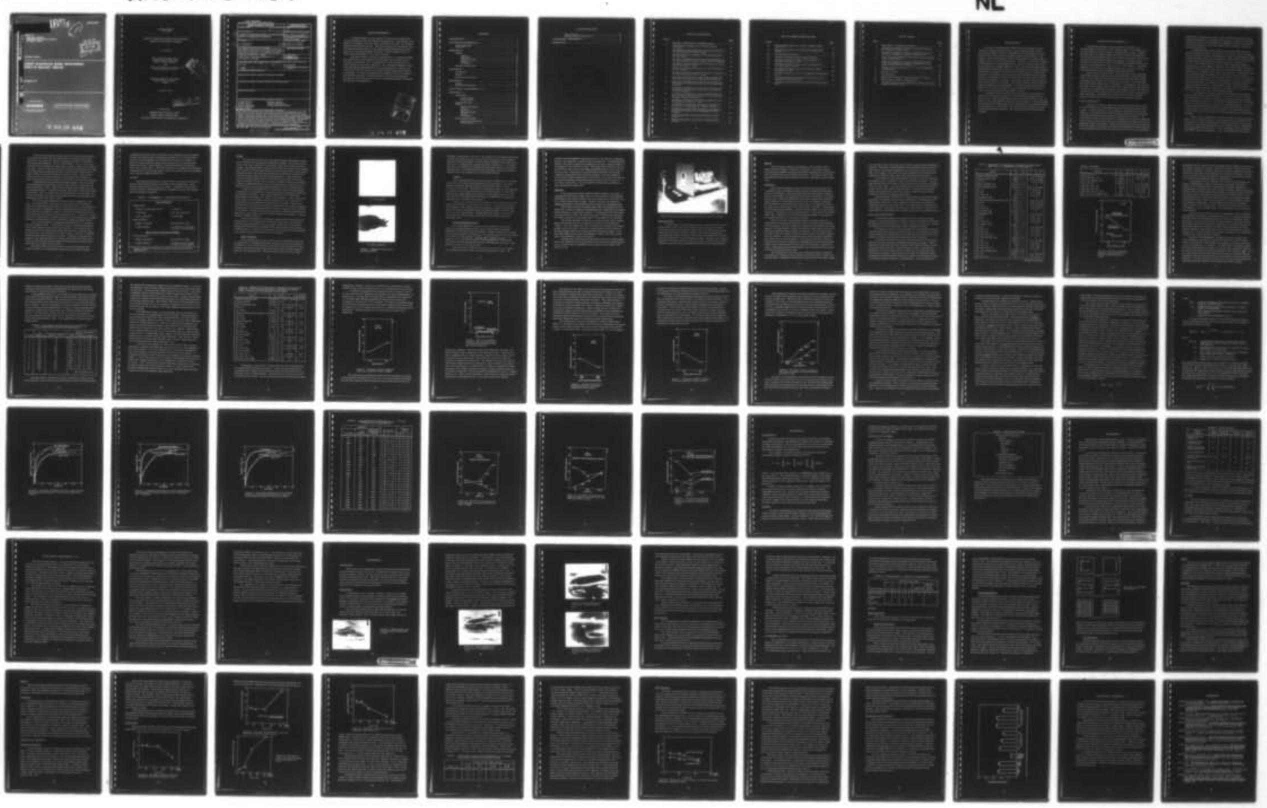
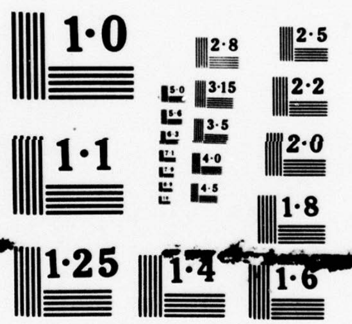
AD-A067 753   HUGHES AIRCRAFT CO CULVER CITY CALIF DISPLAY SYSTEMS LAB   F/G 17/8
TARGET ACQUISITION MODEL DEVELOPMENT: EFFECT OF REALISTIC TERRA--ETC(U)
DEC 76 L A SCANLAN   DAAG53-75-C-0266

UNCLASSIFIED   HAC-P76-484   NL

1 OF 1
ADA
067753



END
DATE
FILMED
6-79
DDC


1·0
2·8
2·5
3·15
2·2
3·5
1·1
2·0
4·0
4·5
1·8
1·25
1·4
1·6

LEVEL II

REPORT NO. P76-484
HAC REF. NO. 4977

Prepared for:

Night Vision Laboratory
United States Army Electronics Command
Fort Belvoir, Virginia

TECHNICAL REPORT

TARGET ACQUISITION MODEL DEVELOPMENT: EFFECT OF REALISTIC TERRAIN

DECEMBER 1976

AEROSPACE GROUPS
HUGHES
HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA

Approved for Public Release – Distribution Unlimited

Technical Report
P76-484

TARGET ACQUISITION MODEL DEVELOPMENT: EFFECT OF REALISTIC TERRAIN

L. A. Scanlan

Man-Machine Systems Section
Display Systems Laboratory
Radar Avionics Division

and

Advanced Programs Laboratory
Tactical Systems Division

HUGHES AIRCRAFT COMPANY
Culver City, California
90230

December 1976

Prepared for

VISIONIONICS TECHNICAL AREA
NIGHT VISION LABORATORY
UNITED STATES ARMY ELECTRONICS COMMAND

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9)
4. TITLE (and Subtitle) (6) TARGET ACQUISITION MODEL DEVELOPMENT: EFFECT OF REALISTIC TERRAIN.		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report.
		6. PERFORMING ORG. REPORT NUMBER D4977, P76-484
7. AUTHOR(s) (10) L. A. Scanlan		8. CONTRACT OR GRANT NUMBER(s) (15) DAAG53-75-C-Ø266
9. PERFORMING ORGANIZATION NAME AND ADDRESS Man-Machine Systems Section, Display Systems Lab., Radar Avionics Div., and Advanced Programs Lab., Tactical Systems Div., Hughes Aircraft Co., Centinela & Teale Sts., Culver City, CA 90230		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Night Vision Laboratory (DRSEL-NV-VI) Ft. Belvoir, VA 22060		12. REPORT DATE (11) December 1976
		13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (14) HAC-P76-484, HAC-REF-D4977 (12) 83p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Target detection
Target search
Target acquisition
Scene complexity
Field-of-view
Imaging sensors
Human performance
Target acquisition model

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The research obtained data on the effect of realistic background scene complexity on tactical vehicle target detection and the interaction of scene complexity with electro-optical sensor and display variables. Data were obtained which verified that performance prediction could be accomplished using selected metrics describing the characteristics of the background scene. The obtained data were used to investigate potential forms of a detection model which includes the influence of the scene. Finally the influence of field-of-view and method of sensor search on detection performance were assessed.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

LB

409 205

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ACKNOWLEDGEMENTS

Contractual support for the reported research was provided by the Visionics Technical Area of the Night Vision Laboratory, United States Army Electronics Command, Contract Number DAAG53-75-C-0266. Mr. Richard Flaherty monitored the contract. The technical effort was conducted by the Man-Machine Systems Section of the Display Systems Department and managed by the Advanced Programs Laboratory, Tactical Systems Division, Hughes Aircraft Company, Culver City, California.

The author wishes to express his appreciation to the following people who contributed to the program and without whom the research could not have been accomplished. Mr. William Stewart, Program Manager, provided excellent and consistant leadership. Mr. Michael Hershberger, Head of the Man-Machine Systems Section and Project Engineer, conducted the initial review of the literature and provided technical leadership throughout the program. Mr. David Craig prepared the imagery used in the research, Ms. Jane Herman and Mr. Clarence Dickson collected the data and performed much of the data analysis, and Mr. David Ketcham wrote the simulation software for Experiment 4.

CONTENTS

CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF TABLES

INTRODUCTION

A number of mathematical models of target acquisition have been developed to predict the performance of electro-optical sensor systems. None of these models, however, adequately treats the influence of the background scene on operator tactical target detection and recognition. Most assume a uniform background of some average luminance; a situation that is unlikely to occur in any realistic mission. The failure to include the influence of backgrounds of varying complexity may result in erroneous predictions of performance that are highly optimistic. Three of the four laboratory experiments reported examined the effect of the background scene on target detection performance. Experiment 1 determined the magnitude of the effect of background scene complexity and its interaction with selected sensor and display variables. Experiment 2 identified potential metrics which could be used to quantify the complexity of a scene. Experiment 3 considered the adequacy of these various metrics as predictors of the time required to detect a tactical target in realistic terrain backgrounds. The data from these three experiments were also used to examine the type of model required to predict the influence of the scene and the target within the scene.

A complete model of target detection performance will also have to account for the sensor search process required to get the target into the field-of-view. Present models only attempt to predict performance when the target is known to be within the field-of-view of the sensor. In actual missions the uncertainty in target location may be sufficiently great to require considerable sensor search. If this is the case, issues of sensor field-of-view and search method become important. The fourth experiment examined these variables.

BACKGROUND EXPERIMENTS 1 - 3

Mathematical models of target acquisition performance, as representations of the world in a more abstract space, can be important and powerful tools for use in the design of imaging sensor systems. Models formally define the relations among objects in the world in mathematical terms and the resulting abstraction can be exercised using the rules of logic and mathematics to test the implied consequences. By using a model, a designer can determine the impact of a contemplated design without the time and cost of actually building a system for test. Alternatively, a strategist might exercise a model of an existing system under a variety of tactical situations to assess the best method of deploying the system.

Whatever the use of the model, correct decisions can be made only if the model accurately predicts real-world performance. Accurate predictions, in turn, require that all parameters that significantly affect system performance be included in the formulation of the model. This means that the characteristics of the sensor, the display, the atmosphere, the observer, the target, the background scene, and the inter-relations among these parameters need to be considered for inclusion in a complete model. The research reported in the present paper directs itself toward the issues of background scene complexity, its quantification, its influence on target search and detection, and its impact on the formulation of a mathematical model.

SCENE COMPLEXITY

An observer's ability to detect and recognize tactical targets located in realistic terrain may be strongly influenced by the characteristics of that terrain. A vehicle located in a broad expanse of open desert sand or in the middle of a large meadow will be easily detected provided that the target has a contrast ratio above the psychophysical threshold of the observer. The same vehicle located in an area of heavy vegetation will be much more difficult to detect because of the large number of competing objects with characteristics apparently similar to the target. Aspects of the background scene and the location of the target in the scene can be expected to influence the method used and the time required to search for the target as well as the time and probability of detecting and recognizing the target. Thus, any effort to predict the target detection performance of an imaging sensor system, whether by means of a mathematical model, or some other more qualitative method, must consider the potential contribution of the background scene on observer search and detection performance.

A number of mathematical models of target acquisition have been developed to predict operator performance as a function of electro-optical sensor and display variables. Stathacopoulos, Gilmore, and Rohringer (1976) reviewed fourteen of these and noted that in all of the models "the descriptions of targets and backgrounds include only the most basic parameters..." (p. 92). As the authors note, none of the subtle target optical signature characteristics such as shadows, highlights, and viewing angle are modeled. Instead, the models consider only target-to-background contrast, based on some measure of radiated target energy, and target size. Those that include any background complexity inputs at all, provide only for the relatively nebulus specification of number of confusing objects. Further, these models base their predictions of the effect of background on the research of Boynton and Bush (1955; 1957) – data obtained using arbitrary abstract geometric forms as stimulus material. Although there is little reason to doubt the findings of Boynton and Bush, there is, at the present time, no way of relating their data to the characteristics of real-world scenes.

The rudimentary treatment of background scene effects in present models results from a number of factors. First, little data exist on which an understanding of the influence of background on target search, detection, and recognition can be based. In the absence of appropriate data, the development of a model represents a formidable, if not impossible, task. Data are needed which will provide direction for the quantitative inclusion of scene characteristics in a model of target acquisition, including consideration of the form of the model, to determine if present approaches have the necessary flexibility.

Another factor inhibiting the inclusion of background effects in present models is the resultant increase in computational effort, particularly if the characteristics of the background interact with other factors such as sensor parameters requiring the calculation of many additional terms. Although the increased computation is feasible, it will be reasonable only if the accuracy of prediction improves significantly. Thus, the magnitude of the effect of background scene needs to be assessed to determine if the increase in model complexity is justified.

The research reported here provides initial data related to each of the needs outlined above. The magnitude of the background scene complexity effect and its interaction with selected sensor and display parameters were assessed using photographs of actual terrain and realistic targets. The feasibility of using quantitative metrics as descriptors of scene complexity and predictors of performance was examined and an initial consideration of the required model form was made using the data obtained.

APPROACH

Three experiments were conducted to examine the effect of background complexity on target search and detection performance. In all experiments the stimulus materials were forward oblique aerial photographs of actual terrain with tactical vehicle targets photographically embedded. Although, any complete model of electro-optical sensor system requires, ultimately, the inclusion of the infrared spectrum as well as the visible, it is expected that the majority of the operator's perceptual characteristics will remain unchanged. Because of this expectation, the initial work undertaken here used imagery representative of that obtained from a high quality television sensor.

The first experiment examined the magnitude of the effect of scene complexity and its potential interaction with various sensor and display parameters. The factorial combination of two levels of background scene complexity, two levels of target-to-background contrast, two levels of display resolution, two levels of target subtense, and three target vehicle types was examined. In addition, three target types, two resolution, and two target subtenses were examined with a uniform background to allow a comparison between this abstract situation and the more realistic background situation.

Experiments 2 and 3 examined the use of quantitative measures or metrics of scene complexity as a means of including background effects in a target acquisition model. Experiment 2 asked subjects to describe those target and scene characteristics they felt would strongly affect detection of the vehicle. Experiment 3 examined the predictive capability of scene metrics using a regression approach. The metrics considered in this study were those suggested by the results of the second experiment.

EXPERIMENT 1

The first experiment examined the effect of high and low background complexity on the time required to detect a tactical vehicle in a realistic scene. Additionally target-to-background contrast, display resolution and angular subtense of the target were manipulated because of the probability that they might interact with scene complexity. A final variable, target type, was also examined to determine if additional consideration of target effects was required. A uniform gray background with zero complexity was also examined in combination with display resolution, target subtense and target type.

On a static monochrome display with a stationary target, detection and recognition can be accomplished along only two dimensions: luminance and spatial. In the case of a target located in a uniform background, the luminance factors will predominate because there is no need to discriminate shape characteristics of the target. The object of the approximate size with a luminance different than the background must be the target. However, a target located in a real-world background cannot be detected as easily on the basis of luminance unless the target-to-background contrast is very high resulting in a target that is much brighter or darker than any other object in the field-of-view. With less extreme values of contrast where other objects are of equal luminance, target detection must rely, to a greater extent, on the spatial characteristic or shape of the target. Detection with complex background scenes thus includes recognition to the point of classifying the object as a tentative member of the target set.

With real-world scenes of low complexity, only a few objects of similar size and luminance to the target will be present and target-to-background contrast may remain a significant cue to the target. However, as the scene becomes more complex a great many objects can have luminances similar to the target, thus making the task more one of shape discrimination. Because of the shift toward a form or shape discrimination task it may be that target-to-background contrast makes little difference with high complexity backgrounds. This potential for an interaction between target-to-background contrast and background complexity requires examination.

The literature discussing contrast generally indicates that detection performance increases with increasing contrast with the largest changes occurring at low contrasts. Because of differing methods of calculating contrast (see Statacopoulos, Gilmore and Rohringer, 1976, p. 119-122 for a partial list) it is difficult to compare actual contrast values across studies. It is clear, however, that contrast can have a large effect whether the display is of abstract figures (Boynton and Bush, 1955; 1956; 1957; Peterson and Dugas, 1972) or real scenes (Bergert and Fowler, 1970). Because of the known effect of contrast and because of the probability of an interaction between contrast and scene complexity, two levels of contrast were examined.

Psychophysical experiments examining threshold contrast detection have demonstrated an interaction between visual subtense of the target and contrast (Blackwell, 1946). Small subtense targets require higher contrast for threshold detection. Although the present study does not deal with threshold detection, there is evidence that an interaction between visual angle subtended by the target and target-to-background contrast will nonetheless exist (Boynton and Bush, 1955; Craig, 1974).

The visual angle subtended by the target can be manipulated in two ways. The field-of-view of the sensor can be reduced and the resulting increase in magnification will produce a larger target size. This method has the disadvantage of changing the extent of the background which indirectly could change its complexity. Further, the number of resolution lines across the target would increase with the smaller field-of-view. As both scene complexity and resolution were variables of interest in the present study, changing subtense by changing field-of-view would have created confounded results.

The second method for manipulating target subtense is to change the subject to display distance. This method causes the visual angle subtended by the entire display to change with target subtense. It was not possible to change the size of the target alone because that would have made the target an inappropriate size compared to the other objects in the scene: a difficulty not present with abstract stimulus materials. Although it is possible that changes in the angle subtended by the display could affect the search portion of the task, this potential problem appeared less serious and subtense was manipulated by changing viewing distance.

Resolution, and more particularly the number of resolution lines across the target, is a variable that is known to have an effect on target recognition performance and is included as a key parameter in some models. In those cases where target to background contrast is low, and target detection must be accomplished by discriminating the shape of the target, the number of lines across the target can have a major effect on performance. With too few lines it may be difficult to distinguish between a target and a tree or other clutter object. With high contrast or low complexity backgrounds resolution may be of less importance because the detection can be accomplished on the basis of brightness differences.

METHOD

The experiment examined the effects of five variables on an operators ability to detect tactical targets in realistic terrain backgrounds. These variables and their levels are given in Table 1. In addition, special cases with non-realistic, plain backgrounds were sampled to provide an additional level of background complexity.

TABLE 1. EXPERIMENT 1: VARIABLES AND LEVELS

Main Experiment	
Within Subject	
Target-to-background contrast	0.7 and 2.0
Target type	APC, tank, and truck
Scene complexity*	Low and high
Between Subjects	
Display resolution	240 and 480 TV lines
Target subtense	0.178 and 0.356 arc grads (9.6 and 19.2 arc minutes)
Special Uniform Background Condition	
Between Subjects	
Display resolution	240 and 480 TV lines
Target subtense	0.178 and 0.356 arc grads (9.6 and 19.2 arc minutes)

*Based on subjective ranking by six judges viewing scenes without target present.

Stimuli

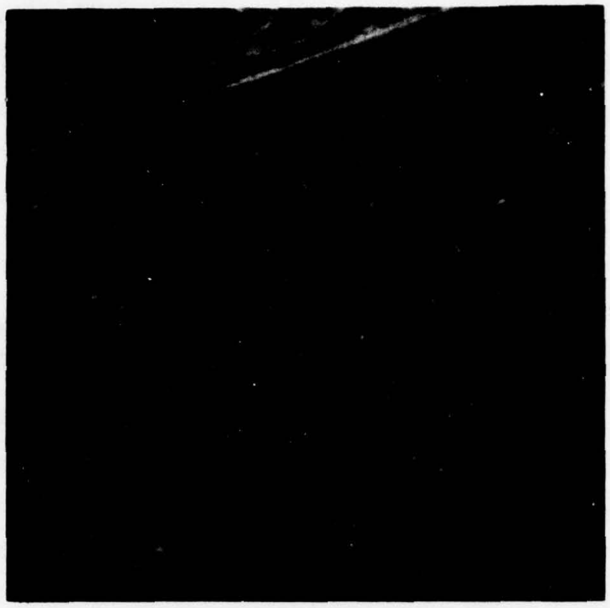
The image scenes used were low altitude oblique photographs of rural New York State with target vehicles optically embedded. The characteristics of the scenes were reasonably representative of the terrain found in Central Europe. The original films, from which the background scenes were selected, were 44.44 by 44.44 grad (40 x 40 degree) forward view aerial reconnaissance photographs taken from approximately 910 meters (3000 feet) altitude with a camera depression angle of 22.22 grads (20 degrees). These latter two values are somewhat greater than is normally characteristic of Army helicopter operations, however, the use of lower altitudes would have introduced a number of additional factors, such as target masking, and excessive scale distortion not germaine to this initial examination of scene characteristics.

Targets were embedded into the background scene by superimposing a transparency of the target on a transparency of the background and optically processing the composite. The target transparencies were obtained by photographing scale models on a featureless background. Using a large print of the target, an artist added a shadow appropriate for the sun angle in the background scene into which the target was to be embedded. The internal contrast of the target was also artificially enhanced so that the final composite image would more nearly approximate the internal modulation characteristics of real targets. Without the artificial contrast the embedded target generally appeared as a dark shape, devoid of internal brightness differences. Typical images are shown in Figure 1.

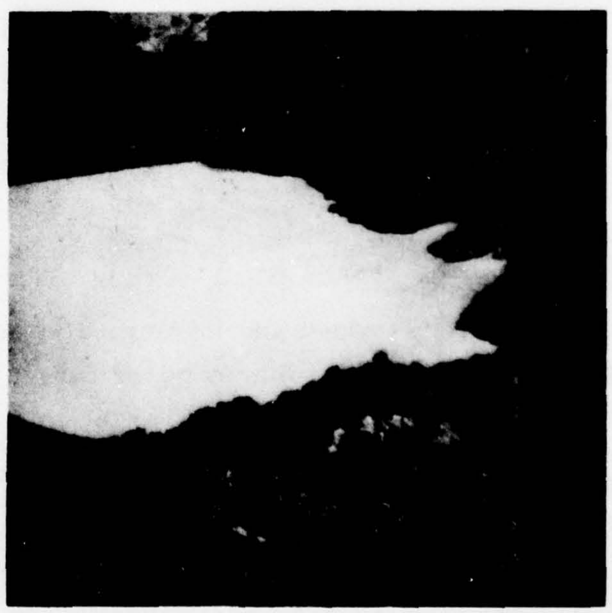
Twelve test images each containing a single tactical target and representing an 8.89 by 8.89 grad (8 by 8 degrees) field-of-view were prepared. Experimental manipulation of scene complexity target type, and target-to-background contrast was accomplished in the construction of these images as described below. In addition, three special images, depicting a single tactical target against a plain gray background, were prepared.

Scene Complexity

Initial quantification of scene complexity was accomplished by having six judges rank 75 candidate photographs according to perceived complexity. "Complexity" was not defined for the raters but left to their subjective interpretation. Some of the backgrounds depicted open fields generally lacking in natural or man-made features. Others contained features such as roads,



a. Low complexity



b. High complexity

Figure 1. Examples of high and low complexity scenes.

buildings, low shrubbery, and trees. Six complex and six simple backgrounds were selected based on mean rank and maximum agreement among judges (minimum standard deviation). The six judges evidenced considerable agreement in their rankings of the simplest and most complex scenes. Much less agreement was found in the rankings of the moderately simple to moderately complex scenes. In addition to the twelve test images, two simple and one complex scene were selected for use as training material.

Targets

Three vehicle types were used in the study: an M-60 tank, a two-and-one-half ton truck, and an armored personnel carrier (APC). The image height of all targets was held constant at two percent of the 8.89 grad (8 degree) field-of-view. This height corresponds to a range of approximately 1.2 kilometers (0.75 mile) and was selected based on pretest results which indicated this target height was necessary to avoid conditions where no subject could detect the target which would make it impossible to ascertain the influence of the variables being examined.

Targets were positioned in the background scene so that their size and location was appropriate to other terrain features. To avoid having targets always appear at the vertical center of the display, the apparent depression angle of the sensor was varied to allow the target to appear anywhere in the center two-thirds of the image. Targets appeared as direct side views or quartering front views with an aspect angle consistent with the apparent depression angle.

Target-Background Contrast

The contrast between the target and its immediate surround was varied by changing the density of the superimposed target image while holding the background film density constant. Target-to-background contrast was calculated using the formula $(B_{max} - B_{min})/(B_{min})$ or because the targets were darker than the surround, $(B_{background} - B_{target})/(B_{target})$. Two contrast values were selected based on the albedos given by Buddenhagen and Wolpin (1961) for dry sand and lush grass with the assumption that target albedo was similar to coniferous forest.

An average target brightness measurement was made on the display with a photometer aperture which covered 80 percent of the target. The average surround brightness measurement used in the contrast computation, however, was actually an average of four measurements. A small photometer aperture was used to measure the brightness at the four cardinal positions of an imaginary circle around the target. The background brightness measurements taken in this way were not highly repeatable and fail to account for the total surround. The problem was particularly troublesome with the high complexity backgrounds where the close proximity of varying brightness terrain features made measurement quite difficult. As a result, the high and low contrast conditions represented ranges of contrast with averages of 2.0 and 0.7 respectively. The contrast of the three targets located in the plain background averaged 0.23.

Apparatus

A rear-projection display apparatus shown in Figure 2 was used. It consisted of a high intensity light source for illuminating the glass mounted film images and an optical system to focus the image on the 12.7 by 12.7 cm (5 by 5 inches) ground glass display. Display resolution was adjusted by changing the position of the projection lens to defocus the displayed image.

The amount of defocus was determined using a Buckbee-Mears resolution chart. Subject to display viewing distance was maintained at a constant value by means of a subject head restraint. Viewing distances of 0.46 meter (18 inches) and 0.91 meter (36 inches) were used to obtain target subtenses of 0.356 arc grad (19.2 arc minutes) and 0.178 arc grad (9.6 arc minutes).

The subject indicated the position of a detected target using a wooden pointer and the experimenter verified the correctness of the designation visually. The time required to make a detection was recorded by the experimenter using a stop watch. Between trials the display could be blanked by placing an opaque cloth between the projection lens and the display. As an aid in identifying salient target characteristics, models of the three target types were positioned in front of the subject.

Ambient room illumination was provided by fluorescent ceiling lights. These were adjusted to the average display luminance of 34.3 candle/(meter)^2 (10 fL) so that luminance adaptation was not required of the subject at the onset of a trial.

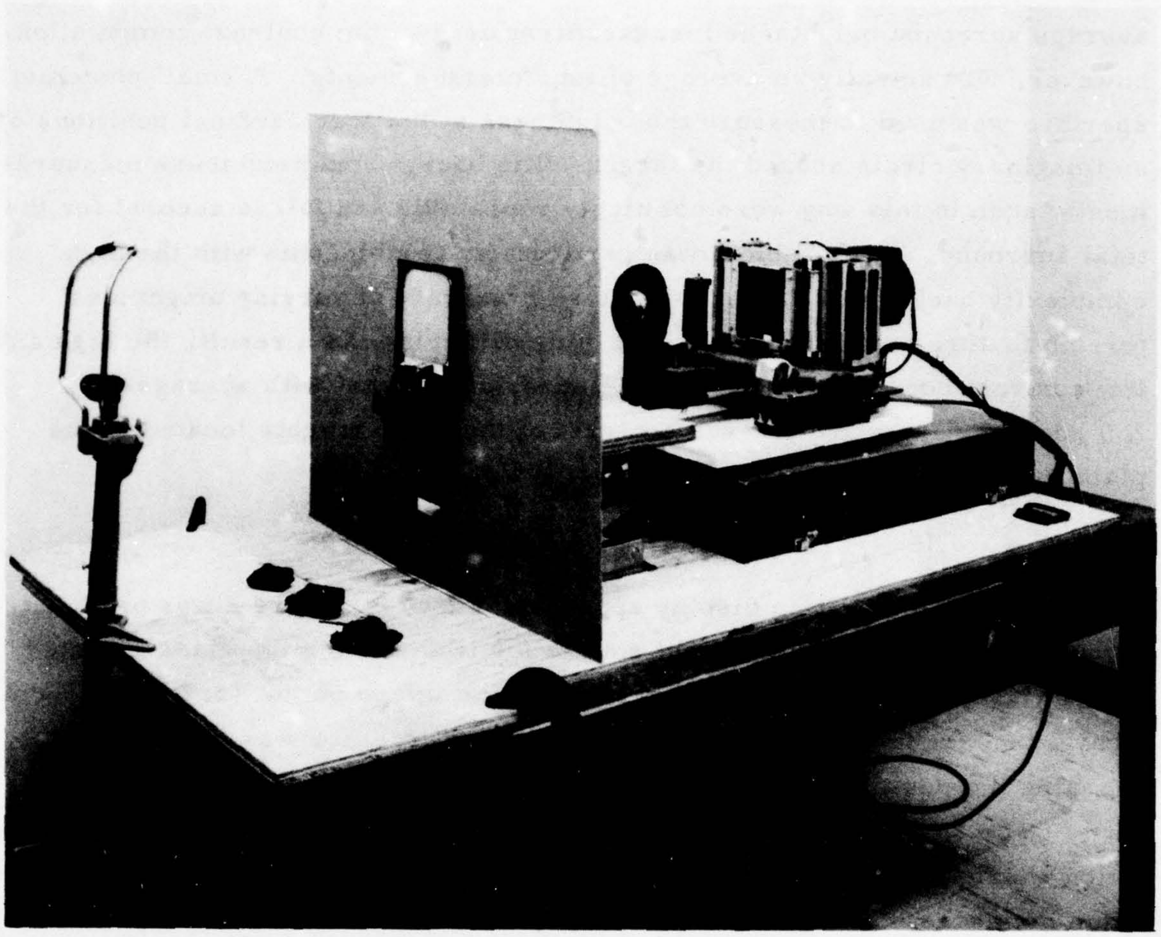


Figure 2. Rear-projection display used to present images to subjects.

Research Design

A mixed-factors factorial design with four variables at two levels and a fifth at three levels was used to examine target detection performance. Two levels of target-to-background contrast and scene complexity and three target types were within subject variables. Each subject experienced all 12 combinations of these variables. Two levels of resolution and target subtense were examined as between-subjects variables. Subjects were randomly assigned to one of four groups of 12 subjects each. Each of the four groups experienced one of the four combinations of resolution and target subtense.

Subjects

All 48 volunteer subjects were members of the technical staff at Hughes Aircraft Company, Culver City, California. In general, subjects were moderately familiar with target detection/recognition tasks using electro-optical sensors prior to the experimental session. All subjects used in the experiment had visual acuity of 20/20 or better as determined using a Snellen chart.

Procedures

Each subject attempted to detect and identify a target located in the 12 realistic terrain backgrounds and three special targets displayed against plain gray backgrounds. Each trial began when the experimenter projected a scene on the display and simultaneously started a stop watch for recording detection time. Subjects searched the scene for the target vehicle and when found, pointed to it on the display and said "there" to indicate a detection. The experimenter noted the time and correctness of the response and asked the subject for an identification of the vehicle type, if possible. If one minute and 50 seconds elapsed without a response from the subject, the experimenter requested a *"best guess"* as to the target location. If the subject was correct in this case, a detection time of 120 seconds was recorded. In all cases, if the response was incorrect a miss was recorded.

Prior to the start of experimental trials, the subject was given a standardized set of written instructions which described the general purpose of the experiment, the nature of the task, and the characteristics of the target images. The experimenter verbally reiterated the major points in the instructions and answered questions posed by the subjects. Several minutes were provided to allow the subject to study the vehicle models and become familiar with their features. During this time subjects were encouraged to examine the models from several orientations. An image with the three targets on a plain background was also displayed to demonstrate the characteristics of the targets at the resolution and size to be encountered during the experimental trials. Following the target familiarization, three training trials were given to clarify the procedures to be used.

The instructions to the subject were carefully worded in an attempt to minimize the response criterion problem discussed by Swets, Tanner and Birdsall (1961). Probability of detection and time to detect are closely related to one another in that a short time can be obtained by sacrificing probability of detection. Obversely, a high probability can be obtained at the cost of time. Subjects were instructed to respond if they were 70 percent certain that the object was a target. The intent was for the probability of detection to have minimum variability so time could be the primary dependent measure.

This method of controlling response criterion is much less desirable and effective than, for example, using a forced-choice procedure. The principle difficulty lies in the subject's interpretation of the 70 percent instruction which can vary widely. As a result some subjects will still be more willing to guess and thus have shorter times to detect. Although not an optimum solution, the inclusion of a criterion level in the instructions was the best technique available in the present experiment.

Because display resolution and viewing distance were between-subjects parameters, the proper levels of these variables were established for each subject before he arrived. Scene complexity, target type, and target-to-background contrast, on the other hand, were within-subject variables fixed in the imagery. All subjects saw the same images, however, the order of presentation was counterbalanced to minimize any effect of presentation order.

RESULTS AND DISCUSSION

The primary dependent variable was the time required for target detection. In those cases where an incorrect detection was made, an arbitrary detection time of 150 seconds was recorded. The time scores, thus, included the influence of incorrect detections, albeit somewhat arbitrarily.

An analysis of variance on the detection time data was performed and a summary of this analysis is presented in Table 2. The analysis of variance revealed that all five main effects were reliable along with six two-way interactions, two three-way interactions and one four-way interaction.

The reliable resolution by target subtense interaction is of particular interest because it indicates a confounding of these two variables. Figure 3 presents the interaction and shows that with small subtense targets, resolution had no effect on the time required to detect a target. One possible explanation is that the visual acuity of the operator was limiting the detail that could be perceived. If this were the case, then increasing the display resolution from 240 to 480 lines would have resulted in no perceptable change in the display from the observers viewing distance. The conditions of the

TABLE 2. SUMMARY OF ANALYSIS OF VARIANCE OF DETECTION TIME FOR ALL CONDITIONS OF EXPERIMENT 1

Source	SS	df	F	P
Display Resolution (DR)	16418.2	1	5.14	0.050
Target Subtense (TS)	88066.0	1	27.56	0.001
DR x TS	18265.7	1	5.72	0.025
Subjects (S)/DR x TS	140574.0	44		
Scene Complexity (SC)	251185.9	1	125.75	0.001
DR x SC	3254.3	1	1.63	NS
TS x SC	255.5	1	0.13	NS
DR x TS x SC	4055.0	1	2.03	NS
SC x (S/DR x TS)	87887.5	44		
Target-to-Background Contrast (TC)	26805.7	1	8.22	0.010
DR x TC	32694.0	1	10.03	0.005
TS x TC	20385.7	1	6.25	0.025
DR x TS x TC	5144.2	1	1.58	NS
TC x (S/DR x TS)	143488.6	44		
Target Type (TT)	39465.7	2	7.20	0.005
DR x TT	3282.0	2	0.60	NS
TS x TT	4400.5	2	0.80	NS
DR x TS x TT	600.5	2	0.11	NS
TT x (S/DR x TS)	241066.6	88		
SC x TC	14774.8	1	7.52	0.010
DR x SC x TC	4465.7	1	2.27	NS
TS x SC x TC	7445.0	1	3.79	NS
DR x TS x SC x TC	2693.2	1	1.37	NS
SC x TC x (S/DR x TS)	86450.4	44		
SC x TT	160064.4	2	50.39	0.001
DR x SC x TT	3764.0	2	1.18	NS
TS x SC x TT	330.1	2	0.10	NS
DR x TS x SC x TT	976.7	2	0.31	NS
TS x TT x (S/DR x TS)	139780.4	88		
TC x TT	207972.1	2	52.60	0.001
DR x TC x TT	2612.6	2	0.66	NS

(Continued next page)

(Table 2, concluded)

Source	SS	df	F	P
TS x TC x TT	15735.4	2	3.98	0.025
DR x TS x TC x TT	5860.0	2	1.48	NS
TC x TT x (S/DR x TS)	173964.4	88		
SC x TC x TT	102079.2	2	37.71	0.001
DR x SC x TC x TT	8450.5	2	3.21	0.050
TS x SC x TC x TT	2497.7	2	0.92	NS
DR x TS x SC x TC x TT	2449.5	2	0.91	NS
SC x TC x TT x (S/DR x TS)	119092.3	88		

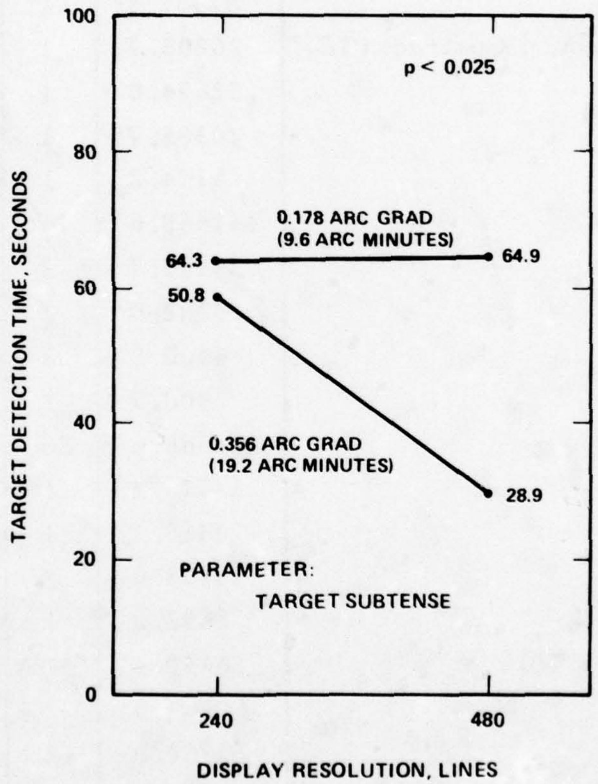


Figure 3. The effects of display resolution and target subtense on the time required to detect a target.

experiment had been selected on the assumption that the observer had an acuity of 0.0185 arc grad (1 arc minute). If such an assumption were true, then in the 480 line resolution, 0.178 arc grad (9.6 arc minute) condition the observer should have been able to just resolve the 9.6 lines across the target. If, however, the observers acuity was less than that assumed, more lines would have been across the target than the observer could perceive.

A post hoc check on the validity of the subject acuity assumption was made by determining the number of lines across the target that could be resolved by the subjects in the four resolution and subtense conditions. A Buckbee-Mears resolution chart was used with the display conditions as nearly identical to the previous experimental conditions as possible with the exception of contrast which is inherently much greater with the resolution chart. The results indicated that, even with the higher contrast, for the small subtense target only 4.6 and 6.5 lines could be resolved in the 240 and 480 line resolution conditions respectively. The calculated number of lines, assuming 0.0185 arc grad (1 arc minute) acuity, would have been 4.8 and 9.6 for the two conditions. Clearly subject acuity was a limiting factor in the small subtense conditions.

The acuity value assumed was entirely reasonable based on published psychophysical data (Patel, 1966; Rogers and Carel, 1973) and was consistant with the results of the Snellen acuity test given to each subject prior to the experiment. The lower acuity found using the resolution chart and actual display suggest differences between the recognition acuity measure provided by the Snellen chart and the resolution measure and/or differences due to the viewing conditions. Regardless of the reason for the differences, it is clear that great care needs to be exercised any time an experiment is conducted with conditions near a psychophysical threshold.

Because of the confounding of resolution and target subtense due to acuity limitations of the subject, the individual effects of the two variables cannot be separated and interpretation of the analysis of variance can be difficult or even erroneous. For example, any effect involving resolution summed across both target subtenses would be artificially small, potentially leading to an incorrect conclusion concerning the magnitude of the resolution effect. Further, because each subject had a different acuity limit, the between-subjects variability would have increased making it more difficult to obtain statistical reliability. For these reasons the two groups of subjects that experienced the small target subtense condition were excluded from further analysis.

A second analysis of variance on the time to detect data, excluding the small target subtense condition, was performed and a summary is presented in Table 3. This analysis showed the main effects of resolution, scene complexity and target type to be reliable ($p \leq 0.01$). Additionally three two-way interactions and two three-way interactions were reliable at the 0.05 level or beyond. As expected several of the effects involving resolution changed in

TABLE 3. SUMMARY OF ANALYSIS OF VARIANCE OF DETECTION TIME FOR LARGE TARGET SUBTENSE CONDITION OF EXPERIMENT 1

Source	SS	df	F	P
Display Resolution (DR)	34658.2	1	17.61	0.001
Subjects (S)/DR	43286.2	22		
Scene Complexity (SC)	117614.2	1	63.52	0.001
DR x SC	25.1	1	0.01	NS
SC x S/DR	40734.5	22		
Target-to-Background Contrast (TC)	219.2	1	0.10	NS
DR x TC	5955.7	1	2.81	NS
TC x S/DR	46593.0	22		
Target Type (TT)	29981.9	2	5.59	0.010
DR x TT	621.8	2	0.12	NS
TT x S/DR	117998.2	44		
SC x TC	21593.7	1	14.75	0.001
DR x SC x TC	108.0	1	0.07	NS
SC x TC x S/DR	32200.2	22		
SC x TT	73107.2	2	27.90	0.001
DR x SC x TT	4094.6	2	1.56	NS
SC x TT x S/DR	57648.6	44		
TC x TT	95492.8	2	50.14	0.001
DR x TC x TT	6435.1	2	3.38	0.050
TC x TT x S/DR	41896.3	44		
SC x TC x TT	42437.5	2	20.00	0.001
DR x SC x TC x TT	4777.8	2	2.25	NS
SC x TC x TS x S/DR	46708.0	44		

importance. The main effect of resolution increased in reliability from 0.05 to 0.001 while the resolution by target-to-background contrast interaction which had been reliable at the 0.005 level became non-significant.

Of the five reliable interactions, four involved target type. The percent of variance accounted for by all effects involving target type, calculated as the sum of the sums-of-squares for the effects divided by the total sum-of-squares, was 29.72 percent. All of the effects in the analysis accounted for 50.51 percent of the variance which means that target type accounted for nearly 60 percent of all of the effects.

So large a contribution from a variable that was expected to be relatively unimportant, required further investigation.

The main effect of target type would indicate that there were major differences in detection performance among the three types of targets. The difference cannot be attributed to changes in target size because it was fixed. Nor are the differences in shape or height-to-width ratio sufficiently great to provide satisfactory explanations. It could be that particular vehicle types were more easily confused with clutter objects but this, too, fails as an explanation.

The two-way interactions of target type with scene complexity and target-to-background contrast and, more particularly, their three-way interaction indicated considerable differences among individual stimulus images. Recall that any given target type and contrast occurred with only a single background scene. Even though the six high complex and six low complex scenes were selected based on highly consistant judgment by six judges, it may be that the location of the target within a scene also strongly influenced detection time. If this were the case, the interactions might have been the result of a combined effect of background scene and target within that scene rather than a difference due to vehicle type per se.

If the large effects of target type are interpreted as reflecting the influence of the placement of a target within a particular scene then why the reliable main effect of target type? Each type of target vehicle occurred in four different scenes and, although not an impressively large number, should have been adequate to make the magnitude of observed main effect unlikely. One explanation might be that the placement of the target in one particular scene resulted in performance highly deviant from the remaining set of images. A single deviant image would have inflated the effects of scene complexity, target-to-background contrast, target type and all interactions with these variables because each image represented the only sample of any particular combination of levels of these three variables.

The possibility that one or more images were unusually easy or difficult was examined by tabulating the mean time and probability of detection for each of the twelve individual images. The results are given in Table 4 where each time and probability entry represents the average over two resolutions times twelve subjects. The time assigned for an incorrect detection was 150 seconds. As can be observed by examining Table 4, Image 11, a high-complexity background with a high-contrast tank target, was unusually difficult for subjects to detect. The mean time to detect was 131.5 seconds; more than three times as long as the overall average of 39.9 seconds and nearly twice as long as the second most difficult image. The probabilities of detection follow a similar pattern.

TABLE 4. MEAN-TIME AND PROBABILITY OF DETECTION FOR EACH STIMULUS IMAGE OF EXPERIMENT 1

Image	Scene Complexity	Target-to-Background Contrast, Percent	Target Type	Time to Detect, Seconds	Probability of Detection
1	Low	70	APC	57.7	0.84
2	Low	70	Tank	24.9	0.96
3	Low	70	Truck	4.9	1.00
4	Low	200	APC	2.9	1.00
5	Low	200	Tank	16.6	0.92
6	Low	200	Truck	11.0	0.96
7	High	70	APC	45.2	0.79
8	High	70	Tank	38.6	0.92
9	High	70	Truck	73.1	0.67
10	High	200	APC	6.2	1.00
11	High	200	Tank	131.5	0.29
12	High	200	Truck	66.0	0.67
				M= 39.9	M= 0.83

Although Image 11 represented a realistic condition that may occur in actual situations, with the relatively limited stimulus sample used in this study, the presence of so deviant an image could have biased the results. To remove the potential for bias, the data obtained using Image 11 were treated as if they had been lost. The missing data procedure given by Winer (1971, p 487-490), a simplified version of the method described by Bennett and Franklin (1954, p 382-383), was used to fill the resulting empty cells. Briefly, this procedure replaced the missing values with estimates based on the best fitting plane surface through adjacent cells. This method replaced the actual detection times for the twelve subjects in the low resolution condition with a value of 26.5 seconds and the data for the subjects in the high resolution condition with a value of 13.2 seconds.

The treatment of Image 11 as missing data was not the ideal way to solve the problem of unusually difficult or easy stimuli. Because all twelve scores within the affected cells were replaced with the same time to detect score, the between-subjects variance in those cells was zero which artificially reduced the experimental error. However, if this were recognized and considered when interpreting the analysis, the likelihood of an erroneous conclusion would be less than if the deviant image were retained. Identification of the unusually difficult image prior to formal data collection, through extensive pre-test, or a considerably larger sample of stimuli would have been two possible ways to avoid the problem under discussion. Both, however, represented major increases in data collection and/or stimulus preparation and were not justified for this initial exploratory work.

A summary of the analysis of variance of the time to detect scores, with Image 11 treated as missing data, is given in Table 5. Also tabulated is Eta^2, the proportion of variance accounted for by each effect, calculated as the sum-of-squares for the effect divided by the total sum-of-squares. A comparison of Tables 3 and 5 revealed several changes as a result of removing Image 11. First, the F-ratios for both display resolution (DR) and scene complexity (SC) changed although the level of significance remained the same ($p < 0.001$). The F-ratio for DR increased reflecting, primarily, the reduction in subject variance due to the estimate of Image 11. The F-ratio for SC decreased considerably because Image 11, a high complexity scene, had been a major source of the previously large difference.

TABLE 5. SUMMARY OF ANALYSIS OF VARIANCE OF DETECTION TIME FOR LARGE TARGET SUBTENSE CONDITION OF EXPERIMENT 1 WITH IMAGE 11 ESTIMATED AS IF MISSING

Source	SS	df	F	P	Eta^2 (%)
Display Resolution (DR)	35998.0	1	20.1	0.001	5.79
Subjects (S)/DR	39392.5	22			
Scene Complexity (SC)	34259.8	1	20.07	0.001	5.51
DR x SC	3.9	1	0.00	NS	0.00
SC x S/DR	37560.2	22			
Target-to-Background Contrast (TC)	29811.5	1	16.49	0.001	4.79
DR x TC	5419.4	1	3.00	NS	0.87
TC x S/DR	39773.2	22			
Target Type (TT)	10042.8	2	1.99	NS	1.61
DR x TT	619.6	2	0.12	NS	0.10
TT x S/DR	111102.7	44			
SC x TC	121.4	1	0.10	NS	0.02
DR x SC x TC	44.6	1	0.04	NS	0.01
SC x TC x S/DR	26245.5	22			
SC x TT	58996.9	2	21.94	0.001	9.49
DR x SC x TT	4035.6	2	1.50	NS	0.65
SC x TT x S/DR	59170.2	44			
TC x TT	27505.5	2	14.40	0.001	4.42
DR x TC x TT	6789.4	2	3.56	0.050	1.09
TC x TT x S/DR	42007.6	44			
SC x TC x TT	3079.3	2	1.49	NS	0.50
DR x SC x TC x TT	4417.0	2	2.14	NS	0.71
SC x TC x TT x S/DR	45500.9	44			

As expected, the main effect of target type (TT) was not reliable once the unusually difficult image was removed. The previously reliable scene complexity (SC) by target-to-background contrast (TC) and the SC x TC x TT interactions were no longer reliable. Instead, the main effect of TC became a reliable ($p \leq 0.001$) effect. Not all of the interactions with target type became insignificant, however. The SC x TT and TC x TT, interactions remained reliable at the 0.001 level and the DR x TC x TT interaction at the 0.05 level.

The effect of background scene complexity shown in Figure 4 was of primary interest in the experiment and was one of the largest determinants of performance. Average detection time with low complexity backgrounds was 19.7 seconds while with high complexity backgrounds the average time increased by a factor of 2.1 to 41.5 seconds. This effect was reliable beyond the 0.001 level of significance which means that a difference this large would be expected to occur by chance less than once in a thousand replications of the experiment. It also accounted for 5.51 percent of the variance in the experiment.

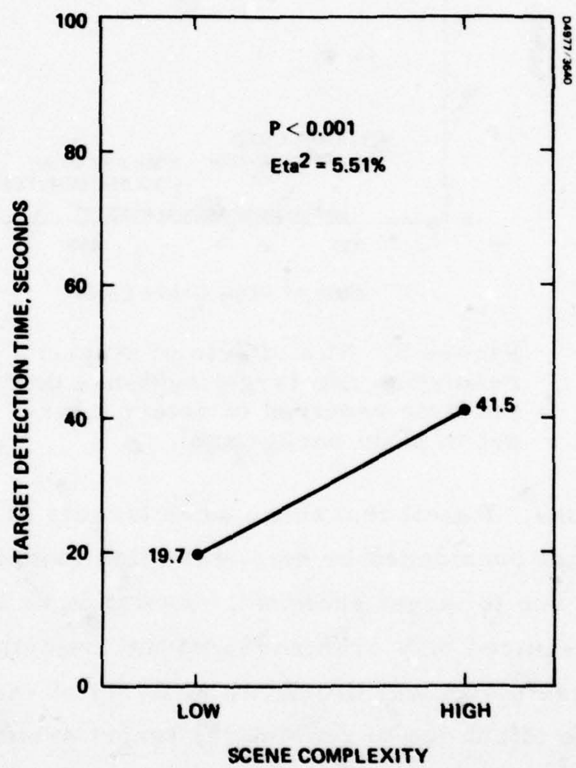


Figure 4. The effect of scene complexity on the time required to detect a target.

The importance of including the effect of scene complexity in a model of target detection performance can be even more dramatically demonstrated by referring to Figure 5 which presents the average detection times for the

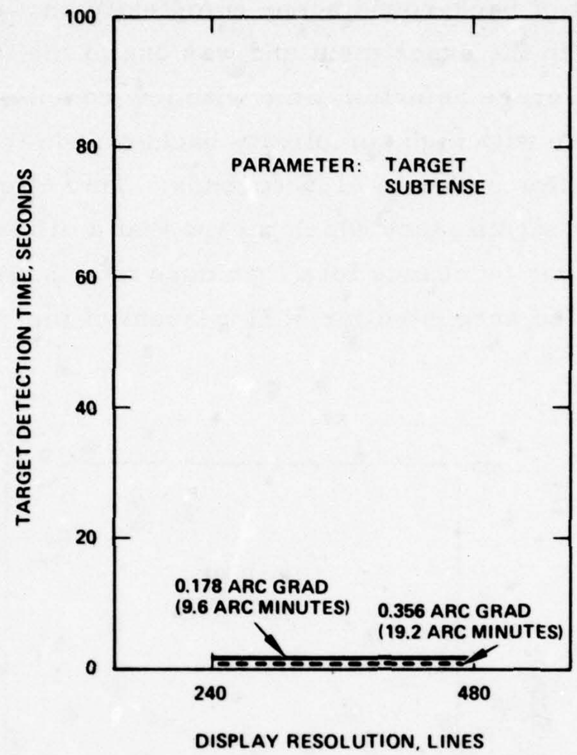


Figure 5. The effects of display resolution and target subtense on the time required to detect a target in plain background.

special test conditions. Recall that these were targets in a plain gray background similar to that considered by most detection models. There was virtually no difference due to target subtense, resolution or target type, and an average detection required only 1.75 seconds; the majority of which was reaction and response movement time. An analysis of variance for this data indicated no reliable effect due to resolution, target subtense, target type or any of their interactions. A model based on a uniform background can underestimate the time required to detect a target in a highly complex background by a factor of nearly 24. For low complexity backgrounds the factor is approximately 11 to 1 which is still substantial, considering that targets were generally located in open fields and only a few clutter objects were present.

The absence of any effect due to subtense with the targets located in the plain background indicates that the targets had sufficient contrast to be well above threshold consistent with Blackwell's (1946) data. Further, the absence of an effect due to resolution indicates that sufficient target detail was available even at the smallest subtense and lowest resolution. The acuity limitation found with the realistic backgrounds was not present, suggesting that subjects could detect a target primarily on the basis of brightness differences between the target and the plain background. Without clutter objects present in the background, targets could be detected with a probability of one regardless of resolution and with contrast as low as 20 percent. The presence of even a simple background dramatically increased the time required for detection and apparently changed the task to one requiring greater spatial detail.

The overall effect of target-to-background contrast is presented in Figure 6. This effect was reliable at the 0.001 level of significance and agrees with previous studies that have shown improved performance with increased contrast. In the present case, increasing contrast from 0.7 to 2.0 resulted in

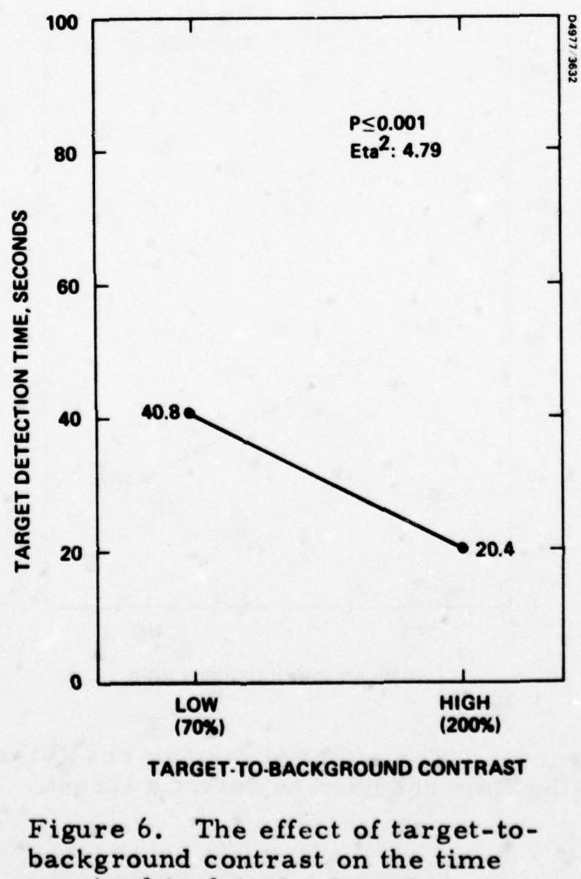


Figure 6. The effect of target-to-background contrast on the time required to detect a target.

a factor of 2 reduction in detection time from 40.8 to 20.4 seconds. Clearly the luminance differences between the target and its immediate background was an important cue to a target located in realistic terrain.

The effect of increasing display resolution from 240 to 480 lines was to halve the detection time from 41.8 to 19.4 seconds as shown in Figure 7. As previously discussed, resolution had no effect when the target was located in a uniform background, however, with realistic backgrounds resolution had a large effect. This would support the hypothesis that as scene complexity increased from zero to the real-world scenes used in the present study the detection task changed from one of luminance detection to one that included a larger component of form perception. The ability to discriminate shape and other spatial factors would be dependent upon the resolution of the display. As the need for spatial detail increased because of the complex background, the importance of resolution increased.

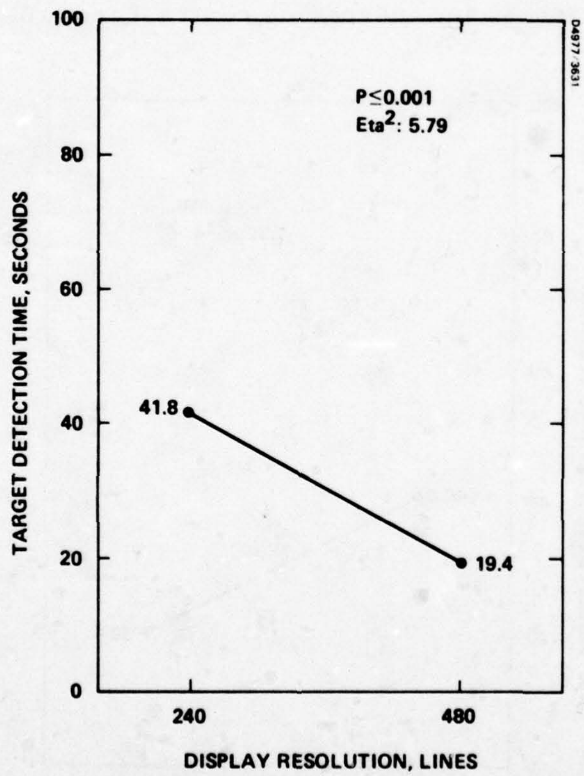


Figure 7. The effect of display resolution on the time required to detect a target.

This hypothesized change in the relative importance of spatial cues can be seen in Figure 8 which plots the time to detect for the two resolutions and all three backgrounds. Resolution made no difference with zero clutter but resulted in more than a 22 second difference for the complex backgrounds. The absence of any difference in the effect of resolution between the two levels of scene complexity suggests that the change in task occurred between zero complexity and the lowest real scene complexity used in this study. Further work will be required to determine where the shift occurs and whether it is abrupt like a threshold or if it is a more gradual change.

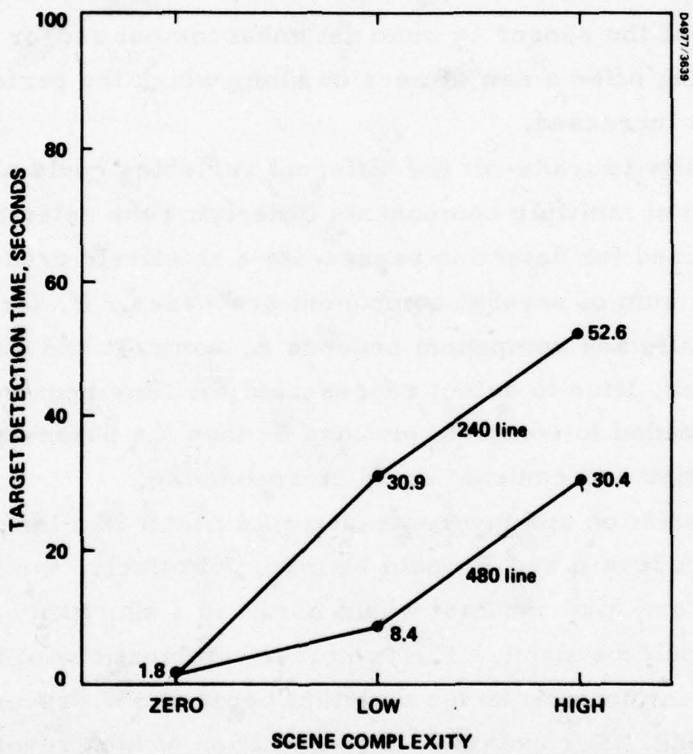


Figure 8. The effects of scene complexity and display resolution on the time required to detect a target.

The absence of realiable SC x TC and DR x TC interactions as well as the non-reliable interaction of DR and SC for realistic scenes shown in Figure 8, may bear on the issue of a shift in task. Each of these relationships was such that if both variables were at the difficult level, performance was poor and if both variables were at the easy level, performance was good.

However, if either variable was at the difficult level and the other at the easy level performance was identical. For example, high resolution combined with low target-to-background contrast and low resolution combined with high target-to-background contrast resulted in nearly identical times to detect. That is, low target-to-background contrast could be compensated for by higher resolution or, alternatively low resolution could be compensated for by higher target-to-background contrast.

These trade-off characteristics can have considerable implications for the design of sensor systems. Rather than attempting to achieve improved performance solely by improving resolution, techniques that result in improved target-to-background contrast can also be effective. Improved use of the dynamic range of the sensor by contrast enhancement and/or other forms of image processing offer a new dimension along which the performance of sensor systems can be increased.

The ability to trade-off the different variables could also be interpreted as an indication of multiple components underlying the detection process. The total time required for detection represents a relatively crude measure which may reflect the sum of several component processes. If, for example, resolution primarily affected component process A, contrast had a major influence on process B and, time to detect represented the time required for process A plus the time needed to complete process B; then the observed relationship between resolution and contrast could be reconciled.

Low resolution and low contrast would result in a long time to detect because both process A and B would be long. Similarly, the combination of high resolution and high contrast would result in a short time to detect because both A and B would be short. The two other combinations of resolution and contrast result in identical times to detect because one process would be short and the other long. For example, a combination of high resolution and low contrast would cause process A to be short and process B to be long. Conversely, low resolution and high contrast would result in a short time for process A and a long time for process B. Although, considerable differences in the relative contributions of the two component processes would exist, the overall performance would appear to be identical.

In terms of these A and B processes, the data shown in Figure 8 would be interpreted as follows. When targets were detected in the uniform background the contrast sensitive process was a major determinant of detection time. However, when detection was accomplished with the real-world scenes a second component process, sensitive to changes in resolution, became a contributor to the total time required for detection.

Consideration of the detection of targets in realistic, complex background scenes as reflecting the influence of two or more component processes can have a major impact on the form of model needed to describe performance. An analysis of the target detection process suggests the presence of several components that may need to be integrated into a complete model.

Prior to the initiation of target search, the observer must spend a short period of time orienting himself. During this period any a priori knowledge, concerning the general viewing conditions is combined with the global features of the scene to form a description of what is to be searched. For example, a priori information such as; oblique view of rural terrain with a 10-degree field of view, 300-foot altitude, center of display equivalent to approximately 4 kilometer range, etc.; is supplemented by global scene information such as; a road beginning at close range and crossing the display from left to right as it extends to long range, dense trees in the upper left region of the scene, open field with scattered trees to the right of the road, etc.

The length of time required for this orientation process will be directly influenced by the extent of the a priori knowledge and its agreement with the scene actually viewed. In most operational situations this information will be extensive and will agree closely with the scene, reducing the orientation time to a very small proportion of the total acquisition time. In experimental situations the a priori knowledge is a function of the instructions given the subject and the type and amount of training provided. In these conditions, the agreement between what is expected and what is observed may or may not be high depending upon the attention given this effect by the experimenter.

Once the observer has oriented himself, a systematic search of the scene is made by rapidly fixating various locations with his eye. The areas to be examined and perhaps the order are influenced by the global description obtained during the orientation process and the observer's knowledge or assessment of likely target locations (Krebs and Graf, 1973; Krebs and Lorence, 1975; Noton and Stark, 1971). For example, tanks are likely to be on roads but certainly will not be in the middle of a lake. Within any area being searched, the attention getting characteristics of objects in the near periphery of the fixated position can also be expected to influence succeeding fixations although the factors that cause an object in the periphery to attract attention are not well understood.

The search process, thus, consists of rapid fixations of points in the scene which migh contain an object with a sufficient number of expected target characteristics to warrant detailed examination. The probability that a fixated candidate object will cause the observer to make this detailed examination will be influenced primarily by target characteristics and secondarily, by the interaction of target and scene.

The first two components, orientation and search to initial detection, are followed in time by the third component of the process. Once an object has been selected as a potential or reasonably suspect target, a more detailed examination takes place to determine if it has enough features or attributes to be definitely a tactical target. This process requires the first stages of recognition; however, it is not necessary that a tank be discriminated from a truck. In the case of a uniform background, this second component of the task is not necessary. With realistic scenes, however, many competing objects may be present in the immediate vicinity of the target and these need to be discriminated from the target. This process will be affected by the target and the scene in the immediate area about the target as well as the interaction of the two. Further, the probability of correctly completing this portion of the task will be a function of time. As time is spent extracting candidate target characteristics a continuing decision process is taking place. The result of this process determines whether the observer continues to examine the candidate target; decides the object is not a target and resumes search or decides the object is a target and selects an appropriate response.

Assuming that the orientation process represents only a small proportion of time compared to the other components and that the decision process can be incorporated in the recognition component, it is possible to derive a two component or two stage model of the target acquisition task. A two parameter model of the search and detection component might be similar to the NVL model. That is:

$$P_D(t) = P_D(1 - e^{-t/\tau_1})$$

where:

$P_D(t)$ = cumulative probability of detecting an object as a candidate target as a function of time.

P_D = probability that an object will be detected as a candidate target in a single fixation.

τ_1 = time constant reflecting the influence of the scene on the search pattern.

t = elapsed time from beginning of search.

A second formulation, similar to the one for detection but effective only after detection has occurred, can be used to describe the recognition and decision process as follows:

$$P_R(t-t_D) = P_R(1 - e^{-(t-t_D)/\tau_2}), \quad t \geq t_D; \; P_R(t-t_D) = 0, \quad t < t_D$$

where:

$P_R(t-t_D)$ = cumulative probability that a detected object will be recognized as a tactical target as a function of time after detection.

P_R = probability that an object will be recognized as a target given infinite observation time.

τ_2 = time constant reflecting the influence of the immediate scene on the time course of recognition.

t = elapsed time from beginning of search.

t_D = elapsed time at which detection occurred.

The probabilities $P_D(t)$ and $P_R(t-t_D)$ are combined in the following manner. Each function is first differentiated with respect to t, yielding $f_D(t) = \frac{d}{dt}(t)$ and $f_R(t-t_D) = \frac{d}{dt}(t-t_D)$, which are respectively the detection probability density function and the recognition probability density function conditional on detection. The product of $f_D(t)$ and $f_R(t-t_D)$ is the joint probability density of detection and recognition which may be integrated to obtain the cumulative probability of detection and recognition as a function of time as follows:

$$P_{RD}(T) = \int_0^T \int_{t_D}^T f_R(t-t_D)\, f_D(t_D)\, dt\, dt_D$$

The cumulative probability of detecting and recognizing at or before time T equals the integral of the probability density of recognition conditional on detection, weighted by the probability density of detection.

The component processes just described are subject to further refinement based on a more detailed examination of the existing literature and additional experimentation. For example, the component form given above does not explicity cover the case where the detection and recognition processes are exercised repetitively before a response is made. The approach, however, appears to be viable and warrants further consideration.

Whether the component model ultimately proves correct or not, an attempt to fit a two parameter model of the form

$$P_D(t) = P_\infty (1 - e^{-t/\tau})$$

to the present data makes it very clear that this form cannot adequately describe the cumulative probability of detection as a function of time for realistic targets and scenes. A least-squares estimation of P_∞ and τ were calculated using a linear-Taylor differential correction technique (McCalla, 1967) and selected results are presented in Figures 9 to 12 and summarized in Table 6. The Chi^2 value given in Table 5 was calculated as a comparison between the observed distribution of time scores and the distribution predicted by the estimated model and represents an assessment of the goodness-of-fit (Hays, 1963, p. 586-588). The null hypothesis of no difference between actual and predicted curves can be rejected with the probability given in the right most column. Note that except for a few isolated cases the two parameter model fails to adequately describe the observed data.

The analysis presented in Table 5 also indicated that three interactions with target type, Figures 13 to 15, remained reliable ($p \leq 0.05$), perhaps indicating the influence of the placement of the target within the scene. The original ratings of scene complexity were made without the target present and it may be that the placement of the target in the scene had a strong influence on detection time. For example, the images requiring the longest times to detect were generally those where the target was located in an area with a number of similar sized objects that formed a pattern. This potential interrelation between target and scene was considered in Experiments 2 and 3 which attempted to quantify scene complexity.

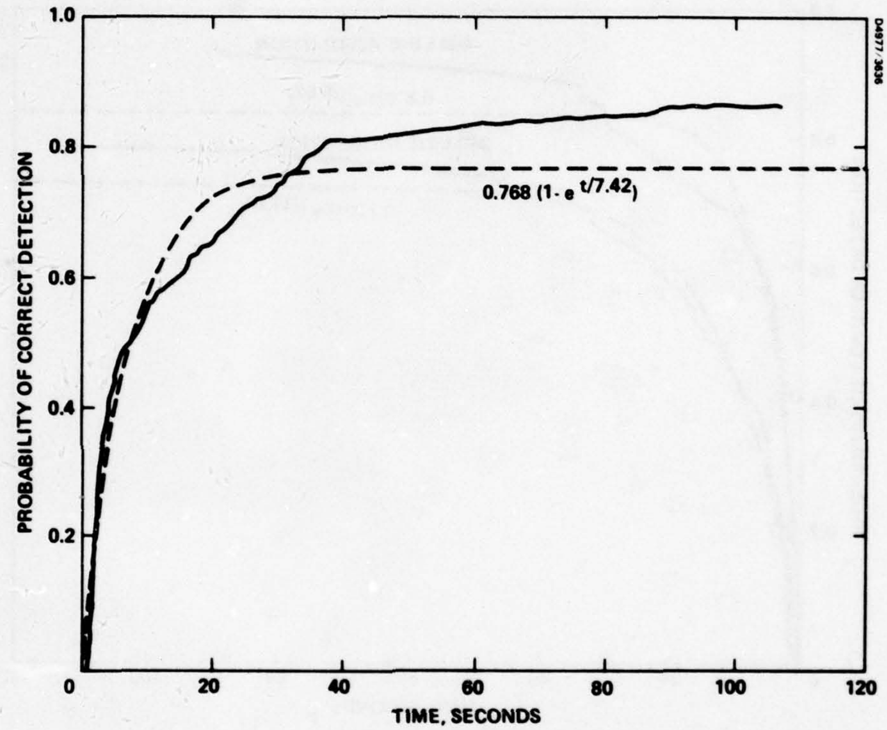


Figure 9. Cumulative probability of correct target detection and best fitting two parameter model for all conditions.

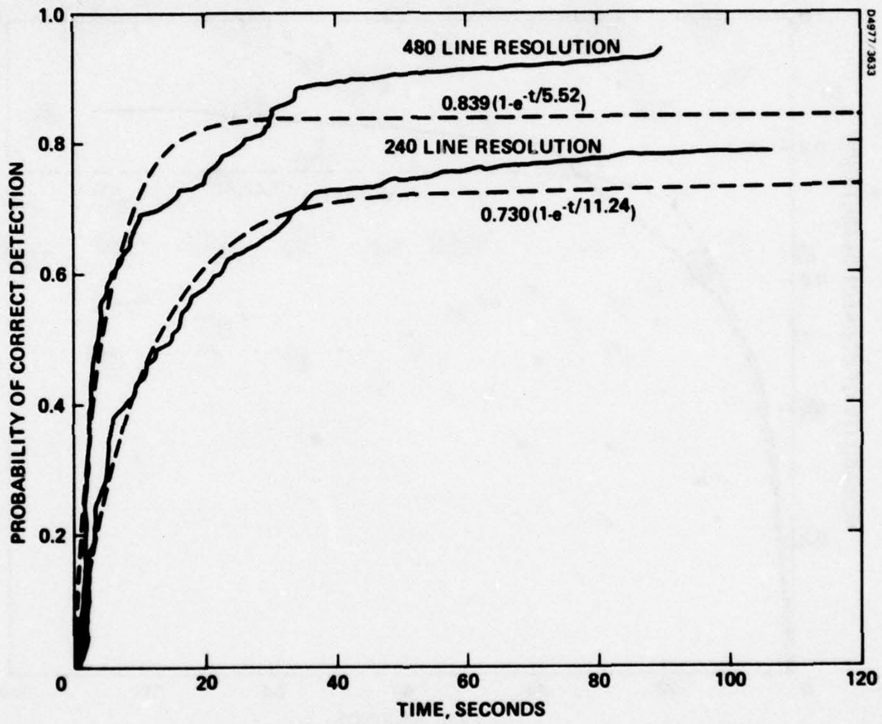


Figure 10. Cumulative probability of correct target detection and best fitting two parameter model for two conditions of resolution.

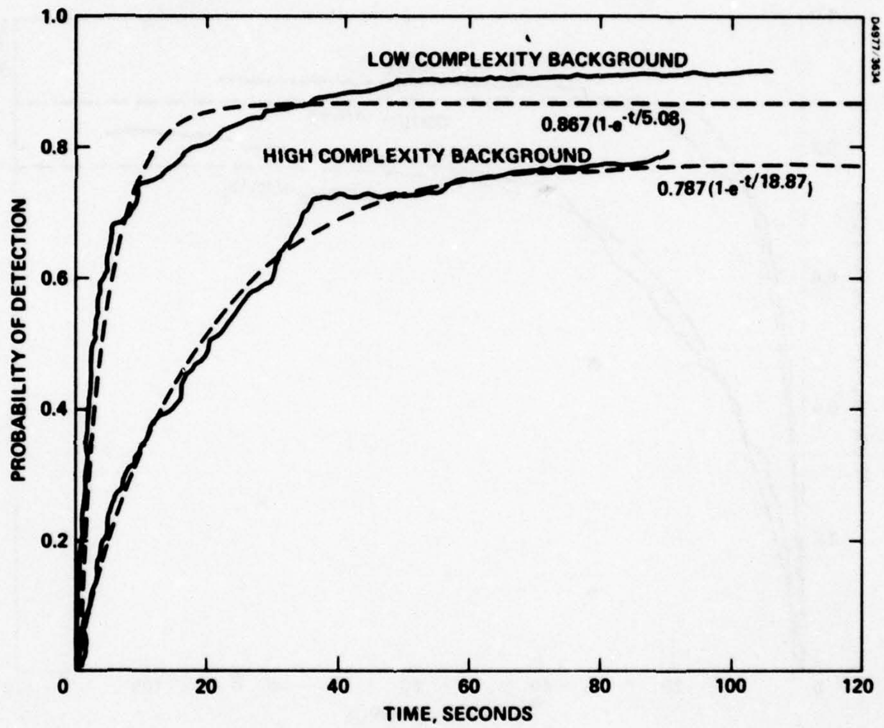


Figure 11. Cumulative probability of correct target detection and best fitting two parameter model for two conditions of scene complexity.

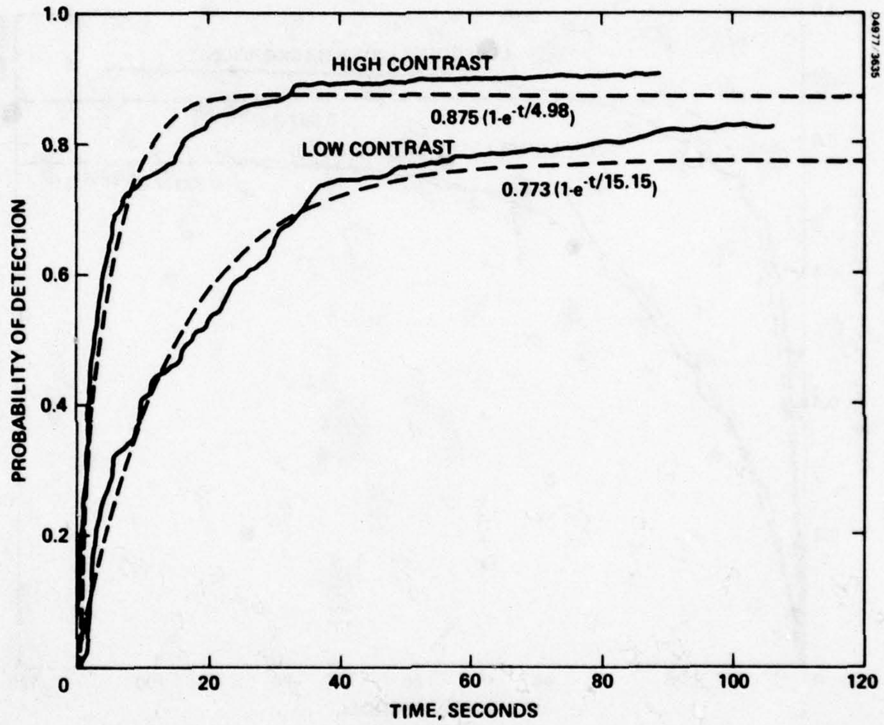


Figure 12. Cumulative probability of correct target detection and best fitting two parameter model for two conditions of target-to-background contrast.

TABLE 6. PARAMETERS FOR THE MODEL $P_\infty (1 - e^{-t/\tau})$ FIT TO THE DATA OF EXPERIMENT 1

Conditions			Parameters		Goodness-of-Fit	
Resolution	Scene Complexity	Target-to-Background Contrast	P_∞	τ	Chi^2	p
All	All	All	0.768	7.42	142.1	0.001
Low	All	All	0.730	11.24	38.7	0.001
High	All	All	0.839	5.52	86.8	0.001
All	Low	All	0.867	5.08	90.9	0.001
All	High	All	0.787	18.87	16.3	0.025
All	All	Low	0.773	15.15	46.8	0.001
All	All	High	0.875	4.98	54.9	0.001
Low	Low	All	0.793	7.04	32.3	0.001
Low	High	All	0.751	24.18	7.9	0.500
High	Low	All	0.974	4.07	57.9	0.001
High	High	All	0.816	13.89	19.3	0.010
Low	All	Low	0.730	23.26	6.8	0.500
Low	All	High	0.844	6.90	19.5	0.010
High	All	Low	0.832	10.00	40.3	0.001
High	All	High	0.951	3.94	43.5	0.001
All	Low	Low	0.785	7.19	60.6	0.001
All	Low	High	1.000	4.24	43.3	0.001
All	High	Low	0.845	31.25	21.5	0.005
All	High	High	0.764	8.20	18.0	0.025
Low	Low	Low	0.702	11.36	19.9	0.010
Low	Low	High	0.977	5.99	19.4	0.010
Low	High	Low	0.772	38.36	7.7	0.500
Low	High	High	0.786	11.36	11.0	0.250
High	Low	Low	0.911	5.49	49.8	0.001
High	Low	High	1.000	2.96	35.7	0.001
High	High	Low	0.932	25.00	9.2	0.250
High	High	High	0.810	6.06	17.6	0.025

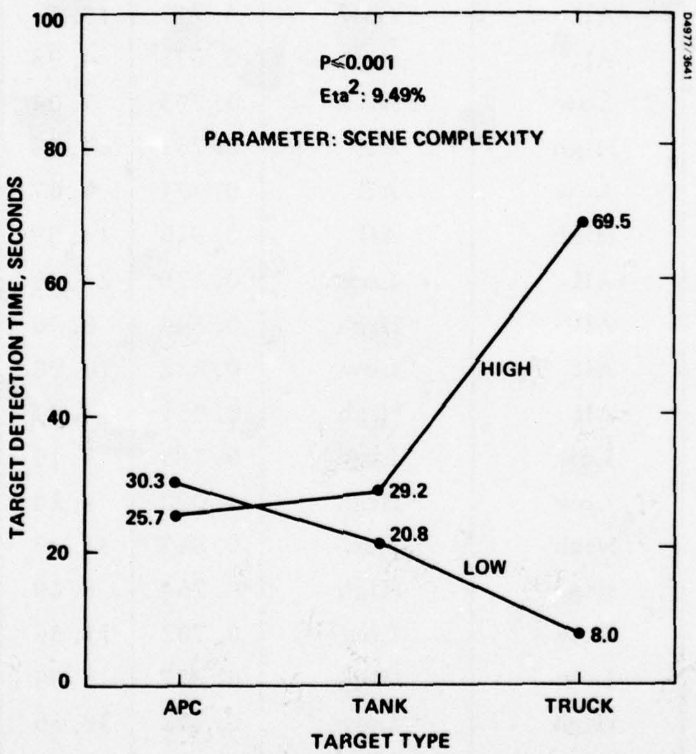


Figure 13. The effects of target type and scene complexity on the time required to detect a target.

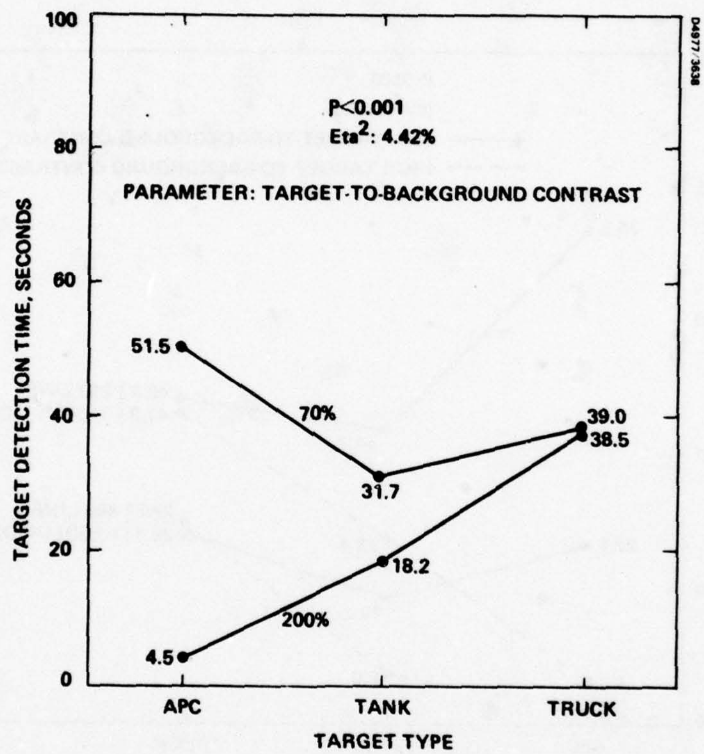


Figure 14. The effects of target type and target-to-background contrast on the time required to detect a target.

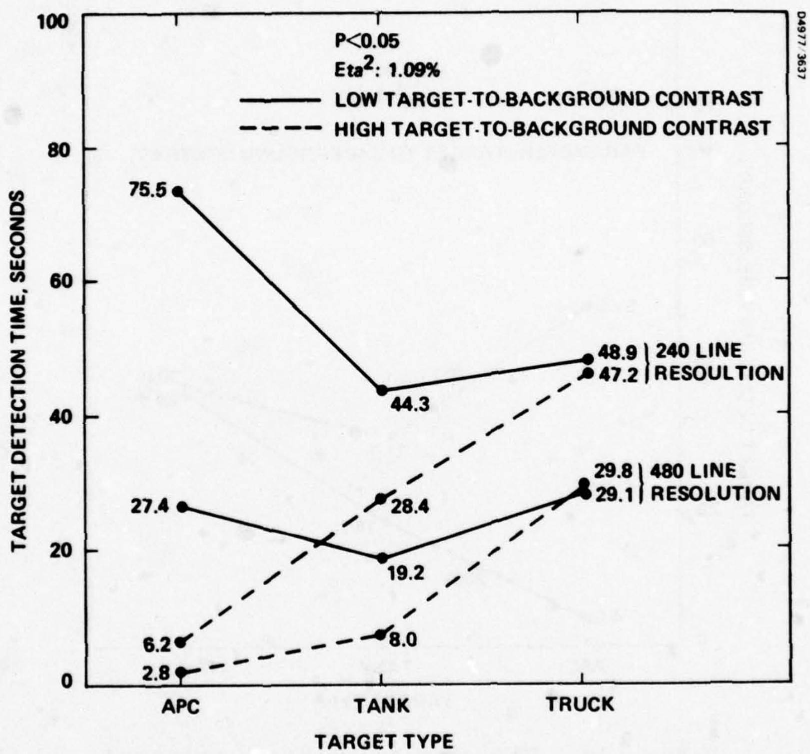


Figure 15. The effects of target type, resolution, and target-to-background contrast on the time required to detect a target.

EXPERIMENT 2

BACKGROUND

Regardless of the form that the final model might take, some number of parameters will have to be estimated as a function of a number of variables, including scene metrics. The most straightforward method for accomplishing the required estimation is to fit an n^{th} order polynomial expression of the independent variables to the obtained data.

As an example, a complete second-order polynomial expression involving K variables or factors would have the form:

$$Y = \beta_o + \sum_{i=1}^{K} \beta_i X_i + \sum_{i=1}^{K} \beta_{ii} X_i^2 + \sum_{i=1}^{K} \sum_{\substack{j=1 \\ j \neq 1}}^{K} \beta_{ij} X_i X_j$$

where: Y is the performance, the β's are weighting factors, and the X's are the scene metrics. The weighting factors in the above equation can be determined using multiple regression techniques. If Y is a function of one of the model parameters, the fit equation will provide the required estimate of the parameters.

However, before the above equation can be obtained, it is necessary to determine which metrics of scene complexity are likely to provide a reasonably descriptive set. The present experiment attempted to identify candidate metrics by having subjects verbalize those characteristics of the scene that made detection easy or difficult. A compilation of the responses was used to suggest some simple metrics of scene complexity that might provide a basis for a more complete target detection model.

METHOD

The stimuli and apparatus were identical to those used in Experiment 1. Twelve members of the Technical Staff at Hughes Aircraft Company who had not participated in Experiment 1 served as subjects. Six viewed the high complexity scenes and six the low complexity scenes. Each subject was asked to verbalize his subjective opinions as to those scene or target characteristics he felt would aid or hinder detection of the target. All responses were tape recorded for subsequent study.

RESULTS AND DISCUSSION

The tape recordings were analyzed to obtain a common set of most mentioned factors with the result shown in Table 7. Although there was considerable agreement among observers, any distillation into a summary list will result in a loss of much of the subtlety of verbal descriptions. For this reason, the data of this experiment were considered only as indicants of potentially important processes.

The data of Table 7 are segregated into three major groupings: background characteristics, target characteristics and subject expectations. Under background characteristics the presence and characteristics of clutter objects were most often mentioned. Subjects noted that when a tree or other object was of a similar size, luminance, and shape as the target it attracted their attention and required considerable examination before a target/non-target decision could be made. The magnitude of the confusion was noted to be greater when clutter was in close proximity to the target. A number of subjects observed that when clutter objects formed a pattern the detection of a target within that pattern was more difficult. A factor mentioned most commonly with the low complexity scenes was the texture of open areas.

Except for target-to-background contrast, all of the target characteristics mentioned related to shape and internal detail. Most observers felt that sharp angular outlines and easily discernible features were facilitating target characteristics. The orientation of the target was thought to be important with a preference for side views. All targets had shadows appropriate for their respective scenes, however, several subjects felt that shadows made detection more difficult because of their interactive effect on the expected target outline.

Subject expectation also appeared to have a fair influence on the ease or difficulty of a particular scene. Subjects noted an expectation for targets to be on or near roads and not in trees. The absence of vehicle tracks was occasionally mentioned as a factor making detection more difficult. The fact that trees and rocks were not likely to be found in the middle of a cultivated field was also mentioned as a factor aiding detection.

TABLE 7. SUBJECTIVE FACTORS

- Background Characteristics
  - Texture
  - Clutter Objects
    - Proximity
    - Size
    - Contrast
    - Pattern
- Target Characteristics
  - Distinctiveness of Outline
  - Visibility of Features
  - Line Structure
  - Target Orientation
  - Internal Modulation
  - Contrast
- Subject Expectation

The subjective factors metioned tend to support the importance of form perception considerations in the detection of a target in a relistic background. They also provide further evidence for an interaction between the background and the target within that background. Finally, they suggest a few quantifiable background characteristics that might be a first step in the development of a model that includes the influence of realistic scenes. The effectiveness of these potential predictor metrics were examined in Experiment 3.

EXPERIMENT 3

Based on the results of Experiments 1 and 2, a number of quantitative descriptors of scene complexity could be identified. The purpose of this third experiment was to determine whether a weighted combination of these descriptors would predict the time required to detect a target in a realistic scene.

Experiment 2 identified the amount of similar clutter, the proximity of the target to clutter, the pattern of clutter and the expectation to find targets on roads as subjective factors influencing detection performance. The amount of similar clutter was operationally defined as the number of objects between 0.5 and 2.0 times the target width in size with a brightness subjectively similar to that of the target. Proximity was measured as the number of target diameters separating the target and the closest clutter. Clutter pattern was crudely assessed by counting the number of clusters of clutter such as a group of trees. Two metrics related to subject expectation were included. These were the percent of the scene occupied by roads, which is approximately equivalent to the length of road because all roads were nearly the same width, and the percent of the scene that might contain a target. This latter metric reflected the expectation that targets would not be located in dense woods or lakes.

Experiment 1 demonstrated the influence of target-to-background contrast and this factor was included. One additional metric was also included, although, it had not been mentioned by subjects in Experiment 2. This was the number of man-made objects in the scene. These seven quantitative metrics are summarized in Table 8 which gives the mean, standard deviation and minimum and maximum values over the 25 test images used.

The same three target types investigated in Experiment 1 were also included in the 25 images of the present study, allowing a second examination of the effect of target type on detection performance. This was accomplished by arbitrarily coding the three types as 1, 2 and 3 and including target type as an eighth predictor variable in the regression analysis. The arbitrary coding of a qualitative variable causes no mathematical problem within the regression analysis. However, should such a coded qualitative variable be

TABLE 8. SCENE METRICS

Predictor Metric	Mean Value	Standard Deviation	Minimum Value	Maximum Value
Number of Clutter Objects (O)	17.8	12.9	2	45
Target-to-Clutter Proximity (P)	2.1	2.7	0	12.0
Number of Cluster Clutter Objects (B)	3.0	1.6	1	6.0
Percent Roads (R)	7.7	5.1	0	20.2
Target-to-Background Contrast (C)	62.2%	17.1%	25%	95.0%
Percent Usable Area (A)	71.1	19.1	33	95.0
Number of Man-Made Objects (M)	1.8	4.1	0	18

a major predictor, some difficulty in interpretation may result. The problem lies in the ill-defined nature of a target type of, say, 1.5 which would be something between an APC and a tank. Because of such ambiguity, the finding of a reliable target type effect could be interpreted only as an indication of the need for quantitative metrics describing the target.

The seven scene metrics given in Table 8 plus coded target type were used as predictor variables to obtain a multiple regression equation relating the weighted combination of the predictors to observed target detection time.

METHOD

The apparatus and procedures were identical to those of Experiment 1. The stimuli were prepared in the same manner as Experiment 1, however, rather than preparing a high and a low complexity group of images the complexity varied over all levels of judged complexity. Similarly, the proximity of the target to clutter and the target-to-background contrast values varied over the ranges given in Table 8.

Ten members of the Technical Staff at Hughes Aircraft Company participated in the experiment. None had served in either of the previous two experiments. Each subject received five practice and 25 test trials using the procedures of Experiment 1. The order of image presentation was randomly determined for each subject.

RESULTS AND DISCUSSION

The time required to detect the target for each of the 10 subjects and 25 scenes along with the predictor metrics for the 25 scenes were initially subjected to a multiple linear stepwise regression analysis which described the data with an equation of the form:

$$Y = \beta_o + \sum_{i=1}^{K} \beta_i X_i$$

where: Y was the time to detect, the X's were the scene metrics, and the β's were the calculated weighting factors. This equation allowed the assessment of the linear effects of each predictor metric. The resulting best fit linear equation relating the metrics to the time required for detection had a multiple regression coefficient of 0.45 which indicated that only 21 percent of the observed variance in the data was accounted for by the equation. Thus, even though the regression equation was reliable at the 0.01 level of significance, only a small proportion of the subjects' behavior was being predicted.

The relatively poor predictive ability of the linear regression suggests either that the selected metrics were inadequate as predictors or that a higher-order equation was needed to describe the data. The latter possibility was explored by fitting the complete second-order equation presented in the introduction to Experiment 2. This second-order equation in eight variables adds 8 squared and 28 linear-by-linear interaction terms to the 8 linear terms of the first-order equation. The total of 47 terms allows a parabolic description of the observed data.

The second-order equation was fit to the observed data using a polynomial stepwise regression analysis. The results indicated that the most important 19 terms produced a multiple regression coefficient of 0.81. These terms accounted for 66 percent of the variance and the equation was reliable at the 0.001 level of confidence. The 19 terms in order of predictive importance, their regression coefficient, and the cumulative variance described are given in Table 9.

Substituting actual scene metric values for the symbols in the left most column of Table 9, performing the appropriate operation, multiplying by the given weighting coefficient, and summing all weighted terms will result in a predicted time to detect. For example, if the number of clutter objects were 50, then O^2 would equal $50^2 = 2500$ and the weighted term would be $-0.0046 \times 2500 = -11.50$. This value summed with the 18 other similarly calculated values plus the intercept would yield a predicted time to detect. It should be noted that all 20 of the terms given in Table 9 need to be included in the calculation of the predicted performance. If an equation with fewer terms is desired, new weighting coefficients need to be calculated which reflect such a reduction.

TABLE 9. SECOND-ORDER STEPWISE REGRESSION ANALYSIS SUMMARY

Term	Coefficient	Cumulative Variance Described, %
Intercept	- 99.00493	
O	11.95191	12.25
O^2	- 0.00460	16.56
AO	- 0.12818	25.15
OC	- 0.14266	29.89
C^2	- 0.04952	33.89
OR	0.46698	36.17
R^2	- 0.31385	41.02
AM	0.16096	43.75
OB	1.52020	46.72
BR	0.48217	49.01
B^2	16.76317	54.97
B	-135.01836	59.27
DM	0.13602	61.36
C	7 50094	61.95
M^2	- 2.17510	63.25
A	0.78235	64.52
MR	- 1.76021	65.16
CM	0.52251	65.65
MP	- 4.31651	66.25

KEY:

O	Number of Clutter Objects
P	Target-to-Clutter Proximity
B	Number of Cluster Clutter Objects
R	Percent Roads
C	Target-to-Background Contrast (%)
A	Percent Usable Area
M	Number of Man-Made Objects

CONCLUSIONS – EXPERIMENTS 1 TO 3

The data from Experiments 1 and 3 provide dramatic examples of the importance of the background scene as a determinant of tactical target detection performance. Any model that fails to account for the influences of realistic background scenes can be expected to grossly overestimate observer performance. Detection of a target in even a simple real background requires 11 times as long for detection as the same target in a uniform background. With high complexity backgrounds the factor becomes 24. Differences of this magnitude make it apparent that the effects of realistic scene characteristics must be understood and modeled.

The absence of an interaction between resolution and scene complexity demonstrated by the data of Experiment 1 suggests that the inclusion of scene characteristics in an advanced model may not necessitate the computation of a large number of scene complexity by sensor characteristic interaction terms. This conclusion is only tentative because the present research examined only two levels of resolution and did not include many of the other potentially important sensor parameters.

The most important results of the present research are those that suggest an alternative model form based on two or more component processes. The data imply one process that might be associated with search for candidate target objects and a second process which follows and includes a determination of whether the object is actually a target. Among the advantages of a component process model is the ability to separate a complex task into behaviorally meaningful parts each of which can be evaluated separately. This may allow considerable economy in data collection because the scene characteristics that influence one process are very likely quite different from those which influence the second process. Rather than including a complete set of metrics and attempting to fit a set of parameters for the entire process, two smaller sub-sets of scene characteristics are used to fit the two distinct sets of parameters. The iterative refinement of the parameters can also be separated because it is possible to identify which component is causing a lack-of-fit in the combined model.

Perhaps the greatest advantage of a component model based on behavioral aspects is the ability to identify which part of the task is causing the most difficulty for the observer, thereby providing important information for the system designer. For example, if it is found that one portion of the task produces a significant decrement in particular situations, the system designer can use this data to modify the existing system or to design the next generation system.

The advantages given for the two component form of the model make it a prime candidate for expanding a detection model to include the influences of realistic terrain features, provided the assumptions made in the derivation of the model are valid. Although each of the assumptions appear reasonable based on existing data and a logical analysis of the detection process, a necessary first step will be the verification of the assumptions and a verification of the model form. For example, it is not totally clear how the two component models presented previously will account for the case where the observer finds a candidate object, examines it in detail, and rejects it as a target only to decide after further search that it was actually a target.

The large number of squared and linear-by-linear terms in Table 9 makes it apparent why a linear equation failed to provide an adequate description of the data. The squared terms add curvature and the linear-by-linear terms represent interactions among various metrics so that the effect of one scene metric depends upon the level of another metric. For example, the effect of number of clutter objects is to increase the time to detect. However, the presence of the A0 term means that the magnitude of the change depends upon the value of A, the useful area of the scene. Larger values of A reduce the influence of the clutter perhaps suggesting that the number of clutter objects per unit useful area may be an alternative metric.

The good prediction obtained with a few simple scene metrics indicates considerable potential for this approach as a method for fitting the parameters of a model which includes scene characteristics. A careful consideration of the relationships among metrics revealed by the regression analysis may provide a means for identifying metrics that are even more efficient and effective than those included in the present study.

Because target type was also included as a predictor variable in the regression analysis a second examination of its influence could be made. Table 9 shows that none of the 19 most important terms included target type indicating virtually no predictive value. The absence of a target type effect in this experiment, which more definitely separated scene and target/scene interaction effects from pure target effects, supports the conclusion of Experiment 1. What had appeared as an effect of target type was actually more an effect of the target within the scene.

The success of the regression approach used on Experiment 3 indicates that this approach is a viable means for fitting the parameters of a model. The principle difficulty with the method is its inefficiency. If the levels of the various predictor metrics are determined by a random sampling, then considerable data are required to assess the weighting factors with any precision. The larger the number of variables and the greater the order of the equation the more serious the problem becomes.

The magnitude of the problem can be reduced if, by careful analysis and the use of screening experiments, the number of variables can be reduced. A first step in this direction would make use of the regression equation obtained from the data of Experiment 3 to identify alternative metrics that are likely to be more descriptive and efficient. This, combined with the use of economical experimental designs such as central-composite designs, will provide an efficient means for fitting the parameters of a model.

EXPERIMENT 4

BACKGROUND

The previous three experiments examined target detection performance when the target was within the 8.89 by 8.89 grad (8 by 8 degree) field-of-view. In most situations, however, the task of positioning the sensor can be a major portion of the detection process. This will be particularly true if the target location is unknown or only poorly specified. In such situations, questions arise concerning the best method of performing the search and the optimum field-of-view or combination of fields-of-view.

Field-of-View

A wide field-of-view, as shown in Figure 16, allows all or a large part of the search area to be examined at one time and the total context perceived. This makes it possible for the operator to use *a priori* information concerning probable target locations to aid him in his search. For example, vehicles are often found on or near roads and an initial search of these features in a scene would be a potentially successful strategy.

The principal disadvantage of a wide field-of-view results from the limited resolution of the sensor system and, to a lesser extent, the observer's visual system. With increased field-of-view comes decreased

Figure 16. Sensor display with a 22.22 grad (20 degree) field-of-view.

displayed target size so that a sufficiently large fields-of-view the target is little more than a speck on the display. For example, a truck viewed from a 910 meter (3000 feet) altitude and a slant range of 2 kilometers (1. 24 miles) with a 22. 22 grad (20 degree) field-of-view will be only 0. 68 percent of the display in height. If the display is a standard television with 480 active lines only 3. 3 of these will be across the target. If the range is increased to 5 kilometers (3. 11 miles), the values become 0. 22 percent and 1. 0 line. In both cases, detection might be possible if the target-to-background contrast were high and the background were very simple. However, with realistic values of contrast and background scenes typical of actual terrain, detection, at best, could be extremely difficult. To obtain an acceptable displayed target image size at the ranges of interest requires that a smaller field-of-view, as shown in Figures 17 to 19, be used.

These observations are substantiated by previous research on the effects of field-of-view. For pre-briefed targets, where the task is primarily one of identifying the appropriate contextual features that identify the target location, field-of-view has little effect (Ozkaptan, Ohmant, Bergert, and McGee, 1968) or larger fields-of-view are superior (Dale, Knudsen, Hawley, Jeffrey, Luninger, and Bliss, 1968). This is because the actual "seeing" of the target is not necessary, rather a correlation of the briefing material and

Figure 17. Sensor display with a 12. 22 grad (11 degree) field-of-view.

Figure 18. Sensor display with a 6.11 grad (5.5 degree) field-of-view.

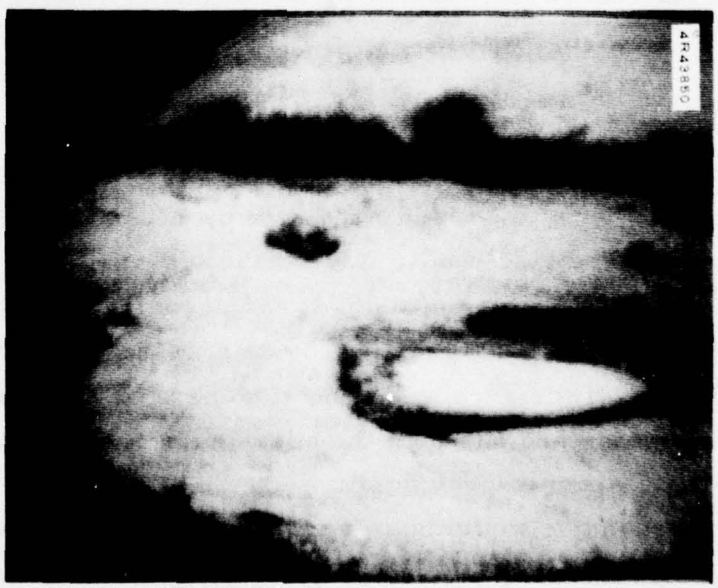


Figure 19. Sensor display with a 3.05 grad (2.75 degree) field-of-view.

the viewed scene must be accomplished. For this task contextual cues are most important and these are best appreciated with a large field-of-view.

Similar results have been obtained in this laboratory in a task requiring operators to find pre-briefed targets with and without the aid of eight-to-one continuous zoom. Both conditions resulted in similar probability of detection, however, the condition with zoom capability required more time for its completion probably because of the time required for the zoom process. Clearly the increased target size and definition resulting from the smaller field-of-view was of little advantage in detecting pre-briefed targets.

For targets that are not pre-briefed, the situation is very different. Decreased fields-of-view improve detection performance, at least down to some optimum field-of-view (Bergert and Fowler, 1970; Fretag and Jones, 1973; Ozkapton, Ohmant, Bergert, and McGee, 1968). In this case, the increased target size and definition are critical to the detection task. Fretag and Jones, (1973) found improved performance as field-of-view decreased down to 5.56 grads (5 degrees) but little improvement below that value, indicating that once a critical target definition threshold or size is reached little further improvement results.

Based on the existing literature, it is apparent that a small field-of-view is desirable provided the target is within that view. If the uncertainty in the target position is high then the desire for a small field-of-view must be mediated by the search method available for getting the target within the field-of-view.

Search Method

For the purposes of the present experiment it was assumed that the target location was known to be within a 22.22 by 22.22 grad (20 by 20 degree) forward view area. The task was to find all tactical targets within this area. There are at least three available methods for accomplishing this task.

The first technique would be to provide a 22.22 grad (20 degree) field-of-view display of the area and allow the operator to search the area with his eyes. This method has both the advantages and disadvantages of a wide field-of-view discussed previously.

A second technique would be to reduce the fieid-of-view and provide the operator with a control for pointing the sensor within the search area. This method has the advantage of allowing a reduced field-of-view, which produces a larger target image and greater target definition. However, with a reduced field-of-view the operator has fewer contextual cues available and may not be able to effectively examine the entire search area. It is possible for the operator to become disoriented and examine some portions of the search area several times while not examining other areas at all. The severity of this problem may be directly related to the field-of-view being used.

A third search technique can be implemented that reduces the problem of keeping track of the area searched. The system could be made to point the sensor in a programmed search pattern guaranteeing that the entire search area is examined. With this type of programmed search, the probability of target detection should improve to the extent that searching the entire area places a larger number of targets within the field-of-view. At the same time, however, it might be expected that the time required to detect a target would increase because the fixed search pattern cannot take advantage of any knowledge concerning likely target positions. The magnitude of this time penalty can be reduced by providing the operator with control over the rate of the search scan. This would allow the operator to move rapidly over areas that could not possibly contain a target.

The discussion to this point has considered only single field-of-view systems. From the operator's standpoint a multiple field-of-view system may prove to be the optimum. A wide view could be provided to allow orientation and maximize the use of a priori knowledge. When a likely target is detected in the wide view, a narrower view could be selected to perform the actual recognition. Two methods of obtaining the narrow fields-of-view can be contemplated. One method would be to provide continuous selectable zoom and the second would be to provide selectable discrete narrow fields-of-view. Both of these implementations increase the complexity of the sensor system and can be justified only if operator performance is improved sufficiently.

EXPERIMENTAL PLAN

The considerations discussed above led to the selection of nine search modes for examination in the present experiment. These are the six factorial combinations of two search methods — sensor pointing and programmed — and three fixed fields-of-view — 12.22, 6.11, and 3.05 grads (11, 5.5 and 2.75 degrees) plus three special cases. These were free eye search with

22.22 grad (20 degree) field-of-view, sensor pointing with continuous selectable zoom between 22.22 grads (20 degrees) and 3.05 grads (2.75 degrees) and sensor pointing with selectable discrete fields-of-view of 22.22, 12.22, and 3.05 grads (20, 11, and 2.75 degrees). The nine conditions summarized in Table 10 were examined in a target search and detection task with realistic background scenes.

TABLE 10. EXPERIMENTAL CONDITIONS – EXPERIMENT 4

Search Method	Field-of-view, grads					
	22.22	12.22	6.11	3.05	Continuous 22.22-3.05	Discrete 22.22, 12.22, and 3.05
Sensor Pointing		X	X	X	X	X
Programmed Scan		X	X	X		
Free Eye	X					

METHOD

Search Techniques

In the following paragraphs, the display characteristics and control functions for each search technique will be described.

Sensor Pointing Search

Five of the nine modes investigated used a sensor pointing search technique where the operator could, by means of a hand control, select a portion of the search area for display. The deflection of the hand control determined the rate and direction of sensor slew. Target designation was accomplished by slewing the sensor to place the target under a set of fixed crosshairs and depressing a thumb actuated lock-on button.

Three of the sensor pointing modes had a single fixed field-of-view and two had variable field-of-view. In these latter cases, field-of-view was controlled with a thumb switch. In the continuous zoom condition, the field-of-view changed continuously as long as the switch was activated. Forward action commanded reduced field-of-view and reverse action increased field-of-view. For the discrete multiple field-of-view condition, each activation of the switch caused a single field-of-view change. Because the simulation was of a rapid lens change, the display was blanked and the elapsed time counter deactivated during the two seconds required for the simulator zoom lens to change the field-of-view. For example, if the system were in a 22.22 grad (20 degree) field-of-view, forward activation of the thumb switch would cause the display to blank for two seconds while the field-of-view was changed to 12.22 (11 degrees). A second forward switch activation would cause a similar change to the 3.05 grad (2.75 degree) field-of-view. Reverse activation of the switch commanded increased field-of-view in a similar manner.

Programmed Search

For the programmed scan modes, the pattern followed depended upon the field-of-view as shown in Figure 20. It can be seen, with a 12.22 (11 degree) field-of-view only two horizontal scan bars were required while four or eight were necessary with the two smaller fields-of-view. In all cases there was a 10 percent overlay between adjacent bars to preclude the possibility of a target occurring at the boundary of two bars and not being imaged. Two patterns are shown in Figure 20 to indicate the complete path followed. The right most figure shows the path followed after completion of the first scan of the search area.

In these modes the hand control served two control functions. During the scanning sequence forward deflection of the hand control increased the rate of the scan up to a maximum of 5.56 field-of-view grads (5 degrees) per second. Pulling back on the hand control slowed the scan rate down to 0.0 grad per second at a minimum.

The hand control was also used to position the target under the crosshairs for target designation. In these programmed search modes it was important from an experimental point of view that the operator not be able to override the scan pattern and operate in a sensor pointing mode. At the same time, it was important that the scene move under fixed crosshairs as occurred in the other modes. These two objectives were reconciled by holding the image in a memory and moving the television raster to allow positioning of the target under the crosshairs. The operator signaled his intention of

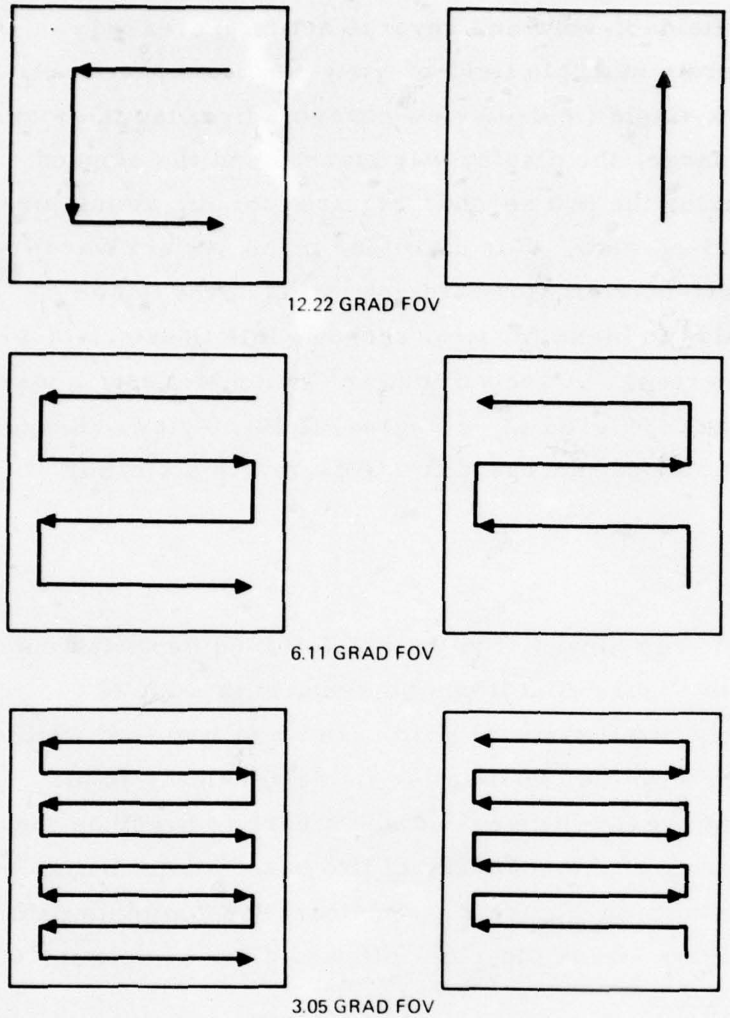


Figure 20. Programmed scan patterns.

designating a target by activating a trigger which held the currently displayed image in memory and changed the function of the hand control to control the image position. With this technique, the designation control task was constant across conditions and the operator could not examine an area beyond that being scanned by the programmed search.

Free Eye Search

The 22.22 grad (20 degree) field-of-view, free eye search condition was essentially a sensor pointing mode where the entire search area could be viewed at one time. Designation was accomplished as in the sensor pointing modes. Because the stimulus photographs were masked to exclude all but the search area, no new terrain was visible during target designation.

Stimuli

The ten test and three training images used in this study were prepared in the same manner as Experiment 1. Each was a 22. 22 by 22. 22 grad (20 by 20 degree) forward oblique view which always included ranges from 1 to 3 kilometers (0. 62 to 1. 86 miles). They were selected to be of moderate complexity and each contained between three and seven embedded targets. All targets were at ranges between 1. 5 and 2. 5 kilometers (0. 93 to 1. 55 miles and had target-to-background contrasts between 0. 6 and 1. 1.

Apparatus

The computer-controlled sensor display simulator shown in Figure 21 was used to present the stimulus images to experimental subjects. Display video originated with a television camera fitted with a servo driven 20 to 1 zoom lens focused on a film transparency mounted in a servo driven platform. The two directions of platform translatory motion were used to simulate the azimuth and elevation sensor pointing degrees of freedom. The zoom lens provided a means of varying field-of-view.

The film transparency was illuminated with a strobe light which was flashed, at 15 frames per second, in synchronization with the television camera scan. The strobe light was used to reduce the image smear that normally occurs with a vidicon camera when the scene is moving. Because the strobe provided very short, high intensity light, the camera only "saw" stationary scenes and the smear was reduced. By reducing smear in this way, the displayed scene had a tendency to jump in discrete steps rather than move smoothly. This characteristic, however, was less objectionable than a smeared image.

An automatic gain control (AGC) circuit on the camera video output assured that the full dynamic range of the display was utilized. This circuit adjusted the video signal so that the minimum signal was displayed as black and the peak signal as maximum white.

After the AGC circuit, the analog video was converted to digital form for storage in a digital scan converter memory. The observer's display was refreshed from this memory at 30 frames per second by reading the memory and converting back to an analog voltage. The memory also provided the means for shifting the displayed video to allow target designation in the

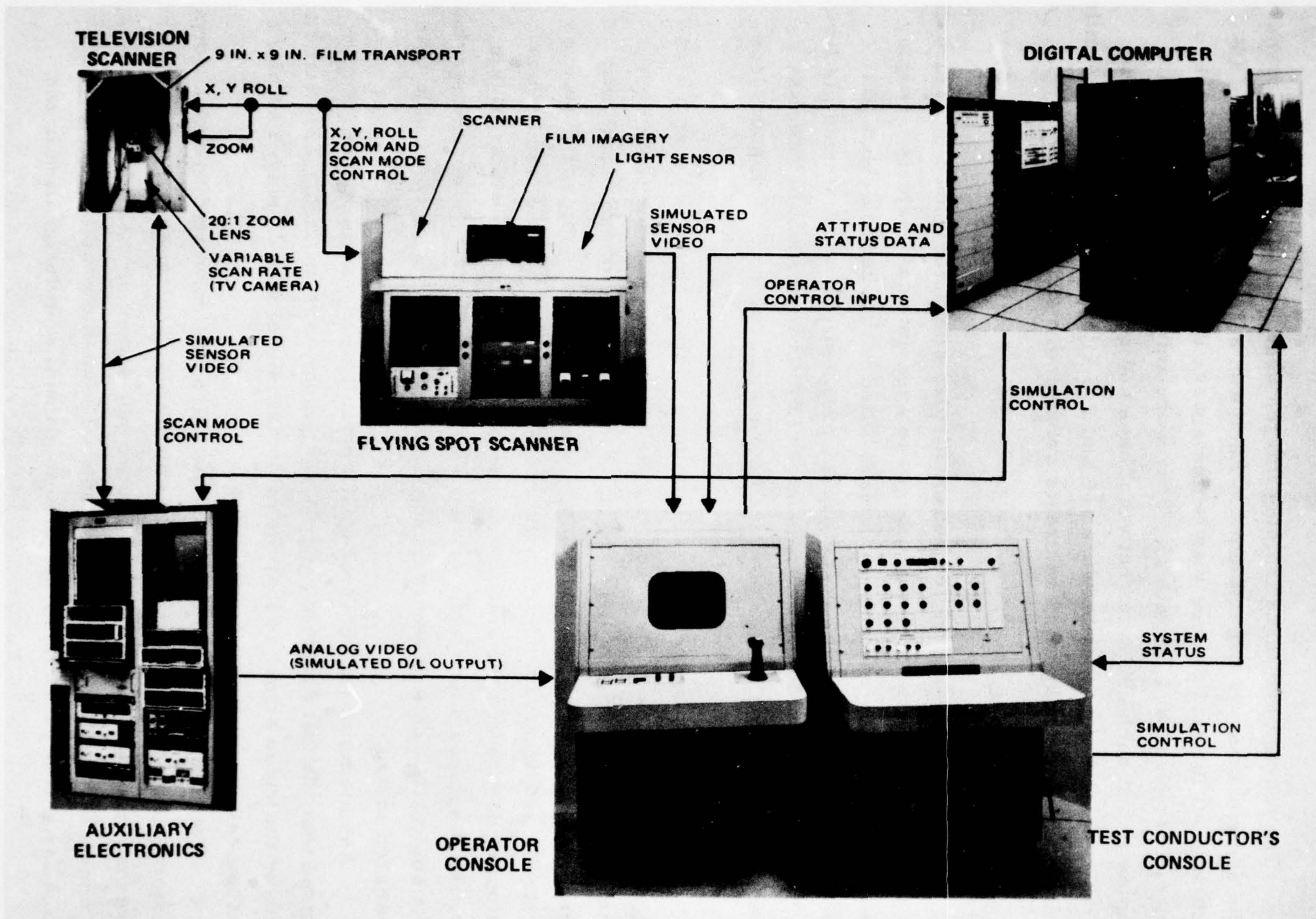


Figure 21. Computer-controlled sensor display simulator.

programmed search modes. This was accomplished by adjusting the memory address from which the image information was to be read. To move the displayed image 100 elements to the left, 100 was added to each horizontal memory address so that the image information presented on the left most position of the display was taken from the 100th memory location.

The subject viewed a 20.3 cm. (8 inch) high standard television display from a distance of approximately 61 cm. (24 inches). The hand control previously described was located for easy activation with the right hand. Next to the subject's display console was the experimenter's station as shown in Figure 21. During experimentation the two positions were separated by a cardboard partition.

The experimenter's console provided various switches and controls which allowed the experimenter to direct the computer simulation. The computer used was a Xerox Data Systems Sigma 5 which was interfaced to the simulation hardware described above. A real-time computer program was written to provide the desired search modes, automatic parameter setting and automatic data collection. The latter capability was achieved by storing the location of each target in the computer and comparing these locations with the position of the sensor when the subject designated a target. This comparison allowed a computation of designation accuracy which was recorded along with the elapsed trial time when the designation was made. All data, including subject number, search mode, date and time, and field-of-view at designation were recorded on magnetic tape and in printed form.

Research Design

A different group of five subjects was used for each of the nine search modes. The resulting between-subjects design was selected to avoid difficulties with changed control operation and search strategy when transferring from one mode to another. With the design used, any given subject was required to learn only a single set of control operations and adopt a single search strategy. Each subject viewed 10 image scenes each with from three to seven targets. In the 10 trials a total of 44 targets were available for detection and designation.

Subjects

Forty-five volunteer subjects participated in the experiment and all were members of the technical staff at Hughes Aircraft Company, Culver City, California. Each was randomly assigned to one of the nine search mode conditions and was unaware of the characteristics of the other eight modes.

Procedures

At the beginning of the experimental session, the subject was given a set of standardized written instructions corresponding to the search mode he was to experience. When the subject had read the instructions the experimenter verbally summarized their salient points and answered any questions from the subject. A 20 minute training session followed which allowed the subject to become familiar with the image and control characteristics.

The 10 experimental trials were begun when the subject was judged to be competent with the task and procedures. Each trial was started with a blank display. When the subject indicated he was ready, the experimenter started the trial which unblanked the display to reveal the image scene. The subject searched for targets for a total trial length of 4 minutes at the end of which the display again blanked. Approximately 2 minutes elapsed between trials, during which time the experimenter changed film images and the subject relaxed. The entire session required approximately 1 hour and 15 minutes.

RESULTS AND DISCUSSION

Performance Measures

Four performance measures were taken to assess the effects of sensor field-of-view and search strategy: 1) number of correct target detections (N_C), 2) number of missed targets (errors of omission, N_O), 3) number of false target detections (errors of invention, N_I), and 4) target detection time (T_D). These measures were used to derive four indices of operator target detection performance: 1) probability of correct target detection [N_C (N_C + N_O)], 2) average time to correctly detect targets (T_D), 3) number of false detections (N_I), and 4) a composite performance score which combined the effects of number of correct detections, number of missed targets and number of false target detections. The equation for the composite score was $N_C/(N_C + N_O + N_I)$.

The correct target detection times were calculated by dividing the total search time by the number of correct target detections. In the event that all targets on a given scene were not detected, the search time was the trial length of 240 seconds. This computational technique resulted in an inflated detection time if a scene contained a difficult target that was not usually detected. For example, if four targets were present and three were found in 90 seconds and the fourth was missed entirely, the computational method used resulted in an average time of 80 seconds even though the first three were found in an average of 30 seconds each. This apparent inequity is another manifestation of the common problem of assigning a latency score when a target is missed.

In the following sections, the effects of field-of-view, search strategy (programmed search versus manual sensor pointing), and zoom and multiple discrete fields-of-view are presented in terms of the four measures of operator performance.

Field-of-View

Figure 22 shows the effect of field-of-view on the probability of correct target detection for 3.05, 6.11 and 12.22 grad (2.75, 5.5 and 11 degree)

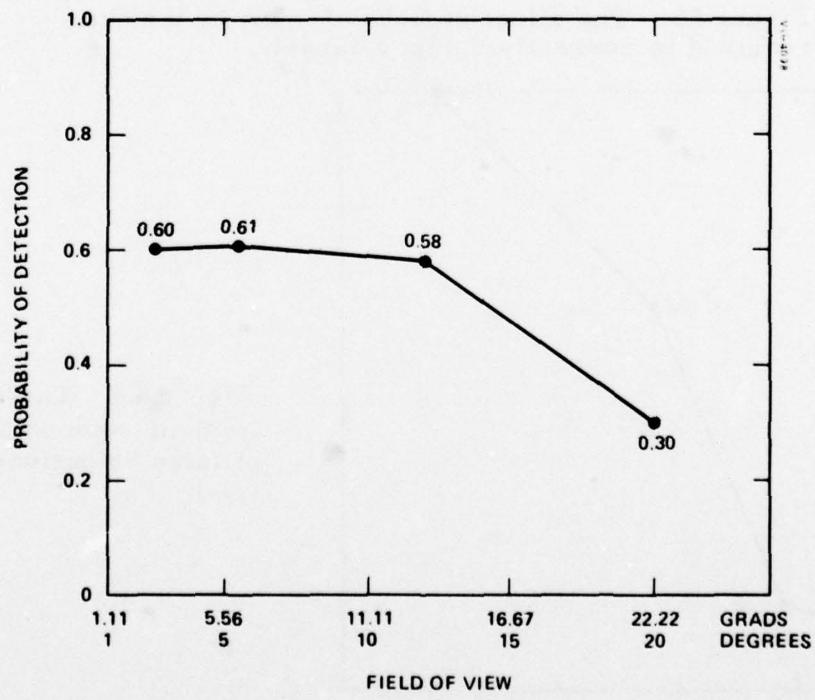


Figure 22. The effect of field-of-view on the probability of correctly detecting a target.

fields-of-view withmanual sensor pointing and 22. 22 grad (20 degree) free eye search conditions. Figures 23 to 25 present similar plots for time to

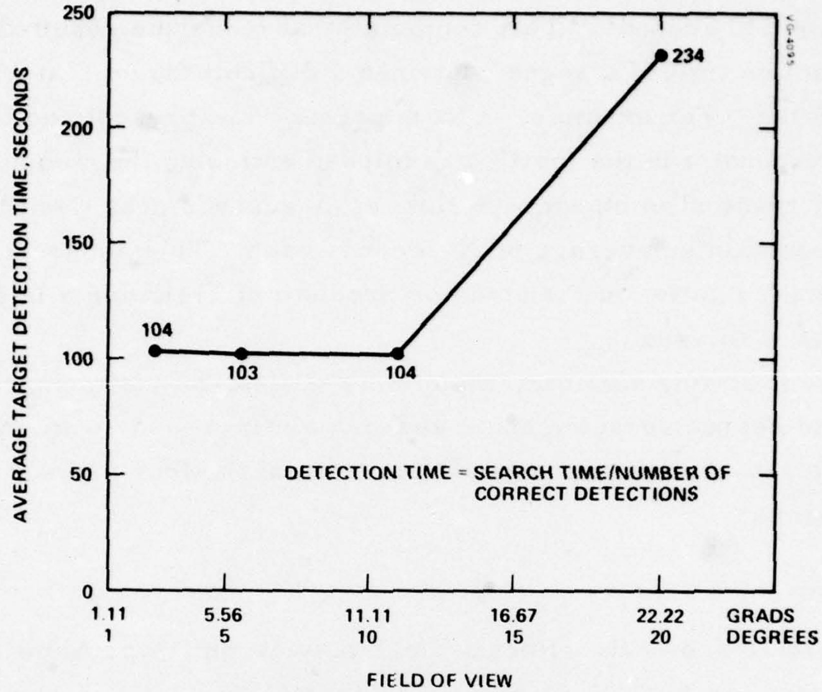


Figure 23. The effect of field-of-view on the time required to correctly detect a target.

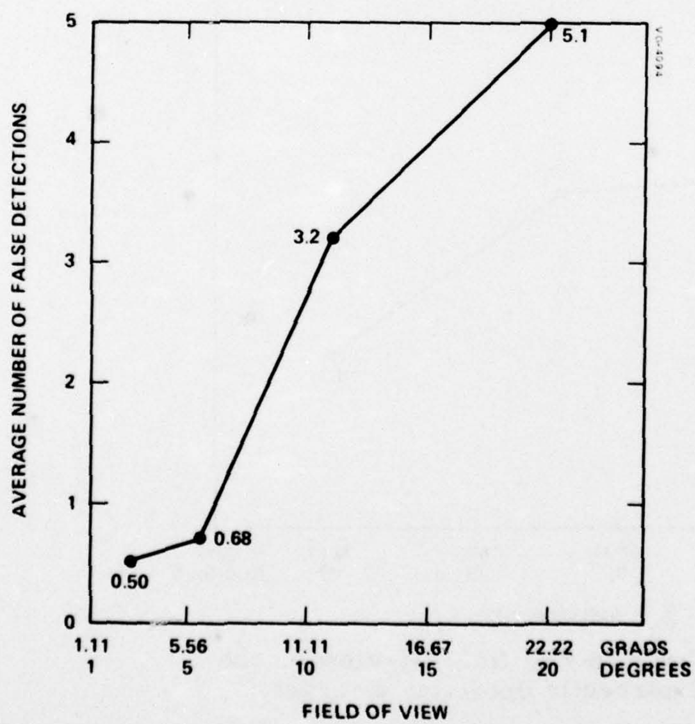


Figure 24. The effect of field-of-view on the number of false detections per trial.

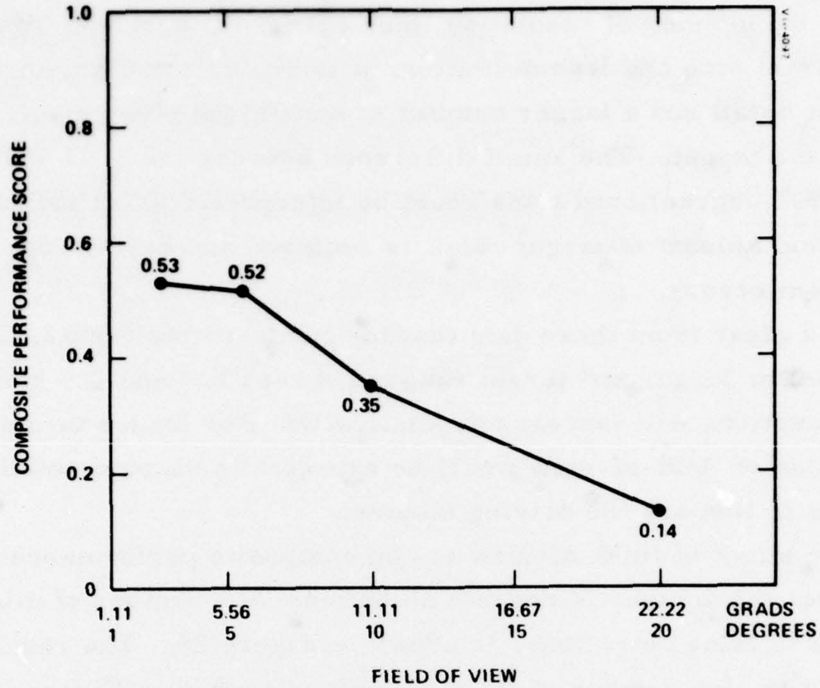


Figure 25. The effect of field-of-view on the composite performance score.

detect, number of false detections and composite performance scores, respectively. From Figure 22 it is apparent that a major decrement in performance occurs in the 22. 22 grad (20 degree) field-of-view condition compared to the other three fields-of-view. The probability of correct target detection was half as large for the 22. 22 grad (20 degree) field-of-view. Analyses of variance indicated that this result was reliable beyond the 0. 01 level. There were virtually no differences between the 3. 05, 6. 11 and 12. 22 grad (2. 75, 5. 5 and 11 degree) field-of-view conditions.

The average time to correct target detection shown in Figure 23 was also reliable at the 0. 01 level of significance. However, because in virtually no case were all of the targets in a given scene detected, no new information is provided by the latency data. If one or more targets in a scene were not detected, the total search time was taken as the trial length. Thus, in most cases the total search time was a constant value and average search time a function only of the number of correct target detections.

The number of false target detections per target scene as a function of field-of-view is shown in Figure 24 and indicates that the number of false detections reliably ($p < 0.01$) increase with field-of-view. This finding is predictable, because as field-of-view increases the displayed size of the target and the number of resolution lines across it decrease. With a smaller target size and less definition, it is difficult to discriminate necessary target detail and a larger number of non-target objects will meet the criterion of a target. The small difference between the 3.05 and 6.11 grad (2.75 and 5.5 degree) conditions could be interpreted as an indication that some critical amount of target detail is required and beyond that little or no improvement occurs.

It is clear from these data that for fields-of-view of 12.22 grads (11 degrees) or larger and target ranges between 1.5 and 2.5 km, the number of false detections will increase dramatically. For longer target ranges, the critical value of field-of-view would be expected to become smaller if target size and definition are the driving factors.

The effect of field-of-view on the composite performance score, which incorporates the number of correct detections, the number of misses, and the number of false detections, is shown in Figure 25. The results are similar to those for the number of false detections plotted in Figure 18 and the interpretations considered previously are also relevant here.

The search for, and detection of, unbriefed tactical targets is influenced by sensor field-of-view because of its influence on the number of resolution lines across the target (target definition), the displayed target subtense, and the amount of terrain displayed. Experiment 1, along with many other experiments, demonstrated the major effect resolution can have on tactical target detection, recognition, and identification. In this study sensor resolution was fixed at 480 active television lines and targets occurred at ranges of 1.5, 2.0 and 2.5 km. Table 11 shows the number of resolution

TABLE 11. TELEVISION LINES ACROSS THE HEIGHT OF A TARGET FOR VARIOUS RANGES AND FIELDS-OF-VIEW

	Field-of-View, grads (degrees)			
Range, km	3.05 (2.75)	6.11 (5.50)	12.22 (11.00)	22.22 (20.00)
1.5	36.4	18.2	9.1	5.0
2.0	23.8	11.9	6.0	3.3
2.5	17.7	8.9	4.4	2.4

lines that were across the height of the targets at the four fields-of-view and and three target ranges. At the two largest fields-of-view, the number of TV lines across the targets ranged from 2.4 to 9.1. Based on the performance obtained with the four fields-of-view, one can speculate that a minimum of 6 to 9 TV lines across a target's height are required to achieve good operator tactical target recognition performance, where recognition is defined as being able to correctly classify an object as a military vehicle, but not necessarily identify the object as, for example, a tank.

For a fixed viewing distance, a decrease in field-of-view will result in an increase in displayed target subtense. As demonstrated in Experiment 1, target subtense can reduce the effective resolution if the target is too small for the observer to resolve all of the information present. Extraction of information concerning the shape of the target can be of considerable importance with high scene complexity. In the absence of high complexity backgrounds, where detection can be accomplished on the basis of luminance differences, the psychophysically demonstrated interaction between target subtense and contrast may influence the selection of field-of-view.

Field-of-view also determines the amount of terrain that will be displayed. A wide field-of-view will display many more clutter objects and may, therefore, be of higher complexity than a smaller portion of the same area. On the other hand, the increased visibility of small terrain features resulting from a narrow field-of-view may also provide a considerable number of clutter objects. If all of a search area is examined with a narrow field-of-view, it is reasonable to assume that the total number of clutter objects will be higher than if that same area were examined with a single wide field-of-view.

The relationship between field-of-view and scene complexity was not examined in the present study so that the magnitude of this factor is not known for the case of target search. However, if the data of Experiment 1 are indicative, the effect will be large. In the absence of knowledge concerning the effect of scene complexity on search, the selection of a particular field-of-view for the search and detection of unbriefed tactical targets should be based on the range at which targets are to be detected, the resolution of the sensor, and on a target definition criterion. The results of this experiment indicate the definition criterion is between six and nine TV lines, provided the target subtense is sufficient for the observer to resolve these lines.

Search Strategy

Six of the nine experimental conditions investigated represented the factorial combinations of two search strategies and three fields-of-view. A separate analysis of these conditions allowed an examination of the relative advantages of programmed search versus sensor pointing search, field-of-view, and their interaction. Analyses of variance on each of the four performance measures indicated that only the composite performance score was reliably ($p < 0.05$) affected by search strategy and field-of-view and that in no case was the interaction reliable.

Figure 26 presents the effect of search strategy and field-of-view on the composite performance score. As can be seen, the manual sensor pointing search method is superior to the programmed search for all fields-of-view. The difference is quite small at the 12.22 grad (11 degree) field-of-view and substantially larger at the other fields-of-view. Although this interaction was not reliable, it is not surprising that performance is nearly the same at the wide field-of-view. Because so much of the search area is viewed at one time the method for pointing the sensor should make little difference.

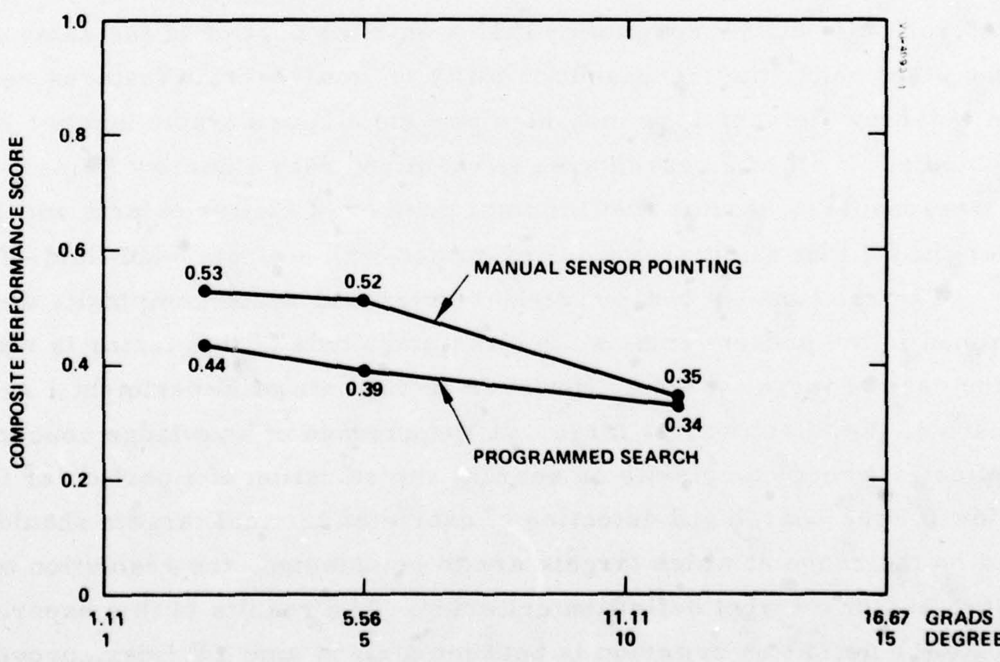


Figure 26. The effect of search strategy and field-of-view on the composite performance score.

The overall superiority of the manual sensor search was not anticipated. It has been expected that the programmed search would improve the probability of detection at the expense of detection time. This was not found; the manual search was as good or better than the programmed search on all measures. A number of factors may have contributed to the superiority of manual search. The manual sensor pointing mode always began a trial pointed at the center of the stimulus image which corresponded to a range of 2 to 3 km. From that starting point, movement of the sensor in any direction had a high probability of bringing a target into view. With the programmed search the scan always began at maximum range and some time was required before a target could possibly be detected. The length of the delay was dependent upon the field-of-view with smaller fields-of-view requiring longer times. An examination of the times required for the first target detection confirms the above.

In the programmed search condition, the subject could stop the scan but he could not reverse it. If a target passed before the subject could stop the sensor, a long wait for the scan to return to that position was required for a second opportunity to designate the target. It is likely that a subject would become reckless and tend to designate everything to maximize the probability of target detection. Such behavior would increase the number of false detections and reduce the composite performance score. Examination of the false detection data supports this supposition for the two larger fields-of-view but not for the narrow field-of-view. This appears to be the case because in the manual sensor pointing and narrow field-of-view condition several subjects interpreted clutter objects at long range as targets and made an unusually large number of false detections.

The stimuli used in the experiment had backgrounds that extended only to the 22.22 by 22.22 grad (20 by 20 degree) search area beyond which was a black mask. In the manual mode it is possible that subjects used the edge of the search area to reduce the uncertainty concerning the location of the sensor. If this were the case, the frequency with which the subjects examined an area previously searched would have been reduced.

Programmed sensor search will result in shorter search time if the scan begins at a range with a high target probability. For example, the 3.05 grad (2.75 degree) programmed search scanned two bars at far range where no targets were located. Examination of the strip chart recordings made during the experiment, reveals that on the average 52.5 seconds were spent on these two bars. If this time is taken into account, the time to first correct detection is reduced from 85.7 to 33.2 seconds. Comparing this value with the 60.6 second average time to first detection for this narrow field-of-view manual sensor pointing mode shows that the programmed search nearly halved the search time.

Perhaps the optimum strategy may be to combine programmed and sensor pointing search. In this mode, the system would perform a programmed search until the operator desired to take command and direct the sensor pointing. Once the manual pointing were complete, the system would resume the programmed search.

Multiple Field-of-View Search

Two conditions in the experiment examined multiple fields-of-view to determine if these would substantially increase target detection performance. In one of these modes continuous zoom was available to the operator so that he could select any field-of-view between 22.22 grads (20 degrees) and 3.05 grads (2.75 degrees). In the second multiple field-of-view condition three discrete fields-of-view of 22.22, 12.22, and 3.05 grads (20, 11, and 2.75 degrees) could be selected by the observer. The composite performance score for these two modes are plotted in Figure 27 along with the scores for the other seven modes. An analysis of variance on the nine modes indicated reliable ($p \leq 0.01$) difference on all of the performance measures. However, a Newman-Keuls post hoc analysis revealed no reliable differences between either of the multiple field-of-view conditions. Continuously selectable field-of-view was no different than discrete step field-of-view.

For the composite performance score, the multiple field-of-view conditions were reliably ($p \leq 0.05$) superior to free eye search, the 12.22 grad (11 degree) field-of-view manual search, and the 12.22 and 6.11 grad (11 and 5.5 degrees) field-of-view programmed search conditions. As can be seen in Figure 27, except for the free eye search, these differences are relatively small and probably do not justify the additional cost of implementing the larger number of fields-of-view.

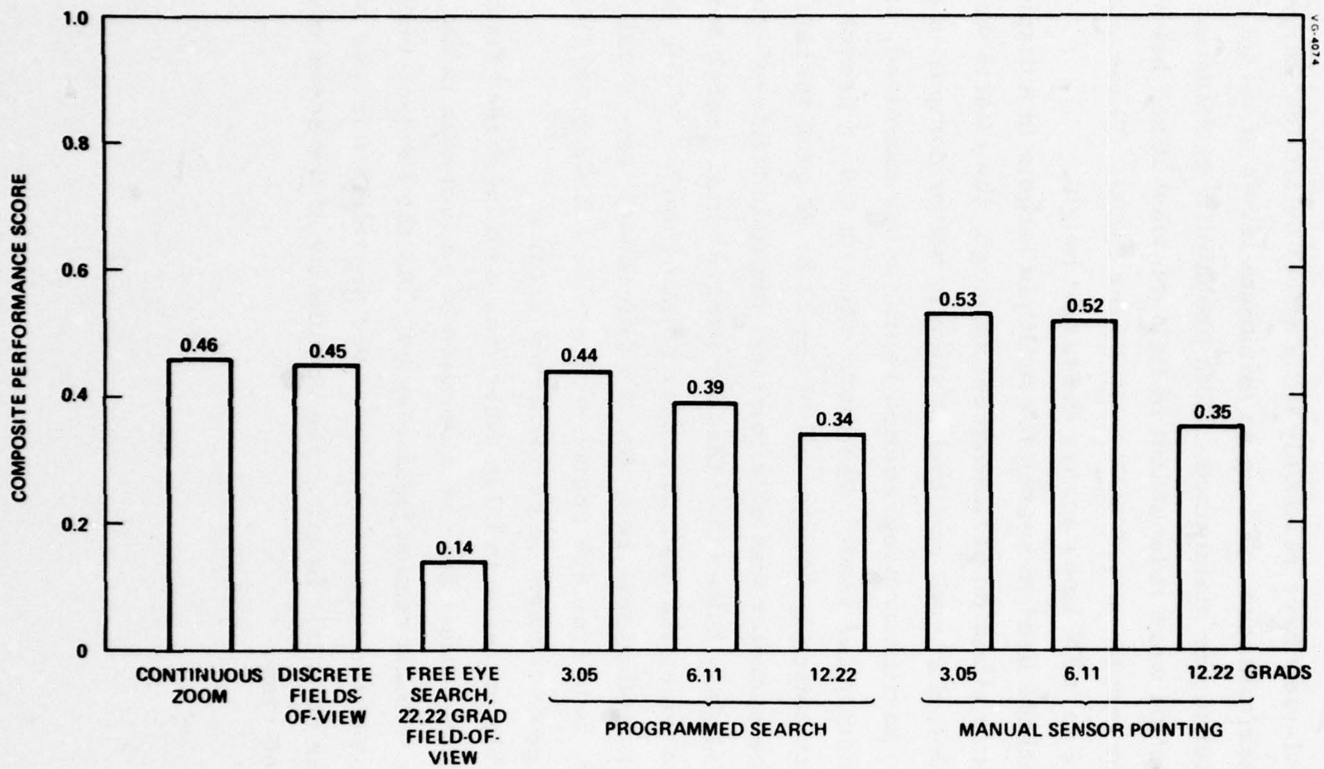


Figure 27. Average composite performance score for the nine search modes.

CONCLUSIONS – EXPERIMENT 4

The results indicate that field-of-view, method of search, and type of multiple field-of-view have relatively little effect on tactical target search and detection performance, as long as minimum levels of target definition and target subtense are maintained. Both probability of detection and latency of detection were independent of field-of-view at and below 12.22 grad (11 degrees) for targets at 1.5 to 2.5 kilometer range. These conditions correspond to 6 to 9 TV lines across the target height.

The methods used to search for multiple targets in a display had surprisingly little effect on performance; although, the effects were reliable. Operator-controlled search resulted in slightly better composite performance scores than the particular programmed search mode examined, indicating that operator-controlled search was more efficient than a preset search pattern which systematically covered all of the 22 by 22 grad terrain area.

When the operator was able to select multiple fields-of-view, performance was slightly better than that for programmed search but not as good as performance under the manual sensor pointing mode. While the overall effect of search method was reliable, the differences were small. Of the nine methods tested, only free eye search with a wide (22.22 grad) field-of-view produced a large decrement in performance scores.

A reliable increase in false detections with increased field-of-view was found, and this effect may be a function of an increase in the number of cluster objects with increased field-of-view. As the field-of-view increases a larger portion of the terrain can be viewed providing a larger number of potential clutter objects. In effect the complexity of the scene may be changing with field-of-view.

REFERENCES

Bennett, C. A. and Franklin, N. L., Statistical analysis in chemistry and the chemical industry. New York: Wiley, 1954.

Bergert, J. W. and Fowler, F. D., Target acquisition studies: Visual angle requirements for TV displayed targets, Tech. Rep. OE10689. Orlando, Florida: Martin Marietta Corporation, May 1970 (AD 706 369).

Bernstein, B. R., Detection performance in a simulated real-time airborne reconnaissance mission. Human Factors, 1971, 13, 1-10.

Blackwell, H. R., Contrast thresholds of the human eye. Journal of the Optical Society of America, 1946, 36, 624-643.

Boynton, R. M. and Bush, W. R., Laboratory studies pertaining to visual air reconnaissance, Part 1, Tech. Rep. WADC TR 55-304. Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, September 1955 (AD 091 874).

Boynton, R. M. and Bush, W. R., Recognition of forms against a complex background. Journal of the Optical Society of America, 1956, 46, 758-764.

Boynton, R. M. and Bush, W. R., Laboratory studies pertaining to visual air reconnaissance, Part 2, Tech. Rep. WADC TR 55-304. Wright-Patterson Air Force Base, Ohio: Wright Air Development Center, April 1957 (AD 118 250).

Bruns, R. A., Bittner, A. C., Jr., and Stevenson, R. C., Effects of target size, target contrast, viewing distance, and scan line orientation on dynamic target detection and identification, Tech. Rep. NMC-TP-72-24. Point Mugu, California: Naval Missile Center, August 1972 (AD 747 983).

Buddenhagen, T. F. and Wolpin, M. P., A study of visual simulation techniques for astronautical flight training, Tech. Rep. WADD-TR-60-756. Wright-Patterson Air Force Base, Ohio: Aerospace Medical Research Laboratory, March 1961.

Craig, G. L., Vehicle detection on television: A laboratory experiment, Tech. Rep. NWC-TC-5636. China Lake, California: Naval Weapons Center, April 1974 (AD 919 898L).

Dale, L. J., Knudsen, K. C., Hawley, E. J., Jeffery, D. W., Leininger, W. E., and Bliss, W. D., Target area orientation simulation, Tech. Rep. no number. Point Mugu, California: Naval Missile Center, 1962.

Freitag, M. and Jones, D. B., Static TV search for linear-cued targets with different fields-of-view. In M. P. Ranc, and T. P. Malone, Proceedings of the Seventeenth Annual Meeting of the Human Factors Society. Santa Monica, California: Human Factors Society, October 1973, 69-73.

Hays, W. L., Statistics. New York: Holt, Rinehart, and Winston, 1963.

Krebs, M. J. and Graf, C. P., Real-time display parameters study, Tech. Rep. RADC TR-73-300. Minneapolis, Minn.: Honeywell Inc., Systems and Research Center, September 1973 (AD 773 850).

Krebs, M. J. and Lorence, L., Real-time display parameters study II, Tech. Rep. RADC-TR-75-43. Griffiss Air Force Base, New York: Rome Air Development Center, February 1975 (AD A007 790).

McCalla, T. R., Introduction to numerical methods and FORTRAN programming. New York: Wiley, 1967.

Noton, D. and Stark, L., Scanpaths in eye movements during pattern perception. Science, 1971, 171, 308-311.

Ozkaptan, H., Ohmart, J. G., Bergert, J. W. and McGee, R. A., Target acquisition studies: Fixed television fields-of-view, Tech. Rep. OR9695. Orlando, Florida, October 1968 (AD 677 322).

Patel, A. S., Spatial resolution by the human visual system. The effect of mean retinal illuminance. Journal of the Optical Society of America, 1966, 56, 689-694.

Petersen, H. E. and Dugas, D. J., The relative importance of contrast and motion in visual detection. Human Factors, 1972, 14, 207-216.

Rogers, J. G. and Carel, W. L., Development of design criteria for sensor displays, Tech. Rep. NR 213-107. Arlington, Virginia: Office of Naval Research, December 1973.

Stathacopoulos, A. D., Gilmore, H. F., and Rohringer, G., Review of mathematical models of air-to-ground target acquisition using TV and FLIR sensors, Tech. Rep. NWC-TP-5840. China Lake, California: Naval Weapons Center, January 1976.

Winer, B. J., Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Zusne, L., Visual perception of form. New York: Academic Press, 1970.