

63-3-2

ESD-TDR-63-136

FORM PERCEPTION IN VIDEO VIEWING:
EFFECTS OF RESOLUTION DEGRADATION AND STEREO
ON FORM THRESHOLDS

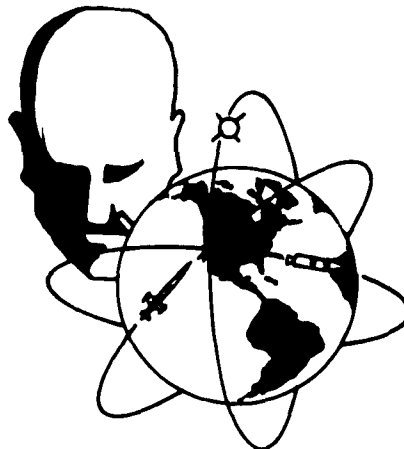
Technical Documentary Report No. ESD-TDR-63-136

December 1962

Norman E. Freeberg

Operational Applications Laboratory
Deputy for Technology
Air Force Electronic Systems Division
Air Force Systems Command

L. G. Hanscom Field
Bedford, Massachusetts



(Prepared under Contract AF 19(628)-328

by

Department of Medical and Biological Physics

AIRBORNE INSTRUMENTS LABORATORY
A DIVISION OF CUTLER-HAMMER, INC.

Great Neck, Long Island, New York

Project 9674, Task 967403

AS AD NO. 401654
CLASSIFIED BY ASTIA

1963
ASTIA

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

Qualified requestors may obtain copies from ASTIA. Orders will be expedited if placed through the librarian or other person designated to request documents from ASTIA.

Do not return this copy. Retain or destroy.

FOREWORD

This report was prepared by Dr. Norman Freeberg, Head of the Human Factors Research Group, Department of Medical and Biological Physics, Airborne Instruments Laboratory, a division of Cutler-Hammer, Inc.

The work was performed under USAF Contract No. AF 19(628)-328, in support of Project 9674, "Information Transmission for Decision," Task 967403, "Information Content of Visual Forms." The contract was administered by the Applications Division, Operational Applications Laboratory, Deputy for Technology of the Electronic Systems Division. Dr. John Coules served as contract monitor and contributed to the planning of the work and the organization of the report.

The author acknowledges the assistance of Mr. Martin Pollack in carrying out the study, and to Mr. Alan Swanson for engineering assistance.

ABSTRACT

Thresholds for video viewing of randomly constructed forms are determined under conditions of pictorial degradation and stereoscopic presentation. Form recognition is not significantly enhanced by stereo viewing but discrimination threshold is reduced. All form thresholds are highly sensitive to video image degradation.

Video acuity is not significantly related to video form thresholds and can be affected by orientation of a pattern in relation to video scan lines.

Reviewed and approved for publication:

FOR THE COMMANDER

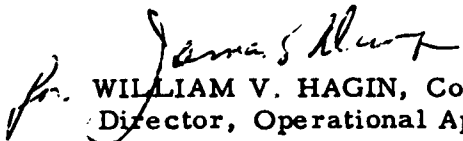

WILLIAM V. HAGIN, Colonel, USAF
Director, Operational Applications Laboratory

TABLE OF CONTENTS

	<u>Page</u>
Foreword	11
Abstract	111
1. Introduction	1
2. Method	4
A. Test Facility	4
B. Stimulus Forms	6
C. Resolution Degradation	9
D. Subjects	11
3. Procedure	13
A. Form Threshold Measures	13
B. Video Acuity Measures	15
4. Results	16
A. Recognition Threshold	16
B. Discrimination Threshold	19
C. Video Acuity	23
D. Form Association	25
5. Discussion	27
6. Summary	33
Appendices	
I. Construction of Stimulus Forms	35
II. Degradation of Video Resolution	37
III. References	41

LIST OF ILLUSTRATIONS

<u>Figure</u>		<u>Page</u>
1	Test Facility	5
2	Non-Stereo Presentation	7
3	Stereo Presentation	7
4	Head-Set for Stereo Viewing	8
5	Non-Stereo Form Faces	10
6	Stereo Form Faces	10
7	Mean Recognition Threshold Scores for Non-Stereo and Stereo Viewing of Four Stimulus Form (N = 7)	17
8	Mean Discrimination Threshold Scores for Non-Stereo and Stereo Viewing (N = 7)	21
I-1	Diagram of Form Construction	36
II-1	Plot of Resolution Degradation vs Equivalent Scan Line Reduction	40

LIST OF TABLES

<u>Table</u>		<u>Page</u>
I	Summary of Analysis of Variance of Non-Stereo Recognition Threshold Scores for Degradation Level and Form Faces	18
II	Summary of Analysis of Variance of Stereo Recognition Threshold Scores for Degradation Levels and Form Faces	19
III	Summary of Analysis of Variance of Discrimination Threshold Scores for Viewing Methods and Degradation Levels	22
IV	Pictorial Resolution	22
V	Summary of Analysis of Variance of Video Acuity Scores for Landolt Ring Gap Positions	24

SECTION 1
INTRODUCTION

The extension of man's visual capabilities by means of electronic imaging techniques has resulted in increasing demands upon observer accuracy in interpreting and utilizing the visual information displayed. Video represents the most widely used of these imaging techniques in many systems where direct visual contact with the operating environment is not feasible, or where a video image serves as an important supplement to direct observation.

Mention of just a few present, or planned, systems applications serves to indicate the scope of dependence upon video presented information. As a primary example, TV has long been an essential feature in the remote handling systems which require the manipulation of fissionable materials ("hot cell" work). Video pictorial information for aerospace systems is to be required in lunar and planetary exploration; for satellites that must be guided visually from earth during terminal phases of rendezvous and for other satellites that provide televised reconnaissance information. Remote control of combat vehicles (for example, tanks and underseas craft) is being extensively studied, with visual information for operators derived primarily from video sensors. In military command and control centers, video pictorial displays are used for more

effective weapon system deployment and damage assessment. The need to interpret video pictorial information accurately can even be extended to the rapidly growing use of video for educational purposes.

Much of the early psychophysical work dealing with the TV presentation was based upon criteria involving pictorial acceptability for an entertainment medium (8). Compromises in image characteristics were tolerable and the specific effects of these upon the transmission of information to the viewer were not of pressing importance. However, given the more stringent demands for accurate visual interpretation imposed upon the operators of a variety of modern systems, questions can be raised regarding the specific visual effects attributable to the video image and to any compromises in its pictorial quality.

In many future systems uses it is likely that the video images available will be far from photographic in quality owing to extremely long range transmission (thousands or millions of miles) coupled with probable limitations in available power and signal bandwidth (3).

Too little research attention has been given to possibilities of systematic visual distortion or to obtaining quantitative measures of visual judgment associated with the TV image. Recognition of many of the problems is not unique. Alterations in the visual field that can accompany a video presentation have been pointed out as a result of initial studies (14, 18). In one of these studies (14) visual effects

were inferred from changes in psychomotor performance - for a remote handling situation - where the performance field was viewed by video cameras at different angular orientations. The nature of such visual effects is relatively unexplored and it is not generally known how objects and patterns viewed on a video screen may be perceptually distorted and, consequently, contribute to such performance changes.

The present study represents a preliminary attempt to measure the perceptual effects resulting from one form of pictorial degradation and the use of stereoscopic video viewing under varying degrees of degradation. In addition, the relation of video visual acuity to video form perception is also considered.

Since the visual effects to be measured in this study are based upon form thresholds, it is necessary to recognize the pitfalls common to the study of form perception itself (see Hake, 11) as well as those unique characteristics inherent in the use of a video imaging technique. Where possible, these considerations will be related to the present study results.

SECTION 2

METHOD

A. Test Facility

The video equipment utilized consisted of a small camera (12" x 12" x 14"), mounted on an electrically powered chassis, which fed RF signals to two standard 17-inch television monitors in an adjoining room. One monitor was mounted at an Experimenter's (E's) Console and the second was located across the room at the Subject's position (Figure 1). A screen which enclosed the Subject's (S's) position blocked his view of the Experimenter's video monitor. Both monitors always displayed the same video image.

The room containing the mobile camera was painted in black for those areas of floor and walls within the visual field of the camera during experimental trials. Lighting in this chamber was provided from 2 banks of three, 200-watt unfrosted globes mounted on each side of the chamber at the juncture of wall and ceiling. The stimulus materials used were positioned in relation to the six lights so as to minimize shadow effects.

Metal strips on the chamber floor were placed along the path of the moving camera and used to provide a measure of cumulative voltage recorded on a voltmeter at E's console. With an appropriately calibrated mask over the voltmeter, it was possible to read off distances of the TV camera from a group of stimulus forms (for later conversion to threshold measures of visual angles subtended by the forms on the TV screen). E could move the camera remotely from his console position.

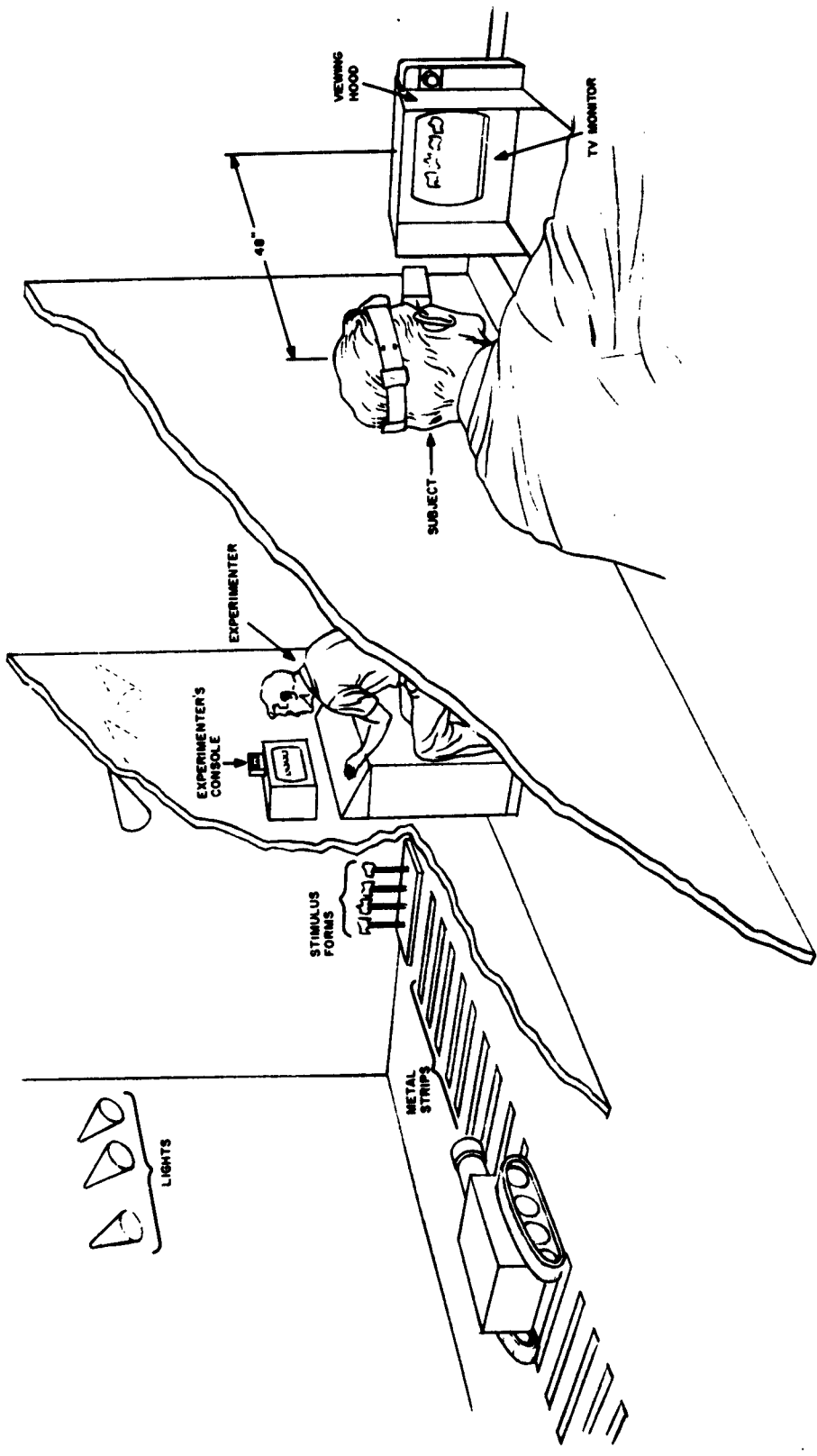


FIGURE 1. TEST FACILITY

A chin rest in front of S's monitor was used to position S's eyes 48 inches from the screen. Two standard 40-watt fluorescent bulbs were lit well behind and above the subject to provide a diffuse low level of illumination. Ambient lighting was necessary since total darkness results in particular problems of glare and visual fatigue for long-term viewing of a TV screen that presents white objects against a black background. The subject's monitor was placed on a black surface and a hood shielded the viewing screen from any noticeable reflections (see Figures 2 and 3).

For normal (that is, Non Stereo) viewing, a 35-mm lens with a visual cone of 30 degrees was used. In order to obtain the stereo presentation a 35-mm binocular lens attachment with a 60-mm inter-objective distance and a visual cone of 30 degrees was used in place of the single lens. This binocular lens projected two side-by-side images on the viewing monitor. A stereo screen consisting of two polarized filters was hinged within the hood on S's monitor and could be swung into place over the stereo images.

A head set containing adjustable right angle prisms, with appropriately oriented polarizing filters, was worn by the subject to achieve the stereo effect (Figure 4).

B. Stimulus Forms

A set of 4 solid stimulus forms were constructed from balsa wood and painted white (Figures 5 and 6). It was intended that these be unfamiliar, non-representational forms in order to avoid the systematic effects upon threshold known to exist with familiar form faces (2, 13).



FIGURE 2. NON-STEREO PRESENTATION



FIGURE 3. STEREO PRESENTATION



FIGURE 4. HEAD-SET FOR STEREO VIEWING

The method of construction for the set of forms is somewhat related to the probability method of Fitts, et al. (9) for their 2-dimensional "metric figures." A detailed explanation of the form construction is presented in Appendix I. Randomization in construction was attempted within the constraints of equivalent figure size and "complexity" (that is, number of contour turns). From this set of 4 solid forms a randomly chosen set of 4 form faces was assigned for Non-Stereo presentation and a second set of 4 faces for the Stereo presentation. Each set of 4 faces constituted the standard stimuli for determining thresholds. On the basis of their construction, the random selection of faces and unbiased assignment to a set, it was assumed that the individual form faces chosen within each set, and the two groups of 4 form faces assigned to each set, were equivalent in "recognition value."

C. Resolution Degradation

The subject's 17-inch monitor was modified so that the video picture could be degraded from 0 percent (optimum resolution) to 40 percent resolution. A calibrated mask was placed over the Fine Tune Control allowing for setting of the resolution level at one of four arbitrarily selected conditions of 0, 10, 20, or 40 percent degradation. A more detailed explanation of the degradation method is given in Appendix II. Also included is a graph showing the relationship between the present method of reducing pictorial information and the equivalent loss of pictorial information that would be achieved by a reduction in scan lines. It can be noted that the

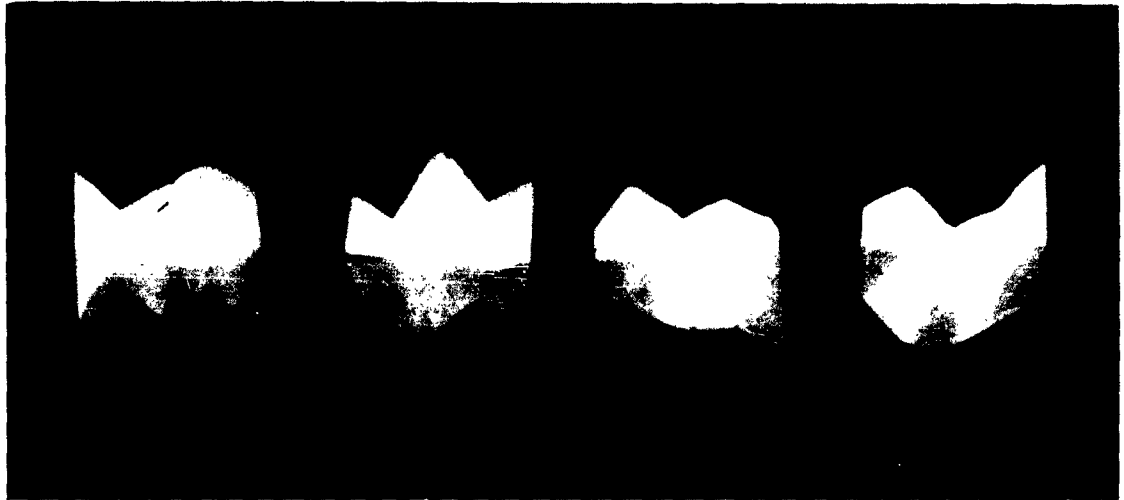


FIGURE 5. NON-STEREO FORM FACES

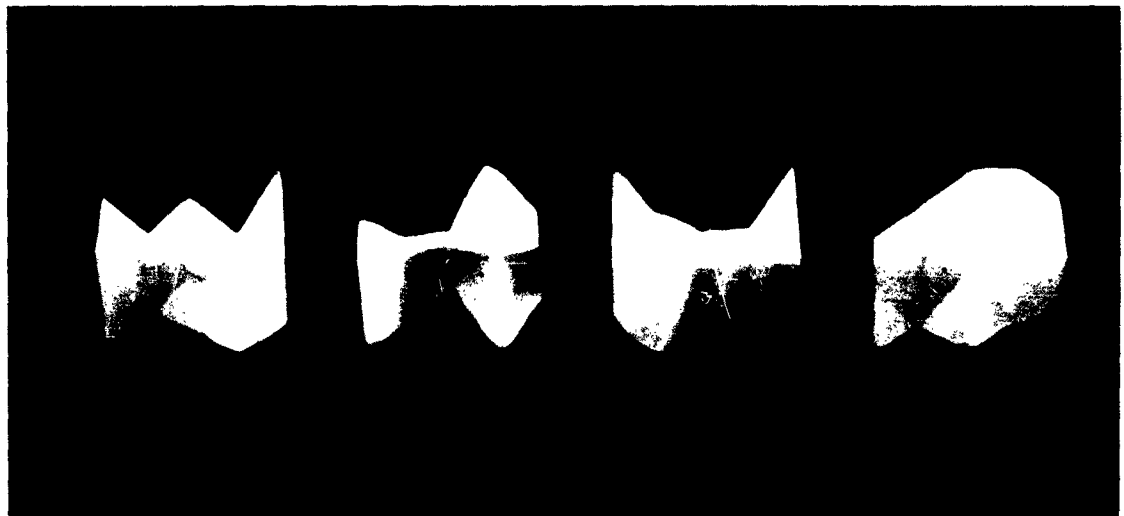


FIGURE 6. STEREO FORM FACES

relationship is approximately linear down to the 40-percent degradation level (the lowest level possible with the present viewing monitor).

Image degradation was based upon a reduction of information per horizontal scan line - by appropriate reductions in bandwidth. The stimulus forms were so constructed and presented, that their contours varied only in the horizontal dimension (see Figures 5 and 6). Consequently, degradation was introduced only in that particular dimension for which visual discrimination was required by the subject (i.e., the horizontal). Thus, if the 4 forms were rotated 90 degrees and scan lines were reduced in number to obtain the degradation effect perceptual results should, theoretically, be the same. (Such a perceptual similarity would still, however, be open to empirical test.)

D. Subjects

Seven adults - 6 males and 1 female - comprised the subject group. One of the S's wore eyeglasses during all of the experimental trials. The purpose of the study was carefully explained and practice trials utilized to ensure subject familiarity with the task. The following is an abstract of the instructions given.

"Your job is to observe the screen, on which you will first be shown a single form. Here is a sample of the type of form you will see. This is called the Test Form. (Show the Test Form.)

You will be permitted to study the form for about a minute so that you can recognize it well. Right after this you will see 4 forms presented on the screen. (Show the 4 forms.)

Among these 4 will always be the Test Form that you had studied. One of the things you'll be asked to do is to pick this form from among the 4. At first the 4 forms will look like "blobs" and you won't be able to distinguish one from the other. But while you are looking at the screen, you'll see the forms move closer in small steps and they'll grow larger on the screen so that gradually you'll notice the difference between them.

When you first see the forms as distinctly different from one another, let me know. This probably will not be the point at which you can pick the test form - but the point at which the 4 forms all appear to have distinctive aspects to them and you can make out differences in their outline.

Your next task, as the forms continue to grow larger on the screen, is to correctly identify the Test Form when you are certain that you can see it. At that point give its location from the left, on the screen (that is, first from the left, second from the left, third, or fourth).

On some trials the 4 forms will be clearly in focus and on others they'll be out of focus. Your job is always the same -- to report when the forms are distinctly different from one another and to identify the Test Form when you are certain that you can pick it out from the group of four.

If that's clear, then here is the first Test Form.

Remember to study it carefully so that you can recognize it later."

(Remove screen cover.)

The instructions were essentially duplicated for the series involving the Stereo presentation, following a training period on the use of the stereo head-set.

SECTION 3

PROCEDURE

A. Form Threshold Measures

The subject was seated with the head positioned on a fixed chin rest and the eyes 48 inches from the TV monitor. A trial consisted of S being shown a standard stimulus form face on the screen for 60 seconds. This was a close-up view with optimum resolution and the form subtending a visual angle of 2° for the Non-Stereo condition (1° for the Stereo condition, owing to the split image*). Following the presentation of the standard, the screen was covered for approximately 12 seconds and then uncovered to reveal the 4 forms which appeared below identification threshold. One of these 4 forms presented the exact face toward the camera and in the exact orientation that S had seen it when the form was presented as the standard.

A modified method of limits was utilized with an ascending series only, so that the camera always moved toward the forms during a trial. As pointed out by Helson (13), a descending series has been found undesirable in such a situation since it begins with the subject's knowledge of the form to be recognized. Two correct consecutive recognition responses

* Equipment limitations made it difficult to equate the size of the standard stimulus forms on the screen. However, the possible biasing effects of presenting the Non-Stereo and Stereo standard stimuli as different in size should have been minimal since all essential form features for recognition seemed as readily visible for the 1° size as for 2° . In addition, the study results will show that form recognition was achieved at about 0.5° or less under any of the viewing conditions.

were always required since, although this tends to raise thresholds somewhat, it has the advantage of stabilizing them (13).

On each trial there were two different threshold responses required of S. One was a Discrimination Threshold consisting of the point at which S reported the four forms as distinct from another in outline or shape. (That is, when they no longer appeared as "shapeless blobs.") The other was a Recognition Threshold, defined as the point at which the standard form was recognized with certainty, by designating its position among the four forms.

Before each trial the faces of the three forms, accompanying the standard form face, were rotated 90 degrees about the longitudinal axis so that the subject did not view identical alternatives to the standard on any trial. The position of the standard form face, in relation to the 3 accompanying forms, was randomized on each trial, while percent degradation (0, 10, 20, and 40 percent) and the form face chosen as the standard (Numbers 1, 2, 3, or 4 of the set) were counterbalanced. (It should be kept in mind that the two sets of four form faces for Non-Stereo and Stereo presentation were randomly selected faces from the same 4 solid forms.)

Thus, each of the 7 subjects ran through a total of 16 trials under Non-Stereo and 16 under the Stereo presentation.

The contrast was held as constant as possible at S's monitor and overall screen brightness was equated for the

Non-Stereo and Stereo condition by varying the video brightness setting to compensate for the reduction in brightness resulting from use of the polaroid filters. An integrated screen illuminance value of 4 FT.-C was used and the contrast ratio for white form to black background was maintained at approximately 5.3:1.0.

B. Video Acuity Measures

Video acuity scores were obtained for each subject by moving the camera in increments toward a white, solid Landolt Ring target. Lighting of the object, screen brightness, and contrast were maintained the same as for the Non-Stereo form threshold situation. The subject was required to report the location of the ring gap at either the 3, 6, 9, or 12 o'clock position. As with the form threshold procedure, a modified method of limits was used with two consecutive correct position identifications required to complete the trial. A total of 8 trials per subject (2 trials for each of the 4 gap positions presented in random order) was used to obtain a video acuity score for each subject. The position variable was isolated on the basis of observations during the course of the study, indicating that orientation of the gap in relation to scan lines seemed to influence acuity.

Subject reports of how they achieved discrimination and recognition were elicited at the end of each session, including details of any form associations that they used to assist in the recognition process.

SECTION 4

RESULTS

A. Recognition Threshold

The curves showing mean Recognition Threshold scores for the different forms and the different resolution degradation levels, under Non-Stereo and Stereo viewing, are plotted in Figure 7. A separate analysis of each viewing condition is shown in Tables I and II. Based upon these results, the equivalence of the form faces within each set cannot be supported, since their threshold values differ significantly.

A single, overall, analysis that groups the forms and tests Viewing Method as a main effect might be open to question. However, if one accepts the possible equivalence of the two sets of forms, owing to the random assignment of faces to each set, then a comparison of the overall mean threshold scores for Non-Stereo ($M = 26.78'$ of visual angle) and Stereo ($M = 24.61'$ of visual angle) viewing would be of interest. The curves in Figure 7 indicate a general threshold decrease for Stereo but the difference between the means of the two Viewing Methods of 2.17 minutes of visual angle is not a significant one.

It can be noted that under Non-Stereo and Stereo viewing a highly significant difference was obtained between Degradation Levels. Inspection of the curves in Figure 7 shows this distinct threshold increase with increasing degradation. In addition a significant interaction effect

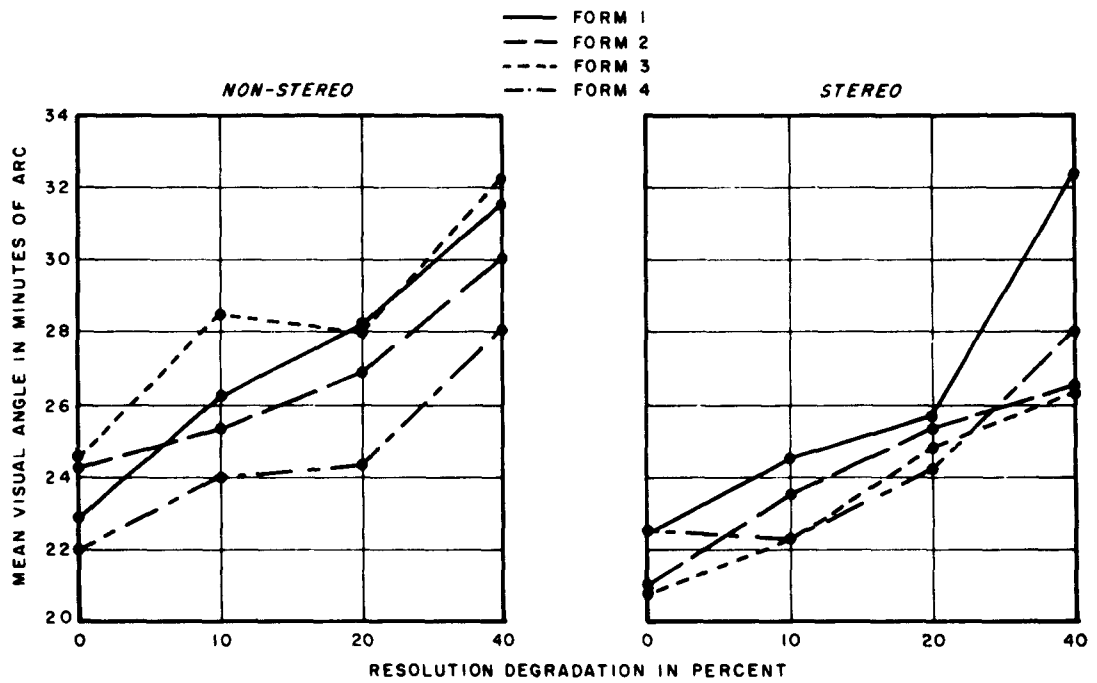


FIGURE 7. MEAN RECOGNITION THRESHOLD SCORES FOR NON-STEREO AND STEREO VIEWING OF FOUR STIMULUS FORM (N = 7)

was obtained, under the Stereo viewing condition, between Degradation Levels and Form Faces (P. = .05).

TABLE I
 SUMMARY OF ANALYSIS OF VARIANCE OF
 NON-STEREO RECOGNITION THRESHOLD SCORES FOR
 DEGRADATION LEVEL AND FORM FACES

Source of Variation	df	Mean Square	F
Degradation Levels (D)	3	243.0	24.3**
Form Faces (F)	3	66.3	4.3*
Subjects (S)	6	147.0	
D x F	9	4.8	.59
D x S	18	10.0	
F x S	18	15.4	
Error (D x F x S)	54	8.1	

* Significant at the 0.05 confidence level.

** Significant at the 0.01 confidence level.

TABLE II
 SUMMARY OF ANALYSIS OF VARIANCE OF
 STEREO RECOGNITION THRESHOLD SCORES FOR
 DEGRADATION LEVELS AND FORM FACES

<u>Source of Variation</u>	<u>df</u>	<u>Mean Square</u>	<u>F</u>
Degradation Levels (D)	3	229.3	23.4**
Form Faces (F)	3	39.3	9.4**
Subjects (S)	6	101.1	
D x F	9	11.0	2.04*
D x S	18	9.8	
F x S	18	4.2	
Error (D x F x S)	54	5.3	

* Significant at the 0.05 confidence level.

** Significant at the 0.01 confidence level.

B. Discrimination Threshold

An analysis of these threshold data allow for a comparison between Viewing Methods and Degradation Levels in a Subjects by Treatments design with replication. Form, as a separate variable, is not pertinent here since the Discrimination Thresholds judgments made were independent of the presence of any particular form face (whereas for the Recognition Threshold identification of the form itself was essential). Thus an overall analysis of Viewing Methods is applicable.

The curves in Figure 8 show the significant increase in threshold with reductions in resolution level. As shown in the summary of the analysis in Table III, the mean threshold value for the Stereo Viewing condition is now found to be significantly lower than for the Non-Stereo condition. A significant interaction is also found between Viewing Method and Pictorial Degradation which is almost entirely the result of a proportionally smaller difference between the two viewing conditions for the optimum (0 percent) degradation level. That is, the Discrimination Threshold is barely affected by a Stereo, or Non-Stereo view when there is no resolution degradation present, whereas this difference is proportionally greater when any degree of degradation is introduced.

Since there might be some question of whether systematic learning would tend to effect subject performance during a series of trials, a comparison was made between threshold scores during the first and second half of each series of trials for Non-Stereo and Stereo Viewing conditions. None of the differences between mean threshold scores of the split halves of any series of trials were found to approach significance.

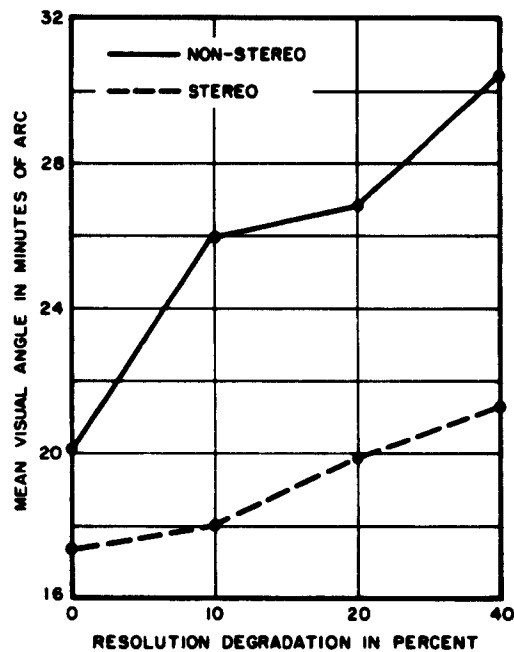


FIGURE 8. MEAN DISCRIMINATION THRESHOLD SCORES FOR NON-STEREO AND STEREO VIEWING (N = 7)

TABLE III
 SUMMARY OF ANALYSIS OF VARIANCE OF
 DISCRIMINATION THRESHOLD SCORES FOR
 VIEWING METHODS AND DEGRADATION LEVELS

Source of Variation	df	Mean Square	F
Viewing Methods (M)	1	2513.0	18.6**
Degradation Levels (D)	3	378.0	19.7**
Subjects (S)	6	55.2	
M x D	3	126.7	13.9**
M x S	6	134.7	
D x S	18	19.2	
M x D x S	18	9.1	2.2**
Within	169	4.1	

** Significant at the 0.01 confidence level.

A comparison of interest is the overall percentage reduction in threshold scores with corresponding percentage reductions in pictorial resolution; which results in the following:

TABLE IV
 PICTORIAL RESOLUTION

	<u>Percent Change</u>		
Pictorial Degradation	10	20	40
Recognition Threshold Reduction	9	16	31
Discrimination Threshold Reduction	18	24	35

While the percentage reductions closely parallel one another the Discrimination threshold shows a greater effect of pictorial degradation. This is understandable, since the Discrimination threshold is more dependent upon the form outline, and is more readily affected by even a minor degree of blurring of the form edges. Recognition on the other hand is probably dependent upon gross contours along with features in the "body" of the form such as discernible shadows and associated form facets.

C. Video Acuity

Mean Video Acuity scores obtained for each subject were correlated with form threshold scores for Recognition and Discrimination thresholds under the Non-Stereo and Stereo viewing conditions. The Rho's are not significant which is in accord with previous work indicating little relationship between acuity and form perception (19).

<u>RANK CORRELATIONS</u>				
<u>Acuity vs Form Thresholds</u>				
	Discrimination		Recognition	
	Non-Stereo	Stereo	Non-Stereo	Stereo
Acuity	.21	.04	.21	.11

Comparison of the gap positions for the Landolt Ring presentation supports the hypothesis that orientation of the gap, in relation to scan lines, influenced video pattern acuity for

these particular viewing conditions. The summary of the analysis in Table V indicates a highly significant difference between gap positions, with the 3 and 9 o'clock positions resulting in significantly lower threshold scores ($M = 2.74$ and $M = 2.58$ minutes of visual angle) than the 6 and 12 o'clock gap positions ($M = 4.94'$ and $6.58'$ respectively).

Mean threshold acuity for the 3 and 9 o'clock positions are approximately twice the size of those found under normal binocular viewing conditions with standard acuity tests (about 1.0' to 1.5' of visual angle). The result is understandable in terms of difficulties in optimizing video contrast and brightness, annoyance of scan lines, and the general flux that appears in any video image. However, the much sharper reduction in threshold for 6 and 12 o'clock positions requires further explanation.

TABLE V
SUMMARY OF ANALYSIS OF VARIANCE OF
VIDEO ACUITY SCORES FOR
LANDOLT RING GAP POSITIONS

Source of Variation	df	Mean Square	F
Between gap positions (G)	3	5.1	23.2**
Between subjects	6	1.5	
Residual	46	0.22	

** Significant at the 0.01 confidence level.

Lighting and shadow contributed little (if anything) to this result since there were no observable shadows cast at any of the gap positions and shifting of the light sources (but maintaining diffuse lighting) fails to change the result. A more likely explanation lies in the nature of the video image with its horizontal scan lines and the tracing of the image by a scanning beam. In effect, the scan line does not break off sharply, between white (signal on) and black (signal cut-off) tones as it sweeps horizontally, but tends to "bleed" over.* On the other hand, scan lines being discrete, break sharply between tonal variations in the vertical plane.

D. Form Association

At the completion of a series of trials the S's had been asked for any form association used. (Many of these were voluntarily given during the trials without being requested.) All S's tended to associate the forms with familiar figures although there was little consistency in the form names applied; ranging from a "Whale" and an "Admiral's Hat" to a "Ski Jump" or the "Taj Mahal" for one form. One subject, although he could give associations, said

* It can be shown that signal rise time in a commercial TV monitor can result in an overlap in tonal change (white to black) of some 1.7 minutes of visual angle at a viewing distance of 48 inches. (Phosphor persistence and blooming of the video spot at signal cutoff could possibly account for another portion of the difference between gap positions.)

that he did not use these for identification purposes. Instead he tried to visualize the standard form as it might be when "distant" (or small) with a "vague outline," and then to memorize the "general shape." This subject also happened to achieve the best Recognition Threshold scores.

SECTION 5
DISCUSSION

The results of the present study indicate that reductions in the perception of visual information can be influenced significantly by reductions in the information elements of the physical (i.e, video) image. Quantitative changes in the achievement of form recognition, or what Hake might term "coherence" (11), appear to be highly sensitive to the signal fidelity of the pattern, or form, to be recognized.

Complex forms, to be recognized on a video screen, should subtend at least 20 minutes of visual angle, when resolution is equivalent to that of a standard, 525 scan line, video monitor. Where even minor pictorial degradation is introduced, this value should be increased to 30 minutes of visual angle or greater. Steedman and Baker (17) using complex forms (photographically presented), had concluded that a visual angle of about 20 minutes would be required for target recognition under operational reconnaissance conditions.

Uniform reductions in information on a video screen may also have a selective effect upon particular shapes or patterns at threshold levels, as shown by the results with the Stereo viewing condition. While there is little research dealing with this effect for unfamiliar forms, the result has been known to occur at threshold levels for familiar geometric patterns (2, 13).

Previous research has dealt with pattern degradation by blurring or distorting of figures (11). Comparison of

such studies with the present results would be tenuous, owing to the fact that the methods of form presentation were based upon very brief exposure intervals and/or required reproduction of the image (by drawing), rather than direct measures of threshold value. However, similarities between findings in these studies regarding how recognition is achieved and the reports given by the subjects of the present study were observed. [that is, the often cited "orderly progression" which leads to visual recognition through such visual stages as "lack of differentiation," "form lability," use of "association," etc. (6, 11)].

The finding that Stereo produced a significant reduction in Discrimination Thresholds (but not in Recognition Threshold) does not appear to be attributable to stereoscopic enhancement in the image itself. A more likely explanation lies in the enhancement of form contours, basic to the Discrimination Threshold judgments, resulting from the use of the stereo polaroid filters. Despite the equating of overall screen brightness and an attempt to maintain approximately the same contrast ratio, the neutral polaroid filters still lent a diffuseness to the video image that tended to reduce the effects of glare. In addition, the appearance of the scan lines which were, on occasion, annoying to several of the subjects under Non-Stereo viewing, were no longer a problem when the dual stereo images were polarized.

Failure to achieve significant stereo enhancement for Recognition Thresholds is a more pertinent result. This identification measure is the customarily used one and is more "objective" in that some verification of the subjects judgment is obtained by his correct identification of a standard stimulus form from among a group of stimulus forms. Unfortunately, there were shortcomings in the stereo viewing system which became apparent during the course of the study and are felt to have reduced the opportunity for clearly demonstrating the properties of stereoscopic viewing. These shortcomings included cumbersome-ness and discomfort produced by wearing of the head set, along with having to maintain a fairly fixed head position. Optimum adjustment by the subject of right angle prisms required to place dual images the proper distance apart and then fusing these to form a single stereo image, requires a reasonable level of skill before it can be mastered (and is probably not accomplished equally well by all subjects). Also common to any stereo system used with video is the visual fatigue that subjects report when required to fuse a changing image with relatively poor pictorial quality, for any extended period or for repeated image fusion over a large number of brief viewing periods. In addition, the use of a white form that appears suspended against a black background provided few of the customary two-dimensional cues that normally enhance depth effects.

Despite the limitations, three of the seven subjects reported that they could detect and did utilize facets of the form, in the stereo plane, for recognition. The other subjects claimed that they did not feel they were using any of the depth cues which they perceived but depended upon the form outline (and shadows created on the form face by the various facets) in the same manner as for the Non-Stereo viewing condition. Different performance measures and changes in the form presentation conditions would seem to be in order for future research on video stereo effects.

Research in visual recognition of forms is said to have value only if one can specify the pertinent perceptual dimensions of the forms used. The present study demonstrates that this can be difficult to achieve on a logical basis, or from the findings of previous form perception studies. Thus, the aspects of the stimulus forms used here, which constituted the bases of their supposed similarity, were equivalence in size, number of contour turns, asymmetry and angularity of contour. These have been shown to account for the major proportion of variance constituting form complexity judgments (1, 4). The rest of the possible biasing form components should have been randomized by the construction method used and the random selection of the form faces.

One important reason that may account for the difficulty in achieving form equivalence is that previous studies depended upon judged complexity variables which may not, in fact, be the most important form variables for predicting

identification thresholds. The nature of the stimulus material may be another reason for the disparity, in that a pictorial TV image presents many qualitative differences from the two-dimensional form silhouette or outlines commonly used in former studies. [Gibson argues strongly for such differences in perceptual units as well as doubting the validity of comparing these abstractions to "real" three-dimensional solids (10).]

Just how much of a role in determining form equivalence may be assigned to association is not clear. It is certainly difficult to control association in a visual presentation which requires recognition of unfamiliar patterns. The tendency of subjects to impart meaning to ambiguous stimulus forms is consistent (15). If the subject anchors too intensely upon some aspect of the form that is systematically affected by pictorial degradation, than recognition of the form at threshold levels could, in turn, be affected. Evidence to support such a contention is contradictory and perceptual distortions have been shown to occur whether specific form associations are used or not (12, 15). There is cause to wonder whether truly "random" non-familiar patterns could ever prove to be perceptually equivalent as long as uncontrolled association may produce a biasing effect.

As stressed previously, when using an electronic imaging technique, for studying form perception, one inherits all of the difficulties found in form perception methodology as well as those little understood visual variables character-

istic of the particular electronic system. Some of the unique video variables, such as the effects of the video presentation on pattern acuity, require further clarification.

Many of the results that are achieved are not only dependent upon the methods of measurement employed but in large measure upon the particular video system used to present the stimuli. Admittedly this would seem to limit generalizations of any visual results derived from one type of video system, to a variety of other systems or techniques. For example, systems that employ greater numbers of scan lines, better resolution, or the use of slow scan techniques could all influence perceptual results. The generalizations when drawn must, of necessity, be rather broad ones or be highly specific to a particular system and a particular technique. In either event, greater knowledge, in the form of quantitative determinations concerning an individual's ability to interpret what is presented under various video pictorial conditions would seem essential before such pictorial visual information can be counted upon as the basis of critical judgment.

SECTION 6

SUMMARY

Form Recognition and Discrimination Thresholds were studied as a function of 4 levels of Video Resolution Degradation and 2 Viewing Methods (Normal Binocular Viewing and Stereo Viewing). Measures of pattern acuity for a video presented target (Landolt Ring) were also obtained and correlated with the Threshold measures.

Four solid stimulus forms were randomly constructed within the constraints of equivalent size and "complexity" (number of contour turns), asymmetry, and angularity of contour. Two sets of form faces were randomly chosen from the four forms - one set for Non-Stereo and one set for Stereo presentation.

Results clearly indicated the association of threshold increases with corresponding percentage degradations in the video image. Recognition Threshold values were not significantly enhanced by the particular stereo system used, whereas Discrimination threshold values - which depended primarily upon contour discrimination - were significantly improved.

Correlations between video pattern acuity and form threshold scores were low and insignificant, generally confirming previous findings for non-video form perception. However, there are unique characteristics of the video image, such as the orientation of the pattern in relation to scan lines, which can affect visual acuity.

The randomly generated and randomly assigned form faces differed significantly from one another within each set of faces chosen. There is serious doubt that forms can be equated for visual threshold values on the basis of form complexity variables determined from previous studies or by some degree of randomization in their construction. The number of significant variables that make up the basic perceptual dimensions of a form and their interacting effects require far more clarification.

APPENDIX I
CONSTRUCTION OF STIMULUS FORMS

The four stimulus forms were developed by first generating 2-dimensional asymmetrical form faces. Each face was constructed by randomly selecting units from 1 to 4 to represent the distance (in squares on graph paper) from the form centerline. These distances from the centerline were taken at 5 equidistant points. Two such form faces were plotted and set at right angles (in "North"- "South," "East"- "West" planes). This procedure is shown in the diagram below (Figure A-1).

It can be seen that there are now four "layers," in each plane which constitute the entire form. By cutting each layer from balsa wood and then gluing these layers together, a fairly complex 3-dimensional and non-representational solid form is obtained. The three-dimensional layer-by-layer cross sections as they would appear viewed from above are shown in the diagram, below the 2-dimensional form faces.

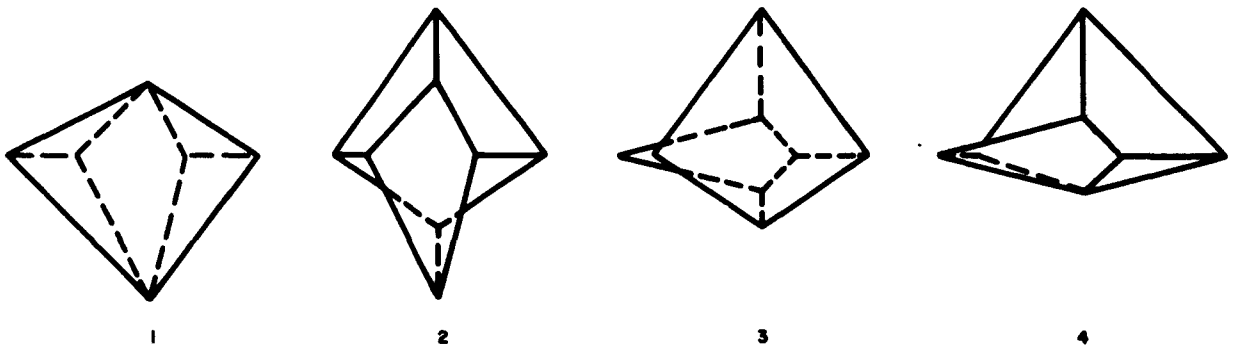
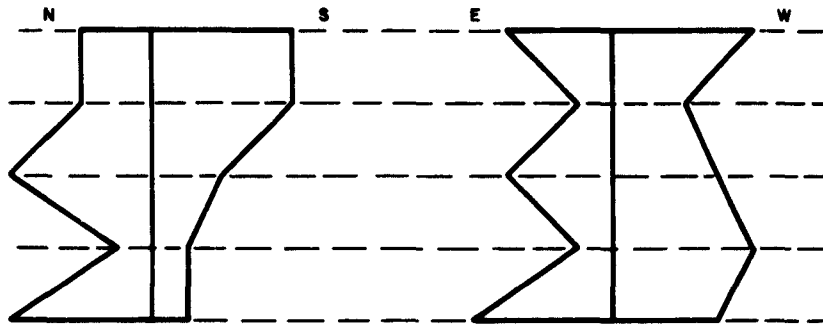


FIGURE I-1. DIAGRAM OF FORM CONSTRUCTION

APPENDIX II

DEGRADATION OF VIDEO RESOLUTION

The resolution of a transmitted television picture may be expressed by the following relationship:

$$N_r = \frac{w}{h} m(kt_v n)^2$$

where

N_r = number of picture elements resolved per frame,

$\frac{w}{h}$ = aspect ratio,

m = ratio of horizontal to vertical resolution,

k = vertical resolution factor, or ratio of number of elements resolved vertically to number of active scan lines,

t_v = ratio of active to total vertical field time,

n = number of scan lines per frame.

The number of elements resolved is not equal to (in fact, it is less than) the number of elements transmitted.

This is due to the fact that elements are transmitted during retrace, and also because k is less than 1.

The total number of elements transmitted per frame, N_t , may be shown to be equal to:

$$N_t = k \left(\frac{t_v}{t_h}\right) \left(\frac{w}{h}\right) mn^2 = N_r / kt_v t_h$$

where t_h = ratio of active to total horizontal scan time.

The bandwidth required to transmit the video signal is expressed by:

$$BW = \frac{1}{2} N_t f = f N_r / 2kt_v t_h = \frac{f}{2} kmn^2 \left(\frac{t_v}{t_h}\right) \left(\frac{w}{h}\right) \text{ cycles/second}$$

It may be seen from the previous equation that the bandwidth is linearly related to the number of elements resolved per frame, N_p . It is also true, although perhaps not so immediately apparent, that the bandwidth is also linearly related to the number of scan lines per frame, n . If n is changed, either w/h , m , or k will change in an inverse manner. For example, if n is doubled either w/h , m , or k will be halved. Therefore, the bandwidth can be linearly related to the number of scan-lines per frame.

By operating on the standard television signal received at the monitor, it is possible to reduce the bandwidth of the displayed signal and thus simulate the condition of degraded resolution (16).

In the video system used, the resolution degradation effect was simulated by detuning the local oscillator slightly so that the higher frequency components of the received signal are eliminated in the mixer. This technique acts to reduce the horizontal resolution. That is, the number of scan lines per frame in the actual displayed television picture will remain at 525 lines, and the bandwidth of the information within a scan line will be decreased.

The method used to detune the local oscillator involved adjusting the bandwidth of the IF signal, out of the VHF tuner, by using the fine tuning control on the receiver. Prior to the beginning of the tests, the fine tuning control was calibrated so that various stages of resolution degradation were easily reproducible.

The degree to which the IF bandwidth was narrowed may be found from the last equation relating bandwidth and resolution. Substituting values for standard broadcast television:

$$BW = \left(\frac{30}{2}\right)(0.707)(1)(525)^2\left(\frac{0.03}{0.85}\right)\left(\frac{4}{3}\right) = 4.5 \times 10^6 \text{ cycles/second}$$

By decreasing the bandwidth in steps of 0.45 Mc, the resolution was degraded down to 40 percent of normal. By use of a standard television test pattern, the validity of this method of resolution degradation can be further demonstrated.

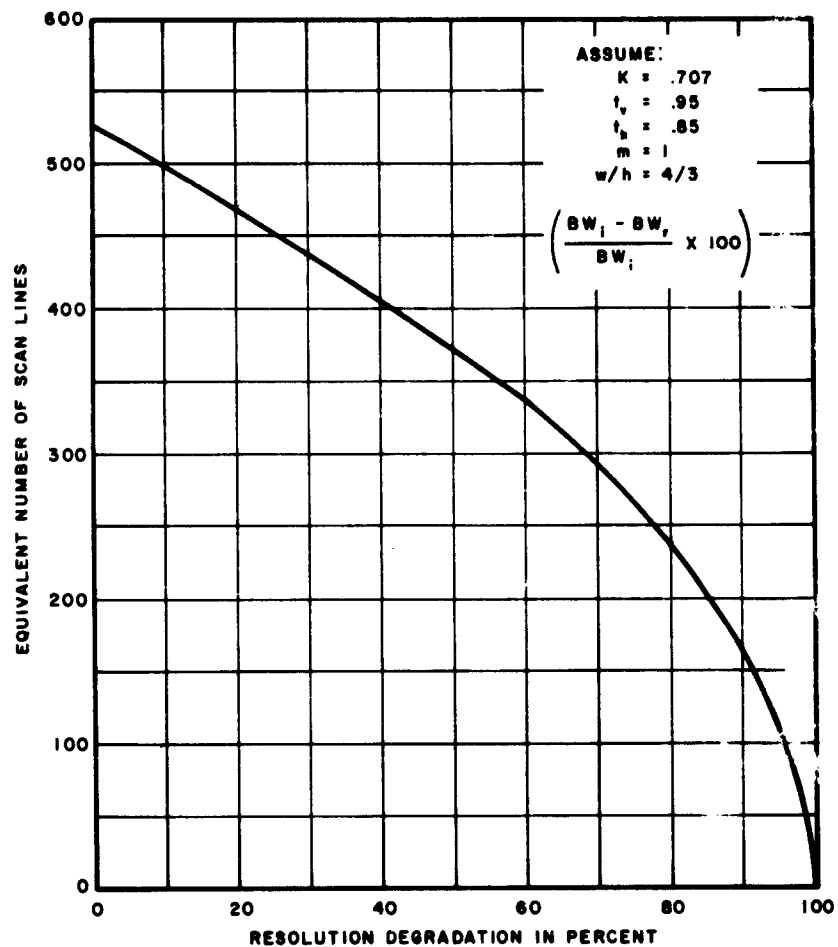


FIGURE II-1. PLOT OF RESOLUTION DEGRADATION VS EQUIVALENT SCAN LINE REDUCTION

APPENDIX III

REFERENCES

1. Attneave, F., Physical determinants of the judged complexity of shapes. J. Exp. Psychol., 1957, 53, 221-227.
2. Casperson, R. C., The visual discrimination of geometric forms. J. Exp. Psychol., 1950, 40, 668-681.
3. Chomet, M., Freeberg, N., and Swanson, A., A Simulation Study of Operator Capability in Robot Vehicle Control. In 1962 IRE International Convention Record, Part 9, pp 213-219, March 26-29, 1962, IRE, New York, N. Y.
4. Coules, John, Duva, J. S., Ganem, G., Effect of Visual Noise on the Judgement of Complex Forms, USAF, Operational Applications Office, AFCCDD-TR60-40, Bedford, Mass., November 1960.
5. Crook, M. N. and Coules, J., The Effect of Noise on the Perception of Forms in Electro-Visual Display Systems: Interim Report No. 7. Reduced contrast and contour degradation as factors in impairment of form recognition, Institute for Applied Experimental Psychology, Tufts University, January 27, 1959.
6. Douglas, A. G., A tachistoscopic study of the order of emergence in the process of perception. Psychol. Monogr., 1947, 61, No. 287.
7. Edwards, A. L., Experimental Design in Psychological Research. New York: Rinehart, 1950.
8. Ingstrom, E. W., A study of television image characteristics. Proceedings of the IRE, 21, No. 12, December, 1933, p 1631-1651.
9. Fitts, P. M., Meyer, W., Rappaport, M., Anderson, N., and Leonard, J., Stimulus correlates of visual pattern recognition: a probability approach. J. Exp. Psychol., 1956, 51, 1-11.
10. Gibson, J. J., What is form? Psychol. Review, 1951, 58, 403, 412.
11. Hake, H. H., Contributions of psychology to the study of pattern vision. USAF, WADC Tech. Rept. 57-621, Dayton, Ohio, October 1957.
12. Hanawalt, N. G., and Demarest, I., The effect of verbal suggestion in the recall period upon the reproduction of visually perceived forms. J. Exp. Psychol., 1939, 25, 159-174.

13. Helson, H. and Fehrer, E., The role of form in perception. Amer. J. Psychol., 1932, 44, 79-102.
14. Martindale, R. and Lowe, W., Use of television for remote control. J. Appl. Psychol., 1959, 43, 122-124.
15. Prentice, W., Visual recognition of verbally labeled figures. Amer. J. Psychol., 1954, 67, 315-320.
16. Schwartz, M., Information Transmission Modulation and Noise, McGraw-Hill, New York, 1959.
17. Steedman, W. C. and Baker, C. A., Target size and visual recognition. WADD, Tech. Rept. 60-93, February 1960, Wright-Patterson AFB, Dayton, Ohio.
18. Smith, W., Smith, K., Stanley, R., and Harley, W., Analysis of performance in televised visual fields: preliminary report. Percept. Mot. Skills, 1956, 6, 195-198.
19. Vernon, M. D., Different types of perceptual ability. Brit. J. Psychol., 1947, 38, 79-89.

Electronic Systems Div. L. G. Hanscom Field, Bedford,
Mass. ESD-TDR-63-136 FORM PERCEPTION IN VIDEO
VIEWING: EFFECTS OF RESOLUTION DEGRADATION
AND STEREO ON FORM THRESHOLDS Feb 63 --p.

Unclassified Report

Thresholds for video viewing of randomly construct-
ed forms are determined under conditions of pictorial
degradation and stereoscopic presentation. Form recog-
nition is not significantly enhanced by stereo viewing
but discrimination threshold is reduced. All form thresh-
olds are highly sensitive to video image degradation.

Video acuity is not significantly related to video
form thresholds and can be affected by orientation of a
pattern in relation to video scan lines.

1. Viewing Screens
2. Perception, Pictures
3. Thresholds

- I Project 9674
- II Contract AF19(628)-328
- III Airborne Instruments Lab.
Cutler Hammer Inc.
Long Island, N. Y.
- IV Norman E. Freeberg
- V In ASTIA Collection

Electronic Systems Div. L. G. Hanscom Field, Bedford,
Mass. ESD-TDR-63-136 FORM PERCEPTION IN VIDEO
VIEWING: EFFECTS OF RESOLUTION DEGRADATION
AND STEREO ON FORM THRESHOLDS Feb 63 --p.

Unclassified Report

Thresholds for video viewing of randomly construct-
ed forms are determined under conditions of pictorial
degradation and stereoscopic presentation. Form recog-
nition is not significantly enhanced by stereo viewing
but discrimination threshold is reduced. All form thresh-
olds are highly sensitive to video image degradation.

Video acuity is not significantly related to video
form thresholds and can be affected by orientation of a
pattern in relation to video scan lines.

1. Viewing Screens
2. Perception, Pictures
3. Thresholds

- I Project 9674
- II Contract AF19(628)-328
- III Airborne Instruments Lab.
Cutler Hammer Inc.
Long Island, N. Y.
- IV Norman E. Freeberg
- V In ASTIA Collection

Electronic Systems Div. L. G. Hanscom Field, Bedford,
Mass. ESD-TDR-63-136 FORM PERCEPTION IN VIDEO
VIEWING: EFFECTS OF RESOLUTION DEGRADATION
AND STEREO ON FORM THRESHOLDS Feb 63 --p.

Unclassified Report

Thresholds for video viewing of randomly construct-
ed forms are determined under conditions of pictorial
degradation and stereoscopic presentation. Form recog-
nition is not significantly enhanced by stereo viewing
but discrimination threshold is reduced. All form thresh-
olds are highly sensitive to video image degradation.

Video acuity is not significantly related to video
form thresholds and can be affected by orientation of a
pattern in relation to video scan lines.

1. Viewing Screens
2. Perception, Pictures
3. Thresholds

- I Project 9674
- II Contract AF19(628)-328
- III Airborne Instruments Lab.
Cutler Hammer Inc.
Long Island, N. Y.
- IV Norman E. Freeberg
- V In ASTIA Collection

Electronic Systems Div. L. G. Hanscom Field, Bedford,
Mass. ESD-TDR-63-136 FORM PERCEPTION IN VIDEO
VIEWING: EFFECTS OF RESOLUTION DEGRADATION
AND STEREO ON FORM THRESHOLDS Feb 63 --p.

Unclassified Report

Thresholds for video viewing of randomly construct-
ed forms are determined under conditions of pictorial
degradation and stereoscopic presentation. Form recog-
nition is not significantly enhanced by stereo viewing
but discrimination threshold is reduced. All form thresh-
olds are highly sensitive to video image degradation.

Video acuity is not significantly related to video
form thresholds and can be affected by orientation of a
pattern in relation to video scan lines.

1. Viewing Screens
2. Perception, Pictures
3. Thresholds

- I Project 9674
- II Contract AF19(628)-328
- III Airborne Instruments Lab.
Cutler Hammer Inc.
Long Island, N. Y.
- IV Norman E. Freeberg
- V In ASTIA Collection