Posted

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF RESEARCH ADMINISTRATION

RESEARCH PROJECT INITIATION

Date: May 26, 1975

Project Title: Image Encoding Subject to a Fidelity Measure

Project No: E-21-659

Principal Investigator: Dr. Barry M. Leiner

Sponsor: National Science Foundation

Agreement Period: From June 1, 1975 Until November 30, 1977
24 month budget period plus 6 months for submission of required reports

Type Agreement: Grant No. ENG75-04992

Amount: $25,000 NSF
19,324 GIT (E-21-335)
$44,324 Total

Reports Required: Annual Letter Technical, Final Report

Sponsor Contact Person (s):
Administrative Matters
thru ORA
Mr. Gaylord L. Ellis
Grants Officer
National Science Foundation
Washington, D. C. 20550
(202) 632-5965

Assigned to: Electrical Engineering

COPIES TO:

Principal Investigator
School Director
Dean of the College
Director, Research Administration
Director, Financial Affairs (2)
Security-Reports-Property Office
Patent Coordinator
Library
Rich Electronic Computer Center
Photographic Laboratory
Project File
Other

RA-3 (6-71)

GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION

Posted

Date: 4/19/78

Project Title: Image Encoding Subject to a Fidelity Measure

Project No: E-21-659

Co-Project Director: Dr. T.P. Barnwell and Dr. R.M. Mersereau

Sponsor: National Science Foundation

Effective Termination Date: 11/30/77

Clearance of Accounting Charges: 11/30/77

Grant/Contract Closeout Actions Remaining:

- _ Final Invoice and Closing Documents
- X Final Fiscal Report
- _ Final Report of Inventions
- _ Govt. Property Inventory & Related Certificate
- _ Classified Material Certificate
- _ Other _____

Assigned to: Electrical Engineering (School/Laboratory)

COPIES TO:

Project Director
Division Chief (EES)
School/Laboratory Director
Dean/Director—EES
Accounting Office
Procurement Office
✓Security Coordinator (OCA)
Reports Coordinator (OCA)

Library, Technical Reports Section
Office of Computing Services
Director, Physical Plant
EES Information Office
Project File (OCA)
Project Code (GTRI)
Other _____

CA-4 (3/76)

N4001.544
E-21-659
Annual Tech Ltr

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF ELECTRICAL ENGINEERING
ATLANTA, GEORGIA 30332

TELEPHONE: (404) 894-2901

June 4, 1976

National Science Foundation
Washington, D. C. 20550

Attn: Dr. Elias Schutzman
Program Director
Electrical and Optical Communications Program
Electrical Sciences and Analysis Section
Division of Engineering

Subject: Change of Principal Investigator on Research Grant ENG75-04992, "Image Encoding Subject To A Fidelity Measure"

Dear Dr. Schutzman:

This letter is written to inform you of the status of research grant ENG75-04992, "Image Encoding Subject To A Fidelity Measure," and to formally request that the principal investigator on the grant be changed from Dr. Barry M. Leiner of Georgia Tech to Drs. Thomas P. Barnwell III and Russell M. Mersereau as co-principal investigators. This request is made since Dr. Leiner has resigned from Georgia Tech to accept a position in industry. Drs. Barnwell and Mersereau are both currently Assistant Professors in the School of Electrical Engineering at the Georgia Institute of Technology, and both have research interests in this area. This grant was initiated in May 1975, and it is scheduled to terminate on November 30, 1977.

A. Progress of the Research

The original proposal for this grant outlined three basic tasks. The first task was to find a fidelity measure for coded images which will allow one number to be associated with each image. This single number is to describe the quality of the image and thereby the quality of the coding scheme by which it was encoded. This fidelity measure must be readily computable from the coded image and must correlate well with human perception. The second task proposed was a theoretical analysis of data compression with this fidelity measure using the methods of rate distortion theory. This analysis was to include a determination of the properties of the distortion-rate function and its relation to other easily calculable distortion-rate functions. The third task was to calculate the effectiveness of various existing image compression techniques. It was expected that this combination of tasks would produce better image compression techniques which have lower data rates than existing techniques, adequate fidelity, and straightforward implementations.

In line with these tasks, a candidate for a fidelity measure has been proposed. This technique takes the supremum over all the components of a vector distortion measure, the components being the weighted mean-squared error in various spatial frequency channels. This measure was chosen on the basis of a model of the human visual system, so there is reason to believe that it should correlate well with human perception. The problem of optimal data compression with this fidelity measure has been extensively analysed and the results have been submitted for publication by Dr. Leiner. Copies of these publications are attached in Appendix B. Thus, the second task outlined above has been completed. Work remains to be done on the remaining two tasks. The fidelity measure must still be correlated with human performance and values for its several parameters must be found to maximize that correlation. Further, if it correlates fairly well, other coding schemes still need to be evaluated. To aid in this regard, considerable software has been developed by Dr. Leiner and a graduate student working with him.

B. The Proposed Principal Investigators

It is proposed that the remainder of this work be completed by Drs. T. P. Barnwell III and R. M. Mersereau of the School of Electrical Engineering. Bio-sketches of these men are included in Appendix A. Dr. Barnwell's recent interests include the efficient coding of speech and the correlation of objective speech quality measures with human evaluations. In this regard an elaborate quality testing facility, consisting of both hardware and software, has been developed. Much of this system can be utilized in the evaluation of images. He has also been involved recently in the problem of displaying visual outputs from an interactive computer. He is currently receiving some NSF support under grant ENG76-02029, "Engineering Research In Very Low Bit Rate Speech Compression Techniques" and also some support from the Defense Communications Agency through the Georgia Tech Post-Doctoral Program for Research on Speech Coding. Dr. Mersereau's primary research interest is in multidimensional digital signal processing and its applications, and the reconstruction of multidimensional signals from their projections. He is to receive support from the Army Research Office for some of this work. The backgrounds and interests of both of these men are well suited to undertake this research on image coding.

C. Facilities

Up to this time, images have been processed using the campus central computer facility. The coded images have then been sent to a commercial firm to be transferred to film and then subjective tests have been conducted using standard 35 mm film projectors. This process has proven to be unsatisfactory in two respects. First, the turn-around time is too long. This discourages any attempt to interactively determine coding parameters, for example. A more serious problem, however, is that degradations in image quality are introduced by the film processing. In some cases those degradations are comparable to the coding errors themselves.

In an effort to overcome these difficulties, we propose to move the remainder of the research to our own computer facility described in Appendix C. This is an interactive facility centered around a NOVA 830 computer. Independent of this proposal we have acquired through Institute funds a COMTAL video display system which will be connected to the NOVA. This display is capable of displaying a 512x512 b/w image with 6 bits per picture element of gray scale resolution. This, we feel, is a better facility for the image evaluations which must be performed. The programs which have been written for the central computer can be easily run on the interactive facility. Furthermore, connected to this computer are facilities for human quality evaluation, which we hope to utilize.

Included in the budget in the next section is a request to divert most of the money currently in a budget for time for the central computer to the purchase of additional memory for the COMTAL to increase the gray scale resolution to 8 bits.

D. Budgets

Below is shown the original budget for the two-year duration of this grant, along with the amounts currently spent.

Item	Approved NSF Support	Approved Georgia Tech Support	NSF Funds Spent	Georgia Tech Funds Spent
1. Personal Services	11475	10775	5418	4950
2. Staff Benefits: Retirement @ 8.936% of Salary for Barry L. Leiner	629	945	269	442
3. Permanent Equipment		600		515
4. Expendable Equip. and Supplies	837		608	
5. Travel	800		435	
6. Publication Costs	800		0	
7. Computer Costs	3000		166	
8. Indirect Costs 68% of S & W	7459	7004	3684	3366
TOTALS	25000	19324	10580	9273

Below is shown a proposed budget for the remaining funds for the remainder of the contract.

Item	NSF Funds	GIT Funds
1. Salaries & Wages		
a. Thomas P. Barnwell III 2 man-months	1031	3052
b. Russell M. Mersereau 2 man-months	886	2624
c. Graduate Student Asst.* 1/3 time, 1 year	4000	
2. Staff Benefits Retirement @ 9.1%	174	517
3. Permanent Equipment**	2700	
4. Expendable Equip. and Supplies	340	
5. Travel	365	
6. Publication Costs	800	
7. Computer Costs**	100	
8. Indirect Costs	4024	3858
TOTALS	14420	10051

Notes: * Subject to our ability to find a capable, interested graduate student. The student who has been working on this project has since graduated.

** This represents our desire to transfer most of the original computer budget to the purchase of extra memory for our COMTAL image display.

E. Schedule of Remainder of Work

Due to earlier commitments for their time, neither of the principal investigators will devote extensive effort to this work during this summer (1976.) Current plans call for two to three man-months of effort by each of the principal investigators during the academic year 1976 - 1977. No problem is forseen in meeting the scheduled termination date.

Sincerely yours,

Barry M. Leiner
Assistant Prof. of E.E.

Russell M. Mersereau
Assistant Prof. of E.E.
(404) 894-2917

Demetrius T. Paris
Prof. and Director
School of Electrical Eng.
(404) 894-2902

Thomas P. Barnwell, III
Assistant Prof. of E.E.
(404) 894-2914

Milton W. Bennett
Office of Contract Administration
(404) 894-4815

Addressee: Two copies

Enclosures: Two copies each
Appendix A - Biographical Sketches
Thomas P. Barnwell, III
Russell M. Mersereau
Appendix B - Publications
Appendix C - Facility Description

NATIONAL SCIENCE FOUNDATION
Washington, D.C. 20550

SUMMARY OF COMPLETED PROJECT

Form Approved
OMB No. 9220013

Please read instructions on reverse carefully before completing this form.

1. INSTITUTION AND ADDRESS	2. NSF PROGRAM	3. PRINCIPAL INVESTIGATOR(S)
Georgia Institute of Technology School of Electrical Engineering Atlanta, GA 30332	Electrical & Optical Communications Engineering Division	T. P. Barnwell, III B. M. Leiner R. M. Mersereau

4. AWARD NUMBER	5. DURATION (MOS)	6. AWARD PERIOD	7. AWARDEE ACCOUNT NUMBER
ENG75-04992	18	from 6-1-75 to 11-30-77	E21-659

8. PROJECT TITLE

Image Encoding Subject to a Fidelity Measure

9. SUMMARY (ATTACH LIST OF PUBLICATIONS TO FORM)

This research concerned finding a numerical quality (or distortion) measure for computer-coded photographic imagery. The goal was to have an objective measure which could be calculated from knowledge of the original and coded images which would provide an estimate of the fidelity of the coded image close to the value a panel of human viewers would give to the same image. Such an objective measure is important for evaluating potential image coding algorithms without the need for expensive subjective tests and for the design of improved coding methods.

Among the major results of this research was a mathematical means for comparing candidate distortion measures and a subjective test for image quality which possessed exceedingly good resolving capability. Using these results a number of new and existing proposed measures for image quality were compared. Two of the measures proposed in the course of this research proved to possess the highest correlation with the subjective results. One of these, in fact, proved to be better than a human viewer at estimating the responses of other human viewers.

One final result was the development of some mathematical bounds for the distortion-rate function associated with one of our fidelity measures. These bounds tell how close any coding method with a fixed number of bits in the digital representation can come to the original image with respect to the distortion measure.

9. SIGNATURE OF PRINCIPAL INVESTIGATOR/ PROJECT DIRECTOR	TYPED OR PRINTED NAME	DATE
	T. P. Barnwell, III	4-13-78

IMAGE ENCODING SUBJECT TO A FIDELITY MEASURE

by

T. P. Barnwell, III
B. M. Leiner
R. M. Mersereau

School of Electrical Engineering
GEORGIA INSTITUTE OF TECHNOLOGY
Atlanta, Georgia 30332

FINAL REPORT E21-659
Grant ENG75-04992
from
NATIONAL SCIENCE FOUNDATION

March 1978

TABLE OF CONTENTS

		PAGE
LIST OF ILLUSTRATIONS		ii
LIST OF TABLES		iii
I.	INTRODUCTION	1
	A. Goals of this Research B. History of this Research C. Major Results of this Research D. Publication Resulting from this Effort	
II.	TESTING OBJECTIVE QUALITY MEASURES USING THE RESULTS OF SUBJECTIVE TESTS	6
III.	THE SUBJECTIVE IMAGE EVALUATION TEST	8
IV.	OBJECTIVE QUALITY MEASURES	19
	A. Fidelity Measures for Images B. Implementation of the Objective Quality Measures	
V.	RESULTS OF THE CORRELATION STUDIES	31
VI.	DISTORTION-RATE FUNCTIONS FOR VECTOR SOURCES AND A MAXIMUM FIDELITY CRITERION	39
VII.	DISCUSSION	46
	A. Open Research Questions	
ACKNOWLEDGEMENT		50
REFERENCES		51
APPENDIX A		53
APPENDIX B		64
APPENDIX C		71
APPENDIX D		77
APPENDIX E		86
APPENDIX F		93

LIST OF ILLUSTRATIONS

FIGURE		PAGE
1	The two original images which were used for this study	9
2	The bandpass filters used to bandlimit the Gaussian noise to the test images to produce distortion classes V, VI, and VII	12
3	A test slide from the subjective test	13
4	A model for the visual detection system	21
5	A block diagram of the computation steps to produce the error energies used for the objective quality measures	23
6	The frequency response of one of the filters in the filter bank shown in Figure 5	25
7	The amplitude weighting functions A(f) for the Mannos and Sakrison quality measure	33

LIST OF TABLES

TABLE		PAGE
1	The distortions	10
2	Results of the subjective quality test for additive white gaussian noise	17
3	Some subjective correlations using the Gray-Leiner metric with differing numbers of filters, B	36
4	Some subjective correlations using optimal linear metric with differing numbers of filters, B	38

I. INTRODUCTION

A. Goals of this Research

This research was motivated by a desire to identify a fidelity measure for coded images which can be computed objectively and which will reflect the evaluations of human subjects. There are many benefits to be derived from having such a measure. First it would allow the subjectively meaningful comparison of algorithms for image coding without the need for extensive subjective tests which are expensive, time-consuming, and generally tedious for the subjects involved. Subjective tests can also be easily biased if they are not carefully administered and it is difficult to relate results from one subjective experiment to another. Objective measures, on the other hand, offer the promise of being cheaper, faster, and free of the human failings of subject and administrator. Furthermore, if a common objective measure is used, the numerical results of one study could readily be compared to those of another. The difficulty of the task is due to the fact that the manner by which humans evaluate distortion in images (or other structured media, such as speech) is not well understood and, as a result, early attempts at finding a fidelity measure have resulted in criteria which do not correlate well with human experience.

As the research began, however, it became evident that we first had to address the question: How can different objective quality measures be compared? This can be restated as the problem of finding a quality measure for the quality measures. Seeking an answer to this problem occupied a major fraction of the total research effort, but having a means to meaningfully compare proposed quality measures was invaluable. We regard the

development of means for comparing quality measures to be a major innovation of this research.

The next research goal was to compare several of the existing quality measures and to see how well each correlated with human observations. Among the measures compared was one proposed in this research which we have called the Gray-Leiner metric.

Another task which was mentioned in the original research proposal, which we did not pursue but which would make a logical follow-on study, is a comparison of existing image coding techniques based on the most promising of the distortion measures.

One peripheral task which was accomplished early in the study concerned the development of bounds for the distortion-rate function associated with the Gray-Leiner distortion metric. Results from rate-distortion theory can be used to determine the minimum number of bits required to reproduce a signal within a given level of fidelity measured by an objective fidelity measure. This is referred to as the rate-distortion function and its inverse is the distortion-rate function. These functions can be used to design a "best" coding strategy consistent with a quantitative fidelity measure.

In the remainder of this section we summarize some of the history of this research project, summarize our major results, and list the publications which have resulted from this work. In the following section, we discuss the problem of correlating subjective and objective quality measures. The subjective test which was used is described in Section III, and objective measures for image evaluation are described in Section IV. Also in this section, the Gray-Leiner metric is motivated. Results of comparison

between the subjective and objective measures are presented in Section V. Section VI summarizes work performed on determining a distortion-rate function for the Gray-Leiner measures and bounds are presented to aid in its evaluation. The last section presents a final discussion of the results and conclusions reached from the whole research effort.

B. History of this Research Effort

The grant for which his document is the final report was proposed by Dr. Barry M. Leiner and was awarded in May 1975. The selection of the research problem, the Gray-Leiner fidelity measure, and the derivation of the distortion-rate functions is solely his. He began processing images for the subjective experiments on the Georgia Tech central computing facility with the actual photographic reproductions performed by an outside firm. This proved to be unsatisfactory because the delays which occurred resulted in reduced flexibility. A better solution proved to be the use of an interactive computer facility where distorted images could be viewed immediately and the resulting degree of distortion adjusted to design good subjective experiments. At the end of the first year of what was to be a two year study, Dr. Leiner left Georgia Tech to accept a position in industry. The research program was then assumed by Drs. Barnwell and Mersereau, who first rewrote the subjective experiments to operate on the interactive Digital Signal Processing Laboratory minicomputer which contained a COMTAL image display. At this time, the research emphasis shifted. The means for evaluating quality measures, the subjective experiments, and the analysis of the data was the responsibility of the latter researchers.

C. Major Results of this Research

Although the remainder of this report discusses our research results

in detail it is perhaps useful to summarize the highlights of this work here.

(1) A means was developed for evaluating fidelity criteria for images. This measure was the correlation coefficient between the results of a subjective test and the prediction of subjective quality made on the basis of objective measurements.

(2) A subjective test was developed which gives statistically significant measures of image quality even for distorted images whose perceived fidelities are cose to one another.

(3) A fidelity criterion for images, the Gray-Leiner metric was proposed, evaluated, and compared with other criteria. It is based on a model for the human visual system. The correlation coefficient between the subjective evaluations and the predictions made by this metric using our data set is significantly better than that obtained through the use of mean-squared error and comparable to the quality measure proposed by Mannos and Sakrison [1]. This correlation can be made still higher if some of the parameters of the metric are optimized.

(4) A distortion-rate function for a vector source with a maximum fidelity criterion was developed and bounds were derived for its use. The Gray-Leiner metric represents a criterion from this class.

D. Publications Resulting from this Effort

To date, there have been three publications which have resulted from this effort. These are attached as Appendices A, B, and C. One of them appeared in a referred journal and the other two were prepared to accompany conference presentations. One more referred journal paper is anticipated. It will consist of the material in Sections II-V of this report.

II. TESTING OBJECTIVE QUALITY MEASURES USING THE RESULTS OF SUBJECTIVE TESTS

One of the major goals of this research was to develop techniques for efficiently testing the effectiveness of objective quality measures by using the results of subjective tests. This procedure involved three distinct steps. First, a subjective quality test needed to be designed which would possess adequate resolving power over an ensemble of distorted images. Second, a collection of candidate objective measures had to be developed which could estimate image fidelity over the same set of distorted images. Finally, the correlation between the subjective and objective estimates of image fidelity had to be determined.

Define a sample space, Ω, consisting of all distorted images to be evaluated. Denote the members of this set by $\{\omega_i\}$. The subjective estimates of image quality can then be interpreted as a random variable $S(\omega_i)$ defined over this sample space. Similarly, for the kth objective (computed) distortion measure, we can define another random variable $C_k(\omega_i)$ over the same sample space. The correlation coefficient ρ_k between the random variables S and C_k then provides a measure of the accuracy by which the values C_k can be used to predict S. If $\rho_k = \pm 1$, then the subjective results can be predicted exactly from the objective estimates. In this case, the kth objective measure would be equivalent to the subjective test. If $\rho_k = 0$, then the estimates of the kth objective measure are uncorrelated with the subjective rankings. In this case, the objective measure is worthless. For a correlation coefficient ρ_k, the minimum least squares estimate of S_i from C_{ki} is given by

$$\hat{S}_i = \frac{\hat{\sigma}_S \hat{\rho}_k}{\hat{\sigma}_{C_k}} (C_{ki} - \overline{C}_k) + \overline{S}$$

If ρ_k is negative, then the scope of the estimate is negative, and $\hat{S}_i$ goes down as C_{ki} get up. This would clearly be the case when C_k is a distortion measure and S is a quality measure. Clearly, the closer $|\rho_k|$ is to 1, the more valuable the objective estimator of quality.

If I images are subjectively evaluated, then the minimum variance estimate of ρ_k, which we will denote by $\hat{\rho}_k$ is given by

$$\hat{\rho}_k = \frac{\sum_{i=1}^{I} (S_i - \overline{S})(C_{ki} - \overline{C}_k)}{[\sum_{i=1}^{I} (S_i - \overline{S})^2]^{\frac{1}{2}} [\sum_{i=1}^{I} (C_{ki} - \overline{C}_k)^2]^{\frac{1}{2}}} . \tag{1}$$

Here $\overline{S}$ and $\overline{C}_k$ denote the means of S and C_k, respectively and S_i and C_{ki} denote the subjective and objective estimates of quality for the ith image under test respectively.

III. THE SUBJECTIVE IMAGE EVALUATION TEST

The subjective test used for this study consisted of a number of sessions in which groups of subjects were shown slides of distorted images and were then asked to numerically rate the image quality. The images used were distorted versions of the two images shown in Figure I. These were prepared as 256 x 256 discrete images and were displayed with up to 256 intensity levels. The two images GIRL and RADOME were chosen to have a relatively low and high information content respectively.

A data base of distorted images was produced by applying 120 distortions to each of the two images. The distortions were divided into eight classes or types of distortion each of which was then represented by fifteen levels of distortion. Within each class the distortion levels were chosen to range from "barely perceivable" to "heavily distorted."

The distortions used in this study are summarized in Table 1. For the first and second classes of distortions, additive white noise was added to the images. The noise was uniformly distributed for class 1 and gaussian for class 2. The next two classes of distortions represented multiplicative noise. For the distortions in class three the distorted picture elements were obtained by

$$u'(m,n) = u(m,n)[1 + C_k N(m,n)] \tag{2}$$

where $u(m,n)$ denotes the undistorted picture element, $N(m,n)$ denotes a sample of uniformly distributed white noise whose amplitude varies between

Figure 1. The two original images which were used for this study. Each contains $(256)^2$ picture elements and the display is capable of accomodating up to 256 intensity levels.

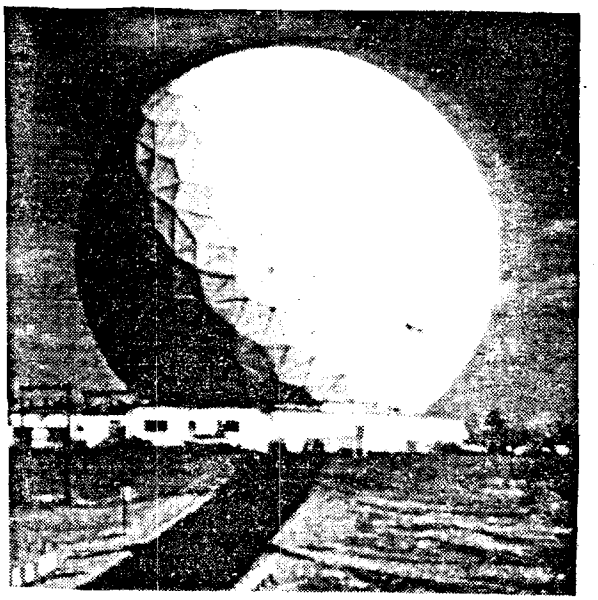


TABLE 1

THE DISTORTIONS

	DISTORTION	C_k RANGE	
1	$U(m,n) + N(m,n) \cdot C_k$	8 - 60.5	ADDITIVE UNIFORM NOISE
2	$U(m,n) + G(m,n) \cdot C_k$	4 - 33.75	ADDITIVE GAUSSIAN NOISE
3	$U(m,n) \cdot [1 + C_k N(m,n)]$	.061 - 1.77	MULTIPLICATIVE UNIFORM NOISE
4	$EXP[\ell n U(m,n) + U(m,n) C_k]$	.04 - 1.16	MULTIPLICATIVE GAUSSIAN NOISE
5	$U(m,n) + BPF_A[G(m,n)] \cdot C_k$	24.8 -209	BAND LIMITED GAUSSIAN NOISE
6	$U(m,n) + BPF_B[G(m,n)] \cdot C_k$	72.7 - 613	BAND LIMITED GAUSSIAN NOISE
7	$U(m,n) + BPF_c[G(m,n)] \cdot C_k$	7.7 - 65.7	BAND LIMITED GAUSSIAN NOISE
8	$LPF_k[U(m,n)]$	21 - 7 (cycles/degree)	BAND PASS FILTER

$N(m,n)$ = WHITE, UNIFORM NOISE

$G(m,n)$ = WHITE GAUSSIAN NOISE

BPF = BAND PASS FILTER

C_k = CONSTANT FOR k^{th} DISTORTION

$0 \leq U(i,j) \leq 255$

LPF = LOW PASS FILTER

-1 and +1 and C_k is a constant which controls the level of distortion. For the images in class four, white gaussian noise was added to the logarithm of the image and the resulting array was then exponentiated. For distortion classes five, six, and seven colored gaussian noise was added to the images. The noise was colored by being passed through one of the bandpass filters shown in Figure 2. For the eighth distortion class, a low pass filtering blur was realized by passing the test images through a two-dimensional circularly symmetric low pass filter. The cutoff frequency of the filter varied with the degree of the distortion. A complete table of distortion parameters is given in Appendix D. No attempt was made to simulate actual coding distortions.

In all cases, the distortions were implemented digitally using a NOVA 830 minicomputer and the resulting distorted images were stored on digital magnetic tape.

The subjective test used in this study was a doubly-anchored isometric quality preference test. For each distortion, a black and white slide was produced which contained three images arranged as shown in Figure 3. In the upper left was a "high anchor" which was the original image. In the upper right was a "low anchor" which was distorted by a number of distortions and which had been prejudged to be worse than, but roughly comparable to, the worst distortion in the test. At the bottom of the slide is the image under test. For all of the distorted versions of the image GIRL the same high and low anchors were used, similarly for the image RADOME. The low anchor for RADOME was produced from the same sequence of distortions as the low anchor for the image GIRL. The slides were photographed from the screen of a CRT controlled by a COMTAL Image Processing System with 512 x 512

Figure 2. The bandpass filters used to bandlimit the Gaussian noise added to the test images to produce distortion classes V, VI, and VII.

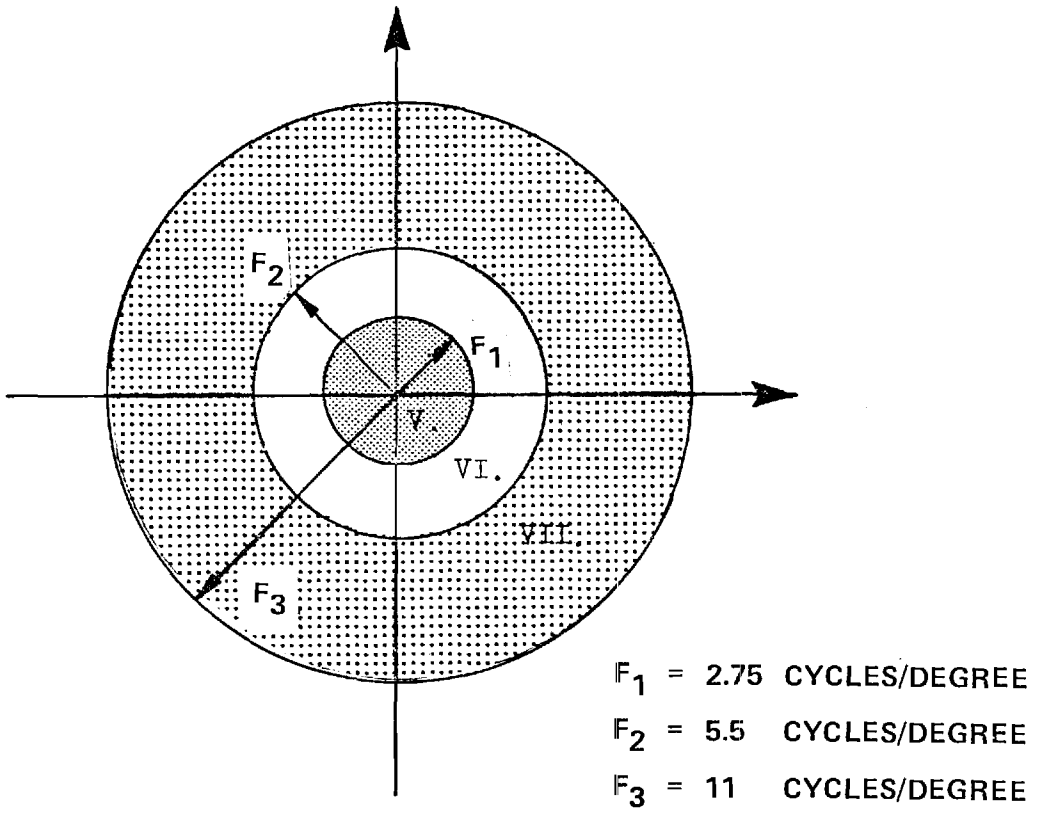
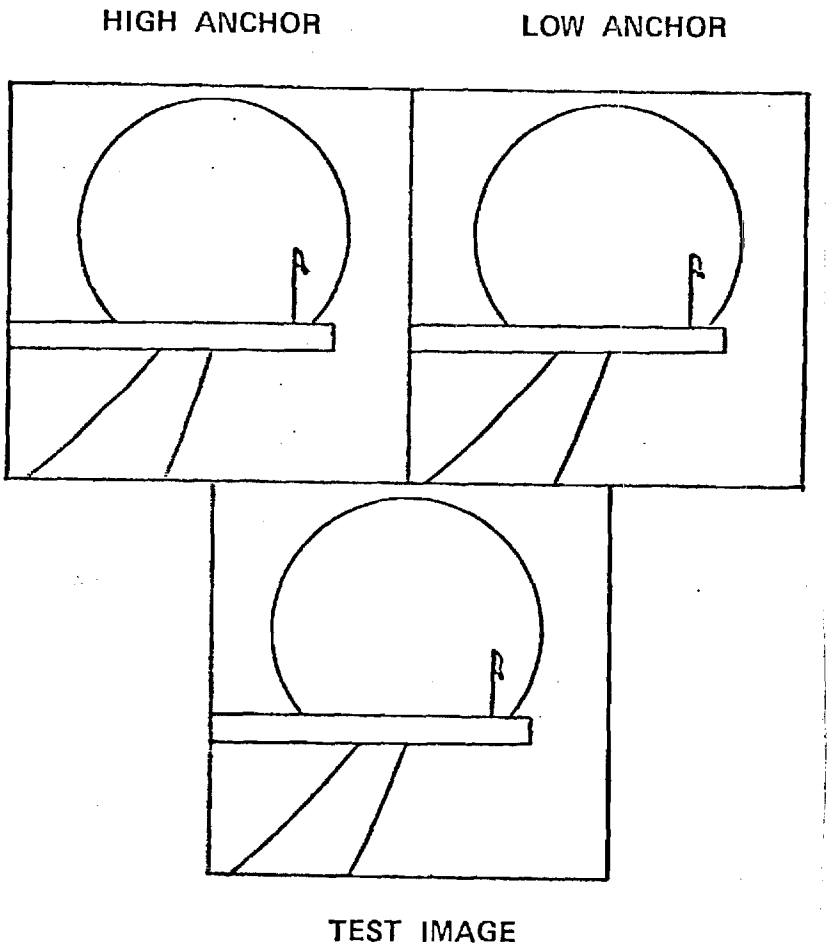


Figure 3. A test slide from the subjective test. The high anchor and low anchor at the top of the slide remained the same. The image under test is presented below.



point resolution and 256 gray levels. The nonlinearities of the CRT and film were approximately linearized using the procedure described in Appendix E. The images were displayed as negative images on negative film. In this fashion the slides could be made directly with only one level of developing.

In the test, subjects were seated so that the image under test subtended an angle of approximately six degrees. The subjects were asked to rate each distorted image on a scale of 0 to 100, and were told that the high anchor deserved a score of 80 and the low anchor deserved a score of 20. In this way, subjects could indicate whether or not a distortion actually enhanced (in their opinion) the quality of an image. The subjects were not trained. For each of the two original pictures, the distortions were randomized and presented at fifteen second intervals in groups of 120. In all twenty subjects participated in the tests for each of the two images.

For the statistical analysis of the results, a test similar to the Newman-Keul test [2] was used. In this test, first the mean across subjects for each distortion is computed.

$$S_i = \frac{1}{L} \sum_{j=1}^{L} S_{ij} \tag{3}$$

where S_{ij} is the response of the jth subject to the ith image and L is the number of subjects (in this case twenty). The quantity S_i was then used as the subjective estimate of quality for the ith image. In addition, the quantities

$$\overline{S} = \frac{1}{M} \sum_{i=1}^{M} S_i \tag{4}$$

$$\hat{\sigma}^2 = \sigma^2_{TOTAL} = \frac{1}{LM-1} \sum_{i=1}^{M} \sum_{j=1}^{L} (S_{ij}-\bar{S})^2 \tag{5}$$

$$\hat{\sigma}^2_{SYS} = \frac{L}{M-1} \sum_{i=1}^{M} (S_i-\bar{S})^2 \tag{6}$$

$$\hat{\sigma}^2_{ERROR} = \frac{1}{M(L-1)} \sum_{i=1}^{M} \sum_{j=1}^{L} (S_{ij}-S_i)^2 \tag{7}$$

are computed where M is the number of distortions (120 for our experiment) and

$$\hat{\sigma}^2 = \frac{M-1}{LM-1} \hat{\sigma}^2_{SYS} + \frac{M(L-1)}{LM-1} \hat{\sigma}^2_{ERROR} \quad . \tag{8}$$

The average distortions S_i are then ranked in descending order. When this is done the statistic q can be defined where

$$q = \frac{S_i - S_{i'}}{\frac{1}{\sqrt{L}} \hat{\sigma}_{ERROR}} \tag{9}$$

and $i' \geq i$. Due to the ranking of the distortions it follows that $q \geq 0$. In fact, q is distributed with the Studentized probability distribution function, $Q_{\alpha,R,f}(q)$. In this distribution, as it is commonly tabulated, R denotes the difference in estimated quality of the two distortions being compared plus one (thus $2 \leq R \leq M$); $f = M(L-1)$, the number of degrees of freedom for $\hat{\sigma}^2_{ERROR}$ and α denotes the desired quantile point. (We have used $\alpha = .01$ and $\alpha = .05$ for our study.) Thus using the q statistic we can determine whether a distortion was perceived to be significantly more objectionable

than another with a statistical significance at the α level. By considering all of the distortion in pairs, we can determine the resolving power of the test.

The results of the subjective statistical analysis for the class 2 distortions (additive white gaussian noise) is shown in Table 2. The complete set of subjective results is given in Appendix F. In these tables the matrix of 1's and 5's gives the significance levels for the various differences in quality score. For example, a 1 appearing in row i and column i' says that the difference of quality scores on distortions i and i' is statistically significant at the .01 level. Had there been a 5 there instead of a 1, the scores would have been statistically significant at the .05 level, but not at the .01 level. If a blank appears at any location, then the difference in scores is not significant at either level.

Several points can be made about the results of these subjective tests. First note that, in the case of the additive noises, the noise amplitude increased linearly with the index of the distortion level. The subjective results showed clearly that equal steps in noise amplitude did not result in equal steps in perceived image quality. It would appear that the perceived image quality is more nearly proportional to the logarithm of the signal to noise ratio.

Another interesting result comes from the lowpass blur distortion results (distortion class 8). Here the distortion on GIRL was not perceived to be significant until the bandwidth of the lowpass filter became less than 10 cycles per degree after which perceived quality dropped sharply as the bandwidth was reduced. This might simply say that GIRL does not have significant energy at frequencies above 10 cycles/degree but that it has

SUBJECTIVE RESULTS

STORTION #	DISTORTION SNR	AVERAGE QUALITY SCORE	DISTORTION # 1	2	3	4	5	6	7	8	9	10	13	11	12	14
1	24.9	69.5	-													
2	21.5	61.2	1	-												
3	19.1	54.2	1	5	-											
4	17.3	50.7	1	1		-										
5	15.7	50.2	1	1			-									
6	14.4	47.1	1	1				-								
7	13.3	44.4	1	1	5				-							
8	12.3	39.0	1	1	1	1	1	5		-						
9	11.4	38.8	1	1	1	1	1	5			-					
10	10.5	38.5	1	1	1	1	1	5				-				
13	8.48	34.0	1	1	1	1	1	1	1				-			
11	9.80	33.8	1	1	1	1	1	1	1	5	5			-		
12	9.12	30.3	1	1	1	1	1	1	1	5	5	5			-	
14	7.89	29.6	1	1	1	1	1	1	1	5	5	5				-
15	7.33	29.2	1	1	1	1	1	1	1	5	5	5				

Table 2. Results of the subjective quality test for additive white Gaussian noise. If a "1" appears at the intersection of two distortion levels this means that the difference in their quality scores is significant at the .01 level. Similarly a "5" means that the difference is significant at the .05 level.

significant energy at frequencies just below this level.

On the whole, the subjective quality test used in this study was judged to be a good test for these correlation studies. The results in Appendix F show that the test consistently found significant differences in perceived quality even for distortion levels which were very close. Furthermore, the corresponding results for the two different images were very similar and all of the quality averages within a distortion class showed monotonic or near monotonic behavior with distortion level. The standard error, given by $\hat{\sigma}/\sqrt{L}$, where $\hat{\sigma}$ is the sample standard deviation for one distortion and L is the number of subjects, ranged from 2.9 to 1.25 corresponding to an average resolving power of about four quality points at the .01 level.

IV. OBJECTIVE QUALITY MEASURES

A. Fidelity Measures for Images

Finding a fidelity measure for images which is both in good accord with subjective evaluations and which is simultaneously mathematically tractable, is a difficult task. While it is generally agreed that mean squared error correlates poorly with subjective evaluations [3,4] much of the previous work on image data compression has used this criterion [5-9] because of its mathematical convenience. A frequency-weighted mean-squared error criterion has also been used [10].

For imagery which is to be viewed by human observers, a logical candidate for a better fidelity measure would be one which is based on a model for the human visual system. Such a measure would be expected to correlate more closely with subjective measures of image fidelity, although it would most certainly not possess the computational simplicity of mean-squared error. Recent psychophysical testing experiments [11-15] seem to indicate that the eye responds to imagery by filtering the scene through a set of spatial channels and then responds if the response on any one of these channels exceeds a certain threshold. There is also evidence that the response of these receptors is pointwise non-linear [1,3,4]. Stockham [4] argues for a nonlinearity of the form $f(x) = \log x$. Mannos and Sakrison [1], in developing a fidelity criterion based on the results of subjective experiments, preferred the nonlinearity $f(x) = x^{.33}$. Actually, for the range of intensities encountered in images, the distinction between these operators is slight. A phenomenological model for the front-end of the human visual

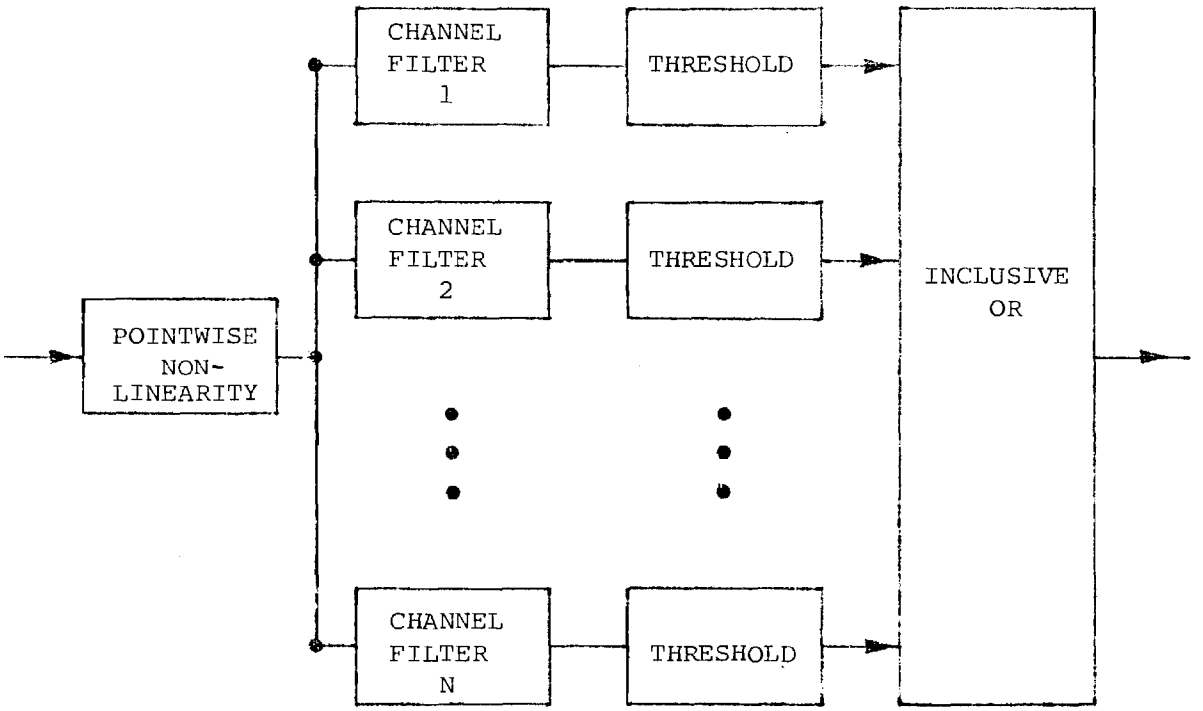
system is thus shown in Figure 4. This model is intended only to crudely represent the response function of the visual system and the components of this model do not represent physiological entities.

Given such a model for the visual system, it can be seen why mean-squared error might not correlate well with subjective measures of distortion. Two images with the same energy could evoke contrary responses from the model. If the energy is distributed throughout the spatial channels, the response would be quite different from the case when all of the energy is concentrated in a single channel. Similarly, two images which are identical except for a moderate difference in intensity between them would, in general, have a large mean-squared error but subjectively would be regarded nearly equivalent. The pointwise nonlinearity in the model along with the thresholds would make the responses equal unless the amplitude of one picture brought the signal energy below its thresholds.

What is therefore desired is a fidelity criterion which, when its value is low, guarantees that the distortion measured by each channel is low. One approach that can be followed is to use a vector-valued fidelity measure [16]. A vector-valued distortion measure in this case would imply the minimization of the error on each of the disjoint spatial-frequency channels. Such a measure, while theoretically mathematically tractable, in practice can be very difficult to use and does not have the physical meaning that we can associate with a scalar measure, e.g. it is difficult to rank images which are evaluated using a vector measure.

In an attempt to develop a scalar measure which nonetheless reflected the structure of the channelized vision model, Mannos and Sakrison [1] used the average weighted distortion over the spatial frequency channels. The

Figure 4. A model for the visual detection system.



frequency weighting which they derived on the basis of subjective experiments was of the form

$$A(f) = [c + (f/f_o)^{k_1}]\exp[-(f/f_o)^{k_2}] \tag{10}$$

with f_o = 8.77 cycles/degree, $k_1 = 1$, $k_2 = 1.1$, and $c = .019$. They also used a cube root nonlinearity for the pointwise operator. Unfortunately, when the Mannos and Sakrison error is small, this does not guarantee that the error will be small on each of the channels of the vision model.

The criterion developed for this research effort, which represents an innovation of this project and which we have referred to as the Gray-Leiner metric selects the maximum weighted distortion over a series of disjoint spatial frequency channels. It thus overcomes the potential deficiency of the Mannos-Sakrison measure while still providing a scalar value.

B. Implementation of the Objective Quality Measures

The calculations involved in evaluating the error energies which were used for the objective fidelity measures are summarized in Figure 5. Let $u(m,n)$ denote the intensity pattern of the original image and $\hat{u}(m,n)$ denote the intensity pattern of the reproduced (distorted) image. A cube-root transformation is first applied to both the image and its reproduction giving $w(m,n) = u^{1/3}(m,n)$ and $\hat{w}(m,n) = \hat{u}^{1/3}(m,n)$. The mean squared errors between u and $\hat{u}$ and between w and $\hat{w}$ are also computed. These errors are denoted E_{MS} and E_{NL}, respectively. They are not used in either the Mannos and Sakrison or the Gray-Leiner measures but are used for comparison purposes since they serve as measures (albeit subjectively poor ones) of

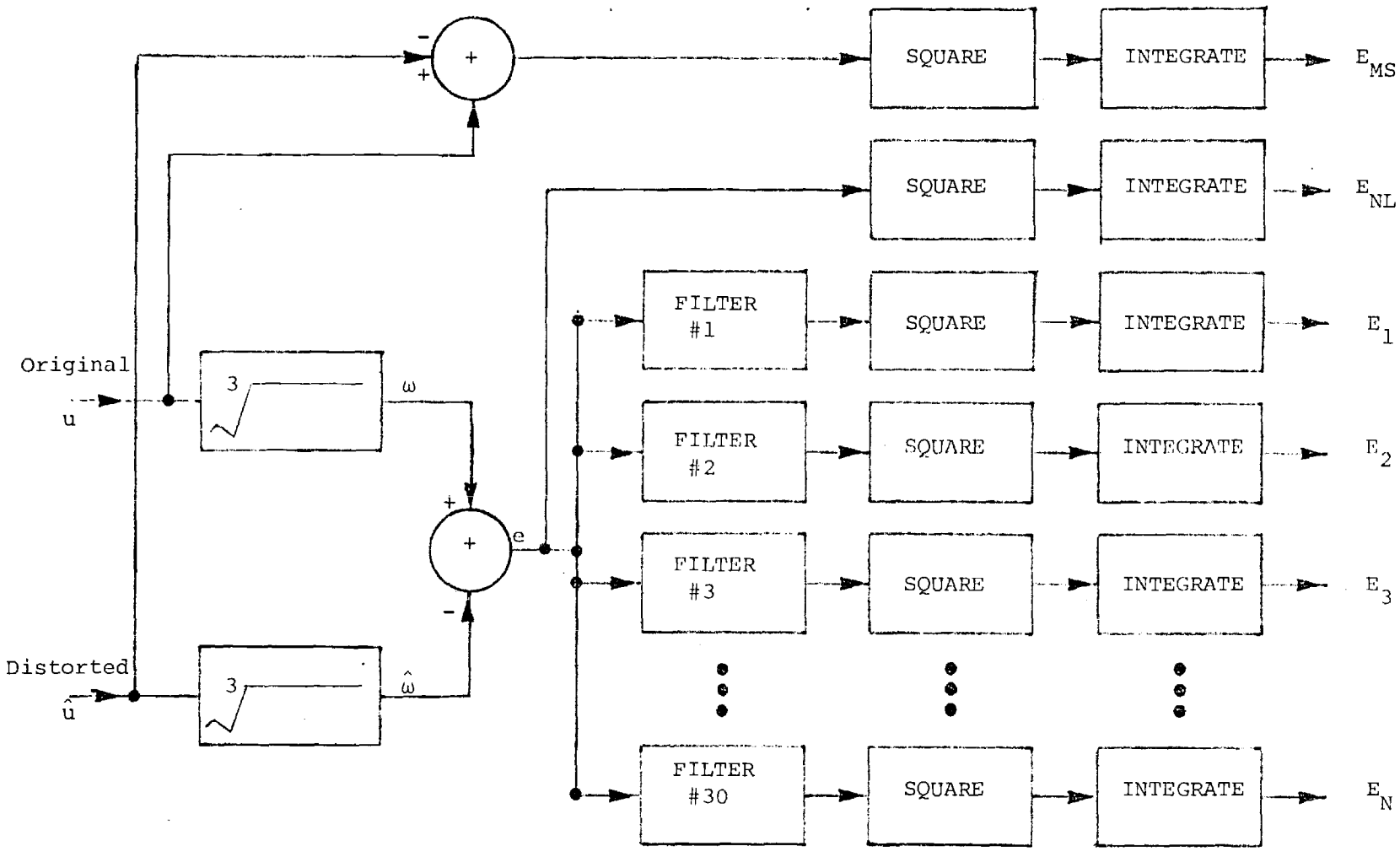


Figure 5. A block diagram of the computation steps to produce the error energies used for the objective quality measures.

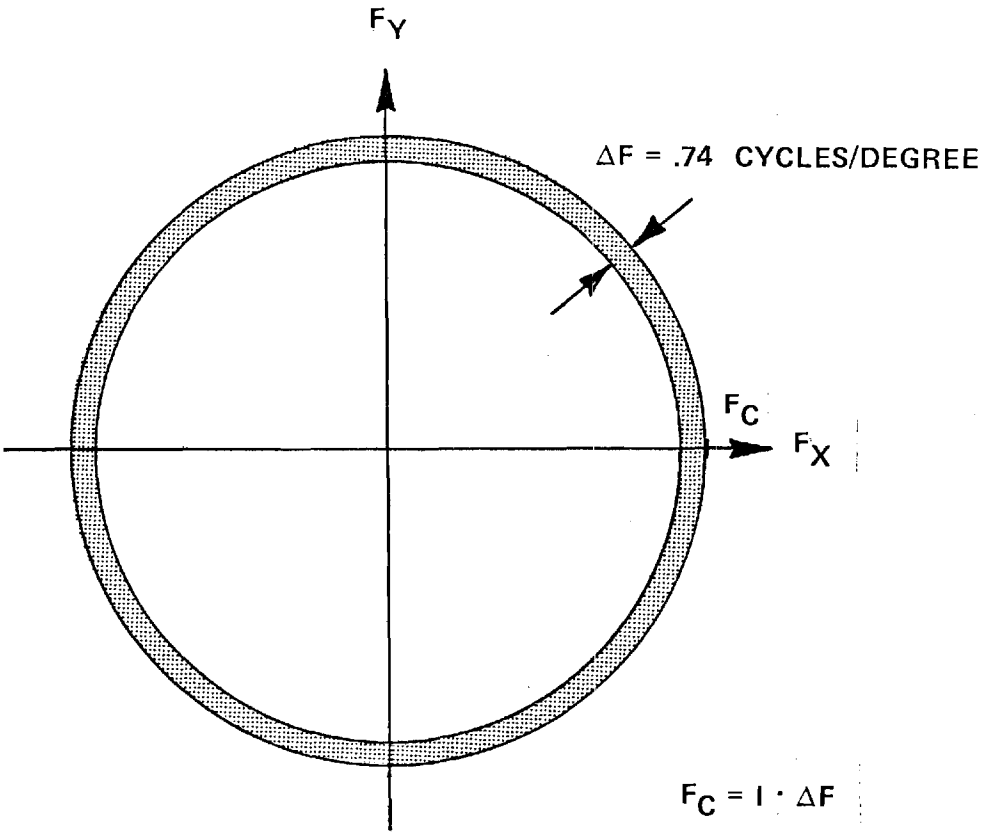
image distortion.

The signals w and $\hat{w}$ are then subtracted to produce the error signal $e(m,n)$. The energy of $e(m,n)$ in disjoint frequency bands is then computed by passing $e(m,n)$ through a filter bank and measuring the energies in the resulting waveforms. These energies are denoted by E_i, $i=1,2,\ldots,N$. The filter bank was designed so as to completely cover the spatial frequency spectrum up to a frequency of about 25 cycles/degree. (The human visual system can detect frequencies up to approximately 40 cycles/degree. Our lower cutoff was limited by our computer and display requirements.) The frequency response of one of the filters in the filter bank is shown in Figure 6. For our arrangement, 30 filters were used. They were chosen to be very nearly disjoint so that quality measures employing smaller numbers of filters could be simulated without the need to re-evaluate the error energies. The impulse response of the filters was 256×256 points. They were designed by the window method using a separable Hanning window. The filters were implemented in the frequency domain using the discrete Fourier transform. The accumulations of error energy were also performed in the frequency domain through the use of Parseval's relation.

All of the quality measures which were simulated utilized linear and nonlinear combinations of these thirty-two error values. Since these quantities were precomputed and stored for the 240 test images, the total data base for evaluating an objective quality measure involved a data base of only $240 \times 32 = 10{,}880$ data values. This is in contrast to the 15×10^6 values required to represent the complete set of distorted images.

It was also possible to realize computational savings in the implementation of the filter bank. In the frequency domain, we can express the

Figure 6. The frequency response of one of the filters in the filter bank shown in Figure 5.



energy in the ith channel of the filter bank as:

$$E_i = \frac{1}{(256)^2} \sum_{k=o}^{255} \sum_{\ell=o}^{255} |E(k,\ell)|^2 |H_i(k,\ell)|^2 \quad . \tag{11}$$

Since all of the filters in the filter bank possess the eight-fold symmetry implied by:

$$H_i(k,\ell) = H_i(\ell,k) = H_i(256-k,\ell) = H(k,256-\ell) \quad , \tag{12}$$

it follows that E_i can be computed from

$$E_i = \frac{1}{(256)^2} \sum_{k=o}^{128} \sum_{\ell=o}^{k} |a(k,\ell)|^2 |H_i(k,\ell)|^2 \tag{13}$$

where

$$a(0,0) = E_i(0,0)$$

$$a(k,0) = E_i(k,0) + E_i(256-k,0) + E_i(0,k) + E_i(0,256-k), \qquad k=1,\ldots,128$$

$$a(k,k) = E_i(k,k) + E_i(256-k,k) + E_i(k,256-k) + E_i(256-k,256-k), \qquad k=1,\ldots,128$$

$$a(k,\ell) = E_i(k,\ell) + E_i(256-k,\ell) + E_i(k,256-\ell) + E_i(256-k,256-\ell) + E_i(\ell,k) + E_i(256-\ell,k) \qquad 0<k<128$$

$$+ E_i(\ell,256-k) + E_i(256-\ell,256-k) \quad . \qquad 0<\ell<k$$

(14)

The array $a(k,\ell)$ is independent of i. Thus from the array $e(m,n)$, $a(k,\ell)$ can be computed and then the error energies E_i can be computed through an inner product calculation with the thirty prestored arrays $|H_i(k,\ell)|^2$.

The distortion measures studied differed in the manner in which they used the error energies. For the Mannos and Sakrison measure an estimate of distortion was produced from a linear combination of the E_i. Specifically, if D_{MS} denotes the Mannos and Sakrison distortion, then

$$D_{MS} = \sum_{i=1}^{30} A^2(f_i)E_i \tag{15}$$

where $A(f_i)$ was defined in Eq. (10) and f_i is the center frequency of the ith filter band ($f_i = 0$ for filter number 0).

With the Gray-Leiner distortion measure the distortion is chosen to equal the maximum of the weighted filter energies. If the full set of filters is used, it would take the form

$$D_{GL} = \max_i [W_i E_i] \tag{16}$$

where the weights W_i are parameters. If a smaller number of frequency bands are used, the E_i would be replaced by energies in those bands.

A third distortion measure studied, like the Mannos and Sakrison measure, formed a distortion estimate from a linear combination of the error energies. Here, however, the weighting coefficients were allowed to be free parameters and were chosen to maximize the correlation coefficient $|\hat{\rho}_k|$ between the quality estimated from the measure and the results of the subjective experiment. If we assume that B bands are used (B may be less

than 30), then an estimate of image quality can be formed as

$$C_j = \sum_{i=1}^{B} a_i E_{ij} + a_o \tag{17}$$

where C_j denotes the quality estimate for test image j and the E_{ij} denote the filter bank energy values for the jth image. If S_j denotes the mean subjective response for this test image, then the total prediction error is given by

$$E = \sum_{j=1}^{240} (C_j - S_j)^2 = \sum_{j=1}^{240} (S_j - \sum_{i=1}^{B} a_i E_{ij} - a_o)^2 \tag{18}$$

If we define $E_{0j} = 1$ for all j and further define

$$\underline{A}^T = [a_0, a_1, \ldots, a_B] \tag{19a}$$

$$\underline{\rho}^T = [\rho_0, \rho_1, \ldots, \rho_B]^T \tag{19b}$$

and

$$\underline{R} = [R_{mn}] \tag{19c}$$

where

$$\rho_i = \sum_{\ell=1}^{240} S_\ell E_{i\ell} \tag{19d}$$

and

$$R_{mn} = \sum_{\ell=1}^{240} E_{m\ell} E_{n\ell}$$

then Equation (18) can be rewritten as

$$E = \underline{A}^T \underline{RA} - 2\underline{\rho}^T \underline{A} + \sum_{\ell=1}^{240} S_\ell^2 \quad . \tag{20}$$

Taking partial derivatives with respect to each of the unknown coefficients $\{a_i\}$ and setting each of them to zero gives the result

$$-2\underline{\rho} + 2\underline{RA} = \underline{0} \quad . \tag{21}$$

Thus the vector $\underline{A}$ which minimizes E is given by

$$\underline{A} = \underline{R}^{-1} \underline{\rho} \quad . \tag{22}$$

If this is viewed as a linear regression analysis, the estimate obtained using Equation (17) with the coefficients determined according to (22) gives the maximum correlation possible for a linear estimator. Hence, this technique represents a bound on the performance of any "linear" distortion measure. (The quotes are due to the fact that the prediction is linear in the E_{ij}, although it must be remembered that the E_{ij} are not linear in the original error $(u(m,n)-\hat{u}(m,n))$. Since the Mannos and Sakrison error is a member of this class, its correlation with subjective results must necessarily be less than the prediction obtained from (17). On the other hand, the prediction of (17) may not be robust. Since the quality predictor is "tuned" on a specific data set, we cannot guarantee that

its predictions will be optimal (or even decent) when used on other distorted images. Also, we cannot draw any physiological conclusions about the coefficients $\{a_i\}$; they represent the solution of a set of linear equations and nothing more.

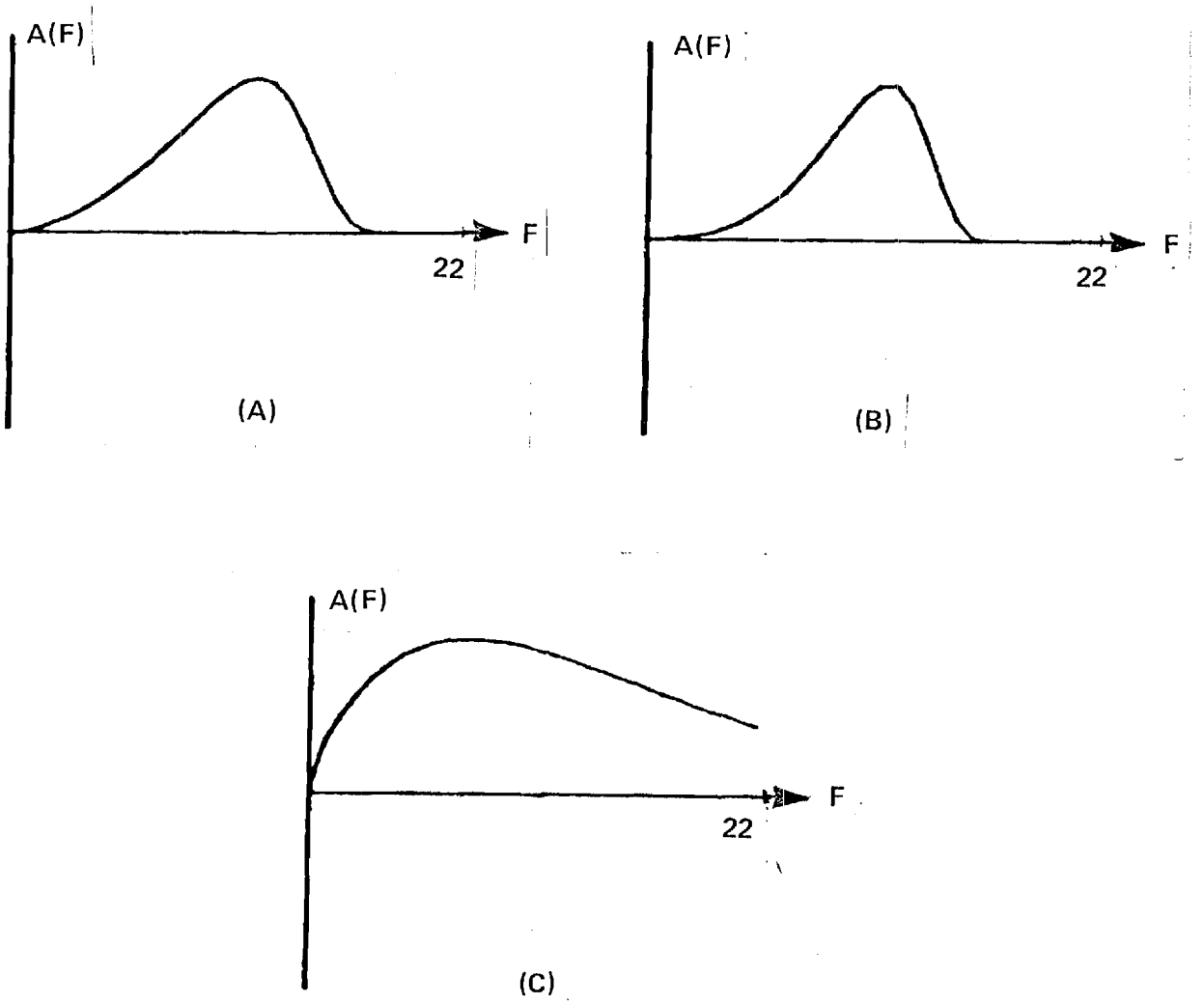
V. RESULTS OF THE CORRELATION STUDIES

In this section the correlation coefficients between candidate distortion measures and the results of subjective tests are presented. The methods by which these figures were calculated have been described in the previous sections. Two preliminary experiments were run in order to generate a feeling for what correlation values should be considered good and what values should be considered a poor correlation. The correlation coefficient between subjective estimates of quality and a mean-squared error estimate of distortion, E_{MS}, was -.16425. Similarly, the correlation coefficient with the nonlinear error, E_{NL}, was -.18979. These results suggest that neither of these measures used alone represents a good measure for image quality. On the other hand, both of these measures are easy to compute and can be readily used in the design of digital coding systems. Thus any proposed quality measures whose correlation values come close to these values is unacceptable as a fidelity measure. At the other extreme, it is to be expected that the best correlation with subjective results would come from the subjective results themselves. Thus the following experiment was performed. The quality evaluations of one subject were used as an "objective" quality measure and the results of the other nineteen subjects were averaged to produce the subjective estimate. This was then repeated three times to assure statistical regularity. The measured correlations were $\hat{\rho}_1 = .954$, $\hat{\rho}_2 = .889$, and $\hat{\rho}_3 = .876$. Thus a single human subject could be expected to have a correlation of .906 with the mean subjective result. Any quality measure which could approach this value would have to be

considered very good since it would then be equivalent to a human observer.

Using the Mannos and Sakrison distortion measure of Equations (10) and (15) with their parameters ($C = .019$, $f_o = 8.77$ cycles/degree, $k_1 = 1$, $k_2 = 1.1$), we observed a subjective correlation coefficient of -.658. Their model was based on the results of some subjective experiments (quite different from ours) to determine a model for the human visual system. Since our goal was simply to generate a quality measure which would correlate well with subjective responses, we felt no need to have a physiologically justifiable amplitude weighting function A(f). Thus, we considered the possibility of changing A(f) to maximize the correlation coefficient between the subjective and objective evaluations. We preserved the functional form of (10) since it would allow the shape of the weighting function to be controlled by four parameters and since the Mannos and Sakrison values could be used as initial values for an optimization. The four parameters of (10) were varied one at a time in an iterative fashion until a (possibly local) maximum value of the correlation coefficient was found. This optimization was performed two times resulting in "optimized" correlation coefficients of -.735 and -.732, respectively. The final parameter values for the two runs were $f_o = 13.3$, $C = .001$, $k_1 = 2.7$, $k_2 = 9.5$, and $f_o = 14.4$, $C = -.009$, $k_1 = 1.7$, $k_2 = 9.20$, respectively. Although these two parameter sets would appear to be quite different if the amplitude weighting functions A(f) are plotted the shapes of the two curves are seen to be virtually identical, but significantly different from the Mannos and Sakrison distribution. These amplitude weighting functions are shown in Figure 7. At this point, two things should be stated. First, the optimized amplitude weighting function is not consistent with most vision models which have been developed [1].

Figure 7. The amplitude weighting functions A(f) for the Mannos and Sakrison quality measure with (a) the first set of optimized parameters, (b) the second set of optimized parameters, (c) the original M&S parameters.



Secondly, the functional form chosen from (10) was chosen for convenience. Other skeleton amplitude weighting functions might have demonstrated higher correlation with subjective results.

It will be recalled that the Gray-Leiner metric estimates image distortion according to

$$D_{GL} = \max_i [W_i E_i] \quad .$$

There are a number of free parameters in this model, among them being the number of filter bands B, the bandwidths, and the weighting coefficients W_i. Due to the inherent nonlinearity of the metric, the determination of the optimum parameter set for the measure is a difficult and, from the point of view of this research project, unsolved problem. If the full thirty filters in the filter bank are used with equal bandwidths for the filters, and weighting coefficients chosen according to the Mannos and Sakrison formula (10), the resulting correlation coefficient was -.45715. This is worse than the Mannos and Sakrison metric with the same parameters. To try to improve on this figure an iterative search was made for an "optimal" parameter set. The Mannos and Sakrison skeleton was used again and the maximum correlation which resulted using 30 filters was -.708.

In another study, a Monte Carlo approach was used to determine the effect of using fewer than 30 filter bands. The bandlimits for the reduced number of filters were determined according to the following algorithm. To clarify the discussion let us denote each of the thirty original bands as subbands. They will be clustered into larger bands which we will simply call bands. The innermost subband is automatically assigned to the first

band. Working out from the center at each subband we can ask a simple question--Should this subband be added to the current band or should it become the first subband in a new band? The decision was made using a simple, binary random process. At each subband the subband would form a new band with probability p and would be added to the current band with probability 1-p. On the average, the number of bands which results will be 30p. With the value p=.2, using the optimized Mannos and Sakrison weighting coefficients, the results contained in Table 3 were obtained. The data set was not large and the weighting coefficients used were not optimized. Despite these facts, two conclusions are evident. Performance can be improved by reducing the number of bands somewhat and the Gray-Leiner metric yields slightly higher correlations than the optimized Mannos and Sakrison measure. These results are quite encouraging and strongly suggest that this metric is worthy of further study, particularly with respect to the question of parameter determination.

The last quality measure studied was the optimal linear measure described in the preceding section with Equation (17). Using all thirty bands a correlation coefficient of .924 was obtained. As described in that section, this value can be considered to be a limit for all "linear" distortion measures. This is a truly remarkable result for the correlation is higher than when a single subject is used to estimate quality.

One possible defect with this measure might be its lack of robustness. That is, since it is optimized for a particular data set, the resulting measure might not perform well on other distorted images. In order to generate a feeling for the robustness of the metric the following experiment was performed. An optimal set of weighting parameters was

B	Bandedges (Subband Index)	Correlation
3	1,2,4	-.731
4	1,3,20,29	-.732
5	1,5,11,12,20	-.734
6	1,6,10,18,24,30	-.726
7	1,5,10,21,24,25,27	-.733
8	1,3,7,9,10,15,18,27	-.720
9	1,4,6,10,17,18,21,27,28	-.680
30	all	-.708

Table 3

Some subjective correlations using the Gray-Leiner metric with differing numbers of filters,B

determined for the image GIRL and then the resulting fidelity measure was used to estimate the quality of distorted versions of the image RADOME. The correlation coefficient with subjective estimates of RADOME was .882. This is still comparable to using a single human observer and suggests that robustness may not be a problem. Such a conclusions cannot be definitely made, however, without a larger data base.

In order to investigate the effect of varying the number of spatial frequency bands on the fidelity measure, the Monte Carlo algorithm described earlier was used. These results are summarized in Table 4. The number of bands can be reduced considerably without a substantial penalty.

B	Bandedges (Subband Index)	Correlation
4	1,13,23,28	.801
5	1,4,8,9,21	.807
6	1,4,6,7,21,27	.831
7	1,9,11,12,24,28,29	.842
8	1,5,6,10,18,21,26,29	.872
10	1,3,9,10,13,17,21,22,28,29	.893
13	1,3,7,8,10,11,12,15,20,22, 28,29,30	.895
17	1,5,6,7,8,10,12,14,16,17, 19,20,21,23,27,28,30	.902
19	1,2,3,4,5,6,7,9,13,14,15, 16,19,22,23,24,25,28,30	.903
20	1,2,4,5,6,7,8,9,13,14,15, 16,18,20,22,24,25,27,28,29	.912
30	all	.924

Table 4

Some subjective correlations using the optimal linear metric with differing numbers of filters, B

VI. DISTORTION-RATE FUNCTIONS FOR VECTOR SOURCES AND A MAXIMUM FIDELITY CRITERION

Rate-distortion theory is concerned with the problem of encoding a signal in as few bits as possible subject to a requirement that the distortion in the reproduced signal be less than some maximum amount. The number of bits depends upon the measure used to evaluate distortion and the statistics of the signal encoded. In this section, we will summarize some results obtained during this research on bounds for distortion-rate functions using a fidelity measure similar to the Gray-Leiner measure. First, however, some preliminaries are in order to lead up to these results.

Let us assume that we have a representation for an image $u(x,y)$ which completely specifies that image. We shall further assume that this representation implies a data rate for transmission which is too high for the channel over which the transmission is to occur, since this is the only interesting case from a source coding point of view. Let the encoded version of the signal be denoted as $u_1(x,y)$. Since are interested in errors due to the source encoding and not those due to channel errors, we shall assume a simplified form for the communications channel, namely that it has a fixed capacity but for transmissions at rates lower than capacity it is error-free. At the receiver the received signal is then decoded to yield an estimate of the original image $\hat{u}(x,y)$. We thus desire that the distortion between u and $\hat{u}$ be as small as possible given that the data rate of the encoded image is less than the maximum data rate of the channel.

This may be stated precisely as follows. We are given a set of possible images, S, along with a probability measure, q, for those images.

The source produces uncorrelated images from S which need to be transmitted at the same rate at which they are produced. We are also given a set of possible reproduction images $\hat{S}$ which are available at the receiver, along with a measure of the distortion between source and reproduction $d(u,\hat{u}) \geq 0$ defined on $S \times \hat{S}$. We desire to reproduce the images at the receiver with as small an average distortion as possible while keeping the data rate below R. To accomplish this we could use a block coding technique. Let $\underline{u} = (u_1, u_2, \ldots, u_n)$ define a set of n images with a similar definition for $\hat{\underline{u}}$. The distortion between the image n-tuples then becomes just the average distortion, i.e.

$$d(\underline{u}_n, \hat{\underline{u}}_n) = n^{-1} \sum_{i=1}^{n} d(u_i, \hat{u}_i) \tag{23}$$

A code of block length n, C_n, is simply a collection of a finite number of reproduction image n-tuples. The size of the code $|C_n|$ is just the number of code words (n-tuples) in C_n. The encoding technique is simply to choose the code word in C_n that has minimum distortion with the source word, i.e. $\underline{u}_n$ is encoded into $\hat{\underline{u}}_n(\underline{u}_n)$ if

$$d(\underline{u}_n, \hat{u}_n(\underline{u}_n)) = \min_{\hat{\underline{u}}_n \in C_n} d(\underline{u}_n, \hat{\underline{u}}_n) \quad . \tag{24}$$

The code, therefore, has an average distortion

$$d(C_n) = E\left[\min_{\hat{\underline{u}}_n \in C_n} d(\underline{u}_n, \hat{\underline{u}}_n)\right] \tag{25}$$

where the expectation is taken with respect to the probability measure on S^n given by

$$q(\underline{u}_n) = \prod_{i=1}^{n} q(u_i) \quad . \tag{26}$$

We note that, if the channel is capable of transmitting R bits per second, then the encoded version of the images can be transmitted over the channel as long as the rate of the code $R(C_n) = n^{-1}\log|C_n|$ is less than R. The minimum distortion at rate R for codes of block length n is then given by choosing the best code. Thus

$$d_n(R) = \min_{C_n : R(C_n) \le R} d(C_n) \quad . \tag{27}$$

Since we are not restricted to any particular length code, the minimum distortion achievable at rate R is just $d(R) = \inf_n d_n(R)$.

This minimum distortion is usually impossible to calculate directly as the class of codes is large and the computations are difficult. Shannon [17], however, showed that d(R) may be related to a well-defined information theoretic minimization in certain circumstances which is amenable to fast computer programming techniques [18]. This famous result, the coding theorem for sources subject to a fidelity criterion, may be stated as follows for memoryless sources. Let $p(\hat{u}|u)$ be a conditional probability measure (test channel) of reproduction images given source images. This test channel may be thought of as a random coding device. Let $I(u,\hat{u})$ be the average mutual information between u and $\hat{u}$ under the joint probability measure induced by q and p and let D(q,p) be the average distortion, i.e.

$D(q,p) = E[d(u,\hat{u})]$. The minimum value for this average distortion is given by the (marginal) distortion-rate function $S_U(R)$ defined by

$$S_U(R) = \inf_{p(u|\hat{u}):I(u,\hat{u})\leq R} E[d(u,\hat{u})] \quad . \tag{28}$$

The marginal distortion-rate function gives the minimum distortion at a given rate when a scalar source is transmitted with no auxiliary information available. Thus $S_U(R)$ is obtained under a random encoding rule while $d(R)$ is the minimum distortion under a deterministic code. The source coding theorem [19] provides further meaning for the distortion-rate function.

> <u>Source Coding Theorem</u>: Let U be a discrete-time memoryless source as defined above and assume there exists a finite set of reproduction images $\{u_j\}$ such that $E[\min_j d(u,\hat{u}_j)]<\infty$. Then $d(R) = S_U(R)$.

The source coding theorem, therefore, guarantees that, as long as there is any finite code having finite average distortion, for long code blocklengths, we can come close to the distortion-rate function. Thus, the distortion-rate function provides a tight lower bound on the average distortion which is achievable by any coding system.

While the distortion-rate function which corresponds to the fidelity measure proposed in Section II is well-defined and can, in theory, be calculated, the size of the possible input and output alphabets makes this effectively an impossible task. Thus, it is desirable to develop bounds to the distortion-rate function which are more easily calculated. These

bounds can then be used to bound the performance of any actual data compression technique. Due to the form of our fidelity measure, we may simply restrict ourselves to the problem of transmitting an image which has been transformed by the pointwise nonlinearity at the front-end of the distortion measure. By determining the statistics of the transformed image, the problem may be restated as simply the determination of the distortion-rate function for a vector source with a distortion measure given by the supremum of the distortion for the individual frequency channels.

The bounds developed here are in the form of inequalities relating the distortion-rate function for a vector source with a supremum-type fidelity measure to that of scalar sources with and without side information. Since the latter distortion-rate functions may be calculated using numerical techniques [18], the resulting bounds may also be calculated and are, therefore, usable bounds.

For simplicity, consider the distortion-rate function for a vector source of dimension 2. When both the source encoder and decoder are allowed to observe a second sequence of information, the minimum distortion attainable is given by the conditional distortion-rate function. Let $U = (X,Y)$ be a memoryless vector source having joint probability density functions $q_U(u)$, marginal densities $q_X(x)$ and $q_Y(y)$ and conditional density $q_{X|Y}(x|y)$. The conditional distortion-rate function of the source X given Y is then defined by

$$\delta_{X|Y}(R) = \inf_{p(\hat{x}|x,y):I(X;\hat{X}|Y)\leq R} E[d(X;\hat{X})] \quad . \tag{29}$$

The conditional distortion-rate function is no more difficult to calculate than the marginal distortion-rate functions.

Now consider the transmission of a vector source when the user has a fidelity criterion on each component of the source. If an additive scalar fidelity criterion $d(u,\hat{u})\geq 0$ is defined for the set of source and decoded images, the fact that the source and reproduction data are vectors is immaterial and the distortion-rate function is still defined by (28).

Let the distortion for the vector source be expressed as $d(u,\hat{u}) = \max[ad_X(x,\hat{x}),bd_Y(y,\hat{y})]$ and the distortion-rate function for the source U, $\delta_U(R)$, be defined by (28). Then, if we denote by $\delta'_U(R)$ the distortion-rate function corresponding to the distortion measure $d(u,\hat{u}) = ad_X(x,\hat{x}) + bd_Y(y,\hat{y})$, we have shown the following relations [20].

$$\delta_U(R) \geq \inf_{\alpha\in[0,1]} \max[a\delta_{X|Y}(\alpha R),b\delta_Y((1-\alpha)R)] \tag{30a}$$

$$\delta_U(R) \leq \delta'_U(R) \leq a\delta_X(\alpha R) + b\delta_Y((1-\alpha)R), \quad \alpha\in[0,1] \tag{30b}$$

Furthermore, if we define the distortion-rate function $\delta_X^\sigma(R)$ for $\sigma\geq 1$ to be the distortion-rate function for X using the criterion $(d_X(x,\hat{x}))^\sigma$ and likewise for Y, we have that

$$\delta_U(R) \leq (a^\sigma\delta_X^\sigma(R) + b^\sigma\delta_Y^\sigma((1-\alpha)R))^{1/\alpha}, \quad \alpha\in[0,1] \tag{31}$$

For identically distributed sources these results may be simplified as

$$\delta_U(R) \leq 2\delta_X(R/2) \quad . \tag{32}$$

If in addition X and Y are independent, then

$$\delta_U(R) \geq \quad {}_X(R/2) \quad . \tag{33}$$

The calculation of the lower bound of (30a) is simplified by noting that, when an $\alpha \epsilon [0,1]$ exists such that $a\delta_{X|Y}(\alpha R) = b\delta_Y((1-\alpha)R)$, the infimum is achieved by that α. Otherwise, it is achieved by one of the endpoints. While these bounds are not as tight as we might desire, neither are they so weak as to be useless as Leiner [20] has shown in some examples (See Appendix B).

VII. DISCUSSION

An important point to bear in mind in evaluting the results of this study is that there were several issues being studied simultaneously. A subjective test was designed and evaluated. The correlation technique for evaluating objective quality measures was developed, and a number of proposed image quality measures were implemented, compared, and refined. These distinct studies impacted upon one another in both obvious and subtle ways.

The doubly anchored subjective test was designed to prevent the "floating bias" effect often noticed when complex, highly structured signals are being judged. With tests which are not anchored, the internal standards by which a subject makes his judgements vary with time and the subjective response for an image depends not only on the distortion of the current image but also on the quality of the immediately preceding images. This phenomenon can be referred to as interdistortion interference. To test for inter-distortion interference, the correlation coefficient for adjacent judgements (in time) was computed and found to be only -.02. Thus, for the doubly anchored test, successive responses are virtually uncorrelated.

The resolving power of the subjective test also appeared to be very good, although no explicit attempt was made to compare this test with others. This test was, however, able to resolve (significantly) distortions which were difficult for a single subject to resolve while viewing both distortions simultaneously.

The clear utility of the correlation method for the comparison of image fidelity measures lies in its ability to study many objective

distortion measures with relative computational ease once the set of error energies E_i are known. Unfortunately, the amount of computation necessary to compute error energies was extensive, as was the effort to obtain subjective results. As a result we tended to squeeze as much data as possible from a small data base. Ideally, the data base would have been larger. Two sample images are clearly not enough and the sample distortions did not cover anywhere near all of the distortions possible with image coding techniques. Image quality perception is an immensely complex process, related not only to the distortions themselves but also to the highly structured nature of the images themselves. Thus, the perception of a distortion depends not only on its numerical measure but also on such factors as whether or not it is localized or spread over the whole picture, whether it affects high-information areas of the image or low information ones, etc. A somewhat similar situation occurs in the perception of speech. Thus, in the long run it may be somewhat simplistic to expect any relatively simple objective measure to be highly correlated with subjective results for all classes of images and distortions.

Some of these issues could have been addressed by increasing the size of the sample under study--by using more images and more distortions. However, although the actual correlation studies are quite compact, the development of the underlying data base is not, and the inclusion of a larger sample set was simply beyond the resources of this study. Similarly, we would have liked to be able to include a comparison with the use of pointwise nonlinearities other than the cube root but this would have increased the amount of computation considerably.

Some comments should also be made concerning the Mannos and Sakrison

distortion measure [1]. First, it should be noted that there were several basic differences between their study and ours. Their images were displayed with a bandwidth of 44 cycles per degree while our bandwidth was 22 cycles per degree. Further, their distortions were all coding distortions which were judged to be barely perceivable. Our were not induced by coding and were on several occasions much more severe. That their measure which was determined by other means could achieve a correlation of -.65 would imply that their measure was a relatively good one, although it may be questioned whether or not this value for the correlation coefficient is high enough for all of the purposes for which an objective measure is needed.

A fourth area of interest in this study was the evaluation of the Gray-Leiner distortion measure. Although the evaluation of this measure was one of the primary goals of the original study, it must be stated that the results obtained can only show trends in the effectiveness of this measure. For example, the results suggest that a relatively small number of bands work best. Further, for the Mannos and Sakrison weighting function, better results can sometimes be obtained with this measure than with the linear measure. A difficulty is that there is no simple method for finding a weighting function which will result in the optimal behavior for this measure. Iterative techniques are unattractive since the correlation coefficient surface as a function of the weighting function will not be a smooth curve. Exhaustive methods are unattractive since the number of cases is prohibitively large. Hence, the results of this study do not represent data on the absolute effectiveness of this measure.

The last area of interest in this study was the optimal linear predictors which gave a correlation of .922. This is an impressive number,

but since this number is obtained from reapplication of the measure to its own training set, the robustness of the measure must still be proven. The weighting function obtained by this method contains both positive and negative values, and its high correlation coefficient can be thought of as the result of subtle effects of the numerical analysis. However, it should be pointed out that when the GIRL was used as the training set, and the RADOME was used as the test set, a correlation coefficient of .88213 was obtained. This suggests that, at least for similar distortions, this may be a relatively robust measure comparable to the .90 obtained from the single subject tests, and a set of weights produced from a larger data base might have a more general application.

A. Open Research Questions

There are several intriguing questions to be answered which would make interesting follow-on studies. First is the question of whether these results would remain the same when applied to a larger base with more test images and more types of distortions, particularly real coding distortions. Of particular interest here is the performance of the optimal linear predictor. Another open question is the assignment of weights for the Gray-Leiner metric. Initial results obtained with this fidelity measure were very encouraging. Finally, the most important open questions associated with objective fidelity measures concern their use. What do these measures say about optimal image coding? How do existing easy to use coding schemes compare? Finally, what does a fidelity measure say about image processing and image restoration?

ACKNOWLEDGEMENT

We would like to acknowledge the able assistance of Mr. James Klassen who developed the procedures for the calibration of the image display devices and prepared the images for the subjective tests.

REFERENCES

[1] J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Trans. Info. Theory, IT-20, pp. 525-536, 1974.

[2] B. J. Winer, Statistical Principles in Experimental Design, McGraw-Hill, 1962.

[3] Z. L. Budrikis, "Visual Fidelity Criterion and Modelling," Proc. IEEE, 60, pp. 771-779, July 1972.

[4] T. G. Stockham, Jr., "Image Processing in the Context of a Visual Model," Proc. IEEE, 60, pp. 828-842, July 1972.

[5] P. A. Wintz, "Transform Picture Coding," Proc. IEEE, 60, pp. 809-820, July 1972.

[6] M. M. Sondhi, "Image Restoration: The Removal of Spatially Invariant Degradations," Proc. IEEE, 60, pp. 854-861, July 1972.

[7] A. Habibi, "Two-dimensional Bayesian Estimate of Images," Proc. IEEE, 60, pp. 878-883, July 1972.

[8] W. K. Pratt, "Transform Image Coding Spectrum Extrapolation," Proc. Seventh Hawaii Int. Conf. on System Sciences, pp. 7-9, 1974.

[9] M. Ishii et al., "Picture Bandwidth Compression by DPCM in the Hadamard Transform Domain," Proc. Seventh Hawaii Int. Conf. on System Sciences, pp. 10-12, 1974.

[10] L. D. Davisson, "Rate-distortion Theory and Application," Proc. IEEE, 60, pp. 771-779, July 1972.

[11] M. Sachs, J. Nachmias, and J. Robson, "Spatial-Frequency Channels in Human Vision," J. Opt. Soc. Am., 61, pp. 1176-1186, 1971.

[12] H. Mostafavi and D. Sakrison, "Structure and Properties of a Single Channel in the Human Visual System," Vision Research, 16, pp. 957-968, 1976.

[13] D. J. Sakrison, "On the Role of the Observer and a Distortion Measure in Image Transmission," IEEE Trans. Communications, COM-25, pp. 1251-1267, 1977.

[14] C. F. Stromeyer, III, and B. Julesz, "Spatial-frequency Masking in Vision: Critical Bands and Spread of Masking," J. Opt. Soc. Am., 62, pp. 1221-1232, 1972.

[15] L. Harman, "Masking in Visual Recognition: Effects of Two-dimensional Filtered Noise," Science, 180, pp. 1194-1197, 1973.

[16] R. M. Gray, "A New Class of Lower Bounds to Information Rates of Stationary Sources via Conditional Rate-Distortion Functions," IEEE Trans. Info. Theory, IT-19, pp. 480-489, 1973.

[17] C. E. Shannon, "Coding Theorems for a Discrete Source with a Fidelity Criterion," in 1959 IRE Nat. Conv. Rec., pt. 4, pp. 142-163.

[18] R. Blahut, "Computation of Channel Capacity and Rate-distortion Functions," IEEE Trans. Info. Theory, IT-18, pp. 460-473, 1972.

[19] R. G. Gallager, Information Theory and Reliable Communication, New York, Wiley, 1968.

[20] B. M. Leiner, "Distortion-rate Functions for Vector Sources and a Maximum Fidelity Criteria," IEEE Trans. Infor. Theory, IT-23, pp. 354-359, 1977.

APPENDIX A

This appendix consists of a reprint of the paper:

B. M. Leiner, "Image Encoding with a Fidelity Criterion," Proc. 13th Ann. Allerton Conf. on Circuit and System Theory, pp. 335-344, 1975.

IMAGE ENCODING WITH A FIDELITY CRITERION

BARRY M. LEINER
Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

The problem of compressing pictorial data for transmission is investigated. It is seen that the classic measures of distortion such as mean-squared error are not indicative of the quality of the picture. A new measure is suggested and some of the theoretical properties related to the optimum coding (i.e., rate-distortion function) are considered.

INTRODUCTION

In transmitting images, the capacity of the channel is often not sufficient to allow perfect reproduction of the image at the receiver without taking an inordinately long time for transmission. One solution to this problem is to encode the images to remove any redundancy, allowing accurate reproduction of the image at a lower bit rate. This bit stream then requires a lower capacity channel or alternatively, less time to transmit each image. The capacity may still, however, not be adequate.

Requiring perfect reproduction is often unrealistic. In a black and white picture, for instance, relative intensity of each point is often more important than the exact intensity. When the statistics of the image are known as well as a quantitative fidelity measure, the minimum number of bits required to reproduce the image within a given fidelity may be calculated using the techniques of rate-distortion theory. We can then try to find various encoding schemes which would hopefully come close to this lower bound.

Unfortunately, the application of rate-distortion theory requires knowledge of both the source (image) statistics and an adequate fidelity measure. While an estimate of the statistics of the image may be obtained using histogram techniques on a large sample of images of the type concerned with, determining the appropriate fidelity measure is a more difficult problem. Classically, the fidelity measure used is a mean-squared error criterion. Although this measure is analytically "nice," it does not correlate well with human perception of the characteristics for a good reproduction.

In this paper, fidelity criteria are discussed for images that are computationally tractable. The distortion measure chosen is motivated in part by recent psychophysical tests and in part by the desire for a tractable measure for images. It is seen that, while this measure leads to a well-defined and theoretically computable distortion-rate function, in practice its computation is difficult. The remainder of the paper deals with the development of bounds to this distortion-rate function, thereby

This work was partially supported by the National Science Foundation under Grant No. ENG75-04992.

developing bounds to the performance of any data compression technique with the chosen fidelity criterion.

DATA COMPRESSION FOR IMAGES

The problem of the transmission of images may be stated as follows. A sender has an image which he would like reproduced for the receiver. The sender has available to him a communication link which operates at some rate. If we allow the sender and receiver to install coding and decoding equipment, what is the best reproduction of the image available to the receiver after some fixed amount of time?

Figure 1 shows the type of system envisioned. The image is first converted to an electrical signal. This would typically be accomplished by a scanning device whose output depends on the pointwise intensity of the image. At the converter output, we then have a signal, U, that completely represents the image. This signal has a data rate which is too high for the channel and must be encoded, giving the signal U_1. To transmit this signal over the channel reliably, it must also be encoded giving the signal U_2 which is sent over the channel. The received signal is then decoded to give $\hat{U}$, an electrical representation of the image. This representation can then be converted to a reproduction of the image. If the channel capacity is not sufficient to transmit a perfect reproduction of the image, the receiver will view a distorted version of the original image. Before proceeding, it should be noted that, in general, the image encoder and channel encoder (and likewise for the decoders) should be shown as a single block. The fact that, under a wide variety of circumstances including the present, the source (image) encoding may be done independent of the channel encoding is a major result of Information Theory [1, p. 3]. Since the interest here is in the required data rate for image reproduction, we shall not be concerned with the channel coding. It will be assumed that the channel has some fixed maximum data rate and never makes any errors.

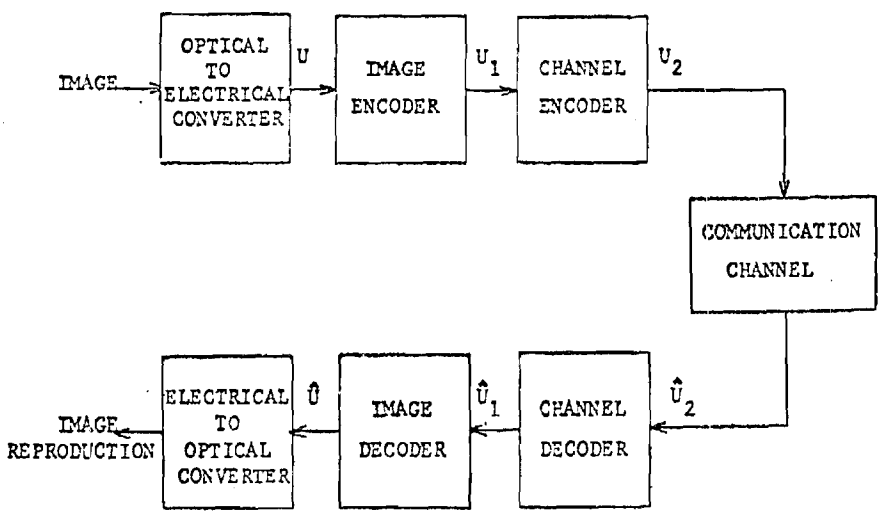


Figure 1. Image Transmission Systems

The system to be analyzed may be redrawn as shown in Figure 2. The representation of the image U is encoded and decoded into a reproduction $\hat{U}$. We desire that the distortion between U and $\hat{U}$ be as small as possible given that the data rate of the encoded image is less than the maximum dat rate of the channel.

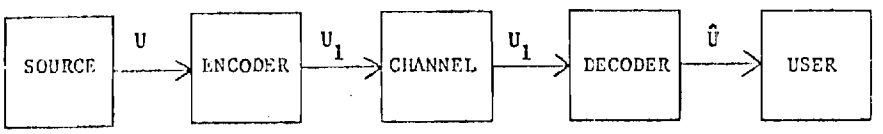


Figure 2. Image Coding System

This may be stated precisely as follows. We are given a set of possible images, S, along with a probability measure, q, for the images. The source produces images from the set S at some rate, say one per secon Each image is produced independently of the other images according to the given probability measure. The set of images may be all possible pages o text, all possible pictures, etc. The probability measure may be allowed to determine the likely images. We may, therefore, take S to be the set all possible intensity patterns on the square $[0, L] \times [0, L]$. We are envisioning a source which produces a set of uncorrelated images that nee to be transmitted at some rate. It should be noted that, by allowing the images to be correlated, we may consider the problem of television transmission in the same framework.

We are also given a set of possible reproduction images $\hat{S}$ which are available at the receiver, along with a measure of the distortion between source and reproduction $d(U, \hat{U}) \geq 0$ defined on $S \times \hat{S}$. We desire to reprod the images at the receiver with as small an average distortion as possibl while keeping the data rate below R. To accomplish this, we will use a block coding technique. Let $\underset{\sim}{U}_n = (U_1, U_2, \ldots, U_n)$, i.e., a set of n ima and similarly for $\underset{\sim}{\hat{U}}_n$. Define the distortion between the image n-tuples a just the average distortion, i.e.,

$$d(\underset{\sim}{U}_n, \underset{\sim}{\hat{U}}_n) = n^{-1} \sum_{i=1}^{n} d(U_i, \hat{U}_i) \quad .$$

A code of block length n, C_n, is simply a collection of a finite number o reproduction image n-tuples. The size of the code $|C_n|$ is just the number of code words (n-tuples) in C_n. The encoding technique is simply to choose the code word in C_n that has minimum distortion with the source word, i.e., $\underset{\sim}{u}_n$ is encoded into $\underset{\sim}{\hat{u}}_n(\underset{\sim}{u}_n)$ if

$$d(\underset{\sim}{u}_n, \underset{\sim}{\hat{u}}_n(\underset{\sim}{u}_n)) = \min_{\underset{\sim}{\hat{u}}_n \in C_n} d(\underset{\sim}{u}_n, \underset{\sim}{\hat{u}}_n) \quad .$$

The code, therefore, has an average distortion

$$d(C_n) = E \min_{\hat{\underset{\sim}{u}}_n \in C_n} d(\underset{\sim}{U}_n, \hat{\underset{\sim}{u}}_n)$$

where the expectation is taken with respect to the probability measure on S^n given by

$$q(\underset{\sim}{u}_n) = \prod_{i=1}^{n} q(u_i) \quad .$$

We note that, if the channel is capable of transmitting R bits per second, then the encoded version of the images can be transmitted over the channel as long as the rate of the code, $R(C_n) = n^{-1} \log |C_n|$, is less than R. The minimum distortion at rate R for codes of block length n is then given by choosing the best code so that

$$d_n(R) = \min_{C_n : R(C_n) \le R} d(C_n) \quad .$$

Since we are not restricted to any particular length code, the minimum distortion achievable at rate R is just $d(R) = \inf_n d_n(R)$.

This minimum distortion is usually impossible to calculate directly as the class of codes is large and the computations very difficult [2]. Shannon [3] showed that d(R) may be related to a well-defined information theoretic minimization in certain circumstances which is amenable to fast computer programming techniques [4]. This famous result, the coding theorem for sources subject to a fidelity criterion, has been generalized several times [1,5,6] and may be stated as follows for memoryless sources (the generalization to sources with correlated images may be found in the references). Let $p(\hat{u}|u)$ be a conditional probability measure (test channel) of reproduction images given sources images. This test channel may be thought of as a random coding device. Let $I(q, p)$ be the average mutual information between U and $\hat{U}$ [1] under the joint probability measure induced by q and p and let $D(q, p)$ be the average distortion, i.e., $D(q, p) = E\, d(U, \hat{U})$. The distortion-rate function for the source with alphabet S, probability measure q, and fidelity measure d is then defined by

$$\delta(R) = \inf_{p : I(q, p) \le R} D(q, p) \quad .$$

Thus, $\delta(R)$ is the minimum distortion under a random encoding rule with rate replaced by mutual information while d(R) is the minimum distortion under a deterministic code. The source coding theorem provides the meaning of the distortion-rate function [1,3,5].

Source Coding Theorem: Let U be a discrete time memoryless source as defined above and assume there exists a finite set of reproduction images $\{\hat{U}_j\}$ such that $E \min_j d(U, \hat{U}_j) < \infty$. Then $d(R) = \delta(R)$.

The source coding theorem therefore guarantees us that, as long as there is any finite code having finite average distortion, we can for long code blocklengths, come close to the distortion-rate function. Thus, the distortion-rate function provides a tight lower bound on the average distortion which is achievable by any real coding system.

Unfortunately, to apply the theory outlined above, it is necessary to have both a probability measure for the source and a fidelity measure which agrees with human perception. By considering histograms of images and calculating various moments, such as the autocorrelation function for the image, we can obtain a reasonable statistical model for the class of images of concern. Fortunately, it is not necessary that the statistical model be perfect as the coding technique once chosen has a rate independent of the source and the difference between the average distortion for the model and the real source may be bounded [24]. Alternately, the distortion for a class of sources having the given statistics may be considered [23]. Thus, an analysis based just on a model derived from first and second order statistics will provide a great deal of useful information.

FIDELITY CRITERIA FOR IMAGES

Finding a meaningful fidelity measure is a more difficult problem. Much of the previous work on image data compression has used either a mean-squared error criterion [7,8,9,10,11] or a frequency-weighted mean-squared error [12,16]. While it was recognized that mean-squared error is a poor criterion for images [13,14], its simplicity of calculation plus a lack of a better criterion has made it popular.

The search for a better criterion must be carried out by turning to models of the visual system. Recent psychophysical testing [20,21,22,25] seems to indicate that the eye detects information by passing the scene through a set of spatial channels and then performs an "or" operation on the output of the channels. There is also evidence [13,16] that prior to channelizing, a pointwise nonlinear operation of the form $f(x) = x^{1/3}$ is performed. A model of the visual detection system that would include these results in shown in Figure 3. Given this model, it is easily seen why a mean-squared criterion, weighted or not, would not correlate with a human perception of distortion. A given level of mean-squared error could be achieved by having a low level of distortion for each channel or a high level of distortion for one channel. The first may not be detected by any channel and would therefore be perceived as a good reproduction while the second would be detected on the one channel and be perceived as a poor reproduction.

What therefore is desired is a criterion that, when its value is low, guarantees that the distortion on each channel is low. Taking the maximum distortion over the given channels would achieve this goal. A class of candidate distortion measures for images is therefore the following. Let $U(x,y)$ be the intensity pattern of the original image and let $\hat{U}(x,y)$ be th intensity pattern of the reproduction image, where x and y vary from 0 to the size of the image. A nonlinear point by point transformation is first performed on both the image and its reproduction, giving $W(x,y) = f(U(x,y)$ and $\hat{W}(x,y) = f(\hat{U}(x,y))$. The transformed images are then passed through a of ideal two-dimensional bandpass filters, $\{H_i(f_x, f_y)\}_{i=1}^{N}$, giving outputs $V_i(x,y)$ and $\hat{V}_i(x,y)$ for each filter. These filters are chosen so as to

completely cover the spatial frequencies up to the human resolution of 40 cycles/deg [13]. The squared error in each channel,

$$d(V_i, \hat{V}_i) = \int_0^L \int_0^L (V_i(x,y) - \hat{V}_i(x,y))^2 \, dxdy$$

is then calculated. The distortion between the original image and its reproduction is then given by $d(U, \hat{U}) = \sup_i (a_i \, d(V_i, \hat{V}_i))$ where the a_i are positive constants. Work is currently underway by the author to verify this distortion measure and determine its associated parameters.

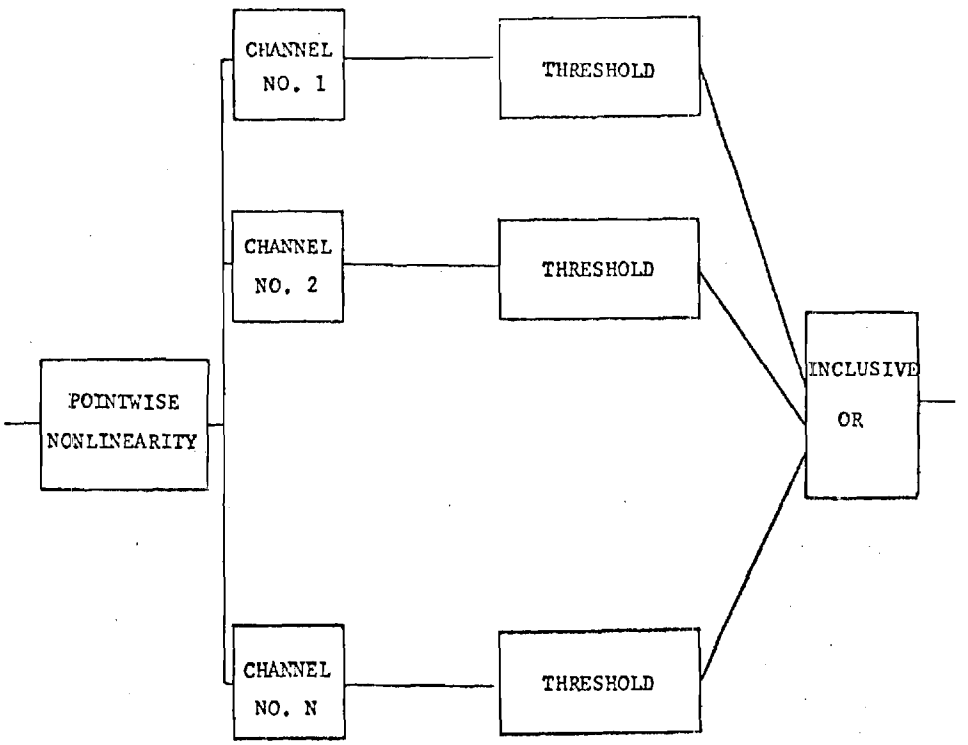


Figure 3. Visual Detection System Model

BOUNDS TO THE DISTORTION RATE FUNCTION

While the distortion-rate function corresponding to the class of fideli measures discussed above is well-defined and can in theory be calculated, the size of the possible input and output alphabets make this a formidable if not impossible task. It is therefore desirable to develop bounds to th distortion-rate function that are more easily calculated. These bounds ma then be used to bound the performance of any actual data compression technique. Since both the linear and non-linear transformations are one-t one, there is no distortion introduced in the transformations. We may therefore concern ourselves with the transmission of the transformed image By determining the statistics of the transformed image, the problem may be restated as simply the determination of the distortion-rate function for a vector source with a distortion measure given by the supremum of the disto tion for the individual components.

The bounds developed here are in the form of inequalities relating th distortion-rate function for a vector source with a supremum-type fidelity measure to that of scalar sources with and without side information. Sin the latter distortion-rate functions may be calculated using numerical techniques [4], the resulting bounds may also be calculated and are therefore usable bounds.

For simplicity here, consider the distortion-rate function for a vect source of dimension 2. This would correspond to having two bandpass filte The extension to arbitrary finite dimension is direct, but would result in unnecessary complexity here. Let (X, Y) be a vector source having pmf q(and marginal pmf's q(x) and q(y), where the per-letter alphabets are A_X, A_Y, and A_{XY} respectively. Let the available reproduction alphabets be $A_{\hat{X}}$, $A_{\hat{Y}}$, and $A_{\hat{X}\hat{Y}}$. Let the distortion between x and $\hat{x}$ be given by $d(x, \hat{x})$ and the distortion between y and y be $d(y, \hat{y})$. The distortion between a vector source letter (x,y) and a vector reproduction letter $(\hat{x}, \hat{y})$ is the given by

$$d((x, y), (\hat{x}, \hat{y})) = \sup (ad(x, \hat{x}), bd(y, \hat{y}))$$

where a and b are some arbitrary positive weighting constants. Define th distortion-rate functions $\delta_X(R)$, $\delta_Y(R)$, $\delta_{XY}(R)$, and $\delta_{X|Y}(R)$ to be the usu marginal and conditional distortion-rate function [1,5,17], e.g.,

$$\delta_X(R) = \inf_{p \in P} E\, d(X, \hat{X})$$

$$P = \{p(\hat{x}|x): E \log (p(\hat{X}|X)/w(\hat{X})) \le R\}$$

$$w(\hat{x}) = \sum_x q(x)\, p(\hat{x}|x)$$

and expectation is taken with respect to the joint pmf $q(x)\, p(\hat{x}|x)$. Note that the standard coding theorem holds for all four distortion-rate functions [1,5,15,17].

Although the distortion-rate function $\delta_{XY}(R)$ may be calculated directly using numerical techniques [4], the size of the joint alphabet

Λ_{XY} often makes this calculation prohibitive. It is, therefore, useful to consider bounding the distortion-rate function by functions which are more readily calculable. Prime candidates for these functions are the marginal distortion-rate functions $\delta_X(R)$ and $\delta_Y(R)$ and the conditional distortion-rate function $\delta_{X|Y}(R)$. By using techniques similar to those used in deriving the relations between the joint, marginal, and conditional rate-distortion functions [15,18 sec. 3], the following upper and lower bounds to the joint distortion-rate function for the class of distortion measure considered here may be shown to hold.

Theorem: Let $\delta_X(R)$, $\delta_Y(R)$, $\delta_{X|Y}(R)$, and $\delta_{XY}(R)$ be the distortion-rate functions as defined above. Then

$$\delta_{XY}(R) \geq \inf_{\alpha\in[0,1]} \{\sup (a\delta_{X|Y}(\alpha R),\ b\delta_Y((1-\alpha)R))\}$$

$$\delta_{XY}(R_1+R_2) \leq a\delta_X(R_1) + b\delta_Y(R_2) .$$

Furthermore, if we define the distortion-rate function $\delta_X^{(s)}(R)$ to be the distortion-rate function for the source and reproduction as above using a distortion measure $d_s(x,\hat{x}) = (d(x,\hat{x}))^s$, we have the upper bound for all $s \geq 1$

$$\delta_{XY}(R_1 + R_2) \leq (a^s\delta_X^{(s)}(R_1) + b^s\delta_Y^{(s)}(R_2))^{1/s} .$$

CONCLUSION

In this paper, the problem of compressing the data representation of an image was considered. A fidelity measure was suggested that has the tightness of a vector measure and also leads to a well-defined distortion-rate function. Bounds to this distortion-rate function that are more easily computable were developed. It is hoped that psychophysical testing of the proposed distortion measure will verify that it is indeed a criterion that reflects human perception and thus may be used in the evaluation of image compression systems.

ACKNOWLEDGEMENTS

The author would like to gratefully thank Dr. Robert M. Gray of Stanford University for his suggestions leading to this research and Roderick E. Thomas for his assistance with its computerized aspects.

REFERENCES

1. R. G. Gallager, Information Theory and Reliable Communication, New Wiley, 1968.

2. R. M. Gray and L. D. Davisson, "A mathematical theory of data compression?" to be published.

3. C. E. Shannon, "Coding theorems for a discrete source with a fidelit criterion," in 1959 IRE Nat. Conv. Rec., pt. 4, pp. 142-163.

4. R. Blahut, "Computation of channel capacity and rate-distortion functions," IEEE Trans. Inf. Theory, vol. IT-18, pp. 460-473, 1972.

5. T. Berger, Rate Distortion Theory, Englewood Cliffs, N.J.: Prentic Hall, 1971.

6. B. M. Leiner and R. M. Gray, "Rate-distortion theory for ergodic sources with side information," IEEE Trans. Inf. Theory, vol. IT-20, pp. 672-675, Sept. 1974.

7. P. A. Wintz, "Transform picture coding," Proc. IEEE, vol. 60, pp. 809-820, July 1972.

8. M. M. Sondhi, "Image restoration: the removal of spatially invariat degradations," Proc. IEEE, vol. 60, pp. 854-861, July 1972.

9. A. Habibi, "Two-dimensional Bayesian estimate of images," Proc. IEEI vol. 60, pp. 878-883, July 1972.

10. W. K. Pratt, "Transform image coding spectrum extrapolation," Proc. Seventh Hawaii Int. Conf. on System Sciences, pp. 7-9, 1974.

11. M. Ishii et. al., "Picture bandwidth compression by DPCM in the Hadamard transform domain," Proc. Seventh Hawaii Int. Conf. on Syst Sciences, pp. 10-12, 1974.

12. L. D. Davisson, "Rate-distortion theory and application," Proc. IEEI vol. 60, pp. 800-808, July 1972.

13. Z. L. Budrikis, "Visual fidelity criterion and modeling," Proc. IEEI vol. 60, pp. 771-779, July 1972.

14. T. G. Stockham, Jr., "Image processing in the context of a visual model," Proc. IEEE, vol. 60, pp. 828-842, July 1972.

15. R. M. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions," IEEE Trans. Inf. Theory, vol. IT-19, pp. 480-489, July 1973.

16. J. L. Mannos and D. J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. Inf. Theory, vol. IT-20, pp. 525-536, July 1974.

17. B. M. Leiner, "Joint, conditional and marginal distortion-rate functions," Proc. Seventh Hawaii Int. Conf. on System Sciences, pp. 40-42, 1974.

18. B. M. Leiner, "Rate-distortion theory for sources with side information," Ph.D. dissertation, Stanford University, Stanford, California, August 1973.

19. D. J. Connor et. al., "Infraframe coding for picture transmission," Proc. IEEE, vol. 60, pp. 828-842, July 1972.

20. M. B. Sachs et. al., "Spatial-frequency channels in human vision," Jour. Optical Society of America, vol. 61, pp. 1176-1186, Sept. 1971

21. C. F. Stromeyer, III and B. Julesz, "Spatial-frequency masking in vision: critical bands and spread of masking," Journ. Optical Society of America, vol. 62, pp. 1221-1232, Oct. 1972.

22. L. Harmon, "Masking in visual recognition: effects of two-dimensional filtered noise," Science, vol. 180, pp. 1194-1197, June 1973.

23. D. J. Sakrison, "The rate of a class of random processors," IEEE Trans. Inf. Theory, vol. IT-16, pp. 10-16, Jan. 1970.

24. R. M. Gray et. al., "A generalization of Ornstein's $\bar{d}$-distance with applications to information theory," Annals of Probability, vol. 3, no. 3, April 1975.

25. H. Mostafavi and D. J. Sakrison, "Structure and properties of a single channel in the human visual system," to be published.

APPENDIX B

This appendix consists of a reprint of the paper:

B. M. Leiner, "Distortion-Rate Functions for Vector Sources and a Maximum Fidelity Criterion," IEEE Trans. Info. Theory, vol. IT-23, pp. 354-359, 1977.

Distortion-Rate Functions for Vector Sources and a Maximum Fidelity Criterion

BARRY M. LEINER, MEMBER, IEEE

Abstract—The problem of distortion-rate functions for vector sources is considered. Two fidelity criteria are treated. The first considers the maximum of the weighted component distortions and then takes the per-letter average. The second takes the maximum of the weighted component per-letter averages. In either case, a well-defined distortion-rate function results, giving the minimum possible distortion achievable at a given rate of transmission. Upper and lower bounds to these distortion-rate functions are derived in terms of more easily calculated functions.

I. Introduction

RATE-DISTORTION functions have been defined for vector sources under a vector fidelity criterion and a weighted average fidelity criterion [1]. The rate-distortion function answers the following question, "What is the minimum rate required to transmit a data stream so that the average distortion of the reproduction is less than some specified amount?" When a vector fidelity criterion is used, each component of the average distortion is constrained, while for a weighted average distortion measure, only the average of the components is considered.

The fidelity requirement of a user is often not the specified quantity. In many situations, the user has a channel available to him of some fixed capacity and would like to know the best reproduction available while maintaining a rate less than capacity. The answer to this question lies in the inverse to the rate-distortion function, the distortion-rate function. While this function is well-defined for a scalar source and for a vector source with a weighted average distortion measure, in the case of a vector fidelity criterion it yields a set of distortion vectors and does not answer the user's question.

A weighted average distortion measure is often too weak a fidelity criterion. For example, recent psychophysical testing for images [2]–[5] seems to indicate that a reasonable criterion for image reproduction fidelity might be the maximum of the distortion over disjoint spatial frequency channels, a stronger requirement than a weighted average.

In this paper, such a maximum fidelity criterion is considered, leading to a well-defined distortion-rate function for vector sources. Bounds to this distortion-rate function are developed in terms of more easily calculable distortion-rate functions, namely, the distortion-rate function for a scalar source with and without side information. These bounds are similar in appearance to the relationships derived for rate-distortion functions of vector sources with vector and weighted average distortion measures [1]. Unlike those relationships, however, the bounds in this paper do not seem to have simple conditions for them to hold with equality. This does not mean that they are so weak as to be useless, as two simple examples will show.

A criterion which seeks to retain the character of a vector fidelity criterion, while still leading to a well-defined distortion-rate function for a vector source, is also considered. This criterion seeks to minimize the maximum of the average distortion, in contrast to the criterion discussed above which considers the average of the maximum distortion. In the case of data streams used by different users, this may be the more meaningful criterion and yet still yields a minimum distortion at a given rate. This distortion-rate function is related to the rate-distortion function of a vector source with vector fidelity criterion. Inequalities relating this distortion-rate function to that for sources with and without side information are also developed.

II. Marginal and Conditional Distortion-Rate Functions

Let U be a scalar memoryless source producing symbols from a source alphabet A_U according to the probability mass function (pmf) $q_U(u)$. It is desired to transmit an encoded version of the source at a rate less than R nats per second, obtaining a reproduction from alphabet $A_{\hat{U}}$ in such a way that the average per-letter distortion $d(u;\hat{u})$ between source and reproduction is minimized. The minimum value for this average distortion is given by the (marginal) distortion-rate function $\delta_U(R)$ defined by

$$\delta_U(R) = \inf_{p(\hat{u}|u):\ I(U;\hat{U})\le R} Ed(U;\hat{U}). \tag{1}$$

By noting that both mutual information and the average distortion are convex functions of the test channel probabilities $p(\hat{u}|u)$, it is readily seen that the distortion-rate function is just the inverse to the rate-distortion function and the appropriate source coding theorems hold [7]–[9]. Thus the marginal distortion-rate function gives the minimum distortion attainable at a given rate when a scalar source is transmitted with no auxiliary information available.

Manuscript received May 15, 1975; revised September 13, 1976. A portion of this work was presented at the IEEE International Communications Conference, San Francisco, CA, June 1975. This work was supported in part by the National Science Foundation under Grant ENG75-04992.

The author was with the School of Electrical Engineering, Georgia Institute of Technology, Atlanta, GA 30332. He is currently with Probe Systems, Inc., Sunnyvale, CA 94086.

When both the source encoder and decoder are allowed observe side information in the form of a sequence re-ted to the one being transmitted, the minimum distortion tainable is given by the conditional distortion-rate nction [11]. Let $U = (X,Y)$ be a memoryless vector urce having joint pmf $q_U(u)$, marginal pmf's $q_X(x)$ and (y), and conditional pmf $q_{X|Y}(x|y)$. The conditional stortion-rate function of the source X given Y is then fined by

$$\delta_{X/Y}(R) = \inf_{p(\hat{x}|x,y):\ I(X;\hat{X}|Y)\le R} Ed(X;\hat{X}). \tag{2}$$

ain, noting that both the average distortion and condi-nal mutual information are convex functions of the test annel $p(\hat{x}|x,y)$, the conditional distortion-rate function the inverse to the conditional rate-distortion function , [9], [10], and therefore the appropriate coding theorems ld.

In what follows, the distortion-rate function of a vector rce with various types of fidelity criteria will be related the marginal and conditional distortion-rate functions the component sources. The utility of these relations lows from the ease of computing the marginal and ditional distortion-rate functions, and the difficulty computing the vector distortion-rate function. Although s not obvious from (2) that the conditional distortion-e function is easily computed, it may be written as the ghted sum of the marginal distortion rate functions for sources $q_{X|Y}(\cdot|y)$, denoted as $\delta_{X|y}(R)$, as

$$\delta_{X|Y}(R) = \sum_y q_Y(y)\delta_{X|y}(R_y), \tag{3a}$$

ere

$$R = \sum_y q_Y(y)R_y, \tag{3b}$$

the $\{R_y\}$ are chosen such that $d/dR\,\delta_{X|y}(R)|_{R=R_y} = s$ each y [10], [11]. Since the slope is usually the param-used in the calculation of the distortion-rate function, er analytically through a form of Lagrange minimiza- [6], [7], or numerically [12], the conditional distor--rate function is no more difficult to calculate than the ginal distortion-rate function.

III. Joint Distortion-Rate Functions

e now consider the transmission of a vector source n the user has a fidelity criterion on each component e source and it is to be transmitted over a single nel. Let $U = (U_1,U_2,\cdots,U_N)$ be a memoryless vector ce having pmf $q_U(u)$, source alphabet $A_U = A_{U_1} \times A_{U_2} \cdots \times A_{U_N}$, and available reproduction alphabet $A_{\hat{U}} = \times \cdots \times A_{\hat{U}_N}$. Note that some elements of the source abet may have $q_U(u) = 0$. If an additive scalar fidelity rion $d(u;\hat{u}) \ge 0$ is defined on $A_U \times A_{\hat{U}}$, the fact that ource and reproduction data are vectors is immaterial the distortion-rate function is defined by (1).

is natural to define the distortion for the vector in s of the distortion for the component data streams. There are several choices for this distortion criterion. Let $d_i(u_i;\hat{u}_i)$ be the additive fidelity criterion for the ith data stream. The distortion between vectors could be defined to be the vector $d(u;\hat{u}) = (d_1(u_1;\hat{u}_1),\cdots,d_N(u_N;\hat{u}_N))$. Thus we are trying to minimize the vector distortion subject to a rate constraint. Although this approach will work in defining the rate-distortion function [1] resulting in the minimization of a scalar (the rate) subject to a vector constraint, it does not lead to a well-defined distortion-rate function in the following sense. We would like to have the distortion-rate function tell us the minimum achievable distortion at a given rate. If we use a vector distortion measure, we will find a set of distortion vectors rather than a single number characterizing the optimum system.

A scalar distortion measure for the vector source U can be defined by taking the weighted average of the component data streams,

$$d(u;\hat{u}) = \sum_{i=1}^{N} a_i d_i(u_i;\hat{u}_i). \tag{4}$$

The problem with the weighted average distortion measure is that it is often too weak a criterion for the source of interest. For example, recent psychophysical testing [2]–[4] seems to indicate that the human eye processes images through independent spatial frequency channels. One might then hypothesize that all the channels must pass an adequate version of their signal for the image to be useful. The weighted average distortion measure would not guarantee this, as a particular average distortion could be achieved by one component having a large distortion with the rest being zero or, alternatively, by all the components having a moderate distortion. We therefore would like to find a scalar distortion measure for the vector source that guarantees the distortion for each component is minimized, as it is for a vector distortion measure. Such a fidelity criterion may be achieved by considering the maximum distortion over all the components. Since some components may be more important than others, we allow a weighting of the component distortions, leading to the distortion measure [5]

$$d(u;\hat{u}) = \max_i a_i d_i(u_i;\hat{u}_i). \tag{5}$$

Thus we have a criterion which is stronger than the weighted average distortion measure, yet leads to a well-defined distortion-rate function as a minimum achievable distortion at the given rate. The appropriate source coding theorem therefore holds.

IV. Distortion-Rate Relations for Two-Dimensional Sources

Although we now have a well-defined distortion-rate function, its computation can be quite tedious. This is a direct result of the source and reproduction letters being vectors and the size of the alphabets therefore growing geometrically with the dimension of the vector. Since the minimization in the definition of the distortion-rate function involves the determination of the minimum

function defined on the joint alphabet $A_U \times A_C$, the computations can become difficult, even using Blahut's algorithm [12], for even moderate size component alphabets and dimension. It is therefore useful to determine upper and lower bounds to the joint distortion-rate function in terms of more easily calculable functions.

In this section, bounds to the distortion-rate function of a two-dimensional source are developed in terms of the marginal and conditional distortion-rate functions of the component subsources. Although these bounds may be generalized immediately to arbitrary finite dimension, considering the two-dimensional case simplifies both the notation and proofs. The relations derived here are similar in appearance to those derived in [1] for the joint rate-distortion function. Proofs for these relations may be found in the Appendix.

Let $U = (X,Y)$ be a two-dimensional vector source as defined above having distortion measures for the components given by $d_X(x;\hat{x})$ and $d_Y(y;\hat{y})$, marginal distortion-rate functions $\delta_X(R)$ and $\delta_Y(R)$, and conditional distortion-rate function $\delta_{X/Y}(R)$. Let the distortion for the vector source and its reproduction be given by $d(u;\hat{u}) = \max\,[ad_X(x;\hat{x}),bd_Y(y;\hat{y})]$, and let the distortion-rate function for the source U, $\delta_U(R)$, be defined by (1). Let the distortion-rate function calculated using distortion measure $d(u;\hat{u}) = ad_X(x;\hat{x}) + bd_Y(y;\hat{y})$ be denoted $\delta'_U(R)$. We then have the following relations.

Theorem 1: The distortion-rate function for the vector source U is related to the marginal and conditional distortion-rate functions of its component subsources by

$$\delta_U(R) \geq \inf_{\alpha\in[0,1]} \max\,[a\delta_{X/Y}(\alpha R),b\delta_Y((1-\alpha)R)] \tag{6a}$$

$$\delta_U(R) \leq \delta'_U(R) \leq a\delta_X(\alpha R) + b\delta_Y((1-\alpha)R), \qquad \alpha \in [0,1]. \tag{6b}$$

Furthermore, if we define the distortion-rate function $\delta_X^\sigma(R)$ for $\sigma \geq 1$ to be the distortion-rate function for X using the criterion $(d_X(x;\hat{x}))^\sigma$ and likewise for Y, we have that

$$\delta_U(R) \leq (a^\sigma\delta_X^\sigma(\alpha R) + b^\sigma\delta_Y^\sigma((1-\alpha)R))^{1/\sigma}, \qquad \alpha \in [0,1]. \tag{6c}$$

For identically distributed sources, Theorem 1 may be simplified as follows.

Corollary: Let X and Y be identically distributed (not necessarily independent) with the same distortion measure (i.e., $d_X(r;\hat{r}) = d_Y(r;\hat{r})$), and let $a = b = 1$. Then

$$\delta_U(R) \leq 2\delta_X(0.5R). \tag{7a}$$

If, in addition, X and Y are independent, then

$$\delta_U(R) \geq \delta_X(0.5R). \tag{7b}$$

The calculation of the lower bound of (6a) is simplified by noting that, when an $\alpha \in [0,1]$ exists such that $a\delta_{X|Y}(\alpha R) = b\delta_Y((1-\alpha)R)$, the infimum is achieved by that α. Otherwise, it is achieved by one of the endpoints.

While the bounds (6) and (7) are not tight, neither are they so weak as to be useless as can be seen from the following examples.

Example: Two-Dimensional Binary Source, Hamming Distortion

Let X and Y be dependent, identically distributed binary sources having $q_X(1) = 0.5$ and $q_U(1,0) = 0.5\rho$, so that $q_U(x,y)$ is given by

$$q_U(x,y) = \begin{cases} 0.5\rho, & x \neq y, \\ 0.5(1-\rho), & x = y. \end{cases}$$

Let $d_X(x;\hat{x})$ and $d_Y(y;\hat{y})$ be the probability of error criterion

$$d_X(r;\hat{r}) = d_Y(r;\hat{r}) = 1 - \delta_{r\hat{r}},$$

and let $a = b = 1$. It is easy to see that, for this case, U is a source having alphabet size equal to four and a probability of error criterion, since $d_U(u;\hat{u}) = 1$ unless $x = \hat{x}$ and $y = \hat{y}$. The distortion-rate function for the source U is then, from Berger [6, sect. 2.9], given parametically as

$$D_t = 1 - S_t + t(N_t - 1)$$

$$R_t = -\sum_{j\in V_t} P_j \log P_j + (1-D_t)\log(1-D_t) + (N_t - 1)\,t\log t,$$

where

$$V_t = \{j : P_j > t\}$$

$$S_t = \sum_{j\in V_t} P_j$$

$$N_t = \sum_{j\in V_t} 1,$$

the $\{P_j\}$ are the four probabilities $q_U(u)$, and t is a parameter that varies from 0 to the second largest of the $\{P_j\}$.

The lower bound to $\delta_U(R)$ is given by (6a) as

$$\delta_{LB}(R) = \inf_{\alpha\in[0,1]} \max\,[\delta_{X|Y}(\alpha R),\delta_Y(1-\alpha)R].$$

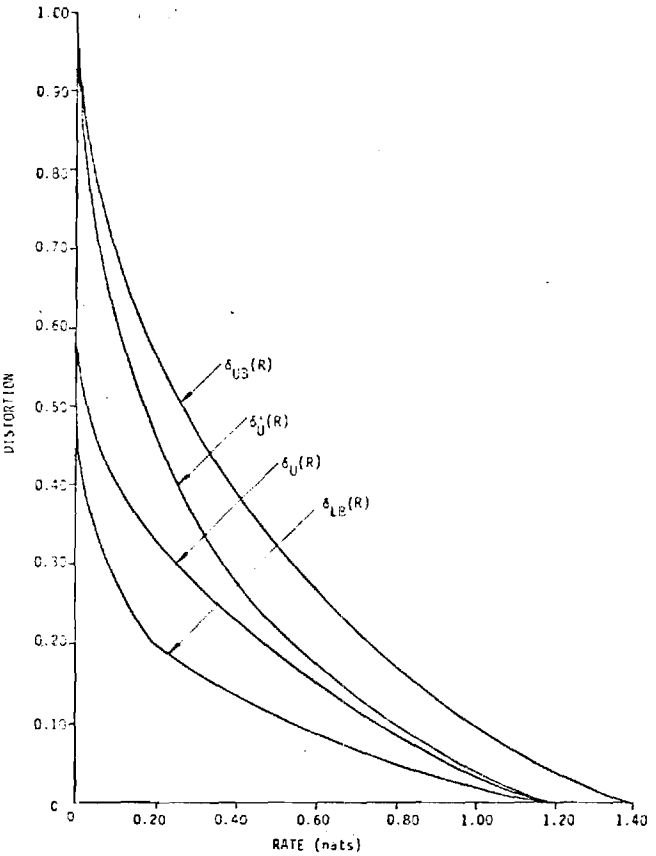
From Berger [2, example 2.7.1], the distortion-rate function for the conditioned sources $q_{X|Y}(\cdot\,|y)$ may be calculated for each y parametically as a function of the slope s. Applying (3) and combining terms yields the conditional distortion-rate function

$$\delta_{X|Y}(R) = h^{-1}(h(\rho) - R),$$

where $h(\cdot)$ is the binary entropy function,

$$h(x) = -x\log x - (1-x)\log(1-x),$$

and $h^{-1}(\cdot)$ is its inverse function defined to have range equal to [0,0.5].



. 1. Distortion-rate function and bounds: binary vector source, Hamming distortion, $\rho = 0.2$.

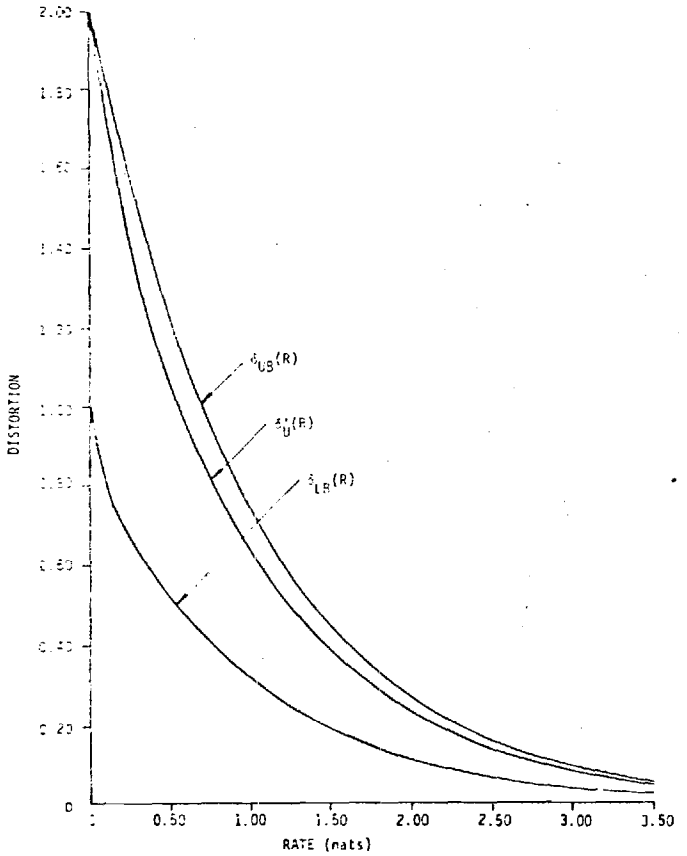


Fig. 2. Bounds to distortion-rate function: Gaussian vector source, mean-squared error, $\rho = 0.5$.

Similarly, the distortion-rate function for Y is given

$$\delta_Y(R) = h^{-1}(H(Y) - R) = h^{-1}(h(0.5) - R).$$

e lower bound is then given by

$$(R) = \inf_{\alpha \in [0,1]} \max\,[h^{-1}(h(\rho) - \alpha R), h^{-1}(h(0.5) - (1-\alpha)R)].$$

The upper bound of (6b), $\delta_U'(R)$, may be easily calcu-ed [2, example 2.7.2] noting that the distortion matrix e is just twice that given in the example. The bound is en implicitly as

$$\begin{cases} \log 2 + h(\rho) - 2h\left(\dfrac{D}{2}\right), & 0 \le D \le 1 - \sqrt{1-2\rho} \\ f(1-\rho) - \dfrac{1}{2}[f(D-\rho) + f(2-D-\rho)], & \\ \quad 1 - \sqrt{1-2\rho} \le D \le D_{\max}, & \end{cases}$$

re $f(x) = -x \log x$.

n this case, the corollary also applies giving the simple er bound

$$\delta_{UB}(R) = h^{-1}(h(0.5) - 0.5R).$$

ig. 1 shows the distortion-rate function and its bounds the case $\rho = 0.2$.

Example: Two-Dimensional Gaussian Source, Mean-Squared Error

As a second example, consider X and Y to be joint Gaussian random variables each having zero-mean and unit variance, and having correlation coefficient ρ. Let the component distortion measures be squared error so that

$$d(u;\hat{u}) = \max\,((x - \hat{x})^2, (y - \hat{y})^2).$$

This distortion measure has the following geometric interpretation. Average mean-squared error results in reproduction points on a circle about the source vector having the same distortion. Maximum mean-squared error causes reproduction points on a square about the source vector to have the same distortion.

The lower bound of (6a) is calculated as follows. Given a value for y, the distribution for X is Gaussian with mean ρy and variance $(1 - \rho^2)$. The conditioned distortion-rate function for each y is therefore [2, p. 99]

$$\delta_{X|y}(R) = (1 - \rho^2)e^{-2R},$$

and, from (3), we therefore have

$$\delta_{X|Y}(R) = (1 - \rho^2)e^{-2R}.$$

Since we also have

$$\delta_Y(R) = e^{-2R},$$

the lower bound is given by

$$\delta_{LB}(R) = \inf_{\alpha \in [0,1]} \max\left[(1-\rho^2)e^{-2\alpha R}, e^{-2(1-\alpha)R}\right].$$

The upper bound of (6b), $\delta'_U(R)$, is readily calculated [2, sect 4.5.2] by noting that the eigenvalues of the correlation matrix for (X,Y) are $\lambda_1 = 1-\rho$ and $\lambda_2 = 1+\rho$, and therefore we have the upper bound parametrically as

$$R_\theta = \sum_{k=1}^{2} \max\left(0, 0.5 \log\left(\frac{\lambda_k}{\theta}\right)\right)$$

$$D_\theta = \sum_{k=1}^{2} \min(\theta, \lambda_k)$$

as θ varies from 0 to $\lambda_{\max}$.

The upper bound of the corollary is given by

$$\delta_{UB}(R) = 2e^{-R}.$$

Fig. 2 shows the upper and lower bounds for $\rho = 0.5$.

It is seen from the above two examples that while the bounds are not tight, they still provide useful information. For more complex sources where the joint distortion rate function cannot be computed due to large alphabet sizes, the relationships at least provide simple upper and lower bounds to the function.

V. Joint Distortion-Rate Functions for Separate Users

The joint distortion-rate function discussed in Section III implies that the receiver is interested in both data streams at the same time, as we consider the maximum distortion on a per-letter basis. Another case of interest is the transmission of correlated data streams to be provided to separate users at the receiver. In this case, each user is interested in the quality of only the data stream provided him. We therefore are led again to a vector-valued distortion measure, giving the average distortion of reproduction provided to each user. The system designer, however, again needs a single number telling him the performance of his system.

The average value of distortion, averaged over the various users, is one possibility for a criterion. Again, as above, this is sometimes too weak as it does not guarantee each user a value of average distortion for his particular data stream. We can guarantee the user a value of average distortion by considering the maximum of the average distortions, weighted according to the importance of each data stream. We again have a well-defined distortion-rate function describing the performance of the optimum system as follows.

Consider the memoryless vector source U as described in Section II with its associated reproduction alphabet and component distortion measures $d_i(u_i;\hat{u}_i)$. The average distortion associated with a particular test channel $p(\hat{u}|u)$ is given by

$$D_i = Ed_i(U_i;\hat{U}_i). \tag{8}$$

The distortion-rate function of the source U having a criterion which minimizes the maximum of the average distortion is then given by

$$\Delta_U(R) = \inf_{p(\hat{u}|u):\ I(U;\hat{U}) \le R} \max_i Ea_i d_i(U_i;\hat{U}_i). \tag{9}$$

Choosing such a criterion guarantees us that the optimum system can provide each user with an average distortion level less than $\Delta_U(R)$, and no system can do better on all the data streams as can be seen from the following reasoning. Let $R_U(D)$ be the rate-distortion function for the vector source U subject to the vector constraint D, i.e., $Ed_i(U_i;\hat{U}_i) \le D_i$ for each i [1]. For the moment, assume that $a_i = 1$ for each i. Then the test channel that achieves $\Delta_U(R^*) = D^*$ satisfies the criterion in the definition for $R_U(D)$ for D having $D_i = D^*$ for each i, and thus $R_U(D) \le R^*$. The coding theorem for $R_U(D)$ then guarantees that there exists a code having rate less than $R^* + \epsilon$ and average distortion for each data stream less than D^*. The converse holds using similar reasoning, and the generalization to arbitrary a_i follows trivially.

The distortion-rate function, $\Delta_U(R)$, of the source U with the constraint that the maximum of the average values of the component distortion be minimized may be defined in terms of the rate-distortion function $R_U(D)$ of the source U subject to the vector constraint $Ed_i(U_i;\hat{U}_i) \le D_i$ as follows. Let $F(r)$ be a set of vector distortion points defined by

$$F(r) = \{D:\ \ R_U(D) \le r\}, \tag{10}$$

and let $B(r)$ be the lower boundary to $F(r)$, i.e.,

$$B(r) = \{D^* \in F(r):\ \ D \in F(r) \text{ implies } D_i \ge D_i^* \text{ for some } i\}. \tag{11}$$

We then see that, by the minimization involved,

$$\Delta_U(r) = \inf_{D^* \in B(r)} \max_i a_i D_i^*. \tag{12}$$

Thus in the sense described above, $\Delta_U(R)$ is an "inverse" function to the rate-distortion function $R_U(D)$, just as the distortion-rate function is the inverse to the rate-distortion function in the scalar case with a scalar distortion measure. While (12) appears complicated to calculate, we note that, if there is a $D^* \in B(r)$ having all equal components, i.e., $a_i D_i^* = D_0$ for each i, then by definition we see that $\Delta_U(r) = D_0$. Furthermore, if this is true, then $R_U(D^*) = r$, substantiating the interpretation of an inverse to the rate-distortion function.

Again, as in Section IV, it is desirable to bound the distortion-rate function $\Delta_U(R)$ by distortion-rate functions which are easier to calculate. As a first step, we note that considering the average of the maximum per-letter component distortion (as in Section III) is a stronger condition than considering the maximum of the average per-letter component distortion (as is done in this section), and we therefore have the relation

$$\delta_U(R) \ge \Delta_U(R), \tag{13}$$

where $\delta_U(R)$ is the distortion-rate function defined by (1) and (5). The upper bounds to $\delta_U(R)$ of Theorem 1 and its

corollary may then be immediately applied to obtain upper bounds to $\Delta_U(R)$.

A lower bound to $\Delta_U(R)$ may be derived for a two-dimensional source $U = (X,Y)$ as in Theorem 1, resulting in the same bound as (6a). We therefore have the following theorem.

Theorem 2: The distortion-rate function for the two-dimensional vector source $U = (X,Y)$ considering a criterion which minimizes the maximum of the average component distortions is related to the distortion-rate functions of its component sources by

$$\Delta_U(R) \geq \inf_{\alpha \in [0,1]} \max\,[a\delta_{X|Y}(\alpha R), b\delta_Y((1-\alpha)R)] \quad (14a)$$

$$\Delta_U(R) \leq \delta_U(R) \leq a\delta_X(\alpha R) + b\delta_Y((1-\alpha)R),\ \alpha \in [0,1]. \quad (14b)$$

ACKNOWLEDGMENT

The author would like to thank Prof. R. M. Gray of Stanford University for his suggestions which led to this research. Thanks are also due R. E. Thomas, Jr., of the Georgia Institute of Technology for his editorial suggestions.

APPENDIX

PROOFS

Theorem 1

Proof of (6a): Standard information-theoretic techniques based on $\ln x \leq x - 1$ yields $I(XY;\hat{X}\hat{Y}) \geq I(X;\hat{X}|Y) + I(Y;\hat{Y})$. Consider any test channel $p(\hat{x}\hat{y}|xy)$ having $I(XY;\hat{X}\hat{Y}) \leq R$. We have that

$$d(u;\hat{u}) = d(xy;\hat{x}\hat{y}) = \max\{ad_X(x;\hat{x}), bd_Y(y;\hat{y})\}$$

for each source reproduction pair, and therefore

$$Ed(U;\hat{U}) = E\max\{ad_X(X;\hat{X}), bd_Y(Y;\hat{Y})\} \geq \max\{aEd_X(X;\hat{X}), bEd_Y(Y;\hat{Y})\}.$$

Letting $R_1 = I(X;\hat{X}|Y)$, $R_2 = I(Y;\hat{Y})$, and minimizing both sides of the equation gives

$$\delta_U(R) \geq \max\{a\delta_{X|Y}(R_1), b\delta_Y(R_2)\}.$$

Since $\delta_{X|Y}(R)$ and $\delta_Y(R)$ are monotonically decreasing, we then have

$$\delta_U(R) \geq \max\{a\delta_{X|Y}(\alpha R), b\delta_Y((1-\alpha)R)\}$$

for some $\alpha \in [0,1]$ proving (6a).

Proof of (6b): Let $p(\hat{x}\hat{y}|xy)$ achieve $\delta'_U(R)$. Then, since $d(u;\hat{u}) = \max\,[ad_X(x;\hat{x}) + bd_Y(y;\hat{y})] \leq ad_X(x;\hat{x}) + bd_Y(y;\hat{y})$, we have

$$\delta_U(R) \leq E[d(U;\hat{U})] \leq \delta'_U(R),$$

proving the left inequality. The right inequality follows from Gray [1, p. 482].

Proof of (6c): Let $\delta^a_X(R)$ and $\delta^a_Y(R)$ be the distortion-rate functions for the sources X and Y and fidelity criteria $(d_X(x;\hat{x}))^\sigma$ and $(d_Y(y;\hat{y}))^\sigma$, respectively. Let $p(\hat{x}|x)$ and $p(\hat{y}|y)$ be test channels such that $I(\hat{X};X) \leq \alpha R$ and $I(\hat{Y};Y) \leq (1-\alpha)R$, and consider $p(\hat{x}\hat{y}|xy) = p(\hat{x}|x)p(\hat{y}|y)$. Again we have that

$$I(XY;\hat{X}\hat{Y}) \leq I(X;\hat{X}) + I(Y;\hat{Y}).$$

Also

$$\begin{aligned}(E[\max\{ad_X(X;\hat{X}), bd_Y(Y;\hat{Y})\}])^\sigma &\leq E[(\max\{ad_X(\hat{X};\hat{X}), bd_Y(Y;\hat{Y})\})^\sigma] \\ &\leq E[a^\sigma(d_X(X;\hat{X}))^\sigma + b^\sigma(d_Y(Y;\hat{Y}))^\sigma],\end{aligned}$$

where the first inequality follows from Gallager [8, p. 523, prob. 4.15d] and the second from the fact that, for any two positive numbers e and f and a $\sigma \geq 1$,

$$e^\sigma + f^\sigma \geq (\max\{e,f\})^\sigma.$$

Since $\delta_U(R)$ is monotonically decreasing and since raising both sides of the equation to the $1/\sigma$ power preserves the order of the inequality, we have that

$$\begin{aligned}\delta_U(R) &\leq E[\max\{ad_X(X;\hat{X}), bd_Y(Y;\hat{Y})\}] \\ &\leq (E[a^\sigma(d_X(X;\hat{X}))^\sigma + b^\sigma(d_Y(Y;\hat{Y}))^\sigma])^{1/\sigma} \\ &= (a^\sigma E[(d_X(X;\hat{X}))^\sigma] + b^\sigma E[(d_Y(Y;\hat{Y}))^\sigma])^{1/\sigma}.\end{aligned}$$

Minimizing the right side yields the desired result.

Corollary: Equations (7a) and (7b) follow immediately from Theorem 1 upon making the appropriate minimization and substitution.

Theorem 2

Proof of (14a): Consider any test channel $p(\hat{x}\hat{y}|xy)$ having $I(XY;\hat{X}\hat{Y}) \leq R$. Letting $R_1 = I(X;\hat{X}|Y)$ and $R_2 = I(Y;\hat{Y})$, we have

$$\max\{aEd_X(X;\hat{X}), bEd_Y(Y;\hat{Y})\} \geq \max\{a\delta_{X|Y}(R_1), b\delta_Y(R_2)\}.$$

Minimizing both sides of the equation as in the proof of (6a) yields the desired result.

Equation (14b) follows from (13) and (6b).

REFERENCES

[1] R. M. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 480–489, July 1973.
[2] M. B. Sachs *et al.*, "Spatial frequency channels in human vision," *J. Opt. Soc. Amer.*, vol. 61, pp. 1176–1186, Sept. 1971.
[3] C. F. Stromeyer, III, and B. Julesz, "Spatial-frequency masking in vision: Critical bands and spread of masking," *J. Opt. Soc. Amer.*, vol. 62, pp. 1221–1232, Oct. 1972.
[4] L. Harmon, "Masking in visual recognition: Effects of two-dimensional filtered noise," *Science*, vol. 180, pp. 1194–1197, June 1973.
[5] R. M. Gray and L. D. Davisson, "A mathematical theory of data compression," to be published.
[6] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
[7] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[8] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *1959 IRE Nat. Conv. Rec.*, pt. 4, pp. 142–163.
[9] R. M. Gray, "Conditional rate-distortion theory," Technical Report No. 6502-2, Information Systems Lab., Stanford University, Oct. 1972.
[10] B. M. Leiner, "Rate-distortion theory for sources with side information," Ph.D. dissertation, Stanford University, Aug. 1973.
[11] ——, "Joint, conditional, and marginal distortion-rate functions," in *Proc. Seventh Hawaii Int. Conf. on System Sciences*, pp. 40–42, Jan. 1974.
[12] R. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 460–473, July 1972.

APPENDIX C

This appendix consists of a reprint of the paper:

T. P. Barnwell, III and R. M. Mersereau, "A Comparison of Some Subjective and Objective Measures for Image Quality," Eleventh Annual Asilomar Conf. on Circuits Systems, and Computers, 1977.

A COMPARISON OF SOME SUBJECTIVE AND OBJECTIVE MEASURES FOR IMAGE QUALITY*

Thomas P. Barnwell, III
Russell M. Mersereau
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia, 30332

In the design of image processors or image coders which operate upon monochrome still images that will ultimately be viewed by human observers, it is important to have a numerical measure of image quality or distortion. Such a quality measure must be readily computable from the original and distorted images and it should correlate well with the results of subjective tests. In this study a comparison is made between several objective distortion measures and the results of subjective tests for several different classes of distortions. None of the objective measures perform as well as we might like.

I. Introduction

Finding a well-defined, objective distortion measure for images which is highly correlated with the results of subjective image quality tests is a difficult task due to the complexity of the visual perception process. Such a measure would be important in the design of image processors or image coders where the final image is presented to human observers. The purpose of this study was to compare several measures for image distortion with the results of a subjective image quality test. The distortion measures considered were motivated by earlier work by Mannos and Sakrison [1] and Leiner [2], but the present work represents an entirely different approach toward measuring the effectiveness of the distortion measures.

The distortion measures were compared by estimating the correlation coefficient between the results of the subjective test and the predictions of the distortion measures. The minimum variance estimate of the correlation coefficient for a particular distortion measure is given by

$$\hat{\rho} = \frac{1}{PK-2} \frac{\sum_{p=1}^{P} \sum_{k=1}^{K} (S_{pk}-\bar{S})(O_{pk}-\bar{O})}{\hat{\sigma}_O \hat{\sigma}_S} \tag{1}$$

where

$$\hat{\sigma}_S^2 = \frac{1}{PK-1} \sum_{p=1}^{P} \sum_{k=1}^{K} (S_{pk}-\bar{S})^2 \tag{2}$$

and

$$\hat{\sigma}_O^2 = \frac{1}{PK-1} \sum_{p=1}^{P} \sum_{k=1}^{K} (O_{pk}-\bar{O})^2 \tag{3}$$

In these expressions P represents the number of distortions considered, K represents the number of different images used, and S_{pk} and O_{pk} represent respectively the average subjective response and the objective measure for the pth distortion applied to the kth image. $\bar{S}$ and $\bar{O}$ are the average subjective and objective measures.

*This work was supported in part by the National Science Foundation under Grant ENG75-04992.

II. The Distortions Used

The test set for this study consisted of the two (256 x 256) sampled images which are shown in Figure 1. To each of these images 120 distortions were applied. These distortions could be subdivided into eight classes - two additive white noises (uniform and Gaussian), two multiplicative noises, 3 additive bandpass noises, and a lowpass filtering blur. In each class fifteen levels of distortion were used which ranged from "barely perceptable" to heavily distorted (but still recognizable)".

III. The Subjective Test

A doubly-anchored isometric quality preference test was chosen as the subjective test in this study. For each distortion, a slide was produced containing three images arranged as shown in Figure 2 - a "high anchor," which was the original 256 x 256 picture; a "low anchor," which was a combination of distortions which had been prejudged to be worse than, but comparable to, the worst distortion in the test; and the distorted test picture. The slides were taken from the screen of a CRT controlled by a Comtal Image processing system with 512 x 512 point resolution and 256 gray levels. The intensities were mapped such that the log energy vs. film density plot had a slope of -1.

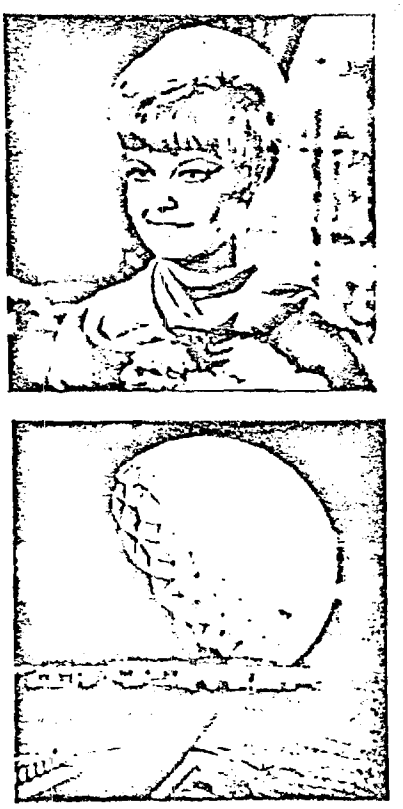


Figure 1

The two test images used for this study.

			1	2	3	4	5	6	7	8	9	10	13	11	12	14
1	24.9	69.5	-													
2	21.5	61.2	1	-												
3	19.1	54.2	1	5	-											
4	17.3	50.7	1	1		-										
5	15.7	50.2	1	1			-									
6	14.4	47.1	1	1				-								
7	13.3	44.4	1	1	5				-							
8	12.3	39.0	1	1	1	1	1	5		-						
9	11.4	38.8	1	1	1	1	1	5			-					
10	10.5	38.5	1	1	1	1	1	5				-				
13	9.80	34.0	1	1	1	1	1	1	1				-			
11	9.12	33.8	1	1	1	1	1	1	1	5	5			-		
12	8.48	30.3	1	1	1	1	1	1	1	5	5	5			-	
14	7.89	29.6	1	1	1	1	1	1	1	5	5	5				-
15	7.33	29.2	1	1	1	1	1	1	1	5	5	5				

Table I: Results of the subjective quality test for additive white Gaussian noise. If a "1" appears at the intersection of two distortion levels this means that the difference in their quality scores is significant at the .01 level. Similarly a "5" means that the difference is significant at the .05 level.

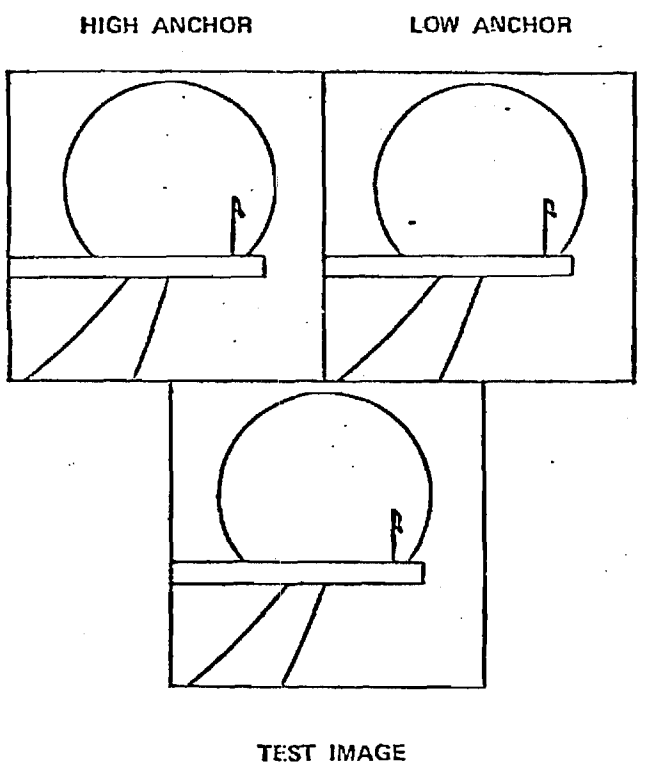


Figure 2: The format of the doubly-anchored subjective test. In the upper left is the original undistorted image ("high anchor"), in the upper right is a badly distorted image ("low anchor") and below is the image under test. All three images were made from the same undistorted image.

In the test, subjects were seated so that the pictures subtended an angle of approximately six degrees. The subjects were asked to rate each distorted picture on a scale of 0 to 100, and were told that the high anchor should get a score of 80 while the low anchor should be scored as 20. For each of the two pictures, the distortions were randomized and presented at 15 second intervals in groups of 120. In all, 20 subjects participated in the test for each of the two pictures.

A Newman-Keul [3] test was applied to the subjective results. In this statistical analysis, the average results for the different distorted images are first ranked, then the Studentized Range Test is applied to these results in pairs to determine whether the differences in ratings are significant. A sample set of results for one class of distortions is shown in Table 1.

The results of this subjective test can be summarized as follows:

1. The test consistently gave significant differences in perceived quality even for distortion levels which were close (2 or 3 levels apart).

2. The corresponding results for the two different original images were very similar.

3. The standard error, given by $\hat{\sigma}/\sqrt{M}$, where $\hat{\sigma}$ is the sample standard deviation and M is the number of subjects, ranged from 2.9 to 1.25, corresponding to an average resolving power of about 4 points at the .01 level.

4. The lowpass filtering distortion on the girl picture showed almost no distortion until the band limit became less than about 10 cycles per degree.

5. All of the other distortions showed fairly linear behavior with distortion level.

BER OF BANDS, N	CORRELATION
4	.34388
5	.46665
6	.50093
7	.50575
8	.50645
9	.52730
10	.52731
11	.53454
12	.55033
13	.54531
14	.54609
15	.55410
16	.55903
17	.56424
18	.56488
19	.57054
20	.57969
21	.57937
22	.57561
23	.58836
24	.59033
25	.58899
26	.58981
30	.60153

Table II: The correlation obtainable with a distortion measure which forms the best linear estimate from the outputs of N bandpass filters vs. N. The filters differed in both bandwidth and center frequency.

IV. Objective Distortion Measures

All of the distortion measures considered in this study can be understood by reference to Fig. 3. The original and distorted images are passed through cube root nonlinearities and then an error sequence e(m,n) was computed. This error was then filtered by a bank of 30 two-dimensional band-pass filters and the energy in each band was computed. The filters in the bank were circularly symmetric and each had a bandwidth of approximately .75 cycles per degree.

The distortion measures differed in the manner by which the image quality was estimated from the filter band outputs. Three different algorithms were used. The first was motivated by the work of Mannos and Sakrison [1]. They filtered their error with a filter whose frequency response is

$$A(f) = [c + (f/f_o)^{k_1}]\exp[-(f/f_o)^{k_2}] \qquad (4)$$

where f is the radial frequency variable. This function can be closely approximated by a series of weighting functions applied to the filter bank outputs.

A second distortion measure, suggested by Leiner [2] calls for weighting the error energy from disjoint bands and using as a distortion measure the maximum of those weighted errors. One question to be addressed is whether a weighting function and band selection can be found such that this distortion measure is more highly correlated with the subjective results than the Mannos-Sakrison measure.

The third distortion measure was a linear combination of the filter bank error energies where the weighting coefficients were selected to minimize the mean squared error between the estimate and the average subjective response. This final distortion measure is of interest for two reasons. First it represents the best linear estimate for the subjective results which can be obtained from the energy measurements of Figure 3. Hence it represents a bound on the expected performance of any "linear" distortion measure. (It is not truly linear, of course, due to the initial nonlinear processing of the images.) Secondly, by combining adjacent bands into larger bands, and again finding the optimum linear predictor, information can be gained on how many bands are necessary for good correlation.

V. The Experimental Study

In the experimental study the same 240 images used in the subjective quality tests were processed using the system of Figure 3. For each distortion, the energy in the individual bands was computed and stored for later analysis. This processing consumed more than 100 hours of computer time on the Nova 830 computer in the Georgia Tech digital signal processing facility.

Four basic experiments were performed. First, the Mannos and Sakrison measure was evaluated with the parameters specified by them [1], (f_o = 8 cycles/degree, c = .019, k_1 = 1, k_2 = 1.1) and the correlation coefficient of (1) was estimated. Second, an automated iterative technique was used to find values for the Mannos-Sakrison parameters which had a higher correlation than their original values. Third, systematic and random groupings of the bands were made and optimal linear fits were made on these groupings. Finally, the Leiner metric was applied to individual bands and groups of bands.

VI. Results

The results of the comparison betweem the objective and subjective distortion measurements are summarized below.

1. The correlation coefficient between the average subjective response and the mean squared error between the original and distorted images was .174. This result was obtained by averaging the results from all 120 distortions for both images.

2. The mean squared error between the cube roots of the original and distorted images produced a correlation coefficient of .245.

3. Using the objective measure of Mannos and Sakrison of eq. (4) with their parameter values (f_0 = 8 cy./deg., C = .019, k_1 = 1.0, k_2 = 1.1) the average correlation was only .115. By assuming a weighting function of the same functional form as (4) but selecting the parameters to maximize the correlation between the objective and subjective measures a correlation of .243 was obtained for the parameter set f_0 = 6 cy./deg., C = .266, k_1 = 11.31, k_2 = 3.68. The normalized weighting curves for the original and perturbed Mannos and Sakrison parameters are shown in Fig. 4.

4. Using Leiner's measure with thirty filters and the original Mannos and Sakrison weighting function

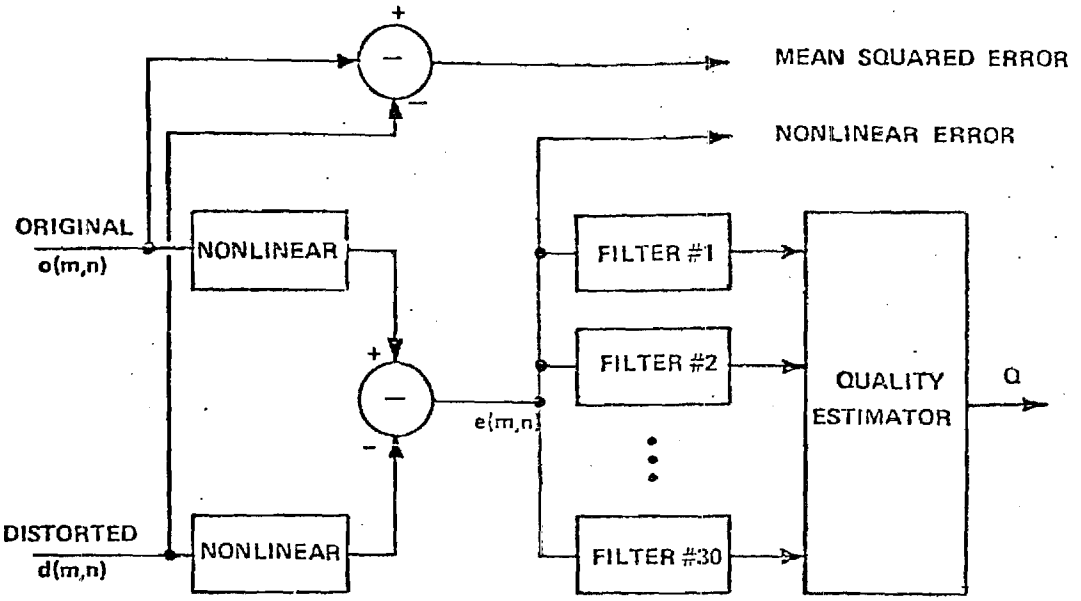


Figure 3: A block diagram which includes several image quality measures. The nonlinear operator was a cube root, and the filters in the bank were circular bandpasses (disjoint in frequency). The estimator either took a weighted sum of its inputs or selected the maximum of the weighted inputs.

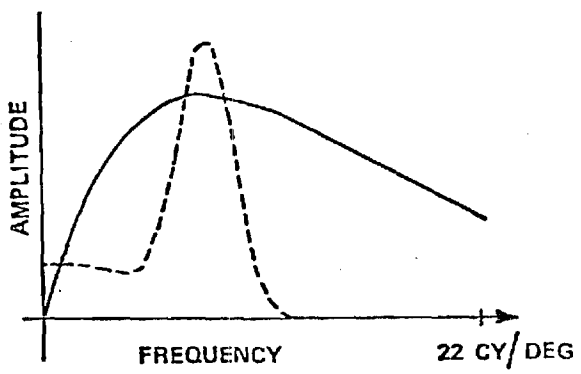


Figure 4: The amplitude weighting functions for the Mannos and Sakrison quality measure. The solid line represents their original function and the dashed curve has been optimized to maximize the correlation with subjective results.

a correlation of $+.30$ was obtained. Although again not a high value this result was felt to be encouraging since with optimization it can hopefully be made higher. No effort has been made to optimize the set of filter bands used or to optimize the weighting coefficients. The later problem can hopefully be solved as a latter program.

5. The final measure used formed a best linear estimate of image quality using as inputs the outputs of the filters in the filter bank (refer to Fig. 3). In this fashion a correlation of .602 can be obtained from the thirty filters used. If fewer filters are used the correlation will decrease but even with as few as six filters a correlation of .50 is possible. These results are summarized in Table 2. Once again these results were obtained by averaging over two pictures.

A possible reason for the discrepancy between our results and those of Mannos and Sakrison is that their tests were performed over a total bandwidth of 44 cycles/degree while ours were performed over a bandwidth of 22 cycles/degree, and due to the nature of the coding distortion used by Mannos and Sakrison, their tests had noise energy in the higher frequencies.

VII. Conclusions

Although it can certainly be argued that an ensemble of two images is a small one and that the distortions used were limited, the fact remains that a good and robust measure should have produced high correlations. The subjective results showed statistically significant measures of quality even for quality scores which were reasonably close together (4 points out of

.) Thus the performance of all of these measures is disappointing and the important problem of finding an objective quality measure which correlates well with human perceptual results remains unsolved.

References

[1] J. L. Mannos and D. J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Trans. Info. Theory, IT-20, pp. 525-536, 1974.

[2] B. M. Leiner, "Distortion-rate Functions for Vector Sources and a Maximum Fidelity Criterion," IEEE Trans. Info. Theory, IT-23, pp. 354-359, 1977.

[3] S. S. Wilks, Mathematical Statistics, J. Wiley and Sons, New York, 1962.

Acknowledgement

We are indebted to the efforts of one of our graduate students, James Klaasen, who developed the procedure for obtaining calibrated images from our display device.

APPENDIX D

THE PICTURE DISTORTIONS

DISTORTION #1

ADDITIVE WHITE UNIFORM NOISE

$$U(m,n) + N(m,n) = C_k \qquad (\bar{N}=0,\ \sigma_N = 1/\sqrt{12})$$

DISTORTION LEVEL	C_k	SNR
1	8	26.6
2	11.75	23.3
3	14.4	21.6
4	19.25	19.0
5	23	17.5
6	26.75	16.2
7	30.5	15.0
8	34.25	14.0
9	38	13.1
10	41.75	12.3
11	45.5	11.6
12	49.25	10.87
13	53	10.24
14	56.75	9.64
15	60.5	9.09

DISTORTION #2

ADDITIVE WHITE GAUSSIAN NOISE

$$U(m,n) + G(m,n) \cdot C_k \qquad (\bar{G}=0, \quad \sigma_G = 1)$$

DISTORTION LEVEL	C_k	SNR
1	4	24.9
2	6.125	21.2
3	8.25	19.1
4	10.375	17.3
5	12.5	15.7
6	14.625	14.4
7	16.75	13.3
8	18.875	12.3
9	21	11.4
10	23.125	10.5
11	25.25	9.80
12	27.375	9.12
13	29.5	8.48
14	31.625	7.89
15	33.75	7.33

DISTORTION #3

MULTIPLICATIVE UNIFORM NOISE

$$U(m,n)[1 + C_k N(m,n)] \quad (\bar{N}=0 \ \sigma_N = 1/\sqrt{12})$$

DISTORTION LEVEL	C_k	$kg(\hat{\sigma}_U)/(\log[HC_k]-\log[1-C_k])$
1	.061	81.0
2	.183	26.9
3	.305	16.1
4	.427	11.4
5	.549	8.8
6	.671	7.1
7	.793	5.9
8	.916	5.0
9	1.038	4.3
10	1.160	3.7
11	1.282	3.3
12	1.404	2.8
13	1.526	2.5
14	1.648	2.1
15	1.770	1.8

DISTORTION #4

MULTIPLICATIVE WHITE GAUSSIAN NOISE

$$EXP[\ell nU(m,n) + G(m,n)\cdot C_k] \quad (\bar{G}=0,\ \sigma_G=1)$$

DISTORTION LEVEL	C_k	$\log(\hat{\sigma}_U)/2C_k\log(e)$
1	.04	61.8
2	.12	20.6
3	.20	12.4
4	.28	8.8
5	.36	6.9
6	.44	5.6
7	.52	4.8
8	.60	4.1
9	.68	3.6
10	.76	3.3
11	.84	2.9
12	.92	2.7
13	1.0	2.5
14	1.08	2.3
15	1.16	2.1

DISTORTION #5

ADDITIVE COLORED GAUSSIAN NOISE

$$U(m,n) + BPF_1(G(m,n)) \cdot C_k \quad (\bar{G}=0,\ \sigma_G=1)$$

(BPF = 0 - 2.75)

DISTORTION LEVEL	C_k	SNR
1	24.8	39.8
2	38.0	36.1
3	51.3	33.5
4	64.5	31.5
5	77.7	29.9
6	90.9	28.5
7	104	27.3
8	117	26.3
9	130	25.3
10	143	24.5
11	157	23.7
12	170	23.0
13	183	22.4
14	196	21.8
15	209	21.2

DISTORTION #6

ADDITIVE COLORED GAUSSIAN NOISE

$U(m,n) + BPF_2(G(m,n)) \cdot C_k$ $(\bar{G}=0,\ \sigma_G=1)$

(BPF=2.75 - 5.5)

DISTORTION LEVEL	C_k	SNR
1	72.7	20.9
2	111	17.2
3	150	14.6
4	188	12.6
5	277	11.0
6	265	9.6
7	304	8.5
8	343	7.4
9	381	6.5
10	420	5.7
11	459	4.9
12	497	4.2
13	536	3.5
14	575	2.9
15	613	2.4

DISTORTION #7

ADDITIVE COLORED GAUSSIAN NOISE

$$U(m,n) + BPF_3(G(m,n)) \cdot C_k \qquad (\bar{G}=0,\ \sigma_G=1)$$

(BPF=5.5 - 11.0)

DISTORTION LEVEL	C_k	SNR
1	7.7	27.6
2	11.9	23.9
3	16.0	21.3
4	20.2	19.3
5	24.3	17.7
6	28.4	16.3
7	32.6	15.1
8	36.7	14.1
9	40.9	13.2
10	45.0	12.3
11	49.1	11.6
12	53.3	10.9
13	57.4	10.2
14	61.6	9.7
15	65.7	9.0

DISTORTION #8

LOW PASS FILTER

DISTORTION LEVEL	BANDLIMIT (cycles/degree)
1	21
2	20
3	19
4	18
5	17
6	16
7	15
8	14
9	13
10	12
11	11
12	10
13	9
14	8
15	7

APPENDIX E

IMAGE REPRODUCTION PROCEDURES

This appendix contains two parts. The first part describes the procedure used to determine the mapping function between intensity values stored in DIGITAL IMAGE REPRESENTATIONS and the (different) intensity values necessary to cause the black and white slides to give correct intensities when viewed by subjects. The second part is a copy of a document entitled "SET-UP PROCEDURE FOR THE COMTAL MONITOR," which describes the procedure used to make the pictures used in the subjective tests from the COMTAL IMAGE PROCESSING SYSTEM.

E.1 The Film Density Correction Procedure

The following procedure was used to determine the correct intensity correction function for the COMTAL IMAGE PROCESSING SYSTEM.

(a) The CRT controls (brightness and contrast) were set so that all intensities were recordable on film; neither end of the scale was in the saturation region of the CRT, and no blooming occurred at high intensity levels. The final values for the control settings were determined iteratively using test pictures consisting of 9 squares (3 x 3) of fixed intensities.

(b) Once an acceptable setting of the controls was found, the following method was used to recreate the same control settings at a later time:

i. Completely darken the room.

ii. Place a constant 512 x 512 image on the screen with all image values set to 255 (the highest intensity).

iii. Record the intensity at the center of the screen (a United Detector Technology, Inc. 40X Opto-Meter is used).

iv. Turn the contrast full counter clockwise, and record the intensity at the center of the screen again.

To recreate the settings later, do the following steps:

i. Completely darken the room.

ii. Place a constant 512 x 512 image on the screen with all image intensities set to 255.

iii. With the contrast set full counter clockwise, turn the brightness up until the light intensity reading from (ii) above is obtained at the center of the screen.

iv. Turn the contrast up until the intensity from (iii) above is obtained at the center of the screen.

(c) Once an acceptable value for contrast and brightness were found, a group of 36 pictures of the CRT screen for images of constant intensity were taken. The intensities were chosen such that the steps in log(I), where I is the intensity, were approximately constant. The film was developed, and a densitometer was used to measure the film density's corresponding to the 36 intensity values. A density vs. LOG(I) function for all 256 intensity values was then approximated using an 8th order polynomial fit. From this function, correction factors were calculated for each of the 256 possible intensities so that the density vs log(I) curve was linear with a scope of -1.

E.2 Set-up Procedure for Comtal Monitor

The procedure given will describe all of the steps necessary to make calibrated slides from the Comtal monitor. The procedure is divided into three parts: wiring, exposure, and processing.

1. Wiring

Connections are made between the Nova and the Comtal memory in the Nova Lab and the Comtal display, camera and terminal in the Optical Information Processing Lab through the wall panels provided for this purpose. In the Nova Lab this panel is located on the west wall about 8 feet from the south wall. At this box the cable connector labeled "Camera Control Pins 2 + 3" should be connected to the box connector labeled "Optics Lab (2,3)" in the right column of connectors. If it is desired to use the terminal in the Optics Lab, an extension must be run from the mini-box on the back of the Nova, or the cable to the other terminal, to the bottom connector marked "Optics" in the left column of the flat connectors in the wall box. Finally, check to see that three coaxial connectors are connected to channels 1, 2, 3 of the Optics Lab column of coaxial connectors. The other ends of these cables should be located on the floor behind the Comtal memory. They should have double male adaptors on them. These are connected to the coaxial cables coming from the back of the Comtal memory being careful to match the labels. Check to see that the remote control cables are unplugged from both tape decks. This completes the necessary wiring in the Nova Lab.

In the Optics Lab the wall box is located on the east wall about ten feet from the north wall. The camera control cable should be connected to the third flat connector from the top in the wall box and the other end plugged into the side of the motor drive on the camera. The first three coaxial channels should be connected to the three coaxial extension cables, the other ends of which are connected to the Comtal monitor at the connector identified by the tag on the cable.

If you are using the CRT terminal in the Optics Lab, it must be connected to the Nova by plugging the connector labeled "Nova Lab" into the back of the terminal at the plug labeled "Modem." The proper baud rate must be selected at the front right of the terminal. The correct rates are 9600 Baud at the mini-box in the Nova Lab or 1200 Baud at the cable to the other terminal. Other switches should be set at PAR-ODD, HDX. The terminal should now be prepared for the Nova. At the end of the picture session, return all connections and switches to their previous positions.

2. Exposure

The following procedures must be carried out very carefully to ensure that the film is exposed properly. The camera must be set up so that the film plane is parallel to the plane of the Comtal monitor screen. I have accomplished this in the past by using the floor tiles as a guide. The height of the camera can be adjusted by centering the monitor in the viewfinder when the camera is far from the monitor, then bring the camera up to the monitor and adjust the height so that the lens is pointing at the center of the screen. Now the camera must be set up 2 yards 27 inches ± ½ inch in front of the monitor screen measured from the front edge of the

monitor cabinet to the front edge of the camera lens. A plumb line may be useful in making this measurement. This measurement is made with the camera focused on the monitor since focusing the lens changes its length.

The Comtal monitor controls are set as follows: On the rear of the monitor the EXT-INT switch should be set on EXT, and the HI-75Ω switch on 75Ω. On the front of the monitor, the HEIGHT should be set for a square picture, the FOCUS for the sharpest raster, and the H + V HOLD controls midway in the range where the picture is locked in sync (not critical). The setting of the BRIGHT and CONTRAST controls is very critical and should be done only after the monitor has been on for about 20 minutes. The following adjustments must be made with the Function Memory OFF. To be sure it is off proceed to the Comtal Memory Control Panel and switch the REMOTE switch off (light off). Then switch white toggle switch number 14 "Funct. Mem Enable" off (down) and press the red SYSTEM RESET and then TRANSFER switches. Return the Comtal to the remote mode by depressing the REMOTE switch (light on). Now run the program COM9GRAY which is located in directory SAVE. It will ask you for nine numbers one at a time. Enter 255 for each number. After it asks for N1 the second time, use CTL A to stop the program. Proceed with the following adjustments in a dark room making the measurements with the United Detector Technology, Inc. 40X Opto-Meter located in the Optical Information Processing Lab. Use the Radiometric Filter with the meter. A spotlight located in the lab may be used to illuminate the meter scale, but be sure to shade the Comtal screen from the small amount of light that is radiated outside of the center spot area. With CONTRAST set fully CCW slowly adjust BRIGHT to obtain a reading of 0.014μW at the center of the screen. Then adjust CONTRAST to obtain a reading of 22.5μW at the center of the screen. The display is now set up

properly. The Function Memory should now be turned on by proceeding to the Comtal Memory Control Panel and switching the REMOTE switch off. Now switch the white toggle switch number 14 "Funct Mem Enable" on (up) and press the red SYSTEM RESET and then TRANSFER switches. Return the Comtal to the remote mode. Now load the Function Memory with the correction data by running the following line from directory SAVE: FUNC FUNC3/I. It is now necessary to load the good and bad reference pictures. The program that accomplishes this is called PICDIR. It and all of the following picture handling programs are located on the disc labeled "2D Filtering & Image Processing MERSEREAU" in the directory IMCODE. The format for running PICDIR is as follows:

PICDIR (GOOD,BAD)/I (1,2)/N

where GOOD is the file name of the good reference picture and BAD is the file name of the bad reference picture.

You should now set the camera aperture to f5.6 and the shutter speed to T. Film is loaded into cassettes by means of the bulk loader in the Optics Lab. It is necessary to load 38 frames to take 30 pictures to allow for leader and a tail. After the camera is loaded be sure to set the frame counter on the motor drive to 36 so that the motor drive will advance through all of the frames. Now if the camera is focused and the film is advanced to frame 1, you are ready to take a series of pictures.

The program PIC60 in directory SAVE can now be run. This program will take the first thirty pictures on a tape and pause for you to reload the camera. Then thirty more pictures will be taken. After running PIC60 it will be necessary to type "RELEASE MTØ" which will rewind the tape. A new tape may then be loaded and more pictures taken.

If individual frames must be taken, the following procedure may be used. First load the tape. Type "INIT MTØ" at your terminal. Then type "PICIN/B MTØ:X" where X denotes the position of the frame on the tape. For example, picture number 3 of the second run would correspond to X=17, if there are 15 pictures per run. Exposures made in this manner must be timed manually. The proper exposure length is 16 seconds.

3. Processing

The exposed film is developed in D-76 stock solution for twelve minutes at a temperature of 68°F. The rest of the procedure is as recommended by Kodak. The processing chemicals and equipment are located in the darkroom in a wooden cabinet next to the west wall. The developer is used once and then disposed of, other chemicals may be reused until they are no longer effective.

The result of the procedure should be rolls of positive slides which are cut and mounted in GEPE TV slide binders (21 x 28 mm) which are available from Crown Camera Exchange. When mounting the slides, use care to keep them free of fingerprints and dust. Dust-Off is helpful in removing dust and is available in the Optics Lab. When mounting the slides, center the picture area inside the mask so that an equal amount of excess picture is removed from each edge. The mounts are assembled with the cover rotated 90° from normal so that the masking produces a square image.

APPENDIX F

COMPUTER ANALYSIS OF THE RESULTS OF THE SUBSECTIVE TESTS

RESULTS OF THE STATISTICAL ANALYSIS

FOR THE "GIRL" IMAGE

NUMBER OF SUBJECTS= 19
AVERAGE=55.632
SAMPLE STANDART DEVIATION=17.769
STANDARD DEVIATION DUE TO DISTORTION=64.555
STANDARD DEVIATION DUE TO SUBJECTS=10.175

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
1	1	68.684	71.784	8.473	-5555555555555
1	2	64.737	106.871	10.338	5555555555555
1	3	55.789	39.620	6.294	--1555555555
1	4	54.737	98.538	9.927	-1555555555
1	5	49.737	140.205	11.841	---5555555
1	6	46.579	133.480	11.553	----55555
1	7	43.842	123.807	11.127	----1555
1	9	41.842	161.696	12.716	----555
1	8	38.947	115.497	10.747	----11
1	10	38.421	216.813	14.725	---11
1	13	35.263	81.871	9.048	----
1	11	33.684	91.228	9.551	---
1	12	30.000	52.778	7.265	--
1	14	28.947	40.497	6.364	-
1	15	27.737	68.205	8.137	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
2	1	75.684	27.784	5.271	---55555555555
2	2	72.895	67.544	8.219	--15555555555
2	3	71.684	47.228	6.872	-15555555555
2	4	67.474	50.152	7.082	--155555555
2	5	63.316	129.450	11.378	----555555
2	6	60.526	60.819	7.799	----55555
2	7	57.368	87.133	9.333	---15555
2	8	56.579	86.257	9.287	--15555
2	9	55.000	86.111	9.280	-15555
2	10	51.842	86.696	9.311	-1155
2	11	46.842	83.918	9.161	----
2	13	43.684	91.228	9.551	---
2	15	43.158	89.474	9.459	--
2	12	40.263	90.205	9.498	-
2	14	39.789	170.509	13.058	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
3	1	74.000	95.778	9.787	55555555555555
3	2	64.737	104.094	10.203	1155555555555
3	3	57.895	81.433	9.024	-11555555555
3	4	56.579	94.591	9.726	15555555555
3	5	50.000	86.111	9.280	---5555555
3	6	48.947	171.053	13.079	--5555555
3	7	43.684	121.784	11.036	-----555
3	8	43.526	141.819	11.909	-155555
3	10	37.368	89.912	9.482	------
3	9	35.789	86.842	9.319	-----
3	11	35.789	106.287	10.310	----
3	12	35.526	121.930	11.042	---
3	15	32.105	78.655	8.869	--
3	14	32.000	118.222	10.873	-
3	13	31.579	89.035	9.436	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
4	1	76.211	43.398	6.588	----5555555555
4	2	74.737	134.649	11.604	---5555555555
4	3	72.632	23.246	4.821	--1155555555
4	4	70.263	176.316	13.278	---55555555
4	5	69.632	65.912	8.119	--55555555
4	6	63.421	47.368	6.882	-----5555
4	7	62.105	45.322	6.732	----5555
4	8	57.368	109.357	10.457	----555
4	10	56.053	51.608	7.184	---115
4	9	55.000	75.000	8.660	--115
4	11	54.474	71.930	8.481	-115
4	12	50.526	74.708	8.643	---
4	14	45.895	202.766	14.240	--
4	13	45.000	83.333	9.129	-
4	15	42.368	112.135	10.589	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
5	1	78.053	19.830	4.453	----------15
5	5	77.737	43.982	6.632	----------15
5	4	77.368	37.135	6.094	----------15
5	6	77.211	52.620	7.254	---------15
5	2	76.316	88.450	9.405	---------5
5	3	75.421	118.368	10.880	--------5
5	7	74.211	59.064	7.685	-------5
5	9	71.842	145.029	12.043	------1
5	11	70.526	83.041	9.113	-----1
5	14	68.947	168.275	12.972	-----
5	8	68.684	77.339	8.794	----
5	10	68.421	86.257	9.287	---
5	12	67.263	181.205	13.461	--
5	13	66.316	119.006	10.909	-
5	15	60.789	209.064	14.459	

SYSTM	DIST	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
6	1	73.421	61.257	7.827	--155555555555
6	2	70.789	61.842	7.864	--15555555555
6	3	66.316	144.006	12.000	--5555555555
6	4	64.211	92.098	9.612	-1155555555
6	5	61.842	128.363	11.330	1555555555
6	7	55.000	108.333	10.408	------115
6	8	54.474	127.485	11.291	-----115
6	6	53.158	145.029	12.043	------5
6	9	51.579	147.368	12.140	-----1
6	10	48.947	148.830	12.200	-----
6	11	48.158	125.585	11.206	----
6	13	45.789	134.064	11.579	---
6	12	44.737	215.205	14.670	--
6	14	44.211	172.953	13.151	-
6	15	40.789	145.175	12.049	

SYSTM	DIST	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
7	1	71.579	52.924	7.275	15555555555555
7	3	65.000	66.667	8.165	--11555555555
7	2	64.211	147.953	12.164	-15555555555
7	5	58.158	108.918	10.436	---55555555
7	4	56.053	84.942	9.216	--11111555
7	6	55.789	134.064	11.579	-11111555
7	7	51.211	122.287	11.058	-----111
7	9	47.105	106.433	10.317	-------
7	8	46.579	155.702	12.478	------
7	10	46.579	175.146	13.234	-----
7	12	46.053	201.608	14.199	----
7	11	45.263	204.094	14.286	---
7	14	41.316	188.450	13.728	--
7	15	40.526	135.819	11.654	-
7	13	39.737	106.871	10.338	

SYSTM	DIST	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
8	5	77.263	22.871	4.782	--------555555
8	1	76.737	150.094	12.251	-------555555
8	4	76.316	88.450	9.405	------555555
8	3	76.053	21.053	4.588	-----555555
8	6	75.789	25.731	5.073	----555555
8	2	75.789	42.398	6.511	---555555
8	7	74.737	101.316	10.066	--555555
8	9	70.947	27.608	5.254	--55555
8	8	69.474	49.708	7.050	-55555
8	10	63.947	84.942	9.216	55555
8	11	54.737	167.982	12.961	5555
8	12	46.053	162.719	12.756	155
8	13	39.211	178.509	13.361	55
8	14	24.737	106.871	10.338	5
8	15	16.053	54.386	7.375	

RESULTS OF THE STATISTICAL ANALYSIS

FOR THE "RADOME" IMAGE

NUMBER OF SUBJECTS= 21
AVERAGE=56.342
SAMPLE STANDART DEVIATION=17.104
STANDARD DEVIATION DUE TO DISTORTION=64.769
STANDARD DEVIATION DUE TO SUBJECTS=9.953

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
1	1	69.524	64.762	8.047	55555555555555
1	2	61.190	97.262	9.862	155555555555
1	3	54.286	58.214	7.630	---155555555
1	4	50.714	73.214	8.557	---55555555
1	5	50.238	76.190	8.729	--55555555
1	6	47.143	158.929	12.607	-11155555
1	7	44.429	75.357	8.681	---55555
1	8	39.048	96.548	9.826	------11
1	9	38.810	119.762	10.944	---111
1	10	38.476	135.762	11.652	--111
1	13	34.048	149.048	12.209	----
1	11	33.810	64.762	9.047	---
1	12	30.333	39.333	6.272	--
1	14	29.571	113.857	10.670	-
1	15	29.190	174.262	13.201	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
2	1	72.143	86.429	9.297	--155555555555
2	2	71.429	57.857	7.606	-155555555555
2	3	66.905	46.190	6.796	----55555555
2	4	62.857	111.429	10.556	----5555555
2	5	60.238	68.690	8.288	----555555
2	6	59.286	100.714	10.036	---555555
2	7	58.571	95.357	9.765	--555555
2	8	55.238	81.190	9.011	----155
2	9	52.143	88.929	9.430	----15
2	11	48.810	122.262	11.057	-----
2	10	48.571	115.357	10.740	----
2	12	47.857	148.929	12.204	---
2	15	45.000	90.000	9.487	--
2	13	42.857	126.429	11.244	-
2	14	40.952	139.048	11.792	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
3	1	74.762	53.690	7.327	---55555555555
3	2	74.524	72.262	8.501	-155555555555
3	3	71.667	48.333	6.952	-15555555555
3	4	66.905	96.190	9.808	-5555555555
3	5	64.048	114.048	10.679	1555555555
3	6	57.381	104.048	10.200	--5555555
3	7	53.333	133.333	11.547	-5555555
3	8	51.190	84.762	9.207	1555555
3	9	43.333	180.833	13.447	----15
3	10	42.143	133.929	11.573	---15
3	11	39.762	96.190	9.808	---1
3	12	37.524	116.162	10.778	---
3	13	35.714	83.214	9.122	--
3	15	33.095	161.190	12.696	-
3	14	30.000	35.000	5.916	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
4	1	76.429	77.857	8.824	-5555555555555
4	2	71.905	53.690	7.327	1555555555555
4	3	64.286	130.714	11.433	155555555555
4	4	57.143	123.929	11.132	--555555555
4	5	52.143	116.429	10.790	-555555555
4	6	51.667	115.833	10.763	555555555
4	9	42.381	126.548	11.249	------15
4	7	42.381	54.048	7.352	---1115
4	10	40.238	143.690	11.987	-----5
4	8	39.048	71.548	8.459	----1
4	11	36.667	103.333	10.165	----
4	13	33.810	147.262	12.135	---
4	12	33.333	58.333	7.638	--
4	14	32.762	118.190	10.872	-
4	15	28.810	57.262	7.567	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
5	1	72.619	54.048	7.352	----------5
5	5	71.190	162.262	12.738	----------1
5	3	71.190	74.762	8.646	----------1
5	4	70.238	151.190	12.296	----------
5	2	70.000	125.000	11.180	---------
5	6	69.286	125.714	11.212	--------
5	7	69.190	113.762	10.666	-------
5	9	68.571	80.357	8.964	------
5	8	67.619	79.048	8.891	-----
5	13	67.143	101.429	10.071	----
5	11	66.429	127.857	11.307	---
5	10	66.429	67.857	8.238	--
5	12	66.190	109.762	10.477	-
5	14	62.857	126.429	11.244	
5	15	60.238	83.690	9.148	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
6	2	73.000	85.000	9.220	----5555555555
6	1	71.095	92.190	9.602	---1555555555
6	3	67.857	148.929	12.204	---115555555
6	4	66.190	72.262	8.501	----1555555
6	5	65.952	96.548	9.826	---1555555
6	6	61.190	109.762	10.477	---155555
6	9	59.286	95.714	9.783	-----115
6	7	58.095	103.690	10.183	-------1
6	8	56.190	127.262	11.281	------
6	13	52.619	126.548	11.249	-----
6	11	52.619	131.548	11.469	----
6	10	52.143	251.429	15.856	---
6	12	49.524	217.262	14.740	--
6	14	49.524	97.262	9.862	-
6	15	47.857	83.929	9.161	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
7	1	72.714	68.214	8.259	-5555555555555
7	2	68.333	88.333	9.399	1155555555555
7	4	61.667	105.833	10.288	----55555555
7	3	60.952	94.048	9.698	---55555555
7	6	55.238	111.190	10.545	------5555
7	5	54.524	124.762	11.170	-----5555
7	7	53.571	85.357	9.239	----1555
7	8	50.000	80.000	8.944	-------
7	9	49.048	114.048	10.679	------
7	10	47.381	81.548	9.030	-----
7	11	46.905	41.190	6.418	----
7	12	43.333	110.833	10.528	---
7	13	41.429	95.357	9.765	--
7	15	41.429	77.857	8.824	-
7	14	40.952	84.048	9.168	

DIST	LEVEL	AVERAGE	VARIANCE	STAND D.	SIGNIFICANCE
8	1	77.143	91.429	9.562	----------1
8	5	76.667	30.833	5.553	----------
8	9	76.333	37.333	6.110	----------
8	7	75.952	54.048	7.352	----------
8	3	75.952	49.048	7.003	----------
8	2	75.714	105.714	10.282	---------
8	10	74.762	36.190	6.016	--------
8	4	74.524	57.262	7.567	-------
8	6	74.238	112.690	10.616	------
8	12	74.048	26.548	5.152	-----
8	13	73.333	90.833	9.531	----
8	8	73.333	100.833	10.042	---
8	11	71.667	100.833	10.042	--
8	14	69.048	104.048	10.200	-
8	15	66.429	117.857	10.856	