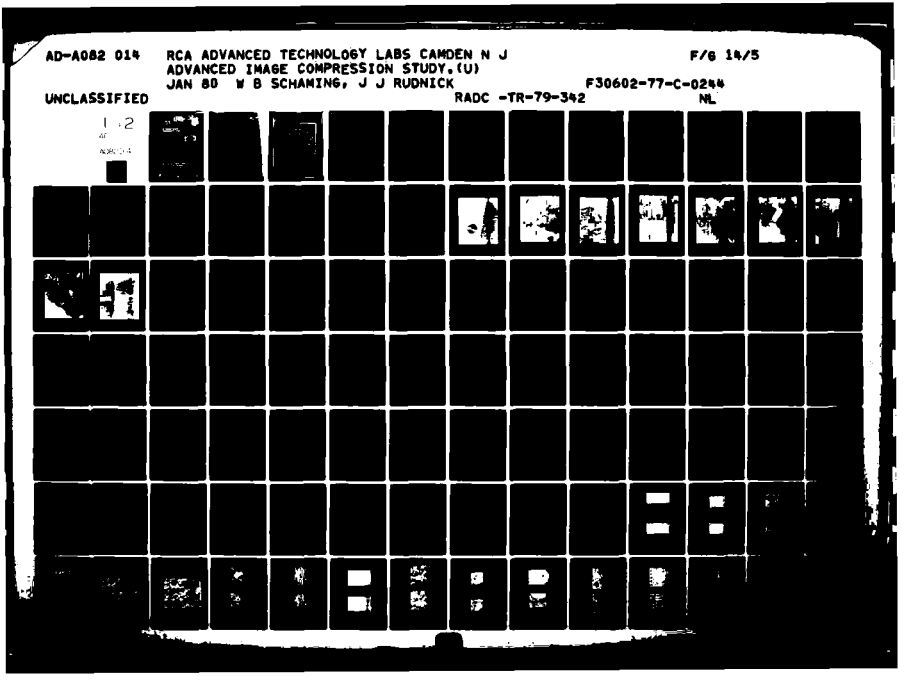
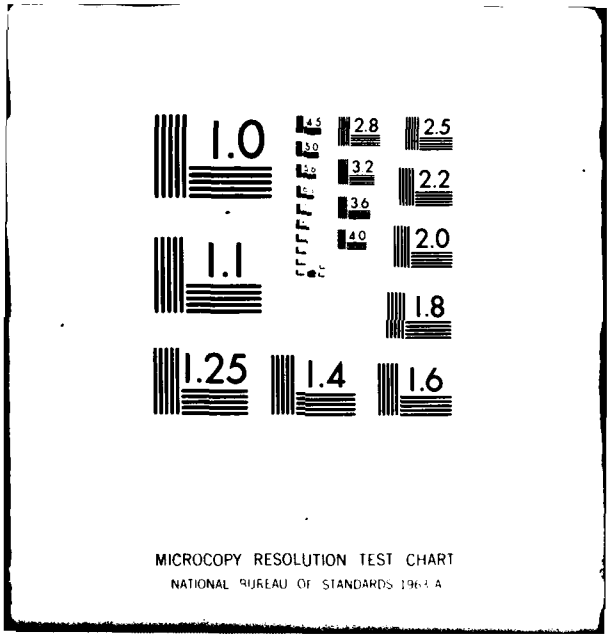
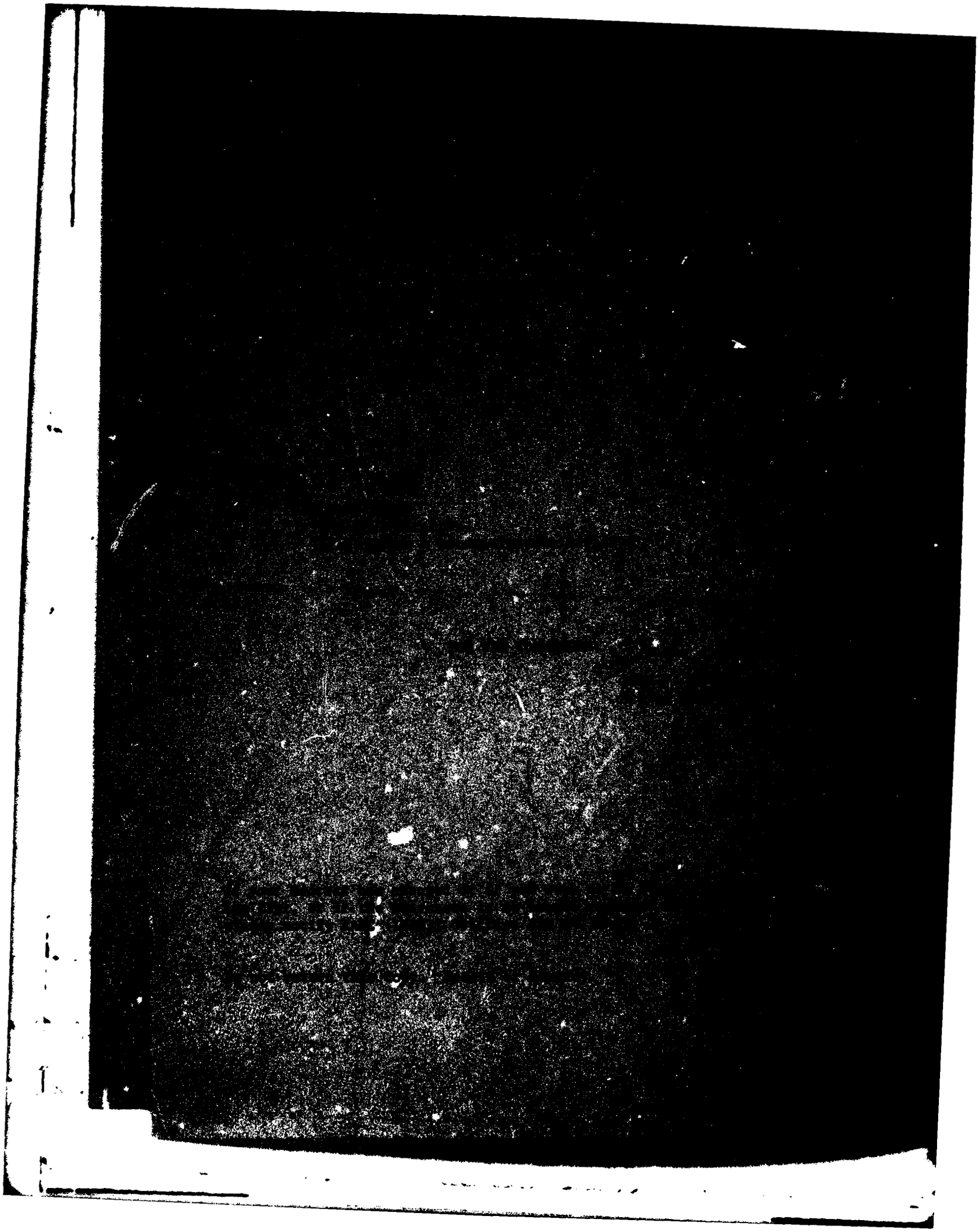
AD-A082 014 RCA ADVANCED TECHNOLOGY LABS CAMDEN N J F/G 14/5
ADVANCED IMAGE COMPRESSION STUDY.(U)
JAN 80 W B SCHAMING, J J RUDNICK F30602-77-C-0244
UNCLASSIFIED RADC -TR-79-342 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1964 A



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(19) REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
(18) 1. REPORT NUMBER: RADC-TR-79-342	2. GOVT ACCESSION NO. / 3. RECIPIENT'S CATALOG NUMBER
(6) 4. TITLE (and Subtitle): ADVANCED IMAGE COMPRESSION STUDY.	(9) 5. TYPE OF REPORT & PERIOD COVERED: Final Technical Report, 30 Sep 77 - 21 May 79
	6. PERFORMING ORG. REPORT NUMBER: N/A
(10) 7. AUTHOR(s): W.B. Schaming, J.J. Rudnick	(15) 8. CONTRACT OR GRANT NUMBER(s): F30602-77-C-0244
9. PERFORMING ORGANIZATION NAME AND ADDRESS: RCA, Advanced Technology Laboratories, Camden NJ 08102	(16) 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 62702F, 62441081 (17) 10
(11) 11. CONTROLLING OFFICE NAME AND ADDRESS: Rome Air Development Center (IRRE), Griffiss AFB NY 13441	12. REPORT DATE: Jan [redacted] 80
	13. NUMBER OF PAGES: 115
(12) 117 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Same	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

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

Same

18. SUPPLEMENTARY NOTES

RADC Project Engineer: Douglas J. Praska, Lt, USAF (IRRE)
Robert LaSalle (IRRP)

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

Bandwidth Compression
Image Features
Image Processing

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

The work described in this report primarily covers two areas. The first effort described is an attempt to group images for bandwidth compression using clustering techniques on statistical measurements. The second effort is an attempt to predict the performance of the cosine transform algorithms from statistics derived from the imagery.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

405561

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

LIST OF TABLES

REPORT SUMMARY

SCOPE OF WORK

The Advanced Image Compression Study was originally intended to investigate the possibility of improving the performance of a bandwidth-compression process by dissecting the original image into smaller pieces having more uniform content and varying the algorithm accordingly.

Considerable effort was expended in trying to determine which of the selected image subsets were statistically similar. This effort requires a data base much bigger than the scope of the program allowed. The thrust of the program was changed to relate the features measured on the image subsets to the actual performance of the two-dimensional cosine transform operating in a non-adaptive mode at the 1.0 bit per pixel level.

IMAGE STATISTICS

After selecting the original imagery, approximately 20 subsets of varying content were digitized. A large group of statistics was computed for each of these subsets. Included in these statistics are measurements computed directly from the brightness values over the image, some computed from the brightness values over a local neighborhood, and those computed from the gradient image. In addition, the Karhunen-Loeve transform was performed on some of the statistics in an attempt to produce a smaller set of features that are optimally decorrelated.

CLUSTERING EXPERIMENTS

Parameters computed from the image statistics, as well as parameters related to the optimum number of bits required to transmit the image, were used as inputs to a clustering algorithm.

Three different clustering algorithms were considered. The first algorithm required that the number of clusters as well as the cluster means be specified. Each sample is then assigned to the nearest cluster. The second approach required that a tolerance be specified for the distance between clusters. Any sample that does not fall within this

tolerance from a previously defined cluster will become the initial point in a new cluster. The third approach assumed that each sample is a separate cluster. The two clusters with the minimum distance between them are combined to reduce the number of clusters by one. This continues until all samples are in a single cluster. A parameter is computed at each stage in the process that is intended to help select the optimum number of clusters for the data.

REGRESSION ANALYSIS

In order to relate the image statistics to bandwidth compression, a regression analysis was performed to try to predict the resultant mean square error of the compressed images. If a set of possible compression algorithms were available in an operational system, the compression would be maximized by applying each algorithm to an image and then selecting the one that provides the most compression at a given performance level. This results in an inordinate amount of computation. However, if each of these algorithms could be predicted for the image from a group of readily measured statistics, the selection of algorithms is much easier.

As an intermediate step toward this goal, an attempt was made to predict the performance of a single compression algorithm, namely the two-dimensional discrete cosine transform. A multiple linear regression analysis was performed, using the jack-knife method, in which all of the images but one are used for training. The resultant image is then used to predict the performance of the algorithm for the remaining image not used in the training set.

EVALUATION

The work performed in this effort has given the Air Force important information which will advance image compression technology and contribute greatly towards accomplishing the goals of technical program objective (TPO) R2C. The effort has demonstrated that assigning various image compression algorithms to subareas within an image for optimum compression of the overall image is a difficult task, and that given the correct image statistics, the performance of a particular image compression technique can be accurately predicted. This knowledge is essential for the development of automated image compression systems.

Douglas J Praska

DOUGLAS J. PRASKA, 1LT, USAF
Project Engineer

Section 1

INTRODUCTION AND SUMMARY

1.1 SCOPE

The Advanced Image Compression Study was originally intended to investigate the possibility of improving the performance of a bandwidth compression process by dissecting the original image into smaller pieces, having more uniform content, and varying the algorithm accordingly. The contractual statement of work was written around this concept. However, after lengthy discussions with RADC, emphasis was shifted. Initially, the thrust was to determine statistics that could be used to partition the image to optimize subsequent compression. Considerable effort was expended to determine which of the selected image subsets were statistically similar. For reasons described later, this effort requires a data base much greater than the scope of the program allowed. The thrust of the program was changed to relate the features measured on the image subsets to the actual performance of the two-dimensional cosine transform operating in a non-adaptive mode at the 1.0-bit-per-pixel level. The cosine transform was selected because it is generally accepted to be nearly optimum. As a result, the work reported herein deviates from the original statement of work.

1.2 IMAGE STATISTICS

The general direction of the efforts on this program can be seen in the diagram of Fig. 1-1 and are briefly outlined below. After selecting the original imagery, approximately 20 subsets of varying content were digitized. A large group of statistics was computed for each of these subsets. Included in these statistics are measurements computed directly from the brightness values over the image, some computed from the brightness values over a local neighborhood, and those computed from the gradient image. The brightness statistics include the mean, variance, dynamic range, skewness, kurtosis, and power spectrum. The local neighborhood statistics are the mean value, the dynamic range, and the maximum to minimum brightness ratio over the neighborhood. The mean and variance of these parameters were computed over all neighborhoods in the image. Neighborhood sizes of 2, 4, 10, 25, and 50 were utilized. The mean and variance of the gradient image were computed to provide edge density information.

As indicated in Fig. 1-1, these statistics were used in a number of ways. An attempt was made to extract information using a two-input scatter diagram. The only useful data derived from these scatter diagrams is shown in Fig. 1-2, which is a plot of the maximum brightness minus the average versus the average minus the minimum brightness. This diagram is, in effect, a measure of the non-symmetry in the brightness histogram. An examination of this data shows that natural subjects such as water, fields, woods, and marshes tend to fall above the "slope = 1/3" line and tend to

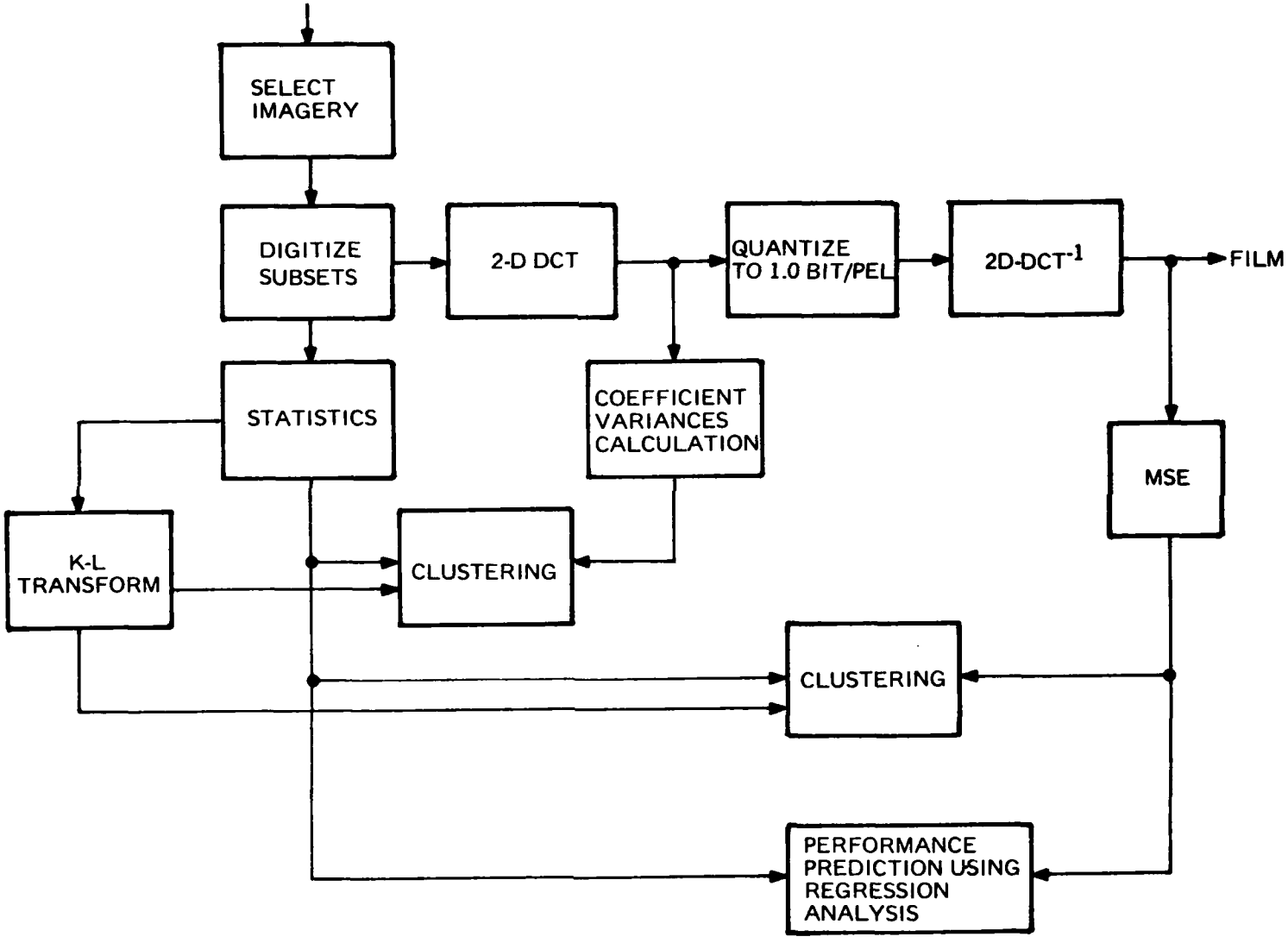


Fig. 1-1. Block Diagram of Advanced Image Compression Study.

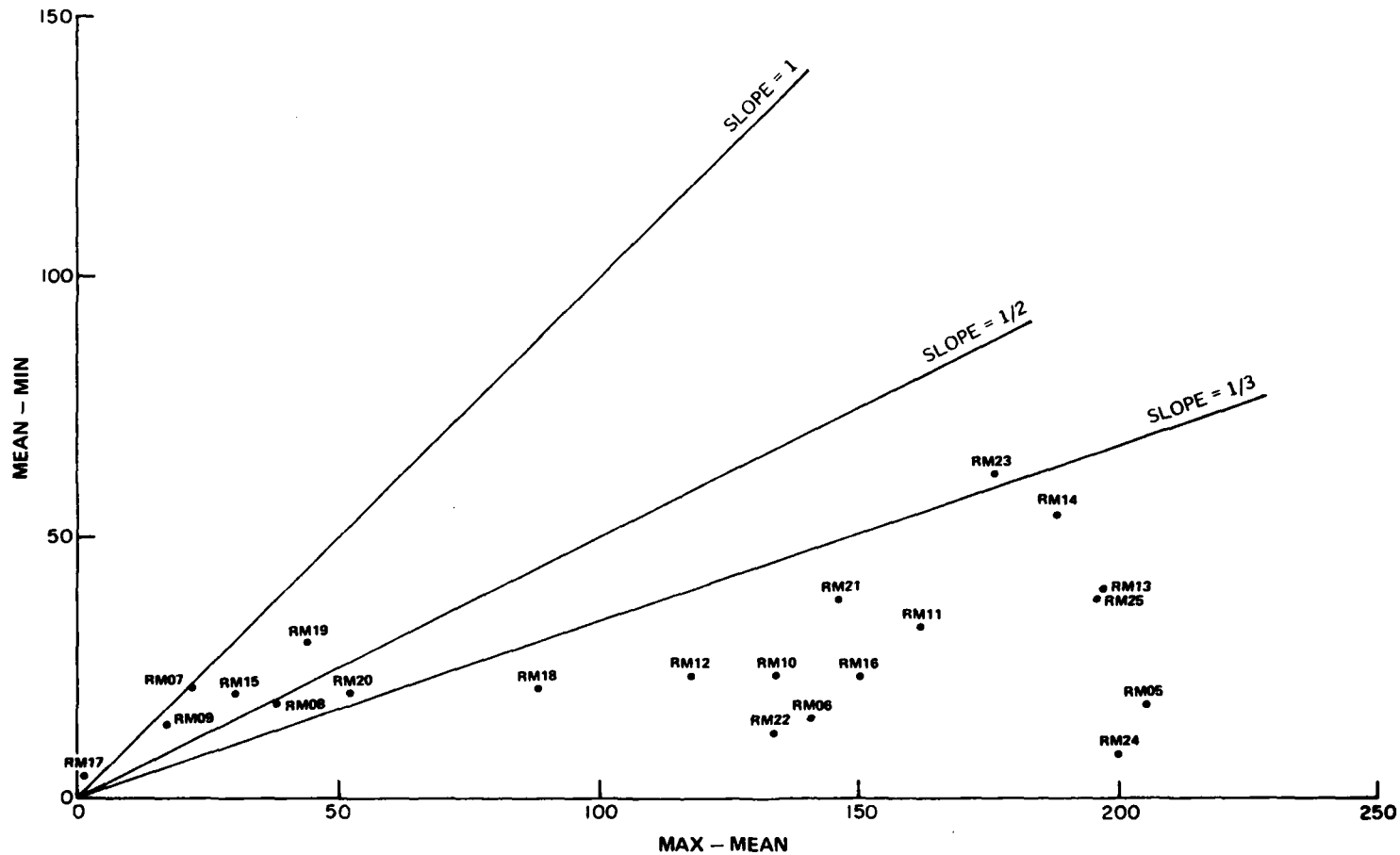


Fig. 1-2. Plot of histogram symmetry characteristics

cluster in the lower left region of the plot, which indicates that the range (MAX-MIN) is also small. On the other hand, subjects with man-made objects such as ships, planes, houses, etc. tend to fall below the "slope = 1/3" line and are generally higher in dynamic range. This same grouping is later verified, using a clustering algorithm on the two plotted parameters.

It became obvious that a more comprehensive method of evaluating the statistical measurements was needed. This led to the evolution of clustering techniques and regression analysis. The clustering techniques were used to group the images and the regression analysis was used to predict the performance of the 2D-cosine transform compression algorithm.

Before applying the clustering technique to the statistics, an attempt was made to reduce the number of statistics required. Since many parameters are being measured for each subject, the set of measurements are surely not independent. One approach suggests that transforming the set of parameters into another pseudo set of parameters that are uncorrelated should reduce the number of inputs required to group the images. The Karhunen-Loeve transform is such a process. Not only does it produce an uncorrelated set of coefficients, but they are in descending order of statistical importance. Consequently, the clustering technique was applied to the truncated output of the K-L transform of the statistics rather than the statistics themselves. In parallel with this work, the two-dimensional DCT was performed on each of the original subsets using a block size of 16 by 16. The variance of each coefficient taken from all the transform blocks in the subset was also computed. From the coefficient variances, a parameter was computed that is related to the optimum number of bits required to transmit the subset using the 2D-DCT. This parameter, in conjunction with data obtained from the statistical analysis of the subsets, was used as input to a two-dimensional clustering algorithm.

The cosine transformed subsets were also quantized to one bit per pixel and inverse-transformed to obtain an approximation to the original subset. The mean-square error was computed for each reconstructed subset with the intention of using it with the statistical data in the clustering algorithm.

1.3 CLUSTERING ALGORITHMS

Three different clustering algorithms were considered. The first algorithm required that the number of clusters as well as the cluster means be specified; each sample is then assigned to the nearest cluster. The second approach required that a tolerance be specified for the distance between clusters, and any sample that does not fall within this tolerance from a previously defined cluster will become the initial point in a new cluster. The third approach assumed that each sample is a separate cluster. The two clusters

---

[1]R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, New York; John Wiley & Sons, 1973, pp. 234,235.

[2]R. O. Duda and P. E. Hart, op. cit.

[3]G. B. Coleman, 'Image Segmentation by Clustering', Univ. of Southern California, Image Processing Institute, Report USCIPI 750, p. 31, July 1977.

with the minimum distance between them are then combined to reduce the number of clusters by one. This continues until all samples are in a single cluster. A parameter is computed at each stage in the process, which is intended to help select the optimum number of clusters for the data.

Each of these clustering algorithms has a limitation. The first algorithm requires prior determination of the number of clusters and the general location of each cluster. The second algorithm requires that a distance be selected. Depending upon the distance chosen, different clusters may result. For example, the sum of squares distance

$$\left[X_1, X_2 \in \mathbb{R}^n, \ \left\| X_1 - X_2 \right\| = \left(\sum_{i=1}^{n} (x_{1i} - x_{2i})^2 \right)^{1/2} \right]$$

used in the experiments performed, is best for dense, clearly separated clusters of odd shapes. A nearest-neighbor distance criterion tends to form long chain-like clusters.[1] A furthest-neighbor distance criterion has a tendency to form compact clusters that are roughly equal in size.[2] The third method seemed to have the most promise. The limitation for this third method was that the optimum number of clusters was often ambiguous. To obtain a unique optimum, it is necessary for a human observer to use judgement as to what comprises good clustering.[3]

The optimum may be ambiguous due to the limited amount of data used in the experiments.

All of the clustering methods considered seem to impose a particular structure on the data, rather than to find structure in the data. Not knowing what this structure should be, i.e. how the images should be grouped for bandwidth compression, was one reason behind the decision to discontinue efforts involving clustering. Another important reason for this decision was that clustering techniques require much more data than was available. At least 100 items should be clustered in order to have reasonable confidence in results obtained from clustering.

1.4 REGRESSION ANALYSIS

Because of these difficulties, a different approach was considered. If a set of possible compression algorithms were available in an operational system, the compression would be maximized by applying each algorithm to an image and then selecting the one that provides the most compression at a given performance level. This results in an inordinate amount of computation. However, if the performance of each of these algorithms could be predicted for the image from a group of readily measured statistics, the selection of algorithms is much easier.

As an intermediate step toward this goal, an attempt was made to predict the performance of a single compression algorithm, namely the two-dimensional discrete cosine transform. A multiple linear regression analysis was performed using the jackknife method, in which all of the images but one are used for training. The last image

is then used to test the process. This is repeated using each one of the images as the one left out for testing. The independent variables in the process are the measured statistics and the dependent variable is the compression performance (i.e. mean square error - MSE). The regression process defines a surface in n-dimensional space that fits the available data points as well as possible. The resultant equation is then used to predict the performance of the algorithm for the remaining image not used in the training set. Only limited success was achieved with this performance prediction. Very few of the coefficients in the resultant regression equation are non-zero. This implies that most of the statistics used in the process are insignificant for predicting the MSE. This does not mean that the technique is unusable, but only that the proper set of statistics may not have been found.

When the dependent variable in this process is changed to the number of frequency coefficients required to limit the MSE due to truncation of the spectrum to 0.25%, the resultant regression equation proves to be a good predictor. This parameter is directly related to the power spectrum, which is also represented by some of the independent variables input to the regression process. The encouraging results obtained when trying to predict this dependent variable indicate that the process is usable when the appropriate statistics are used.

The very limited number of images used in this evaluation is insufficient to provide strong conclusions or high confidence levels in the apparent conclusions.

As an outgrowth of this work, the possibility of using simple statistical features to direct an adaptive 2D-cosine algorithm is attractive. This approach is briefly outlined in Section 9.

Section 2

SELECTION OF IMAGERY

The first task in this study effort was to obtain imagery from Rome Air Development Center (RADC) that contained a wide variety of information. Initially some 49 images were selected from the RADC data base. Positive transparencies were prepared and supplied to RCA. The images provided a wide range of targets and background. When the positive transparencies were viewed, it was apparent that many of the shadows and highlights were saturated. Since the intended use of the pictures required that many statistics be computed from the digitized version of these images, it was decided to select a different data set.

A roll of film was obtained, from RADC, which contained a reasonable range of scene content and which does not appear to be saturated. The negatives were also supplied to RCA for this roll. Nine frames were selected from this imagery and are shown in Figs. 2-1 through 2-9. From these nine frames, 21 subsets were chosen to be digitized. These subsets were chosen to include various terrains and different amounts of detail. These subsets are outlined and identified on the images in Figs. 2-1 through 2-9.

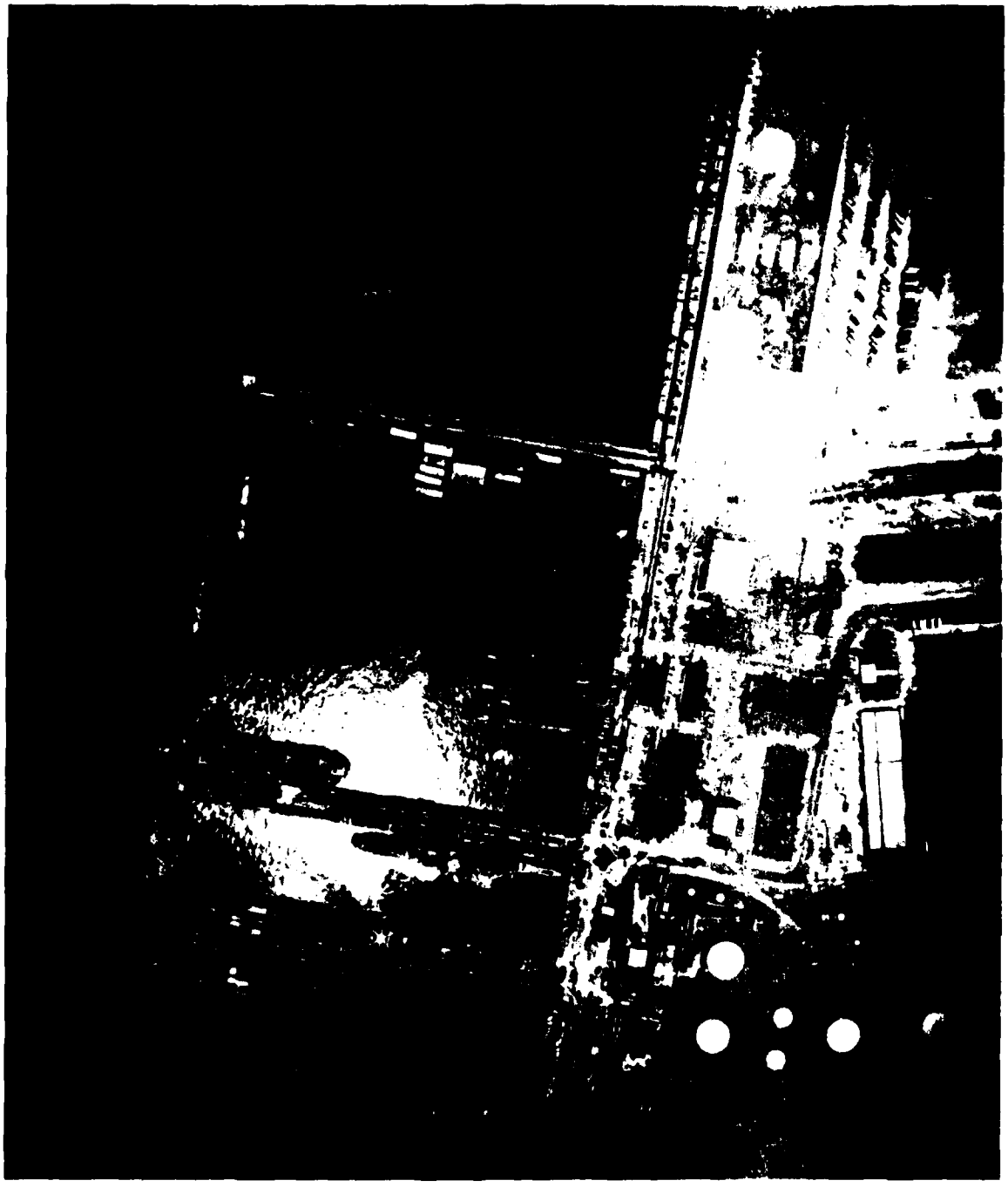


Fig. 2-1. Frame F143.

Fig. 2-2. Frame 4083.

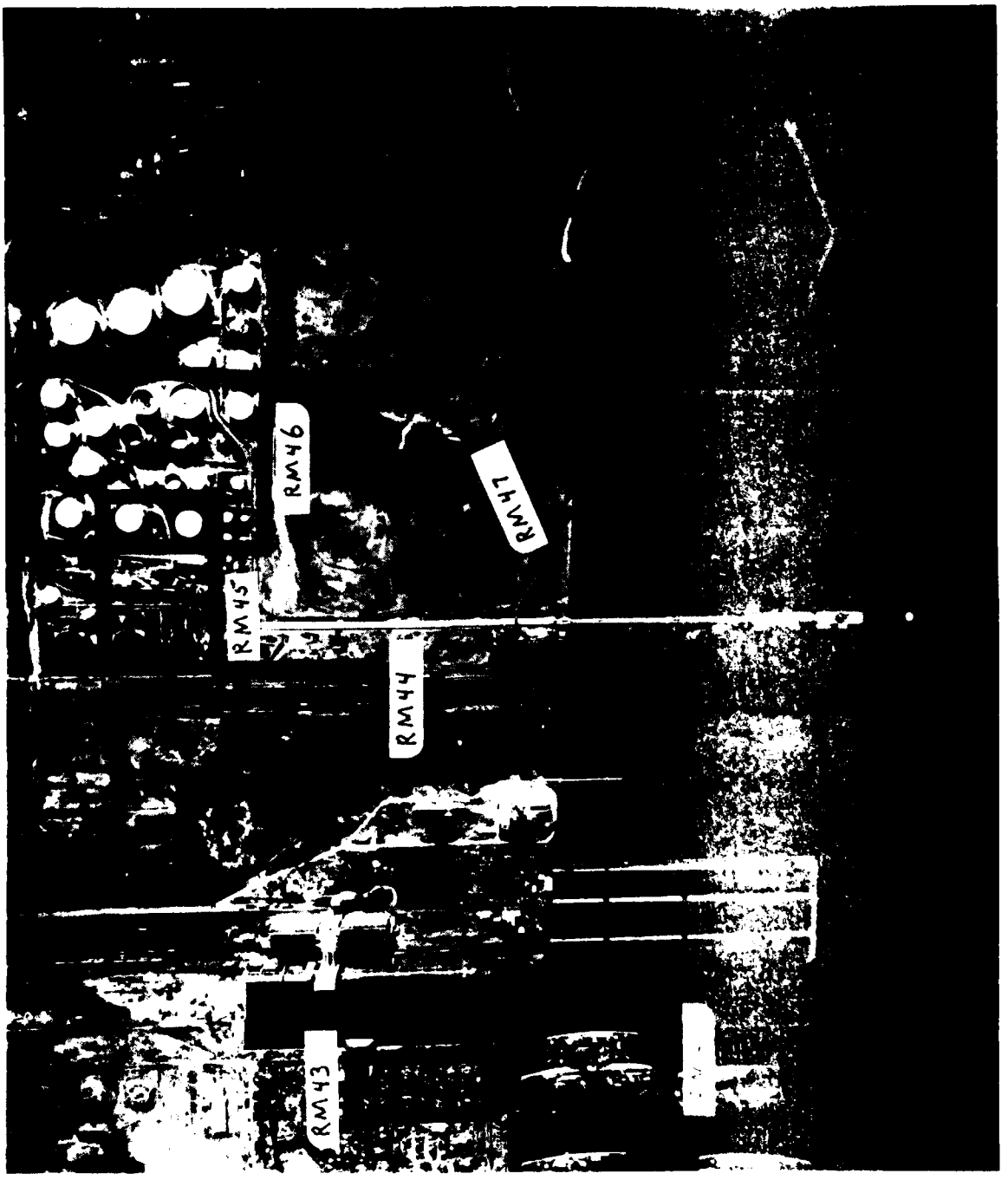


Fig. 2-3. Frame 4109.

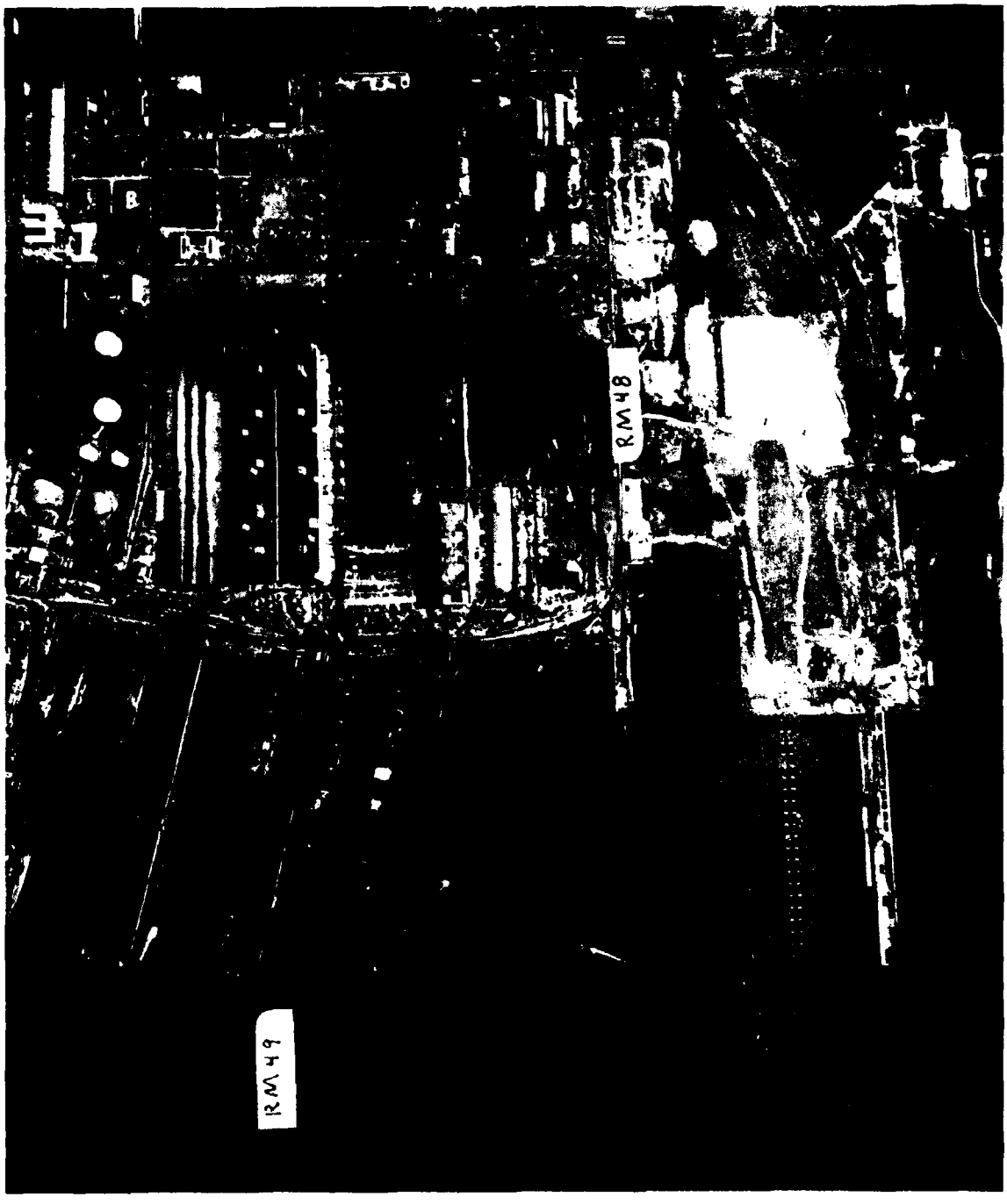


Fig. 2-4. Frame 4120.

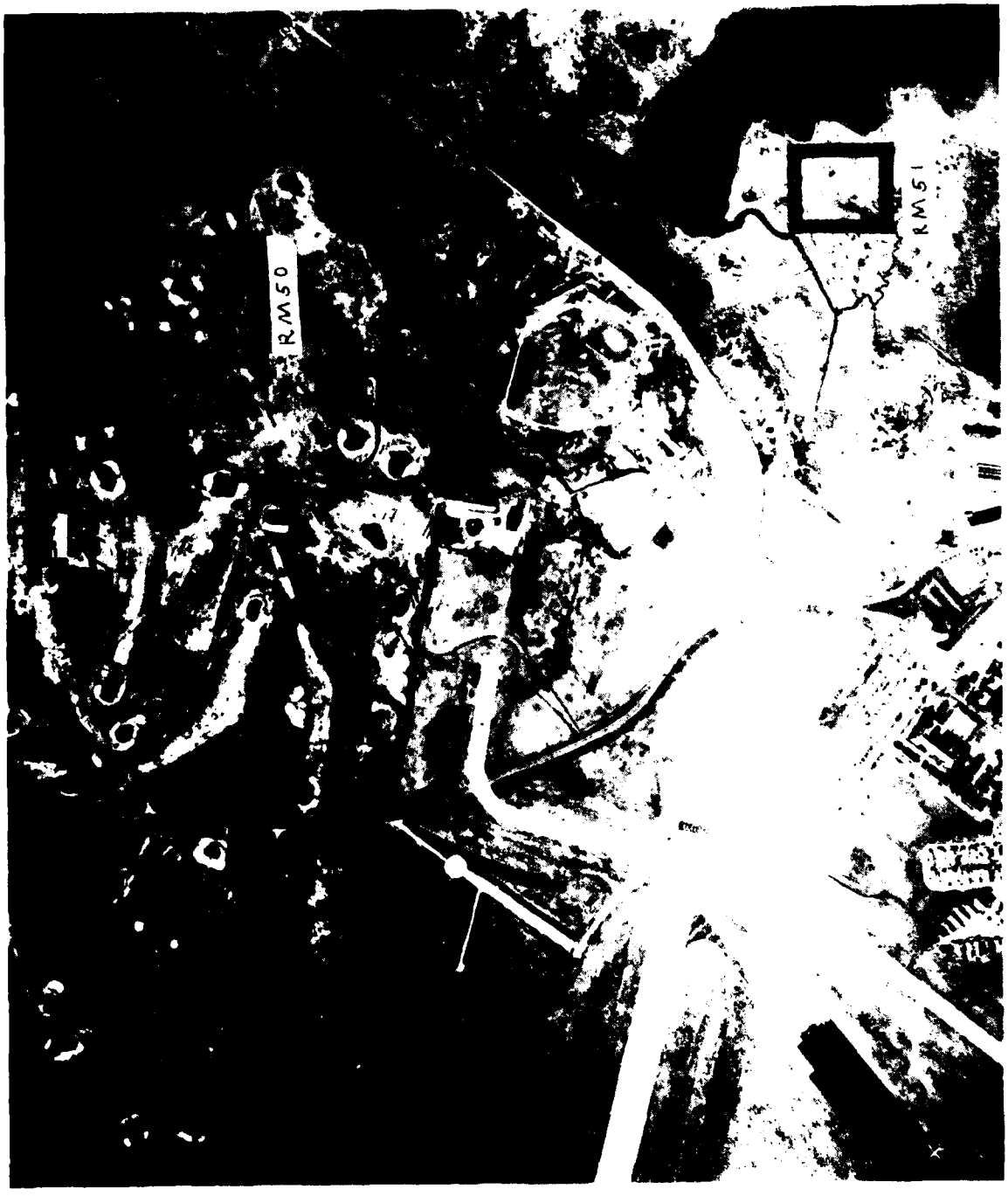


Fig. 2-5. Frame 4128.

Fig. 2-6. Frame 4131.

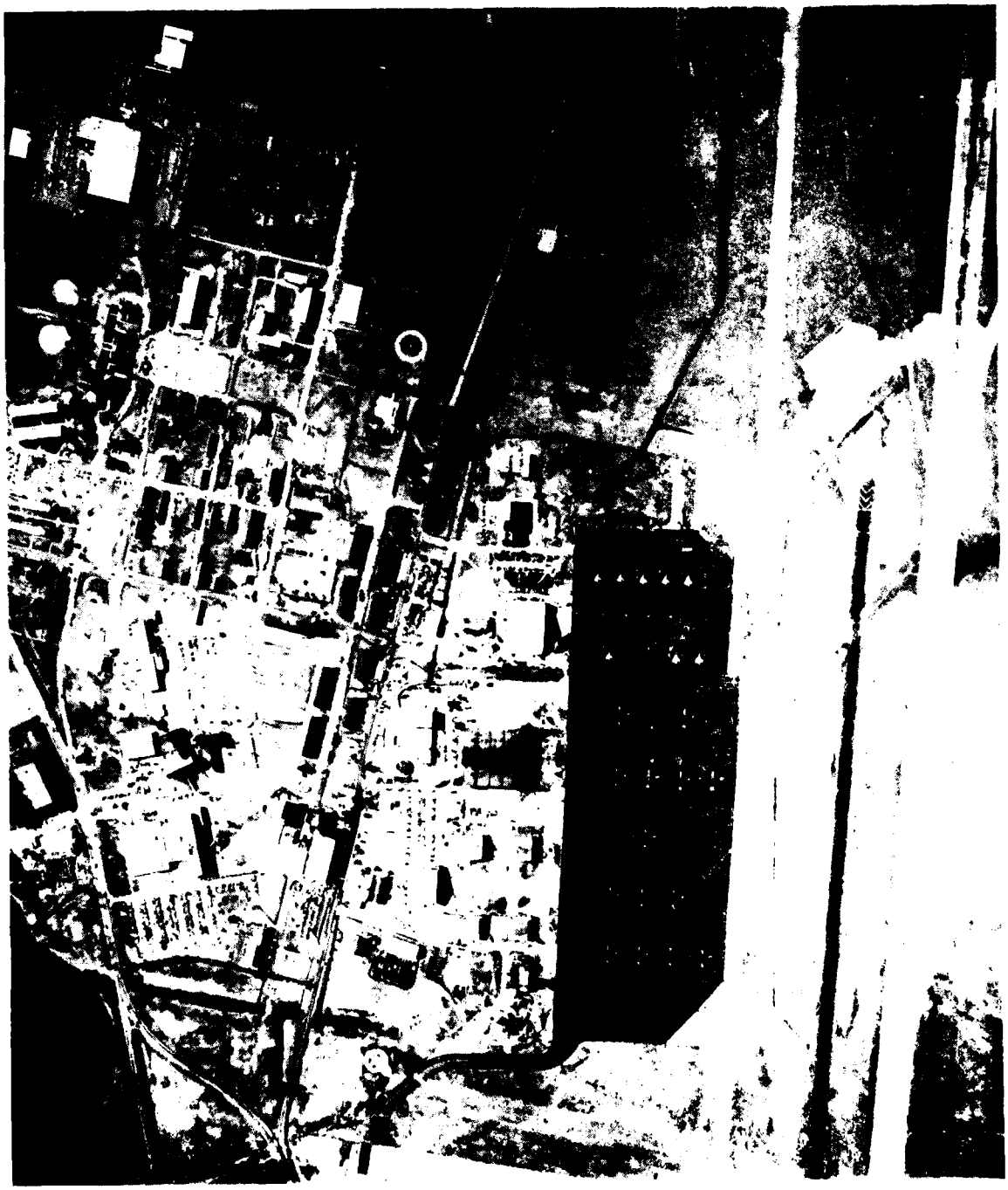


Fig. 2-7. Frame 4134.

Fig. 2-8. Frame 4137.

Fig. 2-9. F141.

Section 3

DIGITIZING THE SUBSETS

In order to obtain a set of digitized images with uniform quality, three topics related to the digitizing process must be considered. These include the scanning spot size, the film density vs. digital number transfer function of the digitizing device, and the MTF of the imagery.

3.1 SCANNING SPOT SIZE AND MTF

The first set of imagery obtained from RADC contained considerable data at a scale of 2000:1. The intention was to scan the images using a five-mil spot, which results in a corresponding spot spacing on the ground of 10 inches. The images in Figs. 2-1 through 2-9, however, have approximately a 6000:1 scale. Scanning with a five-mil spot would represent a sample spacing of 30 inches on the ground. This spot size would result in an extremely poor rendition of the original imagery. In order to determine the correct scanning aperture to use in digitizing the images, the MTF of the original transparencies was computed. Edges at random orientation in the film were scanned on a Joyce-Loebl microdensitometer to provide the edge response. This edge response was used to compute the MTF as described in Appendix A.

The amplitude of the complex modulation transfer function is the familiar MTF along the direction perpendicular to the edge. Figs. 3-1 through 3-6 show some of the MTF curves computed for the randomly oriented edges. Since the plots in Figs. 3-1 and 3-6 show the MTF falling off considerably sooner than the others, an investigation was undertaken to determine if this required corrections. Three additional edges at approximately the same orientation were scanned from the same transparency containing the edge which produced Fig. 3-6. Two of these three scans produced MTF that which were greater than 0.1 out to 25 cycles/mm. It was therefore concluded that the low MTF edges were due to the actual edge being scanned rather than some correctable phenomenon such as aircraft motion.

Figure 3-7 shows the MTF of a circular scanning aperture. With a 5-mil diameter and 5-mil spacing, the sampling frequency is:

$$f_s = \frac{1}{5(0.0254)} = 8 \text{ cycles/mm.}$$

This implies that anything above 4 cycles/mm will not be reproduced in the sampled image. A more appropriate sampling frequency for these images would be above 20 cycles/mm, where the MTF falls below 25%. A circular aperture with 1-2/3-mil diameter corresponding to a sampling frequency of 24 cycles/mm was therefore used to digitize the subsets. No MTF correction was applied to the digitized images.

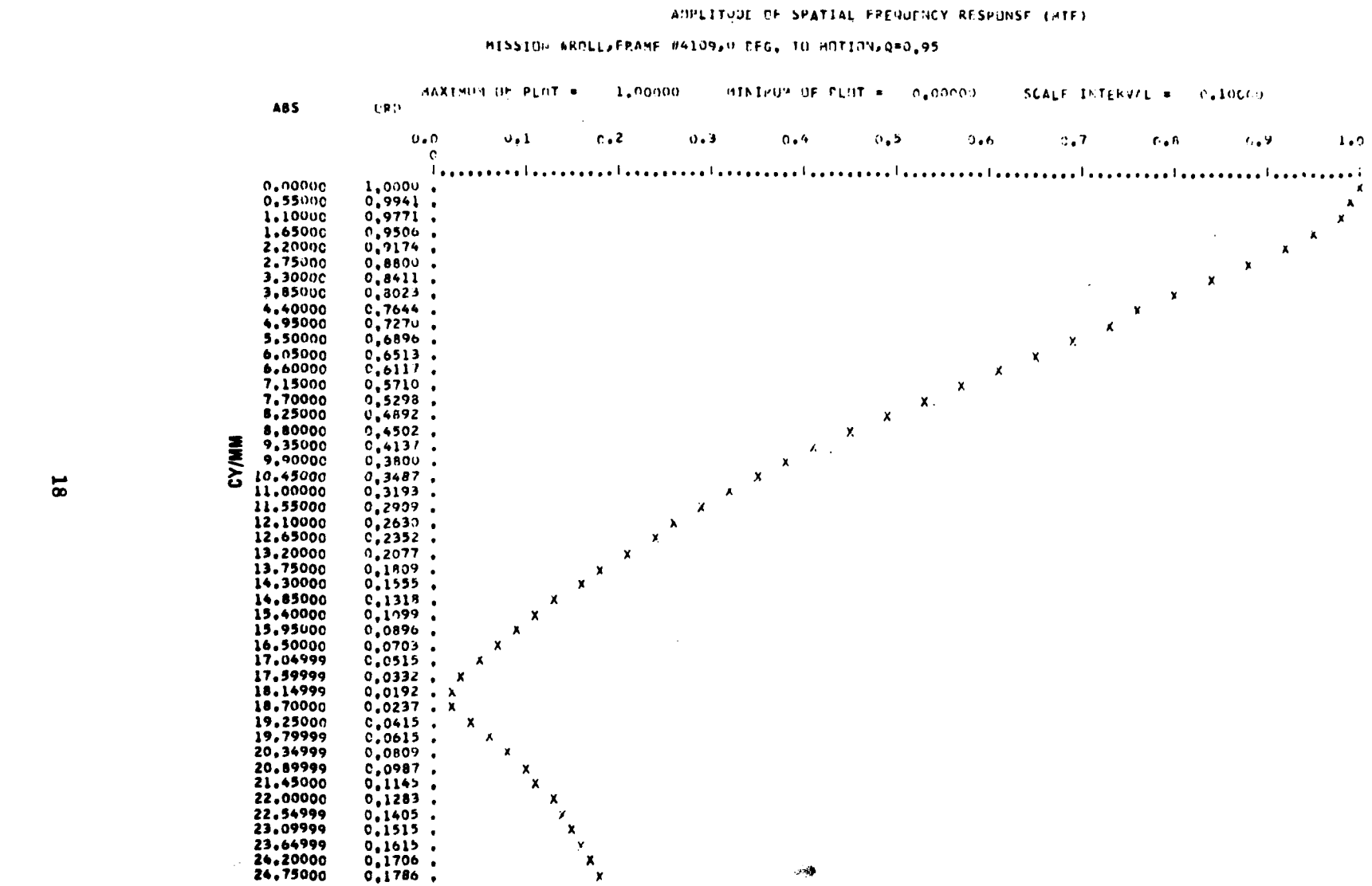


Fig. 3-1. Amplitude of spatial frequency response (MTF). Mission #Roll, Frame 4109, 0 deg. to motion, Q = 0.95.

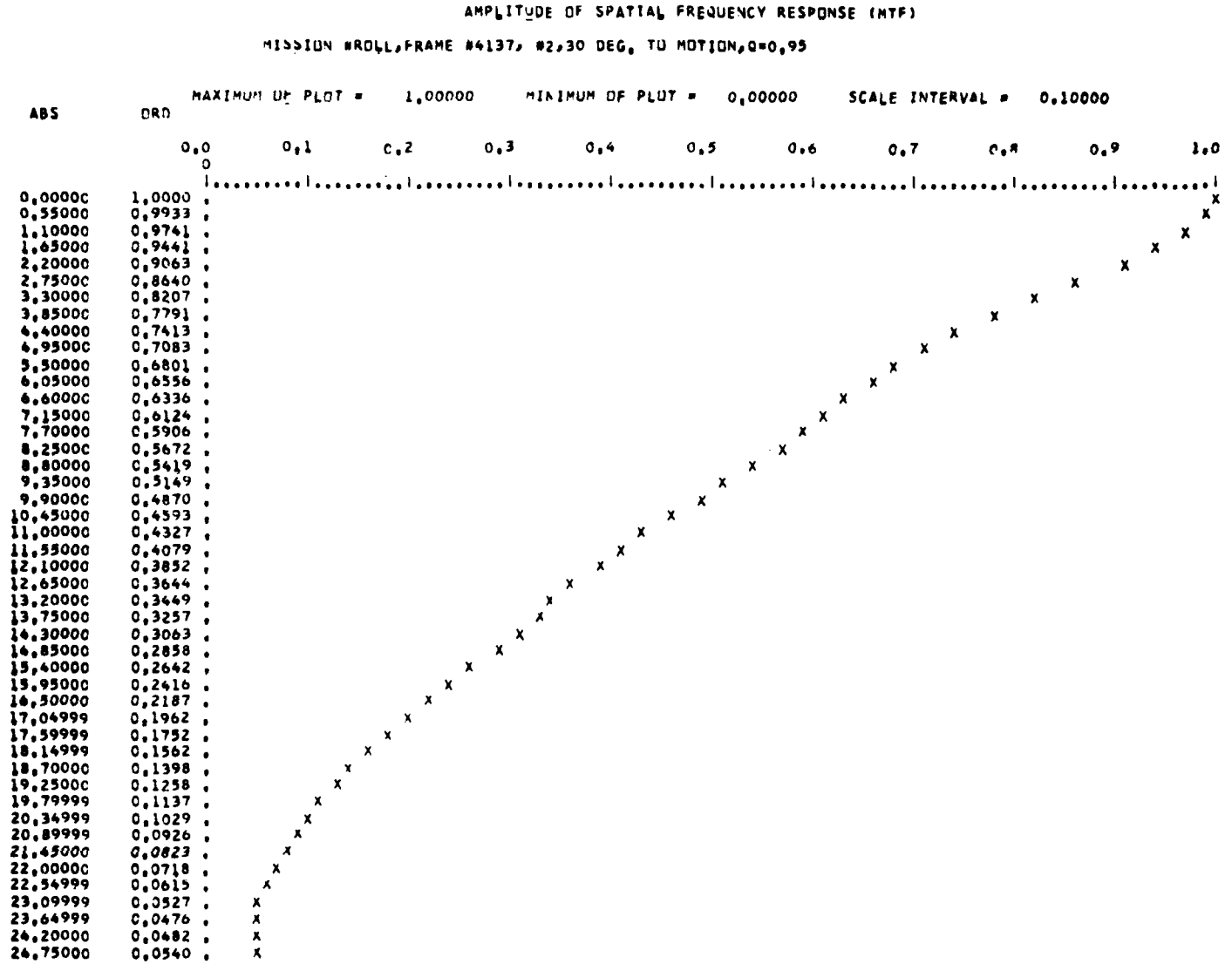


Fig. 3-2. Amplitude of spatial frequency response (MTF). Mission #Roll, Frame 4137, #2, 30 deg. to motion. Q = 0.95.

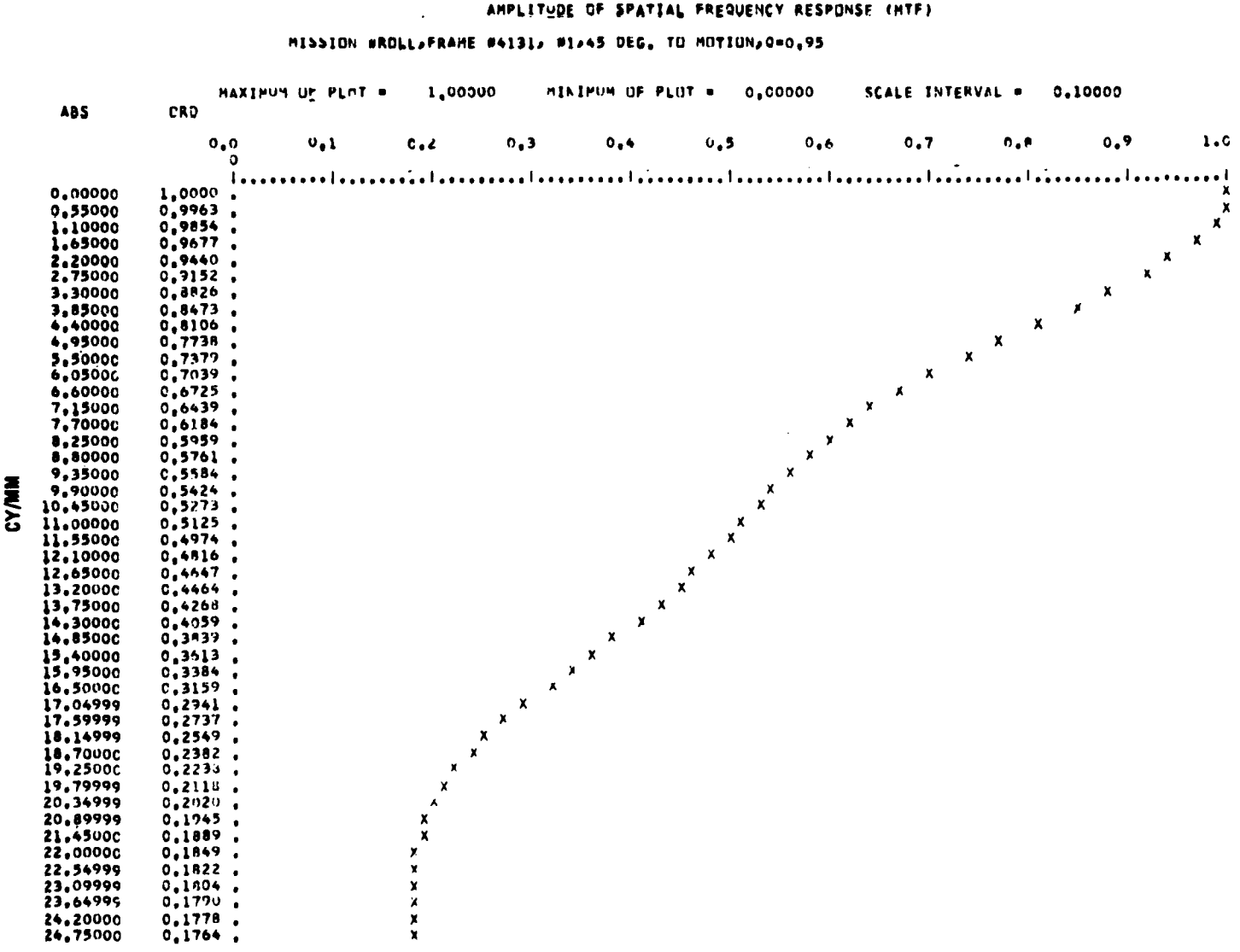


Fig. 3-3. Amplitude of spatial frequency response (MTF). Mission #Roll, Frame #4131, #1, 45 deg. to motion, Q = 0.95.

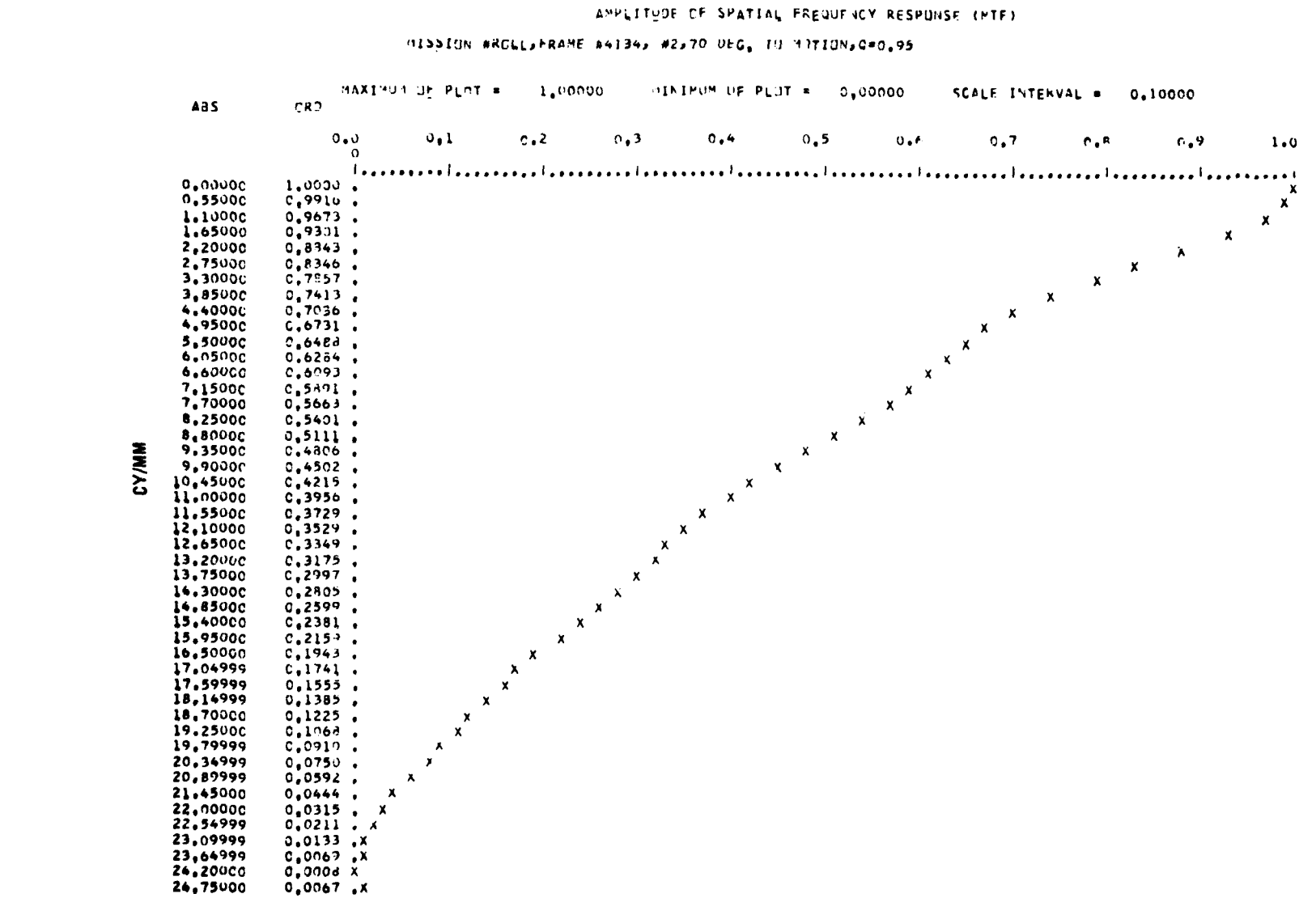


Fig. 3-4. Amplitude of spatial frequency response (MTF). Mission #Roll, Frame #4134, #2, 70 deg. to motion, Q = 0.95.

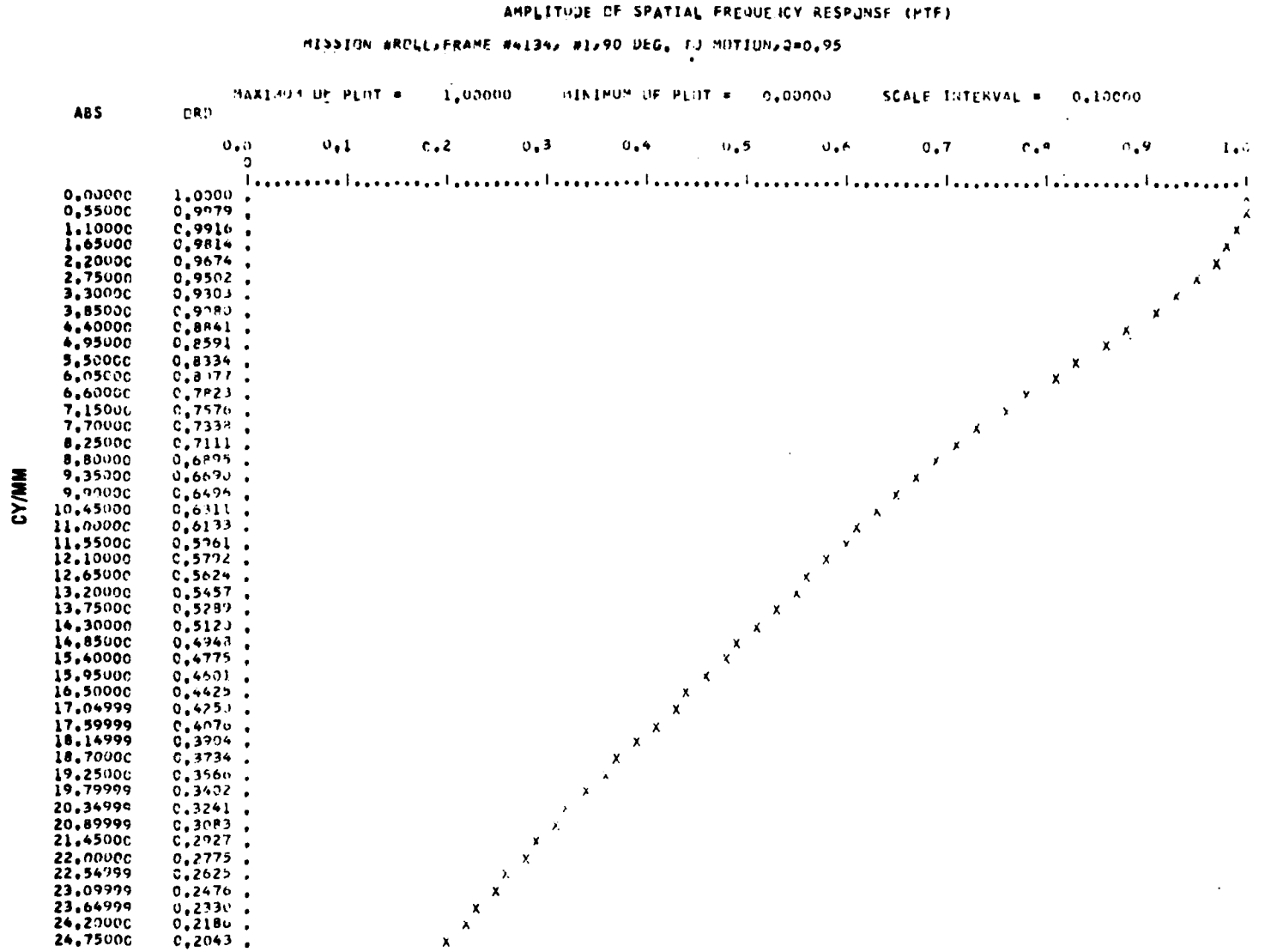


Fig. 3-5. Amplitude of spatial frequency response (MTF).
Mission #Roll, Frame #4134, #1, 90 deg. to motion, Q = 0.95.

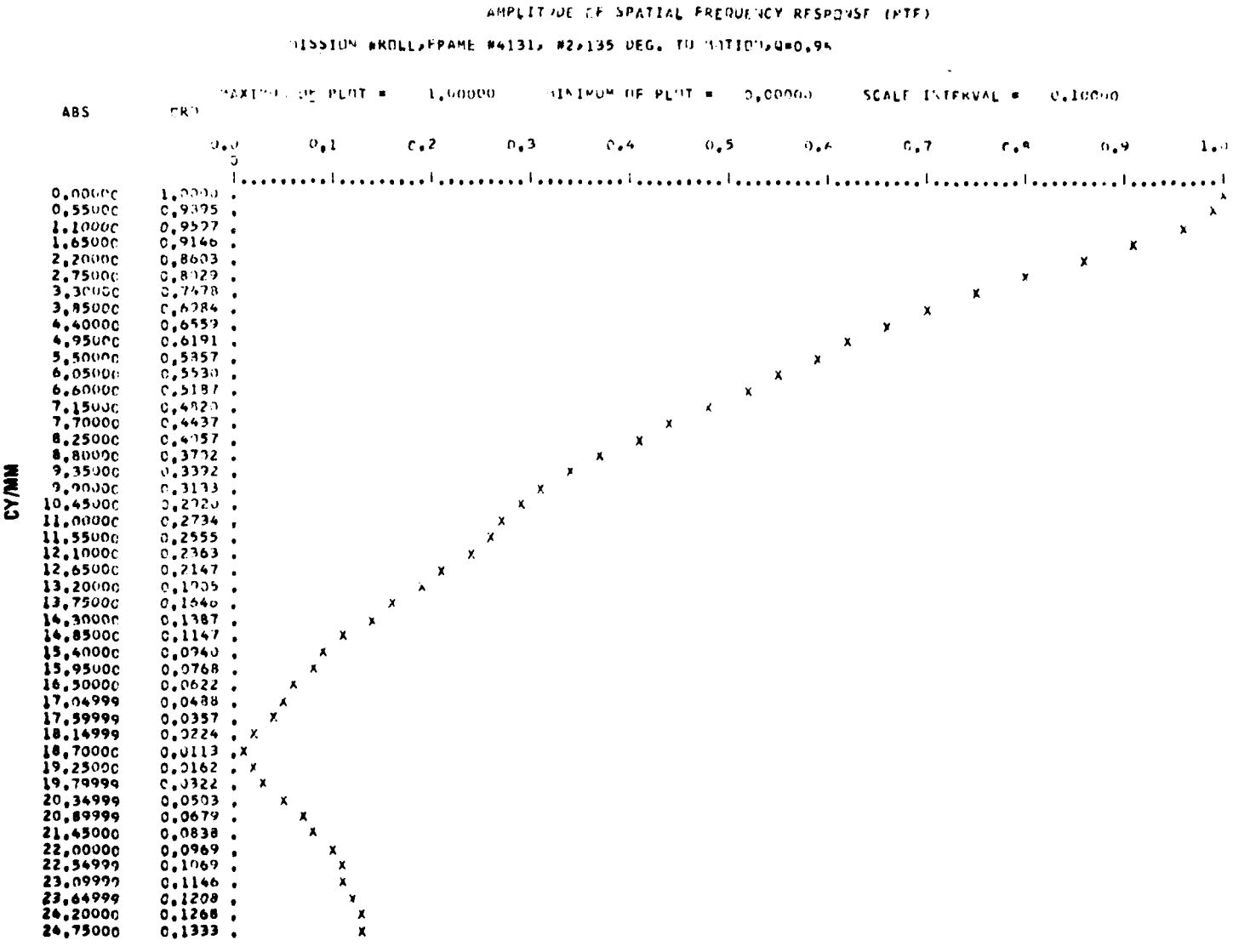


Fig. 3-6. Amplitude of spatial frequency response (MTF). Mission #Roll, Frame #4131, #2, 135 deg. to motion, Q = 0.95.

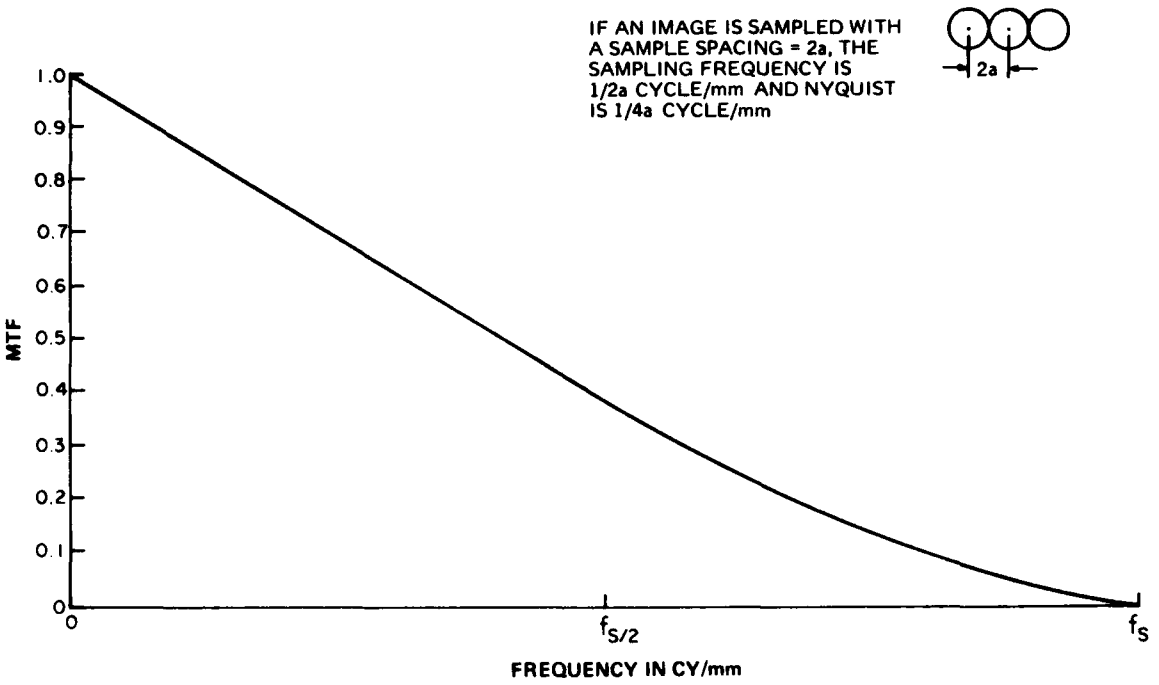


Fig. 3-7. MTF of circular aperture with radius a in millimeters.

3.2 FILM DENSITY VS. DIGITAL DATA

A correction was applied to the digital data in order to maintain a linear (slope=-1) relation between log digital counts from the A/D converter and density on the positive transparency being scanned. The correction that was applied to the data is given by:

$$\text{Corrected Data} = (\text{Scanner Output})^{1/Q} \times \frac{256}{256^{1/Q}}.$$

This form of correction is used, since the scanner output is proportional to transmittance. (See Appendix B for details.)

At the time the positive transparencies were digitized, a standard step wedge (made on the same film type) was also digitized. The resulting plot shown in Fig. 3-8 indicates a value of Q of 1.16. This value was used to correct all the digitized outputs. Failure to perform the Q correction on the digitized imagery will result in a reduced dynamic range output when rewritten on film. This assumes that the write-out process is controlled so that the plot of log digital numbers vs. film density is of unity gamma.

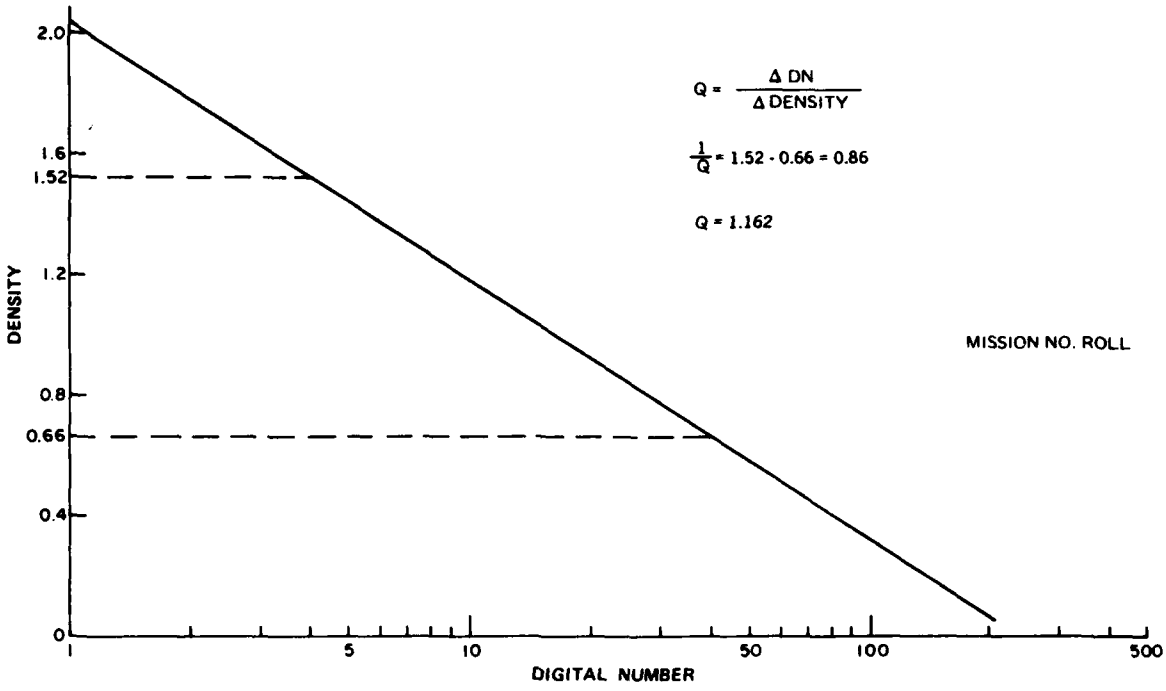


Fig. 3-8. Plot of density vs. digital number for step wedge digitized on scanner.

Section 4

STATISTICS COMPUTED FROM THE SUBSETS

The 21 subsets that were digitized are listed in Table 4-1, which gives their size as well as a brief description of their contents.

Many measurements were made on the subsets in an attempt to find parameters that will facilitate the grouping of images for compression. The measurements can be divided into three different types, namely those computed directly from the brightness values over the image, those computed from the brightness value over a local neighborhood, and those computed from the gradient image. The different statistics are described in the following.

4.1 STATISTICS COMPUTED FROM ENTIRE SUBSET

The measurements made on the brightness values of the pixels are listed in Table 4-2. These statistics were computed from the definitions given below. I (x, y) is the brightness value of the image sample at location x, y in the image, where x goes from 1 to N and y goes from 1 to M.

The mean and variance of the brightness values in the image provide a measure of the average gray level and the spread about that level.

$$\text{Mean} = \frac{1}{NM} \sum_{x=1}^{N} \sum_{y=1}^{M} I(x, y)$$

$$\text{Variance} = \frac{1}{NM} \sum_{x=1}^{N} \sum_{y=1}^{M} I(x, y)^2 - \text{Mean}^2$$

The maximum and minimum provide image dynamic range data.

Maximum = Largest brightness value in image
Minimum = Smallest brightness value in image

TABLE 4-1. LIST OF SUBSETS WITH SIZE AND DESCRIPTION OF CONTENTS

Subset	Samples per Line	Number of Lines	Visual Description	Edge Types, Straight or Curved	Edge* Density	Object* Size	Dynamic Range ≤75 L 75 - 175 M ≥175 H
RM39	382	392	Fields	S	VL		M
RM40	290	600	Woods				L
RM41	320	420	Fields	S	VL		L
RM42	350	675	Ships	S&C	H	S, M, L	H
RM43	260	600	Parking Lot	S	VH	S	H
RM44+	230	600	Railroad	S	M	M	M
RM45	740	450	Storage Tanks	S&C	M	M, L	H
RM46	805	1020	Storage Tanks	C	M	L	H
RM47	500	420	Marshland				L
RM48	530	330	Railroad	S	M	S, M	M
RM49	410	600	Water				L
RM50	530	540	Golf Course	C	L	M	M
RM51	320	330	Marshland				M
RM52	410	540	Water & Marshland				M
RM53	650	420	Residential	S&C	M	M, L	H
RM54	590	600	Aircraft	S	M	S, M	M
RM55	440	600	Industrial	S&C	M	S, M, L	H
RM56	230	420	Aircraft	S	L	L	H
RM57+	230	530	Industrial	S&C	M	S, M	H

* - Subjective Measurements
+ - Out of Focus

TABLE 4-2. MEASUREMENT MADE ON PIXEL BRIGHTNESS VALUES

Subset	Mean	Var.	Max.	Min.	Skewness	Kurtosis	ρ_x	ρ_y	ρ_{ave}
RM39	39.5	55.2	113	9	0.004	0.048	0.929	0.921	0.925
RM40	27.5	69.1	71	4	0.575	0.445	0.950	0.950	0.950
RM41	29.7	13.8	51	10	0.020	0.787	0.947	0.961	0.954
RM42	26.9	552.8	188	2	3.436	3.301	0.954	0.963	0.959
RM43	37.1	779.2	202	0	5.166	6.188	0.872	0.916	0.894
RM44	26.6	199.7	147	0	3.143	6.901	0.967	0.990	0.979
RM45	40.5	979.8	203	2	2.672	2.538	0.960	0.976	0.960
RM46	50.3	1907.6	206	2	2.137	1.646	0.985	0.982	0.984
RM47	25.0	28.9	84	9	0.534	0.867	0.943	0.925	0.934
RM48	26.6	213.8	171	6	5.647	11.621	0.958	0.873	0.915
RM49	9.1	0.12	19	4	58.473	232.4	0.395	0.292	0.343
RM50	32.2	88.3	103	0	8.348	15.147	0.984	0.985	0.984
RM51	48.5	136.9	155	10	0.087	-0.161	0.924	0.929	0.926
RM52	23.5	108.2	116	0	0.525	0.802	0.959	0.938	0.949
RM53	48.4	847.5	220	2	2.713	2.903	0.975	0.965	0.970
RM54	21.5	131.8	154	2	32.074	3.966	0.924	0.948	0.936
RM55	61.2	1839.3	211	0	0.788	-0.290	0.990	0.989	0.989
RM56	16.9	253.9	187	4	44.689	51.10	0.961	0.991	0.976
RM57	41.0	1142.7	195	7	3.585	3.957	0.991	0.992	0.991

The skewness is a measure of the non-symmetry in the brightness histogram about the mean.

$$\text{Skewness} = \frac{\mu_3^2}{\mu_2^3}$$

where: μ_n is the nth central moment defined as:

$$\mu_n = \frac{1}{NM} \sum_{x=1}^{N} \sum_{y=1}^{M} (I(x,y) - \text{MEAN})^n$$

The kurtosis is a measure of the shape of the brightness histogram as compared to a normal distribution. For a normal distribution, the kurtosis is zero. When the kurtosis is less than zero, the distribution is more flat than normal and when it is greater than zero, the distribution is more sharply peaked.

$$\text{Kurtosis} = \frac{\mu_4}{\mu_2^2} - 3$$

The correlation coefficients are a measure of redundancy in the data samples at a fixed pixel separation in the orthogonal directions. For this work, a sample separation of one pixel has been used.

Correlation Coefficients:

$$\rho_x = \left\{ \frac{1}{NM} \sum_{x=1}^{N} \sum_{y=1}^{M} (I_{x,y} - \bar{I})(I_{x+1,y} - \bar{I}) \right\} / \sigma^2$$

$$\rho_y = \left\{ \frac{1}{NM} \sum_{x=1}^{N} \sum_{y=1}^{M} (I_{x,y} - \bar{I})(I_{x,y+1} - \bar{I}) \right\} / \sigma^2$$

where $\bar{I}$ = Mean of image samples

σ^2 = Variance of image samples

$\rho_{ave} = 1/2 (\rho_x + \rho_y)$

Figures 4-1 through 4-3 show the power spectral density (PSD) for each of the 21 subsets. The curves represent the average of the PSD computed for the x and y directions, which was computed along the x and y axes for each subset. A least-square smoothing operation was performed along each axis and the results were then averaged and normalized to obtain an average PSD curve for the subset. One way to utilize the PSD data in a subset discrimination process is to split the curve into three, equal-size frequency bands. The dynamic range of the PSD curve in each of the bands describes the shape of the curve by defining the shape of a linear segment over each band. Table 4-3 shows the dynamic range data for the 21 subsets. Note that a higher dynamic range in a frequency band indicates a steeper slope in the PSD curve. Another way to use this data is simply to take the value of the PSD at specific frequencies.

4.2 STATISTICS COMPUTED FROM LOCAL NEIGHBORHOODS

For each of the available subsets, the average value, max.-min., and max./min. were computed for each n x n area in the subset. The mean and variance of these parameters over the entire subset was then computed. These calculations were repeated for all subsets using values of n equal to 2, 4, 10, 25, and 50. The resulting measurements are normalized and plotted as a function of n in Fig. 4-4 through 4-18.

The choice of normalization factor is arbitrary, but the choices seem to be reasonable. The shape of these curves appears to be influenced by both the relative size of objects and the number of objects in the subset. For example, many of the figures tend to make subsets RM43 and RM47 stand out from the rest. In fact, these two subsets are considerably different from the others. RM43 contains a very high density of small objects of similar shape, while RM47 contains practically nothing.

4.3 GRADIENT STATISTICS

The differentiation of images tends to accentuate edges. Any derivative operator can therefore be used to detect edges, since the value of a point represents the strength of any edge at that point. For digital images, differences are normally used to approximate the derivatives. One such operator is the gradient, which is defined as:

$$\left\{ [f(i,j) - f(i-1,j)]^2 + [f(i,j) - f(i,j-1)]^2 \right\}^{1/2}$$

One approximation to the gradient often used is:

$$\text{Maximum} \left| f(i,j) - f(r,s) \right| ,$$

where r, s ranges over either four or eight neighbors of the point f (i, j). The gradient was computed for all the subsets under study using the maximum difference of the four horizontal and vertical neighbors. Table 4-4 shows the resulting values for the mean and standard deviation of the gradient images.

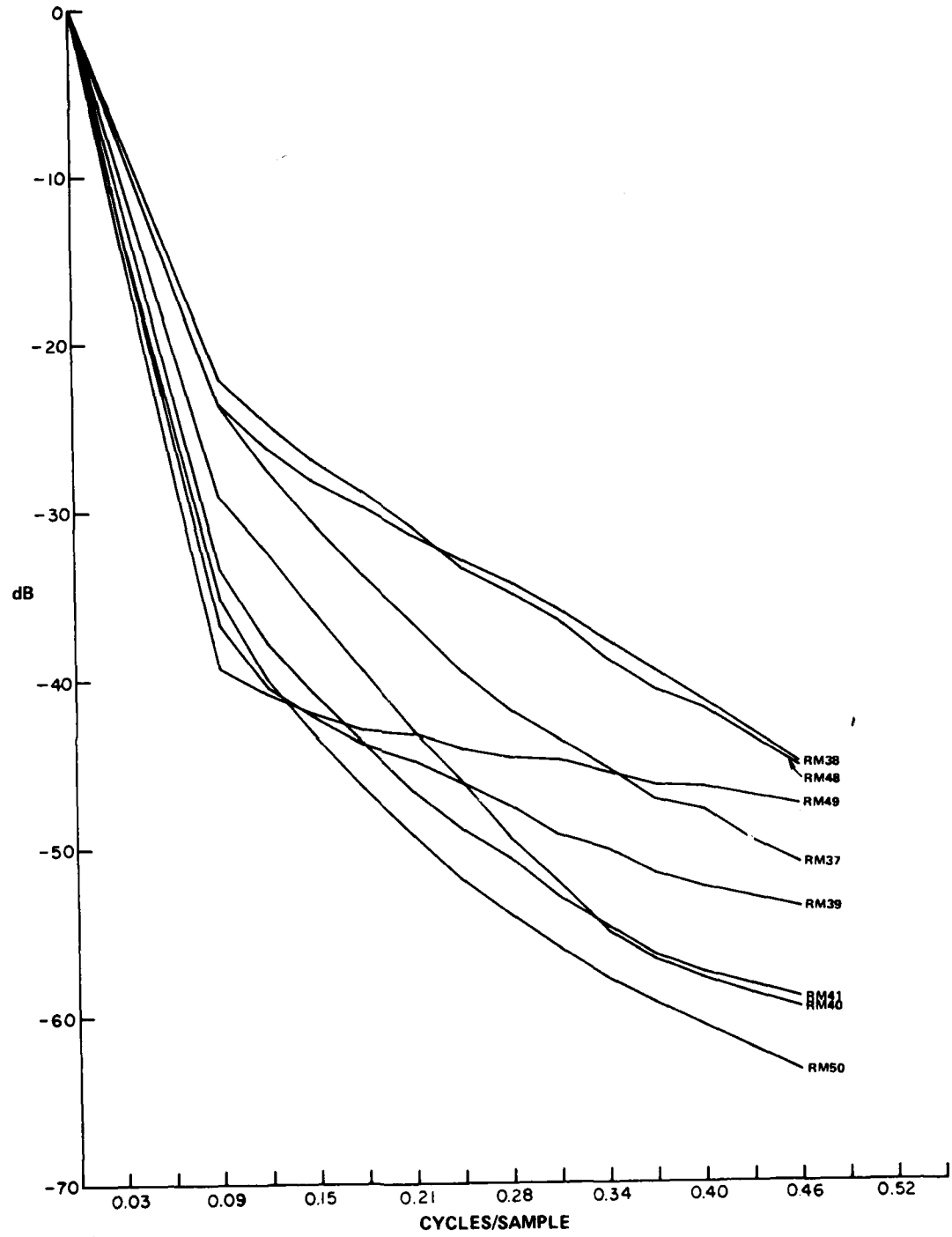


Fig. 4-1. Average power spectral density from the two axes; RM37-RM41, RM48-RM50.

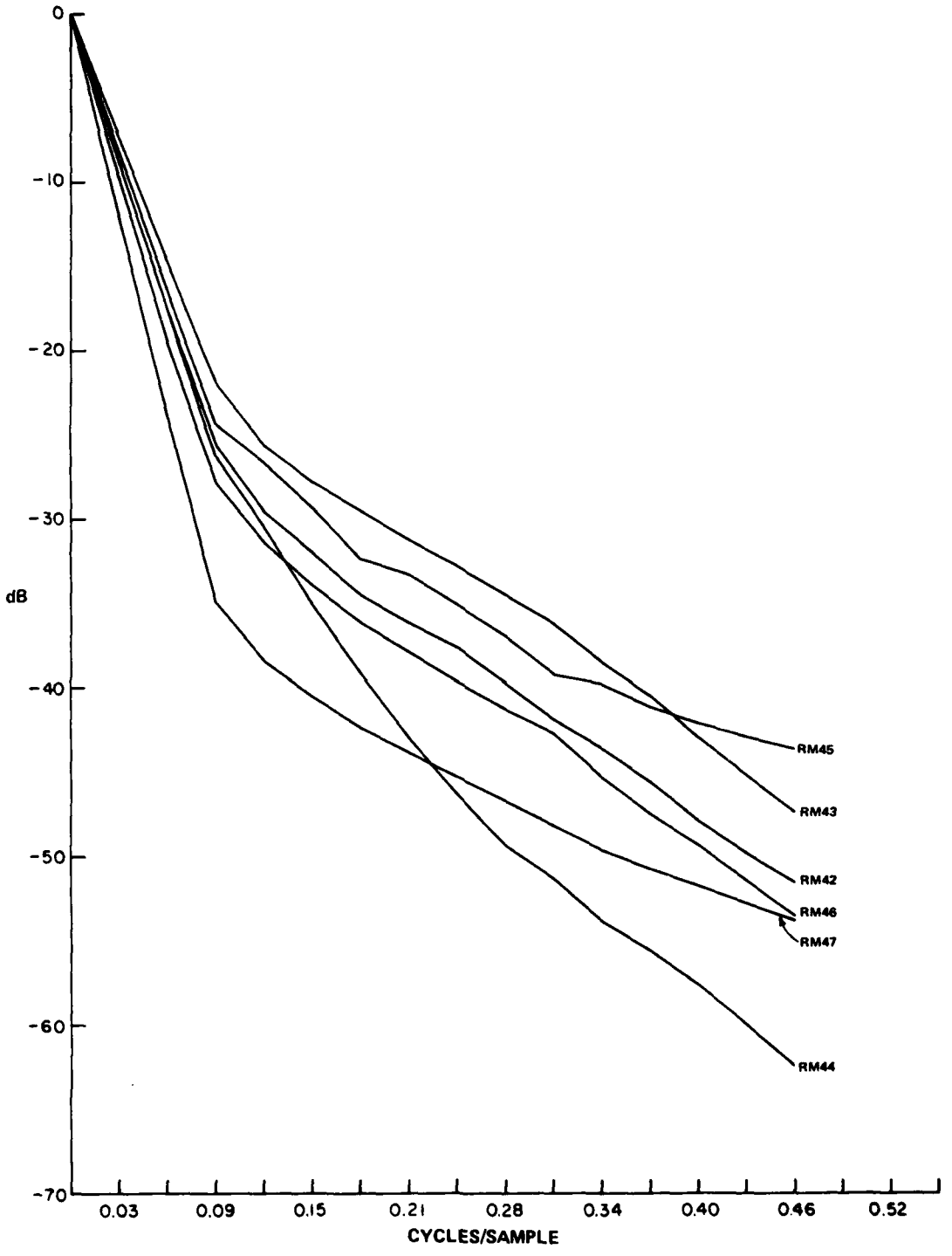


Fig. 4-2. Average power spectral density from the two axes; RM42-RM47.

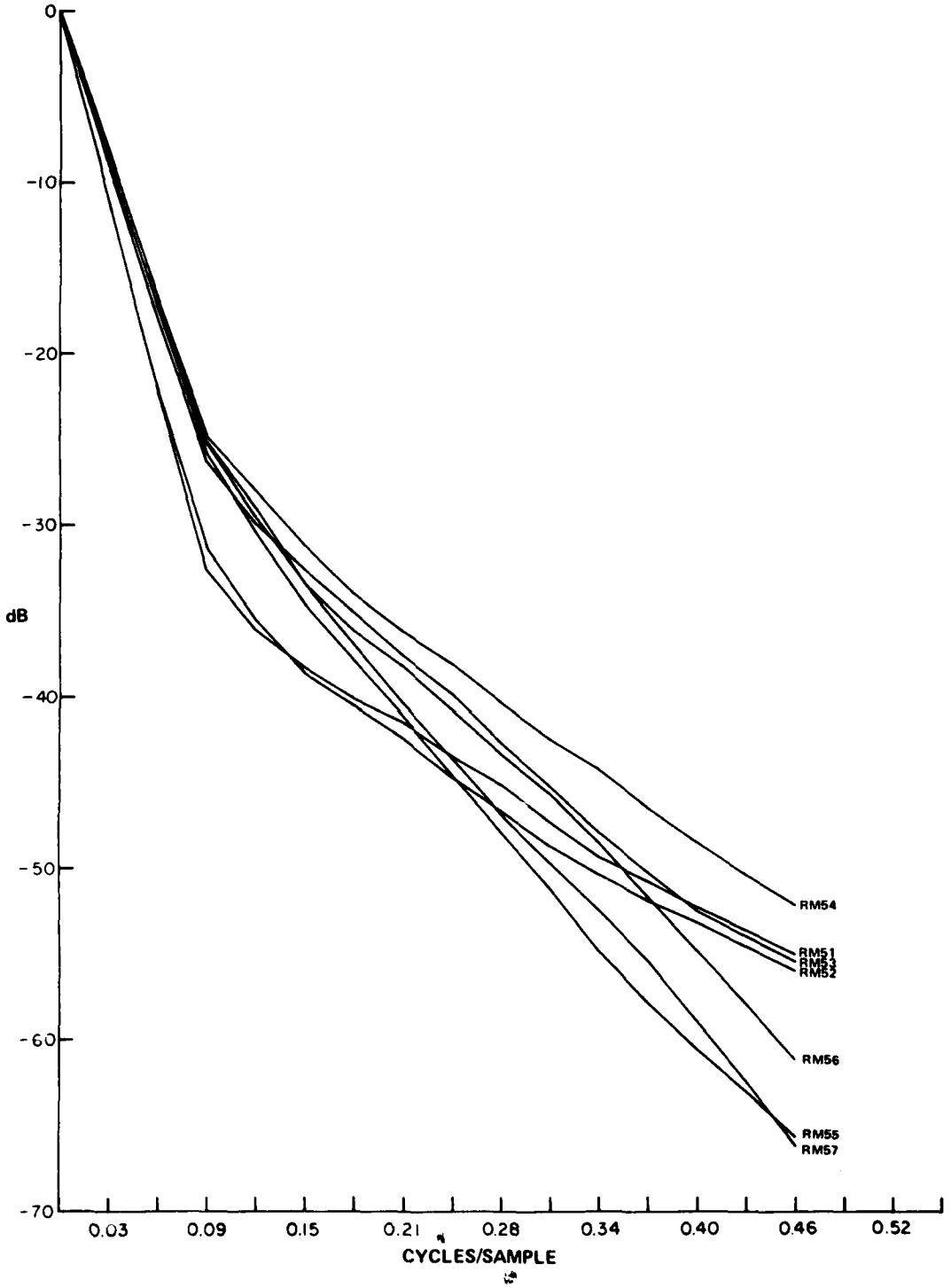


Fig. 4-3. Average power spectral density from the two axes; RM51-RM57.

TABLE 4-3. DYNAMIC RANGE IN EACH OF FOUR FREQUENCY BANDS IN THE POWER SPECTRAL DENSITY

Tape No.	Band 1 [dB]	Band 2 [dB]	Band 3 [dB]	Band 4 [dB]
RM37	30.8	12.7	7.4	51.0
RM38	28.3	7.5	9.1	45.0
RM39	42.3	6.8	4.4	53.6
RM40	36.1	16.0	7.3	59.6
RM41	42.9	10.1	6.0	59.0
RM42	32.0	10.0	9.5	51.5
RM43	27.9	8.3	11.1	47.3
RM44	35.0	16.3	11.1	62.4
RM45	29.2	9.9	4.4	43.6
RM46	33.9	8.8	10.7	53.4
RM47	40.6	7.5	5.5	53.7
RM48	27.1	9.4	8.7	45.3
RM49	42.1	2.7	2.7	47.5
RM50	43.6	12.4	7.4	63.4
RM51	38.5	8.9	7.5	54.9
RM52	38.6	9.9	7.3	55.8
RM53	32.6	12.5	10.2	55.3
RM54	31.4	11.0	9.6	52.0
RM55	34.5	16.5	14.4	65.4
RM56	33.4	12.1	15.5	61.0
RM57	33.4	16.2	16.4	66.0

Band 1: 0-0.15 Cycles/Sample
Band 2: 0.15-0.31 Cycles/Sample
Band 3: 0.31-0.46 Cycles/Sample
Band 4: 0-0.46 Cycles/Sample

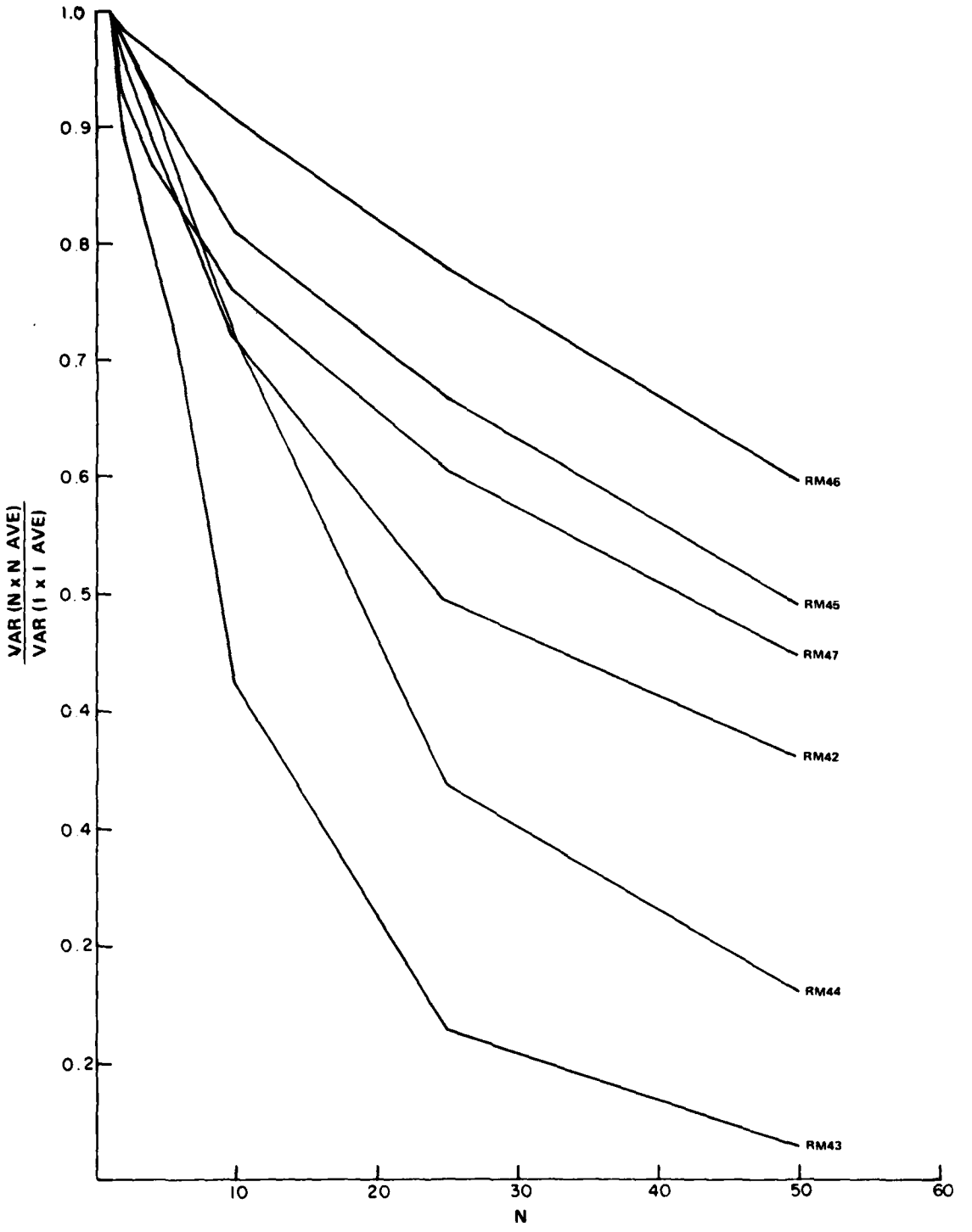


Fig. 4-4. Normalized variance of local average vs. neighborhood size; RM42-RM47.

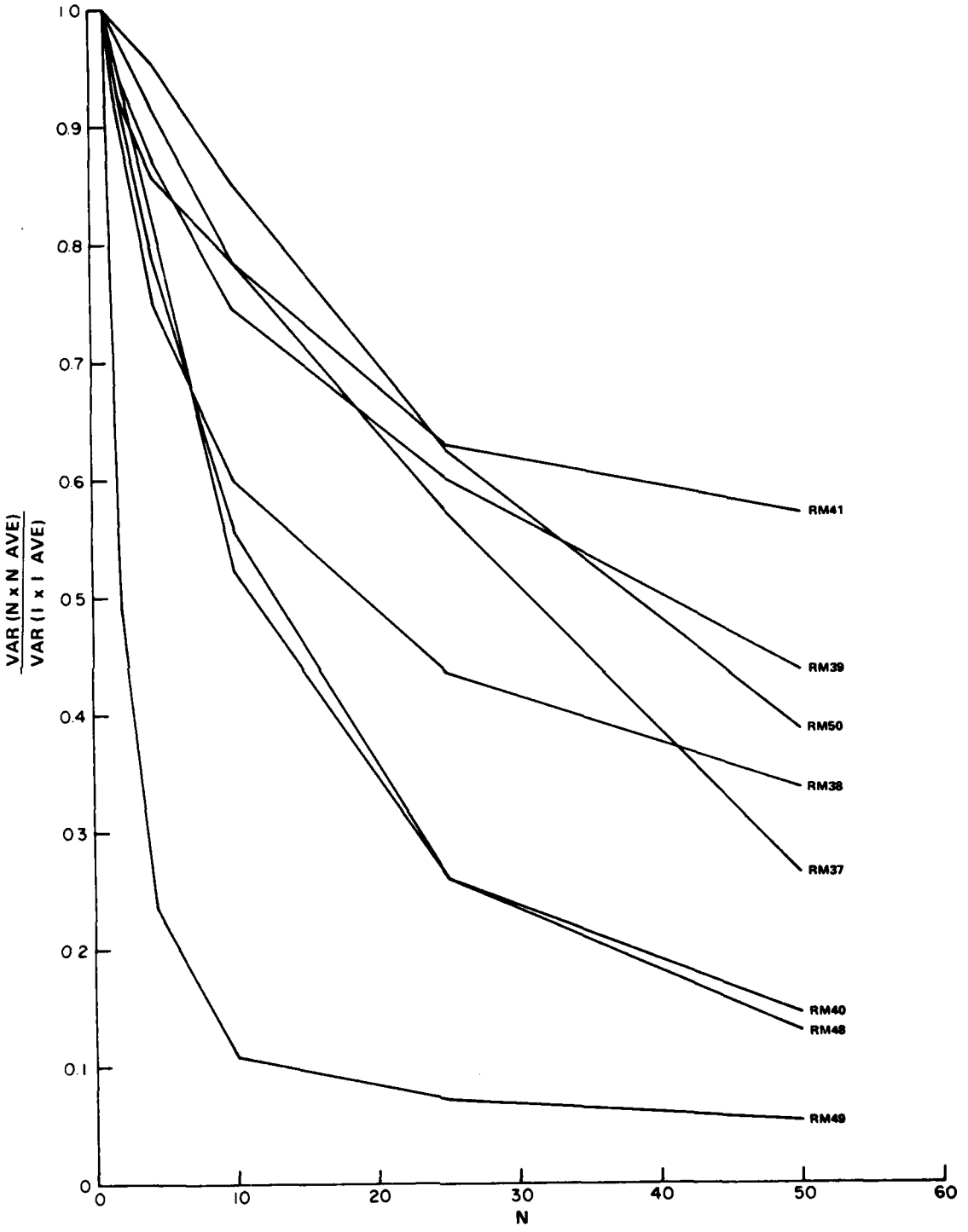


Fig. 4-5. Normalized variance of local average vs. neighborhood size; RM37-RM41, RM48-RM50.

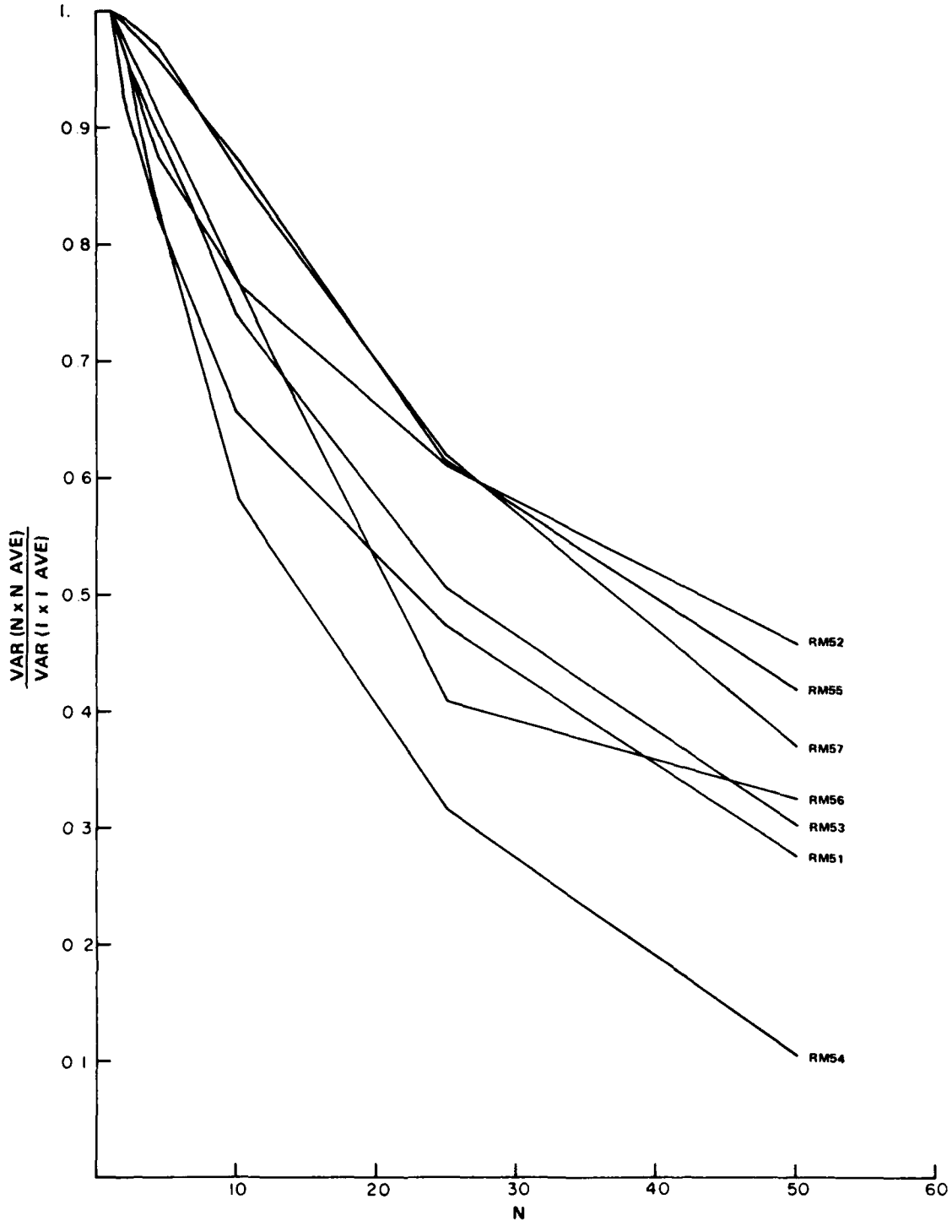


Fig. 4-6. Normalized variance of local average vs. neighborhood size; RM51-RM57.

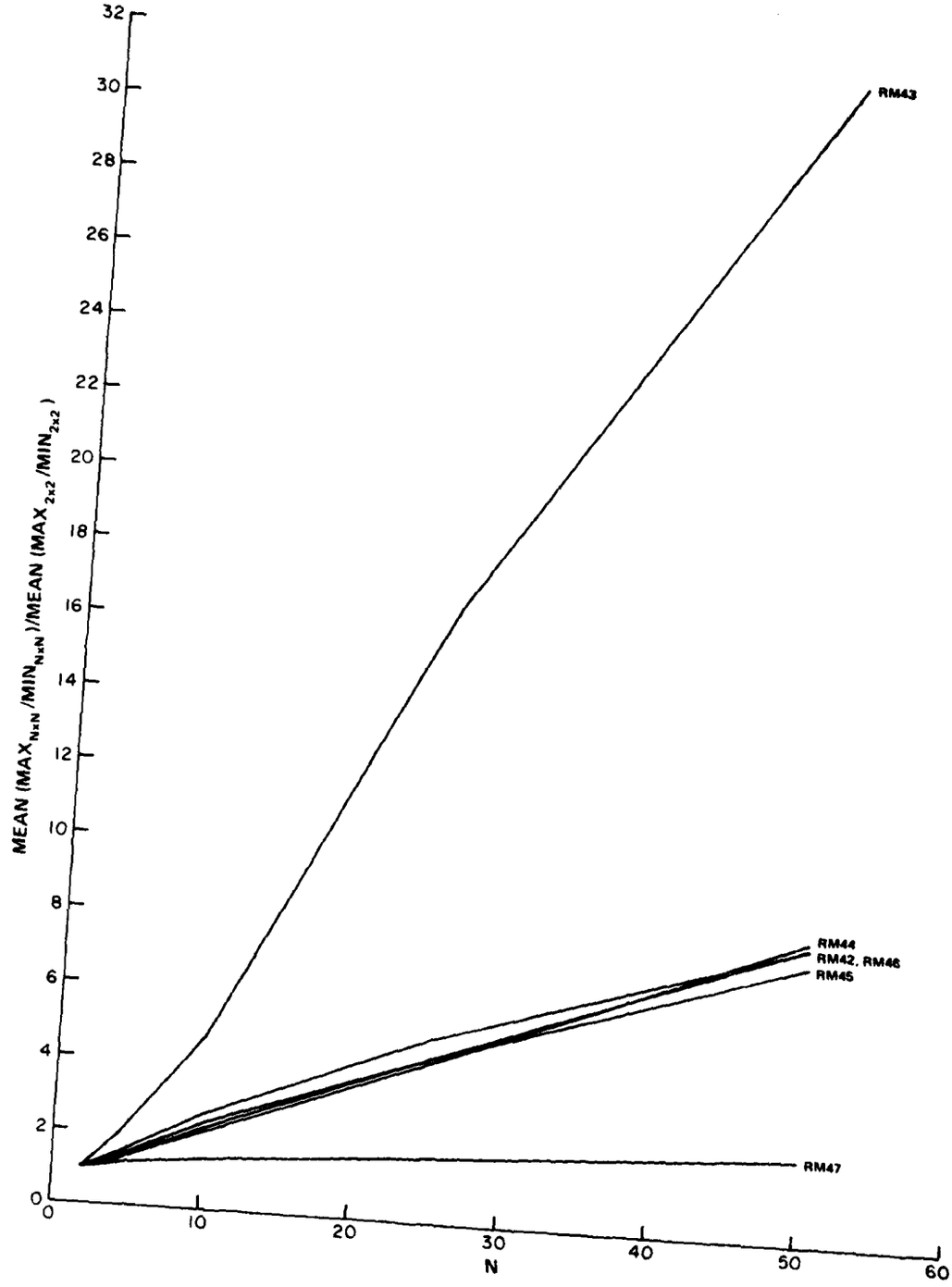


Fig. 4-7. Mean of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM42-RM47.

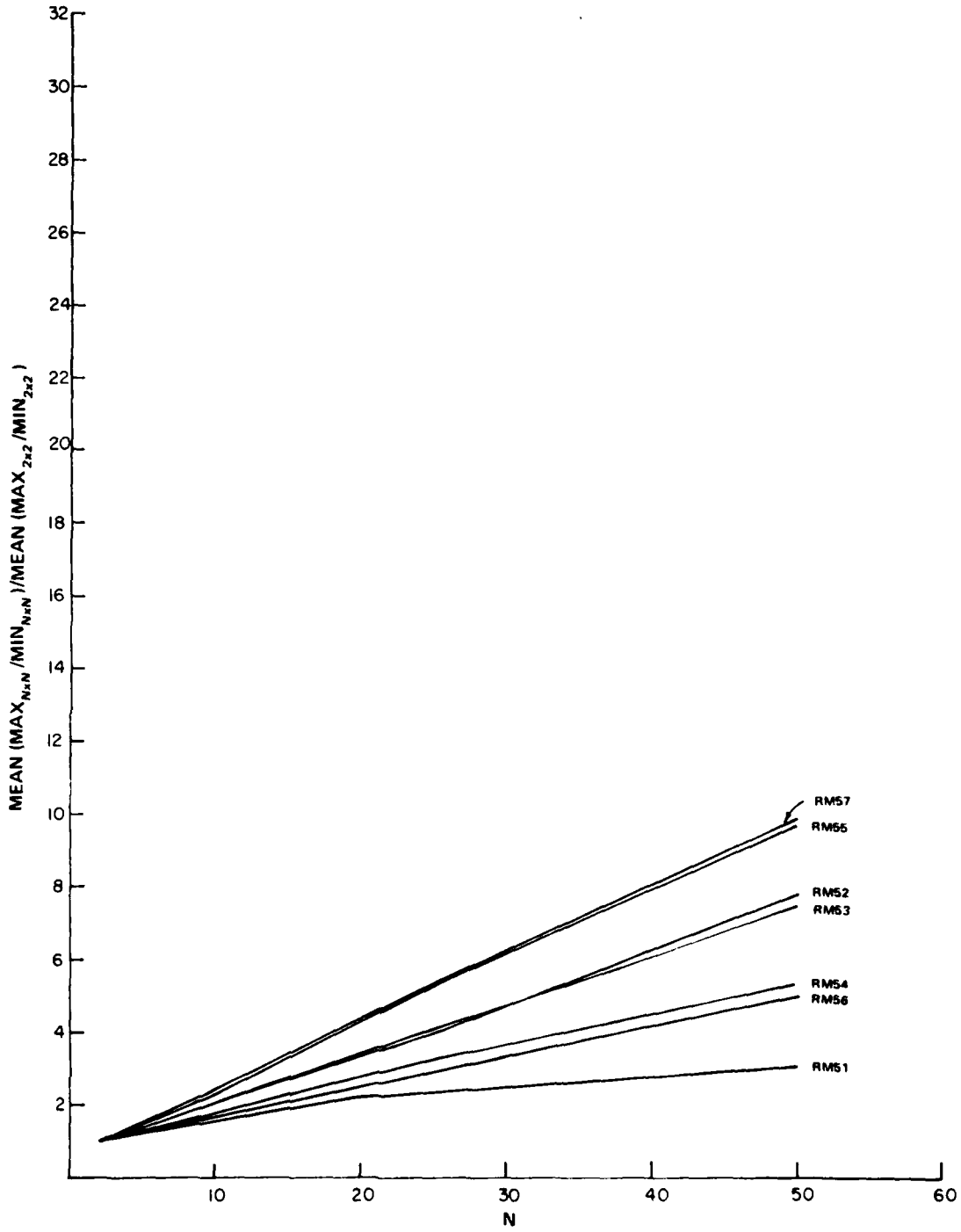


Fig. 4-8. Mean of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM51-RM57.

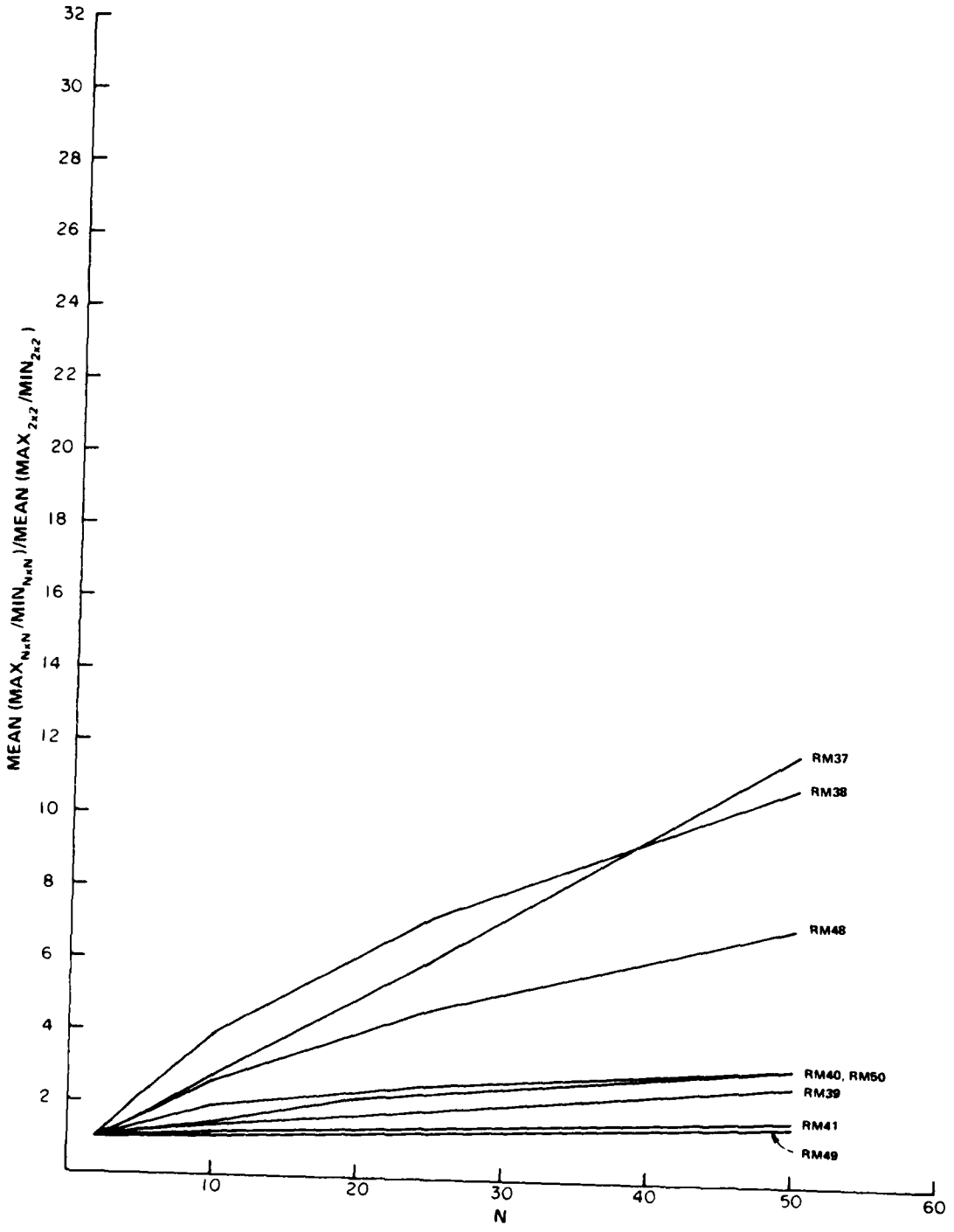


Fig. 4-9. Mean of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM37-RM41, RM48-RM50.

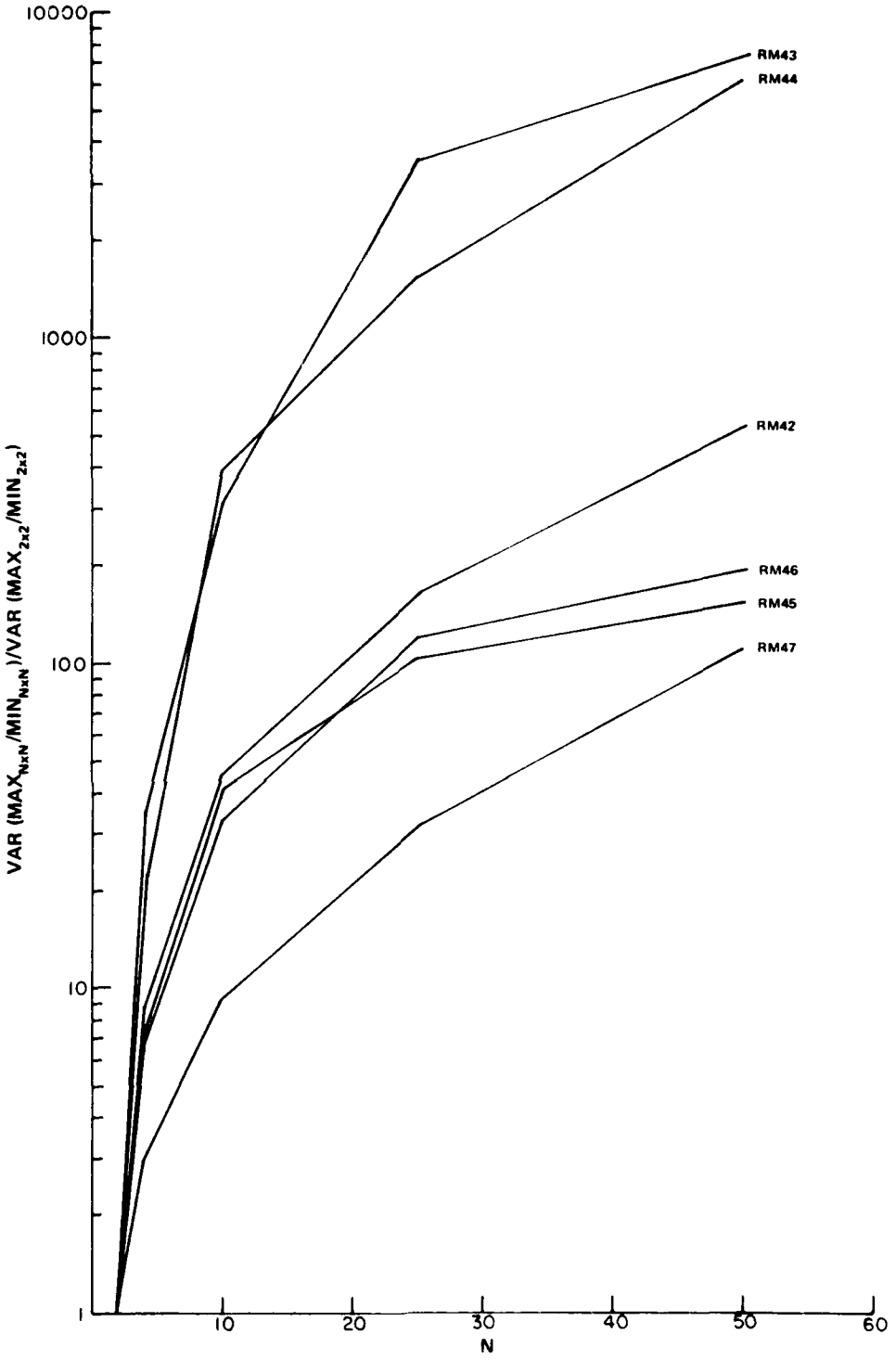


Fig. 4-10. Variance of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM42-RM47.

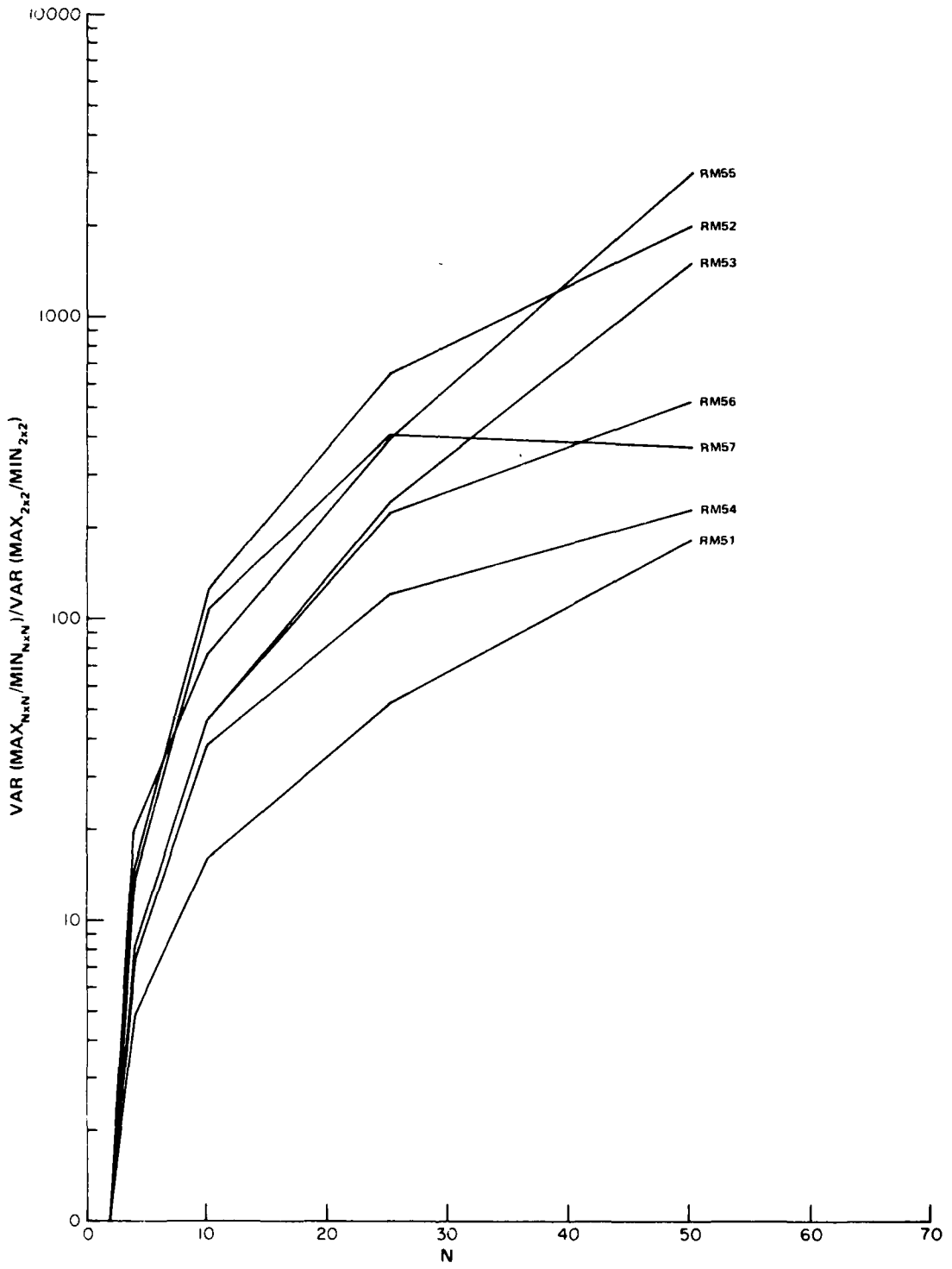


Fig. 4-11. Variance of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM51-RM57.

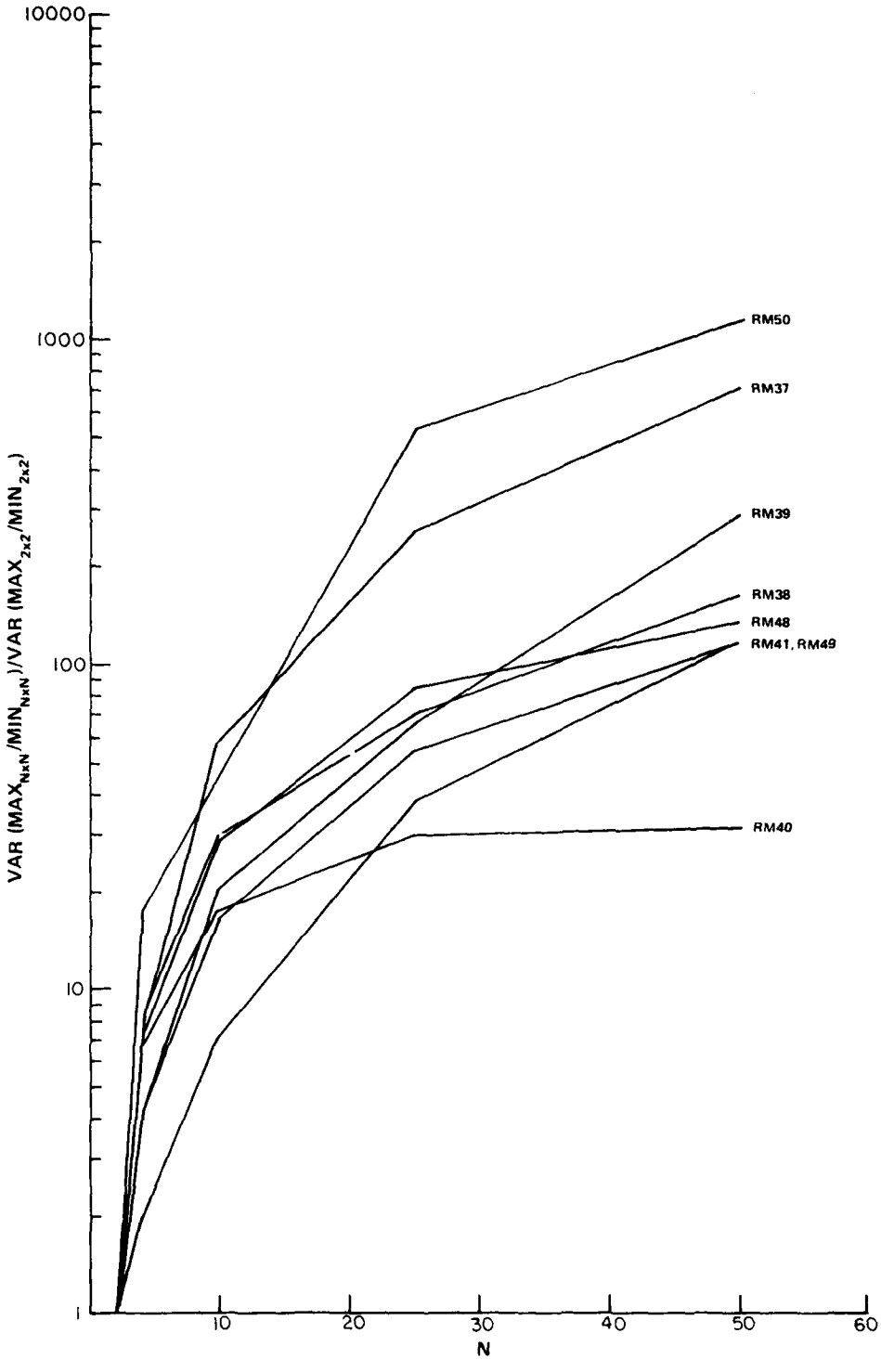


Fig. 4-12. Variance of local brightness ratio (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM37-RM41, RM48-RM50.

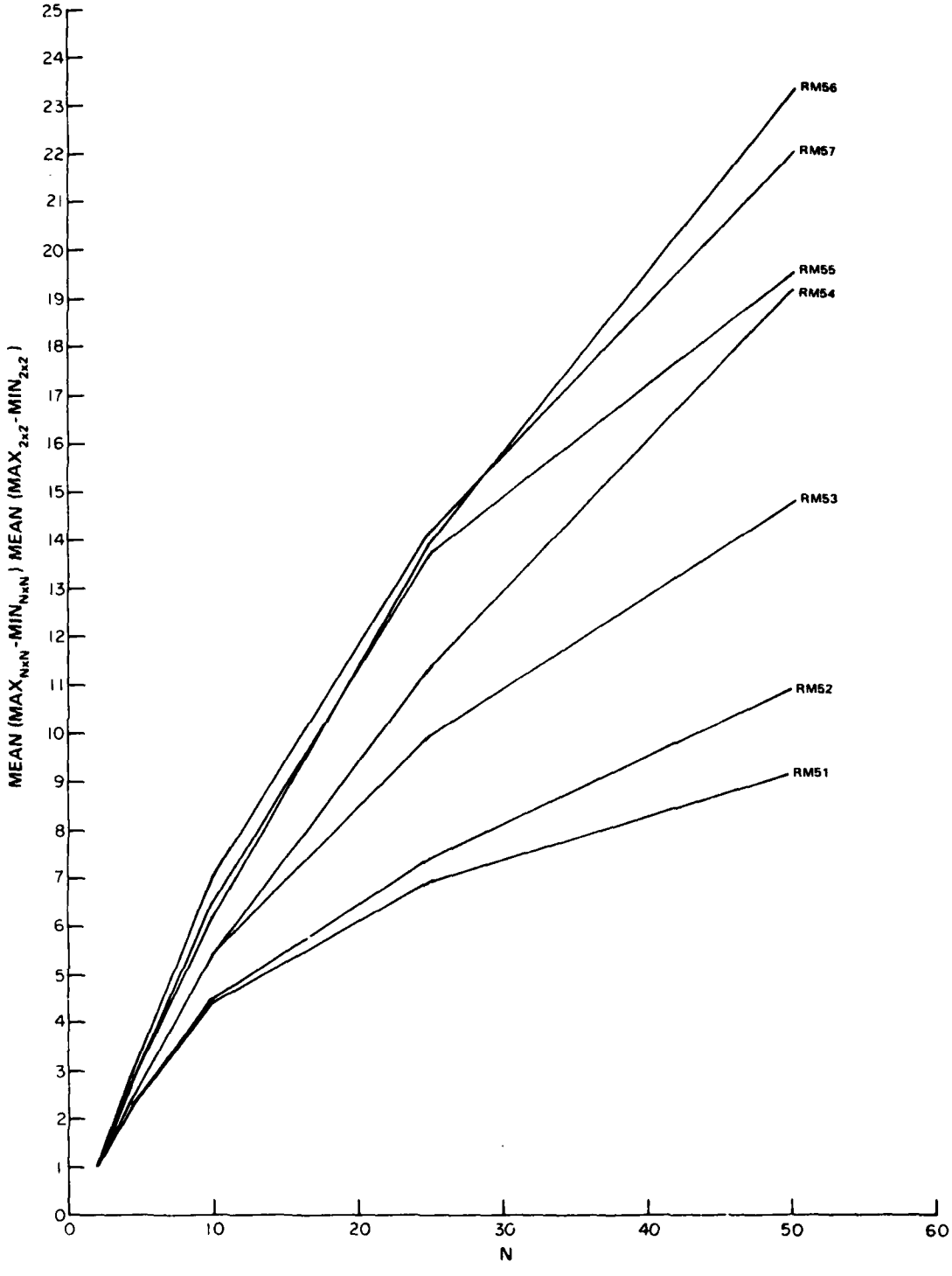


Fig. 4-13. Mean of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM42-RM47.

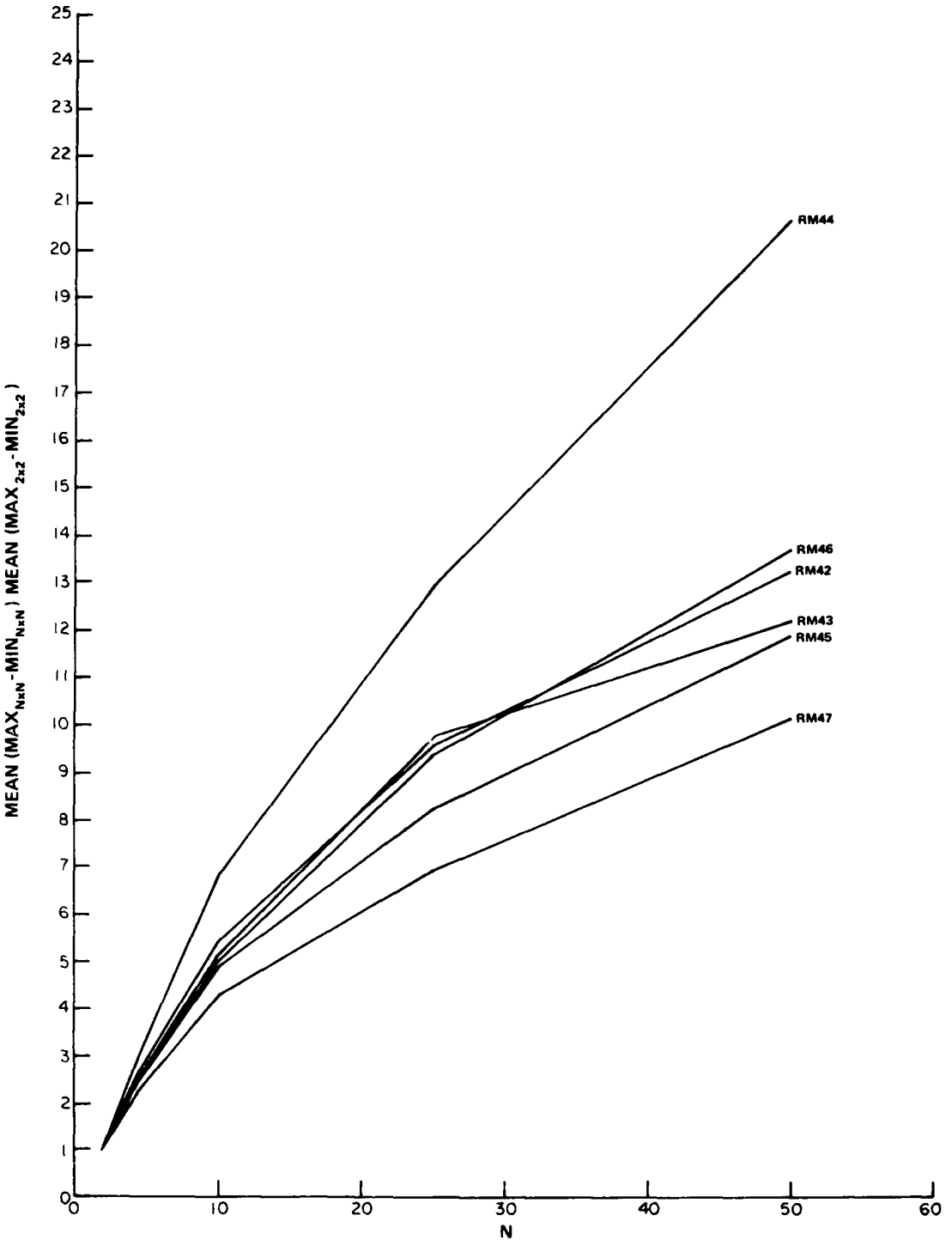


Fig. 4-14. Mean of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM51-RM57.

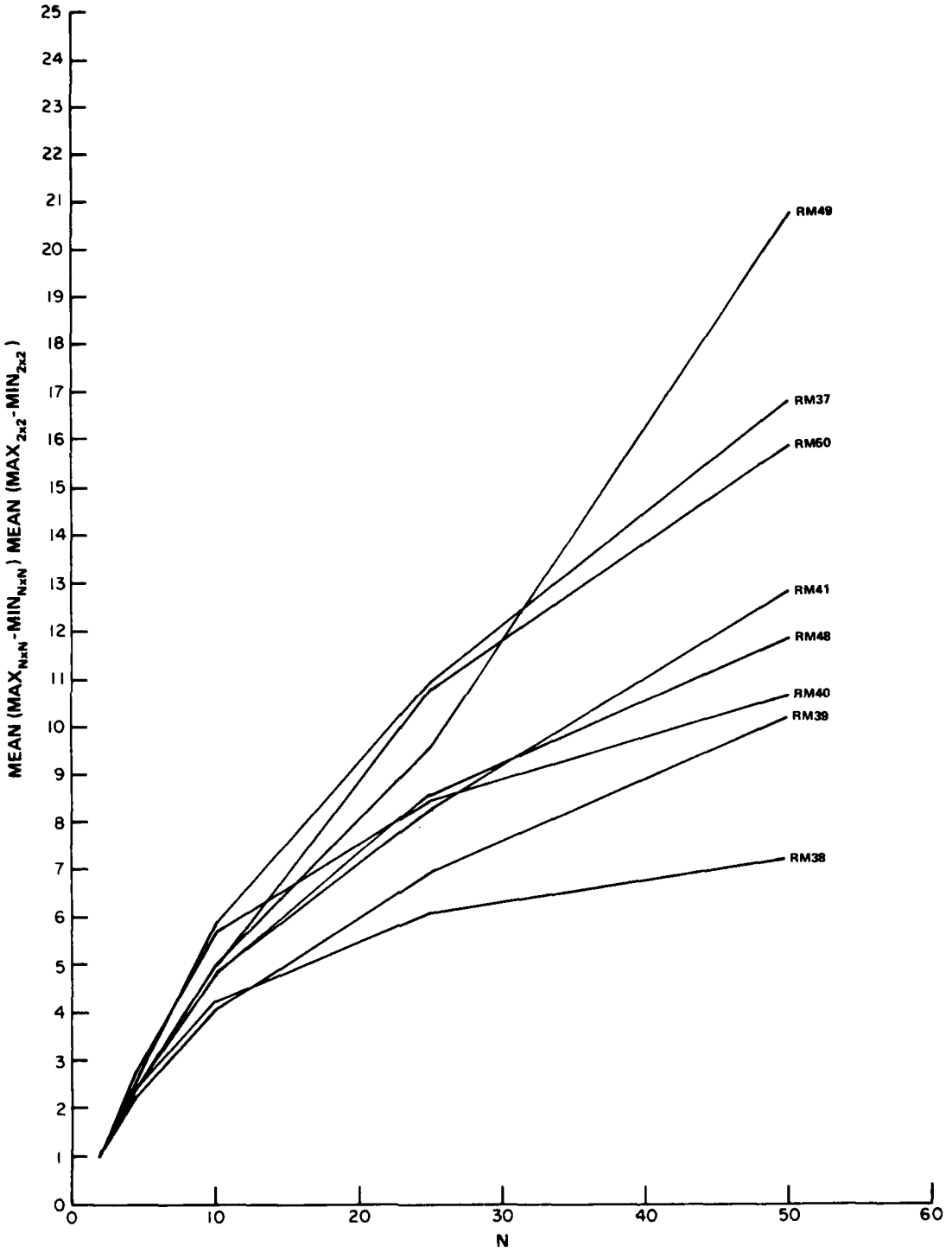


Fig. 4-15. Mean of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM37-RM41, RM48-RM50.

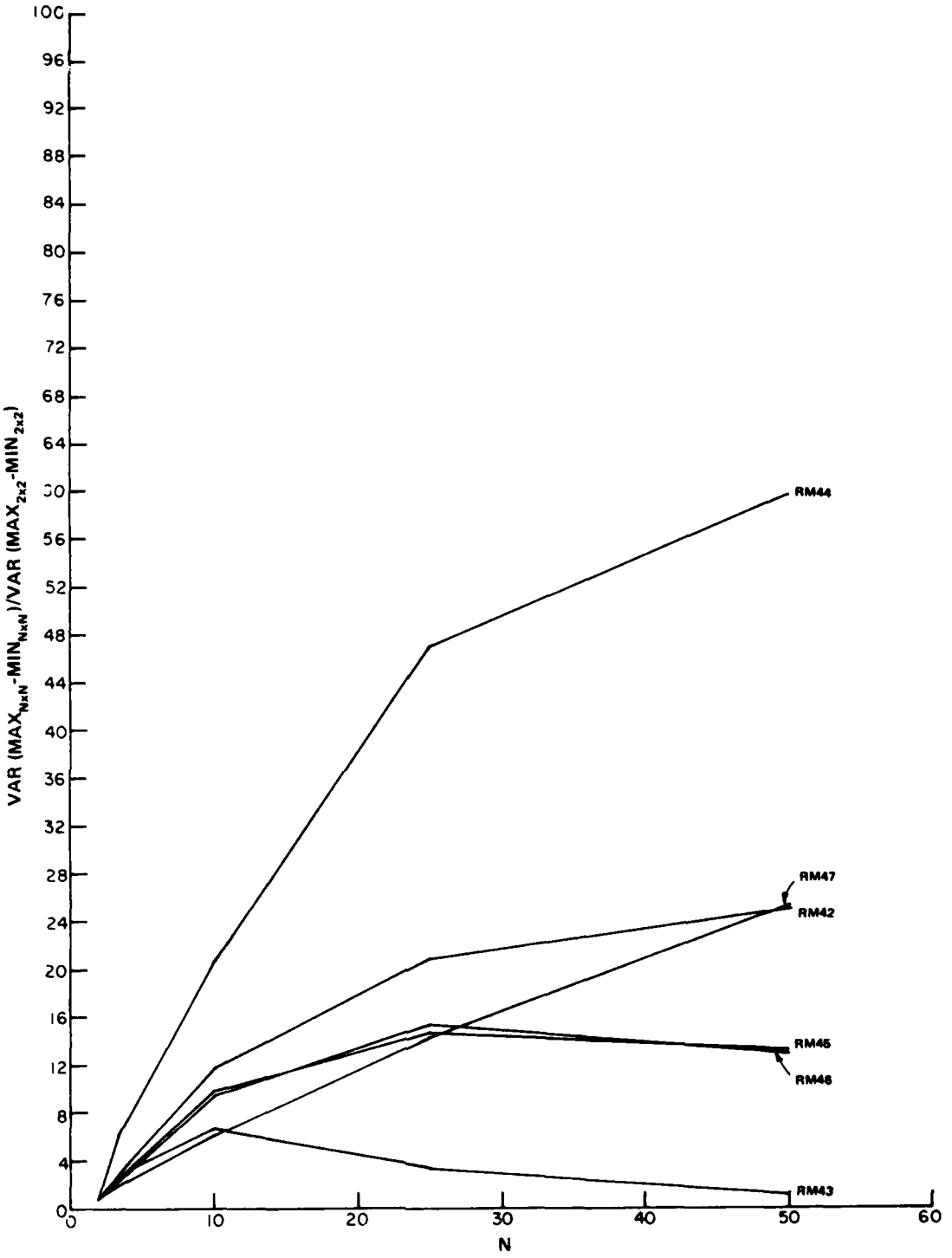


Fig. 4-16. Variance of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM42-RM47.

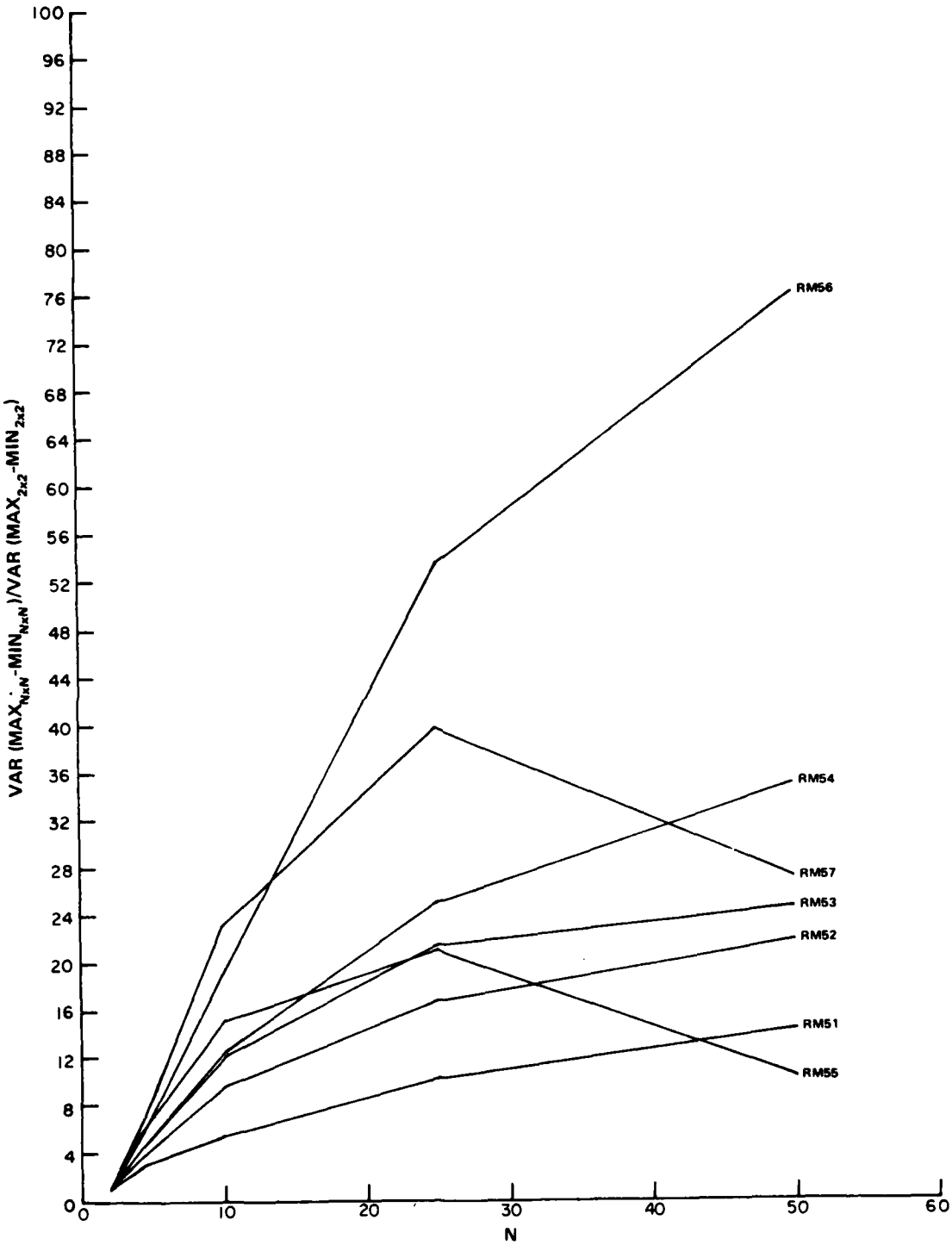


Fig. 4-17. Variance of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM51-RM57.

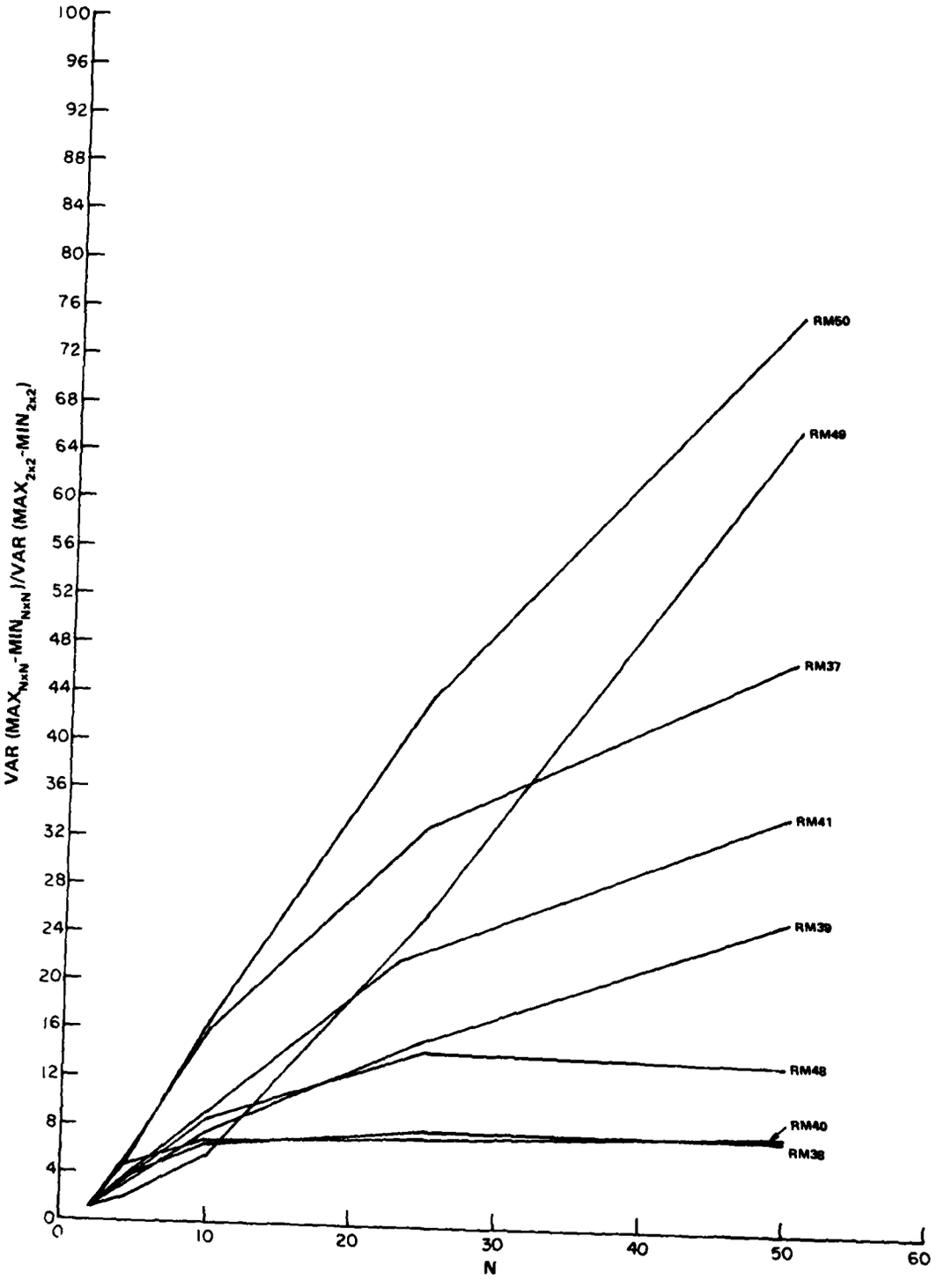


Fig. 4-18. Variance of local brightness range (normalized to 2 x 2 neighborhood) vs. neighborhood size; RM37-RM41, RM48-RM50.

TABLE 4-4. STATISTICS FROM THE GRADIENT IMAGES COMPUTED FROM THE FOUR HORIZONTAL AND VERTICAL NEIGHBORS

	Gradient	
Tape No.	Mean	Standard Deviation
RM37	3.26	6.09
RM38	10.15	14.42
RM39	3.8	2.34
RM40	3.46	1.85
RM41	1.44	1.08
RM42	6.57	9.29
RM43	13.46	16.98
RM44	3.24	3.35
RM45	8.17	10.44
RM46	7.95	10.62
RM47	2.51	1.69
RM48	6.92	7.71
RM49	0.3	0.53
RM50	2.14	1.54
RM51	6.07	3.55
RM52	3.78	3.44
RM53	7.25	9.12
RM54	3.62	6.07
RM55	6.97	8.1
RM56	2.18	5.54
RM57	4.85	5.71

Section 5

K-L TRANSFORM OF STATISTICS VECTOR

The next step in this effort was to try to relate the measurements made on the image subsets to the performance of a specific compression algorithm. The first attempt at doing this involved the use of clustering algorithms to group the subsets. The first problem, however, is to determine what the inputs to the clustering algorithm should be. Since many parameters were measured for each subset, chances are fairly good that some of them are correlated. One approach suggests that transforming the set of parameters into another, uncorrelated set of parameters should reduce the number of inputs required to classify a given subset. The Karhunen-Loeve (K-L) transform is a process that will optimally decorrelate a set of data. The K-L transform matrix was obtained from 20 parameters from each of 19 image subsets. The transform was then performed on the parameters of each scene. Table 5-1 shows the set of image parameters used as input to the transform. The magnitudes of these features differ considerably. For example, the average correlation is by definition no larger than unity, while other features may be several orders of magnitude larger. In order to weigh each feature equally, the average of each feature over the 19 image subsets was normalized to 100. Table 5-2 is the scaled input data for the first eight image subsets. The set of eigenvalues computed for this data is given in Table 5-3. The first eight basis vectors for the K-L Transform are shown in Table 5-4 and the output of the K-L transform operating on the scaled input data is given in Table 5-5 for the same eight subsets. This data was used to derive inputs to a clustering process in an attempt to group the subsets.

TABLE 5-1. ONE SET OF IMAGE PARAMETERS USED AS INPUT TO K-L TRANSFORM

1)	Variance
2)	$\frac{\text{maximum - mean}}{\text{mean - minimum}}$ (MAX-MEAN)/(MEAN-MIN)
3)	Average correlation
4)	Power spectral density at .09 cyc/sample
5)	Power spectral density at .18 cyc/sample
6)	Power spectral density at .31 cyc/sample
7)	Power spectral density at .46 cyc/sample
8)	Variance of the means of 2 X 2 areas
9)	Variance of the means of 10 X 10 areas
10)	Variance of the means of 50 X 50 areas
11)	Average gradient
12)	Variance of the gradient
13)	Variance of the means of 2 X 2 areas normalized by scene variance
14)	Variance of the means of 10 X 10 areas normalized by scene variance
15)	Variance of the means of 50 X 50 areas normalized by scene variance
16)	Average ratio of max to min over 10 X 10 area normalized by average ratio over 2 X 2 area
17)	Average ratio of max to min over 50 X 50 area normalized by average ratio over 2 X 2 area
18)	Variance of the ratio of max to min over 10 X 10 area normalized by variance of ratio over 2 X 2 area
19)	Variance of the ratio of max to min over 50 X 50 area normalized by variance of ratio over 2 X 2 area
20)	Skewness

TABLE 5-2. INPUT DATA FOR FIRST EIGHT SUBSETS
INPUT DATA SCALED
20 PARAMETERS FROM 19 SOURCES (MAX/MIN)

		1 RM39	2 RM40	3 RM41	4 RM42	5 RM43	6 RM44	7 RM45	8 RM46
1	VAR	11.22565	14.04294	2.80082	112.35408	158.37123	40.59273	199.13673	387.69677
2	X-U/U-MN	55.92603	45.42952	24.93518	152.93573	105.89137	105.53517	98.64832	75.18550
3	P AVE	100.27397	102.98436	103.40720	103.91675	96.94560	106.11758	104.91417	106.62715
4	PSD .09	127.56638	101.24152	115.98206	89.41438	76.30397	90.83643	84.17717	96.66325
5	PSD .18	115.87868	104.54387	115.77322	91.02116	77.60391	103.46309	85.27469	95.21239
6	PSD .31	106.22154	112.62329	114.28302	90.52972	78.09265	110.61873	84.53752	92.08167
7	PSD .46	96.96729	107.67923	106.66765	93.02936	85.60507	112.82747	78.77690	96.56992
8	VAR(U) 2	10.81566	13.74228	2.77115	111.34805	145.79303	41.00562	198.68161	393.95190
9	VAR(U) 10	10.54061	9.32638	2.81600	102.61594	85.46164	37.07298	202.82930	446.45166
10	VAR(U) 50	12.34600	5.38830	4.15370	104.94099	12.13674	17.00189	251.78519	596.37451
11	AVE (GRD)	76.22946	69.55287	28.91187	131.82765	269.97094	65.10181	163.86731	159.45633
12	VAR (GRD)	10.87021	6.83854	2.34634	171.43008	572.80615	22.34387	216.56647	224.05214
13	NV(U) 2	100.39612	102.00847	103.08337	103.29836	95.88152	105.23318	104.05080	105.87813
14	NV(U) 10	106.46930	75.30756	114.11548	103.72823	61.02509	104.01677	115.55815	130.70622
15	NV(U) 50	130.92665	46.67001	180.41566	113.38623	9.08343	51.05511	153.79184	186.99333
16	NA(RT) 10	67.20122	90.89084	58.01541	130.05122	232.06172	104.91122	115.54741	107.81198
17	NA(RT) 50	38.77264	45.70644	25.18806	113.48784	442.48925	115.32741	106.41254	113.77084
18	NV(RT) 10	26.67575	22.63911	27.43344	61.35947	423.24560	527.11401	56.74803	44.06332
19	NV(RT 50	23.18379	2.42966	9.18078	41.63423	580.21752	488.79174	12.08031	15.19002
20	SKEWNESS	0.04352	6.26093	0.21980	37.38168	56.21126	34.19682	29.07512	23.25705

TABLE 5-3. EIGEN VALUES

1	83971.12500
2	53959.08200
3	29199.74200
4	11952.01900
5	3622.87620
6	1546.10100
7	1220.70840
8	642.83105
9	218.64494
10	156.82895
11	90.93962
12	47.47408
13	19.60738
14	14.31018
15	2.05236
16	1.07275
17	0.39676
18	0.02022
19	-0.00287
20	-0.00592

TABLE 5-4. K-L TRANSFORM, FIRST EIGHT VECTORS K-L TRANSFORM MATRIX SCALED
20 PARAMETERS FROM 19 SOURCES (MAX/MIN)

	1	2	3	4	5	6	7	8
1	0.38704	-0.01973	0.02037	-0.03340	-0.02042	-0.01044	0.00289	0.39010
2	-0.14572	0.02652	-0.00028	-0.01930	-0.01248	-0.00850	-0.00444	-0.15847
3	0.10810	0.17658	-0.04953	-0.01682	-0.02138	-0.02729	-0.01334	0.10752
4	0.04287	-0.08473	-0.00095	0.03806	0.05880	0.07020	0.07175	0.06423
5	-0.01659	0.86875	0.14802	-0.16087	-0.08118	-0.03510	0.02898	-0.01114
6	-0.18235	-0.06162	0.00955	0.15352	0.10194	0.01747	-0.03874	-0.19461
7	0.17148	0.16707	0.08688	0.03180	0.07377	0.11213	0.15174	0.17734
8	0.12609	-0.20517	-0.02976	0.02138	0.05680	0.07922	0.12081	0.13833
9	-0.26829	0.13642	-0.01292	0.13784	0.07716	0.03536	0.05654	-0.25715
10	0.04422	0.10882	-0.37789	0.20231	-0.05380	-0.23954	-0.29803	0.04312
11	-0.04815	-0.16183	0.26416	0.16417	0.16100	0.11395	0.14755	-0.06092
12	0.07776	0.25505	-0.20809	0.64796	0.39888	0.19953	0.27339	0.06874
13	0.03926	-0.04891	-0.00224	-0.18226	0.15694	0.42510	0.41199	0.03609
14	-0.02844	-0.05870	-0.11475	0.25861	0.15173	-0.10681	-0.31741	0.01038
15	-0.12020	-0.02023	-0.48571	-0.25026	-0.20524	-0.14601	0.54680	-0.13513
16	-0.40548	-0.02105	0.11397	0.04696	0.21151	-0.30130	0.16550	-0.19715
17	0.17277	-0.05957	0.26390	0.32528	-0.24706	-0.62681	0.41122	0.13895
18	0.18618	0.00859	-0.00945	-0.40638	0.75959	-0.40274	-0.00942	0.00542
19	0.63667	-0.00212	0.01326	0.05674	-0.05799	0.02027	0.00124	-0.74885
20	0.05681	-0.00116	-0.61534	-0.00281	0.00978	-0.05098	0.00900	0.05785

TABLE 5-5. K-L TRANSFORM OUTPUT, SCALED, FIRST EIGHT VECTORS K-L TRANSFORM SCALED
20 PARAMETERS FROM 19 SOURCES (MAX/MIN)

	1 RM39	2 RM40	3 RM41	4 RM42	5 RM43	6 RM44	7 RM45	8 RM46
1	62.49403	50.88686	36.39017	297.57446	667.18481	291.43725	488.31054	899.13745
2	17.44600	19.42497	-1.30803	36.98941	787.42260	567.49829	-79.21373	-320.55322
3	-20.17253	-3.84410	-39.58362	114.77890	276.90502	81.14323	141.92355	230.89392
4	11.66085	-4.91451	35.13426	-83.26674	-177.87393	350.13916	-71.22157	52.58801
5	91.94388	74.19873	78.63591	172.65994	66.09819	157.59686	133.37630	107.31207
6	125.09637	43.10143	171.76865	106.63306	124.78183	95.78680	140.01596	156.44505
7	98.32826	71.87593	104.15700	83.33401	101.44614	41.78230	45.95149	29.92934
8	58.55408	58.33151	90.87849	61.17044	78.50719	38.60442	58.79938	42.97171
9	29.90309	46.15826	12.61041	23.98346	36.47836	16.33220	0.94197	49.50922
10	-34.38304	-69.50121	-53.96196	-41.16692	-48.10658	-46.37646	-36.23814	-54.10298
11	179.92467	162.76195	160.89560	155.96519	173.72878	168.25385	174.17172	173.62569
12	96.03474	90.64294	79.81398	95.14587	90.64810	89.38428	77.75229	92.69019
13	132.13376	142.67592	138.02176	144.62457	135.69891	136.16154	139.69148	134.26004
14	59.52754	53.81633	51.84689	65.25711	57.10764	58.76674	60.73578	56.64899
15	-77.07056	-76.07549	-72.45526	-74.83540	-74.07919	-74.09943	-76.02089	-73.05180
16	42.52740	40.16046	42.84276	42.47874	41.48228	41.54596	40.33092	41.33557
17	25.36455	25.01302	25.61128	26.67569	25.25754	25.44138	24.27013	25.84351
18	-24.85956	-24.52473	-24.56743	-24.45270	-24.50966	-24.52025	-24.45186	-24.57187
19	9.85496	9.77803	9.79292	9.74505	9.77877	9.77789	9.77388	9.78451
20	9.67705	9.65950	9.66830	9.67239	9.66140	9.66416	9.64412	9.67615

Section 6

CLUSTERING EXPERIMENTS

As shown in Fig. 1-1 the features computed from the image subsets, as well as the K-L transform of a portion of those features, were used as input to clustering techniques. The purpose of this effort was to determine if the image subsets could be grouped in any way that related to the performance of a bandwidth compression algorithm.

Three clustering algorithms were considered for this work. The clustering algorithm, as described by Fukunaga[4] requires that the number of clusters be known, as well as the mean location of each cluster. Since this information is unknown, the algorithm was discarded. The algorithm explained by Coleman[5] requires that a tolerance be specified for the distance between the clusters. This technique was discarded also, since any information of the nature would be purely a guess. The third approach, described by Duda and Hart, [6] is the one that was used. In this algorithm, each sample is assumed to be a separate cluster.

Successive iterations are performed to combine the two clusters having the minimum distance between them. This is continued until only a single cluster remains.

The minimum-distance criterion used to combine clusters can be computed in a number of ways. Some possible measures include the distance between cluster means, the distance between the two farthest points of the clusters, and the distance between the two nearest points of the clusters. There may be variations in the resulting clusters, depending upon which method is chosen. All of our experiments utilized the distance between cluster means as the criterion for combining clusters. At each interation, a parameter is computed, which is intended to indicate the optimum number of clusters for the data. To compute this parameter (called beta), two matrices must be computed, namely the within-scatter matrix and the between-scatter matrix. The within-scatter matrix is a measure of how scattered the points in a given cluster are. The between-scatter matrix is a measure of how scattered the clusters are in relation to each other.

The within-scatter matrix is defined as the product of a matrix P and its transpose matrix P^T, where P is an M x N matrix defining M points in N-dimensional space after subtracting from each point the mean associated with the cluster to which it has been assigned.

---

[4]K. Fukunaga, Introduction to Statistical Pattern Recognition, New York; Academic Press, 1972, pp. 324-326.

[5]G. B. Coleman, op. cit.

[6]R. O. Duda and P. E. Hart, op. cit.

The between-scatter matrix is defined similarly to the within-scatter matrix, excepting that from each point, the mean of all the cluster means is subtracted rather than the mean of the cluster to which the point has been assigned.

The parameter beta is the product of the traces of these two scatter matrices, where the trace of a matrix is the sum of the diagonal elements. A maximum value for beta implies an optimum number of clusters for the data.

Two different scatter diagrams were plotted in an effort to investigate the relationship between the transform of the measured statistics and the number of bits required to transmit the cosine transform of the image with minimum mean square error (MSE). Since the optimum number of bits for quantizing the cosine transform coefficients is proportional to the logarithm of the variance of the coefficient, the parameter $\Sigma \log \sigma_{ij}^2$ was used as an indicator of the total number of bits required to transmit the image with minimum MSE. In Fig. 6-1, the length of the vector made up of the first six K-L transform coefficients is plotted against the sum of logs. Four clusters were indicated by the beta parameter as being the optimum number. Figure 6-2 is a scatter diagram of the sum of logs plotted against the first K-L transform component. In this case, six clusters are indicated. The trend of the beta parameter for the two figures is given in Fig. 6-3. The content of the subsets grouped by the clusters shown in the figures is presented in Table 6-1. These results are not very encouraging instinctively, since we do not expect images of fields to be similar to images of storage tanks or ships in terms of bandwidth compression.

In retrospect, however, it is not surprising that the results are confusing, since the sum-of-logs parameter does not represent constant quality. In other words, the cluster in Fig. 6-1, which contains both the image of a field (RM41) and the image of ships (RM42), may imply that a similar number of bits is required to represent those images with minimum MSE, but the resulting images may have totally different quality.

It did appear that two simple features related to the shape of the brightness histogram tended to isolate image subsets that contained man-made objects. A two-parameter clustering was performed using the brightness maximum minus the mean of one parameter and the brightness mean minus the minimum as the other parameter. Table 6-2 shows the resulting clusters that were determined with a description of the image content.

It will be noticed that the images assigned to clusters 1 and 2 are all natural scenes (except RM44), while cluster 3 contains only images that have man-made objects in them. RM44 appears to be out of focus, which may explain why it fell within the cluster that it did.

X = NUMBER PROPORTIONAL TO NO. OF BITS REQUIRED TO TRANSMIT COSINE TRANSFORM WITH MINIMUM MSE.

$\Sigma \text{ LOG}_{10} \ \sigma_{ij}^2$

Y = LENGTH OF VECTOR MADE UP OF FIRST SIX K-L TRANSFORM COEFFICIENTS USING SCALED INPUT FEATURES.

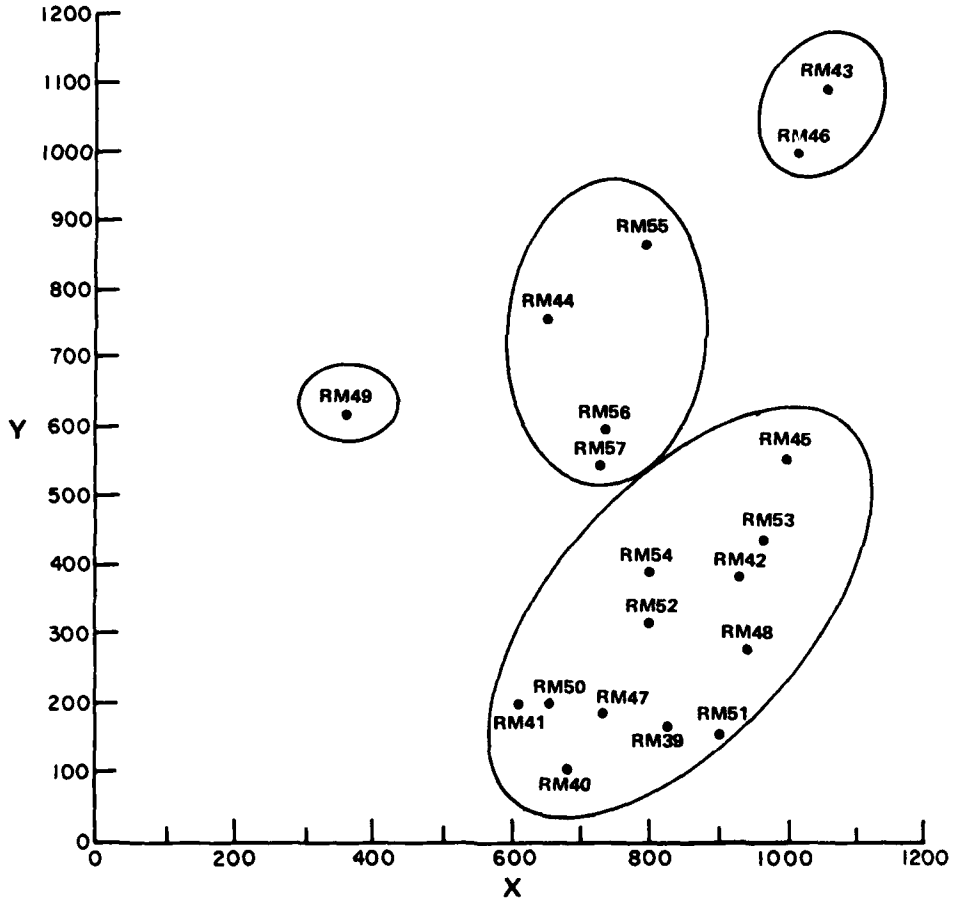


Fig. 6-1. Scatter diagram; first six K-L transform coefficients *vs.* sum of logs.

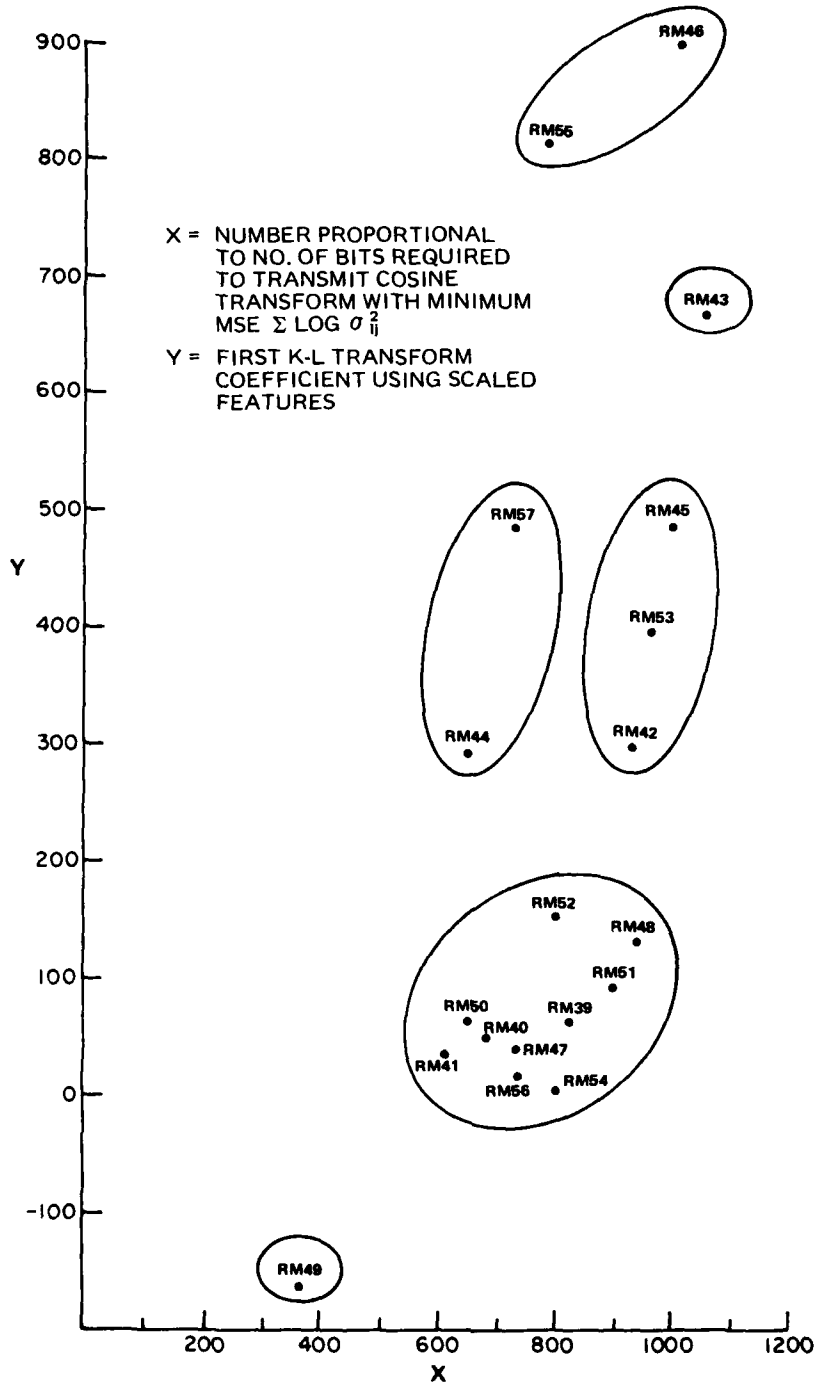


Fig. 6-2. Scatter diagram, sum of logs vs. first K-L transform coefficient.

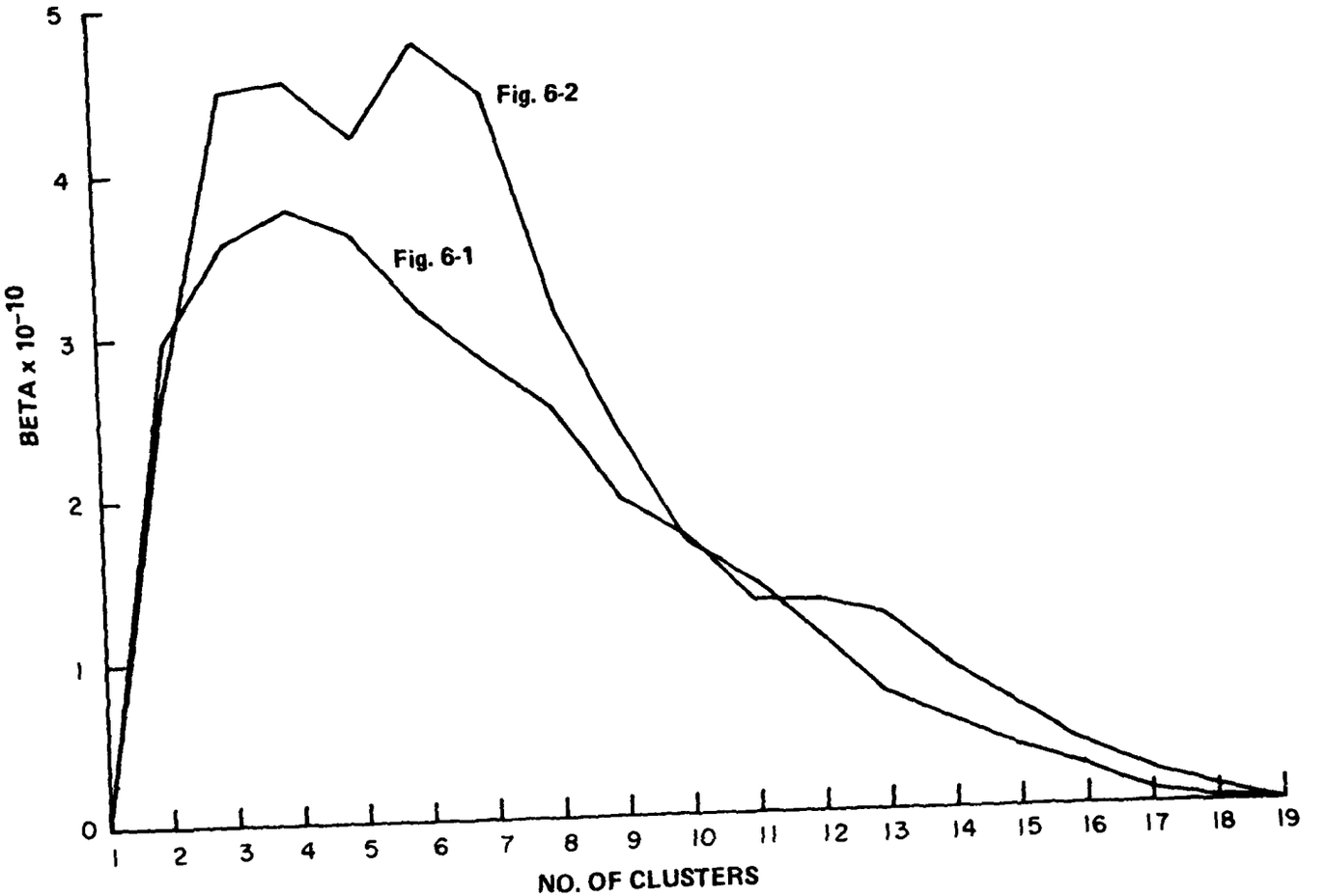


Fig. 6-3. Plot of beta parameter for the clustering experiments in Figs. 6-1 and 6-2.

TABLE 6-1. SUBSET CONTENT GROUPED BY CLUSTERS DETERMINED FROM FIGURES 6-1 AND 6-2.

FIGURE 6-1.		FIGURE 6-2.	
Subset	Content	Subset	Content
RM43	Parking Lot	RM46	Storage Tanks
RM46	Storage Tanks	RM55	Chemical Plant
RM49	Water	RM43	Parking Lot
RM44	Railroad	RM44	Railroad
RM55	Chemical Plant	RM57	Chemical Plant
RM56	Airplanes		
RM57	Chemical Plant	RM42	Ships
		RM45	Storage Tanks
RM39	Fields	RM53	Residential
RM40	Woods		
RM41	Fields	RM39	Fields
RM42	Ships	RM40	Woods
RM45	Storage Tanks	RM41	Fields
RM47	Marsh	RM47	Marsh
RM48	Railroad	RM48	Railroad
RM50	Golf Course	RM50	Golf Course
RM51	Marsh	RM51	Marsh
RM52	Woods and Marsh	RM52	Woods and Marsh
RM53	Residential	RM54	Airplanes
RM54	Airplanes	RM56	Airplanes
		RM49	Water

TABLE 6-2. CLUSTERS GENERATED USING TWO PARAMETERS (MAX-MEAN AND MEAN-MIN).

Cluster	Subset	Content
1	RM41	Fields
	RM49	Water
2	RM39	Fields
	RM40	Woods
	RM44	Railroad
	RM47	Marshland
	RM50	Golf Course
	RM51	Marshland
	RM52	Woods and Marshland
3	RM42	Ships
	RM43	Parking Lot
	RM45	Storage Tanks
	RM46	Storage Tanks
	RM48	Railroad
	RM53	Residential
	RM54	Airplanes
	RM55	Chemical Plant
	RM56	Airplanes
	RM57	Chemical Plant

Section 7

COMPRESSION TO ONE BIT PER PIXEL USING TWO-DIMENSIONAL COSINE TRANSFORM

Seventeen of the nineteen image subsets have been bandwidth-compressed using the two-dimensional discrete cosine transform (2D-DCT) on 16 x 16 element blocks. Figure 7-1 is a processing diagram for each 16 x 16 element block within the image. A non-adaptive process was used, in which each transform block is quantized to an average of 1.0 bit per pixel. Before quantization, the two-dimensional frequency coefficients C_{ij} are multiplied by the corresponding filter value F_{ij}. The purpose of this frequency-plane filter is to standardize the variance of each coefficient to the same variance that the Gaussian quantizer matches. The filter weighting for any given image was generated from the relation:

$$F_{ij} = \frac{\sigma_q}{\sigma_{ij}}$$

where σ_q is the standard deviation that matches the quantizer and σ_{ij} is the standard deviation of the i, jth coefficient. The quantizer used is a MAX quantizer, assuming Gaussian input with zero mean. The number of bits assigned to each of the coefficients is defined in Fig. 7-2.

After reconstructing the images from the quantized coefficients, the normalized mean square error (NMSE) was computed with respect to the original images using the definition:

$$\text{NMSE} = \frac{\sum_{i=1}^{N} \sum_{j=1}^{M} (I'_{ij} - I_{ij})^2}{\sum_{i=1}^{N} \sum_{j=1}^{M} I_{ij}^2} \cdot 100\%$$

The resulting NMSE values are listed in Table 7-1. The resulting images compressed to 1.0 bit/pel along with the corresponding originals are presented in Figs. 7-3 through 7-20.

The NMSE is used as a performance measure for image compression, even though it admittedly has some limitations as an indicator of subjective quality. This NMSE data is later used for training data in a regression analysis to predict the performance of the 2D-DCT compression algorithm from the measured statistics.

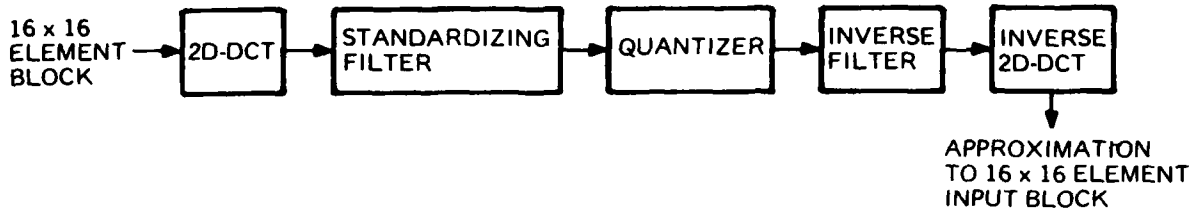


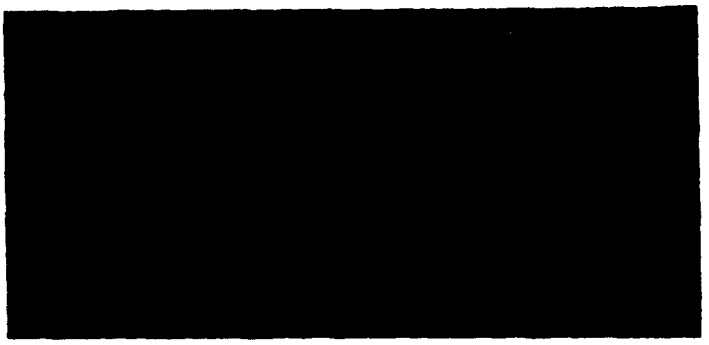
Fig. 7-1. 2D-DCT Processing diagram for 16 x 16 transform block.

8	5	5	5	4	3	3	3	3	2	2	2	2	0	0	0
5	5	5	4	3	3	3	3	2	2	2	2	0	0	0	0
5	5	4	3	3	3	3	2	2	2	2	0	0	0	0	0
5	4	3	3	3	3	2	2	2	2	0	0	0	0	0	0
4	3	3	3	3	2	2	2	2	0	0	0	0	0	0	0
3	3	3	3	2	2	2	2	0	0	0	0	0	0	0	0
3	3	3	2	2	2	2	0	0	0	0	0	0	0	0	0
3	3	2	2	2	2	0	0	0	0	0	0	0	0	0	0
3	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
2	2	2	0	0	0	0	0	0	0	0	0	0	0	0	0
2	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	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	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 7-2. 1.0 bit per pel bit assignment pattern.

TABLE 7-1. NMSE VALUES FOR IMAGE SUBSETS COMPARED TO 1.0 BIT PER PEL

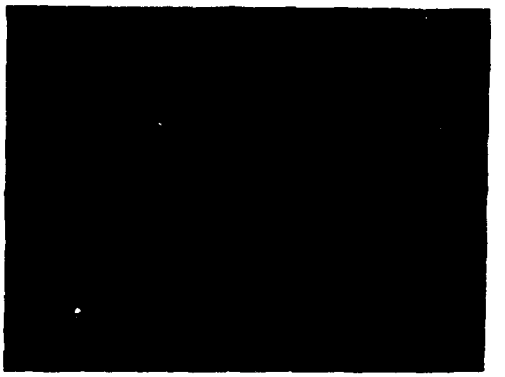
Subset	Percent NMSE
RM39	0.77
RM40	1.42
RM41	1.39
RM42	1.89
RM43	2.08
RM44	1.85
RM45	1.34
RM46	0.902
RM47	1.91
RM48	2.74
RM49	17.69
RM50	1.19
RM51	0.541
RM52	2.176
RM53	0.9
RM54	4.266
RM55	0.328
RM56	8.13
RM57	0.669



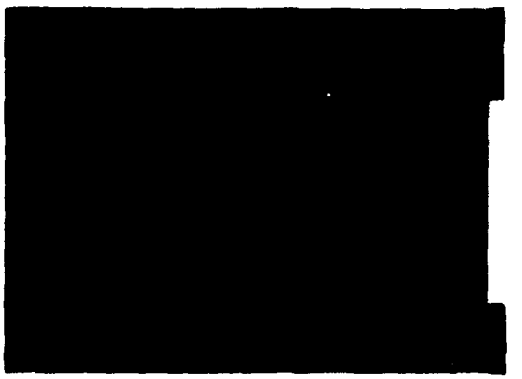
ORIGINAL

1.0 BIT-PEL 2D-DCT

Fig. 7-3. RM40.



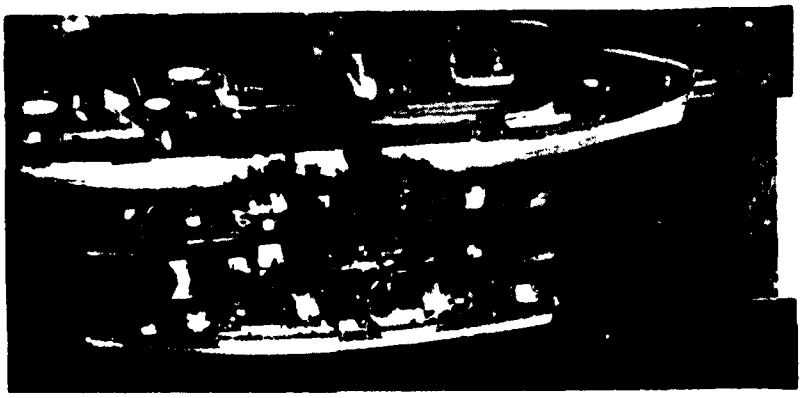
ORIGINAL



1.0 BIT-PEL 2D-DCT

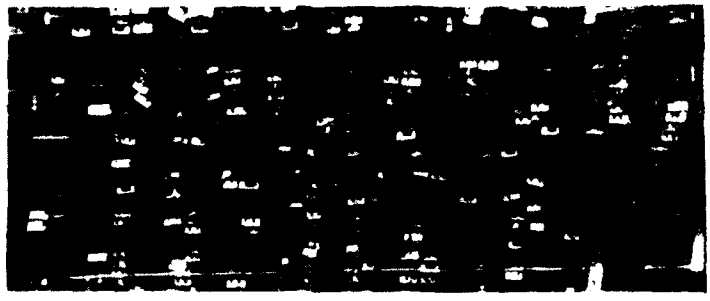
Fig. 7-1. RM41.

ORIGINAL



1.0 BIT-PEL 2D-DCT

Fig. 7-5. RM42.



ORIGINAL

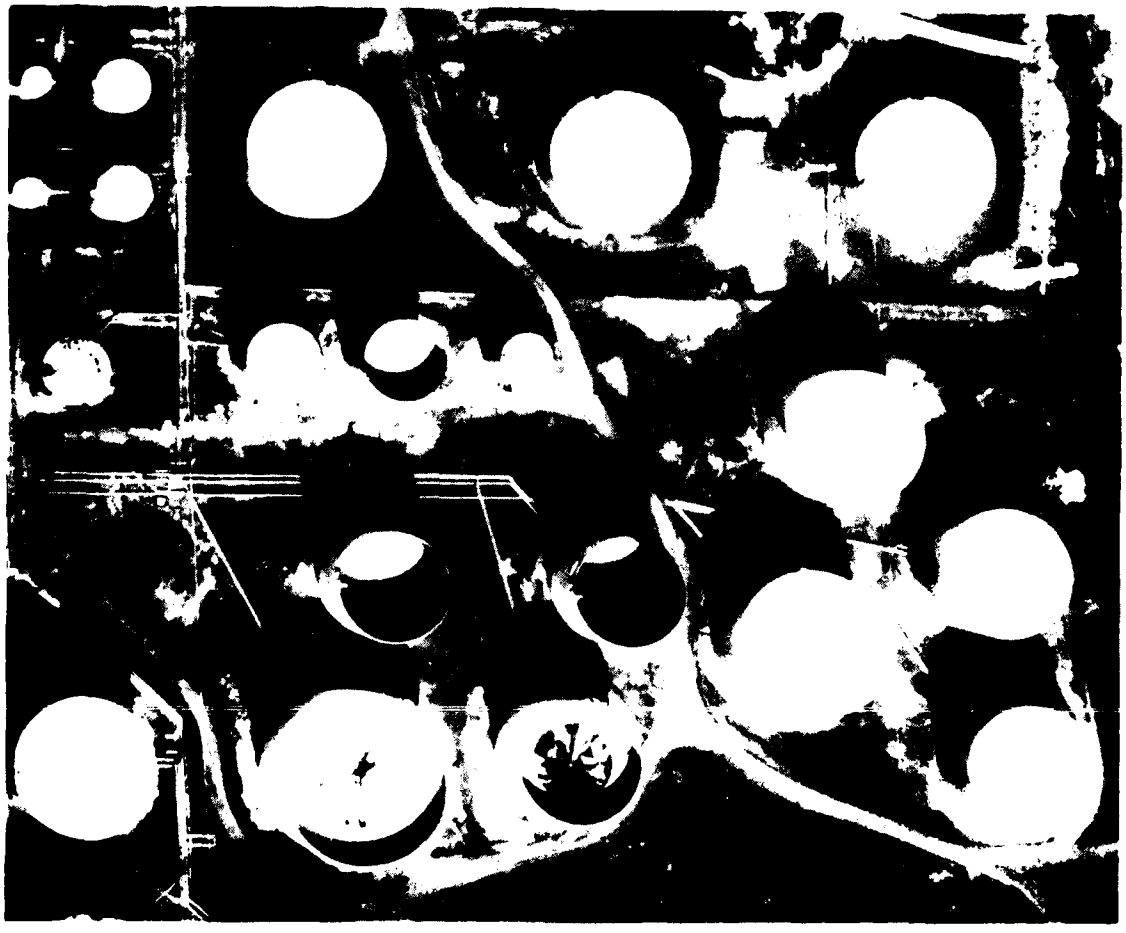
1.0 BIT-PEL 2D-DCT

Fig. 7-6. RM43.

ORIGINAL

1.0 BIT-PEL 2D-DCT

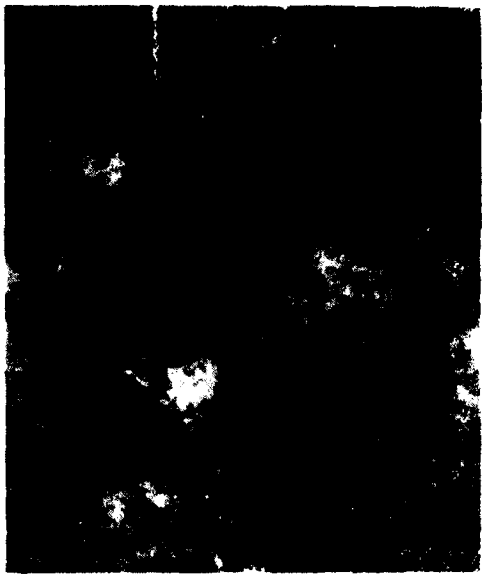
Fig. 7-7. RM44.



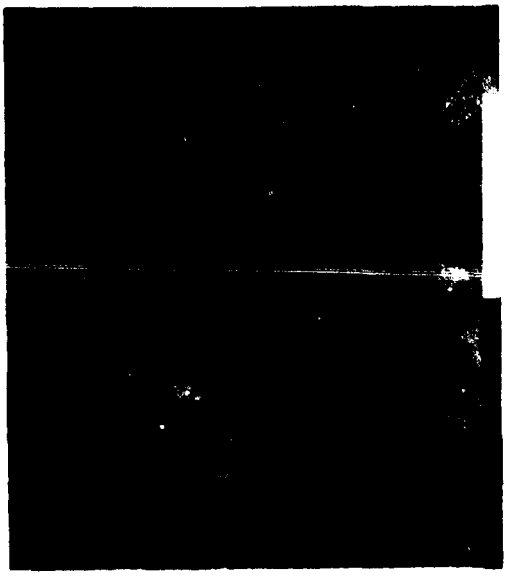
ORIGINAL

Fig. 7-8. RM45 (original).

Fig. 7-9. RM45 (1.0 bit per pel 2D-DCT).



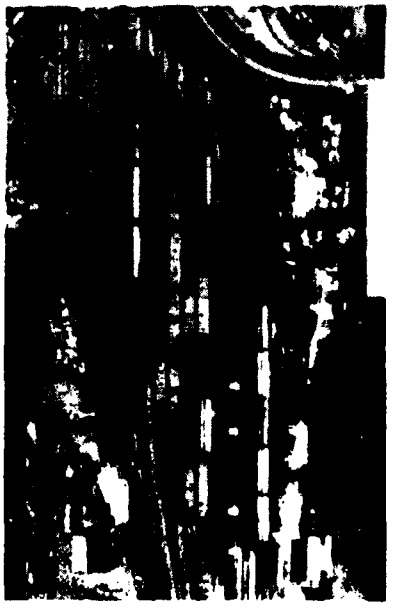
ORIGINAL



1.0 BIT-PEL 2D-DCT

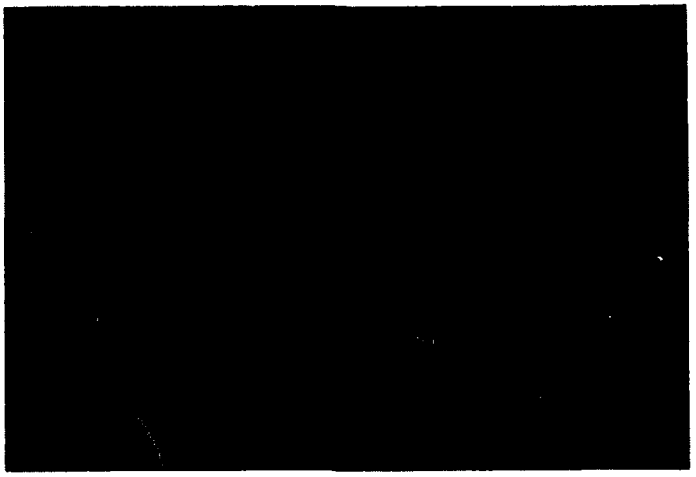
Fig. 7-10. RM17.

ORIGINAL

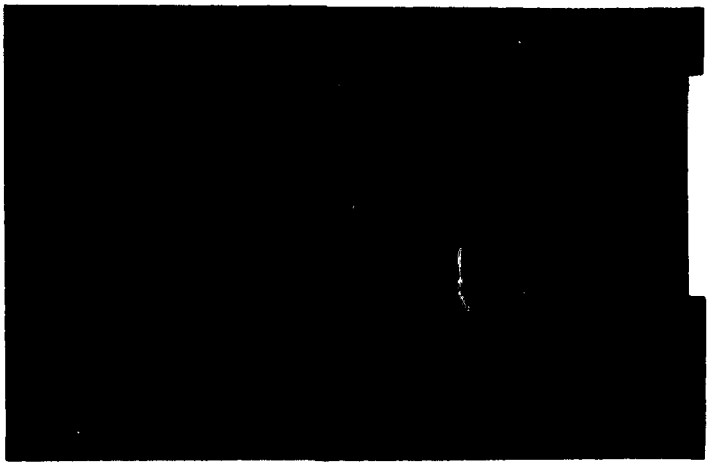


1.0 BIT-PEL 2D-DCT

Fig. 7-11. RM48.



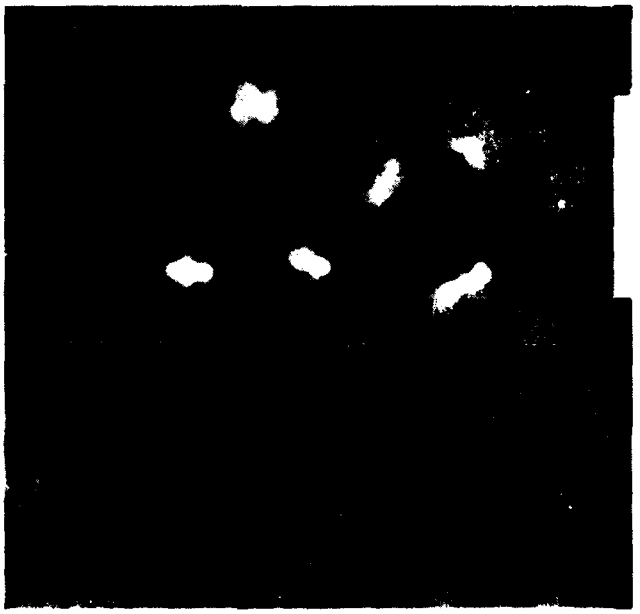
ORIGINAL



1.0 BIT-PEL 2D-DCT

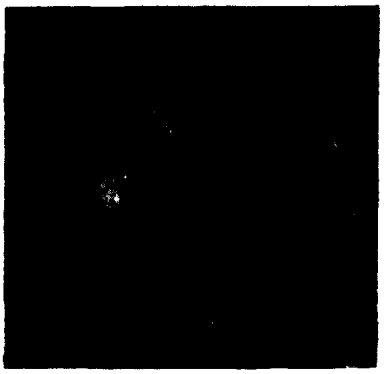
Fig. 7-12. RM49.

ORIGINAL



1.0 BIT-PEL 2D-DCT

Fig. 7-13. RM50.



ORIGINAL

1.0 BIT-PEL 2D-DCT

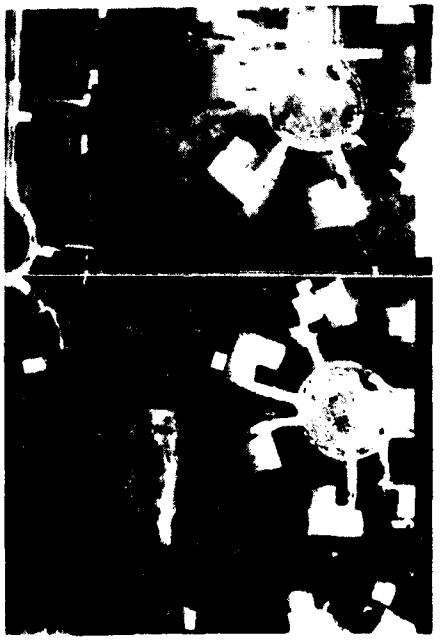
Fig. 7-14. RM51.

ORIGINAL

1.0 BIT-PEL 2D-DCT

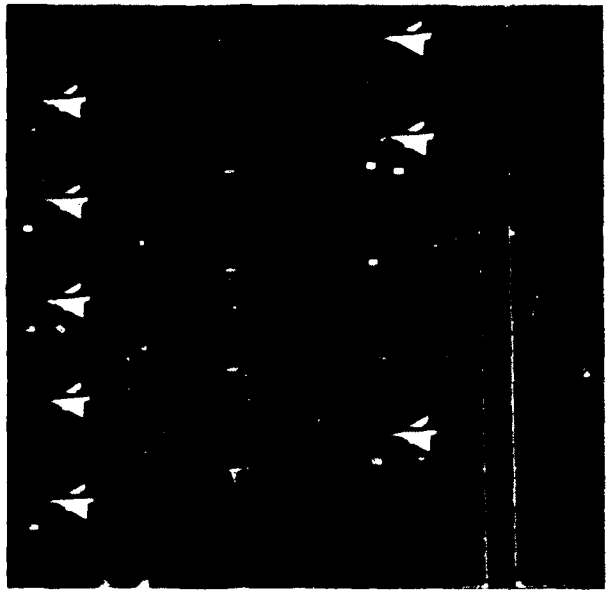
Fig. 7-15. RM52.

ORIGINAL

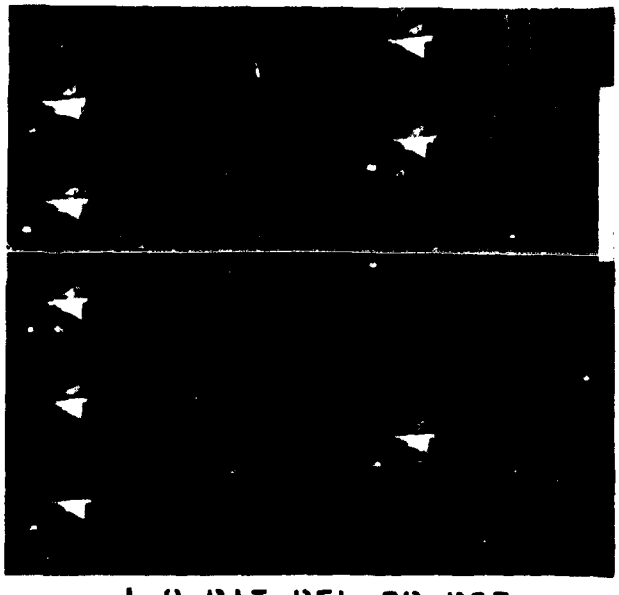


1.0 BIT-PEL 2D-DCT

Fig. 7-16. RM53.



ORIGINAL



1.0 BIT-PEL 2D-DCT

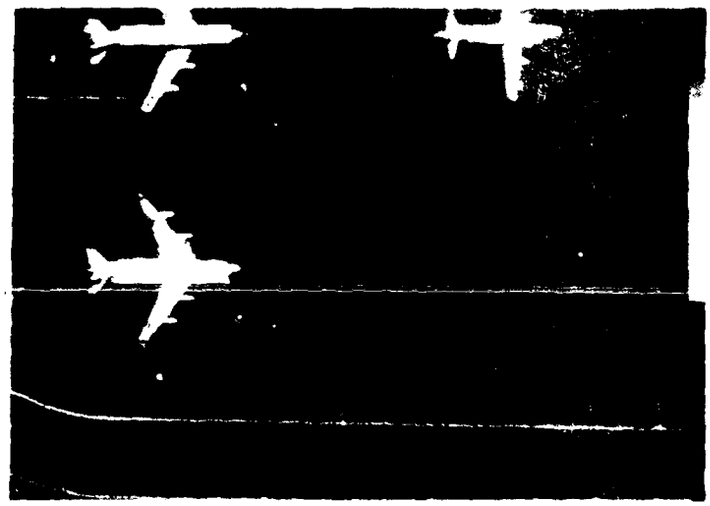
Fig. 7-17. RM54.

ORIGINAL

1.0 BIT-PEL 2D-DCT

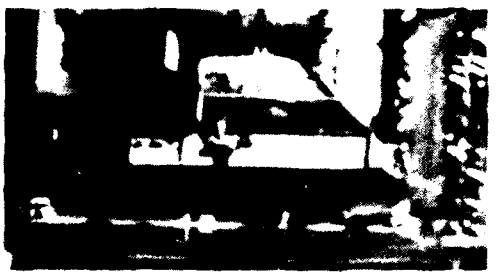
Fig. 7-18. RM55.

ORIGINAL

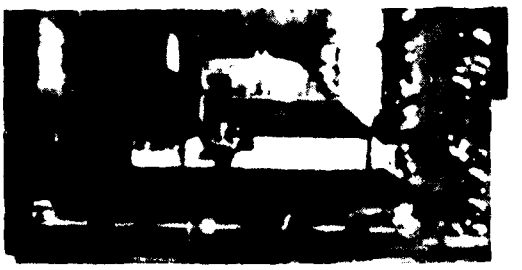


1.0 BIT-PEL 2D-DCT

Fig. 7-19. RM56.



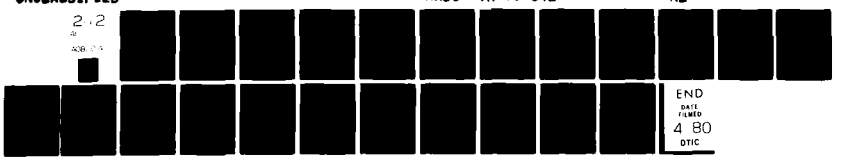
ORIGINAL



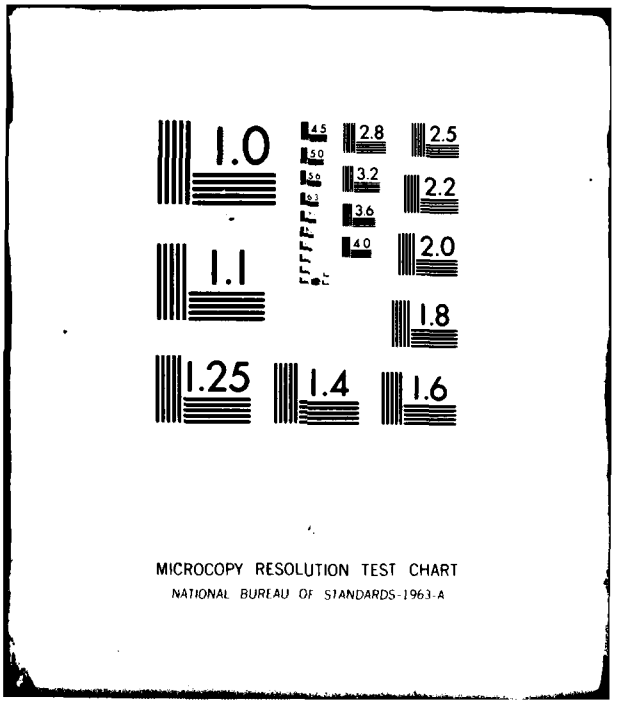
1.0 BIT-PEL 2D-DCT

Fig. 7-20. RM57.

AD-A082 014 RCA ADVANCED TECHNOLOGY LABS CAMDEN N J F/G 14/5
ADVANCED IMAGE COMPRESSION STUDY.(U)
JAN 80 W B SCHAMING, J J RUDNICK F30602-77-C-0244
UNCLASSIFIED RADC -TR-79-342 NL



END
DATE FILMED
4 80
DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Section 8

PREDICTING THE PERFORMANCE OF SINGLE COMPRESSION ALGORITHMS

Regression Analysis[7] can be used to extract the basic characteristics of the relationships hidden or implied in data, and to relate these characteristics by a mathematical equation.

The mathematical equation can be approximated by initially assuming that a linear relationship with unknown parameters exists. A fitted equation is obtained by estimating the unknown parameters for certain assumptions with the help of existing data. Tests can be performed to determine the value of the fitted equation in terms of confidence levels for the parameters and responses. Tests can also be performed to determine whether the underlying assumptions were violated.

If the assumptions are not violated and the required confidence levels are met, then the mathematical equation can be extremely valuable for predicting the value of some of the variables from the knowledge of others.

Suppose a linear regression equation is to be established for a response Y in terms of variables X_1, X_2, ..., X_k (where the Xs are thought to be the complete set of Xs deemed necessary and desirable). Stepwise linear regression[8] is a method to select the "best" regression equation, where best implies choosing as many Xs as possible so that the predictions are reliable, and at the same time choosing as few Xs as possible in order to minimize costs.

The procedure for stepwise linear regression is to insert variables in turn until the regression equation is satisfactory. The order of insertion is determined by using the partial correlation coefficient as a measure of the importance of variables not yet in the equation. Variables previously inserted are re-examined at every stage. A variable that may have been the best single variable to enter at an early stage may become superfluous at a later stage, because of the relationships between the variable and the other variables now in the regression.

---

[7] R. E. Walpole and R. H. Myers, Probability and Statistics for Engineers and Scientists, New York; Macmillan Publishing Co., Inc., 1972, Chapters 8, 9.

[8] N. Draper and H. Smith, Applied Regression Analysis, New York; John Wiley & Sons, Inc., 1966.

At each step of the regression procedure, an analysis of variance is performed that provides the necessary information to determine which variables should be removed from the model and which new variables should be inserted into the regression equation. This process is continued until no more variables can be inserted into the equation and no more can be removed.

8.1 APPLICATIONS OF REGRESSION TECHNIQUES TO IMAGE DATA

This regression analysis technique has been used to predict the performance of a single compression scheme when applied to the available subsets. This approach and its results are described in the following.

The 37 statistics that were computed for each of 19 scenes were used in this analysis. These 37 statistics are listed in Table 8-1. Each of the 19 scenes was compressed, using the 2D-DCT algorithm, and reconstructed; therefore, each scene has an associated MSE computed from the reconstructed image.

Assuming that a linear relationship exists between the MSE for a given reconstructed scene and the 37 statistics associated with that same scene, an attempt was made to characterize this relationship by a mathematical equation using the method of stepwise regression. After generating the regression equation, tests were performed to see if the equation predicted MSE accurately, knowing the values of the statistics in the equation.

Having only 19 scenes available creates certain difficulties, the most important being that concrete conclusions cannot be made from 19 observations.

In addition to this problem of not having enough data to make a conclusion, there is the immediate problem of deriving maximum benefit from the data available (even though it involves only 19 scenes). An approach was used that obtains the maximum amount of information from the available data and at the same time prevents the overlapping of the training and testing sets of data. The approach first eliminates a scene, performs the stepwise regression analysis on the remaining scenes, and then uses the scene that was left out to test the regression equation. Knowing the MSE associated with the scene that was left out, the regression equation can be tested to determine the accuracy of its prediction of MSE. Three selected experiments were performed using this approach of eliminating one scene at a time, finding the regression equation using the remaining scenes, and testing on the left-out scene.

The first experiment utilized all 19 scenes in an attempt to predict the MSE from knowledge of the statistics for a given scene. No further testing, beyond observation of the predicted MSE values, was required to determine that the predictions were unreliable.

TABLE 8-1. THIRTY-SEVEN STATISTICS USED IN REGRESSION ANALYSIS

Statistic No.	Description
1	Variance of brightness samples
2	(MAX-MEAN)/(MEAN-MIN)
3	Average correlation coefficient
4	Value on PSD curve at 0.09 cycle/sample
5	Value on PSD curve at 0.18 cycle/sample
6	Value on PSD curve at .31 cycle/sample
7	Value on PSD curve at .46 cycle/sample
8	Variance of 2 x 2 means
9	Variance of 10 x 10 means
10	Variance of 50 x 50 means
11	Average of MAX-MIN over 2 x 2 areas
12	Average of MAX-MIN over 10 x 10 areas
13	Average of MAX-MIN over 50 x 50 areas
14	Variance of MAX-MIN over 2 x 2 areas
15	Variance of MAX-MIN over 10 x 10 areas
16	Variance of MAX-MIN over 50 x 50 areas
17	Average of gradient image
18	Variance of gradient image
19	Average of MAX/MIN over 2 x 2 areas
20	Average of MAX/MIN over 10 x 10 areas
21	Average of MAX/MIN over 50 x 50 areas
22	Variance of MAX/MIN over 2 x 2 areas
23	Variance of MAX/MIN over 10 x 10 areas
24	Variance of MAX/MIN over 50 x 50 areas
25	Statistic 8/Statistic 1
26	Statistic 9/Statistic 1
27	Statistic 10/Statistic 1
28	Statistic 12/Statistic 11
29	Statistic 13/Statistic 11
30	Statistic 15/Statistic 14
31	Statistic 16/Statistic 14
32	Statistic 20/Statistic 19
33	Statistic 21/Statistic 19
34	Statistic 23/Statistic 22
35	Statistic 24/Statistic 22
36	Skewness
37	Kurtosis

In Fig. 8-1, each column is headed with a scene number (RM39 through RM57), which indicates the scene that was eliminated, while the regression analysis was performed on the remaining scenes to find a regression equation. The parameters for the equation are given in the row for the statistic with which they are associated. The intercept is a constant in the equation which multiplies no statistic. The actual MSE and the predicted MSE are given for comparison. The best case, which is when the predicted MSEs equal the actual MSEs, is illustrated by a straight line with a slope of one when plotted one against the other. The plot of the actual vs. the predicted MSEs (Fig. 8-2) gives additional evidence of the poor predictions, since the points do not form a straight line with slope equal to one.

The second experiment differed from the first in that it eliminated scene RM49 from the set because the MSE associated with this scene has a much greater magnitude than the MSEs of the other scenes. In this experiment, different statistics were retained in the regression equations, but the overall results were not noticeably different (see Fig. 8-3 and 8-4). Eliminating scene RM49 from the regression analysis did not cause the predictions to be more accurate.

The third experiment was performed using the same approach as used for the other two experiments. As in the second experiment, only 18 scenes were used, scene RM49 being eliminated from the analysis. The 18 scenes were being used in an attempt to predict, not the MSE, but the number of coefficients required to produce 0.25% MSE due to truncating the spectrum.

The resulting predictions were remarkably accurate for the size of the data set (refer to Fig. 8-5). A plot of the actual number of coefficients vs. the predicted number of coefficients (see Fig. 8-6) illustrates how near the points fall to the straight line with slope equal to one, which indicates perfect predictions. Reference to Fig. 8-5 shows that the PSD 31 Statistic is the most significant in terms of predicting the number of coefficients that must be retained in order to limit the MSE due to truncation of the spectrum to 0.25%. This is intuitively reasonable, since the variable being predicted is related to the area under the tail of the power spectrum, and hence, somewhat to the shape of the power spectrum.

From observing Case 3, it can be seen that given the correct set of statistics, the stepwise regression method works well to find an equation that predicts some image parameter accurately. The implication to be drawn in the first two cases is not as clear, since equations were not produced that made accurate predictions. The set of image statistics may or may not have been sufficient. Insufficient statistics is certainly a possibility in explaining the poor predictions. Another explanation could be that the relationship between the statistics and MSE needs to be expressed by something other than a linear model. A third explanation mentioned earlier is that Cases 1 and 2 lack a constant parameter in the experiment, while Case 3 is always related to 0.25% MSE.

	Scene / Statistic	Eliminated Scenes																	
		RM39	RM40	RM41	RM42	RM43	RM44	RM45	RM46	RM47	RM48	RM49	RM50	RM51	RM52	RM53	RM54	RM55	RM56	RM57
1	VAR																			0.425
2	X-U/U-MN	0.285	0.402	0.479		0.374	0.251	0.418	0.436		0.485	0.428	0.194	0.389		0.389				
3	PAVE	-18.762	-26.073		-13.489	-20.402	-18.872	-21.217	-13.380				-17.882	-20.886	-13.341		-11.820	-13.453	-17.740	-18.370
4	PSD 0.09																			
5	PSD 0.18																			0.067
6	PSD 0.31					0.074		0.100	0.086					0.0894						
7	PSD 0.46																			
8	VAR(U) 2																			
9	VAR(U) 10																			
10	VAR(U) 50															-0.181				
11	AV(RV) 2	-0.148																		
12	AV(RN) 10																			
13	AV(RN) 50		-0.021																	
14	VR(RN) 2																			
15	VR(RN) 10										-0.001									
16	VR(RN) 50			-0.001								-0.001								
17	AVE(GRD)																			
18	VAR(GRD)		0.009																	
19	AV(RT) 2																			
20	AV(RT)10																			
21	AV(RT) 50																			
22	VR(RT) 2																			
23	VR(RT)10																			
24	VR(RT) 50																			
25	NV(U) 2																			
26	NV(U) 10		4.574																	
27	NV(U) 50					2.404														
28	NA(RN) 10																			
29	NA(RN) 50		0.162																	
30	NV(RN) 10															-0.090				
31	NV(RN 50												0.025							
32	NA(RT) 10																			
33	NA(RT) 50																			
34	NV(RT)10																			
35	NV(RT) 50														0.0					
36	Skewness	0.078	0.033	0.057	0.143	0.090	0.106	0.075	0.072	0.144	0.089	0.075	0.082	0.075	0.145	0.098	0.162	0.142	0.096	0.085
37	Kurtosis			0.008							0.006					0.006				0.003
	Intercept	18.904	20.203	-0.147	13.847	15.828	17.291	15.186	15.557	13.710	0.047	0.091	16.794	16.870	13.857	1.859	12.270	13.885	18.088	6.516
	MSE	0.77	1.42	1.39	1.89	2.08	1.85	1.34	0.902	1.91	2.74	17.89	1.19	0.541	2.176	0.9	4.268	0.339	8.13	0.969
	Predicted MSE	1.629	-0.51	0.29	1.41	2.35	0.859	0.513	0.067	1.31	2.82	5.32	2.41	1.6	1.078	0.828	6.404	0.886	5.65	1.72

NOTE: REGRESSION ANALYSIS UTILIZES ALL 19 SCENES AND MSE.

Figure 8-1. Data for first regression analysis experiment.

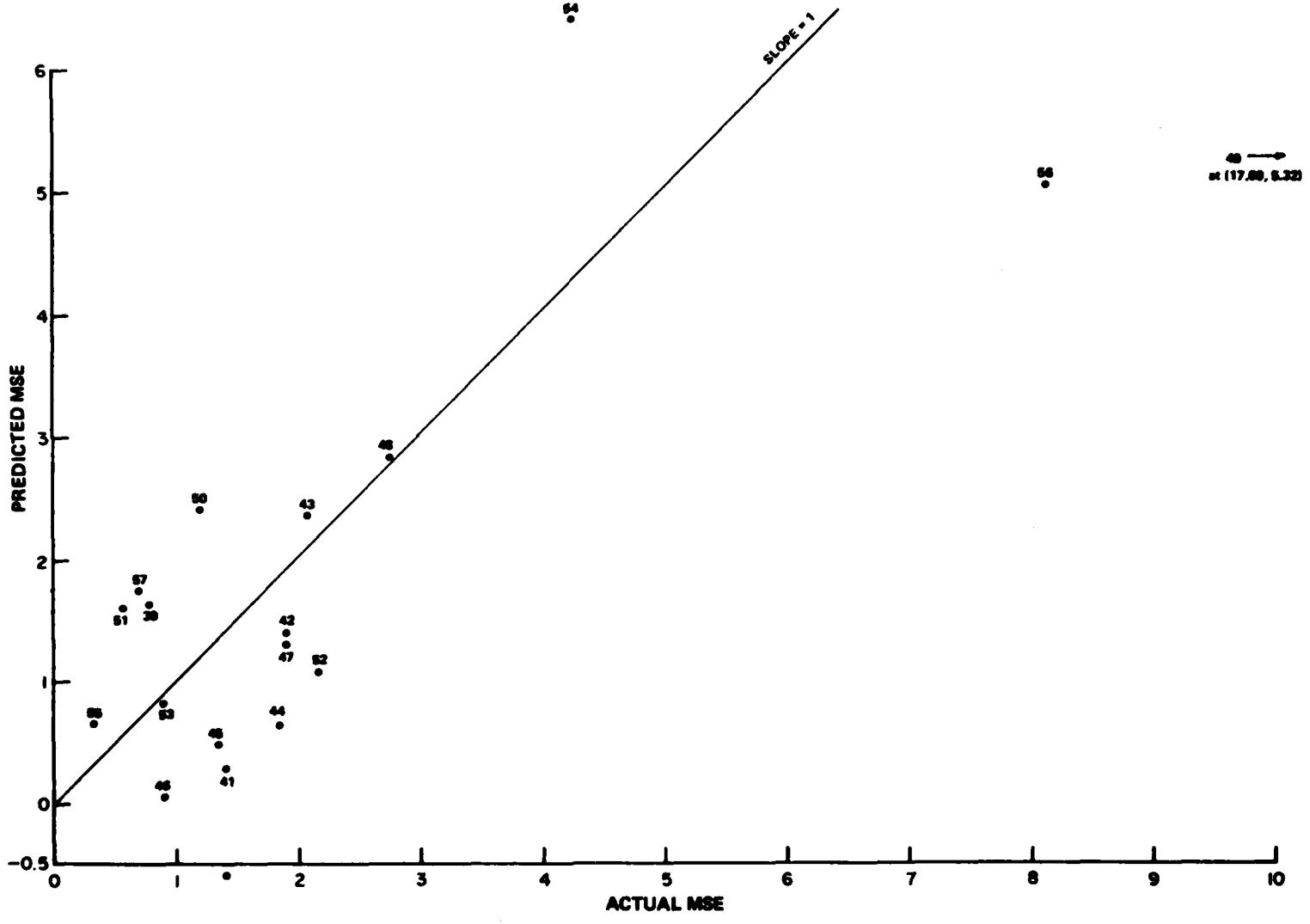


Figure 8-2. Actual vs. predicted MSE, using all 19 scenes.

		Eliminated Scenes																	
	Scene / Statistic	RM39	RM40	RM41	RM42	RM43	RM44	RM45	RM46	RM47	RM48	RM49	RM50	RM51	RM52	RM53	RM54	RM55	RM56	RM57
1	VAR																			
2	X-U/U-MN	0.427	0.585	0.547	0.351	0.382	0.426	0.428	0.315	0.416	0.486		0.404	0.433	0.403	0.448		0.445		0.439
3	ρAVE																-12.583		-12.169	
4	PSD .08																			
5	PSD 0.18																			
6	PSD 0.31																			
7	PSD 0.46																			
8	VAR*U) 2				-0.001															
9	VAR(U) 10								-0.001											
10	VAR(U) 50																			
11	AV(RN) 2																			
12	AV(RN)10					-0.042														
13	AV(RN)50																			
14	VR(RN) 2																			
15	VR(RN)10																			
16	VR(RN)50	-0.001					-0.001	-0.001		-0.001	-0.001		-0.001	-0.001	-0.001	-0.001		-0.001		-0.001
17	AVE(GRD)																			
18	VAR(GRD)																			
19	AV(RT) 2																			
20	AV(RT) 10																			
21	AV(RT) 50																			
22	VR(RT) 2																			
23	VR(RT) 10																			
24	VR(RT) 50																			
25	NV(U) 2																			
26	NV(U) 10																			
27	NV(U) 50																			
28	NA(RN) 10																			
29	NA(RN) 50																	0.000		
30	NV(RN) 10																			
31	NV(RN) 50																			
32	NA(RT) 10																			
33	NA(RT) 50																			
34	NV(RT) 10																			
35	NV(RT) 50																			
36	Skewness	0.075			0.057	0.052	0.075	0.075	0.074	0.077	0.068		0.080	0.074	0.080	0.074	0.103	0.000	0.001	0.021
37	Kurtosis		0.291	0.097																
	Intercept	0.123	17.769	17.327	0.285	0.925	0.090	0.090	0.389	0.082	0.031		0.189	0.157	0.090	0.085	12.920	-0.004	12.769	0.097
	MSE	0.77	1.42	1.39	1.89	2.08	1.85	1.34	0.902	1.91	2.74	17.69	1.19	0.841	2.176	0.9	4.298	0.328	5.13	0.000
	Predicted MSE	0.94	0.28	1.29	2.24	0.237	1.41	0.761	-0.204	1.56	2.82		1.47	1.14	1.44	-0.208	6.46	1.01	4.96	1.16

NOTE: SCENE 49 WAS NOT USED FOR THE REGRESSION ANALYSIS

Figure 8-3. Data for second regression analysis experiment.

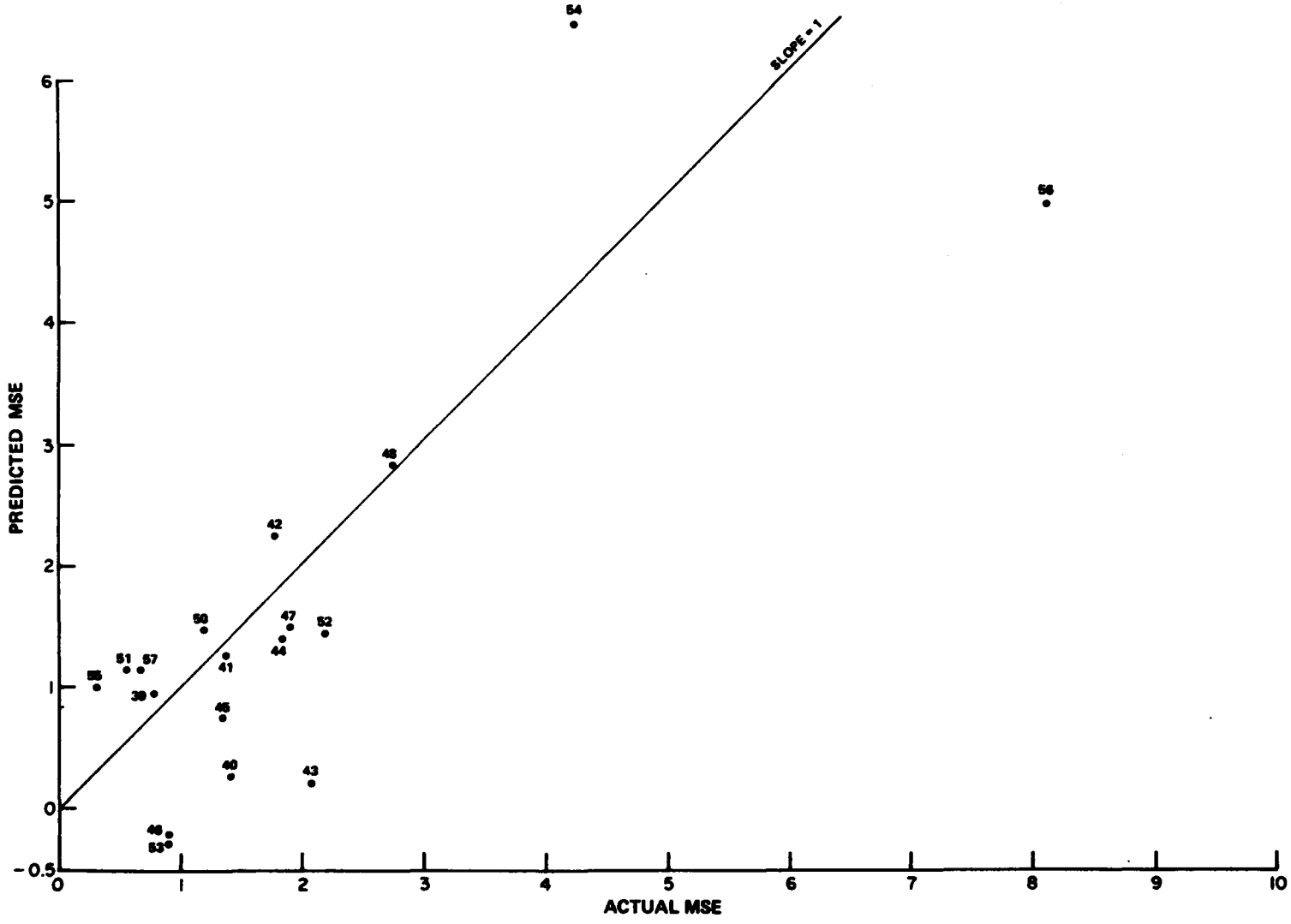


Figure 8-4. Actual vs. predicted MSE, with Scene RM49 eliminated from analysis.

	Scene / Statistic	Eliminated Scenes																		
		RM39	RM40	RM41	RM42	RM43	RM44	RM45	RM46	RM47	RM48	RM49	RM50	RM51	RM52	RM53	RM54	RM55	RM56	RM57
1	VAR																			
2	X-U/U-MN																			
3	P AVE																			
4	PSD 0.9																			
5	PSD 0.18																			
6	PSD 0.31	-6.082	-6.197	-5.891	-5.960	-2.451	-3.419	-6.232	-6.154	-6.085	-6.111		-6.374	-6.101	-5.024	-6.106	-6.157	-3.740	-6.111	-3.668
7	PSD 0.46																			
8	VAR(U) 2																			
9	VAR(U) 10																			
10	VAR(U) 50																			
11	AV(RN) 2																			
12	AV(RN) 10						-0.914											-1.035		-0.767
13	AV(RN) 50																			
14	VR(RN) 2																			
15	VR(RN)10																			
16	VR(RN) 50																			
17	AVE(GRD)																			
18	VAR(GRD)																			
19	AV(RT) 2						307.130								96.548			312.534		275.855
20	AV(RT)10																			
21	AV(RT) 50																			
22	VR(RT) 2					470.774														
23	VR(RT) 10																			
24	VR(RT) 50																			
25	NV(U) 2					-427.238														
26	NV(U) 10																			
27	NV(U) 50																			
28	NA(RN) 10																			
29	NA(RN) 50																			
30	NV(RN) 10																			
31	NV(RN) 50																			
32	NA(RT) 10																			
33	NA(RT) 50																			
34	NV(RT) 10																			
35	NV(RT) 50																			
36	Skewness																			
37	Kurtosis																			
	Intercept	342.295	347.653	334.890	336.197	558.723	-114.488	350.335	346.552	343.220	344.197		355.484	343.680	178.336	344.059	346.721	-103.890	344.499	-70.720
	No.Coeff.	38	33	5	103	130	23	96	76	45	120		14	55	91	66	78	28	60	23
	Predicted No. Coeff.	43.56	23.86	22.55	85.88	322.7	36.89	105.9	83.65	50.23	120.6		-1.65	54.68	47.52	68.3	85.6	18.21	66.14	34.18

NOTE: REGRESSION ANALYSIS EXCLUDED SCENE RM49; NO' OF COEFFICIENTS REPLACED MSE.

Figure 8-5. Data for third regression analysis experiment.

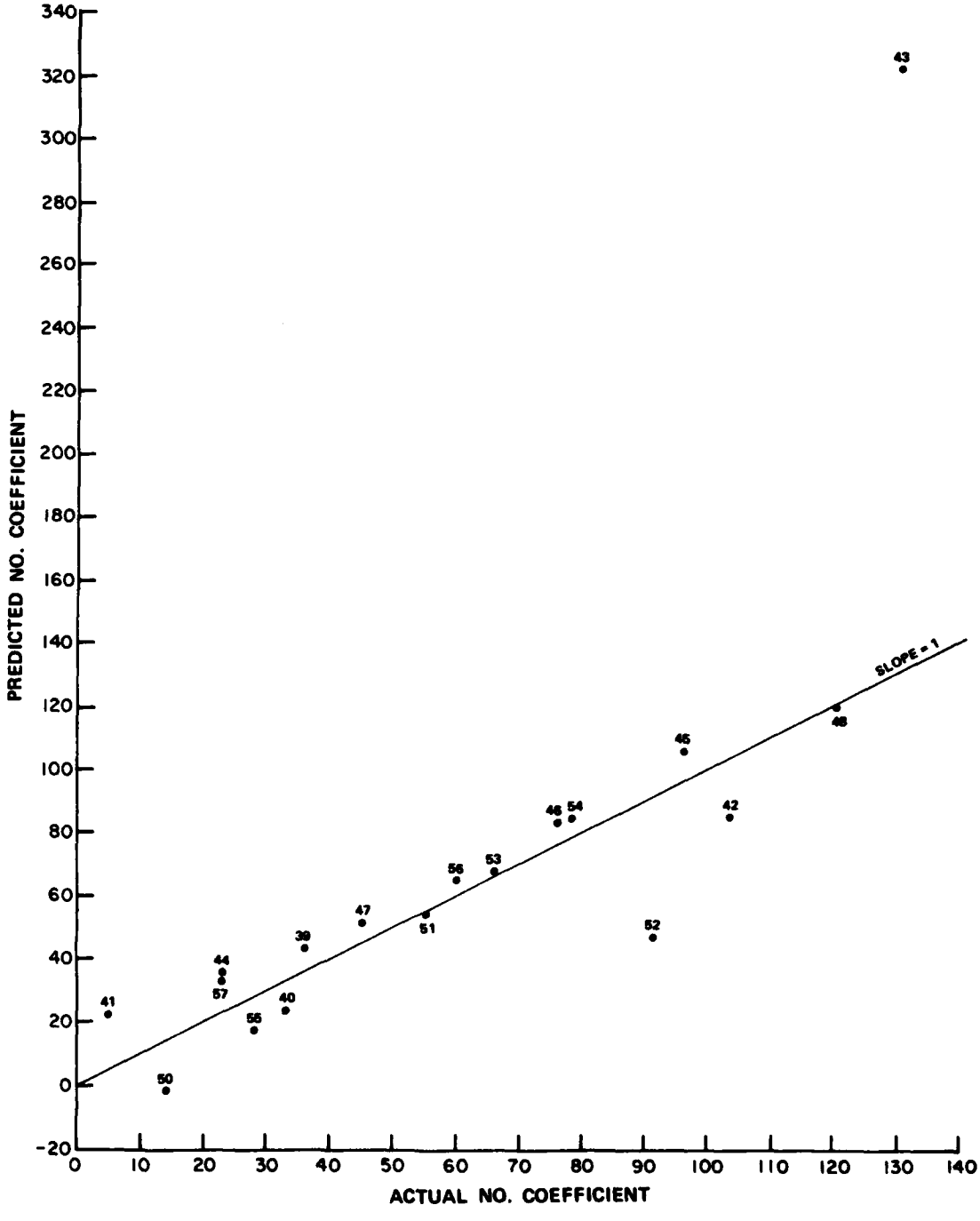


Figure 8-6. Actual vs. predicted number of coefficients required to limit MSE due to truncating spectrum to 0.25%.

Section 9

SUGGESTION FOR FUTURE INVESTIGATION

As a result of the work performed in this study, three particular areas appear to be candidates for additional investigation. These areas are: (1) The performance prediction of single algorithms, (2) The prediction of best algorithm for a single image subset, and (3) The generation of an adaptive compression algorithm driven by statistical measurements. Each of these areas is discussed briefly in the following.

9.1 PERFORMANCE PREDICTION FOR SINGLE ALGORITHMS

Unfortunately it was not until late in this current effort that performance predictions of a single algorithm were started. As is obvious from the previous discussions, this effort was not completed. The results seem to imply (although not necessarily) that the right set of features has not yet been found for this process. Of the 37 measurements used in the regression analysis, only a few appear to be significant for this task. At the same time, the prediction performance is marginal at best. As in any pattern recognition problem, the selection of features is the most difficult and the most significant part of the problem. One entire set of features unused in this effort is that of textural properties. Most textural features are based on statistics of co-occurance matrices that describe how often one gray tone will appear in a specific spatial relationship to another gray tone. As evidenced in recent publications from the Image Understanding Workshop[9], as well as in the open literature, the importance of textural measures is becoming more widely accepted.

Another deficiency in the current attempt to predict performance is the limited number of images used. For a statistical evaluation of this sort, 200 images would be more appropriate than 20. Obtaining, digitizing, and computing all the features on such a large data set is a major undertaking.

9.2 SELECTION OF ALGORITHM BY PERFORMANCE PREDICTION

Once all of the data is gathered and some assurance is obtained that a reasonable set of features is available for analysis, a regression equation could be generated for as many different algorithms as desired. This, however, requires that all the images to be used in the analysis be compressed with each algorithm, and a performance measure computed on each reconstructed output image at the same data rate. An alternative approach would be to locate the minimum bit rate required for each algorithm to produce a reconstructed image with a constant quality. This could be done readily if the

[9]Proceedings of Image Understanding Workshop, November 1978 and April 1979, Science Applications, Inc., Report Numbers SAI-79-814-WA and SAI-80-895-WA.

quality measure is the MSE due to truncating the spectrum. However, there is no guarantee that this MSE correlates with any subjective evaluation by human observers. Locating the minimum bit rate for a constant quality can be a major effort if subjective evaluation and ranking are the quality criteria. Assuming that this data could be obtained, a regression equation could be generated to predict the required bits per pixel necessary to produce a given quality for each algorithm. For an image that is not part of the training set, the bit rate for each algorithm would be predicted. The algorithm for which a minimum bit rate is predicted would be the selected algorithm for that image.

9.3 STATISTICAL DRIVEN ADAPTIVE COMPRESSION ALGORITHM

Many adaptive compression schemes utilize the ac energy as the parameter that controls the adaptivity. It may be profitable to investigate the use of a set of features other than energy to control an adaptive algorithm. For example, if a 2D-DCT algorithm with 16 x 16 picture element blocks is used, a set of features would be computed separately for each 16 x 16 area of the image. These features would be used to select an appropriate data rate for each transform block and to control the way in which bits are distributed within the transform block. One set of features would be used in a regression equation to select an appropriate data rate for each transform block. A second set of features could be used in a separate process to dictate the way in which the bits are distributed within the transform block. The features used for the two functions, (i.e., setting the bit rate and controlling the assignment of bits) are clearly different since the optimization of the bit assignment requires directionality in the features.

The design of such a system should be straightforward even though a sizable effort is required to select the proper features and generate the regression equations. Care would have to be taken to ensure that the features selected for use in such a process were within reason in terms of computation requirements, otherwise a hardware implementation would be impractical.

Appendix A

MTF COMPUTATION FROM EDGE DATA

In order to determine the MTF of the input transparencies, available edges were located on the film. Edges at random orientation in the film were scanned on a Joyce-Loebl microdensitometer to obtain the edge response function E (η). If the edge response function is normalized to unit height and U (η) is the unit step function at the origin, the complex modulation transfer function can be computed from:

$$\text{MTF}(x) = 1 + x \int_{-\infty}^{\infty} [E(\eta) - U(\eta)] \sin \eta x \, d\eta$$

$$- xj \int_{-\infty}^{\infty} [E(\eta) - U(\eta)] \cos \eta x \, d\eta$$

This expression comes from the relations between the complex MTF, the line spread function $S(\eta)$, and the edge response $E(\eta)$ given below:

$$S(\eta) = \frac{d}{d\eta} E(\eta)$$

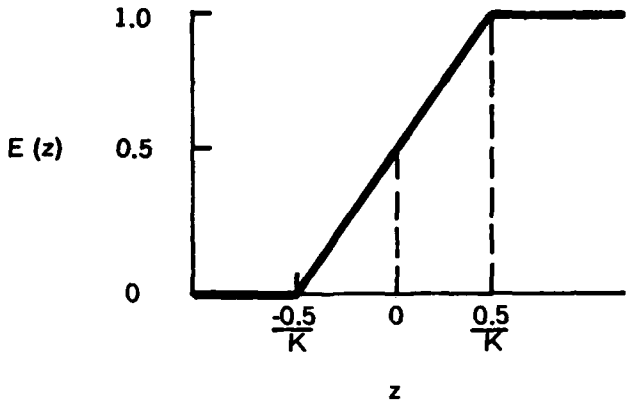
$$\text{MTF}(x) = \int_{-\infty}^{\infty} S(\eta) \, e^{j\eta x} \, d\eta$$

One question that arises in computing these MTF curves is whether the sampling interval used on the microdensitometer is influencing the results. To ensure the validity of the results, a perfect edge with slope K as shown in Fig. A-1 was examined to determine the sample spacing required to compute the correct MTF for the edge. For the edge in Fig. A-1, which is symmetrical about the midpoint, the MTF is given by:

$$\text{MTF}(x) = \frac{2K}{x} \sin \frac{x}{2K}$$

The zero crossings of the MTF occur when

$$\sin \frac{x}{2K} = 0 \quad \text{for } x > 0$$



$E(z) = 0 \qquad z < \frac{-0.5}{K}$

$E(z) = Kz + 0.5 \qquad \frac{-0.5}{K} < z < \frac{0.5}{K}$

$E(z) = 1 \qquad \frac{0.5}{K} < z$

Figure A-1 Perfect edge with slope equal to K.

The first zero crossing occurs when

$$x = 2\pi K$$

or

$$f = K \text{ cycles/mm}$$

Figure A-2 shows the MTF computed for a perfect edge with slope K = 50, using a sampling frequency of 100 cycles/mm, which is twice the frequency at the first zero crossing of the MTF. As can be seen, a sampling frequency equal to 2k cycles/mm is inadequate to compute the MTF. Figure A-3 shows that the same MTF computed with the edge sampled at 4k cycles/mm is very close to the expected result.

The slopes of the edges that were scanned to produce the MTF data in Fig. 3-1 through 3-6 were estimated and are shown in Table A-1. The steepest edge that was used had an approximate slope of 27 and was sampled with a spacing equal to 0.005 mm. The corresponding MTF is shown in Fig. A-4. As an additional check on the process, this MTF was computed using a sample spacing equal to 1.0 mm (half the sampling frequency). The results in Figure A-5 are essentially the same, which implies that the interval 0.005 mm is more than adequate.

A similar check was made for one of the steepest edges for which we used a sampling interval of 0.01. From these data we can conclude that the MTF computations were not significantly affected by the sampling interval used.

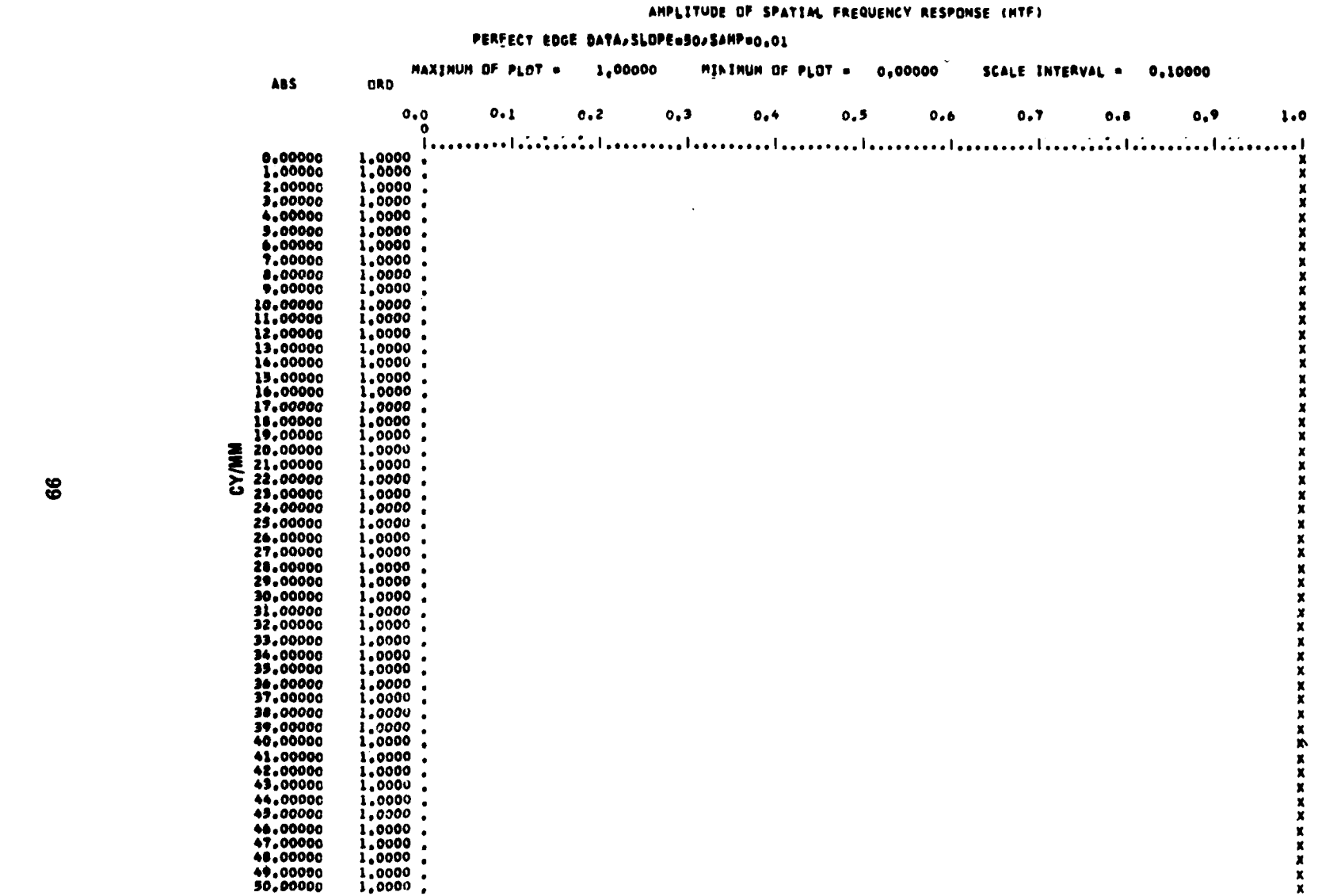


Figure A-2. MTF for perfect edge; slope K=50, sampling frequency = 100 cycles/mm.

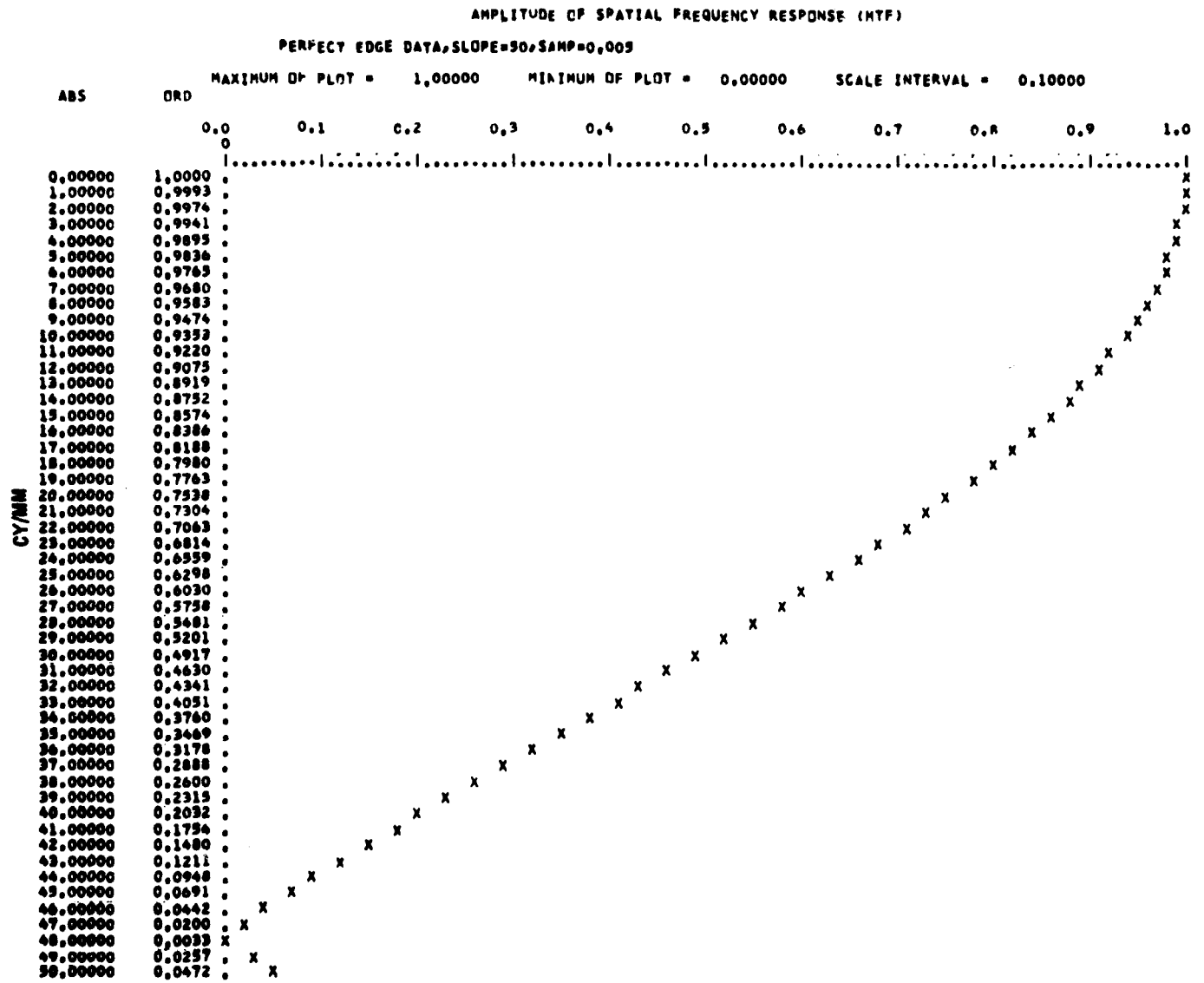


Figure A-3. MTF for perfect edge; slope K = 50, sampling frequency = 4K cycles/mm.

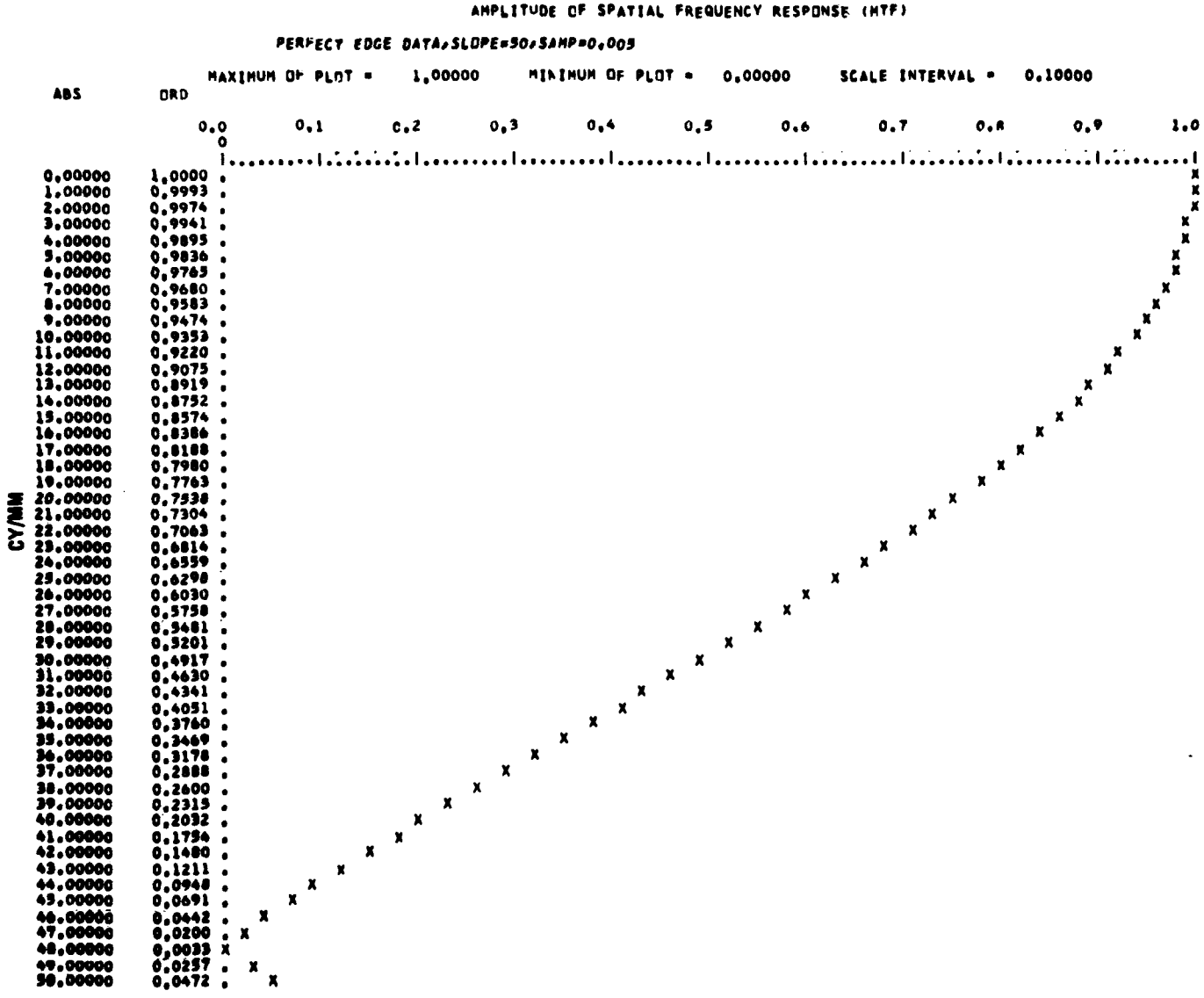


Figure A-3. MTF for perfect edge; slope K = 50, sampling frequency = 4K cycles/mm.

TABLE A-1. ESTIMATED SLOPE OF SCANNED EDGES, WITH SAMPLING INTERVAL USED AND THE MAXIMUM SAMPLING INTERVAL DEFINED BY f_s = 4 x SLOPE (cycles/mm).

Edge Identification	Estimated Slope	Sampling Interval Used	Max Sampling Interval = 1/(4 x slope)
4109	15.58	.01	.016
4120	17.09	.01	.0146
4131 #1	24.39	.005	.0102
4131 #2	13.57	.01	.0184
4134 #1	27.08	.005	.0092
4134 #2	15.89	.01	.0157
4137 #1	16.50	.01	.0151
4137 #2	17.16	.01	.0145
4134 #4 Top	10.51	.01	.0237
4134 #4 Middle	16.89	.01	.0148
4134 #4 Bottom	17.22	.01	.0145
4134 #4 Average	14.77	.01	.0169

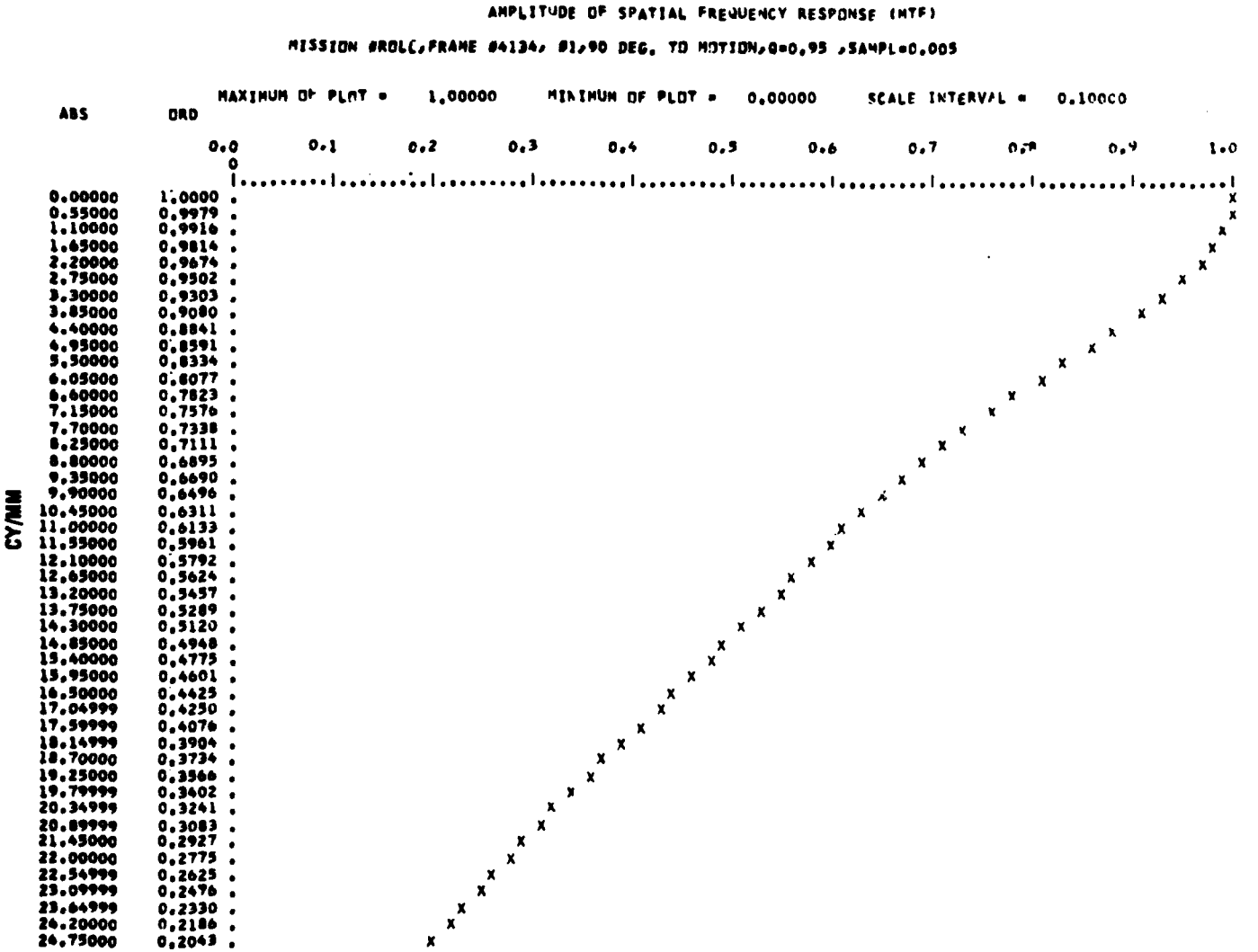


Figure A-4. MTF for edge with approximate slope of 27; sampling spacing = 0.005 mm.

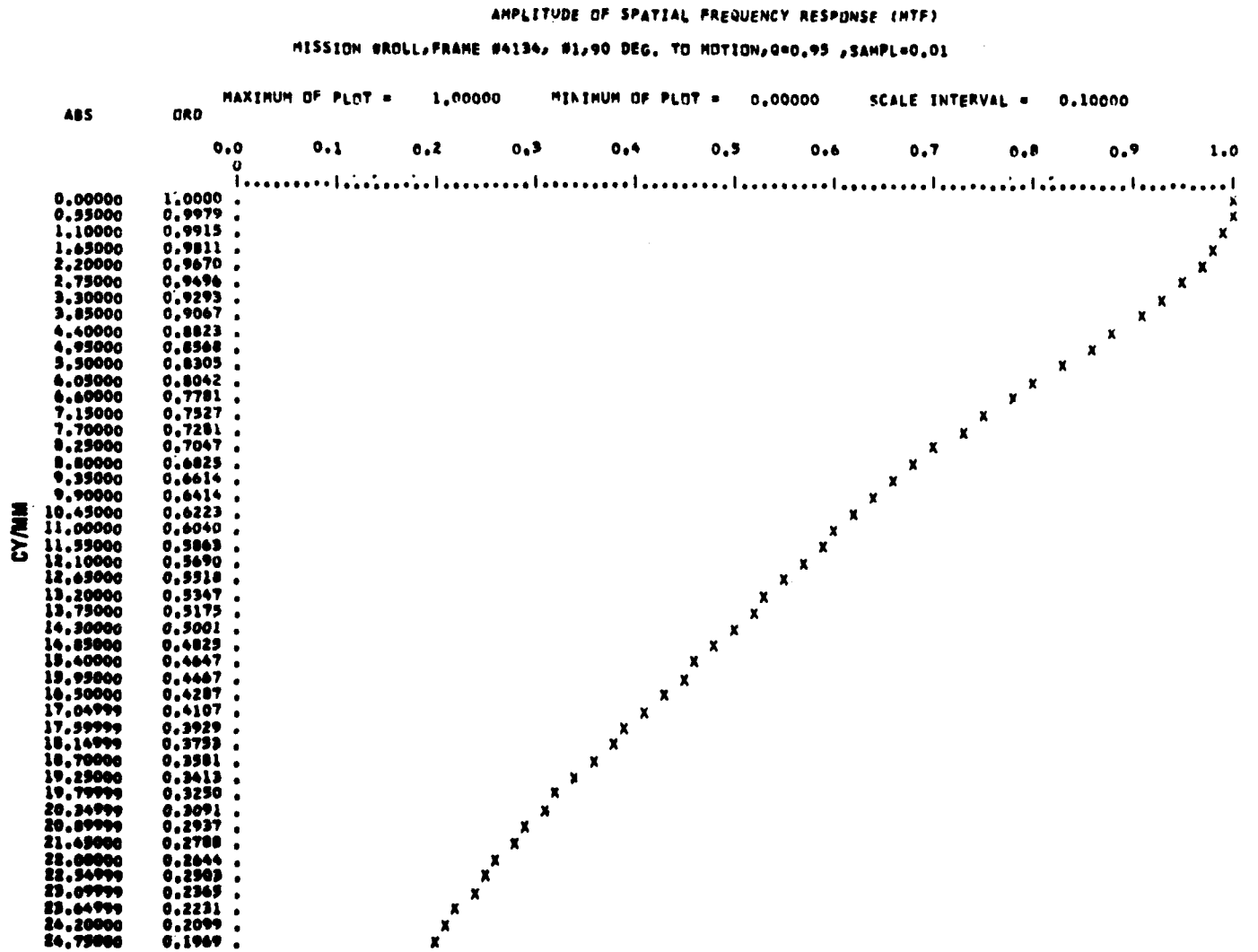


Figure A-5. MTF for edge with approximate slope of 27; sampling spacing = 1.0 mm.

Appendix B

CALLIER Q CORRECTION FOR DATA PROPORTIONAL TO DENSITY OR TRANSMITTANCE

The film density measured by a densitometer is diffuse density. The scanner output is related to specular density, since the optics will not collect all the light passing through the film due to scattering. Thus the necessary modification to the data is a Callier Q correction process, whose purpose is to remove the effects of measurement differences between diffuse density and specular density. The Callier Q of an equipment is determined by scanning a known step wedge with the device. This results in a plot of diffuse density of the input material vs. the specular density measurement at the output of the scanner. The slope of the line is the Q. This output measurement may be digital counts from an A/D converter or it may be relative numbers on a known linear scale as might be obtained from a scanning microdensitometer. It is important to know whether the measuring device is producing outputs proportional to density or to transmittance, since the correction is different for the two cases.

In order to correct the data so that the Q is equal to unity (i.e. the specular density equals diffuse density), the specular density data is divided by Q.

$$D_{diffuse} = D_{spec}/Q$$

The equivalent correction for an equipment output proportional to transmittance is obtained as follows. Since transmittance is related to density by

$$D = 10 \log \frac{1}{T}$$

$$T = \frac{1}{10^{D/10}}$$

The corrected transmittance T_D is given by:

$$T_D = \frac{1}{10^{\frac{D_S/Q}{10}}} = \left[\frac{1}{10^{\frac{D_S}{10}}} \right]^{1/Q} = (T_S)^{\frac{1}{Q}}$$

Consequently, the Callier Q correction for output data proportional to density is obtained by dividing by Q, i.e.:

$$\text{Corrected data} = \frac{\text{output data}}{Q}$$

The correction for output data proportional to transmittance is given by:

$$\text{Corrected data} = (\text{output data})^{1/Q}$$

Renormalization after Q correction must be such that the uncorrected and corrected data are identical for a density of zero.

REFERENCES

1. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, New York, John Wiley & Sons, 1973, pp. 234,235.

2. R.O. Duda and P.E. Hart, op. cit.

3. J.B. Coleman, "Image Segmentation by Clustering", Univ. of Southern California, Image Processing Institute, Report USCIPI 750, p. 31, July 1977.

4. K. Fukunaga, Introduction to Statistical Pattern Recognition, New York, Academic Press, 1972, pp. 324-326.

5. J.B. Coleman, op. cit.

6. R.O. Duda and P.E. Hart, op. cit.

7. R.E. Walpole and R.H. Myers, Probability and Statistics for Engineers and Scientists, New York, Macmillan Publishing Co., Inc., 1972, Chapters 8, 9.

8. N. Draper and H. Smith, Applied Regression Analysis, New York, John Wiley & Sons, Inc., 1966.

9. Proceedings of Image Understanding Workshop, November 1978 and April 1979, Science Applications Report Numbers SAI-79-814-WA and SAI-80-895-WA.