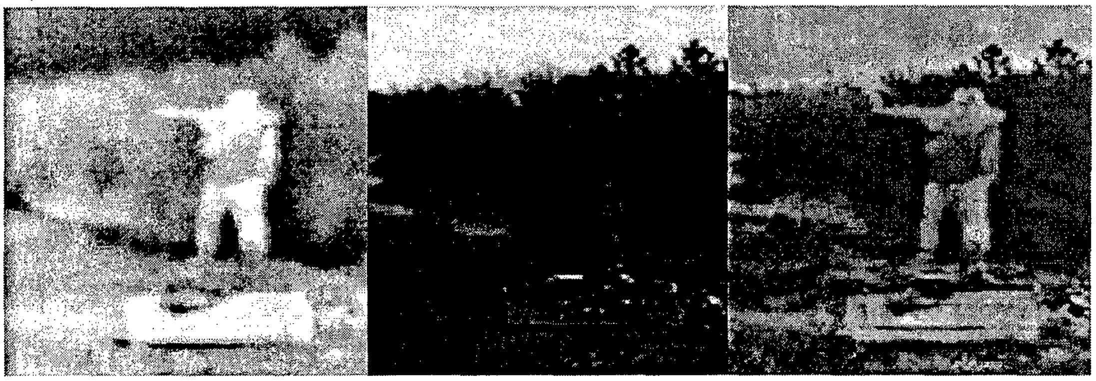

Figure 6. Images of a man with the RPG: visual image, IR image, and fused image.

Test No. 2 was ranking three types of images: visual, IR and fused. The test subjects were shown 32 sets of three images: visual, IR and fused (the total number of images was 96). The RPG was present in every picture. They were asked the question: "Which image is better in terms of detecting the RPG?" They were asked to rank images in the set as # 1, 2 and 3. The noise, brightness and contrast levels in some images were adjusted to make it harder to detect a threat. Figure 7 shows three identical images of a man with the RPG; the noise level in these images was adjusted. The image on the left side is the visual image, the image in the middle is the infra-red (IR) image, and the image on the right is the fused image.

Figure 7. Images of a man with the RPG with the noise level adjusted: visual image, IR image, and fused image.

Figure 8 shows three identical images of a man with the RPG; the brightness level in these images was adjusted. The response time was also measured.

Figure 8. Images of a man with the RPG with the brightness level adjusted: visual image, IR image, and fused image.

Figure 9 shows three identical images of a man with the RPG; the contrast level in these images was adjusted. The objective was to see the relationship between the probability of detection and the entropy metric.

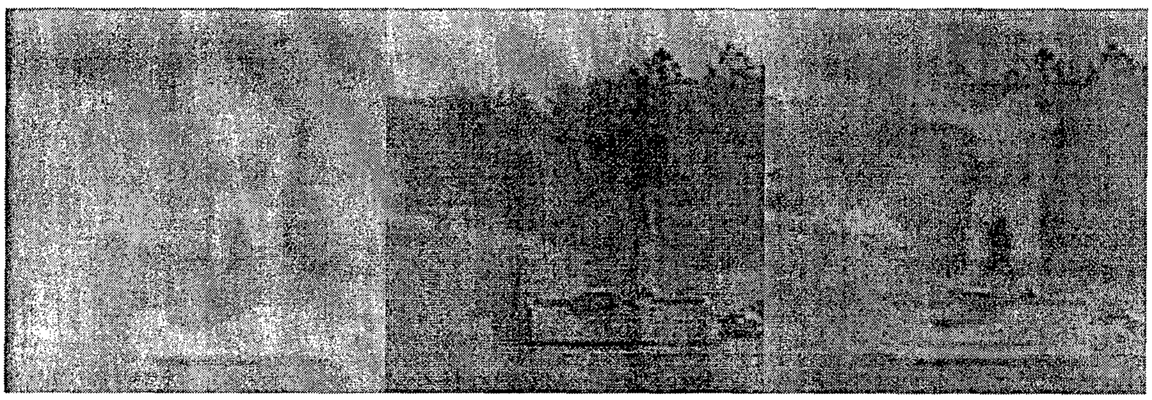

Figure 9. Images of a man with the RPG with the contrast level adjusted: visual image, IR image, and fused image.

2. Statistical Analysis of Test Data.

The statistical results of the RPG test were analyzed. Figure 10 shows the detection rate of the RPG in the image versus the sensor type. As we can see from Figure 11, the highest detection rate was achieved using the IR sensor, and the lowest detection rate was with the visual image. The test subjects also commented on the fact that they preferred the fused image, because it had more details.

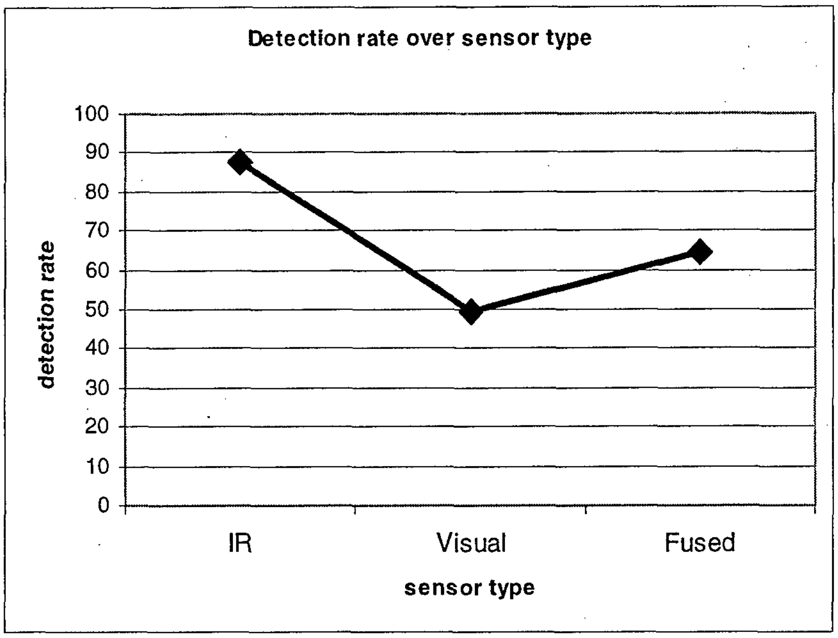

Figure 10. The detection rate of the RPG in the image versus the sensor type.

The results of test No.2 are shown in Figure 11. Test No. 2 was ranking three types of images: visual, IR and fused. The test subjects were sets of three images: visual, IR and fused. They were asked the question: "Which image is better in terms of detecting the RPG?" They were asked to rank images in the set as # 1, 2 and 3. The noise, brightness and contrast levels in some images were adjusted to make it harder to detect a threat.

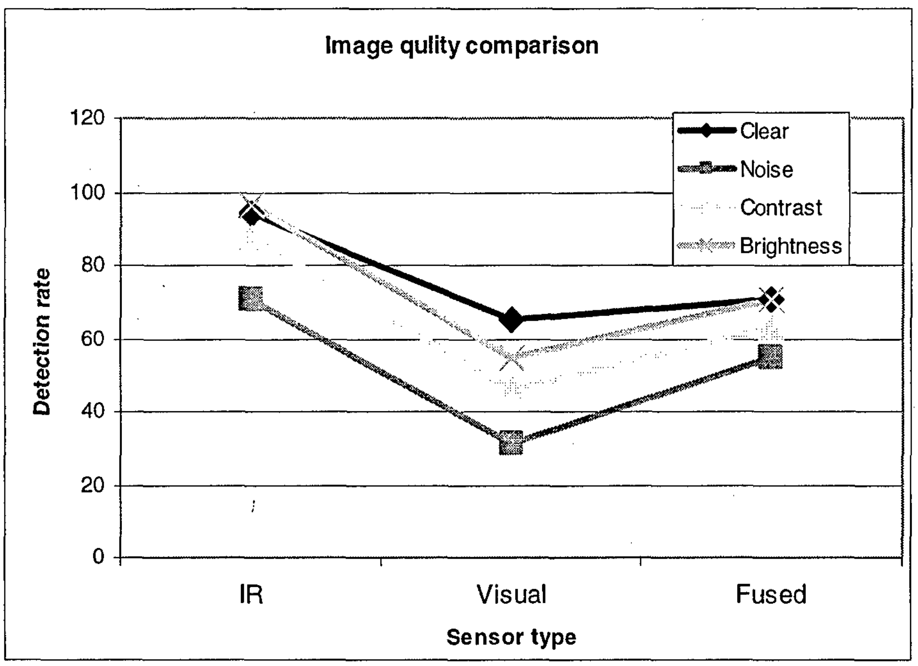

Figure 11. Image quality comparison: detection rate versus the sensor type.

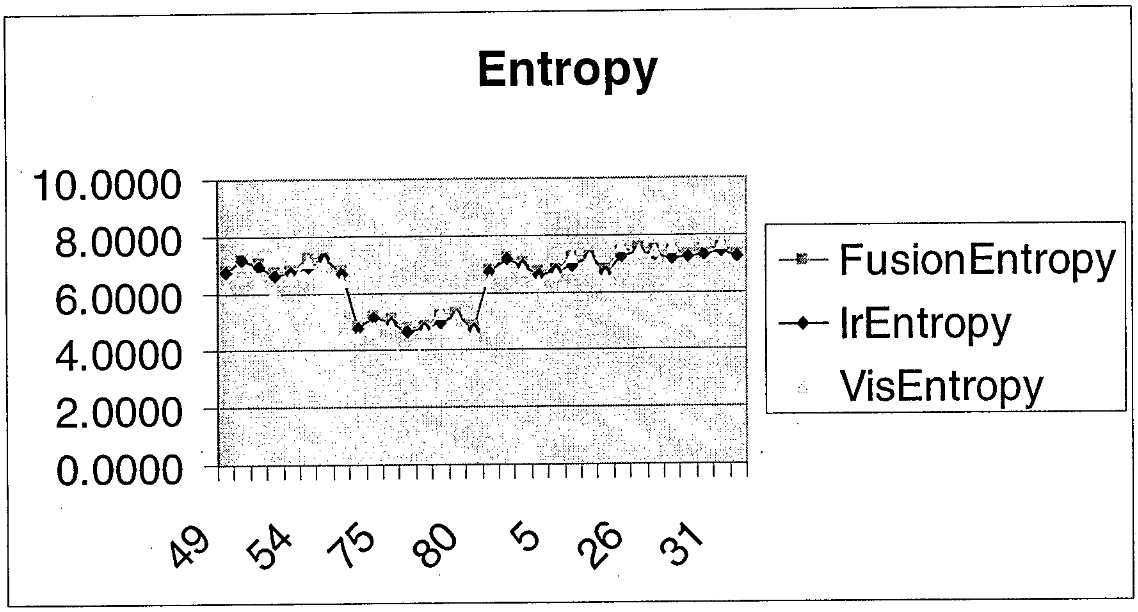

Figure 12. The entropy of the image versus the image type.

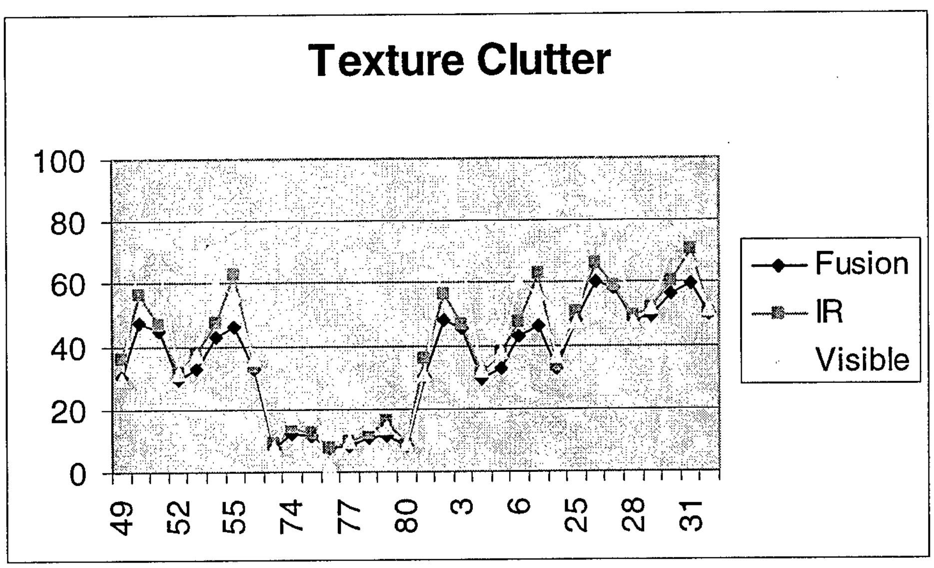

Figure 13. The texture clutter versus the image type.

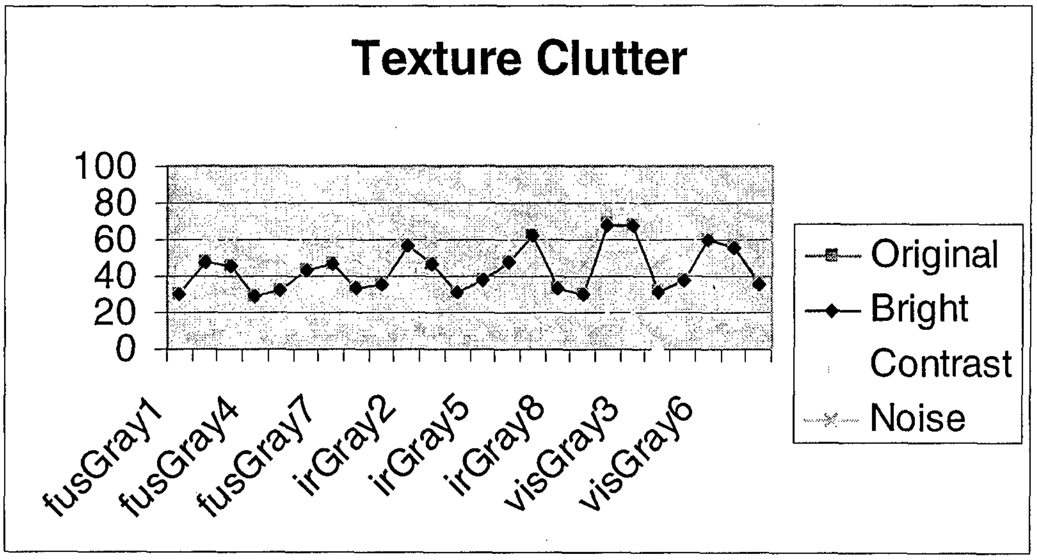

Figure 14. The texture clutter versus the image type.

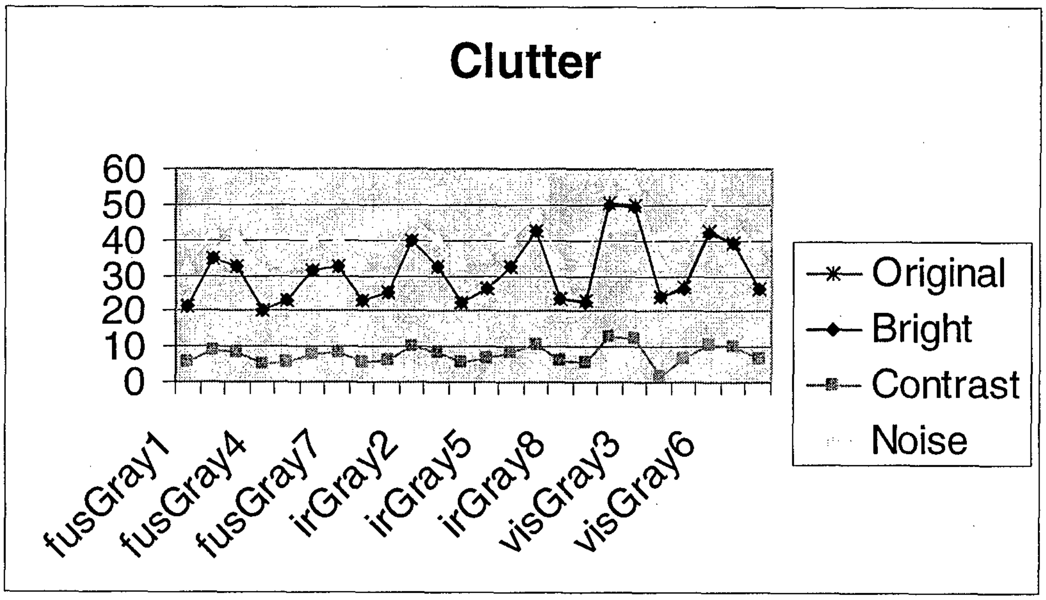

Figure 15. The clutter versus the image type.

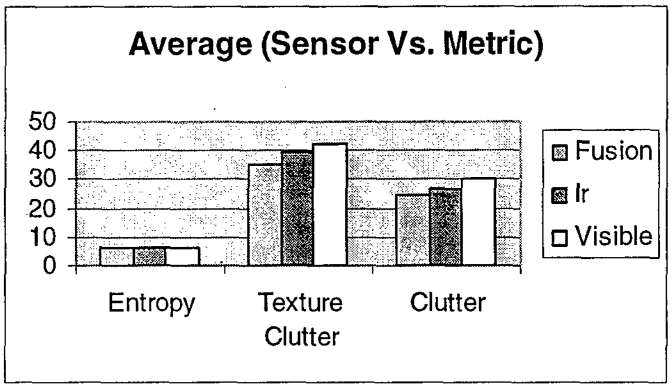

Figure 16. The average sensor versus metric.

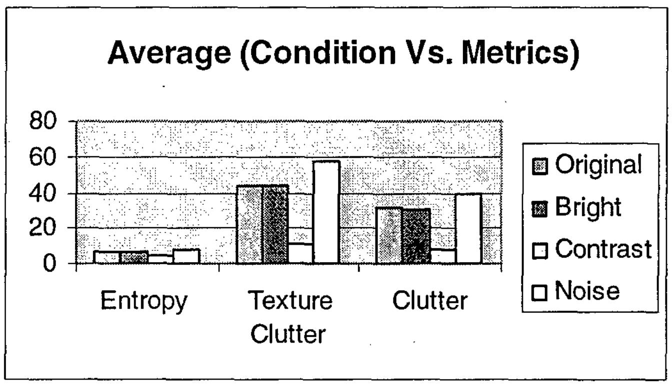

Figure 17. The average condition versus metric.

IV. CONCLUSION.

Entropy correlates to Probability of detection (Pd) for the LAV data set and also the RPG data set. We want to explore other data sets to see if the entropy metric correlates well for them. The statistical results of the RPG test were analyzed. The detection rate of the RPG in the image was analyzed versus the sensor type. People had the highest detection rate of the RPG using the IR sensor, and the lowest detection rate was with the visual image. People taking the RPG test also commented on the fact that they preferred the fused image, because it had more details. We could not find correlation between entropy and Pd in the RPG test (the correlation coefficient was -0.0144). We observed correlation between the entropy and textured clutter in the RPG test. The correlation coefficient was 0.876.

Future experiments will involve using functional magnetic resonance imaging (FMRI) as a window into the visual cortex to develop system models of how the human visual detection system works.

Conclusions

- We want to make the experimental error as small as possible and this is done by removing the variability between subjects from the experimental error. This is accomplished by using the subjects as blocks.
- The importance of randomizing the presentation order is the following: small differences in the detectability will be masked by the bias imposed on the data if everyone sees the same order of imagery.
- A photosimulation laboratory is the ideal environment to compare image metrics against probability of detection.
- Using the lab environment, image sets can be archived for future use and used on different subject pools with different training and instructions. These are some of the benefits of using a laboratory as compared to the field as the location for the test.

16171

OPSEC REVIEW CERTIFICATION

(AR 530-1, Operations Security)

I am aware that there is foreign intelligence interest in open source publications. I have sufficient technical expertise in the subject matter of this paper to make a determination that the net benefit of this public release outweighs any potential damage.

Reviewer: Edward J Pecke | 12 | Susceptibility/Uses
Name | Grade | Title

Edward J Pecke | 20 Aug 03
Signature | Date

Description of Information Reviewed:

Title: BENEFITS OF PHOTOSIMULATION & SENSOR FUSION FOR THREAT DETECTION

Author/Originator(s): Bankowski, Bednarz, Meitzler, et al

Publication/Presentation/Release Date: Aug. 22nd 03

Purpose of Release: technical paper / draft

An abstract, summary, or copy of the information reviewed is available for review.

Reviewer's Determination (circle one):

1. Unclassified Unlimited.

2. Unclassified Limited, Dissemination Restrictions IAW ________

3. Classified. Cannot be released, and requires classification and control at the level of ________.

Security Office (AMSTA-CS-S):

Concur/Nonconcur ________ 25 Aug 03
Signature | Date

Public Affairs Office (AMSTA-CS-CT):

Concur/Nonconcur ________ 25 SEP 03
Signature | Date