

M. Nesenberg

NTIA Report 89-252

Image Data Compression Overview: Issues and Partial Solutions

Martin Nesenbergs



U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary

Janice Obuchowski, Assistant Secretary
for Communications and Information

October 1989



PREFACE

The Naval Sea Systems Command of the Department of the Navy has tasked the Institute for Telecommunication Sciences (ITS) to assist the Navy in developing local, metropolitan, and wide area networks for integration and transmission of CAD/CALS and high-resolution graphics data along with video teleconferencing over satellite networks.

This NTIA Report represents work done on one particular subtask of the program, namely, the one that deals with motion picture data compression.

The views, opinions, and findings contained in this Report are those of the author only. The report does not reflect Navy, NTIA, ITS, or any other agency position, policy, or decision unless so designated by other official documentation. Certain commercial equipment and software products are identified in this report to adequately describe the design and conduct of the research or experiment. In no case does such identification imply recommendation or endorsement by the National Telecommunications and Information Administration, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

Administrative and technical monitoring of this project was performed by Lt. Commander Richard Manning of the Naval Sea Systems Command. Technical leadership and administrative supervision of the program at ITS were provided by the project leader, Dr. Edmund A. Quincy.

CONTENTS

	Page
LIST OF FIGURES	vi
ABSTRACT	1
1. INTRODUCTION	1
2. HUMAN PERCEPTION OF VISUAL INFORMATION	3
3. THEORETICAL REDUNDANCY REDUCTION	8
4. COMPRESSION TECHNIQUES	11
5. MOTION COMPENSATION	16
6. CURRENT SYSTEMS AND THEIR TRENDS	30
7. CONCLUSIONS	38
8. REFERENCES	39
9. BIBLIOGRAPHY ON IMAGE DATA COMPRESSION	41

LIST OF FIGURES

	Page
Figure 1. The visual resolution domain of a human eye.	10
Figure 2. Simplified overview of different image coding classes.	14
Figure 3. Identification of motion (dark) subfields among static subfields.	23
Figure 4. The role of displacement vector (a,b) in block matching.	25
Figure 5. Basic diagram of the scene adaptive coder (from Chen, 1981).	32
Figure 6. Combined interframe and intraframe scene adaptive coding system (as in Chen and Hein, 1987).	35

IMAGE DATA COMPRESSION OVERVIEW: ISSUES AND PARTIAL SOLUTIONS

Martin Nesenbergs*

This report presents an overview of moving image (video) data compression. Because this report deals with a subtask of a major program, a limited number of the most pertinent topics and issues have been addressed. These topics include a listing of human perception effects and parameters for visual information; theoretical aspects of image data compression; a selective review of compression techniques; a closer look at motion detection, prediction and compensation; and a start to survey existing systems and their capabilities. Finally a multipage bibliography, while not exhaustive in any sense, shows the depth and breadth of the rapidly growing technology of image data compression.

Key Words: data compression; digital; images; motion compensation; prediction; visual effects

1. INTRODUCTION

This report presents a brief look into the growing present-day technology of image data compression. Because of the concentrated thrust of the program, neither the depth nor the breadth of this sophisticated and fast-growing technology are fully covered in the following pages. From the many topics pertinent to teleconferencing and video systems, only certain few have been selected.

The topic choices were made according to the following three criteria: (1) to introduce the reader to the image compression field; (2) to identify key problems and potential solutions (or at least, to recognize what appear to be the more active areas of work); and (3) to give a representative overview of some of the promising advanced codecs, their performance, and future trends.

The material presented here is divided into sections as follows. Section 2 deals with the video end-user, i.e., the viewer and what he or she perceives or fails to perceive. An incredible number of effects and parameters are associated with human vision. Moreover, their interactions are numerous, complex, and poorly understood. Section 2 does not attempt to

*The author is with the Institute for Telecommunication Sciences, National Telecommunications and Information Administration, U.S. Department of Commerce, Boulder, CO 80303.

discuss them all, nor to explain the underlying theories in a comprehensive fashion. Instead, the more commonly encountered vision effects are listed, in alphabetical order, and defined. A reader interested in pursuing some or all of the effects in depth is advised to turn to the very extensive existing literature, starting with the partial bibliography included at the end of this text.

Section 3 looks at the basic theoretical principles and results for image compression. The classical approaches of information theory to this kind of problem have been along the lines of rate distortion theory and source encoding. An additional problem associated with image reproduction systems is their dependence on brain-eye-image interaction. This implies that to have a practical impact, among other things, the theory must somehow utilize the existing knowledge about various aspects of human vision (see the parameters listed in Section 2). Additional theoretical background is also found in the bibliography (Section 9).

Sections 4, 5, and 6 describe compression technology. To start, Section 4 gives an overview of various image compression schemes. Different systems answer the needs of different applications and different performance requirements by end-users as they are balanced against the capabilities and practices of existing telecommunication networks.

Section 5 describes motion compensation techniques that have received increased attention and that appear to be quite effective for image compression. Topics such as interframe prediction, motion estimation, gain compensation, effects of bit-error-rate (BER) on compressed image transmission, and the general potential of motion compensation are reviewed.

Section 6 attempts to illustrate the technology of the latest and what appear to be the best video compression systems on the market. Originally, it was the intent for this section to be a comprehensive survey of existing systems. However, changing circumstances dictated the release of this abbreviated report. Data on many omitted codecs (or coders/decoders) remain to be gathered and evaluated.

Section 7 provides a brief conclusion, and Section 8 is the list of references.

Finally, Section 9 is the bibliography. It cannot claim to be complete, nor to represent fairly the different domains of image data compression. Instead, it is simply a list of journal articles and books that we have--even if in some cases ever so briefly--consulted as part of this study.

2. HUMAN PERCEPTION OF VISUAL INFORMATION

This section presents an alphabetic listing of the more common viewer perception effects, both as single parameters and as more complex (multi-parameter) visual phenomena. The effects are far from independent. They interact in an extremely complicated manner to cause or to influence the many vision sensations experienced by humans (Seyler and Budrikis, 1965; Pirenne, 1967; Cornsweet, 1970; Budrikis, 1972; Limb, 1973; Mannos and Sakrison, 1974; Bowen and Limb, 1976; Sakrison, 1977; Boynton, 1979; Miyahara, 1988; Netravali and Haskell, 1988). To save space, only a minimal explanation is associated with each item. For more extensive explanation of these human effects, including the grouping of parameters into subject-related classes, one can consult the rest of the rather extensive bibliography given at the conclusion of this report. Finally, the following should not be viewed as a complete list.

Area busyness	Inordinate, high-frequency variation in the brightness or color content of the image: spatially, temporally, or both.
Aspect ratio	The ratio of width to height for the visual image, such as 4:3 for NTSC standard television and 5:3, or even 16:9, for future HDTV.
Background granularity	Resolution limitation due to rough textures exhibited by phosphor, film, or other media used in image generation.
Binocular and monocular effects	Perceived image degradation due to different emphases placed on one or on both eyes. For people with various visual biases, including partial blindness, in either of the two eyes.

Brightness	Perception effects associated with brightness (said to be caused by image luminance), such as general brightness, its uniformity and saturation, plus eye responses under varied brightness conditions in the entire or partial image and in the background.
Chromatic adaption	Visual response as function of present and past color exposures of the retina. Sometimes, in addition to the real color, complementary colors are noted.
Color sensitivity variation	Sensitivity variation of the human eye as a function of any number of objective, subjective, and psychophysical conditions and parameters.
Complementary images	Pseudo-images induced by certain special image patterns.
Contrast sensitivity	Property of the eye, where particular contrasts appear sharper than others.
Detectability of periodic pattern	Relative ability of the eye to distinguish sinusoidal gratings (horizontal, vertical, consisting of varied brightness or of different colors).
Distorted line effect	Individual's reaction to image streaking, as noted by the bad lines due to interference on television monitors.
Echoes or ghosts	Perceived echoes of the entire image, or specific parts, or specific colors. As observed, for instance, under multipath conditions on TV.
Equivalent BER	For arbitrary distortions, the value of bit-error-rate that in a subjective comparison test would yield comparable image quality.
Eye characteristics	Includes many specifics such as key optic parts (fovea, optic nerve, retina, rods, etc.) and their roles, functions that start with optics

	and end in the brain (chromatic aberration, subjective colors, focussing, movement recognition, joint spatial and temporal frequency response, saturation, target tracking, vacillation, etc.), and eye malfunctions (astigmatism, partial or complete color blindness, etc.).
Feature recognition	Inferences drawn from the expected and the actually observed image information.
Fidelity criteria	General and specific factors in image quality assessment. Specific criteria may be defined for different user groups, different applications, and different image classes.
Flicker	In presence of periodically repeating frames, a visible effect due to the repeat frequency falling within the low-pass bandwidth of the human vision.
Geometric distortion	Nonuniform reproduction of the (x,y) coordinates of objects either over the entire planar image or for specific parts of the image.
Horizontal scan	The noticeable presence of horizontal scan lines, as in television.
Image properties	The common statistical image characteristics (e.g., means, variances, autocorrelation functions) and their visual impact.
Impairment scales	Any of the subjective rating scales used to measure image quality in presence of impairments, such as the 1 to 5, 1 to 10, or -3 to 3 scoring scales. Their averages are called the mean opinion score (MOS) numbers.
Interval vision tests	Specific ranges or classes of quality used for testing a given image. The test score can be either in an interval or out, yes or no, better or worse.
Lateral inhibition	Optic nerve phenomenon, where a strong photoreception in one cell suppresses the signals in neighboring cells.
Light sources	The general effect of different light sources on perceived image quality. The light sources can be industry standards (NTSC, PAL, CCITT) or commercial lighting products.

Luminance properties	The effects of such measured luminance properties as luminous energy, luminous flux, luminous exitance, and the three types of luminous intensity (per solid angle, per area, and per area times solid angle).
Moire effect	Image distortion similar to signal aliasing. Given improper sampling frequency, the effect may be that of a periodic interference pattern. The Moire effect may also arise when poor thresholding techniques are used in picture quantization.
MOS	Mean opinion score, as deduced from a particular impairment scale.
Nominal vision tests	Adjustment of one or more processor parameters in one of two images until they yield comparable subjective quality.
Ordinal vision tests	Comparison tests of several images to rank them according to subjective quality.
Perimeter effects	The distracting effects of peripheral features of a picture when focussing on the main, central object in that picture.
Psychophysical distinction test	Tests on the optical, neurological, and psychological functions that enhance image resolution, or distinction between selected facets of the image.
Ratio tests	Tests of subjective image quality, where a single parameter is scaled (varied) to achieve a target ratio of quality scores, such as increasing the MOS by a factor of two.
Rubber-sheet distortion	A pliant, in and out, optical illusion, seen by the eye as it moves across the image.
Shape recognition	In a noisy background, the ease and speed with which certain marginally bright objects are recognized by individuals.
Shimmering effect	An aftereffect perceived on a blank background, after the image is gone.
Simultaneous contrast	Property of the eye to recognize some contrasts simultaneously, others with delay, and at possibly different intensity levels.

Snow effect	The speckled, snow-like patterns caused by noise, as noted on television screens when the broadcast signal is weak or absent.
Spatial bandwidth	The maximum number of cycles per degree or alternatively the number of image samples or pels per degree, as either determined by the Nyquist criterion or as measured from the viewing distance.
Streaming phenomenon	A visual impression that resembles a "fine dust" flowing across the TV picture, the flow being perpendicular to the scan lines.
Subjective color	Perception of color where it should not be or when it actually does not exist. As in certain rapidly changing black-and-white patterns.
Subjectively weighted errors	Different weights assigned to different image impairments in accordance to subjective assessment by viewers.
Surrounding illumination	The visual effects of the illumination environment that surrounds the picture and its viewer.
Test colors	The primary or standard colors used for system test purposes such as the red, green, blue, as well as the reference white, of the NTSC color television.
Test patterns	The various patterns contained in the IEEE Facsimile Test Chart, the CCITT Standards, and others, used for image quality tests.
Texture	The granular, often nearly periodic, chrominance details that persist through portions of the picture.
Threshold of visibility	Usually applied to edge detection, it denotes the value by which an edge contrast can be perturbed before the perturbation becomes visible. Also called the achromatic detectability.
Time and intensity reciprocity	When viewing short pulses of object illumination, a duration time that is inversely proportional to light intensity appears to have the same effect on the eye. Sometimes called the Bloch's law.

Viewing angles	The effect of different, horizontal and vertical, viewing angles on perceived image quality. There is a maximum permitted angle from the central image axis, within which fine resolution is possible.
Viewing distances	Effect of observer-to-image distance on perceived image quality.
Visual acuity	Human ability to see detail, especially in dim light, as function of age, health, corrective lenses, etc.
Visual ratings experiments	Tests intended to compare or to rate different images.
Waterfall illusion	A perceived motion, opposite in direction to the actual motion, but often at different speed. Also seen as reverse rotation of wheels in moving pictures.

3. THEORETICAL REDUNDANCY REDUCTION

Both statistical communication theory and information theory pertain to image transmission; however, their application is not straightforward. The basic theories typically give performance bounds or benchmark targets that ideally could be achieved with some "optimum" system on images with specified statistical characteristics. The theories usually stop short of saying what these optimum systems are or how they could be implemented. But the theory often indicates the general nature of reasonably good systems, including an assessment of their complexity.

A very basic illustration is found in the rate distortion theory, which is part of information theory and has its roots in the work of Shannon (Shannon, 1959; Davisson and Gray, 1968). In that context, a central role is played by the Data Processing Theorem, which states that the mutual information between any two random variables (e.g., images) cannot be increased by operations on either variable. A key part of the theory pertains to properties, computation, and uses of the rate-distortion function or $R(D)$. For a communications link, $R(D)$ represents the lower bound on channel capacity (relative to the sampling rate) which is required to achieve any particular image distortion not in excess of D .

Different image classes with different statistics and applications may warrant different fidelity (quality) measures or distortion criteria. Three difficulties arise. First, in systems of interest the estimation of source statistics takes too much time and effort. Second, it seems almost impossible to identify a distortion criterion that is both meaningful and mathematically (or perhaps computationally) tractable. And third, given source statistics and the distortion criterion, the rate-distortion function $R(D)$ is usually extremely difficult to compute.

The few known exceptions occur under very ideal assumptions, such as for independent, identically distributed, Gaussian source statistics and mean-square error (MSE) distortion criterion. An example, said to be the Shannon's MSE Lower Bound (Gallager, 1968; Berger, 1971; Blahut, 1972), is

$$R(D) \geq H(b)/N - (1/2) \log_2(2\pi eD).$$

Here $H(b)$ is the source entropy in bits per N -pel block sample and D is the permitted MSE level. Note that, if $D = 1/(2\pi e) = .058$, then the required channel throughput must be no less than $H(b)/N$ times the sampling rate.

Having the interpretation of an average number of information bits per pel, $H(b)/N$ values may be estimated for various types of images. Not surprisingly, the results vary a lot (Huang, 1977; Jain, 1981; Netravali and Haskell, 1988). The average number has been noted to be anywhere from .1 to 6 bits per pel for one-tone (or gray) pictures, and only slightly higher, say .1 to 7 bits per pel, for color pictures seen on broadcast television.

There are several interpretations possible for such "bits-per-pel" numbers, depending on what role is assigned to the observer. Consider, for instance, the typical viewing scenario shown in Figure 1. Let an eye view the object image from a distance R . The dimensions of the rectangular image are H and V , the aspect ratio is H/V . For large R , the viewing angle is small and the eye is limited in its ability to resolve the finer details in the picture. This means that there must exist a lesser area, say $h \cdot v$, within which further details are lost to the eye.

The assignment of bits/pel numbers in the eye-image system thus must depend at least on the pel spacings, horizontally and vertically, plus the

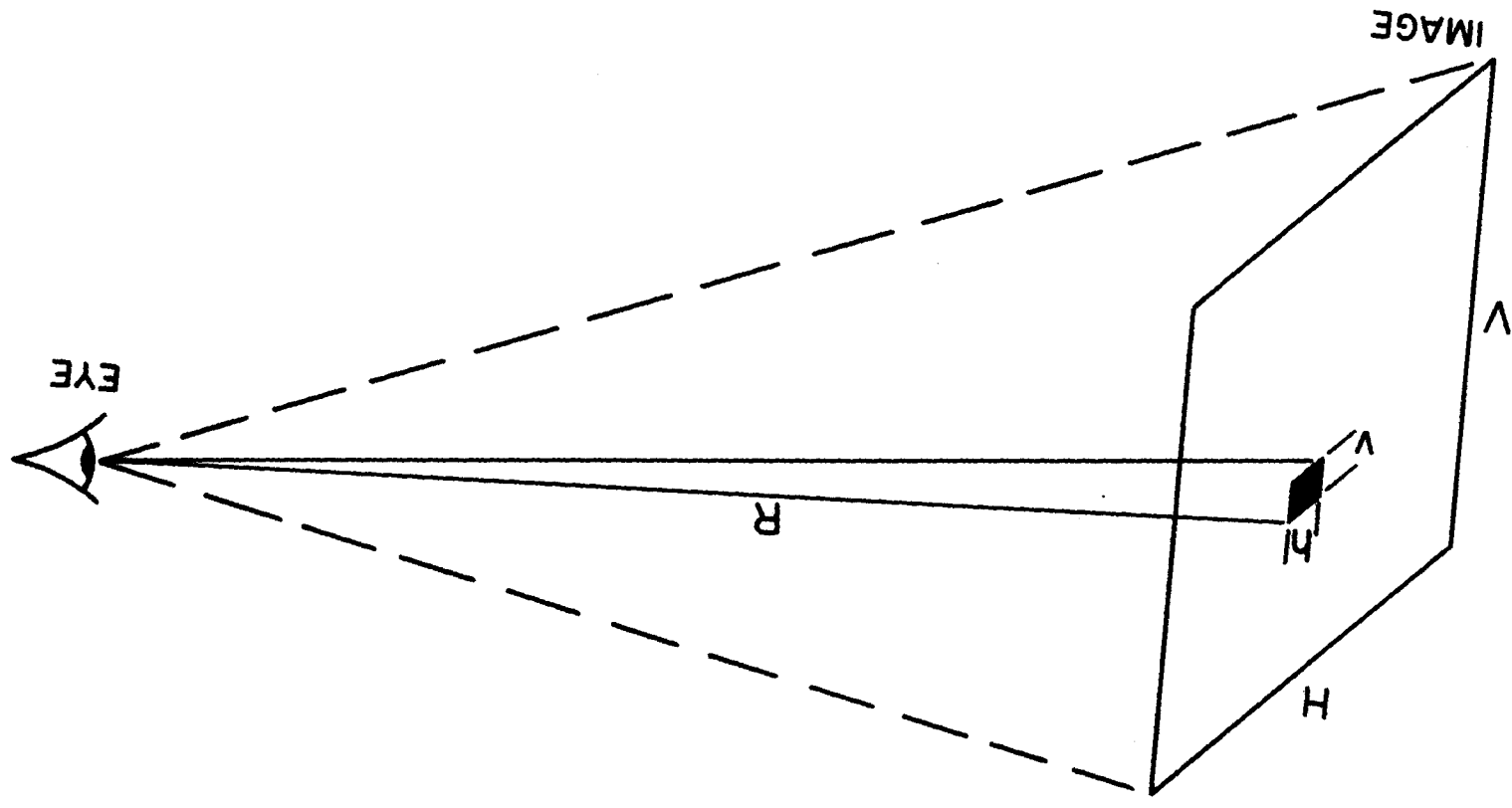


Figure 1. The visual resolution domain of a human eye.

viewing distance R . If the pel spacings are far less than h or v , the number of bits per pel is reduced. If the spacings are larger, the entropy can increase subject to the statistics of the image. A secondary effect that influences the $H(b)/N$ value occurs for N larger than the number of pels per frame. Then the correlation between successive frames leads to further entropy reduction.

Theoretically, one can get rid of the undesired effect of distance R by either of two strategies. First, one can define a "standard distance," such as perhaps $R=2$ meters, for normal viewing of broadcast television. But, that does not account for different screen sizes, different resolution environments, and the practical need for different viewing distances. In a second strategy, one can move the eye closer and closer to the image. In the limit of $R=0$, the eye could be replaced by an ideal optical instrument that has unlimited resolving power. But then the entropy of the visual signal would be overshadowed by picture granularity and random noise. And even for very small pel spacings, this would yield an $H(b)/N$ number that differs from zero and has no real connection to the desired image.

The conclusion is that a direct fixed assignment of link capacity (bits per second) to the eye-image system is far from simple. To arrive at efficient redundancy reduction schemes, complicated, perhaps adaptive, processing and coding schemes must be designed and implemented.

4. COMPRESSION TECHNIQUES

An incredibly large number of image compression schemes have been attempted in different applications.

First, different user requirements apply to facsimile, black-and-white graphics, gray scale, color, broadcast television, high-definition television (HDTV), video teleconferencing, CAD/CAM, and so forth. The nature of the user application (mission) establishes what classes of pictures are to be sent and what their resolution should be as well as other quality objectives.

Second, different telecommunication media, such as VHF/UHF channels, satellite links, optical fibers, terrestrial microwave, cables, and wire lines, offer different bandwidth and data rate capacities through their analog

and digital modems, respectively. In the following discussion only the digital channels will be considered.

A good starting point for an overview appears to be pulse code modulation (PCM). In the case of voice communications, PCM has represented an industry standard for approximately 30 years. At the Nyquist sampling rate and with 8 bits per sample, PCM runs at 64 kb/s. The resultant voice reproduction is considered exceptionally good by industry standards.

Straightforward application of PCM to image transmission leads to capacity problems. As an example, consider a 525 lines/frame TV raster, common to the NTSC systems used in the United States. It is said to work at 30 frames per second; however, the actual implementation uses a 2:1 interlacing with 59.94 fields/second. The resultant analog bandwidth is 4.2 MHz for luminance and 1.9 MHz for chrominance. Nyquist sampling at a 12.2-MHz rate and with 8 bits per sample yields a total transmission rate around 97 Mb/s.

A rate of 97 Mb/s is clearly too expensive for most applications. Several alternatives exist. One can delay the output by not transmitting in real time. But that may be acceptable only for facsimile or graphics, not for television. Or one can bandlimit the analog signal, sample it less frequently, or quantize it to a lower number of levels. At the cost of considerable visual distortion, data rate savings can be so realized. For instance, there have been reports of limiting the total baseband to 3.0 to 3.5 MHz, sampling it at a rate somewhere between 6 and 7 MHz, and using either 6 or 7 quantizing bits. The target transmission rate for such degraded PCM may be 42 Mb/s, which is still high, but the quality is questionable.

A more modern approach is to leave the analog domain and to process the image signal in the digital domain. One samples and quantizes the image readings, both brightness and color, as early as is practical. The individual samples are loosely called pels or pixels. From the entropy point of view, the amount of uncertainty carried by each pel should not be much less than unity (oversampling) nor should it be much larger than unity (undersampling). Many processing schemes, some of which will be discussed below, claim typical $H(b)/N$ values in the .2 to 3 bits/pel range for practical implementation and acceptable quality images.

To compress digitized source images for transmission, the images are encoded in ways that fit the channel and that reduce the number of transmitted bits. At the receiver, decoders reverse the process and deliver the destination image. In principle, two types of redundancy are reduced by image coding: the statistical (picture) redundancy and the subjective (viewer) redundancy.

The statistical redundancy coding is of such character that, in principle, it incurs no information or entropy loss. It merely removes the repetitive and predictable parts from the data transmission process. Given an appropriate decoder and a noiseless channel, the original picture content can be restored in every detail for this kind of compression coding. Tree codes, Huffman codes, run-length codes, and various universal, adaptive, predictive, and conditional codes have been associated with statistical redundancy coding of images.

The subjective redundancy coding produces an approximation of the image. In so doing, it irreversibly destroys some of the fine information content in the picture. This type of coding, especially for television and video teleconferencing, has been justified with claimed characteristics of the human eye. The removed subjective redundancy is beyond normal visual sensitivity and its loss causes no complaints from most, if not all, human observers. The degree of approximation is said to represent in some sense the fidelity criterion for the encoding scheme in question. Entropy coding, source coding, and other compression coding methods are used to reduce the subjective redundancy, given the above fidelity criterion.

The partition of picture coding into information preserving and approximation classes is far from the entire story. As Figure 2 shows, there are many other compression codes that belong to both classes. That is, under certain circumstances they can exhibit properties of either one, or of both classes. Signal domain and transform domain codes are common examples, as illustrated next.

Signal domain coding applies to the luminance and chrominance voltages as derived directly from the photo-detector scan of the picture. The sample values, also called the pel readings, can be treated individually or as groups or vectors. Furthermore, as in the case of TV, distinction must be made

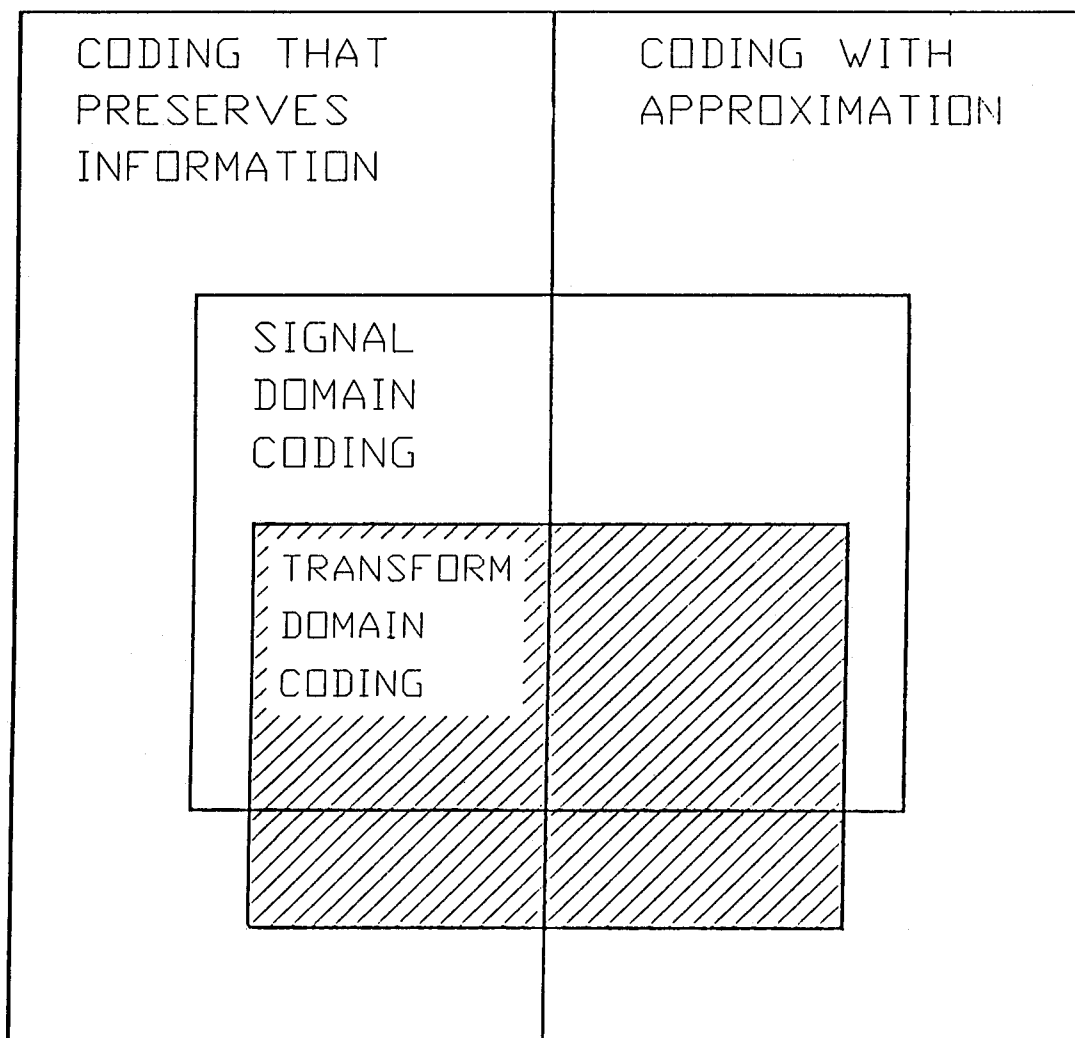


Figure 2. Simplified overview of different image coding classes.

between procedures that take place within the same frame (intraframe) and those that involve several successive frames (interframe). Simple examples of intraframe procedures are the spatial differential or predictive quantization schemes. Interframe functions are apparent in temporal differential and predictive schemes. Motion detection and motion compensation methods provide more sophisticated examples of interframe processes.

Transform domain coding entails signal transformation before encoding. Almost universally, fast discrete or finite field transforms are employed. That implies that the image field must be partitioned (or divided into blocks or subfields) and the analog photo-detector readings must be quantized and thresholded. The most frequently mentioned transforms are the Discrete Fourier Transform (DFT), the Discrete Cosine Transform (DCT), the Walsh-Hadamard Transform (WHT), and the Karhunen-Loeve Transform (KLT) (Andrews, Kane, and Pratt, 1969; Ahmed and Rao, 1975; Green, 1983).

For video images with stationary properties, such as with a fixed autocovariance function, the optimum transform (in the sense of least MSE) is the KLT (Jain, 1976; Clarke, 1985). This transform has uncorrelated real coefficients that happen to be nearly stationary for even rapidly changing scenes. The number of significant nonzero coefficients is minimal. Optimal or near-optimal data compression would be possible with KLT, but it requires that the autocovariance function be known for a frame to be transmitted and received. This is a serious problem for KLT. Adaptive methods that estimate the autocovariance function and transmit the needed updates have been suggested. But that, in turn, entails considerable complexity and processing delays.

The other transforms, namely DFT, DCT, and WHT, do not need any a priori knowledge of the stationary correlation properties of the picture sequence. However, their data compression performance is also inferior to KLT. The familiar DFT exhibits the Gibbs phenomenon, which often can be recognized as an objectionable grating-like distortion on a uniformly subdivided screen. For this and other reasons, neither DFT nor WHT have found much acceptance in recent image compression systems. From a practical point of view, the preferred transform seems to be the DCT (Ahmed, Natarjan, and Rao, 1974; Kato, Mukawa, and Okubo, 1987). It avoids the Gibbs problem, it has the same number

of real coefficients as the DFT, it is easy to implement, and its performance for first-order Markov processes is only slightly inferior to the MSE-optimal KLT (Jain, 1979).

Figure 2 also shows that there are coding schemes that belong both to signal and transform domains. These may be called hybrid schemes. Various versions are possible. Difference signals, derived spatially in the intra-frame format or temporally in the interframe format, are products of the signal domain. They can be quantized and thresholded in the same signal domain, but encoded for redundancy reduction in the transform domain. Or for nondifferential signals, some low frequency, such as near DC, encoding could be carried out before transforms, only to be followed by differencing, thresholding, quantization of the coefficients, and perhaps higher frequency coding in the transform domain. The arguments for and against specific signal processing chains involve, among other factors, ease and economics of implementation plus considerations of how dominant the low-band energy is in contrast to the energy content in the higher frequency bands.

5. MOTION COMPENSATION

Assume that the object motion exhibits a predictable trend. Then to restore the full image, the receiver may require only minor corrections to specify the deviations from the trend or the extent of the trend. If the motion is uniform over long periods of time or over large parts of the image, then significant data compression can be realized by various techniques. In general, the techniques associated with movement recognition, prediction, and correction are known as motion compensation.

To achieve optimal motion compensation for transmitted images, many disciplines of science and engineering should be utilized. As noted, the number of channel bits must be minimized while maintaining a subjectively acceptable visual quality. Thus, such fields as optics, psychophysics (i.e., the interaction between the human eye, the optical nerves, and the brain), color metrics, motion dynamics, signal processing, transmission systems, and display technology may all pertain here.

The rather lengthy list of perception factors (see Section 2) and the associated explanations (see the Bibliography in Section 9) indicate that

their interaction is, unfortunately, both extensive and complex. To keep this section reasonably short and lucid, we shall be concerned here only with the engineering aspects of live (not fixed frame) video signal processing and transmission.

Interframe prediction

Successive video frames usually differ from each other. The differences are caused by several phenomena, such as changes in lighting, changes in shades of color, or emerging or vanishing patterns. New patterns are created when objects come into or go out of focus, or disappear/appear behind each other. One also observes directionless random flutter or agitation of picture elements, as well as discernible movements (translations, rotations) of entire objects or their parts.

In motion compensation, one is usually limited to local treatment of reasonably well-defined moving edges. To benefit from the presence of such manageable or benign motion, a number of basic functions must be performed:

- the initial identification of the edge movement, including edge contrast changes
- the estimation of movement direction and speed
- the extrapolation or prediction of the motion for future frames
- the final correction to account for the real edge movement not being in total agreement with the previously predicted movement

All motion compensation systems incorporate these functions. In particular, the same prediction algorithm is nominally known at both ends of the communications link. That means that after the trends have been set, only the final correction bits need to be sent over the link. If the edge motion is slow and predictable, only a few update bits need to be sent. Therein lies the potential advantage.

The disadvantage lies in implementation complexity (e.g., memory, delays, costs) and the fact that many visual scenes contain an arbitrary number of diffuse edges whose motion may be hard to characterize, estimate, and correct. The difference between systems is found largely in the manner of implementation at the transmitter and receiver terminals, plus their cost and performance.

Systems that transmit several fields per frame, such as the NTSC 2:1 format, need not wait for the end of a current frame to proceed with the prediction and correction functions. Partial information gained from the last available field can be used in an interfield/intraframe fashion to improve the performance (notably, the delay).

As far as chrominance is concerned, the trend seems to be to avoid motion compensation (viz., edge processing) in the full, three-color component domain. Instead, only the luminance (or brightness) is typically subjected to the edge motion compensation algorithms. Near the moving edge, the color components are either attenuated, given a gain proportional to luminance, or simply ignored. This is justified by two arguments. First, the information bits associated with color per se form a relatively small fraction, perhaps 1/8 to 1/4, of the total number of transmitted bits. Second, the fidelity of color reproduction near well-defined moving edges appears not to be essential for the human eye.

Predictors can be variously classified, such as linear, nonlinear, fixed, adaptive, intrafield, intraframe, interframe, and so forth. Most of them have more theoretical than practical significance. A notable exception may be the Graham's predictor (Graham, 1967). Originally proposed for linear intraframe prediction, it follows simple selection (not linear addition) rules from adjacent pels. The distinctive feature is the following. Under linear addition rules, weighted linear sums of previously received pel values constitute the predicted value for a future pel. Graham's predictor is not a linear adder. Given specific logic rules, it selects one particular pel from a domain of received pels as the predicted value for the next pel.

Several extensions of the basic Graham's principle have been studied for interframe operation with moving images. The different versions are a consequence of choosing different selection domains from the product space of

previous pel(s) on the same line, previous line(s) in the same field, previous field(s), same frame, and previous frame(s).

As expected, prediction based on interframe differences appears most efficient when pictures are quite uniform, with little detail, and with little (slow) motion. As the detail and the motion increases, the advantage first shifts from interframe to interfield prediction techniques. But then new and bigger problems arise. Unless the movement is consistently predictable (which is seldom the case), further increase of motion leads to deterioration of all except the simplest intrafield predictors. In the limit of total visual turmoil, such as wideband noise, no predictor can work well.

There are two main reasons for the above trend. First, a high degree of motion implies less correlation between a present pel and all its predecessor pels (Connor and Limb, 1974; Dubois and Sabri, 1984). And second, due to time integration by the video camera, the spatial correlation in the direction of motion increases to the point where it seems to dominate all other redundancies in the television signal (Haskell, Mounts, and Candy, 1972; Gimlett, 1975; Girod, 1987).

The effectiveness of prediction schemes is expressed in terms of their respective prediction errors. A single number, namely the entropy of the residual error signal, shows in bits per pel the amount of information contained in the random process generated by the errors. Entropy represents the lowest residual bit rate that must be transmitted in order to completely restore the image at the receiver.

Table 1 summarizes typical published results on the performance of realistic frame difference (interframe) and pel difference (intraframe) prediction schemes (Netravali and Haskell, 1988). Interfield/intraframe and related hybrid techniques are typically within the bounds set by these two cases. The numbers given represent the entropy in bits per moving area picture element. The values are approximations, as they clearly depend on many parameters whose nature and impact are only vaguely understood.

The three parameters depicted in Table 1 are:

- (1) The nature of the televised moving visual object. Two cases--the human head and a live scene (on stage or in a natural background)--are presented.

Table 1. Estimates of Prediction-error Entropy in Bits per Moving Area Pel

Number of Quantizing Levels	Speed in Pels per Frame Time	Object: Human Head		Object: Live Scene	
		Frame Difference Prediction	Pel Difference Prediction	Frame Difference Prediction	Pel Difference Prediction
35	25	2.0 - 2.4	3.5 - 3.8	2.2 - 2.5	2.5 - 3.2
	1	2.6 - 3.2	2.6 - 3.4	2.8 - 3.6	2.8 - 3.2
	4	3.0 - 3.5	2.0 - 2.6	3.0 - 3.4	1.8 - 2.2
511	25	4.5 - 5.0	6.4 - 6.8	4.5 - 5.4	5.5 - 6.4
	1	5.3 - 6.2	5.3 - 6.5	5.6 - 6.7	5.5 - 6.2
	4	5.8 - 7.0	4.5 - 5.5	5.6 - 7.0	4.5 - 6.0

- (2) The speed of the moving object, measured in number of pel displacements per frame time. Three values, that is, .25, 1, and 4 pels per frame time, are assumed.
- (3) The number of quantizing levels, 35 or 511, for the digitized video. The not infrequently mentioned 6-, 7-, and 8-bit quantizations should yield comparable results.

One would like to conclude something definite about the real pictures encountered in practice. The following conclusion appears valid. Being random and different in texture and motion, no matter how quantized or predicted, the number of bits per moving area pel should be somewhere between 2 and 6.

As an example, consider the standard NTSC television in digital format, which scans 14,430 lines per second. If 512 pels were assigned per line, then nearly 7.4 million pels would be scanned each second. Clearly, if these were all "moving area pels," the prediction situation would call for transmission rates between 15 and 44 Mb/s. One could assign fewer pels per line, say 256, and hope that with high probability no more than 20 percent of the screen will be in motion at the same time. Such and similar scenarios could reduce the required communications rate by a factor of 10. According to Table 1, the prediction technology would still call for 1.5 to 4.4 Mb/s throughput. This is in the general ballpark with other recent ideal targets for digital TV. It remains unclear, however, whether the resultant video quality would be acceptable.

Motion estimation

The first step in motion estimation is to identify regions that contain motion and those that contain none. When a frame is divided into a regular checkerboard of subfields, the task reduces to identification of those subfields that are static and those that are not. The encoder stores the previous frame. Then, either one at a time or in parallel, the pels of each subfield of the current frame are compared with those in the stored frame. A subfield is declared static if there is no difference or if the difference is

sufficiently small. Otherwise, the subfield is said to contain changes associated with motion. The encoder must, of course, tell the decoder which are the motion subfields, followed by subfield-unique specification of the motion.

The identification of the movement subfields is the easy part. As an illustration, Figure 3 shows a frame divided into 36 subfields. Of these, four are shown dark to indicate the presence of some motion. The other 32 cells are static. The encoding of the frame can be a simple listing of static gaps (i.e., the number of contiguous unchanged subfields) interspersed with the individually encoded motion subfields. For example, the frame in Figure 3 could be encoded as

Gap=14, C(15), Gap=0, C(16), Gap=4, C(21), Gap=4, C(26), Gap=10,

where C(j) represents the coding for subfield j. This run-length structure can be variously modified. In particular, the zero gaps and the final gap can be deleted as they appear redundant for most receiver implementations.

The next item to be considered is the motion estimation within a given subfield. Several simplifying assumptions appear both typical and necessary from the start. Only a single moving object can be allowed. Its motion is assumed to be a simple translation in the focal plane. This excludes many complicated motions such as rotations, movements toward and away from the camera, an object moving behind or emerging from behind another, and so forth. These assumptions are particularly poor for large subfields. As the block-size of the subfield is decreased, the validity of the assumptions improves. But then the computational complexity, storage, and delays increase at both ends of the link. The associated transmission overhead can also become significant. For an arbitrary population of motion pictures, the best block-size or its prerequisite tradeoffs are only partially understood.

If $f(x,y,t)$ is the image signal at point (x,y) and at time t , then a simple translation is defined as

$$f(x,y,t) = f(x-a,y-b,t-\tau).$$

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

Figure 3. Identification of motion (dark) subfields among static subfields.

Here, τ is a known frame delay period. Depending on the system used, its value can be 1/15, 1/30, or 1/60 of a second. Vector $D=(a,b)$ represents the displacement during the frame period. It is the unknown to be estimated. Or equivalently, one seeks the velocity vector $V=D/\tau$.

Of the many estimation methods studied, two broad types stand out. They are the Block (or Subfield) Matching and Recursive Methods (Limb, Rubinstein, and Thompson, 1977; Netravali and Limb, 1980). We shall discuss them briefly next.

In the Block Matching Method, one searches for the displacement vector $D=(a,b)$, which matches a new block to the aggregate of blocks stored from the previous frame. Normally, a "match" is declared when the block correlation exceeds a preset threshold. Figure 4 shows that nine adjacent previous blocks may be involved in a process that tests various choices:

$$\begin{aligned}(a,b) = & (0,0), (0,\pm 1), \\ & (\pm 1,0), (\pm 1,\pm 1), \\ & (\pm 2,0), (\pm 2,\pm 1), (\pm 2,\pm 2), \\ & (\pm 3,0), (\pm 3,\pm 1), (\pm 3,\pm 2), (\pm 3,\pm 3), \\ & \dots ,\end{aligned}$$

where a and b denote counts of displacement pels.

The search sequence need not be consecutive, as shown, but can be adaptive by going (a) first in the direction of the displacement found for the previous frame and (b) generally following the direction of largest current correlation. Several specific algorithms have been proposed to reduce the number of steps in the search, such as the 2D-logarithmic, the three step, and the modified conjugate direction searches (Netravali and Haskell, 1988; Netravali and Robbins, 1980). Also, several methods have been studied to either avoid or to simplify the rather lengthy pel-by-pel block correlation for each shift.

The Recursive Methods put more emphasis on strategy (a) noted under Block Matching. That is, they estimate local displacements iteratively by utilizing displacement estimates from previous frames. An argument favoring this approach relies on the premise that a subfield may contain only a single

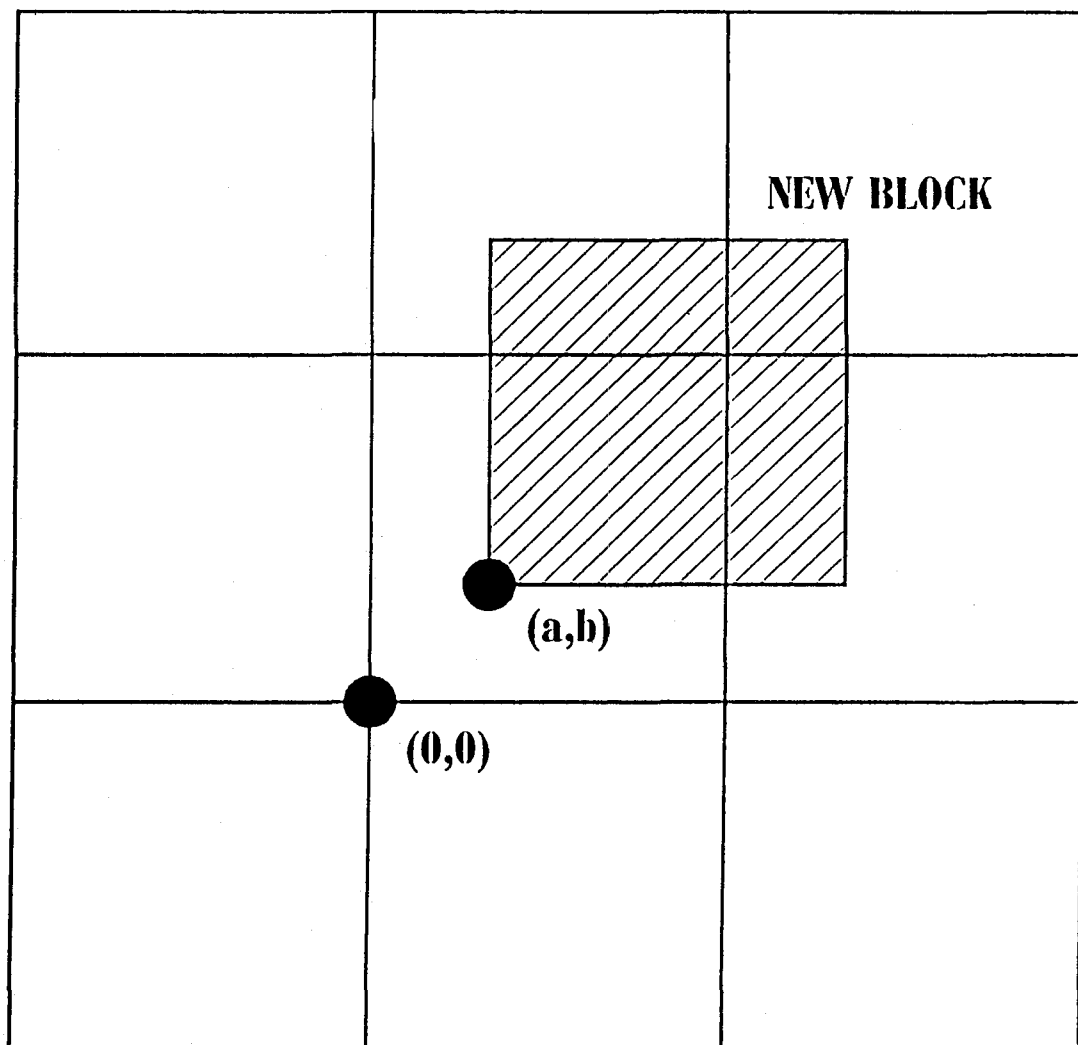


Figure 4. The role of displacement vector (a,b) in block matching.

moving edge. But a straight edge can only move in the direction perpendicular to itself. Thus, a search over the entire surrounding neighborhood may be wasteful.

Different algorithms result from different definitions of the recursion domain. Such domains could be individual pels, small groups of pels, parts of subfields (blocks), or entire subfields. Moreover, the iteration can traverse the frame along different paths such as following the scan lines, or across lines, block to block, or even frame to frame.

All Recursive Methods face the problem of convergence time, in addition to the small probability of wrong convergence. If it were not for these issues, the Recursive Methods could have the theoretical capability of permitting multiple moving objects, as noted by some authors (Netravali and Haskell, 1988).

Gain compensation

Gain compensation is a technique wherein the displacement or more broadly, the motion compensation, is augmented by a variable gain factor (G). Formally, the displaced image intensity is represented as

$$f(x,y,t) = G f(x-a,y-b,t-\tau),$$

where G stands for spatially and/or temporally slowly varying gain. Physically, G can be thought to come from illumination, shadowing, and related effects that gradually evolve as frames follow one after the other. Gain variations can be tracked by gradient-type algorithms that seek to minimize the square of the compensation error. The complexity of such methods is claimed to be a fraction of that for motion compensation. By itself, gain compensation yields compression numbers (in bits per pel) that are superior to frame difference prediction (see Table 1), but inferior to stand-alone displacement compensation. The best performance will apparently occur when one combines motion and gain compensations. But then, the complexity of the implementations may become a burden to the transmitting and receiving facilities.

Bit-error-rate effects on quality of motion compensated transmissions

Relatively more damage to reconstructed picture quality is caused by transmission errors whenever the data are compressed. The most familiar examples occur for differential pulse code modulation (DPCM) and related coding schemes, where previous signal values are used as references to predict future values (Abbott, 1971; Pratt, 1972; Brainard and Othmar, 1987). That means that any bit (pel) error will affect an indeterminate number of subsequent bits (pels). This is known as output error propagation and, unless one takes appropriate precautions, it can have serious effects on all intraframe and interframe prediction systems discussed earlier.

There are two main solutions. One is to ensure that the channel BER is sufficiently low, so that error-caused image blurring or streaking is practically never seen. This can be engineered by providing sufficiently high signal-to-noise ratio or by one of many error-control means (such as forward error-correction coding). The second solution is to provide predictor and motion compensator designs that limit the spread of error propagation.

Consider the communication links typically provided by long-distance carriers, with BER in the range of 10^{-6} to 10^{-5} . If the compressed video data are sent at 1.5 Mb/s, then one could expect on the average some 1.5 to 15 error streaks per second. Even if these streaks were limited individually to a single scanning line on the screen, their effect would be quite detrimental to good-quality television. By the way, the above numbers would imply the occurrence of one streak, on the average, in 2 to 20 frames, assuming a 30-Hz frame rate. For higher video bit rates, the number of error-caused streaks would be proportionally higher.

To reduce the number of output errors to something like one event per minute, the channel BER in the 1.5 Mb/s example should not exceed 10^{-8} . This could be achieved with forward-acting error correction. There are many such error-correcting codes known. The more promising ones have been implemented and tested. Some code equipment is available for information rates above 15 Mb/s. Needless to say, such channel-coding strategies add redundancy. They require increased transmission rates. A typical code, like the 1/2-rate convolutional code (Viterbi and Omura, 1979), would require a doubling of the transmission speed, or 3 Mb/s to carry the aforementioned 1.5-Mb/s information

rate. For more details on the topic, see the rather extensive technical literature on specific error-correction and detection coding methods or on the general topic of channel error control (Clark and Cain, 1982).

The technical issues for two-dimensional interframe prediction and error truncation are complex. Some schemes are known to be unstable, that is, the after effects of a transmission error may propagate through at least one screen. Good criteria for video predictor stability are said to be "not known." On the other hand, certain "reasonably stable" two-dimensional predictors are known (Connor, Pease, and Scholes, 1971). They exhibit a characteristic rate of decay, where the aftereffect of an error event vanishes outside a small fraction of the screen.

One can further divide the error effects into two classes: those that represent incorrect pel values, and those that select wrong prediction or compensation rules. For interframe techniques, the second class seems to represent a more serious threat. Unfortunately, it also seems to be the least understood. At this time there appears to be only one easy way out. When a rule violation is detected, one stops the algorithm in progress and returns to "start". The screen display may briefly be left blank or it may repeat the previous frame or parts thereof.

Practical potential for motion compensation

A key to practical implementation and marketing of compression systems discussed above must be in finding an acceptable compromise between performance and complexity. Performance has two dimensions that run counter to each other and require another compromise. Those dimensions are the end-user perception of image quality (see Section 2) and the data rate compression achieved. System complexity eventually boils down to the cost of encoding and decoding devices.

In Table 1, which under qualified circumstances pertains to pel and frame difference predictors, a typical compression range from 2 to 6 bits/pel is indicated. Application of adaptive motion and gain compensation techniques can reduce the data rate further. Depending on the classes of processors and transforms involved, average values between 1.0 and 1.5 bits/pel have been alleged for "excellent" color picture reproduction. Others infer that .2 to

1.0 bits/pel may eventually suffice for "practical or tolerable" picture quality.

For purposes of discussion, consider an illustrative motion compensation scheme that does indeed achieve the 1.0 bits/pel compression. If the television raster consists of slightly more than 500 lines, each line between 350 and 400 pels, and the frame repetition rate is 30 Hz, then one expects the encoder to generate around 6 Mb/s on the average. To reduce this to the more practical DS1 (also known as T1) rate of 1.544 Mb/s, further compression is necessary. The present research and development efforts seem to leave the number of pels per raster essentially unchanged (except for high-definition television, which may have a larger pel count, and teleconferencing, which uses roughly half the lines per frame and 256 pels per line). One seeks further gains through signal compression, such as .5 bits/pel, and interlacing.

Interlacing is a common processing technique where a frame is divided into several (usually two) fields. Successive fields consist of alternate pel patterns. For 2:1 interlacing, a 30-Hz field rate corresponds to a 15-Hz frame rate. Under the above assumptions, the motion compensation systems may have a near-term potential for broadcasting or point-to-point distribution over DS1 network facilities.

Unfortunately, the above technical argument is quite one-sided. It neglects the previously mentioned, essential tradeoff between complexity and performance. One knows that digitized pictures undergo degradation as the degree of compression is increased. Given the best, realistic algorithms, one could proceed to decrease the data rate (in Mb/s) from 15 to 6, to 3, to 1.5. As this is done, the received image quality must gradually deteriorate. Exactly what this quality deterioration represents as a function of compression is not well understood at this time. In fact, there are even no objective methods or tools available to measure video quality.

The only workable methodology in current use involves subjective viewer tests. Subjective tests have several drawbacks, such as being expensive, time consuming, not repeatable, often biased (i.e., affected by user panel statistics), and so forth. Subjective tests reveal very different performance scores for different picture classes (e.g., humans, inanimate objects, nature

objective methods, if we had such, would also show comparable scoring variance for the different picture classes. If so, then the system application, general usage, and even marketing could influence the practical selection of compression and motion compensation schemes, and vice versa.

6. CURRENT SYSTEMS AND THEIR TRENDS

The theoretical trends of image data compression have spread to industrial laboratories and, to somewhat lesser degree, to systems applications. Practical developments, such as high-definition television (HDTV), video teleconferencing (VTC), and the general emphasis on the integrated services digital networks (ISDN), have stimulated the implementation of real image data compression products. This section presents a brief introduction to the current status of image data compression systems. We begin with the Rembrandt series of codecs (i.e., coders/decoders), which are typical state-of-the-art products available today.

Rembrandt series

Compression Labs, Incorporated (CLI), of San Jose, CA, produces and markets video transmission equipments called Rembrandt and Rembrandt 56. While the main part of the equipment consists of the video codecs, the CLI product series also incorporates high-speed data ports and provisions for acceptable resolution graphics.

Internally, the system breaks into two large, functionally separate, parts: the codec and the data terminal equipment (DTE). The codec contains quantizers, signal processors, encoders and/or decoders, buffer memories, and so on. We shall return to a more thorough discussion of the codecs later in this subsection.

The DTE provides compatible access to a variety of standard and customized telecommunication interfaces. Many options are available. It suffices to mention that the Rembrandt and Rembrandt 56 systems are distinguished on the basis of their communication channel data rates. Thus, the higher rate Rembrandt typically works over 384 kb/s, 768 kb/s, 1.544 Mb/s (or DS1, or T1), or even multiple T1 lines. Less commonly used seem to be the incremental rates specified by $384 + nx64$ kb/s, where $n = 1, 2, 3, \dots$. The

lower-rate Rembrandt 56 was originally targeted for the common carrier 56-kb/s lines. However, claims are now made that Rembrandt 56 works comparably well over a series of channel rates, such as 40, 48, 56, 64, and so forth, up to 384 kb/s. Needless to say, the video reproduction quality of the DS1-rate Rembrandt is visibly superior to the lower rate Rembrandt 56's.

The basic technology of the Rembrandt series appears to have its origins in the Scene Adaptive Coder (SAC) (Chen, 1981; Chen and Pratt, 1984). The SAC technique applies to both monochrome and color images. It is an intraframe process capable of generating fixed-rate data for transmission.

In the basic SAC concept, the image is divided into many small blocks or subfields of pels. A typical block may be 16x16 pels. A clean and simple raster design should ideally consist of whole multiples of these small blocks. Let the SAC image be 512 by 512 pels. Then one screen contains 1,024 blocks. Each block is separately transformed by a fast, two-dimensional Discrete Cosine Transform (DCT) (Chen, Smith, and Fralick, 1977). Figure 5 shows this cosine transform being performed on the image signal, $f(j,k)$, of the 16x16 subfield as the first of a number of processing steps at the transmitter.

The transform generates 256 real DCT coefficients, but only a small fraction of them need to be sent to the receiver. First, in a thresholding device, all coefficients below an adaptive threshold are set to zero. The threshold can start with some a priori reasonable value, but need not stay fixed very long. Adaptive changes in the threshold setting can be controlled by feedback from the rate buffer. This is a buffer with variable input rate. Its objective is to generate a constant output rate by varying the buffer contents with the aid of feedback. It turns out that a vast majority of the coefficients are "in the noise" and so fall below the threshold. Being zero, they need not be individually normalized, quantized, or encoded. This achieves a substantial part of the intraframe compression.

Next, the amplitudes are normalized. This is a scaling process, also under feedback control from the rate buffer, that adjusts the range of coefficients among consecutive blocks so that subsequent quantization can generate an appropriately decreased or increased number of bits.

Quantization maps the analog (floating point) numbers directly into rounded-off integers. No special conversion rules, tables, or feedback

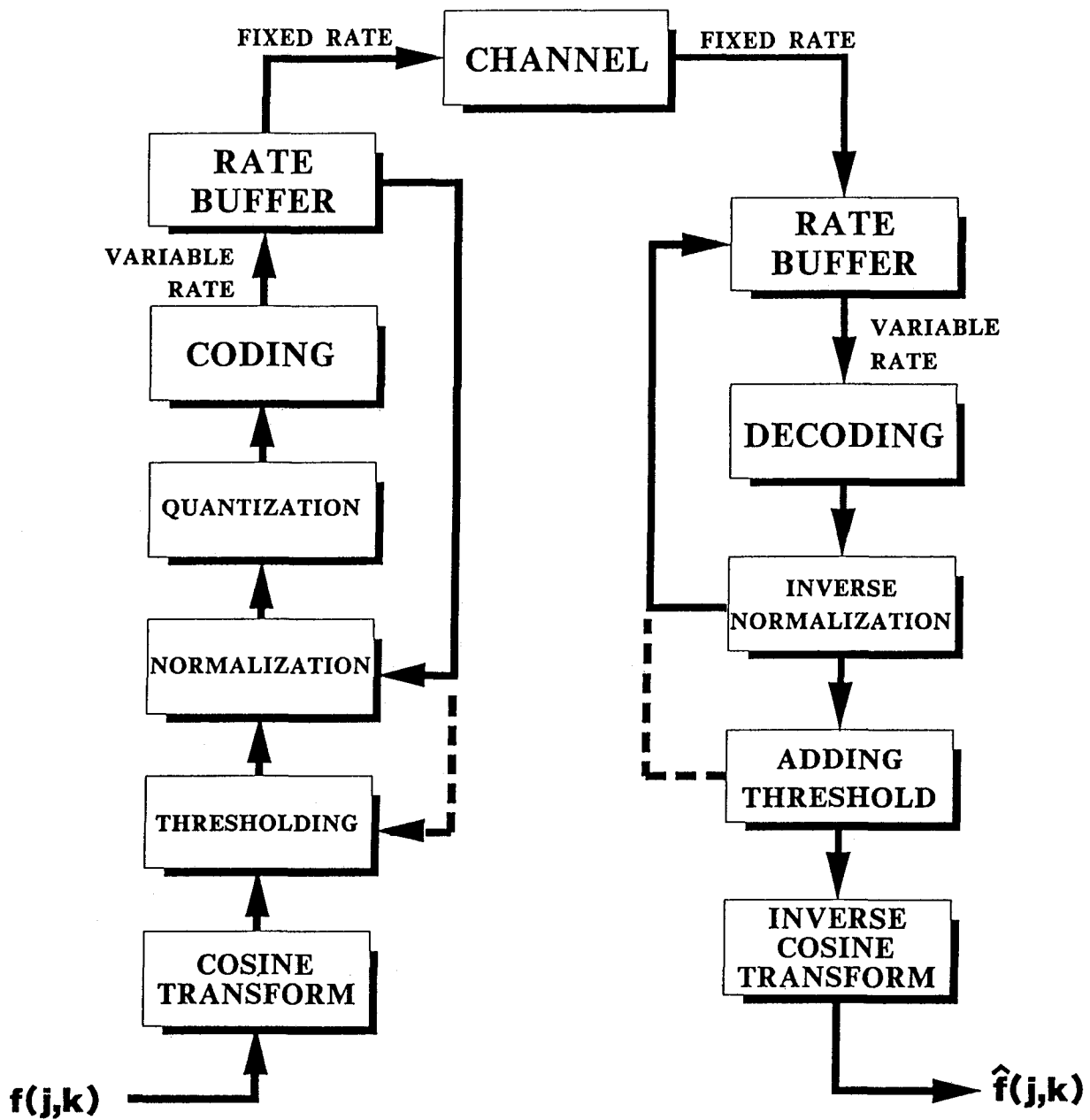


Figure 5. Basic diagram of the scene adaptive coder (from Chen, 1981).

controls are involved. Some of the normalized fractional coefficients may be rounded off to zero. This achieves further data compression.

The box labeled "coding" performs a limited amount of further compression by doing several things. First, the high-energy (which are the lowest frequency or dc) coefficients are treated separately from the higher frequency (or ac) coefficients. The dc term itself is truncated into nine bits. The amplitudes of the larger-than-zero ac terms can be encoded in a low redundancy Huffman code (Huffman, 1962). The zero sequences can be encoded in another efficient code, namely, the run-length code with the appropriate prefix properties. Or, perhaps even more effectively, joint hybrid or combination coding can be used for the product space of amplitudes and run lengths. The hybrid code can also be a Huffman code or one of many, so called, block codes. Any of the Bose-Chaudhuri-Hooquenghem (BCH) codes offers a good example (Peterson and Weldon, 1972). Forward error correction can be naturally incorporated with the BCH codes.

The final box shown at the transmitter (in Figure 5) is the rate buffer. Clearly, depending on the statistical characteristics of the image, its input is an encoded, variable data rate, signal. Its output, on the other hand, should have a fixed rate to take full advantage of the channel transmission capability, such as perhaps the commonly available 384 kb/s. The rate buffer performs this variable-to-fixed rate conversion by simultaneously (a) monitoring its own contents and (b) sending feedback signals to threshold and normalization devices to either slow down or speed up the bit generation process.

One can easily visualize a system that performs two or more passes on the same 16 x 16 block, where the initial pass(es) estimate the potential number of bits and where, after an adjustment of thresholding and normalizing scales, the final pass generates a practically constant rate transmitter output. However, it turns out that by paying careful attention to the first-in-first-out principle in the codec design, fixed rate outputs can be obtained by SAC in a single pass.

The SAC intraframe compression numbers are as follows. To encode 512 x 512 frames with 30-Hz repetition rate into a 1.544-Mb/s channel rate calls for around .2 bits/pel. This seems to be a design challenge for the

present state of the art. SAC proponents claim that color video can be compressed to .4 bits/pel, while black-and-white TV can be squeezed as far as .2 bits/pel. Thus, as described and in actual use, SAC appears to fall a little short of being a T1-compatible solution for NTSC color. Format changes such as frame rate or pel density reductions may provide some relief, but only at the cost of almost assured quality degradation.

But, as stressed elsewhere, the most relevant problem remains unsolved. It is simply not known quantitatively how image degradation depends on SAC algorithms and compression numbers. If one had a curve, a table, or even a very approximate rule of thumb for this degradation function, then more specific judgments could be made about SAC's performance over many practical communication channels (e.g., data links at 384, 768, 1544, etc., kb/s).

The receiver performs the inverse functions in the inverse order. First, the rate buffer receives fixed-rate data and outputs variable rate Huffman code. After Huffman decoding, inverse normalization and threshold restoral functions take place. Because of the adaptive thresholding and normalization done earlier at the transmitter, control data must be sent to the receiver to track the randomly adaptive process. Control bits also generate feedback to control the output rate of "hat" the rate buffer. Finally, the inverse cosine transform restores the replica, $\hat{f}(j,k)$, of the original image block $f(j,k)$.

Motion detection, prediction, and related interframe compression techniques are also applicable to the scene adaptive intraframe coders. Figure 6 outlines a color video SAC system that has been augmented with these techniques.

The red color (R-Y) and the blue color (B-Y) are sampled at a lower rate than the luminance (Y). The rate reduction is 1/4 in the horizontal, as well as in the vertical direction. That means that the total color samples add $2 \times 1/4 \times 1/4 = 1/8$ to the luminance sampling. The number of bits per sample becomes an issue only at the quantization stage. Therefore, if the three components follow the same signal path in Figure 5 preceding and through quantization, then the number of bits per color sample is the same as for the gray scale sample. In the Rembrandt series, color requires roughly 12.5 percent more transmission capacity as does monochrome.

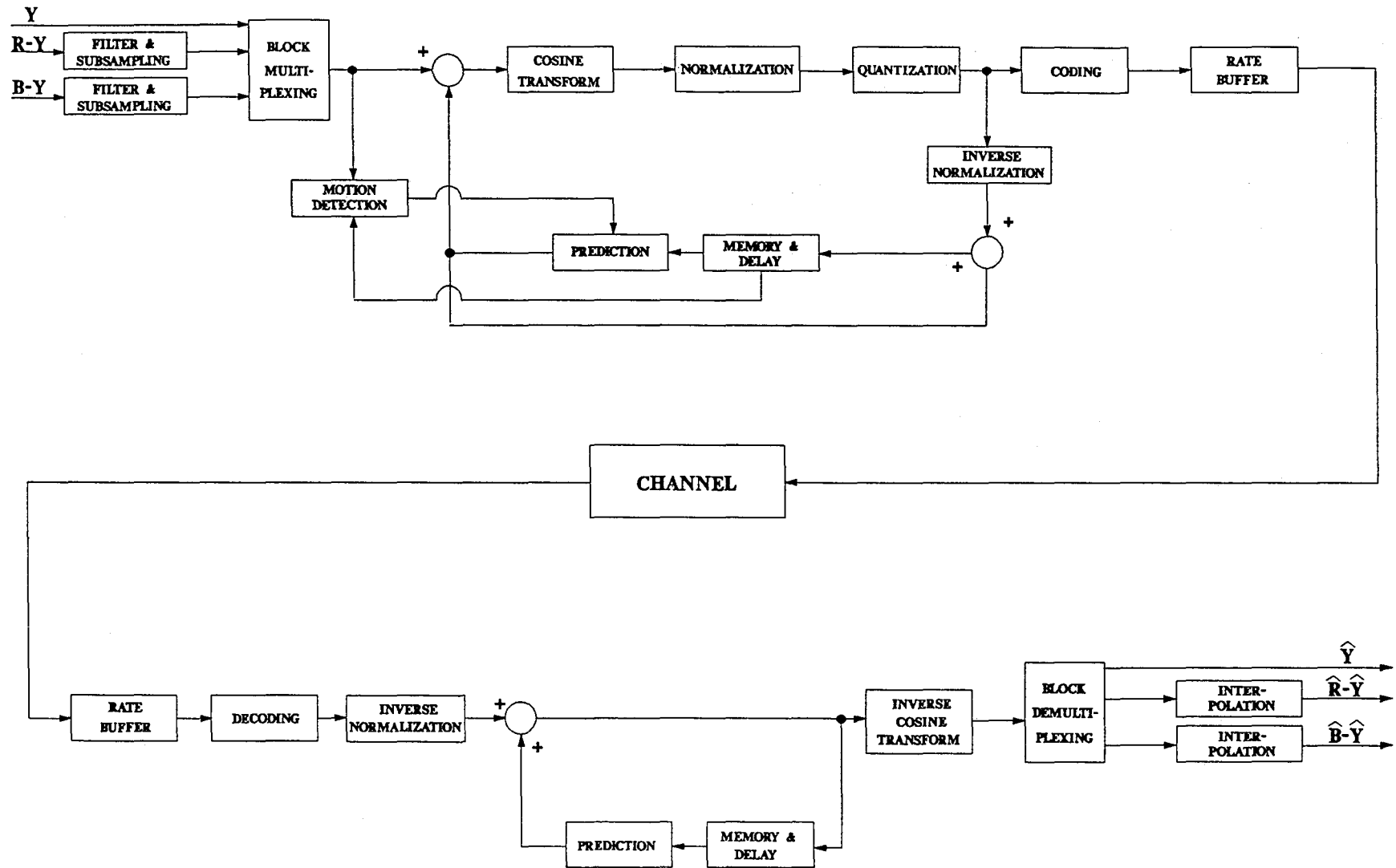


Figure 6. Combined interframe and intraframe scene adaptive coding system (as in Chen and Hein, 1987).

Note that the SAC system of Figure 5 forms the main forward chain of elements in Figure 6. The key new element in Figure 6 is the memory and delay unit, which stores the blocks from the previous frame. A special motion detection device compares every new (i.e., present) block with the previously stored block. According to established rules, the motion detector decides whether there is or is not excessive motion (changes with directional trend) in the block. If there is too much motion, the previously described SAC intraframe functions are used to encode the block in detail. If there is sufficiently little motion, the predictor unit is called into action. It produces frame-to-frame differences for the block in question. These relatively small differences are then DCT transformed, encoded, tagged by single bit motion indicators, and sent to the data terminal equipment (DTE).

The interframe coding results in a further compression advantage due to the presence of motion detection and compensation. An approximate 1/2 compression factor allegedly holds for 30 frames/second color video. This should enable potential transmissions over 1.544 Mb/s lines. Another 1/2 factor is said to be valid if the frame rate is lowered to 15 frames/second. Now one may be in a position to send compressed color NTSC signal over 768 or even 384 kb/s. But here, again, the image quality issue raises its head. At the present time it seems difficult to tell what encoding schemes will yield what performance and whether that is suitable for a given viewer class.

At the receiver, block-by-block decoding takes place again in the reverse order. The motion indicator bits resolve the uncertainty of which blocks get the ordinary SAC treatment, and which receive the storage, prediction, and motion compensation indicated by the feedback loop in Figure 6.

Other Systems

There are other systems, system techniques, and existing (and proposed) standards that should be discussed to convey a more complete picture of current technology.

Numerous organizations, here and abroad, are said to be active in the development and eventual implementation of image data compression systems. The list of organizations includes, but is not limited to: Stanford Research

International (SRI) or Sarnoff Laboratories, Zenith, MIT Advanced Television Group, AT&T Bell Laboratories, New York Institute of Technology, Del Rey Group, Faroudja Labs, North American Phillips, Nippon Electronic Corporation (NEC), Fujitsu, GEC, Sony, The Japanese Broadcasting Company NHK, and perhaps several others. A very active product line appears to be represented by the Multiple Sub-Nyquist Encoding (MUSE) schemes of NHK. While most of the MUSE versions are of high bandwidth, as typified by high definition television (HDTV) and related markets, there are also reports of the "Narrow MUSE," which is said to incorporate considerable signal compression.

Color video technologies, which extend beyond or in addition to the SAC scheme reported earlier, are alleged to include intrafield differential pulse-code modulation (DPCM), interfield adaptive DPCM, and various interfield transform codes, to mention a few (Bage, 1986; Chen and Hein, 1987; Gerken and Schiller, 1987). Remarkably high image data compression is achieved by still-frame codecs in video teleconferencing over narrow-band communication links.

Standardization work is done in the United States under the auspices of the American National Standards Institute (ANSI). Internationally, the main player is the International Consultative Committee for Radio (CCIR) and, to a slightly lesser extent, the International Consultative Committee for Telegraph and Telephone (CCITT). Standards are being developed for image data compression suitable for video teleconferencing, just like they were developed decades earlier for broadcast television. Examples of the latter are the standard TV systems called NTSC--in North and Central America and Japan; PAL--in parts of Europe and Africa, South America, Australia, Middle East, and Asia; and SECAM--in the remaining parts of Europe, Africa, and Asia (including France and USSR).

Shortage of time has prevented us from collecting data on these existing, potentially important, compression systems and their properties. The task, however, is essential and necessary to fully understand the video data compression field. Additional data collection should be undertaken in the future.

7. CONCLUSIONS

Research and development of image compression techniques has been an active field over the last few decades. The most significant and the most sophisticated work has been in the digital domain, where fortuitous convergences of information theory, signal processing, coding, integrated circuitry (e.g., very large scale integration or VLSI), software, and computer-related advances, have taken place. The theory predicts many potential options for significant data compression. Recent laboratory work at various institutions confirms the existence of many good alternative techniques. Without undue exaggeration, it now appears possible to implement any compression ratio one desires. For video teleconferencing, code rates from multiples of 1.544 Mb/s down to 40 kb/s can be tailored to suit the transmission rate of almost any existing communications link. Given all that capability, what is then the problem?

The major problem seems to be the performance of image compression systems. Here the term "performance" must be interpreted in more than its technical sense. The term must include end-user performance specification in addition to system parameters (such as pel resolution numbers, line and frame scan rates, compression algorithms, transmission rates, bandwidths, implementation complexity, memory requirements, device costs).

It is clear that viewer satisfaction or dissatisfaction with video image quality depends on subjective (human) factors. Starting with Section 2 of this study, it has been noted how important these subjective processes and parameters are in assessing image quality. Yet, subjective testing is an expensive and cumbersome task. The repeatability of results is often challenged.

What this industry really needs is an objective quality measurement method. The method should be an automated, preferably standardized, laboratory tool. It must serve relative system tests, where two or more image compression systems are compared against each other. And it must perform absolute tests on stand-alone systems with or without reference charts.

The development of such an objective method is likely to be quite difficult. However, its benefits are expected to be more considerable than the costs. Multitudes of system variations would be promptly and reliably evaluated. The best codecs would be identified and implemented.

8. REFERENCES

- Abbott, R.P. (1971), DPCM codes for video telephony using 4-bits per sample, IEEE Trans. Commun. Technol. COM-19, pp. 907-913.
- Ahmed, N., T. Natarjan, and K.R. Rao (1974), On image processing and a discrete cosine transform, IEEE Trans. Comput. C-23, pp. 90-93.
- Ahmed, N., and K.R. Rao (1975), Orthogonal Transforms for Digital Signal Processing (Springer-Verlag, New York, NY).
- Andrews, H.C., J. Kane, and W.K. Pratt (1969), Hadamard transform image coding, Proc. IEEE 57, pp. 58-68.
- Bage, M.J. (1986), Interframe predictive coding of images using hybrid vector quantization, IEEE Trans. Commun. COM-34, pp. 411-415.
- Berger, T. (1971), Rate Distortion Theory - A Mathematical Basis for Data Compression (Prentice-Hall, Englewood Cliffs, NJ).
- Blahut, R.E. (1972), Computation of channel capacity and rate-distortion functions, IEEE Trans. Inform. Theory IT-18, pp. 460-473.
- Bowen, E.G., and J.O. Limb (1976), Subjective effects of substituting lines in a video-telephone signal, IEEE Trans. Commun. COM-24, pp. 1208-1211.
- Boynton, R.M. (1979), Human Color Vision (Holt, Rinehart and Winston, New York, NY).
- Brainard, R.C., and J.H. Othmar (1987), VLSI implementation of DPCM compression algorithm on digital TV, IEEE Trans. Commun. COM-35, pp. 854-856.
- Budrikis, Z.L. (1972), Visual fidelity criterion and modeling, Proc. IEEE 60, pp. 771-779.
- Chen, W.H. (1981), Scene adaptive coder, Proceedings of IEEE Intern. Conf. on Commun. (ICC'81), 2, pp. 22.5.1-22.5.6.
- Chen, W.H., and D. Hein (1987), Recursive temporal filtering and frame rate reduction for image coding, IEEE J. Select. Areas Commun. SAC-5, pp. 1155-1165.
- Chen, W.H., and W.K. Pratt (1984), Scene adaptive coder, IEEE Trans. Commun. COM-32, pp. 225-232.
- Chen, W.H., C.H. Smith, and S. Fralick (1977), A fast computational algorithm for the discrete cosine transform, IEEE Trans. Commun. COM-25, pp. 1004-1009.
- Clark, G.C., and J.B. Cain (1982), Error-Correction Coding for Digital Communications (Plenum Press, New York, NY).

- Clarke, R.J. (1985), Transform Coding of Images (Academic Press, London, UK).
- Connor, D.J., and J.O. Limb (1974), Properties of frame-difference signals generated by moving images, IEEE Trans. Commun. COM-22, pp. 1564-1575.
- Connor, D.J, R.F.W. Pease, and W.G. Scholes (1971), Television coding using two-dimensional spatial prediction, BSTJ 50, pp. 1049-1061.
- Cornsweet, T.N. (1970), Visual Perception (Academic Press, New York, NY).
- Davisson, L.D., and R.M. Gray (1976), Data Compression (Dowden Hutchinson & Ross, Stroudsburg, PA).
- Dubois, E., and S. Sabri (1984), Noise reduction in image sequences using motion-compensated temporal filtering, IEEE Trans. Commun. COM-32, pp. 826-831.
- Gallager, R.G. (1968), Information Theory and Reliable Communication (John Wiley and Sons, New York, NY).
- Gerken, P., and H. Schiller (1987), A low bit-rate image sequence coder combining a progressive DPCM on interleaved rasters with a hybrid DCT technique, IEEE J. Select. Areas Commun. SAC-5, pp. 1079-1089.
- Gimlett, T.I. (1975), Use of activity classes in adaptive transform image coding, IEEE Trans. Commun. COM-23, pp. 785-786.
- Girod, B., (1987), The efficiency of motion-compensating prediction for hybrid coding of video sequences, IEEE J. Select. Areas Commun. SAC-5, pp. 1140-1154.
- Graham, D.N. (1967), Image transmission by two-dimensional contour coding, Proc. IEEE 55, pp. 336-346.
- Green, W.B. (1983), Digital Image Processing, A Systems Approach (Van Nostrand Reinhold, New York, NY).
- Haskell, B.G., F.W. Mounts, and J.C. Candy (1972), Interframe coding of video-telephone pictures, Proc. IEEE 60, pp. 792-800.
- Huffman, D.A. (1962), A method for the construction of minimum redundancy codes, Proc. IRE 40, pp. 1098-1101.
- Jain, A.K. (1976), A fast Karhunen-Loeve transform for a class of stochastic processes, IEEE Trans. Commun. COM-24, pp. 1023-1029.
- Jain, A.K. (1979), A sinusoidal family of unitary transforms, IEEE Trans. Pattern Analysis & Machine Intell. PAM-1, pp. 356-365.
- Kato, Y., N. Mukawa, and S. Okubo (1987), A motion picture coding algorithm using adaptive DCT encoding based on coefficient power distribution classification, IEEE J. Select. Areas Commun. SAC-5, pp. 1090-1099.

- Limb, J.O. (1973), Picture coding: The use of a viewer model in source encoding, BSTJ 52, pp. 1271-1302.
- Limb, J.O., C.B. Rubinstein, and J.E. Thompson (1977), Digital coding of color video signals - a review, IEEE Trans. Commun. COM-25, pp. 1349-1385.
- Mannos, J.L., and D.J. Sakrison (1974), The effects of a visual fidelity criterion on the encoding of images, IEEE Trans. Inform. Theory IT-20, pp. 525-536.
- Miyahara, M. (1988), Quality assessments of visual services, IEEE Communications Magazine 26, pp. 51-60 (October).
- Netravali, A.N., and B.G. Haskell (1988), Digital Pictures (Plenum Press, New York, NY).
- Netravali, A.N., and J.O. Limb (1980), Picture coding: A review, Proc. IEEE 68, pp. 366-406.
- Netravali, A.N., and J.D. Robbins (1980), Motion-compensated coding: Some new results, BSTJ 59, pp. 1735-1745.
- Peterson, W.W., and E.J. Weldon (1972), Error-Correcting Codes (MIT Press, Cambridge, MA).
- Pirenne, M.H. (1967), Vision and the Eye (Chapman and Hall, New York, NY).
- Pratt, W.K. (1972), Binary symmetric channel error effects on PCM color image transmission, IEEE Trans. Inform. Theory IT-18, pp. 636-643.
- Sakrison, D.J. (1977), On the role of the observer and a distortion measure in image transmission, IEEE Trans. Commun. COM-25, pp. 1251-1267.
- Seyler, A.J., and Z.L. Budrikis (1965), Detail perception after scene changes in television image presentations, IEEE Trans. Inform. Theory IT-11, pp. 31-43.
- Shannon, C.E. (1959), Coding theorems for a discrete source with a fidelity criterion, IRE Nat. Conv. Record, Part 4, pp. 142-163.
- Viterbi, A.J., and J.K. Omura (1979), Principles of Digital Communication and Coding (McGraw-Hill Book Co, New York, NY).

9. BIBLIOGRAPHY ON IMAGE DATA COMPRESSION

- Anderson, G.B., and T.S. Huang (1971), Piecewise Fourier transformation for picture bandwidth compression, IEEE Trans. Commun. Technol. COM-19, pp. 133-140.

- Andrews, H.C. (1970), Computer Techniques in Image Processing (Academic Press, New York, NY).
- Andrews, H.C., and C.L. Patterson (1976), Singular value decomposition (SVD) image coding, IEEE Trans. Commun. COM-24, pp. 425-432.
- Arnold, J.F., and M.C. Cavenor (1981), Improvements to the constant area quantization bandwidth compression scheme, IEEE Trans. Commun. COM-29, pp. 1818-1823.
- Arps, R.B. (1974), Bibliography on digital graphic image compression and quality, IEEE Trans. Inform. Theory IT-20, pp. 120-122.
- Arps, R.B. (1980), Bibliography on binary image compression, Proc. IEEE 68, pp. 922-924.
- Barba, J., N. Scheinberg, D.L. Schilling, J. Garodnick, and S. Davidovici (1981), A modified adaptive delta modulation, IEEE Trans. Commun. COM-29, pp. 1767-1785.
- Bello, M.G., A.S. Willsky, B.C. Levy, and D.A. Castanon (1986), Smoothing error dynamics and their use in the solution of smoothing and mapping problems, IEEE Trans. Inform. Theory IT-32, pp. 483-495.
- Bodson, D., and R. Schaphorst (1980), Compression and error sensitivity of two-dimensional facsimile coding techniques, Proc. IEEE 68, pp. 846-853.
- Bosworth, R.H., and J.C. Candy (1969), A companded one-bit coder for television transmission, BSTJ 48, pp. 1459-1480.
- Brainard, R.C., and J.C. Candy (1969), Direct-feedback coders: Design and performance with television signals, Proc. IEEE 57, pp. 776-786.
- Brainard, R.C., F.W. Mounts, and B. Prasada (1967), Low resolution TV: Subjective effects of frame repetition and picture replenishment, BSTJ 56, pp. 261-271.
- Brofferio, S., and F. Rocca (1977), Interframe redundancy reduction of video signals generated by translating objects, IEEE Trans. Commun. COM-25, pp. 448-455.
- Budrikis, Z.L., J.L. Hullet, and D.Q. Phiet (1971), Transient-mode buffer stores for nonuniform code TV, IEEE Trans. Commun. Technol. COM-19, pp. 913-922.
- Burt, P.J., and E.H. Adelson (1983), The Laplacian pyramid as a compact image code, IEEE Trans. Commun. COM-31, pp. 532-540.
- Cafforio, C., and F. Rocca (1976), Methods for measuring small displacements of television images, IEEE Trans. Inform. Theory IT-22, pp. 573-579.

- Candy, J.C., and R.H. Bosworth (1972), Methods for designing differential quantizers based on subjective evaluation of edge busyness, BSTJ 51, pp. 1495-1516.
- Candy, J.C., M.A. Franke, B.G. Haskell, and F.W. Mounts (1971), Transmitting television as clusters of frame-to-frame differences, BSTJ 50, pp. 1889-1917.
- Capon, J. (1959), A probabilistic model for runlength coding of pictures, IRE Trans. Inform. Theory IT-5, pp. 157-163.
- Chang, N.S., and K.S. Fu (1981), Picture query languages for pictorial data-base systems, IEEE COMPUTER 14, pp. 23-33 (November).
- Chang, S.K., and T.L. Kunii (1981), Pictorial data-base systems, IEEE COMPUTER 14, pp. 13-21 (November).
- Chellappa, R., and G. Sharma (1985), A model-based approach for estimation of two-dimensional maximum entropy power spectra, IEEE Trans. Inform. Theory IT-31, pp. 90-99.
- Chen, W.H., and C.H. Smith (1977), Adaptive coding of monochrome and color images, IEEE Trans. Commun. COM-25, pp. 1285-1292.
- Chiariglione, L. (1988), IVICO--A European effort toward economic video coding, CSELT Technical Reports XVI, pp. 75-78 (March).
- Chiariglione, L., S. Fontolan, M. Guglielmo, and F. Tommasi (1987), A variable resolution video code for low bit-rate applications, IEEE J. Select. Areas Commun. SAC-5, pp. 1184-1189.
- Chock, M., A.F. Cardenas, and A. Klinger (1981), Manipulating data structures in pictorial information systems, IEEE COMPUTER 14, pp. 43-50 (November).
- Chow, M.C. (1971), Variable-length redundancy removal coders for differentially coded video telephone signals, IEEE Trans. Commun. Technol. COM-19, pp. 923-926.
- Clark, J. (1980), A VLSI geometry processor for graphics, IEEE COMPUTER 13, pp. 59-68 (July).
- Connor, D.J., R.C. Brainard, and J.O. Limb (1972), Intra-frame coding for picture transmission, Proc. IEEE 60, pp. 779-791.
- Coueignoux, P., and R. Guedj (1980), Computer generation of colored planar patterns on TV-like rasters, Proc. IEEE 68, pp. 909-922.
- Danielsson, P.E., and S. Levialdi (1981), Computer architectures for pictorial information systems, IEEE COMPUTER 14, pp. 53-67 (November).

- Davisson, L.D. (1968), Data compression using straight line interpolation, IEEE Trans. Inform. Theory IT-14, pp. 390-394.
- De Lameillieure, J.L.P. (1987), A heuristic algorithm for the construction of a code with limited word length, IEEE Trans. Inform. Theory IT-33, pp. 438-443.
- Dukhovich, I.J., and J.B. O'Neal (1978), A three-dimensional spatial non-linear predictor for television, IEEE Trans. Commun. COM-26, pp. 578-582.
- Ehrman, L. (1967), Analysis of some redundancy removal bandwidth compression techniques, Proc. IEEE 55, pp. 278-287.
- Elnahas, S.E., and J.G. Dunham (1987), Entropy coding for low-bit-rate visual telecommunications, IEEE J. Select. Areas Commun. SAC-5, pp. 1175-1183.
- Eng, K.Y., and O.C. Yue (1981), Time compression multiplexing of multiple television signals in satellite channels using chirp transform processors, IEEE Trans. Commun. COM-29, pp. 1832-1840.
- Ericsson, S. (1985), Fixed and adaptive prediction for hybrid predictive/transform coding, IEEE Trans. Commun. COM-33, pp. 1291-1302.
- Farvardin, N., and J.W. Modestino (1986), On overflow and underflow problems in buffer-instrumented variable-length coding of fixed-rate memoryless sources, IEEE Trans. Inform. Theory IT-32, pp. 839-845.
- Fink, D.G. (1955), Color Television Standards: Selected Papers and Records of the National Television Systems Committee (McGraw-Hill, New York, NY).
- Frank, A.J., J.D. Daniels, and D.R. Unangst (1980), Progressive image transmission using a growth-geometry coding, Proc. IEEE 68, pp. 897-909.
- Fujio, T. (1981), High-definition television systems: Desirable standards, signal forms, and transmission systems, IEEE Trans. Commun. COM-29, pp. 1882-1890.
- Gallagher, N.C. (1978), Quantizing schemes for the discrete Fourier transform of a random time-series, IEEE Trans. Inform. Theory IT-24, pp. 156-163.
- Gharavi, H., and A.N. Netravali (1983), CCITT compatible coding of multilevel pictures, BSTJ 62, pp. 2765-2778.
- Goblick, T.J., and J.L. Holsinger (1967), Analog source digitization: A comparison of theory and practice, IEEE Trans. Inform. Theory IT-13, pp. 323-326.
- Goldberg, M., P. Boucher, and S. Shlien (1986), Image compression using non-adaptive spatial vector quantization, IEEE Trans. Commun. COM-34, pp. 180-187.

- Goldberg, M., and H.F. Sun (1986), Image sequence coding using vector quantization, IEEE Trans. Commun. COM-34, pp. 703-710.
- Goldberg, M., and H. Sun (1988), Frame adaptive vector quantization for image sequence coding, IEEE Trans. Commun. COM-36, pp. 629-635.
- Goodall, W.M. (1951), Television by pulse code modulation, BSTJ 30, pp. 33-49.
- Goyal, S.K., and J.B. O'Neal (1975), Entropy coded differential pulse-code modulation systems for television, IEEE Trans. Commun. COM-23, pp. 660-666.
- Habibi, A. (1971), Comparison of nth order DPCM encoder with linear transformation and block quantization techniques, IEEE Trans. Commun. COM-19, pp. 948-956.
- Habibi, A. (1974), Hybrid coding of pictorial data, IEEE Trans. Commun. COM-22, pp. 614-624.
- Habibi, A. (1977), Survey of adaptive image coding techniques, IEEE Trans, Commun. COM-25, pp. 1275-1284.
- Habibi, A. (1981), An adaptive strategy for hybrid image coding, IEEE Trans. Commun. COM-29, pp. 1736-1740.
- Habibi, A., and P.A. Wintz (1971), Image coding by linear transformation and block quantization, IEEE Trans. Commun. COM-19, pp. 50-62.
- Hang, H.M., and B.G. Haskell (1988), Interpolative vector quantization of color images, IEEE Trans. Commun. COM-36, pp. 465-470.
- Hang, H.M., and J.W. Woods (1985), Predictive vector quantization of images, IEEE Trans. Commun. COM-33, pp. 1208-1219.
- Harashima, S., T. Kumamoto, and T. Kitamura (1981), Japanese VIDEOTEX system "CAPTAIN" - Experimental service and users reaction outline, IEEE Trans. Commun. COM-29, pp. 1959-1968.
- Harrison, C.W. (1952), Experiments with linear prediction in television, BSTJ 31, pp. 764-783.
- Haskell, B.G. (1969), Computation and bounding of rate-distortion functions, IEEE Trans. Inform. Theory IT-15, pp. 525-531.
- Haskell, B.G. (1974), Frame-to-frame coding of television pictures using two-dimensional Fourier transforms, IEEE Trans. Inform. Theory IT-20, pp. 119-120.
- Haskell, B.G. (1975), Entropy measurements for nonadaptive and adaptive frame-to-frame, linear-predictive coding of videotelephone signals, BSTJ 54, pp. 1155-1174.

- Haskell, B.G., P.L. Gordon, R.L. Schmidt, and J.V. Scattaglia (1977), Interframe coding of 525 line, monochrome television at 1.5 Mbits/s, IEEE Trans. Commun. COM-25, pp. 1339-1348.
- Haskell, B.G., and R.L. Schmidt (1975), A low bit-rate interframe coder for video-telephone, BSTJ 54, pp. 1475-1495.
- Hassan, M.H., and P. Siy (1988), Thresholding based on learning theory, Proc. IEEE 76, pp. 1379-1381.
- Hayes, J.F., A. Habibi, and P.A. Wintz (1970), Rate distortion function for a Gaussian source model of images, IEEE Trans. Inform. Theory IT-16, pp. 507-509.
- Healy, D.J., and D.R. Mitchell (1981), Digital video bandwidth compression using block truncation coding, IEEE Trans. Commun. COM-29, pp. 1809-1817.
- Heiman, A., and K. Rose (1987), A look-up-based universal real-time transformer for image coding, IEEE J. Select. Areas Commun. SAC-5, pp. 1207-1213.
- Hochman, D., H. Katzman, and D.R. Weber (1967), Application of redundancy reduction to television bandwidth compression, Proc. IEEE 55, pp. 263-266.
- Huang, J., and P. Schultheiss (1963), Block quantization of correlated Gaussian random variables, IEEE Trans. Commun. COM-11, pp. 286-296.
- Huang, T.S. (1977), Coding of two-tone images, IEEE Trans. Commun. COM-25, pp. 1406-1424.
- Huang, T.S., and H.U. Koller (1975), Coding of multilevel graphics, IEEE Trans. Commun. COM-23, pp. 598-606.
- Huang, T.S., and A.B. Shahid-Hussain (1975), Facsimile coding by skipping white, IEEE Trans. Commun. COM-23, pp. 1452-1466.
- Huang, T.S., and O.J. Tetriak (1968), A pseudorandom multiplex system for facsimile transmission, IEEE Trans. Commun. Technol. COM-16, pp. 436-438.
- Huang, T.S., and O.J. Tetriak (1972), Picture Bandwidth Compression (Gordon and Beach, New York, NY).
- Hunter, R., and A.H. Robinson (1980), International digital facsimile coding standards, Proc. IEEE 68, pp. 854-867.
- Hwang, K., and K.S. Fu (1983), Integrated computer architectures for image processing and database management, IEEE COMPUTER 16, pp. 51-60 (January).

- Iinuma, K., Y. Iijima, T. Ishiguro, H. Kaneko, and S. Shigaki (1975), Interframe coding of 4-MHz color television signals, IEEE Trans. Commun. COM 23, pp. 1461-1466.
- Ishiguro, T., and K. Iinuma (1982), Television bandwidth compression transmission by motion-compensated interframe coding, IEEE Commun. Mag. 20, pp. 24-30 (November).
- Ismail, M.G., and R.J. Clarke (1981), Adaptive block/location coding of facsimile signals using subsampling and interpolation for pre- and postprocessing, IEEE Trans. Commun. COM-29, pp. 1925-1934.
- Jain, A.K. (1981), Image data compression: A review, Proc. IEEE 69, pp. 349-389.
- Jain, A.K. (1981), Advances in mathematical models for image processing, Proc. IEEE 69, pp. 502-528.
- Jain, J.R., and A.K. Jain (1981), Displacement measurement and its application in interframe image coding, IEEE Trans. Commun. COM-29, pp. 1799-1808.
- Johnsen, O., and A.N. Netravali (1981), Progressive transmission of two-tone images, IEEE Trans. Commun. COM-29, pp. 1934-1941.
- Johnsen, O., J. Segen, and G.L. Cash (1983), Coding of two-level pictures by pattern matching and substitution, BSTJ 62, pp. 2513-2543.
- Judice, C.N. (1977), Digital video: A buffer-controlled dither processor for animated images, IEEE Trans. Commun. COM-25, pp. 1433-1440.
- Kamangar, F.A., and K.R. Rao (1981), Interfield hybrid coding of component color television signals, IEEE Trans. Commun. COM-29, pp. 1740-1753.
- Kaneko, M., Y. Hatori, and A. Koike (1987), Improvements of transform coding algorithm for motion-compensated interframe prediction errors - DCT/SQ coding, IEEE J. Select. Areas Commun. SAC-5, pp. 1068-1078.
- Kaneko, H., and T. Ishiguro (1980), Digital television transmission using bandwidth compression techniques, IEEE Commun. Mag. 18, pp. 14-22 (July).
- Kang, C.H., H. Ryu, and C.H. Lee (1988), Realization of multiplierless coherent detection scheme in digital domain, Proc. IEEE 76, pp. 1394-1395.
- Kappagantula, S., and K.R. Rao (1985), Motion compensated interframe prediction, IEEE Trans. Commun. COM-33, pp. 1011-1015.
- Kidode, M. (1983), Image processing machines in Japan, IEEE COMPUTER 16, pp. 68-80 (January).

- Kim, T., and D.L. Neuhoff (1988), Delta codes for line drawings, IEEE Trans. Inform. Theory IT-34, pp. 400-416.
- Kishimoto, T., and Y. Sato (1981), Simultaneous transmission of voice and handwriting signals: "Sketchphone system," IEEE Trans. Commun. COM-29, pp. 1982-1986.
- Knowlton, K. (1980), Progressive transmission of grey scale and binary pictures by simple, efficient, and lossless encoding scheme, Proc. IEEE 68, pp. 885-896.
- Koga, T., Y. Iijima, K. Iinuma, and T. Ishiguro (1981), Statistical performance analysis of an interframe encoder for broadcast television signals, IEEE Trans. Commun. COM-29, pp. 1868-1876.
- Koga, T., and M. Ohta (1987), Entropy coding for a hybrid scheme with motion compensation in subprimary rate video transmission, IEEE J. Select. Areas Commun. SAC-5, pp. 1166-1174.
- Kretz, F. (1975), Subjectively optimal quantization of pictures, IEEE Trans. Commun. COM-23, pp. 1288-1292.
- Kretzmer, E.R. (1952), Statistics of television signals, BSTJ 31, pp. 751-763.
- Kunt, M., and O. Johnsen (1980), Block coding of graphics: A tutorial review, Proc. IEEE 68, pp. 770-786.
- Kutz, R.L., and J.A. Sciulli (1968), The performance of an adaptive image data compression system in the presence of noise, IEEE Trans. Inform. Theory IT-17, pp. 273-279.
- Landau, H.J., and D. Slepian (1971), Some computer experiments in picture processing for bandwidth reduction, BSTJ 50, pp. 1525-1540.
- Lee, C., and M. Nadler (1987), Predictive image coding with pseudo-Laplacian edge detector, IEEE J. Select. Areas Commun. SAC-5, pp. 1190-1206.
- Lei, T.R., N. Scheinberg, and D.L. Schilling (1977), Adaptive delta modulation systems for video encoding, IEEE Trans. Commun. COM-25, pp. 1302-1314.
- Lempel, A., and J. Ziv (1986), Compression of two-dimensional data, IEEE Trans. Inform. Theory IT-32, pp. 2-8.
- Levy, B.C., and J.N. Tsiksiklis (1985), A fast algorithm for linear estimation of two-dimensional isotropic random fields, IEEE Trans. Inform. Theory IT-31, pp. 635-644.
- Li, C.E., and K.R. Rao (1984), Composite-component transformation and predictive coding of the component color TV signal, IEEE J. Select. Areas Commun. SAC-2, pp. 353-358.

- Limb, J.O. (1967), Source-receiver encoding of television signals, Proc. IEEE 55, pp. 364-380.
- Limb, J.O. (1969), Design of dither waveforms for quantized visual signals, BSTJ 46, pp. 2555-2582.
- Limb, J.O. (1972), Buffering of data generated by the coding of moving images, BSTJ 51, pp. 239-259.
- Limb, J.O., and F.W. Mounts (1969), Digital differential quantizer for television, BSTJ 48, pp. 2583-2599.
- Limb, J.O., and H.A. Murphy (1975), Measuring the speed of moving objects from television signals, IEEE Trans. Commun. COM-23, pp. 474-478.
- Limb, J.O., and R.F.W. Pease (1971), A simple interframe coder for video telephony, BSTJ 50, pp. 1877-1888.
- Limb, J.O., R.F. Pease, and K.A. Walsh (1974), Combining intra-frame and frame-to-frame coding for television, BSTJ 53, pp. 1137-1173.
- Limb, J.O., and C.B. Rubinstein (1974), Plateau coding of the chrominance components of color picture signals, IEEE Trans. Commun. COM-22, pp. 812-820.
- Limb, J.O., and C.B. Rubinstein (1978), On the design of quantizers for DPCM coders: A functional relationship between visibility, probability, and masking, IEEE Trans. Commun. COM-26, pp. 573-578.
- Lippel, B., and M. Kurland (1971), The effect of dither on luminance quantization of pictures, IEEE Trans. Commun. COM-19, pp. 879-888.
- Lippel, B., M. Kurland, and A.H. March (1971), Ordered dither patterns for coarse quantization of pictures, Proc. IEEE 59, pp. 429-431.
- Lloyd, R. (1982), Least squares quantization in PCM, IEEE Trans. Inform. Theory IT-28, pp. 129-137.
- Mark, J.W., and J.O. Limb (1984), Integrated voice/data services on Fasnnet, AT&T Tech. J. 63, pp. 307-336.
- Matsui, K., M. Achida, and T. Fukinuki (1981), High-speed transmission of sequential freeze-pictures by extracting changed areas, IEEE Trans. Commun. COM-29, pp. 1977-1981.
- Matsumoto, M., H. Ochi, and S. Yoshino (1981), A high-speed facsimile apparatus for satellite communication, IEEE Trans. Commun. COM-29, pp. 1952-1958.
- Max, J. (1960), Quantization for minimum distortion, IRE Trans. Inform. Theory IT-6, pp. 7-12.

- Maxemchuck, N.F., and J.A. Stuller (1979), An adaptive intraframe DPCM code based upon nonstationary image model, *BSTJ* 58, pp. 1395-1413.
- Maxemchuck, N.F., and J.A. Stuller (1979), Reduction of transmission error in adaptively predicted, DPCM encoded picture, *BSTJ* 58, pp. 1413-1425.
- Meiri, A.Z., and E. Yudilevich (1981), A pinned sine transform image coder, *IEEE Trans. Commun.* COM-29, pp. 1728-1735.
- Mertz, P., and F. Gray (1934), A theory of scanning and its relation to the characteristics of the transmitted signal in telephotography and television, *BSTJ* 13, pp. 464-515.
- Meyr, H., H.G. Rosdolsky, and T.S. Huang (1974), Optimum run length codes, *IEEE Trans. Commun.* COM-22, pp. 826-835.
- Millard, J.B., and H.I. Maunsell (1971), Digital encoding of the video signal, *BSTJ* 50, pp. 459-480.
- Miller, K.S., R. Raghavan, and M.M. Rochwarger (1985), Invariance in moving target detection, *IEEE Trans. Inform. Theory* IT-31, pp. 69-80.
- Mitchell, O.R., and E.J. Delp (1980), Multilevel graphics representation using block truncation coding, *Proc. IEEE* 68, pp. 868-873.
- Mitchell, O.R., and A.J. Tabatabai (1981), Channel error recovery for transform image coding, *IEEE Trans. Commun.* COM-29, pp. 1754-1762.
- Miyahara, M. (1975), Analysis of perception of motion in television signals and its application to bandwidth compression, *IEEE Trans. Commun.* COM-23, pp. 761-766.
- Modestino, J.W., and V. Bhaskaran (1981), Robust two-dimensional tree encoding of images, *IEEE Trans. Commun.* COM-29, pp. 1786-1798.
- Modestino, J.W., V. Bhaskaran, and J.B. Anderson (1981), Tree encoding of images in the presence of channel errors, *IEEE Trans. Inform. Theory* IT-27, pp. 677-697.
- Modestino, J.W., and D.G. Daut (1979), Combined source-channel coding of images, *IEEE Trans. Commun.* COM-27, pp. 1644-1659.
- Moorhead, R.J., S.A. Rajala, and L.W. Cook (1987), Image sequence compression using a pel-recursive motion-compensated technique, *IEEE J. Select. Areas Commun.* SAC-5, pp. 1100-1114.
- Morita, H., and S. Arimoto (1983), SECT - A coding technique for black/white graphics, *IEEE Trans. Inform. Theory* IT-29, pp. 559-570.
- Mounts, W.F. (1969), A video encoding system using conditional picture element replenishment, *BSTJ* 48, pp. 2545-2554.

- Mounts, F.W., E.G. Bower, and A.N. Netravali (1979), An ordering scheme for facsimile coding, *BSTJ* 58, pp. 2113-2128.
- Mounts, F.W., A.N. Netravali, and B. Prasada (1977), Design of quantizers for real-time Hadamard transform coding of pictures, *BSTJ* 56, pp. 21-48.
- Mukawa, N., H. Kuroda, and T. Matsuoka (1984), An interframe coding system for video teleconferencing signal transmission at a 1.5 Mb/s rate, *IEEE Trans. Commun.* COM-32, pp. 280-287.
- Murakami, H., H. Hashimoto, and Y. Hatori (1988), Quality of band-compressed TV services, *IEEE Commun. Mag.* 26, pp. 61-69 (October).
- Musmann, H.G., P. Pirsch, and H.J. Grallert (1985), Advances in picture coding, *Proc. IEEE* 73, pp. 523-548.
- Musmann, H.G., and D. Preuss (1977), Comparison of redundancy reducing codes for facsimile transmission of documents, *IEEE Trans. Commun.* COM-25, pp. 1425-1433.
- Myers, W. (1980), Computer graphics: A two-way street, *IEEE COMPUTER* 13, pp. 49-58 (July).
- Nagao, M. (1980), Data compression of Chinese character patterns, *Proc. IEEE* 68, pp. 818-829.
- Nam, M.K., and W.D. O'Neill (1987), Adaptive linear predictive coding of time-varying images using multidimensional recursive least squares ladder filters, *IEEE J. Select. Areas Commun.* SAC-5, pp. 1115-1126.
- Natarajan, T.R., and N. Ahmed (1977), On interframe transform coding, *IEEE Trans. Commun.* COM-25, pp. 1323-1329.
- Neagoe, V.E. (1988), Predictive ordering and linear approximation for image data compression, *IEEE Trans. Commun.* COM-36, pp. 1179-1182.
- Netravali, A.N. (1977), On quantizers for DPCM coding of picture signals, *IEEE Trans. Inform. Theory* IT-23, pp. 360-370.
- Netravali, A.N. (1977), Interpolative picture coding using a subjective criterion, *IEEE Trans. Commun.* COM-25, pp. 503-508.
- Netravali, A.N., and E.G. Bowen (1981), A picture browsing system, *IEEE Trans. Commun.* COM-29, pp. 1968-1976.
- Netravali, A.N., and E.G. Bowen (1982), Improved reconstruction of DPCM-coded pictures, *BSTJ* 61, pp. 969-979.
- Netravali, A.N., and F.W. Mounts (1980), Ordering techniques for facsimile coding: A review, *Proc. IEEE* 68, pp. 796-807.

- Netravali, A.N., and B. Prasada (1977), Adaptive quantization of picture signals using spatial masking, Proc. IEEE 65, pp. 536-548.
- Netravali, A.N., B. Prasada, and F.W. Mounts (1977), Some experiments in adaptive and predictive Hadamard transform coding of pictures, BSTJ 56, pp. 1531-1547.
- Netravali, A.N., and J.D. Robbins (1979), Motion compensated television coding, Part I, BSTJ 58, pp. 631-670.
- Netravali, A.N., and C.B. Rubinstein (1977), Quantization of color signals, Proc. IEEE 65, pp. 1177-1187.
- Netravali, A.N., and C.B. Rubinstein (1979), Luminance adaptive coding of chrominance signals, IEEE Trans. Commun. COM-27, pp. 703-710.
- Netravali, A.N., and R. Saigal (1976), An algorithm for the design of optimum quantizers, BSTJ 55, pp. 1423-1435.
- Netravali, A.N., and J. Salz (1985), Algorithms for estimation of three-dimensional motion, AT&T Tech. J. 64, pp. 335-346.
- Netravali, A.N., and J.A. Stuller (1979), Motion compensated transform coding, BSTJ 58, pp. 1703-1718.
- Neuhoff, D.L., and K.G. Castor (1985), A rate and distortion analysis of chain codes for line drawings, IEEE Trans. Inform. Theory IT-31, pp. 53-68.
- Ninke, W.H. (1985), Design considerations of NAPLPS, the data syntax for videotex and teletext in North America, Proc. IEEE 73, pp. 740-753.
- Ninomiya, Y., and Y. Ohtsuka (1982), A motion-compensated interframe coding scheme for television pictures, IEEE Trans. Commun. COM-30, pp. 201-211.
- O'Brien (1970), Color Television (IEEE Press, New York, NY).
- Ohira, T., M. Hayakawa, and K. Matsumoto (1978), Orthogonal transform coding system for NTSC color television signals, IEEE Trans. Commun. COM-26, pp. 1454-1463.
- O'Neal, J.B. (1966), Delta modulation quantizing noise - Analytic and computer simulation results for gaussian and television input signals, BSTJ 45, pp. 117-141.
- O'Neal, J.B. (1966), Predictive quantizing systems (differential pulse code modulation) for the transmission of television signals, BSTJ 45, pp. 689-722.
- O'Neal, J.B. (1971), Entropy coding in speech and television differential PCM systems, IEEE Trans. Inform. Theory IT-17, pp. 758-761.

- O'Neal, J.B. (1976), Differential pulse-code modulation (DPCM) with entropy coding, IEEE Trans. Inform. Theory IT-21, pp. 169-174.
- O'Neal, J.B., and T.R. Natarajan (1977), Coding isotropic images, IEEE Trans. Inform. Theory IT-23, pp. 697-707.
- Ouellet, J., and E. Dubois (1981), Sampling and reconstruction of NTSC video signals at twice the color subcarrier frequency, IEEE Trans. Commun. COM-29, pp. 1823-1832.
- Pearl, A.N., H.C. Andrews, and W.K. Pratt (1972), Performance measure for transform data coding, IEEE Trans. Commun. COM-20, pp. 411-415.
- Pearlman, W.A., and P. Jakatdar (1984), The effectiveness and efficiency of hybrid transform/DPCM interframe image coding, IEEE Trans. Commun. COM-32, pp. 832-838.
- Pearson, D.E. (1975), Transmission and Display of Pictorial Information (Pentech Press, London, UK).
- Pease, R.F.W. (1972), Conditional vertical subsampling - A technique to assist in the coding of television signals, BSTJ 51, pp. 787-802.
- Pease, R.F.W., and J.O. Limb (1971), Exchange of spatial and temporal resolution in television coding, BSTJ 50, pp. 191-200.
- Pirsch, P. (1981), Design of DPCM quantizers for video signals using subjective tests, IEEE Trans. Commun. COM-29, pp. 990-1000.
- Pirsch, P. (1982), Adaptive intra-interframe DPCM coder, BSTJ 61, pp. 747-764.
- Potter, J.L. (1983), Image processing on the massively parallel processor, IEEE COMPUTER 16, pp. 62-67 (January).
- Prabhu, K.A., and A.N. Netravali (1983), Motion compensated composite color coding, IEEE Trans. Commun. COM-31, pp. 216-223.
- Pratt, W.K. (1967), A bibliography on television bandwidth reduction studies, IEEE Trans. Inform. Theory IT-13, pp. 114-115.
- Pratt, W.K. (1971), Spatial transform coding of color images, IEEE Trans. Commun. COM-19, pp. 980-992.
- Pratt, W.K. (1978), Digital Image Processing (Academic Press, New York, NY).
- Pratt, W.K. (1979), Image Transmission Techniques (Academic Press, New York, NY).
- Pratt, W.K., P.J. Capitant, W.H. Chen, E.R. Hamilton, and R.H. Wallis (1980), Combined symbol matching facsimile data compression system, Proc. IEEE 68, pp. 786-796.

- Pratt, W.K., W.H. Chen, and L.R. Welch (1974), Slant transform image coding, IEEE Trans. Commun. COM-22, pp. 1075-1093.
- Pratt, W.K., O.D. Faugeras, and A. Gagalowicz (1981), Applications of stochastic texture field models to image processing, Proc. IEEE 69, pp. 542-551.
- Pratt, W.K., J. Kane, and H.C. Andrews (1969), Hadamard transform image coding, Proc. IEEE 57, pp. 58-68.
- Pugsley, P.C. (1981), Pre-press picture processing in the graphic arts industry, IEEE Trans. Commun. COM-29, pp. 1891-1897.
- Ramachandran, K. (1980), Coding method for vector representation of engineering drawings, Proc. IEEE 68, pp. 813-817.
- Reininger, R.C., and J.D. Gibson (1983), Distributions of the two-dimensional DCT coefficients for images, IEEE Trans. Commun. COM-31, pp. 835-839.
- Roberts, L.G. (1962), Picture coding using pseudo-random noise, IRE Trans. Inform. Theory IT-8, pp. 145-154.
- Rocca, F., and S. Zanutelli (1972), Bandwidth reduction via movement compensation on a model of the random video process, IEEE Trans. Commun. COM-20, pp. 960-965.
- Roese, J.A., W.K. Pratt, and G.S. Robinson (1977), Interframe cosine transform image coding, IEEE Trans. Commun. COM-25, pp. 1329-1339.
- Rosenfeld, A. (1983), Parallel image processing using cellular arrays, IEEE COMPUTER 16, pp. 14-20 (January).
- Rosenfeld, A., and A.C. Kak (1982), Digital Picture Processing (Academic Press, New York, NY).
- Rubinstein, C.B., and J.O. Limb (1972), Statistical dependence between components of a differentially quantized color signal, IEEE Trans. Commun. COM-20, pp. 890-899.
- Rzeszewski, T.S. (1983), Color Television (IEEE Press, New York, NY).
- Rzeszewski, T.S. (1983), A compatible high-definition television system, BSTJ 62, pp. 2091-2111.
- Sabri, S. (1984), Movement compensated interframe prediction for NTSC color TV signals, IEEE Trans. Commun. COM-32, pp. 954-968.
- Sabri, S., and B. Prasada (1984), Coding of broadcast TV signals for transmission over satellite channels, IEEE Trans. Commun. COM-32, pp. 1323-1330.

- Sabri, S., and B. Prasada (1985), Video conferencing systems, Proc. IEEE 73, pp. 671-688.
- Sakrison, D.J., and V.R. Algazi (1971), Comparison of line-by-line and two-dimensional encoding of random images, IEEE Trans. Inform. Theory IT-17, pp. 386-398.
- Samet, A., C. Munoz, and N. Garcia (1984), Approximation quality improvement techniques in progressive image transmission, IEEE J. Select. Areas Commun. SAC-2, pp. 359-373.
- Santago, P., and S.A. Rajala (1987), Using convex set techniques for combined pixel and frequency domain coding of time-varying images, IEEE J. Select. Areas Commun. SAC-5, pp. 1127-1139.
- Sawada, K., and H. Kotera (1978), A 32 Mbit/s component separation DPCM coding system for NTSC color TV, IEEE Trans. Commun. COM-26, pp. 458-465.
- Schmidt, R.L. (1983), A comparison of line difference predictions for time-frequency multiplexing of television signals, BSTJ 62, pp. 1955-1975.
- Schreiber, W.F. (1956), The measurement of third-order probability distributions for television signals, IRE Trans. Inform. Theory IT-2, pp. 94-105.
- Schreiber, W.F., and R.R. Buckley (1981), A two-channel picture coding system: II - Adaptive companding and color coding, IEEE Trans. Commun. COM-29, pp. 1849-1858.
- Sharma, D.K., and A.N. Netravali (1977), Design of quantizers for DPCM coding of pictures, IEEE Trans. Commun. COM-25, pp. 1267-1274.
- Southall, J.P.C. (1937), An Introduction to Physiological Optics (Oxford University Press, London, UK).
- Srinivasan, R., and K.R. Rao (1985), Predictive coding based on efficient motion estimation, IEEE Trans. Commun. COM-33, pp. 888-896.
- Srinivasan, R., and K.R. Rao (1987), Motion-compensated coder for video conferencing, IEEE Trans. Commun. COM-35, pp. 297-304.
- Stenger, L. (1977), Quantization of TV chrominance signals considering the visibility of small color differences, IEEE Trans. Commun. COM-25, pp. 1393-1406.
- Stoffel, J.C., and J.F. Moreland (1981), A survey of techniques for pictorial reproduction, IEEE Trans. Commun. COM-29, pp. 1898-1925.
- Stuller, J.A., and B. Kurz (1977), Interframe sequential picture coding, IEEE Trans. Commun. COM-25, pp. 485-495.
- Stuller, J.A., and A.N. Netravali (1979), Transform domain motion estimation, BSTJ 58, pp. 1673-1702.

- Stuller, J.A., A.N. Netravali, and J.D. Robbins (1980), Interframe television coding using gain and displacement compensation, *BSTJ* 59, pp. 1227-1240.
- Takikawa, K. (1981), Simplified 6.3 Mbit/s code for video conferencing, *IEEE Trans. Commun.* COM-29, pp. 1877-1882.
- Tanimoto, M., N. Chiba, H. Yasui, and M. Murakami (1988), TAT (Time-Axis Transform) bandwidth compression system of picture signals, *IEEE Trans. Commun.* COM-36, pp. 347-354.
- Tasto, M, and P.A. Wintz (1971), Image coding by adaptive block quantization, *IEEE Trans. Commun. Technol.* COM-19, pp. 957-971.
- Teramura, H., K. Ono, S. Ando, Y. Yamazaki, S. Yamamoto, and K. Matsuo (1981), Experimental facsimile communication system on packet switched data networks, *IEEE Trans. Commun.* COM-29, pp. 1942-1951.
- Thompson, J.E. (1971), A 36-Mbit/s television coder employing pseudorandom quantization, *IEEE Trans. Commun. Technol.* COM-19, pp. 872-879.
- Thompson, J.E. (1981), European collaboration on picture coding research for 2 Mbit/s transmission, *IEEE Trans. Commun.* COM-29, pp. 2003-2004.
- Thompson, J.E., and J.J. Sparkes (1967), A pseudo-random quantizer for television signals, *Proc. IEEE* 55, pp. 353-355.
- Ting, D., and B. Prasada (1980), Digital processing techniques for encoding of graphics, *Proc. IEEE* 68, pp. 757-769.
- Townsend, G.B. (1970), *Pal Colour Television* (Cambridge University Press, Cambridge, UK).
- Troxel, D.E. (1981), Application of pseudorandom noise to DPCM, *IEEE Trans. Commun.* COM-29, pp. 1763-1767.
- Troxel, D.E., W.F. Schreiber, P. Curlander, A. Gilkes, R. Grass, and G. Hoover (1979), Image enhancement-encoding systems using pseudorandom noise processing, *Proc. IEEE* 67, pp. 972-973.
- Troxel, D.E., W.F. Schreiber, R.P. Bishop, R.R. Buckley, G.J. Bunza, T. Eguchi, S.M. Goldwasser, J.E. Jakobson, J.S. Kung, D. Ng, S. Takemoto, and E. Yokoyama (1981), A two-channel picture coding system: I - Real-time implementation, *IEEE Trans. Commun.* COM-29, pp. 1841-1848.
- Usubuchi, T., T. Omachi, and K. Iinuma (1980), Adaptive predictive coding for newspaper facsimile, *Proc. IEEE* 68, pp. 807-813.
- Van de Polder, L.T., D.W. Parker, and J. Roos (1985), Evolution of television receivers from analog to digital, *Proc. IEEE* 73, pp. 599-612.
- Walker, D.R., and K.R. Rao (1984), Improved pel recursive motion compensation, *IEEE Trans. Commun.* COM-32, pp. 1128-1134.

- Wang, L., and M. Goldberg (1988), Progressive image transmission by transform coefficient residual error quantization, IEEE Trans. Commun. COM-36, pp. 75-87.
- Weiss, L., I.M. Paz, and D.L. Schilling (1975), Video encoding using an adaptive digital delta modulator with overshoot suppression, IEEE Trans. Commun. COM-23, pp. 905-920.
- Wilkins, L.C., and P.A. Wintz (1971), Bibliography on data compression, picture properties, and picture coding, IEEE Trans. Inform. Theory IT-17, pp. 180-199.
- Wintringham, W.T. (1951), Color television and colorimetry, Proc. IRE 39, pp. 1135-1173.
- Wintz, P.A., (1972), Transform picture coding, Proc. IEEE 60, pp. 809-820.
- Woods, J.W. (1975), Stability of DPCM coders for television, IEEE Trans. Commun. COM-23, pp. 845-846.
- Wyle, H., T. Erb, and R. Barow (1961), Reduced time facsimile transmission by digital coding, IRE Trans. Commun. Systems CS-9, pp. 215-222.
- Yamada, T. (1979), Edge-difference coding - A new, efficient redundancy reduction technique for facsimile signals, IEEE Trans. Commun. COM-27, pp. 1210-1217.
- Yamamoto, H., Y. Hatori, and H. Murakami (1981), 30 Mbit/s code for the NTSC color TV signal using an interfield-intrafield adaptive prediction, IEEE Trans. Commun. COM-29, pp. 1859-1867.
- Yan, J.K., and D.J. Sakrison (1977), Encoding of images based on two-component source model, IEEE Trans. Commun. COM-25, pp. 1315-1323.
- Yasuda, Y. (1980), Overview of digital facsimile coding techniques in Japan, Proc. IEEE 68, pp. 830-845.
- Yasuda, Y., M. Dubois, and T.S. Huang (1980), Data compression for check processing machines, Proc. IEEE 68, pp. 874-885.
- Yasuda, H., H. Kuroda, H. Kawanishi, F. Kanaya, and H. Hashimoto (1977), Transmitting 4-MHz TV signals by combinational difference coding, IEEE Trans. Commun. COM-25, pp. 508-516.
- Zobrist, A.L., and G. Nagy (1981), Pictorial information processing of Landsat data for geographic analysis, IEEE COMPUTER 14, pp. 34-41 (November).
- Zschunke, W. (1977), DPCM picture coding with adaptive prediction, IEEE Trans. Commun. COM-25, pp. 1295-1302.

BIBLIOGRAPHIC DATA SHEET

	1. PUBLICATION NO.	2. Gov't Accession No.	3. Recipient's Accession No.
4. TITLE AND SUBTITLE Image Data Compression Overview: Issues and Partial Solutions		5. Publication Date October, 1989	
		6. Performing Organization Code ITS.N1	
7. AUTHOR(S) M. Nesenbergs		9. Project/Task/Work Unit No.	
8. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Department of Commerce NTIA/ITS.N1 325 Broadway Boulder, CO 80303		10. Contract/Grant No.	
11. Sponsoring Organization Name and Address Naval Sea Systems Command Department of the Navy Washington, DC 20362-5101		12. Type of Report and Period Covered	
		13.	
14. SUPPLEMENTARY NOTES			
15. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.) This report presents an overview of moving image (video) data compression. Because this report deals with a subtask of a major program, a limited number of the most pertinent topics and issues have been addressed. These topics include a listing of human perception effects and parameters for visual information; theoretical aspects of image data compression; a selective review of compression techniques; a closer look at motion detection, prediction and compensation; and a start to survey existing systems and their capabilities. Finally a multipage bibliography, while not exhaustive in any sense, shows the depth and breadth of the rapidly growing technology of image data compression.			
16. Key Words (Alphabetical order, separated by semicolons) Key Words: Data compression; digital; images; motion compensation; prediction; visual effects			
17. AVAILABILITY STATEMENT <input checked="" type="checkbox"/> UNLIMITED. <input type="checkbox"/> FOR OFFICIAL DISTRIBUTION.		18. Security Class. (This report) Unclassified	20. Number of pages 70
		19. Security Class. (This page) Unclassified	21. Price:

