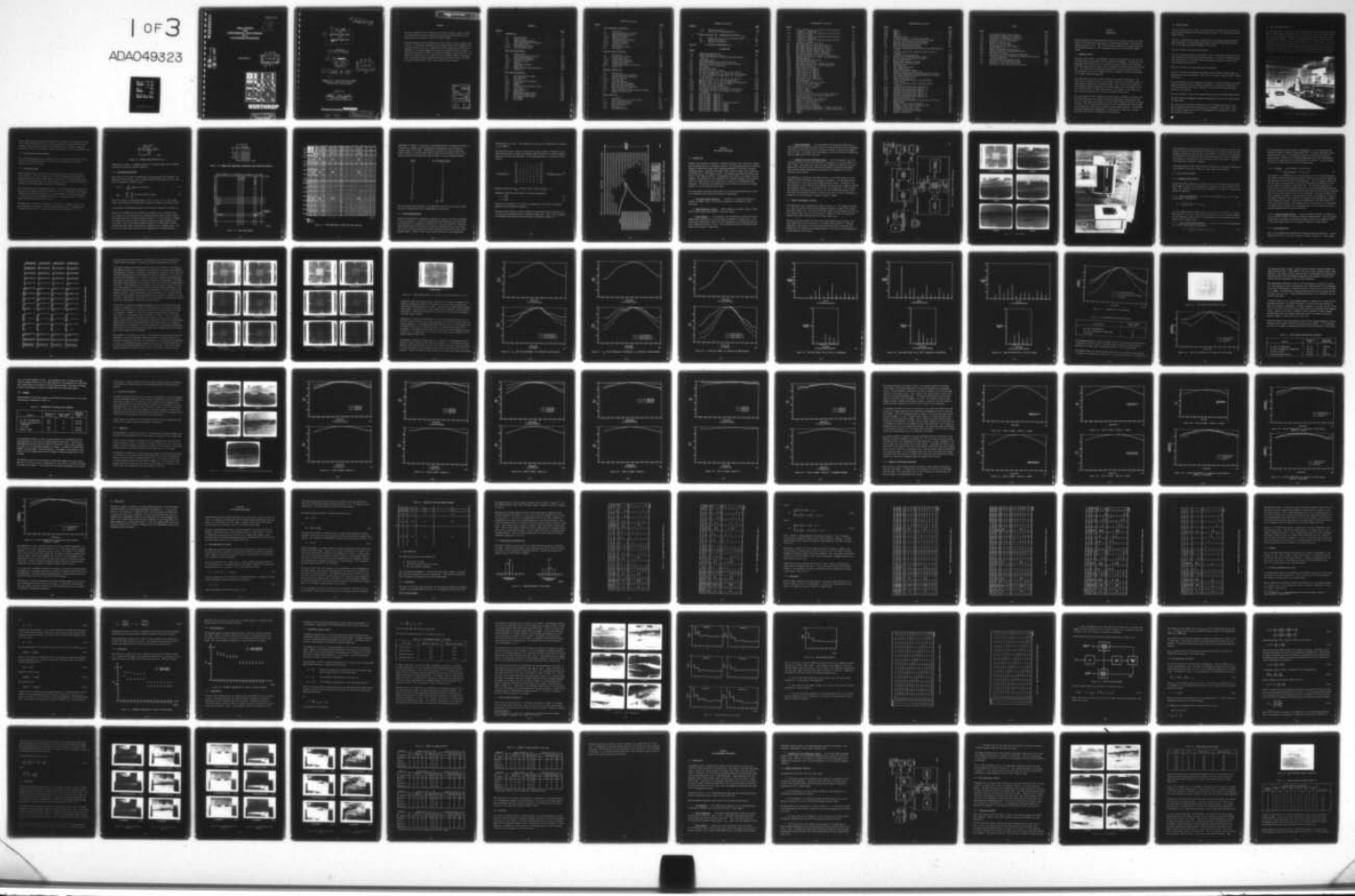
AD-A049 323 NORTHROP CORP ANAHEIM CALIF ELECTRO-MECHANICAL DIV F/G 17/5
EXPERIMENTAL DEVELOPMENT OF A FLIR SENSOR PROCESSOR.(U)
SEP 77 T LEIBOFF, H HENNING, T NODA, B DEAL DAAG53-76-C-0188
UNCLASSIFIED NORT-77Y196 NL

1 OF 3
ADA049323



AD A049323

AD No.
DDC FILE COPY

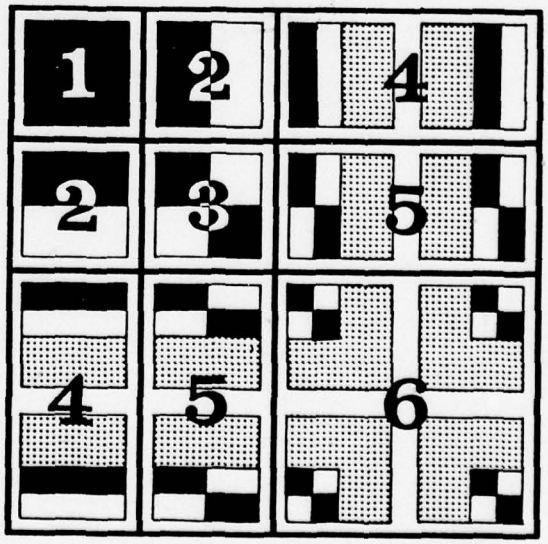
REPORT NO. 77Y196

21

FINAL REPORT FOR EXPERIMENTAL DEVELOPMENT OF A FLIR SENSOR PROCESSOR

DDC
RECEIVED
JAN 30 1978
A

30 SEPTEMBER 1977



NORTHROP

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

(14) NORT
Report No. 77Y196

(9) FINAL REPORT,
FOR
(6) EXPERIMENTAL DEVELOPMENT
OF A
FLIR SENSOR PROCESSOR.

(15) Contract
DAAG53-76-C-0188

(11) 30 Sep[redacted]77

(12) 210 p.

Prepared by

(10) T. Leiboff, T. Noda
H. Henning, B. Deal

DDC
RECEIVED
JAN 30 1978
A

Prepared for U.S. Army Electronics Command,
DRSEL-NV-V1, Night Vision Laboratory,
Fort Belvoir, Virginia 22060

Approved by

Ross M. Chiles

Ross Chiles, Manager
Automation Systems Engineering

Northrop Corporation Electro-Mechanical Division
500 East Orangethorpe Avenue, Anaheim, California 92801

NORTHROP

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

388 834

PRECEDING PAGE NOT FILMED
BLANK

FOREWORD

The work accomplished in the "FLIR Sensor Processor" program, results of which are contained in this report, promises to provide significant benefits in the operational effectiveness of future FLIR sensor systems.

Northrop Corporation, Electro-Mechanical Division wishes to express appreciation to those U.S. Government organizations who sponsored and actively supported this important technical activity. Sponsorship for this program was provided by the Department of the Army's Advanced Concepts Team, headed by Charles H. Church, under the Deputy Chief of Staff for Research, Development, and Acquisition. The program was managed by the U.S. Army Electronics Command's Night Vision Laboratory. Exceptional technical guidance was provided by John S. Dehne, who heads the Image Processing Group, and G. David Singer, Project Engineer.

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

CONTENTS

Section			Page
1	INTRODUCTION		1-1
	1.1	Quarterly Review	1-1
	1.2	Report Outline	1-2
	1.3	Image Processing Facility	1-3
	1.4	Fast Rationalized Haar Transform	1-4
	1.4.1	Basic Definitions	1-4
	1.4.2	Two-Dimensional Formats	1-5
	1.4.3	Matrix Representation	1-8
2	AUTO FOCUS EXPERIMENTS		2-1
	2.1	Introduction	2-1
	2.2	General Experimental Procedure	2-2
	2.3	Test Pattern Experiment	2-6
	2.3.1	Candidate Focus Measures	2-6
	2.3.2	Focus Measure Data	2-7
	2.3.3	Summary	2-22
	2.4	FLIR Image Experiment	2-23
	2.4.1	Image Data	2-23
	2.4.2	Summary of FLIR Image Experiments	2-31
	2.5	Conclusions	2-37
3	AUTO DEBLUR EXPERIMENTS		3-1
	3.1	One Dimensional Blur Model	3-1
	3.2	Blur Operations	3-3
	3.2.1	Focus Error	3-3
	3.2.2	Diffraction Limited Aperture	3-4
	3.2.3	Translation	3-7
	3.2.4	Summary	3-12
	3.3	A Diagonal Representation of Blur	3-12
	3.3.1	Focus Error	3-14
	3.3.2	Limited Aperture	3-15
	3.3.3	Image Motion	3-15
	3.4	Approximate Diagonal Deblur	3-16
	3.5	Motion Detection/Correction	3-18
	3.6	Two Dimensional Blur Model	3-25
	3.7	Test Images	3-27
	3.8	Conclusion	3-32

CONTENTS (Continued)

Section			Page
4	EDGE ENHANCEMENT EXPERIMENTS		4-1
	4.1	Introduction	4-1
	4.2	General Experimental Procedure	4-2
	4.3	Edge Enhancement Criteria	4-4
	4.3.1	Definition of Edges	4-4
	4.3.2	Weighting Factors	4-11
	4.4	Edge Enhancement Experiment	4-16
	4.4.1	Constant Multiplier	4-16
	4.4.2	Variable Multiplier	4-20
	4.4.3	Descriptors	4-21
	4.5	FLIR Image Experiments	4-23
	4.6	Conclusions	4-34
5	ADAPTIVE SIGNAL PROCESSING		5-1
	5.1	Introduction	5-1
	5.2	DC Restoration Problem	5-1
	5.2.1	Problem Definition	5-1
	5.2.2	DC Restoration Solution	5-5
	5.3	Dynamic Range Compression	5-23
	5.3.1	Problem Definition	5-23
	5.3.2	Dynamic Range Compression Solution	5-24
	5.4	Conclusions	5-26
6	BANDWIDTH COMPRESSION		6-1
	6.1	Introduction	6-1
	6.2	General Experimental Procedure	6-1
	6.3	Reduced Resolution Compression	6-3
	6.3.1	Scan Line Sampling	6-3
	6.3.2	Area Sampling	6-6
	6.3.3	Intensity Sampling	6-6
	6.3.4	Reduced Resolution Summary	6-6
	6.4	Haar Transform Compression	6-9
	6.4.1	Sequency Domain Bit Assignments	6-9
	6.4.2	Coefficient Elimination	6-9
	6.5	Image Descriptors for Haar-Compressed Images	6-21
	6.6	Summary and Conclusions	6-29
7	IMAGE DESCRIPTORS		7-1
	7.1	Introduction	7-1
	7.2	Image Description in the Haar Domain	7-2
	7.2.1	Local Image Descriptors	7-2
	7.2.2	Global Image Descriptors	7-9
	7.3	Descriptor Results	7-13
	7.3.1	Focus Results	7-14
	7.3.2	Enhancement Results	7-14

CONTENTS (Continued)

Section			Page
	7.3.3	Compression Results.	7-15
	7.4	Conclusions and Recommendations.	7-16
8		SENSOR PROCESSOR UNIT - DEMONSTRATION PROTOTYPE, and . . .	8-1
	8.1	System Description, Sensor Processing Unit	8-1
	8.2	System Block Diagram	8-2
	8.3	Hardware Configuration	8-4
Addendum		AUTOFOCUS DEMONSTRATION.	A-1

ILLUSTRATIONS

Figure		Page
1-1	Image-Processing Facility.	1-3
1-2	General Haar Function h(r,m,x)	1-5
1-3	General Two-Dimensional Rationalized Haar Production Function	1-6
1-4	Haar Image Domain.	1-6
1-5	Two-Dimensional Rationalized Haar Functions.	1-7
1-6	Format of Northrop's Real-Time 2-D Haar Processor.	1-10
2-1	Experimental Test Configuration.	2-3
2-2	Test Imagery	2-4
2-3	Autofocus Test Setup	2-5
2-4	Portion of Lens Cycle Printout	2-8
2-5	Test Pattern Images as a Function of Lens Position	2-10
2-6	S_n, Sum of AC Magnitudes, as a Function of Lens Position . .	2-13
2-7	S_n^m, Sum of Normalized AC Magnitudes, as a Function of Lens Position.	2-14
2-8	C, Sum of AC Power, as a Function of Lens Position	2-15
2-9	RMS Noise Levels for S_n, Sum of AC Magnitudes.	2-16
2-10	RMS Noise Levels for S_n^m, Sum of Normalized AC Magnitudes . .	2-17
2-11	RMS Noise Levels for C, Sum of AC Power.	2-18
2-12	A Comparison of Focus Measures	2-19
2-13	Test Pattern Descriptor Input Areas.	2-20
2-14	Activity Descriptors as a Function of Lens Position.	2-20
2-15	FLIR Images with Input Areas Flagged by Increased Brightness	2-24
2-16	Total AC Power - Image No. 1	2-25
2-17	Total AC Power - Image No. 2	2-26
2-18	Total AC Power - Image No. 3	2-27
2-19	Total AC Power - Image No. 4	2-28
2-20	Total AC Power - Image No. 5	2-29
2-21	Total AC Power - Image No. 5 (Alternate Sectors)	2-30
2-22	Total AC Power - Image No. 1 Target.	2-32
2-23	Total AC Power - Image No. 2 Target.	2-32
2-24	Total AC Power - Image No. 3 Target.	2-33
2-25	Total AC Power - Image No. 4 Target.	2-33
2-26	Total AC Power - Image No. 5 Target.	2-34

ILLUSTRATIONS (Continued)

Figure		Page
2-27	Activity Descriptors as a Function of Lens Position - Image No. 1 Background	2-34
2-28	Activity Descriptors as a Function of Lens Position - Image No. 1 Target	2-35
2-29	Activity Descriptors as a Function of Lens Position - Image No. 4 Background	2-35
2-30	Activity Descriptors as a Function of Lens Position - Image No. 4 Target	2-36
3-1	Impulse Responses of Focus Error	3-4
3-2	Haar Domain Operator Focus Error (Case 1)	3-5
3-3	Haar Domain Operator Focus Error (Case 2)	3-6
3-4	Haar Domain Operator Aperture Blur (Case 1)	3-8
3-5	Haar Domain Operator Aperture Blur (Case 2)	3-9
3-6	Haar Domain Operator One-Pel Translation	3-10
3-7	Haar Domain Operator Two-Pel Translation	3-11
3-8	Sequency Attenuation as a Result of Focus Error	3-14
3-9	Sequency Attenuation as a Result of Limited Aperture	3-15
3-10	FLIR Thermograms	3-19
3-11	Average Magnitude Haar Spectra	3-20
3-12	Haar Spectrum Prototype	3-21
3-13	Haar Domain Operator for a One-Pel Left Shift	3-22
3-14	Haar Domain Operator for a One-Pel Right Shift	3-23
3-15	Motion Detection Method	3-24
3-16	Deblur Results for Image #1	3-28
3-17	Deblur Results for Image #5	3-28
3-18	Deblur Results for Image #6	3-29
3-19	Deblur Results for Image #7	3-29
3-20	Deblur Results for Image #8	3-30
3-21	Deblur Results for Image #10	3-30
4-1	Experimental Test Configuration	4-3
4-2	FLIR Thermograms	4-5
4-3	Haar Functions, Groups 4 through 15	4-7
4-4	All AC Coefficients, Image 1	4-9
4-5	Groups 9, 11, and 12, Image 1	4-9
4-6	Groups 10 and 13, Image 1	4-10
4-7	Group 14, Image 1	4-10
4-8	Group 15, Image 1	4-11
4-9	All AC Coefficients	4-12
4-10	AC Coefficients for all Six Test Images (Composite)	4-13
4-11	Groups 9 through 15 with Identical Amplitude	4-14
4-12	Group Normalization	4-15
4-13	Effects of Different Edge Thresholds	4-17
4-14	Effects of Different Multipliers	4-18
4-15	Effects of System Noise	4-19
4-16	Effects of TV 2:1 Interface	4-20
4-17	Nonlinear Transfer Functions	4-22
4-18	Variable Multiplier Enhancement - Transfer Function No. 1	4-23
4-19	Variable Multiplier Enhancement - Transfer Function No. 2	4-24
4-20	Image 1	4-26
4-21	Image 5	4-27

ILLUSTRATIONS (Continued)

Figure		Page
4-22	Image 6.	4-28
4-23	Image 7.	4-29
4-24	Image 8.	4-30
4-25	Image 10	4-31
4-26	Parking Lot Scene.	4-33
5-1	AC Coupling Effects in Parallel Scan FLIRS	5-2
5-2	Vertical Correlation Loss in Parallel Scan FLIRS	5-3
5-3	Ideal DC- and AC-Coupled Infrared Imaging.	5-4
5-4	Weighting Function of AC Coupling Circuit.	5-4
5-5	Multiscan Droop Characterization	5-6
5-6	Multiscan Droop Correction Analog.	5-8
5-7	Conservative Representation of Parallel-Scan FLIR High-Pass Filter	5-10
5-8	Processor Algorithm for Absolutely Stable Corrective Network.	5-11
5-9	Gray-Scale Centering Algorithm	5-18
5-10	Non-Artifacting Brightness Control Scheme.	5-20
5-11	Contrast Linearization Analytical Model.	5-21
5-12	Haar Contrast Control.	5-22
5-13	Dynamic Range Compression Example.	5-25
6-1	Experimental Test Configuration.	6-2
6-2	*Transform Format*	6-4
6-3	Under Sampled Scan Lines	6-5
6-4	Under Sampled Areas.	6-7
6-5	Reduced Quantization Levels.	6-8
6-6	Example of Amplitude Distributions	6-10
6-7	Test Pattern Compressed by Sequency Domain BIT Limiting.	6-11
6-8	FLIR Image Compressed by Sequency Domain BIT Limiting.	6-12
6-9	Zonal Filtered Test Pattern.	6-14
6-10	Zonal Filtered FLIR No. 1.	6-15
6-11	Effects of Threshold Levels in Sequency Domain, Test Pattern	6-17
6-12	Effects of Threshold Levels in Sequency Domain, FLIR No. 1	6-18
6-13	Threshold and N-Coefficients Test Pattern.	6-19
6-14	Threshold and N-Coefficients (FLIR No. 1).	6-20
6-15	Threshold and N-Coefficients (NVL No. 4)	6-22
6-16	Threshold and N-Coefficients (NVL No. 6)	6-23
6-17	Threshold and N-Coefficients (NVL No. 7)	6-24
6-18	Threshold and N-Coefficients (NVL No. 8)	6-25
6-19	Threshold and N-Coefficients (NVL No. 10).	6-26
7-1	Group Edge Weighting, α_{rp}	7-5
7-2	Representative Histograms of Coefficients of j^{th} Group	7-6
7-3	Sample Images.	7-7
7-4	Examples of Local Tone and Activity.	7-8
7-5	Examples of Local Texture and Edge Activity.	7-9
7-6	Histogram of Local Tone of Girl.	7-11
7-7	Histogram of Local Tone of Moon.	7-11
8-1	Top-Level Functional Block Diagram	8-3
8-2	Hardware Configuration	8-5

TABLES

Table		Page
2-1	Focus Measure Signal-to-Noise Ratios	2-19
2-2	Focus Measure Computational Complexity	2-21
2-3	Performance of Candidate Focus Measures	2-22
3-1	Grouping of the Haar Domain Operator	3-3
3-2	MSE Observed Using D^{-1} to Deblur	3-17
3-3	Summary of Deblur Results	3-31
4-1	Descriptors for Test Images	4-6
4-2	Average Adaptive Threshold (Image 1)	4-7
4-3	Percent Edge Terms	4-8
4-4	Number of Terms Exceeding θ_1 (Image 1)	4-8
4-5	Selected Parameters for Constant Multiplier	4-20
4-6	Descriptors for Image 1	4-25
4-7	Descriptors of Selected Enhancement Algorithms	4-32
5-1	Computer Simulations for Matched and Mismatched AC Coupling Corrective Networks	5-14
5-2	Rationalized Haar Components of g(x)	5-26
6-1	Descriptors for Haar-Compressed Images	6-27
7-1	Statistics of Local Tone and Activity	7-10
7-2	Statistics of Local Edge and Texture Activity	7-13

SECTION 1
INTRODUCTION

Northrop Corporation, Electro-Mechanical Division, is pleased to submit the final report of the Experimental Development of a FLIR Sensor Processor. The program's funding started on July 1, 1976, and this report summarizes the technical performance through September 30, 1977, marking the completion of Contract No. DAAG53-C-0188.

1.1 QUARTERLY REVIEW

During the first quarter, the Experiment and Study Plan was written, the test system configuration was assembled and the software was debugged. In addition to the Experiment and Study Plan, the first quarterly report included a brief Advanced Scout Helicopter mission scenario used to characterize the FLIR operational environment. The report also included basic definitions of the Haar functions used in the experiments and a set of image descriptors, designed to facilitate evaluation of the image processing algorithms.

The second quarterly report primarily covered the autofocus experiments. The object was to determine the use of the 2D Haar processor as an accurate sensor of infocus/defocus conditions in the FLIR systems. The results of the experiments and subsequent analysis showed that using FLIR imagery, with some amount of activity within the view, a simple scalar value can be computed in real time, which has a sensitivity greater than the eye in determining best focus.

The third quarterly report summarized the experiments performed to determine how the 2D Haar processor can be used to enhance FLIR imagery by highlighting edge details. Algorithms were developed to extract edge information and then to enhance the image. These algorithms were tested on several thermograms provided by Night Vision Laboratory. Results were analyzed and the imagery included. A fixed edge threshold algorithm was reommended for the extraction of edges and a variable multiplier algorithm was recommended for edge emphasis.

1.2 REPORT OUTLINE

For the convenience of the reader, this final report includes much of the material which previously appeared in the three quarterly reports. An outline of the final report follows.

Section 1, Introduction, includes a description of the image processing facility and a discussion of the Fast-Rationalized Haar transform, which forms the basis for all the image processing presented in this report.

Section 2 contains the autofocus experiments.

Section 3 describes the deblur experiments. This includes blurs caused by defocusing, diffraction limitations and scene translations. A method of formulating these blur characteristics in the Haar domain and operations leading to simple deblurring algorithms are presented.

Section 4 consists of the edge enhancement experiments.

Section 5 includes the experiments performed in DC restoration dynamic range compression and, gain and brightness control under the general heading "Adaptive Signal Processing".

Section 6 presents the data compression experiments. Several techniques that reduce image encoding data rates by removing redundancy are described and compared with filtering techniques that reduce resolution. The processed pictures and image statistics are included.

Section 7 provides a recap of the Image Descriptors used to evaluate image quality.

Section 8 contains a suggested hardware implementation of some of the most successful algorithms.

The results of the autofocus demonstration, held at Wright-Patterson Air Force Base in September 1977, will be presented as an addendum to this report. Here, the 2-D Haar processor was coupled to an operating FLIR to demonstrate its real-time autofocus capabilities.

1.3 IMAGE PROCESSING FACILITY

The image-processing facility used in the performance of this contract was developed under an IR&D project that began in 1971. After initial studies and computer modeling of transform image processing, the need for real-time image processing capability became very apparent. In 1973, a one-dimensional (1-D) real-time processor was designed, built, tested, and successfully demonstrated. Following this, a 2-D real-time processor was fabricated and tested. To use its full potential, this 2-D real-time image processor was combined with an instrumentation system and with various image input and output devices to form the complete image-processing facility shown in Figure 1-1. This image-processing facility is capable of receiving real-time



Figure 1-1. IMAGE-PROCESSING FACILITY

sensor imagery and of digitizing and converting the images into an equivalent ensemble of Haar transform coefficients. Real-time and off-line processing can be performed on these coefficients prior to transformation back to images. Image and transform statistics are assembled, computed, and printed out for further analysis.

1.4 FAST RATIONALIZED HAAR TRANSFORM

The following paragraph covers the basic definition of Haar functions, their extension to two-dimensional formats for image processing applications, and their expression in terms of matrices.

1.4.1 Basic Definitions

Haar transforms provide the basis for all image-processing techniques presented in this report. Like other transforms, they can assign quantitative measures to qualitative aspects of the original image, or replace an involved measurement technique by an easier and faster one. These measures, in turn, identify the important variables for processing image restoration and enhancement.

Haar transforms are also comparatively simple. Each transform component effectively multiplies the original function by either 0, +1, or -1. Furthermore, they combine the abilities of local transforms (e.g., Block transforms) to touch up given areas of a picture, with those of global transforms (e.g., Walsh transforms) to render broad comprehensive interpretation of data.

The general Haar function $h(r,m,x)$ appears in Figure 1.2. Sequency r bears a certain resemblance to frequency. Offset m ranges between 1 and 2^r. These mutually orthogonal functions can be rationalized by reducing all amplitudes $2^{r/2}$ to a

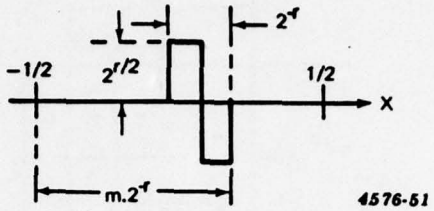


Figure 1-2. GENERAL HAAR FUNCTION h(r,m,x)

common value of unity. A weighing function of 2^r, however, must then be included in either the forward or inverse transform.

1.4.2 Two-Dimensional Formats

Haar functions can be used to expand general one-dimensional (1D) waveforms. Two dimensional (2D) waveforms, such as image intensity distributions I(x,y) can similarly be expanded in a product series:

$$I(x,y) = \sum_{r,m,p,n} a_{rmpn}\, h(r,m,x)h(p,n,y) \tag{1}$$

$$a_{rmpn} = \int_{-1/2}^{1/2}\int_{-1/2}^{1/2} I(x,y)h(r,m,x)h(p,n,y)dxdy \tag{2}$$

While this applies to preliminary domain $-1/2 \le x \le 1/2$, $-1/2 \le y \le 1/2$, simple linear scale changes will qualify these relations for images of any finite extent.

A rationalized two-dimensional product h(r,n,x)h(p,n,y) appears as in Figure 1-3.

When all such diagrams have been located in rectangular (r,p) coordinate space as in Figure 1-4, the resulting array represents the Haar image transform domain. A detailed picture of the array is given in Figure 1-5. It is possible to segment this domain into regions in which the Haar functions are simply translations of one another. Each of the largest possible regions of the Haar domain in which the Haar functions exhibit this property is referred to as a sequency group. The set of 16 x 16 Haar basis functions can be segmented into 25 sequency groups.

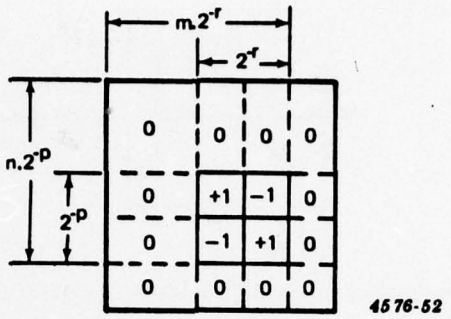


Figure 1-3. GENERAL TWO-DIMENSIONAL RATIONALIZED HAAR PRODUCTION FUNCTION

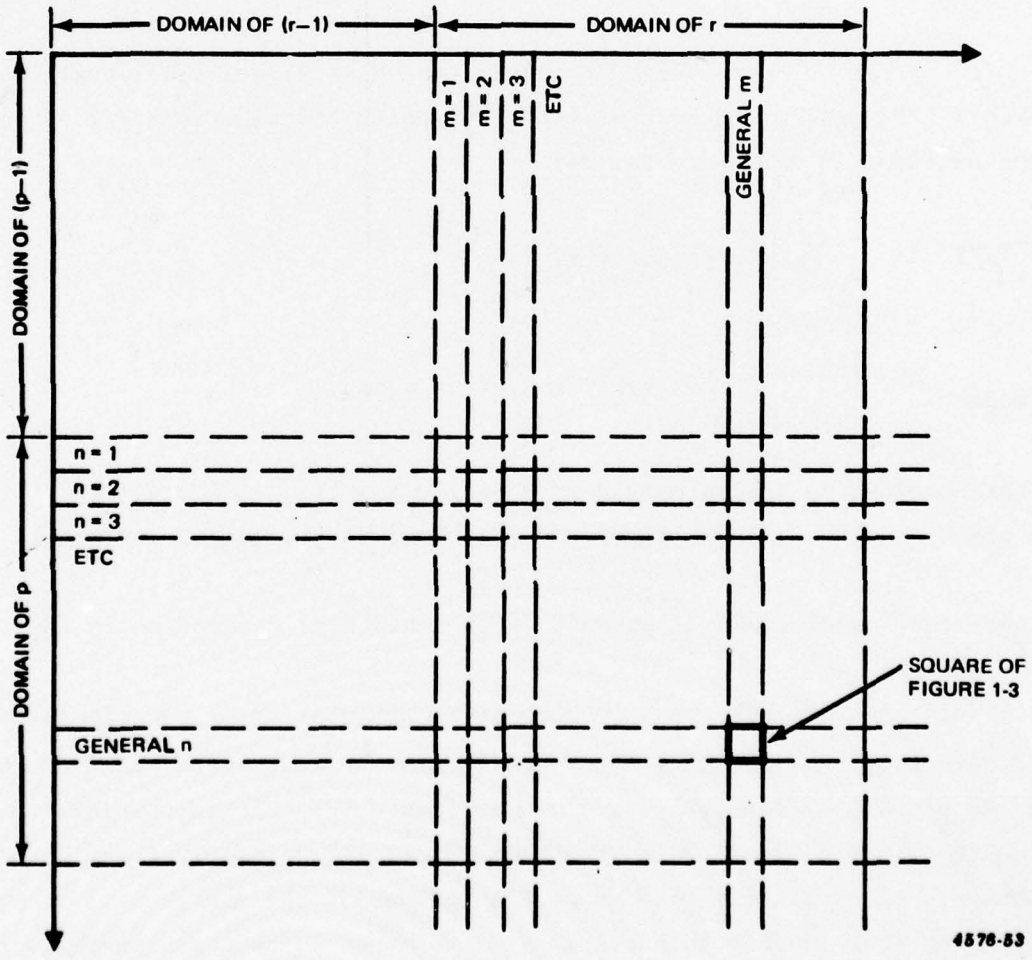


Figure 1-4. HAAR IMAGE DOMAIN

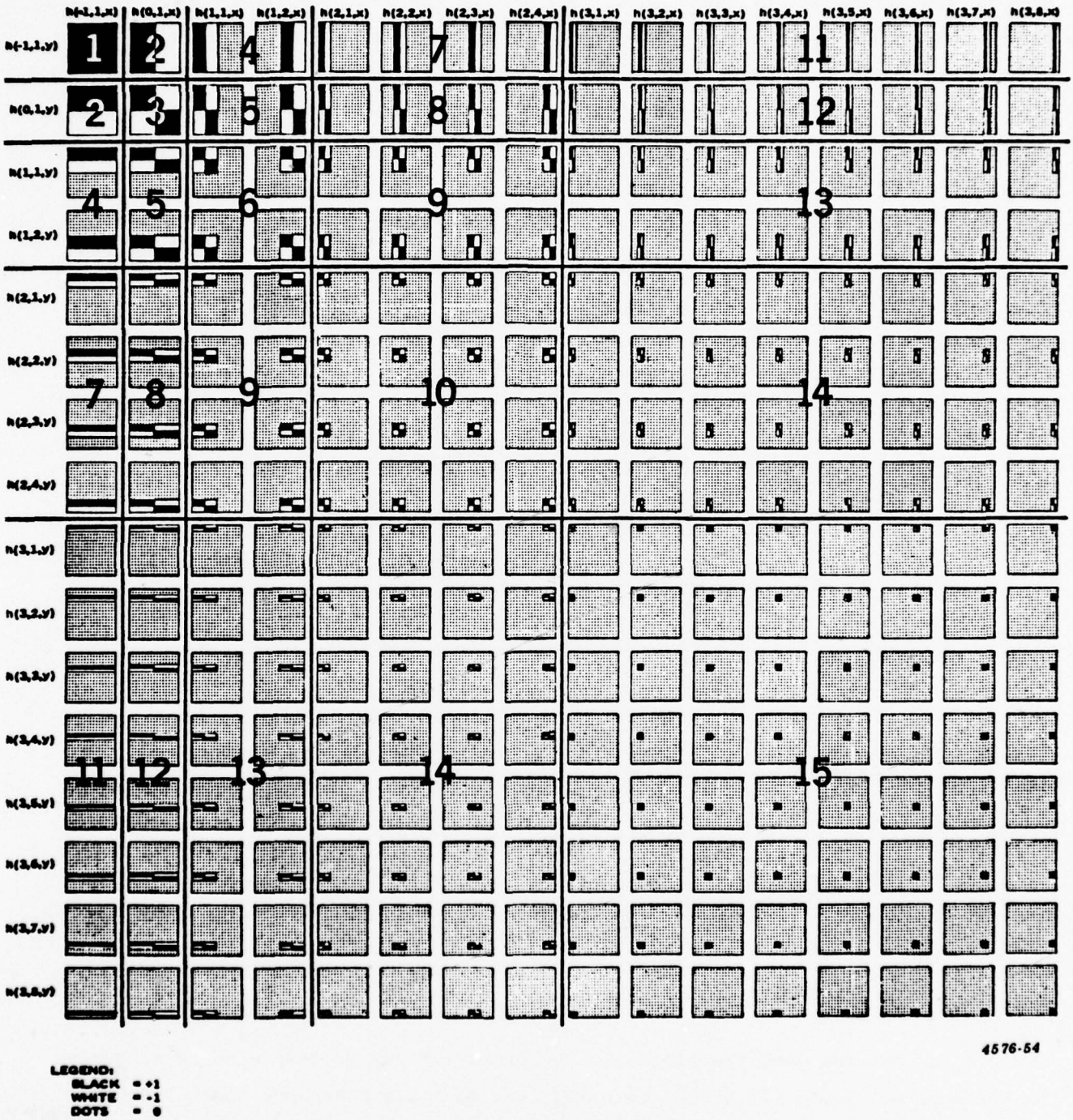


Figure 1-5. TWO-DIMENSIONAL RATIONALIZED HAAR FUNCTIONS

Examination of Figure 1-5 shows that, by grouping the Haar coefficients as shown, it is possible to separate those coefficients corresponding to the horizontal or vertical edges in the image. Assuming that horizontal and vertical edges are equally likely in a typical image, these 25 groups can be combined to get 15 new groups as shown by the numbering scheme in Figure 1-5 and as enumerated below:

Group	No. of Terms in Group
1	1
2	2
3	1
4	4
5	4
6	4
7	8
8	8
9	16
10	16
11	16
12	16
13	32
14	64
15	64
	Sum = 256

With the Northrop Haar processor, it is possible to select lower and upper thresholds and a multiplication factor independently for each group.

1.4.3 Matrix Representation

Forward and inverse Haar transforms in equations (1) and (2) can also be expressed in terms of matrices. It should be noted that matrix expressions substitute discrete summation for integration in equation (2). With this provision, continuous image distribution $I(x,y)$ reduces to the discrete expression $I(t\Delta x, l\Delta y)$ where Δx and Δy typically assume values smaller than the resolution spot size of the imaging system. Integers t and l span a range sufficient to cover maximum and

minimum values of x and y. This expression in turn can be represented by the matrix $I_{t,1}$ or simply I.

Haar function h(r,m,x) reduces to the matrix h_{ts} where integer s designates successive Haar functions in a particular sequency and then moves on to the next sequency. Using only the first few sequencies in a rationalized format to clarify the illustration:

$$H_{\text{rationalized}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} = \left[h_{ts}\right]_{\text{rationalized}} \quad (3)$$

Expansion coefficient a_{rmpn} likewise reduces to matrix $\left[a_{sq}\right] = A$.

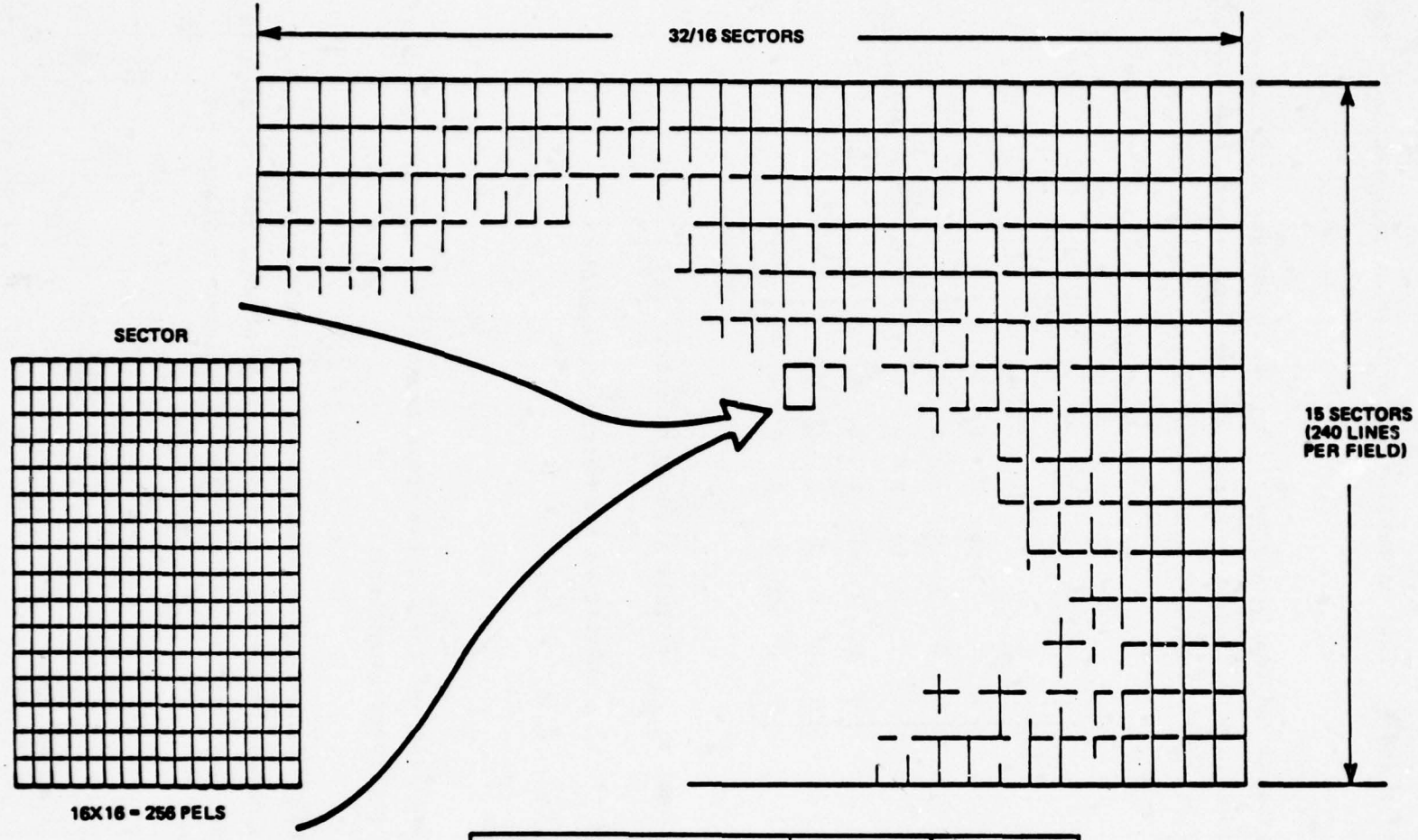
Equations (1) and (2) then convert to the matrix expressions

$$I = H^T A H \quad (4)$$

$$A = H I H^T \quad (5)$$

Because of the simplicity of (3) and the compactness of (4) and (5), the matrix approach will be preferred in general.

Northrop's real-time Haar processor performs the rationalized Haar transform on 16 x 16 pel non-overlapping blocks of a digitized image array. The precise arrangement of transform blocks and data processing rates is depicted in Figure 1-6.



SAMPLING RATE	8.19 MHZ	4.095 MHZ
TOTAL NUMBER OF SECTORS	480/FIELD	240/FIELD
	960/FRAME	480/FRAME
TOTAL NUMBER OF PELS	123K/FIELD	61K/FIELD
	246K/FRAME	123K/FRAME

4697-2

Figure 1-6. FORMAT OF NORTHROP'S REAL-TIME 2-D HAAR PROCESSOR

SECTION 2
AUTO FOCUS EXPERIMENTS

2.1 INTRODUCTION

Because severe demands are placed on a two-man helicopter team, the need for manual FLIR refocusing can pose a formidable problem. FLIR defocusing arises from continual changes in observation range, small thermal deformations of the lens, and other incidental causes. Automatic refocusing offers a means to combat these tendencies and relieve part of the manpower burden. The apparatus must sense the defocus condition of the image and produce a corresponding error signal for automatic lens adjustment. The objective of this experiment is to develop a measure of the state of focus of an image based on the distribution of energy in its rationalized Haar transform. The resultant focus measure may then be used as the basis for the generation of a focus-correcting error signal.

The focus measure developed should exhibit the following characteristics in order to be compatible with system and operational requirements:

a. Low Computational Complexity - The amount of computation required to evaluate an image's state of focus must be amenable to real-time (TV rates) processing.

b. High Sensitivity to Focus - Small changes in an image's state of focus should produce significant changes in the focus measure.

c. Noise Immunity - In the present test configuration, noise in the video signal before digitization is the overriding contributor to system noise. In an operational configuration, small (one pel) translations of the image may also produce a significant amount of noise. The focus measure should be relatively insensitive to both types of image noise.

d. Scene Independence - The focus measure should perform equally well on both background and targets; however, there are some limitations to this requirement. There must be sufficient activity in the area of interest to evaluate the image's state of focus. For example, any focus measure will not be expected to perform well when the image contains only clear sky.

e. Adaptability in an Operational System - The focus measure in its final form must be able to serve as the basis for the generation of an error signal in a servo system. This signal must have sufficient bandwidth and stability to be compatible with FLIR image quality parameters and ASH/AAH operational requirements. These specific system requirements are discussed in the Appendix to the first quarterly report (76Y133).

The experiment is divided into two major sections. In the first section the performance of several candidate focus measures is examined. The test image used in this section is a video test pattern (RETMA resolution chart). Tradeoffs of the characteristics of the candidate focus measures are made and two of them are selected for subsequent analysis. In the second section, the capabilities of the selected focus measures are evaluated on the set of five FLIR thermograms supplied by NVL. Both target and background objects are used as input data in order to determine the behavior of the focus measures on general classes of images.

2.2 GENERAL EXPERIMENTAL PROCEDURE

The experimental test configuration is shown in Figure 2-1. Test imagery consisted of a video test pattern and five FLIR thermograms supplied by Night Vision Laboratory (NVL) (Figure 2-2). The focus of the video camera was set manually with the aid of 13 equally spaced marks on the lens barrel. Figure 2-3 shows the optical rail apparatus and a more detailed view of the video camera lens.

The data collection sequence for each of the focus measures was as follows: The video camera lens was set initially at one extreme (lens position A). The Haar transforms of both fields of four previously selected sectors of the image were acquired and the candidate focus measure computed. The result was then printed out on the Centronics Line Printer.

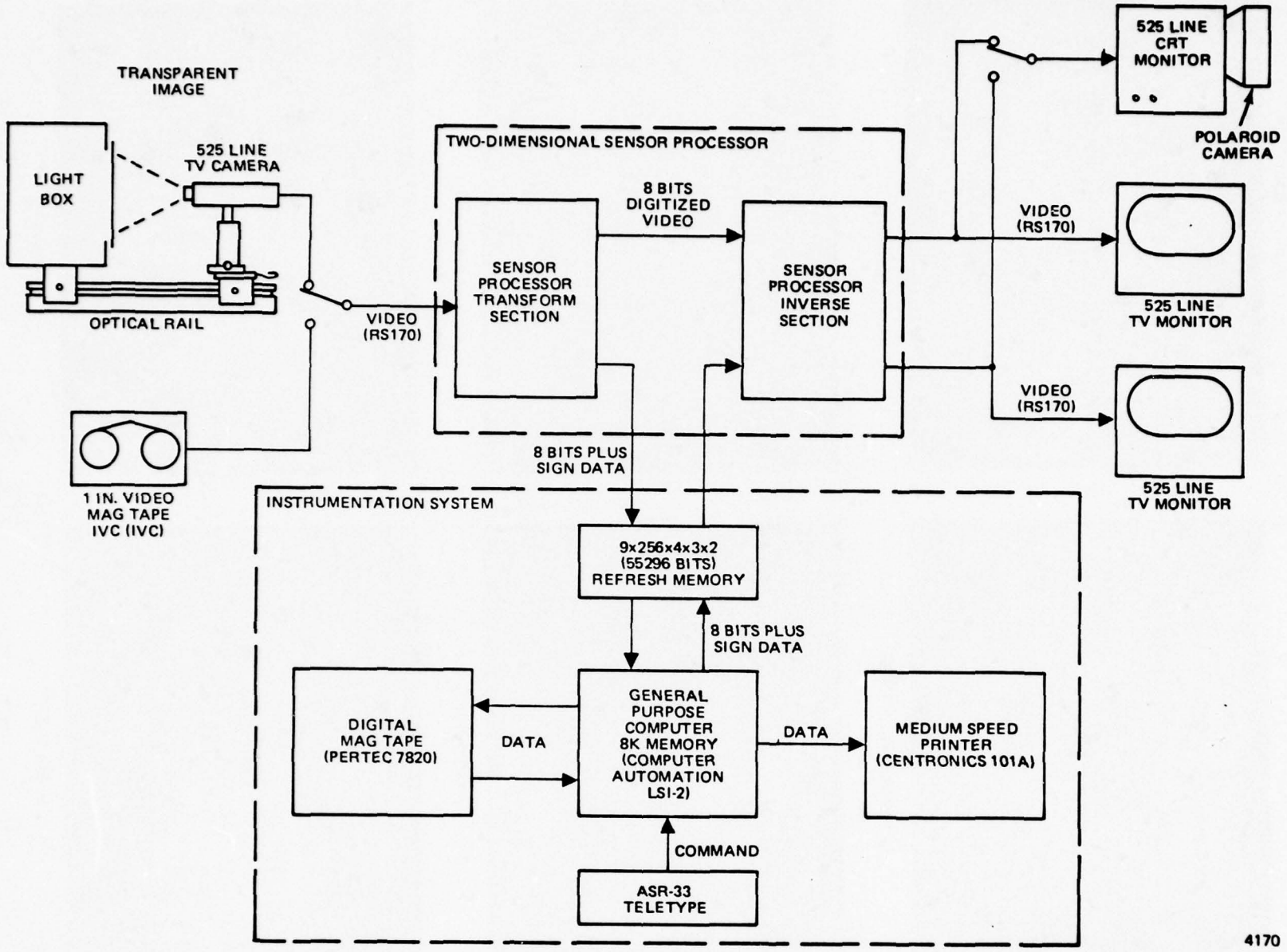
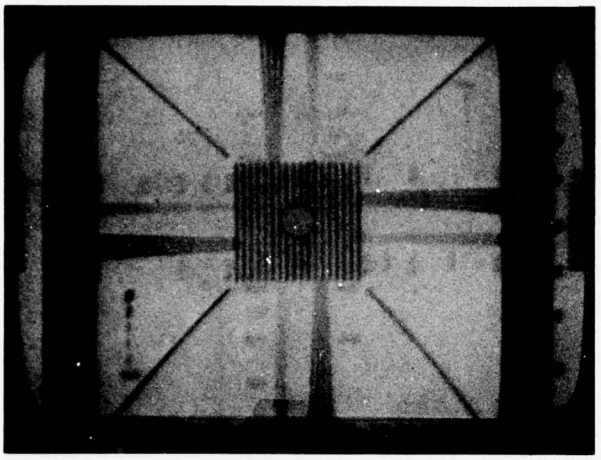
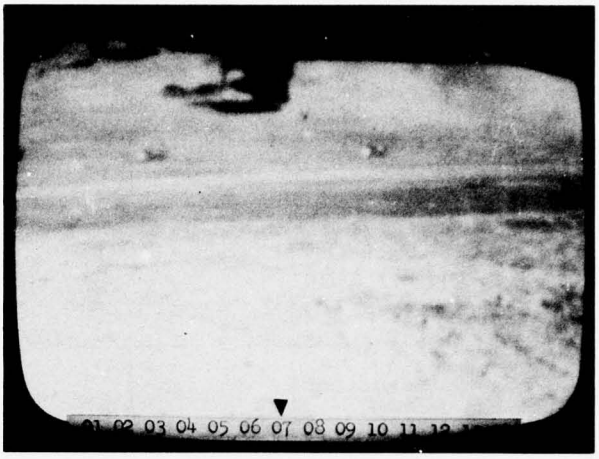


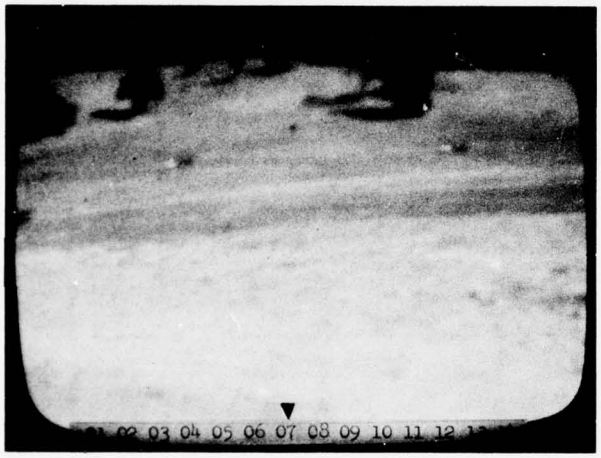
Figure 2-1. EXPERIMENTAL TEST CONFIGURATION



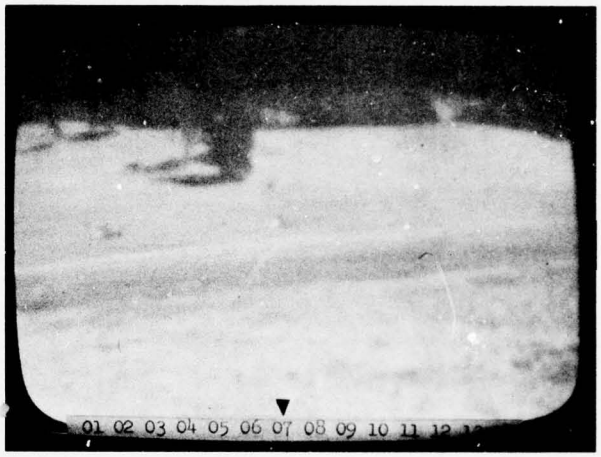
(a) VIDEO TEST PATTERN



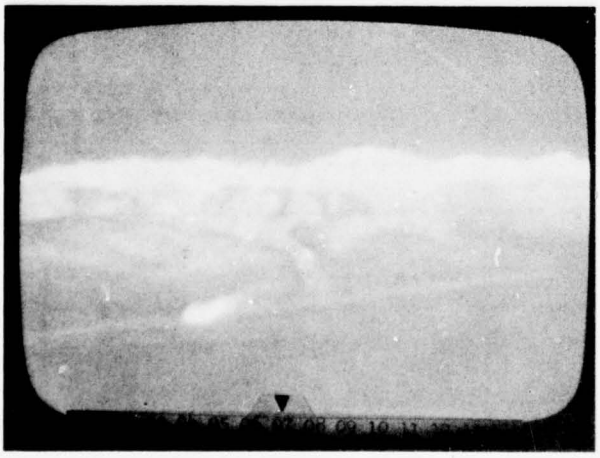
(b) IMAGE 1



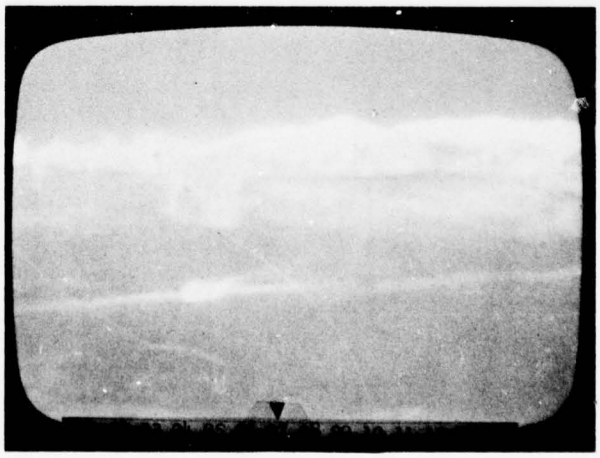
(c) IMAGE 2



(d) IMAGE 3



(e) IMAGE 4



(f) IMAGE 5

Figure 2-2. TEST IMAGERY

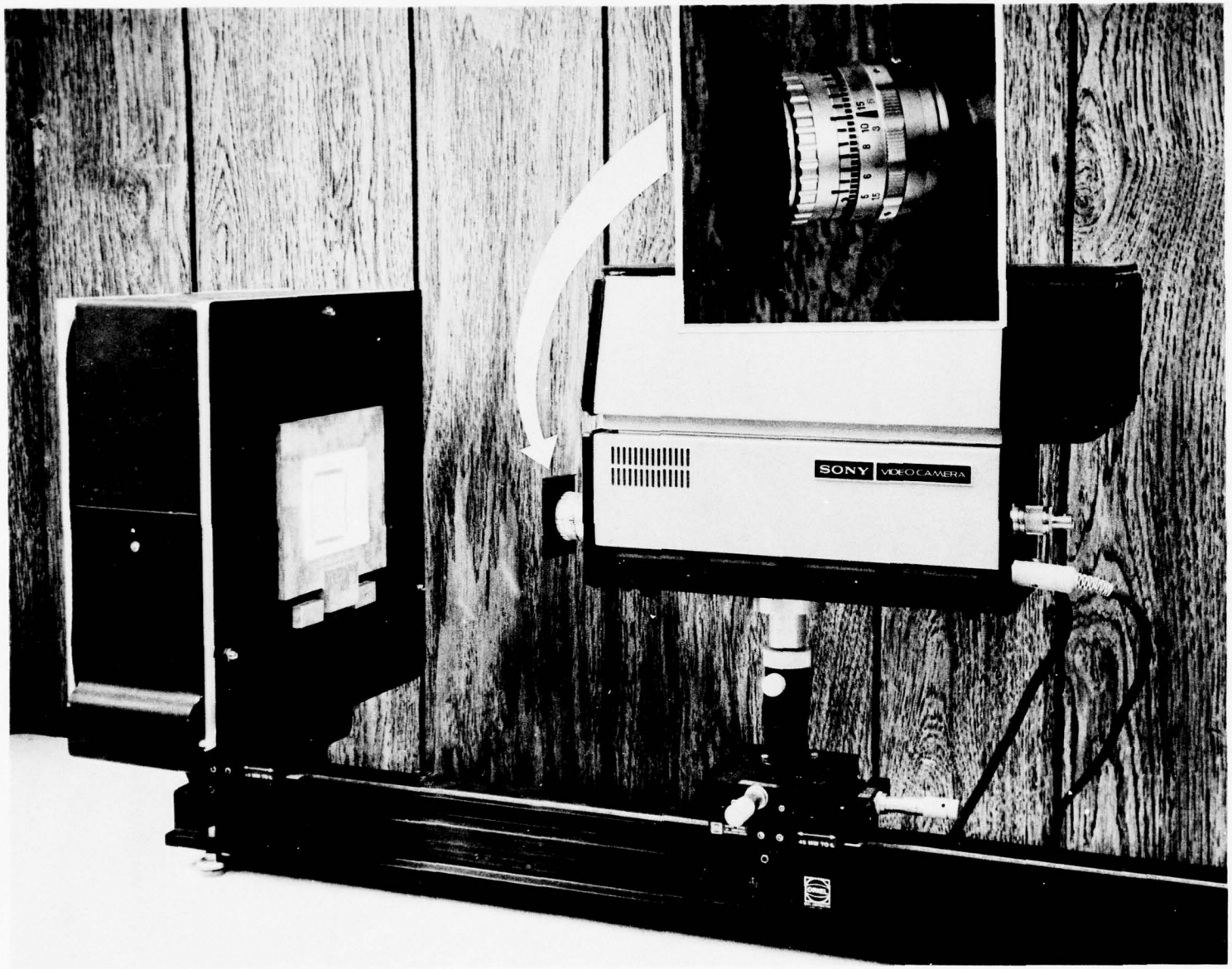


Figure 2-3. AUTOFOCUS TEST SETUP

This was repeated, advancing the lens position one step each time, until the other extreme lens position (lens position M) was reached. The entire process was then repeated, stepping the lens backward from position M to position A. This process of advancing and retracting the lens position through one cycle while evaluating a focus measure at each position will be referred to as a "Lens Cycle." One lens cycle takes about 6 minutes to complete.

Noise measurements of each focus measure were performed by constantly computing the focus measure's value while leaving the video camera lens stationary.

2.3 TEST PATTERN EXPERIMENT

2.3.1 Candidate Focus Measures

Several preliminary measures of an image's state of focus were selected for study. Each measure is based on the distribution of energy in the Haar transform of the image. The candidate focus measures are presented here in order of increasing computational complexity.

2.3.1.1 Sum of the Magnitudes - The sum of the magnitudes, S_m, of the ac Haar coefficients is expressed as

$$S_m = \sum_r \sum_m \sum_p \sum_n |a_{rmpn}| - a_{-1,1,-1,1} \qquad (1)$$

S_m is the computationally simplest index of the total ac information in an image's Haar transform. In order to ascertain the distribution of ac information throughout the transform as well, totals of the magnitudes of the coefficients were also taken for each sequency group.

2.3.1.2 Sum of the Normalized Magnitudes - The sum of the normalized magnitudes, S_n, of the ac coefficients is expressed as

$$S_n = \sum_r \sum_m \sum_p \sum_n f_{rp} |a_{rmpn}| - a_{-1,1,-1,1} \qquad (2)$$

The two-dimensional sensor processor's implementation of the rationalized Haar transform contains an inherent high sequency boost. To remove this boost, a normalization factor, f_{rp}, is introduced to weight the information from the rp^{th} sequency group. The sum of the normalized magnitudes more closely approximates the sum of the magnitudes of an unrationalized Haar transform. S_m was evaluated for all the ac terms as well as for each sequency group.

2.3.1.3 AC Power - The ac power, C, is defined as

$$C = \sum_r \sum_m \sum_p \sum_n (a_{rmpn})^2 - (a_{-1,1,-1,1})^2 \qquad (3)$$

The parameter C is equivalent to the local activity of a 16 by 16 pel image sector. The image's activity should provide an index of focus that is more noise immune and more sensitive to focus than the previous measures. The noise immunity arises from the fact that, by squaring terms, a few large fluctuations in amplitude may still be seen despite the presence of many small ones. The increased sensitivity to focus is due to a similarity between the Fourier and unrationalized Haar transforms. Since both transforms are unitary and possess equivalent dc terms, Parseval's theorem tells us that the ac energy in each is also equivalent. Defocusing an image effectively subjects the image to a low-pass filter, thereby decreasing the ac energy in its Fourier transform. The ac energy in the image's unrationalized Haar transform would likewise be decreased. Since the rationalization of the Haar transform involves only simple scaling of each sequency group, the direction, if not the amount, of change in the ac energy of the rationalized Haar will be equivalent to that of the unrationalized Haar and Fourier transforms.

2.3.1.4 Texture and Edge Activity - Texture and edge activities, as defined in the first quarter report, are the most expensive computationally. They are included in the study to determine if variations in image activity due to changes in state of focus will appear primarily as variations in texture or edge activity.

2.3.2 Focus Measure Data

Each of the foregoing focus measures was computed through one lens cycle. A representative portion of a lens cycle printout is shown in Figure 2-4. Since during

AUTO FOCUS GROUP POWER LENS POSITION A

SECTOR	GRP1	GRP2	GRP3	GRP4	GRP5	GRP6	GRP7	GRP8	GRP9	GRP10	GRP11	GRP12	GRP13	GRP14	GRP15	AC
2290	325	82		19		1	14	7	6	16	25	5	23	96	116	410
229E	325	82		31		2	22	7	8	8	39	7	26	70	147	449
2300	329	65	1	4	1	2	13	2	7	7	9	5	16	42	85	259
230E	330	64	1	15	1	2	23	3	10	11	19	5	21	56	136	367
2310	327	29	1	6	3	2	7	3	7	14	4	4	14	66	107	267
231E	327	29	1	3	2	3	6	6	5	7	5	5	18	81	132	303
2320	320	5		6	2			3	6	10	4	4	19	45	68	172
232E	320	5		5	1	2	1	3	10	7	5	9	16	54	84	202
	325	45		11	1	1	10	4	7	10	13	5	19	63	109	303

AUTO FOCUS GROUP POWER LENS POSITION B

SECTOR	GRP1	GRP2	GRP3	GRP4	GRP5	GRP6	GRP7	GRP8	GRP9	GRP10	GRP11	GRP12	GRP13	GRP14	GRP15	AC
2290	325	82		12	1	2	16	5	16	16	14	7	27	80	98	376
229E	325	82		36	1	2	29	4	10	18	28	8	20	86	97	421
2300	329	65	1	6	1	3	7	3	11	7	12	4	15	47	65	247
230E	329	50	1	10	1	6	23	4	14	8	32	7	13	65	76	310
2310	326	20	1	11	3	6	5	5	10	13	3	4	18	84	126	309
231E	327	29	1	6	3	2	7	2	12	7	6	3	9	75	129	291
2320	320	5		5	1	1	3	2	4	7	5	4	18	41	94	190
232E	320	5		2	3	2	3	3	6	14	2	7	22	59	112	240
	325	42		11	1	3	11	3	10	11	12	5	17	67	99	298

AUTO FOCUS GROUP POWER LENS POSITION C

SECTOR	GRP1	GRP2	GRP3	GRP4	GRP5	GRP6	GRP7	GRP8	GRP9	GRP10	GRP11	GRP12	GRP13	GRP14	GRP15	AC
2290	325	101		12	1	2	13	3	16	9	28	15	26	62	147	435
229E	325	82		39	2	2	20	8	9	11	27	5	22	81	117	425
2300	329	65	1	3	1	4	6	3	9	9	9	3	18	34	91	256
230E	329	50	1	9	2	6	16	4	11	10	31	5	24	48	76	293
2310	327	26	1	6	6	1	8	3	10	16	7	10	22	63	118	297
231E	327	26	1	6	5	3	11	4	7	13	9	9	14	54	146	308
2320	320	5		5	1	1	3	2	4	6	5	3	22	57	84	198
232E	320	2		2	2	1	3	2	3	6	6	4	13	78	71	193
	325	44		10	2	2	10	3	8	10	15	6	20	59	106	300

AUTO FOCUS GROUP POWER LENS POSITION D

SECTOR	GRP1	GRP2	GRP3	GRP4	GRP5	GRP6	GRP7	GRP8	GRP9	GRP10	GRP11	GRP12	GRP13	GRP14	GRP15	AC
2290	325	101		12	1	2	18	8	13	17	22	15	30	112	104	455
229E	325	82		22	1	2	21	3	13	13	32	10	26	59	122	406
2300	329	65	1	7	1	3	11	4	11	16	16	6	24	56	113	334
230E	329	50	1	14	3	6	21	2	16	20	11	5	29	73	102	353
2310	327	26	1	6	6	1	6	2	8	9	7	7	26	70	104	279
231E	327	26	1	3	4	1	10	4	9	16	4	6	16	92	118	310
2320	320	5		3	1		2	1	5	7	7	3	20	45	82	181
232E	320	5		5	2	4	3	3	6	11	9	5	20	46	93	212
	325	45		9	2	2	11	3	10	13	13	7	23	69	104	316

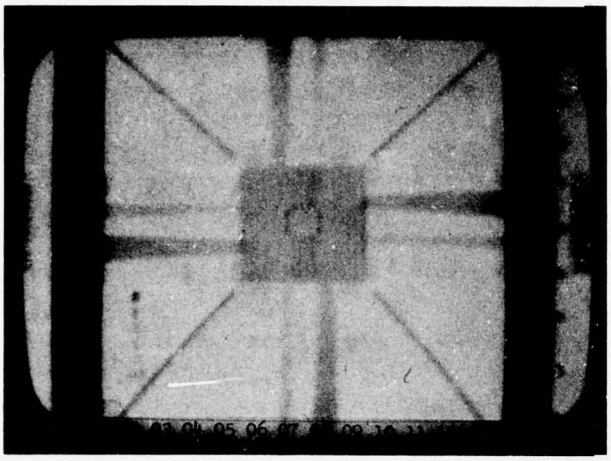
Figure 2-4. PORTION OF LENS CYCLE PRINTOUT

each lens cycle each lens position is encountered twice, the two resultant focus measure values are averaged to facilitate graphical presentation of the data.

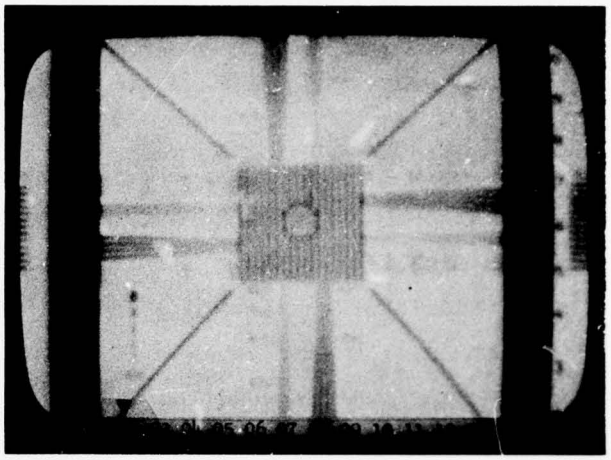
The range of defocus encountered through one lens cycle is shown in the series of photographs in Figure 2-5, A through M. The behavior of the first three candidate focus measures is shown as a function of lens position in Figures 2-6, 2-7, and 2-8. The noise characteristics of each measure are shown in Figures 2-9, 2-10, and 2-11. As can be seen in Figures 2-6, 2-7, and 2-8, each focus measure tested exhibited a peak value at lens position G. In order to determine if lens position G is indeed the sharpest focus position, the fine detail of the test pattern was observed under a hand lens. Close examination of the wedges in the test pattern (Figure 2-5) shows that lens positions F, G, and H are the ones that yield images with sharper focus. Visually discriminating the difference in focus between F, G, and H is next to impossible, however. Since lens position G is in the center of F, G, and H, and the peaks in the focus measure curves are distinct, lens position G is deemed to be the "in focus" lens position. The ability of the focus measure to evaluate the state of focus of an image appears to rival that of the observer.

The division of energy into sequency groupings for each focus measure can be seen in part A of Figures 2-6, 2-7, and 2-8. The changes in the ac energy of the Haar transform due to movement of the lens are localized in sequency groups 11 through 15, and groups 7 through 10. Also, both sequency groupings tend to behave in a like fashion. It is possible, then, to evaluate the state of focus of an image using only the information contained in one sequency grouping. Such an approach, however, would increase the focus measure's sensitivity to image noise (particularly image motion) as well as allow some possibly significant changes in ac energy in another group to go unobserved. For these reasons, the focus measures studied here will sum all of the ac energy within the image's Haar transform and will not single out any one sequency group or grouping.

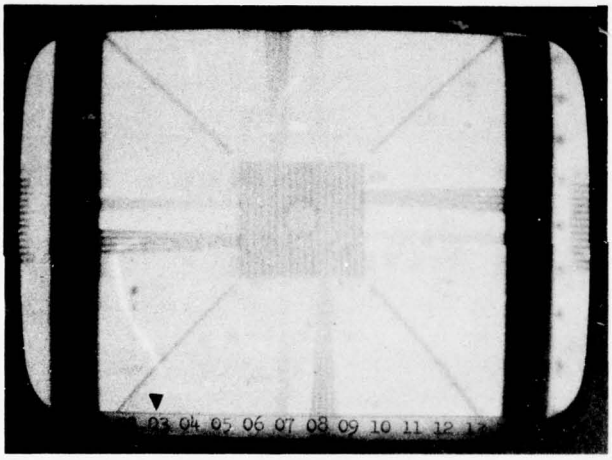
The sharpness, or sensitivity to focus, of the candidate focus measures varies substantially. To aid in comparison, each focus measure is plotted in Figure 2-12. The curve resulting from the sum of the magnitudes of ac coefficients is comparatively flat, its performance being only slightly poorer at its end points than that of the sum of the normalized magnitudes of the ac coefficients. The relative insensitivity exhibited by these two curves is due to the uniform weighting of each



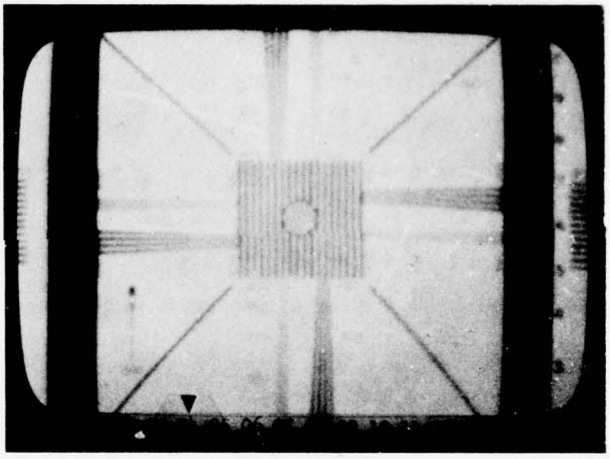
(a) LENS POSITION A



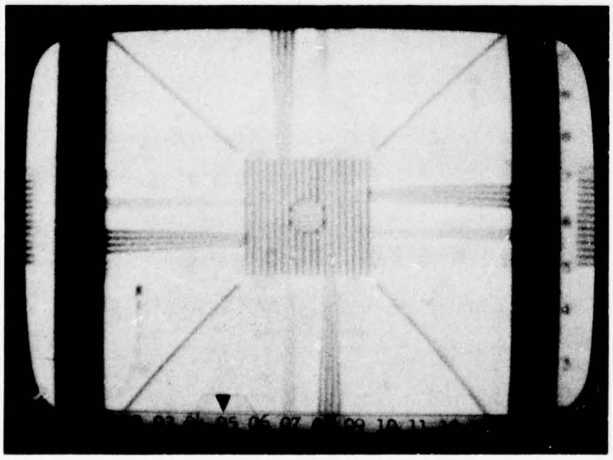
(b) LENS POSITION B



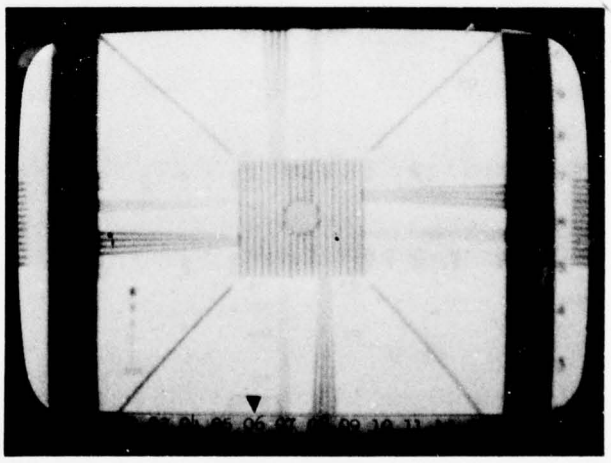
(c) LENS POSITION C



(d) LENS POSITION D

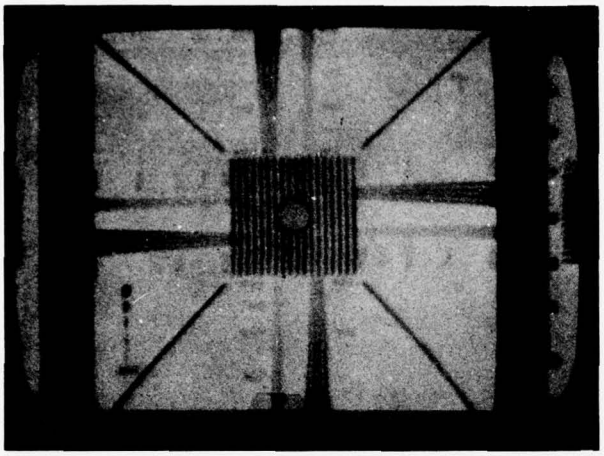


(e) LENS POSITION E

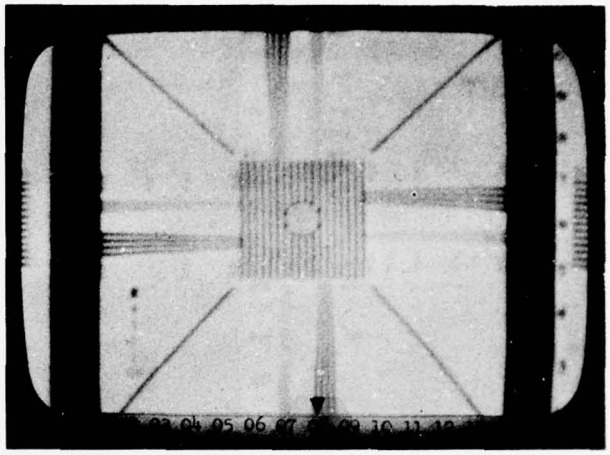


(f) LENS POSITION F

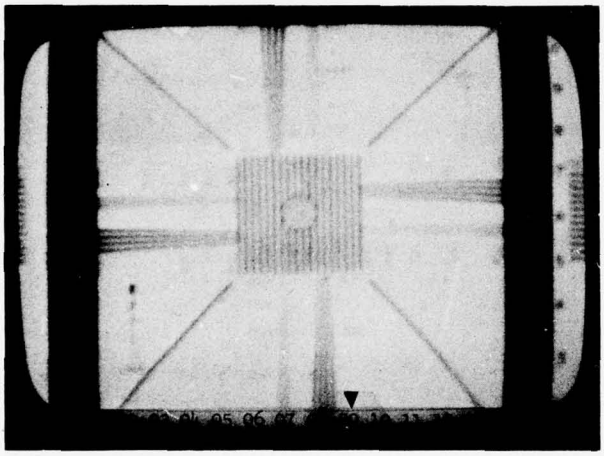
Figure 2-5. TEST PATTERN IMAGES AS A FUNCTION OF LENS POSITION (Sheet 1 of 3)



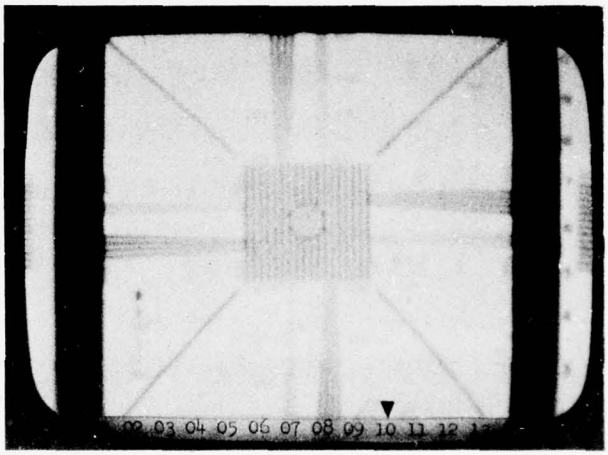
(g) LENS POSITION G



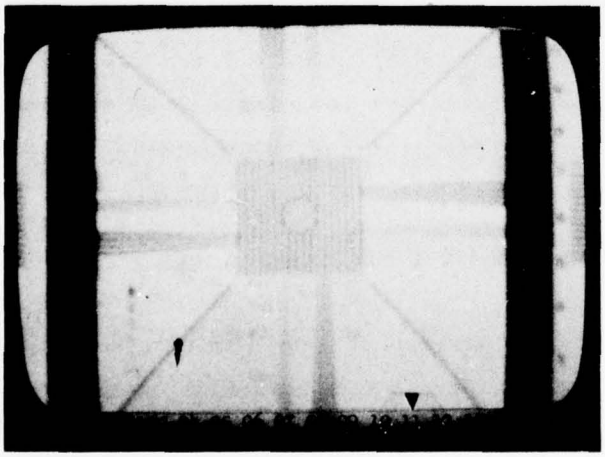
(h) LENS POSITION H



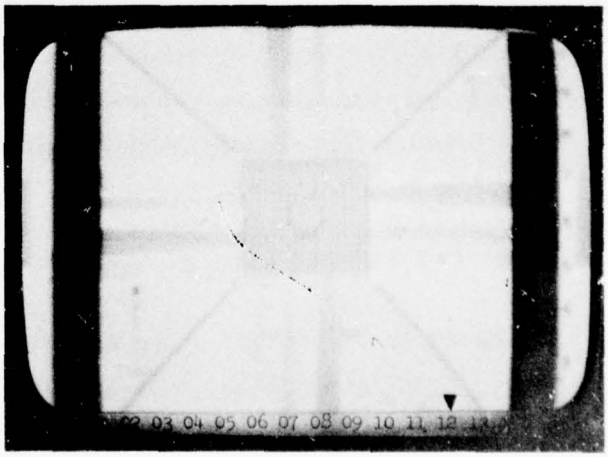
(i) LENS POSITION I



(j) LENS POSITION J

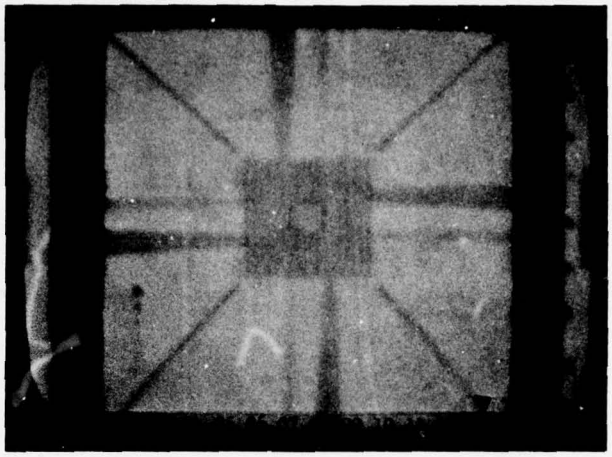


(k) LENS POSITION K



(l) LENS POSITION L

Figure 2-5. TEST PATTERN IMAGES AS A FUNCTION OF LENS POSITION (Sheet 2 of 3)



(m) LENS POSITION M

Figure 2-5. TEST PATTERN IMAGES AS A FUNCTION OF LENS POSITION (Sheet 3 of 3)

coefficient within any sequency group. When any set of numbers is summed in this fashion, a substantial change in a few members of the set often is obscured by effectively averaging them with the others. By squaring the numbers before summing them, however, the sensitivity to small changes may be increased. This increased sensitivity is exhibited in Figure 2-12 by the curve for the ac power contained in the image's Haar transform.

Examination of the noise characteristics of each focus measure (Figures 2-9, 2-10, and 2-11) shows that their noise levels, expressed as a fraction of peak value, are roughly equivalent. By assuming the maximum observed value of a focus measure minus its minimum value to be the available signal, the signal-to-noise ratio can be determined. The signal-to-noise ratios of the first three focus measures are shown in Table 2-1. As expected, C, the ac power of the image's Haar transform, has the highest signal-to-noise ratio.

The behavior of texture and edge activity as a function of lens position was determined in a fashion similar to the other focus measures, i.e., evaluating the focus measure at each point through one lens cycle. However, the area of the test pattern for which the descriptors were computed was different than that used for the other measures. Figure 2-13 shows the test pattern with the descriptor input areas flagged

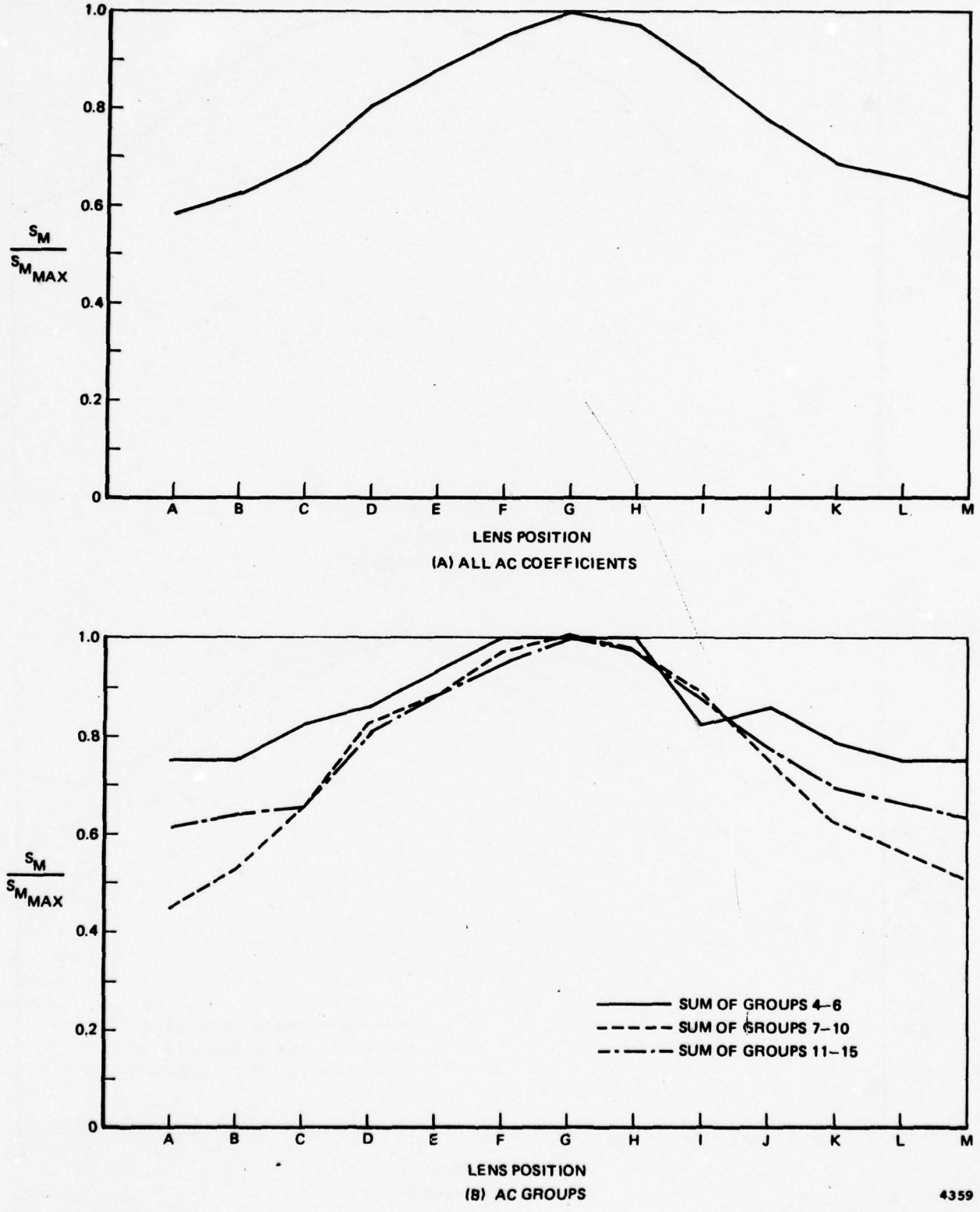
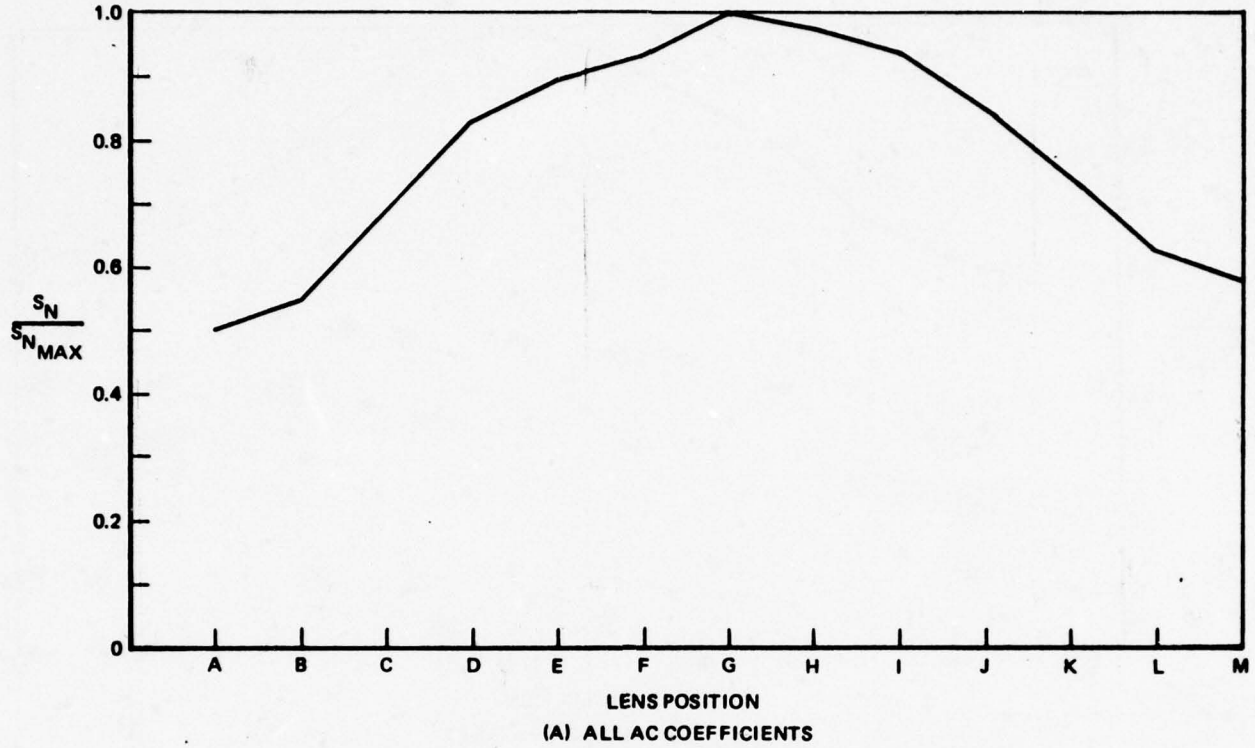


Figure 2-6. S_m, SUM OF AC MAGNITUDES, AS A FUNCTION OF LENS POSITION



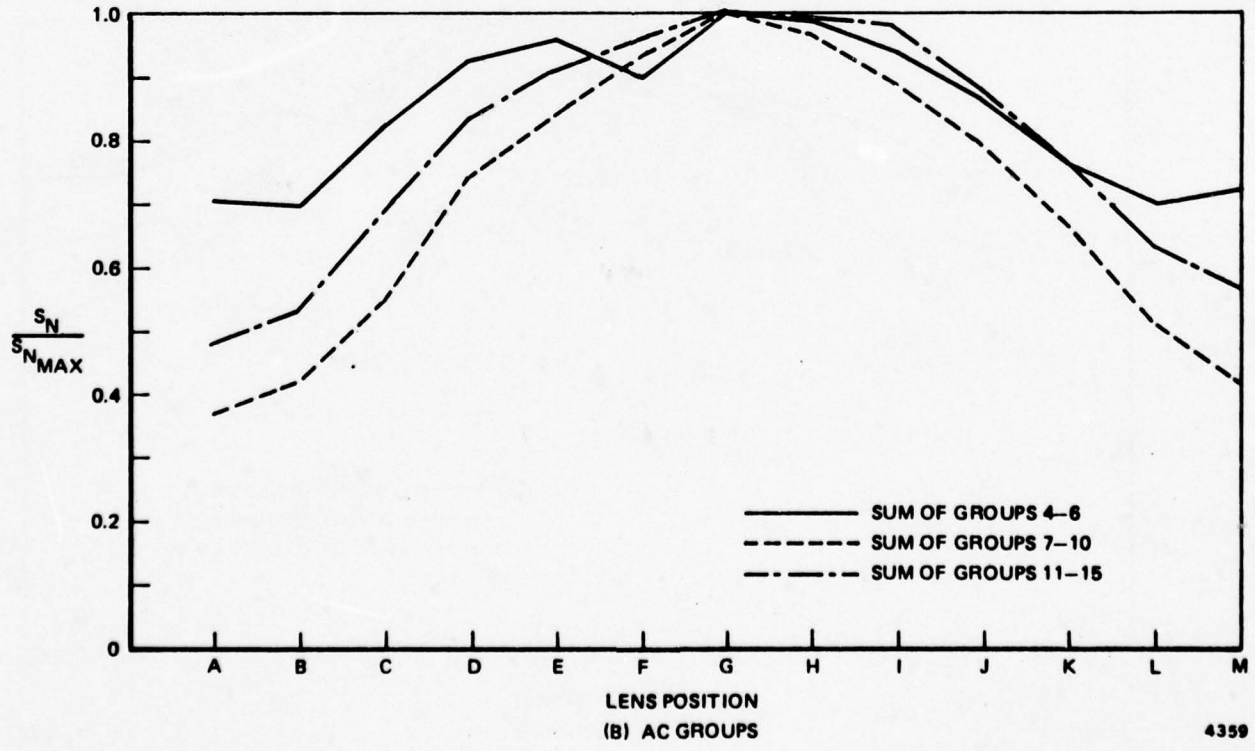


Figure 2-7. S_n, SUM OF NORMALIZED AC MAGNITUDES, AS A FUNCTION OF LENS POSITION

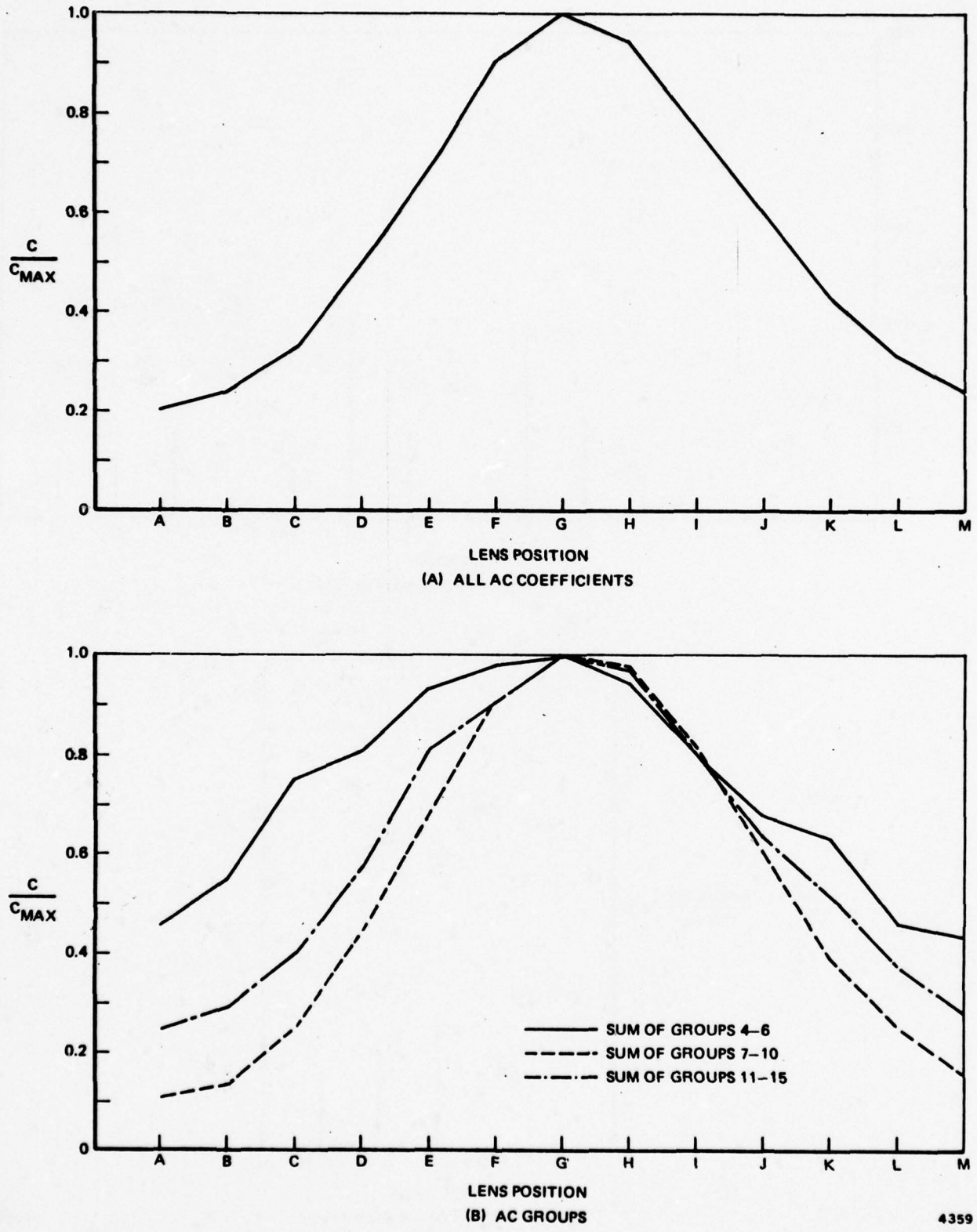


Figure 2-8. C, SUM OF AC POWER, AS A FUNCTION OF LENS POSITION

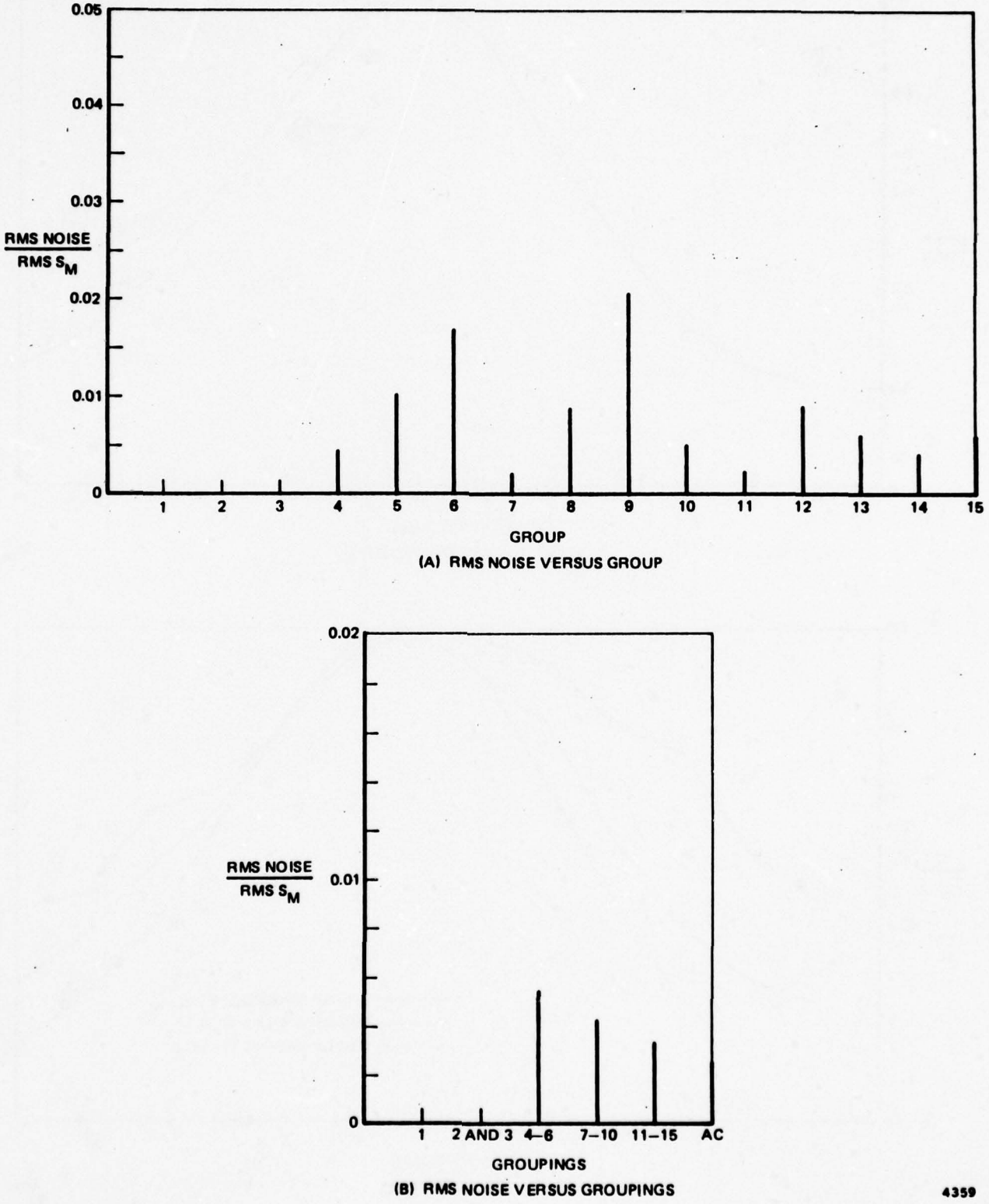


Figure 2-9. RMS NOISE LEVELS FOR S_m, SUM OF AC MAGNITUDES

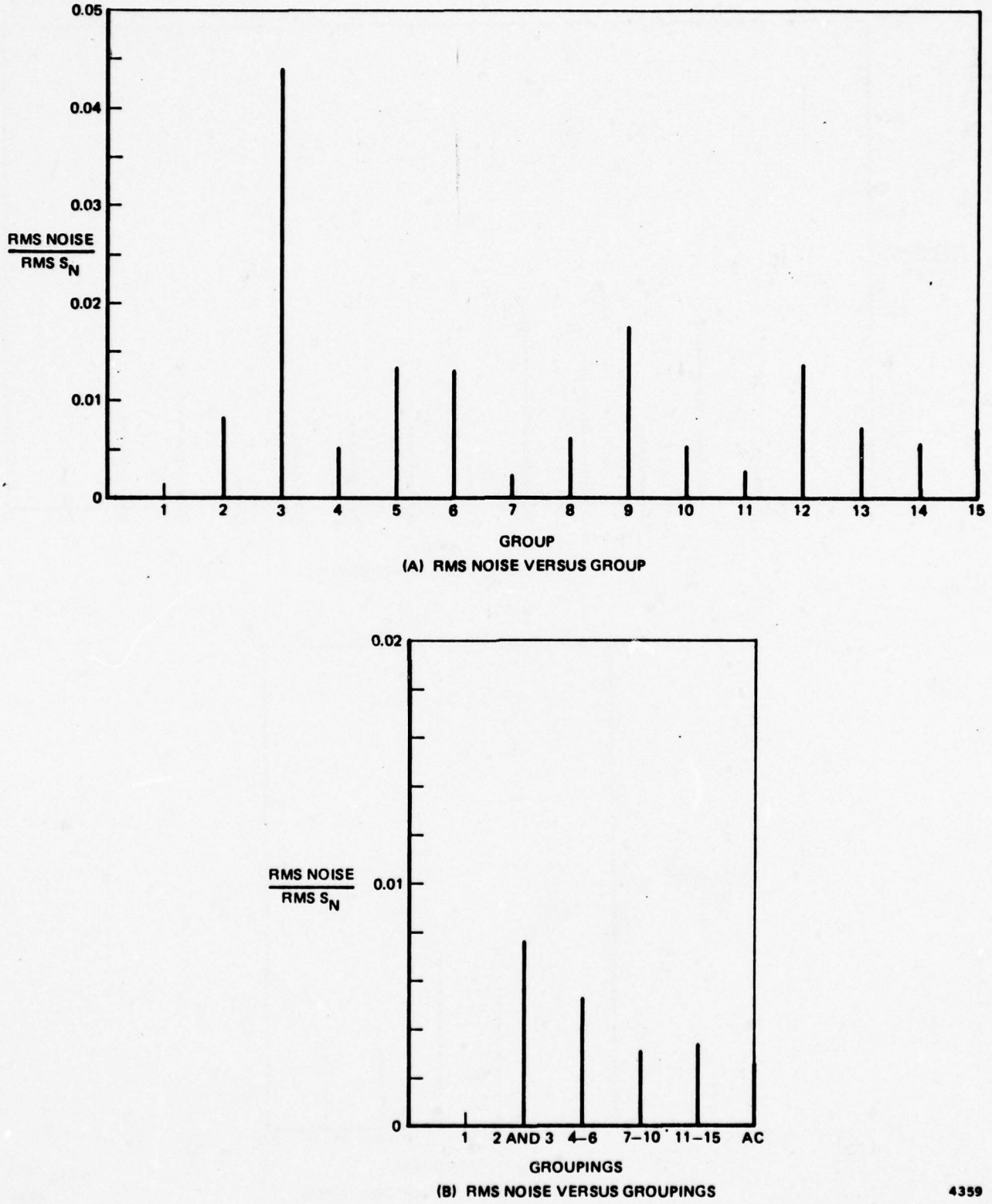


Figure 2-10. RMS NOISE LEVELS FOR S_n, SUM OF NORMALIZED AC MAGNITUDES

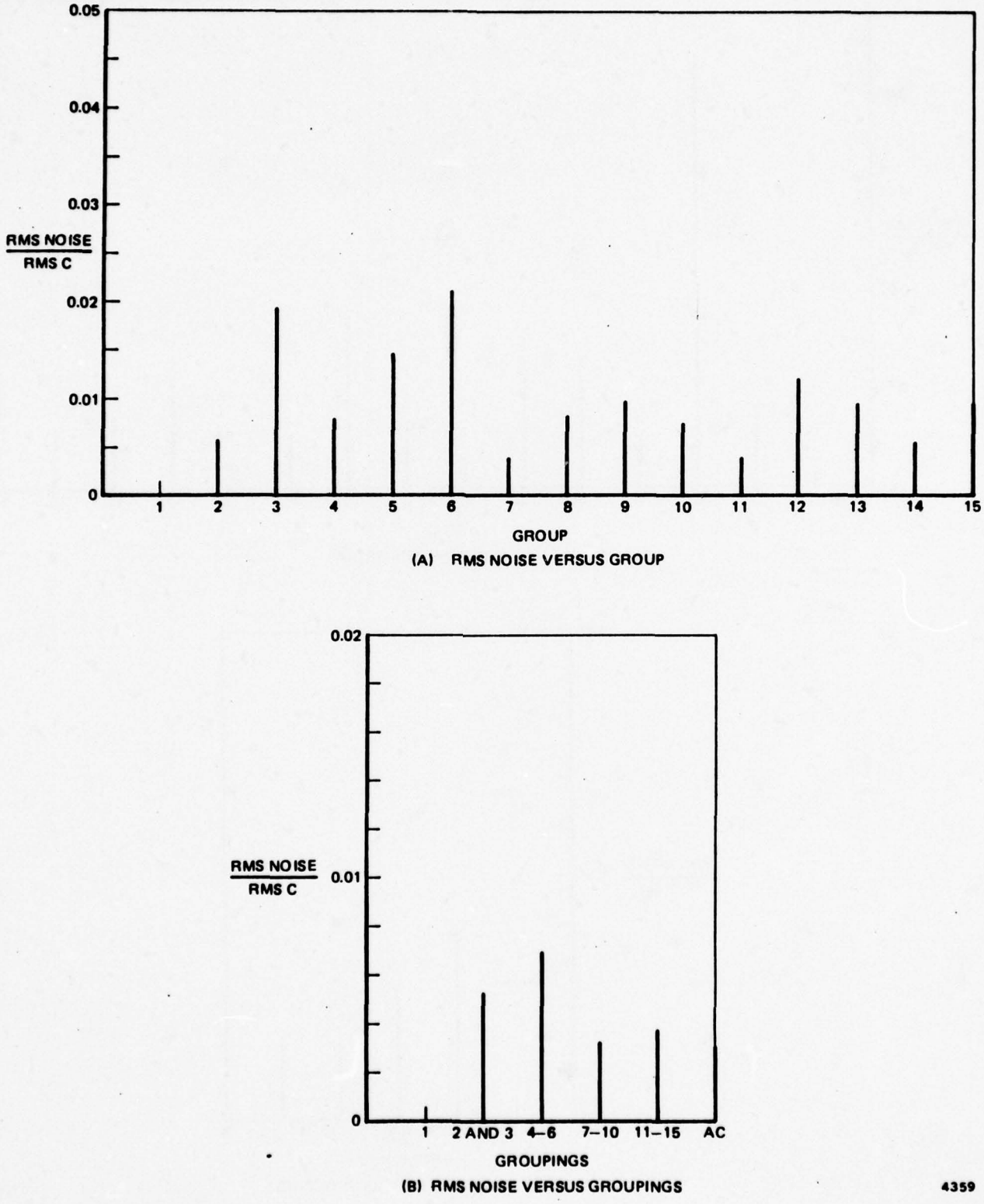


Figure 2-11. RMS NOISE LEVELS FOR C, SUM OF AC POWER

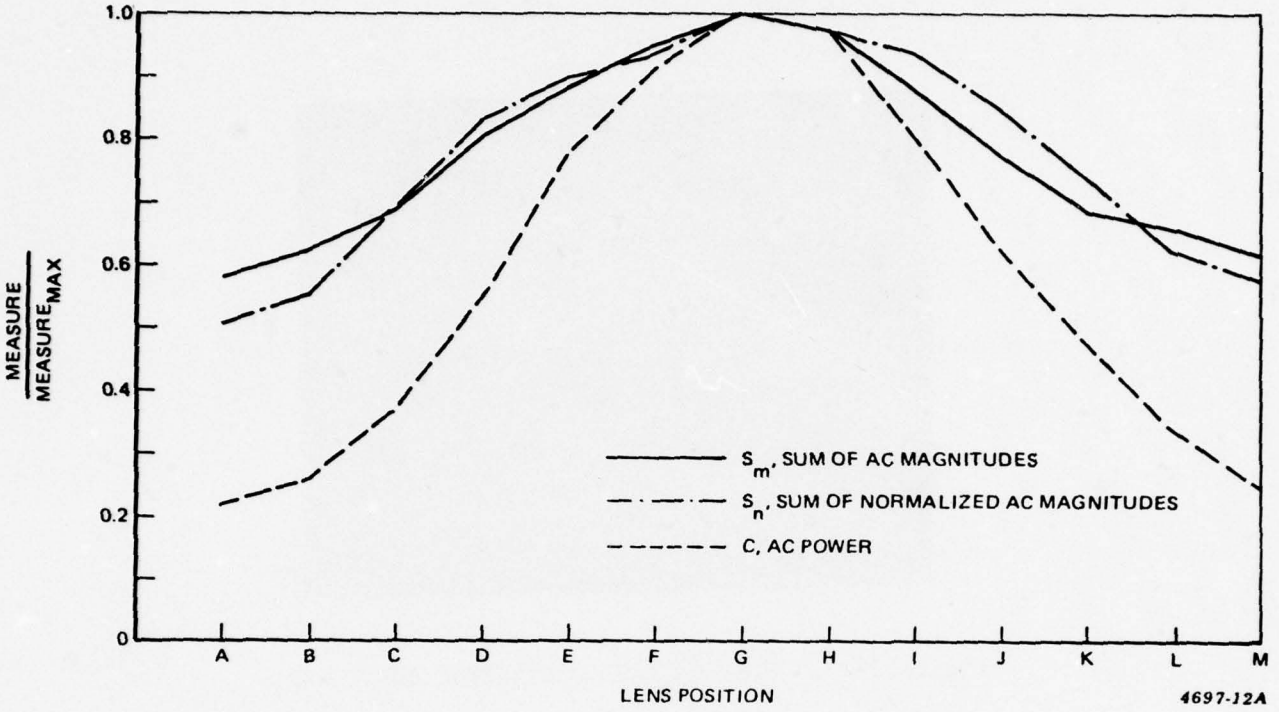


Figure 2-12. A COMPARISON OF FOCUS MEASURES

Table 2-1. FOCUS MEASURE SIGNAL-TO-NOISE RATIOS

Focus Measure	Signal-to-Noise Ratio (dB)
S_m, Sum of AC Magnitudes	43
S_n, Sum of Normalized AC Magnitudes	44
C, AC Power	49

by increased brightness. These areas were chosen so that texture and edge would be computed for both the center of the image, which is primarily texture, and the borders of the image, which are composed primarily of two singular edges.

The computed values of texture activity, edge activity, and their sum are shown as a function of lens position in Figure 2-14. As depicted in Figure 2-14, edge activ-

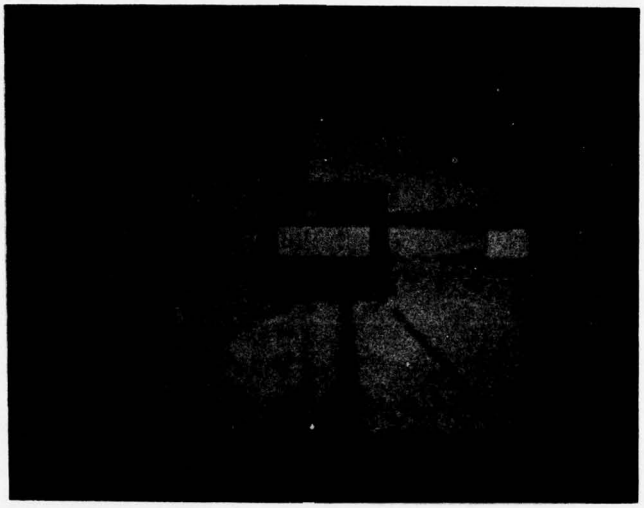


Figure 2-13. TEST PATTERN DESCRIPTOR INPUT AREAS

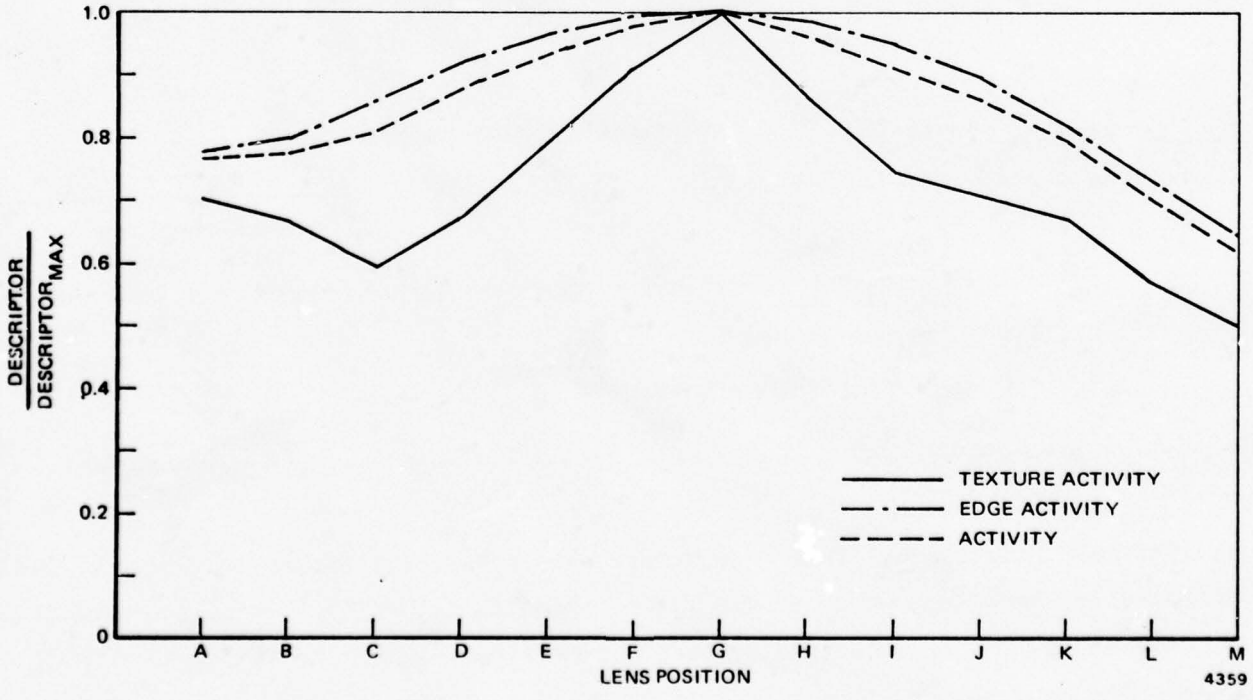


Figure 2-14. ACTIVITY DESCRIPTORS AS A FUNCTION OF LENS POSITION

ity behaves much like ac power. Texture activity, however, exhibits markedly different characteristics. While similar to the other curves in having a peak at lens position G, the texture activity curve also has a local minimum at lens position C. The precise effect that defocusing an image has on the division of activity into texture and edge information is not yet fully understood.

The computational complexity of each of the focus measures is tabulated in Table 2-2. The computation time of each measure is the time required for the instrumentation system's minicomputer to evaluate the state of focus of one 16 by 16 pel block of the image. The minicomputer (Computer Automation LSI-2) has an average instruction cycle time of 1 μsec. System overhead operations are not included in the computation time values.

As shown in Table 2-2, S_m is the simplest measure to compute as it evolves only 254 additions operations. S_n requires 190 power of two multiplications (shifts) as well as the additions operations. In order to compute C, the total ac power of the 16 by 16 pel block, each of the 255 ac coefficients must be squared before summation occurs. Because of the LSI-2's arithmetic format, squaring, or simple integer multiplication, uses two computer instruction cycles. The number of cycles required to compute C then is (255 x 2) + 254 = 765.

Unlike the previous three focus measures, which involve integer arithmetic, texture and edge activity require the use of floating point arithmetic. As a consequence, not only is the number of requisite instruction cycles much greater for each opera-

Table 2-2. FOCUS MEASURE COMPUTATIONAL COMPLEXITY

Measure	Computation Time	Real Time Compatibility
S_m, Sum of AC Magnitudes	254 μsec	Excellent
S_n, Sum of Normalized AC Magnitudes	464 μsec	Good
C, Total AC Power	765 μsec	Good
Texture and Edge Activity	~204 msec	Very Poor

tion, it is data dependent as well. The computation time for texture and edge activity listed in Table 2-2 is the time observed for the computation of these measures, scaled down by an estimate of percent system overhead. The texture and edge activity focus measure is clearly not compatible with a real-time system.

2.3.3 Summary

The performance of each focus measure in evaluating the state of focus of the video test pattern is summarized in Table 2-3.

Table 2-3. PERFORMANCE OF CANDIDATE FOCUS MEASURES

Measure	Sensitivity to Focus	Signal-to-Noise Ratio (dB)	Required Computation Time
S_m, Sum of AC Magnitudes	Fair	43	254 μsec
S_n, Sum of Normalized AC Magnitudes	Fair	44	464 μsec
C, AC Power	Good	49	756 μsec
Texture vs Edge	Poor	--	204 msec

Of the measures studied, the total ac power provides the focus measure with the best sensitivity to focus and the highest signal-to-noise ratio. The only disadvantage associated with using ac power as a measure of focus is that, compared to S_m and S_n, it is relatively expensive to compute. In terms of TV frame rates, however, its performance is still acceptable. For example, the ac power of about forty 16 by 16 pel image blocks could be computed during the elapsed time of one TV frame.

The measures of focus to be evaluated on the NVL FLIR imagery in the next experiment are C, the total ac power or image activity, and X and E, the texture and edge activities of the image. C was selected for its superior performance on the video

test pattern. Texture and edge activity are being evaluated further to determine their behavior on images that are of a more general nature than the video test pattern.

2.4 FLIR IMAGE EXPERIMENT

In order to ascertain the ability of ac power to discriminate focus on FLIR imagery, the five FLIR images supplied by NVL were used as input data in the test configuration shown in Figure 2-1. In each image, areas containing only background and areas containing both background and target information were considered. For the background data, the ac power of three fixed sets of eight 16 by 16 pel blocks was evaluated. These input areas are flagged by increased brightness in Figure 2-15. Target data were collected by restricting the input data to only the sector(s) that contained one of the targets in the image under consideration.

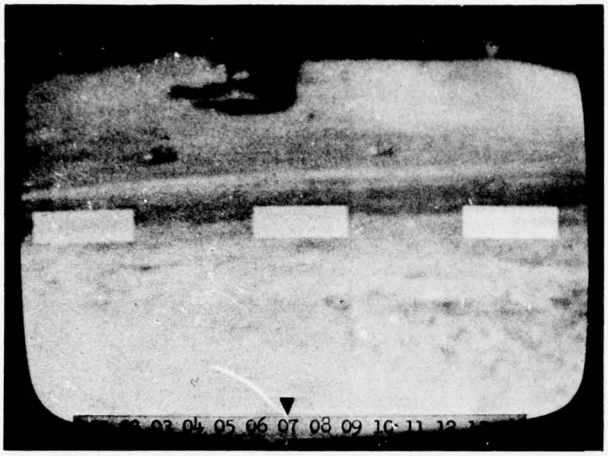
In each case, the ac power of the image input area was evaluated through four lens cycles and the results averaged for graphical presentation.

2.4.1 Image Data

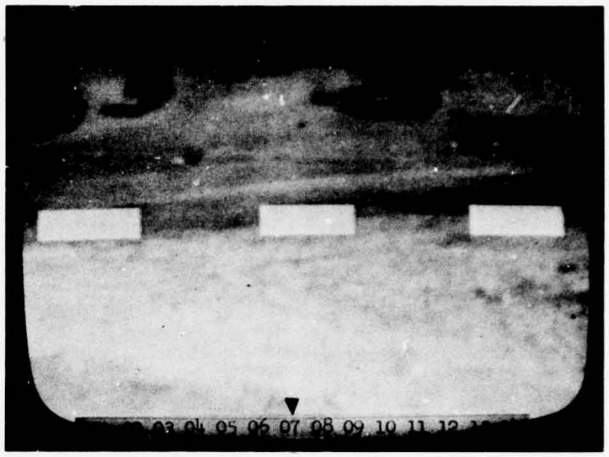
The variations in ac power as a function of lens position are shown in Figures 2-16 through 2-21 for background data, and in Figures 2-22 through 2-26 for target data.

As can be seen in Figures 2-16 through 2-21, the background data curves are relatively insensitive to focus. The curve for Image 5 (Figure 2-20), as a worst case example, is almost completely insensitive to changes in lens position. This effect is due primarily to the lack of activity in the input sectors of the image.

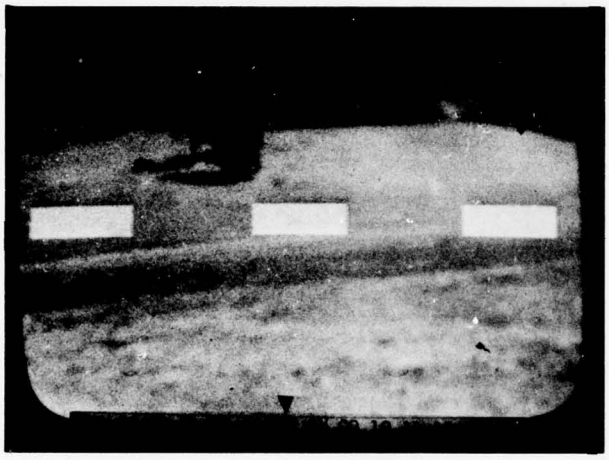
By choosing an alternate set of input sectors for Image 5 and evaluating the resultant fluctuations in ac power, a curve whose sensitivity to focus is comparable to those of the other images is obtained (Figure 2-21). The alternate sectors used are those immediately above the ones shown in Figure 2-15. The average activity increase obtained by moving the input areas in this instance is 33 percent. The central input area experiences a two-fold increase in activity.



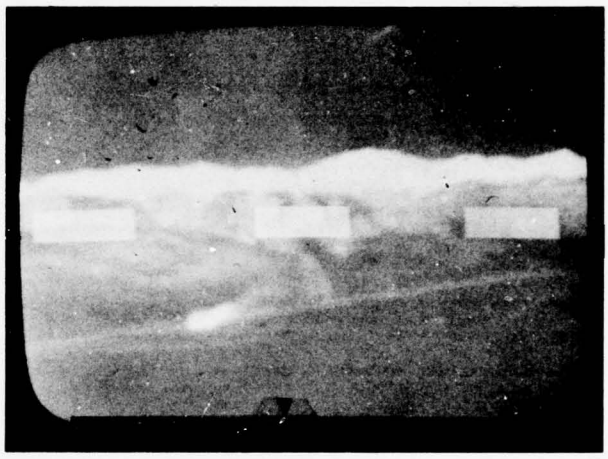
(a) IMAGE 1



(b) IMAGE 2



(c) IMAGE 3

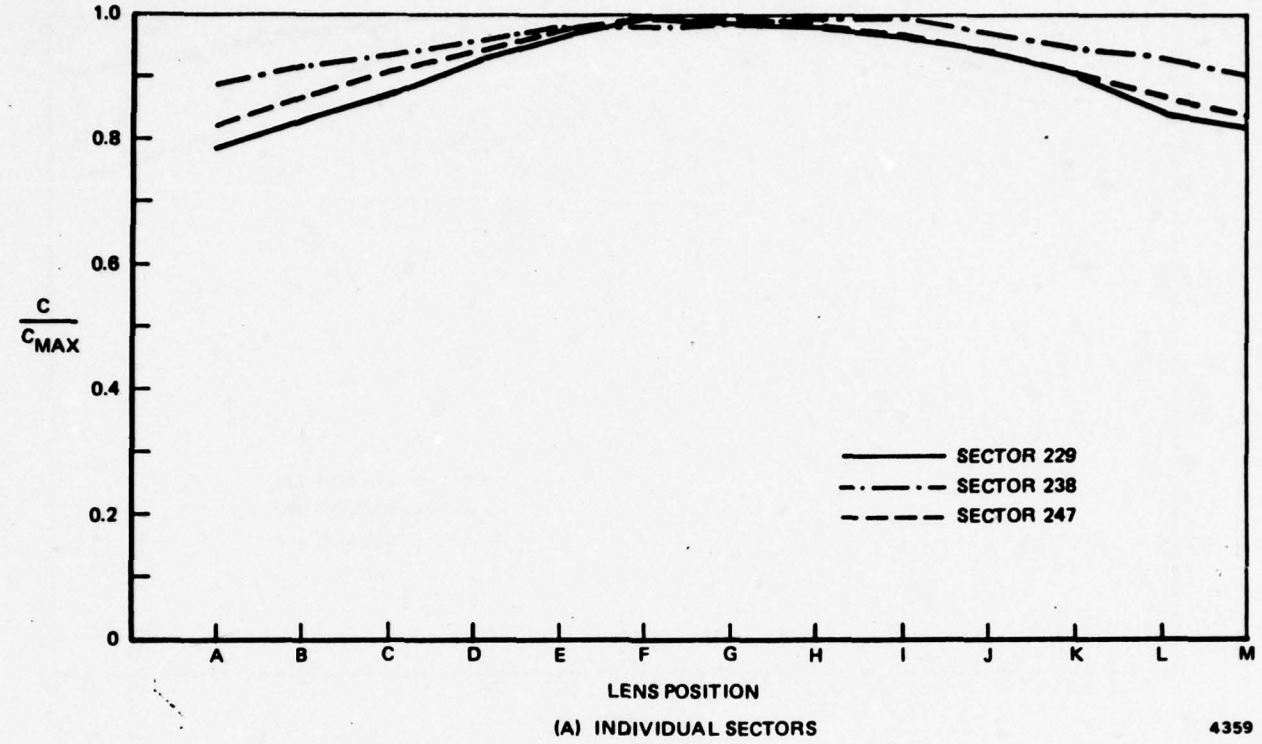


(d) IMAGE 4



(e) IMAGE 5

Figure 2-15. FLIR IMAGES WITH INPUT AREAS FLAGGED BY INCREASED BRIGHTNESS



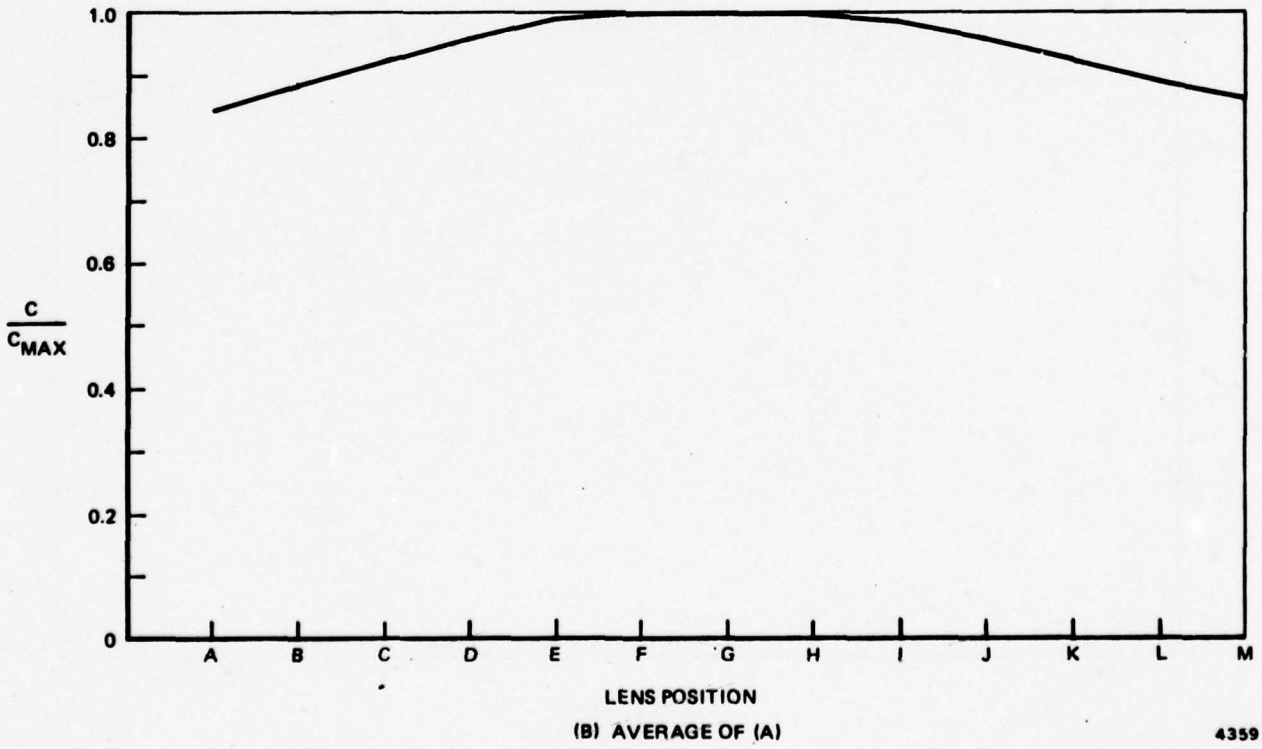


Figure 2-16. TOTAL AC POWER - IMAGE NO. 1

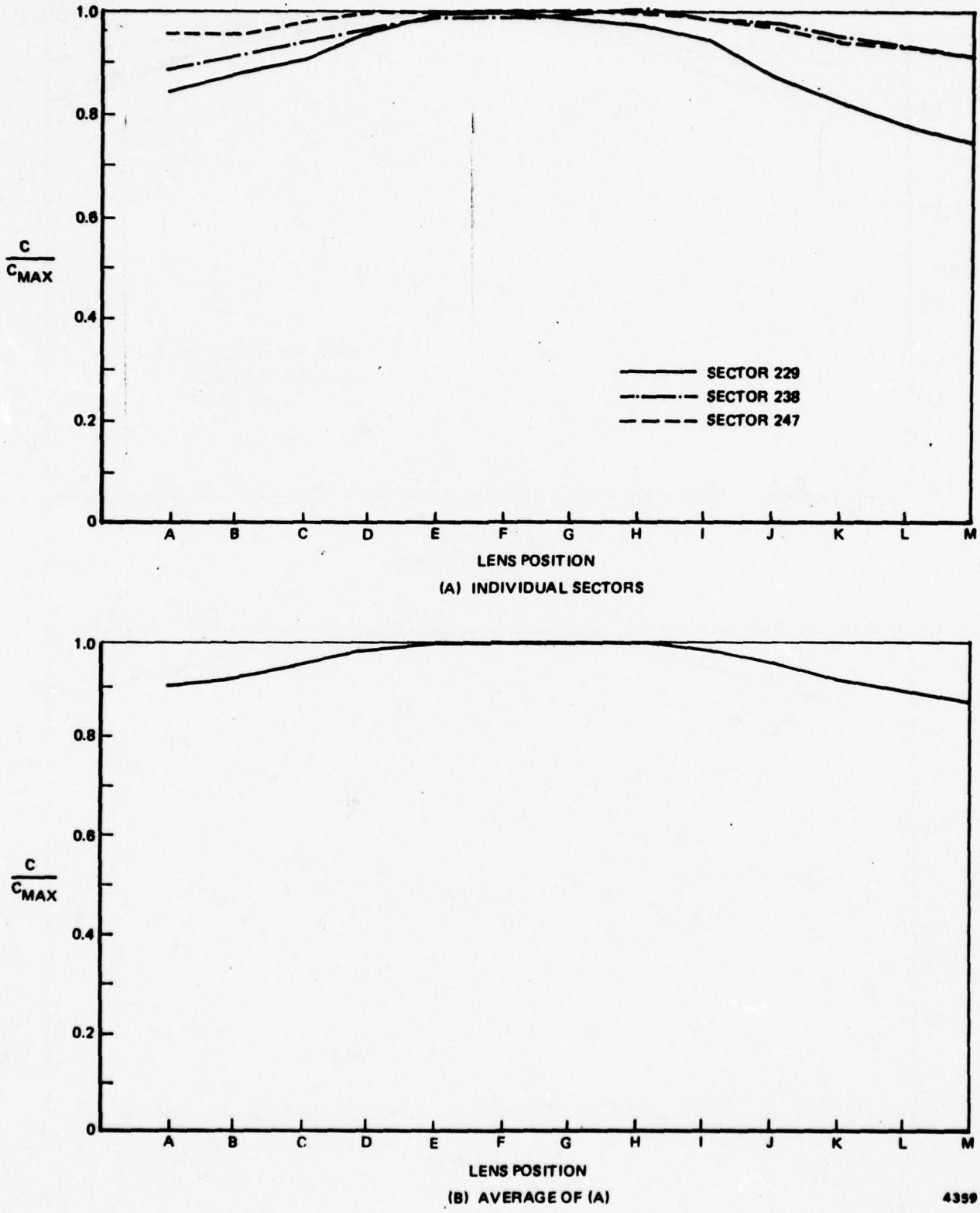


Figure 2-17. TOTAL AC POWER - IMAGE NO. 2

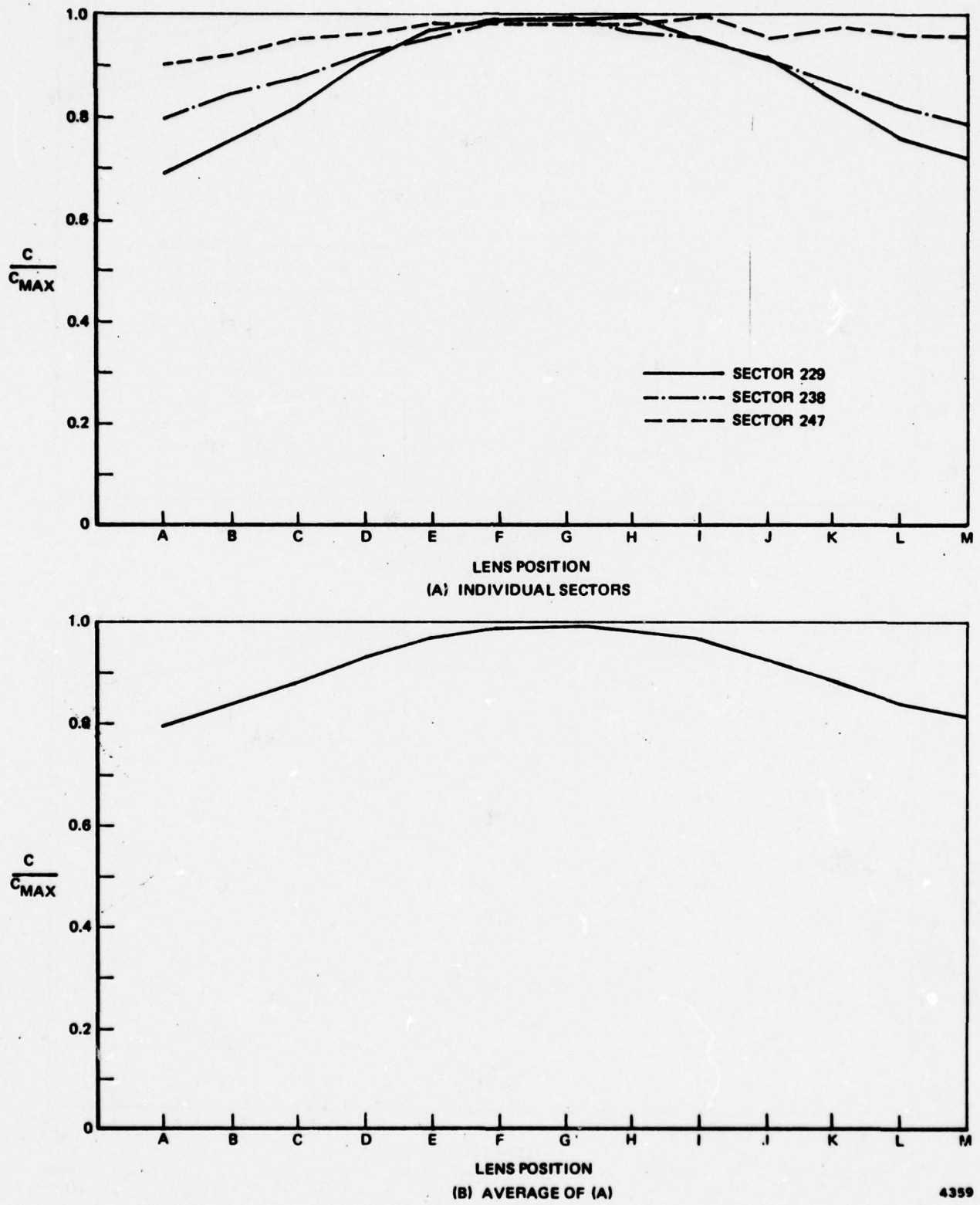


Figure 2-18. TOTAL AC POWER - IMAGE NO. 3

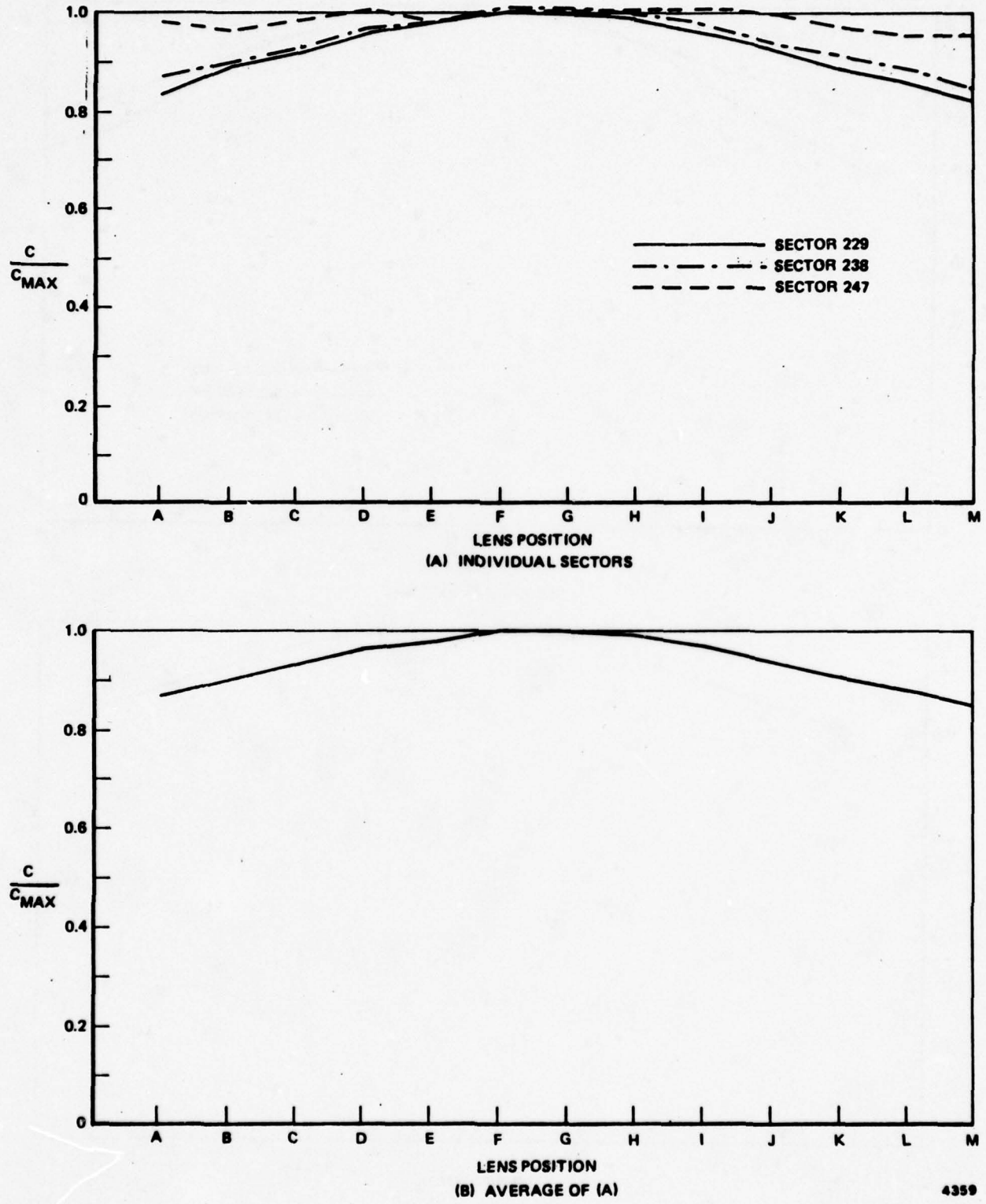


Figure 2-19. TOTAL AC POWER - IMAGE NO. 4

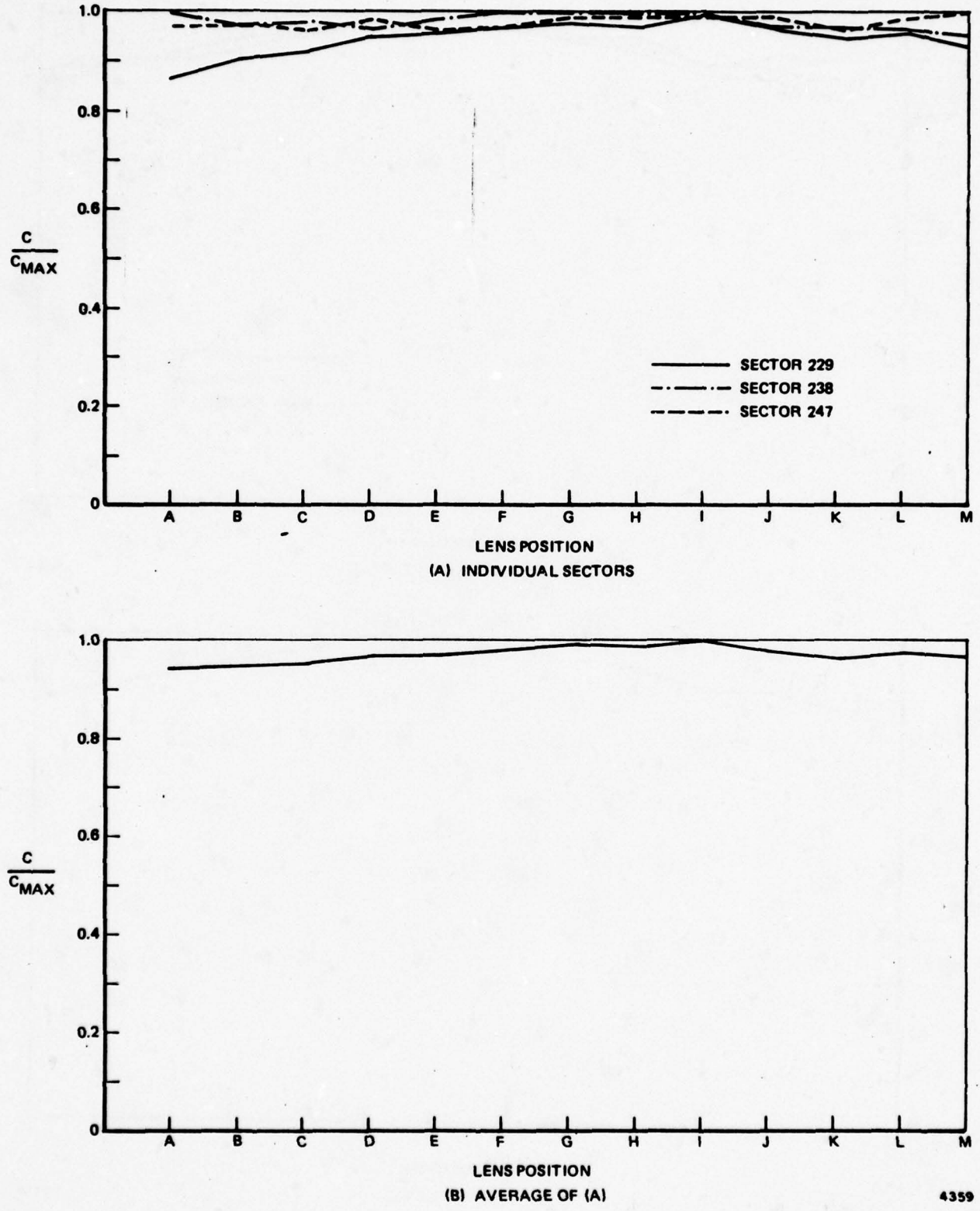
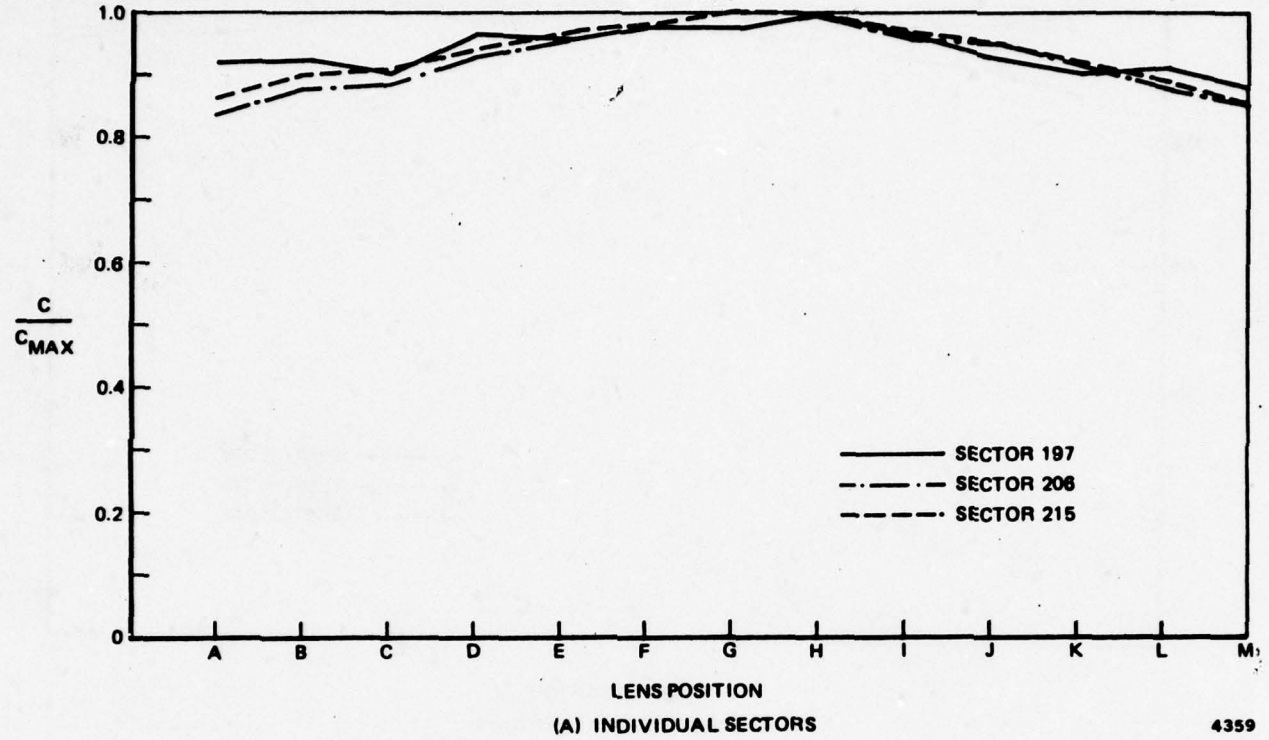


Figure 2-20. TOTAL AC POWER - IMAGE NO. 5



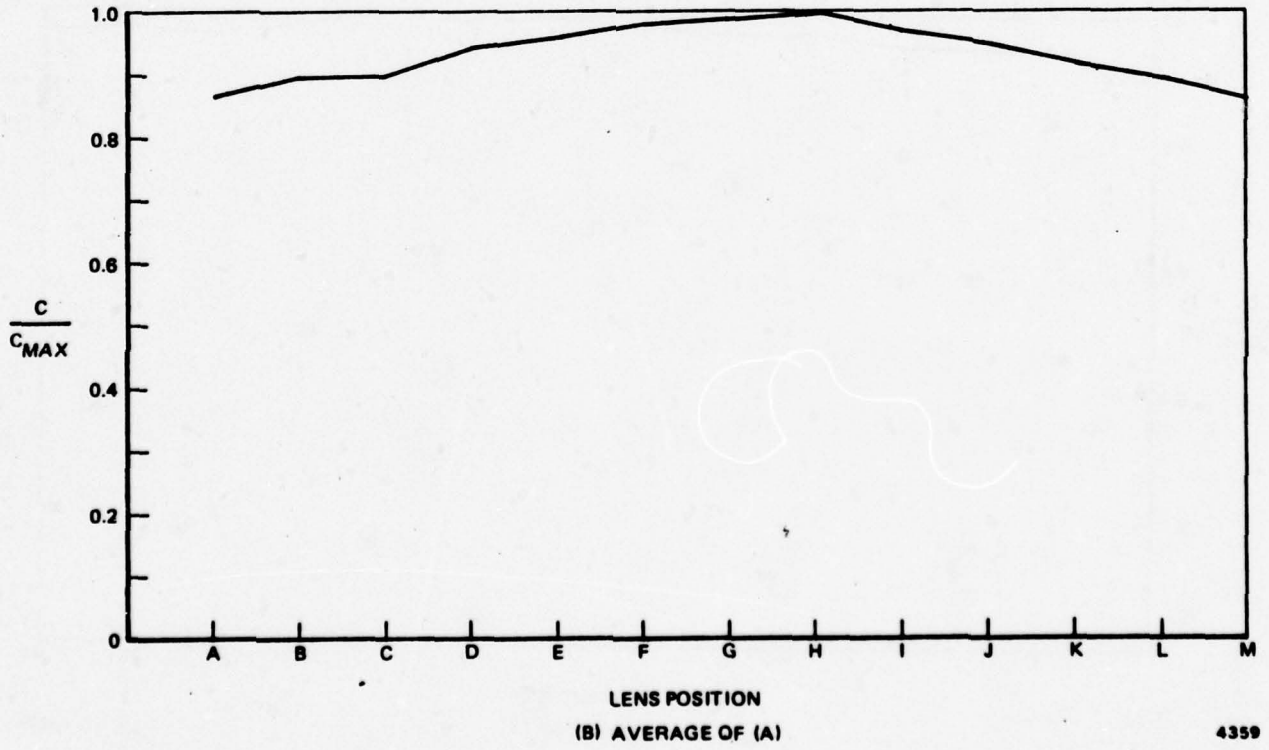


Figure 2-21. TOTAL AC POWER - IMAGE NO. 5 (ALTERNATE SECTORS)

The curves relating to the state of focus of the targets (Figures 2-22 through 2-26) exhibit an increased sensitivity to focus when compared to the background curves. This improvement in sensitivity is, again, primarily due to the increased activity of the target sectors. The ac power of the target sectors is about twice that of the average background sector. The target curves showing the most sensitivity to focus are those from Images 1, 2, and 3. The targets in each of these images are small, detailed, high-contrast objects. Objects of this type appear to provide the best information on which to base a state of focus evaluation.

To determine the effect of focus on the division of ac energy into texture and edge activity, the image descriptors were evaluated on representative areas of selected FLIR images. The test FLIR imagery shown in Figure 2-2 falls into two major groups. One group (Images 1, 2, and 3) exhibits small targets on a relatively bright background, while the other group (Images 4 and 5) consists of larger targets against a relatively dark background. One image from each group (Images 1 and 4) was selected to represent its group for the evaluation of the image descriptors. Both target and background areas of the two images were examined. The background areas are identical to those shown in Figure 2-15. For target input information, a 3 by 4 block of image sectors was centered about one of the targets in each image.

The curves relating the behavior of texture and edge activity as a function of lens position are shown in Figures 2-27 through 2-30. It can be seen that while texture and edge activity show approximately the same sensitivity to focus, regardless of the type of input data, the texture curve is much noisier. Averaging the values of texture activity over several lens cycles would probably reduce this noise level substantially, but the increase in computation time for such an approach would be prohibitive in a real-time system. For the background type data seen in the sample FLIR images provided by NVL, the division of activity into texture and edge information provides little to no additional information about the image's state of focus.

2.4.2 Summary of FLIR Image Experiments

The ac power of selected image sectors can provide a useful index of the state of focus of a FLIR image. The sensitivity of the measure is less for the FLIR images than that seen on the test pattern, but is still sufficient to form the basis of an error signal in an operational servo system.

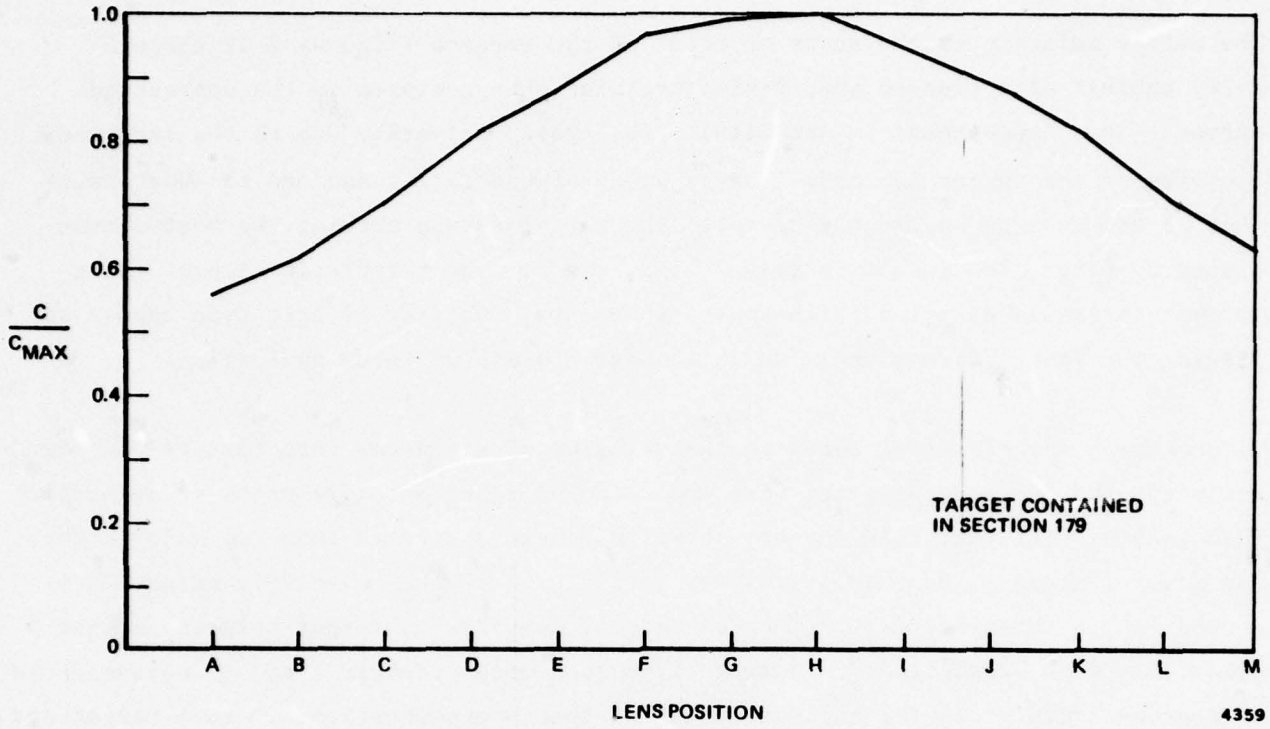


Figure 2-22. TOTAL AC POWER - IMAGE NO. 1 TARGET

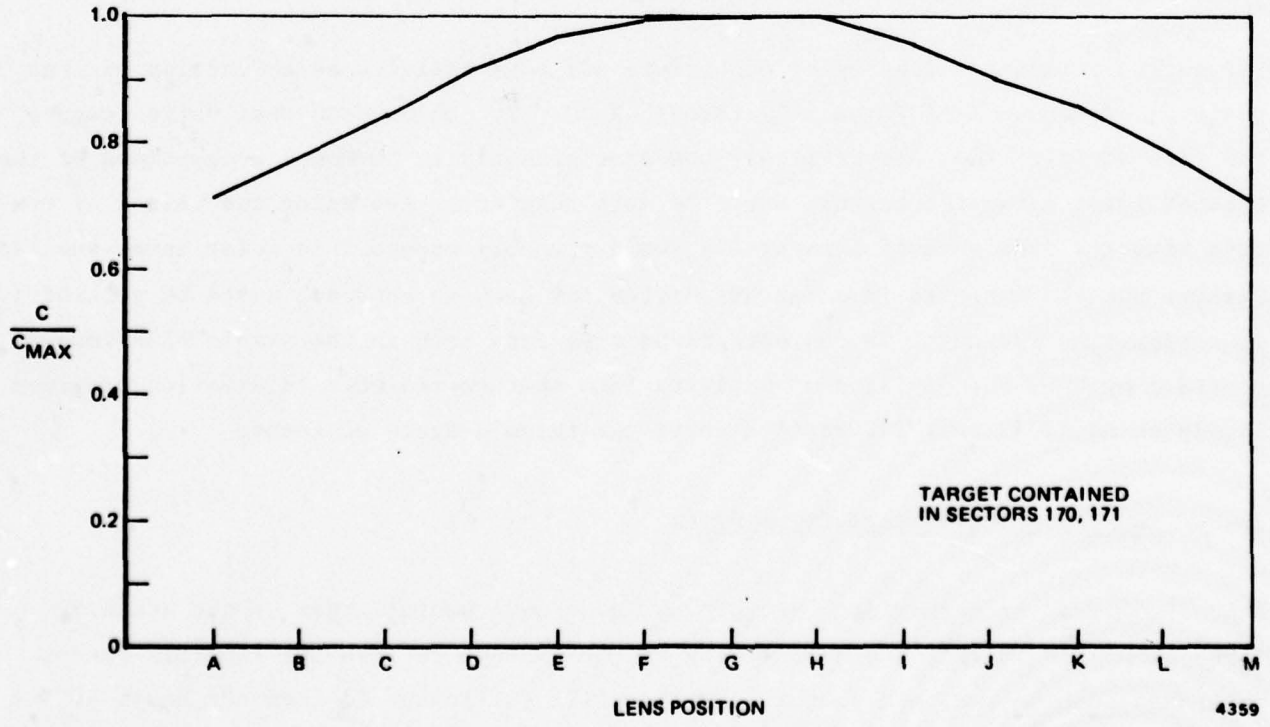


Figure 2-23. TOTAL AC POWER - IMAGE NO. 2 TARGET

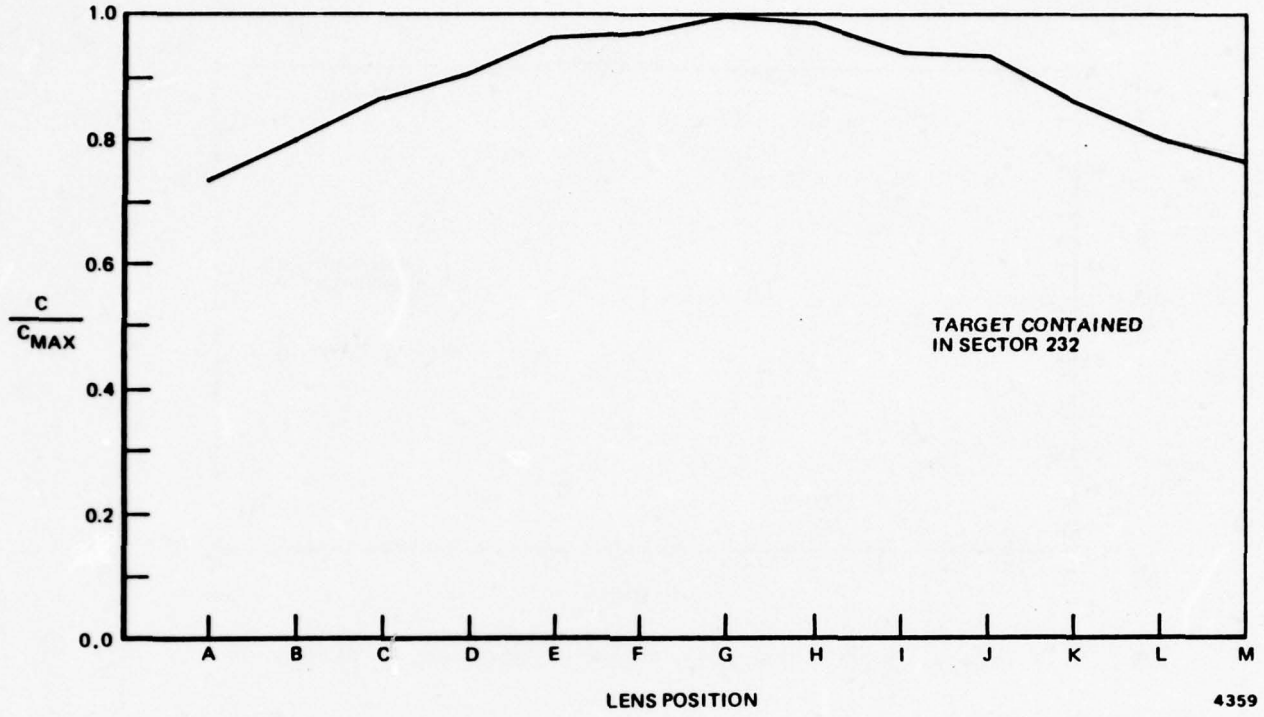


Figure 2-24. TOTAL AC POWER - IMAGE NO. 3 TARGET

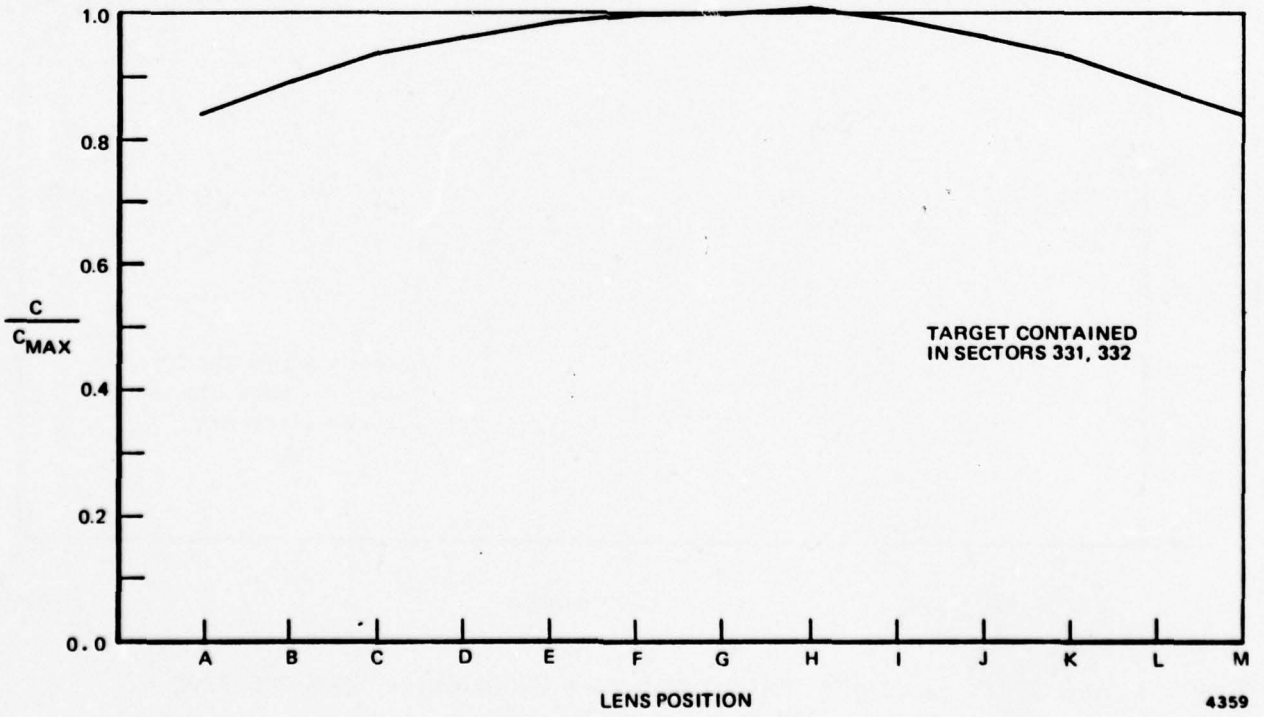


Figure 2-25. TOTAL AC POWER - IMAGE NO. 4 TARGET

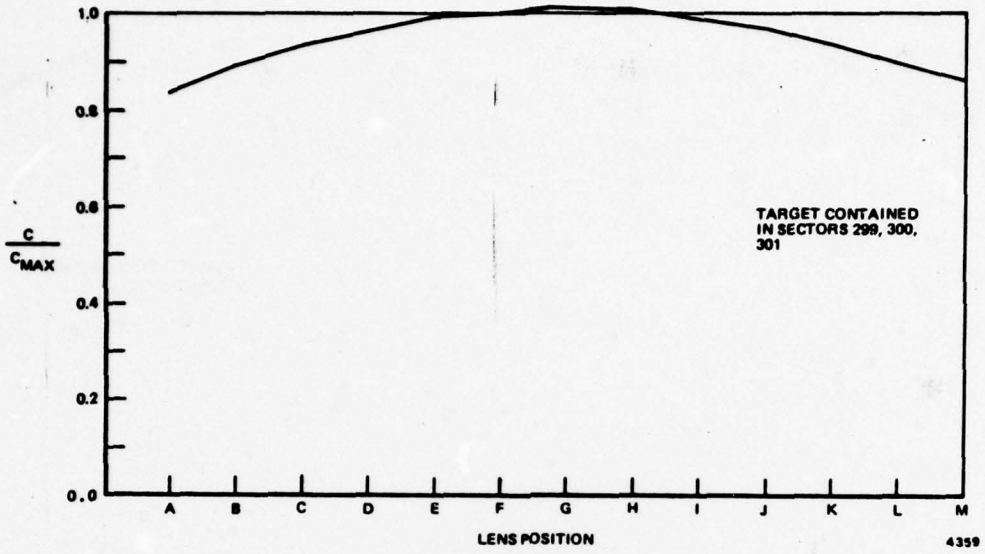


Figure 2-26. TOTAL AC POWER - IMAGE NO. 5 TARGET

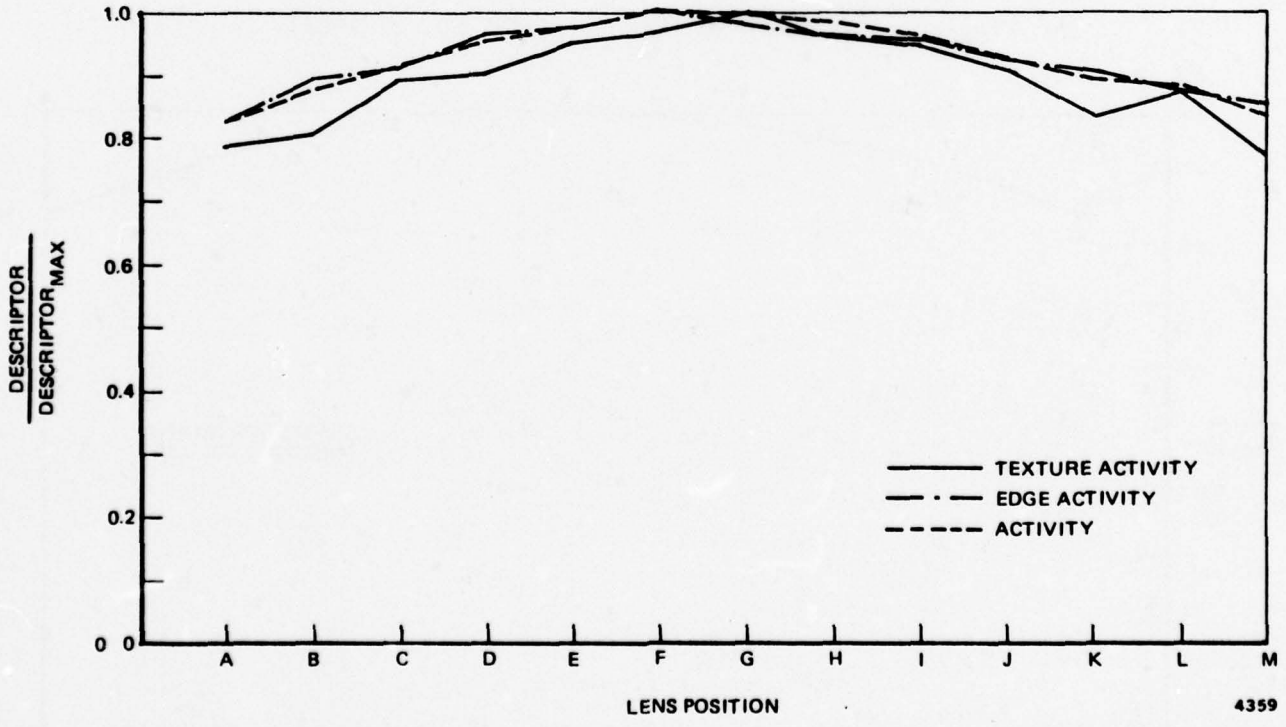


Figure 2-27. ACTIVITY DESCRIPTORS AS A FUNCTION OF LENS POSITION - IMAGE NO. 1 BACKGROUND

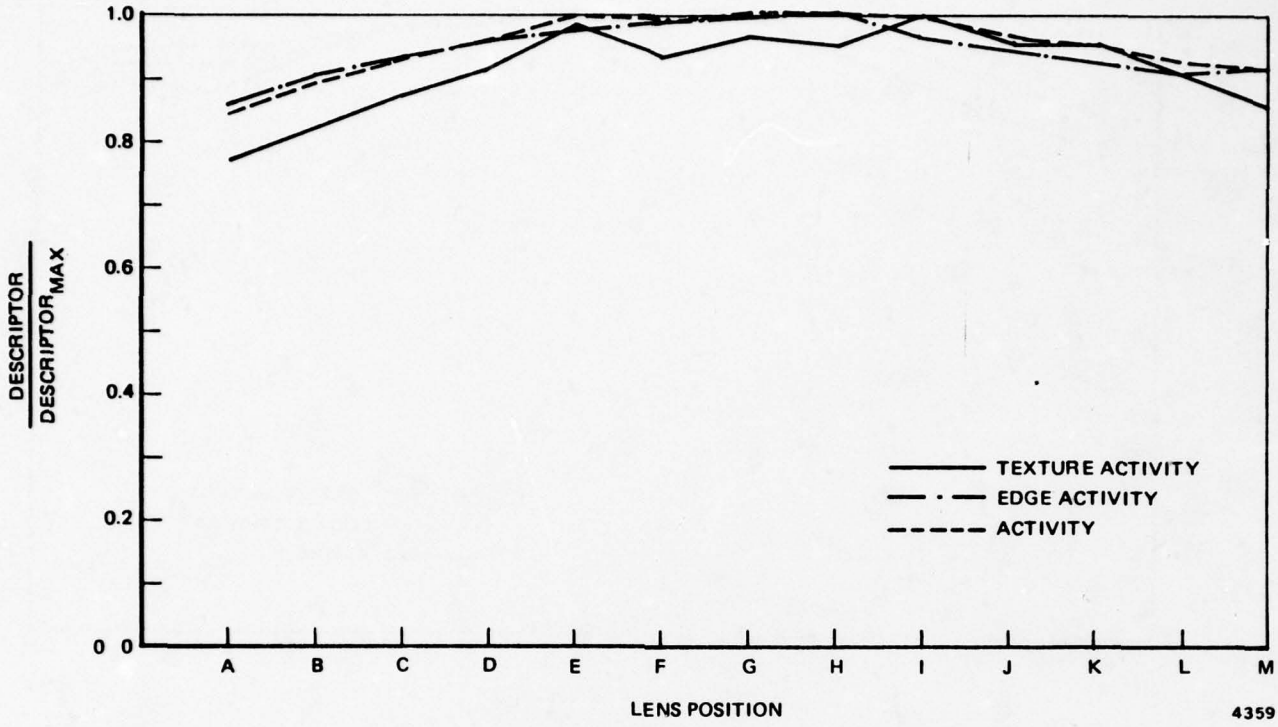


Figure 2-28. ACTIVITY DESCRIPTORS AS A FUNCTION OF LENS POSITION - IMAGE NO. 1 TARGET

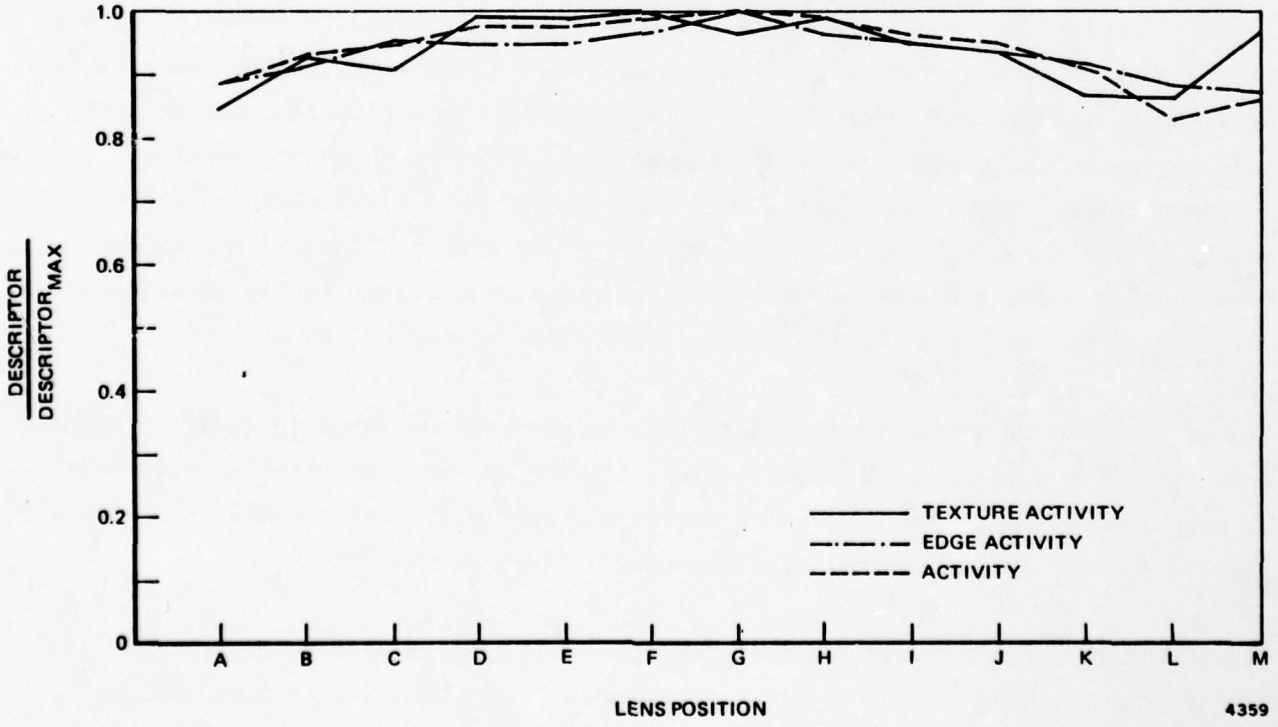


Figure 2-29. ACTIVITY DESCRIPTORS AS A FUNCTION OF LENS POSITION - IMAGE NO. 4 BACKGROUND

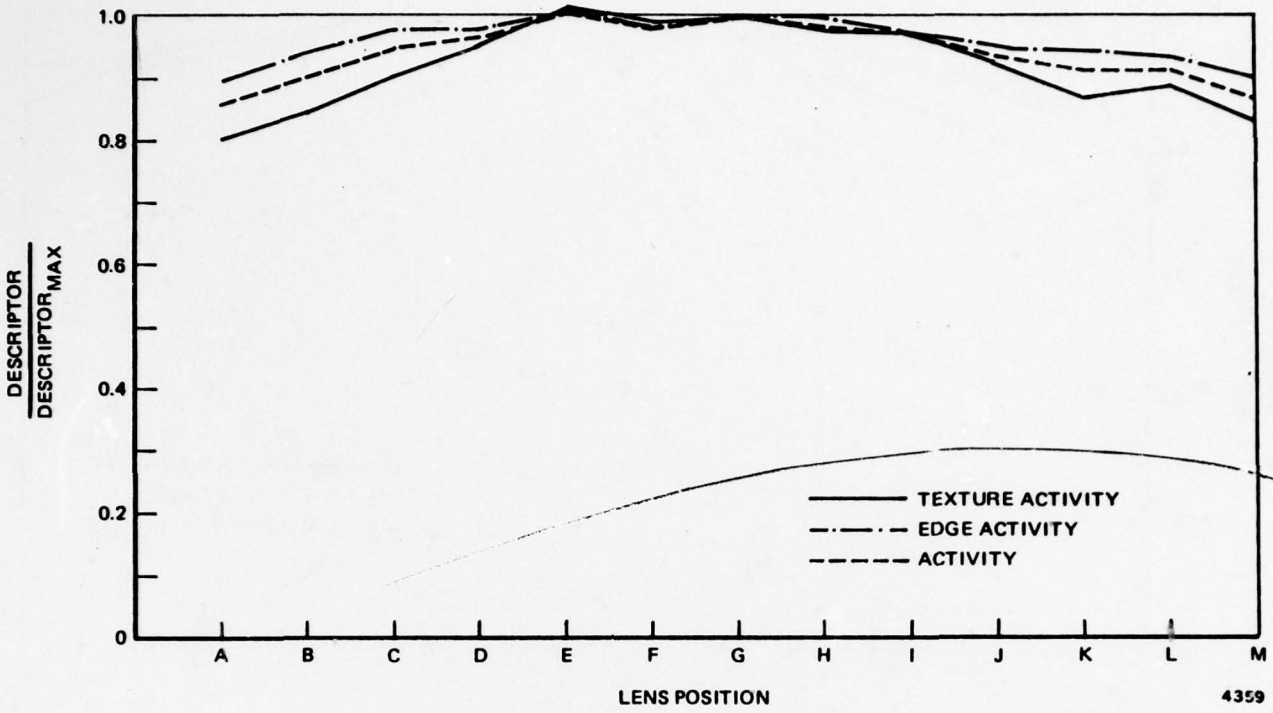


Figure 2-30. ACTIVITY DESCRIPTORS AS A FUNCTION OF LENS POSITION - IMAGE NO. 4 TARGET

The sensitivity of such a measure to focus is seen to be activity dependent; i.e., the ac power of a sector does not perform well as a focus measure when the maximum activity of the sector, regardless of focus, is low. This result of not focusing well on low activity image areas agrees with the manner in which an operator would focus an image. Rather than examining, say, the sky or a dim background while adjusting the lens, he would look at the same sharp edge or fine texture to determine focus. In both the human and automatic cases, sufficient ac high-frequency energy must be locally present to evaluate an image's state of focus.

If, in the case of automated focusing, each subarea of the image in question cannot be examined in a reasonable time, then a criterion of minimal activity would have to be established. Only those subareas satisfying such a criterion would then be considered as useful input data to the focus evaluating device.

The division of activity information into texture and edge appears to be of little use in evaluating the state of focus of an image. The behavior of both texture and edge activity as well as those of the other image descriptors will continue to be evaluated in subsequent experiments.

2.5 CONCLUSIONS

The local ac power, or activity, of an image demonstrates each of the five characteristics desired of a focus measure presented in paragraph 2.1. It can be computed in a time span that is compatible with a real time system; its sensitivity to focus can rival that of an observer: it has a good signal-to-noise ratio; it exhibits scene independence, subject to the limitations set forth a priori; and, since it is a simple scalar value that may be computed for each image frame, it is easily adaptable to an operational system. A real time imaging system that has an auto focus capability is certainly feasible and easily implementable using the two dimensional sensor processor.

SECTION 3
AUTO DEBLUR EXPERIMENTS

Forward Looking Infrared (FLIR) images may suffer from forms of blur not related to focus. For example, platform motion and the limited optical aperture may also contribute to the blurring of the FLIR image. Automated deblurring of FLIR imagery would therefore be a desirable feature of an advanced FLIR system.

Automatic image deblurring requires a characterization of the blur process and a method of inverting the blur. Blur characterization is treated first. A method of formulating an equivalent blur operation in an image's Haar domain is then developed, and the resulting operations presented. Approximations of the blur operations and deblur operations are computed and tested.

3.1 *ONE DIMENSIONAL BLUR MODEL*

The image and blur models presented here are one dimensional. Many of the results presented are easily extended to two dimensions because of separability considerations. However, others will require more effort. The extension to two dimensions will therefore be treated separately.

Let $\underline{x}$ be an arbitrary 16 x 1 data vector. A shift invariant operation (blur) is to be performed on $\underline{x}$. The sampled values of the impulse response of the operation are given by the 16 x 1 vector $\underline{b}$. If $\underline{y}$ is the output vector then

$$\underline{y} = \underline{x} * \underline{b} \quad \text{or} \quad y_i = \sum_j x_j b_{j-i} \tag{3-1}$$

where the subtraction of the subscripts is performed modulo 16. Equation (3-1) may also be expressed in more convenient matrix format by

$$\underline{y} = B\,\underline{x} \tag{3-2}$$

where the elements of B are given by $B_{ij} = b_{j-i}$.

The operator matrix B may represent any shift invariant operation (motion blur, focus blur, diffraction limiting, laplacian operator, etc.). The question we are addressing is: How does B manifest itself in the Haar domain?

By taking the Haar transform of equation (3-2) we arrive at

$$H\underline{y} = H B \underline{x}$$

or

$$H\underline{y} = \left[H B H^{-1}\right] H \underline{x} \qquad (3\text{-}3)$$

The quantity in brackets in equation (3-3) is a matrix that operates on the Haar transform of the input (a vector) to produce the Haar transform of the output (also a vector). This operation is independent of $\underline{x}$ and $\underline{y}$ and is a function of only B. Let

$$M = H B H^{-1} \qquad (3\text{-}4)$$

denote this mapping of the Haar transform of unblurred original into the Haar transform of the blurred image. Analysis of the forms of M obtained by the above relation may be facilitated by dividing M into submatrices or groups. A natural grouping to use with the Haar transform is shown in Table 3-1. The groupings shown in the table are those usually seen for the two-dimensional Haar transform. However, the interpretation of "group," as used here, is much different. As a consequence, the group numbering scheme is also different.

The matrices that result from equation (3-4) are operators on the one-dimensional (1-D) Haar domain. The groups of elements within the operator will therefore be referred to as "operator groups." The numbering system shown in Table 3-1 relates the effect of one of the five sequency groups in the input vector upon another sequency group in the output vector. For example, if an operator group is numbered 3,5 then the elements of that operator group represent the contribution of sequency group 3 in the input to sequency group 5 in the output.

In the experiments that follow, several forms of B and their resultant values of M are computed. Also, some of the effects of M on the Haar domain are discussed.

Table 3-1. GROUPING OF THE HAAR DOMAIN OPERATOR

$$M = \begin{bmatrix} 1,1 & 2,1 & 3,1 & 4,1 & 5,1 \\ 1,2 & 2,2 & 3,2 & 4,2 & 5,2 \\ 1,3 & 2,3 & 3,3 & 4,3 & 5,3 \\ 1,4 & 2,4 & 3,4 & 4,4 & 5,4 \\ 1,5 & 2,5 & 3,5 & 4,5 & 5,5 \end{bmatrix}$$

3.2 BLUR OPERATIONS

The image operations that were examined are:

a. Focus error (2 cases)
b. Diffraction limited aperture (2 cases)
c. One- and two-pel* translation.

Each is examined independently. In any actual FLIR system, however, they would most likely be encountered in various combinations. The translation operations blur the display by jittering the successive image frames.

3.2.1 Focus Error

Two cases of focus error were evaluated. The characteristic shape of the impulse response of a one-dimensional focus error is a rectangular pulse with unit area.

*pel = picture element

The sampled values of the two impulse responses used are shown in Figure 3-1. The Haar domain operators that result from these impulse responses are shown in Figures 3-2 and 3-3.

Examination of the Haar domain operators reveals many terms of significance off the main diagonal of the matrix. In fact, all of the main diagonal terms of group 5,5 are zero (i.e., the value of a high-sequency term in the output transform is independent of the value of its corresponding term in the input transform). Also note that operator groups not residing on the main diagonal contain relatively large terms. These results have an impact on automatic image deblurring in several ways. First, since the Haar domain operator M is poorly approximated by a diagonal matrix, blur detection algorithms may tend to be more complex than those in operating other transform domains. Second, deblurring approaches in the Haar domain will exhibit complexity comparable to that of M.

3.2.2 Diffraction Limited Aperture

Two impulse responses characteristic of a diffraction-limited optical system were evaluated. Squared synchronization (sync) functions were used in order to simulate the effects of an incoherently illuminated aperture. Equation (3-5) gives the sampled values of both impulse responses.



Figure 3-1. IMPULSE RESPONSES OF FOCUS ERROR

1.000	-0.000	-0.000	-0.000	0.0	-0.000	-0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,1	2,1	3,1			4,1						5,1				
-0.000	0.833	-0.000	0.000	-0.042	0.042	0.042	-0.042	-0.042	0.0	0.0	0.042	0.042	0.0	0.0	-0.042
1,2	2,2	3,2			4,2						5,2				
0.0	-0.000	0.750	-0.083	0.042	0.042	-0.042	-0.042	-0.042	0.083	0.083	-0.042	-0.042	0.0	0.0	-0.042
1,3	2,3	3,3			4,3						5,3				
0.0	0.000	-0.083	0.750	-0.042	-0.042	0.042	0.042	-0.042	0.0	0.0	-0.042	-0.042	0.083	0.083	-0.042
0.0	-0.167	0.083	-0.083	0.500	-0.083	0.0	-0.083	0.083	0.083	-0.083	0.0	0.0	0.0	0.0	-0.083
0.0	0.167	0.083	-0.083	-0.083	0.500	-0.083	0.0	0.0	-0.083	0.083	0.083	-0.083	0.0	0.0	0.0
1,4	2,4	3,4			4,4						5,4				
0.0	0.167	-0.083	0.083	0.0	-0.083	0.500	-0.083	0.0	0.0	0.0	-0.083	0.083	0.083	-0.083	0.0
0.0	-0.167	-0.083	0.083	-0.083	0.0	-0.083	0.500	-0.083	0.0	0.0	0.0	0.0	-0.083	0.083	0.083
0.0	-0.333	-0.167	-0.167	0.167	0.0	0.0	-0.167	0.0	-0.167	0.0	0.0	0.0	0.0	0.0	-0.167
0.0	0.0	0.333	0.0	0.167	-0.167	0.0	0.0	-0.167	0.0	-0.167	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.333	0.0	-0.167	0.167	0.0	0.0	0.0	-0.167	0.0	-0.167	0.0	0.0	0.0	0.0
0.0	0.333	-0.167	-0.167	0.0	0.167	-0.167	0.0	0.0	0.0	-0.167	0.0	-0.167	0.0	0.0	0.0
1,5	2,5	3,5			4,5						5,5				
0.0	0.333	-0.167	-0.167	0.0	-0.167	0.167	0.0	0.0	0.0	0.0	-0.167	0.0	-0.167	0.0	0.0
0.0	0.0	0.0	0.333	0.0	0.0	0.167	-0.167	0.0	0.0	0.0	0.0	-0.167	0.0	-0.167	0.0
0.0	0.0	0.0	0.333	0.0	0.0	-0.167	0.167	0.0	0.0	0.0	0.0	0.0	-0.167	0.0	-0.167
0.0	-0.333	-0.167	-0.167	-0.167	0.0	0.0	0.167	-0.167	0.0	0.0	0.0	0.0	0.0	-0.167	0.0

4483-19A

Figure 3-2. HAAR DOMAIN OPERATOR FOCUS ERROR (CASE 1)

1.000	-0.000	-0.000	-0.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,1	2,1	3,1		4,1				5,1							
-0.000	0.700	-0.000	0.000	-0.075	0.075	0.075	-0.075	-0.025	0.0	0.0	0.025	0.025	0.0	0.0	-0.025
1,2	2,2	3,2										5,2			
0.0	-0.000	0.550	-0.150	0.075	0.075	-0.075	-0.075	-0.025	0.050	0.050	-0.025	-0.025	0.0	0.0	-0.025
1,3	2,3	3,3		4,3				5,3							
0.0	0.000	-0.150	0.550	-0.075	-0.075	0.075	0.075	-0.025	0.0	0.0	-0.025	-0.025	0.050	0.050	-0.025
0.0	-0.300	0.150	-0.150	0.100	-0.150	0.0	-0.150	0.050	0.050	-0.050	0.0	0.0	0.0	0.0	-0.050
0.0	0.300	0.150	-0.150	-0.150	0.100	-0.150	0.0	0.0	-0.050	0.050	0.050	-0.050	0.0	0.0	0.0
1,4	2,4	3,4				4,4						5,4			
0.0	0.300	-0.150	0.150	0.0	-0.150	0.100	-0.150	0.0	0.0	0.0	-0.050	0.050	0.050	-0.050	0.0
0.0	-0.300	-0.150	0.150	-0.150	0.0	-0.150	0.100	-0.050	0.0	0.0	0.0	0.0	-0.050	0.050	0.050
0.0	-0.200	-0.100	-0.100	0.100	0.0	0.0	-0.100	0.0	0.100	0.0	0.0	0.0	0.0	0.0	0.100
0.0	0.0	0.200	0.0	0.100	-0.100	0.0	0.0	0.100	0.0	0.100	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.200	0.0	-0.100	0.100	0.0	0.0	0.0	0.100	0.0	0.100	0.0	0.0	0.0	0.0
0.0	0.200	-0.100	-0.100	0.0	0.100	-0.100	0.0	0.0	0.0	0.100	0.0	0.100	0.0	0.0	0.0
1,5	2,5	3,5				4,5						5,5			
0.0	0.200	-0.100	-0.100	0.0	-0.100	0.100	0.0	0.0	0.0	0.0	0.100	0.0	0.100	0.0	0.0
0.0	0.0	0.0	0.200	0.0	0.0	0.100	-0.100	0.0	0.0	0.0	0.0	0.100	0.0	0.100	0.0
0.0	0.0	0.0	0.200	0.0	0.0	-0.100	0.100	0.0	0.0	0.0	0.0	0.0	0.100	0.0	0.100
0.0	-0.200	-0.100	-0.100	-0.100	0.0	0.0	0.100	0.100	0.0	0.0	0.0	0.0	0.0	0.100	0.0

4483-20

Figure 3-3. HAAR DOMAIN OPERATOR FOCUS ERROR (CASE 2)

Case 1

$$b_i = \begin{cases} \text{sync}^2 \left[i\,\pi/2 \right], & i = 0, 7 \\ \text{sync}^2 \left[(8 - i)\,\pi/2 \right], & i = 8, 15 \end{cases} \tag{3-5a}$$

Case 2

$$b_i = \begin{cases} \text{sync}^2 \left[0.8\,i\,\pi/2 \right], & i = 0, 7 \\ \text{sync}^2 \left[(8 - i)\,0.8\,\pi/2 \right], & i = 8, 15 \end{cases} \tag{3-5b}$$

Case 1 simulates a system sampled at the Nyquist frequency. Case 2 represents a system that is sampled 25 percent faster than the Nyquist frequency. The Haar domain operators that result from these two image operations are shown in Figures 3-4 and 3-5.

Examination of Figures 3-3 and 3-4 shows that while the number of nonzero terms in the matrices is greater than in the case of focus error, the terms of the main diagonal are generally greater than their neighbors. Thus, the forms of M exhibited in Figures 3-4 and 3-5 appear to more closely approximate a diagonal operator than those shown for focus error.

Comparison of the two operators shows Case 1 to have a greater concentration of energy along its diagonal than Case 2. This is true particularly in the highest sequences. The possibility of easily correcting for Case 1 diffraction limiting is much greater than for Case 2.

3.2.3 Translation

The Haar domain operators resulting from one- and two-pel rigid translations are shown in Figures 3-6 and 3-7, respectively. In each case, translation was to the right in the image-row vector (or down in a column vector).

1.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000
1,1	2,1	3,1		4,1				5,1							
0.0	0.826	-0.000	0.000	-0.027	0.027	0.027	-0.027	-0.025	-0.002	0.002	0.025	0.025	0.002	-0.002	-0.025
1,2	2,2	3,2		4,2				5,2							
0.0	-0.000	0.760	-0.066	0.027	0.027	-0.027	-0.027	-0.021	0.048	0.048	-0.021	-0.024	-0.003	-0.003	-0.024
1,3	2,3	3,3		4,3				5,3							
0.0	0.0	-0.066	0.760	-0.027	-0.027	0.027	0.027	-0.024	-0.003	-0.003	-0.024	-0.021	0.048	0.048	-0.021
0.0	-0.109	0.055	-0.055	0.605	-0.045	-0.002	-0.045	0.046	0.046	-0.043	-0.003	-0.001	-0.001	-0.003	-0.043
0.0	0.109	0.055	-0.055	-0.045	0.605	-0.045	-0.002	-0.003	-0.043	0.046	0.046	-0.043	-0.003	-0.001	-0.001
1,4	2,4	3,4		4,4				5,4							
0.0	0.109	-0.055	0.055	-0.002	-0.045	0.605	-0.045	-0.001	-0.001	-0.003	-0.043	0.046	0.046	-0.043	-0.003
0.0	-0.109	-0.055	0.055	-0.045	-0.002	-0.045	0.605	-0.043	-0.003	-0.001	-0.001	-0.003	-0.043	0.046	0.046
0.0	-0.204	-0.083	-0.098	0.092	-0.005	-0.002	-0.085	0.305	-0.115	-0.016	-0.006	-0.004	-0.006	-0.016	-0.115
0.0	-0.015	0.192	-0.011	0.092	-0.085	-0.002	-0.005	-0.115	0.305	-0.115	-0.016	-0.006	-0.004	-0.006	-0.016
0.0	0.015	0.192	-0.011	-0.085	0.092	-0.005	-0.002	-0.016	-0.115	0.305	-0.115	-0.016	-0.006	-0.004	-0.006
0.0	0.204	-0.083	-0.098	-0.005	0.092	-0.085	-0.002	-0.006	-0.016	-0.115	0.305	-0.115	-0.016	-0.006	-0.004
1,5	2,5	3,5		4,5				5,5							
0.0	0.204	-0.098	-0.083	-0.002	-0.085	0.092	-0.005	-0.004	-0.006	-0.016	-0.115	0.305	-0.115	-0.016	-0.006
0.0	0.015	-0.011	0.192	-0.002	-0.005	0.092	-0.085	-0.006	-0.004	-0.006	-0.016	-0.115	0.305	-0.115	-0.016
0.0	-0.015	-0.011	0.192	-0.005	-0.002	-0.085	0.092	-0.016	-0.006	-0.004	-0.006	-0.016	-0.115	0.305	-0.115
0.0	-0.204	-0.098	-0.083	-0.085	-0.002	-0.005	0.092	-0.115	-0.016	-0.006	-0.004	-0.006	-0.016	-0.115	0.305

4483-21

Figure 3-4. HAAR DOMAIN OPERATOR APERTURE BLUR (CASE 1)

1.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000	0.000	-0.000	-0.000	0.0	0.0	-0.000	-0.000	0.000
	1,1	2,1	3,1			4,1					5,1				
0.0	0.785	-0.000	0.000	-0.035	0.035	0.035	-0.035	-0.029	-0.001	0.001	0.029	0.029	0.001	-0.001	-0.029
	1,2	2,2	3,2			4,2					5,2				
0.0	-0.000	0.700	-0.086	0.035	0.035	-0.035	-0.035	-0.027	0.058	0.058	-0.027	-0.030	-0.001	-0.001	-0.030
	1,3	2,3	3,3			4,3					5,3				
0.0	0.000	-0.086	0.700	-0.035	-0.035	0.035	0.035	-0.030	-0.001	-0.001	-0.030	-0.027	0.058	0.058	-0.027
0.0	-0.139	0.069	-0.069	0.504	-0.057	-0.004	-0.057	0.057	0.057	-0.054	-0.003	0.000	0.000	-0.003	-0.054
	1,4	2,4	3,4			4,4					5,4				
0.0	0.139	0.069	-0.069	-0.057	0.504	-0.057	-0.004	-0.003	-0.054	0.057	0.057	-0.054	-0.003	0.000	0.000
0.0	0.139	-0.069	0.069	-0.004	-0.057	0.504	-0.057	0.000	0.000	-0.003	-0.054	0.057	0.057	-0.054	-0.003
0.0	-0.139	-0.069	0.069	-0.057	-0.004	-0.057	0.504	-0.054	-0.003	0.000	0.000	-0.003	-0.054	0.057	0.057
0.0	-0.235	-0.109	-0.119	0.113	-0.006	0.001	-0.108	0.177	-0.101	0.010	0.006	-0.000	0.006	0.010	-0.101
0.0	-0.010	0.232	-0.003	0.113	-0.108	0.001	-0.006	-0.101	0.177	-0.101	0.010	0.006	-0.000	0.006	0.010
0.0	0.010	0.232	-0.003	-0.108	0.113	-0.006	0.001	0.010	-0.101	0.177	-0.101	0.010	0.006	-0.000	0.006
0.0	0.235	-0.109	-0.119	-0.006	0.113	-0.108	0.001	0.006	0.010	-0.101	0.177	-0.101	0.010	0.006	-0.000
	1,5	2,5	3,5			4,5					5,5				
0.0	0.235	-0.119	-0.109	0.001	-0.108	0.113	-0.006	-0.000	0.006	0.010	-0.101	0.177	-0.101	0.010	0.006
0.0	0.010	-0.003	0.232	0.001	-0.006	0.113	-0.108	0.006	-0.000	0.006	0.010	-0.101	0.177	-0.101	0.010
0.0	-0.010	-0.003	0.232	-0.006	0.001	-0.108	0.113	0.010	0.006	-0.000	0.006	0.010	-0.101	0.177	-0.101
0.0	-0.235	-0.119	-0.109	-0.108	0.001	-0.006	0.113	-0.101	0.010	0.006	-0.000	0.006	0.010	-0.101	0.177

4483-22

Figure 3-5. HAAR DOMAIN OPERATOR APERTURE BLUR (CASE 2)

1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,1	2,1	3,1		4,1				5,1							
0.0	0.750	0.125	-0.125	0.0	0.125	0.0	-0.125	0.0	0.0	0.0	0.125	0.0	0.0	0.0	-0.125
1,2	2,2	3,2		4,2				5,2							
0.0	-0.250	0.625	-0.125	0.250	-0.125	0.0	-0.125	0.0	0.250	0.0	-0.125	0.0	0.0	0.0	-0.125
1,3	2,3	3,3		4,3				5,3							
0.0	0.250	-0.125	0.625	0.0	-0.125	0.250	-0.125	0.0	0.0	0.0	-0.125	0.0	0.250	0.0	-0.125
0.0	-0.500	-0.250	-0.250	0.250	0.0	0.0	-0.250	0.500	-0.250	0.0	0.0	0.0	0.0	0.0	-0.250
0.0	0.0	0.500	0.0	-0.250	0.250	0.0	0.0	0.0	-0.250	0.500	-0.250	0.0	0.0	0.0	0.0
1,4	2,4	3,4		4,4				5,4							
0.0	0.500	-0.250	-0.250	0.0	-0.250	0.250	0.0	0.0	0.0	0.0	-0.250	0.500	-0.250	0.0	0.0
0.0	0.0	0.0	0.500	0.0	0.0	-0.250	0.250	0.0	0.0	0.0	0.0	0.0	-0.250	0.500	-0.250
0.0	-1.000	-0.500	-0.500	-0.500	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	0.0	-0.500
0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	1.000	0.0	-0.500	-0.500	0.0	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0
1,5	2,5	3,5		4,5				5,5							
0.0	1.000	-0.500	-0.500	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	-0.500	-0.500	0.0	0.0
0.0	0.0	0.0	1.000	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	-0.500	-0.500	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	-0.500	-0.500

4483-23

Figure 3-6. HAAR DOMAIN OPERATOR ONE-PEL TRANSLATION

1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,1	2,1	3,1			4,1						5,1				
0.0	0.500	0.250	-0.250	0.0	0.250	0.0	-0.250	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,2	2,2	3,2			4,2						5,2				
0.0	-0.500	0.250	-0.250	0.500	-0.250	0.0	-0.250	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,3	2,3	3,3			4,3						5,3				
0.0	0.500	-0.250	0.250	0.0	-0.250	0.500	-0.250	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	-1.000	-0.500	-0.500	-0.500	0.0	0.0	-0.500	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	1.000	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1,4	2,4	3,4			4,4						5,4				
0.0	1.000	-0.500	-0.500	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.000	0.0	0.0	-0.500	-0.500	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0	0.0
1,5	2,5	3,5			4,5						5,5				
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0	0.0
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.000	0.0

4483-24

Figure 3-7. HAAR DOMAIN OPERATOR TWO-PEL TRANSLATION

Examination of Figure 3-6 shows that some of the highest sequency terms in the output Haar spectrum use significant information from the lowest ac sequency in the input Haar spectrum. This is shown immediately by the ± 1's in operator group 2,5. This form of M shows that translation invariant measures that apply equally to all of the high-sequency Haar coefficients are at best very complex. The possibility of detecting and correcting image motion in the Haar domain is quite good, however. This is demonstrated by the sparseness of the two Haar domain operators.

Figure 3-7, the Haar domain operator resulting from a two-pel translation, shows an interesting result. The form of operator group 5,5 is simply that of a one-element translation operation. Thus, translating the image by two pels translates the highest sequency group of the image's Haar transform by one term. Also note that the upper left quarter of Figure 3-7 is just a scaled down copy of Figure 3-6. Thus, translating two pels operates on the middle sequencies similar to the manner by which translating one pel operates on the higher sequencies.

3.2.4 Summary

Each of the Haar domain operators presented in this series of experiments is non-diagonal in form. The net impact of this characteristic is to make both blur detection and inversion processes more difficult. A diagonal representation of the blur process is required. Such diagonal operators would lend themselves more directly to simple detection and inversion methods.

3.3 A DIAGONAL REPRESENTATION OF BLUR

A straightforward method of approximating a Haar domain operation, M, with a diagonal matrix is to simply ignore those terms in M that are off the main diagonal. This approach has been shown[1] to not function well.

Another approach to developing a diagonal representation of a Haar domain operator is to require that, on the average, the effect of using M is equivalent to using its diagonal representation.

---

[1] D. J. Ketcham, et al , Image Enhancement for Cockpit Displays, Office of Naval Research, ONR-CR213-124-2.

Let

$$\underline{y}_i = M \underline{x}_i \tag{3-6}$$

represent the blur operation. $\underline{x}_i$ is a Haar transform of an image before blurring, M is the Haar domain operator representing the blur, and $\underline{y}_i$ is the resultant Haar transform of the blurred image. We wish to find a positive-valued diagonal operator, D, so that if

$$\hat{\underline{y}}_i = D \underline{x}_i \tag{3-7}$$

the average magnitude of the elements of $\hat{\underline{y}}_i$ will be equivalent to those of $\underline{y}_i$; i.e.,

$$E\{|\hat{\underline{y}}_i|\} = E\{|\underline{y}_i|\} \tag{3-8}$$

where E is the expectation operator over all i and $|\underline{y}_i|$ is a vector whose elements are the magnitudes of $\underline{y}_i$. Using the requirement of equation (3-8), D may then be derived from equation (3-7).

$$|\hat{\underline{y}}_i| = D|\underline{x}_i| \tag{3-9}$$

Taking the expected value,

$$E\{|\hat{\underline{y}}_i|\} = D\, E\{|\underline{x}_i|\} \tag{3-10}$$

From equation (3-8),

$$E\{|\underline{y}_i|\} = D\, E\{|\underline{x}_i|\} \tag{3-11}$$

Equation (3-11) is saying that the diagonal operator, D, simply multiplies the j^{th} element of the average magnitude of the original transform by a scalor, d_{jj}, to arrive at a value for the j^{th} element of the average magnitude of the blurred transform. The values of the elements of D along its diagonal, d_{jj}, are thus given by

$$d_{jj} = \frac{E\{|\underline{y}_i|\}_j}{E\{|\underline{x}_i|\}_j} \quad \text{or} \quad d_{jj} = \frac{E\{|M\underline{x}_i|\}_j}{E\{|\underline{x}_i|\}_j} \tag{3-12}$$

Preliminary results of a series of experiments are presented in which the diagonal representations of the Haar domain operators presented earlier were computed.

The one-dimensional images used as input for these experiments were generated by a first-order Markov process ($\rho = 0.03$). Two thousand such vectors were used to evaluate the expected values of Haar spectra required in equation (3-12).

3.3.1 Focus Error

The values of d_{jj} observed for Case 1 and Case 2 focus error are shown in Figure 3-8. As shown in the figure, the attenuation introduced by d_{jj} is greater for the 5-pel-wide focus error than for the 3-pel-wide focus error. The Haar sequency

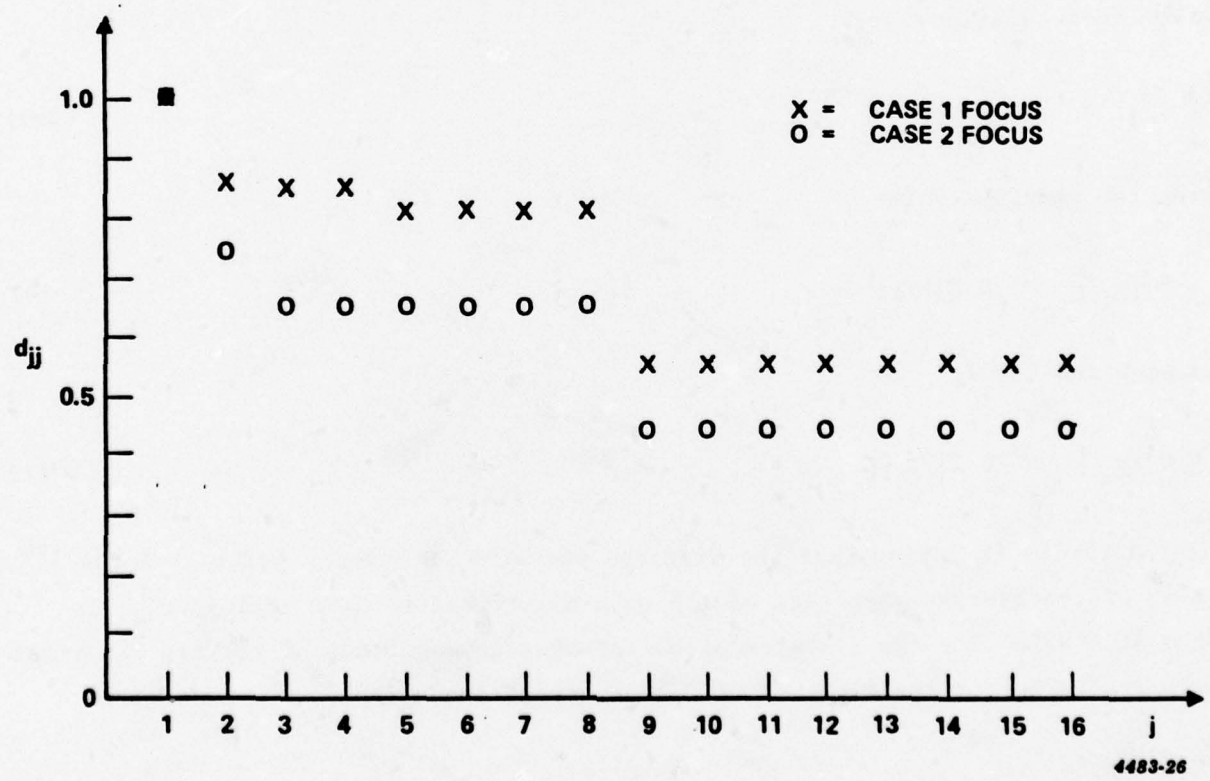


Figure 3-8. SEQUENCY ATTENUATION AS A RESULT OF FOCUS ERROR

spectrum is thus altered by a focus error in a manner similar to attenuation that would be observed in the frequency spectrum.

3.3.2 Limited Aperture

The diagonal operator elements computed for Case 1 and Case 2 aperture limiting are shown in Figure 3-9, which shows that this result of reducing the extent of the aperture by 25 percent is to attenuate the sequency spectrum by only a small amount. No "cutoff sequency" analogous to a cutoff frequency is observed.

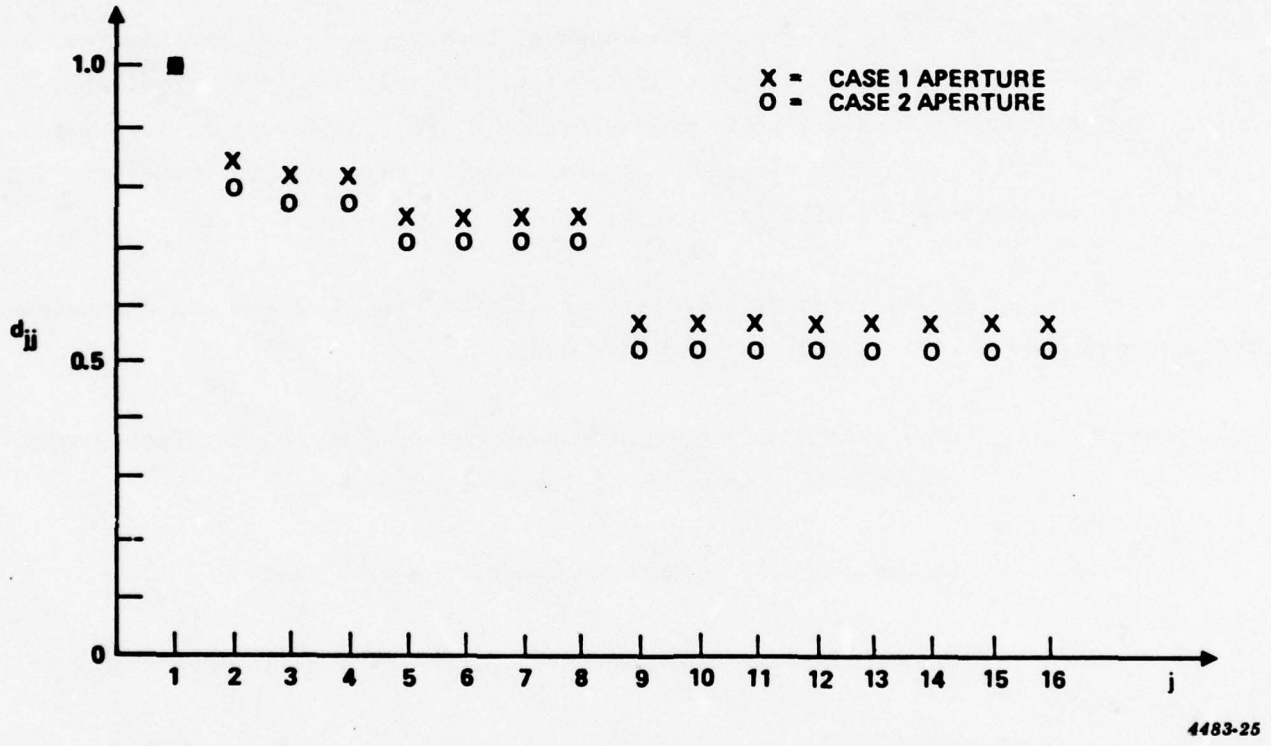


Figure 3-9. SEQUENCY ATTENUATION AS A RESULT OF LIMITED APERTURE

3.3.3 Image Motion

For both cases of image motion, the elements along the diagonal of D are all equal to unity. Although translating the image introduces no sequency attenuation, this result is still consistent with the blur model. Image motion is seen as a blur, not within a specific frame of a display, but rather is observed as a blur temporally from frame to frame. The intraframe effects of motion then are not

expected to alter the average magnitude of a single image's Haar transform. As a consequence, image motion will be considered apart from the other blur types.

3.4 APPROXIMATE DIAGONAL DEBLUR

A diagonal representation, D, of the blur operations has been derived from the average Haar spectra of both the original and blurred image. A diagonal representation of a corresponding deblurring operator is now required. The inverse of D is the most likely candidate to perform the requisite deblurring.

The computation of D and D^{-1} required knowledge of both the original and blurred image. However, in an automated deblurring system, the original is not available. Before some of the real-world aspects of evaluating D are considered, it is first necessary to examine the performance D^{-1} in deblurring. This will show whether the pursuit of exact values of D is worthwhile.

The performance of D^{-1} as a deblurring operator will be tested on the one dimensional Markov model used before. Recalling the blur model,

$\underline{y} = M\underline{x}$ Denotes the mapping of the Haar spectrum of the original image into the Haar spectrum of the blurred image,

$\hat{\underline{y}} = D\underline{x}$ is the diagonal representation of the blur, and

$\hat{\underline{x}} = D^{-1}\underline{y}$ is the diagonal representation of the deblurring operation.

Since we are dealing with one dimensional data, the Mean Squared Error (MSE) between two Haar vectors is used to characterize the effects of the various blurring and deblurring operations.

Let

$$\alpha = \frac{100}{16} \sum_{j} (x_j - y_j)^2$$

be the MSE due to blurring and

$$\beta = \frac{100}{16} \sum_{j} (x_j - \hat{x}_j)^2$$

be the residual MSE observed after deblurring.

The results of deblurring with D^{-1} are shown in Table 3-2.

Table 3-2. MSE OBSERVED USING D^{-1} TO DEBLUR

Blur Type	MSE Due to Blurring	MSE After Deblurring
Focus (Case 1)	0.35	0.38
Focus (Case 2)	0.62	0.72
Aperture (Case 1)	0.20	0.15
Aperture (Case 2)	0.29	0.24

As shown in the table, the residual MSE is comparable to that due to blurring. A slight improvement is seen in the aperture blur examples while the focus blurs show a small degradation. The relative performance of D^{-1} in undoing these two types of blurs is consistent with the observed characteristics of the exact Haar domain implementations computed previously. For example, from the comparison of Figures 3-3 and 3-4, one would expect the operator representing the aperture blur to be more closely approximated by a diagonal operator than the operator representing focus blur. The results shown in Table 3-2 bear this out.

Although little to no improvement in an MSE sense is realized by using D^{-1} to deblur one dimensional images, a subjective enhancement should be obtained. This prediction is based on edge enhancement experiments (Section 4) showing that a marked crispening effect is observed when the high sequencies of an image's Haar spectrum are amplified. The use of D^{-1} to deblur will therefore be investigated using some two dimensional images. The resultant improvements, if any, may then be evaluated with the MSE as well as subjectively and using image descriptors.

The problem of evaluating D and its inverse still remains. From equation (3-12) we see that the average Haar spectrum of the original must be known in order to assess the average sequency attenuation resulting from the blur. Since we are given only the blurred image, a prototype original must be assumed. By then calculating the average magnitude of the Haar spectrum of blurred image and comparing it to that of the prototype an estimate of D may be made. This method of substituting a prototype spectrum for that of the original is analogous to power spectrum equalization method of blind deconvolution. This approach has been used with good results by Stockham et al [2].

The definition of the prototype Haar spectrum will establish the amount of sequency attenuation observed in any one image. In the one-dimensional examples presented thus far, the average magnitude of the input Haar spectrum follows directly from the Markov model utilized to generate the image vectors. When dealing with real images, however, the uniformity and stability of the average magnitude of the Haar spectrum must still be investigated before a prototype substitution may be safely made.

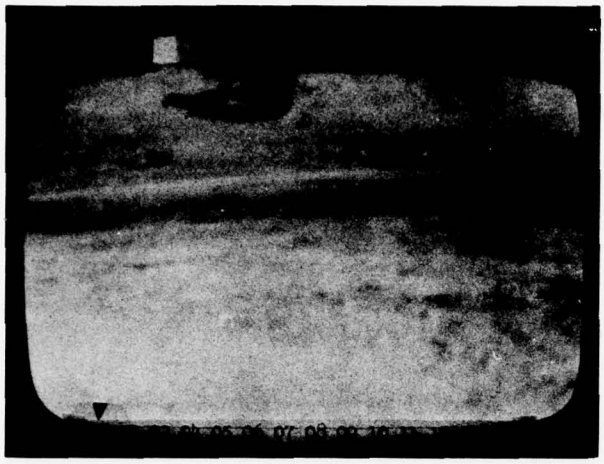
For the spectrum prototype to be meaningful, the average magnitude of the Haar coefficients must not vary substantially from image to image. Figure 3-10 shows six FLIR images displaying distinctly different scenes. The corresponding average magnitude Haar spectra for scan lines of the image are shown in Figure 3-11. Each of the Haar spectra shown exhibit the same characteristic shape; i.e., relatively high amplitude, low sequency terms with a gradual tapering off with increasing sequency. By forming the average of each of the curves in Figure 3-11 and averaging across sequency groupings, the Haar spectrum prototype can be formed. The prototype is shown in Figure 3-12. Comparison of the individual image curves with the prototype shows that the images vary by about ± 20 percent from the prototype. In light of the marginal change in MSE observed by deblurring with D^{-1}, 20 percent tolerance in prototype definition should be acceptable.

3.5 MOTION DETECTION/CORRECTION

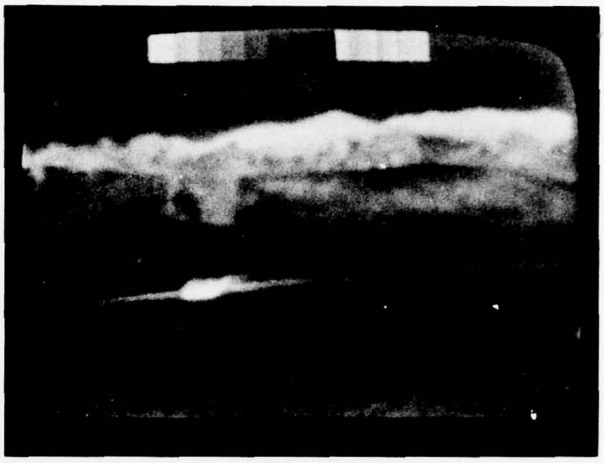
Motion blur, as considered here, is a uniform translation or "jump" of the image from one image frame to the next. While this effect does not blur any individual image frame, displaying translated frames at 30 hertz produces a subjective smearing or jitter effect.

---

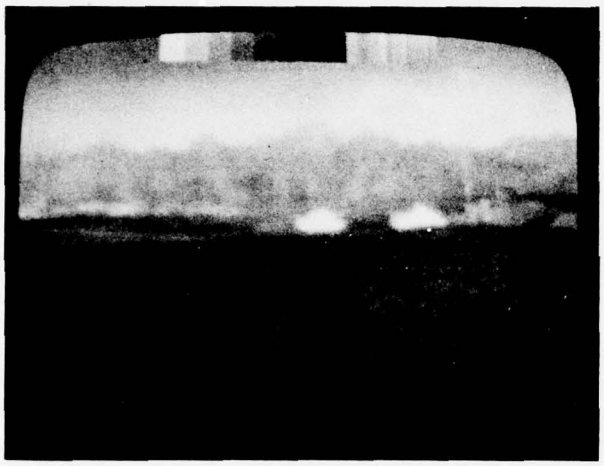
[2] T. G. Stockham, T. M. Cannon, R. B. Ingebretsen "Blind Deconvolution Through Digital Signal Processing," Proc. IEEE, 4-75, V63, No.4.



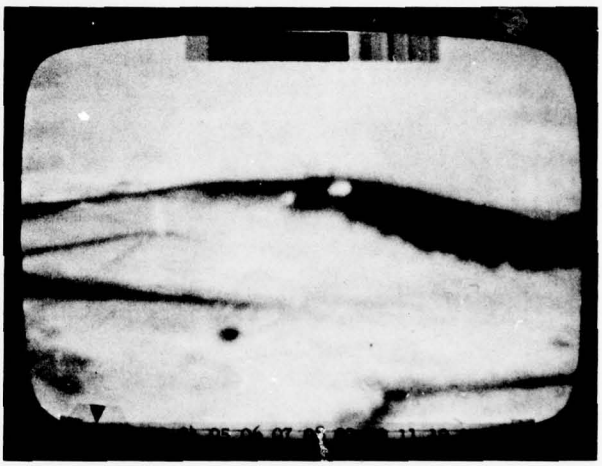
(a) IMAGE 1



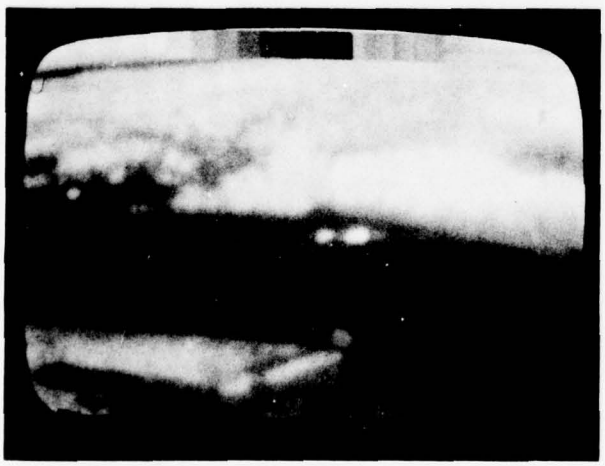
(b) IMAGE 5



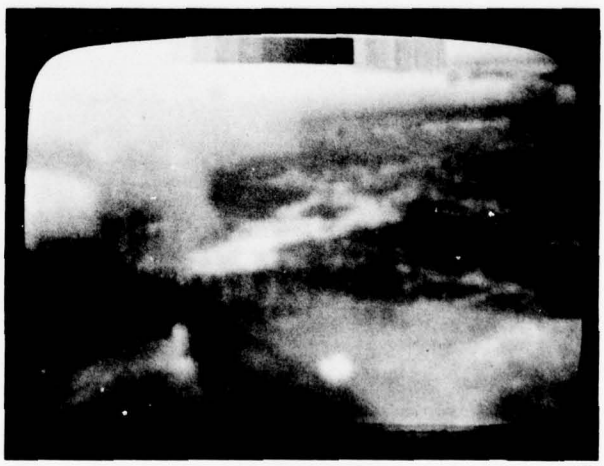
(c) IMAGE 6



(d) IMAGE 7



(e) IMAGE 8



(f) IMAGE 10

4483

Figure 3-10. FLIR THERMOGRAMS

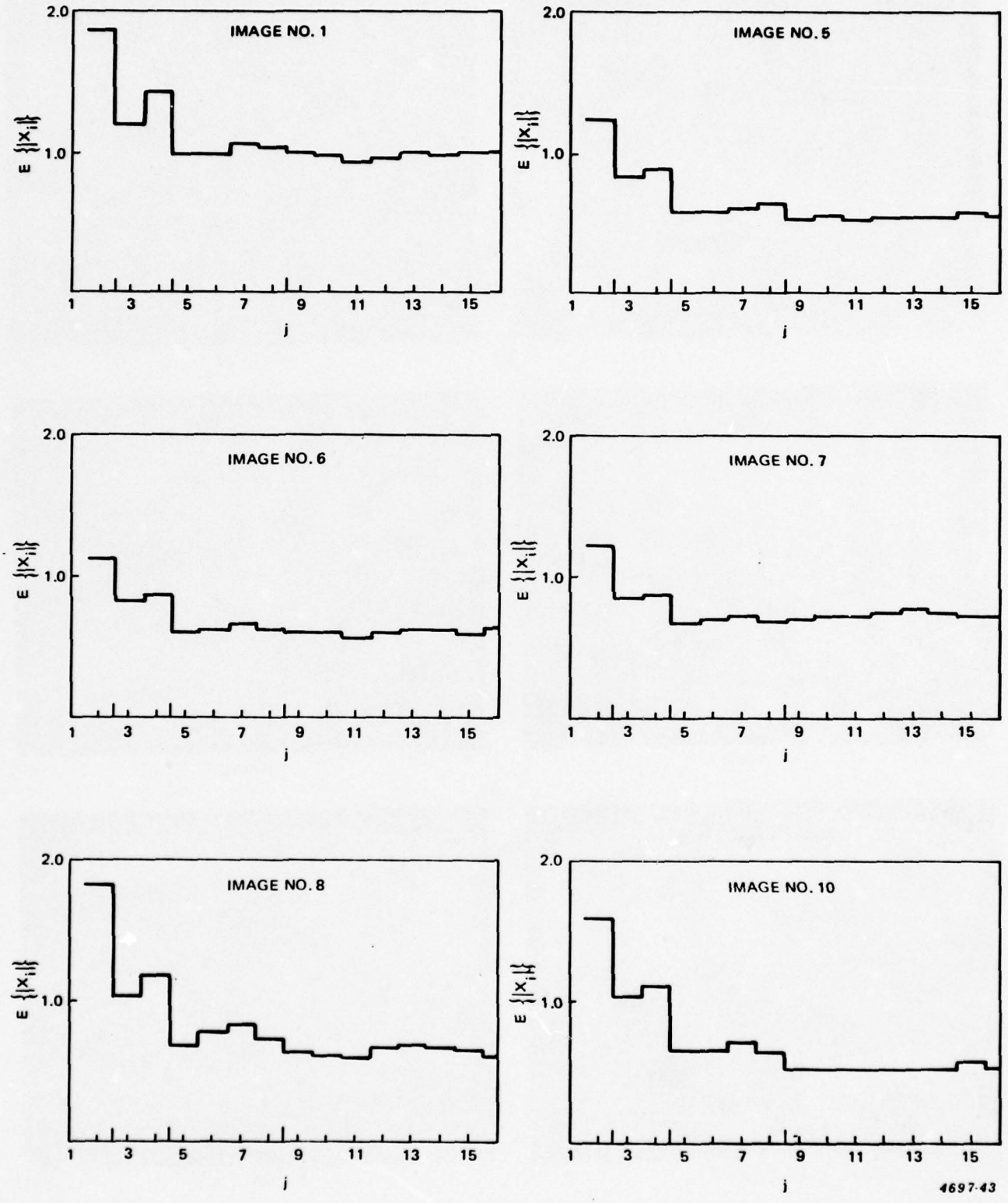


Figure 3-11. AVERAGE MAGNITUDE HAAR SPECTRA

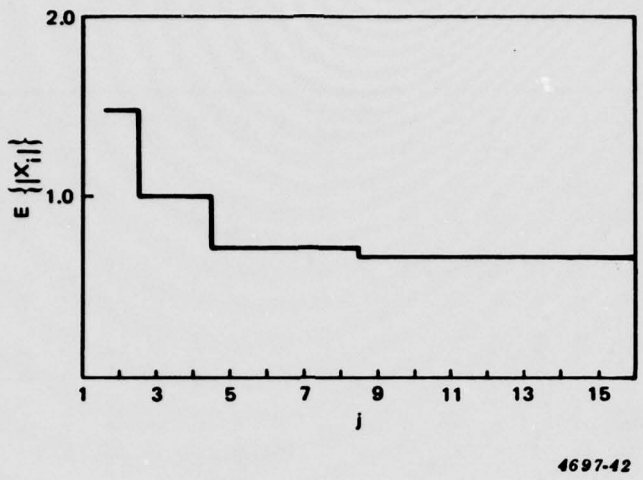


Figure 3-12. HAAR SPECTRUM PROTOTYPE

Since no one frame is blurred, image motion must be detected by comparing the Haar spectra of successive image frames. In one dimension, two images are related by the Haar domain operator for motion. The two forms of this operator resulting from left and right image translations are shown in Figures 3-13 and 3-14. Examination of Figures 3-13 and 3-14 reveals the following characteristics:

a. Four of the bottom eight rows of each operator have only three nonzero elements (i.e., they exhibit a low complexity).

b. The values of these nonzero elements are 1 and -0.5 and can thus be easily implemented in digital hardware.

c. The rows with these properties are the odd-numbered rows for a left shift and even numbered rows for a right shift. The effects of shifting either left or right are therefore separable.

Row No.																
1	1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.75	-0.13	0.13	-0.13	0.0	0.13	0.0	-0.13	0.0	0.0	0.0	0.13	0.0	0.0	0.0
3	0.0	0.25	0.63	-0.13	-0.13	0.25	-0.13	0.0	-0.13	0.0	0.25	0.0	-0.13	0.0	0.0	0.0
4	0.0	-0.25	-0.13	0.63	-0.13	0.0	-0.13	0.25	-0.13	0.0	0.0	0.0	-0.13	0.0	0.25	0.0
5	0.0	0.0	0.50	0.0	0.25	-0.25	0.0	0.0	-0.25	0.50	-0.25	0.0	0.0	0.0	0.0	0.0
6	0.0	0.50	-0.25	-0.25	0.0	0.25	-0.25	0.0	0.0	0.0	-0.25	0.50	-0.25	0.0	0.0	0.0
7	0.0	0.0	0.0	0.50	0.0	0.0	0.25	-0.25	0.0	0.0	0.0	0.0	-0.25	0.50	-0.25	0.0
8	0.0	-0.50	-0.25	-0.25	-0.25	0.0	0.0	0.25	-0.25	0.0	0.0	0.0	0.0	0.0	-0.25	0.50
9	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	0.0	1.00	0.0	-0.50	-0.50	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0
12	0.0	1.00	-0.50	-0.50	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0
14	0.0	0.0	0.0	1.00	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50
16	0.0	-1.00	-0.50	-0.50	-0.50	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	-0.50

4697-32

Figure 3-13. HAAR DOMAIN OPERATOR FOR A ONE-PEL LEFT SHIFT

Row No.																
1	1.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.75	0.13	-0.13	0.0	0.13	0.0	-0.13	0.0	0.0	0.0	0.13	0.0	0.0	0.0	-0.13
3	0.0	-0.25	0.63	-0.13	0.25	-0.13	0.0	-0.13	0.0	0.25	0.0	-0.13	0.0	0.0	0.0	-0.13
4	0.0	0.25	-0.13	0.63	0.0	-0.13	0.25	-0.13	0.0	0.0	0.0	-0.13	0.0	0.25	0.0	-0.13
5	0.0	-0.50	-0.25	-0.25	0.25	0.0	0.0	-0.25	0.50	-0.25	0.0	0.0	0.0	0.0	0.0	-0.25
6	0.0	0.0	0.50	0.0	-0.25	0.25	0.0	0.0	0.0	-0.25	0.50	-0.25	0.0	0.0	0.0	0.0
7	0.0	0.50	-0.25	-0.25	0.0	-0.25	0.25	0.0	0.0	0.0	0.0	-0.25	0.50	-0.25	0.0	0.0
8	0.0	0.0	0.0	0.50	0.0	0.0	-0.25	0.25	0.0	0.0	0.0	0.0	0.0	-0.25	0.50	-0.25
9	0.0	-1.00	-0.50	-0.50	-0.50	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0	-0.50
10	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	1.00	0.0	-0.50	-0.50	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0
13	0.0	1.00	-0.50	-0.50	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0	0.0
15	0.0	0.0	0.0	1.00	0.0	0.0	-0.50	-0.50	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50	0.0
16	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.00	0.0	0.0	0.0	0.0	0.0	0.0	-0.50	-0.50

4697-33

Figure 3-14. HAAR DOMAIN OPERATOR FOR A ONE-PEL RIGHT SHIFT

d. Each even-numbered row for the right shift is identical to a neighbor odd-numbered row for the left shift (e.g., Row 9 in Figure 3-14 is identical to Row 10 in Figure 3-13). This results in a beneficial redundancy.

These properties lead to the motion detection method shown in Figure 3-15.

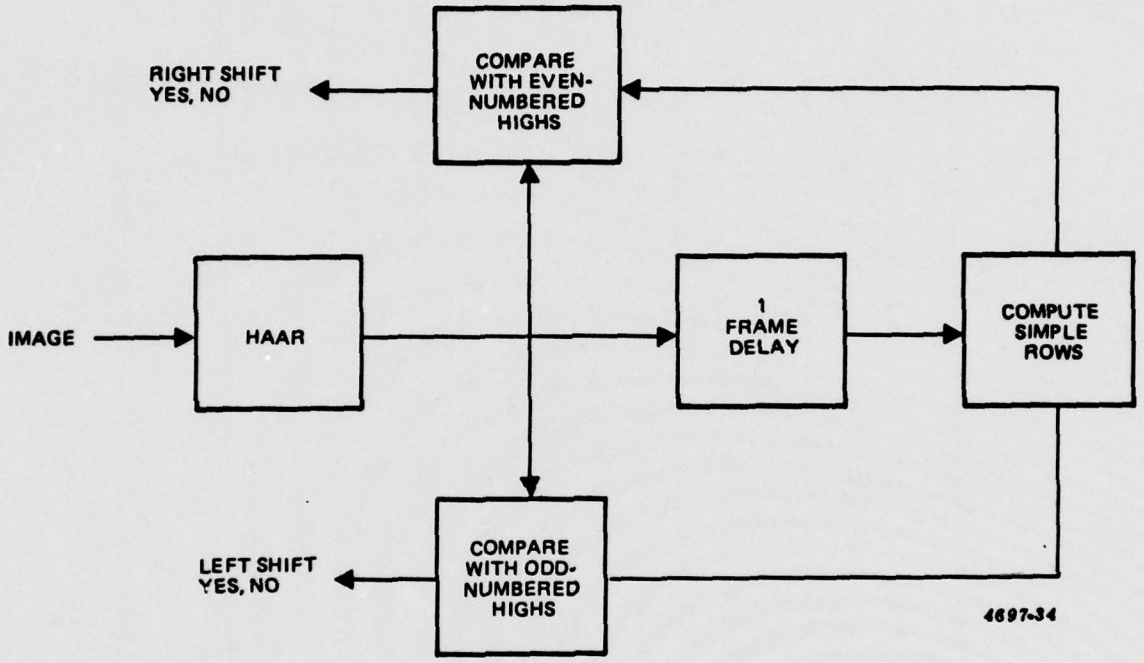


Figure 3-15. MOTION DETECTION METHOD

The box "Compute Simple Rows" performs the following function:

$$T_{[m/2]} = a_{r-1,\,[m/2]} - \frac{1}{2}\left(a_{r,m} + a_{r,m-1}\right) \qquad (3\text{-}13)$$

where r and m are the sequency and offset of the Haar function and [m/2] means "the integer part of m/2."

The comparators then compare values of $T_{[m/2]}$ with the corresponding odd and even $a_{r,m}$. If they are equal (within a certain tolerance), a shift in the appropriate direction is indicated.

This method of motion detection involves 16 adds and 8 comparisons to detect a left or right shift. An equivalent method in the image domain would require 32 comparisons. A potential computational savings could thus be realized by using this approach.

Once an image shift has been detected, the motion may be easily reversed by an appropriate inverse shift in the image domain.

3.6 TWO DIMENSIONAL BLUR MODEL

The results presented so far have been one-dimensional. Real life imagery is, of course, two-dimensional, thus requiring a two-dimensional formulation of the blurring and deblurring process. Let X be a 16 x 16 sampled original and Y its blurred image. The mapping of X into Y can be expressed as a two-dimensional convolution

$$\{Y\}_{ij} = \sum_k \sum_l b_{k,l} \{X\}_{k-i,\ l-j} \tag{3-14}$$

where b_{kl} is the impulse response of the blur. If the blur is separable (i.e. can be expressed as two one-dimensional blurs) then a more convenient matrix expression may be used, namely

$$Y = B\ X\ B^T \tag{3-15}$$

where B is the matrix representation of a one-dimensional blur. We will assume for the moment that the blur is indeed separable.

By taking the two dimensional Haar transform of (3-15) we get

$$HYH^T = H\ B\ X\ B^T\ H^T$$

or, since $H^{-1} = H^T$

$$H\,Y\,H^T = \left[H\,B\,H^T\right]\left[H\,X\,H^T\right]\left[H\,B^T\,H^T\right]$$
$$= \left[H\,B\,H^T\right]\left[H\,X\,H^T\right]\left[H\,B\,H^T\right]^T \qquad (3\text{-}16)$$

Recalling equation (3-4), equation (3-16) may be written as

$$H\,Y\,H^T = M\left[H\,X\,H^T\right]M^T \qquad (3\text{-}17)$$

where M is the Haar domain operator for the blur in question. The forms of M resulting from several typical blurs were presented in paragraph 3.2. Each operator is highly non-diagonal in nature and therefore complicates the blur detection and inversion process. However, the diagonal representation of M may be substituted, thereby simplifying the expression.

$$H\,\hat{Y}\,H^T = D\left[H\,X\,H^T\right]D \qquad (3\text{-}18)$$

To implement equation (3-18), a new matrix D^* can be formed where

$$\left\{D^*\right\}_{ij} = \left\{D\right\}_{ii}\left\{D\right\}_{jj} \qquad (3\text{-}19)$$

and the estimate of the blurred image is given by

$$H\,\hat{Y}\,H^T = D^* \cdot \left[H\,X\,H^T\right] \qquad (3\text{-}20)$$

where "." implies element-by-element multiplication, the matrix D^* relates the average sequency attenuation observed due to the blur. In cases where the blur function is not separable into two one-dimensional blurs, the elements of D^* may be derived directly from the two-dimensional Haar spectra of the original and blurred images.

$$D^*_{ij} = \frac{\left\{\overline{H\,Y\,H^T}\right\}_{ij}}{\left\{\overline{H\,X\,H^T}\right\}_{ij}} \qquad (3\text{-}21)$$

where $\overline{H\,Y\,H^T}$ denotes the average of the magnitude of the blurred images Haar spectrum. The assumption of separability is therefore not required to evaluate D^*.

The reciprocal of the elements of D^* can now be used to approximately deblur two-dimensional images in the same manner that D^{-1} was used on one-dimensional images. The average magnitude of the input Haar spectrum is assumed to be known a priori. The Haar spectrum prototypes computed in the same fashion as in paragraph 3.5 may also be substituted for the input spectrum.

This deblurring process is given by

$$\left[H\ X\ H^T\right] = D^{*-1} \cdot \left[H\ Y\ H^T\right] \tag{3-22}$$

where

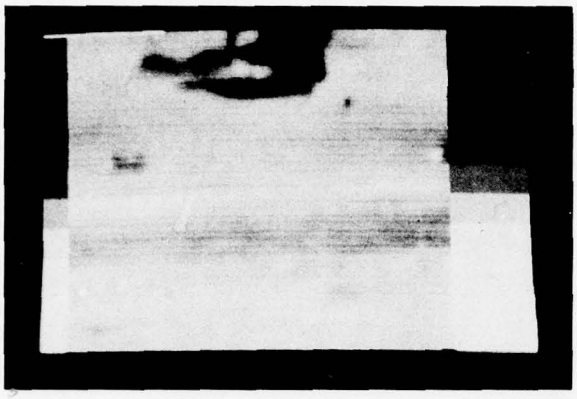
$$\left\{D^{*-1}\right\}_{ij} = \frac{1}{\left\{D^*\right\}_{ij}}$$

3.7 TEST IMAGES

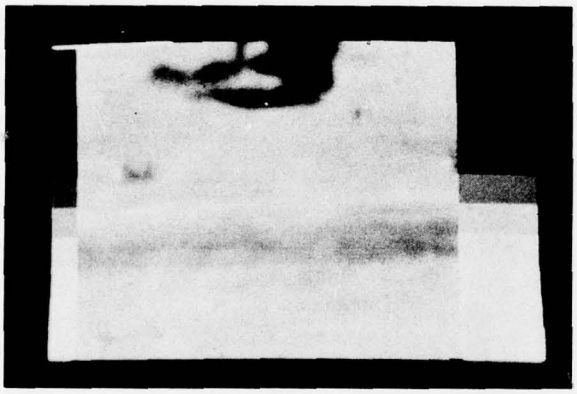
A 256 x 256 pel region of each of the six FLIR images presented in Figure 3-10 was blurred by the blur functions presented in paragraph 3.2. The resultant blurred images were then subjected to the deblurring process given by equation (3-22). The original, blurred, and deblurred images are presented in Figures 3-16 through 3-21 for the case 2 focus blur. This blur was chosen for pictorial demonstration since it is the most severe in an MSE sense. A summary of the changes in MSE and image descriptors for all the blurs for each image is presented in Table 3-3.

As shown in Table 3-3, deblurring an image using this approach results in an MSE degradation of at least three times that observed from the blur itself. This changes in image descriptors, however, indicate a substantial increase in edge activity as a result of deblurring. These edges are evidenced in the deblurred images (Figures 16-21) not as sharper edge detail, but rather as ringing and noise across the entire image.

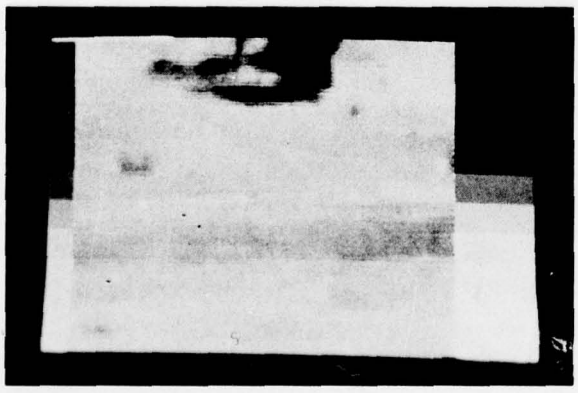
This subjective enhancement that was predicted as a result of deblurring was not observed. This is due primarily to the linear nature of the deblurring process.



A. IMAGE NO. 4 ORIGINAL 4697-44

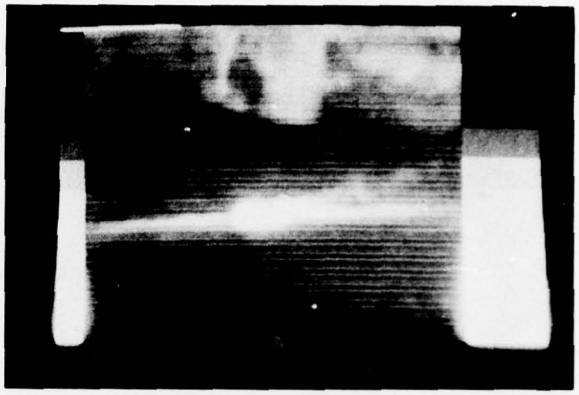


B. BLURRED FOCUS, CASE 2 4697-45

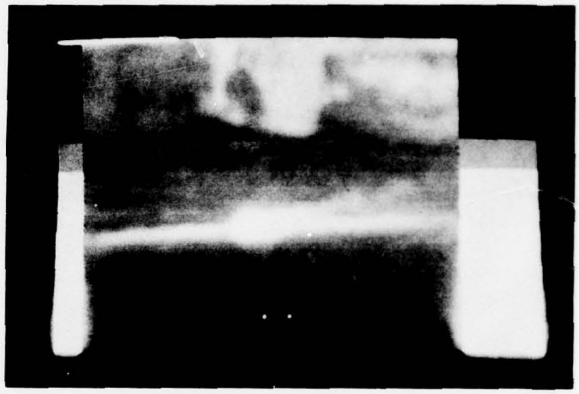


C. DEBLURRED 4697-46

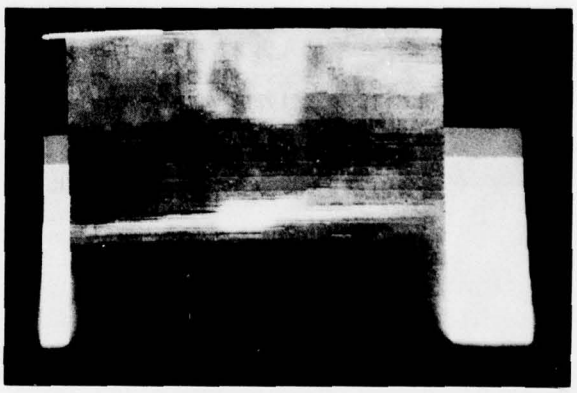
Figure 3-16. DEBLUR RESULTS FOR IMAGE #1.



A. IMAGE NO. 5 ORIGINAL 4697-47

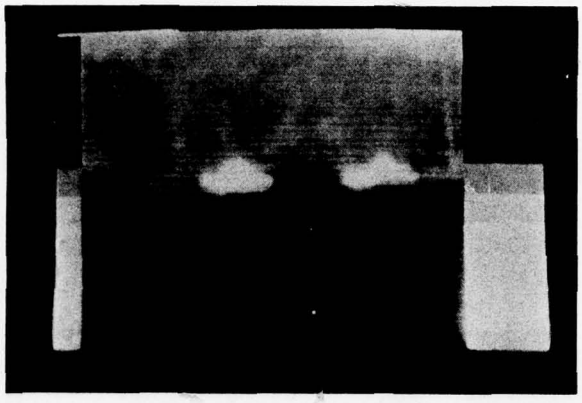


B. BLURRED FOCUS, CASE 2 4697-48

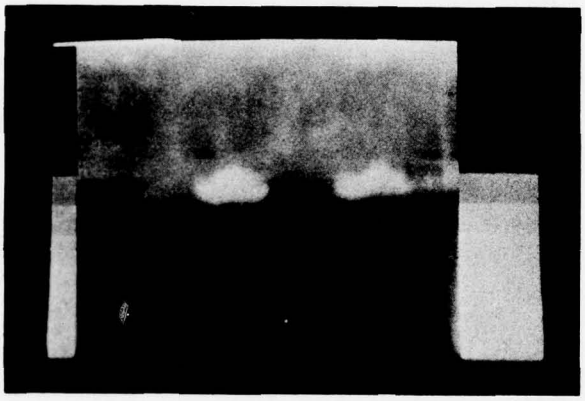


C. DEBLURRED 4697-49

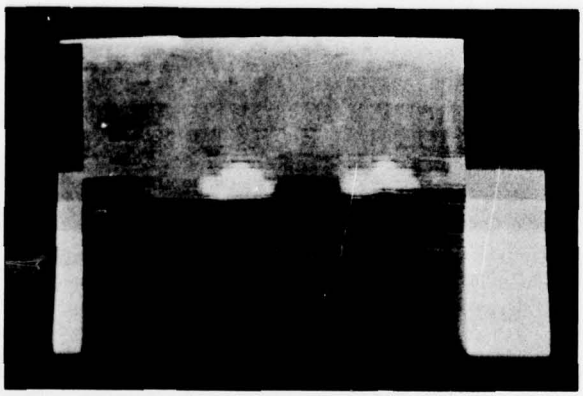
Figure 3-17. DEBLUR RESULTS FOR IMAGE #5.



A. IMAGE NO. 6 BLURRED 4697-50

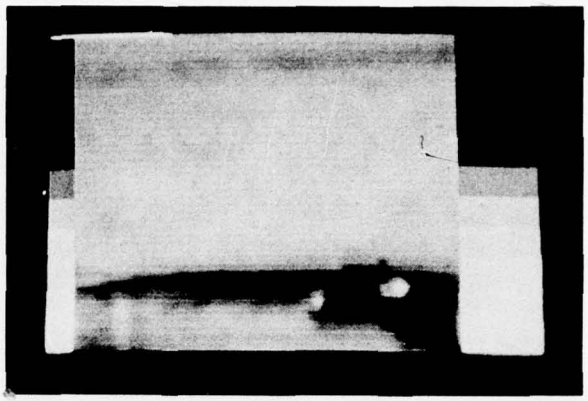


B. BLURRED FOCUS, CASE 2 4697-51

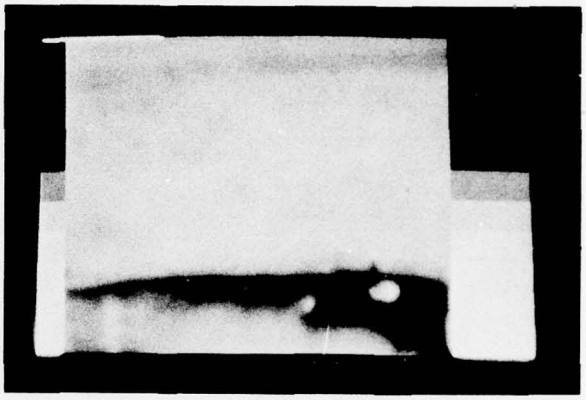


C. DEBLURRED 4697-52

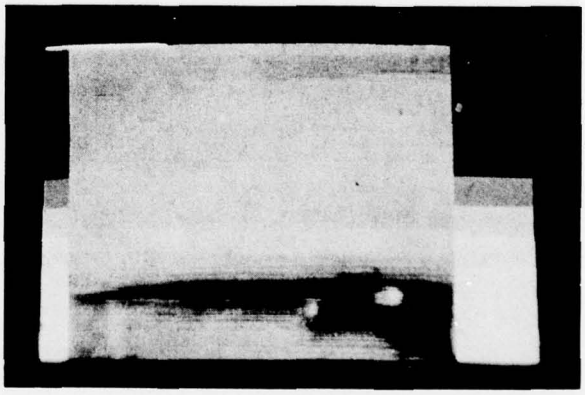
Figure 3-18. DEBLUR RESULTS FOR IMAGE #6.



A. IMAGE NO. 7 ORIGINAL 4697-53

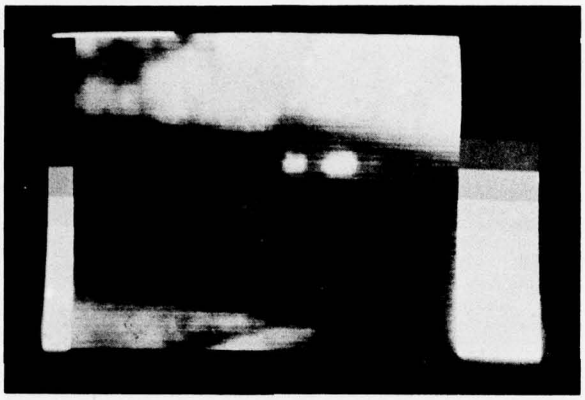


B. BLURRED FOCUS, CASE 2 4697-54

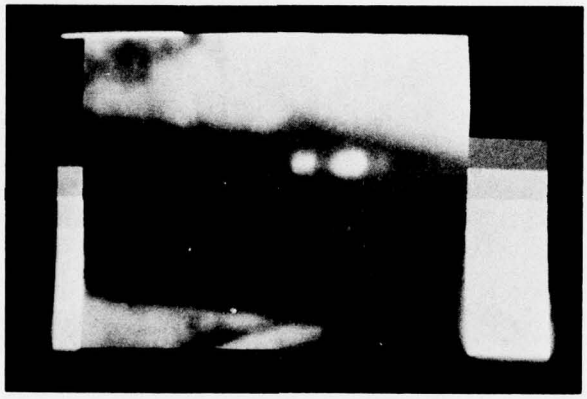


C. DEBLURRED 4697-55

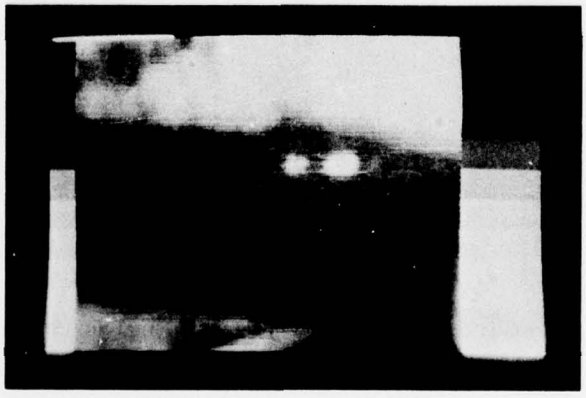
Figure 3-19. DEBLUR RESULTS FOR IMAGE #7.



A. IMAGE NO. 8 ORIGINAL 4697-56

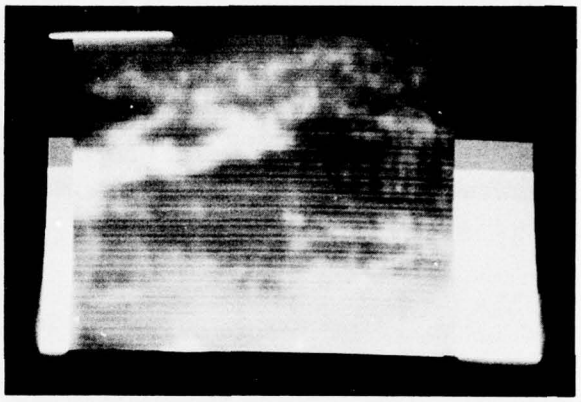


B. BLURRED FOCUS, CASE 2 4697-57

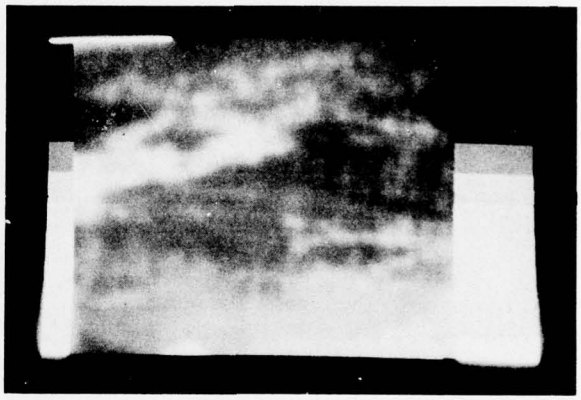


C. DEBLURRED 4697-58

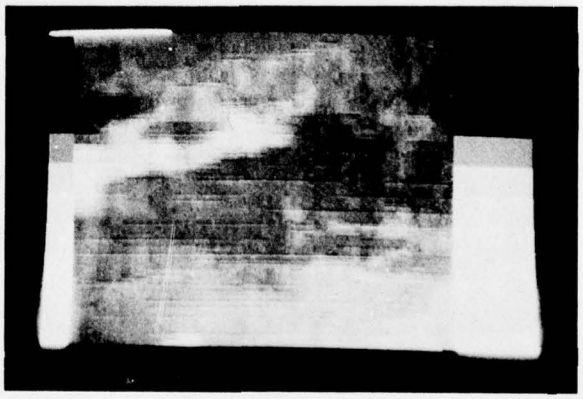
Figure 3-20. DEBLUR RESULTS FOR IMAGE #8



A. IMAGE NO. 10 ORIGINAL 4697-59



B. BLURRED FOCUS, CASE 2 4697-60



C. DEBLURRED 4697-61

Figure 3-21. DEBLUR RESULTS FOR IMAGE #10.

Table 3-3. SUMMARY OF DEBLUR RESULTS*

Image #1 Edge Activity, e = 143 Texture Activity, x = 84

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	2.1	9.1	116	339	16	60
Focus 2	5.5	17.2	91	574	10	65
Aperture 1	0.6	6.0	133	234	31	77
Aperture 2	1.8	7.7	114	339	16	60

Image #5 Edge Activity, e = 146 Texture Activity, x = 37

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	1.4	6.1	104	234	16	40
Focus 2	4.5	15.2	84	413	9	72
Aperture 1	0.5	4.8	123	193	22	40
Aperture 2	1.3	5.6	105	240	16	40

Image #6 Edge Activity, e = 130 Texture Activity, x = 38

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	1.4	6.6	86	235	16	43
Focus 2	4.4	17.5	68	456	8	85
Aperture 1	0.5	5.0	107	188	22	44
Aperture 2	1.3	5.9	89	237	15	40

Image #7 Edge Activity, 3 = 113 Texture Activity, x = 51

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	1.7	7.6	78	261	11	42
Focus 2	4.6	14.3	57	432	6	56
Aperture 1	0.5	5.5	93	183	22	55
Aperture 2	1.5	6.7	76	259	11	42

*Subscript α refers to blurred image, β to "deblurred" image

Table 3-3. SUMMARY OF DEBLUR RESULTS* (Continued)

Image #8 Edge Activity, e = 143 Texture Activity, x = 51

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	1.4	6.5	111	236	19	42
Focus 2	4.3	13.5	93	363	13	82
Aperture 1	0.5	5.1	126	204	25	46
Aperture 2	1.3	5.9	110	242	19	45

Image #10 Edge Activity, e = 109 Texture Activity, x = 25

	MSE (%)		Edge Activity		Texture Activity	
Blur	α	β	e_α	e_β	x_α	x_β
Focus 1	1.2	6.2	74	180	13	31
Focus 2	3.5	12.1	62	271	7	49
Aperture 1	0.4	5.2	88	151	18	33
Aperture 2	1.0	5.9	76	176	12	33

*Subscript α refers to blurred image, β to "deblurred" image

When deblurring, all sequency terms regardless of amplitude are amplified linearly by the appropriate element of the deblur matrix. The results of the edge enhancement experiments show that subjective enhancement is best obtained by a nonlinear increase of only some above-threshold sequency terms.

3.8 CONCLUSION

The results of deblurring with a diagonal operator in the Haar domain show a degradation in image quality (both subjective and MSE) instead of the desired improvement. This result is due primarily to the basic non-diagonal nature of blurs in the Haar domain and the rough approximation that a diagonal blur operator provides. The Haar domain is, in general, not well suited for the implementation of scalar filtering functions that correspond to spatial convolution.

The Haar transform can, however, provide a measure of the image loss due to blurring. By comparing the observed average magnitude spectrum with the Haar transforme prototype, some information about the magnitude, if not the nature, of the blur may be obtained. In an automated deblurring process, this information would be required before any further processing was initiated.

SECTION 4
EDGE ENHANCEMENT EXPERIMENTS

4.1 INTRODUCTION

A fundamental task of imaging sensor systems, whether thermal or visual, is to facilitate target detection and classification. The target, being a man-made object, usually consists of a greater number of straight lines (edges) than most natural objects. Therefore, one of the more promising techniques for extracting the target from natural background clutter is edge enhancement. Edge enhancement will also highlight roads, overhead wires, paths, and other such cues as to where a target may be found. The object of the experiment described in this section is to determine how the edge-correlating coefficients of the rationalized Haar transform can be effectively modified in real time so that these edges are highlighted and image details are made more visible.

During this period, a set of edge-enhancement algorithms were developed and evaluated both subjectively and in terms of image descriptors.

Edge enhancement algorithms should exhibit the following characteristics:

a. Low Complexity - The computational complexity of the enhancement algorithms must be amenable with real-time processing of video imagery.

b. Scene Independence - The selected edge-enhancement scheme must perform equally well for different types of scenes. Edge-enhancement algorithms should tolerate both high- and low-contrast situations Also, performance should not be significantly altered by image content; e.g., target size and background clutter.

c. Noise Immunity - System noise can be categorized into two classes: (1) low-frequency noise, where the grey shade levels of an image vary from one TV frame to another, and (2) random, high-frequency noise (transient or shot noise).

Provisions should be made in the edge-enhancement algorithm to minimize, if not eliminate, enhancement of these types of system noise.

d. Adaptability in the Operational System - The selected edge-enhancement scheme must be capable of being hardware implemented with sufficient resolution, dynamic range, stability, and speed to be compatible with the FLIR image quality parameters as specified in the classified appendix to the First Quarterly Report (76Y133).

4.2 GENERAL EXPERIMENTAL PROCEDURE

The experiment was divided into three major tasks:

a. Preliminary studies and experiments were conducted to establish the basic criteria for edge enhancement. The goal of this effort was to determine which of the 15 Haar sequency groups require alteration and the amount of alteration necessary in order to enhance edges.

b. The performance of several different methods of edge enhancement on a representative FLIR image were examined.

c. The performance of the edge enhancement methods selected in the second task were evaluated on a set of six FLIR thermograms supplied by NVL.

The experimental test configuration is shown in Figure 4-1. To ensure the validity of the experimental test results, a number of test setup procedures were initiated as follows:

a. To ensure that the test image was in focus, the lens of the video camera was manually adjusted using the autofocus concept described in Section 2.

b. The f-stop of the video camera lens was opened to its extreme (f/1.4). This minimizes degradation of the system Modulation Transfer Function (MTF) by the optics. Also, to prohibit the video signal from being overdriven into saturation, an appropriate neutral-density filter was inserted between the test image and the camera lens.

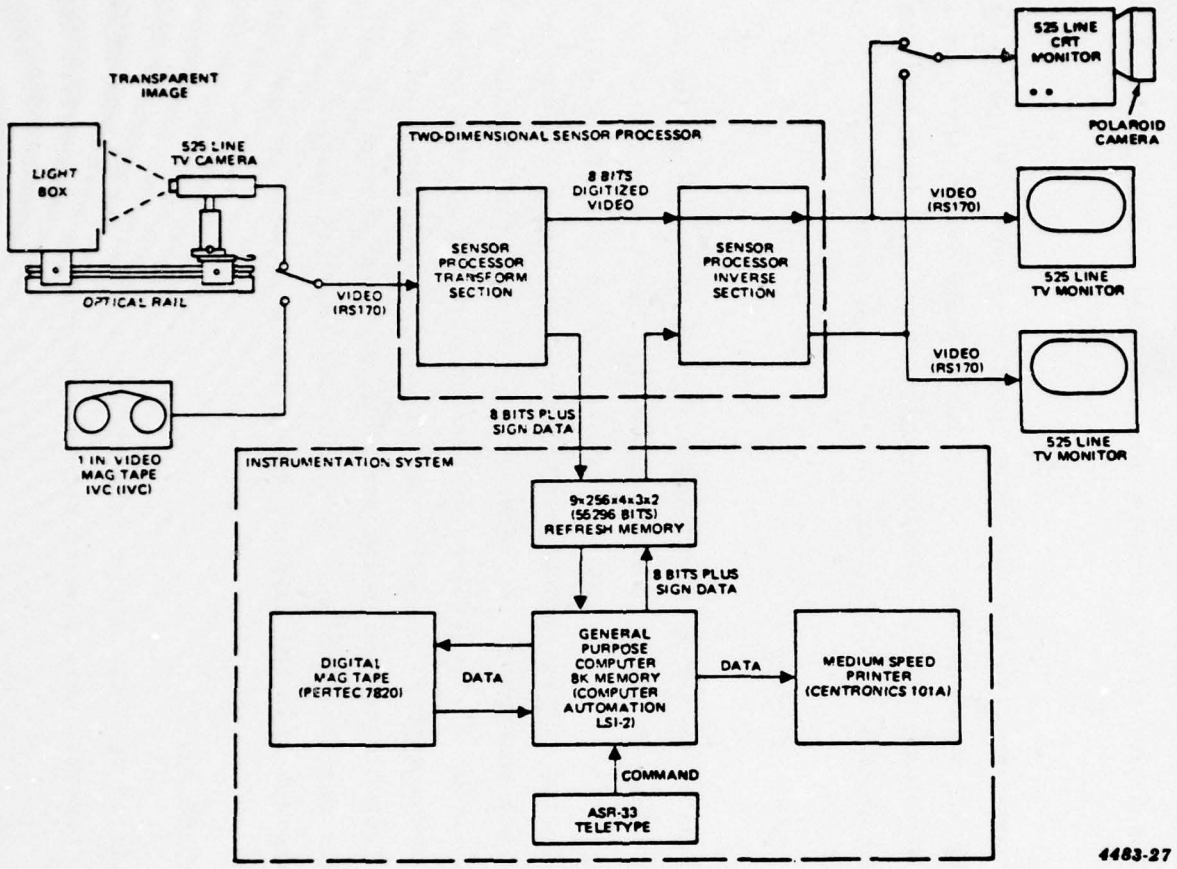


Figure 4-1. EXPERIMENTAL TEST CONFIGURATION

c. The offset bias and the video gain were adjusted to obtain the maximum attainable dynamic range for each FLIR test image.

The Image Descriptors (ID's) of each original and enhanced image used in this study were computed and recorded as a measure of enhancement. The ID's corresponding to each image of a series of experiments are presented at the end of the appropriate presentation.

The six FLIR thermograms used for the enhancement experiments supplied by Night Vision Laboratory are shown in Figure 4-2. Their image descriptors are presented in Table 4-1. This subset of the available data base was selected to demonstrate performance of edge-enhancement features over a range of scenes.

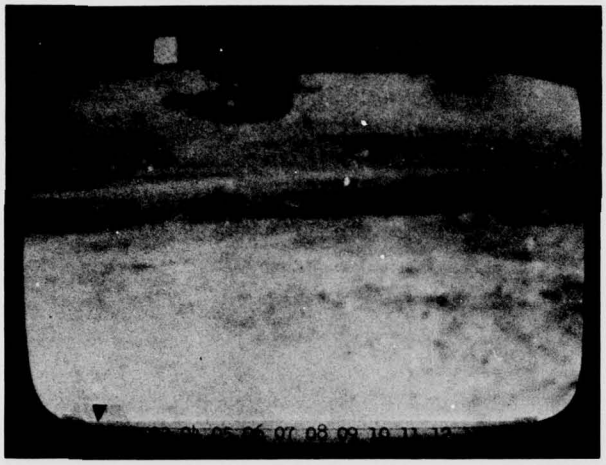
4.3 EDGE ENHANCEMENT CRITERIA

The edge terms of an image's Haar transform are those that satisfy the edge criterion function as defined in the section on image descriptors in the First Quarterly Report. The edge criterion function adaptively sets an edge threshold, θ, based on the power distribution within each sequency group. AC terms, the amplitudes of which are greater than the threshold, are identified as edges. Computation of θ, however, requires a large number of floating-point operations. The present hardware configurtion of the two-dimensional sensor processor coupled with the requirement that edge-enhancement algorithms operate in real time preclude the use of such an adaptive threshold for edge definition. In performing the FLIR image experiments described in paragraph 2.5, the adaptive thresholds were used in non-real time.

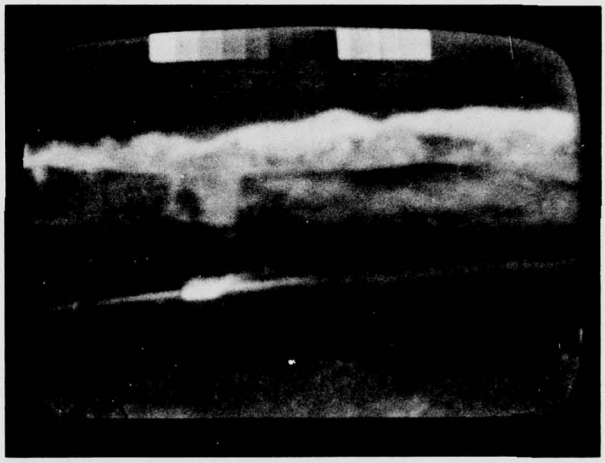
4.3.1 Definition of Edges

Edge terms identified by θ can appear in most of the sequency groups of an image's Haar transform. However, only some of sequency groups are appropriate for edge enhancement.

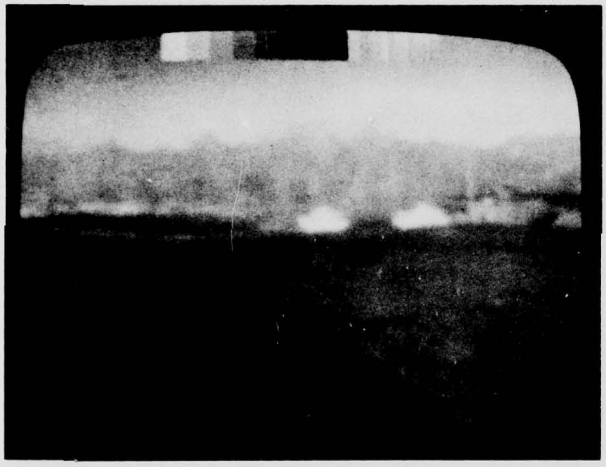
The Haar coefficient groups selected for edge enhancement must meet two basic criteria. First, the area on which the Haar function has an effect should not be any larger than the smallest object of interest in the scene. If this condition is not met, the object of interest will simply take on the shape of the Haar function when enhanced. The second criterion is that the area of the Haar functions selected



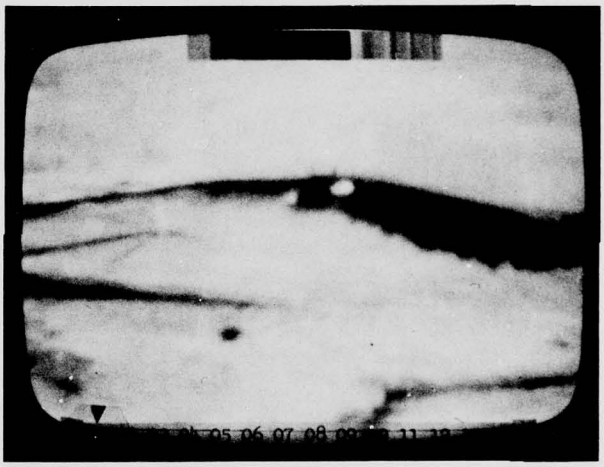
(a) IMAGE 1



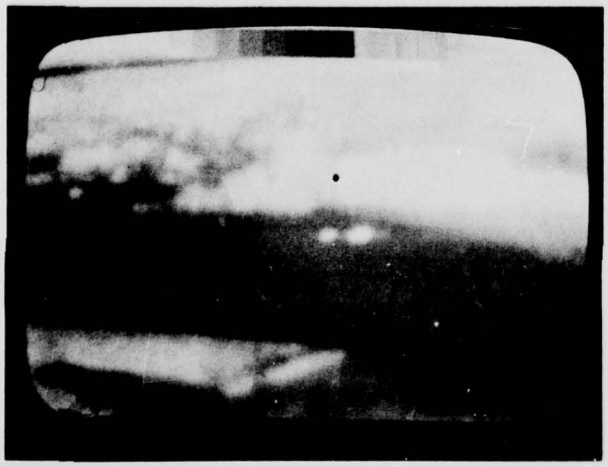
(b) IMAGE 5



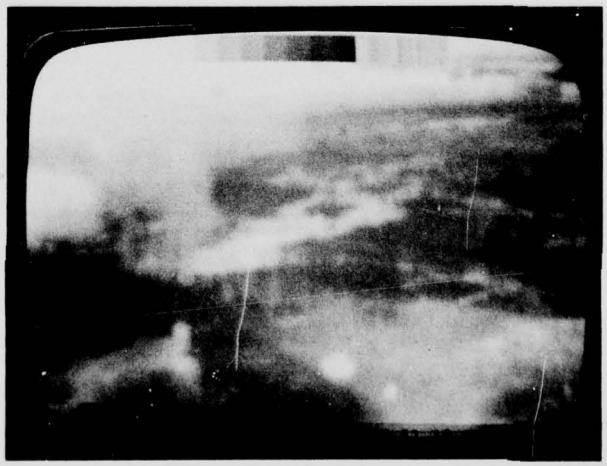
(c) IMAGE 6



(d) IMAGE 7



(e) IMAGE 8



(f) IMAGE 10

4483

Figure 4-2. FLIR THERMOGRAMS

Table 4-1. DESCRIPTORS FOR TEST IMAGES

Figure	Tone	Edge Activity	Texture Activity
2-2a	21	455	322
2-2b	13	630	301
2-2c	23	434	266
2-2d	26	658	453
2-2e	17	563	277
2-2f	19	516	263

should be large enough that random high-frequency noise in the system will not be enhanced, creating random speckles on the display. The two criteria thus place upper and lower sequency bounds on the range of Haar functions suited to edge enhancement.

Figure 4-3 illustrates the size and shape of each Haar function in group 4 (located at upper left corner) through group 15 (lower right corner). Subjective evaluation of the size and shape of the Haar functions with that of the scene in the background led to the selection of groups 9 through 15, which can be altered for edge enhancement.

Separation of the edge terms from non-edge terms in groups 9 through 15 necessitates the derivation of an edge threshold. The sensor processor has the capability of setting a different fixed-edge threshold for each sequency group, j, of an image's Haar transform. An immediate goal in design of an edge-enhancement algorithm is to determine a set of fixed-edge thresholds θ_j that can approximate the adaptive one.

As a first step in obtaining a fixed threshold, the values of the adaptive edge threshold derived by the image descriptor for each individual sector were averaged across each of the six test images. The resultant average thresholds for groups 9 through 15 are shown in Table 2-2. As shown in the table, the average value of the adaptive edge threshold, $\bar{\theta}$, varies substantially from group to group, but is relatively constant from image to image. This consistency in the behavior of $\bar{\theta}$ shows a beneficial scene independence in the definition of an edge. The values of $\bar{\theta}$ for some groups, however, are quite high. This is because of the dual function that θ performs. When θ is low, it serves as an edge-texture separator. When



Figure 4-3. HAAR FUNCTIONS, GROUPS 4 THROUGH 15

Table 4-2. AVERAGE ADAPTIVE THRESHOLD (IMAGE 1)

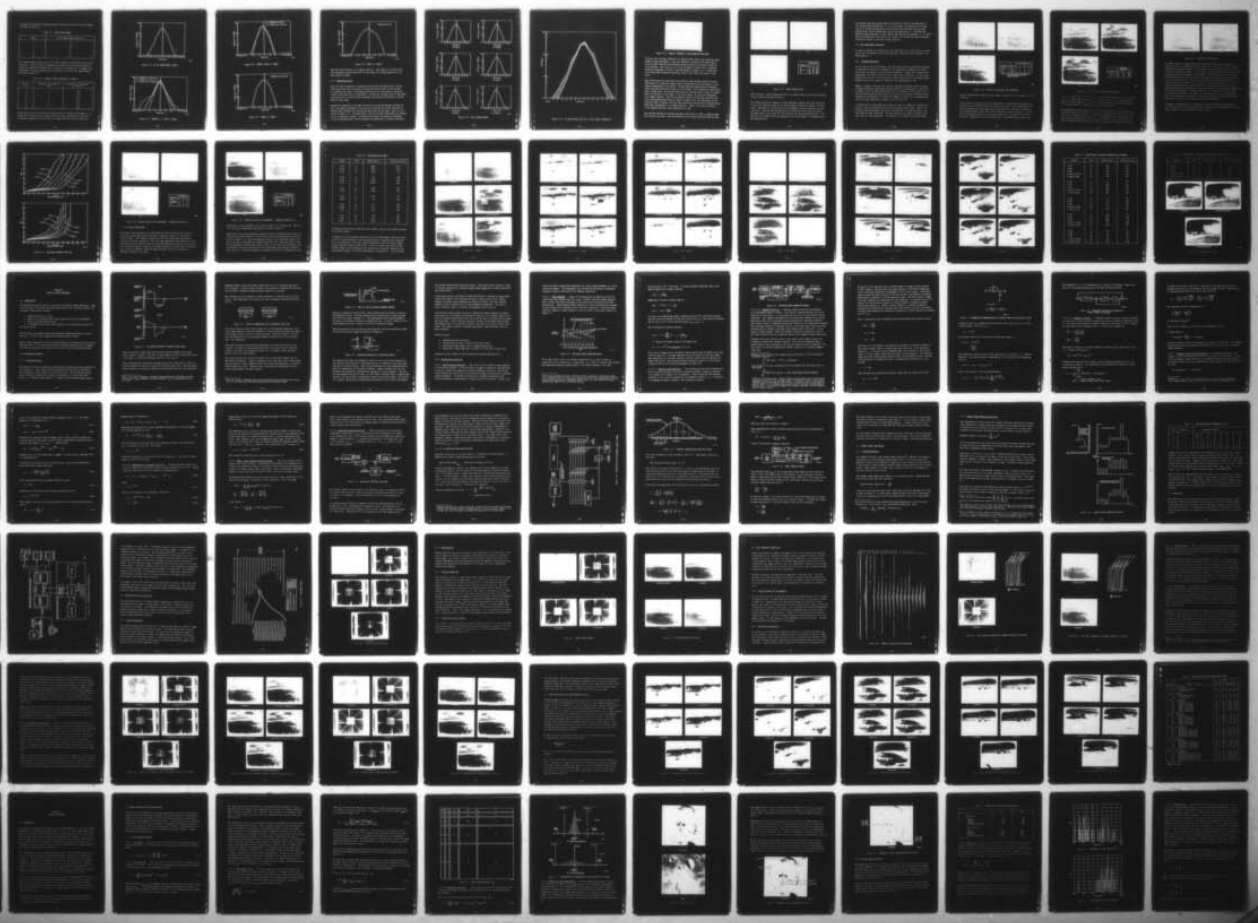
Sequency Group	Average Adaptive Threshold $\bar{\theta}$						$\bar{\theta}$ Average
	Image 1	Image 5	Image 6	Image 7	Image 8	Image 10	
9	4.2	3.9	2.8	4.2	1.9	1.8	3.1
10	1.9	1.4	1.2	1.9	0.9	0.9	1.4
11	18.6	26.7	14.1	23.3	17.6	11.4	18.6
12	3.1	3.4	1.9	3.3	1.6	1.4	2.4
13	4.6	4.1	3.4	5.1	3.6	2.7	3.9
14	9.3	7.0	6.5	10.5	6.3	6.2	7.6
15	10.1	7.8	7.4	12.4	7.6	7.5	8.8

θ takes on large values, it is preventing high-power texture terms from being considered as edges. To fulfill this second function θ frequently is quite large as reflected in the large values of $\bar{\theta}$. Since $\bar{\theta}$ can exhibit values greater than the maximum observed coefficient amplitudes, it is therefore difficult to derive a fixed-edge threshold θ_j from the values of $\bar{\theta}$ shown in Table 4-2.

Another approach to determining valid fixed-edge thresholds is to require that, on the average, θ_j identifies an equal number of edge terms as θ. The percentages

AD-A049 323 NORTHROP CORP ANAHEIM CALIF ELECTRO-MECHANICAL DIV F/G 17/5
EXPERIMENTAL DEVELOPMENT OF A FLIR SENSOR PROCESSOR.(U)
SEP 77 T LEIBOFF, H HENNING, T NODA, B DEAL DAAG53-76-C-0188
UNCLASSIFIED NORT-77Y196 NL

2 OF 3
ADA049323



of edge terms observed in sequency groups 9 through 15 in each of the test images are given in Table 4-3.

Table 4-3. PERCENT EDGE TERMS

Image	Percent Edge Terms Groups 9-15
1	13
5	23
6	22
7	12
8	18
10	21

The average number of edge terms in the higher sequencies identified by θ_j is about 18 percent. Values of θ_j with comparable performance can be determined from the coefficient amplitude histogram shown in Figures 4-4 through 4-8. A summary of the percentages of terms exceeding some intermediate values of θ_j that approximate $\bar{\theta}$ are shown in Table 4-4.

Table 4-4. NUMBER OF TERMS EXCEEDING θ_j (IMAGE 1)

Group No. (j)	Coefficient with Amplitude $\leq(\theta_j)$	Probability of Exceeding θ_j (Percent)
9	2	19.4
10	2	19.6
11	2	23.0
12	2	10.5
13	2	16.1
14	2	28.7
15	3	20.1

These values of θ will be used as a starting point in subsequent experiments. Although these values of θ_j are based on only one image, examination of the total ac energy histograms for all six FLIR test images (Figures 4-9 and 4-10) shows

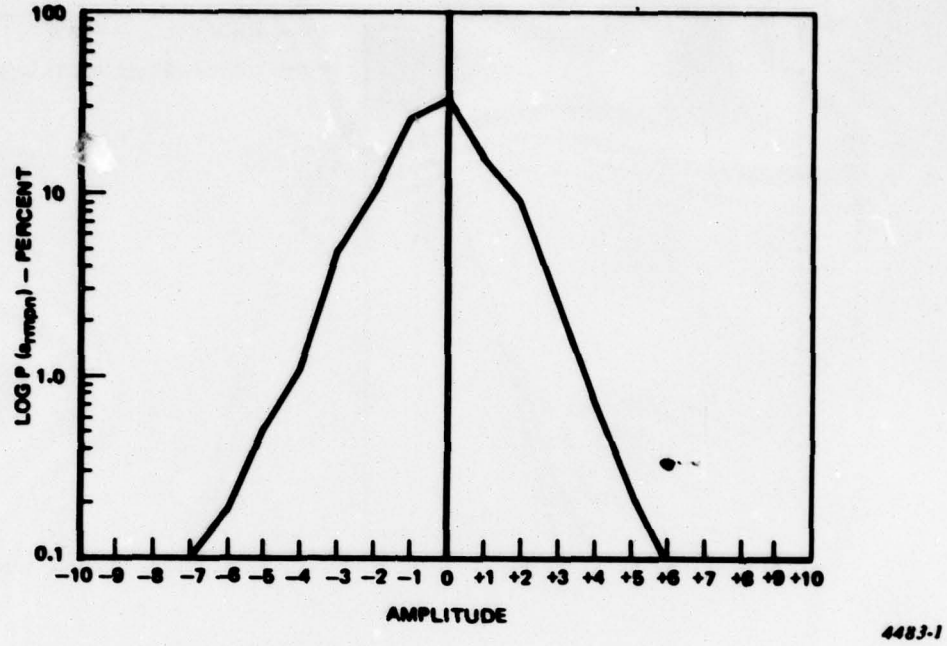


Figure 4-4. ALL AC COEFFICIENTS, IMAGE 1

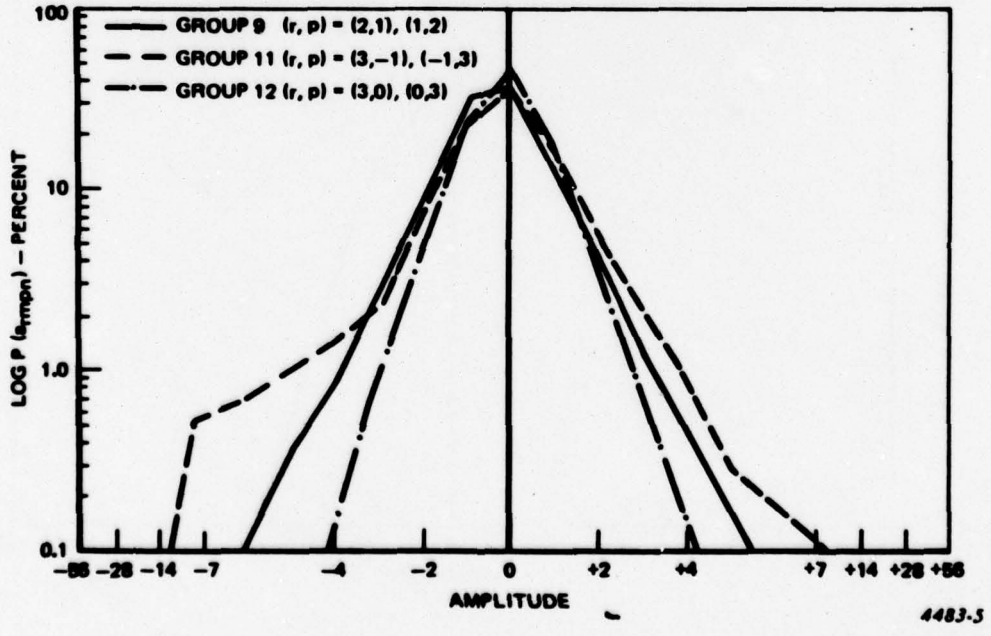


Figure 4-5. GROUPS 9, 11, AND 12, IMAGE 1

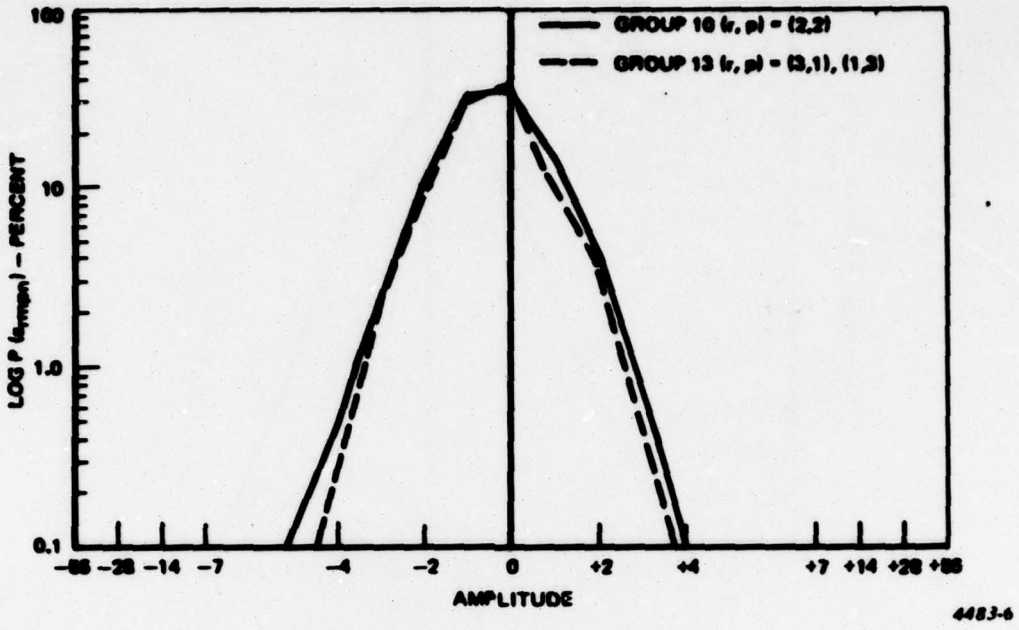


Figure 4-6. GROUPS 10 AND 13, IMAGE 1

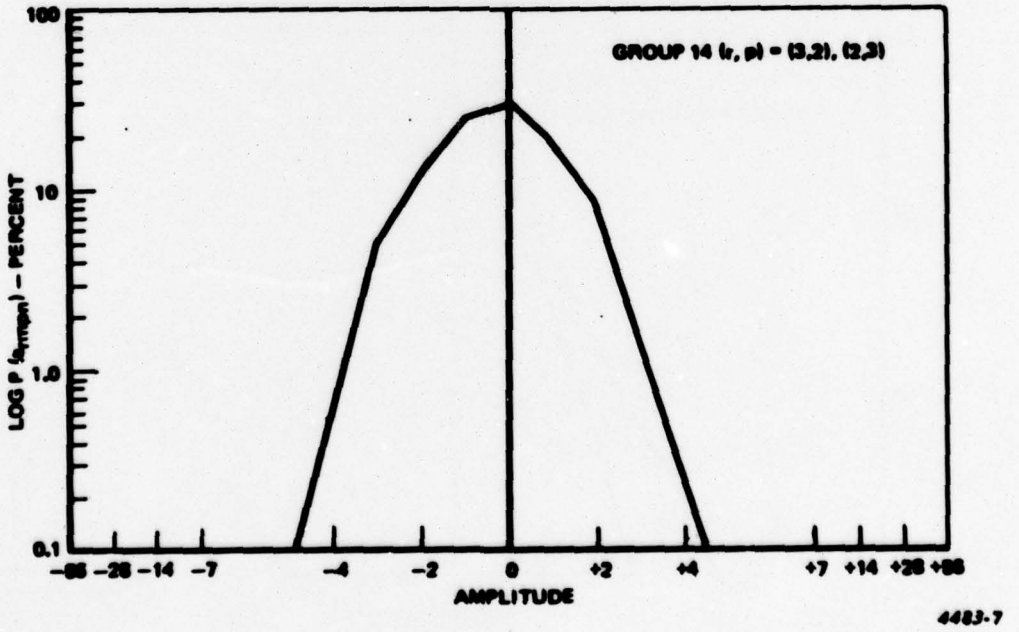


Figure 4-7. GROUP 14, IMAGE 1

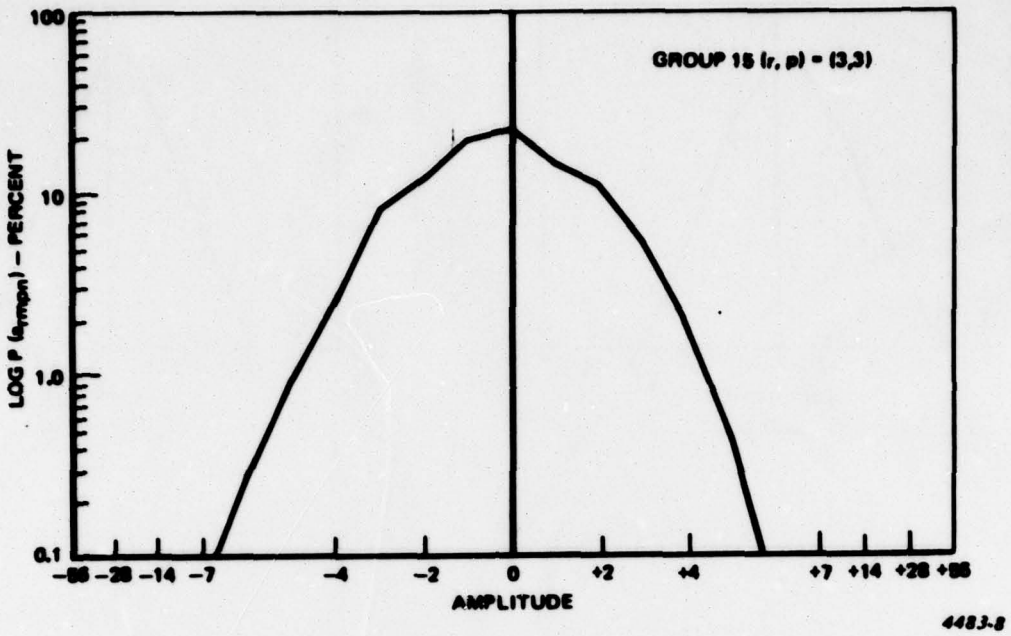


Figure 4-8. GROUP 15, IMAGE 1

amplitude distributions to be roughly identical. This leads us to believe that the enhancement algorithm selected for one image will perform equally well for other types of scenes.

4.3.2 Weighting Factors

After the edge information is separated from texture and system noise, the Haar coefficients must then be altered to obtain the desired effect on the display. To accomplish this goal, the weighting factors for the edge terms in each group of Haar coefficients must be determined. In addition, an upper limit of edge amplitude must be set to avoid further enhancement and possible distortion of already highly visible edges.

Because the dimensions of the Haar functions to be altered (Groups 9 through 15) vary, the apparent brightness of one group of functions may differ from that of another group although their coefficient amplitudes are the same. This is illustrated in Figure 4-11. The amplitudes of each of the Haar coefficients are equally set to an arbitrary constant. As shown in the figure, the apparent brightness of the corresponding Haar functions degrades progressively as sequency increases.

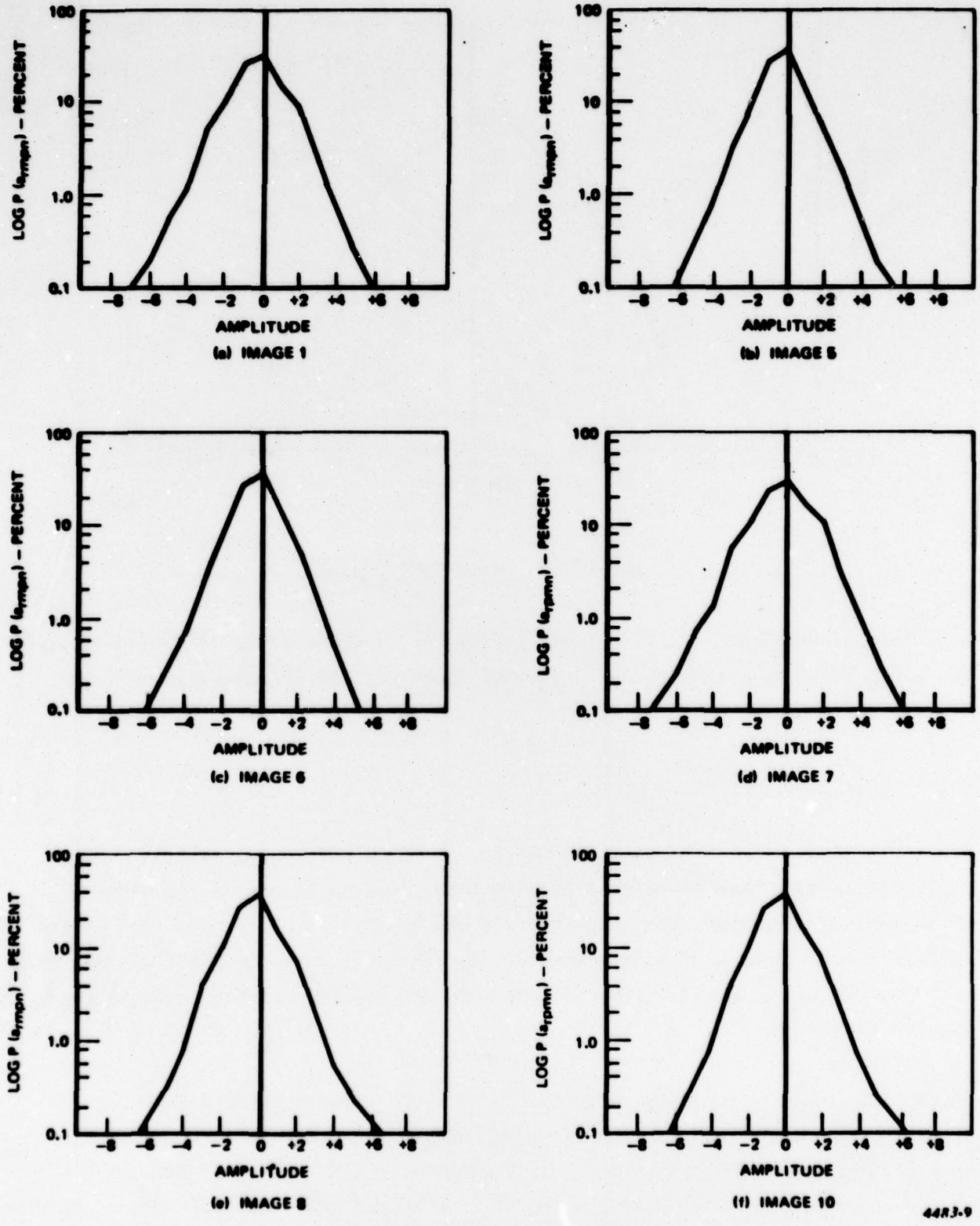


Figure 4-9. ALL AC COEFFICIENTS

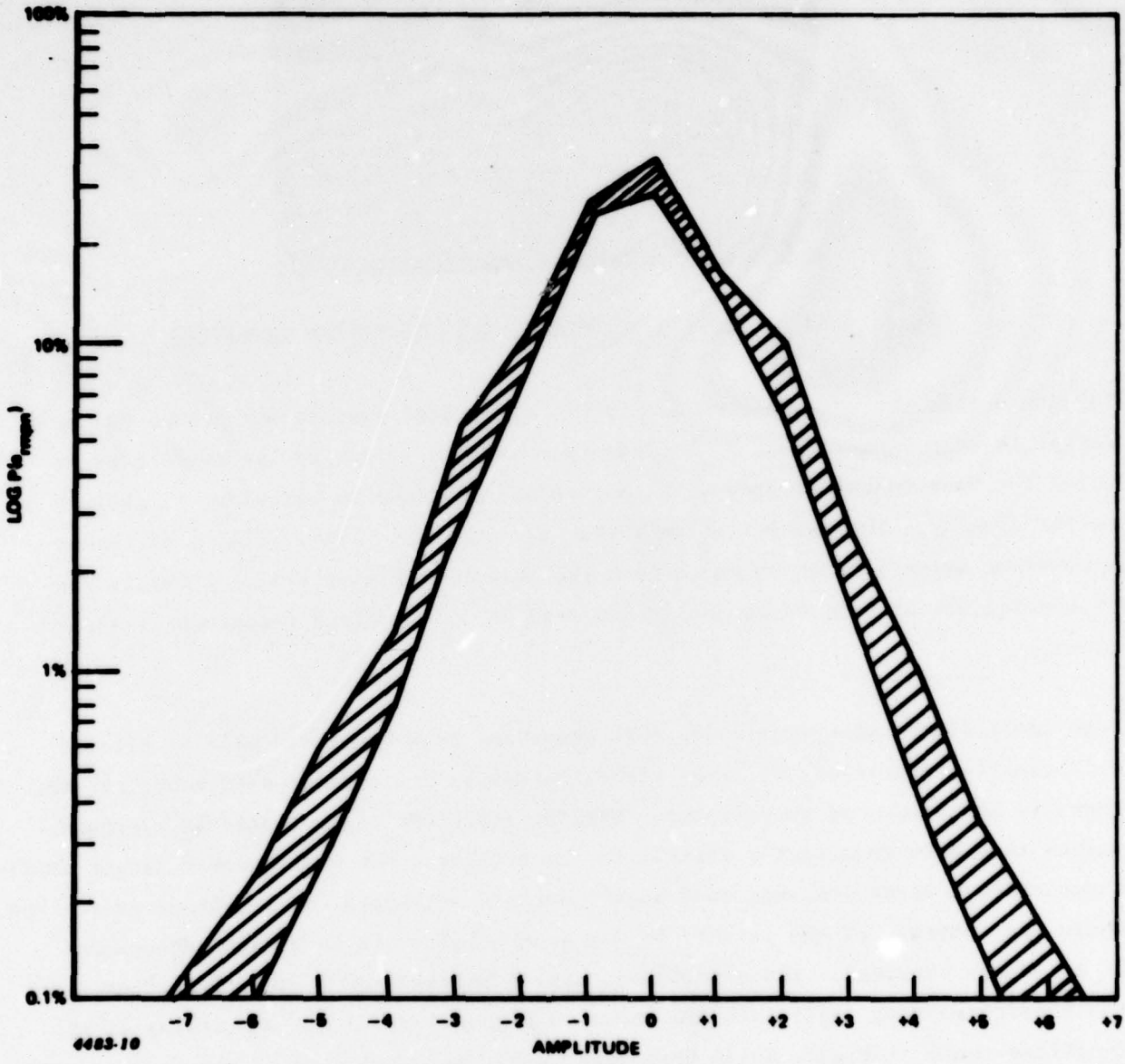


Figure 4-10. AC COEFFICIENTS FOR ALL SIX TEST IMAGES (COMPOSITE)

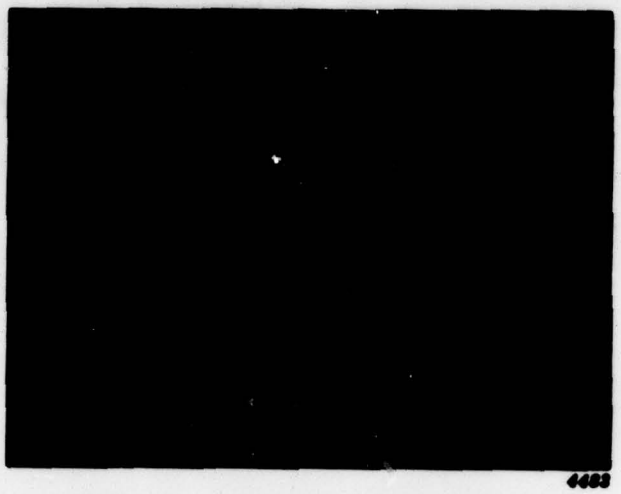


Figure 4-11. GROUPS 9 THROUGH 15 WITH IDENTICAL AMPLITUDE

To ensure that the edge information in different Haar coefficient groups has equal weight in edge enhancement, a test was conducted to determine the amplitudes at which the Haar functions from different groups produced an equivlent visibility on the display. The results of this test are shown in Figure 4-12; a different equivalent brightness is shown in each photograph of Figure 4-12. A tabulation of the equivalent brightness amplitudes used in each coefficient group is also given.

Edge coefficients with amplitudes just above the edge threshold must be altered sufficiently to crispen the image without creating distortions detrimental to the cosmetic appearance of the display. Minimum amplitude edges should be assigned values that make them barely visible in the display. For subsequently larger amplitudes of edge terms, the assigned amplitudes are increased until a point is reached where the contrast of the target, or any other object, is such that enhancement is no longer required. The selected groups of barely visible Haar functions with their corresponding amplitudes are shown in Figure 4-12a. The amplitudes of the functions shown in Figure 4-12b have been tentatively selected as the maximum amplitudes where enhancement is no longer preferred. Figure 4-12c shows the effect on the imagery when much larger amplitudes are used.

Two different methods for altering the Haar coefficients in order to enhance edges were selected for study. The two methods are the constant multiplier and the vari-



(a)



(b)



(c)

GROUP NO.	AMPLITUDES a	b	c
9	8	12	24
10	8	16	32
11, 12 & 13	10	20	40
14	12	24	48
15	16	32	64

4483

Figure 4-12. GROUP NORMALIZATION

able multiplier. Both methods make use of the edge-threshold results and weighting-factor concepts developed previously.

The constant multiplier approach to edge enhancement employs only shift operations to alter the amplitude values of the selected Haar coefficients deemed to be edge terms. Each group of coefficients may be programmed to have a different multiplying factor. Once the multiplying factor is selected for that particular group, all edge terms in that group will be altered by the same multiplying factor. This method is the simpler of the two approaches from the standpoint of hardware design and implementation.

The variable multiplier method adapts the multiplier based on the amplitude of the incoming edge information; i.e., if in one group, the amplitude of the edge term is 3, the multiplying factor would be some value different than the multiplying factor for an incoming edge term with an amplitude of 4. Although this method of edge enhancement is more complex than the previous approach, it is limited only by the amplitude resolution of the system. The variable multiplier approach effectively creates a nonlinear transfer function.

4.4 EDGE ENHANCEMENT EXPERIMENT

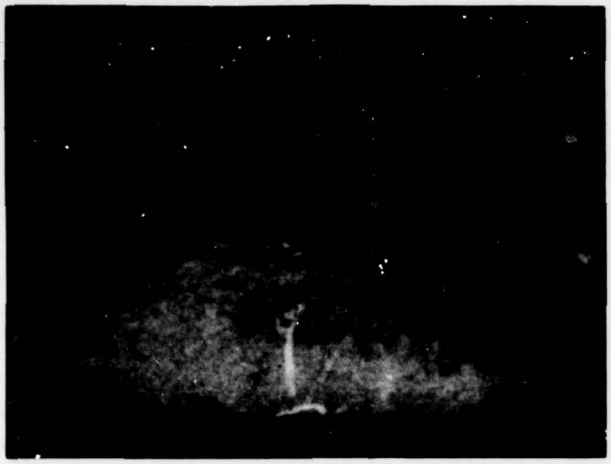
In order to evaluate the capability of the edge-detection concepts and the performance of the two methods for altering edge terms, experiments were conducted using FLIR Image No. 1.

4.4.1 Constant Multiplier

In the first series of experiments, the edge thresholds were altered while keeping the multipliers constant to determine if the edge/texture boundaries do indeed occur where the analysis of the histograms indicates. Figure 4-13 contains three photographs, each with a different set of edge/texture boundaries. In each photograph, the multipliers for each group were kept constant with the following values: groups 9 and 10 = X2, groups 11 through 14 = X4, and group 15 = X8.

Images in Figures 4-13b and 4-13c resulted from edge thresholds that were higher (thereby creating less edge terms) than the analysis in paragraph 4.3.1 indicated. In Figure 4-13a, the edge threshold was set in accordance with the study in 4.3.1. Seven untrained observers were asked to evaluate three separately enhanced images displayed on a CRT monitor. Of the three images in Figure 4-13, a consensus preferred the edge enhancement of image b over the other two.

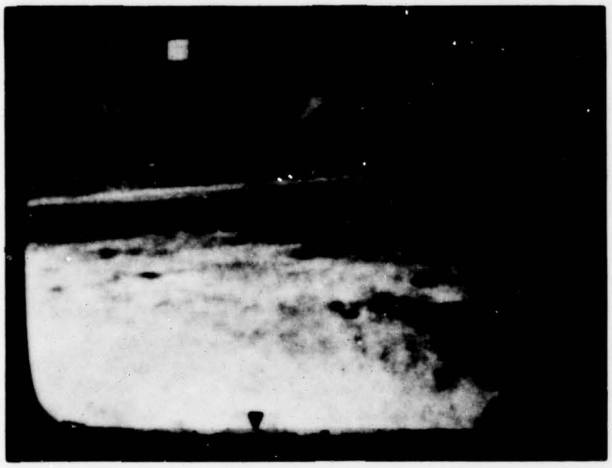
A second series of experiments was conducted on the test imagery to evaluate the effects of various weighting factors for each group of edge terms. Figure 4-14 contains three photographs, each with a different set of weighting factors, but with identical edge thresholds. The edge thresholds for image b in Figure 4-13 were used for these experiments. A majority of the viewers selected the image



(a)



(b)



(c)

GROUP NO.	MULT.	EDGE THRESHOLD*		
		a	b	c
9, 10	X2	2	3	4
11, 13, 14	X4	2	3	4
12	X4	1	2	3
15	X8	3	4	5

*NOTE: ALL VALUES EQUAL TO OR GREATER THAN THE VALUES SHOWN ARE DEFINED AS EDGES

4483

Figure 4-13. EFFECTS OF DIFFERENT EDGE THRESHOLDS

labeled b when asked to grade the three images in Figure 4-14 in accordance to their preference.

In the course of conducting these two experiments, two undesirable phenomena were observed on the CRT monitor that are not very discernible on the photographs. The first phenomenon was a glittering effect on the monitor, created by enhancement of the high-frequency noise (shot noise) prevalent in groups 14 and 15. To eliminate this problem, alteration of the Haar coefficients in groups 14 and 15 were inhibited. Although this action prevents quite a few high-frequency edge terms

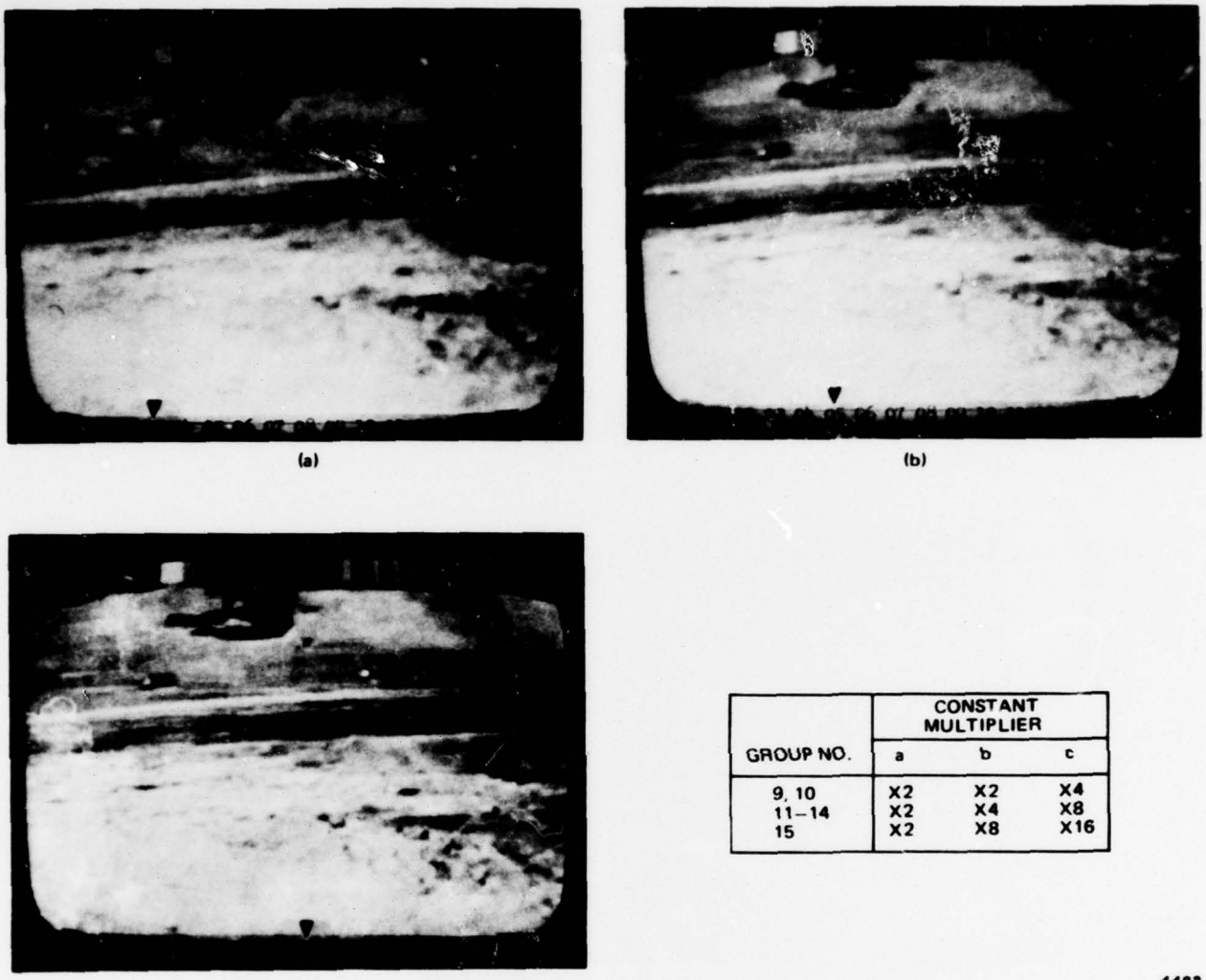


Figure 4-14. EFFECTS OF DIFFERENT MULTIPLIERS

from being enhanced, it had very little, if any, detrimental effect on the performance of the edge-enhancement properties on the display. This is illustrated in Figure 4-15 where image a is a result of enhancing groups 9 through 15 and image b is the end product of enhancing only groups 9 through 13.

The second undesired phenomenon observed on the CRT monitor was the appearance of numerous horizontal lines where a significant horizontal edge existed in the original; e.g., the horizon of the test image. This phenomenon is created mainly by the interlace format of the TV raster scan. Northrop's sensor processor does not store a complete field of video data and then operate on 16 consecutive lines of

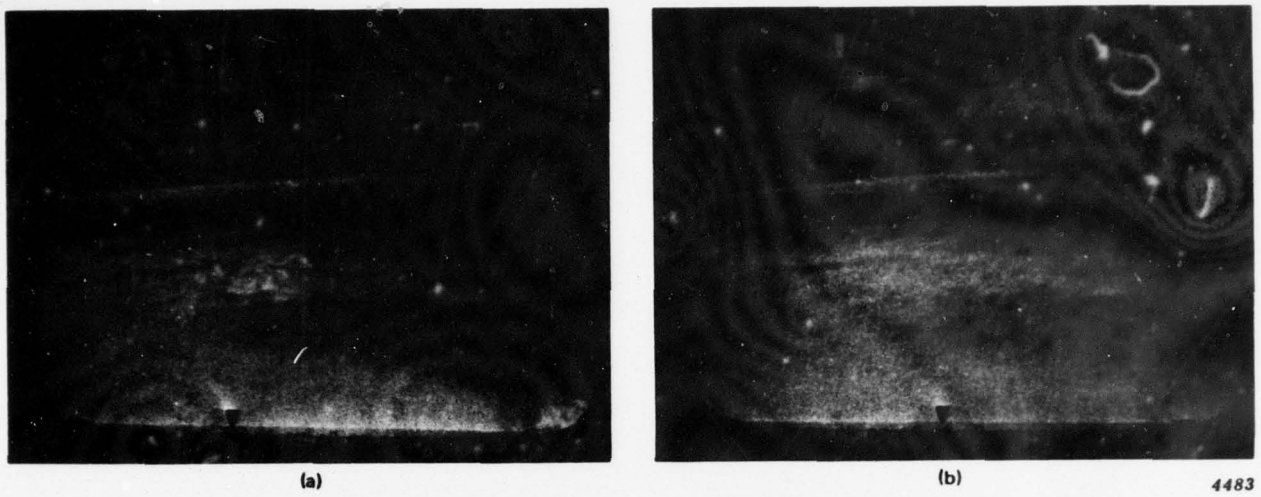


Figure 4-15. EFFECTS OF SYSTEM NOISE

the image. Rather it computes the Haar transform on the 16 consecutively presented TV lines of a single field (every other line of the scene). Thus, the edge-enhancement algorithm interprets a single edge in the horizontal direction as being two edges (one edge per field). Over a frame time when these two edges are enhanced, multiple horizontal lines appear on the display. This problem can be minimized by using weighting factors that are smaller for edges in the horizontal direction than in the vertical direction. The sensor processor has the capacity to separate one group into two in its present form. Group 11, which contributes the most in creating the undesirable effects of horizontal lines, was therefore divided into two separate groups and the multiplier for the half of group 11 that measures differences in the vertical direction was reduced from X4 to X2. The effects of this action are illustrated in Figure 4-16; image a is a photograph of the test image before group 11 was separated, and image b is the result of group 11 separation and reduction of the multiplier for the vertical direction differences.

Im summary, the parameters of the constant multiplier algorithm that produced the most desired effects of edge enhancement for the majority of the observers are tabulated in Table 4-5 and illustrated in Figure 4-16b.

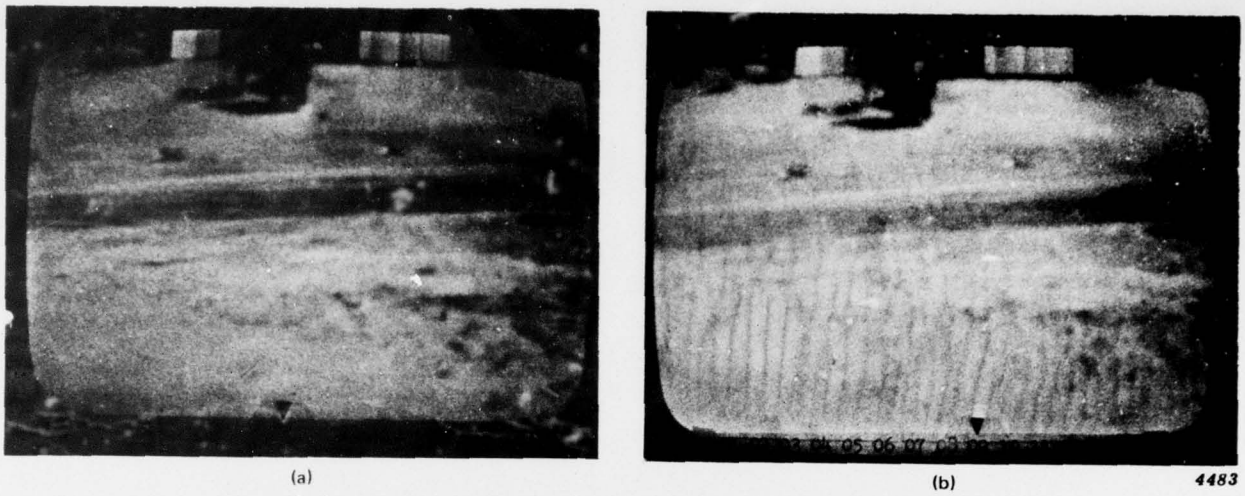


Figure 4-16. EFFECTS OF TV 2:1 INTERFACE

Table 4-5. SELECTED PARAMETERS FOR CONSTANT MULTIPLIER

Group No.	Amplitude Values		Multipiier
	From	To	
9	3	12	X2
10	3	16	X2
11 (Vert.)	3	20	X2
11 (Hor.)	3	20	X4
12	2	20	X4
13	3	20	X4

4.4.2 Variable Multiplier

The majority of the basic parameters that affect the quality of edge enhancement using the variable multiplier algorithm have been determined in previous experiments. The remaining question to be addressed for this algorithm is the shape and form of the preferred nonlinear transfer function. Two basic forms of nonlinear transfer functions with varying degrees of slope were selected for study. The difference in the two basic forms is that in one transfer function the maximum output amplitude was curtailed to a smaller value than the other.

The curves in Figure 4-17 for nonlinear transfer function No. 1 were programmed in programmable read-only memories (PROMS) and the sensor processor modified to allow the PROM, rather than the constant multipliers, to alter the selected edge terms.

The representative test image was subjected to varying degrees of edge enhancement by selection of different curves for the various groups of Haar coefficients to be altered and subjectively evaluated. The effects of this is shown in Figure 4-18 along with a tabulation of curves used for each of the three photographs. It is shown that the edges of the targets, trees, cold horizontal band across the center of the image gets progressively sharper and discernible from image a to b to c although the artifacts (shaped like Haar functions) are also increased. The majority of the evaluators selected image b as their preference.

The procedures and tests performed on nonlinear transfer function No. 1 was also conducted on the curves for nonlinear transfer function No. 2. The resultant effects on the display (shown in Figure 4-19) was similar to the test conducted for transfer function No. 2. The majority of the evaluators again selected image b as their preference.

The evaluators were not able to arrive at a consensus in selecting the preferred nonlinear transfer function, because the resultant effects on the display were so similar. As a result, the curves used to derive an enhanced image in Figure 4-19b were selected for further testing.

4.4.3 Descriptors

The Image Descriptors (ID's) for the enhanced versions of FLIR Image 1 are given in Table 4-6, which shows that edge-enhanced images exhibit substantially larger edge activities with only a marginal increase in texture activity. However, some anomalies do exist in the relative levels of edge activity between subjectively equivalent images.

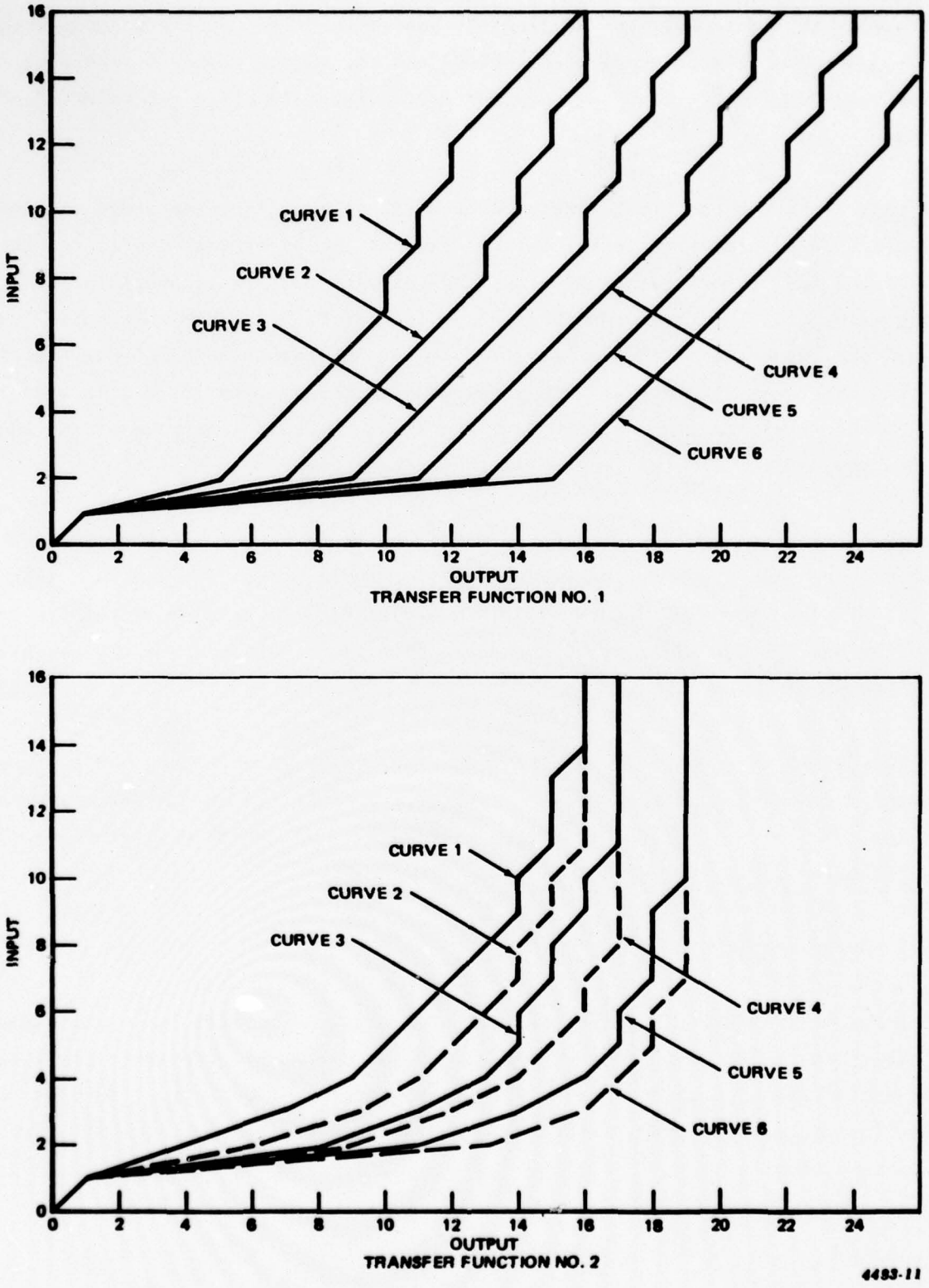
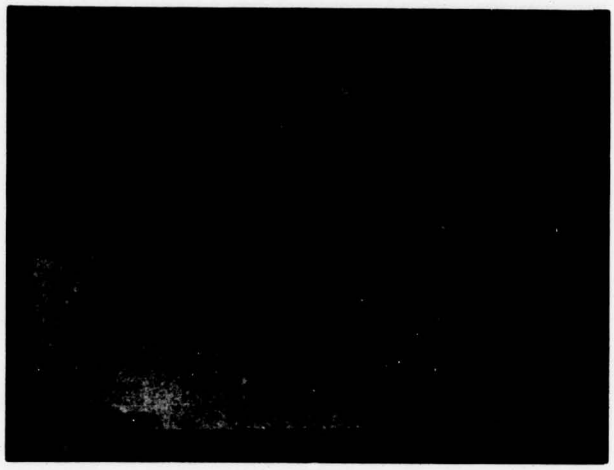
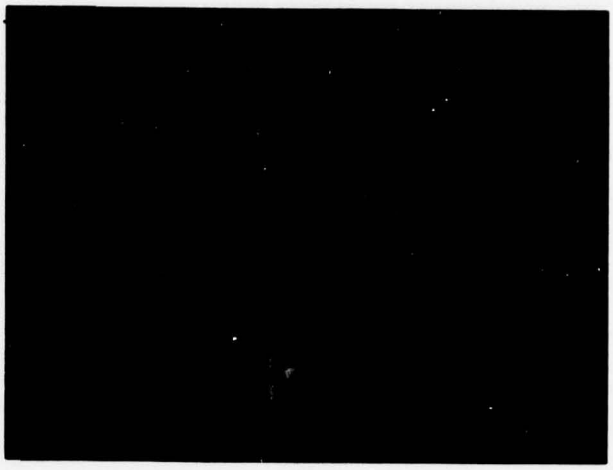


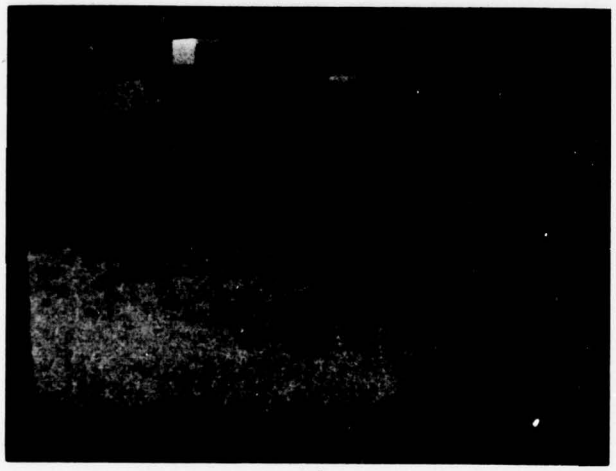
Figure 4-17. NONLINEAR TRANSFER FUNCTIONS



(a)



(b)



(c)

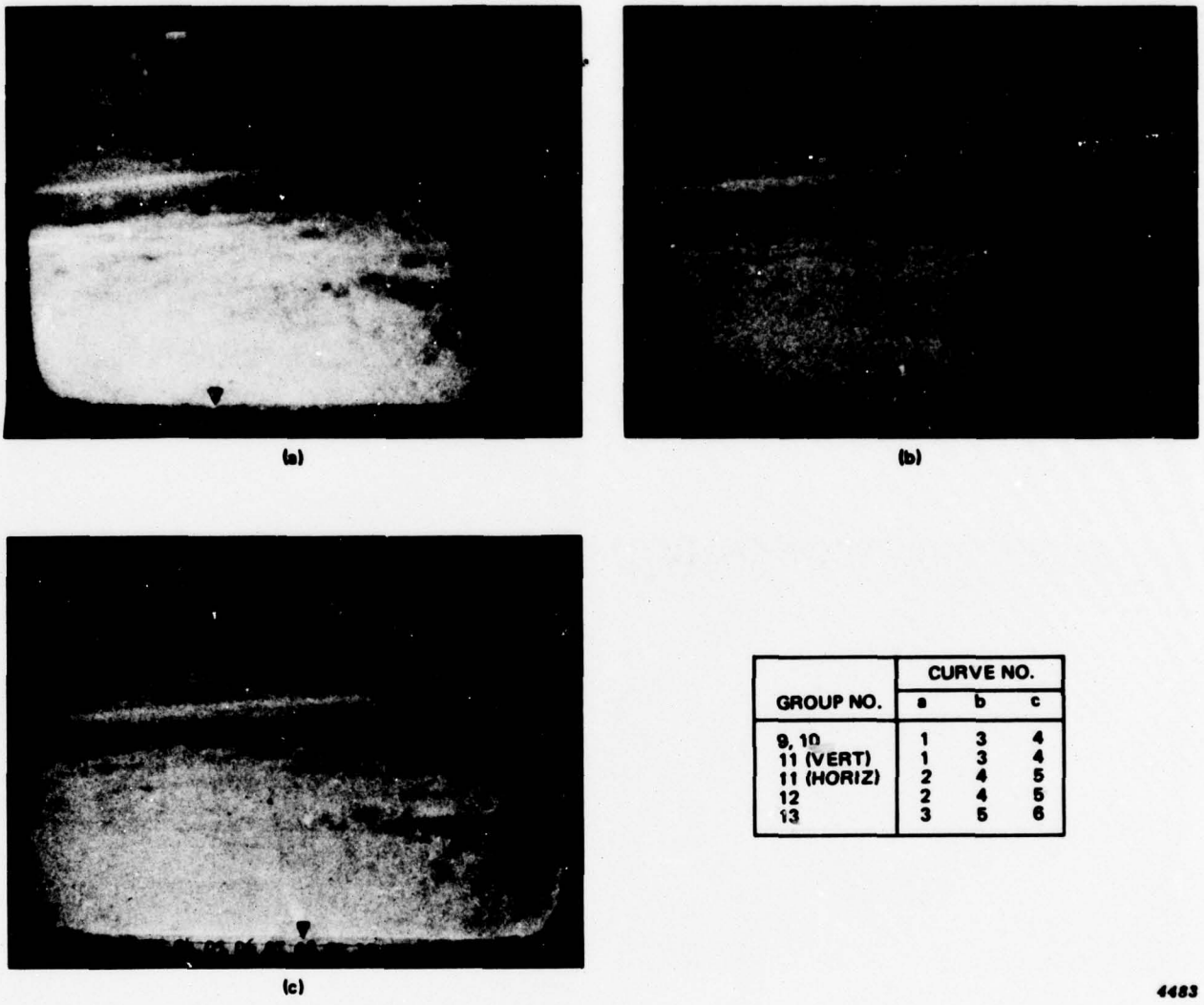
GROUP NO.	CURVE NO. a	b	c
9, 10	1	3	4
11 (VERT)	1	3	4
11 (HORIZ)	2	4	5
12	2	4	5
13	3	5	6

4483

Figure 4-18. VARIABLE MULTIPLIER ENHANCEMENT - TRANSFER FUNCTION NO. 1

4.5 FLIR IMAGE EXPERIMENTS

Experiments were conducted on six FLIR images in order to evaluate the ability of the selected edge enhancement algorithms to emphasize edges in a wide range of diverse images. Each image was subjected to the three enhancement algorithms, selected for further evaluation. The three edge enhancement algorithms are (1) the constant-multiplier approach using the fixed-ege threshold criterion, (2) the variable multiplier approach using the fixed-edge threshold criterion, and (3) the variable multiplier approach, using the Image Descriptor (ID) as a criterion for determining edges in the scene.



GROUP NO.	CURVE NO. a	b	c
9, 10	1	3	4
11 (VERT)	1	3	4
11 (HORIZ)	2	4	5
12	2	4	5
13	3	5	6

Figure 4-19. VARIABLE MULTIPLIER ENHANCEMENT - TRANSFER FUNCTION NO. 2

Photographs of these experiments are shown in Figures 4-20 through 4-25. Descriptors of selected enhancement algorithms are listed in Table 4-7.

Experiments for the edge-enhancement algorithm using the image descriptor were conducted on the sensor processor in conjunction with the instrumentation system (depicted in Figure 4-1). The instrumentation system, with its general purpose computer, was used to implement the complex algorithm of the image descriptor. Since the instrumentation system cannot process video in real time, each rectangular area (containing 12 sectors) processed by the ID enhancement algorithm was integrated over eight runs. This was performed to simulate the effects of integration by the

Table 4-6. DESCRIPTORS FOR IMAGE 1

Figure	Tone	Edge Activity	Texture Activity
2-13a	19	8883	2619
2-13b	22	5551	559
2-13c	22	2324	475
2-14a	22	937	408
2-14b	23	5859	648
2-14c	22	12,700	1459
2-15a	22	6050	645
2-15b	22	1094	596
2-16a	22	1106	599
2-16b	22	779	430
2-18a	22	647	389
2-18b	22	913	500
2-18c	21	1107	563
2-19a	22	656	392
2-19b	23	975	507
2-19c	22	1081	562

CRT display and eye-brain combination upon slightly varying video signals produced by the video camera.

In addition to the experiments conducted on the six FLIR images, identical experiments for constant-multiplier and variable-multiplier approaches were conducted on a live parking-lot scene (not a phototransparency) using the vidicon camera (visual spectrum) shown in Figure 4-26. It can be readily seen that the edge-enhanced images of the parking lot were significantly improved, compared to the edge-enhanced images of the FLIR's photographic transparencies. It was deduced

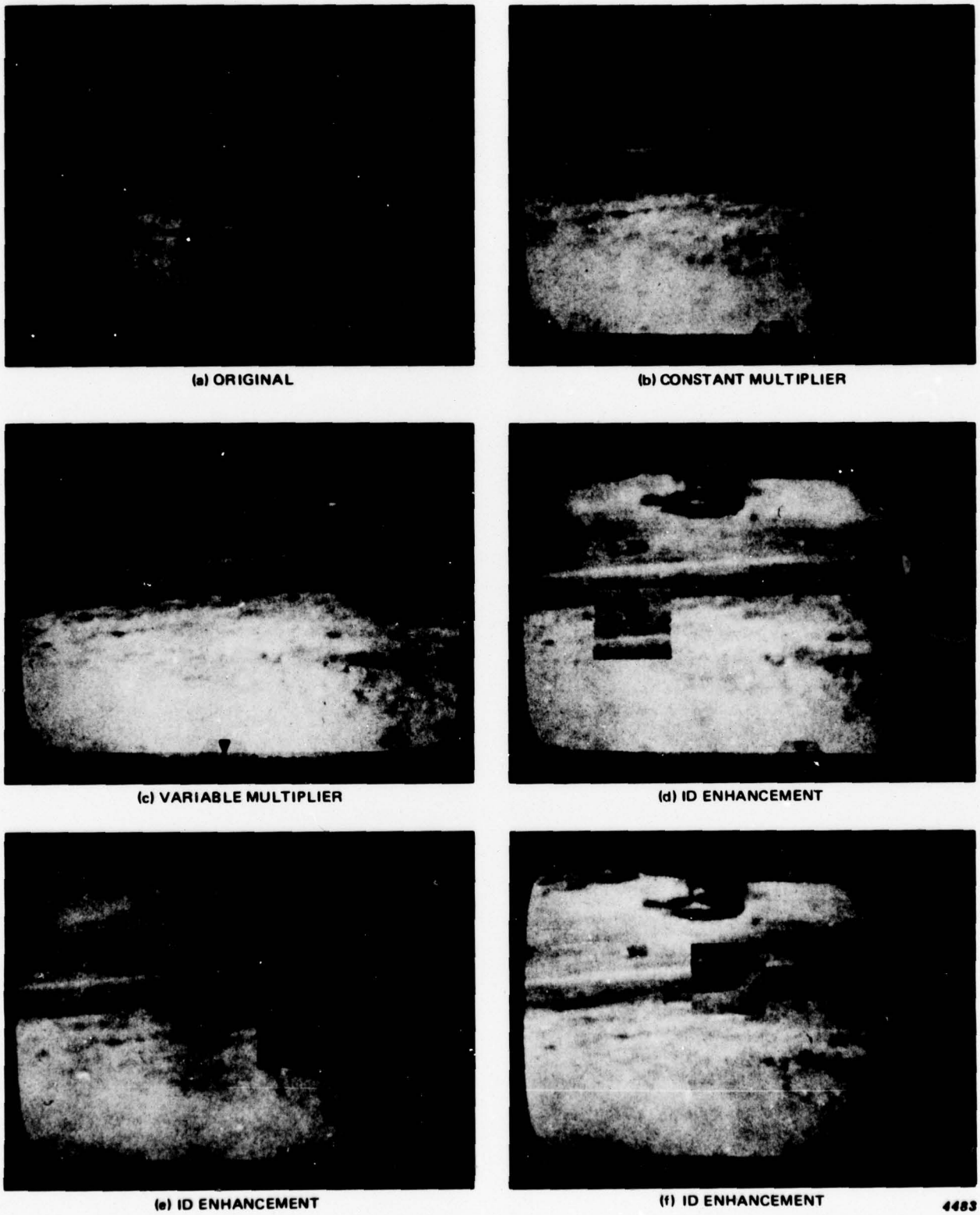


Figure 4-20. IMAGE 1

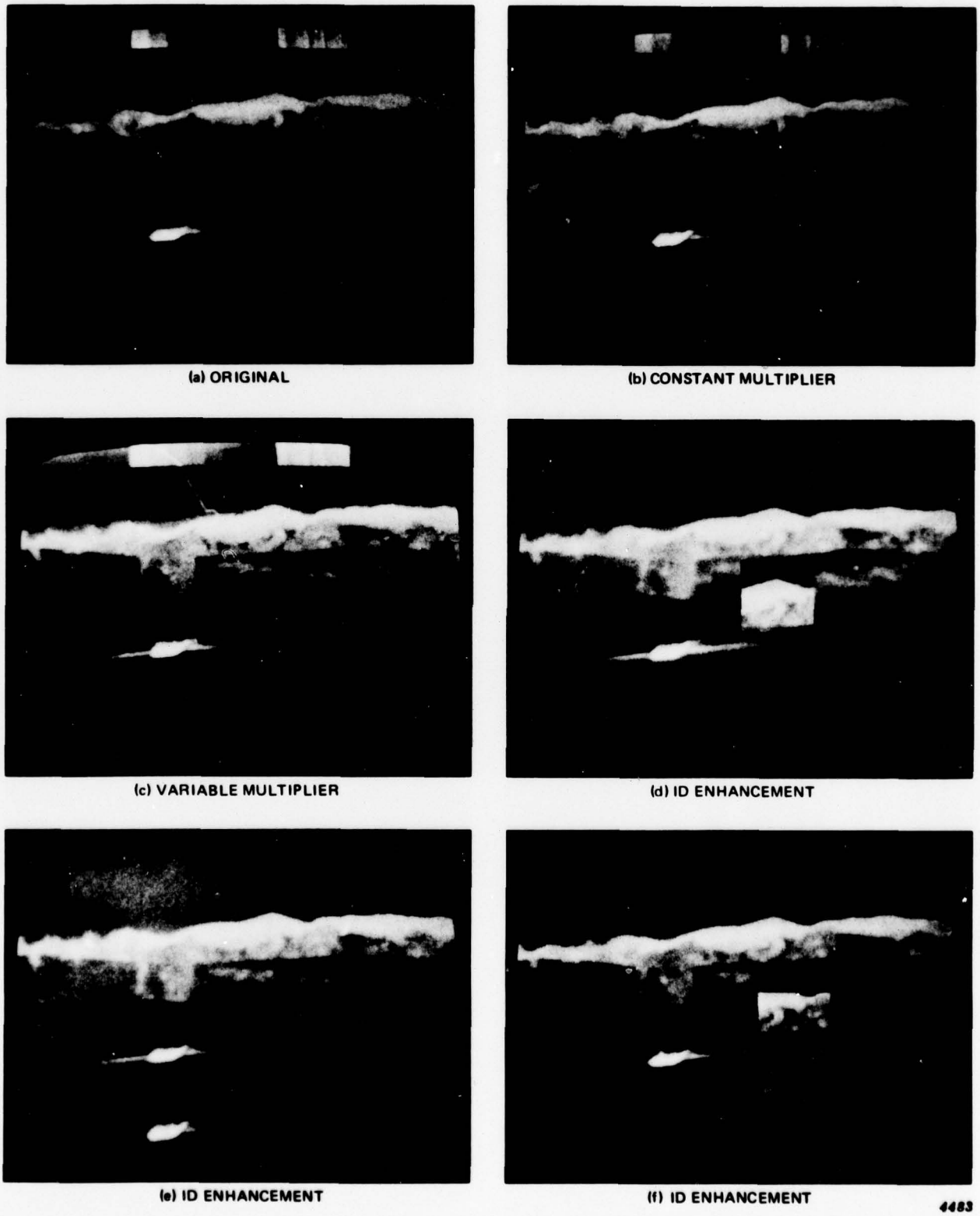
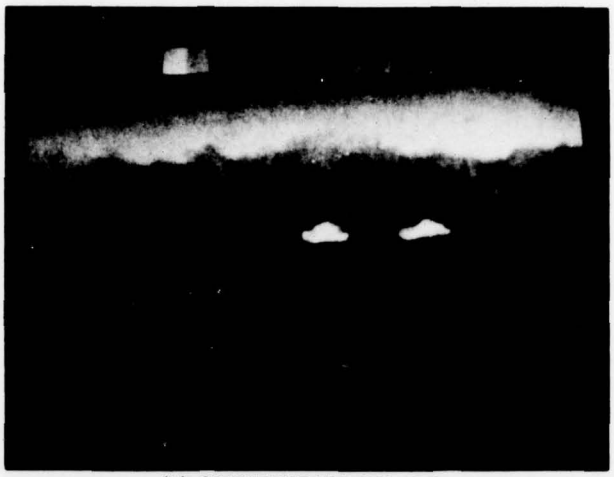


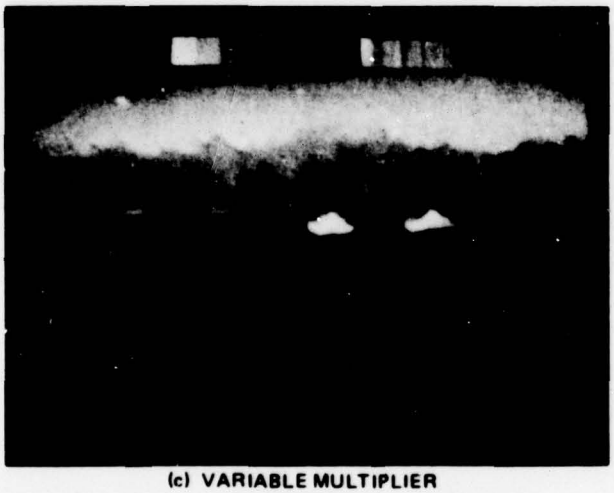
Figure 4-21. IMAGE 5



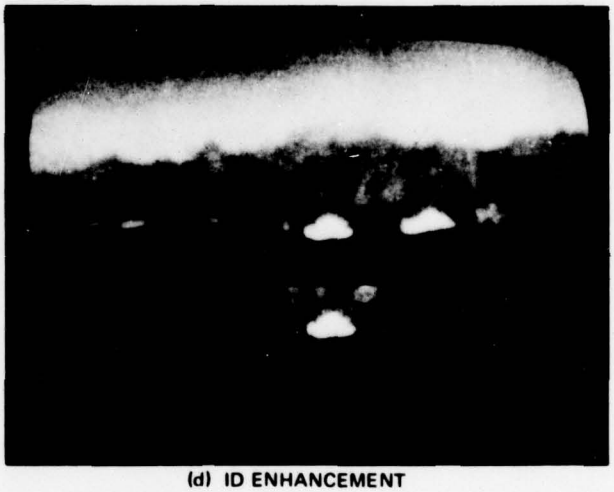
(a) ORIGINAL



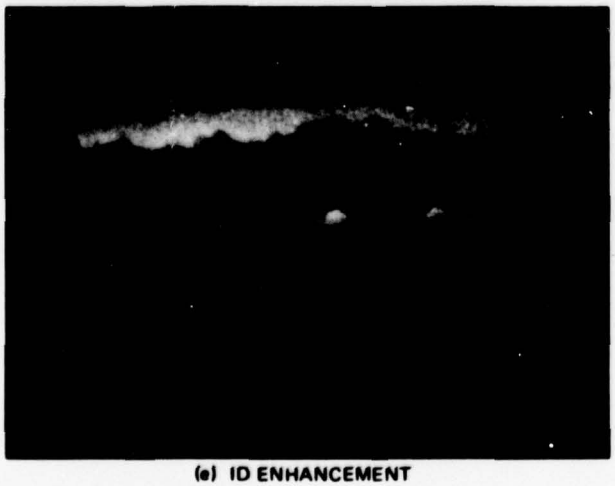
(b) CONSTANT MULTIPLIER



(c) VARIABLE MULTIPLIER



(d) ID ENHANCEMENT



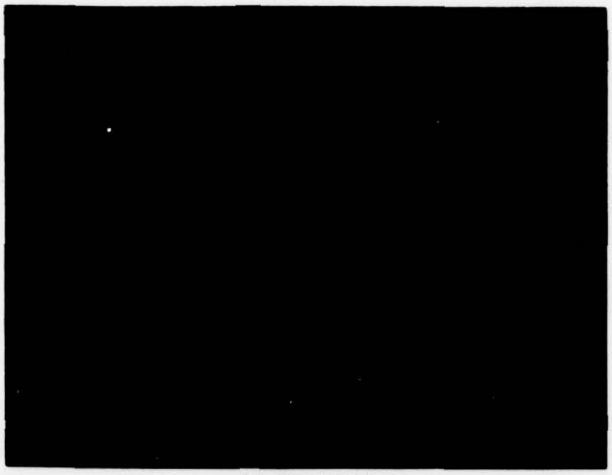
(e) ID ENHANCEMENT



(f) ID ENHANCEMENT

4483

Figure 4-22. IMAGE 6



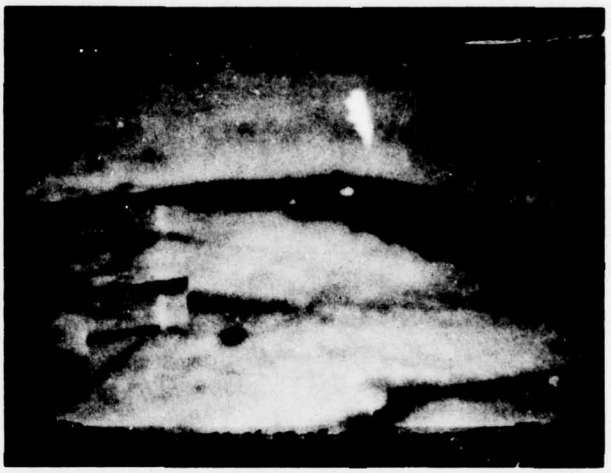
(a) ORIGINAL



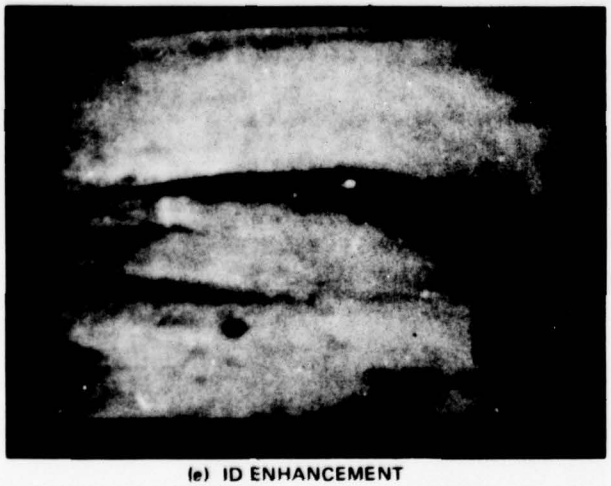
(b) CONSTANT MULTIPLIER



(c) VARIABLE MULTIPLIER



(d) ID ENHANCEMENT



(e) ID ENHANCEMENT



(f) ID ENHANCEMENT

4483

Figure 4-23. IMAGE 7

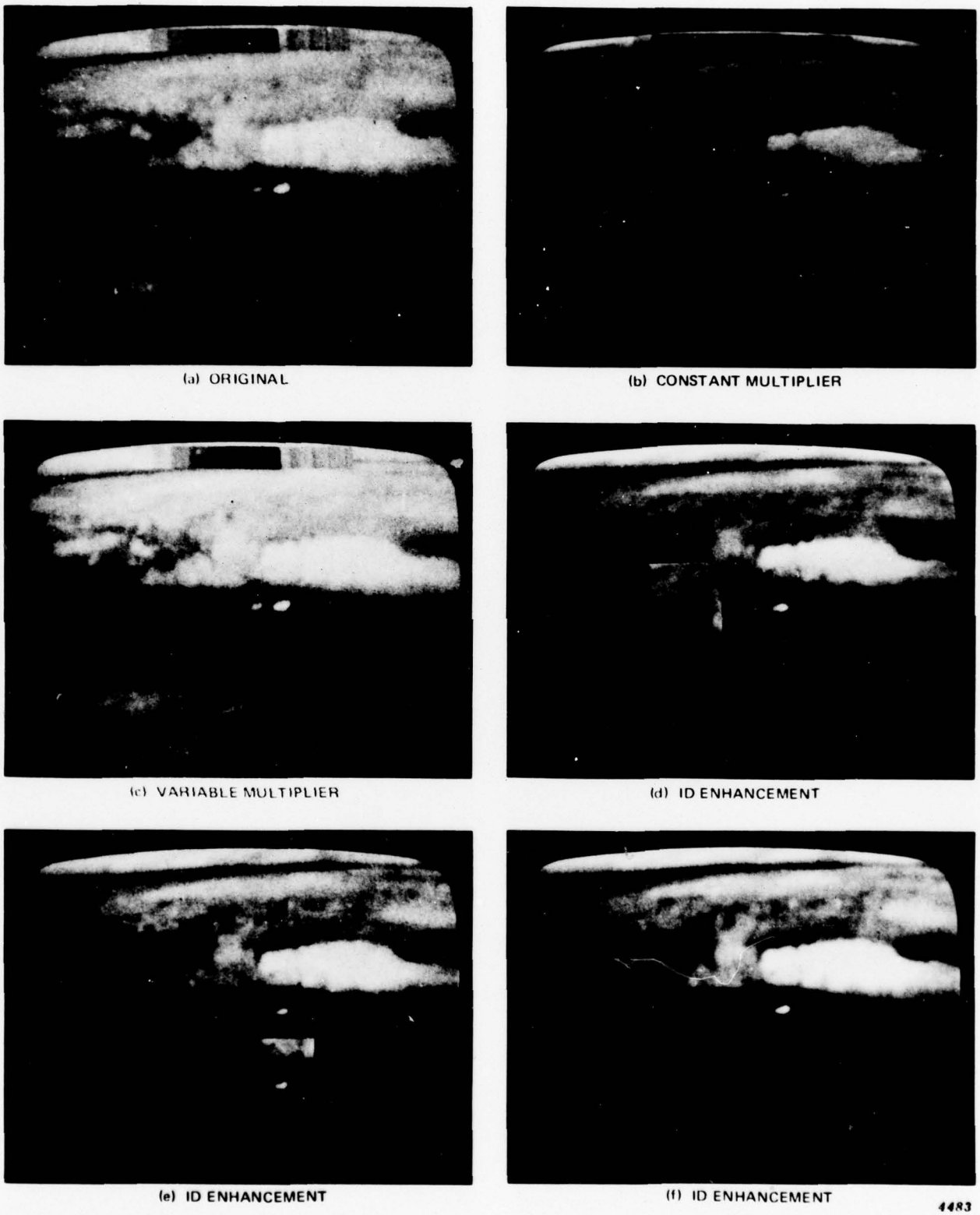
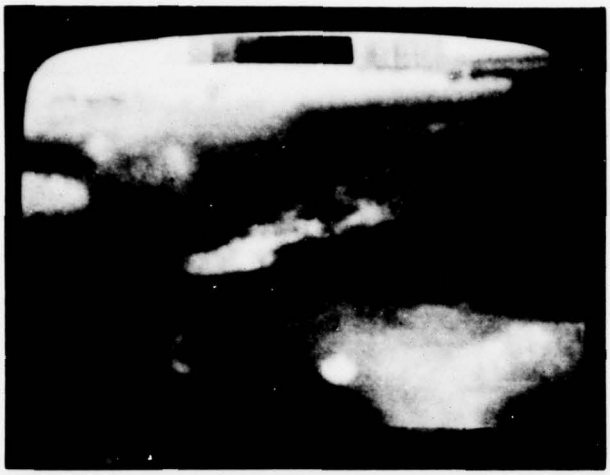


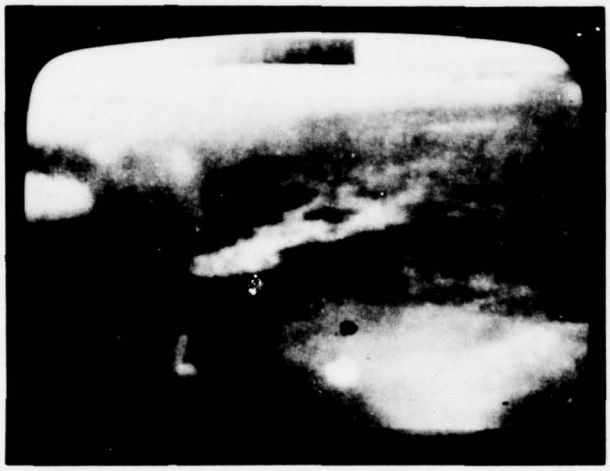
Figure 4-24. IMAGE 8



(a) ORIGINAL



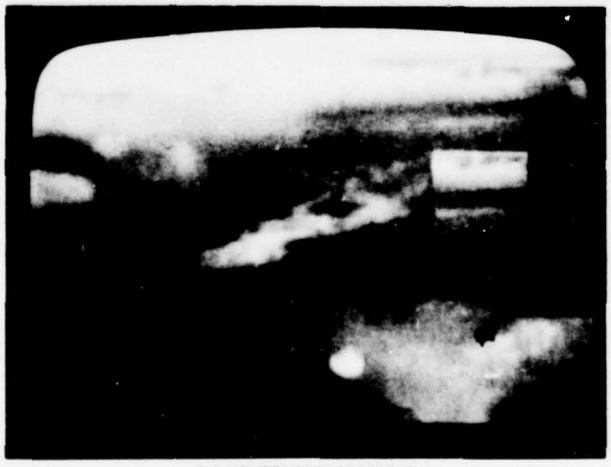
(b) CONSTANT MULTIPLIER



(c) VARIABLE MULTIPLIER



(d) ID ENHANCEMENT



(e) ID ENHANCEMENT



(f) ID ENHANCEMENT

4483

Figure 4-25. IMAGE 10

Table 4-7. DESCRIPTORS OF SELECTED ENHANCEMENT ALGORITHMS

Figure	Tone	Edge Activity	Texture Activity
4-20a	21	455	322
4-20b	21	704	404
4-20c	21	801	436
4-20d,e,f Orig	22	506	369
4-20d,e,f Enh	22	829	379
4-21a	13	631	301
4-21b	13	886	464
4-21c	13	880	464
4-21d,e,f Orig	28	1052	494
4-21d,e,f Enh	28	1685	685
4-22a	23	434	266
4-22b	22	574	384
4-22c	22	615	384
4-22d,e,f Orig	20	714	392
4-22d,e,f Enh	20	1362	631
4-23a	26	658	453
4-23b	26	1008	688
4-23c	26	1080	783
4-23d,e,f Orig	20	756	488
4-23d,e,f Enh	20	1227	596
4-24a	17	563	277
4-24b	18	715	412
4-24c	18	812	504
4-24d,e,f Orig	26	614	317
4-24d,e,f Enh	26	1218	582

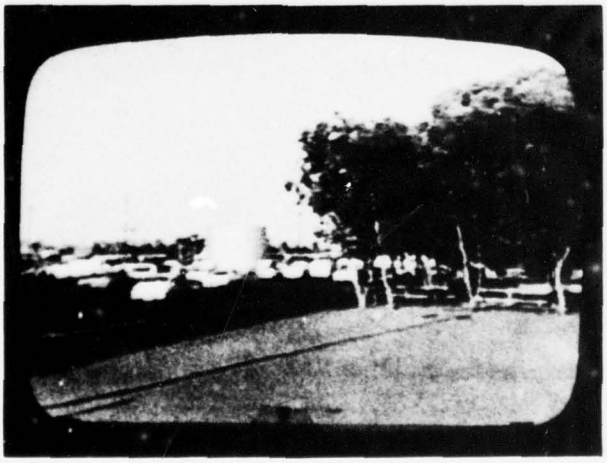
Table 4-7. DESCRIPTORS OF SELECTED ENHANCEMENT ALGORITHMS (Continued)

Figure	Tone	Edge Activity	Texture Activity
4-25a	19	516	263
4-25b	19	665	350
4-25c	19	739	425
4-25d,e,f Orig	27	589	326
4-25d,e,f Enh	27	1141	595



A. ORIGINAL

B. CONSTANT MULTIPLIER



C. VARIABLE MULTIPLIER

4483

Figure 4-26. PARKING LOT SCENE

that the combination of smaller dynamic range (reduced gray shade resolution) and lower Modulation Transfer Function (MTF) of FLIR transparencies contributed to the difference in the performance of the edge-enhancement algorithms on the two types of test images. These two system parameters (dynamic range and MTF) have a great impact on the performance of the edge-enhancement algorithm since it determines the algorithm's ability to separate edges from texture information and noise; i.e., if these two system parameters are low, the smaller edge terms will not be distinguishable from texture or noise terms.

Seven observers were seated approximately 30 inches away from an 11-inch CRT monitor and shown the results of the selected edge-enhancement algorithms for all six FLIR images. On the average, five of seven of the evaluators selected the image created by the variable multiplier algorithm (b) as their preference. However, preferences for specific images by individual evaluators varied.

4-6. CONCLUSIONS

The image descriptor based on the edge criteria described in the first quarterly report provides a measure for edges. In order to meet the goal of simple processing at real-time video rates, the fixed edge threshold algorithm is recommended for the extraction of edges. The variable multiplier algorithm is recommended for edge emphasis.

The recommended algorithms are somewhat scene dependent. The dynamic range, contrast, and amount of activity influence the effectiveness of the edge enhancement to a degree. At this time, we feel that dynamic range compression (gray-scale stretch), brightness, and gain control performed before edge enhancement will mitigate this problem. Unfortunately, this could not be included within the scope of this experiment, but would be an interesting continuation in the future.

These algorithms, which indicate a high level of noise immunity, are basically impervious to low-frequency noise. Random high-frequency noise tends to be accentuated in an edge-enhanced picture, but on the whole, it should not interfere with target recognition. Noise caused by scan-interlace processing can be mitigated by moderating the horizontal edge Haar coefficients.

Finally, it is felt that the edge enhancement techniques could be built into a real-time 2D Haar image processor as an integral part of a next generation or upgraded common module FLIR using presently available technology.

SECTION 5
ADAPTIVE SIGNAL PROCESSING

5.1 INTRODUCTION

The following discussion covers dc restoration and dynamic range compression. Under dc restoration, the desirable and undesirable effects of ac coupling in FLIR's are discussed. These effects include:

a. Exponential decay, or droop
b. Ability to distinguish small targets against low-contrast backgrounds
c. Drift in average waveform level
d. No Automatic Gain Control (AGC) capability to maintain good contrast for all signal levels.

This discussion investigates Haar-transform-based algorithms designed to correct effects a, c, and d, and to improve the tradeoff between a and b.

Dynamic range compression involves nonlinear methods for optimizing use of the available gray scale. This discussion outlines an equivalent approach that rearranges the nonlinear operations in a manner better suited to Haar transform processing.

5.2 DC RESTORATION PROBLEM

5.2.1 Problem Definition

AC-coupled electronics is generally characterized by exponential decays toward zero, as in Figure 5-1. This figure typifies Infrared (IR) line scans of uniform hot and cold objects, which largely fill the field of view. The illustrated decays result from an ac coupling that blocks all frequencies (f) below some cutoff point f_o = (coupler RC time constant)$^{-1}$. Larger f_o values produce greater droop.

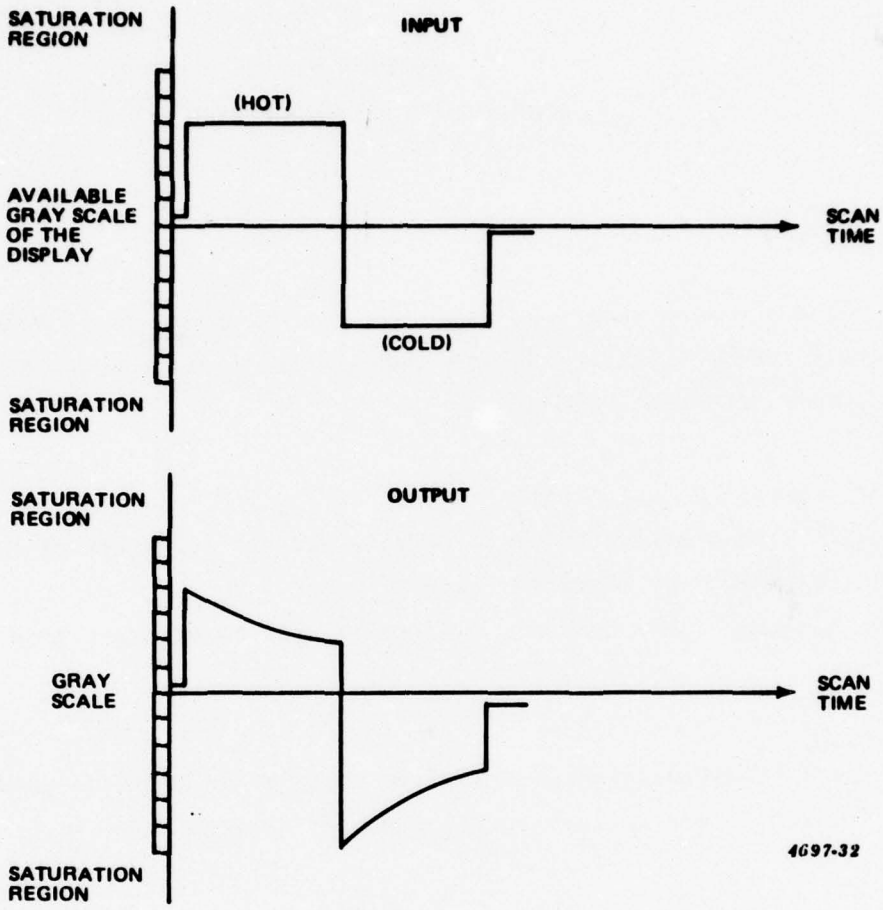


Figure 5-1. AC COUPLING EFFECTS IN PARALLEL SCAN FLIRS

They also provide a larger stop band for preventing the passage of 1/f noise. Recent studies[1], however, claim that 1/f-noise immunity can be achieved with relatively small f_o, producing negligible droop per scan time (T).

These investigations consequently disregard single-scan droop effects. They do, however, address the problem of cumulative droop over many scans. Cold bars represent one method of coping with cumulative or multiscan droop, but their operation

[1] Tack, D. H., and P. Narendra, "Automated Image Enhancement Techniques, 2nd Generation FLIR" Report #DAA653-76-C-0195, Honeywell, Inc., Systems Research Center, 1976.

sometimes leads to saturation and/or inefficient use of the available gray scale of the display.[2] Present studies therefore seek methods to eliminate multiple-scan droop via Haar transform processing without recourse to cold bars.

One consequence of this approach to droop elimination is improved vertical correlation. Such improvement would relieve the type of problems characterized by Figure 5-2.

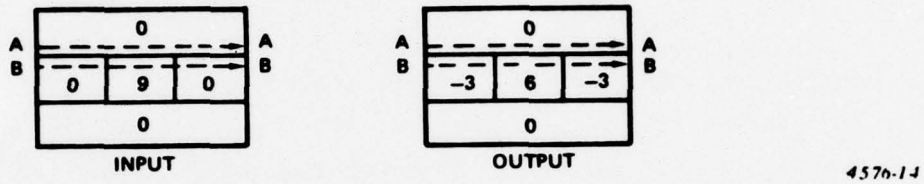


Figure 5-2. VERTICAL CORRELATION LOSS IN PARALLEL SCAN FLIRS

This figure shows how a hot central object on a cold background degenerates into cooler image against a distorted background. Parallel scans AA and BB each produce zero-average outputs by virtue of the ac coupling. Unlike AA, however, scan BB can only achieve this result through deliberate distortion of its dc level (and hence, the background).

Other goals include preservation and enhancement of the desirable effects of ac coupling. Straight dc coupling gives relatively little play to small amplitude signals. The solid-line waveform in Figure 5-3, for example, might represent a match in front of a blast furnace.

The central pip may not span even one gray level in the output display, and the match may go undetected. Ideal ac coupling, on the other hand, would assign larger spectral gains to signal components with a more spike-like appearance, as depicted by the broken-line curve. The match would now stand out unmistakeably from the blast furnace despite the small contrast.

[2] With the cold bar, extremely hot or cold scenes push the horizontal axis in Figure 5-1 to either end of the available gray scale--possibly beyond.

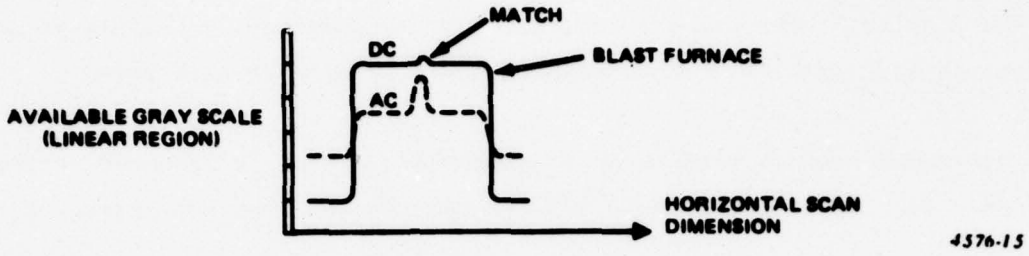


Figure 5-3. IDEAL DC- AND AC-COUPLED INFRARED IMAGING

Actual ac coupling (via an RC filter) cannot achieve quite so perfect a result as the broken-line curve. The blast furnace profile would distort in much the same way as the hot object in Figure 5-1. The superimposed match waveform meanwhile will grow in size and will acquire an exponential tail. In short, typical ac couplings improve pictorial contrast at the cost of skewing the image.

These descriptions can be verified by convolving the solid-line waveform of Figure 5-3 with the RC filter weighting function of Figure 5-4.

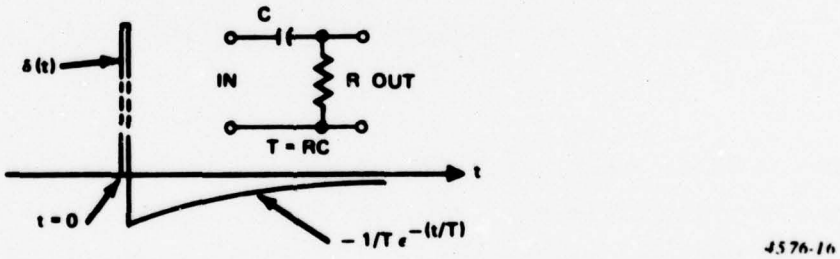


Figure 5-4. WEIGHTING FUNCTION OF AC COUPLING CIRCUIT

The negative area under the exponential portion of the curve exactly balances the unit positive area under $\delta(t)$, indicating that the dc component of any input waveform will be totally blocked. The δ-portion reproduces the input waveform exactly, while the exponential tail introduces drooping. Large T flattens this tail and minimizes the droop while simultaneously lowering the cutoff frequency. The present case assumes sufficiently large T so that droop is minimal within a single-frame interval, and only becomes a problem when viewed over multiple frames. This choice of T, however, minimizes the differentiation properties of the circuit and reduces

the contrast between match and blast furnace. The present study, therefore, seeks to use Haar transforms for producing high-contrast images while removing objectionable droop effects.

A third goal concerns the ability of Haar-aided FLIR's to maintain average signal level at the middle of the display gray scale. Without this ability, sudden changes in average scene brightness can saturate the FLIR. The other extreme, however, leads to complete loss of scene brightness information. This area of study, therefore, aims at a flexible intermediate level of control.

A fourth goal would uniformly stretch or compress the signal waveform to improve gray level use. The broken-line waveform of Figure 5-3, for example, might be vertically stretched (accentuated) until the tip of the match coincided with the uppermost gray level. Such a nonlinear operation, however, would depend upon minimum and maximum values of the display intensity, and finds no solution via Haar transforms. This study, therefore, considers approximate linear solutions suited to Haar transforms.

These four goals,

- Compensating multiscan droop
- Accentuating the match in front of the blast furnace
- Adjusting average signal level to midgray display level
- Adjusting ac signal amplitude in relation to total span of gray scale,

represent the four aspects of the dc restoration problem addressed here.

5.2.2 DC Restoration Solution

5.2.2.1 Multiscan Droop Solution - This section first provides a basic approach to solving this problem. It then describes the impact of stability considerations, as well as a method for maintaining transient (absolute) stability. The section further details the sensitivity of the correction scheme to fluctations in coupling network parameters. It presents the results of a computer simulation of the corrective network in suppressing ac coupling circuit droop effects. Finally, the

section provides a quantitative expression for a basic design parameter, K, of the corrective network. (An intermediate step involves the derivation of subsidiary constant R, which is also used in the computer simulation.)

5.2.2.1.1 Basic Approach - Figure 5-5 characterizes the nominal dc level of a linear scene that remains fixed over scans 1 through 5 and then changes abruptly. Such a linear scene would correspond to the output from one parallel-scan FLIR photodetector. Higher harmonics/Haar sequencies of each scan waveform are assumed to pass through the ac coupling circuit relatively unaffected.[3] Multiscan droop, therefore, appears to impact primarily on the dc levels of successive scans as in Figure 5-5.

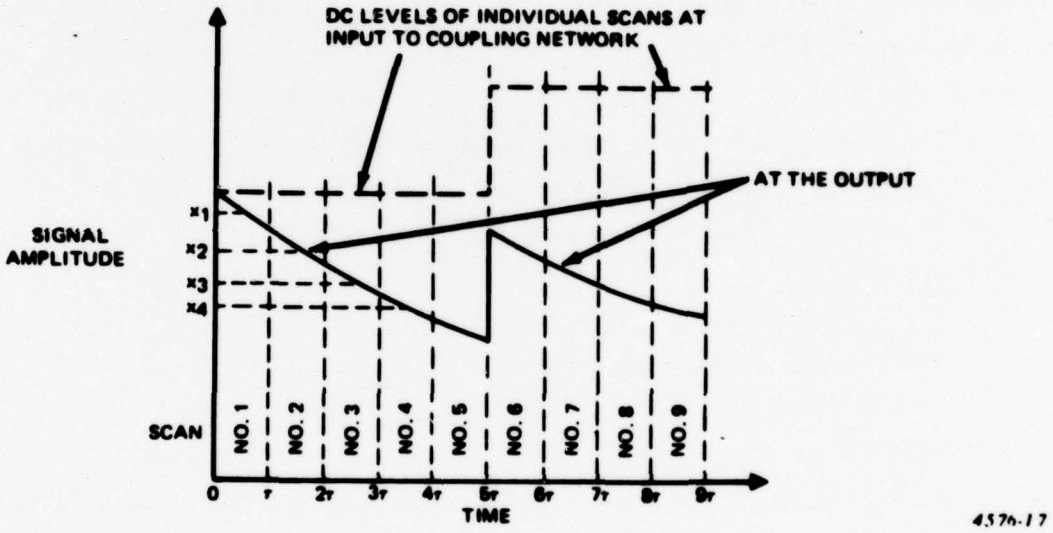


Figure 5-5. MULTISCAN DROOP CHARACTERIZATION

The dc Haar term of scan 1 will assume an amplitude of x_1, scan 2 yields $x_2 = x_1\exp(-t/T)$, scan 3 produces $x_3 = x_1\exp(-2t/T)$, etc., where T = network time constant. Such amplitudes represent sample values of the network response to the input

[3] Scan waveform harmonics suffer relatively little attenuation, particularly when the ac coupling calls for large time constants and low cutoff frequencies. All Haar terms contain some spectral components below cutoff, but the proportion becomes progressively smaller with increasing sequency.

function $x_1 u(t)$; $u(t)$ = unit step. In terms of Laplace transforms, they can be characterized by the transfer function

$$A(s) = \frac{Ts}{1 + Ts}$$

suggesting a corrective network based on

$$B(s) = A^{-1}(s) = 1 + \frac{1}{Ts}$$

$$b(t) = \delta(t) + \frac{u(t)}{T}$$

In terms of the sampled-data model, weighting function b(t) effectively requires that the corrective network (1) take the current value x_n of the distorted waveform, and (2) add to it K • (Σ all previous values, x_{n-1}, x_{n-2} ...).

This corresponds to discrete operator

$$b(z) = 1 + K \sum_{n=1}^{\infty} z^n = 1 + \frac{Kz}{(1 - z)}$$

z ~ Delay of d seconds, equal to the sample time

$K = 1 - e^{-d/T}$ (see paragraph 5.2.2.1.4)

which can be integrated into the general Haar-processed FLIR according to the algorithm of Figure 5-6. The same algorithm can also be applied to the next higher Haar sequencies, but the small increase in picture fidelity may not warrant the added complication. This procedure should effectively counter the lost horizon problem as well as the distorting phenomenon of Figure 5-2.

5.2.2.1.2 Stability Considerations - The preceding discussion has presented the basic corrective procedure for the dc restoration problem. Stability considerations however still remain to be investigated. These include both the absolute stability of the corrective network and its parametric behavior; i.e., its ability to accommodate parameter changes in the original ac coupling.

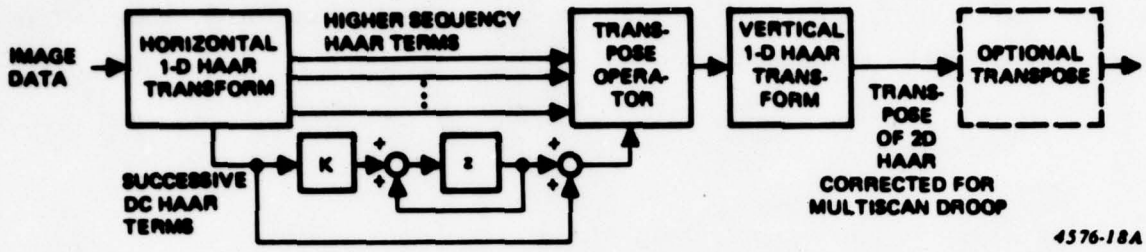


Figure 5-6. MULTISCAN DROOP CORRECTION ANALOG

5.2.2.1.2.1 Absolute Stability - Absolute stability concerns the avoidance of an increasing transient response. The transient response describes the filter reaction to a null input; i.e., zero input for all $t > o$, but with nonzero initial conditions. This appears superimposed on the steady-state response of the filter. Absolutely stable filters generally produce exponentially decreasing transients which quickly reduce to negligible proportions compared to the steady-state response. The increasing transients of unstable filters, on the other hand, often arise from unwanted positive feedback involving active elements within the filter. In the Laplace transform domain, stable and unstable transients correspond to poles in the right (positive) and left (negative) halves of the complex s-plane respectively.

Proposed filter B(s) contains a pole at the origin of the complex s-plane, and represents a potentially unstable condition. One corrective approach using pseudo-inversion was investigated, but does not seem particularly promising.[4] For this reason, an alternate approach was adopted, based on the following pragmatic conditions:

---

[4] Pseudo-inversion replaces the original specification $B(s) = A^{-1}(s)$ by the more reasonable condition

$$\int_{-j\infty}^{j\infty} |B(s)\, A(s) - 1|^2 \, ds \longrightarrow \text{Minimum}$$

This minimum is in turn constrained by the requirement that the noise output of the system

$$\int_{-j\infty}^{j\infty} |B(s)|^2 \, ds = N_o; \; N_o = \text{Input white-noise spectral density}$$

holds to a low fixed value. Analysis based on variational calculus effectively replaces the above pole by two weaker ones in the positive half of the s-plane (unstable) and in the negative half (stable). Synthesis calls for deleting the partial fraction term representing the positive pole, and does indeed produce a stable filter. Its corrective capabilities however seem weak, and the approach was not pursued further.

The object of the high-pass filter in FLIR systems is to emphasize the presence of small targets, especially when they appear against a low-contrast background (this is discussed extensively in paragraph 5.2.2.2 of this report). Since small targets occupy only small portions of the scan interval, they generate relatively high-frequency components in the photodetector output. The filter seeks to emphasize these components by suppressing the low frequencies. It achieves this by rolling off the low-frequency response at 10 dB per decade of frequency decrease below break-frequency f_2. Below a certain frequency f_1, however, progressive attenuation has reached such an extreme state that further roll-off is not needed. A further break-frequency can therefore be inserted at f_1 which flattens the response for all $f < f_1$.

Thus, a conservative form of high-pass filter would possess the transfer function

$$A'(s) = \frac{s+\omega_1}{s+\omega_2}$$

$$\omega_1 = 2\pi f_1$$

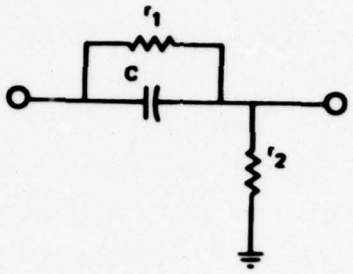
$$\omega_2 = 2\pi f_2$$

Note that A'(s) corresponds to the physically realizable rc network of Figure 5-7. This network differs from the original high-pass filter only in the addition of resistor r_1. It provides a more reasonable model since non-zero r_1 will always shunt C, if for no other reason than the finite Q of the capacitor. The polarized capacitors used in FLIR systems (typically about 33 microfarads) suffer a dissipation specified by K (between 0.01 and 0.1, depending upon temperature) in the relation

$$I = KCV$$

Thus, KC specifies the equivalent parallel conductance and leads to the result

$$\omega_1 = K \sim 0.063$$



$$\omega_1 = (r_1 C)^{-1}$$

$$\omega_2 = \left(\frac{r_1 r_2}{r_1 + r_2} \cdot C\right)^{-1} = \left(\frac{r_1 + r_2}{r_2}\right) \omega_1$$

4697-29

Figure 5-7. CONSERVATIVE REPRESENTATION OF PARALLEL-SCAN FLIR HIGH-PASS FILTER

A typical value for ω_2 meanwhile follows from the usual choice of f_2 (2 to 5 Hz) in FLIR systems. Thus, for $f_2 = 4$

$$\omega_2 = 25.133$$

The transfer function of the corrective network now changes to

$$B'(s) = (A'(s))^{-1}$$

$$= \frac{s+\omega_2}{s+\omega_1}$$

This expression reveals no poles in the right half of the s-plane; i.e., absolute stability. Via inverse Laplace transformation, B'(s) converts to the weighting function

$$b'(t) = \delta(t) + (\omega_2 - \omega_1)\, \epsilon^{-\omega_1 t}$$

which in turn leads to the z-transform operator

$$b'(z) = 1 + \left[1 - \epsilon^{-d(\omega_2 - \omega_1)}\right] \sum_{n=1}^{\infty} \epsilon^{\left[-\omega_1 dn\right]} z^n$$

See paragraph 5.2.2.1.5 for determination of constant multiplying Σ. Figure 5-8 gives the corresponding algorithm for the Haar-transform processor.

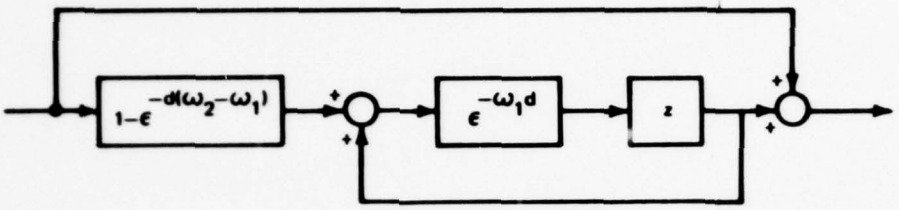


Figure 5-8. PROCESSOR ALGORITHM FOR ABSOLUTELY STABLE CORRECTIVE NETWORK

5.2.2.1.2.2 Parametric Behavior - Assuming absolute stability, potential problems can still arise if the corrective network uses inaccurate numerical values for ω_1 and ω_2. Such inaccuracies might be caused by temperature variations in the ac coupling or by any other form of parameter drift. In particular, suppose that

$$\hat{A}(s) = \frac{S+\omega_1 + \Delta\omega_1}{S+\omega_2 + \Delta\omega_2}$$

$$\hat{B}(s) = \frac{S+\omega_2}{S+\omega_1}$$

The impulse response of the tandem combination changes from an ideal value of $\delta(t)$ (i.e., perfect correction) to

$$\delta(t) = \Delta\omega_1 \epsilon^{-\omega_1 t} - \Delta\omega_2 \epsilon^{-\omega_2 t}$$

Parameter drift may therefore introduce some performance degradation but it will not produce instability. This degradation can be expressed in terms of the mean square transient error

$$M = \frac{1}{2\pi i} \int_{-i2\pi W}^{i2\pi W} \left| \hat{A}(s)\, \hat{B}(s) - 1 \right|^2 \left| E(s) \right|^2 ds$$

W = Signal bandwidth in Hz
E(s) = Laplace transform of input signal

To simplify calculation, take E(s)=1, indicating a flat power spectrum of unit strength inside the (0,ω) Hz bandwidth, and zero strength outside. Substitution of the preceding expressions for A" and B' then shows that

$$M = \frac{(\Delta\omega_1)^2}{\omega_1} \frac{\tan^{-1}\left(\frac{2\pi W}{\omega_1}\right)}{\pi} + \frac{(\Delta\omega_2)^2}{\omega_2} \frac{\tan^{-1}\left(\frac{2\pi W}{\omega_2}\right)}{\pi}$$

This expression attains a maximum value of

$$\frac{1}{2}\left(\frac{(\Delta\omega_1)^2}{\omega_1} + \frac{(\Delta\omega_2)^2}{\omega_2}\right)$$

for arbitrarily large W.[5]

Thus, for $\omega_1 = 0.063$, $\omega_2 = 25.133$, $\Delta\omega_1 = \omega_1/10$ and $\Delta\omega_2 = \omega_2/10$,

an input power of

$$(1 \text{ watt/Hz}) \left(\frac{\omega_2-\omega_1}{2\pi}\right) = 3.99 \text{ watts}$$

produces a waveform distortion component with a power of 0.126 watts. This suggests signal/distortion ratios as high as 15 dB even for a 10 percent error in network design parameters.

5.2.2.1.3 Simulation of Correction Algorithm - A computer simulation of the original corrective technique based on b(z) has been successfully applied. In the following examples, sample time d (A in the computer printout) has been equated to the interval between successive computations of Haar terms for the same 16-by-16 pel block:

$$(30 \text{ frames/sec})^{-1} = 0.03333 \text{ sec}$$

[5] Actually, $W = (\omega_2-\omega_1)/2\pi$ in the expression for M. The limiting expression above however gives a clearer indication of the effects of parameter drift.

This figure is typical of parallel-scan FLIRs. These examples also employ a nominal time constant of

$$T = \frac{1}{2\pi (4\ \text{Hz})} = 0.0398\ \text{sec}$$

Example 1 in Table 5-1 assumes perfect agreement between the time constant of the high-pass filter and the corresponding K-value of the corrective network. Column X represents a step-function type waveform input to the high-pass filter from the FLIR photodetector; Y represents the filter output and input to the corrective network; Z gives the output of the corrective network. Note that Z represents a perfect restoration of X.[6]

Example 2 repeats the same general X-Y-Z scenario, but assumes that the high-pass filter time constant has drifted by 1 percent from its nominal value. Example 3 assumes 5 percent drift and example 4 assumes 10 percent.

The computer program permits the simulation of five examples per run, though only four have been used here. Note that corrected output Z shows progressively poorer fidelity for increasing mismatch in T. For 10 percent mismatch (T=0.00438), a typical error of 6.43 percent over the bandwidth between ω_1 and ω_2 (3.99 Hz) translates to an equivalent M-value of $(0.0643)^2 \times (3.99) = 0.0165$, or about 18 dB signal/distortion ratio. This approximate calculation seems in rough agreement with the theoretically derived 15 dB figure cited earlier.

5.2.2.1.4 Determination of Constant K - This determination first characterizes the sampled-data equivalent of the RC high-pass filter A(s). It then examines the manner in which the train of ouput samples might drive a corrective network. In particular, it looks for the particular value of K which allows the corrective network to recreate the original sampled data input to A(s).

[6] Because of the finite time interval over which X is specified and the exclusion of noise, questions of stability do not arise.

Table 5-1. COMPUTER SIMULATIONS FOR MATCHED AND MISMATCHED AC COUPLING CORRECTIVE NETWORKS

```
ENTER CORRECTIVE TIME INTERVAL -A- AND TIME CONSTANT -T-.
 NOTE  IF A=0, JOB RESTARTED.  IF A+T=0, JOB TERMINATED
?
.0333 .0398

ENTER SAMPLING TIME INTERVAL -A- AND FIVE VARIATIONS FOR -T-.
 NOTE: IF A=0, RUN RESTARTED.  T=0 TERMINATES SERIES
?
.0333 .0398 .04020 .04179 .04378 0

WAVEFORM INPUT -X-, FILTERED OUTPUT -Y-, CORRECTED OUTPUT -Z-
```

		0.0398		0.0402		0.0418		0.0438	
I	X	Y	Z	Y	Z	Y	Z	Y	Z
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	1.0000	0.4331	1.0000	0.4368	1.0036	0.4508	1.0176	0.4674	1.0342
6	1.0000	0.1876	1.0000	0.1908	1.0052	0.2032	1.0255	0.2184	1.0502
7	1.0000	0.0813	1.0000	0.0833	1.0059	0.0916	1.0291	0.1021	1.0577
8	1.0000	0.0352	1.0000	0.0364	1.0062	0.0413	1.0307	0.0477	1.0612
9	1.0000	0.0152	1.0000	0.0159	1.0063	0.0186	1.0315	0.0223	1.0628
10	1.0000	0.0066	1.0000	0.0069	1.0064	0.0084	1.0318	0.0104	1.0636
11	1.0000	0.0029	1.0000	0.0030	1.0064	0.0038	1.0319	0.0047	1.0640
12	1.0000	0.0012	1.0000	0.0013	1.0064	0.0017	1.0320	0.0023	1.0641
13	1.0000	0.0005	1.0000	0.0006	1.0064	0.0008	1.0320	0.0011	1.0642
14	1.0000	0.0002	1.0000	0.0003	1.0064	0.0003	1.0320	0.0005	1.0642
15	0.0	-0.9999	0.0000	0.9999	0.0064	-0.9998	0.0320	-0.9998	0.0643
16	0.0	-0.4331	0.0000	0.4367	0.0028	-0.4507	0.0144	-0.4673	0.0300
17	0.0	-0.1876	0.0000	-0.1907	0.0012	-0.2031	0.0065	-0.2184	0.0140
18	0.0	-0.0813	0.0000	-0.0833	0.0005	-0.0916	0.0029	-0.1021	0.0066
19	0.0	-0.0352	0.0000	-0.0364	0.0002	-0.0413	0.0013	-0.0477	0.0031
20	0.0	-0.0152	0.0000	-0.0159	0.0001	-0.0186	0.0006	-0.0223	0.0014
21	0.0	-0.0066	0.0000	-0.0069	0.0000	-0.0084	0.0003	-0.0104	0.0007
22	1.0000	0.9971	1.0000	0.9970	1.0000	0.9962	1.0001	0.9951	1.0003
23	1.0000	0.4319	1.0000	0.4354	1.0036	0.4490	1.0177	0.4651	1.0344
24	1.0000	0.1871	1.0000	0.1902	1.0052	0.2024	1.0256	0.2174	1.0503
25	1.0000	0.0810	1.0000	0.0831	1.0059	0.0912	1.0291	0.1015	1.0577
26	1.0000	0.0351	1.0000	0.0363	1.0062	0.0411	1.0307	0.0475	1.0612
27	1.0000	0.0152	1.0000	0.0158	1.0063	0.0185	1.0315	0.0222	1.0628
28	1.0000	0.0066	1.0000	0.0069	1.0064	0.0084	1.0318	0.0101	1.0636
29	1.0000	0.0029	1.0000	0.0030	1.0064	0.0038	1.0319	0.0048	1.0640
30	1.0000	0.0012	1.0000	0.0013	1.0064	0.0017	1.0320	0.0023	1.0641

To this end, consider the inverse Laplace transform of A(s); i.e., the impulse response of the high-pass filter:

$$A(s) = 1 - \frac{1}{1+Ts} \tag{5-1}$$

$$a(t) = \delta(t) - (1/T)\,\epsilon^{-t/T} \tag{5-2}$$

Expression a(t) indicates that the sampled filter output y_n should represent the sum of the present input sample x_n and a scaled, exponentially attenuated train of previous input samples x_{n-1}, x_{n-2}, . . . Specifically,

$$y_n = x_n - R\left(x_{n-1}\epsilon^{-d/T} + x_{n-2}\epsilon^{-2d/T} + \ldots\right) \tag{5-3}$$

constant $R \neq 1/T$, as (5-2) might seem to suggest. It does, however, approach (d/T) as d approaches zero.

To determine the correct value of R, take all $x_m = 1$, sum the resulting geometric series within the brackets and obtain

$$y_n = 1 - R\left(\frac{\epsilon^{-d/T} - \epsilon^{d/T}}{1-\epsilon^{-d/T}}\right) \tag{5-4}$$

Since Y_n approaches zero as n becomes arbitrarily large,

$$R = \epsilon^{d/T} - 1 \tag{5-5}$$

Substitution of this value into the general expression yields

$$y_n = \epsilon^{-(n-1)\,d/T} \tag{5-6}$$

These samples constitute the driving function for a corrective network characterized by

$$b(z) = 1 + K\sum_{n=1}^{\infty} z^n \tag{5-7}$$

Network output corresponds to

$$w_r = y_r + K\,(y_{r-1} + y_{r-2} + y_{r-3} + \ldots + y_1) \tag{5-8}$$

Replacing the bracketed y's by their exponential equivalents in (5-6) and summing the resulting geometric series,

$$w_r = \bar{\epsilon}^{(r-1)d/T}\left(1 + \frac{K}{\bar{\epsilon}^{d/T_{-1}}}\right) - \frac{K}{\bar{\epsilon}^{d/T_{-1}}} \tag{5-9}$$

The last fraction in 5-9 represents w_r for arbitrarily large r, and should equal $x_r = 1$ for perfect correction. This yields

$$K = 1 - \bar{\epsilon}^{d/T}$$

Note that substitution of this value into (5-9) renders $w_r = 1$ not only for large values of r but for all values.

5.2.2.1.5 Determination of Constant in b' (z) - The same rationale which led to (5-3) also applies to the absolutely stable network characterized by b' (z). Taking y_n and w_n as the input and output sequences of sampled values,

$$w_n = y_n + Q\,(\alpha y_{n-1} + \alpha^2 y_{n-2} + \ldots + \alpha^{n-1} y_1) \tag{5-10}$$

where

$$\gamma = \epsilon^{-\omega_1 d} \tag{5-11}$$

and Q is the constant to be determined. From (5-6)

$$y_n = \bar{\epsilon}^{(n-1)\,d/T} = p^{n-1} \tag{5-12}$$

$$p = \bar{\epsilon}^{d/T} \tag{5-13}$$

Substituting (5-12) into (5-10) and summing the geometric series within the bracketed term,

$$w_n = p^{n-1}\left(1 + \frac{Q\gamma}{p-\gamma}\right) - \frac{Q\gamma^n}{p-\gamma} \qquad (5\text{-}14)$$

This expression gives a corrective network output which attempts to restore the train of unit pulses which characterized the original input to the high-pass filter. To the extent that α is slightly less than unity, it can only approximate this result. Note, however, that, with γ temporarily taken equal to unity, the first term in (5-14) varies with n while the second term gives the desired constant condition over all n. Thus, W_n = constant if the coefficient of p^{n-1} is zero; i.e.,

$$Q = 1 - (p/\gamma) = 1 - \epsilon^{-d(\omega_2 - \omega_1)} \qquad (5\text{-}15)$$

This yields the coefficient used in the expression for b' (z).

5.2.2.2 Match - Blast Furnace Contrast Solution - Large time constants tend to suppress this type of contrast. Assuming the droop problem to be adequately corrected by the algorithm in Figure 5-6, contrast can be reinstated by boosting the values of the higher sequency Haar coefficients. Such a process may in many ways resemble a new ac coupling with reduced T, but it does not reintroduce droop.

One algorithm that achieves this goal would increase all Haar coefficients of the droop-corrected image in proportion to their sequencies. Thus, a 2D image:

$$I(x,y) = \sum_{r,m}\sum_{p,n} a_{rm}a_{pn}h(r,m,x)h(p,n,y)$$

$$\sum_{r,m} = \sum_{r=-1}^{r_{max}}\sum_{m=1}^{2^r}, \quad \sum_{p,n} = \sum_{p=-1}^{p_{max}}\sum_{n=1}^{2^p}$$

would convert to

$$J(x,y) = C\sum_{r,m}\sum_{p,n}(r+1)(p+1)\,a_{rm}a_{pn}h(r,m,x)h(p,n,y)$$

where C = any normalization constant consistent with the limited display gray scale. Note that h(-1,1,x) represents the dc term. Other algorithms might replace (r+1)(p+1) by $(r+1)^s$ $(p+1)^t$ where the exponents might be empirically derived for the class of scenes under consideration.

5.2.2.3 Gray-Scale Centering Solution - Perfect centering can be achieved by replacing dc coefficient $a_{-1,1,-1,1}$ by whatever value corresponds to the middle of the display gray scale. This process, however, would disregard any significant changes in scene brightness. In removing this difficulty, the algorithm of Figure 5-9 converts $a_{-1,1,-1,1}$ to (midgray value) + m · $a^2_{-1,1,-1,1}$ where m = some small positive constant.

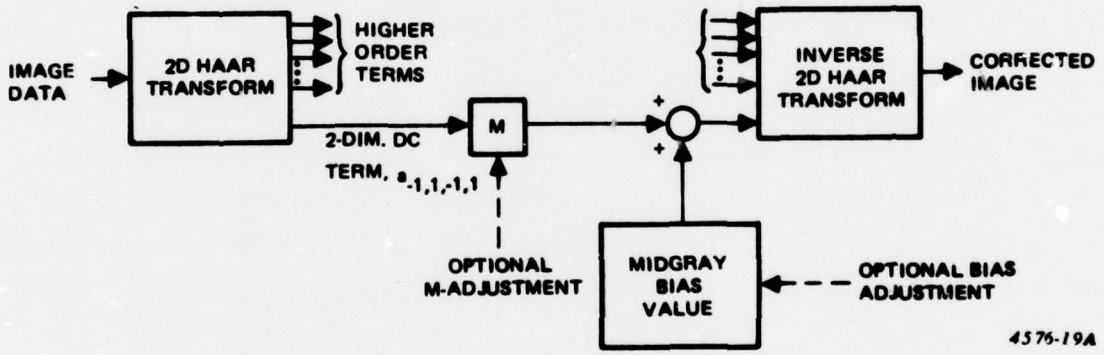


Figure 5-9. GRAY-SCALE CENTERING ALGORITHM

This scheme, however, would respond to the brightness levels of individual 16-by-16 pel blocks within the scene, without any regard to smoothing the transition from one block to its neighbors. Such discontinuities result in the appearance of artifacts.

To eliminate these artifacts, an improved correction scheme manipulates only the average value of all individual-block brightness levels. For each block, the deviation between individual brightness level and the overall average is preserved as a fixed value. Manipulations of the average value then will change overall

scene brightness, but will not affect the relative variation of brightness from block to block. The result will be a brightness control scheme free of artifacts. Figure 5-10 outlines the basic process. Successive dc Haar terms are stored in the left-hand shift register just long enough to reduce each one by their common average value. Each processed term is then increased by a voltage corresponding to mid-gray, with a small fraction (m) of the average superimposed. Read-out to the inverse 2D Haar transformer proceeds through a second shift register which preserves the original order in which the dc terms were read in. The higher order terms, meanwhile, undergo a compensating delay; those terms corresponding to a particular 16-by-16 pel block arrive at the inverse transformer at the same time as their associated $a_{-1,1,-1,1}$ term.

5.2.2.4 AC AMPLITUDE STRETCHING SOLUTION

Amplitude stretching, or contrast control, is a basically nonlinear process. It adjusts the ac gain of image waveform g(x,y) such that

$$\left|g(x,y)-g(\text{average})\right|_{max} = \text{Half-span of display gray scale}$$

Certain approximations, however, can linearize this process, adapting it to Haar-transform processing. The rationale for these approximations follows from Figure 5-11. This figure assumes a normal distribution of intensity deviations about mid-gray.[7] As Δx increases, a larger total interval of gray scale deviations will be covered, but the average value associated with each gray level tends to decrease, (i.e., each gray level zone tends to slide down the Gaussian curve to one side or the other, with increasing x). Quantitatively expressed,

$$\text{Expected Information of Display} = D = \sum_{n=o}^{N} \epsilon^{-\left(\frac{\Delta x}{2} + n\Delta x\right)^2 / 2\sigma^2}_{\Delta x}$$

σ = Standard Deviation

[7] Intensity deviations of a given scene may or may not follow a normal distribution. Averaged over the many scenes that might apply to a given scenario, however, the central limit theorem suggests a progressive approach to the normal model.

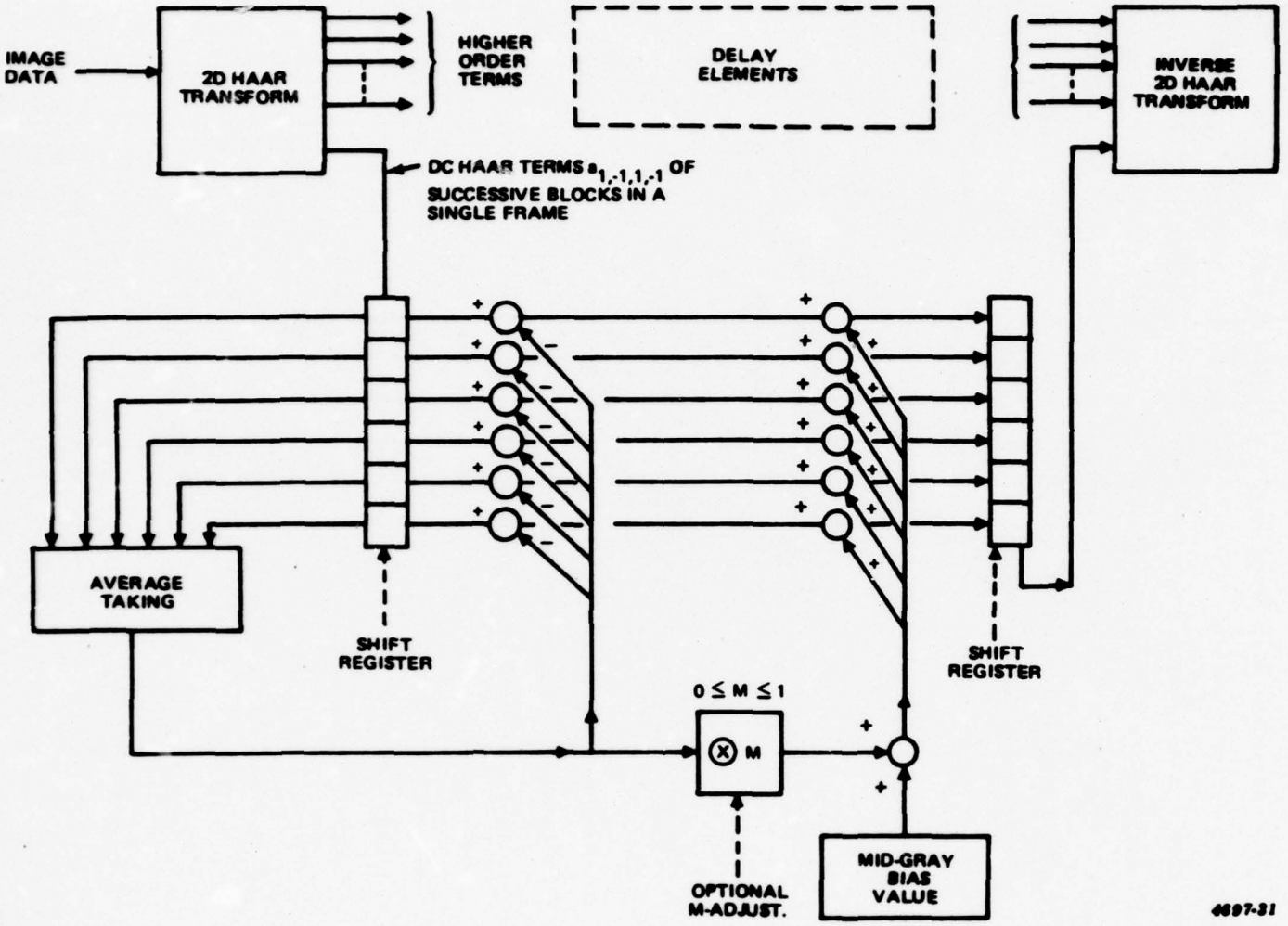


Figure 5-10. NON-ARTIFACTING BRIGHTNESS CONTROL SCHEME

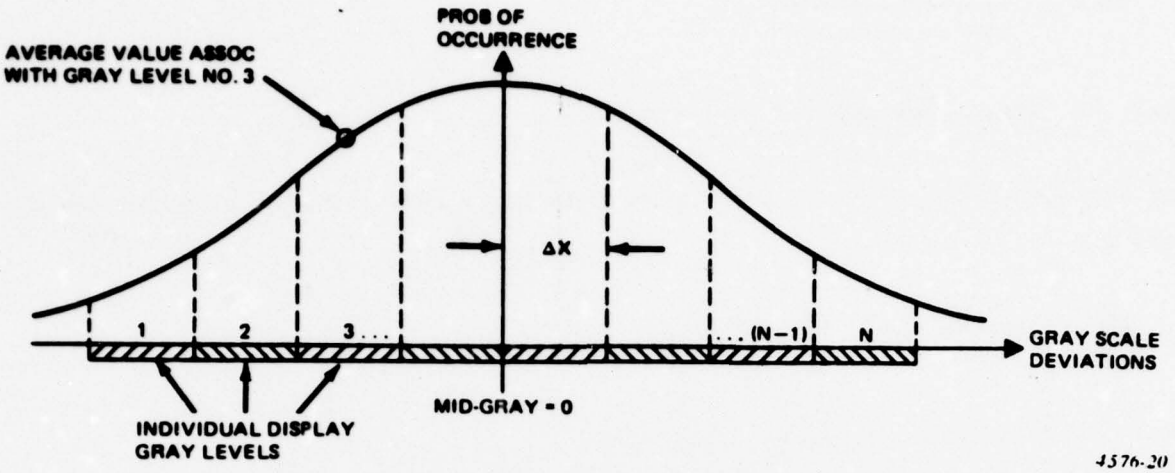


Figure 5-11. CONTRAST LINEARIZATION ANALYTICAL MODEL

will reach a maximum for some appropriate value of x. Approximate calculations show that

$$N\Delta x \text{ (Optimum Gray-Scale Span)} \doteq \sigma\sqrt{2}$$

particularly for N greater than about 10. In effect, gain adjustments should not force the gray levels to cover the total span of fluctuations but only about 1.5 times the standard deviation of all the expected scenes. Fluctuations that fall below level 1 or above N may be lost because of saturation, but this sacrifice enhances gray-level resolution and maximizes overall display information.

Following is the approximate calculation used to derive the optimizing condition:

$$D \doteq \sum_{n=o}^{N} \left[1 - \frac{n^2(\Delta x)^2}{2\sigma^2}\right] \Delta x$$

$$\frac{dD}{d(\Delta x)} \doteq \sum_{n=o}^{N} \left(1 - \frac{3n^2}{2\sigma^2}(\Delta x^2)\right) = \left(\sum_{n=o}^{\infty} 1\right) - \frac{3(\Delta x)^2}{2\sigma^2}\left(\sum_{n=o}^{\infty}\right) n^2$$

$$= N - \frac{3(\Delta x)^2}{2\sigma^2}\ \frac{N^3}{3} + \frac{N^2}{2} + \frac{N}{6} = 0$$

$$N\Delta x = \frac{2\sigma N}{\sqrt{2N^2 + 3N + 1}} \rightarrow \sigma\sqrt{2}$$

where the limit value applies to large N.

Haar implementation of these principles associates activity with standard deviation, i.e.,

$$\sigma^2 = \text{Activity} = \sum_{r,m} \sum_{p,n} a^2_{rmpn}$$

Figure 5-12 outlines the general algorithm

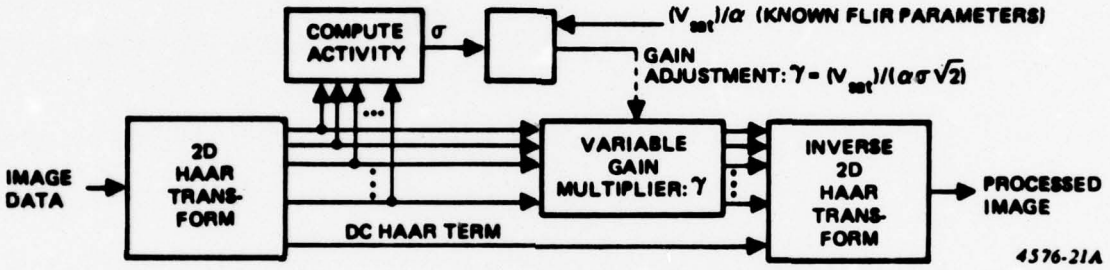


Figure 5-12. HAAR CONTRAST CONTROL

This algorithm takes V_{sat} as the voltage span between upper and lower saturation limits. Thus, with N gray levels in the display, $(V_{sat})/N$ represents the incremental voltage needed to move from one gray level to the next. This voltage should produce a display intensity change of Δx, indicating a required volts-per-unit-intensity-change conversion gain of

$$\frac{V_{sat}}{N\Delta x} = \frac{V_{sat}}{\sigma\sqrt{2}}$$

In actuality, however, the conversion gain for a given FLIR display may correspond to some fixed parameter α (volts per unit intensity change). A variable gain (adaptive gain) multiplier γ is therefore needed to make up the discrepancy:

$$\gamma\alpha = \frac{V_{sat}}{\sigma\sqrt{2}}$$

$$\gamma = \frac{V_{sat}}{\alpha\sigma\sqrt{2}}$$

Note that relatively active scenes with large intensity fluctuations in gray shades correspond to large σ which in turn invoke small γ. In this instance, adaptive Haar Contrast Control serves as a gray-shade compressor. Similar reasoning for quiet scenes (low σ) show that the algorithm then effectively expands intensity fluctuations.

As in the case of automatic Haar brightness control, individual contrast adjustment for each 16-by-16 pel block will tend to create artifacts. This can be avoided by using the rms value of σ (computed over the entire frame) in the proposed algorithm.

5.3 DYNAMIC RANGE COMPRESSION

5.3.1 Problem Definition

On a somewhat different tack, dynamic range compression[8] addresses the problem of inefficient use of the available display gray levels. It stretches each small amplitude interval of the waveform in proportion to the occupancy of that interval. A long interval of small waveform variations about a given mean value will be accentuated. Large spikes, on the other hand, contribute little to the occupancy of any level and will tend to shrink.

True dynamic range compression, however, is a nonlinear process. Taking g(x) and h(x) as the initial and final waveforms,

$$\text{Required Small-Signal Gain} = \frac{dh}{dg}$$

varies as a function of g rather than x and cannot be readily implemented in terms of Haar transforms.[9] A further goal of adaptive signal processing, consequently, consists in replacing this process by an approximation suited to the Haar domain.

---

[8] This process has also been designated Histogram Equalization in the literature.

[9] Consider x_n = roots of $g(x) = g_o$ = some specified amplitude. Then,

$$(dh/dg)_{g_o} \sim \sum_{\text{all } x_n} (dg/dx)^{-1}_{x_n} = \text{function of } g_o.$$

5.3.2 Dynamic Range Compression Solution

Haar implementation requires that the original (nonlinear) description of dynamic range compression, "A small-signal gain assigned to every amplitude which varies in proportion to the occupancy of that level," be replaced with, "A gain value assigned to each sequency and proportional to the occupancy of that sequency."

Sequency occupancy is defined as $\sum_{m=1}^{2^r} a_{mr}$.[10]

In the one-dimensional case, both descriptions generally promote the amplification of long intervals of low-amplitude fluctuations about some constantmean. Both also tend to deemphasize sharp spikes.

Figure 5-13 illustrates a case in point. Original waveform g(x) occupies various amplitude levels according to the numbers to the left of the g-axis. In the usual nonlinear approach to dynamic range compression, these numbers specify slopes (small-signal gains) required of the transducer, as shown. Transducer output appears as h(x).

The linear equivalent output meanwhile appears as j(x). Construction of this waveform involves derivation of the Haar components of g(x) as listed in Table 5-2. Multiple-entry sequencies arrange terms in order of increasing offset. Each component is then amplified by the gain associated with its sequency, also as in Table 5-2.

These gains are obtained by multiplying all terms of given sequency r by $2^{(r/2)}$[11], and adding the absolute values of the resulting numbers.[12] The result is then processed through an inverse Haar transform to yield j(x) as shown in Figure 5-13.

---

[10] Other candidate definitions include $\sum_{m=1}^{2^r} a_{mr}^2$, $\sum_{m=1}^{2^r} a_{mr}^4$ etc., but these tend to minimize the roles played by the smaller a's too severely.

[11] An exception is the (-1) or dc Haar term. This term does not alter waveform shape, but only regulates its average value (brightness). Numerical calculations simplify when this particular gain is chosen equal to unity.

[12] Such a procedure actually yields the numbers 31, 21.21, 30, and 42.44 for sequencies 0, 1, 2, and 3, respectively. For simplicity, the numbers have been divided by 10 and rounded to the nesreat integer to yield the values in Table 5-2.

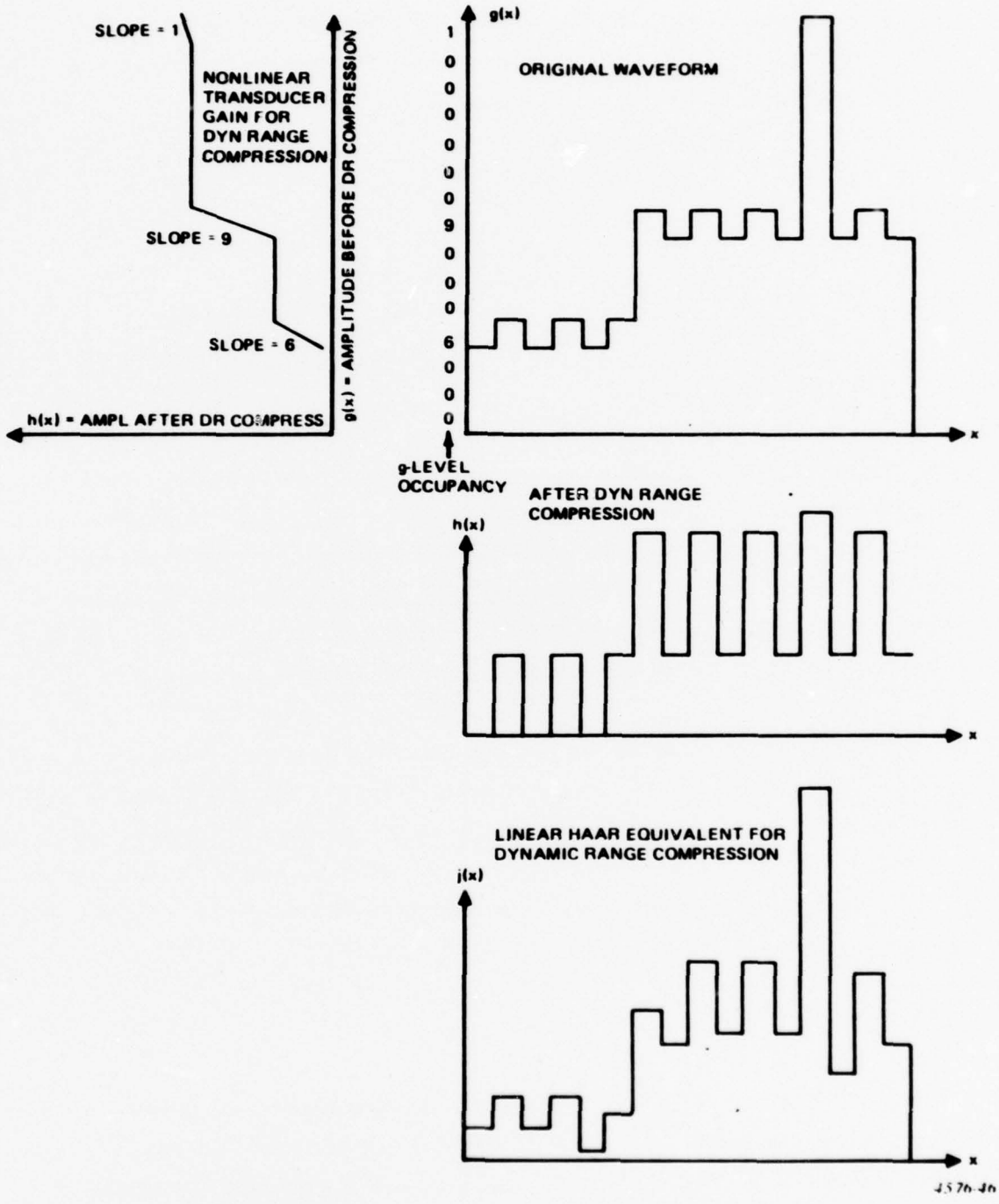


Figure 5-13. DYNAMIC RANGE COMPRESSION EXAMPLE

Table 5-2. RATIONALIZED HAAR COMPONENTS OF g(x)

Gain	Sequency, r	Component Strengths							
1	-1	103							
3	0	-31							
2	1	8	-7						
3	2	0	-8	0	7				
4	3	-1	-1	-1	1	1	1	8	1

Average values of the ripple-groups are no longer constants, but seem to wander somewhat. This, however, does not seem particularly objectionable. Note, too, that the jump between the ripple groups has dropped from four times the ripple amplitude to only three times. Likewise, the single spike has shrunk from eight times the amplitude of the surrounding ripples to something less than six times. Finally the average ripple amplitude has increased from (0.08) x ($g_{max} = g_{min}$) to (0.11) x ($h_{max} - h_{min}$). All three changes represent dynamic range compression.

Graph j(x) has not compressed the waveform as much as h(x) and does not, therefore, use the gray scale quite as efficiently. Note, however, that h(x) barely yields a discernible spike. This suggests that, beyond a certain point, compression may degrade intelligibility. The j(x) (Figure 5-13) could nevertheless benefit from further compression; various modifications of the linear compression algorithm are possible.

5.4 CONCLUSIONS

Haar transforms can serve as an effective tool in all aspects of adaptive signal processing. Under the subheading of dc restoration, the best procedure seemed to be the identification of four main problem areas and a separate corrective Haar algorithm for each one. In applying these algorithms, the feasibility of converting a nonlinear aspect of dc restoration in an equivalent linear one is also confirmed.

Dynamic range compression represents a still more complex problem involving non-linear operations in the image domain. Here, equivalent Haar domain operations had to proceed via an analogy between gray-level occupancy and sequency occupancy. Results show obvious imperfections but also suggest means for further improvement.

SECTION 6
BANDWIDTH COMPRESSION

6.1 INTRODUCTION

In sampling TV rate imaging sensors (512 lines, 30 frames per second) at the Nyquist rate (8 megasamples/sec), using eight bits to encode the image intensity (256 shades of gray), a data rate of 64 megabits/sec is required. Although difficult to achieve, processing at this high data rate in real time is now practical. This has been demonstrated by the operations performed using the Northrop 2-D Haar processor in fulfilling this study contract. Unfortunately, magnetic recording of digital images at this data rate is not practical at this time. Since data recording has many advantages over analog video recording, it would be highly desirable to reduce the data bandwidth by such an order of magnitude that small, low-cost recording devices could be used in conjunction with the airborne sensor. This would be particularly significant in a "pop-up" mode where a vulnerable vehicle such as a helicopter can pop-up, snatch several frames of video and rapidly remask while the data is processed and analyzed. A reduction in data bandwidth also greatly reduces the difficulty of transmitting the data to a remote location by means of a radio link if this becomes part of an operational requirement.

This study is an evaluation of several techniques that reduce data rates by removing redundancy, comparing them with simpler filtering techniques that reduce resolution. By using the Haar transform and a combination of "thresholding" and the "N coefficient" algorithm, a data rate of one bit per pel (bandwidth reduction of 8 to 1) is easily attainable. This does not take into consideration other techniques,such as frame rate reduction, change detection recording, companding, and source encoding schemes, which, when incorporated into the system, allow for even greater compression ratios.

6.2 GENERAL EXPERIMENTAL PROCEDURE

The experimental test configuration is shown in Figure 6-1. The video received from the TV camera is sampled at 8.19 MHz and then digitized to eight bits (256

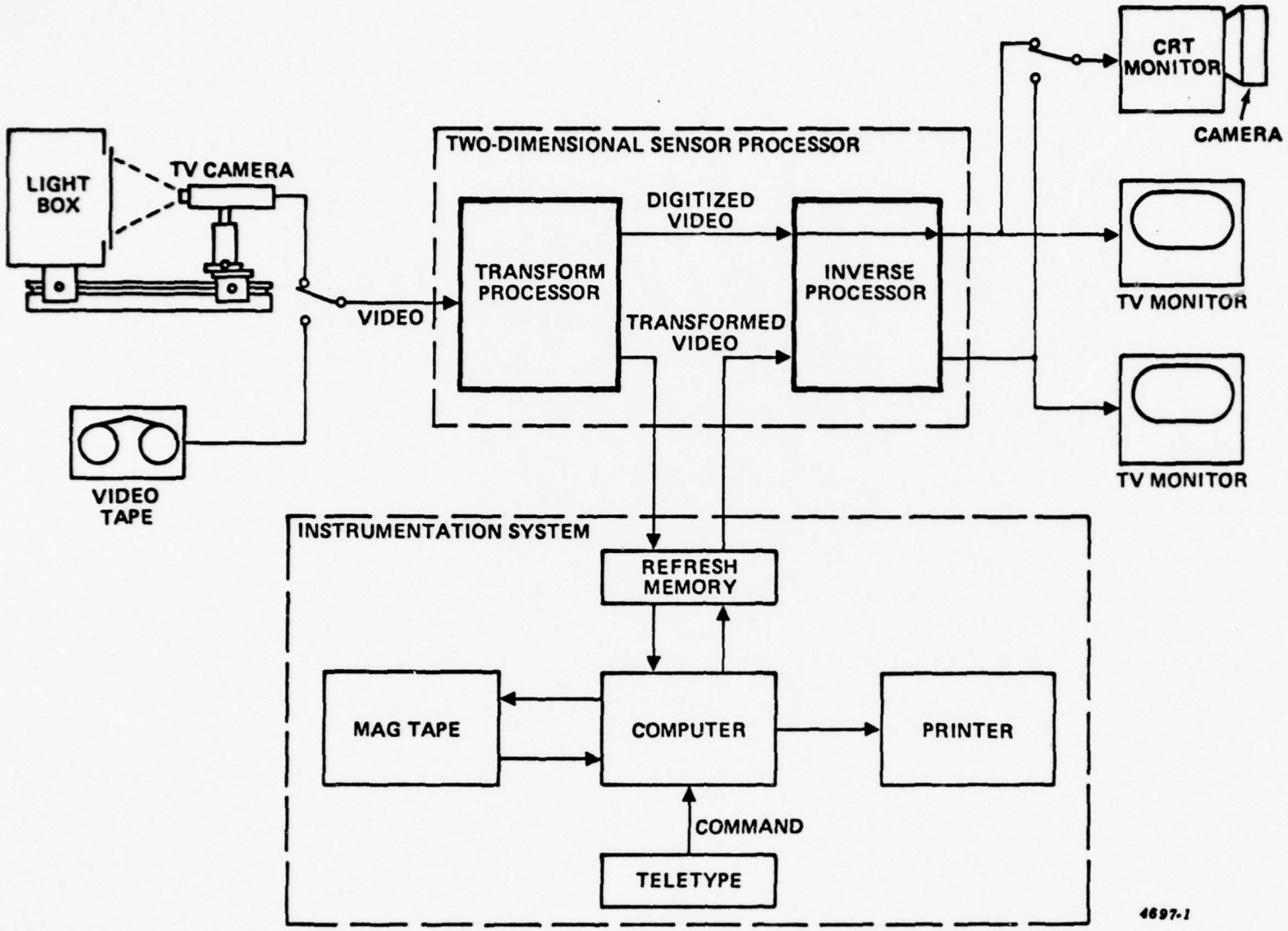


Figure 6-1. EXPERIMENTAL TEST CONFIGURATION

gray shades) on a linear scale. This digital image is used as the non-compressed image and will be referred to as the "full-resolution" image. It exhibits an information content of 8 bits/pel. The data rates determined as a result of the application of experimental compression algorithms are referenced to the full-resolution image. For example, after compression the total number of bits required to represent the image is divided by the number of pels in the full-resolution image. The full-resolution image is subdivided into 16x16 pel sectors in the transform processor. The Haar transform of each section is then computed and transmitted to the inverse processor where the digital image is reconstructed. The transform format of the processor is illustrated in Figure 6-2. Image data compression algorithms may be applied either to the Haar transform of the image sector or, in some cases, to the image section directly.

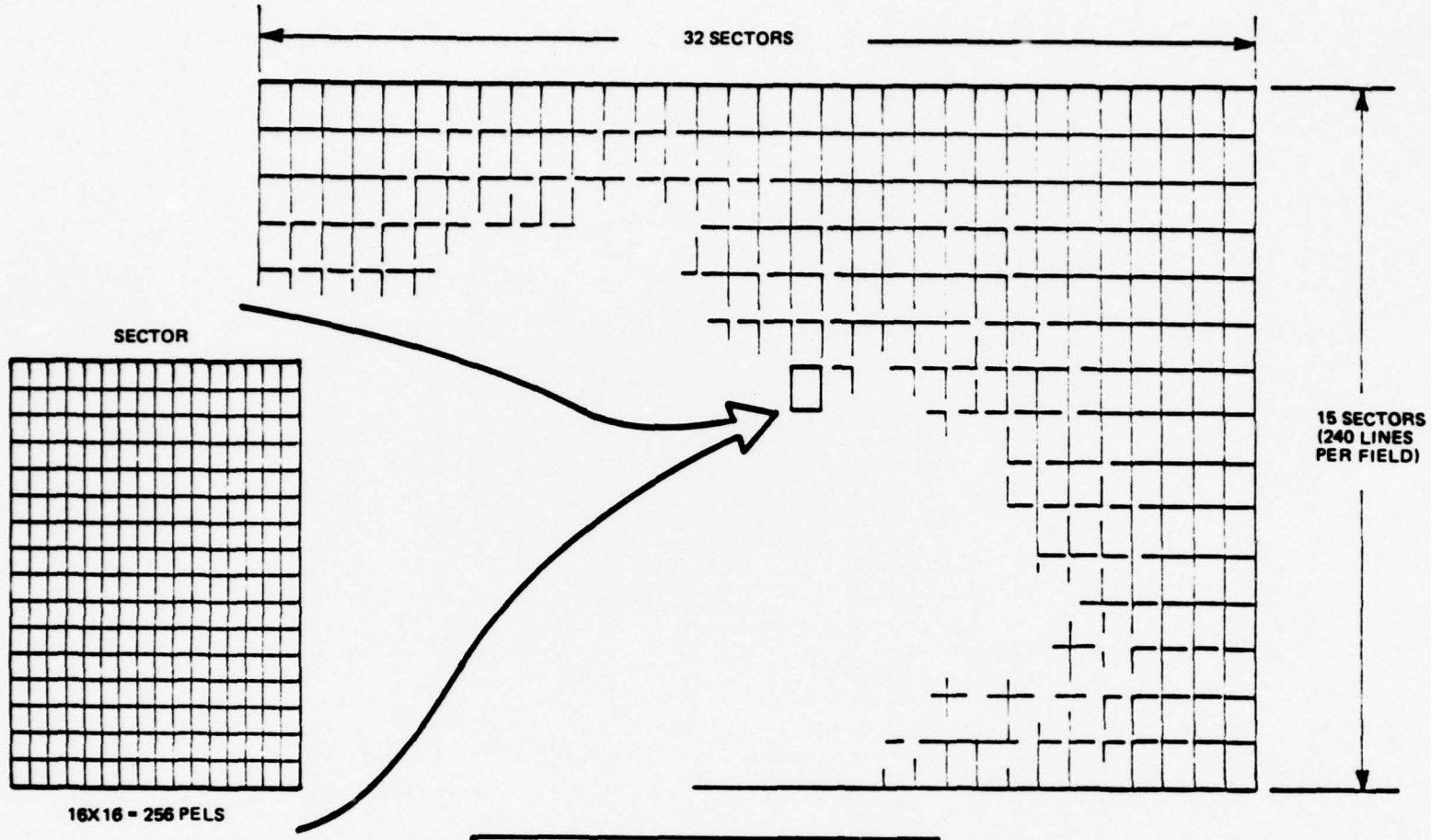
Two images, an Electronic Industries Association (EIA) resolution chart, and a FLIR thermogram, were used as test imagery for the initial experiments. The resolution chart provides spatial resolution information while the FLIR image provides a more realistic gray-scale distribution.

6.3 REDUCED RESOLUTION COMPRESSION

The information content of a sampled digital image may be compressed directly by reducing either the number of image samples or the number of shades of gray used to represent each sample. This approach to bandwidth reduction is referred to as reduced resolution compression. Reduced resolution compression is the simplest form of bandwidth reduction and provides a baseline for the evaluation of more involved image-compression schemes.

6.3.1 Scan Line Sampling

In a scanned imaging system, such as TV or a parallel-scan FLIR, the number of image samples may be reduced on a line-by-line basis. This approach sacrifices spatial resolution in only the scan direction. The results of one dimensional sample reduction are shown in Figure 6-3, where the images shown were made by reducing the clock frequency to the output register of the inverse Haar processor. As is shown in the figure, the horizontal resolution of the system decays fairly rapidly starting with the four bit/pel compression rate.

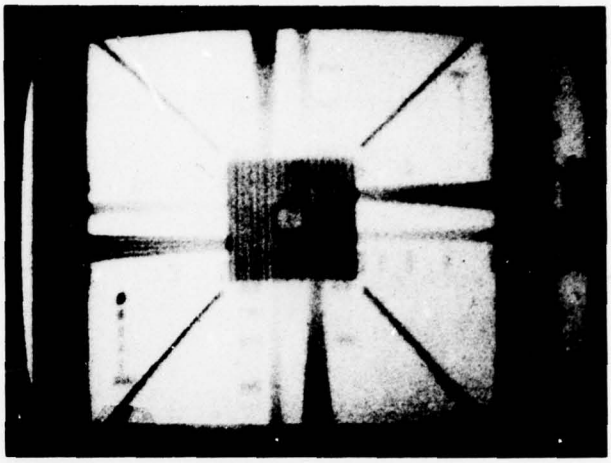


SAMPLING RATE	8.19 MHZ
TOTAL NUMBER OF SECTORS	480/FIELD
	960/FRAME
TOTAL NUMBER OF PELS	123K/FIELD
	246K/FRAME

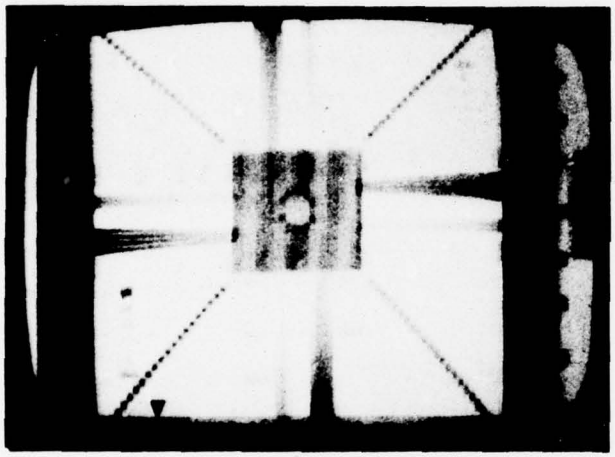
Figure 6-2. TRANSFORM FORMAT



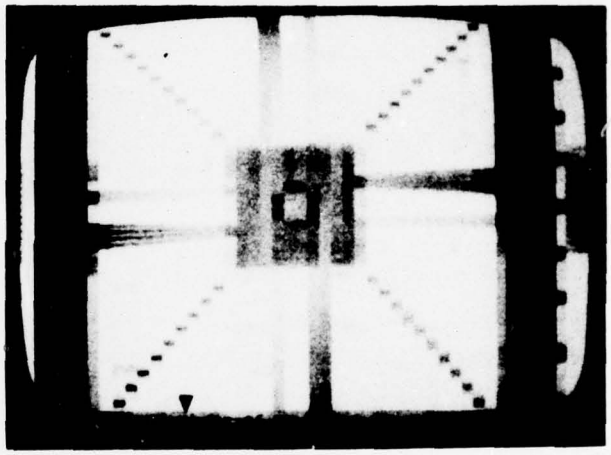
(a) ORIGINAL 8 BIT/PEL



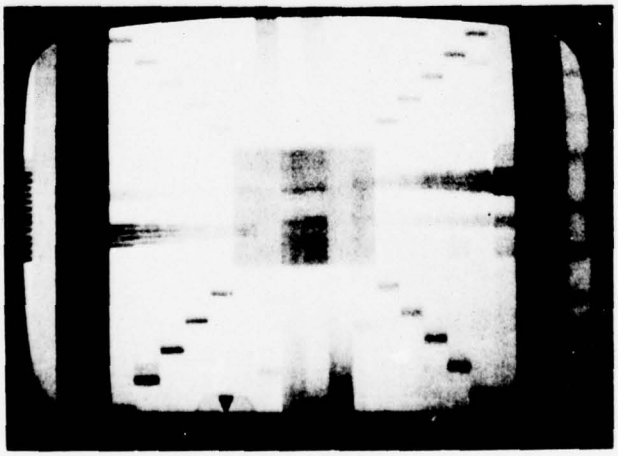
(b) 4 BIT/PEL



(c) 2 BIT/PEL



(d) 1 BIT/PEL



(e) 1/2 BIT/PEL

Figure 6-3. UNDER SAMPLED SCAN LINES

6.3.2 Area Sampling

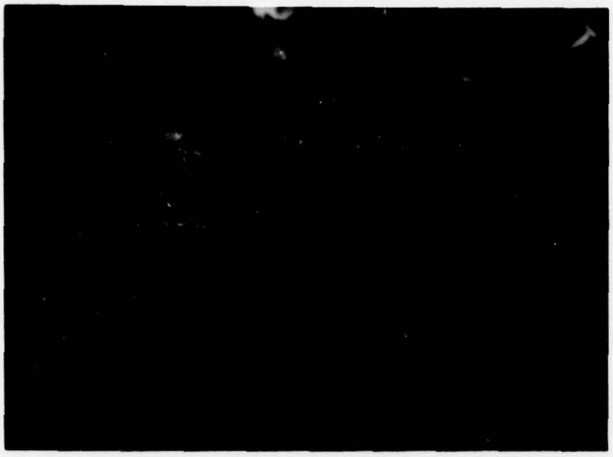
Spatial resolution may also be reduced in two dimensions or on an area basis. Figure 6-4 shows the results of averaging square areas of pels together to form the output image. As is shown in Figure 6-4, two bits/pel is about the limit of acceptable degradation that can be tolerated using this approach. Note the the compression rate achieved by area sampling is the square of the rate obtained by reduced line sampling.

6.3.3 Intensity Sampling

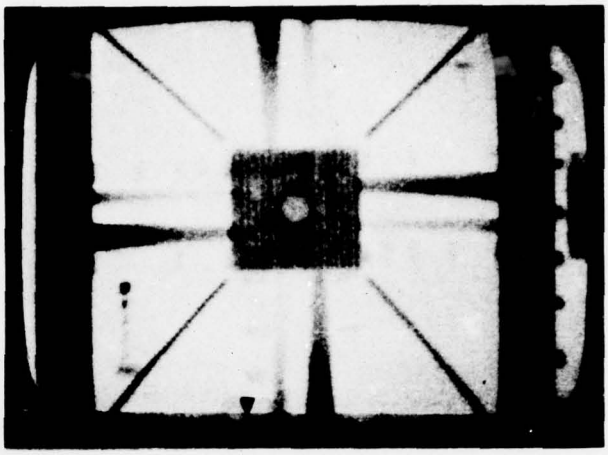
The information content of image may be reduced simply by using fewer bits to represent the intensity of each image sample. The effects of various levels of quantization of intensity are illustrated in Figure 6-5. A FLIR image was used here because it displays more shades of gray than the video test pattern. Note that reducing the intensity quantization from eight bits/pel (256 gray shades) to six bits/pel (64 gray shades) produces only a small amount of degradation in the image. At four bits/pel, however, false intensity contours become evident in the background in the vicinity of the targets, while at two bits/pel, much scene information is lost. The number of bits required to represent the image intensity will also vary with the type of imager, future processing requirements for the image, and the type of display used. FLIR's, for example, exhibit a very wide dynamic range, thereby possibly requiring more than eight bits of intensity information for adequate resolution. Also, if the image information is to be processed in any fashion (e.g., targets detected) a relatively high number of gray shades would be required.

6.3.4 Reduced Resolution Summary

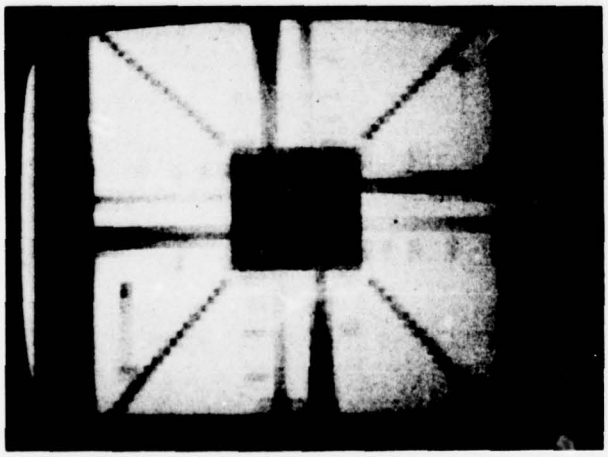
The results of reduced resolution compression indicate that a full resolution image may be compressed up to two bits/pel with a small degradation in quality. The goal of subsequent image compression effort then is to exceed this bandwidth reduction with little additional quality losses.



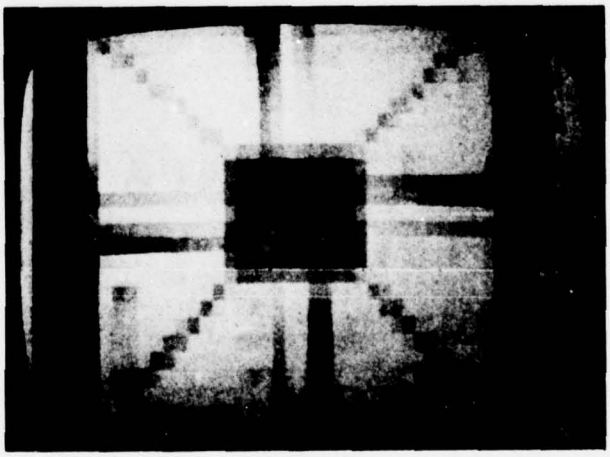
(a) ORIGINAL 8 BIT/PEL



(b) 2 BIT/PEL



(c) 1/2 BIT/PEL



(d) 1/8 BIT/PEL

Figure 6-4. UNDER SAMPLED AREAS



(a) ORIGINAL 8 BIT/PEL



(b) 6 BIT/PEL



(c) 4 BIT/PEL



(d) 2 BIT/PEL

Figure 6-5. REDUCED QUANTIZATION LEVELS

6.4 HAAR TRANSFORM COMPRESSION

Rather than attempt to compress the image directly, a more sophisticated approach to compression is to transform the image, using an orthogonal transform, and then compress the result. In order to be useful for compression, the transform used in this approach should isolate the relevent image properties from the pel-to-pel redundancies most images exhibit. Also, the computational complexity of the transform must be compatible with the data rate of the system. The two dimensional Haar transform satisfies both of these requirements.

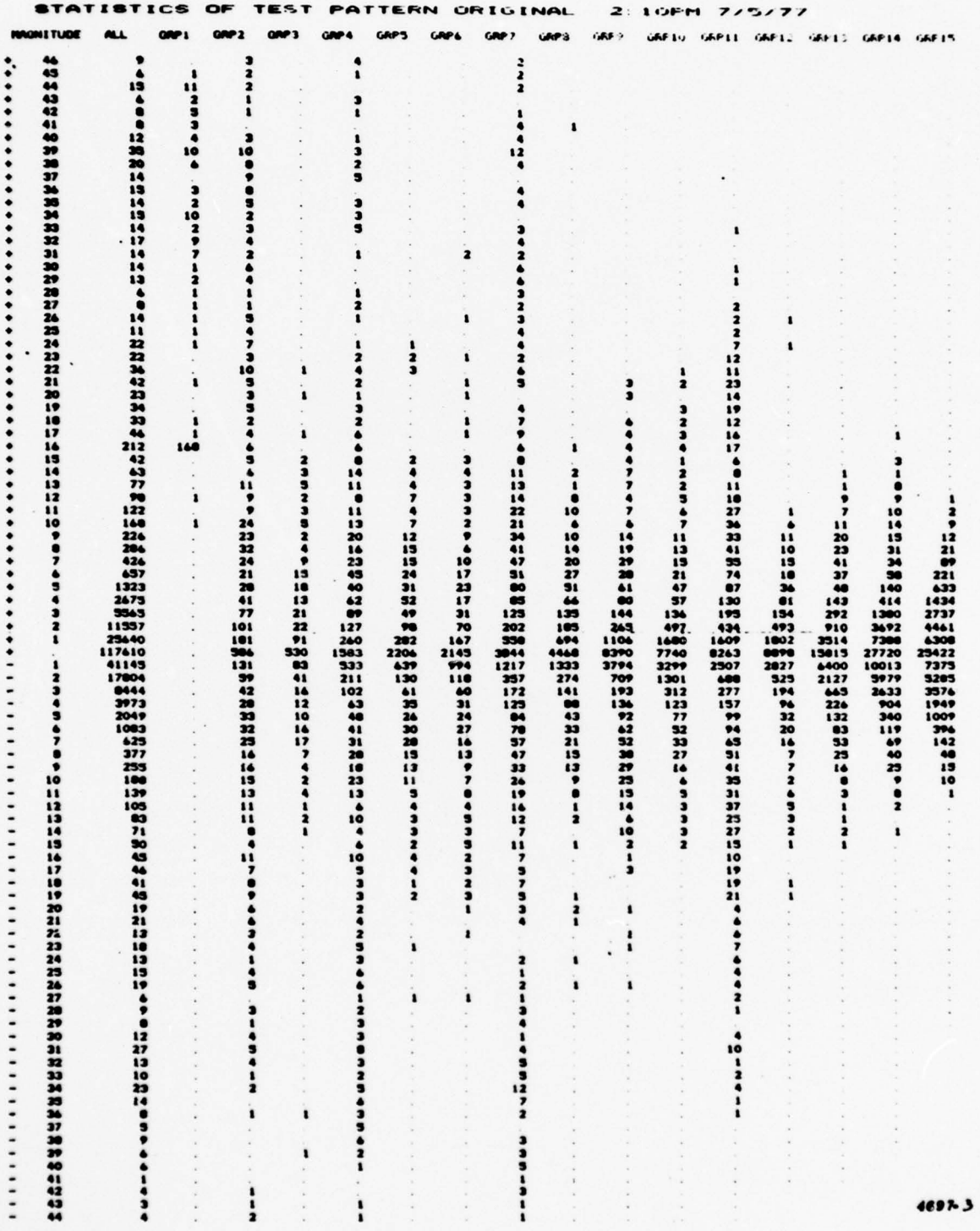
The Haar transform of an image may be compressed in two basic ways. The first approach is to reduce the number of bits used to represent each Haar coefficient. The second method of compressing is to reduce the number of Haar coefficients used to reconstruct the original image. Both approaches to compression are explored here.

6.4.1 Sequency Domain Bit Assignments

The 2-D sensor processor divides the Haar transform sequency domain into 15 regions or sequency groups. A sample amplitude distribution table for each group is shown in Figure 6-6. Examination of the amplitude distributions shows that in most groups, fewer than the full-precision nine bits (eight bits plus sign) are required to represent the elements of the group. By bit limiting the coefficient amplitudes in each group, about a two-to-one reduction in bandwidth may be achieved with no visible degradation in image quality. Figures 6-7 and 6-8 both show an original eight-bit/pel image along with an image compressed by assigning fewer bits to each sequency group. A table showing the bit assignments used is also given. The data rate realized by this approach is 4.14 bits/pel.

6.4.2 Coefficient Elimination

As a result of the information compaction properties of the Haar transform, many of the transform coefficients contribute little or no information about the image. As a consequence, these terms need not be used to reconstruct the original. Thus, the image information may be compressed by eliminating some coefficients from consideration. Several simple forms of coefficient elimination are considered here.



STATISTICS OF TEST PATTERN ORIGINAL 2:10PM 7/5/77

MAGNITUDE	ALL	GRP1	GRP2	GRP3	GRP4	GRP5	GRP6	GRP7	GRP8	GRP9	GRP10	GRP11	GRP12	GRP13	GRP14	GRP15
46	9		3		4			2								
45	6	1	2		1			2								
44	15	11	2					2								
43	4	2	1		1											
42	8	5	1		1			1								
41	8	3						4	1							
40	12	4	3		1			4								
39	35	10	10		3			12								
38	20	6	8		2			4								
37	14		9		5											
36	15	3	8					4								
35	14	2	5		3			4								
34	15	10	2		3											
33	14	2	3		5			3				1				
32	17	9	4					4								
31	14	7	2		1		2	2								
30	14	1	6					6				1				
29	13	2	4					6				1				
28	6	1	1		1			3								
27	8	1	1		2			2				2				
26	14	1	5		1		1	3				2	1			
25	11	1	4					4				2				
24	22	1	7		1	1		4				7	1			
23	22		3		2	2	1	2				12				
22	36		10	1	4	3		6			1	11				
21	42	1	5		2		1	5		3	2	23				
20	23		3	1	1		1			3		14				
19	34		5		3			4			3	19				
18	33	1	2		2		1	7		6	2	12				
17	44	1	4	1	6		1	9		4	3	16			1	
16	212	168	6		6			6	1	4	4	17				
15	42		5	2	8	2	3	8		4	1	6			3	
14	63		4	3	14	4	4	11	2	7	2	8		1	1	
13	77		11	5	11	4	3	13	1	7	2	11		1	8	
12	98	1	9	2	8	7	3	14	8	4	5	18		9	9	1
11	122		9	3	11	4	3	22	10	7	6	27	1	7	10	2
10	168	1	24	8	13	7	2	21	6	6	7	36	6	11	14	9
9	226		23	2	20	12	9	34	10	14	11	33	11	20	15	12
8	284		32	4	16	15	6	41	14	19	13	41	10	23	31	21
7	426		24	9	23	15	10	47	20	29	15	55	15	41	34	89
6	657		21	15	45	24	17	51	27	28	21	74	18	37	58	221
5	1323		28	18	40	31	23	80	51	41	47	87	36	48	140	633
4	2675		41	13	62	52	17	85	66	80	57	130	81	143	414	1434
3	5565		77	21	89	49	31	125	135	144	134	195	134	292	1380	2737
2	11557		101	22	127	98	70	202	185	265	497	434	493	910	3692	4461
1	25440		181	91	260	282	167	558	694	1106	1680	1609	1802	3514	7388	6308
	117610		986	530	1583	2206	2145	3844	4468	8390	7740	8263	8898	15815	27720	25422
1	41145		131	83	533	639	994	1217	1333	3794	3299	2507	2827	6400	10013	7375
2	17804		99	41	211	130	118	357	274	709	1301	688	525	2127	5979	5285
3	8444		42	14	102	61	60	172	141	193	312	277	194	665	2633	3576
4	3973		28	12	63	35	31	125	88	136	123	157	96	226	904	1949
5	2049		33	10	48	26	24	84	43	92	77	99	32	132	340	1009
6	1083		32	14	41	30	27	78	33	62	52	94	20	83	119	396
7	625		25	17	31	28	16	57	21	52	33	65	16	53	69	142
8	377		16	7	28	15	13	47	15	38	27	51	7	25	40	48
9	255		16	4	18	13	9	33	13	29	16	41	7	16	25	15
10	188		15	2	23	11	7	24	9	25	6	35	2	8	9	10
11	139		13	4	13	5	8	19	8	15	5	31	6	3	8	1
12	105		11	1	6	4	4	16	1	14	3	37	5	1	2	
13	83		11	2	10	3	5	12	2	6	3	25	3	1		
14	71		8	1	4	3	3	7		10	3	27	2	2	1	
15	50		4		6	2	5	11	1	2	2	15	1	1		
16	45		11		10	4	2	7		1		10				
17	44		7		5	4	3	5		3		19				
18	41		8		3	1	2	7				19	1			
19	45		9		3	2	3	5	1			21	1			
20	19		6		2		1	3	2	1		4				
21	21		6		4			4	1			6				
22	12		2		2		1			1		6				
23	18		4		5	1				1		7				
24	13		1		3			2	1			6				
25	15		4		6			1				4				
26	19		5		6			2	1	1		4				
27	6				1	1	1	1				2				
28	9		3		2			3				1				
29	8		1		3			4								
30	12		4		3			1				4				
31	27		5		8			4				10				
32	13		4		3			5				1				
33	10		1		2			5				2				
34	23		2		5			12				4				
35	14				6			7				1				
36	8		1	1	3			2				1				
37	5				5											
38	9				6			3								
39	6			1	2			3								
40	6				1			5								
41	1							1								
42	4		1					3								
43	3		1		1			1								
44	4		2		1			1								

4697-3

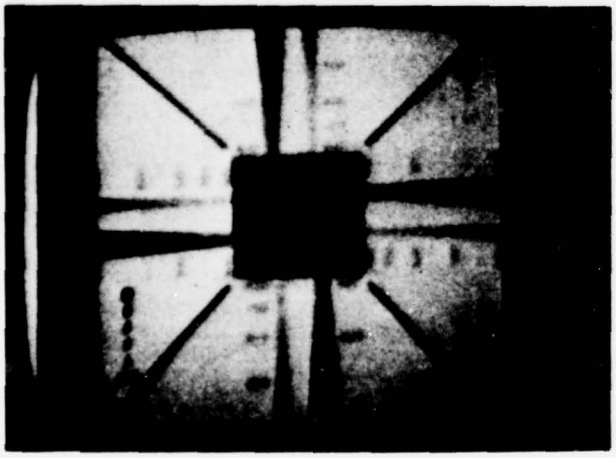
Figure 6-6. EXAMPLE OF AMPLITUDE DISTRIBUTIONS



ORIGINAL 8 BIT/PEL

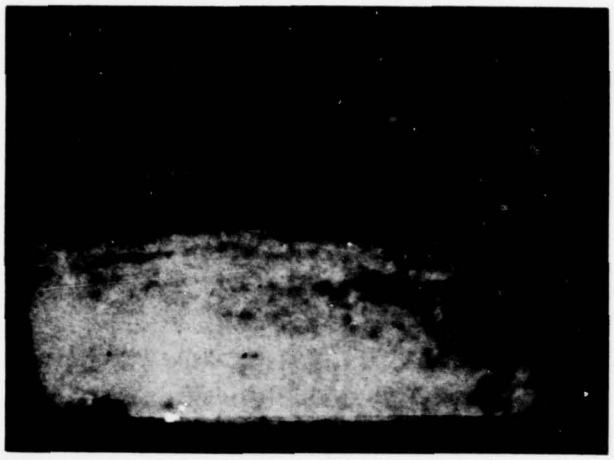
GROUP NO.	NO. OF COEF/ GROUP	NO. OF BITS/ COEF	TOTAL NO. OF BITS/GROUP	
1	1	8	8	
2	2	8	16	
3	1	5	5	
4	4	7	28	
5	4	5	20	
6	4	6	24	
7	8	6	48	
8	8	4	32	
9	16	5	80	
10	16	5	80	
11	16	5	80	
12	16	4	64	
13	32	4	128	
14	64	4	256	
15	64	3	192	
	256		1061	TOTAL

$\frac{1061}{256}$ = 4.14 BIT/PEL



4.14 BIT/PEL

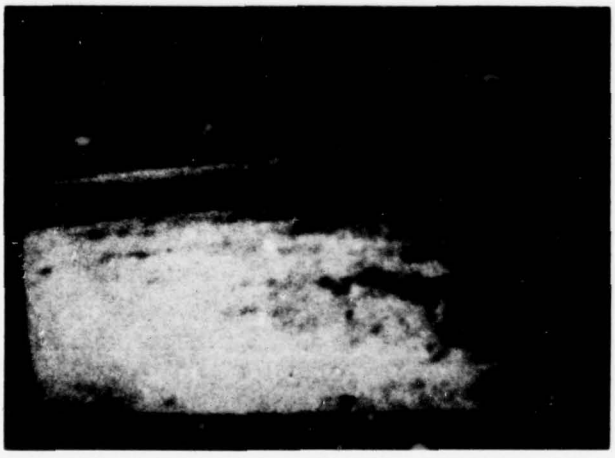
Figure 6-7. TEST PATTERN COMPRESSED BY SEQUENCY DOMAIN BIT LIMITING



ORIGINAL 8 BIT/PEL

GROUP NO.	NO. OF COEF/ GROUP	NO. OF BITS/ COEF	TOTAL NO. OF BITS/GROUP
1	1	8	8
2	2	8	16
3	1	5	5
4	4	7	28
5	4	5	20
6	4	6	24
7	8	6	48
8	8	4	32
9	16	5	80
10	16	5	80
11	16	5	80
12	16	4	64
13	32	4	128
14	64	4	256
15	64	3	192
	256		1061 TOTAL

$\frac{1061}{256}$ = 4.14 BIT/PEL



4.14 BIT/PEL

Figure 6-8. FLIR IMAGE COMPRESSED BY SEQUENCY DOMAIN BIT LIMITING

6.4.2.1 Zonal Filtering - Zonal filtering is performed simply by zeroing entire sequency groups or zones in the Haar plane. The sequency groups that are retained (or passed through the zonal filter) are those that are deemed a priori to contain the important scene information.

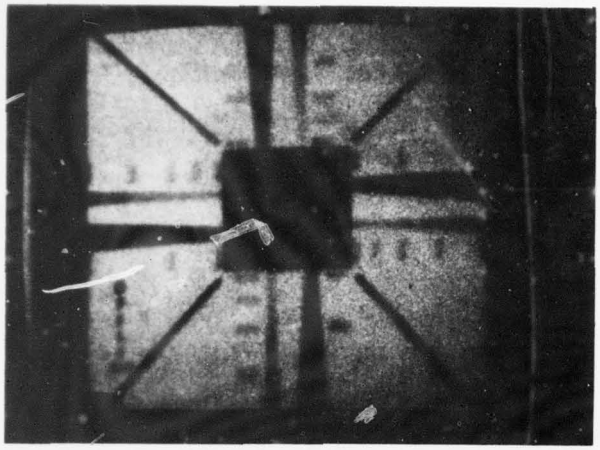
Several zonal filtered images are shown in Figures 6-9 and 6-10. In each case, the filters used are essentially low-pass filters. However, by selecting sequency groups in a judicious manner, a majority of the edges within the scene may still be retained. The specific edges retained are also determined by the sequency group selection. This edge-isolating property of the Haar transform can be seen by comparing Figures 6-9c and 6-9d. While both images are compressed to the same level, the diagonal edges of Figure 6-9d are more accurately represented than those of Figure 6-9c. This affinity for diagonal edges, however, is obtained by reducing the resolution of horizontal and vertical edges.

6.4.2.2 Thresholding - Thresholding eliminates transform coefficients by requiring them to exhibit an amplitude greater than some specific threshold value in order to be included in the image reconstruction. This approach to coefficient elimination is more scene adaptive than zonal filtering in that relatively high amplitude coefficients may be passed regardless of sequency. However, since the position of the above-threshold coefficients is not known a priori as it was in zonal filtering, some additional positional information must be included with each coefficient retained.

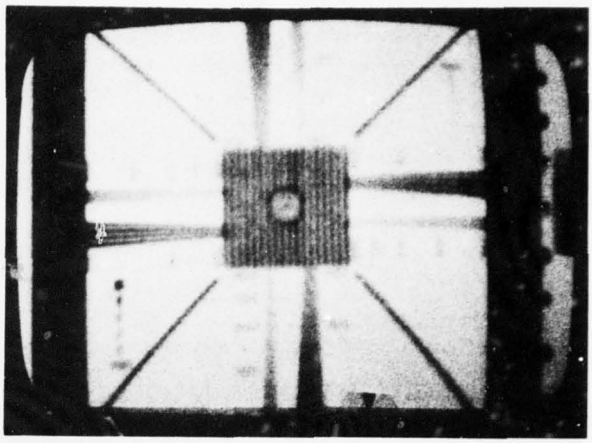
The number of bits required to provide coefficient position information can vary depending on the coding scheme used. For example, if the row and column address is used to locate each nonzero coefficient, a total of eight bits per coefficient is required. This requirement may be reduced by scanning the Haar plane in a specified serial fashion and encoding the number of zeroes between coefficients. This approach is known as run-length encoding. A bit-efficient comma-free Huffman code can then be derived from the statistics of the run lengths. By using this type of encoding while scanning the Haar plane in order of increasing sequency, only five bits of position information (worst case) are required per coefficient.[1]

---

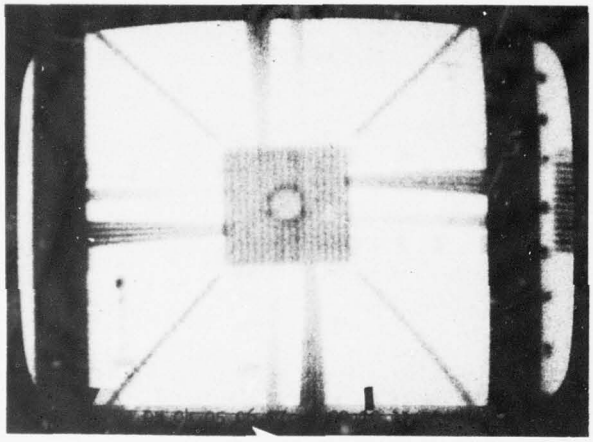
[1] Powers, J. W., Nort. 72-331, Image Data Compression Using Digital Holography (U)
Powers et al, Nort. 73-293, Imagery/Bandwidth Compression RPV Study (S)



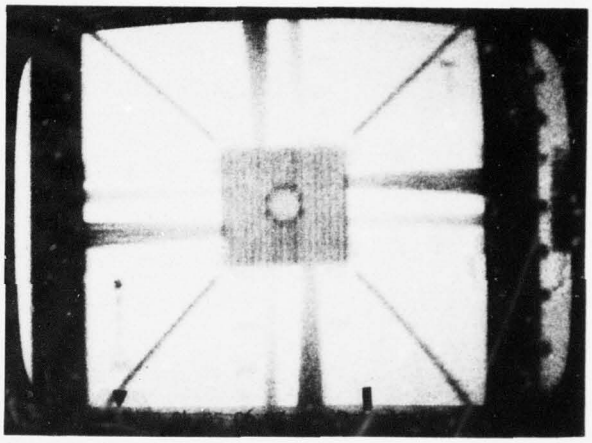
(a) 8 BIT/PEL



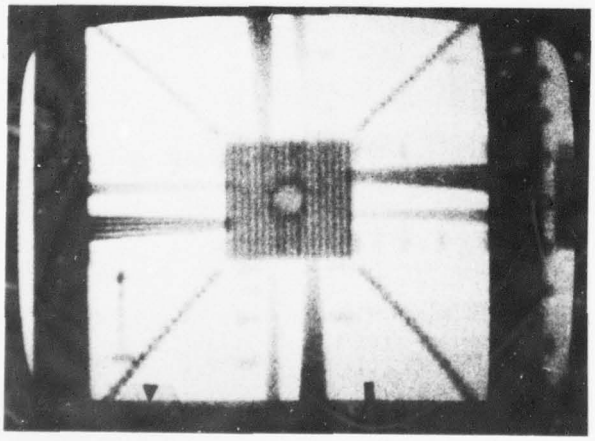
(b) GROUPS 1–13; 2.4 BIT/PEL



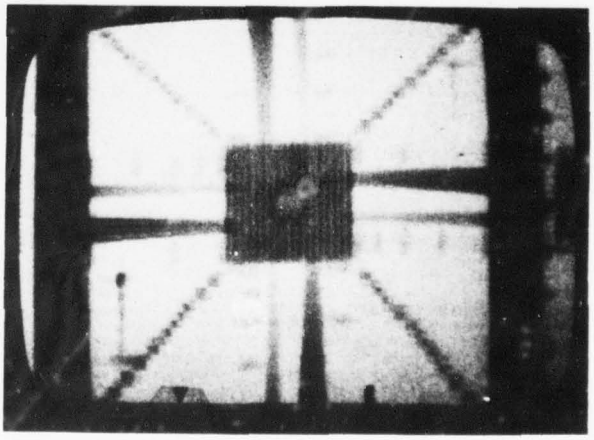
(c) GROUPS 1–9, 11, 12; 1.6 BIT/PEL



(d) GROUPS 1–10, 12; 1.6 BIT/PEL

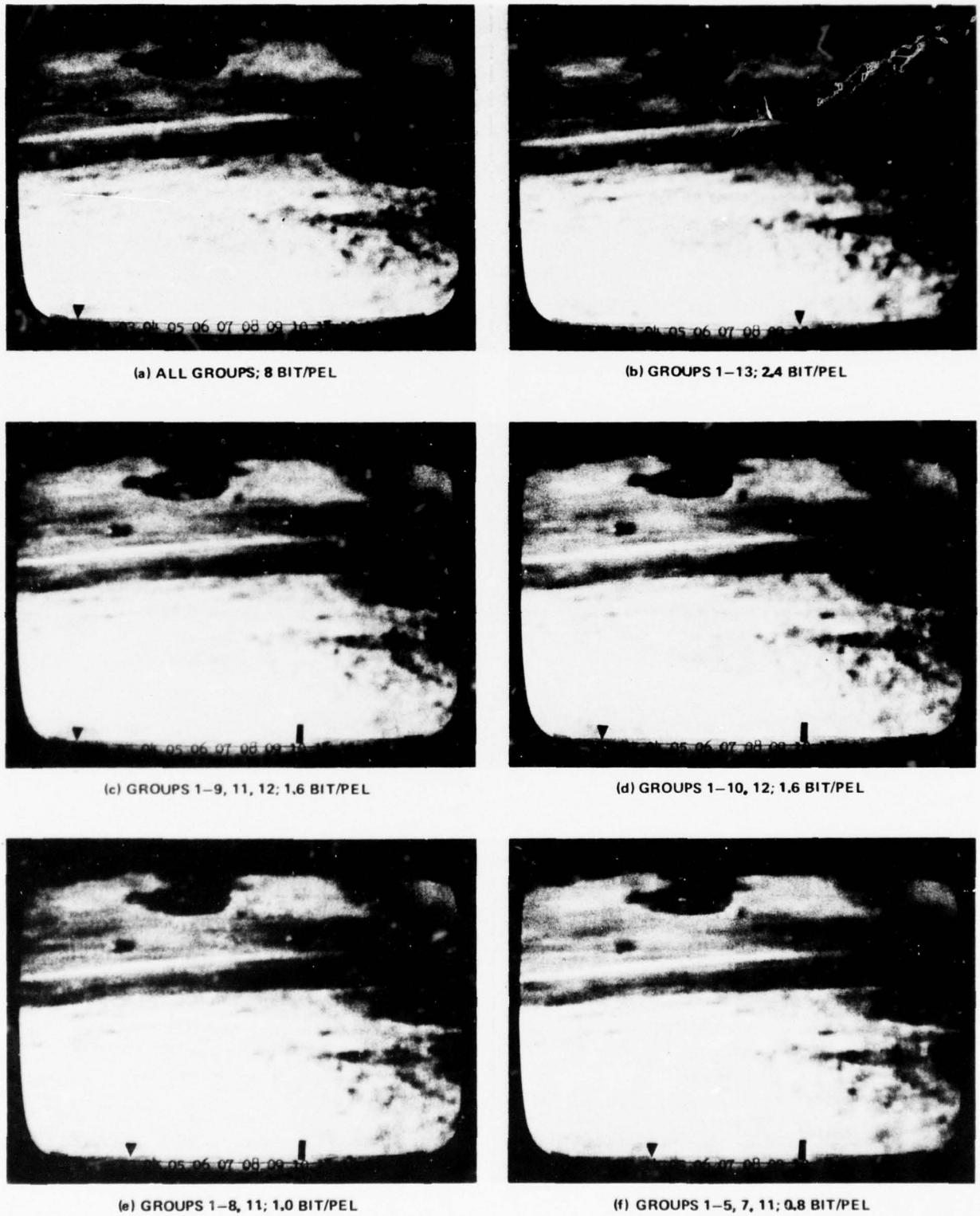


(e) GROUPS 1–8, 11; 1.0 BIT/PEL



(f) GROUPS 1–5, 7, 11; 0.8 BIT/PEL

Figure 6-9. ZONAL FILTERED TEST PATTERN



(a) ALL GROUPS; 8 BIT/PEL

(b) GROUPS 1–13; 2.4 BIT/PEL

(c) GROUPS 1–9, 11, 12; 1.6 BIT/PEL

(d) GROUPS 1–10, 12; 1.6 BIT/PEL

(e) GROUPS 1–8, 11; 1.0 BIT/PEL

(f) GROUPS 1–5, 7, 11; 0.8 BIT/PEL

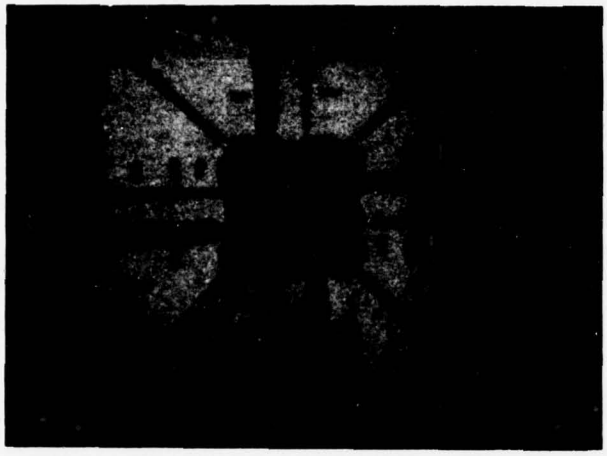
Figure 6-10. ZONAL FILTERED FLIR NO. 1

Once the positional information requirements are set, the compression rates for different levels of thresholding may be determined. Figures 6-11 and 6-12 show the effects of various threshold levels in the sequency domain. The corresponding data rates shown in the figures were determined by multiplying the number of above-threshold coefficients in each group by the group's bit assignment (Figure 6-7) plus five bits of position information. The resultant values are then summed over the entire image and scaled by the number of pels to arrive at the number of bits per pel. Computing the data rate in this fashion, however, assumes the use of a large data buffer to accommodate the variations in frequency of above-threshold coefficients.

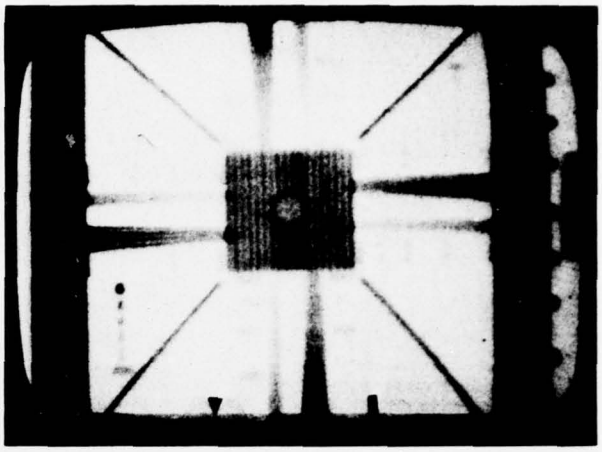
While the compression ratios obtainable using thresholding are much greater than those observed in zonal filtering, the storage requirements make the practical use of thresholding unattractive. A compromise approach that exhibits much of the scene adaptability of thresholding while providing the advantages of a relatively constant data rate will be examined next.

6.4.2.3 N Coefficient Selection - A method of assuring a constant data rate with minimal memory requirements is to require that for each image sector, exactly N coefficients are retained. The N coefficients are selected by setting a threshold as before and, while scanning the Haar domain in order of increasing sequency, retaining the first N above-threshold terms. Thus, if there are more than N significant coefficients in an image sector's Haar transform, the ones relating the higher sequency information will be excluded from the reconstruction process. Conversely, if fewer than N above-threshold coefficients are present in a given sector, an appropriate number of zeroes may be included. This compression approach requires storage of only N elements rather than a majority of the compressed image frame.

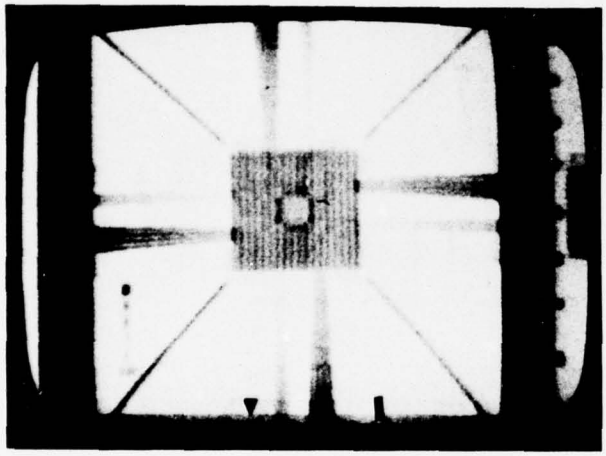
Figures 6-13 and 6-14 show the effects of this type of compression on the two test images. Note that for fairly similar image quality, the compression rates obtained are comparable to those observed by thresholding only. Also, comparison of Figures 6-13c and 6-14c shows that N coefficient-selection provides an image with better resolution than an under-sampled one while requiring only one-half as many bits per pel.



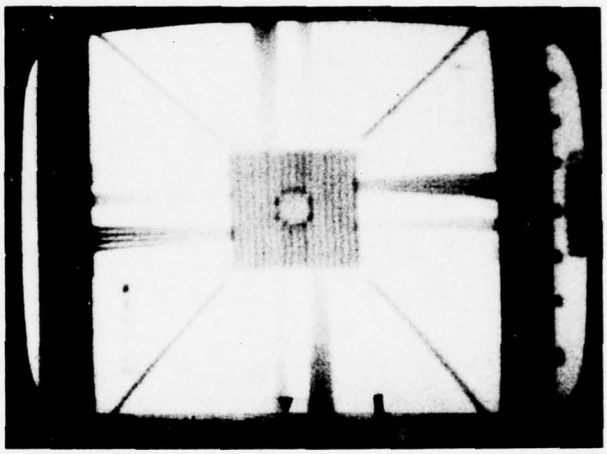
(a) ORIGINAL



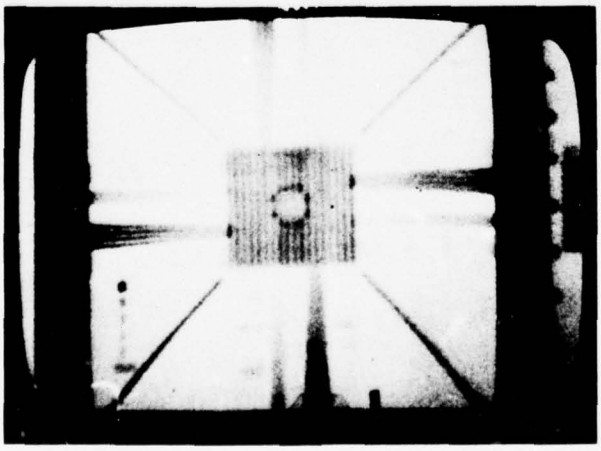
(b) 1.87 BIT/PEL; $\theta = 2$



(c) 0.55 BIT/PEL; $\theta = 3$



(d) 1.78 BIT/PEL; $\theta = 1,1,2,1,2; 2,2,1,1,1; 2,2,2,2,3$



(e) 0.83 BIT/PEL; $\theta = 27,1,2,1,2,4,2,2,2,2,3,3,3,3,3$

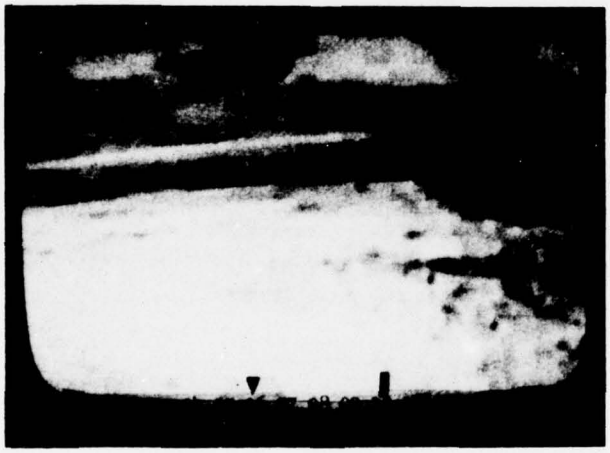
Figure 6-11. EFFECTS OF THRESHOLD LEVELS IN SEQUENCY DOMAIN, TEST PATTERN



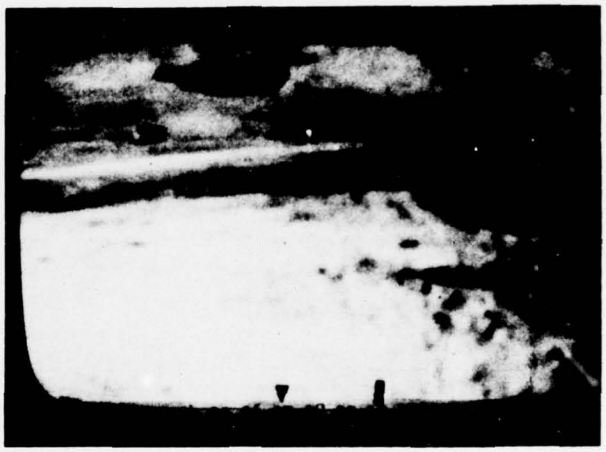
(a) ORIGINAL



(b) 1.87 BIT/PEL; THRESHOLD AT 2



(c) 0.24 BIT/PEL; THRESHOLD AT 3



(d) 1.78 BIT/PEL; θ = 1,1,2,1,2; 2,2,1,1,1; 2,2,2,2,3

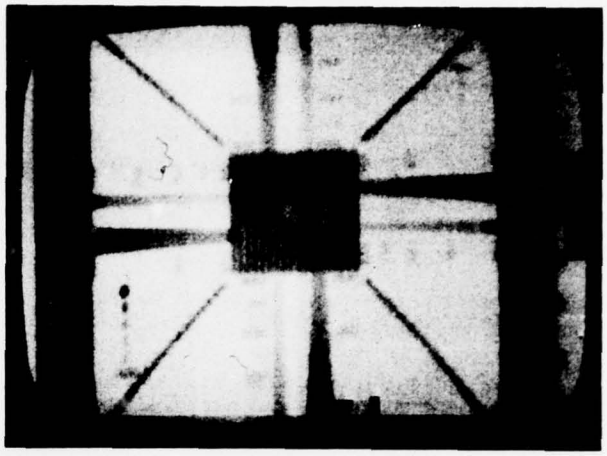


(e) 0.62 BIT/PEL; θ = 27,1,2,1,2,4,2,2,2,2,3,3,3,3,3

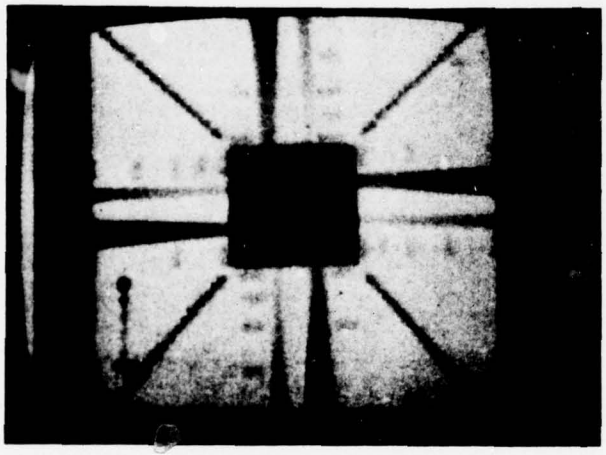
Figure 6-12. EFFECTS OF THRESHOLD LEVELS IN SEQUENCY DOMAIN, FLIR NO. 1



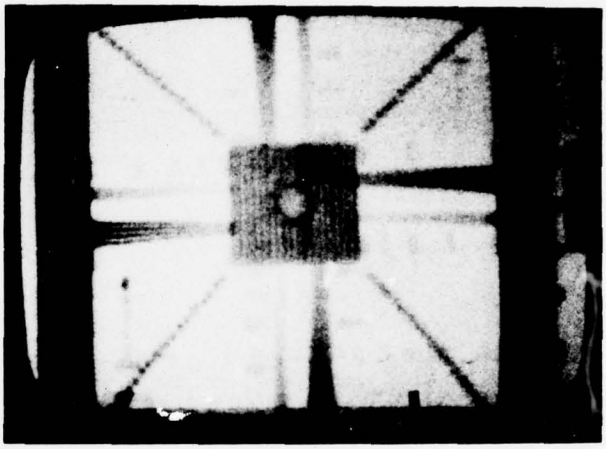
(a) ORIGINAL



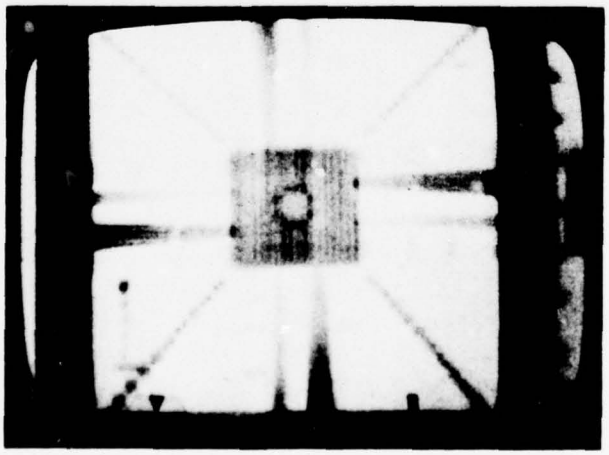
(b) 1.47 BIT/PEL; N = 51



(c) 0.99 BIT/PEL; N = 31



(d) 0.83 BIT/PEL; N = 22



(e) 0.55 BIT/PEL; N = 15

Figure 6-13. THRESHOLD AND N-COEFFICIENTS TEST PATTERN



(a) ORIGINAL



(b) 1.47 BIT/PEL; N = 51



(c) 1.15 BIT/PEL; N = 31



(d) 0.83 BIT/PEL; N = 22



(e) 0.55 BIT/PEL; N = 15

Figure 6-14. THRESHOLD AND N-COEFFICIENTS (FLIR NO. 1)

The performance, simplicity and system requirements of the N coefficient selection algorithm make it the optimal compression algorithm within the sensor processor's immediate capability. Five FLIR images that have been compressed to various levels using this algorithm are shown in Figures 6-15 through 6-19. The figures show that a data rate of 1.15 bits/pel is easily obtainable for each image.

6.5 IMAGE DESCRIPTORS FOR HAAR-COMPRESSED IMAGES

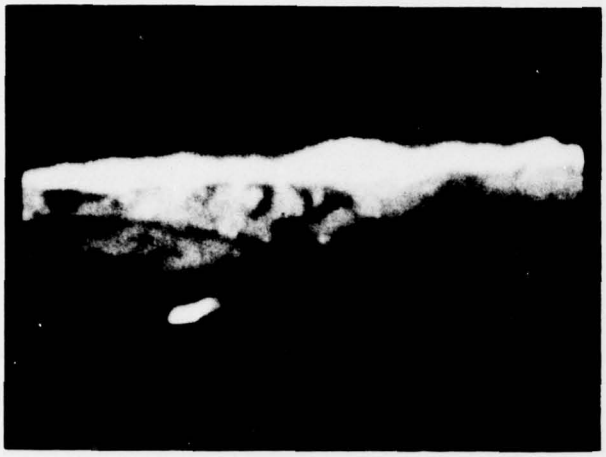
The set of image descriptors was calculated for each of the processed images included in this section (see paragraph 7.2 for a derivation of the "image descriptor"). Table 6-1 lists the average tone, texture, and edge values computed for the entire image. In each case, the tone value remains approximately the same throughout processing. It is included as a check for each image/computer set-up. Since targets are separated from clutter primarily by their edge content, successful bandwidth compression would try to preserve edges at the expense of texture. Therefore, in a perfect bandwidth reduction scheme, the ratio of the average edge to the average texture divided by the bandwidth reduction would remain constant. (This presupposes ideal edge and texture descriptors.) As can be seen from the table, the edge descriptors show a minimum amount of decrease, while the texture falls off rapidly with bandwidth reduction.

A figure of merit (f) has been devised to offer a rough grading of the effects of compression on the different image samples:

$$f = \frac{(\bar{e}/\bar{x})\ (b/8)}{\bar{e}_o/\bar{x}_o}$$

where b is bits/pel and $\bar{e}_o$ and $\bar{x}_o$ are edge and texture average descriptors of original.

As is noted in the table, f decreases as the compression increases in almost all cases. Looking at the images, one can see deterioration increasing with compression, which tends to validate the values of f. This relationship holds for the test pattern and each of the FLIR images. If 0.5 were arbitrarily selected for a cut-off value, one bit per pel would be obtainable, using the N-coefficient algorithm with threshold.



(a) ORIGINAL



(b) 1.47 BIT/PEL



(c) 0.99 BIT/PEL

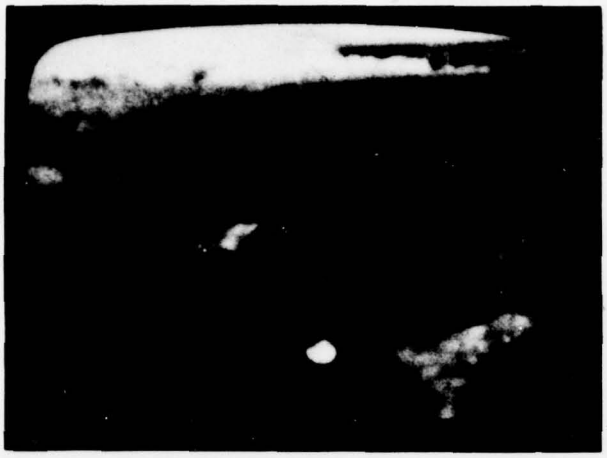


(d) 0.75 BIT/PEL



(e) 0.52 BIT/PEL

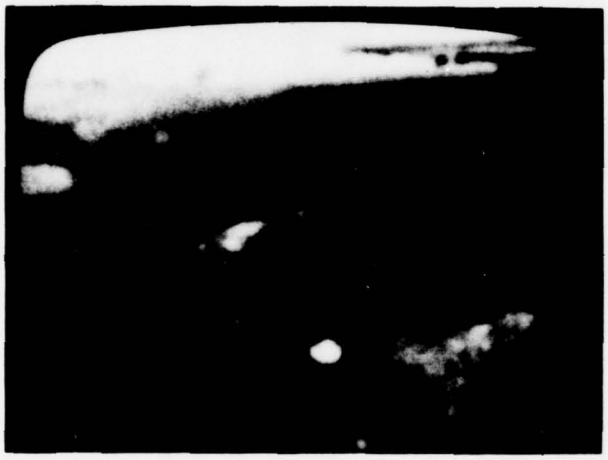
Figure 6-15. THRESHOLD AND N COEFFICIENTS (NVL NO. 4)



(a) ORIGINAL



(b) 1.47 BIT/PEL



(c) 0.99 BIT/PEL

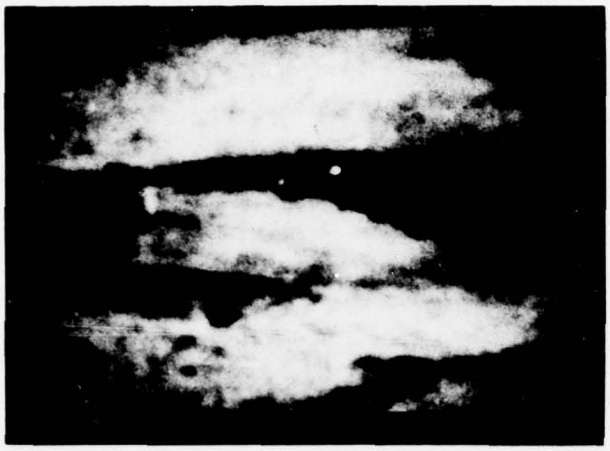


(d) 0.75 BIT/PEL

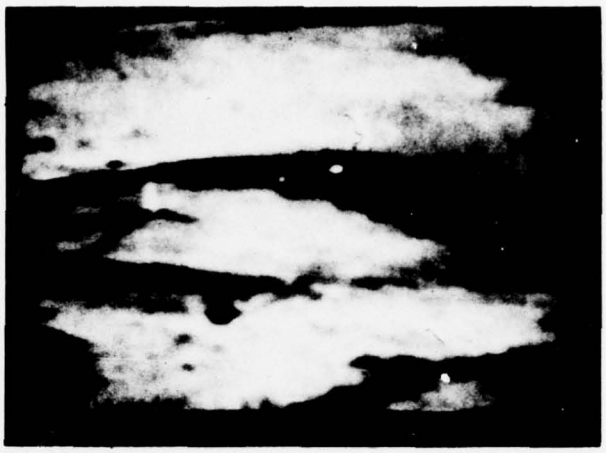


(e) 0.52 BIT/PEL

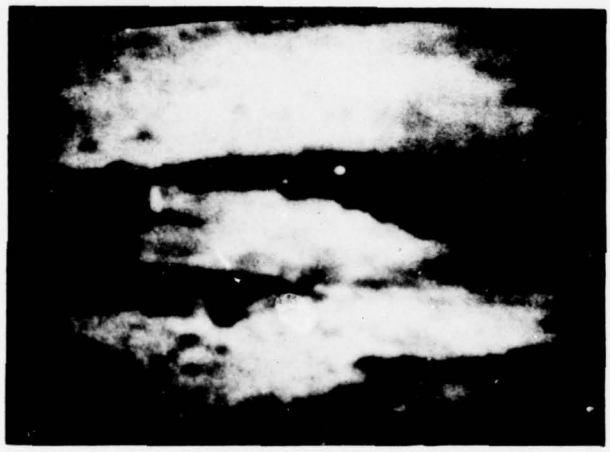
Figure 6-16. THRESHOLD AND N-COEFFICIENTS (NVL NO. 6)



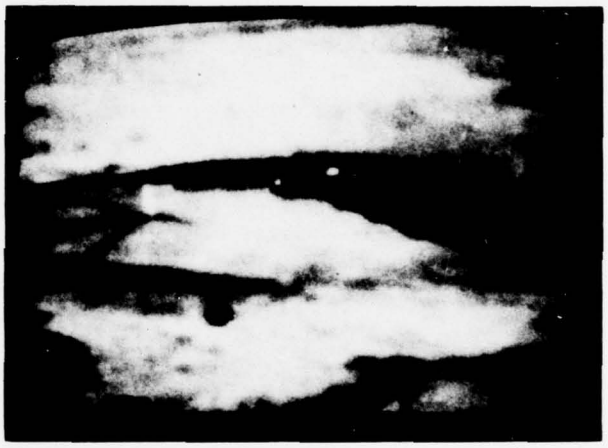
(a) ORIGINAL



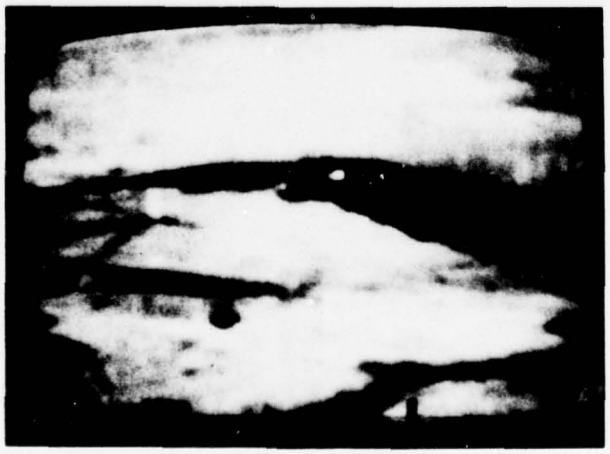
(b) 1.47 BIT/PEL



(c) 0.99 BIT/PEL

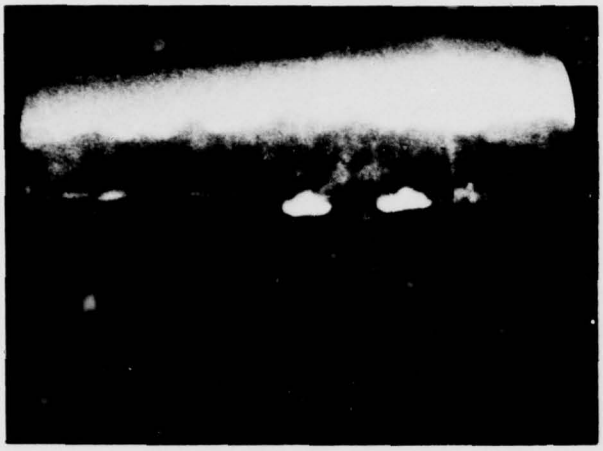


(d) 0.75 BIT/PEL

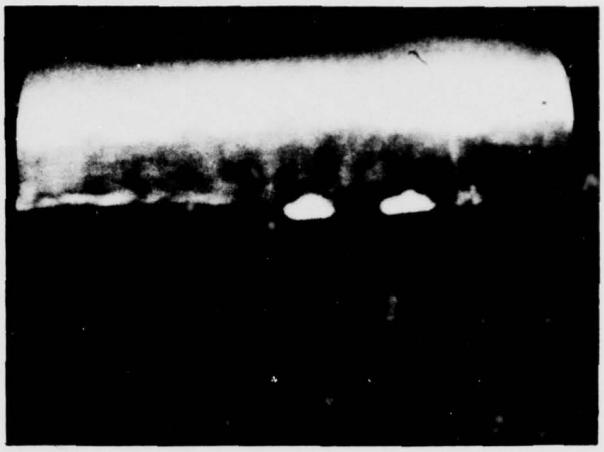


(e) 0.52 BIT/PEL

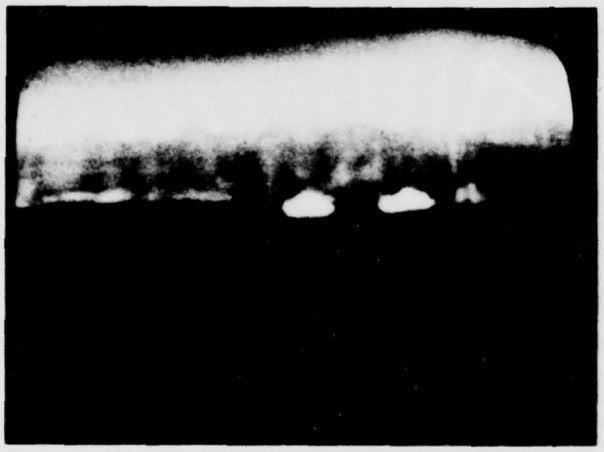
Figure 6-17. THRESHOLD AND N COEFFICIENTS (NVL NO. 7)



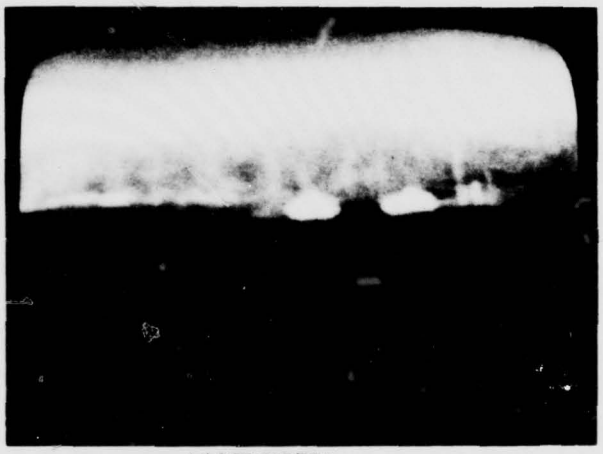
(a) ORIGINAL



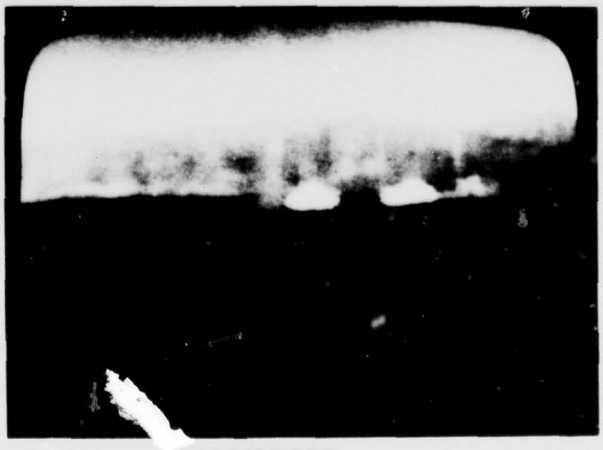
(b) 1.47 BIT/PEL



(c) 0.99 BIT/PEL

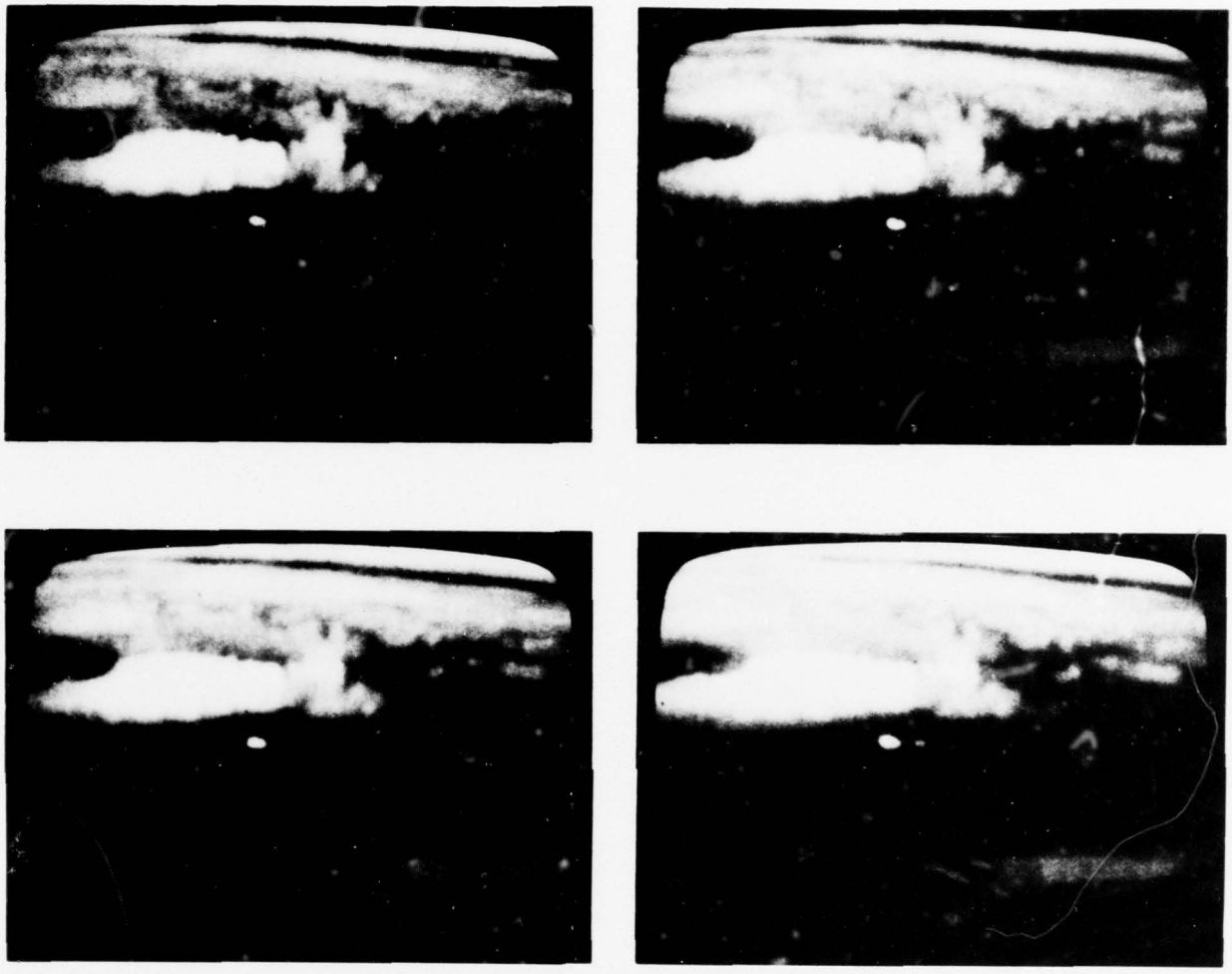


(d) 0.75 BIT/PEL



(e) 0.52 BIT/PEL

Figure 6-18. THRESHOLD AND N COEFFICIENTS (NVL NO. 8)



(c) 0.99 BIT/PEL

(d) 0.75 BIT/PEL

(e) 0.52 BIT/PEL

Figure 6-19. THRESHOLD AND N COEFFICIENTS (NVL NO. 10)

Table 6-1. DESCRIPTORS FOR HAAR-COMPRESSED IMAGES

Figure	Image - Process	Average			f*
		Tone	Text.	Edge	
6-7	Test Pattern				
a	8-bit/pel (original)	32	568	1747	1.10
b	4.14-bit/pel	32	524	1239	0.40
6-8	FLIR #1				
a	8-bit/pel(original)	22	416	516	1.00
b	4.14-bit/pel	22	399	501	0.52
6-9	Test Pattern				
a	Original	32	602	1677	1.00
b	Zonal 2.4-bit/pel	32	214	1462	0.73
c	Zonal 1.6-bit/pel	32	213	1398	0.46
d	Zonal 1.6-bit/pel	32	159	1371	0.61
e	Zonal 1.0-bit/pel	32	177	1364	0.34
f	Zonal 0.8-bit/pel	32	158	1336	0.30
6-10	FLIR #1				
a	Original	22	483	582	1.00
b	Zonal 2.4-bit/pel	22	206	472	0.57
c	Zonal 1.6-bit/pel	22	169	438	0.43
d	Zonal 1.6-bit/pel	22	130	430	0.55
e	Zonal 1.0-bit/pel	22	178	414	0.24
f	Zonal 0.8-bit/pel	22	121	391	0.27
6-11	Test Pattern				
a	Original	32	637	1612	1.00
b	Threshold 1.87-bit/pel	32	339	1615	0.45
d	Threshold 1.78-bit/pel	32	287	1624	0.50
e	Threshold 0.83-bit/pel	34	186	1633	0.38
c	Threshold 0.55-bit/pel	34	120	1665	0.36
6-12	FLIR #1				
a	Original	22	454	546	1.00
b	Threshold 1.87-bit/pel	22	164	673	0.80
d	Threshold 1.78-bit/pel	22	144	648	0.83
e	Threshold 0.62-bit/pel	22	75	607	0.52
c	Threshold 0.24-bit/pel	22	50	595	0.30
6-13	Test Pattern				
a	Original	32	731	1712	1.00
b	N-coefficient N=51 (1.47)	31	221	1481	0.54
c	N-coefficient N=31 (0.49)	32	142	1388	0.52
d	N-coefficient N=22 (0.75)	32	103	1332	0.53
e	N-coefficient N=15 (0.52)	32	58	1236	0.60

*Figure of merit.

Table 6-1. DESCRIPTORS FOR HAAR-COMPRESSED IMAGES (Continued)

Figure	Image - Process	Average			f*
		Tone	Text.	Edge	
6-14	FLIR #1				
a	Original	22	449	546	1.00
b	N-coefficient N=51 (1.47)	23	103	505	0.75
c	N-coefficient N=31 (0.99)	23	80	422	0.54
d	N-coefficient N=22 (0.75)	23	62	388	0.48
e	N-coefficient N=15 (0.52)	23	41	355	0.47
6-15	FLIR #4				
a	Original	18	286	723	1.00
b	Threshold & N-coefficient 1.47-bit/pel	19	87	729	0.62
c	Threshold & N-coefficient 0.99-bit/pel	18	69	630	0.45
d	Threshold & N-coefficient 0.75-bit/pel	18	52	565	0.41
e	Threshold & N-coefficient 0.52-bit/pel	18	42	524	0.32
6-16	FLIR #6				
a	Original	23	340	567	1.00
b	Threshold & N-coefficient 1.47-bit/pel	24	64	533	0.90
c	Threshold & N-coefficient 0.99-bit/pel	24	47	437	0.68
d	Threshold & N-coefficient 0.75-bit/pel	23	38	390	0.57
e	Threshold & N-coefficient 0.52-bit/pel	23	29	367	0.48
6-17	FLIR #7				
a	Original	30	656	1029	1.00
b	Threshold & N-coefficient 1.47-bit/pel	31	200	959	0.55
c	Threshold & N-coefficient 0.99-bit/pel	30	163	861	0.41
d	Threshold & N-coefficient 0.75-bit/pel	30	116	813	0.41
e	Threshold & N-coefficient 0.52-bit/pel	30	87	776	0.36
6-18	FLIR #8				
a	Original	16	353	573	1.00
b	Threshold & N-coefficient 1.47-bit/pel	15	94	536	0.65
c	Threshold & N-coefficient 0.99-bit/pel	15	83	475	0.44
d	Threshold & N-coefficient 0.75-bit/pel	15	62	431	0.41
e	Threshold & N-coefficient 0.52-bit/pel	15	45	407	0.37
6-19	FLIR #10				
a	Original	18	299	444	1.00
b	Threshold & N-coefficient 1.47-bit/pel	18	61	416	0.84
c	Threshold & N-coefficient 0.99-bit/pel	18	51	332	0.43
d	Threshold & N-coefficient 0.75-bit/pel	18	38	308	0.51
e	Threshold & N-coefficient 0.52-bit/pel	18	28	279	0.43

6.6 SUMMARY AND CONCLUSIONS

The Haar transform provides a practical method for removing redundancy from FLI images in real time. Four techniques for reducing data in the Haar domain have evaluated. A judicious assignment of coding bits for Haar sequency groups more than halves the bandwidth required to represent the transformed image. Zonal filtering provides a simple method for further reducing bandwidth by omitting e groups of coefficients, but zonal filtering is scene-dependent and tends to fav either edges or diagonals. Thresholding is a more efficient means of coefficie reduction. It is highly adaptive to different scenes; unfortunately, it requir variable length memory buffer and a complex encoder. A good compromise is the coefficient selection algorithm, which permits thresholding with a fixed-size b and a simple encoder. As a result of the experiments performed, Northrop recom- mends a combination of reducing coding bits assigned to Haar sequency groupings limited zonal filtering and the N-coefficient selection algorithm. When used w other reduction methods, i.e., frame rate reduction, this should provide a band reduction of 25 to 1 or greater for representative FLIR images without losing s targets in the background clutter.

SECTION 7
IMAGE DESCRIPTORS

7.1 INTRODUCTION

It is generally accepted that evaluation of processed imagery is a difficult problem. Mean square error has been used as a fidelity criterion for most image coding techniques and is believed not to compare well with human evalution. Mean square error does not make sense in the areas of image enhancement, for example. Also, in most cases of image restoration, it cannot be obtained since no reference image is available. Quantification of the effect of various image enhancment and restoration techniques requires a measure that can easily be obtained from the image and agrees well with the human subjective evaluation.

The development of a set of Image Descriptors (ID's) that are based on the statistics of an image's two-dimensional rationalized Haar transform is presented (see Section 1). The ID's define a division of pictorial information into two major categories: luminance (dc information) and activity (ac information). Luminance is further subdivided into brightness, given by the average gray level, and tone, which is determined from the variance of a local ensemble average gray level. A division of activity into texture and edge information is accomplished by examining an edge criterion function for each ac term of the image's Haar transform.

The ID's can be used to quantify the characteristics of images both before and after the application of various image processing algorithms. The resultant changes in the ID's can then be used to detect the net effect of processing on the image. When each ID is viewed as an element of a descriptor vector, an image description space is then defined. The evaluation of images within this description space was a part of this study.

The effects on the ID's observed, due to the application of various algorithms, are reviewed in this section. Some conclusions and recommendations based on the descriptors' performance are also discussed.

7.2 IMAGE DESCRIPTION IN THE HAAR DOMAIN

To facilitate the computation of the image descriptors, their definition must be compatible with the data handling format of the sensor processor. The transform format of the sensor processor is such that the Haar transform is evaluated for individual, nonoverlapping, 16-by-16 pel blocks. Local descriptors are, therefore, obtained for each block and then combined to form global descriptors that apply to the entire image or areas of interest within the image. The definitions of the local descriptors and their rules of composition to form the global- or image-descriptors are presented here.

7.2.1 Local Image Descriptors

7.2.1.1 Local Tone - The local tone of the i^{th} block of the digitized image is the average intensity of the block, and is represented by the dc term in the Haar domain:

$$t_i = a_{-1,1,-1,1} = \frac{1}{N^2} \sum_{y=0}^{15} \sum_{x=0}^{15} I(x,y) \qquad (7\text{-}1)$$

7.2.1.2 Local Activity - The local activity of the i^{th} block is defined as the ac energy of the block. This is given by a weighted sum of the sum of the squares of the ac terms of its rationalized Haar tranform:

$$c_i = \sum_{rmpn}\sum\sum\sum w_{rmpn} (a_{rmpn})^2 - (a_{-1,1,-1,1})^2 \qquad (7\text{-}2)$$

The coefficients $w_{r,m,p,n}$ are weights to measure the importance of transform coefficients $a_{r,m,p,n}$. For initial analysis, the weights were taken to be unity. The quantity c_i can be used as a measure of activity. If the subpicture has no detail, then $c_i = 0$. If it has a lot of detail, then c_i is large.

The term c_i by itself is insufficient to characterize the information in block i. Two blocks may, for example, have comparable activity when one contains significant edges (outline of an object, etc.) while the other contains only high-power white noise. For this reason, activity is further subdivided into texture activity and edge activity.

The separation of activity into texture and edge takes advantage of the local properties of the Haar basic functions. Within a sequency group, the Haar functions are simple translations of one another. By viewing each Haar function as an image primitive and the corresponding Haar sequency component as a measure of that function's contribution to the image, the concept of texture being a repetition of local image primitives may be extended to the Haar domain. When the total power of a sequency group is distributed relatively evenly among the elements of the group, each of the primitives contribute an equal amount of information and a texture is represented. When the group contains a few high-power "spikes," they are deemed to be the result of edges in the image domain. For example, consider the upper portion of Group 11 in Figure 1-5 (Section 1). If the power of each of the coefficients is equal, a texture of vertical stripes is represented. If there is one coefficient with significantly greater power than its neighbors, then an edge similar to the corresponding Haar function is present. The goal of the edge-texture separation procedure is to essentially sort out the edge terms in a group. The remaining power in the group is then, by default, due to texture.

7.2.1.3 Local Edge Activity - The local edge activity of the i^{th} block is the ac energy that satisfies the edge criterion function. In subjective terms, the edge criterion function dictates that if the power of an ac coefficient contributes more than its "fair share" to the total power of its sequency group, then that term is the result of an edge in the image domain. More specifically, if ψ denotes the set of edge terms, then $a_{r,m,p,n}$ is an element of ψ when

$$\frac{(a_{rmpn})^2}{\sum\sum_{mn} (a_{rmpn})^2} \geq \alpha_{rp} \; S_{rp} \qquad (7\text{-}3)$$

where α_{rp} is an empirical edge power threshold for the rp^{th} sequency group and S_{rp} is the relative entropy of the group. If $P(a_{rmpn})$ is the probability that a_{rmpn} will occur within the group, then

$$S_{rp} = \frac{\sum\sum_{mn} P(a_{rmpn}) \log_2 P(a_{rmpn})}{\log_2 (\text{number of terms in } rp^{th} \text{ group})} \qquad (7\text{-}4)$$

The values of α_{rp} are shown in Figure 7-1. As can be seen in the figure, only 15 groups are used. This reflects the assumption that horizontal and vertical edges are equally likely. The values of α_{rp} were chosen empirically, but approximate $1/1+K_{rp}$ where K_{rp} is the number of Haar functions in the rp^{th} group to which a straight edge would contribute power. This approximate relationship for α_{rp} allows diagonal edges to be detected more readily.

The purpose of weighting α_{rp} with the group's relative entropy can be seen by examining the two representative histograms of coefficients of the rp^{th} group shown in Figure 7-2.

The power that the edge terms would require in order to differ significantly from the other terms is greater in the flat distribution of Figure 7-2B than in the distribution of Figure 7-2A. A good measure of the distribution flatness is its entropy. The entropy is, therefore, used to weight the edge-power threshold.

If we let e_i be the local edge activity, then

$$e_i \triangleq \sum\sum\sum\sum_{rmpn} (a_{rmpn})^2, \; a_{rmpn} \in \psi \qquad (7\text{-}5)$$

The term $a_{-1,1,-1,1}$ can be excluded from equation (7-5) by ensuring that it will never satisfy the edge-criterion function.

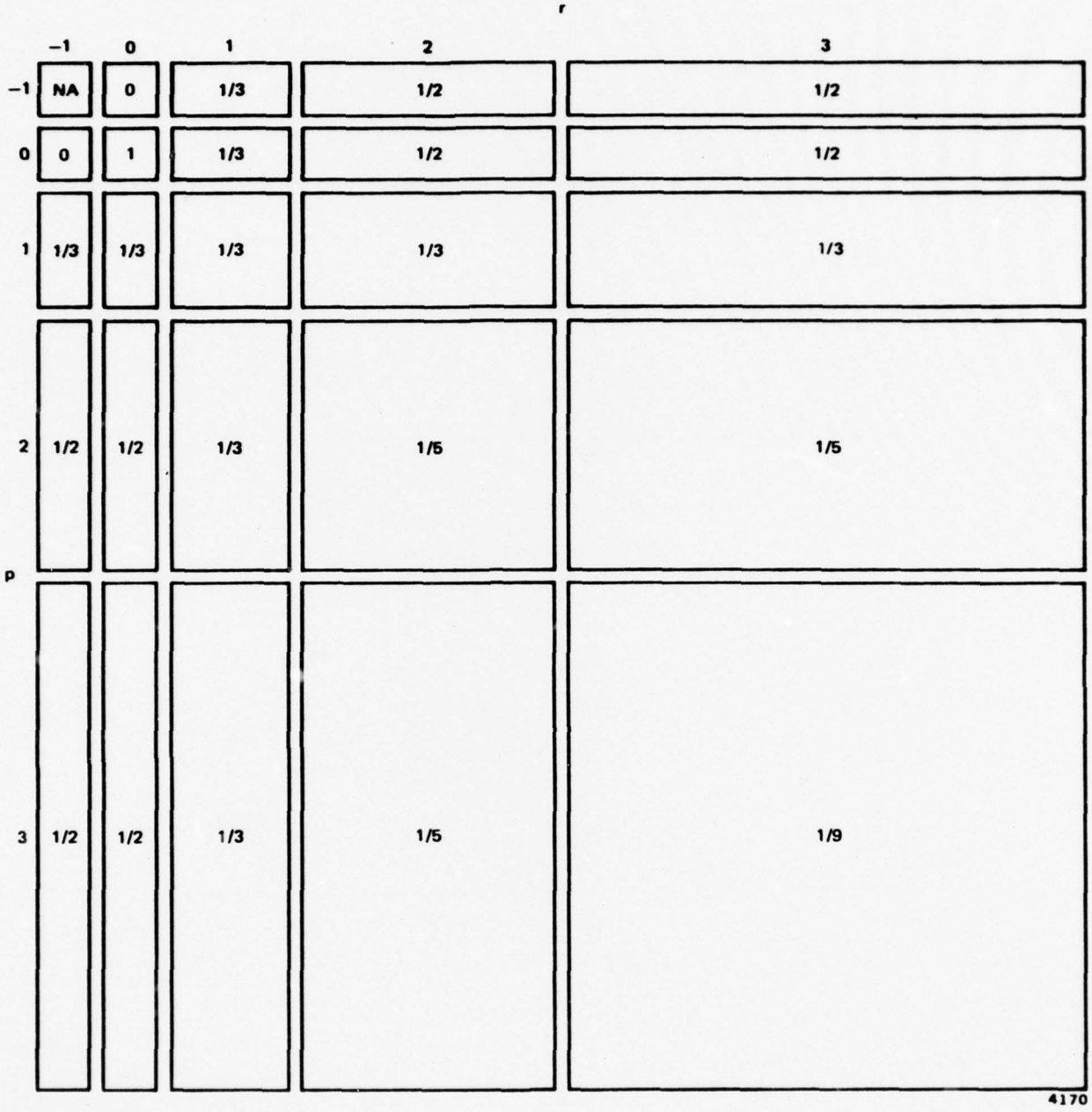
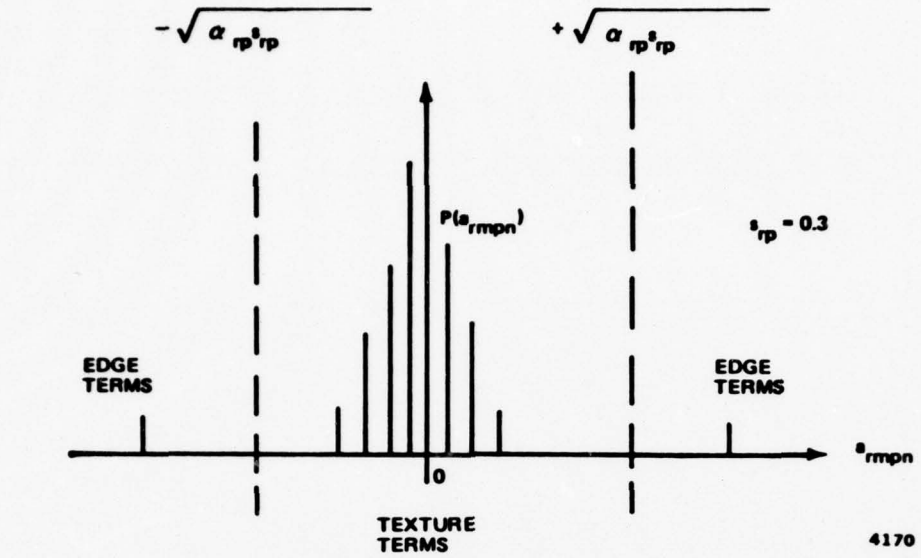


Figure 7-1. GROUP EDGE WEIGHTING, α_{rp}

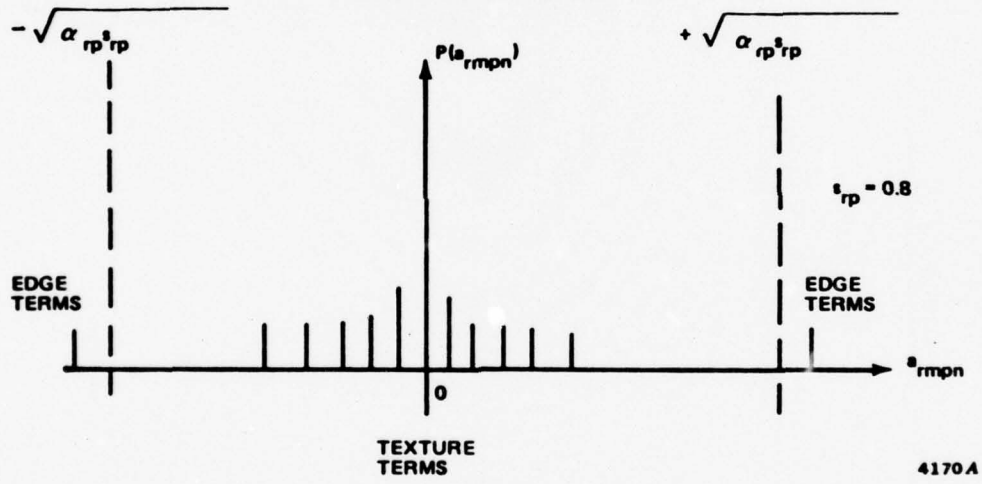
7.2.1.4 Local Texture Activity - Local texture activity is, by definition, the ac energy remaining in the block's Haar transform after all the edge terms have been removed.

If x_i is the local texture activity for the i^{th} block, then

$$x_i \triangleq \sum_{rmpn} \sum\sum\sum (a_{rmpn})^2 - (a_{-1,1,-1,1})^2, \; a_{rmpn} \; \forall \psi \qquad (7\text{-}6)$$



A. PEAKED DISTRIBUTION



B. FLAT DISTRIBUTION

Figure 7-2. REPRESENTATIVE HISTOGRAMS OF COEFFICIENTS OF j^{th} GROUP

7.2.1.5 Examples of Local Descriptors - The two sample images shown in Figure 7-3, "USC Girl" and "Moon," form a good data set for the expository development of the descriptors. It is desirable for the test data set to be kept as small as possible for demonstration purposes and yet be sufficiently diverse to exhibit a reasonable cross section of image characteristics. It is felt that the two selected images, Girl and Moon, provide enough local diversity individually and global diversity between each other to provide an insight into the behavior of the descriptor set.



U.S.C. GIRL



MOON

Figure 7-3. SAMPLE IMAGES

Each image consists of 256 x 256 pels that are subdivided into a 16 x 16 array of 16 x 16 pel blocks. The local descriptors were computed for each block of both sample images. Representative results for the Girl image are shown in Figures 7-4 and 7-5.

Examination of the blocks relating the local tone of the Girl image (Figure 7-4) shows that the results obtained are in good agreement with subjective evaluation. This is to be expected since the local tone is nothing more than the average gray level of the 16 x 16 pel block. Dark areas of the picture have a correspondingly lower local tone than bright areas. Subjective agreement is also good for local activity. Busier areas of the picture have correspondingly higher values of activity.

Figure 7-5 provides a good example of the division of edge and texture activity. The two blocks shown exhibit comparable activity and yet contain significantly different information. The block containing the edge of the Girl's face exhibits high edge activity (edge:texture ratio is 3.5:1) while in the block containing the flower, texture activity is the majority contributor (edge:texture ratio is 1:1.3).

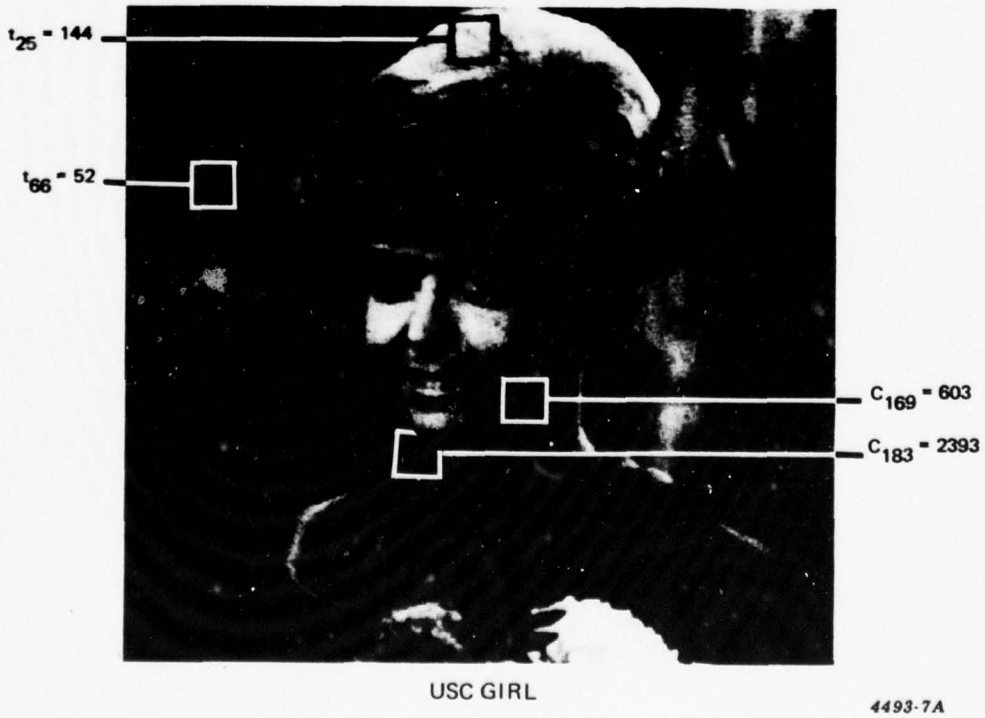


Figure 7-4. EXAMPLES OF LOCAL TONE AND ACTIVITY

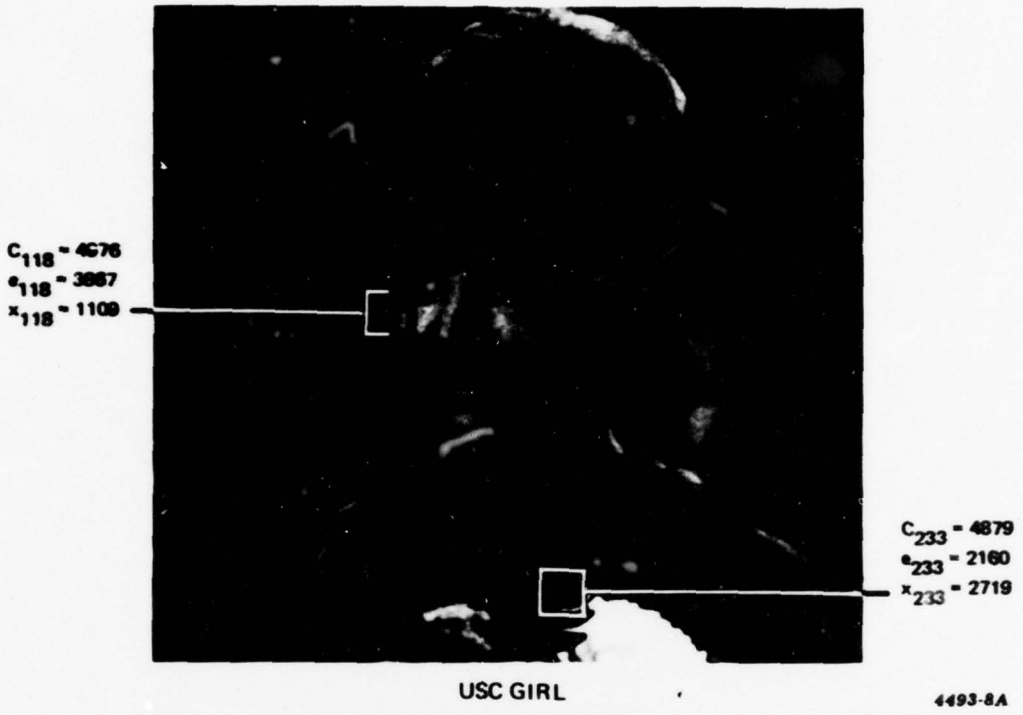


Figure 7-5. EXAMPLES OF LOCAL TEXTURE AND EDGE ACTIVITY

7.2.2 Global Image Descriptors

Obtaining global descriptors of tone, edge activity, and texture activity requires a suitable combination of the quantities computed on the subpictures of the image. Since the edge and texture activity add up to local activity numerically, only the results of local tone and activity will be developed herein. The rules for the combination of local edge and texture activity will be equivalent to those for the combination of local activity.

The local tone and activity have been computed on all of the 256 blocks of sample images Girl and Moon. A global index must be obtained from the set of local values. To aid in the comparison of the images, some first-order statistics of the local descriptors were computed as shown in Table 7-1.

Table 7-1. STATISTICS OF LOCAL TONE AND ACTIVITY

	Girl	Moon
Local Tone		
Mean	73	127
Minimum	10	30
Maximum	205	178
Standard Deviation	35	22
Local Activity		
Mean	2,367	2,065
Minimum	130	489
Maximum	23,606	9,975
Standard Deviation	3,189	1,391

7.2.2.1 Image Tone - As can be seen from Table 7-1, the mean of the local tone (average brightness) is greater for Moon than for Girl. Subjective evaluation, however, indicates that the Girl image has the most global tone. Image tone should relate the diversity of the local brightness rather than the brightness itself. The standard deviation of the local tone then is taken as a measure of the global, or image, tone. Let the image tone be denoted by T; then

$$T \triangleq \sigma_{t_i} = \left(\frac{\sum_i (t_i - B)^2}{N - 1} \right)^{1/2} \tag{7-7}$$

where t_i is the local tone of the i^{th} block, B is the average brightness of the image, and N is the number of 16 x 16 pel blocks in the image. The source of the difference in standard deviations of the local tone of the two images can be seen readily by examining the histograms of local tone shown in Figures 7-6 and 7-7.

The preceding expression for image tone exhibits the following characteristics: When the local tone of each image block is approximately equal, i.e., the total picture is of uniform average intensity, the image tone is near zero. When the local tone changes significantly from block to block, the image tone is large.

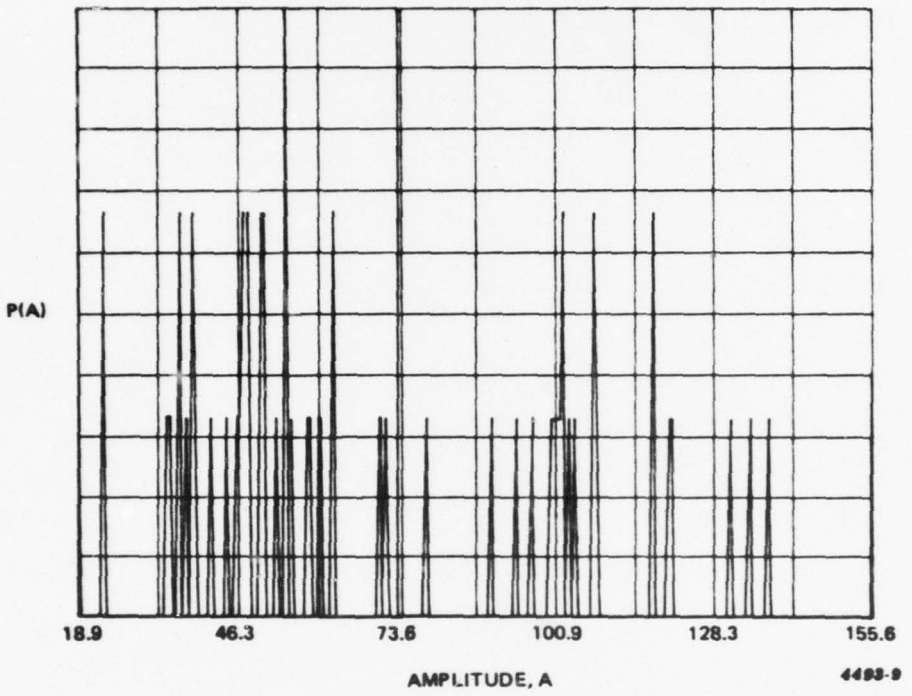


Figure 7-6. HISTOGRAM OF LOCAL TONE OF GIRL

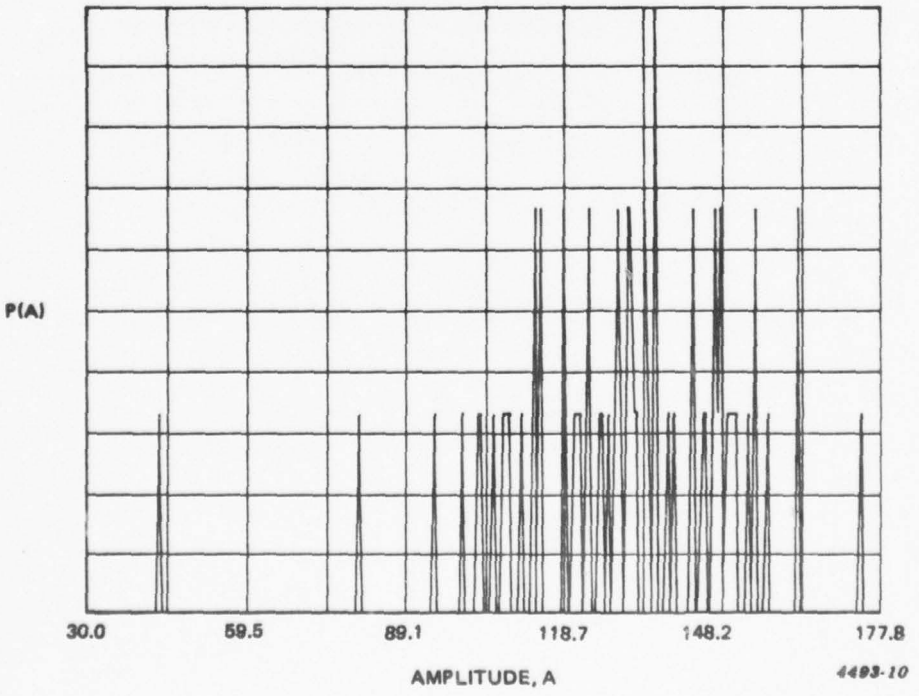


Figure 7-7. HISTOGRAM OF LOCAL TONE OF MOON

7.2.2.2 Image Activity - A comparison of the statistics of the local activity of the Girl and Moon images (Table 7-1) reveals a significant difference in their standard deviations. The standard deviation of the local activity is not, however, taken as a measure of image activity. The reason for this decision becomes clearer when we consider the following example: In a given image assume that each 16 x 16 pel block contains either sufficient edge or texture information to give it a local activity of c_o, a constant. Each block may be completely different from its neighbors; the only similarity is their local activity. If the standard deviation of the local activity were the measure for image activity, the image activity would be zero. But, if one-half of the blocks of this image were changed to a uniform gray, i.e., made to contain no ac energy, the standard deviation of the local activity would increase. Such a result is counterintuitive to a measure of image activity. The standard deviation of the local activity does, however, impart some information about the total image. This parameter is dubbed "image diversity" but, along with image brightness, is not presently considered a necessary element of the global image descriptors.

The measure to be used for image activity is the mean of the local activity. Image activity thus measured would rise and fall in agreement with subjective evaluation in the example given. Examination of Table 7-1 shows that the Girl image is about 15 percent more active than the Moon image. Let C denote the global, or image, activity,

$$C = \frac{1}{N} \sum_i c_i \qquad (7\text{-}8)$$

where c_i is the activity of the i^{th} block. Similar expressions apply for X, the image texture activity and E, the image edge activity; i.e.,

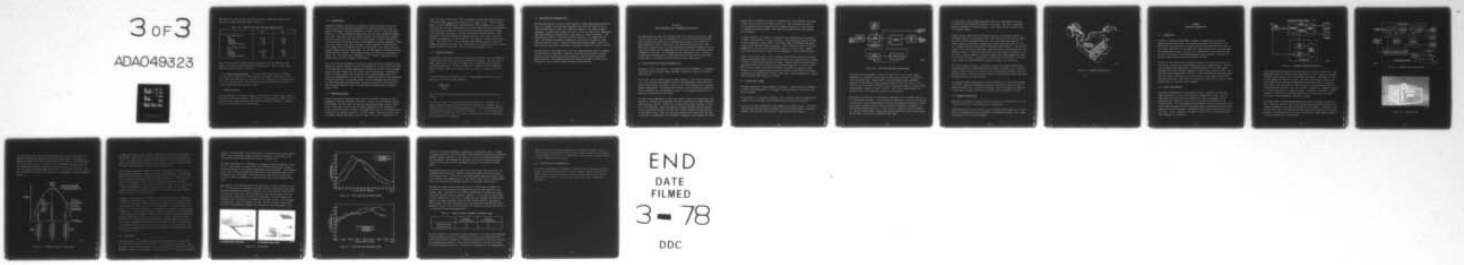
$$X = \frac{1}{N} \sum_i x_i$$

$$E = \frac{1}{N} \sum_i e_i$$

where x_i and e_i are the local texture and edge activity respectively.

AD-A049 323 NORTHROP CORP ANAHEIM CALIF ELECTRO-MECHANICAL DIV F/G 17/5
EXPERIMENTAL DEVELOPMENT OF A FLIR SENSOR PROCESSOR.(U)
SEP 77 T LEIBOFF, H HENNING, T NODA, B DEAL DAAG53-76-C-0188
UNCLASSIFIED NORT-77Y196 NL

3 OF 3
ADA049323



END
DATE FILMED
3-78
DDC

The statistics of the division of local activity into edge and texture activity for the two sample images are shown in Table 7-2.

Table 7-2. STATISTICS OF LOCAL EDGE AND TEXTURE ACTIVITY

	Girl	Moon
Local Edge Activity		
Mean	1,191	850
Maximum	15,147	4,937
Minimum	99	195
Standard Deviation	1,754	737
Local Texture Activity		
Mean	1,176	1,214
Maximum	14,675	5,783
Minimum	5	223
Standard Deviation	1,635	776

Table 7-2 shows that the global texture activities of the two images are comparable, but the Girl image displays more edge activity. This is in good agreement with initial subjective analysis.

7.2.2.3 Global Descriptor Vector The three global features used as image descriptors are T, image tone; E, image edge activity; and X, image texture activity. These three descriptors are defined to be the components of the image descriptor vector, $\underline{ID}$, $\underline{ID}=[T,E,X]$. Thus, the descriptors of an image define a point in a three-dimensional image descriptor space.

7.3 DESCRIPTOR RESULTS

In each experiment the image descriptors have been used to evaluate image content both before and after processing. Specifically, the descriptors of images of various states of focus, edge enhancement, and compression have been computed. A summary of each of these results is presented below.

7.3.1 Focus Results

During the autofocus experiment, the behavior of texture and edge activity as a function of lens position was determined for the video test pattern as well as several FLIR images. This was performed in a fashion similar to the other focus measures, i.e., by evaluating the focus measure at each point through one complete lens cycle. The area of the test pattern for which the descriptors were computed was more inclusive than that used for the other measures. Figure 2-13 shows the test pattern with the descriptor input areas flagged by increased brightness. These areas were chosen so that texture and edge would be computed for both the center of the image, which is primarily texture, and the borders of the image, which are composed primarily of two singular edges. On the FLIR iamges, both target and background areas were examined. The background areas are identical to those shown in Figure 2-15. For target input information, a 3 by 4 block of image sectors was centered about one of the targets in each image.

The curves relating the behavior of texture and edge activity as a function of lens position are shown in Figures 2-14 and 2-27 through 2-30. It can be seen that while texture and edge activity show approximately the same sensitivity to focus, regardless of the type of input data, the texture curve is much noisier. Averaging the values of texture activity over several lens cycles would probably reduce this noise level substantially, but the increase in computation time for such an approach would be prohibitive in a real-time system. For the background type data seen in the sample FLIR images provided by NVL, the division of activity into texture and edge information provides little to no additional information about the image's state of focus.

7.3.2 Enhancement Results

The ID's were used in conjunction with subjective analysis to evaluate the edge enhancement algorithms developed in Section 4. The ID's for the enhanced versions of FLIR Image 1 given in Table 2-6 show that edge-enhanced images possess substantially larger edge activities than non-enhanced imaged with only a marginal increase in texture activity. However, some anomalies do exist in the relative levels of edge activity between subjectively equivalent images. Figures 2-15a and 2-15b for example, exhibit edge-texture ratios of about 10:1 and 2:1 respectively, yet

appear the same on the display. This inconsistency in edge description is due to the uniform weight w_{rmpn} given each Haar term in equation (7-2). With a uniform weight, certain sequency groups containing mostly noise information (e.g., Groups 14 and 15) may have their contribution to the total image description overemphasized This problem was circumvented to some extent by not subjecting the noise groups to the enhancment algorithms. The noise contribution from these groups is then equivalent in both the original and enhanced versions of the image. Using this approach, an edge-texture ratio of about 2:1 was regularly preferred by the evaluators. This result will be altered slightly when alternate sets of coefficient weights are developed that will more closely match the contribution of each sequency group to the MTF of the system. This is a current development effort.

7.3.3 Compression Results

The behavior of the image descriptors resulting from different levels of bandwidth reduction was evaluated as part of the compression experiment. The basic goal of image compression is to suppress redundant texture information while retaining pertinent edge features. As shown in Table 6-1, increased levels of compression reduce an image's edge content slightly while the texture descriptor falls off rapidly.

A figure of meter (f) was devised to offer a rough grading of the effects of compression on the different image samples:

$$f = \frac{(\bar{E}/\bar{X})\ (b/8)}{\bar{E}_o/\bar{X}_o}$$

where b is bits/pel and $\bar{E}_o$ and $\bar{X}_o$ are edge and threshold average descriptors of original.

In almost all cases, f decreased with increasing compression. Looking at the images (Figures 6-7 through 6-19), one can see deterioration increasing, with the compression level further validating the values of f. This relationship holds for the test pattern as well as each of the FLIR images. If f = 0.5 were arbitrarily selected for a cut-off value, a data rate of one bit per pel would be obtainable, using the N-coefficient algorithm.

7.4 CONCLUSIONS AND RECOMMENDATIONS

The image descriptors can provide useful measures of image change caused by processing. For example, the enhancement experiments have shown that edge-to-texture ratios increase when the edges of an image are enhanced. Also, when an image is compressed, a quality index, based on the ID's, can give an index of the compression algorithm's ability to conserve edges at the expense of texture information. The edge-texture division is relatively insensitive, however, to an image state of focus. The autofocus experiments demonstrated that activity in general, rather than just texture or edge, is reduced by mildly defocusing an image. Blurs of a large, spatial extent must be examined to determine the effects of focus on the ID's.

The contribution of each Haar term to the total image description must be weighted to better match the system MTF. When the uniform coefficient weights currently defined are used, system noise can be incorrectly characterized as significant edge information. By assigning weights that minimize the effect of noise on the descriptors, a more realistic image descriptor may be obtained.

SECTION 8
SENSOR PROCESSOR UNIT - DEMONSTRATION PROTOTYPE

The image-processing algorithms developed under this contract and Northrop's IR&D projects form the basis for the development of a test and demonstration prototype Sensor Processor Unit (SPU) and for applying these concepts to the present U.S. Army Common Module FLIR as configured for the Advanced Attack Helicopter (AAH). Northrop was selected by the U.S. Army Development and Readiness Command for development of the Target Acquisition/Designation System and a Pilot Night Vision System (TADS/PNVS) for the AAH. The TADS/PNVS uses the common module FLIR. Therefore, a demonstration "brassboard" of a FLIR Sensor Processor Unit would be both timely and pertinent to the Army's Advanced Attack Helicopter Weapon System.

8.1 SYSTEM DESCRIPTION, SENSOR PROCESSING UNIT

The SPU will serve as an adjunct to the Common-Module FLIR (MODFLIR). Its primary objectives include automatic image enhancement and multiple-target detection and classification.

To this end, specific design goals have been enumerated. They include dc restoration, local area gain and brightness control, edge enhancement, target detection, target classification, and cueing. These, in turn, have prompted several important design considerations: interfacing with the FLIR, real-time operating capabilities such as automatic focus, freedom from artifacts, and acceptable probabilities of detection, classification, and cueing.

In terms of major operations, the SPU first seeks to enhance the image before interrogating it. Effectively, enhancement increases accessibility of image information before actually using it. To this end, it employs Haar image transforms. In general, Haar transforms simplify the approach to image enhancement by dividing the picture into global and local functions that somewhat resemble desired structures in the enhanced scene. Thus, dc restoration uses multiple dc Haar functions from a horizontal one-dimensional (1-D) transform, brightness control employs the single

dc Haar from a horizontal and vertical two-dimensional (2-D) transform, while gain control and auto focus use all Haar terms but dc. Haar transforms are used to an even greater extent in target detection, where the ability of local functions to represent restricted regions of the image while ignoring other areas becomes quite important.

In target detection, the object is to apply multiple algorithms that discriminate in favor of likely points or areas of interest. These algorithms make use of high relative temperature, sharp corners, and straight edges to distinguish man-made targets from a nominally natural background. Their combined use helps to increase probability of detection while decreasing probability of false alarm. Weighted sums of individual detector outputs will be compared to a preset threshold level.

Target classification operates only on objects selected during the detection phase. Relatively simple operations on this transform enable size, translation, and rotation invariance in distinguishing different targets. Target-recognition output overlays appropriate display cues by means of the flag-symbol generator and scan converter. The output also feeds back within the function to minimize occurrence of false alarms, and to the target-detection function to enable a self-constructed library. In addition, the target-detection output feeds back to image enhancement as a further means of keeping false alarms to a minimum.

8.2 SYSTEM BLOCK DIAGRAM

The basic system block diagram appears in Figure 8-1. Image data from the MODFLIR is digitized and presented in a rectangular row-column format. Each row is then Haar-transformed.

At this point, dc restoration is applied. This operation corrects the degradation caused by the ac coupling of each photodetector in the parallel-scan FLIR.

A vertical 1-D Haar transformation then completes the 2-D Haar matrix. This operation must simultaneously draw on information from consecutive pairs of interlaced sweeps of the race-track scan and uses a field store for this purpose.

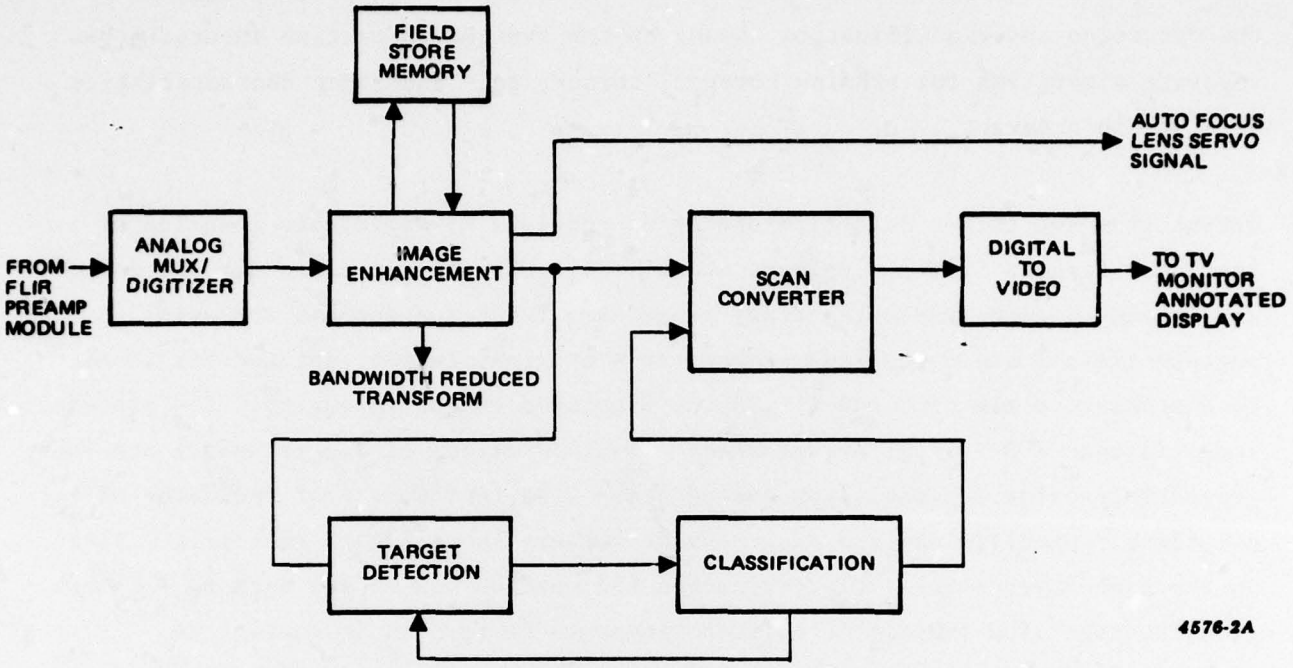


Figure 8-1. TOP-LEVEL FUNCTIONAL BLOCK DIAGRAM

Local area gain and brightness control now takes place in the 2-D Haar domain. Brightness is controlled by the dc Haar term while gain (contrast) is regulated by the remaining terms. Edge-extraction, edge-enhancement and/or bandwidth-reduction is then applied. These operations apply to each local area, or 16 x 16-pel sector, into which the overall frame has been partitioned. A special averaging algorithm is applied to reduce artifacts. Here, servo command signals for automatic focus of the FLIR lens system are also generated.

The improved 2-D Haar data is subsequently converted back to the image domain via the inverse Haar transform. It will produce an image that is free of dc-restoration-type problems. The resulting image will further show moderated responses to brightness changes and scene contrast variations that are designed for optimum use of the available display gray scale. A scan converter changes the image information to a 525-line serial format. Appropriate target-designation flags are superimposed and the result converted to an analog TV-type signal.

A second output from the image enhancement functions simultaneously activates the detection and classification phases of the system. Detection incorporates separate algorithms for sensing hotspot, corner, edge and other characteristics of a possible target.

Output from the target detection operation consists of coordinate location of available targets in the displayed scene. This effectively calls out the immediate locale of each available target as an area requiring further analysis. A pattern limiter and normalizer expands each original target spot and its locale to a standard scale size and limits the displayed target intensity. The processed image is then 2-D Fourier transformed. Absolute values of the transform are integrated to provide a translation and rotation invariant numerical indicator of target class. Specific numbers or groups of numbers then trigger appropriate flags in the symbol generator. Classification information can be fed back to the detection, storage, and limiter-normalizer stages to improve their operations.

The overall system, therefore, presents an enhanced FLIR image with appropriate flags designating and classifying particular targets within the scene. Different options, such as target motion detection with a Moving Target Indicator (MTI) and multiple limiters, and their practical effects in reducing false alarms and improving detection will be studied further.

The proposed hardware implementation involves maximum use of Large Scale Integration (LSI). High-speed microprocessor modules have been selected that offer a powerful capability for flexibility and growth. This technology is an extension of many years of design, fabrication, and testing of microprocessor systems.

8.3 HARDWARE CONFIGURATION

The sensor processor unit (Figure 8-2) will consist of an enclosure, five 17- by 7-inch circuit cards, cooling fans, and a wiring harness.

Although this unit will not be built to meet military specifications, it will be built to withstand flight test and expected environmental extremes. Best commercial practices will be used throughout.

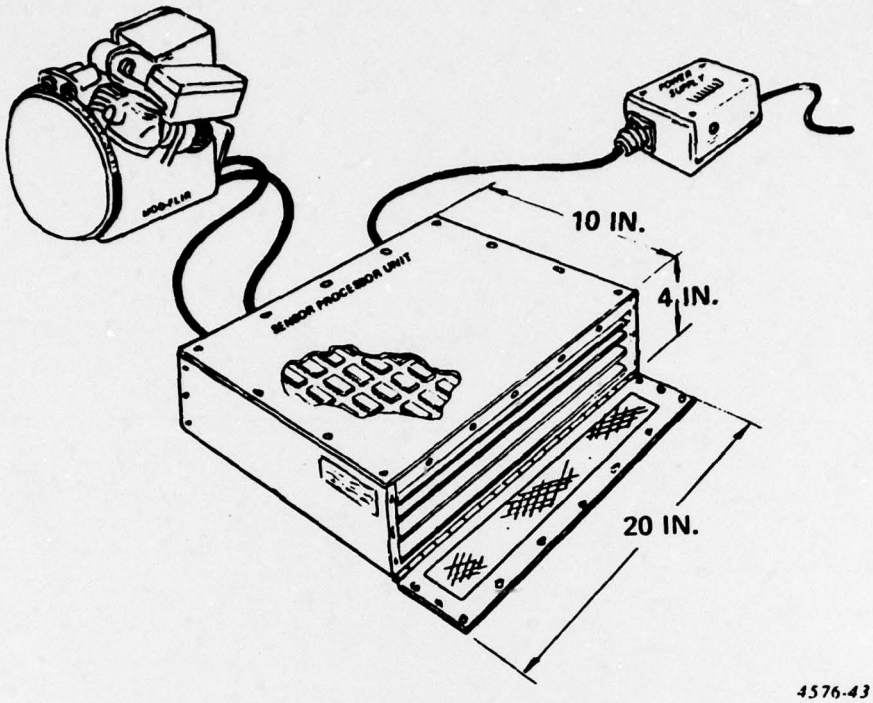


Figure 8-2. HARDWARE CONFIGURATION

ADDENDUM
AUTOFOCUS DEMONSTRATION

A-1. INTRODUCTION

In Section 2 (Autofocus Experiments) a method of automatically focusing an imaging system was developed. The procedure developed for detecting and correcting the focus condition of an image consists of measuring the total power of the ac Haar coefficients of the image while controlling the lens focus adjustment by electromechanical means. As the lens is defocused (in either direction from optimum focus position), the observed sequency power decreases. This results in an error signal determining the optimum focus position.

In order to fully evaluate this method of autofocus, an autofocus unit was constructed and interfaced with the two-dimensional Haar sensor processor. The autofocus system was then demonstrated at Wright Patterson Air Force Base during September 1977 using a Honeywell Chaparral Forward Looking Infrared (FLIR). The FLIR and autofocus system were located in a tower overlooking the base and beyond. Tests were run at different hours of the day and at night, under varied thermal radiating conditions. The details of the autofocus system and the results of the demonstration are presented in this addendum.

A-2. SYSTEM CONFIGURATION

The overall configuration of the autofocus system is presented in Figure A-1. Image information is output from the FLIR as standard RS-170 video. The two-dimensional Haar sensor processor digitizes the video and computes the two-dimensional Haar transform of each 16 x 16 pel sector in each field of the digitized image. Both the digitized and transformed video are then routed to the inverse Haar processor where the transformed image is reconstructed and displayed along with the digitized image. Transformed video is also accepted by the autofocus unit. Here, a measure of the image's state of focus is evaluated and lens servo commands are originated.

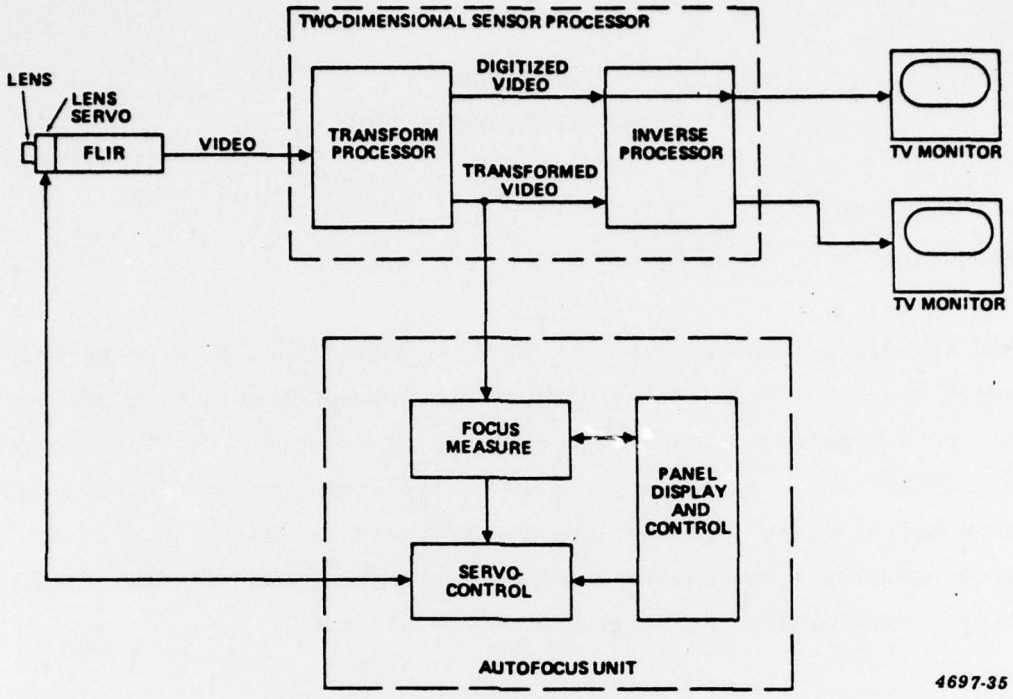


Figure A-1. AUTOFOCUS SYSTEM CONFIGURATION

A more detailed system diagram of the autofocus unit is given in Figure A-2. The transformed video, in the form of Haar coefficients, is first input to a hardware multiply circuit where all the ac coefficients are squared. The square ac terms are then summed over the entire sector. The number and location of the sectors to be summed in this fashion are determined by front panel control settings on the autofocus unit, see Figure A-3 (any location on the video monitor and up to 15 sectors may be selected). After the ac power is summed over a number of sectors, it is then integrated over a number of fields. The amount of temporal, or field, integration is also set on the front panel (up to 15 fields may be selected). The output of the field memory is the focus measure.

The focus measure is both displayed and used as input to the servo control function. The servo control in turn focuses the system by continually moving the lens by small increments in an effort to maximize the focus measure. Although the system will never come to rest precisely at the optimum focus position utilizing this approach, the lens excursions about that position may be small enough to present no visible degradation in image quality.

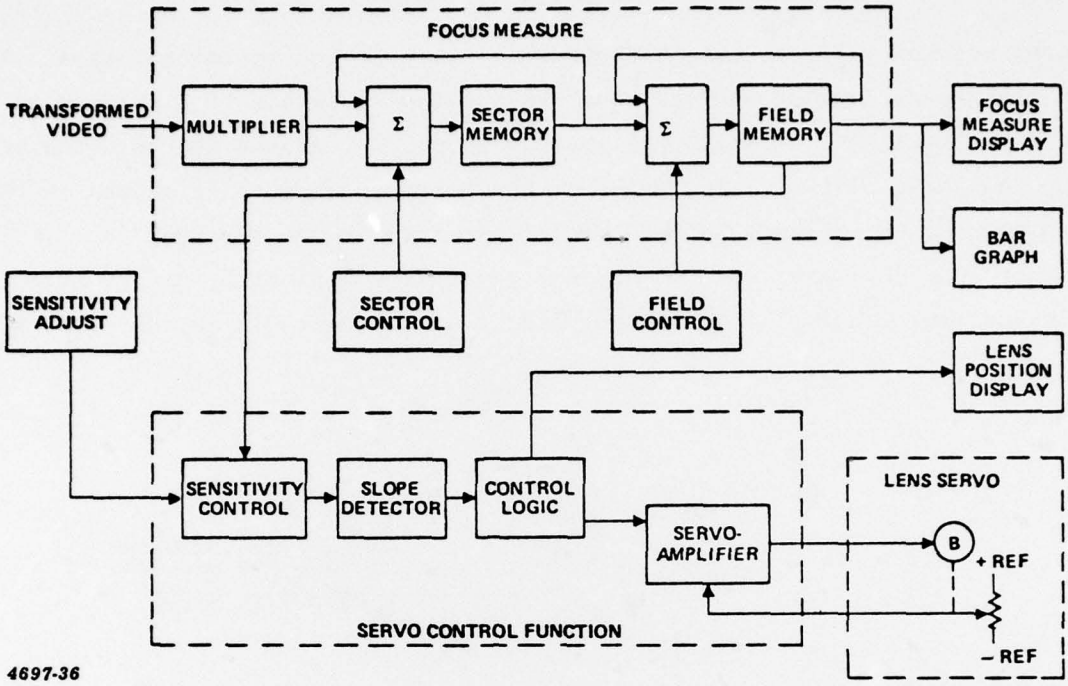


Figure A-2. AUTOFOCUS UNIT BLOCK DIAGRAM

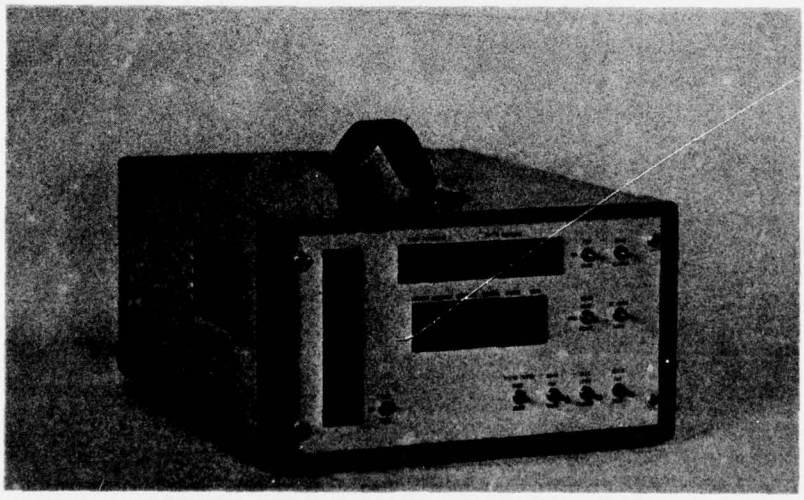


Figure A-3. AUTOFOCUS UNIT

The servo control may seek out the optimum focus position in several ways. A method that keeps lens excursions about the optimum position to a minimum is depicted in Figure A-4. Small oscillations are first induced in the focus adjustment by the servo control. The phase of the corresponding oscillations in the focus measure then indicates which side of the focus curve the lens is on. The focus is adjusted towards optimum and the phase is reevaluated for the next step. When the optimum focus position is reached, a second harmonic in the focus measure appears.

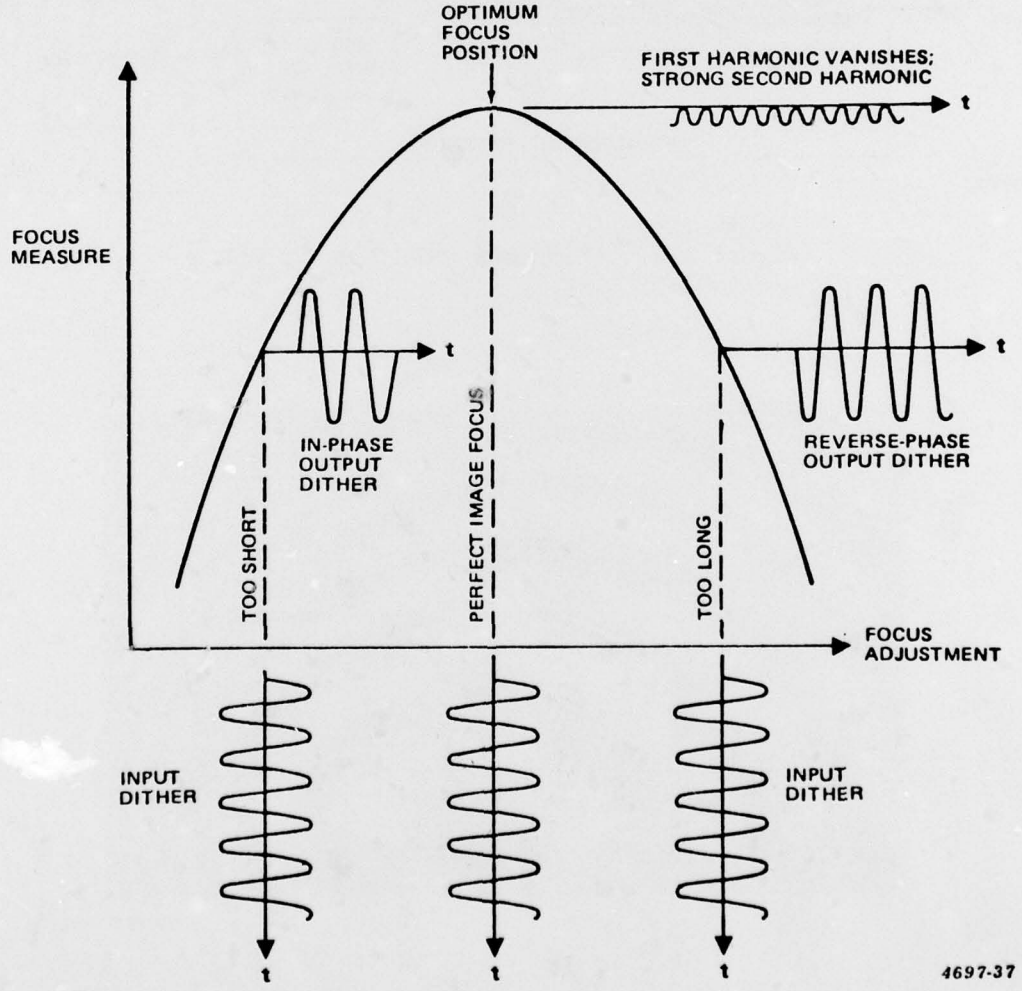


Figure A-4. DITHER SENSITIVITY OF IMAGE FOCUS

Although this method of focus control is efficient in terms of noise rejection and economy of lens travel, it places several restrictions on the focus servo system. For example, since small rapid oscillations of a lens are required, the mass of the focusing lens is limited. Also, backlash of the focus gear train may introduce large nonlinearities into the focus control system.

A more generally applicable method of seeking the optimum focus position is to move the focus adjustment in one direction at a uniform rate and, if a downward trend in the focus measure is observed, change directions. The system will therefore be constantly dithering across the optimum focus position. This approach moves the lens slightly farther from the optimum position than the previous method. By using small focus adjustment increments, however, it is still possible to keep any degradation in image quality below an observable level. Previous autofocus experiments indicate that a 15 percent reduction in the focus measure cannot be detected on the video monitor.

A degree of noise rejection is achieved by controlling the system's sensitivity to changes in the focus measure. The servo control logic will reverse the direction of lens travel only if three consecutive decreases in focus measure are experienced and the decreases are of sufficient magnitude. The thresholding of the amount of focus measure degradation is performed by the sensitivity control shown in Figure A-2. The sensitivity control functions by adding a fixed threshold value to each incoming focus measure. The amount of focus measure degradation must then exceed the threshold in order to be indicated as a negative slope by the slope detector. This method of reversing direction on focus degradation was used for the autofocus demonstration. This was necessary since the lens servo mechanism of the FLIR consisted of an unspecified mass driven by an unspecified motor. The gear train connecting the motor and its optical element also exhibited a 4 percent of full-scale backlash.

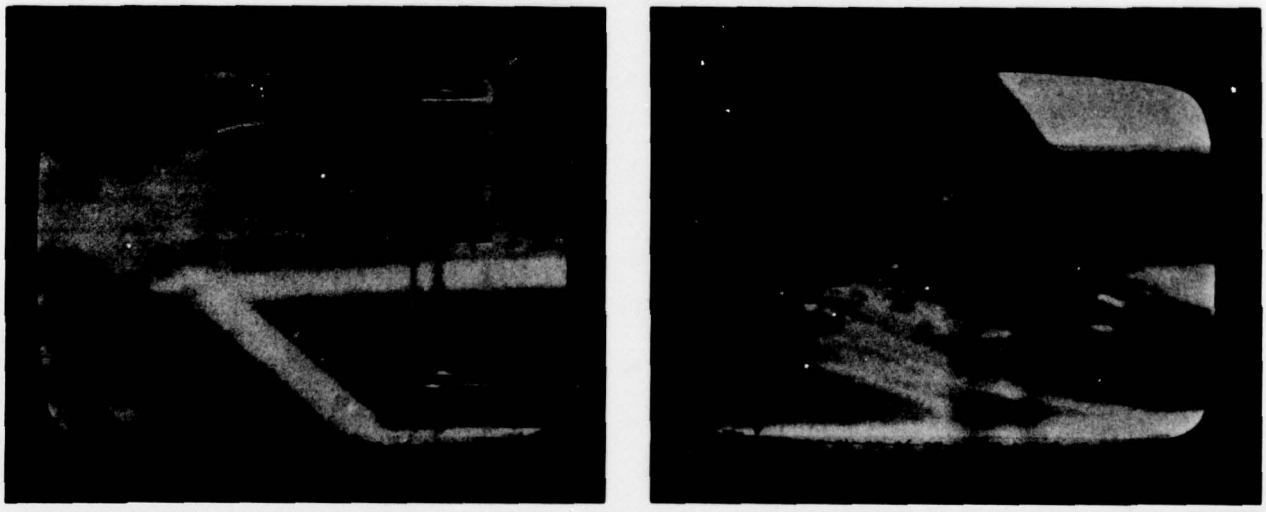
A-3. SYSTEM DATA

As a preliminary to the automated operation of the focusing system, static data was collected to evaluate the focus measure as a function of lens position. This was done to validate the laboratory concept of autofocus and to determine the system parameters relevant to the automatic operation of the system. Such system parameters

include: 1) noise levels in the focus measure, 2) sensitivity and scaling requirements to detect meaningful changes in the focus measure, 3) uncertainty of the lens position, and 4) the desired lens position increment size.

Two image areas chosen for consideration are flagged by reduced brightness in Figure A-5. They consist of a radio tower in the foreground (approximate range 50 meters) and garage in the background (approximate range 300 meters). The background focus scene is focused at about 20 percent of full-scale adjustment away from the infinity setting. The degree of focus difference between these two ranges may be seen by observing the portion of defocused radio tower in the background focus scene.

Focus measure curves derived from the two image areas are shown in Figures A-6 and A-7. The focus measure curves were generated by manually increasing or decreasing the lens adjustment servo command in discrete steps, and observing a sample of the focus measure at each step. As a starting point, the step size was set such that 100 steps were required for full-scale focus adjustment. The focus measure was integrated over four sectors of the image and two frames of video. The corresponding update rate of the focus measure is 15 Hz. Two sets of data, increasing and decreasing focus adjustment command, are presented for each image segment. This



A. FOREGROUND FOCUS—RADIO TOWER

B. BACKGROUND FOCUS—GARAGE

4697-38

Figure A-5. FOCUS AREAS

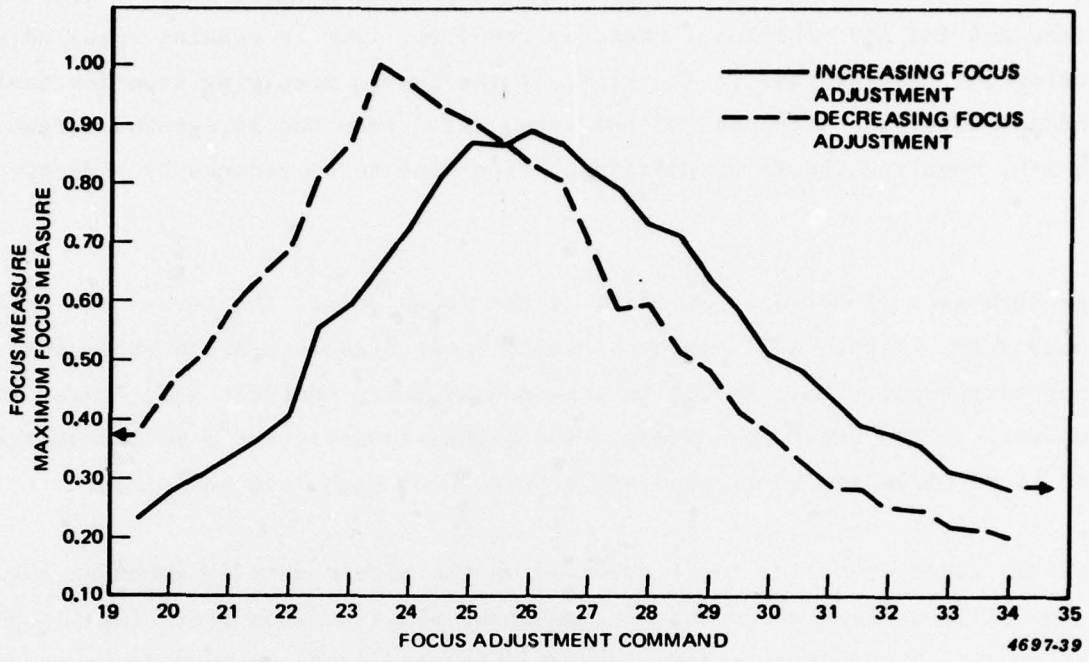


Figure A-6. FOCUS CURVE FOR FOREGROUND SCENE

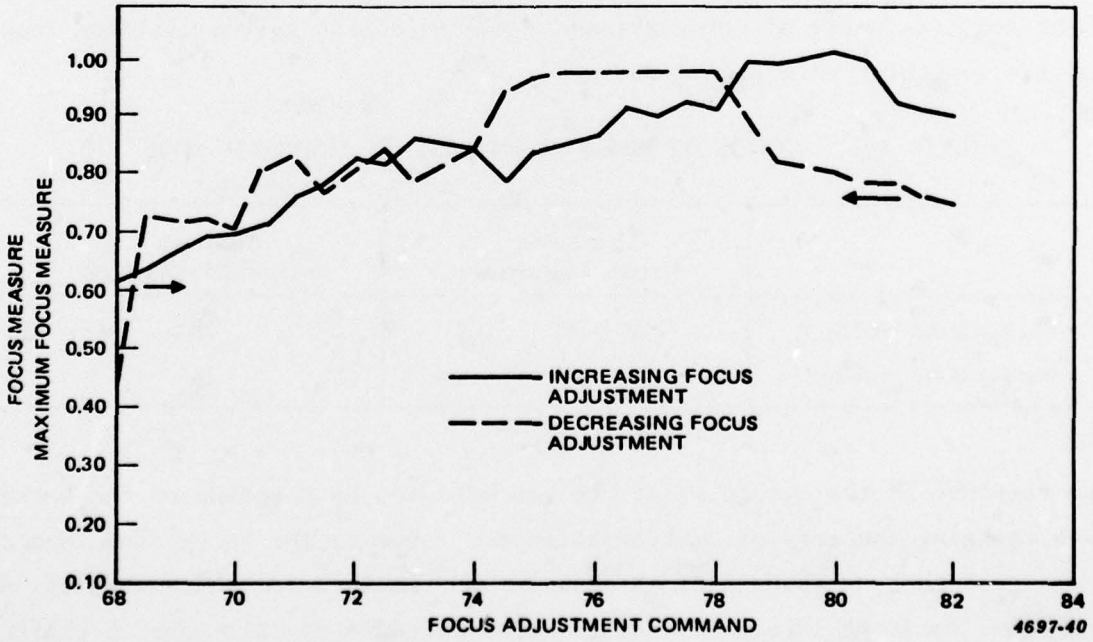


Figure A-7. FOCUS CURVE FOR BACKGROUND SCENE

was done to illustrate the degree of backlash in the focus gear train. As shown in Figures A-6 and A-7, the focus measures resulting from increasing focus adjustment commands exhibit a peak well to the right of the curves resulting from decreasing focus adjustment. The steepness of the focus curve from the foreground target (Figure A-6) required the focus adjustment step size to be reduced by a factor of two.

Although both sets of curves show peaks at the focus point, the curves due to the background scene (Figure A-7) exhibits a much lower signal-to-noise ratio than the foreground scene. This is due to the reduced image contrast and, hence, lower focus measure in the background scene. The signal-to-noise ratio in the background scene is 14 dB while the foreground exhibits a 20 dB signal-to-noise ratio.

Although the signal-to-noise level observed in the static data is somewhat low, the system still dithers about the optimum focus position when operating in the automatic mode. The limits of focus dither coresponding to maximum focus measure sensitivity for both scenes are given in Table A-1. A lens travel of about 8 positions was experienced in each case. Although the focus adjustment varied by these amounts, the image's subjective state of focus remained constant. By using the median value of the focus adjustment as the focus point and calibrating the focus adjustment scale in units of focus distance, the autofocus system can then function as a passive monocular autoranging device.

Table A-1. LIMITS OF FOCUS ADJUSTMENT IN AUTOMATIC MODE

	Maximum Focus Adjustment	Minimum Focus Adjustment
Foreground Scene	82	74
Background Scene	30	24

The step response of the autofocus system was measured by focusing on the background scene and changing the area of consideration for focus to the foreground target. It was observed that an average of 12 seconds was required for the system to identify the foreground focus position. The system step response in this case is limited by the lens adjustment step rate. The system, as configured for the demonstration,

adjusts the lens in half-step increments once every 350 milliseconds. This is done to minimize the workload of the lens servo. However, since the focus measure is being updated at 15 Hz, a 5 to 1 improvement in step response could be realized by using a tighter lens servo system.

A-4. CONCLUSIONS AND RECOMMENDATIONS

The autofocus demonstration has shown the feasibility of an operational autofocus system. The observed noise values of the focus measure identify the need for a degree of noise filtering before the focus is evaluated. The speed of response and positional stability of the lens servo system also requires significant improvement.