

AD-A108 183 TECHNOLOGY SERVICE CORP SANTA MONICA CA
LOW ALTITUDE SIMULATION/PREDICTION TECHNIQUES.(U)

TECHNOLOGY SERVICE CORP SANTA MONICA CA

F/G 8/5

LOW ALTITUDE SIMULATION/PREDICTION TECHNIQUES. (U)

JUN 81 S A KNOBEL, A J STENGER, W R STONE

F30602-78-C-0360

UNCLASSIFIED

TSC-PD-B623-1

RADC-TR-81-138

NL

1 OF 1

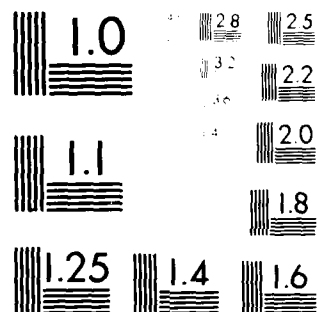
ΔD A

118 118

AD A

$$\frac{1}{2} + \frac{1}{2} = 1$$

END
DATE
FILMED
01-82
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AD A108183

(12) LEVEL II

RADC-TR-81-138
Final Technical Report
June 1981



LOW ALTITUDE SIMULATION/ PREDICTION TECHNIQUES

Technology Service Corporation

**Susan A. Knobel
Anthony J. Stenger
Warren R. Stone**

**DTIC
ELECTE
DEC 8 1981
S B**

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

*Original contains color
plates: All DTIC reproduct-
ions will be in black and
white*

DTIC FILE COPY

**ROME AIR DEVELOPMENT CENTER
Air Force Systems Command
Griffiss Air Force Base, New York 13441**

81 12 08 247

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

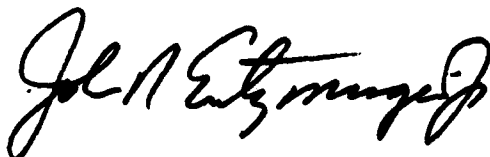
RADC-TR-81-138 has been reviewed and is approved for publication.

APPROVED:



ROBERT H. LASALLE
Project Engineer

APPROVED:



JOHN N. ENTZMINGER
Technical Director
Intelligence & Reconnaissance Division

FOR THE COMMANDER:



JOHN P. HUSS
Acting Chief, Plans Office

If your address has changed or if you wish to be removed from the mailing list, or if the addressee is no longer employed by your organization, please notify RADC (IRRP) Griffiss AFB NY 13441. This will assist us in maintaining a current mailing list.

Do not return this copy. Retain or destroy.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER RADC-TR-81-138	2. GOVT ACCESSION NO. AD-A108 183	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) LOW ALTITUDE SIMULATION/PREDICTION TECHNIQUES		5. TYPE OF REPORT & PERIOD COVERED Final Technical Report 30 Sep 78 - 30 Dec 80
		6. PERFORMING ORG. REPORT NUMBER TSC-PD-B623-1
7. AUTHOR(s) Susan A. Knobel Anthony J. Stenger Warren R. Stone		8. CONTRACT OR GRANT NUMBER(s) F30602-78-C-0360
9. PERFORMING ORGANIZATION NAME AND ADDRESS Technology Service Corporation 2950 31st Street Santa Monica CA 90405		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62702F 45941806
11. CONTROLLING OFFICE NAME AND ADDRESS Rome Air Development Center (IRRP) Griffiss AFB NY 13441		12. REPORT DATE June 1981
		13. NUMBER OF PAGES 60
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A
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) Same		
18. SUPPLEMENTARY NOTES RADC Project Engineer: Robert LaSalle		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Image predictions Level of scene detail Image generation Experiment on image utility Image content Resolution Quantization Reduction Image features Edge detection		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Preparation of sensor predictions at the present time is a manual process requiring several hours of effort by highly trained personnel. This extremely slow response capability is unacceptable within the framework of most Air Force missions, particularly those of a tactical nature. The problem is accentuated in low-altitude missions where the effective sensor resolution places a greater burden on the existing sensor image prediction system and the supporting data bases.		

DD FORM 1473
1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

(Cont'd)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Item 20 (Cont'd)

The purpose of the study was to analyze the use of image predictions in order to find a more compact set of prediction techniques that would result in a more efficient image prediction system.

An experimental approach was used to determine the relationship between navigator performance and the type and amount of information in the prediction image. A number of subjects were given image predictions containing different image features and then asked to identify the predicted targets in corresponding dynamic flight sequences over scenes of cultural content.

Three techniques were used to compress the amount of information contained in the image predictions. The techniques employed are edge representation, a reduction in resolution, and gray level quantization. The results of the experiments showed the performance to be essentially comparable for all techniques, with the resolution degradation technique having slightly higher performance. In addition, while performance is degraded when either the target or background is reduced, the degradation is less when the target area is reduced than when the background is reduced.

The experiments continue to support the results of the previous study, namely that image prediction based on reduced information content are a more effective means of generating image predictions. The processing and storage requirements of the image generation system and the requirements of the supporting data are significantly lowered when image predictions of reduced information content are employed.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

CONTENTS

<u>Section</u>	<u>Page</u>
1. INTRODUCTION	1
1.1 OBJECTIVE	1
1.2 BACKGROUND	3
1.3 SUMMARY	6
1.4 REPORT OUTLINE	6
2. EXPERIMENTAL DESIGN	8
2.1 INFORMATION REDUCTION TECHNIQUES	8
2.1.1 Quantization Reduction	8
2.1.2 Resolution Averaging	9
2.1.3 Edge Thresholds	10
2.2 AMOUNT OF INFORMATION REDUCTION	11
2.2.1 Quantization Reduction	12
2.2.2 Resolution Averaging	14
2.2.3 Edge Thresholds	14
2.3 INDEPENDENT MANIPULATION OF TARGET AND BACKGROUND	15
2.4 EXPERIMENTAL HYPOTHESES	16
2.5 DESIGN	18
2.5.1 Stimulus Conditions (Predictions)	18
2.5.2 Scene Effects	18
2.5.3 Order Effects	22
2.5.4 Dependent Variables	24
2.6 EXPERIMENTAL MATERIALS	24
2.7 PROCEDURE	34
2.7.1 Preparation	34
2.7.2 Practical Trials	34
2.7.3 Test Trials	39
3. RESULTS	40
3.1 SCENE ANALYSIS	40
3.2 REDUCTION ANALYSIS	43
4. CONCLUSIONS	52

LIST OF ILLUSTRATIONS

<u>Figure</u>	<u>Page</u>
1. A randomized complete block design.	21
2. A Latin-Squared Design	23
3. Triangle Building	27
4. Quonset Huts	27
5. Round Tank	28
6. Tall Building	28
7. Athletic Field	29
8. Department Store	29
9. Hilton	30
10. Pillar Building	30
11. Reduced Version of Figure 8, B/W Target - 9 Pixel Background	35
12. Reduced Version of Figure 8, 9 Pixel Target and Background	35
13. Reduced Version of Figure 8, B/W Target - 25 Pixel Background	36
14. Reduced Version of Figure 8, B/W Target - 4-Gray Level Background	36
15. Reduced Version of Figure 8, B/W Target - 2- Gray Level Background	37
16. Reduced Version of Figure 8, B/W Target - Edge Background	37
17. Reduced Version of Figure 8, Edge Target and Background	38

LIST OF ILLUSTRATIONS (continued)

<u>Figure</u>		<u>Page</u>
18.	Hit and False Alarm Rates for each of the Eight Scenes, summed across reduction manipulations	42
19.	The percentage of hits, relative to the baseline, for each of the reduction manipulations	46
20.	The functional relationship between estimated information reduction and target identification for the independently-manipulated targets and backgrounds	48
21.	The effects of mild versus extreme background degradation on identification for full B/W targets	50
22.	The effects of independent (full B/W) versus non-independent target degrading for two reduction methods	51

Accession For	
NTIS GRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A	

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. Type of Prediction	19
2. Scene Descriptions	31
3. Scene Analysis Data	41
4. Reduction Analysis Data	44

1. INTRODUCTION

1.1 OBJECTIVE

Preparation of sensor predictions at the present time is a manual process requiring several hours of effort by highly trained personnel. This extremely slow response capability is unacceptable within the framework of most Air Force requirements, particularly those of a tactical nature. Distinct needs exist for the ability to generate image predictions within time spans of several minutes.

Thus far, the development of equipment for generating predictions has been based on the assumption that hard-copy predictions should replicate all the detail and characteristics of the on-board sensor display. This approach is proving to be overly complex, inefficient, and too costly. However, by reducing prediction-image content to only essential information, significant reductions in processing, storage, and systems cost can be realized--considerations which become increasingly significant as prediction capabilities are extended to general sensor classes.

An initial study^{*} investigated the use of image predictions of reduced information content. The results of the study showed that image predictions of reduced information content are a more efficient means of generating image predictions. In some of the cases reported, the prediction performance with images of reduced scene content is comparable to that using scenes of full scene detail. In addition the data base storage

^{*}Stenger, A. J., et. al., Sensor Image Prediction Techniques, Final Report PADC, Contract F30602-78-C-0054, Technology Service Corporation.

and image prediction system required to generate predictions of reduced information content are of much less complexity than the present prediction generation system and its supporting data base.

An objective of this study is to extend the results of the previous investigation to other forms of reduced image content. Previously, the selected target(s) and background were treated in a homogeneous manner. The intent of this study is to perform different operations on the target and background and to embed the target in a background of different image content.

A second objective of this study is to extend the mission to missions conducted at low altitude. Strategic and tactical mission plans increasingly emphasize the use of low altitude penetration techniques. As a result, prediction and simulation of sensor images, at the lowest mission altitudes, have gained an increased degree of importance. Generation of sensor predictions for the altitude range of 200 to 2000 feet introduces problems which do not exist, or are less severe, at higher altitude levels. For example, sensor resolution translates into greater ground level image detail. This increased level of detail easily exceeds the amount of information available in the Defense Mapping Agency's Digital Landmass System (DLMS) data base. DLMS terrain data consists of a gridded format where each grid point is assigned an altitude value. The grid of data is produced at either or both of two point spacings. These are referred to as levels 1 and 2, corresponding to approximate distances of 300 and 100 feet. Thus, if it becomes necessary to portray terrain infor-

mation in greater detail, the data must be extended by interpolation or the data base itself must be expanded to the new level of detail. The problem is more critical with regard to cultural data. As object detail increases, more and more stored data is required and an increased processing burden results. Consequently, response time falls and system costs rise. Thus, it is imperative that only essential image information be stored and processed.

To repeat, then, the objectives of this study are twofold: to extend the investigation of the previous study to (1) other forms of reduced information content and (2) image predictions used in low altitude missions.

1.2 BACKGROUND

As noted above, this study is related to a previous study that had the similar purpose of investigating the relationship between the amount of information contained in an image and the resultant mission performance. For the purposes of background information, it is worthwhile to briefly summarize the results of the previous study titled, "Sensor Image Prediction Techniques."

To understand how prediction images are used, the navigator's role and his performance of mission tasks were analyzed. Although image predictions have a great applicability, their use by the navigator in performing his search, detection, and recognition tasks typifies their wider usage.

Two major conclusions were drawn from this analysis. The first was that the elements of the navigator's job are the same in their essentials regardless of the sensor, even though different terms are used with the various sensors. Secondly, the operations required in navigating to the target (e.g., at the check points) are the same as those required at the target area. In other words, the behavioral operations required for locating the check points are the same as those required for locating the target, even though there may be major differences in the subject matter of the supporting images.

The approach used in this study was to experimentally determine the minimum amount of information that a prediction must contain in order for a navigator to perform acceptably. In a larger sense, the purpose of the experiment was to determine the relationship between navigator performance and the amount of information contained in the prediction image. Basically, the experiment was as follows: The subject is presented with prediction images (one at a time) that have various levels of information. After each presentation, he views a dynamic sequence and attempts to identify objects within the sequence that have been pointed out in the prediction image. His performances for the predictions are then compared as a function of the information level contained in the predictions.

The experiment used four sets of imagery. The first set, test set A, contained five image features: color, black and white, 4 gray levels, 2 gray levels, and outline only. All prediction images of test set A were at full scene detail. Three other sets--test sets B, C, and D--

were at reduced scene detail and consisted of the target areas alone or the target areas and up to eight other scene elements. Test set B used 2-gray-level predictions, and test set C used an outline representation of the target areas. Test set D used a symbolic representation (i.e., only the positions of the target areas and other scene elements were denoted). Each test set also contained three images of different scene content: one cultural, one terrain, and one mixed (cultural and terrain). The dynamic scenes and predictions were derived from video tape sequences and photographs of urban, mountain, coastal, and desert areas in and around Los Angeles, California.

The data from the experiment was analyzed and a number of significant conclusions were reached.

- 1) The performance is maximum when the target or navigational points are in scenes of mixed cultural and terrain content rather than in scenes of only cultural or terrain content.
- 2) As a general conclusion, the performance is degraded to varying levels as the information levels are reduced. However, the pattern of decreasing accuracy is mainly due to the use of symbolic predictions.
- 3) Performance is only slightly reduced when either contrast or outline predictions are used.
- 4) The performance resulting from the use of 4-gray-level prediction was approximately the same as that using the full detail color predictions.
- 5) The impact on the image prediction system and supporting data bases is significant. Processing speeds can be greatly increased and can be accomplished through image processing of low dynamic range.

1.3 SUMMARY

Three techniques were used to compress the amount of information contained in the image predictions. The techniques employed are edge representation, a reduction in resolution, and gray level quantization. The results of the experiments showed the performance to be essentially comparable for all techniques, with the resolution degradation technique having slightly higher performance. In addition, while performance is degraded when either the target or background is reduced, the degradation is less when the target area is reduced than when the background is reduced.

The experiments continue to support the results of the previous study, namely that image predictions based on reduced information content are a more effective means of generating image predictions. The processing and storage requirements of the image generation system and the requirements of the supporting data base are significantly lowered when image predictions of reduced information content are employed.

1.4 REPORT OUTLINE

Section 2 of this report contains a discussion of the experimental design. It begins with a discussion of the image reduction techniques and continues with a discussion of the experimental hypothesis, i.e., a statement of why the tests were conducted and why the particular tests were chosen. The design of the experiment is then discussed, followed by

a description of the tests themselves. The section ends with a description of the scenes and data used in the experiment and describes how they were generated.

The results of the experiment are discussed in detail in Section 3 of the report. The results are keyed to the different types of image content used in the predictions and to the scenes themselves. The conclusions pertaining to the types of image predictions and to the image generation system appear in Section 4 of the report.

2. EXPERIMENTAL DESIGN

This section gives a complete description of the experiments. First the techniques of information reduction are described and a qualitative measure of the amount of reduction is presented. A key feature of this study is the independent reduction of the target area and its background. This feature is discussed next. The experimental hypothesis, i.e., what the experiment is trying to accomplish, is then discussed. This is followed by describing the design of the experiment. The generation of the materials used in the experiment is the topic of the next subsection. This section concludes with a listing of the procedures followed in the experiment.

2.1 INFORMATION REDUCTION TECHNIQUES

2.1.1 Quantization Reduction

In the field of image processing, there are several different methodologies that can be applied to reduce the information required to produce an image. One such technique is a quantization reduction i.e., reducing the number of gray level intensities into which a black-and-white (B/W) picture is quantized. Typically, 64-gray levels are used to represent a B/W picture. In the last study, a tapered quantization method was used to reduce the B/W scene to a 4-gray level representation. This reduction proved highly successful in the previous study, and was repeated in the current study to examine the resulting effect when independent manipulation of the target and background are employed.

2.1.2 Resolution Averaging

Another technique that can be used to reduce the required information for representation is that of resolution degradation, or averaging. The resolution of a picture is determined by how finely graded its representation is. An area the size of a standard television screen is typically graded into 512×512 "picture elements" or pixels. Each pixel covers an area of approximately $8.4 \times 10^{-4} \text{ in.}^2$ and contains a single value of gray level intensity information. If the same total area is more coarsely graded into 256×256 elements (yielding larger elements, each covering an area of $\sim .0034 \text{ in.}^2$) or 128×128 elements (each element covering 0.0134 in.^2), the total number of elements required to represent the scene is reduced, but the resulting resolution is also reduced.

Given a 512×512 pixel representation, the typical method of reduction is through pixel averaging. Resolution reduction is accomplished by averaging together some number of pixels within a square and representing the intensity of those pixels by the average value. Thus, a reduction of 512×512 pixels to a 256×256 representation would be performed by averaging together each of the 4-pixel squares (2×2) and using one intensity value for all four pixels. Similar averaging of larger square areas would be performed to create more coarsely-graded reductions.

This method of pixel averaging was used in the current study, as the human perceptual system is particularly equipped to handle resolution degradations under certain circumstances. The human mind itself is an enormous averager, and is amply equipped to reconstruct information based

on averaging. It was particularly attractive to consider resolution degradation in this study since the purpose of the study is to investigate target identification under low-altitude conditions. The effect of resolution degradation upon target identification becomes worse as the target object and other scene objects are smaller in area. For a given target, the higher the altitude at which it is perceived, the smaller it will appear. Thus, in a high-altitude situation, where the target and other scene cues will tend to be smaller, resolution degradation could have severe effects upon the ability of the perceptual system to reconstruct the appropriate information for identifying the target. However, in a low-altitude situation, both target and scene elements tend to be larger in area; thus resolution degradation has a less serious effect on the ability of the perceptual system to reconstruct the necessary information for correct identification. For this reason, resolution degradation was employed in this study as one of the types of prediction imagery.

2.1.3 Edge Thresholds

A third method of reducing the image information is to represent the entire scene in an outline form. In the previous study, an attempt was made to accomplish this by using hand-drawn outlines of the relevant scene elements. This method is somewhat unsatisfying, however, as it does not relate to an automatic method of reducing scene information. Therefore, a different methodology for generating outlines was sought.

The technique used in the current study involved developing software to detect contrast changes in the scene based on a threshold. The

information in the scene is then reduced to information containing changes in contrast. This results in edge representations. In areas where the contrast between elements was low, edge information disappeared or became very faint. In areas where there was a sharp contrast, there was a prominent edge representation. Thus, an outline-like drawing of the scene was generated.

In terms of the human perceptual mechanism, an edge representation should provide extremely salient cues for identification. The reason for this is that the human perceptual system itself tends to sharpen edges or contrast information. Psychophysical experiments (e.g., Mach bands, lateral inhibition, etc.) have determined that contrast differences are one of the major factors used by the perceptual system to interpret scene information. The perceptual system seems to be more sensitive toward contrast information and edge detection than any other form of visual input. Therefore, input consisting mainly of edge information should allow for high identification. The current study, then, employed three methods of information reduction: quantization reduction, resolution averaging, and edge thresholds.

2.2 AMOUNT OF INFORMATION REDUCTION

The three techniques employed in this study for reducing the information necessary to represent a scene have been presented. The specific amount by which the information is reduced and the degree to which it is employed varies greatly with each technique. Unfortunately,

for some of these techniques, the amount by which the information is reduced also varies as a function of the scene content. Thus, any attempt to estimate the amount by which a technique reduces the information, independent of the particular scene, is at best an approximation. An approximation is attempted to provide some idea of the different amounts of information reduction, and to allow for an analysis of resulting target identification as a function of information reduction.

2.2.1 Quantization Reduction

The quantization method is the hardest of the three to estimate information reduction since it is the most dependent on particular scene content. To store the original 64-gray level intensity information, a 6-bit word of intensity information is required for each pixel ($2^6 = 64$). To reduce this information to 2-gray levels, only a 1-bit word is required (e.g., 1 = white, 0 = black). Thus, this involves a reduction from 6 bits to 1 bit for each pixel, or a 6:1 reduction in information storage space. Similarly, to reduce the information to 4-gray levels, only 2 bits are required per pixel, yielding a 3:1 reduction.

However, it would be most inefficient and unnecessary to store the information for each pixel individually in a reduced gray-level representation. Across a raster (row) of 512 pixels, the same information may be repeated in a long stream of contiguous pixels, particularly in a homogeneous portion of the scene. For example, in a 2-gray level representation of a scene which included some sky area, a particular raster across the sky may contain 512 pixels all of which have the same

value of 1. In that case, it would be much simpler to represent the information in two words, one indicating the gray level value (1), the other indicating the number of contiguous pixels containing that value (512). For this raster, then, the reduction in information would be from 512 words to 2 words, or 256:1, above the already obtained 6:1 reduction in bit-space.

The number of contiguous pixels containing the same value will vary as a function of homogeneity of the scene. The terrain scenes of the previous study contained the most homogeneity thereby yielding the greatest reduction of information. The cultural scenes were the least homogeneous, yielding the least reduction of information. For a busy cultural scene there may be as many as 50 changed gray levels across a raster, requiring 50 pairs of words to represent the information in that raster. The reduction in this case, then, would be from a 512-word raster to a 100-word raster, or approximately a 5:1 reduction above the 6:1 reduction in bit space already obtained.

The current study did not employ terrain scenes at all, rather the scenes were essentially cultural. A reasonable estimate therefore is that the information reduction per raster varies between 5:1 and 10:1. The reduction of information per raster is greater for the 2-gray level representation than that of 4-gray levels, as two contiguous areas of different values in the 4-gray level representation may map into the same value for 2-gray levels. Reasonable estimates for the overall average raster reduction are 8:1 for the 2-gray level representations and 7:1 for

the 4-gray levels. Combining this with the already obtained bit information reduction (6:1 for 2-gray levels, 3:1 for 4-gray levels) yields a total reduction of approximately 50:1 for 2-gray levels and 20:1 for 4-gray levels.

2.2.2 Resolution Averaging

Unlike quantization, the reduction of information using resolution averaging is essentially independent of scene content, and thus straightforward to calculate. Information is reduced by averaging a square area of pixels together and representing the square of pixels with the one average value. For a 2 x 2 square area, all 4 pixels in the square would be represented by one single averaged value, yielding a 4:1 reduction in information. For a 3 x 3 square area, 9 pixels would be represented by one value, yielding a 9:1 reduction. Similarly, a 4 x 4 averaging yields a 16:1 reduction, 5 x 5 averaging yields 25:1, etc.

Clearly, there is a limit to the degree of averaging that can be done and still yield a perceptible image. Four and nine pixel averaging (2 x 2 and 3 x 3) yield reasonable perceptible scenes. Anything beyond 25-pixel averaging (5 x 5) yields practically indistinguishable images. The current study employed both a mild degree of pixel averaging (9 pixels averaged) and a more extreme degree (25 pixels averaged).

2.2.3 Edge Thresholds

In the edge representation, each pixel is either part of an outlined edge or it is not. Since only 2-gray levels are necessary to represent edges (black outlines on a white background) the 6-bit word for each pixel can be reduced to a 1-bit word yielding a 6:1 reduction in information.

Unlike 2-gray quantization reduction, however, value changes across a raster of edges will be far more frequent. This is particularly true of cultural scenes, in which there is little homogeneity and many edges. It thus becomes less attractive to attempt further reduction of pixel information within rasters. Representing raster information as described for the quantization method can require a larger overhead in software processing. This is due to the fact that the raster records will be variable rather than fixed-length records. In the quantization method this overhead is inconsequential compared to the tremendous reduction obtained. However, the edge representation will not yield nearly as large a reduction in raster information, and thus it is not clear that it is worth the extra overhead involved. For this reason, a full raster representation is assumed for edge reductions, leaving the information reduction estimate at 6:1.

2.3 INDEPENDENT MANIPULATION OF TARGET AND BACKGROUND

In the previous study, both the target and the background were degraded to the same degree. In the current study, the target degradation differed from that of the background. To the extent that target and background are independently manipulated, the target degradation should be equal to or less than that of the background. That is, the target should be more salient and have a higher information content than that of the background.

Since the target covers a small area relative to the entire scene, the higher information content of it will have very little effect on the

overall amount of information reduction produced by the background degradation. Therefore, to study the effect of a highly salient target embedded in a degraded background, targets are represented by the full, 64-gray level, unaveraged information (B/W). The small target areas are embedded in backgrounds degraded by the methods described in Section 2.1. In addition, since two of these methods (resolution averaging and edge thresholds) were not used in the previous study, the effects of non-independently degrading both target and background by these two methods were included in this study. The results were used as a baseline to compare the independently manipulated situation. This allowed the comparison of the effects of independent and non-independent manipulation of target and background for the different degradation methods.

2.4 EXPERIMENTAL HYPOTHESES

Both the previous and current study investigated the effects of image degradation produced by reduced information content upon subsequent identification. This study differed from the previous one in several major ways: 1) all scenes were taken from low altitude levels, in contrast to the higher altitudes used in the previous study; 2) all scenes were essentially cultural in nature; 3) the image degradation methods differed from the previous study; and 4) there was independent manipulation of the target and background degradation. In addition, there were many design and procedural differences.

In the current study, three major questions were investigated:

1. How is subsequent target identification affected by the different methods of image degradation? How does it vary as a function of information reduction?

This question refers to the effects of the reduction methods on identification. It is asking, for example, how well do subjects identify targets in a 9-pixel averaged scene versus a 4-gray level scene? The second question refers to the functional relationship between the estimated amount of information reduction and the identification results.

2. When the image quality of the target is held constant for a given reduction method, what is the effect of varying the background degradation from a mild to a more extreme degree? Does this effect vary across different reduction methods?

This question refers to fixing a target, for example full B/W, choosing a reduction method, as resolution averaging, and investigating the effects on identification when the reduction method is applied to a mild degree (e.g., 9 pixels averaged) versus a more extreme degree (e.g., 25 pixels averaged). The second question refers to determining if the effect is the same for a different reduction method, e.g., for quantization, comparing the milder degree of 4-gray levels versus the more extreme degree of 2-gray levels.

3. For a fixed reduction method and degree of background degradation, what is the effect of reducing the target information? Does this effect vary for different background reduction methods?

This question refers to the effects of independent versus non-independent manipulation of target and background. For example, when the background

is degraded by averaging 9 pixels, how well is the target identified when it is similarly degraded (9 pixels averaged) versus when it is not degraded (full B/W)? The second question refers to whether the effect is the same when a different background degradation is used, e.g., edge thresholds.

2.5 DESIGN

2.5.1 Stimulus Conditions (Predictions)

Eight different combinations of target/background information were employed (see Table 1). For the first five reduction conditions (M1 through M5), a full, 64-gray level (B/W) target image was embedded in a degraded background image. For the next two conditions (M6 and M7), the target was degraded to the same level as the background, i.e., M6 indicates edge threshold was used to reduce the entire scene, both target and background; M7 indicates the entire scene was reduced by averaging 9 pixels together across target and background. M8 is a full color picture, which is used as a baseline for comparing the other seven candidates.

2.5.2 Scene Effects

There are unfortunately many other factors which affect target identification outside of the reduction manipulations investigated in this study. Scene-dependent factors such as homogeneity and complexity of scene, size of target, placement of target relative to other scene objects, type of scene objects, etc., can all have effects upon identification.

TABLE 1. TYPE OF PREDICTION

LABEL - NUMBER		BACKGROUND REDUCTION	TARGET TYPE
EDGE	M1	EDGE THRESHOLD	64-GRAY LEVEL (B/W)
9 PIX	M2	RESOLUTION DEGRADATION - 9 PIXELS AVERAGED	64-GRAY LEVEL (B/W)
25 PIX	M3	RESOLUTION DEGRADATION - 25 PIXELS AVERAGED	64-GRAY LEVEL (B/W)
4 GRAY	M4	QUANTIZATION - 4 GRAY LEVELS	64-GRAY LEVEL (B/W)
2 GRAY	M5	QUANTIZATION - 2 GRAY LEVELS	64-GRAY LEVEL (B/W)
EDGE	M6	EDGE THRESHOLD	EDGE THRESHOLD
9 PIX	M7	RESOLUTION DEGRADATION - 9 PIXELS AVERAGED	9 PIXELS AVERAGED
COLOR	M8	COLOR (BASELINE CONDITION)	

Therefore, it is necessary for the design to accomplish two objectives with respect to these effects. First, regardless of the particular scenes used in the experiment, the results should be generalized beyond this specific set to all types of low-altitude, cultural scenes. Second, conclusions should be made concerning the reduction manipulations that are independent of any consideration of scene types.

The first objective, generalization, was met by using several scenes spanning different types and degrees of difficulty within the low-altitude, cultural domain being investigated. By varying the nature and difficulty of the scenes, the results are not dependent on any one type, but can be generalized to the entire domain of cultural scenes.

The second objective, unbiased results, is accomplished through the use of a randomized complete block design. This design forces assignment of each scene to each reduction manipulation. As this requires an equal number of scenes and manipulation conditions, eight scenes of varying difficulty were chosen, to match the eight target/background reduction combinations. Each scene (C) was paired with each reduction manipulation (M), yielding 64 C-M combinations (see Figure 1.) Each subject was presented with eight C-M pairs. So, for example, Subject 1 (S1) saw scene C1 with manipulation M8 (color), scene C2 with manipulation M2 (T B/W-9 pix), scene C3 with manipulation M5 (2-gray levels), etc. Each subject saw all eight scenes and all eight manipulations, in one of the 8 sets of 8 C-M pairs. Across a block of 8 subjects, all 64 C-M pairs were presented and counterbalanced with respect to each other. This type of counterbalancing eliminates any biases in the identification measures or appearances between manipulations and scenes.

		Target/Background Manipulations (M)							
Scenes (C)		M1	M2	M3	M4	M5	M6	M7	M8
	C1	S3	S4	S7	S5	S6	S8	S2	S1
	C2	S6	S1	S5	S3	S8	S4	S7	S2
	C3	S4	S3	S2	S7	S1	S5	S6	S8
	C4	S8	S2	S3	S6	S5	S1	S4	S7
	C5	S1	S5	S6	S2	S7	S3	S8	S4
	C6	S2	S7	S1	S8	S4	S6	S5	S3
	C7	S7	S6	S8	S4	S3	S2	S1	S5
	C8	S5	S8	S4	S1	S2	S7	S3	S6

S = Subject Number

Figure 1. A randomized complete block design. Each subject(S) is presented with each of the scenes (C) paired with a different manipulation (M). Across a block of eight subjects, each of the eight scenes has been paired with each of the eight manipulations, and all 64 combinations of scene manipulation pairings have been presented.

2.5.3 Order Effects

In addition to the C-M counterbalancing, there is one more factor which must be taken into account--the order of manipulation presentations. In any new perceptual-cognitive task, there tends to be a learning effect across trials. That is, the subject's performance improves as a function of continued practice with the task. To avoid having this improvement bias the manipulation data, it is necessary to counterbalance the presentation order of the eight manipulations across the block of 8 subjects.

This counterbalancing is achieved by employing a Latin Square design. The Latin Square design is a type of randomized block in which two factors are scene-manipulation (C-M) pairing, and the manipulation-trial number (or order) assignment. Figure 2 presents the Latin Square used in the current experiment.

In addition to presenting the C-M pairings for each subject, Figure 2 also shows the order in which these pairings occurred. Across the 8 subjects, each manipulation appears once and only once in each of the trial number positions. Thus for Subject 1 (S1), manipulation 1 (M1) was presented first. For S2, M5 was presented first; for S3, M2 was first; and so forth for each subject and each trial number. Note that the order of scenes (C) is not counterbalanced across trials. Only one of these two factors (C or M) can be counterbalanced with respect to order, and the manipulation factor is the critical variable under investigation.

To summarize, the design consists of a Latin Square configuration, counterbalancing scene-manipulation pairing, and manipulation order across

	Trial Number							
	1	2	3	4	5	6	7	8
Subject Number	S1	- C5 - M1	- C1 - M8	- C7 - M7	- C4 - M6	- C2 - M2	- C8 - M4	- C3 - M5
	S2	- C8 - M5	- C1 - M7	- C2 - M8	- C3 - M3	- C6 - M1	- C4 - M2	- C7 - M6
	S3	- C3 - M2	- C7 - M5	- C5 - M6	- C8 - M7	- C4 - M3	- C1 - M1	- C2 - M4
	S4	- C4 - M7	- C8 - M3	- C7 - M4	- C6 - M5	- C5 - M8	- C2 - M6	- C1 - M2
	S5	- C1 - M4	- C3 - M6	- C2 - M3	- C7 - M8	- C4 - M5	- C6 - M7	- C8 - M1
	S6	- C5 - M3	- C7 - M2	- C2 - M1	- C4 - M4	- C6 - M6	- C8 - M8	- C3 - M7
	S7	- C4 - M8	- C3 - M4	- C6 - M2	- C7 - M1	- C2 - M7	- C5 - M5	- C1 - M3
	S8	- C1 - M6	- C4 - M1	- C2 - M5	- C8 - M2	- C6 - M4	- C7 - M3	- C3 - M8

Figure 2. A Latin-Square Design. Each subject(S) is presented with a unique pairing of scene (C) with manipulation (M). In addition, the order (trial number) of manipulations is unique for each subject-trial number combination.

a block of 8 subjects. A total of 3 blocks were used in the study, yielding 24 subjects.

2.5.4 Dependent Variables

The dependent variables of accuracy and latency were the same as those of the previous study. The accuracy of response was determined by measuring the percentage of correct responses (hit rates). The speed of the response was determined by measuring the latencies. Latency is defined as the time elapsed between presentation of the stimulus and the subsequent identification. A third variable, the number of incorrect identifications (false alarms) was also measured. Confidence ratings were not taken for this experiment, as that information did not prove very useful in the last study. Also, confidence ratings require a certain amount of time and extra thinking on the part of the subject.

2.6 EXPERIMENTAL MATERIALS

This subsection describes the method used to gather the necessary raw materials and process them into a usable form for the experiment. The end product was to be the dynamic image sequences and prediction imagery that would be viewed by the subjects during the experiment.

As with the previous experiment, the prediction imagery and dynamic sequences were taken with visual sensors. The predictions were based on 35 mm photographs of the scenes and the dynamic sequences were video tape recordings from TV imagery taken during flights over the scene. The use of visual sensors and imagery taken specifically for this program are justified since sensor imagery (e.g., radar or FLIR), when obtainable, is

not suitable for this study of an experimental nature. First, the problem with using live sensor imagery for an experimental program is that the imagery is fixed to the conditions for which it was taken and thus cannot be extended to other conditions. For example, the flight path parameters of bearing and altitude and the setting of display contrast are fixed. In addition, the missions of the live imagery correspond to navigational or targeting missions and thus either the way point or the target areas are always in the center or near-center of the display. Since the experiment deals with target detection accuracy, the target must wander into the observer's field of view at random points during the sequence instead of being constantly in view.

Three flights were taken in a Bell helicopter to select target scenes and gather the sequences and prediction imagery. The flights, during which the dynamic sequences were recorded and photographs of the scenes taken, were flown at an altitude of approximately 500 ft, at a speed of approximately 70 kn. The first flight was used to scout for potential target areas, check out the TV and photographic equipment, and verify that the flight dynamics and resultant imagery were suitable for the purposes of the experiment. Material for the dynamic sequences was gathered on the second flight. The flight paths were repeated on a third flight to obtain 35 mm photographs of the target scenes to be used for the various prediction imagery.

All scenes were located in the San Fernando Valley and west side sections of Los Angeles, California. All scenes were of a cultural

content as opposed to the scenes of the previous study that were of cultural, natural, and mixed content. The effect of scene content was answered in the previous study (performance is best in scenes of mixed content). The scenes are given in Figures 3 through 10. The targets are listed in Table 2 for each of the scenes.

A psychologist viewed the raw videotapes and selected segments whose content was appropriate for the experiment. After this selection process was completed, each sequence of the video tape was edited for a compact run time. The sequences corresponding to the 8 target areas (and two practice sequences) were each put on a separate tape so they would be easy to show during the experiment. Table 2 lists the length of each sequence and the time at which the target appears on the tape. The length of time the target stays in the tape varies, i.e., it may not remain for the entire length of the sequence.

A total of 64 distinct prediction images were prepared corresponding to the 8 types of predictions and the 8 scenes. Since some of the predictions use a target at full detail embedded in a background of lower information content, an additional image (at full detail) was required for each scene. Thus, the 64 predictions were derived from 72 images of the 8 scenes.

Since the subjects viewed the video tapes on a 512-line display which provided for six bits of either black-and-white or color (two bits each for red, green, and blue) information, all color and black-and-white imagery presented to the subjects as prediction imagery had to be converted

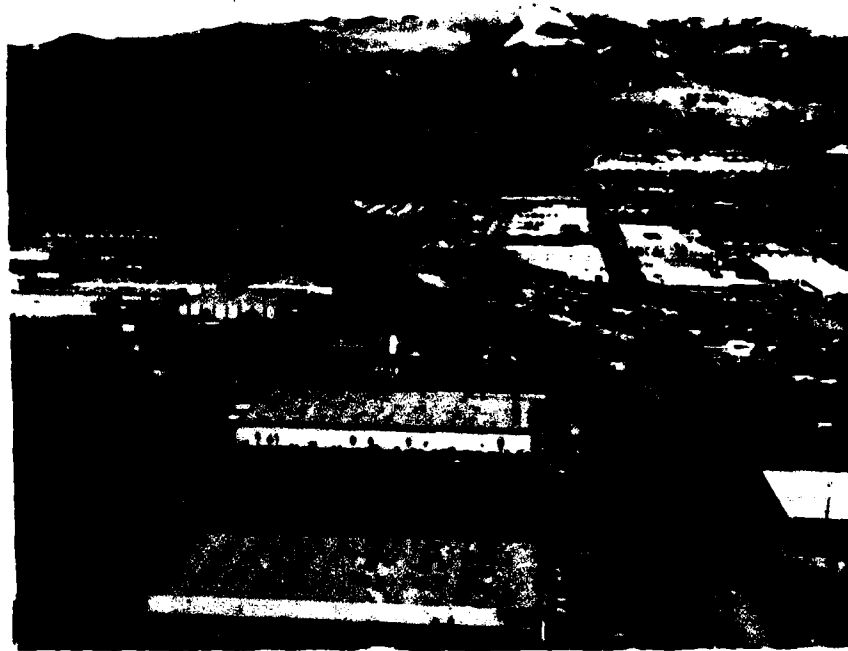


Figure 3. Triangle Building.



Figure 4. Quonset Huts.



Figure 5. Round Tank.



Figure 6. Tall Building.

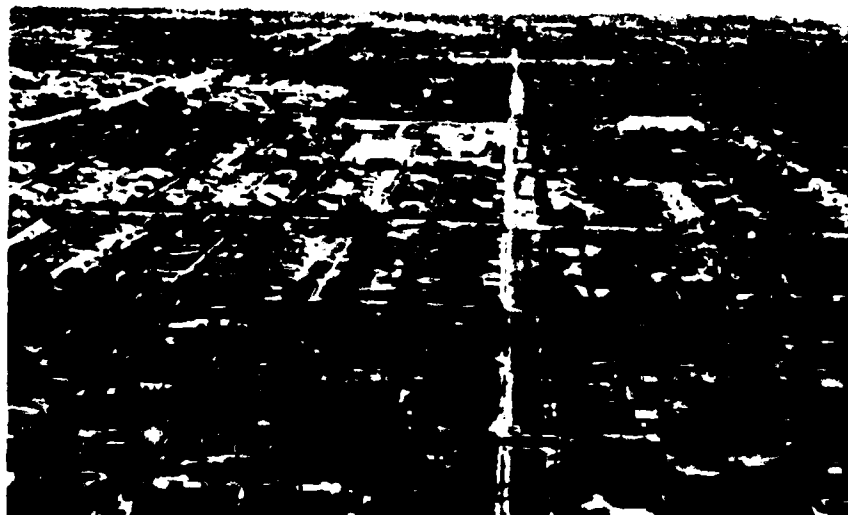


Figure 7. Athletic Field.



Figure 8. Department Store.



Figure 9. Hilton.

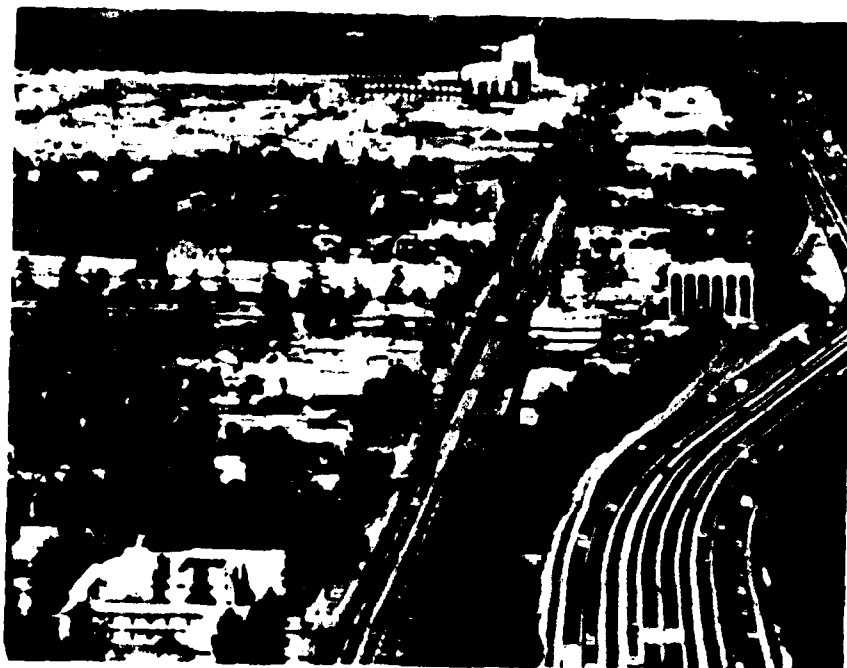


Figure 10. Pillar Building.

TABLE 2. SCENE DESCRIPTIONS

Scene (C)	Total Elapsed Time for Video Segment (sec)	Total Elapsed Time from Beginning of Video Segment to Appearance of Target (sec)
<u>Test:</u>		
C1 - Triangle building	63.0	41.6
C2 - Quonset huts	73.2	60.8
C3 - Round tank	31.2	18.7
C4 - Tall building - right of road	31.8	2.0
C5 - Athletic field	75.0	6.2
C6 - Department Store	64.8	16.0
C7 - Hilton building	71.4	13.6
C8 - Building with pillars - left of freeway	62.4	37.1

to the same format. Therefore, each photograph had to be color digitized three times (corresponding to red, green, and blue filters) and then recombined on a computer using two bits for each color. The process involved lowering the intensities of the three colors, so that the three intensities preserved the dynamic range of the original image. These uniform red, green, and blue images were then combined to form the color images used as prediction imagery in the experiment. A new black-and-white image was made by taking the average of the picture elements or red, green, and blue images. Thus, the black-and-white image has a full dynamic range of 64 intensity levels or 6-bits. The full black-and-white image was used only for the particular target area which was embedded in a background of reduced information.

The 9- and 25-pixel images were created from the full black-and-white image by averaging pixel values in a 3 x 3 or 5 x 5 pixel window. The 4-gray and 2-gray level predictions were automatically created from the full black-and-white image by using thresholds resulting in equal probability. In other words, there are an equal number of pixels at each gray level in the 4-gray and 2-gray level images.

The edge prediction was created automatically by using a standard edge detection algorithm. The edge detection algorithm used in this study was based on the Sobel/Robinson operator. Although this operator calculates both edge magnitude and direction, only the edge magnitude was used in the final result. The Sobel operator is defined by use of the gradient masks $[S_x]$ and $[S_y]$:

$$[S_x] = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \quad [S_y] = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

Let $[F_{ij}]$ be the sub-image around pixel (i,j) . Then the edge magnitude, $G_s(i,j)$, at pixel (i,j) is

$$G_s(i,j) = \left[([F_{ij}] * [S_x])^2 + ([F_{ij}] * [S_y])^2 \right]^{1/2}$$

where $*$ denotes the discrete convolution function. After calculating the edge magnitude, extraneous edge noise was eliminated to render a cleaner edge map. This was accomplished by using a locally adaptive thresholding technique. Given a 3×3 (or a general $n \times n$, n odd) neighborhood around pixel (i,j) in image F , $[F_{ij}]$, and the edge magnitude $G_s(i,j)$, define a new edge magnitude, $G_T(i,j)$ as

$$G_T(i,j) = \begin{cases} G_s(i,j) & \text{if } G_s(i,j) \geq \text{AVE}([F_{ij}])\tau \\ 0 & \text{otherwise} \end{cases}$$

where $\text{AVE}([F_{ij}])$ is the average value of the data in the 3×3 neighborhood about (i,j) . In this manner, an edge magnitude at (i,j) is retained if it is large with respect to the local data about pixel (i,j) . Largeness is a relative value with respect to the threshold τ . A threshold of $\tau = 0.5$ was used in the study.

The predictions used for Scene 6 (Figure 8) are given in Figures 11 through 17.

2.7 PROCEDURE

2.7.1 Preparation

Before the subject(s) arrived, the experimenter (E) prepared the materials. The appropriate packet of ordered scene-manipulation photographs was chosen, and the video tapes were placed in corresponding order. The two packets of practice trials were readied, along with the data sheets for recording. Mechanical equipment, stopwatches, etc. were also checked and readied.

2.7.2 Practical Trials

When S arrived, E first gave S some general information about the purpose of the experiment. A script was provided to E for this purpose, so that the same general information was given to each S. S was then instructed in the procedure. Following the instructions, S could ask questions, after which the two practice trials were given.

For each practice trial, E presented S with the color version of the scene. E pointed to the target and used a one-word description (building, tank, or field) in the sentence "The target is this _____". E then retired for two minutes while S studied the picture. At the end of two minutes, E removed the picture from S, and simultaneously turned on the video sequence. Whether or not S successfully identified the target, the video sequence was played through to the end. At the conclusion of the video, E turned off the recorder, and presented to S each of



Figure 11. Reduced Version of Figure 8, B/W Target -
9 Pixel Background.



Figure 12. Reduced Version of Figure 8, 9 Pixel
Target and Background.



Figure 13. Reduced Version of Figure 8, B/W Target -
25 Pixel Background.



Figure 14. Reduced Version of Figure 8, B/W Target -
4-Gray Level Background.



Figure 15. Reduced Version of Figure 8, B/W Target -
2-Gray Level Background.



Figure 16. Reduced Version of Figure 8, B/W Target -
Edge Background.



Figure 17. Reduced Version of Figure 8, Edge Target and Background.

the seven other versions of that scene, one-at-a-time, while pointing out the target in each. No feedback was given to S concerning the accuracy of his performance. The second trial followed the exact same procedure. At the end of this practice, testing began.

2.7.3 Test Trials

For each of the eight test trials the following steps were performed by E:

- 1) Select video tape, ready monitor (do not show yet)
- 2) Present scene to S. Point at target and use the one-word descriptor phrase.
- 3) Allow S 2 minutes to study the scene. Remove at the end of 2 minutes.
- 4) Immediately start both the tape segments and the stopwatch.
- 5) When S correctly identifies target, stop the stopwatch (continue showing the video tape until the end).
- 6) At the end of the video segment, stop the video. Record S's time and number of false alarms (FA), if any.
- 7) Rewind and remove the video tape, ready the next segment.

Again, no feedback was provided to S concerning his performance. At the end of the last test trial, S was told the experiment was over and thanked for his participation. S was also told that the experiment was hard and he did well, and was asked not to talk about the experiment with anyone else until all the data was collected.

Subjects were obtained by requesting participation from company employees. Subject selection was random, without regard to any particular subject characteristics. Data collection took approximately one week.

3. RESULTS

The results of the experiment are organized and presented along two basic factors. First, a scene analysis is presented showing the results for each scene, averaged across the reduction techniques. Then the results are presented to show the effect of the various reduction techniques, averaged across all scenes.

3.1 SCENE ANALYSIS

Basically, the results of interest here are those relating to manipulations or reduction methods, not scenes. However, because the design is counterbalanced, it is also possible to study the identification results for each scene, averaged across reduction manipulations. These data yield some interesting results.

Table 3 presents the data for each scene. Although there appears to be a relationship (inverse) between the Hit and FA rates, the mean latencies do not appear to relate to either of these two. It seems that the latency of response may be due to some other individual characteristic of the scenes themselves.

Figure 18 is a graphic representation of the Hit and FA data. It is easy to see that some trend toward an inverse relationship exists between the Hits and FA. As the Hit rates reduce, the FA rates tend to increase. This is not surprising in that failure to identify the target may be due to identifying false targets. This failure could be induced by two possible situations:

- 1) there are objects in the scene that are similar to the target and thereby are possible sources of FA, and
- 2) the target object itself is difficult to perceive, thereby encouraging false reconstructions.

From examining the scenes in which the FA rates are high, it seems more likely that the latter situation applies.

Table 3. Scene Analysis Data

Scene	Number of Hits	Number of FA	Mean Latency (sec)
C1 - Triangle Building	11	14	9.1
C2 - Quonset Huts	12	12	8.3
C3 - Round Tank	17	1	26.5
C4 - Tall Building	17	3	22.1
C5 - Athletic Field	23	10	45
C6 - Department Store	23	2	9.2
C7 - Hilton	24	2	28.8
C8 - Pillar Building	24	2	9.8

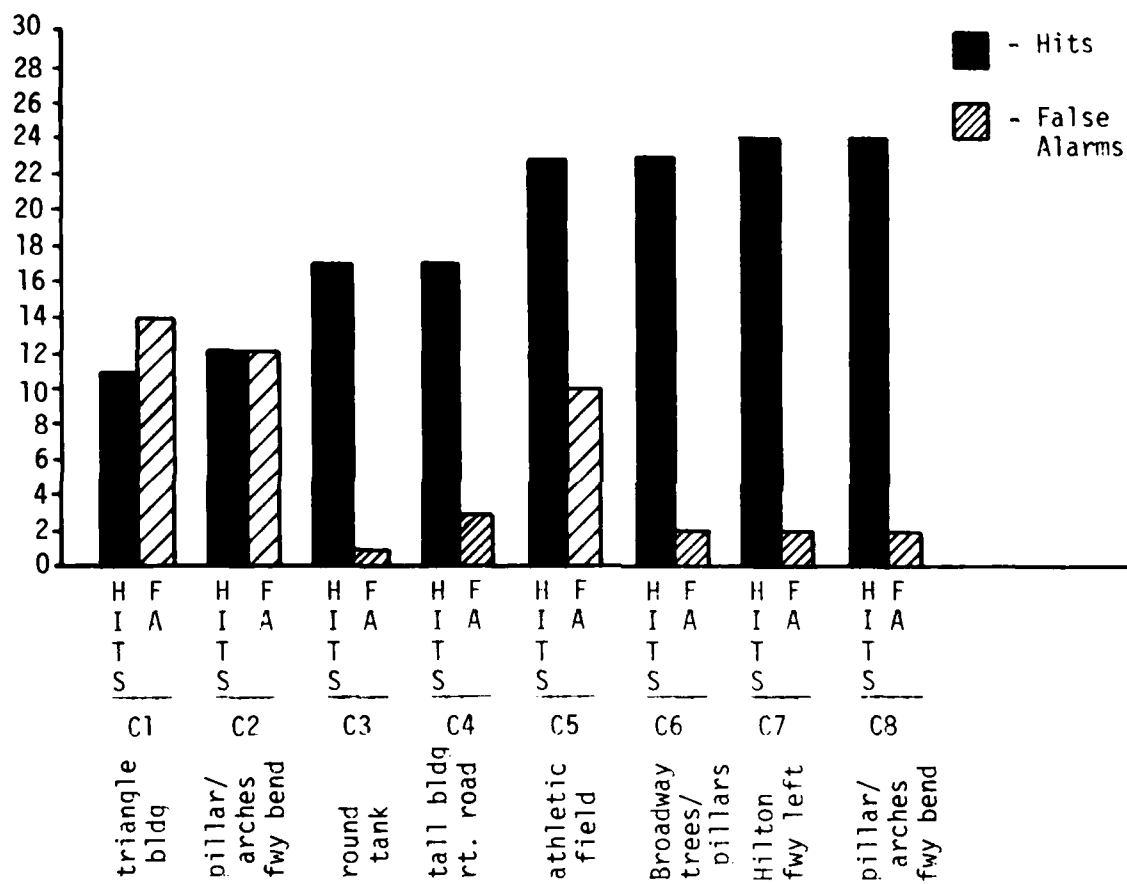


Figure 18. Hit and False Alarm Rates for each of the Eight Scenes, summed across reduction manipulations.

The scenes in Figure 18 have been ordered according to their identification results, from poorest to best (note that the numbering of the scenes here and in Table 2 was done after the fact). The question of interest then is, along what dimension is this order varying that may be related to performance results? The factors that are increasing across these scenes are the target-to-background and within-target contrasts. In scene C1, which has the poorest performance, there is little of either type of contrast. The target of C1, the triangle building, is a low, flat all-white building (no within-target contrast). It is situated behind and next to other low, flat, white buildings, and in front of a homogeneous, flat dirt road (little target-to-background contrast). The contrast is similarly poor for the quonset huts in scene C2. At the other extreme, both the Hilton and the department store (C6 and C7) are tall buildings that stand out in a field of trees and contain some pillars or lattice-work on the face (high contrast). The C8 target is a building with thin, bright-white columns containing dark areas between them (very high within-target contrast) situated by some trees near a bend in a highly-salient, wide, flat freeway (very high target-to-background contrast). Thus, identification appears to be aided greatly by contrast, both the contrast within the target itself, and the contrast between the target and nearby background elements.

3.2 REDUCTION ANALYSIS

Table 4 presents the data for each of the reduction manipulations. In this analysis, the color condition (M8) was used as a baseline for each of the other conditions. The percentage of hits was computed for each

Table 4. REDUCTION ANALYSIS DATA

Reduction Manipulation	% Hits Relative Baseline	Number of False Alarms	Mean Latency (Sec)
Edge	83	7	23.2
9 pix	87	7	22.7
25 pix	74	7	18.4
4 gray	83	6	22.7
2 gray	74	6	22.2
Edge	78	6	23.0
9 pix	78	6	19.9
Color	100	1	17.8
		$\sigma^2 = .29$ (excluding color)	$\sigma^2 = 4.83$

condition and compared to that of color. The resulting relative percent hits was then computed and is being reported in this table (color naturally has a 100% hit rate relative to itself).

Examining the mean latency data in this table versus that of Table 3 indicated some interesting differences. Mean latencies vary greatly across scenes ($\sigma^2=175$) but vary little across reduction mainipulations ($\sigma^2=4.83$). This implies that the speed with which one can identify a target is dependent upon characteristics of the scene itself, and, not upon the degradation characteristics. This may explain why the latency data of the previous study was difficult to interpret.

A similar situation holds for the false alarm data. In Table 3, the FA vary greatly ($\sigma^2=28.2$) and appear to be inversely related to hit rates. In Table 4, the FA vary extremely little ($\sigma^2=.29$ excluding color), and appear to be independent of hit rates. This result is to be expected. One would expect FA to be much more related to the characteristic of a scene and have little relationship to the degradation manipulation.

Figure 19 presents the relative percent of hits in a graphic arrangement. These data compare the effects on identification of different reduction methods, and allow for an analysis of the first hypothesis. The effects group into three relative classes: high, medium and low identification. The high-identification class includes the edge threshold, 9 pixel averaging, and 4 gray level backgrounds, with the full B/W target. The medium-identification class includes the edge threshold and 9 pixel averaging across both background and target. The lower-identification class

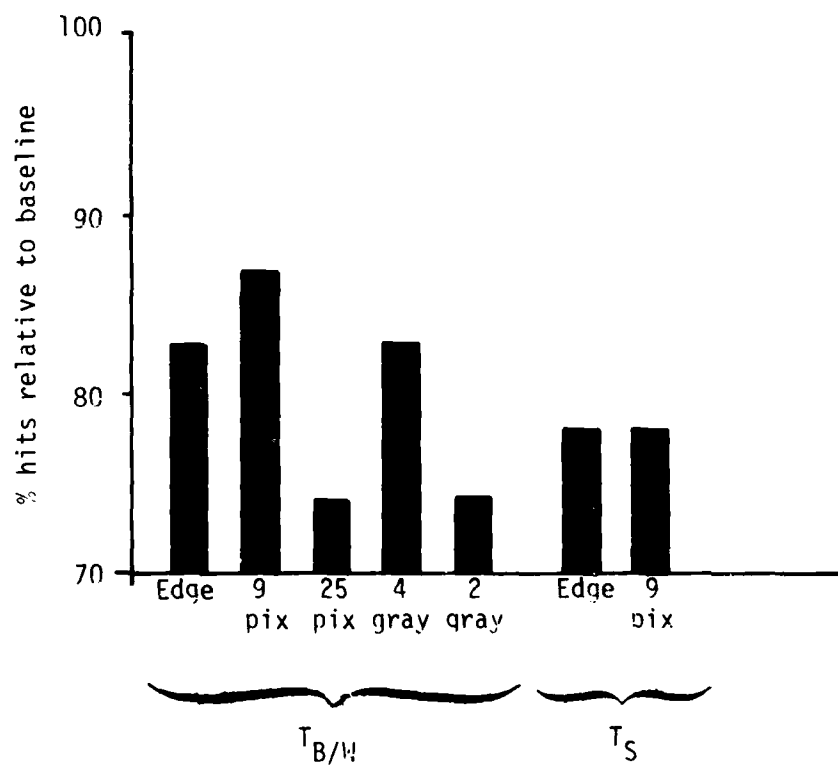


Figure 19. The percentage of hits, relative to the baseline, for each of the reduction manipulations.

includes the 2-gray level and 25 pixel averaging backgrounds with the full B/W target.

Several conclusions can be drawn from these results. It appears that all three reduction methods yield high identification when applied to only a mild degree and contain a full B/W target. Degrading the target to the same mild degree as the background, reduces the subsequent identification somewhat, but not very severely. However, even with a full B/W target, degrading the background to a more extreme degree, has a stronger effect upon subsequent identification. This implies there are trade-offs between background degradation and target degradation. There are limits to the degree that improving one will compensate for decreases in the other.

Figure 20 shows the function relating identification to reduction of information. There is an obvious inverse relationship, as the reduction of information increases, the subsequent identification decreases. Furthermore, this relationship is non-linear. There is a sharp drop in identification from the 4-gray level to the 25 pixel methods, despite the fact that these two are fairly close in terms of their degree of information reduction. In contrast to that, there is no change in identification between the 25 pixel and 2-gray level methods, despite the fact that they differ greatly in terms of their information reduction. In a sense, this is beneficial. It implies that a great savings of information does not necessarily mean a great loss of identification. It is possible to find ways to significantly reduce the image information without seriously

hampering the ability to identify targets as demonstrated by the 4-gray level quantization. Note that these conclusions are based on the rather qualitative assessment of the amount of reduction of image information as computed in Subsection 2.2. However, the general trends should be valid under a more rigorous assessment of the image compression.

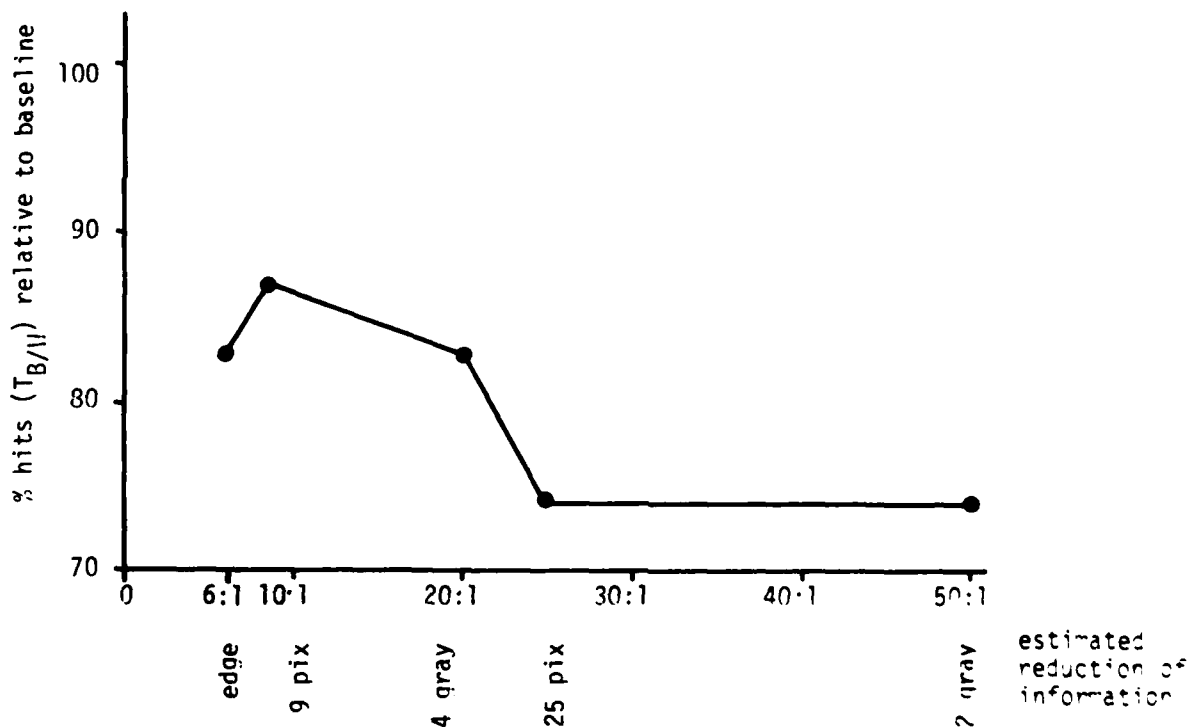


Figure 20. The functional relationship between estimated information reduction and target identification for the independently-manipulated targets and backgrounds.

The second hypothesis posed concerns the effect of mild versus extreme background degradation. Figure 21 displays the data for these conditions. Clearly, there is a decrement in identification with the more extreme degradation. Furthermore, the effect appears to be the same across the two different methods. Despite the fact that quantization reduction and resolution averaging are very different methodologies that yield quite different perceptual results, the same effect is observed for each with respect to degree of degradation. This implies that regardless of the reduction technique employed, there are limits to the degree of background degradation that can be tolerated, even with a full B/W target.

The third hypothesis considered the effect of manipulating target degradation for a given background reduction. Figure 22 presents the data for these cases. The results show that reducing the target information does reduce the subsequent identification, although the drop is not as great as that of Figure 21. This implies that reducing the target from a full B/W to a mild degradation has a less severe effect than reducing the background from a mild degradation to an extreme one when the target is a full B/W. Again, although the two reduction methods (resolution-9 pixels averaged and edge threshold) are very different, the effect is virtually the same across both methods.

Thus, regardless of the background reduction method, reducing the target information results in somewhat of a reduction in target identification.

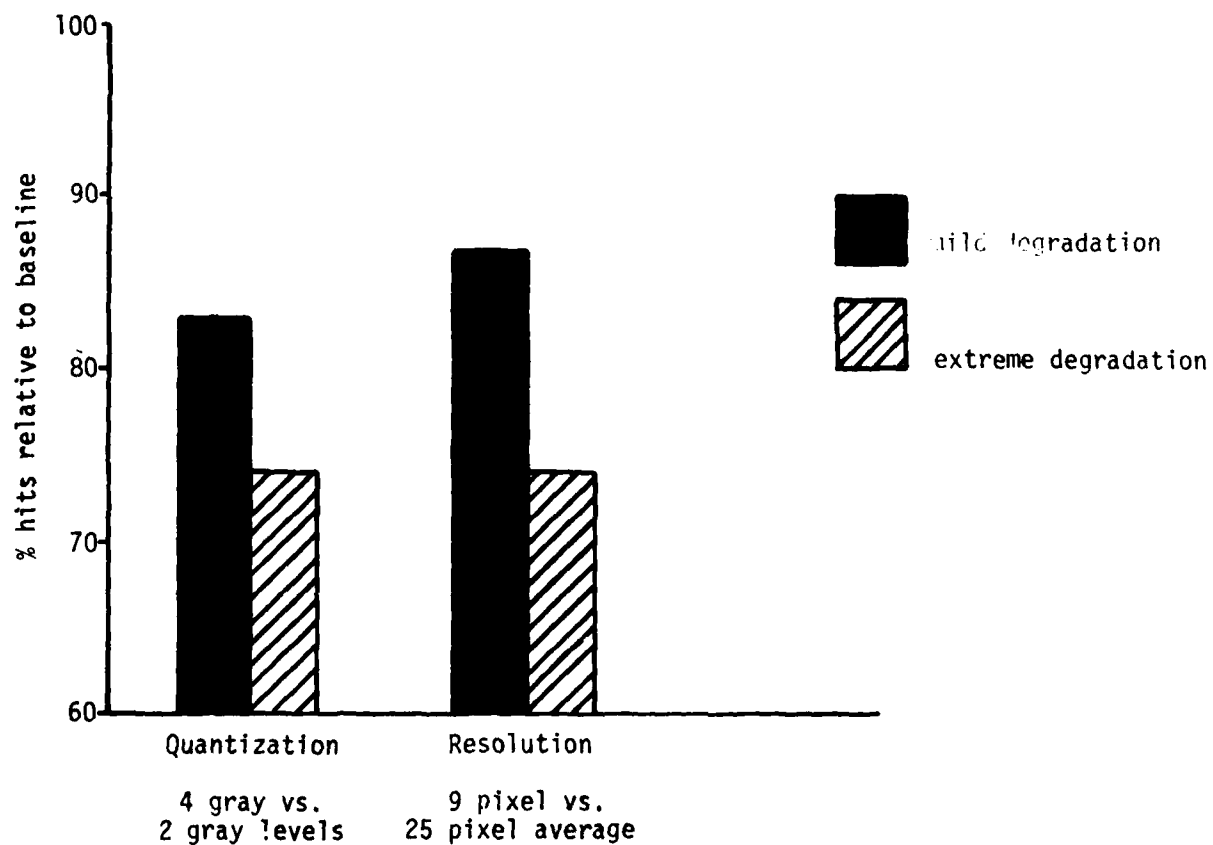


Figure 21. The effects of mild versus extreme background degradation on identification for full B/W targets.

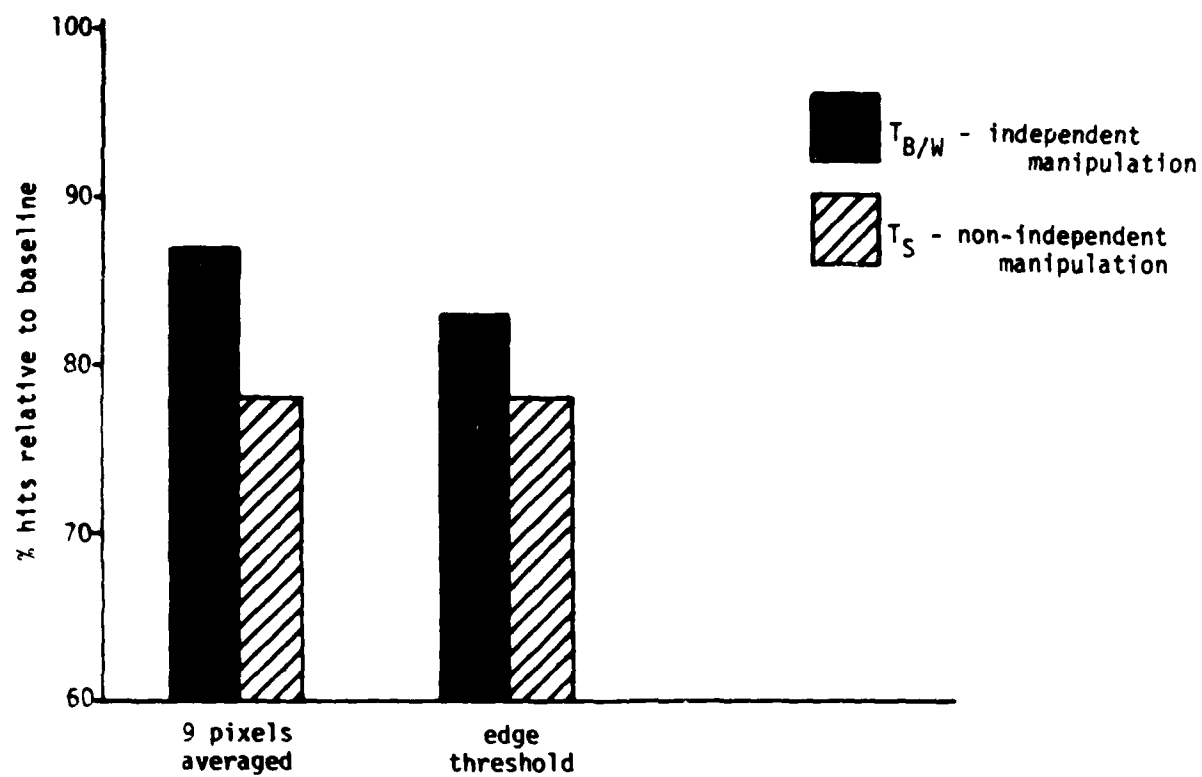


Figure 22. The effects of independent (full B/W) versus non-independent target degrading for two reduction methods.

4. CONCLUSIONS

The major results of the experiments are summarized here. The three methods of reducing the amount of information in a scene (resolution, quantization and edge representation) all result in comparable detection performance, even though they are different methodologies that yield different perceptual results. When the target area is represented at full detail (i.e., 64-gray level) and the background is moderately degraded, the performance is acceptable to that with the full detail, color baseline considering the much reduced image generation system that is required to produce them. When the resolution cell size is reduced by a factor of 9 to 1, the resulting performance is 87% of the baseline performance. The 9-pixel averaging method yielded the highest performance of the reduction techniques. The 4-gray level technique that showed promise in the previous study resulted in a performance level of 83%. The same performance resulted from predictions using the edge representations.

The resultant detection performance levels are further reduced when either the representation of the target area or the background are degraded. However, it is an interesting result that when the target is reduced to that of the background, the performance degradation (to 78%) is less than when the background is further degraded. The performance level is then 74%. These trends hold regardless of the information reduction technique.

The results of the experiments have a great impact on both the image generation system and the supporting data base. In all cases, the processing required by the image generation system is greatly reduced. It need only process data at one-ninth the rate (9-pixel average technique) of the existing systems, use only 2-bit processing (4-gray level technique), or process only edge information (edge technique) rather than the full detail as is currently done. All techniques significantly reduce the amount of storage required.

The requirements placed on the supporting data base are significantly relaxed. All techniques suggest that the data base can be generated at a much lower resolution. In addition, if the edge technique is employed, then the data base does not need to maintain any information on the internal structure of objects.

Two final comments should be made. All performance has been compared to the baseline condition, which uses a color photograph of the actual scene for the prediction. However, existing image prediction systems do not generate predictions of the quality used for the baseline experiment. Thus, the performance resulting from predictions of reduced information content, when compared to that based on existing image predictions, should be higher than that stated here. The experiments continue to support the results of the previous study, namely that image predictions based on reduced information content are a more effective means of generating image predictions.

*MISSION
of
Rome Air Development Center*

RADC plans and executes research, development, test and selected acquisition programs in support of Command, Control Communications and Intelligence (C³I) activities. Technical and engineering support within areas of technical competence is provided to ESD Program Offices (POs) and other ESD elements. The principal technical mission areas are communications, electromagnetic guidance and control, surveillance of ground and aerospace objects, intelligence data collection and handling, information system technology, ionospheric propagation, solid state sciences, microwave physics and electronic reliability, maintainability and compatibility.

TE
MED

82