SIGNAL COMPRESSION AND RECONSTRUCTION USING MULTIPLE BASES REPRESENTATION

by

Felix G. Safar

A. A. (Louis) Beex, Chairman

Electrical Engineering

(ABSTRACT)

The problem of efficient signal communication at low data rates involves, in general, the encoding of the source for maximum data compression at the transmitter end, and the reconstruction using the received information and all the available a priori or side information at the receiver end. In this thesis, we propose an adaptive signal representation scheme, based on the use of multiple orthogonal basis sets, that exhibits very good potential in both the source encoding and the signal reconstruction problems. This representation leads naturally to the splitting of the signal into additive components, each of which has a simpler description than the original process. In addition, it exhibits a structure similar to that of codebook based coding. As a result, a very compact signal representation can be achieved.

A splitting procedure called recursive residual projection is proposed, and its performance evaluated for the separation of imagelike signals into basis-defined "edge" and "texture" components. The coding of these components leads to lower rates than those for transform coding methods. In reconstruction, the representation can be considered as a well-behaved constraint. This allowed for the development of the corresponding unique projection operator, with application to general iterative reconstruction methods. In particular, we also proposed a noise tolerant version of the operator, a so-called soft projection operator, capable of achieving convergence under noisy measurement conditions. Computer simulations in the representation, coding, and reconstruction applications corroborate the usefulness of this proposed representation.

Acknowledgements

My foremost thanks go to my advisor , Dr. A. A. (Louis) Beex, who has given me the opportunity to pursue the degree of Master of Science in Electrical Engineering at Virginia Polytechnic Institute and State University. I feel indebted to him for his continual and knowledgeable guidance in the course of my graduate education. His assistance, attitude, patience, and human qualities cannot be overstated.

My gratitude extends also to Dr. F. W. Stephenson for his technical and moral support through all the stages of this degree, and for his friendship and incredible hospitality. To the members of my advisory commitee, Dr. A. A. Beex (chair), Dr. R. L. Moose, Dr. F. W. Stephenson, and Dr. Kai-Bor Yu, I extend my sincere thanks for their contribution and assistance in the various aspects of the entire course of my study.

I would like to thank the Rotary Foundation for offering me the opportunity to come over here. I consider myself fortunate to have been part of such a rewarding program. In addition, I would like to recognize those many fellow students and staff members, too numerous to mention, for their assistance and friendship, and for making this experience all the more enjoyable.

A Julieta y Santiago, mis hijos, les debo mi presencia, el tiempo compartido, el amor sensible, los juegos.... Espero poder compensarles por las carencias en este período, y que este logro de papá redunde en mejores oportunidades futuras para ustedes. A Claudia, mi esposa, mi total

agradecimiento por haberme estimulado a aceptar este reto, aún cuando ello significaba posponer su propia carrera, y asumir temporalmente toda la carga familiar. Nadie mejor que yo conoce el valor de tu sacrificio. Listo para ponerte el hombro, confío en tiempos mejores por venir.

Deseo finalmente agradecer a mi familia, especialmente a mis padres, quienes me ayudaron en ocasiones y formas harto numerosas como para mencionar. A mi padre Felix A., quien con su honestidad, mentalidad progresista y pujanza a toda prueba, es fuente contínua de inspiración y aliento en mi proceso de superación profesional y personal. A mi madre Nely L., de quien aprendí la importancia de trabajar por una vida armónica, que excediendo el marco de las satisfacciones profesionales, se nutre de los momentos con la familia, los amigos y la música.

Table of Contents

List of Illustrations

List of Tables

1.0 Introduction

1.1 Overview

One of the most basic problems in communication theory is the transmission of information over a channel whose capacity is lower than the average information rate generated by the source. In such a case, unavoidable distortion at the receiver end occurs. The main role of rate-distortion theory is to determine, given the description of the source and the channel capacity, the minimum achievable distortion over all possible source encoding schemes. These bounds are generally given for the case of infinitely long signals. Any practical system will have a performance inferior to the bound imposed by this branch of information theory.

If the transmission rate is specified, there are two basic approaches to reduce the distortion at the receiver end: 1) encode the source efficiently, to make the distortion as close as possible to the theoretical minimum, and 2) incorporate side information at the receiver end. This information, known as a priori knowledge of the signal, can be combined with the actually received information to reconstruct the signal and to reduce the resulting average distortion.

The source encoding or data compression problem, investigated for more than two decades, has been motivated mainly by speech and image coding applications. When performance close to

the theoretical bounds is desired, there are mainly two high-complexity methods available for encoding a finite-length signal block: a linear transformation followed by independent (scalar) quantization of the transformed coefficients, called Transform Coding (TC), and the encoding using a codebook, called Vector Quantization (VQ).

Transform coding is a classical ad-hoc procedure that involves the decorrelation of the source followed by scalar quantization. It is motivated by the fact that, for a Gaussian source with memory, the optimal linear transformation followed by the optimal scalar quantization results in performance very close (within 1/4 bit/sample) to the theoretical rate-distortion bound. Speech and video signals can be modelled as highly correlated processes with moderate success. As a result, the particular statistical structure of the source cannot be exploited fully. A typical example is the poor edge coding performance in transform coders. Nevertheless, TC is a robust procedure that has proven effective for moderate compression ratios. In addition, its implementation is dramatically simplified by using fast orthogonal transformations.

The more recently developed codebook encoders, or vector quantizers, are based on the idea of assigning to every signal block a codeword drawn from a codebook, such that the representation of that codeword will constitute the best possible match for that signal. By having a large number of appropriately chosen codewords, small distortions can be achieved. The main advantage of vector quantizers is their structure, which resembles the optimum system that achieves the rate-distortion bound, regardless of the probability density function of the source. The main disadvantage comes from the very high computational and storage costs.

The problem of incorporating a priori information at the receiver, to help in the reconstruction process, usually involves the use of some kind of iterative technique. In this setting the available information is enforced sequentially, and the process hopefully converges to a signal that satisfies all the participating pieces of information or constraints. The particular structure of each constraint is, in general, responsible for the convergence performance of these algorithms. Well-behaved constraints or pieces of information constitute the necessary (and sometimes sufficient) condition for guaranteed convergence.

In this research work, we propose a signal representation scheme based on the use of multiple orthogonal bases, and show that it has good potential in both coding and reconstruction applications. This representation can be viewed as the result of using a codebook search (basis vector selection) and a gain selection. This results in the adaptive splitting of the signal in terms of each orthogonal set.

When applied to source coding of composite signals, the ability of the representation to split the signal into simpler components allows for a better overall performance than obtained with transform coding. In the reconstruction phase, the multiple bases representation is shown to be a well-behaved (convex) constraint, corresponding to a unique projection operator and resulting in convergent iterations.

1.2 Organization of the material

The remainder of this thesis is organized as follows. In Chapter 2, we review the mathematics associated with data compression of information sources, the information theoretic bounds, and different quantization schemes. We summarize the properties of the fast transforms most commonly used in signal processing in Chapter 3. In addition to the classical comparison with the optimal Karhunen-Loeve transform, the basis vectors are analyzed in terms of their feature extraction power, which will play a key role in the multiple bases setting.

A form of signal representation based upon the combination of several orthogonal basis sets is proposed in Chapter 4. The compactness of the representation and its applicability to source coding are analyzed, particularly for the case of imagelike signals. Some supporting computer simulations are included at the end of the chapter. In Chapter 5, the application of the multiple bases representation to the problem of iterative signal reconstruction in a Hilbert space is analyzed. After reviewing different iteration schemes, the application of the multiple bases representation as a constraint is introduced, and the corresponding projection operator derived. A modified operator,

consistent with the practical case of noisy measurements, is introduced next. The usefulness of these operators is corroborated by some computer simulations. Finally, in Chapter 6, the observed advantages and limitations of the proposed representation are discussed, and some extensions and suggestions for future research work given.

2.0 Mathematical basis for data compression

2.1 Rate-distortion theory

The foundations of Information Theory were cleverly laid by C. E. Shannon in his celebrated journal article, "A Mathematical Theory of Communication," published in 1948. He pointed out that any communication problem can be separated into two component problems, namely 1) "What information should be transmitted," and 2) "How should it be transmitted." The bulk of the work by Shannon was devoted to the second problem, that of encoding the set of possible messages in such a way that they can be transmitted reliably over a noisy communication channel. The importance of his coding theorem for a noisy channel justified this emphasis.

In 1959, Shannon broke the silence regarding the first problem with the publication of the article, "Coding Theorems for a Discrete Source with a Fidelity Criterion." In that paper, Shannon defined the rate-distortion function of an information source with respect to a fidelity criterion and established the fundamental theorems relative to it, that represent a milestone in the subsequent development of Rate-Distortion Theory.

As a branch of Information Theory, Rate-Distortion Theory is devoted to those cases in which the entropy or average information content of the source exceeds the capacity of the transmitting

channel. In such situations some distortion inevitably results. In order to keep this distortion to a minimum, it is necessary to first order the data produced by the source in accordance with its importance at the eventual destination, and then either condense or delete the less important information prior to actual transmission. This represents the basic idea of source coding for data compression.

The rate-distortion function R(D) of a source determines the minimum channel capacity required to transmit the source output as a function of the desired minimum average distortion, where the distortion function (or fidelity criterion D) is a user-specified measure of agreement between the source and the received output. Its inverse D(R), called distortion-rate function, gives the minimum attainable average distortion as a function of the average transmission rate.

2.1.1 A communication system model

Consider the block diagram model of a communication system depicted in Fig. 1. This model has been shown to possess sufficient generality to encompass most practical communication systems and, at the same time, it proves amenable to analysis. We proceed to discuss briefly the various blocks that comprise the diagram.

The source generates the information that the system has to communicate to the user located at the receiver end. The channel is a physical medium that links the source to the user. In the context of this thesis, no emphasis will be placed on channels, being described then simply by their capacity C. The encoder is a device that transforms the source input into a form suitable to be transmitted over the channel. Likewise, the decoder converts the channel output into a form that can be interpreted by the user. Designing a communication system will generally involve designing an encoder and decoder for a given source, channel, and user.

It is convenient to divide the encoder into the cascade of the source encoder and the channel encoder. The former maps the output X produced by the source over some time interval (block encoding) into one of a (finite) set of preselected messages U. That is, the space of possible source

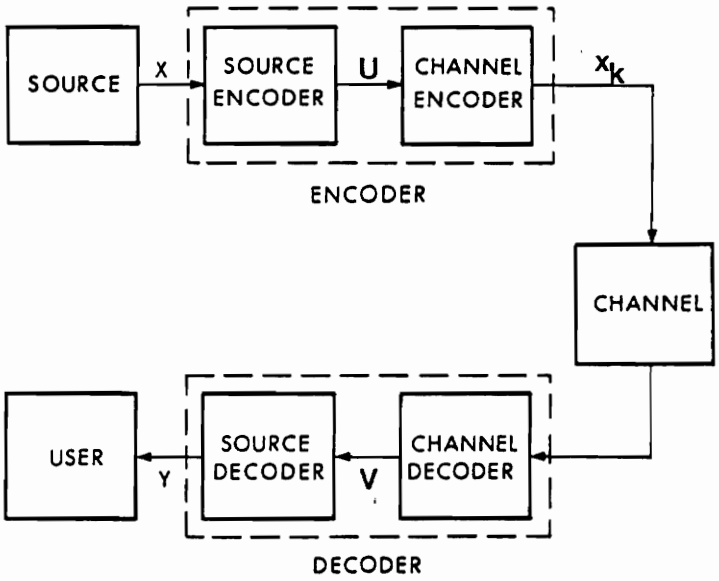

Figure 1. Communication system model. (from Berger [9].)

outputs is divided into a set of equivalence classes, and a partition-equivalence class correspondence is assigned by the encoder. It is worth noting that, when the source alphabet has more symbols than there are equivalence classes in the encoder, a many-to-one mapping occurs. Since such a mapping has no inverse, the encoding process involves information loss. This is always the case when coding analog sources, or when data compression of a finite entropy source is desired. The main job of the source encoder is to remove source redundancy.

The channel encoder receives the k-th equivalence class assigned by the source encoder, and transmits the corresponding waveform $x_k(t)$ for that class. At the receiver end, the channel decoder examines the received waveform over an appropriate time interval and makes a decision regarding what message was sent (message V). For a given channel, the criteria for designing the channel encoder (i.e. the transmitter waveforms x(t)) and the channel decoder (decision device) is to maximize immunity to channel noise. This is implemented by inserting controlled redundancy. We will not treat this problem any further.

An estimate Y of the original signal X is given to the user by combining, in the general case, two sources of information. One is provided by the source decoder, which maps each decoded message back into an appropriately selected point of the source space. The other one is the optionally available side information, usually in the form of a priori knowledge about X, that can be incorporated into the recovery process to improve the quality of the estimate Y.

2.2 *The rate distortion function R(D)*

From the introductory discussion, it should be clear that R(D) represents the effective rate at which the source produces information, subject to the constraint that the user can tolerate an average distortion of D. If distortionless reproduction is required, then the rate at which a source produces information is called the entropy of the source. It follows that the rate-distortion function is a generalization of the concept of entropy, and they coincide (when defined) for $D = 0$.

The main advantage of knowing R(D) for a given source and fidelity criterion, is having a theoretical bound against which practical communication systems can be compared. The system designer can perform a feasibility study for a given set of specifications in a communication system, before going into what could end up being an impossible task. For existing practical methods (or systems), one can assess whether it is worth trying to improve them by consulting R(D).

Nevertheless, the definition of R(D) implies full statistical knowledge of the source, and the existence of a well-defined fidelity measure. When it comes to coding signals like real images or speech, this requirement can only be partially satisfied, since neither detailed statistics are available nor a completely reliable fidelity measure has been developed yet for such real world signals.

2.2.1 Definitions and properties

Consider again the communication-system block diagram of Fig. 1. Suppose that the source is generating a sequence of discrete-time discrete-amplitude values, which are encoded in successive blocks of length N. Each block then is described by one of a enumberable set of messages $\{X_i\}$ with probability function $P(X_i)$. Now, any given system can be mathematically described by the conditional probability $F(Y_j \mid X_i)$ of message Y_j being output by the decoder given the source output was X_i.

The average information content of the N-block source is given by the Nth-order entropy $H_N(X)$. Consequently, at least $H_N(X)$ bits are required for optimum error-free coding. Similarly [9], the N-block mutual information $I_N(X,Y)$ is given by

$$I_N(X,Y) = H_N(X) - H_N(X \mid Y) \tag{2.1}$$

where $H_N(X \mid Y)$ is the entropy of the source given the observed decoded output Y. As entropy is a measure of uncertainty, the mutual information between X and Y equals the uncertainty in the source output X, minus the uncertainty in the source output X given the decoder output Y. The source entropy is an upper bound to the mutual information (in the extreme case when X is

uniquely determined by Y). When Y contains no information about X, the mutual information is zero.

Let d(X,Y) be a measure of distortion between X and Y. Then the average distortion per source symbol is

$$D(F) = \frac{1}{N} E[d(X,Y)] \tag{2.2}$$

The N-block rate-distortion function $R_N(D_*)$ is the minimum of the average mutual information per symbol, for D(F) less than some value D_*. For the given source, the minimization is then over the conditional probability F(.|.) , defining

$$R_N(D_*) = \inf_{F:D(F)\leq D_*} \frac{1}{N} I_N(X,Y). \tag{2.3}$$

The limiting value of the N-block rate-distortion function is called the rate-distortion function,

$$R(D_*) = \lim_{N\to\infty} R_N(D_*) \leq R_N(D_*) \tag{2.4}$$

That is, the rate-distortion gives minimum mutual information over all possible transmission systems described by F(.|.) for an infinitely long transmission block. The N-block rate-distortion can then be considered as the rate-distortion of a source that produces statistically independent blocks of N symbols each. Depending on the logarithm base used in the definition of H(.), the rate-distortion function units are bits/sample (base 2), nats/sample (base e), or Hartleys/sample (base 10).

Definining the channel capacity C from maximizing the average mutual information between the channel input U and the channel output V,

$$C = \lim_{N\to\infty} \sup \frac{1}{N} I_N(U,V). \tag{2.5}$$

the following information transmission theorem, a combination of the Shannon coding theorem and source coding theorem [36], results in

$$C \geq R(D) \quad \text{for reliable transmission and fidelity D.} \tag{2.6}$$

This theorem states that waveform reconstruction with fidelity D is also possible after transmission over a noisy channel, provided that its capacity is greater than R(D). This justifies the separation of source coding and channel coding operations. Thus the only property of the channel that concerns the source coder is the single parameter C.

Starting from proving the existence of R(D), Berger [9] describes and gives proofs for its most relevant properties. Among them, R(D) satisfies

$$0 \leq R(D) \leq I(X \mid Y) \leq H(X) \leq \log M \tag{2.7}$$

where M is the number of letters in the source alphabet. In the case of continuous-amplitude sources where the absolute entropy is infinite, H(X) can then be substituted for in most relations by the differential entropy h(X). Fig. 2 shows a typical rate-distortion function, from which several properties can be observed. Perhaps the most important one is that R(D) is a convex function. This guarantees continuity and strictly decreasing behavior for $D < D_{max}$. The value D_{max} is the average distortion associated with the best guess one can make when only the statistics of the source are known.

Another important consideration is the degree of memory present in the source process. For two stationary sources having the same marginal probability function $P(X_i)$, it can be shown that [36]

$$H(X)\mid_{\text{with memory}} < H(X)\mid_{\text{without memory}} \leq \log M \tag{2.8}$$

and

$$R(D)\mid_{\text{with memory}} < R(D)\mid_{\text{without memory}}. \tag{2.9}$$

The source redundancy, which is the positive difference between log M and the source entropy H(X), is due to two reasons in general: a nonuniform distribution of probabilities $P(X_i)$, and the

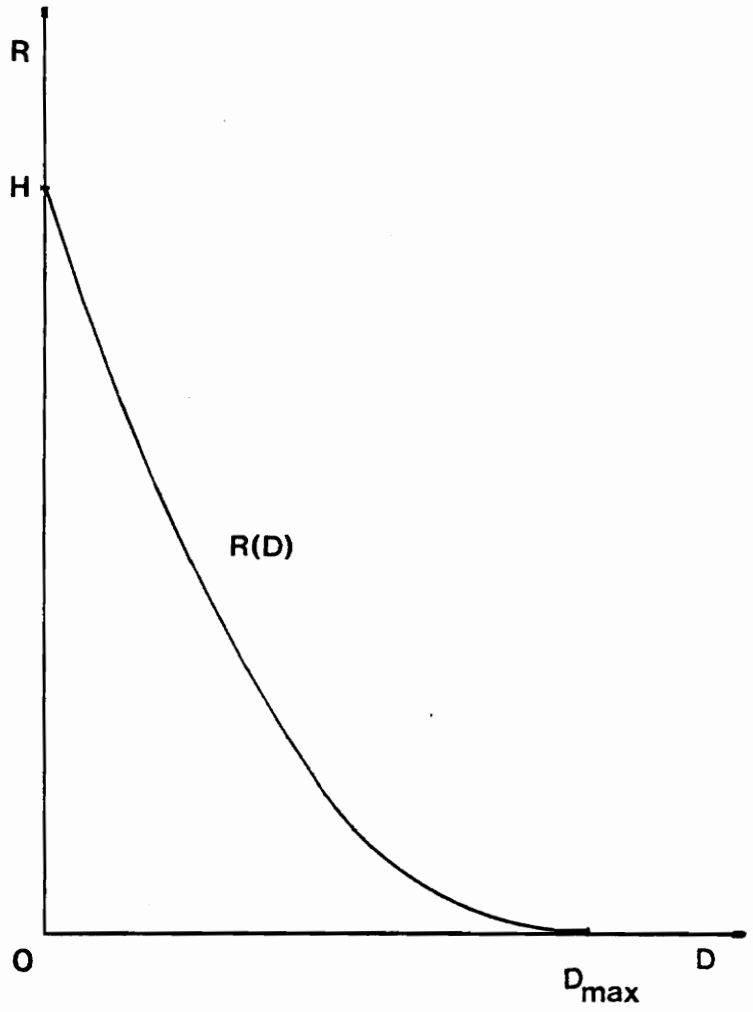

Figure 2. Typical rate-distortion function

presence of memory. The main point in data compression systems is to efficiently take advantage of these two factors.

In the computation of R(D) a variational problem has to be solved. Except for a limited set of distortion measures and source statistics in which analytical solutions have been derived [9], finding R(D) represents a formidable if not impossible task. Blahut [9] proposed an algorithm for numerically computing the rate-distortion function. General solutions exist for Gaussian statistics and some difference distortion measures. Since for a given variance the Gaussian distribution is the most unpredictable one, $R_G(D)$ gives an upper bound for R(D) among all possible distributions when the same fidelity criterion is used [9]. In addition, the use of the MSE (or weighted MSE) criterion, while dramatically simplifying the computations involved, gives a reasonable nonspecific distortion measure for waveform coding problems.

2.3 *Bounds on R(D) for a stationary source and MSE error criterion*

In general, it is not possible to obtain closed form solutions for R(D) for any stationary process under the MSE distortion criterion. Nevertheless, we can compute tight upper and lower bounds on R(D), valid for arbitrary source statistics. By recognizing that, for given second order statistics, the Gaussian probability function has maximum entropy [36], $R_G(D)$ becomes an upper bound for R(D).

For discrete-time continuous-amplitude memoryless sources $R_G(D)$ (in bits/sample) is given by [36]

$$R_G(D) = \max \{0, \frac{1}{2} \log_2 \frac{\sigma_x^2}{D}\} \tag{2.10}$$

or

$$D_G(R) = 2^{-2R}\sigma_x^2 \tag{2.11}$$

One interesting conclusion is that the MSE is reduced by a factor of 4 for each additional bit spent in transmission. This behavior will also roughly correspond to any practical data compression system. The statistics of the optimal mapping that realizes R(D) are given explicitly in Fig. 3 [9]. After scaling by the parameter $\beta = 1 - D/\sigma_x^2$, the input is corrupted by additive zero-mean X-independent Gaussian noise of variance βD. This model gives conditions for optimality of the reconstruction error, which should be reproduced by any close-to-optimal real quantization system.

Additionally, all non-Gaussian memoryless sources are bounded from below by the Shannon lower bound [9]

$$R_S(D) = \max\{0, \frac{1}{2}\log_2\frac{Q}{D}\} \tag{2.12}$$

or

$$D_S(R) = 2^{-2R}Q \tag{2.13}$$

where Q, called the entropy power of the source,

$$Q = (2\pi e)^{-1}2^{2h(X)} \tag{2.14}$$

equals the variance of a Gaussian process having the same differential entropy as the process X. Consequently, it has a maximum value, equal to σ_x^2 for a memoryless Gaussian source, with independent identically distributed (i.i.d.) outputs. In this particular Gaussian case, both the upper and lower bounds collapse into R(D). It can be also shown that, for small distortion D, $R_S(D)$ converges to the true R(D) for a rather broad class of distributions [36].

For the case of processes with memory, both R(D) and the bounds are reduced accordingly. It is convenient to define the Spectral Flatness Measure of a zero mean stationary process X having the spectrum $S_{xx}(e^{j\omega})$ by [36]

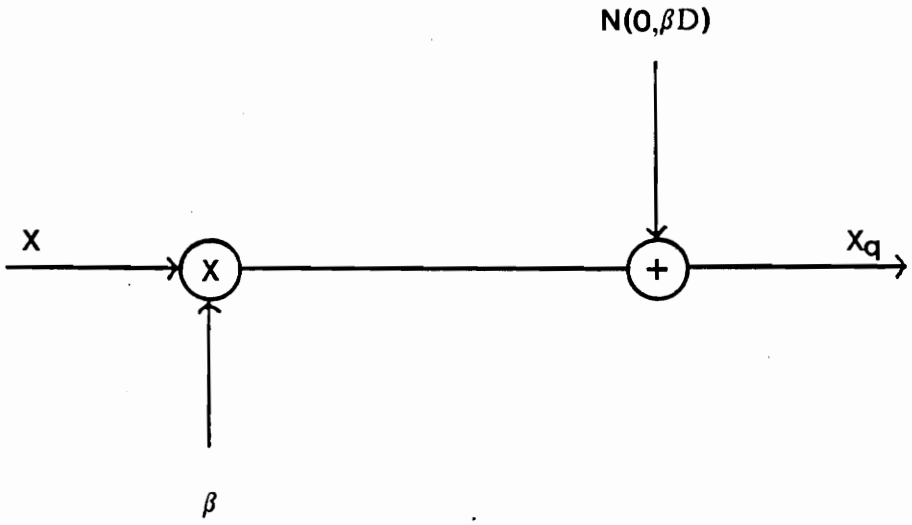

Figure 3. Optimal mapping for a memoryless source

$$SFM(X) = \frac{\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log_e S_{xx}(e^{j\omega})d\omega\right)}{\sigma_x^2} \tag{2.15}$$

The SFM(X) measures redundancy as manifested in the structure and shape of the power spectral density of X.

The N-block rate-distortion function for a stationary Gaussian process with memory is given by the following parametric expression [36]

$$D_{N,G}(\theta) = \frac{1}{N}\sum_{k=0}^{N-1} \min\{\theta, \lambda_k\} \tag{2.16}$$

$$R_{N,G}(\theta) = \frac{1}{N}\sum_{k=0}^{N-1} \max\{0, \frac{1}{2}\log_2 \frac{\lambda_k}{\theta}\} \tag{2.17}$$

where λ_k is the k-th eigenvalue in the eigenvalue-eigenvector expansion for the N-th order autocorrelation matrix R_{xx}^N of the process X. These expressions have the following important connotation: given the expansion of a colored Gaussian process in terms of a linear combination of its autocorrelation matrix eigenvectors, the coding scheme that achieves $R_{N,G}(D)$ should discard all coefficients whose eigenvalues are less than θ, and code the remaining coefficients with equal mean-square error.

The rate-distortion function R(D) can be obtained by going into a limiting process when $N \to \infty$. Upon applying the Toeplitz distribution theorem [9], the resulting expressions follow

$$D_G(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \min\{\theta, S_{xx}(e^{j\omega})\}d\omega \tag{2.18}$$

$$R_G(\theta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \max\left\{0, \frac{1}{2}\log_2 \frac{S_{xx}(e^{j\omega})}{\theta}\right\}d\omega \tag{2.19}$$

This result has the interpretation of the "water-filling" procedure [36] shown in Fig. 4. The frequency axis can be divided into two sets A and B:

$$\omega \in A \text{ if } S_{xx}(e^{j\omega}) \geq \theta, \text{ and } \omega \in B \text{ if } S_{xx}(e^{j\omega}) < \theta \tag{2.20}$$

As in the finite-N case, optimum encoding will discard all components belonging to the set B (stopbands), and code the remaining components (passband, set A) with constant error power spectral density equal to θ (shaded areas in Fig. 4). This error spectrum can be realized by combinations of source X filtering plus a non-white Gaussian source-independent additive noise. Note that, by making θ a function of ω, a weighted MSE criterion can be accommodated easily.

In the particular case of low distortions ($D = \theta$) and Gaussian statistics, the rate-distortion function is given by [36]

$$D_G(R) = SFM(X)2^{-2R}\sigma_x^2 \tag{2.21}$$

This indicates that, for high transmission rates, the distortion when coding processes with memory can be made SFM(X) times smaller than in the case of coding memoryless sources.

For non-Gaussian stationary sources with memory, a lower bound to R(D) is again given by the generalized Shannon lower bound [9]. It has exactly the same form as in the memoryless case. The main problem associated with it is the computation of the entropy power of the source X for the non-Gaussian case. An even tighter lower bound can be obtained by first computing a lower bound to the N-block rate-distortion function using the parametric equations [36]

$$D(\theta) = \frac{1}{N}\sum_{k=0}^{N-1} \min\{\theta, \lambda_k\} \tag{2.22}$$

$$R(\theta) \geq \frac{1}{N}\left\{h(X_0,..,X_{N-1}) - \sum_{k=0}^{N-1} \frac{1}{2}\log[2\pi e \min(\theta, \lambda_k)]\right\} \tag{2.23}$$

In a manner similar to the case of the Gaussian upper bound, this lower bound can be extended to infinitely long blocks, resulting in corresponding equations in terms of $S_{xx}(e^{j\omega})$. It is worth noting

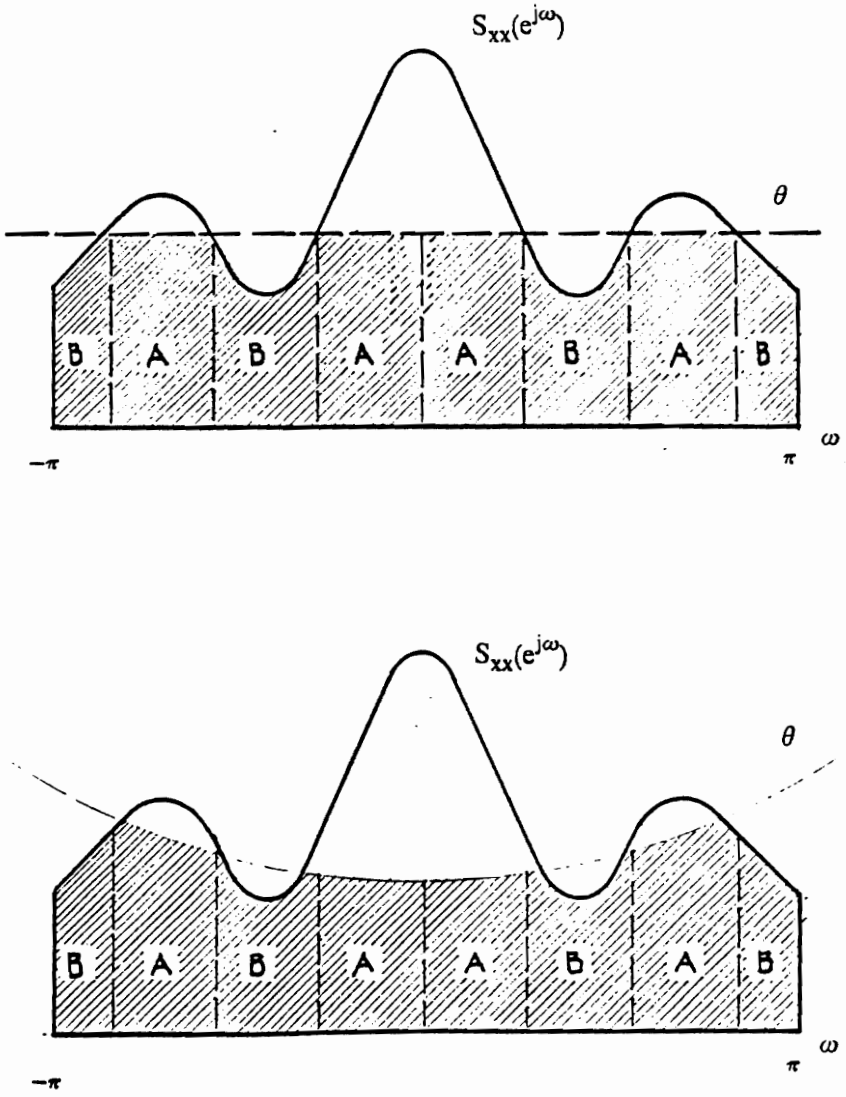

Figure 4. "Water-filling" procedure

that, in the case of low distortions ($D = \theta$), both lower bounds coincide with the true rate-distortion function R(D). This holds independently of the probability density function of X.

2.4 *Application of rate-distortion theory to image encoding*

The Shannon rate-distortion function provides a potentially useful lower bound against which to compare rate-versus-distortion performance of practical image encoding-transmission systems. In order to successfully apply the results of information theory to the efficient storage or transmission of images, it is necessary to find an appropriate statistical model for the source along with a meaningful measure of distortion.

Traditionally, the measure of distortion has been the squared error criterion. While simple to treat mathematically, it gives an overall good performance in nonspecialized applications. In image representation, however, squared error is far from exhibiting a reasonable correspondence with the subjective evaluation of the human observer or interpreter. Finding distortion measures that take into consideration the visual system would provide a better basis for comparing the performance of different systems.

Modelling the source as a colored Gaussian process has the main advantages of mathematical tractability and simple determination of the statistical parameters from sampled data. Actual images however, are far from being represented well by a simple Gaussian random field. As a result, the minimum rate predicted by the Gaussian lower bound is significantly higher than the true rate of the source. In addition, systems designed to optimally code Gaussian sources are inherently suboptimal for real images. Finding a better source model, and the corresponding rate-distortion function, would give a more realistic bound along with insight for designing more efficient image coders.

2.4.1 A distortion measure for monochrome still images

The problem of finding a distortion measure that is in good accord with subjective evaluation is in its infancy. In spite of that, some facts are well known: that the human observer is more sensitive to some spatial frequencies than others; that one is more sensitive to intensity errors in grey areas than in white ones. It is well known [51] that the visual system is not linear. However, the results of studies [51] seem to indicate that, after an initial nonlinear transformation, the remainder of the visual system may be considered linear over a moderate range of intensities.

A simplified model that takes these two effects into consideration was proposed by Mannos and Sakrison [51]. Its block diagram is depicted in Fig. 5. The sampled image X is first distorted by a log-type nonlinear function f(X), and the resulting field is then linearly distorted by a shift-invariant isotropic point spread function $A(f_r)$, where f_r is the radial distance from the 2D spatial frequency origin.

After carrying out a series of subjective testing experiments, Mannos proposed the following functions f(.) and A(.)

$$f(u) = u^{0.33} \tag{2.24}$$

$$A(f_r) \simeq 2.6[0.0192 + 0.114f_r] \exp[-(0.114f_r)^{1.1}]. \tag{2.25}$$

Fig. 6 shows the shape of the corresponding curves. The bandpass form of A(.) with a central peak at 8 cycles/degree and a rapid decrease on either side of this peak, is typical of the contrast sensitivity functions obtained from psychophysical experiments. After applying the nonlinear-linear transformation to both fields we want to compare, a standard sum of squared errors can be used as a distance measure.

This processing only accounts for an oversimplified model of the human visual system. Nevertheless, its main result is that, by simply preprocessing with the best model before transmission and postprocessing with its inverse after reception, the subjective quality of an image

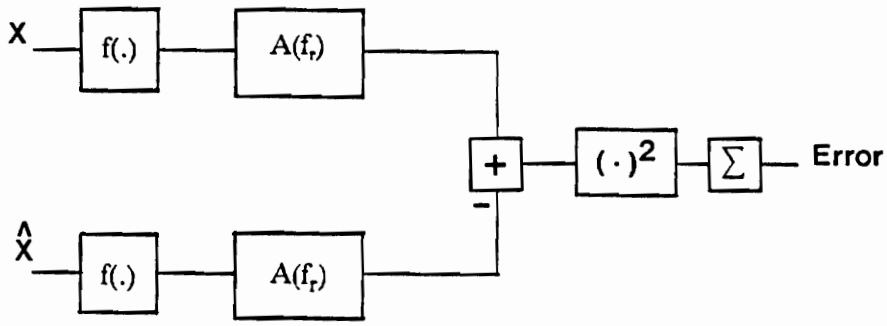

Figure 5. Distortion model for images

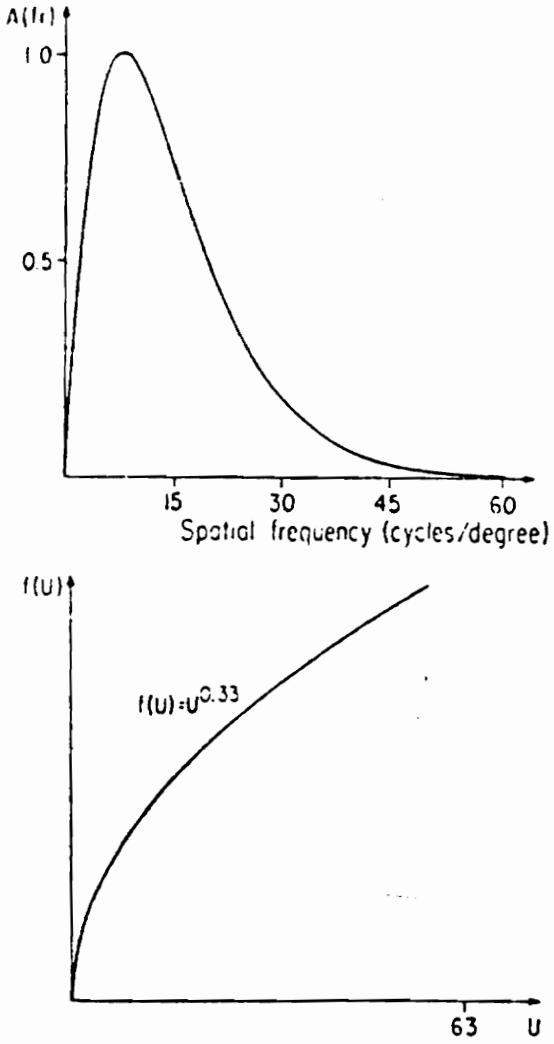

Figure 6. **The f(.) and A(.) functions.** (from Mannos and Sakrison [51].)

can be markedly improved for the same transmission rate. This is a form of "pre-emphasis", a technique used in several other signal processing applications.

2.4.2 Two-component source model

When evaluating the statistics of an image source it is useful, in line with (2.24), to consider the source output as log-intensity (or cube-root) rather than intensity itself. This accounts for the sensitivity of the observer to relative changes, which was already mentioned above. Fig. 7 shows a cube-root horizontal scan line of a typical image field. Such a field appears to exhibit very abrupt changes of root-intensity occurring at the edges of objects. Superimposed on these changes are much smaller variations due to local variations in reflectivity (texture). This obviously does not resemble the realization of a homogeneous Gaussian field having a continuous covariance function.

To assume that a source is Gaussian, when in fact it is not, entails a loss of coding effectiveness since the Gaussian source has the largest entropy of any source with a given covariance function. In fact, transform methods, while inherently close-to-optimum for Gaussian sources with memory, exhibit poor performance in coding the discontinuities or edges, giving blurred contours in general.

With a view to solving the problem, Yan and Sakrison [77] proposed a source model based on two additive components. The first or discontinuous component represents the presence of distinct objects within a scene and the nearly linear changes of root-intensity due to incident lighting. Approximations to the first component by using straight line segments [77] or higher order polynomial segments [54] have been reported in the literature. The second or remainder component, defined by substracting the approximated discontinuous component from the original root-intensity image, should then have the appearance of a sample field from a Gaussian random field. This model has proven to be very successful for image coding purposes. In fact, it is one of the bases for Second-Generation Image-Coding Techniques [42], in which intraframe compression ratios above 40:1 can be achieved with reasonable image quality.

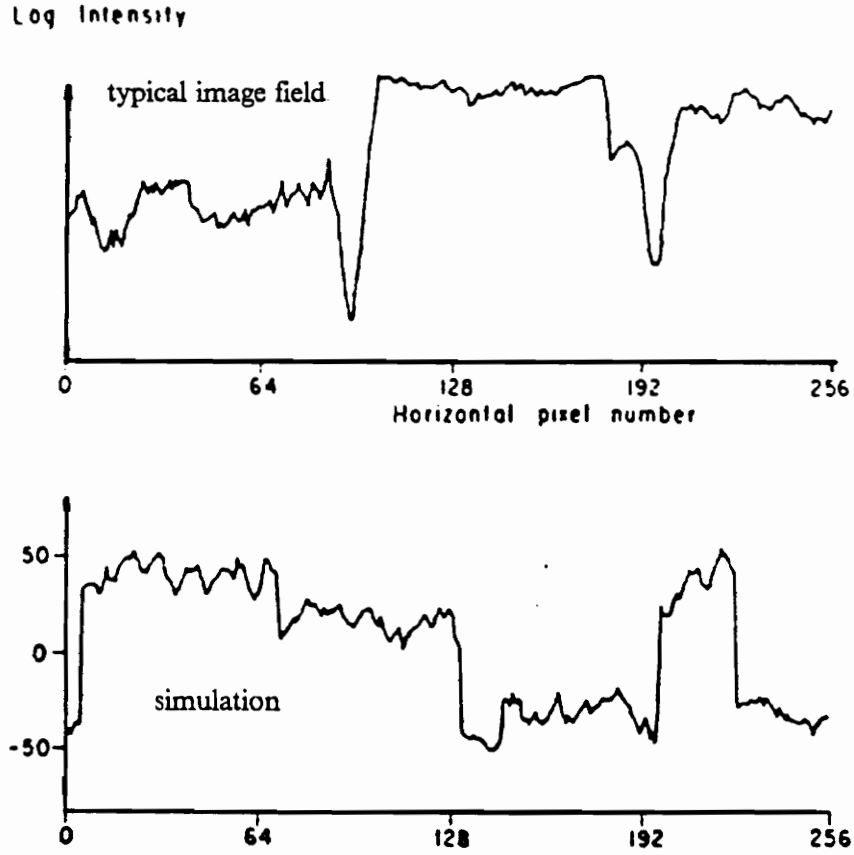

Figure 7. Log-intensity of a typical image. (from Anastassiou and Sakrison [4].)

A theoretical analysis of the performance of "two-component coding" of images compared to transform coding was published by Anastassiou and Sakrison [4]. By considering the source as the sum of two independent random entities X_1 and X_2 , each of which has a simpler description, they derived useful upper and lower bounds for $R_X(D)$ in terms of the individual $R_{X1}(D)$ and $R_{X2}(D)$ functions. If any (or both) of the two components have non-Gaussian statistics, these bounds will always be below $R_G(D)$ (which is the lower bound for transform coding systems). This leaves room for significant improvement over conventional one-component methods for coding of non-Gaussian signals in general , and images in particular. We will return to this model in Chapter 4.

2.5 Source coding for data compression

2.5.1 Scalar quantization

Scalar quantization is a lossy mapping Q(.) that transforms a (in general) mixed type random variable x into a discrete random variable x_q of less entropy. Since Q(.) is a many-to-one operator, x_q is a distorted (with distortion D) version of x. For a fixed total number L of output levels (i.e. fixed rate R), the goal of scalar quantizer design is to determine Q(.) such that D is minimum, as close as possible to the absolute "yardstick" that R(D) imposes. This results in a probability density function (PDF) optimized nonuniform quantizer.

In the quantization process the amplitude of x is compared to a set of decision levels. If the sample amplitude falls between two decision levels, it is quantized to a fixed reconstruction level lying in the quantization band. The scalar quantization problem entails specification of a set of decision levels d_j and reconstruction levels r_j such that if $d_j \leq x < d_{j+1}$ then the sample is quantized to a reconstruction value r_j.

If the selected distortion measure is the mean-square error (MSE), and the optimal decision levels have already been fixed somehow, then it is a classical estimation theory result [12] that the reconstruction levels that minimize the MSE are given by

$$r_{j,opt} = E(x \mid x\epsilon(d_{j,opt}, d_{j+1,opt}]) \; ; \quad j = 1,2,...,L \tag{2.26}$$

that is, the best estimate of the original signal x (in MSE sense) is given by its conditional mean in the corresponding j-th subinterval between two consecutive decision levels.

When looking for the optimum decision levels d_j, the necessary condition for a minimum MSE is given by

$$\frac{\partial \sigma_q^2}{\partial d_j} = 0 \; ; \quad j = 2,3,...,L \tag{2.27}$$

These two conditions define at least a local minimum. Nevertheless, if the log of the PDF is concave (a condition satisfied by most common PDF's), then global optimality is guaranteed [52]. Applying the conditions for optimality, Lloyd [47] and Max [52] derived the important result

$$d_{j,opt} = \frac{1}{2}(r_{j,opt} + r_{j-1,opt}) \; , \quad j = 2,3,...,L \tag{2.28}$$

and

$$d_{1,opt} = -\infty \; , \quad d_{L+1,opt} = \infty \tag{2.29}$$

In order to simultaneously solve for d_j and r_j, they also derived an efficient iterative technique. This PDF-optimized quantizer is then called the Lloyd-Max quantizer.

When compared with the rate-distortion function for a Gaussian random variable the Lloyd-Max quantizer performs within 1 bit/sample of the ideal limit for the Gaussian case. It is sometimes useful to split the quantizer performance in the form

$$\sigma_q^2 = \varepsilon_*^2 2^{-2R} \sigma_x^2 = \varepsilon_*^2 D_G(R) \tag{2.30}$$

When quantizing a Gaussian random variable, the correction term ε^2 is greater than one, and quantifies the departure from the ideal rate-distortion bound. For a non-Gaussian case, there are two components in the correction term: a positive one due to the nonideal quantization process, and a negative one due to the expected error reduction for non-Gaussian sources. This could render the correction factor less than one in some cases.

A somewhat improved performance can be obtained at the expense of extra encoding effort. The r_j MSE-optimized output levels in the Lloyd-Max quantizer will have, in general, unequal probability of occurrence. This means that the entropy of the quantized output x_q is less than the $R = \log_2 L$ bits resulting from coding the L levels in a straight binary format. A coding scheme that assigns shorter words to the more probable levels and vice versa is called Entropy Coding [36]. When applied to the Lloyd-Max quantized output, about 0.25-0.5 bits per quantized sample can be saved. By reformulating the optimum quantizer design problem in terms of minimizing the entropy of the quantized output subject to a distortion D [9], a rate within 1/4 bit of the rate-distortion function can be achieved, regardless of the PDF of the source.

When coding memoryless stationary processes, these scalar quantizers give the same performance as in the random variable case, i.e. the samples are quantized independently. Looking into the statistical properties of the quantization error for scalar quantizers, the following assumptions hold for most operating conditions [36]: 1) the PDF of the quantization error is uniform 2) the quantization error and the input signal are uncorrelated, and 3) the quantization error is white. In most cases, a Gain-Plus-Additive-Noise model can be used to describe them. In fact, this model has exactly the same structure as the optimal model of Fig. 3 that realizes the R(D) bound. The main difference however, is the fact that the PDF of the required quantization error in the optimum model is not the same as the actually flat PDF of the scalar quantizer. This limitation explains the gap of about 0.25 bit/sample still present in the most sophisticated minimum entropy scalar quantizers. As we will see later, this gap can be reduced by using vector quantization techniques.

In addition to the optimum scalar quantizers described below, another suboptimal quantizer is worth mentioning. It is the logarithmic quantizer in which the decision levels are spaced according

to a logarithmic law [36], i.e. the interval width for the decision levels is approximately proportional to the midpoint value of x in the interval [36]. Consequently, the signal-to-noise ratio (SNR) is almost independent of the amplitude of the source. This means that the performance of the quantizer is almost insensitive to changes in the statistics of the source. When compared to the SNR of the PDF-optimized quantizer, the SNR of the log-quantizer is significantly smaller. This is the price paid for the added flexibility.

2.5.2 Block quantization - Transform coding

We just showed that, for memoryless sources, the optimized scalar quantizers can perform very close to (within 1/4 bit) the R(D) limit. When the source has memory, the corresponding R(D) is reduced due to the redundancy or predictability of neighboring samples. Since a scalar quantizer codes every sample independently of the rest, the redundancy cannot be detected and removed. As a result, scalar quantization always gives the memoryless rates, far off from the corresponding true rates.

In order to code the source closer to the R(D) limit, blocks of N samples have to be coded as a whole. By doing that, redundancy can be detected and significantly reduced. Also, by choosing N sufficiently large (compared to the correlation depth of the source), consecutive blocks will be almost uncorrelated (i.e. there will be no mutual information between them), resulting in an overall performance close to the theoretical limit corresponding to an infinitely long sequence.

One way of implementing such a coder is to apply an invertible (i.e. information preserving) linear transformation to decorrelate the source prior to the scalar quantization process. At the receiver end, the inverse linear operation can be applied to the quantized transmitted signal, to obtain an approximation to the original signal. By decorrelating the source the scalar quantizer will again operate in the memoryless environment, where its performance is optimized. Unlike the case of memoryless sources however, the transformed samples will have rather different variances. This indicates that some criteria for allocating the available bits among the samples is necessary.

This whole process is called Transform Coding (TC). Fig. 8 depicts the block diagram of a generic transform coder. This is also called the Basis-Restricted Structure [56]. We will briefly discuss the components for an optimal basis restricted coder.

For a given zero mean process X having an N-th order autocorrelation matrix R_{xx}^N, there is a unique orthogonal transformation A that decorrelates the source. This is the discrete Karhunen-Loeve Transform (KLT)[36]. Besides its unique decorrelation property, it minimizes the geometric mean of the variances of the transform coefficients [36] (i.e. it has the best energy packing efficiency). It is worth noting that the KLT does not change at all the information content of the source. It only prepares the data for an efficient scalar quantization. For example, if the total number of bits were uniformly allocated in the transform domain, no compression gain would be achieved [74].

The problem of optimal bit allocation can be stated as follows: find the individual rates R_k corresponding to the k-th coefficient in the transform domain such that the average coefficient error variance

$$\sigma_q^2 = \frac{1}{N}\sum_{k=0}^{N-1}\sigma_{q,k}^2 \tag{2.31}$$

is minimized, with the constraint of a given average rate

$$R = \frac{1}{N}\sum_{k=0}^{N-1} R_k = \text{constant} \tag{2.32}$$

By using the Lagrange multiplier method, and considering the same quantizer performance (i.e. the same correction term ε^2) in both the original and the transform domains, it follows that [36]: the optimum bit allocation results in identical error variances in coefficient quantization. As a result, the average reconstruction error variance is proportional to the geometric mean of the transform domain coefficients. Therefore, the gain in SNR due to transform coding is given by the ratio of arithmetic mean (original domain) to geometric mean (optimized transform domain). When applying this optimized bit allocation scheme to a real coding scheme, it has to be slightly modified

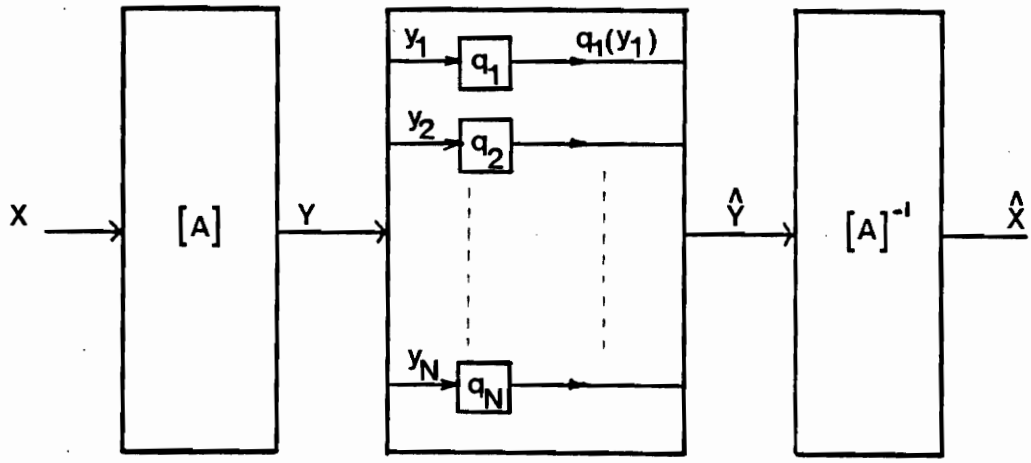

Figure 8. Transform coding model

to force the individual rates to be non-negative integers. Huang and Schultheiss [31] proposed a re-optimization procedure.

In applications, the KLT is generally replaced by a fast orthogonal transform, whereas the optimum bit allocation is typically replaced by a zonal sampling [60]. Nevertheless, the performance of transform coders for highly correlated Gaussian sources is close to the N-block rate-distortion function. When applied to image coding, acceptable quality can be obtained at 1-2 bits/pel.

Originating from the basis restricted structure, this scheme has a main drawback: the lower bound for TC is the Gaussian (with memory) N-block rate-distortion function. Real world signals are far from being Gaussian: the PDF of a typical image tends to be uniform, and speech signals are Laplace or Gamma distributed [36]. This results in a considerable increase in rate for a given distortion D.

Also, while it is true that the limit of $R_N(D)$ for infinitely long blocks is the true R(D), even tighter bounds can be found for N-block schemes [76]. This indicates that TC is suboptimal even in the Gaussian case. A significant improvement over TC can be achieved by using other block quantization schemes, based on codebook coders (or vector quantizers).

2.5.3 Block quantization using codebooks - Vector quantization (VQ)

An L-level N-dimensional vector quantizer is a mapping Q(.) that assigns to each input vector $X = (x_0,x_1,...,x_{N-1})$ a reproduction vector $X_q = Q(X)$, drawn from a finite reproduction alphabet, $C = \{y_i; i = 1,...,L\}$. The quantizer Q(.) is described completely by the reproduction alphabet (or codebook) together with the partition $S = \{S_i; i = 1,..,L\}$ of the N-dimensional vector space into the sets $S_i = \{X:Q(X) = y_i\}$ of input vectors mapping into the i-th reproduction vector (or codeword).

Like in the case of scalar quantizers, the design problem consists of determining the L decision regions and reconstruction values in an N-dimensional Hilbert space [26]. The optimum reconstruction values are the conditional means or centroids, over the corresponding decision region, of the N-th order PDF. By extending the results of Lloyd-Max quantizers to this

multidimensional case, Linde et al. [46] developed a similar iterative procedure to determine the optimum decision regions and reconstruction values. Due to the general unavailability of higher order statistics, they modified the procedure to work with a long training sequence.

Once designed, the quantization of each input vector imposes the computation of (in general L) N-th order distortions. This is in general a high computational load. Nevertheless, the main advantage of optimized vector quantizers is their performance being closer to the information theoretic bounds. Even for the memoryless case, VQ gives better performance than scalar quantization [36]. The VQ has the natural structure of the ideal system that implements the R(D) realization. Yamada [76] et al. derived tight bounds for Block quantizers that outperform the $R_N(D)$ bounds for TC, and are called VQ lower bounds.

3.0 Fast orthogonal transforms and their properties

3.1 Sinusoidal unitary transforms

In a remarkable paper [33], Jain introduced in 1979 a new family of unitary transforms. He showed that the well-known discrete Fourier, cosine, sine, and Karhunen-Loeve transforms (for first-order stationary Markov processes) are members of this family. Although they exhibit significant differences in performance when coding finite-length blocks, all members of the family have the following characteristics in common: (1) the basis vectors are sinusoidal sequences, (2) as the block length approaches infinity, they are all asymptotically equivalent to the Karhunen-Loeve Transform (KLT), (3) each member is the Karhunen-Loeve transform of a unique, first order Markov process, (4) poor edge coding performance due to the sinusoidal basis vectors.

For a first order stationary Markov process, the eigenvectors of the covariance matrix are analytically known to be of sinusoidal form. Their frequencies however, which are solutions of a transcendental equation, are nonharmonic in the general case. This fact precludes the existence of an FFT-like algorithm for the KLT. Fortunately,under limiting conditions on the correlation matrix, the KLT turns out to converge to a sine or cosine transform.

Of this family, we will analyze only the DCT and DST further, being the two most important fast members of the family. Although historically important, the DFT has inferior coding performance and requires complex arithmetic. Consequently, it is used less and less in most data compression applications.

3.1.1 Discrete cosine transform (DCT)

The Discrete Cosine Transform (DCT) has been widely used in image and speech processing. Several variations and versions of the DCT have been developed, and realized in hardware for real-time applications. Its remarkable properties and fast algorithm implementations have both contributed to its popularity.

The DCT was originally developed in 1974 [1]. It was then experimentally shown that the DCT performs very close to the optimal KLT for highly correlated sources. Fig. 9 shows its basis-restricted rate-distortion function ($\rho = .9$), which is practically indistinguishable from the optimal KLT one.

In 1982, Flickner and Ahmed [21] derived the DCT as a limiting case, as the correlation coefficient approaches unity, of a first-order Markov source. It was also shown analytically that the DCT outperforms the DFT in decorrelation power for all positive values of the correlation coefficient [29], and that it is relatively immune to statistical changes compared to the DFT. In addition, the DCT is much more effective in reducing the block-to-block edge effect than other discrete transforms [2]. As this transform is separable, multidimensional versions can be implemented by applying the 1D-DCT along each dimension. Fig. 10 shows the 1D-DCT basis vectors for $N = 16$.

Several fast DCT algorithms have been developed, all having the common computational requirement of $O(N\log_2 N)$ operations. While a fast DCT algorithm can be derived in terms of FFT algorithms [48], the full advantage of using real arithmetic can be achieved only with specialized

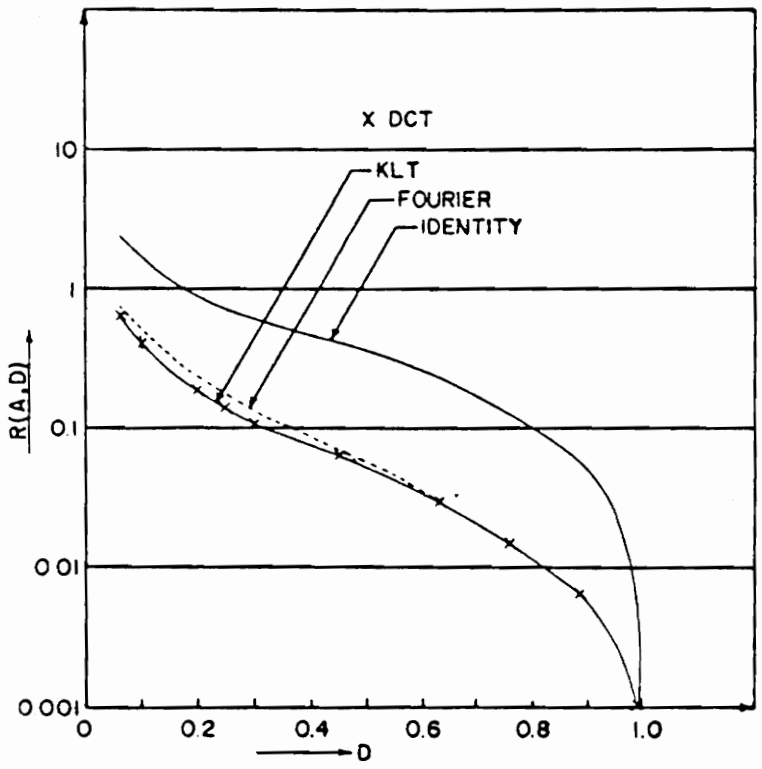

Figure 9. Basis-restricted rate-distortion function comparison. (from Ahmed and Rao [1].)

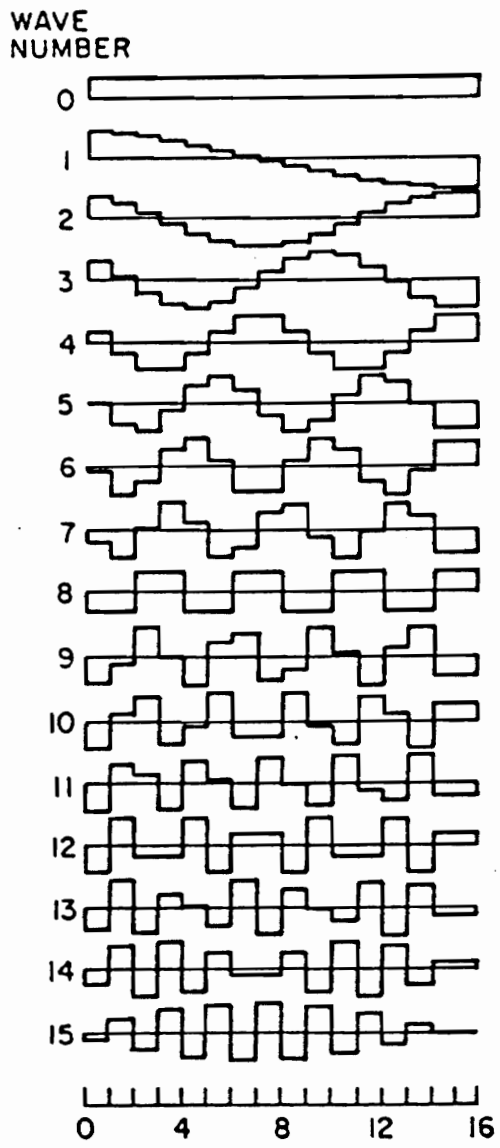

Figure 10. DCT basis vectors. (from Pratt [60].)

DCT algorithms [15]. A very close approximation to the DCT for N = 16, called the C-matrix transform (CMT), has proven excellent properties for fast hardware implementation [69].

3.1.2 Discrete sine transform (DST or DST1)

The DST represents the limiting case of the KLT of a first-order Markov process as the correlation coefficient approaches zero [33]. It has been shown that, for correlation coefficients up to +/-.5, the DST performs as well as the DCT [33].

The importance of the DST however, comes from the fact that, for a first-order Markov source, the KLT reduces to the DST if the boundary values of the sequence are fixed or known [35]. Considering the extra information required to code the boundary values, this fast KLT algorithm has an overall performance close to the DCT performance [40].

Fig. 11 shows the corresponding basis vectors for N = 15. Note the absence of a constant vector. This suggests that the DST performance can be improved considerably if the mean of the sequence is removed and processed separately. The DST admits an FFT-based fast algorithm [33]. Still faster algorithms, using only real arithmetic, were developed [79].

3.2 Walsh-Hadamard transform (WHT)

The Walsh-Hadamard Transform (WHT), also called Hadamard Transform or Bipolar Fourier Transform (BIFORE) is a unitary transform based on the complete set of two-valued orthonormal Walsh functions, which may be defined in terms of difference equations, or from Rademacher functions. The known one-to-one relationship between certain Hadamard matrices of rank 2^m (m integer) and the sampled Walsh functions, gives an alternative way of defining and generating the orthogonal set of $N = 2^m$ N-point discrete basis functions, via Kronecker products

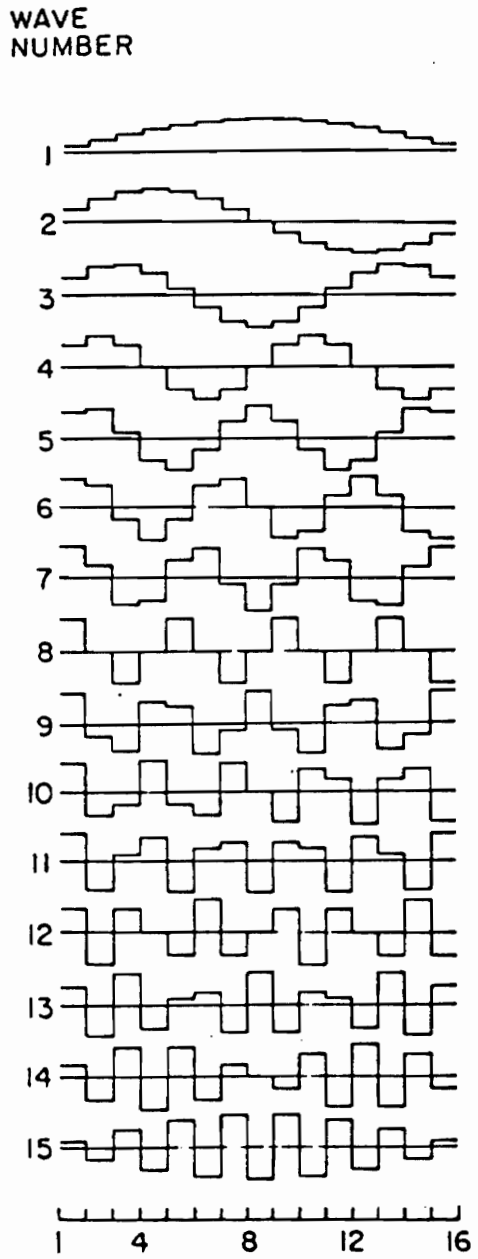

Figure 11. DST basis vectors. (from Pratt [60].)

of half-order matrices. While Walsh functions exist for any integer N, this is not in general the case for Hadamard matrices. To avoid ambiguous definitions, the term WHT generally refers to the common case of N being an integer power of two. In conjunction with Fourier harmonics, Walsh functions probably represent the most used orthogonal set in communications. Excellent developments of Walsh function theory and their applications are available [10,30].

Walsh-Hadamard signal analysis closely resembles Fourier harmonic analysis in both geometrical and analytical characteristics. Fig. 12 shows the first sixteen Walsh-Hadamard and Fourier basis vectors superimposed, revealing a remarkable similarity between them. Even though the zero crossings of the two functions do not correspond exactly, the Walsh functions look like an infinite clipped version of the Fourier sinusoids. This correspondence led to the definition of sequency ("one half of the average number of zero crossings per unit time interval") as an extension of the trigonometric frequency concept. In addition, the ideas of spectral decomposition, phase, correlation, and filtering have been extended to this generalized frequency context, and successfully applied to communications, signal processing, and pattern recognition problems [10,30].

The WHT basis functions allow for a Cooley-Tukey type of fast algorithm. Several algorithms have been reported [41,58], displaying differences in speed and memory requirements, and in the ordering of the transform coefficients ("sequency" ordered or "Hadamard" ordered). However, any fast WHT algorithm has a computational complexity of $O(N\log_2 N)$ operations. Thanks to the orthogonality and symmetry properties of the Hadamard matrices, the direct and inverse WHT are identical.

The main advantage of using the two-valued Walsh basis functions is that multiplications are replaced by add/subtract operations, and only real arithmetic is required. As a result, approximately a tenfold speed increase can be achieved over regular FFT software algorithms, and digital hardware implementations are dramatically simplified.

The WHT is separable and it can easily be extended to multiple dimensions. Another unique property of the WHT is that the 2D-WHT of a matrix is equivalent to the 1D-WHT of a vector resulting from a lexicographic ordering of the elements of the matrix [6]. Fig. 13 shows the

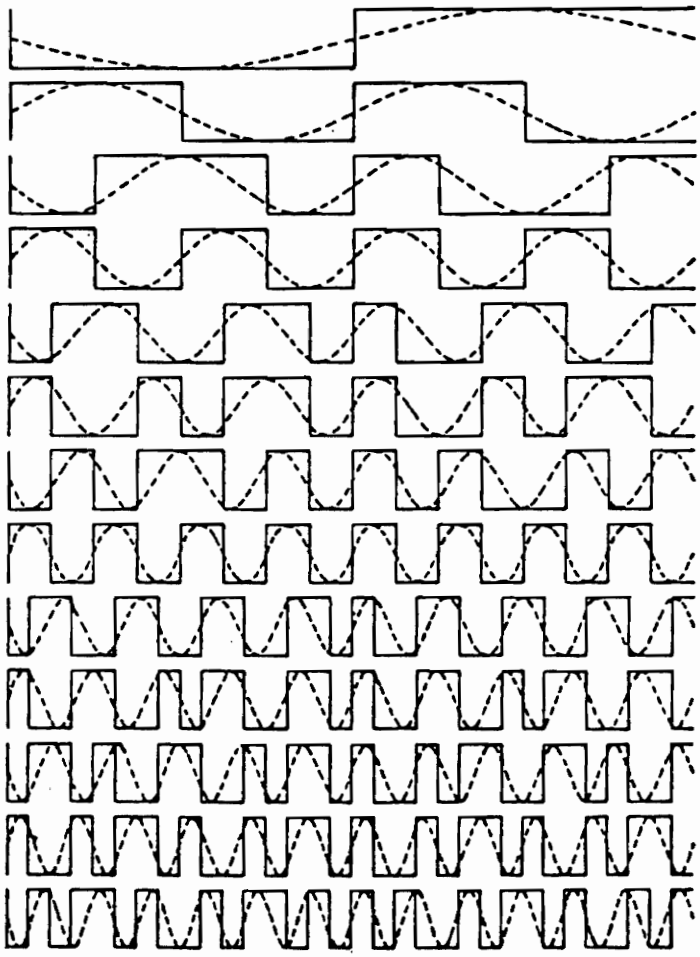

Figure 12. Walsh and Fourier basis vectors. (from Harmuth [30].)

two-dimensional 8x8-point basis functions. Note that the (i,j) component can be obtained by performing the logical exclusive-OR between the (i,0) and the (0,j) components.

When considering coding performance, the WHT is not as effective as transforms with sinusoidal basis vectors, like the DFT or DCT. In fact, WHT performance is not even asymptotically optimum with respect to block size. Fig. 14 shows typical coding MSE as a function of the block size for several unitary transforms when coding a highly correlated ($\rho = .95$) separable first-order Gauss-Markov source, using zonal sampling with 4:1 coefficient reduction. It can be seen that there is almost no gain for blocks larger than 8 x 8 points. However, for block sizes less than 16 x 16, the WHT performs as good as or better than the DFT. The 8 x 8 block size then represents the best complexity versus performance tradeoff for most WHT hardware image processing implementations [58].

The energy-packing efficiency (EPE) of the WHT has been analytically evaluated by Kitajima [39] for zero mean wide sense stationary processes, later being extended to the generalized discrete transforms by Yip and Rao [78]. Again, the WHT performs very close to the DFT when EPE is considered. Due to its performance to complexity ratio, the WHT is accepted as a good substitute for the DCT in real applications. In addition, the rectangular basis functions in the WHT suggest that somewhat better relative performance should be achieved in real images having many edges and discontinuities, while such features are not present in simple first order stationary test sources.

The WHT and other transforms reviewed in this chapter, except the Haar transform, have one-half even and one-half odd basis vectors, and are consequently called even/odd transforms (EOT). It has been shown that the conversion from one even/odd transform to another requires only multiplication by a sparse matrix, having at least one half of its elements equal to zero [37]. Since among all the EOT's, the WHT is by far the easiest to implement in hardware (or software), it has become the natural first step in the computation of any other EOT. For small N (up to 16), and whenever several transforms are required, the computation via the WHT has real advantages over separate implementations for each transform. Very fast implementations of the close to DCT performing CMT Transform [69] and the Slant Transform, via WHT and using ROM lookup tables to perform the sparse multiplication, are feasible in the actual technological context.

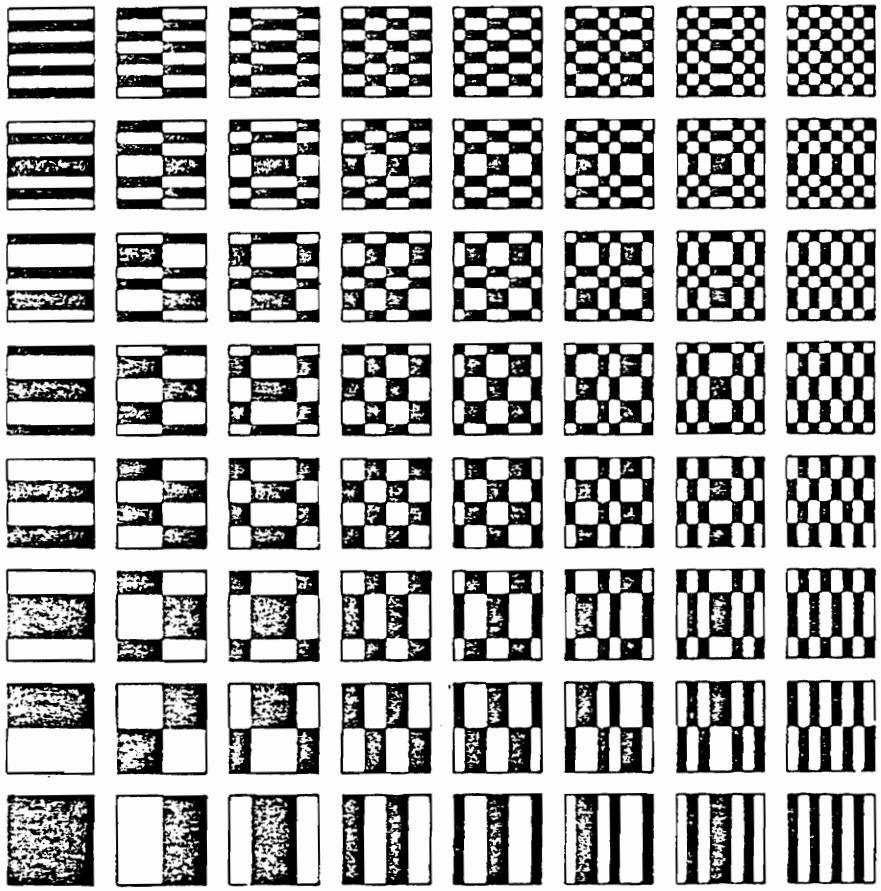

Figure 13. 2D-WHT basis functions. (from Pratt [60].)

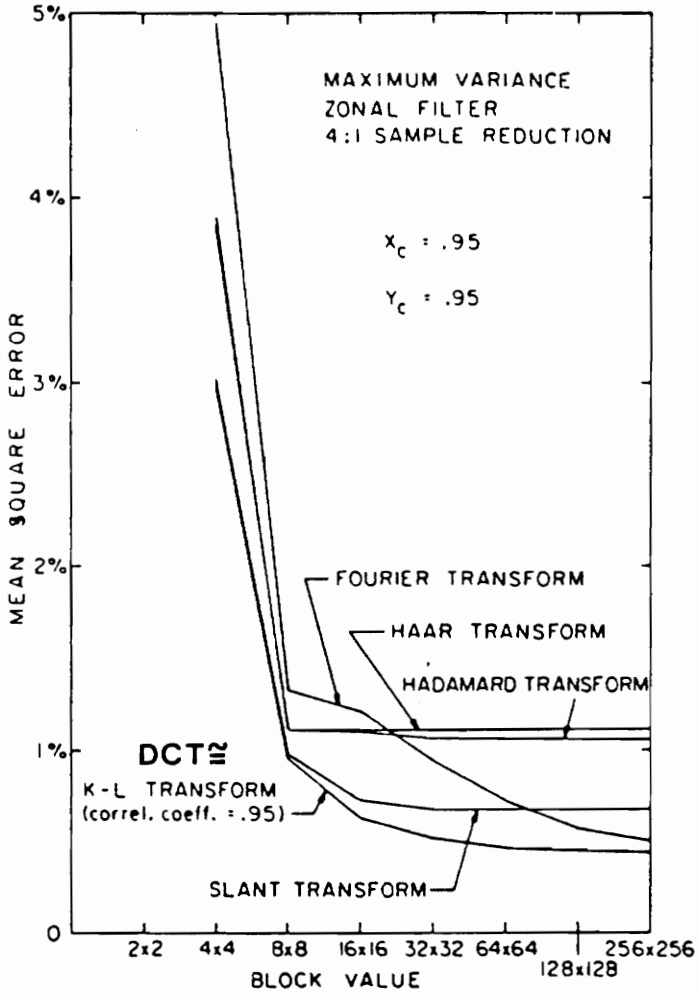

Figure 14. Typical transform coding performance. (from Pratt et al. [59].)

3.3 *Haar transform (HT)*

The Haar Transform (HT) is a unitary transform based on the complete set of orthonormal three-valued Haar functions, which defined in a recursive form, only exist for N an integer power of two [10,30]. Its first two vectors are identical to the first two WHT ones. After that, new vectors are recursively generated by squeezing and shifting the previous ones. Fig. 15 shows the resulting vectors for N = 16. The essential characteristic of the Haar functions is seen to be a constant value everywhere except in one sub-interval where a double step takes place. Note that, depending on the sub-interval width, the corresponding amplitudes vary accordingly, in order to satisfy the unit norm condition.

Like other transforms, the HT is separable and easily extendible to several dimensions. The two-dimensional Haar basis functions are shown in Fig. 16. Unlike other transforms however, the HT has both global (the first two) and local basis vectors. It can be likened to a sampling process in which rows of the transform matrix sample the input data sequence with finer and finer resolution, increasing in powers of two.

In image processing applications, the HT provides a transform domain in which a type of differential energy is concentrated in localized regions. Due to this property, the HT can code edges much better than other transforms. Some good examples of edge extraction using the HT were reported in the literature [19,61]. On the other hand, the HT has the poorest performance in coding stationary sources. Pearl et al. [57] showed that, for a pure second order Gauss-Markov process, almost no gain in decorrelation is observed in the HT domain. Similarly poor performance can be expected for higher order processes. Reasonable whitening can only be achieved when the process exhibits considerable correlation between adjacent samples. Fig. 14 shows the HT performance for the highly correlated first order source case. Its behavior is almost indistinguishable from that for the WHT.

Since the Haar matrix is not symmetric, the direct and inverse HT require separate algorithms. The HT admits a fast implementation. In fact, due to the sparseness of its basis vectors, a fast HT

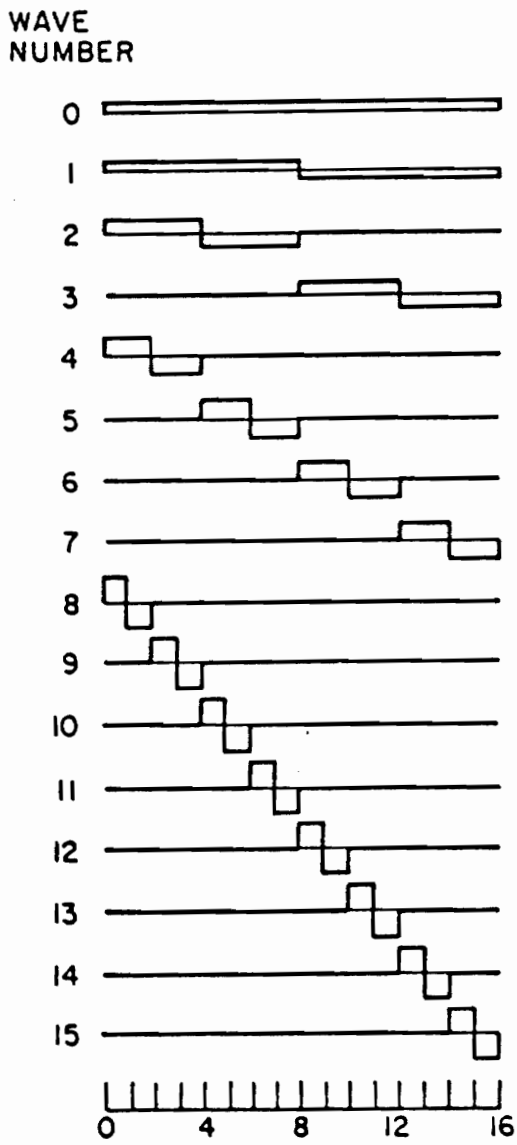

Figure 15. Haar transform basis vectors. (from Pratt [60].)

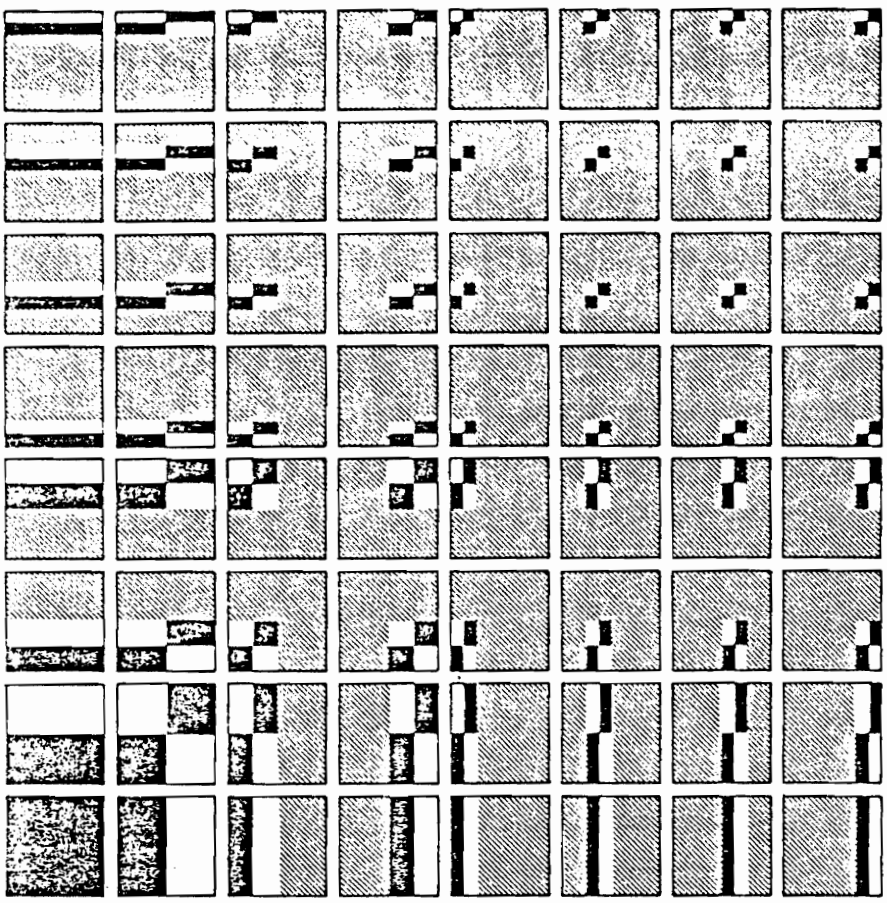

Figure 16. 2D-Haar basis functions. (from Pratt [60].)

algorithm requires only 2(N-1) add/subtract operations, being the fastest linear transformation presently available [10]. It has, however, integer powers of $\sqrt{2}$ as multipliers. By adopting a normalization process, these numbers are converted to integer powers of two, resulting in what is called the rationalized HT (RHT).

The HT behaves quite differently from any other transform described in this chapter, mainly because of its unique ability to efficiently code local nonstationarities such as edges and discontinuities. This property makes the HT an obligatory choice for combination with other transforms in a multiple bases image processing setting.

3.4 Slant transform (ST)

The Slant Transform (ST) was originally developed in 1971 by Enomoto and Shibata [20] for the single case of $N = 8$. This was later extended by Pratt, Chen, and Welch [59] for N any power of two using certain recursive properties. The ST is a transform designed to possess the following properties: (1) orthonormal set of basis vectors, (2) one constant basis vector, (3) one slant basis vector (monotonically decreasing in constant size steps from a maximum to a minimum amplitude, (4) sequency property, (5) fast computational algorithm, (6) variable size transformation, and (7) high energy compaction. Similar to other orthogonal transforms, the ST is separable, permitting two-dimensional processing by using successive one-dimensional transformations.

As the ST is an even/odd transform (EOT), it can be expressed via the WHT using a suitable conversion matrix. In a paper about a computational algorithm for the ST [75], Wang revealed the relationship between the WHT and the ST, and showed that the ST may be approached by a series of steps which gradually change the WHT basis vectors into the ST ones. By doing this, a total of $N\log_2 N + N - 2$ additions and subtractions, together with 2(N-2) multiplications, are required to compute it. In addition, this approach is suitable for efficient hardware implementation.

Specifically designed for image coding, the ST possesses a discrete sawtoothlike basis vector which efficiently represents linear brightness variations along an image line. Fig. 17 shows the ST basis vectors for $N = 16$.

In terms of signal compaction, the ST exhibits excellent performance. Fig. 14 shows typical coding MSE vs. block size for the previously described source. It can be seen that the ST outperforms both the HT and the WHT for any block size, being close to the optimal KLT. It also outperforms the DFT for block sizes less than 64 x 64. For very large block sizes (not generally used in image processing), the DFT outperforms the ST, becoming asymptotically equivalent to the KLT.

The ST has been successfully used in monochrome and color TV signal coding, because of its combination of very good energy compaction performance and very fast computational algorithm, suitable for hardware implementation. Although the importance of the ST has been overshadowed by the discovery of the DCT, which is widely accepted as the best substitute for the KLT for highly correlated signals like images or speech, its slant basis vector and its computation via WHT make it a very attractive set for the multiple bases representation analyzed in this research work.

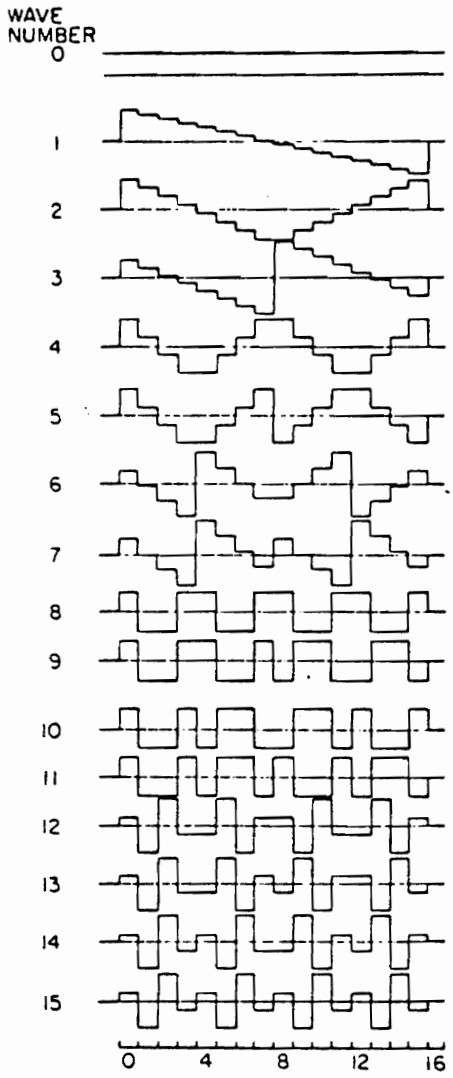

Figure 17. Slant transform basis vectors. (from Pratt [60].)

4.0 Representation and coding of signals using multiple bases

4.1 Signal representation using linear transformations

A discrete time N-point signal can be expressed as a vector in an N-dimensional Hilbert space. The natural basis that represents it is given by the canonical or natural basis set [12] in correspondence to the coordinate axes of a Euclidean space. In this representation, each time sample is projected onto a separate coordinate axis.

Let B be a matrix formed columnwise by any complete set of orthonormal basis vectors that span the whole space. A vector Y_I can now be expressed in terms of the new basis set as follows

$$Y_B = B^T Y_I \tag{4.1}$$

The new signal Y_B is obtained by applying a linear transformation to the original Y_I. In terms of metric spaces, both Y_I and Y_B represent the same vector, expressed in terms of different coordinate systems (B for Y_B, and identity I for Y_I). This linear transformation is nothing but a change of basis [12].

Since B is orthonormal, B^{-1} exists and is given by its transpose. Then, the inverse transformation follows

$$Y_I = (B^T)^{-1} Y_B = B Y_B \tag{4.2}$$

In addition, due to the orthonormality of B, the Parseval relation applies, i.e.

$$\|Y_I\|^2 = \|Y_B\|^2 \tag{4.3}$$

This indicates that the linear transformation is a magnitude preserving multidimensional rotation.

By appropriately choosing the basis vectors in B in accordance with the statistics of the signal to be represented, the most relevant features of the signal, usually hidden in the canonical form, will become more visible and separated in the transformed signal Y_B. Each basis vector, ranked in importance according to the average projected power onto it, will represent a feature of the original signal Y_I. This basic idea of multidimensional rotations has been successfully exploited not only in the context of source coding for data compression, but also in pattern recognition problems [5].

For given signal statistics and stationary processes, the optimum orthogonal basis set will pack on the average the most energy in the least number of coefficients (or features). Since the total power is preserved, this means that the geometric mean of the coefficient variances is minimized. The resulting optimal transformation is the already mentioned Karhunen-Loeve Transform (KLT). In addition, the coefficients in the transform domain are uncorrelated (linear dependencies are removed).

The process of data compression using transforms involves the truncation of the signal expansion, followed by some bit allocation scheme for the remaining coefficients. In terms of vector spaces, the N-dimensional signal is projected onto an M-dimensional subspace ($M < N$). Let B_M be an M-vector subset of B. Then, the M-truncated transform representation results

$$Y_{B_M} = B_M^T Y_I \tag{4.4}$$

Due to the truncation involved, Y_I cannot be recovered exactly from Y_{B_M}. Since Y_I was projected orthogonally, the following least-squares approximation results

$$\hat{Y}_I = B_M Y_{B_M} \simeq Y_I \tag{4.5}$$

At this point, for a uniform bit assignment among the M retained coefficients, the compression ratio would be N/M:1. For a given distortion measure, basis set B, and rate R, the overall quality of the approximation will be determined by two factors: a) truncation strategy (sampling methods), and b) bit allocation.

There are two basic strategies for sample selection: zonal sampling and threshold sampling. In zonal sampling, the selection of the retained samples is fixed a priori, based on the ranking of the coefficient variances in the transformed domain. In most cases, the truncation affects the high frequency components, resulting in some loss of signal detail. In any case, the estimated variance and expected probability of occurrence of each component are used to determine the number of quantization levels and the code word length in the quantization phase. Summarizing, zonal sampling is a form of fixed frequency filtering in which the "passband" is determined by the statistics of the signal under analysis.

With threshold sampling the coding is performed on the subset of transform domain samples that are larger than a certain threshold. The subset of retained samples is not fixed, and depends on the particular realization of the process. The amplitude of the retained coefficients is again quantized and coded. It is also necessary to code the position of each significant sample (book-keeping information). This is efficiently done by coding the number of nonsignificant samples between significant ones (run-length coding [36]).

Threshold sampling performs somewhat better than zonal sampling due to its adaptive nature. It can keep track of more statistical detail than just the linear dependencies detected and (partially) removed by using zonal sampling. The procedure itself is nonlinear: the sampling of the sum of two signals does not correspond (in general) to the sampling of the individual components.

The application of an orthogonal transform followed by threshold sampling can be considered as a timid step in the direction of the codebook type of coders, or vector quantizers. Vector quantizers have the ability to remove higher order dependencies (or redundancies) than just the linear dependencies done by transform methods followed by scalar quantization. In the transform

plus threshold sampling approach however, the set of features (or shape codebook) is constrained to have just N members, and to satisfy conditions of orthogonality. With such a restricted scheme, the performance improvement over zonal sampling is only marginal.

4.2 *Multiple bases representation (MBR) - Rationale*

As we saw in the previous section, in threshold sampling the signal is approximated by a linear combination of M selected basis vectors. Over all possible M-tuples, the selected vectors minimize the approximation error for that particular signal realization. In addition, over all orthogonal basis sets, the average approximation error is minimized for the KLT basis sets. No better approximation can be obtained under this setting.

Consider the more general case of having L > N unit norm vectors in an N-dimensional Euclidean space (using squared error criterion). They can be designed such that every possible N-tuple of vectors (resulting from picking N vectors out of the total of L) forms a linearly independent set. Let B_j be the matrix having the j-th N-tuple as column vectors. These vectors form a (in general nonorthogonal) basis that spans the N-dimensional space. Consequently, the signal Y_I admits the exact representation

$$\begin{aligned} Y_I &= B_j Y_B \\ Y_B &= B_j^{-1} Y_I = R_j Y_I \end{aligned} \tag{4.6}$$

where

$$R_j B_j = I \tag{4.7}$$

The matrix R_j, called the reciprocal basis, will coincide with B_j^T for the case of an orthonormal basis. Similarly, if a generic k-th M-vector subset out of the total of L vectors is considered, then the best approximation in MSE sense is given by the orthogonal projection of Y_I onto the subspace spanned by the k-th M-vector subset [12]. This corresponds to the least squares solution for the overdetermined system of equations

$$\hat{Y}_I = B_{M,k} Y_{B_{M,k}} \simeq Y_I \tag{4.8}$$

The solution to this system, using the pseudo-inverse [12], is then

$$Y_{B_{M,k}} = (B_{M,k}^T B_{M,k})^{-1} B_{M,k}^T Y_I = R_{M,k} Y_I \tag{4.9}$$

resulting in the following squared error

$$MSE_{M,k} = Y_I^T (Y_I - B_{M,k} Y_{B_{M,k}}) \tag{4.10}$$

This gives the optimum expansion of Y_I in terms of the k-th M-vector subset. Now, let S be the set whose elements are all possible M-tuples of vectors picked from the total L. This set will have a total of $\binom{L}{M}$ competing M-dimensional subspaces. Subject to the unique constraint of using M vectors drawn from the set of L vectors, the optimal representation (in MSE sense) gives a minimum error

$$MSE_{M,opt} = \min_{k \in S} MSE_{M,k} \tag{4.11}$$

i.e. we choose the subspace that gives the closest projected point. This can be interpreted as the natural extension of threshold sampling to the more general case of having L > N vectors in an N-dimensional space.

The main idea behind this representation is the possibility to pack the energy of the signal in a reduced number of features (or basis vectors). By appropriately choosing the L participant vectors, the N-dimensional space will be subdivided into a dense grid of vectors and hyperplanes. As a result, the average projection-onto-closest-hyperplane error will be reduced with respect to the

N-orthogonal basis case. In addition, this compact representation will allow for efficient quantization over a variety of statistical source models.

The computation of this unrestricted representation involves the application of full search procedures which, in the general case, require the computation of $\binom{L}{M}$ distortion measures. Since each computation of distance involves the solution of a linear system of equations, this search becomes a formidable task even for small values of M and L. By restricting the search to a subset of the total set S, we will see that very good suboptimal representations can be achieved with a dramatic reduction in computational cost. Also, by imposing certain structure on the basis vectors, substantial further reductions are achieved. In this work, we constrain the representation vectors to be the basis vectors of several fast orthogonal transforms.

It is worth mentioning that optimality in the representation does not extend to optimality in quantization performance. In fact, quantization conditions will reduce the search over an even smaller subset of S. This will be reflected in the quantization schemes proposed for MBR coding, having additional impact on the computational cost reductions.

4.3 Application of MBR to source coding

In this section, the application of multiple bases representation of signals to source coding problems is analyzed. Starting from a nonspecific signal model, the application of MBR to source coding is interpreted within the framework of vector quantization. This gives some useful insight in the expected performance of MBR coding, and criteria for designing the basis sets and quantization schemes. In particular, the combination of several fast orthogonal transforms is proposed, to generate the codebook. Based on this highly structured codebook, a recursive full search technique called recursive residual projection (RRP) is proposed.

The application of MBR to the case of image-like signals is then analyzed. By appropriately choosing the participating basis sets, an efficient two-component source decomposition is devised.

This decomposition, supported by the two-component source model theory and the corresponding rate-distortion bounds, will allow for substantial improvements over transform coding methods. Some ad-hoc splitting techniques are proposed also.

The selection of the participating fast transforms, based on the source characteristics, is discussed. Finally, the feasibility and advantages of this representation are substantiated by performing some computer simulations. The performance of the proposed schemes is evaluated in terms of rate-distortion and computational complexity, and compared with transform coding methods.

4.3.1 MBR coding as a vector quantization process

Vector quantization is intrinsically superior to transform coding, and other suboptimal and ad-hoc procedures since, regardless of the statistical structure of the source, it achieves rate-distortion performance subject only to a constraint memory or block length of the observable signal segment to be encoded [22]. For signals of length N and a codebook size L, the main task in the quantizer design phase is the partitioning of the N-dimensional space into L disjoint regions, and assigning to each one a codeword and a representation vector.

Once the codebook has been designed, the quantization of each input vector (with minimum distortion) is performed by finding, among all possible reproduction vectors, the reproduction word closest to the vector to be coded. This procedure, called full search quantization, requires the computation of L distance measures, each having a typical complexity of N operations. If, in addition, the codewords have equal length (i.e. $R = \frac{\log_2 L}{N}$ bits/sample), it can be easily shown that, for both codebook design and full search quantization, the computational and storage costs are proportional to

$$N2^{RN} \tag{4.12}$$

that is, the computational and storage costs are exponential in the number of dimensions and the number of bits per dimension. This is the main drawback of vector quantization. Nevertheless, by imposing some structure in the codebook design and search procedures, substantial reductions in the computational and/or storage costs can be achieved, at the expense of somewhat degraded performance [26].

Consider now the set of vectors resulting from collecting the basis functions of P fast orthonormal transforms. This set will have a maximum of L = PN different vectors. Then, for a zero-mean unit-variance signal, the 1 vector MBR (M = 1), and the vector quantization with the L-vector codebook are the same thing. The immediate advantage in computational cost is given by the fact that each distortion computation takes only $\log_2 N$ operations. With respect to storage requirements, when in-place fast algorithms are used, a total of just L locations is necessary. Nevertheless, due to the imposition of a fixed structure in the codebook, some loss in rate-distortion performance will result.

For given source statistics and squared error distance measure, an N-point vector quantizer is optimal if it minimizes the mean-square Euclidean distance between the signal and its quantized version. It can be shown [36] that the optimal quantizer satisfies the following two necessary conditions for optimality: 1) for a given partitioning of the N-dimensional space, each optimal reproduction vector is the centroid (conditional mean or center of gravity) of the partition, and 2) for a given set of reproduction vectors, the optimal partitioning is determined by the so-called nearest neighbor rule or Voronoi partition [26]. That is, every point belonging to the j-th partition is closer to the j-th reproduction vector than to any other reproduction vector.

Being given the joint statistics of the source (if available) or a long training sequence of data, Linde, Buzo, and Gray [46] developed a vector quantizer design scheme. Based on an iterative fixed point type of algorithm that enforces the two conditions for optimality, an (at least locally) optimal design is achieved. This algorithm, a multidimensional version of the Lloyd-Max procedure for scalar quantizers, is called the clustering or LBG algorithm.

The reproduction vectors in MBR are the (scaled) basis vectors of the participating fast transforms. By applying the nearest neighbor rule to them (2-nd condition), the partitions are

automatically determined. The first condition for optimality however, will not be satisfied in the general case, leading to a resulting suboptimal codebook design. Nevertheless, for large L and N, the theory of random codebooks [26] predicts asymptotic optimality when the reproduction vectors are likely realizations of the source process. Therefore, by increasing the number of participating fast transforms, and choosing (or designing) them such that their basis vectors resemble likely realizations of the source, good suboptimal performance can be expected.

There is however a basic limitation in this approach, that arises when medium to low average distortion is desired. In such a case the required number L of representation vectors increases significantly. As a result, the number of fast transforms available and possible fast hybrid versions designed from them become insufficient to populate the codebook. By treating the resulting quantization error as a new signal to be quantized and coded, a cascade or multistage vector quantization scheme results [38]. This is the basic idea of the proposed recursive residual projection (RRP) procedure.

4.3.2 Recursive residual projection (RRP)

The recursive residual projection (RRP) procedure, applied to both source representation and/or quantization, is (in principle) a way to obtain small representation/coding errors, starting from a reduced set of basis vectors. In RRP, the original signal is first represented/quantized by projecting it onto the best (nearest neighbor) approximating vector of the set. The representation/quantization error signal, called residual, is again projected and treated like the original signal was, resulting in a second (and smaller) residual. This error reducing procedure can be recursively repeated until some error threshold is reached. Fig. 18 shows a flow-graph of the RRP procedure.

When applied to signal representation, the M-stage RRP procedure results in a simple way of choosing the M-vector representation subspace. It is a tree search procedure: having fixed the previous k-1 vectors, RRP picks the k-th vector in an optimal way. Since not all possible

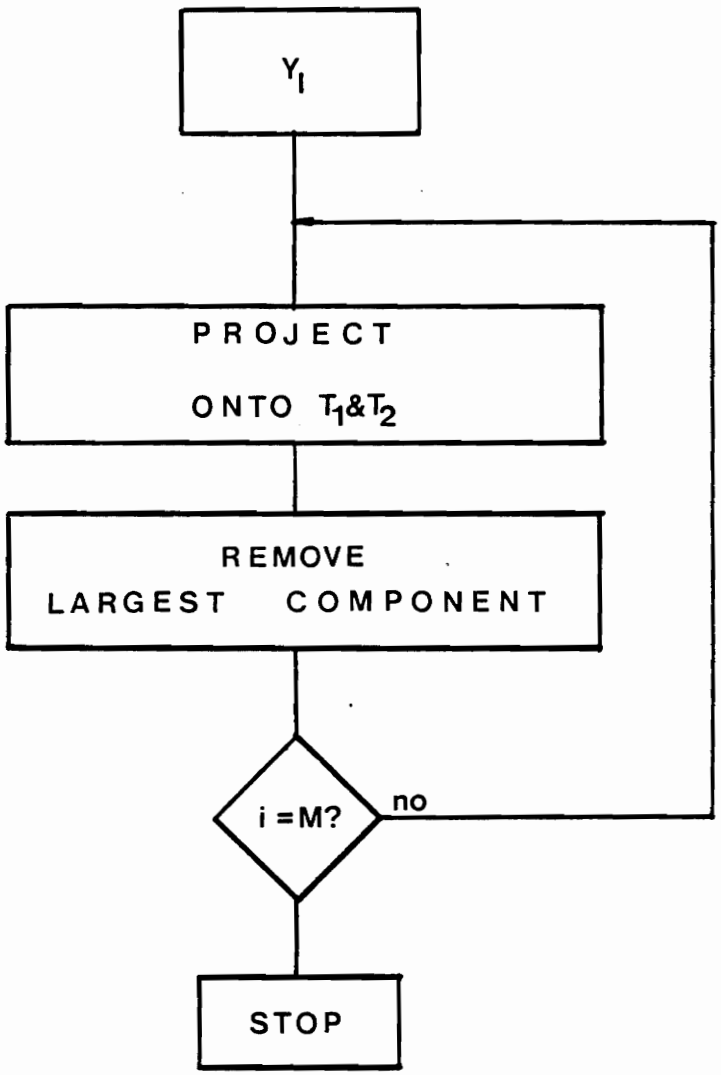

Figure 18. Recursive residual projection (RRP)

M-dimensional spaces are considered at once (full search), RRP results in a suboptimal representation performance. Nevertheless, it will be shown by simulations that the performance gap is relatively small. In addition, among all possible M-dimensional subspaces, RRP will discard most subspaces formed by close-to-dependent vectors, due to its procedure of projecting onto the nearest neighbor. This results in advantage over the full search approach when quantization is considered.

When used in a quantization context, RRP can be considered as a cascaded or multistage vector quantizer [38]. That is, a sequence of VQ stages, each operating on the residual of the previous stage. The set of participating vectors at each stage will be the same. Nevertheless, since the variances of the residuals vary from stage to stage, a gain factor has to be included in the quantization scheme. At each stage, the resulting codebook will be the Cartesian product of a vector codebook describing the shape of the reproduction vector, and a scalar codebook describing the gain or energy. This quantizer is called "shape-gain VQ" [63]. In our case, once the participating fast transforms have been selected, the shape codebook is fixed. The only freedom we have is in the design of the scalar gain quantizer. The optimized Lloyd-Max procedure implemented with a training sequence can be successively applied at each stage. Moreover, each shape vector could be assigned an individual gain quantizer, optimized with respect to its probability density function (PDF).

Multistage VQ is a suboptimal procedure. The reduction in performance with respect to the optimal VQ comes from the fact that, after each stage, the residuals are pooled together to form the input to the next stage. Since, in general, the residual PDF's from the different clusters will be different, pooling them together will result in a single PDF that will lose many of the dependencies that existed in the initial clusters. That is, the residuals tend to have flat spectrum and uniform PDF. As a result, the performance gap with respect to optimal VQ will increase with the number of stages. The main advantage of multistage VQ however, is the dramatic reduction in computational and storage costs [26].

In the RRP case with $L = PN$, P fast transform computations are in principle required at each stage. Having M stages, a total of order $MPN\log_2 N$ computations would be necessary for the shape codebook. Nevertheless, by precomputing and storing the representation of each and every basis

vector in all other transform domains, no further transform computations will be necessary after the first stage, resulting in a total of order $PN(\log_2 N + M - 1)$.

The other basic interpretation and application of RRP is in terms of signal splitting. Considering that each different transform has basis vectors (or features) tailored to certain statistical signal characteristics, RRP represents a way to split the signal process into additive components, each one of them having a simpler statistical description. This will have important impact when coding imagelike signals or other composite sources.

4.3.3 Application of MBR to composite source coding - RRP splitting

In many cases of practical interest, modelling the statistical properties of the source signals using a single source model leads either to a very poor description of the source or to a very complicated often nontractable formulation. The typical example is the case of image sources. Modelling images by using simple highly correlated Gaussian autoregressive sources has been a common practice in image processing. In using these, performance has usually been compared with that of transform coding methods. Such a simple model however, does not take into consideration that the log-intensity (or cube root-intensity) of real gray-scale images, in a still simplified model, is composed of two nearly independent basic features: a discontinuous one due to distinct objects (edges or contours) and a continuous one due to such effects as texture [77]. As a result, methods like transform coding, close-to-optimal for single source models, have poor edge coding performance, resulting in blurred contours.

The idea of coding images based on a two-component source model was introduced by Yan and Sakrison [77]. This concept was followed by several applications and second-generation methods [42,54]. In those methods, the texture component is processed by using standard transform coding, whereas the edge component is treated and coded by implementing ad-hoc nonlinear procedures. In any case, compared to transform coding, better coding performance and overall subjective quality is achieved. This performance improvement is supported theoretically by the

work of Sakrison [4]. As mentioned in Chapter 2, the proposed upper and lower bounds to the rate-distortion function for additive sources. In particular, he assumed a first order Gauss-Markov model for the texture component and a zero-mean stationary first-order Markov uniformly distributed jump sequence model for the edge component. The resulting additive process, while resembling a typical observed scan line sequence of log-intensity function of an image, was shown to exhibit a much better rate-distortion bound than the single model Gaussian source bound.

The basic point in the application of an m-component source model is being able to separate the available signal into m components, each of which has a simpler (non-Gaussian in general) statistical description than the original source signal, and to code each component efficiently by using vector or even scalar quantization techniques. In image processing, the goal will be to split the log-intensity signal into, for example, well-defined texture and edge components.

By selecting the participating fast transforms according to the features of the source we want to split, MBR can be used to efficiently represent such composite sources as follows

$$X = \sum_{k=0}^{M1} a_k \alpha_k + \sum_{k=0}^{M2} b_k \beta_k + \ldots + \sum_{k=0}^{M_P} p_k \rho_k \tag{4.13}$$

where the coefficients a, b, ..., p and the selected vectors $\alpha, \beta, \ldots, \rho$ should be determined for each signal realization by the splitting procedure such that the expansion of each component corresponds to the sought statistical description.

For imagelike signals and a two-component source model, the natural choice for the fast transforms is the DCT for the texture component and the FHT for the edge component. The former is the best substitute for the Karhunen-Loeve transform of a correlated first-order Gauss-Markov process which is a good model for the texture component, whereas the latter is the fast transform having the best edge coding properties.

When used in combination with the DCT and FHT transforms, the RRP procedure can be applied with success in separating the two sought components. The computer simulations that were performed, and their experimental results at the end of this chapter, corroborate the usefulness of

this technique. The two separated components can then be coded by using any efficient quantization scheme, with overall advantage over transform coding methods.

4.3.4 Basis set selection

When applying MBR to a representation/coding problem, the selection of the participating fast transforms will have, in general, significant impact on the representation/coding performance. While no general technique has been developed in this thesis, good choices can be made for some particular cases. As we mentioned already, for the case of imagelike signals and using just two transforms, the DCT and the FHT are the best choice. If more basis vectors are to be added, then the Slant transform with its sawtoothlike basis vectors is the next best candidate, due to its ability to code linear brightness variations, very common in real life images.

If the process to be coded is a Gauss-Markov process, then the combination of the DCT and the DST, optimal for ρ close to + 1 and zero respectively, should provide consistently good coding performance, more independent of the statistical changes in the source.

4.3.5 Examples and supporting computer simulations

The following three computer simulation examples were designed to evaluate the performance of the MBR in representation and source coding problems. They are intended to demonstrate the feasibility of MBR in representation and source coding, rather than to propose a specific coding scheme.

4.3.6 Example 4.1

In this experiment, the energy packing characteristics of MBR, the use of the RRP procedure versus the optimal (with respect to representation) full hyperplane search, and the incidence of basis set selection is evaluated. The source is given by a first-order Markov process with $\rho = .9$, with block length $N = 8$. This process is then approximated by several truncated expansions, and the average (over 20 realizations) mean square approximation error is plotted versus the number of coefficients in the expansion (Fig. 19). The observed behavior of the single transforms is as expected, the KLT and DCT being the best performers. With respect to the MBR curves, we can observe that: 1) The superdimensionality involved in MBR allows for an energy packing characteristic that outperforms the KLT one, 2) The RRP procedure performs very close to the optimal full search procedure, and 3) The mix of similar basis vectors (DCT and WHT) performs poorer than the combination of very different ones (DCT and HT).

4.3.7 Example 4.2

In this example the representation of an imagelike signal using the RRP procedure and the DCT and HT basis sets is compared with a DCT representation. The source model, given by Sakrison [4], is the addition of two random processes: 1) A first order Gauss-Markov texture process with $\rho = .85$ and Power = 60, and 2) A zero-mean stationary first-order Markov jump sequence, such that each sample is a continuous random variable, uniformly distributed in the interval [-47,47]. Given that the sample value at time i is z_i, at time i + 1 the value will be $z_{i+1} = z_i$ with probability a = .04, or an independent value with probability (1-a). Fig. 20 shows a normalized realization of the same additive process with N = 64. Figs. 21, 22, and 23 show the result of applying the RRP splitting procedure (DCT + HT, its 8 HT components only, its 2 DCT components only, respectively) keeping 10 coefficients. The representation error is 5.7%. Fig. 24

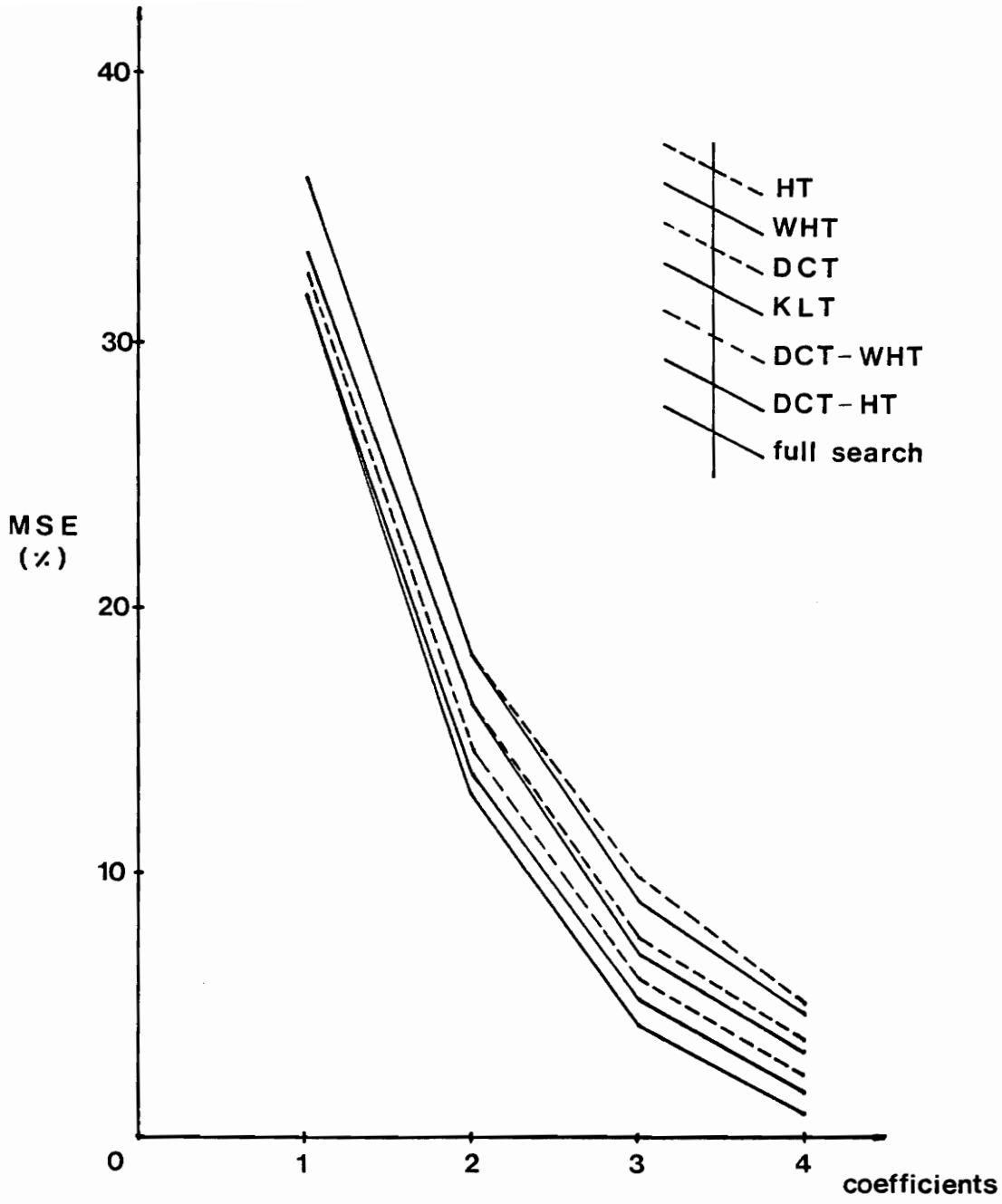

Figure 19. Energy packing MBR performance

shows the corresponding approximation using 10 of the DCT basis functions only, having an error of 13.1% and much poorer subjective quality. Twice as many coefficients (20) are necessary to achieve the MBR error performance when using strictly the DCT (Fig. 25). This experiment shows the improved representation using MBR.

4.3.8 Example 4.3

In this experiment, the two-component imagelike source of the previous example is encoded for data compression using RRP MBR and DCT transform coding (threshold sampling). The following bit allocation scheme was used for both the DCT transform coder and the MBR RRP coder: 1) Coefficient address: a 100 realization training sequence was used and the signals analyzed to determine the sample probability of occurrence of each coefficient. An entropy (variable word-length) coder was then assumed to convey this address information. 2) Coefficient value: the coefficient values were uniformly quantized and the resulting error computed. Again, a 100 realization training sequence was used to determine the probability of occurrence of each quantization interval for each coefficient. According to the resulting probabilities, the intervals are entropy coded assuming a variable word-length encoder. By varying the quantization interval, different rates are obtained. In the RRP MBR case, the quantization intervals for the DCT and HT components are allowed to be different. This permits control over the proportion of the total rate assigned to each component.

Fig. 26 shows the resulting rate-distortion performance for this ad-hoc scheme compared with the theoretical bounds given by Sakrison [4]. Although the rate-distortion performance is relatively far from the bounds, the MBR RRP outperforms the DCT transform coder in an amount close to the gap between the two bounds. Since the same bit allocation strategy was used in both cases, this indicates that the performance improvement is relevant. However, to take full advantage of this two-component representation provided by MBR, the quantization and bit allocation problem should be investigated further. Summarizing, this experiment shows the feasibility of MBR in

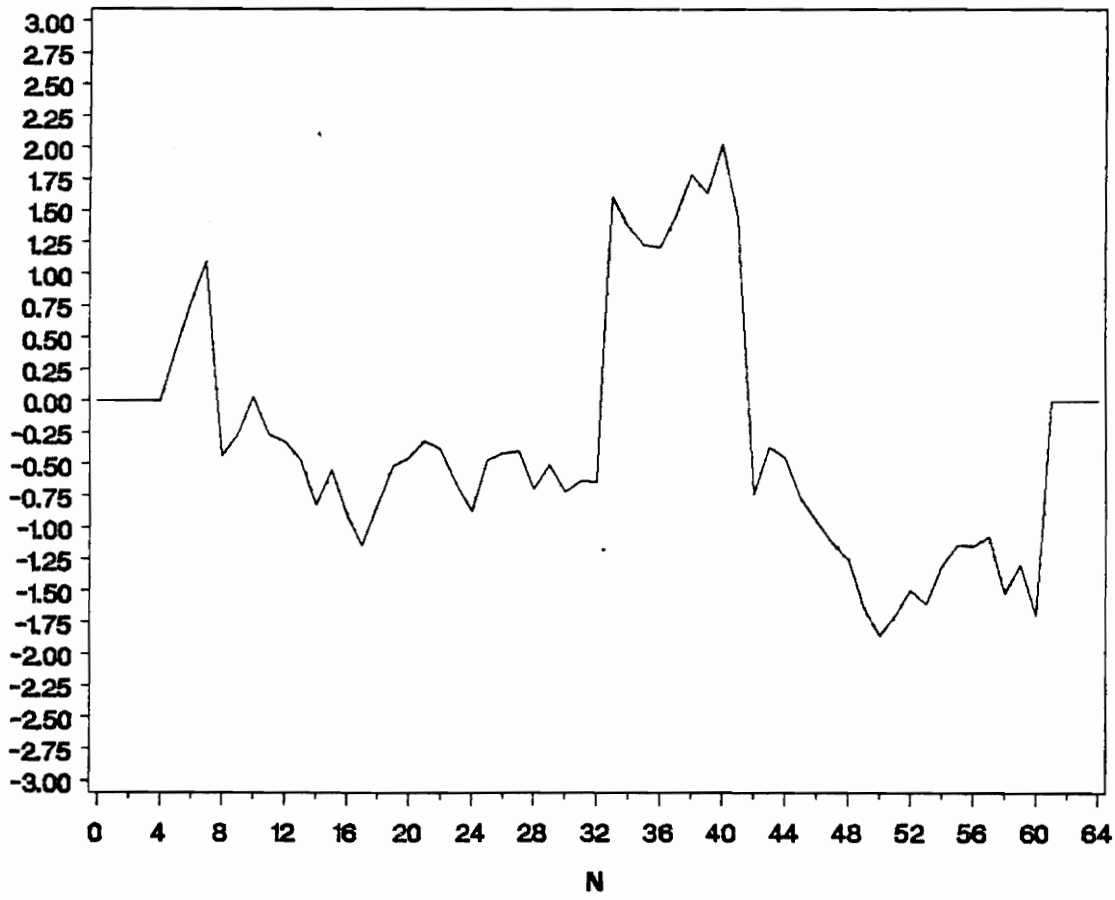

Figure 20. Simulated imagelike signal

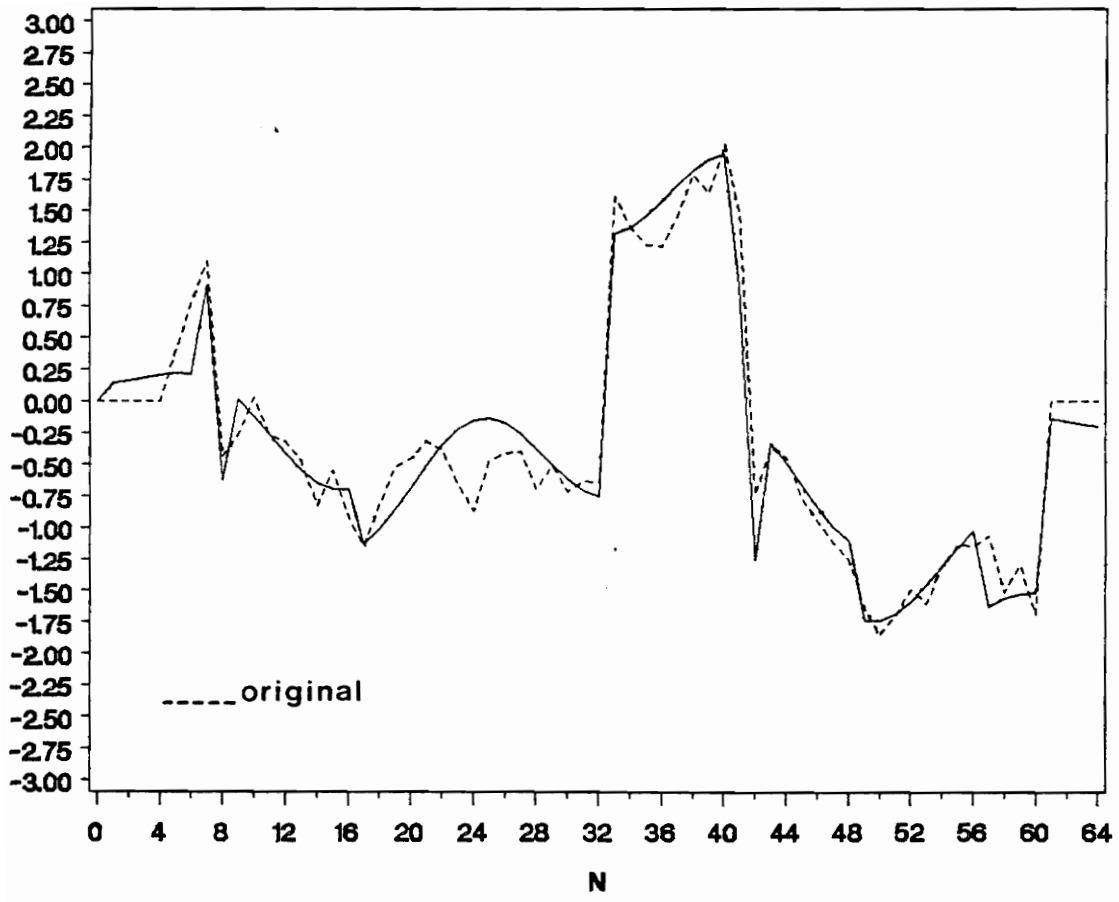

Figure 21. DCT + HT 10 coefficient MBR signal representation: MSE = 5.7%.

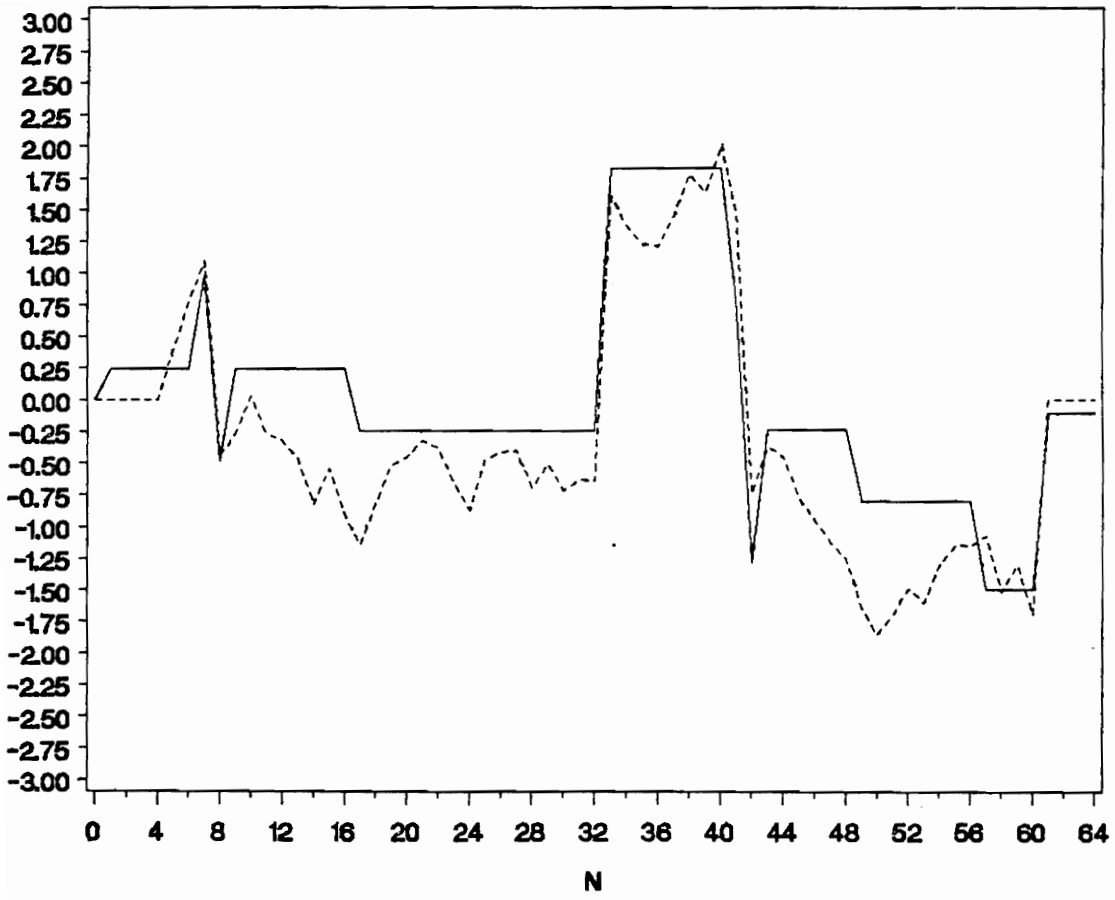

Figure 22. MBR HT component: 8 coefficients.

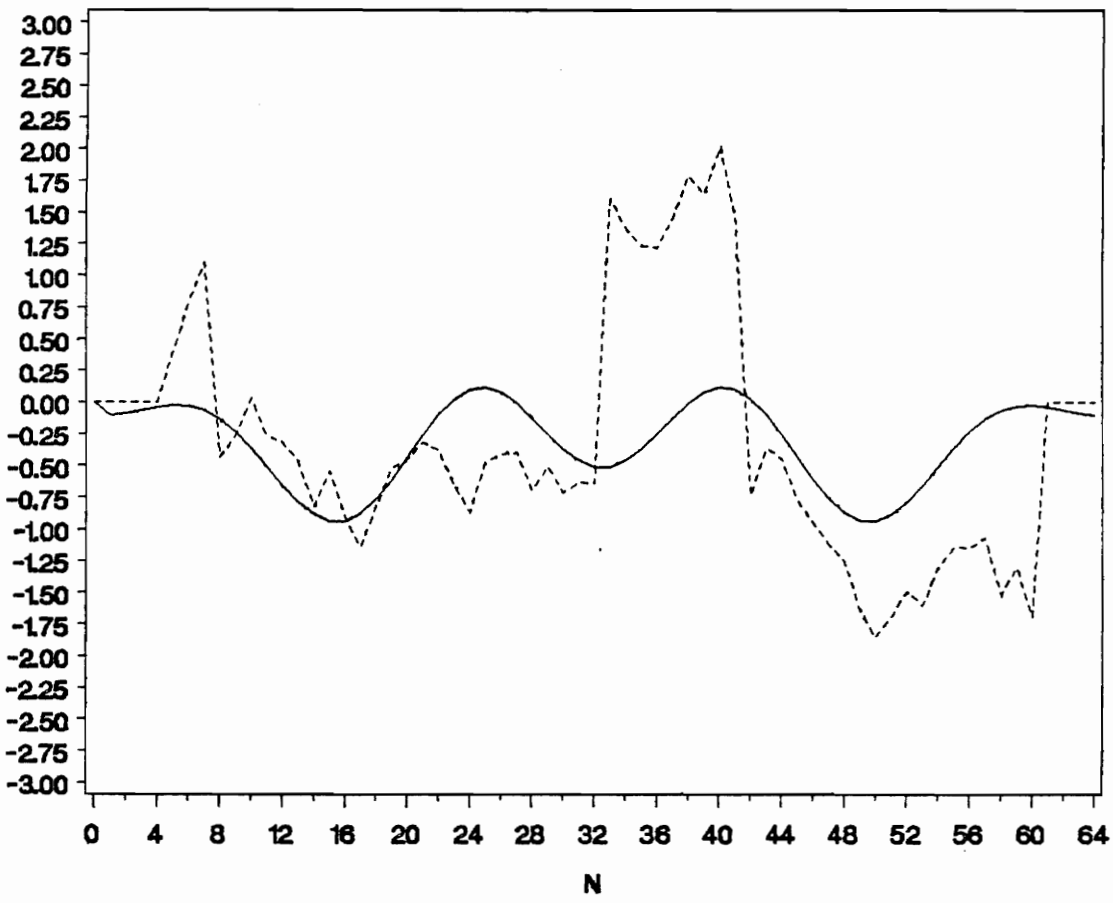

Figure 23. MBR DCT component: 2 coefficients.

Figure 24. DCT only 10 coefficient representation: MSE = 13.1%.

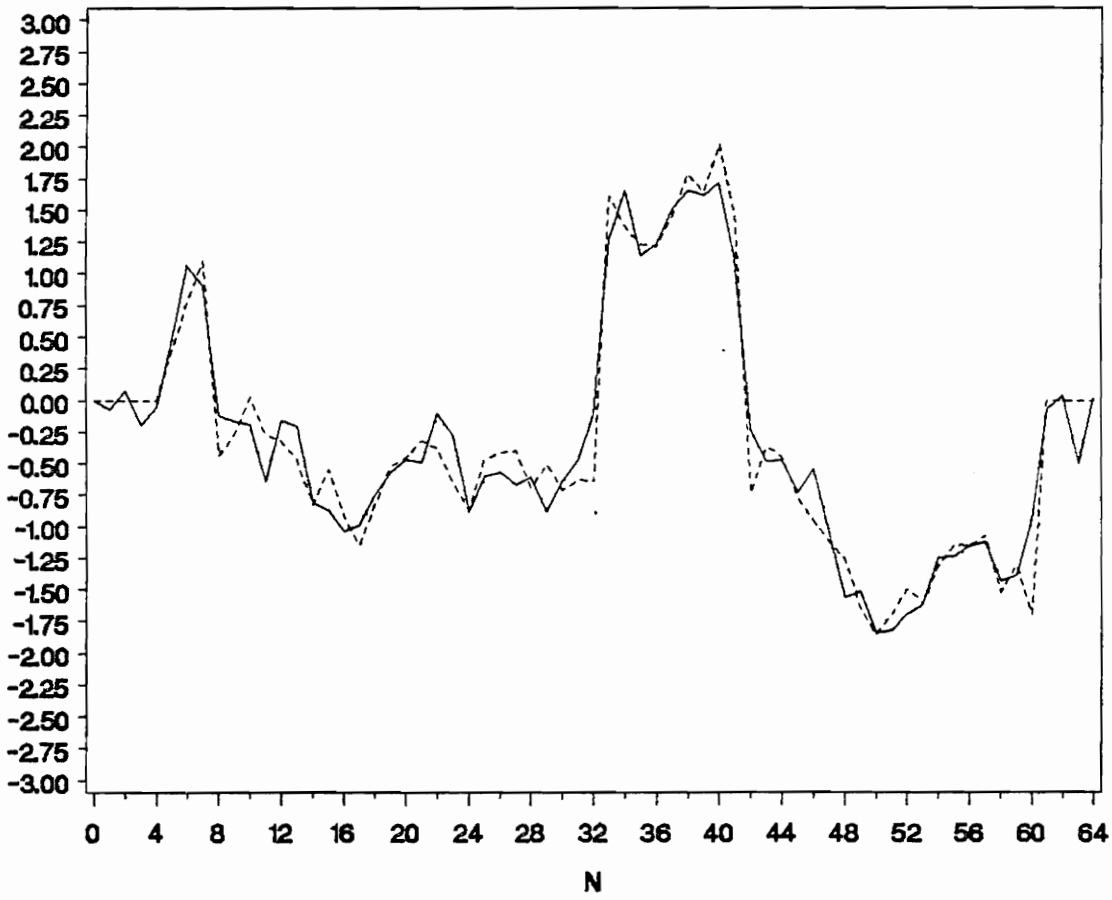

Figure 25. DCT only 20 coefficient representation: MSE = 5.6%.

source coding. Important improvements in error performance and subjective quality over transform coding can be expected in image processing applications.

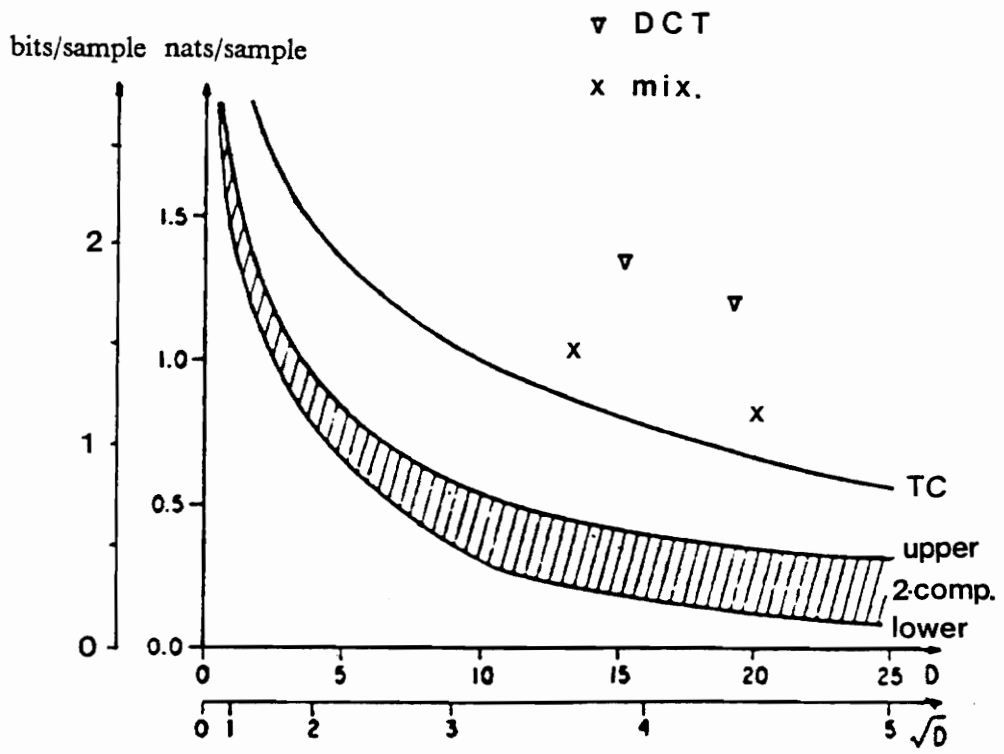

Figure 26. **MBR vs. transform coding rate-distortion performance.** R(D) bounds from Anastassiou and Sakrison [4].

5.0 Application of MBR to signal restoration

5.1 Signal recovery in a Hilbert space setting

The general signal recovery problem is to estimate the original signal from a degraded and/or noise corrupted signal. Some examples are the recovery of the input to a linear shift-invariant system from its output (deconvolution), the extrapolation of a signal from a finite observation segment, the recovery of the input to a nonlinear or shift-varying system from its output, and the restoration of a multidimensional signal from some of its projections (which represents the application in the context of this research work).

In any real world case, we cannot know the degradation operator exactly. Also, as a result of a measurement or a quantization process, we are given data corrupted by random noise. Thus, when comparing restoration methods, it becomes essential to evaluate their robustness, i.e. the sensitivity of the estimation process to those sources of error or that limited knowledge. We will see that knowledge of some of the statistical properties of those errors plays an important role in obtaining robust restoration algorithms, i.e. with better restoration properties, and in determining some confidence measures for the solutions.

The general signal restoration problem admits a formulation in terms of operators in a Hilbert space. Consider the Hilbert space H with elements f,g,h,x,y, etc., a zero vector ϕ, and an inner product (x,y). By definition, $\|f\| = \sqrt{(f,f)} \geq 0$ is the "length" of f, and the sequence $\{f_k\}$ is said to converge to f, denoted $f_k \rightarrow f$, if $\lim_{k\rightarrow\infty} \|f_k - f\| = 0$. This is called weak convergence or convergence in norm. For the case of finite-dimensional vector spaces, convergence in norm actually implies strong convergence [70].

Let T (and C,D,L,F, etc.) be a set-to-point transformation, called an operator, defined on a Hilbert space. An operator maps each element of a set S_1 to a unique element of S_2, where both S_1 and S_2 are subsets of the parent Hilbert space H. The set of all elements for which an operator T is defined is called its domain D(T), whereas the set of all possible elements generated by T is its range, denoted by R(T). The set of elements x for which $Tx = 0$ is called the null-space of T, N(T). Whenever an operator T is a one-to-one mapping, it has an inverse T^{-1} whose domain includes R(T). An element x which, under the application of T, is mapped into itself ($Tx = x$), is called a fixed point of T. Linearity, boundedness, compactness, continuity, rank, and spectrum, are some other important operator properties [70], though we will not concentrate on these in the present context.

Considering the noiseless case first, we can express the degradation D operating on the sought signal x and generating the measured signal y, as an operator equation of the form $y = Dx$, called the measurement operator equation. The distortion operator D is (in principle) known. In case we have a priori information about x available, it can be expressed in terms of the constraint operator equation $x = Cx$, where C is called the constraint operator.

One approach to solving for x is to find the inverse operator D^{-1}. In most cases of practical interest however, it may be difficult or even impossible to calculate it, for example when we have only an approximation to D. In addition, if D is not a one-to-one mapping, the inverse operator does not exist. In such cases, we want to incorporate some additional information about x into the problem solving, in order to remove the ambiguity.

For the reasons mentioned above, alternative methods are considered usually. Many of these try to solve the inverse problem by solving an approximating sequence of direct problems that will

hopefully converge to the unique solution, if it exists. If the solution does not exist or is not unique, convergence is desired to the best possible approximation compatible with all the available information and the selected criterion of distance.

One remarkable technique for generating such a sequence is the method of successive substitutions or fixed point iteration. Given an initial value x_0 and a certain operator F, the iterative procedure is given by $x_{k+1} = Fx_k$. It can be shown that, if the sequence $\{x_k\}$ converges (in norm) to some point x where F is continuous, then it converges to a fixed point of F. Although the fixed point iteration does not always define a convergent sequence, certain conditions on F and on the initial value x_0 can warrant convergence [65].

For some x_i and x_j in the same closed subspace of H, consider the inequality $\|Fx_i - Fx_j\| \leq r.\|x_i - x_j\|$. If $0 \leq r < 1$, the operator is said to be a contraction in that subspace. If $r = 1$ the operator is nonexpansive, and if $r = 1$ and the equation holds with equality only if $x_i = x_j$, then the operator is strictly nonexpansive. If the operator is a contraction in some subspace, then it has a unique fixed point x in that subspace, and the fixed point iteration converges to x (in norm) for every choice of the starting signal x_0 in that subspace, as $k \to \infty$. This result is the well known contraction mapping theorem [65]. A further consequence is that every sequence of iterations will converge at a geometric (linear) rate.

If the operator is only nonexpansive, the situation is not as nice. Unlike contractions, nonexpansive mappings may have any number of fixed points. Strictly nonexpansive mappings, however, have at most one fixed point. If, in addition, their range is compact (closed and bounded), then the results from the contraction mapping theorem apply again [65]. Many operators in signal restoration problems belong to this class.

The importance of the fixed point iteration comes from the fact that, for almost every signal restoration problem, solutions can be expressed in terms of finding fixed points of appropriately designed composite operators. Practical interest in this iterative scheme stems from the great flexibility which is available for mixing signal constraints and distortions. The resulting composite mapping may be a contraction or not, depending upon the specific properties of D and C.

5.2 *Constrained iterative restoration algorithms*

The combination of the measurement and the constraint operator equations, gives a general operator equation of the form

$$Fx = Cx + \Lambda\{y - DCx\} \tag{5.1}$$

The term between brackets is called the residual signal z, and the linear weighting operator Λ modifies the characteristics of the mapping F, and has an impact on the rate of convergence of the fixed point iteration. In their survey paper, Schafer et al. [65] showed that most of the known iterative restoration methods can be formulated as finding the fixed point of the operator F.

If the operator F is a contraction, then it has a unique fixed point and this fixed point is the original signal. For nonexpansive F, the fixed point is not unique, and the solution depends on the initial condition. If x is a fixed point of F however, then either the residual signal is zero or x does not satisfy the constraints. The latter case can be easily identified as a wrong solution. Thus, practically, when the iterations converge the residual signal is zero.

The sensitivity of the solution to noise contamination can easily be seen by including additive noise n in the degradation model, which results in the equation $y = Dx + n$. A solution, making the residual signal zero, is $x' = x + (DC)^{-1}n$. The additional noise term may dominate this solution. Although the conditioning of the problem can be improved by using constraints, the solution is still sensitive to noise. For the case of linear inverse problems, the classical regularization theory of Miller-Tikhonov can be applied successfully [27]. Roughly speaking, this is the theory of continuous approximations to the discontinuous inverse - or generalized inverse - of the linear operator D.

Another possible cure is to define a new convergence criterion and stop the iterations accordingly. Trussell proposed convergence criteria [72] based on the observation of the residual signal z. Convergence is achieved when the residual signal matches a likely realization of the corrupting noise. The method has proven to be successful.

At the same time, Beex proposed an elegant approach to accommodate noisy measurements [7,8]. Recognizing and showing that the application of noisy measurements as if they were absolute knowledge can cause constraint incompatibilities, he defined a "soft" measurement constraint operator. Such an operator forces the measurements only within some tolerance consistent with the expected noise power. Since weighted norms can be accommodated into the tolerances, frequency and/or time (or scene) dependent measurement noise information may readily be incorporated. This has points in common with the idea of Trussell to use the statistics of the residual signal for defining constraint operators in the signal space [73].

5.3 Constraint enforcement: method of convex projections (POCS)

It is a well known fact that reconstruction quality can be significantly improved by incorporating more a priori information about the sought signal x. Each known property about the signal x can be modelled as a set (a subspace) in the Hilbert space H. When the constraints are consistent, the sought signal x satisfies all of them. In other words, the signal x is confined to the subspace or set corresponding to the nonempty intersection of all individual constraint sets involved. Any signal in this intersection set is called a "feasible solution."

The main problem with modelling a priori information as fixed points of a composite operator is that often the operator is not simple. It is in general very difficult to investigate properties such as nonexpansitivity of complex composite operators. This problem can be drastically simplified when all the constraint sets involved are closed and convex [80]. In this case, each set θ uniquely determines an associated operator P_θ , called a projection operator onto a convex set (POCS). It assigns to every point in H, its unique-nearest-neighbor in θ, being then unambiguously defined by means of the minimality criterion $\|x - P_\theta x\| = \min_{y \in \theta} \|x - y\|$. Fig. 27 shows a geometric

interpretation of this definition. The segment connecting the point x and its projection $P_\theta x$ is normal to the tangent hyperplane to θ at the projection $P_\theta x$. Having θ and x on opposite sides, it is a separating hyperplane. In the event that the point x belongs to θ, the operator leaves it unaffected.

It can be easily shown that a projection operator onto a convex set is a nonexpansive operator [80]. Therefore, POCS shares the main properties of orthogonal linear operators projecting onto closed linear manifolds (CLM). In fact, the class of closed convex sets contains all CLM as a subset. Since the nonexpansiveness of POCS is guaranteed, the composite operator $C = P_m P_{m-1} \ldots P_1$ is also nonexpansive, and the fixed point iteration converges to a feasible point [71]. The location of the feasible point however, will depend on the initial condition and the particular ordering of the projection operators [80]. Fig. 28 shows an example of the application of the fixed point iteration for the case of 3 convex sets and the corresponding projections.

Composite operators for projection onto closed convex sets with nonempty intersection constitute an exceptionally well-behaved subclass of nonexpansive transformations known as "reasonable wanderers" [80]. When used in a fixed point iteration context, convergence at geometric rate is guaranteed. This regular behavior has encouraged the use of some good results in fixed point theory, and more research into the development of acceleration techniques for POCS is expected.

Many signal properties can be modeled as closed and convex sets. In their original paper, Youla et al. [80] derived some projection operators corresponding to useful signal constraints, including limited support, amplitude and energy bounds, phase information, subspace projections, positivity, etc..

As stated earlier, POCS converges to a point in the feasible set. In some cases in which the feasible set is small enough, any feasible point may be an acceptable estimate for the restoration problem. In other cases, this set may provide a restricted region over which subsequently some functional may be optimized in order to produce a unique solution. This setting corresponds to the basic goal of mathematical programming. Consequently, many of its powerful tools can be successfully applied to signal restoration problems.

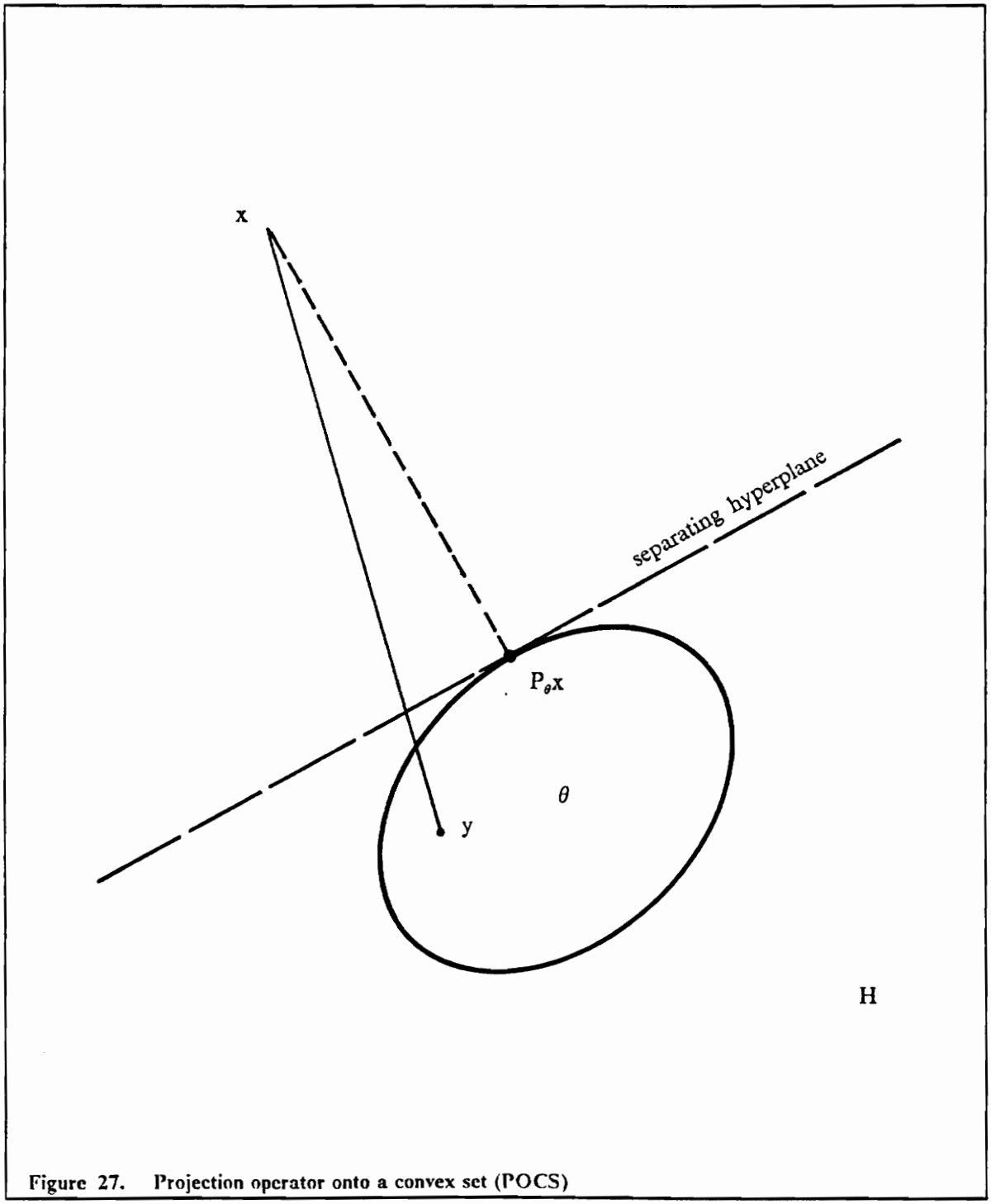

Figure 27. Projection operator onto a convex set (POCS)

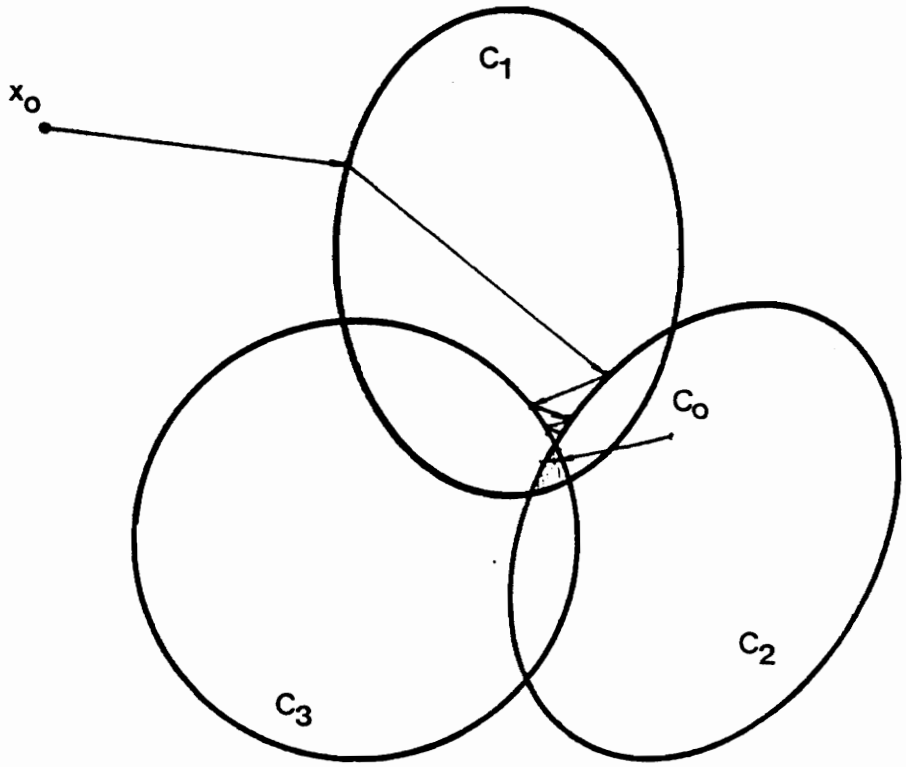

Figure 28. Iterative projection procedure

In recent work, Biemond et al. [11] developed POCS-constrained steepest descend and conjugate gradient algorithms. These methods minimize the well-known Miller-Tikhonov functional [27], which constitutes the basic regularizer for linear inverse problems, like deconvolution. In a somewhat different approach, Mammone et al. [50] derived a constrained optimization procedure based on the unconstrained pseudoinverse solution. After relaxing the minimum energy condition on the solution vector and finding a point on the optimal surface, POCS is used to search for a feasible point along that surface. The resulting solution satisfies both the constraints and the minimum residual norm condition.

In some cases, the set of elements that belong to some criterion-based optimal surface can be expressed as a convex set. It is then possible to perform constrained optimization by the use of convex projections exclusively. Sezan et al. [66] presented a regularized deconvolution procedure in the context of POCS, where an "extended Wiener set" is added to other deterministic constraints, in an attempt to filter out the observation noise. This set corresponds to all signals that satisfy the Wiener solution at all frequencies but those in the neighborhood of the zeroes of the distortion operator, where the dependence of the solution on the observations is discontinuous. When the appropriate deterministic constraints are added, the procedure converges to a unique signal called the constrained Wiener solution.

Unfortunately, not all possible signal properties admit representation as a closed and convex set. A typical case is the set of signals having some specified Fourier magnitude [70]. The method of generalized projections [44], which extends POCS to handle nonconvex sets, has more restrictions and weaker convergence properties, thus producing a limited answer to the problem at hand. Some other promising set-theoretic methods have been introduced recently, extending POCS to the case of imprecisely known information [17](using fuzzy set theory), and point-to-set mappings [14].

5.4 *MBR as a (hard/soft) constraint*

Assume we are given an incomplete set α of M basis vectors that span the linear subspace $U(\alpha)$ of a parent N-dimensional space H. Any specific linear combination of these M vectors determines a point y in $U(\alpha)$. Now, let $B(\alpha,y)$ be the set of all vectors in H whose projections onto $U(\alpha)$ equal y. From the projection theorem [12] $B(\alpha,y) = y + U^{\perp}(\alpha)$, where $U^{\perp}(\alpha)$ is the orthogonal complement space of $U(\alpha)$.

This subspace $B(\alpha,y)$ that results from displacing the closed subspace $U^{\perp}(\alpha)$ a distance y from the origin ϕ is called a linear variety [80]. Linear varieties are closed and convex. As a result, MBR can be used as a constraint in a POCS restoration setting. This involves the derivation of the corresponding projection operator(s).

Depending on the degree of confidence about the location of the point y, the implemented projection operator will be called "hard" (absolute knowledge) or "soft" (uncertainties considered). The soft constraint set can prevent the occurrence of constraint incompatibilities, which could result when noisy measurements are taken and imposed as the true output of the system under evaluation.

In addition, since MBR represents a linear constraint, the resulting operator can be used in a broad spectrum of linear restoration methods that exceed the POCS context. One significant advantage of restricting the restoration to be linear is that acceleration techniques can be applied successfully to speed up convergence [43,53].

5.4.1 Hard constraint operator

When y is considered to be absolute knowledge, the hard projection operator P_B results from the solution of the following minimization problem:

$$\|z - P_B z\| = \min_{x \in H} \|z - x\| = \min_{x \in H} \|d\| \tag{5.2}$$

subject to

$$P_U(x) = y \tag{5.3}$$

where $P_U(x)$ is the projection of the sought signal onto the measurement space U. Since the Euclidean norm is used, this results in a least-squares problem [12]. Expressing the correction term as $d = d_U + d_{U\perp}$, it follows that the distance is minimized by setting $d_{U\perp} = \phi$, (Fig. 29). This results in the projection operator

$$P_B z = y + z_{U^\perp} = z + (y - z_U) = z + c_U \tag{5.4}$$

The main task in the computation of P_B is obtaining the projection z_U from the original signal z. This is equivalent to finding the least-squares solution to an overdetermined linear system of equations [12]. Let A be the NxM matrix formed by the M MBR column basis vectors a_k. Then, the ij-th element of the square MxM matrix A^TA is given by the dot product of the i-th and j-th MBR basis vectors. The projection z_U is then given by, since it is in the column space of A

$$z_U = Aw_{opt} \tag{5.5}$$

where w_{opt} is obtained by solving the MxM system

$$A^T A w_{opt} = A^T z \tag{5.6}$$

Since all the elements of A^TA can be computed in advance, the implementation of this operator has in principle a computational complexity of O(NM) operations. This compares unfavorably with single fast transform implementations.

A substantial reduction in computational costs can be achieved however, by taking advantage of the orthonormal nature of each individual participant basis. Without loss of generality, let us assume that two fast transform sets, say T_1 and T_2, are mixed in the MBR process. The term A^Tz can be computed with significant advantage by applying the fast transforms T_1 and T_2 to the signal z, and then retaining the corresponding M coefficients in the transform domains.

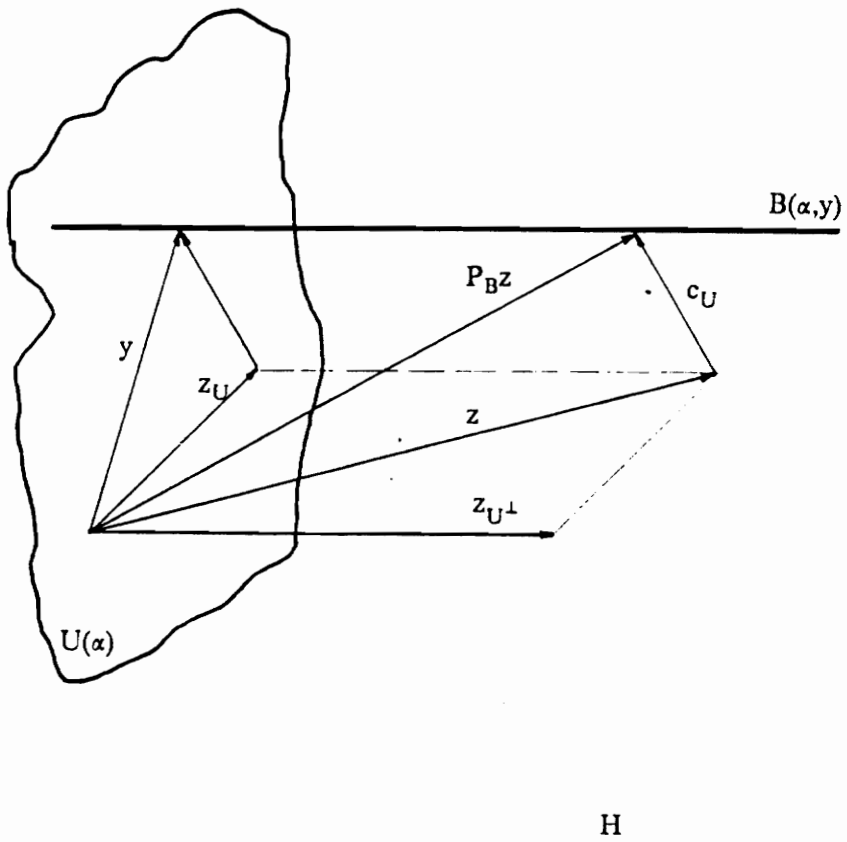

Figure 29. Hard constraint operator

In addition, a closer inspection of the matrix A^TA reveals that the system of equations can be partitioned as follows

$$\begin{bmatrix} I & D \\ D^T & I \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} A_1^T z \\ A_2^T z \end{bmatrix} \tag{5.7}$$

$1 \equiv$ Transform 1 $\quad 2 \equiv$ Transform 2

When the number of T_1 and T_2 components is about the same, this partitioned system can be solved in roughly half the number of operations required for the original system [12].

5.4.2 Soft constraint operator

An important point that arises in this context of signal coding is the consideration of quantization noise. Due to it, the measurement vector y will represent the original MBR only within certain tolerances determined by the noise component. A soft operator will force the projection of the iterate to be within the tolerances around the measurements. In Fig. 30, the uncertainties about the measurement are represented by an ellipsoid centered at the measurement point y. For any two signals whose projections are within the ellipsoid, it can be easily shown that any convex combination [80] of them will also be inside the ellipsoid. Consequently, the soft constraint constitutes a convex set. Note that in the case of quantization noise, the noise component is confined to the subspace U. Without loss of generality, we will consider uniformly distributed noise of constant power e_i to be added independently to each MBR component corresponding to the i-th participating transform.

Suppose again that only two fast transforms are involved in the MBR process. Both the measurement y and the projection z_U admit expansions in terms of T_1 and T_2, $y = y_{T1} + y_{T2}$ and

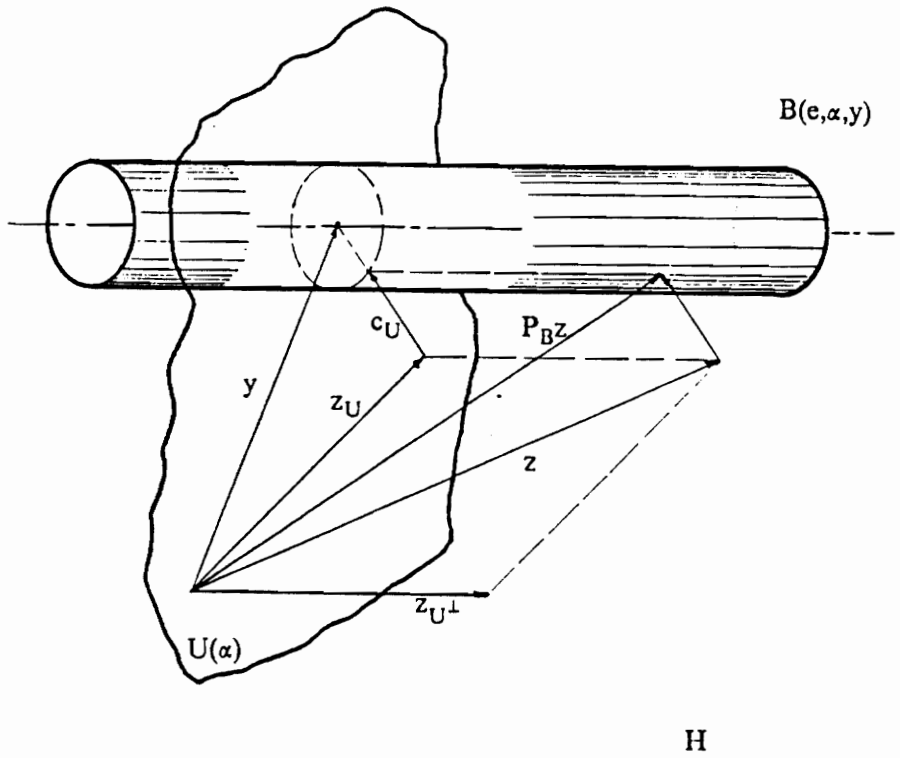

Figure 30. Soft constraint operator

$z_U = z_{UT1} + z_{UT2}$. Those expansions are available from the data y itself, and as a by-product of the projection procedure to obtain z_U . The correction term c_U of the hard constraint is then modified to consider the measurement noise as follows [8]

$$c_U = s_1(y_{T1} - z_{UT1}) + s_2(y_{T2} - z_{UT2}) \tag{5.8}$$

where s_1 and s_2 are given by

$$s_i = \begin{cases} 0 & \text{if } \|y_{Ti} - z_{UTi}\| < e_i \\ 1 - \dfrac{e_i}{\|y_{Ti} - z_{UTi}\|} & \text{if } \|y_{Ti} - z_{UTi}\| \geq e_i \end{cases} \qquad i = 1,2 \tag{5.9}$$

By defining a weighted norm, individual component tolerances can be accommodated easily. In addition, by varying the global noise power tolerance e in the correction term relative to the real measurement noise power, different degrees of softness can be achieved.

5.5 *Examples*

The following experimental examples were devised to show the usefulness of MBR in the context of signal recovery from partial (and noisy) information. A synthetic random source signal is first generated and then coded for data compression by using the RRP procedure for MBR. In order to simulate the whole process of source coding, the MBR coefficients are optionally contaminated with synthetic quantization noise.

The information provided by the MBR source coder, as well as deterministic properties the original signal is known to possess, is then iteratively enforced by using the method of convex projections (POCS). In example 5.1, under noiseless conditions, the hard MBR projection operator is used to show convergence to a feasible point that satisfies all the a priori information, resulting

in a better quality estimate of the original signal. This feasible solution shows the advantage of including the extra a priori information in the decoding process [55], over just using the provided MBR information. In many cases, some a priori information about the signal can be assumed at the decoder end without requiring any transmission. This can be included in the restoration process for an improved final result.

Example 5.2 evaluates the effect of quantization noise on the restoration quality. Both hard and soft constraint MBR operators are applied, and their average relative performance is compared for a fixed signal over several realizations of the contaminating noise. As expected, the soft constraint operator gives better performance than the hard operator, particularly when a good estimate of the contaminating noise power is available.

5.5.1 Example 5.1: hard MBR constraint and noiseless measurements

The source signal is a 64-point realization of a zero mean first order Gauss-Markov process with $\rho = .4$, truncated to have a centered support of 75% of the total interval, and normalized to unit power. Fig. 31 shows the resulting source signal.

The signal is then approximated by using a truncated multiple basis representation of 10 coefficients. In this example, the mix of DCT and FHT basis vectors was implemented by using the standard RRP procedure. Fig. 32 shows the resulting truncated representation, having a quadratic approximation error of 24 %.

In addition to the MBR information, other a priori information is imposed in terms of limited support, zero crossing locations, amplitude bounds, and energy bounds, through implementation of the corresponding projection operators. Fig. 33 shows the reconstructed signal after 10 iterations of the POCS procedure. The improvement in quality is evident, and the approximation error is reduced to 12%.

The absence of distortion operators with memory or continuous nonlinearities allows the iteration to converge fast. It was found that almost no signal change is noticed when going beyond

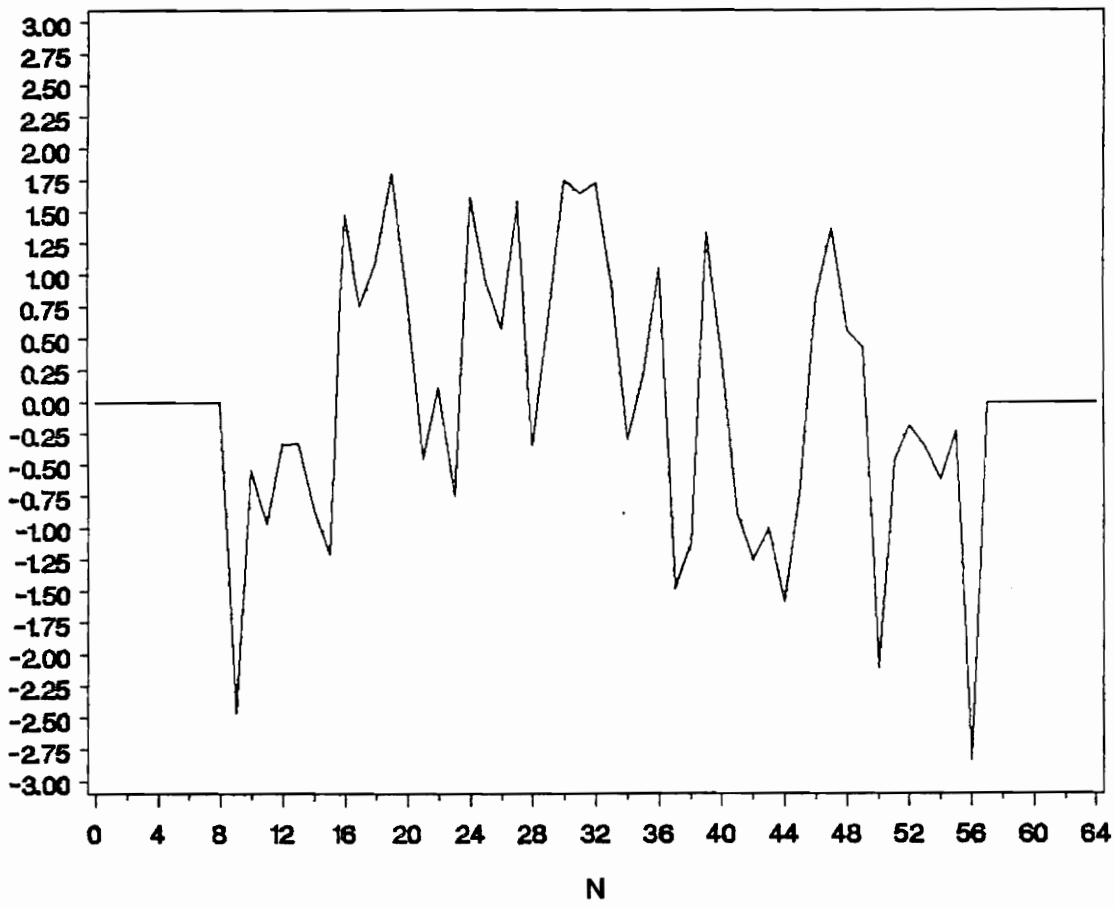

Figure 31. Gauss-Markov source signal

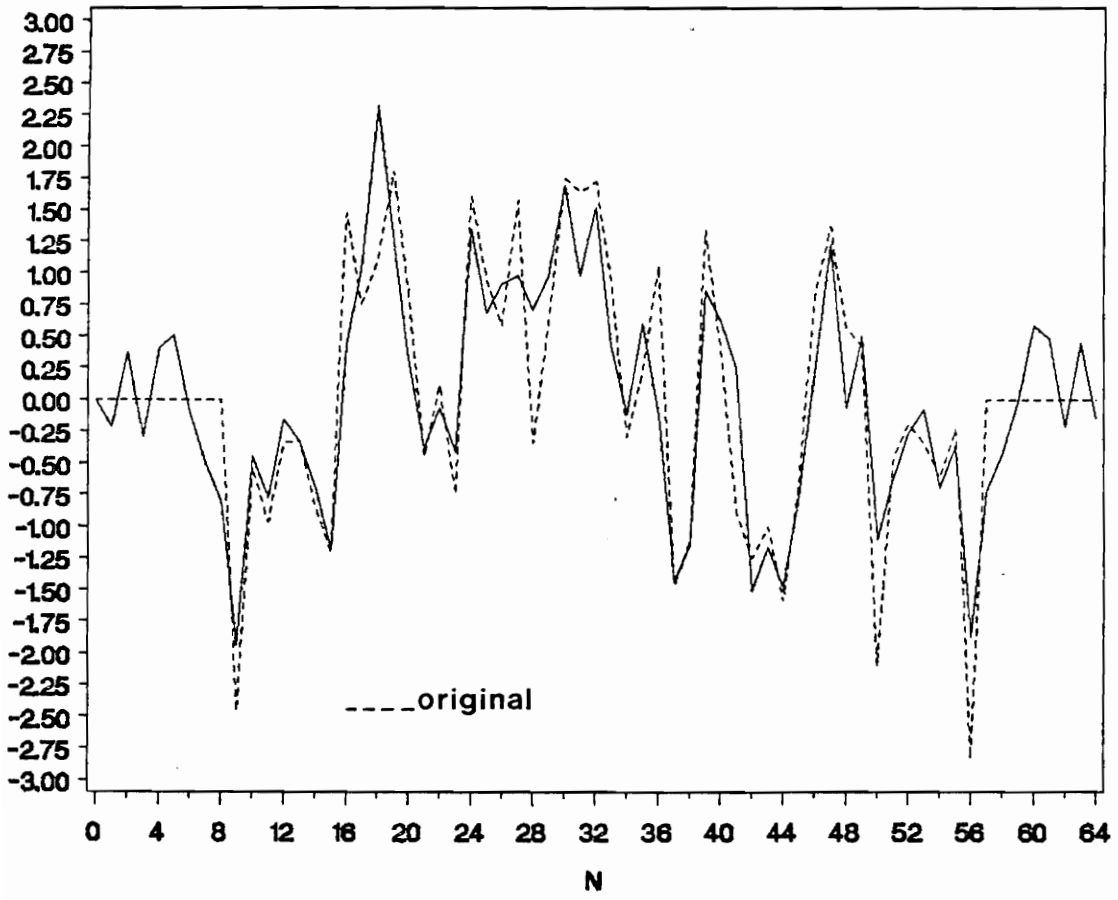

Figure 32. RRP truncated representation: M = 10, MSE = 24%.

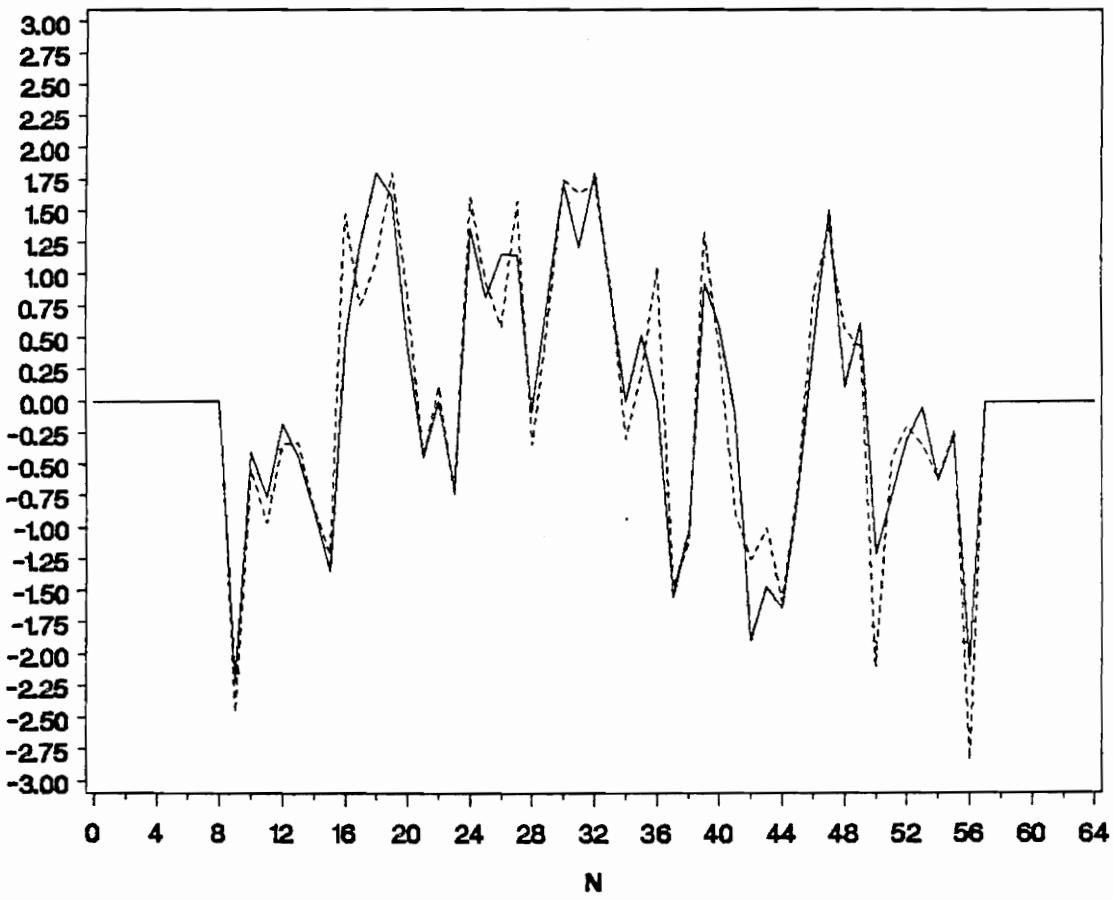

Figure 33. Reconstructed signal: MSE = 12%.

5 iterations. This is a real advantage when considering the computational cost of iterative decoding schemes for real time applications.

5.5.2 Example 5.2: soft vs. hard MBR constraint in a noisy case

The same source signal of example 5.1 is generated and then approximated by the same truncated multiple basis representation. The resulting MBR coefficients are then contaminated by a uniformly distributed noise of power e, independently applied to each coefficient, simulating quantization noise. The resulting signal-to-noise ratio is 6.88 decibels.

The resulting measurement is then combined with the same a priori information used in example 5.1, and POCS restoration is then applied for both hard and soft MBR constraint operators. This procedure is repeated for 100 different realizations of the quantization noise and the reconstruction errors are averaged.

Fig. 34 and Table 1 summarize the results of this experiment as the softness of the operator varies. It can be seen that, for this high noise case, the soft constraint reconstruction outperforms the hard constraint reconstruction. In addition, as the projection operator hardens, the iteration is more frequently nonconvergent, because of constraint incompatibilities. In addition, if the operator is softened too much, the error starts increasing again. This is due to an excessive tolerance, which decreases the enforcing properties of the operator. Non-convergence due to incompatibility of constraints typically results in oscillatory behavior of the iterates. Upon detecting this phenomenon one can increase measurement operator softness until convergence is reached.

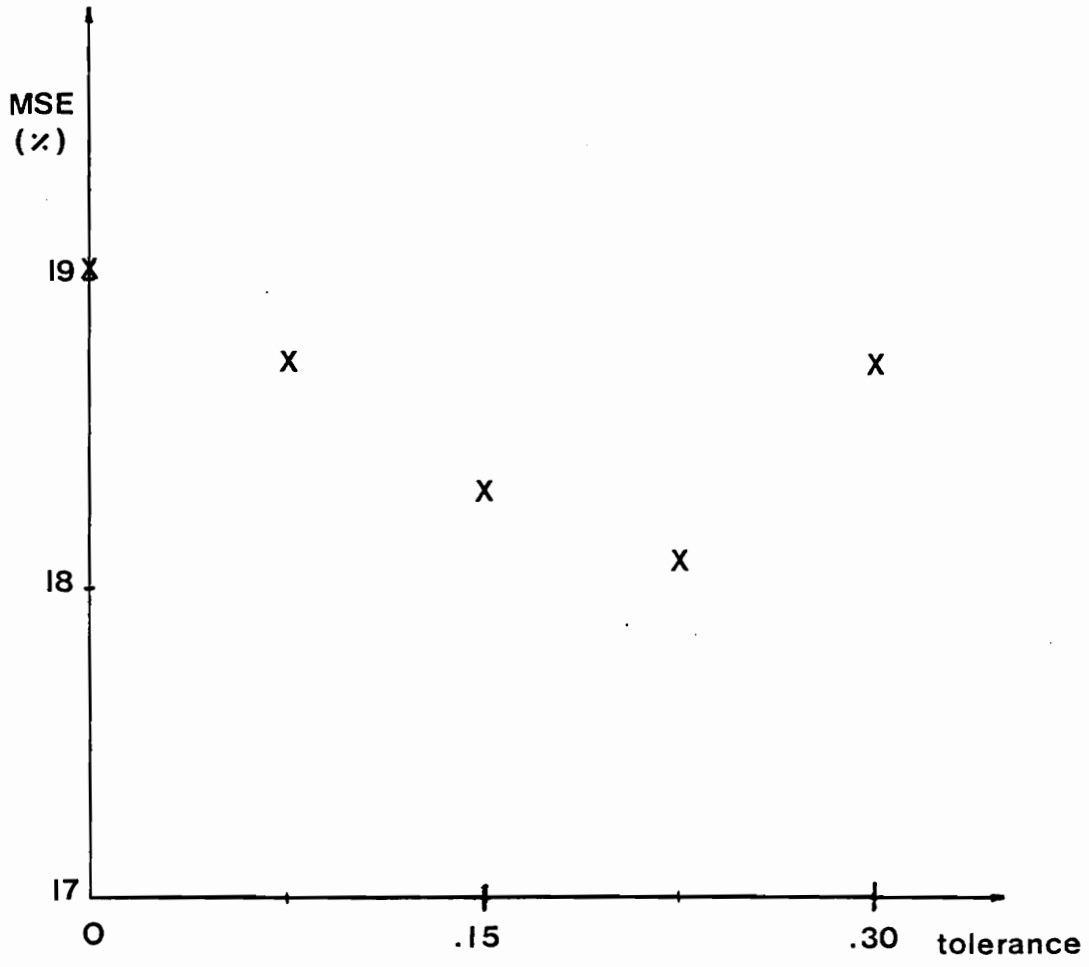

Figure 34. Operator performance vs. softness

Table 1. Operator performance vs. softness

TOLERANCE	NON CONVERGENCE (%)	ERROR (%)
0	22	19.0
.075	11	18.7
.150	4	18.3
.225	2	18.1
.300	0	18.7

6.0 Conclusion

In this thesis, we have proposed a signal representation scheme based on the combination of multiple fast orthogonal transform bases. This representation leads naturally to the idea of signal splitting, i.e. the separation of the signal into several components, in such a way that each one is efficiently represented by the corresponding transform involved. In particular, we proposed the RRP splitting procedure, which works especially efficiently in separating the edge (Haar transform basis) and texture (Discrete Cosine basis) components of imagelike signals. The resulting representation requires a reduced total number of coefficients when compared to single basis representation.

When applied to data compression, MBR can be viewed as a combination of transform and codebook coding methods. Depending on the quantization method used, the RRP procedure results in a multistage vector quantizer or just a signal splitter. The latter was applied to imagelike signals, and the separated components were quantized using scalar quantization. The resulting performance improvement over a DCT transform coder that uses the same bit allocation strategy is approximately the same as the difference between the corresponding theoretic rate-distortion bounds. This indicates the excellent potential of MBR for efficiently coding non-Gaussian additive sources, at some extra computational cost.

The application of MBR to signal restoration was also investigated. Considered as a constraint, the MBR can be expressed as a convex set. As a result, it can be successfully included in most iterative restoration algorithms. In this work, we derived the projection operator for MBR. As an extension to the noisy case, we proposed a soft constraint projection operator, which enforces the MBR coefficients only within a tolerance compatible with the measurement noise. Computer simulation results confirm the usefulness of these projection operators in iterative reconstruction.

One important problem with this representation is the limited number of fast transforms available. This restricts the number of features, i.e. the chance of matching the statistical properties of the source. In addition, when used as a vector quantizer, the number of resulting codebook words is in general too small for an acceptable distortion. Therefore, a suboptimal multistage scheme must be used.

Future research should be concentrated in three areas: 1) generating criteria for selecting and/or designing the best fast transforms, in accordance with the statistics of the source, if available explicitly, or a training sequence of data, 2) designing and analyzing alternative signal splitting procedures, and 3) developing quantization and bit allocation schemes for better rate-distortion performance. Another interesting research point is the application of RRP to progressive transmission of images [74] in which a rough version of the image is transmitted first, and quantized versions of the residuals are transmitted afterwards, as the user wishes.

We conclude that the MBR exhibits excellent potential in signal representation, coding, and reconstruction problems. With more investigation, the MBR should prove to be a comprehensive and powerful tool for use in a variety of signal processing problems.

References

1. Ahmed, N., Natarajan, T., and Rao, K.R.: "Discrete cosine transform," *IEEE Trans. Comput.*, vol. C-23, pp. 90-93, 1974.

2. Ahmed, N., and Flickner, M.D.: "Some considerations of the discrete cosine transform," *IEEE 16th Asilomar conference on circuits, systems, and computers*, Pacific Grove, CA., pp. 295-299, 1982.

3. Algazi, V.R.: "On the optimality of the Karhunen-Loeve expansion," *IEEE Trans. Inf. Theory*, pp. 319-321, March 1969.

4. Anastassiou, D., and Sakrison, D.J.: "New bounds to R(D) for additive sources and applications to image encoding," *IEEE Trans. Inform. Theory*, vol. IT-25, no. 2, pp. 145-155, March 1979.

5. Andrews, H.C.: "Multidimensional rotations in feature selection," *IEEE Trans. Comput.*, vol. C-20, pp. 1045-1051, Sept. 1981.

6. Arazi, B.: "Two-dimensional digital processing of one-dimensional signal," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-22, pp. 81-86, 1974.

7. Beex, A.A.: "Iterative reconstruction of space-limited scenes from noisy frequency-domain measurements," *IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'83)*, pp.147-150. Boston, MA, April 14-16, 1983.

8. Beex, A.A.: "Soft constraint iterative reconstruction from noisy projections," *IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP'84)*, pp. 12A.6.1-4, San Diego, CA, March 19-21, 1984.

9. Berger, T.: *Rate distortion theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.

10. Beauchamp, K.G.: *Walsh functions and their applications*, London: Academic Press, Inc., 1975.

11. Biemond, J., and Lagendijk, R.L.: "Regularized iterative image restoration in a weighted Hilbert space," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'86)*, pp. 1485-1488, Tokyo, Japan, 1986.

12. Brogan, W.L.: *Modern control theory*. Englewood Cliffs, NJ: Prentice Hall, 1985.

13. Buzo, A., Gray, A., Gray R.M., and Markel, J.D.: "Speech coding based upon vector quantization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 5, pp. 562-574, Oct. 1980.

14. Cadzow, J.A.: "Signal enhancement-a composite property mapping algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 1, pp. 49-62, Jan. 1988.

15. Chen, W., Smith, H., and Fralick, H.S.: "A fast computational algorithm for the discrete cosine transform," *IEEE Trans. Commun.*, vol. COM-25, no. 9, pp. 1004-1009, Sept. 1977.

16. Civanlar, M.R., and Santago, P.: "An improved transform coder for image sequences using attributes of difference pictures," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'85)*, pp. 343-346, Tampa, Florida, 1985.

17. Civanlar, M.R., and Trussell, H.J.: "Digital signal restoration using fuzzy sets," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 4, pp. 919-936, Aug. 1986.

18. Davisson, L.D.: "Rate-distortion theory and application," *Proceedings of the IEEE*, vol. 60, no. 7, pp. 800-808, July 1972.

19. Dixit, V.V.: "Edge extraction through Haar transform," *IEEE 14th Asilomar Conference on Circuits Systems, and Computers*, Pacific Grove, CA, pp. 141-143, 1980.

20. Enomoto, H., and Shibata, K.: "Orthogonal transform coding system for television signals," *IEEE Trans. Electromag. Compat.* vol. EMC-13, pp. 11-17, 1971.

21. Flickner, M.D., and Ahmed, N.: "A derivation for the discrete cosine transform," *Proc. of the IEEE*, vol. 70, pp. 1132-1134, 1982.

22. Gersho, A.: "Asymptotically optimal block quantization," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 4, pp. 373-386, July 1979.

23. Gersho, A.: "On the structure of vector quantizers," *IEEE Trans. Inf. Theory*, vol. IT-28, no. 2, pp. 157-166, March 1982.

24. Gersho, A., and Ramamurthi, B.: "Image coding using vector quantization," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'82)*, pp. 428-431, Paris, France, 1982.

25. Gray, R.M., and Linde, J.: "Vector quantizers and predictive quantizers for Gauss-Markov sources," *IEEE Trans. Commun.*, vol. COM-30, no. 2, pp. 381-389, Feb. 1982.

26. Gray, R.M.: "Vector quantization," *IEEE ASSP magazine*, pp. 4-28, April 1984.

27. Groetsch, C.W.: *The theory of Tikhonov regularization for Fredholm equations of the first kind*, London: Pitman Publishing Inc., 1984.

28. Habibi, A., and Wintz, P.: "Image coding by linear transformation and block quantization," *IEEE Trans. on Commun. Tech.*, vol. COM-19, no. 1, pp. 50-62, Feb. 1971.

29. Hamidi, M., and Pearl, J.: "Comparison of the Cosine and Fourier transforms of Markov-1 signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-24, pp. 428-429, 1976.

30. Harmuth, H.F.: *Transmission of information by orthogonal functions* , Berlin: Springer-Verlag, 1972.

31. Huang, J.J.Y., and Schultheiss, P.M.: "Block quantization of correlated Gaussian random variables," *IEEE Trans. on Commun. Systems*, pp. 289-296, Sept. 1963.

32. Jain, A.K.: "A fast Karhunen-Loeve transform for a class of random processes," *IEEE Trans. Commun.*, vol. COM-24, pp. 1023-1029, 1976.

33. Jain, A.K.: "A sinusoidal family of unitary transforms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, pp. 356-365, 1979.

34. Jain, A.K.: "Image data compression: a review," *Proceedings of the IEEE*, vol. 69, no. 3, pp. 349-389, March 1981.

35. Jain, A.K.: "Advances in mathematical models for image processing," *Proc. of the IEEE*, vol. 69, no. 5, pp. 502-528, May 1981.

36. Jayant, N.S., and Noll, P.: *Digital coding of waveforms*, Englewood Cliffs, NJ.: Prentice-Hall, 1984

37. Jones, H.W., Hein, D.N., and Knauer, S.C.: "The Karhunen-Loeve, discrete cosine, and related transforms obtained via the Hadamard transform," *Int. Found. Telemetering Conf. Proc.*, vol. 14, pp. 87-98, 1978.

38. Juang, B.: "Multiple stage vector quantization for speech coding," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'82)*, pp. 597-600, Paris, France, 1982.

39. Kitajima, H.: "Energy packing efficiency of the Hadamard transform," *IEEE Trans. Commun.*, vol. COM-24, pp. 1256-1258, 1976.

40. Kitajima, H., and Shimono, T.: "Some aspects of the fast Karhunen-Loeve transform," *IEEE Trans. Commun.*, vol. COM-28, pp. 1773-1776, 1980.

41. Kunt, M.: "On computation of the Hadamard transform and the R transform in ordered form," *IEEE Trans. Comput.*, vol. C-24, pp. 1120-1121, 1975.

42. Kunt, M., Ikonomopoulos, A., and Kocher, M.: "Second-generation image-coding techniques," *Proceedings of the IEEE*, vol. 73, no. 4, pp. 549-574, April 1985.

43. Lagendijk, R.L., Mersereau, R.M., and Biemond, J.: "On increasing the convergence of regularized iterative image restoration algorithms," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'87)* , pp. 1183-1186, Dallas, Texas, 1987.

44. Levi, A., and Stark, H.: "Image restoration by the method of generalized projections with application to restoration from magnitude," *J. Opt. Soc. Am.*, vol. 1, no. 2, pp. 932-943, 1984.

45. Linch, R., and Reis, J.J.: "Haar transform image coding," *IEEE 1976 National Telecommunications Conference*, Dallas, TX., pp. 44.3.1-44.3.5, 1976.

46. Linde, Y., Buzo, A., and Gray, R.M.: "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, no. 1, pp. 84-95, Jan. 1980.

47. Lloyd, S.P.: "Least square quantization in PCM," *Bell Telephone Lab. Memo.*, Murray Hill, N.J. (unpublished).

48. Makhoul, J.: "A fast cosine transform in one and two dimensions," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, no. 1, pp. 27-34, Feb. 1980.

49. Makhoul, J., Roucos, S., and Gish, H.: "Vector quantization in speech coding," *Proceedings of the IEEE*, vol. 73, no. 11, pp. 1551-1588, Nov. 1985.

50. Mammone, R.J., and Rothacker, R.J.: "General iterative method of restoring linearly degraded images," *J. Opt. Soc. of America*, vol. 4, no. 1, pp. 208-215, Jan. 1987.

51. Mannos, J.L., and Sakrison, D.J.: "The effects of a visual fidelity criterion on the encoding of images," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 525-536, July 1974.

52. Max, J.: "Quantizing for minimum distortion," *IRE Trans. on Inf. Theory*, vol. IT-6, pp. 7-12, March 1960.

53. Mitin, A.V.: "Linear extrapolation in an iterative method for solving systems of equations," *U.S.S.R. Comput. Maths. Math. Phys.* vol. 25, no. 2, pp. 1-6, 1985.

54. Mitrakos, D.K., and Constantinides, A.G.: "Nonlinear image processing for optimum composite source coding," *IEE Proceedings*, vol. 130, Pt. F, no. 5, pp. 441-451, August 1983.

55. Nobakht, R.A., and Rajala, S.A.: "An image coding technique using a human visual system model and image analysis criteria," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'87)*, pp. 1358-1361, Dallas, Texas, 1987.

56. Pearl, J.: "Basis-restricted transformations and performance measures for spectral representation," *IEEE Trans. Information Theory*, pp. 751-752, Nov. 1971.

57. Pearl, J., Andrews, H.C., and Pratt, W.K.: "Performance measures for transform data coding," *IEEE Trans. Commun.*, pp. 411-415, June 1972.

58. Pratt, W.K., Kane, J., and Andrews, H.C.: "Hadamard transform image coding," *Proceedings of the IEEE*, vol. 57, no. 1, pp. 58-68, Jan. 1969.

59. Pratt, W.K., Chen, W., and Welch, L.R.: "Slant transform image coding," *IEEE Trans. Commun.*, vol. COM-22, pp. 1075-1093, 1974.

60. Pratt, W.K.: *Digital image processing*, New York: John Wiley & Sons, 1978.

61. Rosenfeld, A., and Thurston, M.: "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol C-20, pp. 562-569, 1971.

62. Rosenfeld, A., and Kak, A.C.: *Digital picture processing, vol. 1* , New York: Academic Press, Inc., 1982.

63. Sabin, M.J., and Gray, R.M.: "Product code vector quantizers for waveform and voice coding," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 3, pp. 474-488, June 1984.

64. Sanz, J.L.C., and Huang, T.S.: "Unified Hilbert space approach to iterative least-squares linear signal restoration," *J. Opt. Soc. of America*, vol. 73, no. 11, pp. 1455-1465, Nov. 1983.

65. Schafer, R.W., Mersereau, R.M., and Richards, M.A.: "Constrained iterative restoration algorithms," *Proceedings of the IEEE*, vol. 69, no. 4, pp. 432-450, April 1981.

66. Sezan, M.I., Tekalp, A.M., and Chen, C.: "Regularized signal restoration using the theory of convex projections," *IEEE Intl. Conf. on Acoust., Speech, Signal Processing (ICASSP'87)*, pp. 1565-1568, Dallas, Texas, 1987.

67. Shannon, C.E.: "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379-423, 1948.

68. Shannon, C.E.: "Coding theorems for a discrete source with a fidelity criterion," *IRE Nat'l. Conv. Rec., part 4*, pp. 142-163, 1959.

69. Srinivasan, R., and Rao, K.R.: "An approximation to the discrete cosine transform for N = 16," *Signal Processing*, vol. 5, pp. 81-85, 1983.

70. Stark, H., editor: *Image recovery: theory and application*, Orlando, FL.: Academic Press, Inc., 1987.

71. Tom, V.T., Quatieri, T.F., Hayes, M.H., and McClellan, J.: "Convergence of nonexpansive signal reconstruction algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 5, pp. 1052-1058, Oct. 1981.

72. Trussell, H.J.: "Convergence criteria for iterative restoration methods," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, no. 1, pp. 129-136, Feb. 1983.

73. Trussell, H.J., and Civanlar, M.R.: "The feasible solution in signal restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 2, pp. 201-212, April 1984.

74. Wang, L., and Goldberg, M.: "Progressive image transmission by transform coefficient residual error quantization," *IEEE Transactions on Communications*, vol. COM-36, no. 1, pp. 75-87, January 1988.

75. Wang, Z.: "New algorithm for the Slant transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, pp. 551-555, 1982.

76. Yamada, Y., Tazaki, S., and Gray, R.M.: "Asymptotic performance of block quantizers with difference distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 1, pp. 6-14, January 1980.

77. Yan, J. K.,and Sakrison, D.J.: "Encoding of images based on a two-component source model," *IEEE Trans. Commun.*,vol. COM-25, no. 11, pp. 1315-1322, Nov. 1977.

78. Yip, P., and Rao, K.R.: "Energy packing efficiency for the generalized discrete transform," *IEEE Trans. Commun.*, vol. COM-26, pp. 1257-1262, 1978.

79. Yip, P., and Rao, K.R.: "A fast computational algorithm for the discrete sine transform," *IEEE Trans. Commun.*, vol. COM-28, pp. 304-307, 1980.

80. Youla, D.C., and Webb, H.: "Image restoration by the method of convex projections: Part 1-Theory," *IEEE Trans. Med. Imaging*, vol. MI-1, no. 2, pp. 81-94, Oct. 1982.

Vita

Felix G. Safar was born in La Plata, Argentina on July 26th. 1957. He studied at *Universidad Nacional de La Plata (UNLP)* where, on March 27th. 1980, he obtained the degree of *Ingeniero en Telecomunicaciones*. Upon graduation, he worked simultaneously at *UNLP* as a teaching assistant and at *CeTAD-Facultad de Ingeniería-UNLP* as a research assistant for *Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET)*. While at *CeTAD,* he was involved in research on signal processing, and the development of complex digital systems for real-time digital signal processing and/or industrial applications.

In September 1986 Mr. Safar enrolled in the Master of Science in Electrical Engineering program at Virginia Polytechnic Institute and State University. After graduation he will return to *UNLP* to continue his teaching and research career working on signal processing, communications, and computer architecture for real-time signal processing.