IMAGE QUALITY: An Overview

Edward M. Granger, Lionel R. Baker
Chairmen/Editors

Proceedings of SPIE—The International Society for Optical Engineering

Volume 549

Image Quality: An Overview

Edward M. Granger, Lionel R. Baker
Chairmen/Editors

Cosponsored by
Sira Ltd.—The Research Association for Instrumentation

April 9-10, 1985
Arlington, Virginia

Published by
SPIE—The International Society for Optical Engineering
P.O. Box 10, Bellingham, Washington 98227-0010 USA
Telephone 206/676-3290 (Pacific Time) • Telex 46-7053

SPIE (The Society of Photo-Optical Instrumentation Engineers) is a nonprofit society dedicated to advancing engineering and scientific applications of optical, electro-optical, and optoelectronic instrumentation, systems, and technology.

The papers appearing in this book comprise the proceedings of the meeting mentioned on the cover and title page. They reflect the authors' opinions and are published as presented and without change, in the interests of timely dissemination. Their inclusion in this publication does not necessarily constitute endorsement by the editors or by SPIE.

Please use the following format to cite material from this book:
Author(s), "Title of Paper," *Image Quality: An Overview,* Edward M. Granger, Lionel R. Baker, Editors, Proc. SPIE 549, page numbers (1985).

Library of Congress Catalog Card No. 85-061483
ISBN 0-89252-584-3

Printed in the United States of America.

IMAGE QUALITY: AN OVERVIEW

Volume 549

Conference Committee

Chairmen
Edward M. Granger
Eastman Kodak Company

Lionel R. Baker
Sira Limited, United Kingdom

Co-Chairmen
Patrick S. Cheatham
Pacific-Sierra Research Corporation

A. F. J. Cox
Ministry of Defense, United Kingdom

Peter G. Engeldrum
Image Technology

J. P. Jennings
Royal Armament Research & Development Establishment, United Kingdom

Michael J. Kidger
Kidger Optics Limited, United Kingdom

John R. Schott
Rochester Institute of Technology

Thomas L. Williams
Sira Limited, United Kingdom

Session Chairmen
Session 1—Image Quality Merit Function
Patrick S. Cheatham, Pacific-Sierra Research Corporation

Session 2—Image Quality: Systems and Simulations
A. F. J. Cox, Ministry of Defence, United Kingdom

Session 3—Sensor and Display Performance
M. J. Kidger, Kidger Optics Limited, United Kingdom

Session 4—Remote Sensing
John R. Schott, Rochester Institute of Technology

IMAGE QUALITY: AN OVERVIEW

Volume 549

INTRODUCTION

The SPIE Conference on Image Quality: An Overview had as an objective to convene an international meeting of those who are currently involved with conventional optics systems and also those who are working in new areas of scanners, false color, thermal images, etc. The meeting was intended to raise the awareness of individuals in the imaging community to the needs of specifying quality for systems where the human observer has little experience. This is particularly true in medical applications where the images are no longer just X rays. Similar problems of image specification exist for thermal images where many new gray scale and chromatic schemes are being tried to enhance the interpretability of the imagery.

An attempt was made at the meeting to obtain a mix of new and old faces in the image evaluation community. The conference was a success from this point of view and more should be done in the future to bring new ideas to this format.

This conference was timely in that image quality studies seemed to have fallen on hard times, but with the advent of new sensors, computers, lasers, scanners, etc., the time has come to start work on the specification of image quality for these unconventional systems. Possibly a number of these meetings are required at more frequent intervals to help lay the foundation for international agreement on specifications in these new fields of optics and optical displays.

Edward M. Granger
Eastman Kodak Company

Lionel R. Baker
Sira Ltd.—The Research Association for Instrumentation

IMAGE QUALITY: AN OVERVIEW

Volume 549

Contents

IMAGE QUALITY: AN OVERVIEW

Volume 549

Session 1

Image Quality Merit Function

Chairman
Patrick S. Cheatham
Pacific-Sierra Research Corporation

Image quality considerations in transform coding

Andrew G. Tescher and John A. Saghri

Electronics and Optics Division, Engineering Group, The Aerospace Corporation
P.O. Box 92957, Los Angeles, California 90009

Abstract

A review of adaptive image transform coding techniques is given. Adaptive techniques are superior to nonadaptive procedures; however, their implementation requires a significant degree of added complexity. A major aspect of adaptive transform coding is the understanding and utilization of the local statistical structure of the image. Rate adaptive communication and buffer feedback are also important aspects of adaptive coding procedures. The discussion emphasizes the image quality aspect of the coding process.

Introduction

Transform image coding was developed in the late 1960s and early 1970s.[1-6] Initial concepts were based on the Fourier, Hadamard, Karhunen-Loeve (K-L), and the cosine transforms. Since then many transform coding algorithms have been developed and analyzed; however, practical implementations have been considered only recently.

The primary purpose of this paper is to review various aspects of transform coding and their impact on image quality. The presentation is based on an earlier paper by one of the authors.[7]

Transform techniques are more complex than other conventional and classical data compression algorithms. One should expect excellent performance considering the complexity of transform algorithms. It is demonstrated that, for nonadaptive transform technique implementations, only small advances are realized.[8] Additional advances are expected primarily in adaptive techniques.

Basic concepts

Transform or block coding is a data compression technique by which a set of source elements is coded as a unit. The term "transform" indicates that the original set of elements is processed by an invertible mathematical transformation prior to encoding. Figure 1 contains the basic block diagram representation of a transform coding-decoding

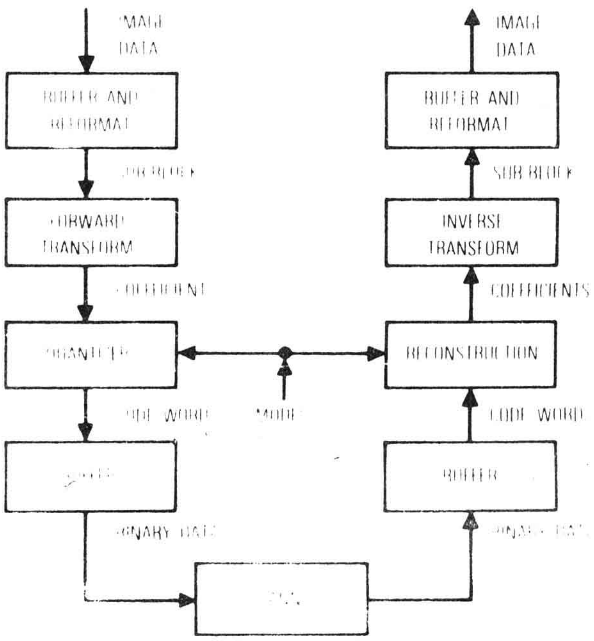

Figure 1. Basic transform coding system

system. The coding unit consists of a reformatting memory followed by the transformation and, finally, the actual coding process. The receiver is the mirror image of the decoder. Motivation for this system is provided by the statistical characterization of typical images.

Huang and Schultheiss[9] demonstrated that, for a correlated Gaussian source, optimum encoding comprises two steps. These are the decorrelation process by means of linear transformation, followed by the application of a memoryless quantizer on the output of the transformation. If the Gaussian assumption can be extended to image sources, then the optimum encoding should always first involve a decorrelation procedure.

The optimum transform for decorrelation is usually identified as the Karhunen-Loeve transformation or the method of principal components. Thus, the early concepts indicated the desirability of a decorrelating transform and provided motivation for transform coding.

Other coding aspects can be related to the relationship between compression rate and distortion, concerning which the available literature is quite extensive.[10-13] The primary and obvious consideration is that transform coding introduces a distortion that is a function of the desired compression rate. The understanding of the trade-off between rate R and distortion D will lead to practical procedures; i.e., for an at least locally stationary procedure, the relationship between rate and distortion leads to useful algorithms. It is demonstrated that the rate-versus-distortion formalism is fundamental in the proper design of a highly adaptive coding system. The appropriate formalism is developed as needed for the proper adaptive techniques.

Unfortunately, most of the research work has directly related to transform algorithms and not to what should be done once the transformation is accomplished.

Early research on transform coding concentrated primarily on the transform algorithms. The Fourier transform was found to be a good approximation to the K-L transform.[1,3,14]

The discontinuity problem associated with the discrete Fourier transform in early applications of transform coding has been significant. The Fourier transform approximation, when used over small blocks, produces undesirable blocking effects.

A symmetrized approach to Fourier coding was discovered by Ahmed, et al.[15] Recent studies further demonstrated that the cosine transform is virtually identical to the K-L transform for numerous practical conditions.[16]

The cosine transform, unlike the K-L transform, can be implemented numerically through a "fast" transform. More important, the cosine transform virtually approaches the K-L transform performance with neither utilization nor knowledge of the source correlation. For the K-L transform, the source covariance model must be available to derive the actual transform matrix.

The cosine transformation is a strictly deterministic transform. Conversely, the K-L transform is a class of transformation, and for each application it is a function of the appropriate covariance matrix. It is a significant practical benefit that the single deterministic transformation closely approaches the entire class of theoretical optimum transformation in performance.

Two other considerations of transform algorithms are computability and size. Under computability, the complexity of the potential device for implementation of the transform is defined. Fortunately, except for the general K-L transform, most of the useful transformations can be implemented by fast algorithms such as the fast Fourier transform algorithm. The suboptimal Hadamard and Haar can be implemented without multiplication. For specialized applications, a computationally more efficient transform, even at the expense of performance loss, may be preferred. On the other hand, various hybrid implementations for the cosine transform; e.g., through charge-coupled-device (CCD) technology, minimize the advantages of the suboptimal transform.[17]

Transform size is an important practical consideration. The argument is often made that no benefit is obtained by choosing transforms larger than the image correlation distance, assuming that it is known. This approach is artificial and ignores the fact that a transform decorrelates only the pixels within a sub-block. It will not decorrelate pixels among sub-blocks. A typical sub-block consists of 8 × 8 or 16 × 16 pixels. The reasoning is that the image correlation is not likely to exceed 8 or 16 pixels, respectively. Although this reasoning has not been challenged, it is not valid. Even if all pixels within the transform block become decorrelated by the transformation, the pixels on the border remain correlated with respect to pixels on the borders of adjacent sub-blocks. However, for practical reasons, it is advantageous not to exceed a 16 × 16 or 32 × 32 size.

An additional argument against the larger size is the concept of adaptivity. It is still unresolved how to optimize the transform size to achieve maximum decorrelation and yet include adaptivity consideration dependent on local image structure. These two concepts are conflicting and the solution is not obvious. To obtain maximum decorrelation, increasing block size is beneficial. Conversely, to adapt to the local image structure, a smaller block size is preferred. In addition, the overhead associated with an adaptive transform algorithm is likely to become more important with decreasing transform block sizes.

In summary, three primary considerations relate to transform size. Larger sizes minimize block-to-block correlation. Adaptive procedures are likely to favor smaller sizes. Finally, the overhead information per sub-block should not exceed a reasonable fraction of the available bandwidth.

Quantization

The general effect of quantization of the source is a primary consideration for all source-coding procedures. Quantization is the noninvertible mapping from the analog source to its quantized equivalent. Thus, the originally continuous parameters will be represented by integers. In general, the quantizer is the major error source in image coding. All other error sources, such as numerical round-off, are negligible in most instances. If not, they can be minimized or even eliminated by increasing register sizes and by using integer arithmetic. However, quantization is a fundamental distortion, and in most cases it is unavoidable.

A quantizer is a mapping from the continuous variable domain of transform coefficients into the domain of integers. These integers become the code words that are transmitted through the channel. An alternative superior approach is a two-phase coding technique in which the integers are the secondary input into an entropy coding processor. The output of this coder generates the final code words to be transmitted through the channel.

The quantizer output is an integer that is also the code word for one-step coding. The relevant optimization procedure is the minimization of mean-square error between original and quantized coefficients. Since all practical transform coding systems utilize unitary transforms, the mean-square error is preserved under the transformation. Thus, quantizer optimization in the transform domain guarantees optimum spatial domain performance in the minimum-mean-square error sense.

For a given transform domain model, the quantization parameters are easily obtained with the Max[18] quantizer algorithm. Knowledge of the probability density function of the random variable for a specified number of quantization bins permits computation of the required thresholds and reconstruction levels through the Max algorithm.

Limitations of nonadaptive transform coding

For a nonadaptive technique with a specified rate, local degradations are essentially unavoidable. Except for changing the bit rate, these degradations are deterministic. A nonadaptive algorithm is designed to be a fixed coding algorithm operating identically for all images and all image regions.

For a nonadaptive technique, the image to be coded is assumed to be a stationary source. Were the stationarity assumption valid, a nonadaptive coder could be an optimal image coder. In general, the stationarity assumption is invalid.

Images usually have varying statistical structures, both from image to image as well as within an image. Image nonstationarity results not only from scene nonstationarity but also from the imaging process itself. An imaging device projects a three-dimensional scene onto a two-dimensional plane. Consequently, those parts of the image in focus require high fidelity and, equivalently, a substantial portion of available bandwidth.

This qualitative discussion suggests that the stationarity model for image coding might be a severe design limitation. The real gain is in adaptivity.

Adaptivity considerations

The assumption made is that imagery as a source should be modeled as a nonstationary representation. Consequently, coding procedures, which to a large extent freeze the algorithm, are not appropriate. In the following discussions, the image-coding model is part of the information to be utilized by the decoder. This fact does not necessarily indicate that new or additional modeling information must be transmitted to the decoder. For an efficient system, the modeling information can and should be derivable from previously decoded information.

The basic assumption made is that the image coding procedure is the transmission of fluctuation information about a preassigned model. However, this model is part of the information to be transmitted. Ideally, the model would permit local changes through parameterization. More important, it would assist in meaningful redistribution of the available bandwidth among partitions of the image.

Although image coding is applicable for nonreal-time applications such as storing pictorial information on magnetic tape, the primary application is real-time image transmission through communication channels. Thus, the image coder must be consistent with the constraints of a realistic communication model.

With increasing adaptivity, the coder output will fluctuate in rate if the adaptivity permits a variable-rate compressor. The relevant problem is how to interface the variable-rate compressor with the fixed-rate channel. For a variable-rate compressor, communications model constraints become a necessary consideration. In principle, the solution is simple. The compressor (source encoder) must be interfaced with the communication channel through a rate-equalizing buffer. This buffer permits the deviation in bits from the average rate as required. However, the practical solution is somewhat involved.[18,19]

Another practical consideration for a communication channel relates to channel errors. The various applicable channel coding techniques can be classified into two groups; in one case, through algebraic coding, channel errors are essentially eliminated. In the other case, a fixed, non-negligible rate of channel errors, including catastrophic errors, is permitted. However, periodic reinitialization of the algorithm is performed.

In general, it is difficult to evaluate the impact of channel errors for adaptive procedures. Erroneous bits may be significantly important, particularly for overhead information. On the other hand, erroneous bits that result only in incorrect reconstruction of a single transform coefficient may result in relatively minor image degradation. Consequently, the primary impact of a communication system is the need to interface the inherently fixed-rate transmission system with a locally fluctuating-rate source-encoding system. The appropriate problem is to design a global fixed-rate system that is highly rate adaptive locally.

Types of adaptivity

Adaptivity procedures can be classified into two broad categories. For the first, the source-coder rate is constant. However, various other coder parameters are changing. For the second, in addition to other parameters, the local compression rate is also variable. The second adaptivity category involves implementation difficulty because of the varying rate.

In addition to classification according to rate, other considerations are appropriate. For a highly adaptive system, the degree of adaptivity or the number of different ways the coder can operate may be either a large number or essentially a continuous parameter. In this case, the overhead information needed by the decoder to determine the operation mode should be decodable from previously decoded data. Otherwise, the transmitted overhead would require a greater fraction of the available bandwidth.

Buffering concepts

Variable-rate source-coding techniques require rate-equalizer buffers. Because the communication channel operates at a fixed rate, the source coder rate fluctuation must be absorbed. A general model of a variable-rate system is displayed in Figure 2.

The rate-equalizing buffer is implemented between the source coder and the channel. The important consideration is how the source-coding system controls gross parameters of the decoding. In a realistic system, the source coder cannot operate at variable rates without some externally controlled mechanism. An open-loop source-coding system is not acceptable. This statement is consistent with the assumed philosophy of nonstationarity. Even for a variable-rate system, image nonstationarity may be more extreme than what the rate-equalizer buffer could allow without an additional controlling mechanism.

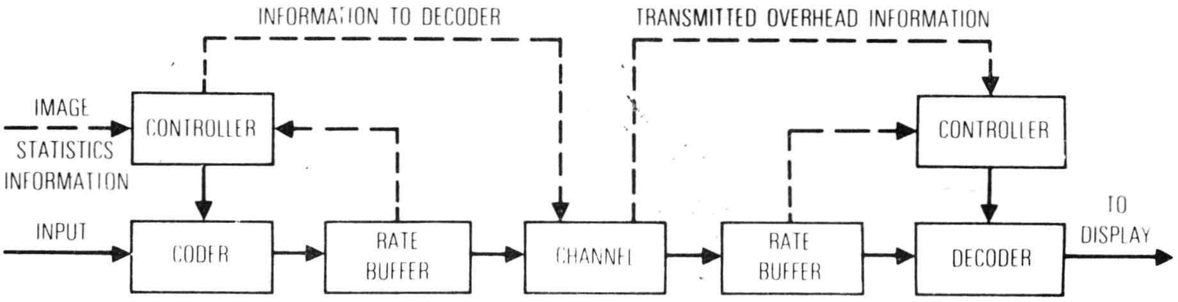

Figure 2. Variable-rate coding system

In any transform coding system some form of a "prebuffer" is required to perform the image reformatting into sub-blocks. For a variable-rate system, an additional buffer is required to accommodate source-coder fluctuations. The basic problem is to determine how to control the overall coding parameters in terms of utilization of the various buffers. Two general solutions are available. The first approach uses the reformatting buffer to control the coding parameters, while the second uses the rate-equalizer buffer in a feedback loop. The two techniques are not equivalent.[16] The first technique is logically easy to understand; however, it is less efficient. Here, the reformatting buffer serves as a control mechanism.

A recent application

A recently developed system (by Widcom, Inc. of San Jose, California) for teleconferencing at 56 kbits/sec is an excellent demonstration of the power of transform coding.[20] The required compression is 1440:1 and is achieved by considering spectral and temporal as well as spatial correlation (Table 1). The conceptual implementation is shown in Figure 3.

Table 1. Image redundancy consideration

CORRELATION TYPE	IMAGE COMPRESSION FUNCTION	COMPRESSION RATIO
SPECTRAL (REDUNDANCY OF COLOR COMPONENTS)	FILTERING AND SUBSAMPLING OF CHROMINANCE COMPONENTS	2.5:1
SPATIAL (REDUNDANCY IN LUMINANCE OR AMPLITUDE)	2:1 SUBSAMPLING IN EACH PRINCIPAL DIRECTION	4:1
	2-d COSINE TRANSFORM CODING FOR PIXEL-TO-PIXEL DECORRELATION	6:1
TEMPORAL (REDUNDANCY IN FRAME-TO-FRAME VALUES)	FRAME SKIPPING AND INTERPOLATION	3:1
	CONDITIONAL REPLENISHMENT	8:1
TOTAL ACHIEVABLE BANDWIDTH COMPRESSION		1440:1

Conclusions

Important aspects of adaptive image transform coding have been reviewed. Although more complex, adaptive techniques significantly outperform nonadaptive procedures. Rate adaptive communication is required to handle the local image structure fluctuations.

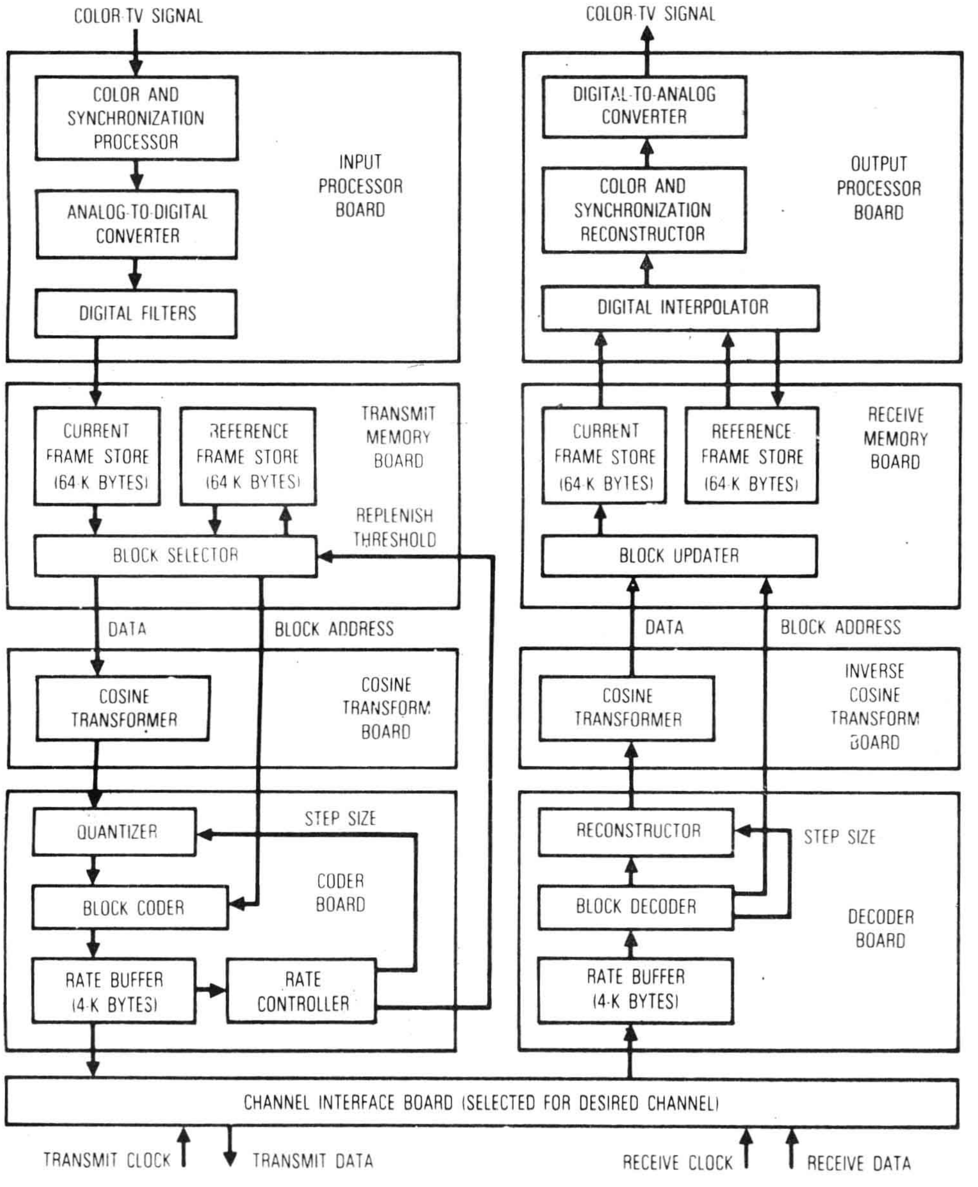

Figure 3. Teleconferencing implementation

References

1. W. K. Pratt, J. Kane, and H. C. Andrews, Proc. IEEE 57, 58-68 (1969).
2. W. K. Pratt and H. C. Andrews, Picture Bandwidth Compression (T. S. Huang and O. J. Tretiak, eds.), Gordon & Breach, New York (1972), pp. 515-554.
3. A. Habibi and P. A. Wintz, IEEE Trans. Commun. Technol. COM-19, 957-972 (1971).
4. G. B. Anderson and T. S. Huang, IEEE Trans. Commun. Technol. COM-19, 133-140 (1971).
5. P. A. Wintz, Proc. IEEE 60, 809-820 (1972).
6. W. K. Pratt, W. H. Chen, and L. R. Welch, IEEE Trans. Commun. Technol. COM-22, 1075-1093 (1974).
7. A. G. Tescher, IEEE Conference Record of the Sixteenth Asilomar Conference on Circuits, Systems and Computers (1983), pp. 50-54.
8. A. Habibi, IEEE Trans. Commun. Technol. COM-19, 948-956 (1971).
9. J. J. Y. Huang, and P. M. Schultheiss, IEEE Trans. Commun. Syst. CS-11, 289-296 (1963).
10. S. T. Huang, and O. J. Tretiak, eds., Picture Bandwidth Compression, Gordon & Breach, New York (1972).
11. T. S. Huang, W. F. Schreiber, and O. J. Tretiak, Proc. IEEE 59, 1586-1609 (1971).
12. H. J. Landau and D. Slepian, Bell Sys. Tech. J. 50, 1525-1540 (1971).

13. A. K. Jain, Image Science Mathematics (O. C. Wilde and E. Barrett, eds.), Western Periodicals Company, North Hollywood, California (1976), pp. 210-223.

14. J. W. Woods and T. S. Huang, Picture Bandwidth Compression (T. S. Huang and O. J. Tretiak, eds.), Gordon & Breach, New York (1972), pp. 555-573.

15. N. Ahmed, T. Natarajan, and K. Rao, IEEE Trans. Comput. C-23, 90-93 (1974).

16. A. G. Tescher, ed., "Advances in Image Transmission Techniques," SPIE Proc. 87, Bellingham, Washington (1976).

17. H. J. Whitehouse, R. W. Means, and E. H. Wrench, SPIE Proc. 66, 36-47 (1975).

18. J. Max, IRE Trans. Inform. Theory IT-6, 7-12 (1960).

19. A. G. Tescher, "Transform Image Coding," Image Transmission Techniques, Academic Press, New York (1979).

20. J. Anderson, et al., "Codec Squeezes Color Teleconferencing Through Digital Telephone Lines," Electronics (26 January 1984).

A Psychophysical Approach to Image Quality

Bernice E. Rogowitz

Vision and Displays Group, IBM Thomas J. Watson Research Center
Post Office Box 218, Yorktown Heights, New York 10598

Abstract

Whether an image is produced on a CRT display, a printer, or a photographic plate, that image can be described as a two-dimensional modulation of luminance and color across space. A psychophysical approach to image quality is one which considers the detectability and discriminability of these images in terms of the spatial, temporal, and color-sensitive mechanisms of perception. The aim is to develop a technology-independent method for characterizing and predicting image quality.

Introduction

Many technologies exist today for making visual images. Printing technologies put ink or electrostatic toner on paper, display technologies bombard phosphors with electrons or excite gasses to emit photons, photographers expose and develop emulsion, etc. Although these technologies vary greatly, the same language can be used to describe the images they create. In all cases, the picture created is a two-dimensional variation of luminance and color across space. The fact that the images created by diverse methods can all be described by the same vocabulary marks the first step to developing a technology-independent approach to characterizing and improving image quality. The second step involves the realization that the way an image "looks" depends on the visual system looking at it. The luminance of an object, for example, can be measured quite easily. The apparent "darkness" or "lightness" of that object, however, depends on the workings of human perception. The Hermann Grid, for example, is an array of black squares.

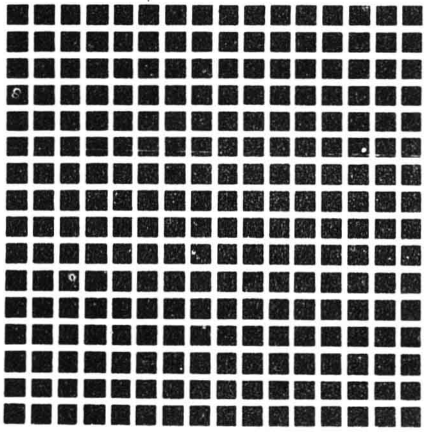

Figure 1. The Hermann Grid: The gray circles are created by perception

Although the white grid has equal luminance throughout, small gray circles can be seen at each of the intersections. These differences in perceived luminance cannot be measured by a photometer; they are not present in the physical stimulus. We see the gray circles because of the way the visual system processes adjacent areas of dark and light. The science which studies the mechanisms which cause us to see the physical world as we see it is called "visual psychophysics".

Spatial Frequency Representations

It has been known since Fourier that a complex periodic signal can be represented as the sum of its sinusoidal components. This theorem, originally developed to describe the decomposition of heat waves, has been broadly applied to electrical waves, auditory waves, and visual images. Figure 2 shows a sinusoidal luminance distribution and its luminance profile.

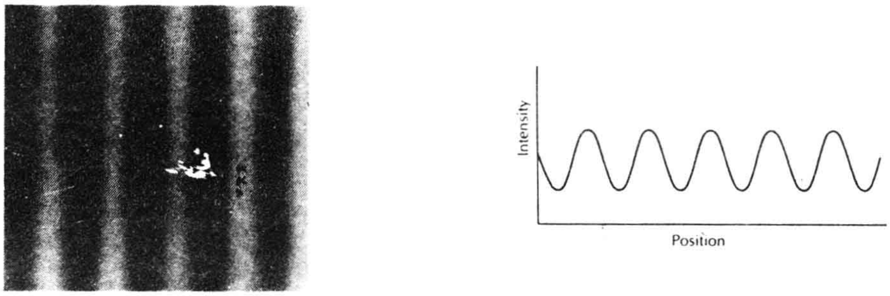

Figure 2: A sine-wave grating: The luminance varies sinusoidally

Any complex visual image can be represented as the sum of sinusoidal modulations of luminance (i.e., sine-wave gratings), varying in their spatial frequency, contrast, phase, and orientation. This provides a convenient way of representing images. This characterization of visual images is particularly interesting to the visual scientist because both electrophysiological and psychophysical evidence support the hypothesis that special mechanisms exist in vision which are particularly sensitive, "selectively" sensitive, to narrow ranges of spatial frequency information. It is as if we have specialized spatial frequency filters.

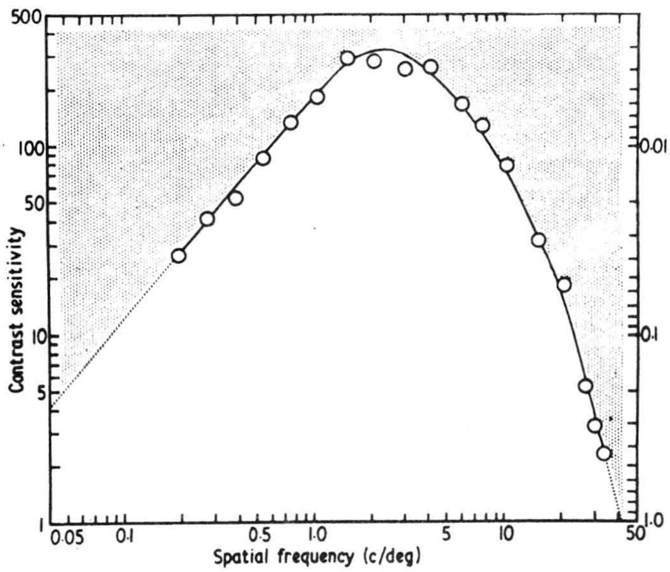

Figure 3: Spatial-frequency sensitivity

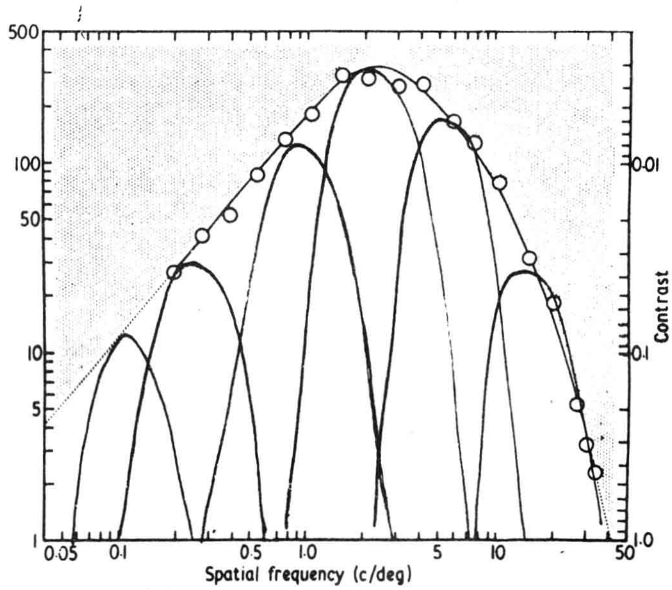

Figure 4: Spatial-frequency "channels"

Figure 3 shows our overall sensitivity to spatial-frequency modulated grating patterns. Figure 4 shows spatial-frequency "channels," the underlying narrow-band spatial filters. We are most sensitive to patterns having 2-4 cycles per degree (c/d) of visual angle. Sensitivity drops off for lower and higher spatial frequencies. (For reference, the vertical spatial frequency of these horizontal lines of print, when held at 20 inches, is roughly 3 c/d.)