AFRL-HE-WP-TR-2000-0157

UNITED STATES AIR FORCE RESEARCH LABORATORY

VISUAL PERFORMANCE MODEL ANALYSIS OF HUMAN PERFORMANCE IN IR TARGET RECOGNITION FOR THIRD GENERATION FLIR

Michael D. Dunkel
Fredrick G. Smith
Allyn W. Dunstan

OPTIMETRICS, INC.
3115 PROFESSIONAL DRIVE
ANN ARBOR, MI 48104-5131

JUNE 1999

FINAL REPORT FOR THE PERIOD MAY 1998 TO JUNE 1999

Approved for public release; distribution is unlimited.

Human Effectiveness Directorate
Crew System Interface Division
2255 H Street
Wright-Patterson AFB OH 45433-7022

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Air Force Research Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
8725 John J. Kingman Road, Suite 0944
Ft. Belvoir, Virginia 22060-6218

TECHNICAL REVIEW AND APPROVAL

AFRL-HE-WP-TR-2000-0157

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER

MARIS M. VIKMANIS
Chief, Crew System Interface Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1999	Final Report - May 1998 to June 1999

4. TITLE AND SUBTITLE
Visual Performance Model Analysis of Human Performance in IR Target Recognition for Third Generation FLIR

5. FUNDING NUMBERS
C: F41624-94-D-6000
PE: 62202F
PR: 7184
TA: 10
WU: 44

6. AUTHOR(S)
Michael D. Dunkel
Fredrick G. Smith
Allyn W. Dunstan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Optimetrics, Inc.
3115 Professional Drive
Ann Arbor, MI 48104-5131

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory, Human Effectiveness Directorate
Crew System Interface Division
Air Force Materiel Command
Wright-Patterson AFB, OH 45433-7022

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARFL-HE-WP-TR-2000-0157

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The eventual goal of this research is to develop a method that emulates human vision and hence is consistently able to predict human performance in target detection/recognition/identification tasks. The model used for simulating human visual performance is the OptiMetrics' developed Visual Performance Model (VPM). With the inclusion of proper calibration parameters, any visual or thermal spectrum sensor can be modeled with VPM. This modeling capability will yield a design tool, which economically allows comparison of various types of sensors and displays, and eliminates the need for many operator tests of prototype devices. This will result in better system designs and may shorten the acquisition cycle for target acquisition systems. A previous research effort compared detectability (d') values resulting from laboratory target detection experiments with those predicted by VPM for first generation (1st Gen) FLIR imagery. Good correlation between VPM predictions and human operator experiments was obtained. This report documents VPM's d' predictions for third generation (3rd Gen) FLIR imagery. Laboratory d' results for the 3rd Gen imagery were not available at the time main text was written, however, those experimental values are documented in an appendix. The d' values calculated by VPM agree well with the 3rd Gen experimental values. Based on that comparison, it appears that VPM is a valuable tool for prediction sensor and display performance.

14. SUBJECT TERMS
3rd Gen FLIR, Visual Performance Model, Target Detection, Target Recognition, Operator Visual Performance, Forward Looking Infrared

15. NUMBER OF PAGES
26

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNL

This page intentionally left blank.

PREFACE

This effort was accomplished under Contract F41624-94-D-6000, Delivery Order 0007 for the Air Force Research Laboratory's Human Effectiveness Directorate, Crew System Interface Division, Information Analysis and Exploitation Branch (AFRL/HECA). It was completed for the prime contractor, Logicon Technical Services, Inc. (LTSI), Dayton, Ohio, under Work Unit No. 71841044: "Target-Related Studies." Mr. Don Monk was the Contract Monitor and Mr. Gilbert Kuperman was the Technical Monitor.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

This page intentionally left blank.

1.0 INTRODUCTION

This report documents work performed by OptiMetrics, Inc. under subcontract P.O. 8137 to prime contract F41624-94-D-6000. The eventual goal of this project is to develop for AFRL a method that emulates human vision and hence is consistently able to predict human performance in detection/recognition/identification tasks for targets under all conditions encountered in the real world. The model used for simulating human visual performance is the OptiMetrics' developed Visual Performance Model (VPM). With the inclusion of proper calibration parameters, any visual or thermal spectrum sensor can be modeled with VPM. This modeling capability will yield a design tool, for the development of crew displays and target acquisition systems, that economically allows comparison of many types of sensors and displays, and eliminates the need for many operator tests of prototype devices. If successful, this will result in better system designs and may shorten the acquisition cycle for target acquisition systems.

Earlier efforts under this contract compared detectability (*d'*) values resulting from laboratory target detection experiments with those predicted by VPM for first generation forward-looking infrared (FLIR) imagery. VPM simulates early human vision processing, after accounting for the processing characteristics of any sensor or optical aid. Calibration parameters, developed from human operator detection experiments, are input to the model to approximate the higher levels of human vision processing. The VPM model was initially developed for the evaluation of the detectability of camouflaged and low signature ground targets with visual spectrum sensors. The model's capabilities have evolved through several additional applications, including evaluation of the conspicuity of automobiles and adaptation for processing FLIR imagery. In a previous project, model calibration parameters were evaluated from first generation FLIR imagery. Good correlation between VPM predictions and human operator performance was obtained. This work is documented in the paper, The Visual Performance Model: Predicting Target Recognition Performance With First Generation FLIR Imagery[1], delivered in March, 1998.

This report documents the next phase in development of the human performance modeling capability. For this effort, images were gathered with an airborne third generation FLIR sensor. A set of these images selected by Logicon staff was processed through VPM to calculate the predicted *d'* parameter. Laboratory detection experiment results for the third generation imagery were not available at the time this report was written. However, *d'* values calculated by VPM based on the earlier calibration appear reasonable. Further evaluation of the VPM results versus those of the operator experiments will be accomplished when the laboratory results are available.

[1] Fredrick G. Smith, Joe T. Riegler, Gilbert G. Kuperman, The Visual Performance Model: Predicting Target Recognition Performance with First Generation FLIR Imagery, IRIS Specialty Meeting on Passive Sensors, March, 1998.

2.0 VISUAL PERFORMANCE MODEL

2.1 Overview

The Visual Performance Model (VPM) was described fully in Section 2.0 of the report documenting the evaluation of first generation FLIR images[2]. In consideration of readers who are not in possession of that document, or any earlier document describing the use of VPM, some of the description is repeated here. Those readers familiar with the VPM methodology and input/output protocols need not read the remainder of Section 2.0.

VPM is an evolution of a human visual performance model developed by OptiMetrics, Inc. VPM development has largely been funded by the U.S. Army through the Tank-Automotive RDE Center (TARDEC). The original version of the model, called TARDEC Vision Model (TVM), was developed to address the detectability of military vehicles. A later version, NAC-VPM, is a modified version that was used for various applications, including examination of the conspicuity of automobiles. That version is essentially the version that was used for evaluation of both first and third generation FLIR imagery under this project. NAC-VPM is documented in the OptiMetrics report OMI-577[3].

VPM is more than a rigid code that implements one specific representation of human vision processing. It actually consists of a number of C++ modules that simulate various processes in the visual systems. The modules are then connected together to perform the series of operations necessary to simulate the overall vision process. The instructions that determine how the various "atomic" modules function together are contained in "Map" files. An understanding of what VPM is doing can be obtained by examining the Map files that define the various component processes.

VPM's utility has been demonstrated in various applications. The Army Materiel Systems Analysis Agency (AMSAA) compared three search and target acquisition models against field data on visual detection of military targets[4]. AMSAA concluded:

> "In general, this comparison shows that with proper calibration, ACQUIRE, ORACLE and TVM NAC-VPM perform about the same when compared to the summer 1994 visual data.
> ...
> the increased complexity of the TVM NAC-VPM model may be of more benefit in the prediction of observer performance against more difficult (i.e., signature managed) targets."

A second comparison was the result of a joint effort between TARDEC, General Motors Research Laboratory, and OptiMetrics. That effort compared and calibrated VPM to experimental data on the ability of drivers to detect oncoming traffic in complex background scenes. That analysis contained 736 cases and a correlation coefficient of 0.79 was found between the VPM d' predictions and observer results obtained by TARDEC. The RMS error between the predicted and observer d' values was 0.56. In that study, the VPM predictions tracked the experimental results over a broad set of environmental and observational conditions ranging from highly visible through barely detectable automobile visibility situations.

2.2 VPM Early Vision Processing Model

VPM simulates the complex image processing chain representing early-vision. Since VPM was designed for detection analysis of visible targets, its modeling includes representation of the effects of color vision. It also includes a capability to model the effects of target movement on detection. The overall early-vision processing represented in VPM is shown in Figure 1.

---

[2] Frederick G. Smith and Allyn W. Dunstan, Visual Performance Model Analysis of Human Performance in IR Target Recognition (AFRL-HE-WP-TR-1998-0065), WPAFB, OH, AFRL.

[3] G. Witus, TARDEC National Automotive Center Visual Perception Model (NAC-VPM), Final Report: Analyst's Manual and User's Manual, OMI-577, (Ann Arbor, MI: OptiMetrics, Inc., September, 1996). Release point for NAC-VPM is Thomas Meitzler of TARDEC, (810) 574-7530.

[4] AMSAA, "SUBJECT: Search and Target Acquisition Model Comparison," MEMORANDUM for Deputy Under Secretary of the Army for Operation Research and Director Assessment and Evaluation, 6 November 1996.

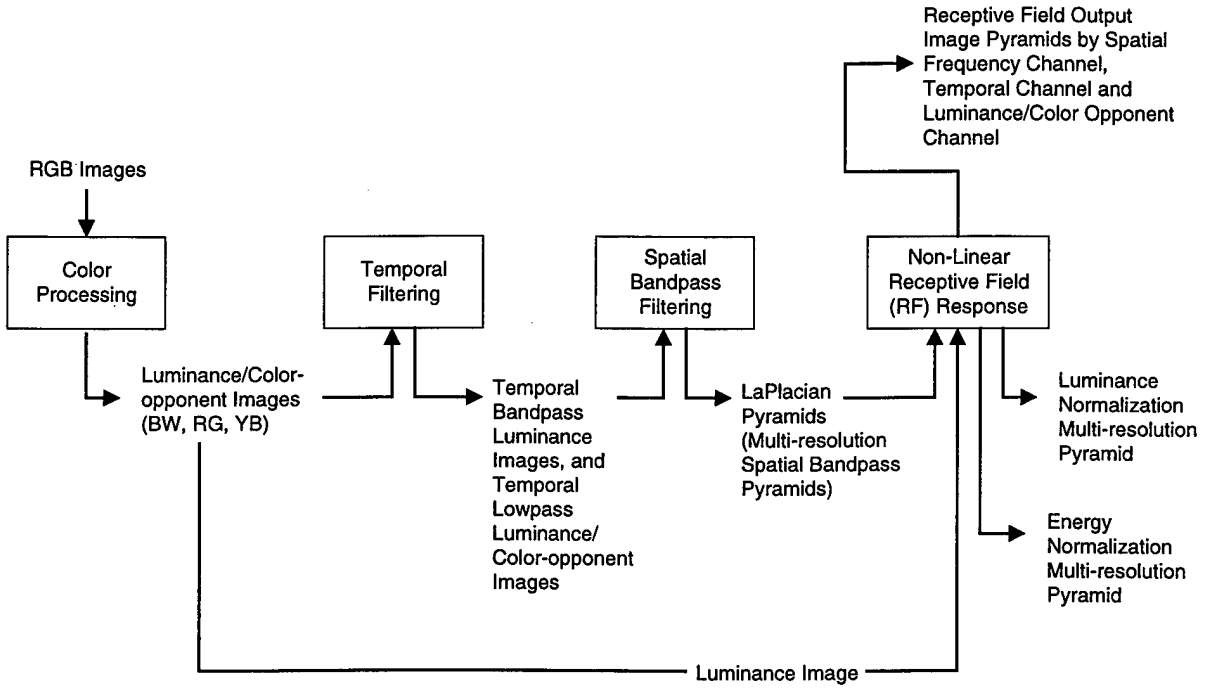


Figure 1. Early-vision Processing Modeling Included in VPM.

For the present project, where the FLIR imagery is presented as grayscale displays, the color processing is not relevant; hence only the luminance images are used. Similarly, the temporal filtering step is not relevant. The temporal processing is significant if the target image on the display is moving at a rate greater than a few tenths degrees-per-second. Because the targets are stationary in the FLIR imagery used for this study, the temporal processing step in VPM was eliminated.

2.3 Target Metric Summation and Predicted *d'*

The result of the early-vision simulation processing is a set of multi-resolution images that represent the receptive field (RF) responses to the input image. The RF images initially include the entire scene; what is needed is a means to select out only the information that the presence of the target contributes to the scene. In VPM the target information is separated through the following process:

1. A target outline is created by the user with the EdTarget utility provided with VPM.
2. The target is "cut-out" of the image.
3. The surrounding scene is blended in, using extrapolation of the surrounding background textures at each level of the multi-resolution images.
4. The blended multi-resolution background images created in step 3 are subtracted from the multi-resolution images of the full scene. The result is that background features are cancelled, leaving the scene components mainly resulting from the target's presence.

The results of the above processes are the TLBWTargetRFDetectability multi-resolution images created by VPM. An example of these multi-resolution images is shown in Figure 2. Those images are decompositions of the raw target image into components. This decomposition process is thought to be consistent with the processing in the human vision processing chain.

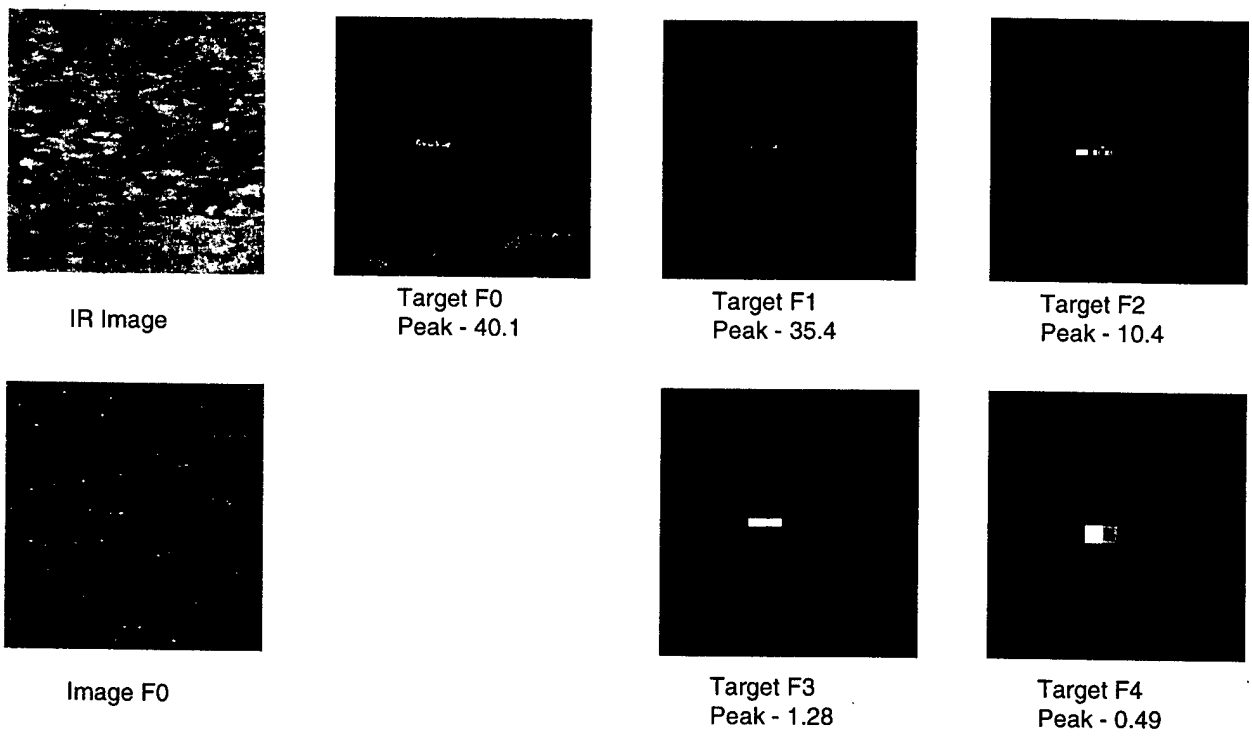


Figure 2. Illustration of RF Detectability Multi-resolution Images of an IR Target Created in VPM.

To provide a single measure (metric) of target detectability, VPM integrates the signal energy represented in the RF detectability multi-resolution images. The general approach used by VPM to compute an aggregate target metric is shown in Figure 3. For the FLIR application discussed here, only the Temporal-Lowpass, Black-White multi-resolution images are included in the summation. VPM also includes the capability to weight the various spatial channels in the sum. The result of this summation is sometimes called the image "energy." The natural logarithm (*ln*) of the energy is the "raw" target metric, or "raw *d'*." Experience has shown that a linear function of the raw *d'* can be correlated to the image detectability values obtained from observer experiments. The determination of the slope, *a*, and intercept, *b*, relating the raw *d'* to detectability is the calibration process for VPM. It is expected that the calibration coefficients (*a*, *b*) will depend on the task that the human observer is asked to perform as well as other details of VPM implementation (e.g., weighting of the various image planes). The calibration coefficients derived for the automobile detection task were slope, 1.02, and intercept, -4.26. VPM had not been previously used for a complex detection/recognition task, as examined in the first generation FLIR imagery report. The derivation of the calibration parameters from the first generation FLIR imagery is described in Section 4.0.

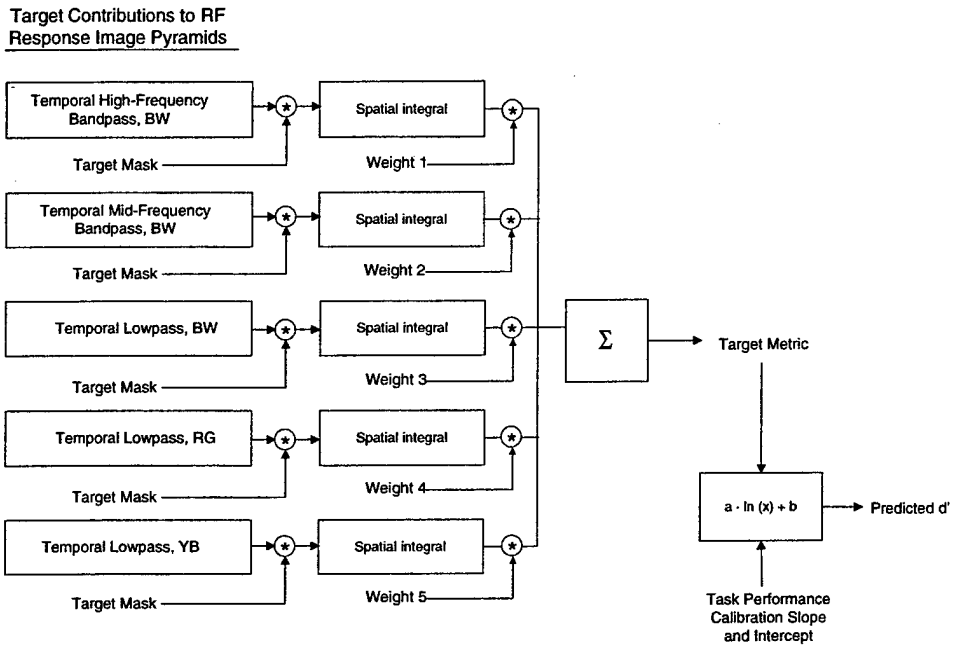


Figure 3. Target Metric Summation and Predicted *d'*.

2.4 VPM Inputs and Outputs

The top-level inputs and outputs of VPM are illustrated in Figure 4. As seen in the figure, the first two inputs are a digital representation of the image, as displayed to the observer, and photometric parameters that allow absolute calibration of that image in radiometric units. Factors also need to be input to describe the characteristics of the human visual system, i.e., for the mean observer. The lower box indicates input of the task performance calibration parameters described in the above paragraph. The final inputs are the definition of the target region and the angle from the observer's viewpoint. The target region is defined by an outline of the target developed by the user with VPM's EdTarget utility. VPM can represent target detection either at, or away from, the center of the eye's focus. The angle from viewpoint is the angle that the target of interest is from the eye's center of focus. In the FLIR imagery used in both the first and third generation device studies, the observer was cued to the object of interest, thus the angle from viewpoint is zero for these cases.

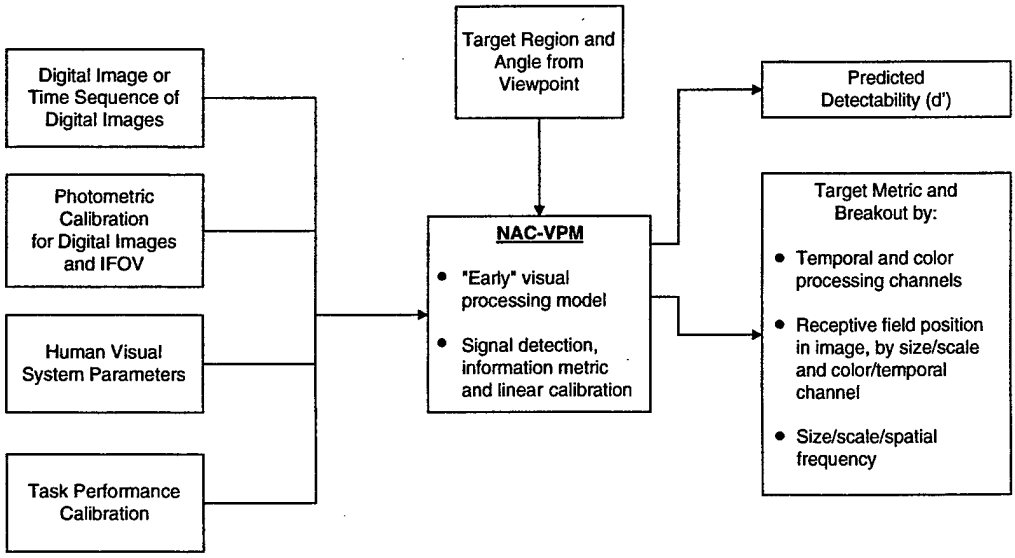


Figure 4. Top Level Illustration of VPM Inputs and Outputs.

The main output of VPM is the predicted detectability, *d'*, of the target. In addition to the *d'* value, VPM also produces a number of intermediate outputs that are useful for verifying correct operation of the model and/or for analysis purposes. Those outputs include images of the "energy" contained in various color, temporal, and spatial channels, as illustrated above in Figure 2. The integrated energy for each of the channels is also output. That output allows analysis of the relative contributions of each spatial channel to the overall predicted detectivity.

The specific VPM data flow is illustrated in Figure 5. Shown there are each of the input/output files that are used/generated by the various components of the model. All of these files are described in the reference cited in footnote 3. The first VPM module is the Convert utility. That utility simply converts a Silicon Graphics format RGB image into the internal data format (Channel Data Packet format) used by VPM. All of the subsequent images generated by VPM are in this CDP format. VPM provides the utilities View and ViewB to display CDP images on the SGI. The second VPM module is the EdTarget utility. Given the input image, EdTarget allows the user to outline the target area within the image.

The two main modules of VPM are the Static Spatial Vision Analyzer and the Static Metric Analyzer. The Static Spatial Vision Analyzer simulates the linear early-vision processes and applies them to the input image. The outputs are the decomposed images representing early-vision effects on the various spectral and spatial components. These files can be viewed using the View or ViewB utilities. The Static Metric Analyzer simulates the non-linear vision processes and sums the contributions from the various image components to determine a single target metric, *d'*. Text files that provide intermediate results of the metric calculations are also created. Other outputs from the Static Metric Analyzer are the Receptive Field (RF) "energy" images: ImageRFDetectability and TargetRF Detectability. Those outputs graphically illustrate how the various image components contribute to the overall target detectability.

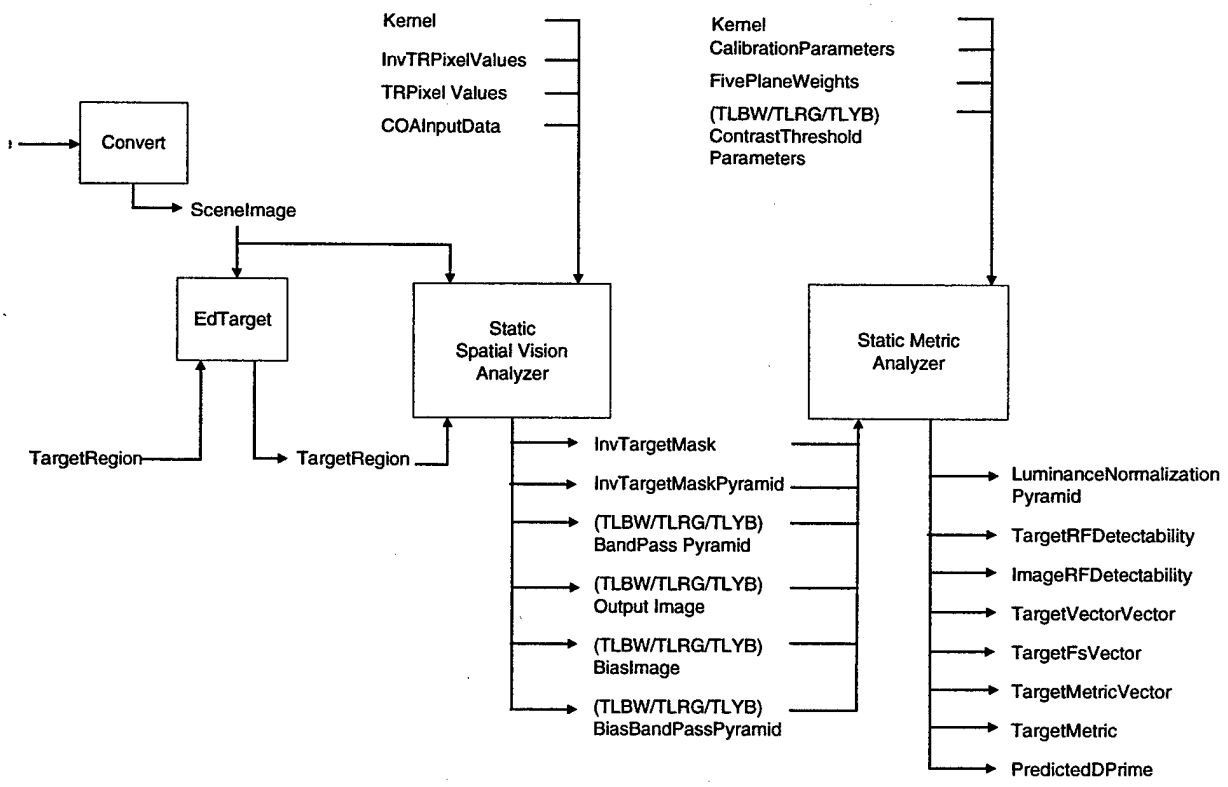


Figure 5. Input and Output Files Used or Created by VPM.

3.0 VPM ANALYSIS PROCESS FOR THIRD GENERATION FLIR IMAGERY

Selected frames of the third generation FLIR imagery, taken from an airborne sensor, were supplied to OptiMetrics by Logicon[5] for analysis using VPM. Images were provided from three missions, all flown during daylight hours. During each mission, an array of three targets was approached from various aspect angles. Data was recorded from ranges of 23 kilometers to overflight. A different background setting (open, sparse treeline, and treeline) was extant for each of the three missions. The three targets were a Scud-B mobile missile transporter-erector-launcher (TEL), a ZiL 131 communication van, and a MAN 4-axle all-wheel drive truck with an air compressor unit. The FLIR imagery was recorded on digital tape for later analysis. 76 frames of the imagery were supplied to OptiMetrics. These images included frames from several aspect angles grouped into eight different range bins (8, 10, 12, 14, 16, 18, 20, and 22 km) for each of the three missions. Each range bin partition encompasses a one kilometer range variation, e.g., the 8 km bin included data from 8 to 9 km actual range.

Before analysis with VPM, the FLIR imagery had to be transformed to be consistent with the VPM assumptions and requirements. The steps performed to convert the imagery to the required VPM formats are outlined in Table 1, below. Most of the conversion processes were performed using a shareware image processing tool for the SGI system called XV[6].

TABLE 1. IMAGERY PREPARATION PROCESS

Step	Purpose	Comment
Convert from raw to rgb format.	VPM process starts with rgb image.	Images in raw format were supplied by Logicon.
Resample image.	Scale to provide square pixel of known size.	This provides 1/32 degree square pixels given the display viewing distance of 76 cm.
Crop image.	VPM requires square image.	A 180x180 image is extracted from each scene. When possible, the image was centered on the TEL.
Save as rgb image in a separate directory.	Save intermediate results.	One directory is created for each image and all subsequent intermediate results are also saved in that directory.

In addition to the image conversions, some calibration parameters also need to be input into various set-up files. Generally, most parameters can be left as the defaults provided with the VPM model. Table 2 displays the modified parameters for the present analysis. These values are the same as used in the first generation FLIR study.

One issue that has not been resolved from the theoretical or empirical data on early-vision is the relative weight to be placed on the various spatial and color channels. Since here we are not analyzing color imagery, the color channel weighting is assumed to be zero; a weight of 1 is given to the luminance (i.e., black-white) channel. However, the weighting of the spatial channels is an issue. Reasonable spatial weighting options are constant, 1/f or $1/f^2$ weighting. For the present study and the earlier first generation FLIR study, we have chosen to use a zero weight for the highest frequency channel and to weight the other channels as 1/f. This weighting is heuristically justified as follows: 1) there is clutter from the background and eye noise on the high frequency channel so we suggest it contributes little to target detection/recognition; 2) the 1/f weighting provides reasonable weight on the intermediate frequencies where most of the target information is visible; and 3) the results are reasonable. This same weighting scheme was employed during analysis of the first generation FLIR imagery, where it also yielded reasonable results.

[5] Bill Janson, Private Communication to Fred Smith, 17 Dec 98 and 30 April 99.
[6] XV © John Bradley. Can be obtained by anonymous ftp on ftp.cis.upenn.edu.

TABLE 2. INPUT PARAMETERS FOR VPM

File	Parameter(s) Modified	Values
COAInputData	RGBExponent	0.75, 0.75, 0.75
TLBWContrastThresholdParameters	IFOV	0.03125
FivePlaneWeights	THBW TMBW TLBW TLRG TLYB	0 0 1 0 0
DprimeParameters	Slope* Intercept*	1 0.453 0 - 0.041
SpatialWeights **	Weights	0,2,4,8,16,0,0,0

* 1 and 0 are used for "raw *d'* calculation." For calibrated *d'*, values of 0.453 and -0.041 are used, as derived from the first generation FLIR imagery. See Section 4.0.
** New file created for this application

3.1 Examples of VPM Processing

Figures 6 illustrates examples of the third generation FLIR images analyzed in this effort. The three images shown are from mission 6370, which was flown while the targets were positioned in an open background site. All images are from aspect angle F6, approximately 22° rotation from a direct side view of the targets.

8 km Range Bin 14 km Range Bin. 22 km Range Bin.

Figure 6. Third Generation FLIR Images Taken During Mission 6370.

Figure 7 shows the VPM weighted energy metrics as a function of spatial channel number (inverse to frequency) for the TEL target in the three images presented in Figure 6. These curves are typical of the data for the third generation FLIR images in that:

- The highest energy metric values usually occur in channel 1 or channel 2. The values for the longest-range bins (20-22 km) always peak in channel 1.
- Many of the energy metric values for the shortest-range bins (8-10 km) show nearly equal values in channels 1 through 3.
- While the total received energy decreases with range, a significant percentage of the energy received at short ranges is still present at long ranges.

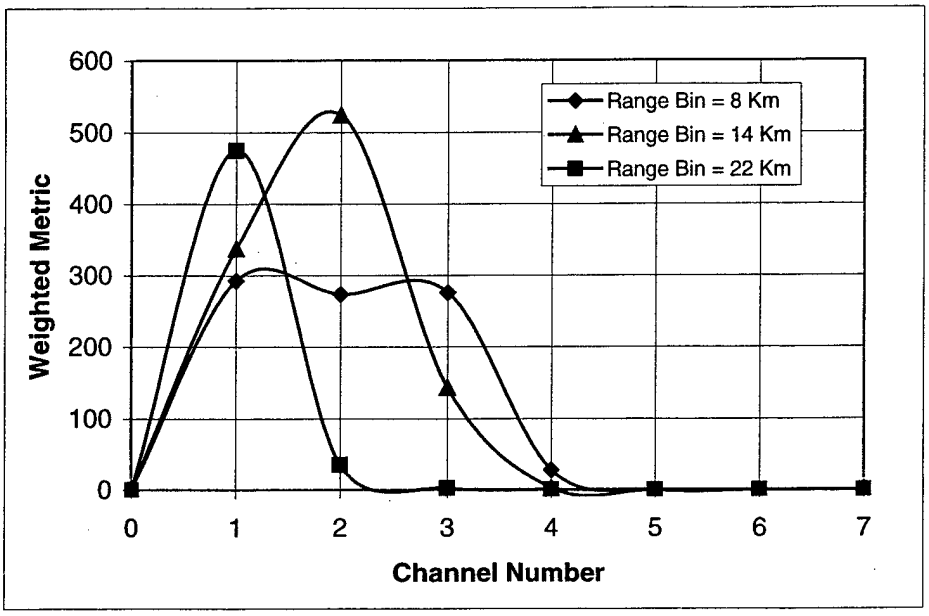


Figure 7. Weighted Energy Metrics for the Images Presented in Figure 6.

4.0 PREDICTION OF DETECTABILITY USING VPM RESULTS

As described in Section 2, VPM calculates a *d'* value using a summation of the weighted energy terms from the image analysis. The general form of the equation used is the following:

$$d' = a * ln\,(\Sigma \text{weighted energy}) + b$$

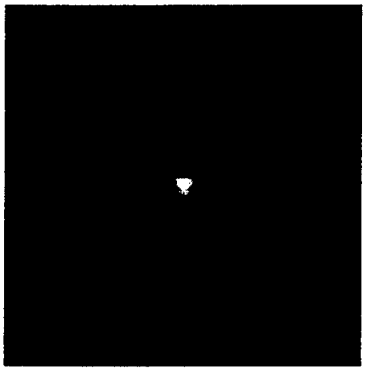
Before the model is calibrated, a value of 1 for a and 0 for b results in a "raw *d'*." Once a and b have been determined, the output is the predicted *d'*. The raw *d'* values computed by VPM for the third generation FLIR imagery are given in Table 3. Also shown in the table are the standard deviations for each data set. The standard deviations in the table result from the fact that raw *d'* values were computed for imagery from several aspect angles for each mission/range bin combination. Those values were then averaged to give the mean raw *d'* values listed. The number of images comprising each data set is also presented in the table.

TABLE 3. RAW *d'* VALUES COMPUTED USING VPM

Mission Designation	6370			6371			6372		
Background Description	Open			Sparse Treeline			Treeline		
Range Bin	# Images	Raw *d'*	SD	# Images	Raw *d'*	SD	# Images	Raw *d'*	SD
8	4	6.468	0.671	4	6.754	0.415	4	5.968	0.658
10	4	6.285	0.423	4	6.783	0.113	4	5.673	0.254
12	4	6.764	0.397	4	6.128	0.322	4	5.436	0.426
14	4	6.831	0.196	4	6.033	0.355	4	5.566	0.357
16	2	6.791	0.360	2	5.707	0.389	3	5.521	0.358
18	2	6.739	0.100	2	5.456	0.709	2	5.522	0.415
20	2	6.516	0.058	2	5.834	0.082	3	5.117	0.310
22	2	6.255	0.020	2	5.490	0.038	4	4.883	0.471

In general, the raw *d'* value decreases with range, as expected. This is not true for the shortest-range bins, especially as evidenced by the mission 6370 analysis results. It is believed that this occurrence is due to the large variation in presented warm target area, and thus in detectability, that occurs at short ranges for different aspect angles. Figure 8 presents the four images that comprised the mission 6370, 8 km range bin data set, along with the individual *d'* values calculated by VPM for the images. Note that aspect angle F1 is an approach towards one end of the TEL, while F7 is an approach approximately 22° off the other end of the TEL. These images produced significantly lower raw *d'* values then those for aspect angles F4 and F6, which were approximately 22 (in opposite directions) from a direct side view of the TEL. This change in raw *d'* with aspect angle is not as evident at greater ranges. Figure 9 shows images from the 14 km range bin for the same 4 aspect angles. At that range, all four raw *d'* values are nearly equal. Images at greater ranges are not presented because generally only two images were

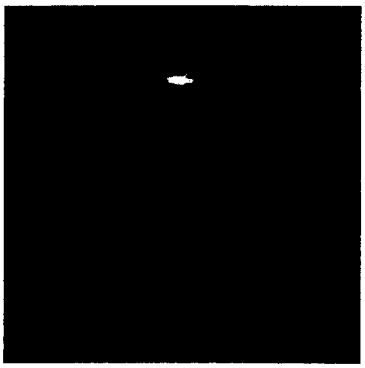
provided for analysis at those ranges, and a valid comparison to the four images of Figure 8 could not be made. Four images were collected in the 22 km range bin during mission 6372, but they were at aspect angles F1, F2, F7, and F8, all of which are towards the ends of the TEL. The statistical uncertainty associated with having only two images for evaluation may also account for the two situations where raw d' did not decrease with range for certain range bins between 14 km and 22 km (mission 6371, range bin 20 and mission 6372, range bin 18).



Aspect Angle: F1, $d' = 6.111$

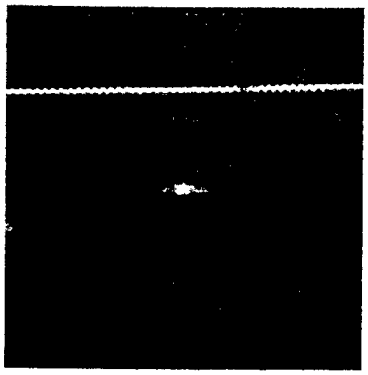
Aspect Angle: F4, $d' = 7.249$

Aspect Angle: F6, $d' = 6.768$

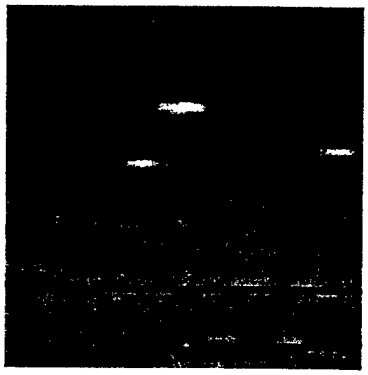


Aspect Angle: F7, $d' = 5.745$

Figure 8. Images from Four Different Aspect Angles at a Range of 8 km.

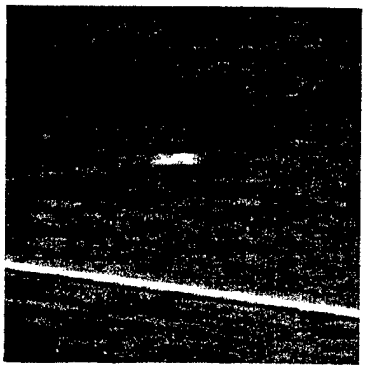


Aspect Angle: F1, $d' = 6.537$



Aspect Angle: F4, $d' = 6.928$

Aspect Angle: F6, $d' = 6.914$



Aspect Angle: F7, $d' = 6.943$

Figure 9. Images from Four Different Aspect Angles at a Range of 14 km.

4.1 Calibration of the VPM *d'* Results

In order to determine the constants a and b in the above equation, it is necessary to correlate the raw *d'* data with actual detection/recognition task data, collected through laboratory experiments with trained observers. Such data was not available for the third generation imagery. As a first look at corrected *d'* calculations, the a and b values obtained from the study of first generation FLIR imagery were used with the third generation raw *d'* values. The use of earlier coefficients is justified because the calibration is to the task and the character of the imagery presented. In the present case, as in the first generation case, the task is detection-recognition of a TEL and the imagery character is infrared imagery in similar background conditions. Thus it is reasonable to apply the earlier derived calibration constants for the third generation performance predictions. Derivation of calibration constants in the first generation study, and also used here, is summarized in the following paragraphs[7].

The first generation FLIR image study used imagery taken from three background sites, for range bin distances from 4 to 8 km, under both day and night conditions. Twelve subjects viewed a series of 240, 5.5 second duration flight sequences, replayed on a high resolution monitor that duplicated the display used in the F-15E. When presented with a crosshair over the intended target (the TEL) or over a background object (one of the other vehicles or a background terrain feature), the operator was asked to indicate target or non-target and rate his confidence of that identification.

[7] Smith, op. cit.

The data collected from the subjects was analyzed by Logicon personnel using the theory of signal detection[8]. Hit and false alarm rates were used to derive the perceptual sensitivity and response bias. Perceptual sensitivity measures the subject's ability to distinguish the signal (target) from noise. The response bias reflects the subject's willingness to identify the existence of the signal. The perceptual sensitivity was measured in terms of the *d'* detectability index. These *d'* values were used for correlation/calibration of the VPM results.

To find a reasonable calibration for VPM for this particular detection/recognition task, various statistical correlations between the VPM raw *d'* values obtained from the first generation FLIR imagery and the empirical *d'* values resulting from the laboratory detection experiments were explored. Multiple linear regressions of the empirical *d'* values as a function of the 4 variables: raw *d'*, time-of-day, site, and background raw *d'* were computed. The only significant predictive variables (i.e., those with T-ratios greater than 0.80) were found to be the raw *d'* and time-of-day. The regression equation found for those two variables is:

$$d' = 1.40 + 0.179 * \text{raw } d' + 0.458 * \text{time-of-day} \quad \text{(time of day is 0 for day and 1 for night)}$$

The above equation results in an adjusted r^2 of 61.5%. The standard deviation about the regression is estimated as 0.30. If time-of-day is considered as a valid independent variable, this is a valuable result. It is our belief that the improved performance at night results from more lenient observer detection criteria for the low clutter nighttime conditions. The above equation can be used to predict *d'* values for conditions outside those where empirical operator results are available. Such predictions require a VPM analysis of the measured imagery and specification of a day or night condition.

In theory, if the operators always used the same criteria for their detection/recognition task, the VPM output should not need to be supplemented by the auxiliary time-of-day variable to give accurate *d'* predictions. To further explore this possibility, the data was segregated into daytime and nighttime data. A regression of just the daytime data to the raw *d'* gave the following result. The coefficients in this equation are recommended as a conservative method for prediction of FLIR performance using VPM.

$$d' = -0.041 + 0.453 * \text{raw } d'$$

If all experimental conditions are equal between the first generation and the third generation FLIR experiments, then the calibration coefficients represented in the above equation should be useful for both sets of experiments. By all things being equal, we mean the uncontrolled factors in the experiments like the inherent target signatures, the atmospheric attenuations, and the task criteria of the observers, are on the average equal. As discussed in the first FLIR study, we believe that the observer criteria may have changed when they viewed the nighttime data. However, for comparing the first generation results with the thrid generation data, which is entirely daytime data, we believe that use of the above calibration coefficients provides a valid approach.

Applying the above formula to the raw *d'* values generated for the third generation FLIR images (Table 3) yields the calibrated *d'* values presented in Table 4. These values seem reasonable for the images analyzed. It may be noted that they are of the same order as the corrected *d'* values calculated for the first generation FLIR imagery, although the present images were taken at greater distances.

[8] See, J.E., et. al., Unaided Target Acquisition Performance with First Generation Forward-Looking Infrared Imagery: A Signal Detection Theory Analysis (AL/CF-TR-1996-0094), WPAFB, OH, AFRL.

TABLE 4. CALIBRATED *d'* VALUES .

Mission Designation	6370	6371	6372
Background Description	Open	Sparse Treeline	Treeline
Range Bin	*d'*		
8	2.889	3.019	2.663
10	2.806	3.032	2.529
12	3.023	2.735	2.422
14	3.053	2.692	2.480
16	3.035	2.544	2.460
18	3.012	2.431	2.460
20	2.911	2.602	2.277
22	2.793	2.446	2.171

4.2 Comparison of First Generation and Third Generation FLIR Performance Predictions

Using the results in Table 3 and the earlier calibrated *d'* predictions from the first generation study, we can estimate the performance improvement that would be gained by moving to the third generation system. This comparison is show in Figure 10.

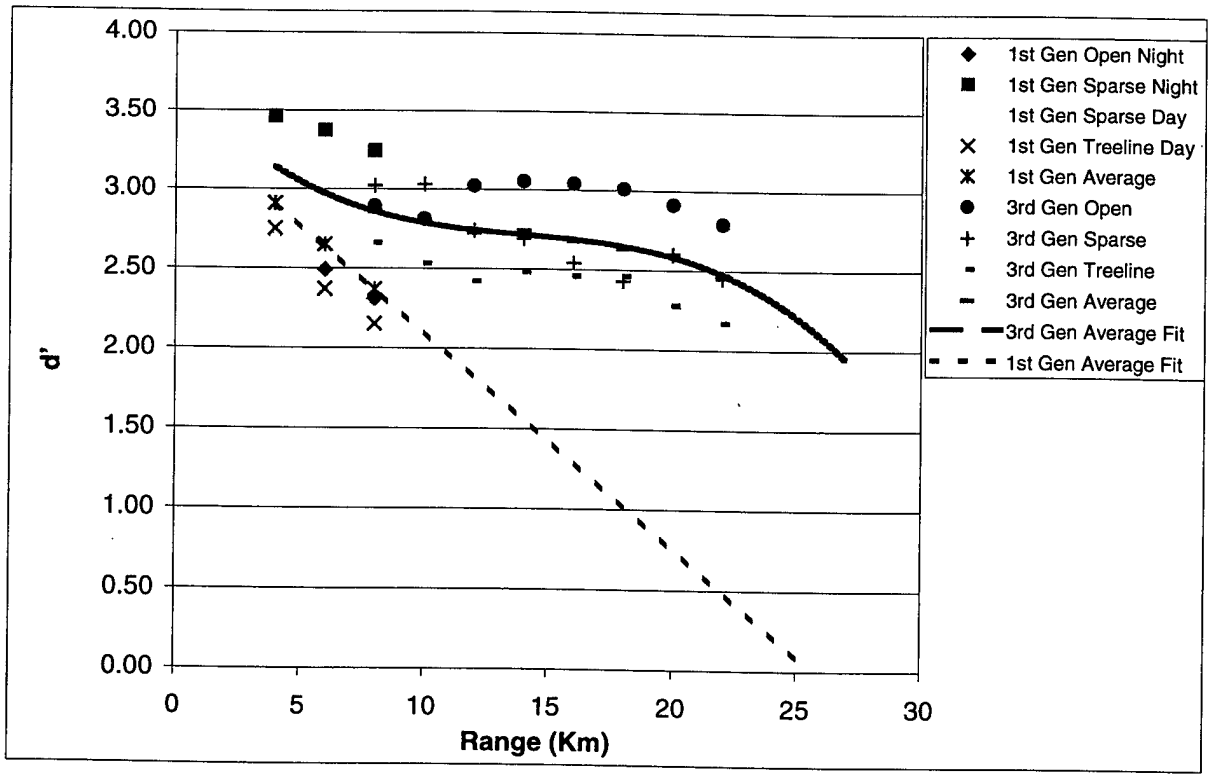


Figure 10. Comparison Between First and Third Generation FLIR Performance Based on the Calibrated VPM Results

That figure includes points from all four datasets from the first generation study and the three datasets from the third generation study. In addition, the averages for each study and curves fit to the averages are shown. To facilitate comparisons between the first generation and third generation FLIR's, the curve fits are extrapolated beyond the measured results.

Inspection of the graphs in Figure 10 shows that the third generation FLIR is always superior to the first generation FLIR, except for the first generation nighttime-open case which we had previously flagged as anomalous. Note that we have no nighttime third-generation results to compare with that result. Comparing the average curve fits shows the overall superiority of the third generation system. The third generation advantage is small at ranges less than 8 km, but if the reasonable extrapolations are believed, the superiority grows rapidly beyond that point.

Since detectability values, *d*'s, are not intuitive to many readers, we will briefly discuss the meaning of these results. A detectability value of "2" can be considered to represent a "moderately difficult" visual task. We assume that the detectability value of "2" represents the maximum effective range of the FLIR for the visual task and target of interest. Given this assumption, the average effective range of the first generation FLIR is approximately 11 km, while the average effective range of the third generation FLIR is estimated as 26 km. This comparison gives a reasonable measure of the superiority of the third generation FLIR system.

5.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary and Conclusions

No absolute conclusions can yet be drawn from this analysis, since the laboratory observer experiment data was not available to confirm the model calibration and for direct comparison to the VPM predictions. However, the relative differences in the VPM *d'* predictions, and results of the preliminary calibrations, give strong indications of the superiority of the third generation FLIR system.

5.2 Recommendations

Recommendations can be made for near-term and long-term development of the visual modeling techniques demonstrated in this study. Near-term recommendations include confirming and utilizing the initial results of this study. Long-term recommendations address generalizing the VPM process and incorporating it into future AFRL design and evaluation tools.

Near term recommendations are:

- Complete the third generation FLIR operator studies and use them to validate the methods and results of the present study
- Generalize the results of this study and the operator results, and incorporate them into models used for AF systems trade studies for this mission application

Long term recommendations are:

- Perform additional operator testing and VPM analysis aimed at:
  - Clarifying the relationships between target characteristics (size, shape, camouflage, time-of-day, etc.) and detectability
  - Developing separate calibrations of the model for the operator tasks of detection, recognition and identification
- Integrate VPM capabilities into a complete engineering level FLIR modeling capability that includes atmospheric effects, sensor simulation, ATR processing, cueing and search, and human performance representation

Execution of the above recommendations would provide AFRL with a complete design and evaluation capability for future imaging sensor and display systems.

RERERENCES

AMSAA (1996, November 6). *"SUBJECT: Search and Target Acquisition Model Comparison,"* MEMORANDUM for Deputy Under Secretary of the Army for Operation Research and Director Assessment and Evaluation.

See, J.E., Riegler, J.T., Fitzhugh, E. (1996). *Unaided Target Acquisition Performance with First Generation Forward-Looking Infrared Imagery: A Signal Detection Theory Analysis* (Report No. AL/CF-TR-1996-0094). Wright-Patterson Air Force Base, OH: Armstrong Laboratory.

Smith, F. G., & Dunstan, A. W. (1998). *Visual Performance Model Analysis of Human Performance in IR Target Recognition* (Report No. AFRL-HE-WP-TR-1998-0065). Wright-Patterson Air Force Base, OH: Air Force Research Laboratory.

Smith, F. G., Riegler, J. T, Kuperman, G. G. (1998, March). *The Visual Performance Model: Predicting Target Recognition Performance with First Generation FLIR Imagery*, IRIS Specialty Meeting on Passive Sensors.

Witus, G. (1996, September) *TARDEC National Automotive Center Visual Perception Model (NAC-VPM), Final Report: Analyst's Manual and User's Manual*, OMI-577, Ann Arbor, MI: OptiMetrics, Inc.

APPENDIX

Comparison of 3rd Gen FLIR Visual Performance Model Predictions to Observer Results

The results of the observer experiments have recently become available[9]; in this appendix, we compare them with the VPM predictions from the body of this report.

A previous effort compared values resulting from laboratory target recognition experiments with those predicted by VPM for 1st Generation FLIR imagery. Good correlation between VPM predictions and human operator performance was obtained there. As component of that study, two VPM parameters were "calibrated" to match the 1st Gen recognition results. Those same parameter values were used for the 3rd Gen predictions presented here; the only input parameter modified was associated with the optics of the sensor, i.e., the IFOV of the 3rd Gen system.

3rd Gen Comparison

The 3rd Gen predictions are compared to the "No Empty Scene Data (Hacker and Ratcliff)" experimental results supplied by AFRL/HEC. The data reduction assumptions used to derive those results are consistent with the assumptions used for VPM. The comparisons are shown in Figure A-1. In that figure, data from the experimental results are labeled as "Obs" (Observer); the other data sets are VPM predictions at comparable conditions. Generally, the VPM predictions for the individual conditions are somewhat below the experimental results. However, the differences are quite small with a standard deviation of less than 0.4 over all conditions. This represents a small difference in probability of recognition, and is likely within the uncertainties of the experiment.

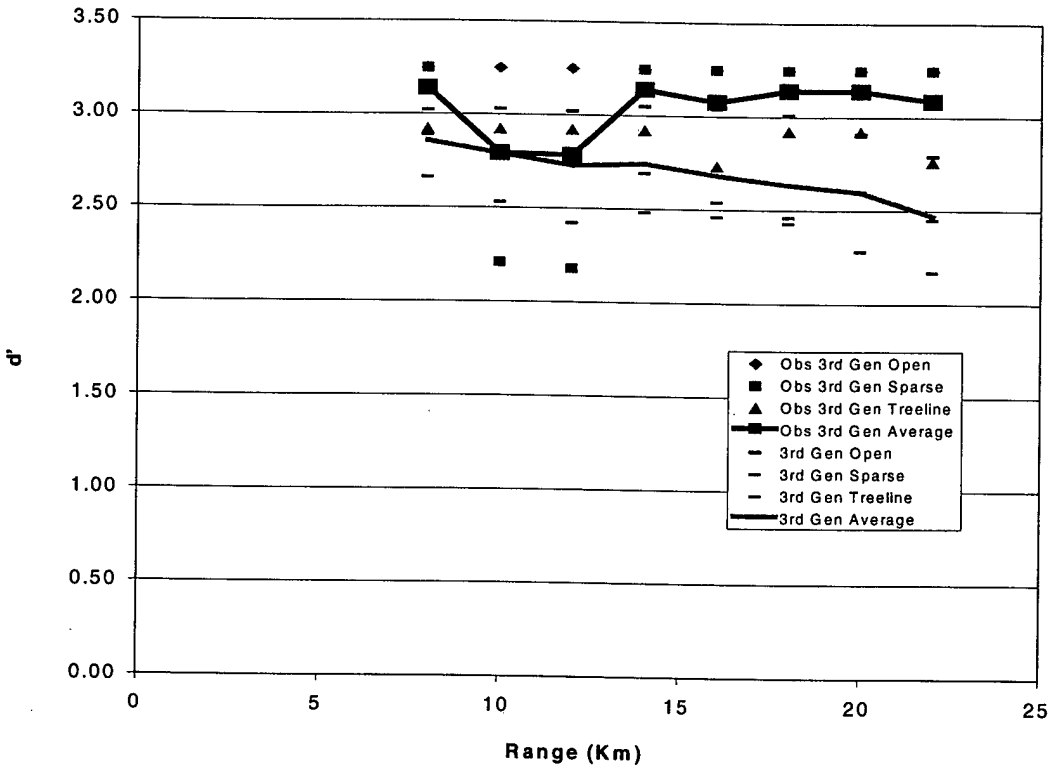


Figure A-1. Comparison of Experimental (Observer) Results with VPM Predictions

The averages over all field conditions at a given range are also shown. The standard deviation between the prediction and experimental curves is approximately 0.2. A correlation between the curves is not meaningful,

[9] Jeff D. Cress, Logicon (AFRL/HEC Support Contractor), Private communication to Frederick Smith, May, 2000.

because both the prediction and the experimental curves are so flat. The experimental data would have to be extended to longer range (lower recognizability) to allow meaningful correlation. The fact that the predictions are within 10% of the experimental *d'* values provides strong confirmation of the model's utility.

In most cases, the predictions are slightly below the observations. This is consistent with the "conservative" calibration used for the model [see references 2 or 3 in main text]. By conservative, we mean the actual system performance is expected to be equal or better then the model predictions. Thus VPM can provide a useful baseline to guide future system designs.

Conclusions and Recommendations

The agreement of the VPM predictions with the experimental results confirms its value as a predictive tool. Both sets of results indicate the strong superiority of the 3rd Gen FLIR over the 1st Gen FLIR system.

Recommendations made in the main text are confirmed by these results; VPM, supplemented with other modeling components, could provide AFRL/HEC with an integrated evaluation capability to assess future imaging sensor and display concepts.

GLOSSARY

AMSAA	Army Materiel Systems Analysis Agency
FLIR	Forward-looking Infrared
NAC-VPM	TARDEC National Automotive Center Visual Perception Model
TARDEC	Tank-Automotive RDE Center
TEL	Transporter-erector-launcher
TVM	TARDEC Vision Model
VPM	Visual Performance Model