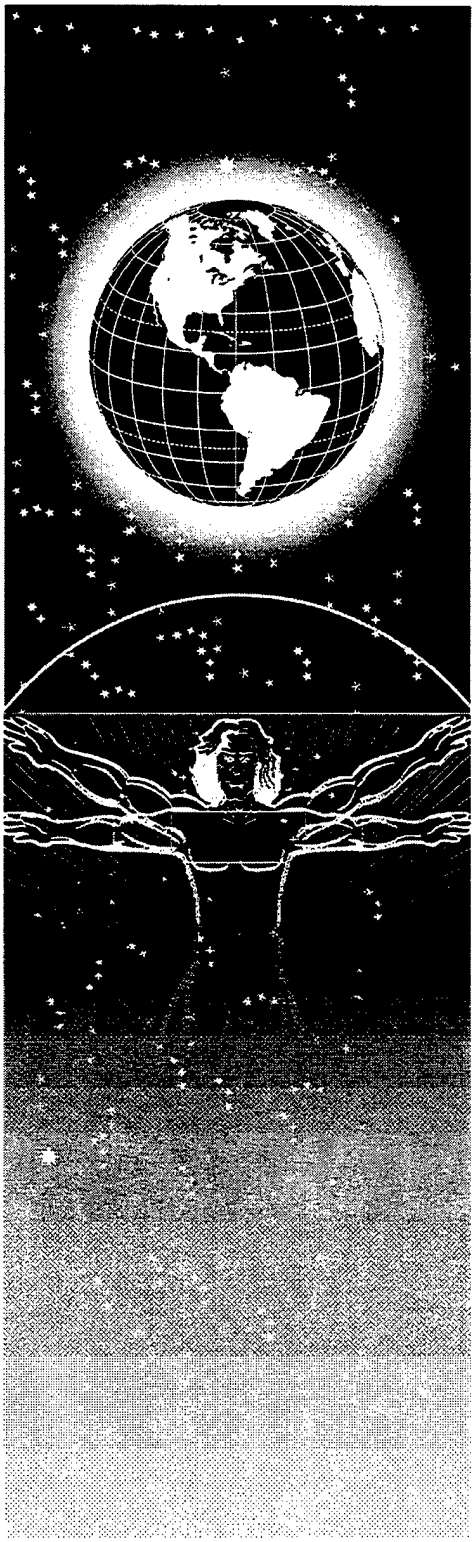
AFRL-HE-WP-TR-1998-0065



UNITED STATES AIR FORCE RESEARCH LABORATORY

VISUAL PERFORMANCE MODEL ANALYSIS OF HUMAN PERFORMANCE IN IR TARGET RECOGNITION

Frederick J. Smith
Allyn W. Dunstan

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

JUNE 1998

INTERIM REPORT FOR THE PERIOD APRIL 1997 TO JUNE 1998

Approved for public release; distribution is unlimited.

Human Effectiveness Directorate
Crew System Interface Division
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 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-1998-0065

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, including foreign nations.

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

FOR THE COMMANDER

HENDRICK W. RUCK, PhD
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 1998	Interim, April 1997 to June 1998

4. TITLE AND SUBTITLE
Visual Performance Model Analysis of Human Performance in IR Target Recognition

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

6. AUTHOR(S)
Frederick J. 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
AFRL-HE-WP-TR-1998-0065

11. SUPPLEMENTARY NOTES

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

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This technical report describes the application of a human-visual system simulation model, a computational vision model, as a method for prediction of operator target detection or recognition performance with infrared imaging sensors. The present report correlates the simulation model's predictions with laboratory data from human observers. The eventual goal of this work is to develop a methodology that will allow reliable predictions of imaging sensor system performance without the need for repeated laboratory testing with human observers.

In this report, the visual performance model (VPM) has been applied to a set of airborne, 1st generation FLIR imagery of mobile ground targets (including Scud-B mobile transporter-erector-launchers [TELs]). The detectability/recognizability metrics (d' values) obtained from VPM have been compared with similar laboratory data obtained using human operators. Good correlations between the VPM detectability/recognizability predictions and the human operator results were obtained during the present study (i.e., r values greater than .70). Future efforts are planned to examine VPM's utility for camouflaged targets and for 3rd generation FLIR imagery.

14. SUBJECT TERMS
Forward Looking Infrared, FLIR, Target Detection, Target Recognition, Sensor Imagery, Operator Performance, Visual Performance Model, VPM

15. NUMBER OF PAGES
24

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	UNLIMITED

This page intentionally left blank.

PREFACE

This technical report describes the application of a human-visual system simulation model, a computational vision model, for prediction of operator target detection or recognition performance with infrared imaging sensors. The project was completed for Air Force Research Laboratory Information Analysis and Exploitation Branch (AFRL/HECA) under Air Force Contract F41624-94-D-6000 for prime contractor, Logicon Technical Services, Inc. (LTSI). Work was accomplished under Work Unit Number 71841046, "Crew Systems for Information Warfare." Mr. Donald Monk was the Contract Monitor.

OptiMetrics, Inc. offers special thanks to Mr. Gilbert Kuperman, of AFRL/HECA, the Work Unit Manager, for his initial interest, and ongoing support and direction which made this effort possible. Acknowledgement also is offered to Mr. Robert L. Stewart of LTSI for management of the project, and to the following members of his staff: Mr. Joseph Riegler, who provided the source imagery as well as ongoing research direction; Dr. Judi See, who contributed the statistical analysis; and Ms. Elisabeth Fitzhugh, for technical editing services. OptiMetrics staff contributing to this effort were Mr. Frederick Smith, the Principal Investigator, assisted by Mr. George Lindquist, who assisted with VPM calibration, and by Mr. Allyn Dunstan, who contributed to the computer analysis.

With minor modifications, the computational vision model used in this effort was the National Automotive Center–Visual Performance Model.(NAC–VPM). This is a third-generation computational vision model developed by OptiMetrics Inc., Ann Arbor, Michigan, under contract to the US Army Tank and Automotive Research, Development, and Engineering Center (TARDEC), Warren, Michigan. The initiator of computational vision model development at TARDEC was Dr. Grant Gerhart. The current point of contact at TARDEC for NAC–VPM related efforts is Dr. Thomas Meitzler. The current version of NAC–VPM has been developed by OptiMetrics, Inc. under contract DAAE07-94-C-R111. The authors and sponsor of the present research wish to express their appreciation for the cooperation and insights of their TARDEC counterparts in the planning and execution of this effort.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1.0 INTRODUCTION

A key factor in designing the crew displays involved in target acquisition is quantifying the capabilities of the human visual system in the detection, recognition, and identification of targets. In the past, various semi-empirical rules (e.g., Johnson criteria) have been used to predict target detection as a function of the mean target-background contrast and the target size. While those rules have shown reasonable success when applied to unstructured targets, they are known to be deficient for many real-world cases. In particular, those rules break down when applied to structured (e.g., camouflaged) targets with a mean contrast near zero and when the targets must be detected in complex background scenes.

The eventual goal of the present project is to develop a method that emulates human visual performance and will therefore consistently predict human target detection/recognition/identification performance for real world conditions. A key factor is the ability to predict detection/recognition of camouflaged (including netted) targets as a function of range and electro-optical system performance for all engagement conditions.

The goals of the present project are to:

- Demonstrate the utility of an existing computational visual performance model (NAC-VPM[1], or simply, VPM) for predicting target detectability/recognizability (Witus, 1996).
- Compare and calibrate the VPM model against baseline first generation forward-looking infrared (FLIR) imagery of uncamouflaged targets.
- Install the VPM capability at Air Force Research Laboratory.

The Visual Performance Model (VPM) is described in more detail in the next section. The VPM development has been funded by the U.S. Army for evaluating the detectability of camouflaged and low signature ground targets. The VPM attempts to emulate some of the early-vision processing functions for the retina and the neural receptive fields. The early-vision modeling is based on a good deal of neurophysiological and psychophysical data on humans and other primates.

Unfortunately, less is known about mid- and late-stage neural vision processes and decision making. Hence a complete, "first-principles" model of human vision is not yet possible. To bridge that gap, the VPM follows the early-vision modeling with a statistical decision model which must be calibrated to measured task performance data. Once calibrated, the VPM is then able to predict the standard psychophysical measure of signal detectability, d', for additional target types and engagement situations (Green & Swets, 1966; MacMillan & Creelman, 1991).

This initial investigation applies the VPM to a series of FLIR images of large ground targets (e.g., Scud-B mobile transporter-erector-launchers [TELs]). The detectability/recognizability metrics from the VPM are then compared with laboratory detectability/recognizability data obtained when human operators viewed the FLIR images. The raw VPM detectability values were then correlated to the experimental results. A baseline calibration for FLIR data was also developed for the VPM.

The final task accomplished on this project was to provide a calibrated version of the VPM for use by Air Force Research Laboratory. Thus, a modified version of the VPM, incorporating parameters appropriate for FLIR imagery and utilizing the calibration results, has been developed and installed on government computers in the Human Effectiveness Directorate of Air Force Research Laboratory.

---

[1]For more detail, see *TARDEC National Automotive Center Visual Perception Model (NAC-VPM); Final Report: Analyst's Manual and User's Manual*, OMI-577, prepared by OptiMetrics, Inc., Sept 1996. Release point for NAC-VPM is Thomas Meitzler of the Tank and Automotive Research and Development Engineering Center (TARDEC), (810) 574-7530.

2.0 VISUAL PERFORMANCE MODEL

2.1 Overview

The Visual Performance Model is an evolution of a human visual performance model developed by OptiMetrics, Inc. The VPM development has largely been funded by the Army through TARDEC. The original version of the model, called the TARDEC Visual Model (TVM), was developed to address the detectability of military vehicles. The latest version, the NAC-VPM, is a modified version that has been used for various applications including examination of the conspicuity of automobiles. That version is essentially the one used in the current study; the NAC-VPM is documented in a recent OptiMetrics report (Witus, 1996).

The 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 the VPM is doing can be obtained by examining the Map files that define the various component processes.

The NAC-VPM is organized into an image-processing front-end model of spatio-temporal "early" visual processing, followed by a "back-end" statistical decision model. The front-end model computes the expected output response of individual neural receptive fields in a color/temporal/spatial multi-channel model of visual processing. The back-end model computes an aggregate measure of the perceptible visual information from a target, and predicts d', a standard psychophysical measure of signal detectability.

The 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 (AMSAA, 1996). 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.
>
> (AMSAA, 1996, p.7)

The second comparison is the result of a joint effort between TARDEC, General Motors Research Laboratory, and OptiMetrics. That effort compared and calibrated the VPM to experimental data on the ability of drivers to detect oncoming traffic in complex background scenes. Those results are reported in reference 1 and illustrated in Figures 1 and 2 included below. Figure 1 shows the overall scatter diagram comparing the experimental d' values with the calibrated d' predictions from the VPM. The analysis contains 736 cases and the correlation coefficient is 0.79. The root mean square (RMS) error between the predicted and experimental d' values is 0.56. Figure 2 shows the agreement between the d' predictions and the experimental results for the various conditions represented in this data set. It can be seen that the VPM results track the experimental results over the broad range of conditions represented.

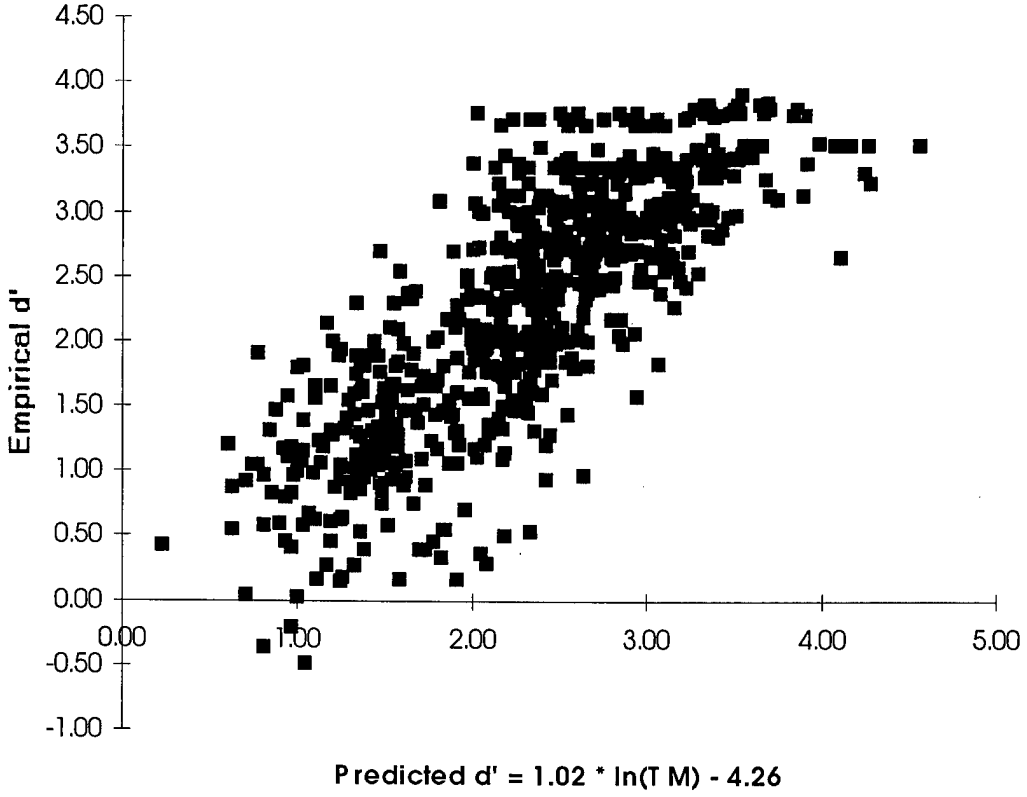


Figure 1. VPM predicted *d'* versus empirical *d'* for automobile conspicuity experiment.

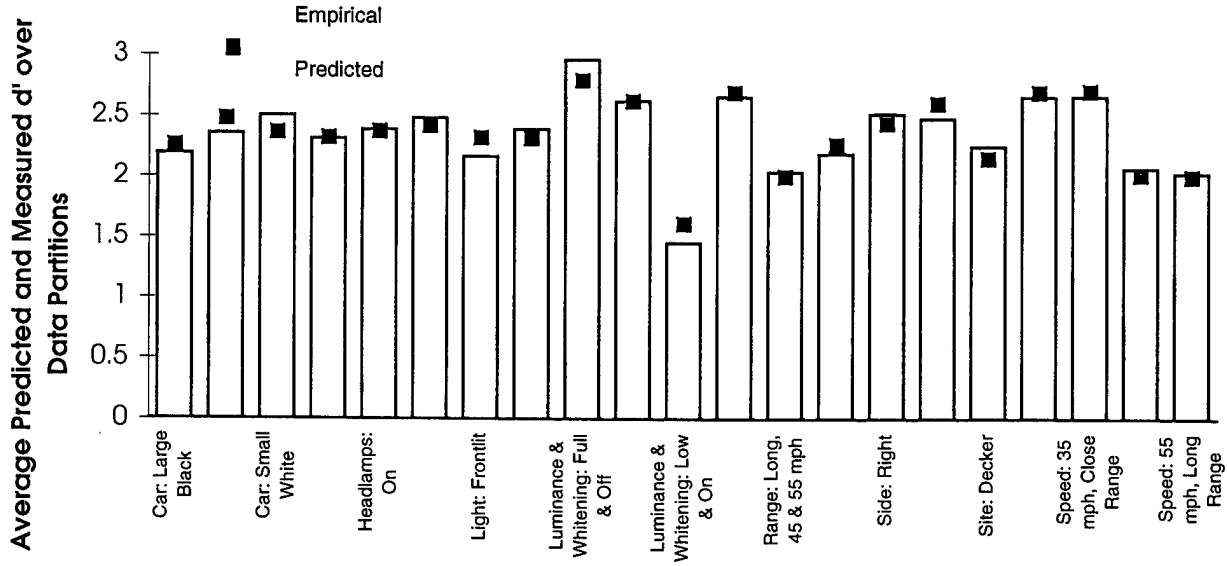


Figure 2. Predicted versus empirical *d'* for various conditions of the data from the automobile conspicuity experiments. (The upper bar of each pair represents the predicted *d'* value.)

2.2 VPM Early Vision Processing Model

VPM simulates the complex image processing chain representing early-vision. Since the 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 the VPM is shown in Figure 3.

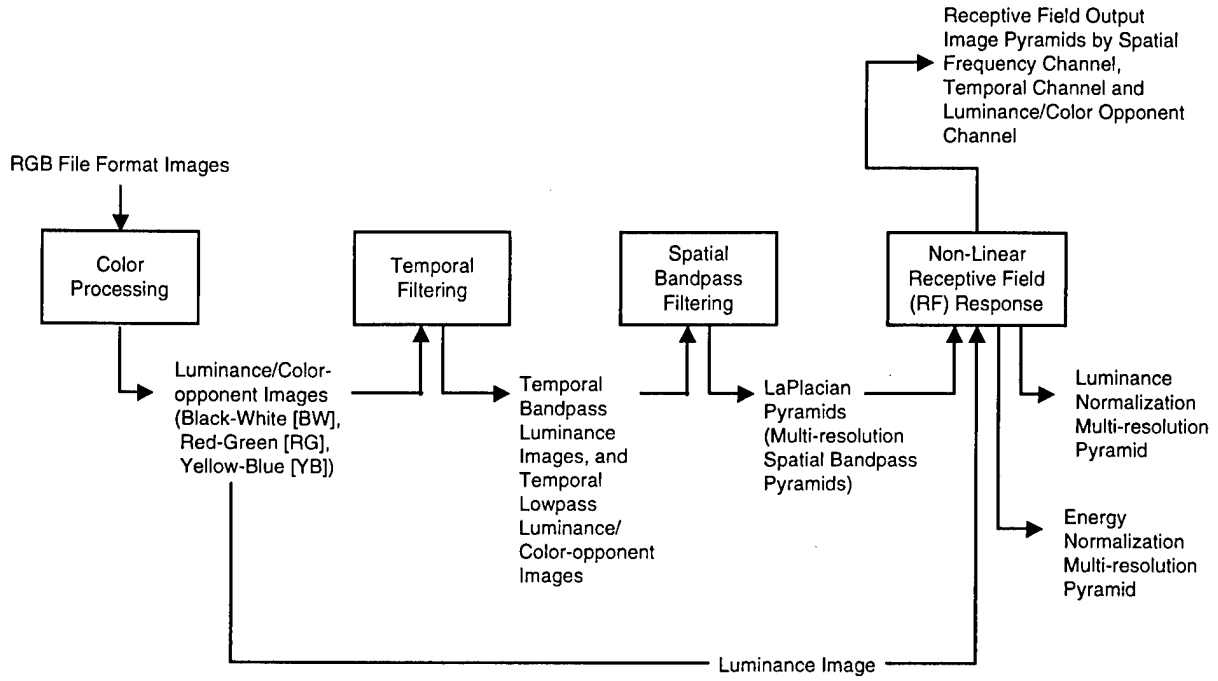


Figure 3. Early-vision processing modeling included in the VPM.

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 also not relevant and has not been used for this project since the targets in the FLIR imagery were stationary. The temporal processing is significant only if the target image on the display is moving at a rate greater than a few tenths degrees-per-second.

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 the VPM the target information is separated through the following process:

1. The user outlines the target with the EdTarget utility provided with the 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[2] multi-resolution images created by the VPM. Examples of these multi-resolution images are shown in Figure 4. 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.

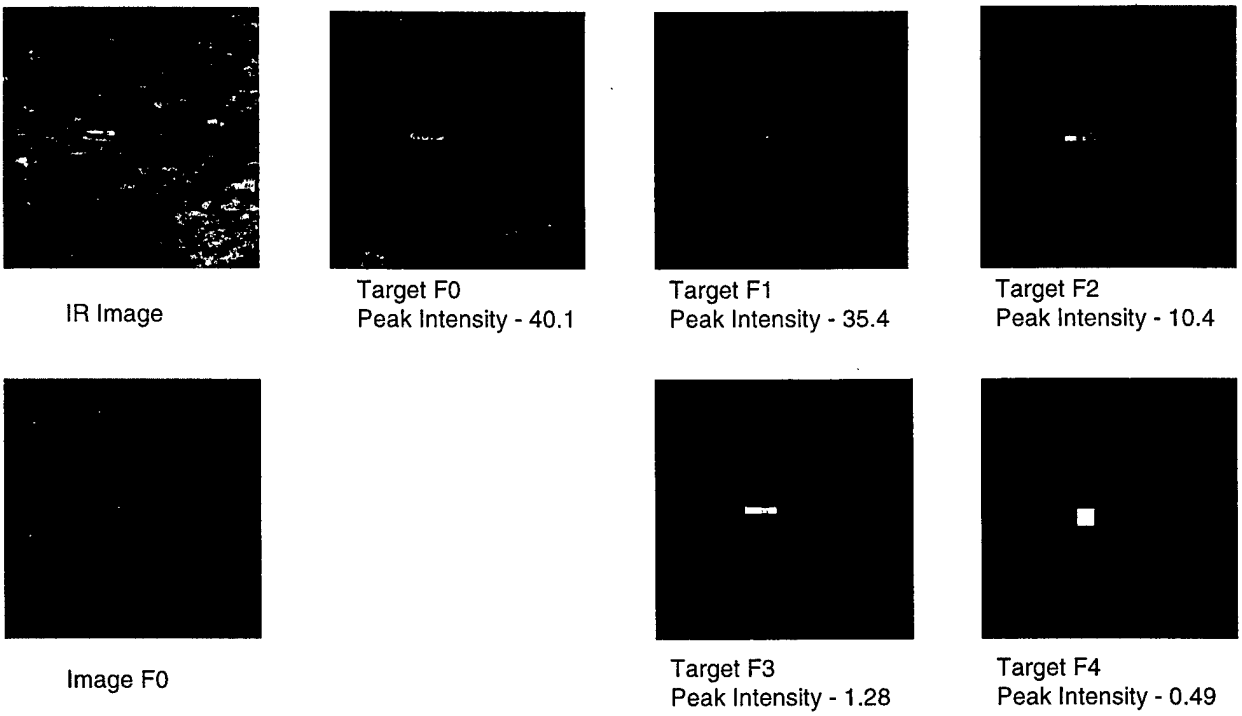


Figure 4. Illustration of RF Detectability multi-resolution images of an IR target created in the VPM.

To provide a single metric of target detectability, the VPM integrates the signal energy represented in the RF detectability multi-resolution images. The general approach used by the VPM to compute an aggregate target metric is shown in Figure 5. For the FLIR application discussed here, only the Temporal-Lowpass, Black-White multi-resolution images are included in the summation. The 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 detectivity values obtained from observer experiments (Witus, 1996; Cook, 1995). The determination of the slope, a, and intercept, b, relating the raw d' to detectability is the calibration process for the 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 the 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. Since the VPM has not been previously used for a complex *detection/recognition* task, as examined in this report, calibration coefficient values for that task are not available.

[2] The VPM output, TLBWTargetRFDetectability multi-resolution image is defined as Temporal-Lowpass, Black-White Target, Receptive Field Detectability multi-resolution image.

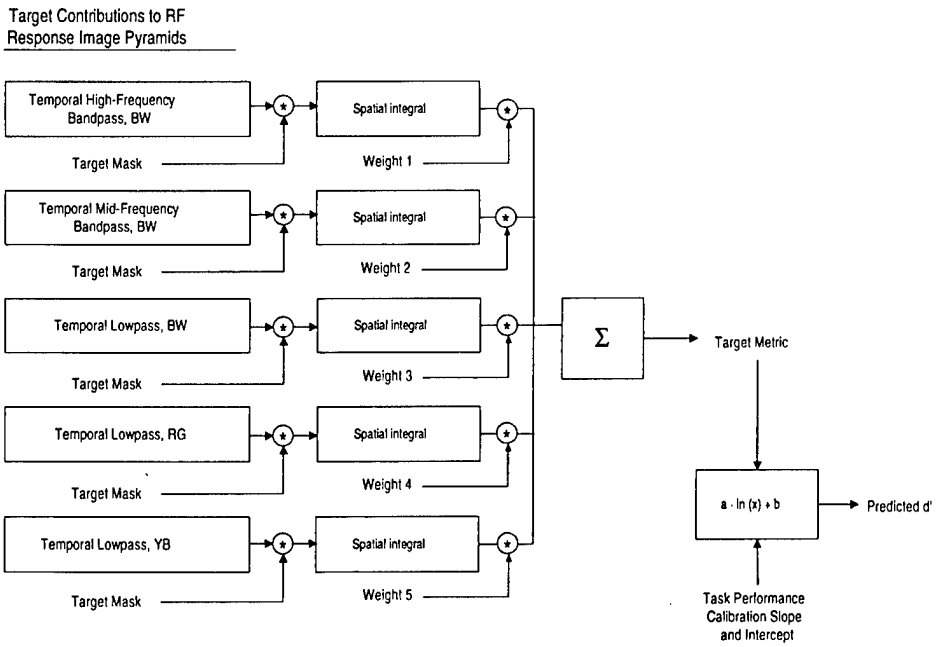


Figure 5. Target metric summation and predicted *d'*.

2.4 VPM Inputs and Outputs

The top-level inputs and outputs of the VPM are illustrated in Figure 6. 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 for the average observer. The lower box indicates input of the task performance calibration parameters described earlier. 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 the VPM's EdTarget utility. The 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. For the FLIR imagery used here, the observer was cued to the object of interest, hence, the angle from viewpoint for this case is zero.

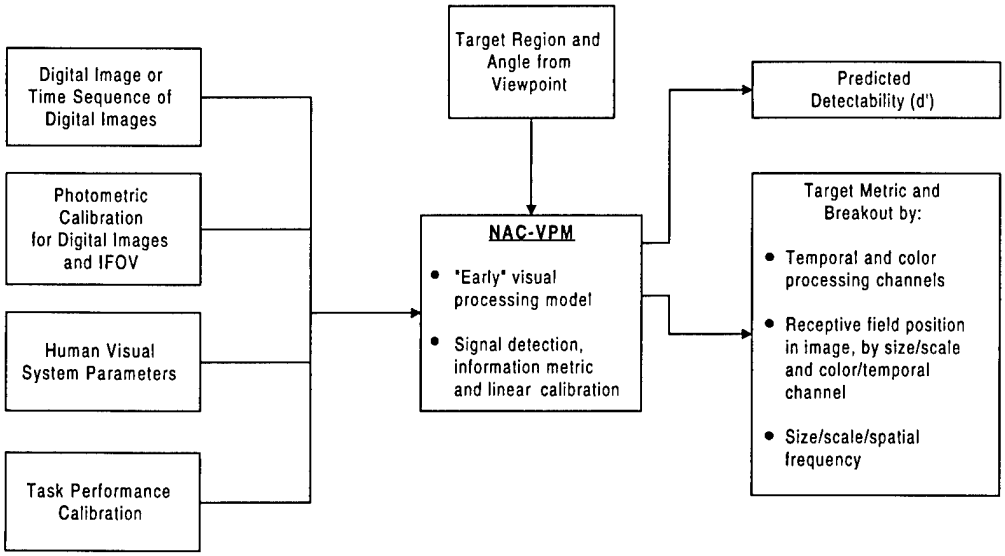


Figure 6. Top level illustration of VPM inputs and outputs.

The main output of the VPM is the predicted detectability, d', of the target. In addition to the d' value, the VPM also produces a number of intermediate outputs that are useful for verifying correct operation of the model and for analysis purposes. Those outputs include images of the "energy" contained in various color, temporal, and spatial channels, as illustrated in Figure 4. 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 7. 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 reference 1. The first VPM module is the Convert utility. That utility simply converts a Silicon Graphics RGB file format image into the internal data format (Channel Data Packet [CDP]format) used by the VPM. All of the subsequent images generated by the VPM are in this CDP format. The VPM provides the utilities View and ViewB to display CDP images on a Silicon Graphics Inc. (SGI) system. 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 the 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 TargetRFDetectability. Those outputs graphically illustrate how the various image components contribute to the overall target detectability.

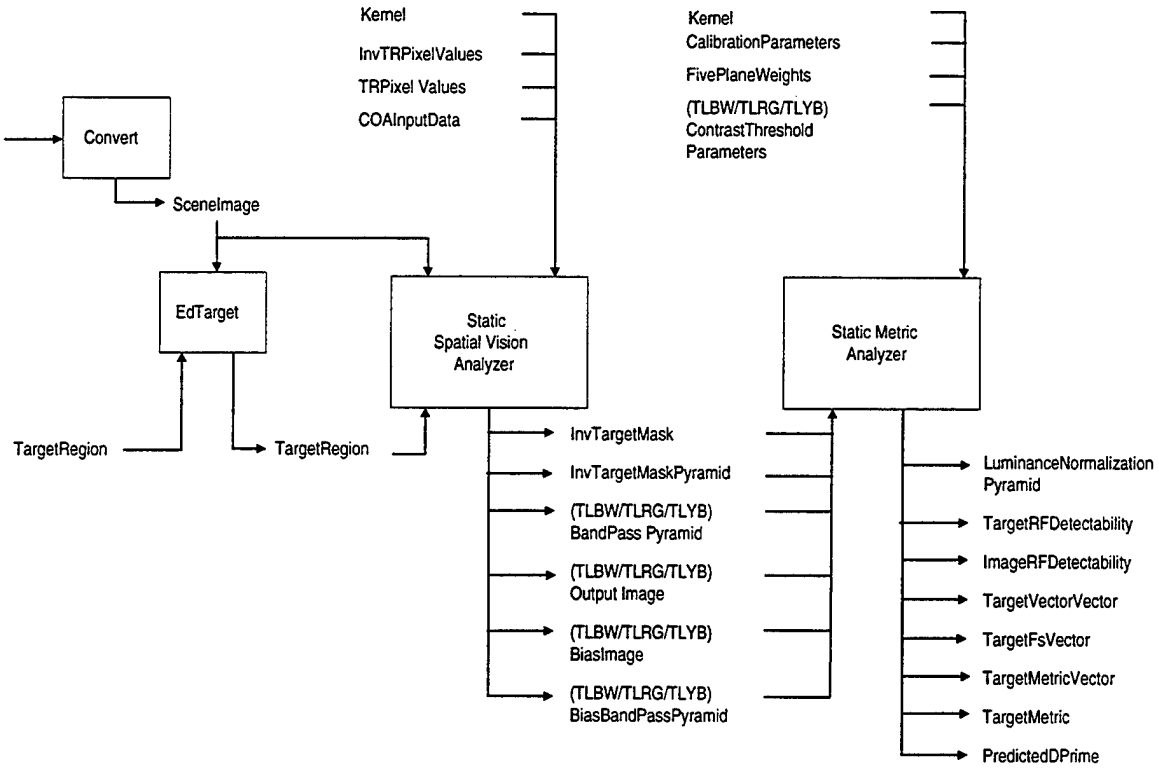


Figure 7. Input and output files used or created by the VPM.

3.0 AF HUMAN EFFECTIVENESS DIRECTORATE TARGET RECOGNITION EXPERIMENTS

Data provided by the AF Human Effectiveness Directorate were used for analysis and calibration with the VPM. The source of the actual FLIR imagery used in the study was from the Theater Missile Defense (TMD) Eagle Smart Sensor and Automatic Target Cueing (ATC; TESSA) program (Pryce, 1995). In that program, FLIR imagery was collected and recorded from a LANTIRN system for a total of nine missions, flown during both daylight and night, over several background sites. The data were collected from 10,000 feet altitude and from ranges of 18.5 km to overflight. The three TESSA targets were a Scud-B mobile missile transporter-erector-launcher (TEL), a ZiL 131 communications van, and a MAN 4-axle all-wheel drive truck equipped with an air compressor unit. The FLIR imagery was recorded on digital tape for later analysis.

The Air Force Research Laboratory's Crew Aiding and Information Warfare Laboratory (CIWAL) used the TESSA data to conduct a study of the operator's ability to perform unaided target detection/recognition with dynamic FLIR imagery (See, Riegler, Fitzhugh & Kuperman, 1996). The study used imagery taken from three background sites under daylight and nighttime conditions and for range bin distances of 4, 6, and 8 kilometers. 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 terrain feature, operators were asked to indicate target or non-target and rate their confidence in their decision.

The data collected from the observers were analyzed by Logicon Technical Services, Inc. (LTSI) personnel using the theory of signal detection. Hit and false alarm rates were used to derive 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'* detectivity index. These *d'* values were used for correlation/calibration of the VPM results.

3.1 Air Force Research Laboratory CIWAL Empirical Detectivity Results

The variables of the study were the site, time of day, and range bin. For analysis purposes, the variables were partitioned into 18 bins: site (open, sparse, treeline), time of day (day, night) and range bin (8 km, 6 km, 4km). The range bin partition includes data taken at various aspect angles and over one kilometer range variation, e.g., the 8 km bin included data from 8 to 7 km actual range. Table 1 summarizes the results of the CIWAL study (See et al., 1996)[3]. That table gives the detectivity values, *d'*, and the standard deviations of those values determined in the study.

Table 1. Empirical detectivity, *d'*

Mission	d'	d'	d'	SD	SD	SD	Site	Time
Bin (Km)	4	6	8	4	6	8		
4370-N	3.10	3.00	2.97	0.52	0.44	0.38	Open	Night
4686-D	2.54	2.13	1.99	0.80	0.65	0.76	Sparse	Day
5434-N	2.95	3.24	2.95	0.23	0.20	0.23	Sparse	Night
4685-D	3.08	2.49	1.85	0.18	0.41	0.87	Treeline	Day

As can be seen from the table, the *d'* values range from just under 2 to slightly greater than 3. The data represent the range from relatively easy to detect/recognize to very easy to detect/recognize targets. For example, a mean detectivity of 2.0 could correspond to a 70% probability of detection with 8% probability of false alarm. At the high end, a detectivity of 3.0 could correspond to a 94% probability of detection with 3% probability of false alarm. Thus, difficult to detect targets are not represented in this data set. The empirical *d'* values are plotted in

[3] The results shown are a subset of the data reported in See et al. (1996). These results were selected by Logicon as the most suitable for comparison with the VPM predictions.

Figure 8 below as a function of range bin. The daytime values have the expected decrease in d' with increasing range. It is interesting to note that the nighttime values show very little range dependence.

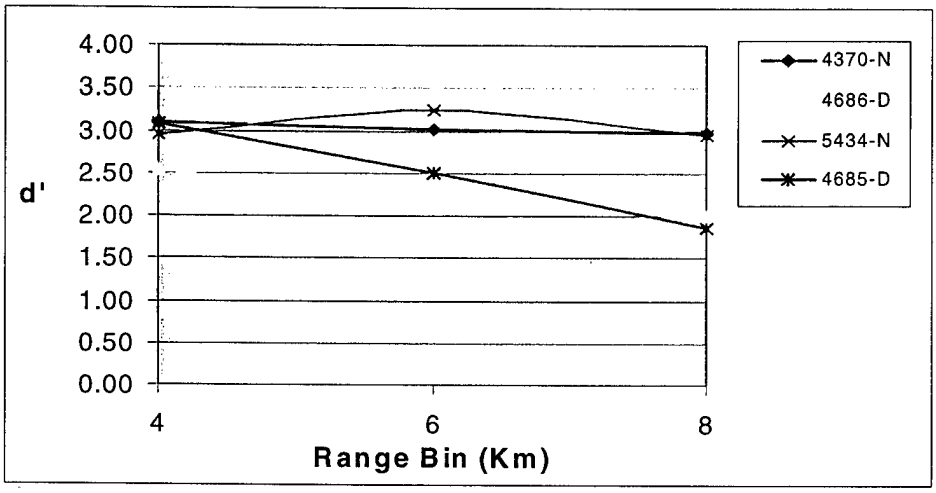


Figure 8. Empirical d' values plotted as a function of range bin.

4.0 VPM ANALYSIS PROCESS FOR TESSA IMAGERY

LTSI supplied selected frames of the TESSA imagery to OptiMetrics for analysis using the VPM. A total of 36 images were analyzed, representing 3 viewing angles for each of the 12 conditions listed in Table 1. Some additional images were also analyzed to determine background statistics and the sensitivities to target outline detail.

Before analysis with the VPM, the TESSA 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 2 below. Most of the conversion processes were performed using a shareware image processing tool for the SGI system called XV[4].

Table 2. Imagery preparation process

Step	Purpose	Comment
Convert from TIFF to RGB file format	VPM process starts with RGB file format image	Images in TIFF file format were supplied by LTSI
Resample image	Scale to provide square pixel of known size	This provides 1/32 degree square pixels given the viewing distance of 76 cm.
Crop image	VPM requires square image	A 256x256 image is extracted from the scene. When possible the image was centered on the TEL.
Save as RGB file format 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 in the defaults provided with the VPM model. Table 3 displays the modified parameters for the present analysis.

Table 3. Input parameters for the 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 0.417 0 0.126
SpatialWeights **	Weights	0,2,4,8,16,0,0,0

* 1 and 0 used for "raw *d'* calculation". For calibrated *d'* values as derived in this report, values 0.417 and 0.126 are used.
** New file created for this application

The actual VPM component models can be run sequentially as shown in Figure 7, or the RunImage script file can be used. To use RunImage, a separate directory is created for each image and the RGB file format image to

[4] XV © John Bradley. Can be obtained by anonymous FTP on *ftp.cis.upenn.edu*, in the directory *pub/xv*. John Bradley's official XV webpage address is http://www.trilon.com/xv/. More information on XV can be found at the Sun Microsystems Products & Solutions XV webpage, at http://www.sun.com/software/catlink/xv/xv.html.

be processed is stored in that file. If RunImage is called from that directory, the sequence shown in Figure 7 will be automatically run and the results saved in the image directory (the metrics may also be automatically printed). This saves a good deal of the manual file movement required if the procedures of reference 1 are followed.

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 we are not analyzing color imagery here, the color channel weighting is assumed to be zero; a weight of 1 is given 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 we have chosen to use a zero weight on 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 energy is visible; and 3) the results are reasonable.

4.1 Examples of VPM Processing

Figures 9, 10 and 11 show examples of the inputs and intermediate spatial metrics computed by the VPM. Figure 9 is an expanded representation of the FLIR imagery for a daytime and a nighttime mission. Both missions were run over the open site at 0 degree aspect. Figure 10 shows both the VPM energy metrics and the weighted energy metrics as a function of spatial channel number (inverse of frequency) for daytime mission. The results shown for this case are very much as might be expected. The shapes of the three target metric curves are similar, with the shortest range condition showing the greatest signal for all frequencies. The background signal is significantly lower than the target signal except on the highest frequency channel. As will be shown later, the d' predictions for this daytime data set correlate very well with the empirical results.

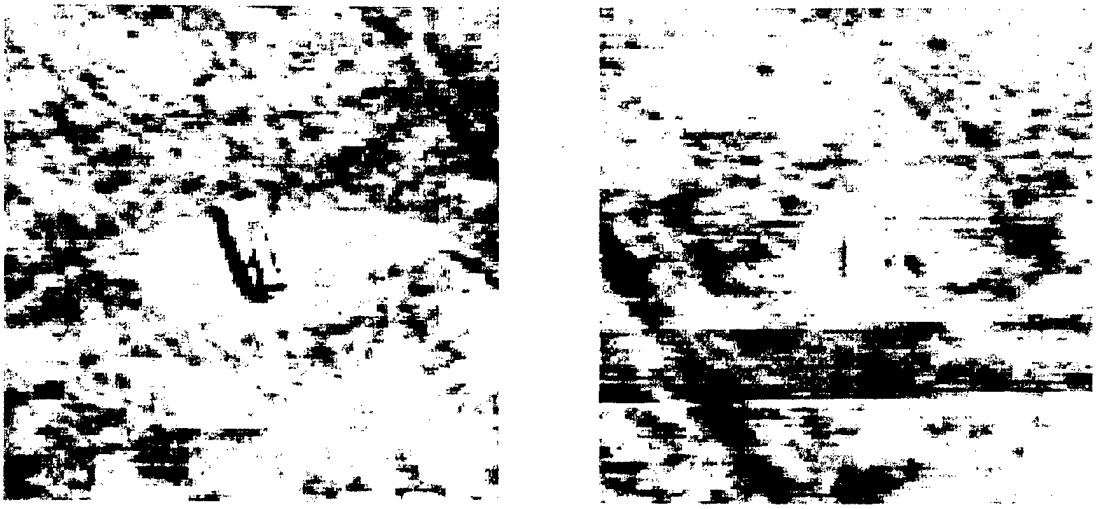


Figure 9. Examples of IR imagery from approximately 4 km range. The left image is a daytime image, while the right image is a nighttime image.

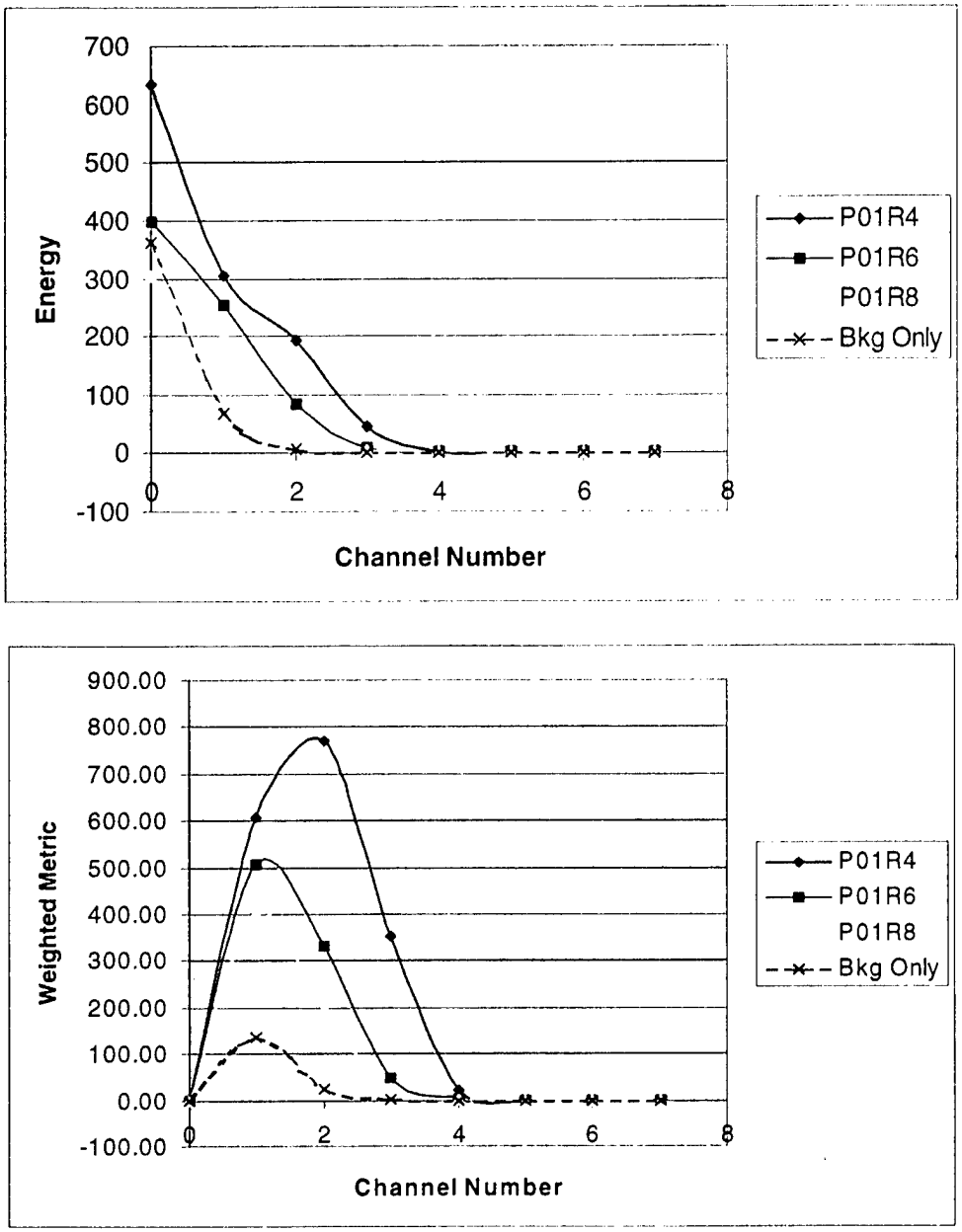


Figure 10. Intermediate Metric Results for a Daytime Mission Over the Open Site at Zero Degree Aspect

Figure 11 shows the intermediate metric results for the comparable nighttime case. There it can be seen that the metrics are lower for comparable channels as compared to the daytime mission. Also the separation between the curves for the different target ranges is not as clear. The smaller signal metrics imply that the detection/recognition detectivities for this case should be lower than seen in the daytime mission; however, that implication is not consistent with the empirical results. In fact, the nighttime mission results show the largest discrepancy between the VPM predictions and the empirical results.

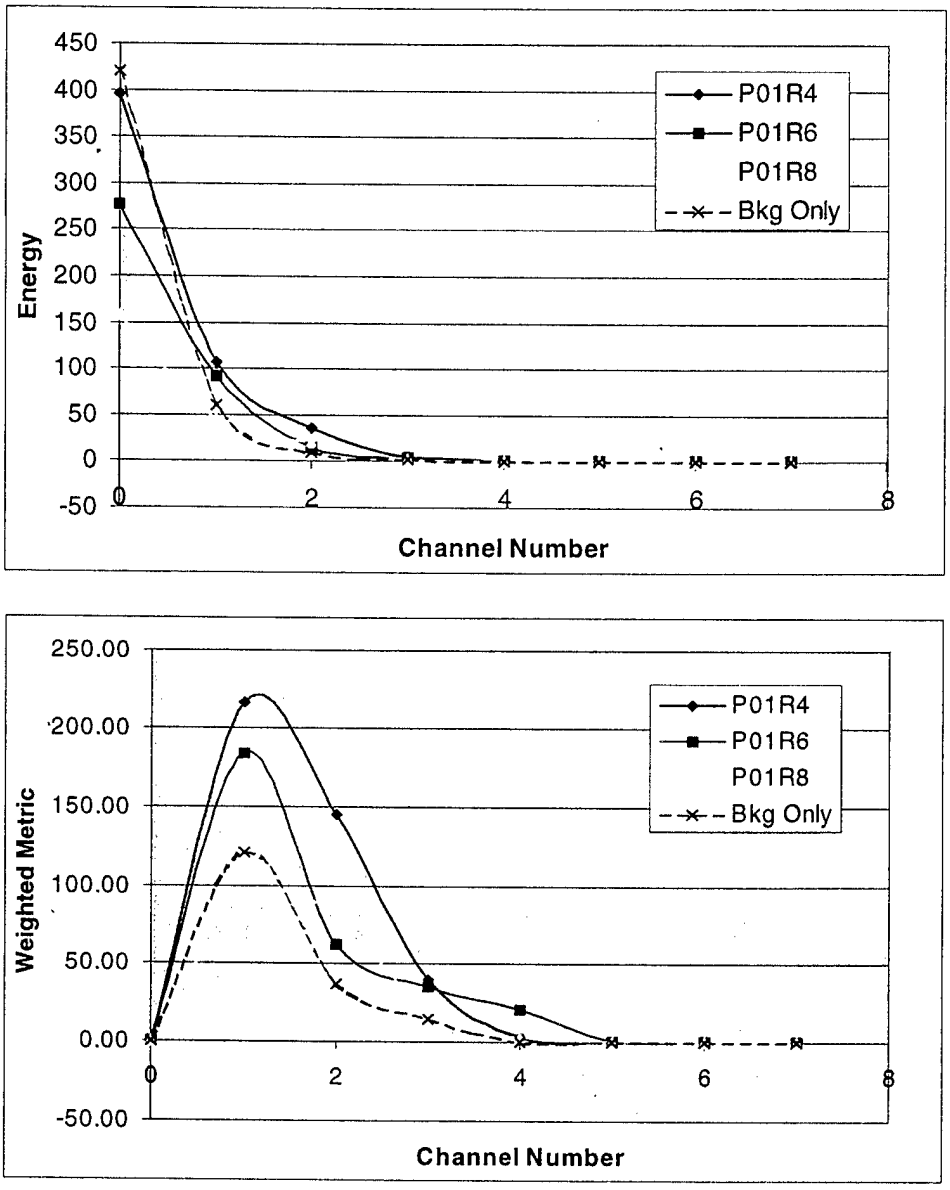


Figure 11. Intermediate Metric Results for a Nighttime Mission Over the Open Site at Zero Degree Aspect

5.0 PREDICTION OF DETECTIVITY USING VPM RESULTS

As described in section 2, the 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 with the VPM for the various missions and conditions are shown in Table 4. Also shown in the table are the standard deviations for each set.

Table 4. Raw *d'* values computed using the VPM

Range Bin	4686 Sparse Day	4685 Tree-line Day	4370 Open Night	5434 Sparse Night
4	6.08 (0.23)	6.16 (0.25)	6.06 (0.09)	7.74 (0.27)
6	5.30 (0.24)	5.31 (0.26)	5.58 (0.17)	7.52 (0.27)
8	3.99 (0.65)	4.76 (0.96)	5.19 (0.06)	7.24 (0.42)

The standard deviations in the table result from the fact that raw *d'* values were computed for imagery from three aspect angles for each mission and range. Those values were then averaged to give the mean raw *d'* values listed. A few observations can be made from the data. The first is that there are significant differences among the various cases. A second observation is the decrease in values with increasing range. Finally, it can be seen that there is a stronger range dependence for the daytime values as compared to the nighttime values.

To find a reasonable calibration for the VPM for this particular detection/recognition task, we explored various statistical correlations between the VPM raw *d'* values of Table 4 and the empirical *d'* values reproduced in Table 1. 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 were found to be the raw *d'* and time-of-day. The *t* ratios and *p* values for those two variables are $t(9) = 1.83$, $p < .10$ and $t(9) = 2.16$, $p < .06$, respectively. The regression equation found with those two independent 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* of .78. 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 good result. The above equation can be used to interpolate or predict *d'* values for conditions outside those where empirical operator results are available. Such predictions simply require a VPM analysis of the measured imagery and specification of a day or night condition.

In theory, if the operators always use 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.

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

The above regression equation gives an adjusted *r* of .77 and a standard deviation of 0.29. A comparison of the predictions to the empirical values is shown in Figure 12.

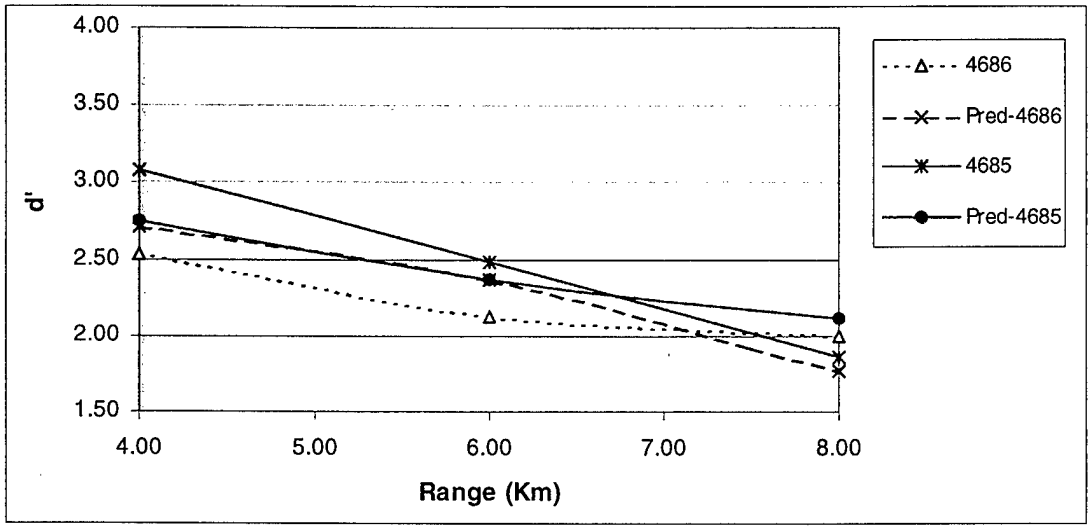


Figure 12. Comparison of Empirical and Predicted *d'* value for daytime cases.

To explore the discrepancies between the daytime and nighttime results, we used the daytime derived equation to predict the results for the nighttime cases. The comparison is plotted below. For the nighttime predictions the adjusted *r* is .44 and the standard deviation is 0.45. By inspection of the results plotted in Figure 13, it is clear that the predictions are systematically below the empirical results for the Mission 5434. The main disagreement is for Mission 4370. For that mission, the predicted detectivity is consistently 0.5 units below the empirical results. Or put into words, the model predicts that the TEL should be harder to detect/recognize than was inferred from the operators' performance. From inspection of the imagery from Mission 4370 it is easy to hypothesize why the observers did so well. For that mission there is almost no background clutter and the TEL is simply the largest, brightest object in the scene. Even though there is little of the target detail visible that would normally be required for a recognition task, the operator could simply guess that the largest, brightest target was the TEL and he/she would be correct a high percentage of the time.

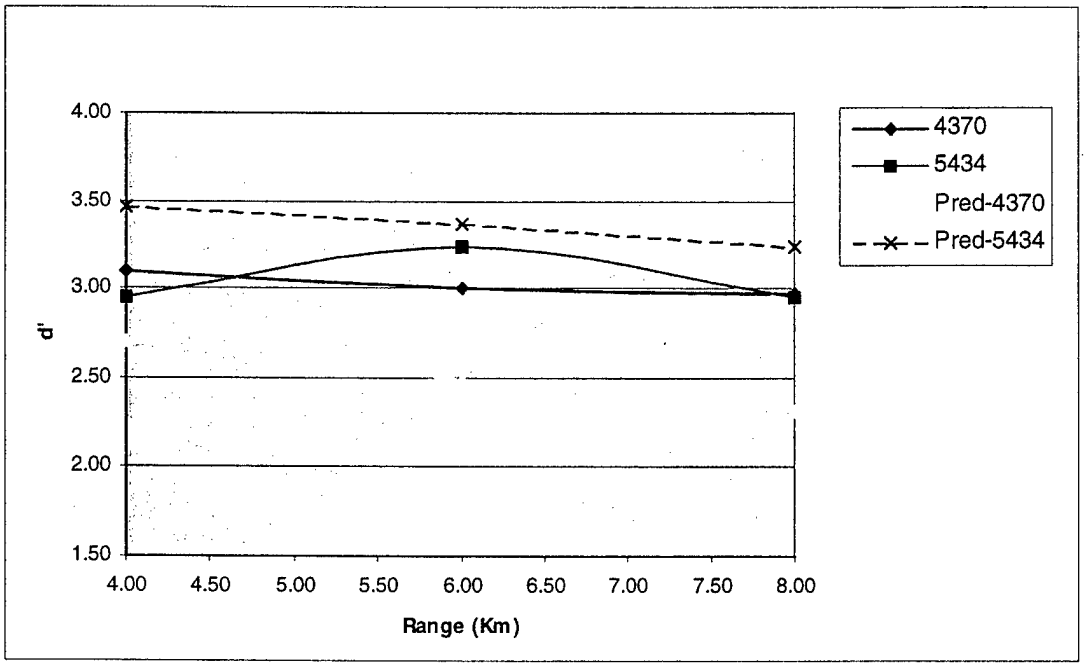


Figure 13. Comparison of Predicted and Empirical *d'* values for nighttime cases, using daytime regression equation.

6.0 SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

6.1 Summary and Conclusions

The Visual Performance Model was used to analyze LANTIRN FLIR imagery for a total of 36 scenario conditions. The VPM "raw d'" predictions were correlated to empirical d' results obtained from operator experiments. For the infrared imagery taken in daytime scenarios, the correlations with the empirical results were quite good, providing an adjusted r of .77. When the same regression equation is used to predict the nighttime d' values, reasonable agreement is found for one mission, but a systematic offset is seen for a second mission. From inspection of the imagery for the second nighttime mission, it appears that special conditions prevailed which allowed the observers to perform the recognition task with greater ease than would normally be possible. Based on these results, we conclude that the following equation can be used to conservatively predict detection/recognition task performance, as defined in the CIWAL experiment.

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

The key factor in this equation, of course, is the "raw d'" value. This can be obtained in at least two ways. The first way is to run the VPM on any infrared image representing the scenario condition of interest, as described in Section 2. The VPM will then compute the raw d' and the actual d', using the calibration parameter given above. A second method, that can be used for the various scenario conditions that have already been analyzed, is to interpolate or extrapolate from previously computed raw d' results. For the four missions analyzed in this report, Table 4 can be used to interpolate raw d' values as a function of range. The interpolated raw d' values would then be inserted into the above equation to give a d' prediction.

6.2 Recommendations

The present results provide a calibration of the VPM for a constrained target detection/recognition task under a limited set of conditions. Probably the greatest limitation of this data set, and hence the validity of the VPM calibration, is that for all of the conditions the target was highly detectable and easily recognizable. Hence the modeling has not been tested for difficult to detect/recognize targets, as could be seen for camouflaged or low-observable treated targets. Thus, the highest priority recommendation is to test/calibrate the model for such low detectable/recognizable targets.

Some other areas for possible near-term investigation are determining the model's utility for analysis of synthetic aperture radar (SAR) imagery and for representation of data fusion among multiple domains (i.e., FLIR, Visible, SAR). An early study indicated that the VPM may be applicable for representing SAR image analysis, but that study should be repeated with the current version of the VPM and a reasonably sized, empirically-calibrated data set. The fusion issue has two sides. The first side is to determine if the VPM detection/recognition measures in two or more domains can be combined to give a measure of human performance in data fusion. That is, can we predict the effectiveness of an operator who is given both FLIR and SAR data? If so, we might turn the question around. Given FLIR and SAR data, can the VPM methods be used to create a merged image that will allow an operator to make better decisions than he would given the two separate images?

An area for longer term investigation is the development of a more complete representation of the human visual processing and the cognitive decision chain. The VPM only claims to represent the early-vision processes. It is known that there are also "middle-vision" transformations as well as cognitive decision processing at work. Taking the next step, to model the middle-vision processes, appears feasible. A more flexible, although probably still heuristic, representation of the cognitive decision functions might also be fruitfully investigated.

REFERENCES

Army Materiel Systems Analysis Agency (AMSAA) (6 November 1996). *SUBJECT: Search and target acquisition model comparison* (Memorandum for Deputy Under Secretary of the Army for Operation Research and Director Assessment and Evaluation). Washington, DC: US Army.

Cook, T. H. (1995). *TARDEC visual model version 2.0, perception testing and calibration report.* (Report No. OMI-554). Ann Arbor, Mi: Optimetrics, Inc.

Green, D. M. & Swets, J. A. (1996). *Signal detection theory and psychophysics.* New York: Wiley & Sons.

Macmillan N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Pryce, J.D. (1995). *TESSA SAR/FLIR data collection report, volume 1.* (Report No. OMI-560). Ann Arbor, MI: OptiMetrics, Inc.

See, J., Riegler, J., Fitzhugh, E. & Kuperman, G. (1996). *Unaided target acquisition performance with first generation forward-looking infrared imagery: A signal detection theory analysis.* (Report No. AL/CF-TR-1996-94). Wright-Patterson AFB, OH: Armstrong Laboratory, Human Engineering Division, Crew Systems Directorate.

Witus, G. (Sept 1996). *TARDEC National Automotive Center Visual Perception Model (NAC-VPM) final report: Analyst's manual and user's manual.* (Report No. OMI-577). Ann Arbor, MI: OptiMetrics, Inc.

ACRONYM LIST

AMSAA	Army Materiel Systems Analysis Agency
ATC	Automatic Target Cueing
CDP	Channel Data Packet
CIWAL	Crew Aiding and Information Warfare Analysis Laboratory
FLIR	Forward looking infrared
LANTIRN	Low Altitude Navigation and Targeting Infrared for Night
LTSI	Logicon Technical Services, Inc.
NAC-VPM	National Automotive Center Visual Performance Model
RF	Receptive Field
RMS	Root Mean Square
SAR	Synthetic Aperture Radar
SGI	Silicon Graphics Inc.
TARDEC	Tank and Automotive Research and Development Engineering Center
TEL	Transporter-erector-launcher
TESSA	Theater Missile Defense Eagle Smart Sensor Automatic Target Cueing
TMD	Theater Missile Defense
TVM	TARDEC Visual Model
VPM	Visual Performance Model