Approximate Signal Reconstruction from Partial Information

by

Phillip J. Moose

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

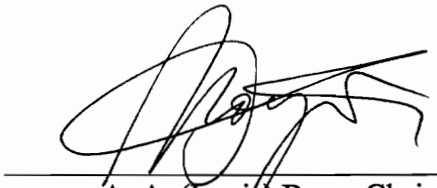
in partial fulfillment of the requirements for the degree of

Master of Science

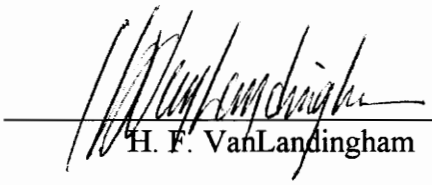
in

Electrical Engineering

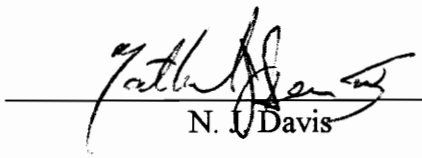
APPROVED:



A. A. (Louis) Beex, Chairman



H. F. VanLandingham



N. J. Davis

May, 1994

Blacksburg, Virginia

Approximate Signal Reconstruction from Partial Information

by

Phillip J. Moose

A. A. (Louis) Beex, Chairman

Electrical Engineering

(ABSTRACT)

It is known that transform techniques do not represent an optimal way in which to code a signal in terms of theoretical rate distortion bounds. A signal may be coded more efficiently if side information is included with the signal during transmission. This side information can then be used to reconstruct the image at some later time.

In this thesis, the type of transform coding used is Multiple Bases Representation (MBR). This coding scheme is known to perform better than transform coding that uses a single basis. The method of Projection Onto Convex Sets (POCS) is used to reconstruct an approximation to the MBR signal using the side information. Thus, any number of constraints may be used as long as they form closed and convex sets and the side information is *a priori* knowledge required to implement projections onto the defined closed and convex sets.

Several closed and convex sets are examined including the MBR, positivity, sign, zero crossing, minimum increase, and minimum decrease constraints. Constraints that tend to limit energy are not as effective as constraints that introduce energy into the signal especially when the observed image is used as the initialization vector.

When a different initialization vector is used, the POCS reconstruction performs considerably better. Two initialization vectors are proposed; the observed signal plus white noise and the observed signal plus a constant. The performance of POCS with initialization by the observed signal plus a constant is superior to that when using the observed signal only.

One nonconvex constraint is considered. The Laplacian histogram constraint requires other convex constraints to help ensure convergence of the reconstruction algorithm, but produces good quality images.

Acknowledgements

I would like to thank my parents Paul and Sally Moose. In many ways I am indebted to them because they are my starting point. I would also like to thank John, my twin brother, for his patience and support. Possibly, I can make up to John the time I spent attaining this goal. Also, I would like to thank Uncle Bob and Aunt Louise Shoener who funded my undergraduate education; they have given a great gift for which I am indebted.

Thanks goes to Dr. A. A. (Louis) Beex, a professor who has earned my respect because of his high expectations from himself as well as his students. I hope that in some way I have lived up to them. Also, I would like to thank my other committee members, Dr. H. F. VanLandingham and Dr. N. J. Davis.

I thank Jeff who has provided me companionship and much support while I worked towards my degree. His patience with my impatience has been greatly appreciated. Although we will go our separate ways soon, I know we will always stay in touch and be best of friends.

Finally, I would like to dedicate this work to Ayn Rand and the group Pink Floyd, without which my sanity would have left me long ago.

Table of Contents

1.0 Introduction 1

1.1 Iterative Signal Reconstruction 2

1.2 Chapter Contents 4

2.0 Theoretical Development 7

2.1 Multiple Bases Representation (MBR) 7

2.2 Reconstruction of the MBR Signal 10

2.3 Definitions 12

2.3.1 Properties of Sets 12

2.3.2 Projection Operators 14

2.3.3 Mappings 15

2.4 Constrained Iterative Restoration 16

2.5 Projection Onto Convex Sets 19

2.5.1 Softened Constraint 22

2.5.2 Prototype Signal-Based Constraint 23

2.5.3 Global versus Block Constraint 25

2.6 Image Fidelity Measures 25

2.7 Error Signal Characteristics 26

2.8 Quantization of a priori Information 28

3.0 Constraint Definitions and Projection Operators 32

3.1 MBR Constraint 33

3.2 Positivity Constraint 36

3.3 Extremum Bound Constraint 37
3.4 Maximum Energy Constraint 38
3.5 Zero Crossing Constraint 38
3.6 Sign Constraint 41
3.7 Spike Constraint 42
3.8 Minimum Absolute Deviation Constraint 44
3.9 Minimum Increase Constraint 45
3.10 Minimum Decrease Constraint 47
3.11 Mean Constraint 48
3.12 Histogram Constraint 49
4.0 Simulation Results Using the Different Constraints 59
4.1 Closed and Convex Sets with MBR and Positivity Constraints 60
4.1.1 Zero Crossing Constraint 61
4.1.2 Sign Constraint 62
4.1.3 Spike Constraint 63
4.1.4 Minimum Absolute Deviation Constraint 64
4.1.5 Minimum Increment and Minimum Decrement Constraints 64
4.1.6 Extremum Bound and Maximum Energy Constraints 65
4.1.7 Summary 66
4.2 Initialization Vector 73
4.2.1 Observed Signal Plus White Noise 74
4.2.2 Observed Signal Plus a Constant 76
4.3 Histogram Constraint 82
4.3.1 Cause of Divergence 82
4.3.2 Convex Constraints that Prevent Divergence 83
4.4 Image Quality 89
4.4.1 1 bit/pixel or Less Required for a Composite Projection 90

4.4.2 More than 1 bit/pixel Required for the Composite Projection 92
4.5 Other Images 99
5.0 Conclusion 102
5.1 Convex Constraints 102
5.2 Initialization Vector 103
5.3 Laplacian Histogram Constraint 104
References 105
Vita 108

List of Figures

Figure 1: Distortion and approximate signal reconstruction 6

Figure 2: Projection Onto Convex Sets 31

Figure 3: Histogram specification 57

Figure 4: MSE of POCS reconstruction using, $P = P_{MBR} P_P$ 68

Figure 5: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{Z_1} P_P$, $P_2 = P_{MBR} P_{Z_2} P_P$, $P_3 = P_{MBR} P_{Z_3} P_P$, criterion 1 69

Figure 6: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{Z_1} P_P$, $P_2 = P_{MBR} P_{Z_2} P_P$, $P_3 = P_{MBR} P_{Z_3} P_P$, criterion 2 69

Figure 7: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{S_8} P_P$, $P_2 = P_{MBR} P_{S_{12}} P_P$, $P_3 = P_{MBR} P_S P_P$ 70

Figure 8: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{S_5} P_P$, $P_2 = P_{MBR} P_{S_{10}} P_P$, $P_3 = P_{MBR} P_{S_{15}} P_P$ 70

Figure 9: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{Spk_1} P_P$, $P_2 = P_{MBR} P_{Spk_2} P_P$ 71

Figure 10: Comparison of continuous and quantized spike constraint, $P_1 = P_{MBR} P_{Spk_1} P_P$ (quantized), $P_2 = P_{MBR} P_{Spk_1} P_P$ (continuous) 71

Figure 11: MSE of POCS reconstruction using, $P = P_{MBR} P_{MA} P_P$ 72

Figure 12: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{MI} P_P$, $P_2 = P_{MBR} P_{MD} P_P$, $P_3 = P_{MBR} P_{MD} P_{MI} P_P$ 72

Figure 13: Comparison of MSE of block and global minimum increase constraint, $P_1 = P_{MBR} P_{MI} P_P$ (global), $P_2 = P_{MBR} P_{MI} P_P$ (block) 73

Figure 14: MSE of reconstruction (50 realizations) using, $P_1 = P_{MBR} P_P$, ($P_2 = P_{MBR} P_P$ with I_n) 78

Figure 15: MSE of reconstruction (50 realizations) using, $P_1 = P_{MBR}P_SP_P$, ($P_2 = P_{MBR}P_SP_P$ with I_n) 79

Figure 16: MSE of POCS reconstruction using, $P = P_{MBR}P_SP_P$ with I_n 79

Figure 17: MSE of POCS reconstruction using, $P_1 = P_{MBR}P_P$, $P_2 = P_{MBR}P_{S_8}P_P$, $P_3 = P_{MBR}P_SP_P$ with I_K 80

Figure 18: MSE of POCS reconstruction using, $P = P_{MBR}P_{MD}P_{MI}P_P$ with I_K 80

Figure 19: Difference in MSE between zeroth and first iteration, $P = P_{MBR}P_P$ with I_K 81

Figure 20: Difference in MSE between zeroth and first iteration, $P = P_{MBR}P_SP_P$ with I_K 81

Figure 21: Histogram of $e_{x,k}$ after the first iteration using, $P = P_{MBR}P_P$ 86

Figure 22: Histogram of $e_{x,k}$ after the first iteration using, $P = P_{MBR}P_{MA}P_P$ 86

Figure 23: Histogram of $e_{x,k}$ after the first iteration using, $P = P_HP_{MBR}P_{MA}P_P$ 87

Figure 24: MSE of POCS reconstruction using, $P = P_HP_{MBR}P_{MA}P_P$ 87

Figure 25: MSE of POCS reconstruction using, $P = P_HP_{MBR}P_SP_P$ and I_k 88

Figure 26: MSE of POCS reconstruction using, $P = P_HP_{MBR}P_{S_8}P_{EB}P_P$ and I_k 88

Figure 27: POCS reconstructions of Mandrill's eye at 1 bit/pixel or less 95

Figure 28: POCS reconstruction of Mandrill's eye at more than 1 bit/pixel 97

List of Tables

Table 1: Histogram specification 58

Table 2: Number of signs constrained by zero-crossing constraint 62

Table 3: Comparison of projections containing convex projections 67

Table 4: MSE decrease and bit/pixel required for projections containing C_H 85

Table 5: % MSE decrease of reconstructions using convex constraints 100

Table 6: % MSE decrease of reconstructions using C_H, I_k, or I_n 101

1.0 Introduction

There are many practical examples in which we would like to reconstruct the original signal which has been distorted, sometimes intentionally, from the observed signal. For example, a technique called interferometric imaging is used in astronomy to increase the resolution of images obtained from telescopes. However, this technique requires the magnitude and phase of the stellar image. The phase is often difficult to obtain so astronomers would like to reconstruct the phase from the magnitude obtained from the telescopes. Also, in computerized tomography, only a limited amount of data may be available from scans of a patient's body and extrapolation of the rest of the data is desired.

In this thesis, we will consider the problem of recovering the original signal from an observed signal which has been compressed or distorted to decrease the transmission time of the signal. Along with the compressed signal, side information is transmitted and used to reconstruct the signal when it is received. Better performance may be achieved by transmitting side information about the signal in place of more transform coefficients and incorporating this side information into a reconstruction process. Of course, this procedure is beneficial only if the reconstruction produces a signal that is closer to the original than had the extra transform coefficients been transmitted instead.

Transform coding using unitary transforms is known to yield suboptimal rate-distortion bounds. The type of compression considered in this thesis is a form of transform coding known as Multiple Bases Representation (MBR) [Safar 1988, Khanna

1990]. The method of Multiple Bases Representation yields lower rates than typical transform coding by better approximating the signal. Instead of selecting vectors from one set of basis vectors as in transform coding, vectors are selected from multiple bases. The bases are unitary and allow fast transform algorithms so that the procedure is fast. Vectors can be chosen that better approximate the signal than vectors from a single basis set, because there are more vectors to choose from. However, orthogonality of the chosen vectors is lost.

Finding the optimum MBR signal requires a search of all permutations of the participating bases vectors. The method of Recursive Residual Projections (RRP) finds a slightly sub-optimum combination but introduces some noise into the process [Safar 1988, Khanna 1990].

1.1 Iterative Signal Reconstruction

In 1972, Gerchberg and Saxton [Gerchberg 1972] introduced an iterative algorithm that reconstructs an image from its magnitude and the magnitude of its Fourier transform. It consists of three steps. The first step involves computation of a random phase which is combined with the known magnitude. The second step involves computing the Fourier transform and setting the magnitude of the Fourier transform equal to the known magnitude while keeping the phase. The third step is to compute the inverse Fourier transform and impose spatial domain constraints to form a new estimate of the signal. The process then repeats at the second and subsequent steps.

In separate papers, Gerchberg and Papoulis [Gerchberg 1974, Papoulis 1975] applied the Gerchberg-Saxton algorithm to restoration of a bandlimited signal known only over a finite region. The basic steps are as follows: bandlimit the signal, then correct the known samples and recurse. The algorithm is proven to converge.

In 1978, a paper written by Youla, entitled "Generalized Image Restoration by the Method of Alternating Orthogonal Projections" was published [Youla 1978]. In this paper, he considers the problem of reconstructing a signal, known to exist in subspace $\mathcal{A}$ of a Hilbert space, from the orthogonal projection of the signal onto another known subspace $\mathcal{B}$ of the same Hilbert space. Necessary and sufficient conditions under which the signal is uniquely determined by the distorted signal are given as well as conditions under which the reconstruction can operate when noise is present.

The reconstruction algorithm involves projecting the image onto the original subspace $\mathcal{A}$, then onto the orthogonal complement subspace of $\mathcal{B}$, and then adding the image back in. This process is repeated until convergence occurs, which is guaranteed as long as certain conditions are satisfied. This reconstruction algorithm includes Papoulis's algorithm mentioned earlier as a specific case.

A survey of iterative restoration algorithms in 1981 shows that many reconstruction schemes can be generalized to a single algorithm [Schafer 1981]. Schafer's algorithm includes a distortion operator and a constraint operator and by defining these two operators appropriately many seemingly disparate reconstruction schemes can be subsumed under one general algorithm. Also, any number of constraints may be involved in the reconstruction process.

The mapping used in Schafer's algorithm must be proven to form a contraction which is a fairly strong requirement. Youla and Webb [Youla 1982] introduced an algorithm called Projection Onto Convex Sets (POCS) that replaced the requirement of a contraction mapping with a weaker one of closed and convex sets. The method of POCS includes Youla's previous algorithm as a specific case because closed linear manifolds are closed and convex.

In the POCS algorithm, the signal to be reconstructed is made to satisfy each constraint successively and if all constraints form closed and convex sets, convergence

is guaranteed. Thus the algorithm consists solely of constraints and no distortion operator is required. Youla defined several constraints which were then used in simulations performed by Sezan [Youla 1982, Sezan 1982]. The simulations show that the POCS algorithm performs better than the Gerchberg-Papoulis algorithm for image restoration by including other *a priori* information.

In this way, constraints may be defined that require *a priori* information to implement the corresponding projection during reconstruction. This *a priori* information is transmitted along with the distorted signal and used in the reconstruction process. To guarantee convergence, the defined constraints must form closed and convex sets, but not all constraints form closed and convex sets. However, there may be advantages to including nonconvex constraints. Even though the guarantee of convergence is lost if nonconvex constraints are used, the final MSE of those reconstructions that do converge may be significantly lowered.

Figure 1 diagrams the reconstruction process considered in this thesis. The MBR representation of the original signal x is computed. An approximation to the MBR signal is transmitted and received. The observed signal is decompressed using the inverse MBR transformation and is iteratively projected onto convex sets C_i, that are known *a priori* to partially define the original signal. The algorithm is guaranteed to converge to a signal that satisfies all the constraints involved. It is the aim of this thesis to determine which constraints are useful in reconstructing the original signal.

1.2 Chapter Contents

The second chapter, *Theoretical Development*, contains an introduction to the approximation of the Multiple Bases Representation (MBR) of a signal as a type of distortion. An introduction to iterative algorithms is presented and the method of Projection Onto Convex Sets (POCS) is proposed to reconstruct the MBR signal. The

algorithm is based in set theory and a summary of the necessary background to set theory is thus warranted. Fidelity measures of images are also discussed.

The next chapter, *Constraints*, lists constraints that are suitable to be used with the method of POCS for the reconstruction of the MBR signal. Each constraint is defined as well as its associated projection operator and the necessary *a priori* information requirements. All constraints are convex except the Laplacian histogram constraint.

The fourth chapter, *Simulations*, begins with a comparison of the constraints, defined earlier, in terms of mean square error of the POCS reconstructions. Next, the importance of a good initialization vector is demonstrated. Two initialization vectors are considered that both rely on *a priori* information. Performance of the nonconvex constraint is then discussed. Finally, the subjective image quality resulting from reconstructions using POCS with different sets of constraints is considered. In the last chapter, *Conclusion*, observed results from the simulations are discussed.

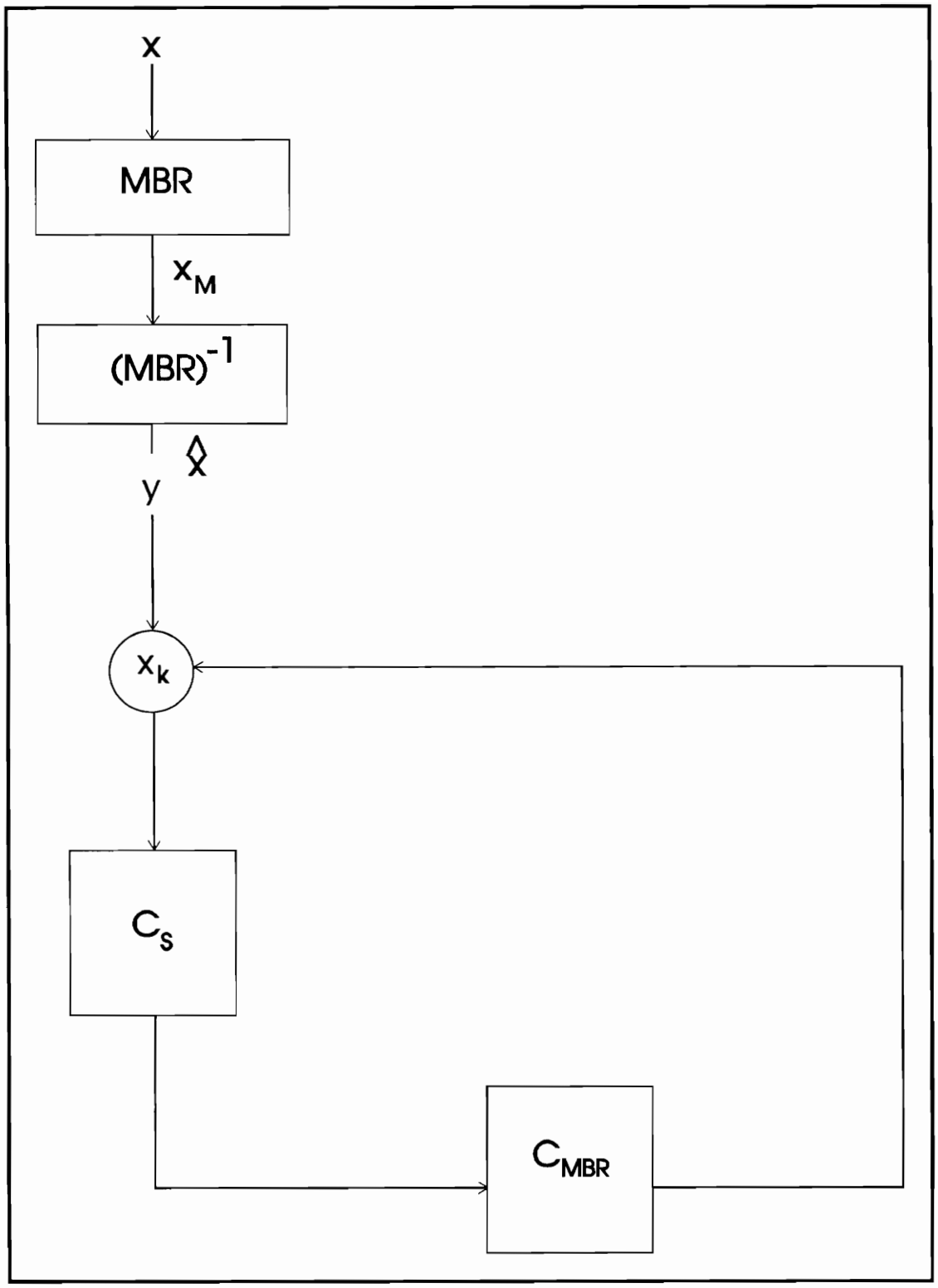


Figure 1: Distortion and approximate signal reconstruction

2.0 Theoretical Development

Signal reconstruction is the process of obtaining the original signal from a distorted version of that original. The distortion is often imposed by the real world in the form of unwanted degradation effects, but may also be purposely imposed for purposes of more efficient transmission. It is the objective of image reconstruction to, at the least, find a better approximation to the original image through the use of *a priori* information.

The type of distortion and *a priori* information available strongly influences the method of reconstruction. Other questions should also be answered. Is exact reconstruction required or is approximate reconstruction enough? Is the distortion operator known exactly or only approximately? Is the observed signal noisy? These questions should be answered before an appropriate reconstruction algorithm can be chosen. Thus, approximation to the Multiple Bases Representation (MBR) of a signal as a type of distortion is addressed first. The method of Projection Onto Convex Sets (POCS) is then proposed as an appropriate reconstruction algorithm. The algorithm is based on set theory and therefore an introduction is provided to set theory.

2.1 Multiple Bases Representation (MBR)

Let $\mathcal{H}$ be a N-dimensional, real Hilbert space with norm,

$$\|x\| = \left(x^T x\right)^{\frac{1}{2}} \tag{2.1}$$

and inner product defined as,

$$(x,y) = x^T y \tag{2.2}$$

where $x, y \in \mathcal{H}$. Elements of $\mathcal{H}$ may be thought of equivalently as a set of points in N dimensional space or a set of signals of length N. Consider L unit norm vectors in $\mathcal{H}$ with L > N. These L vectors can span $\mathcal{H}$ but cannot form a basis because the L vectors cannot form a linearly independent set.

If N vectors are chosen from the set of L vectors such that a linearly independent set is formed, then this N-tuple forms a basis for $\mathcal{H}$. Each signal x in $\mathcal{H}$ can be represented by a unique linear combination of these N vectors,

$$x = Ax' \tag{2.3}$$

where A is an N x N matrix whose columns are the vectors of the N-tuple and x' is a unique representation of x in $\mathcal{H}$.

Now consider K vectors chosen from the set of L vectors such that the K vectors again form a linearly independent set and K < N. Let $\mathcal{M}$ be the K-dimensional subspace of $\mathcal{H}$ formed by these K vectors. The K vectors can only form an approximation to x.

Let $\hat{x}$ represent this approximation and $e_x = x - \hat{x}$ be the approximation error. Given $x \in \mathcal{H}$, it is desired to find $\hat{x} \in \mathcal{H}$ that minimizes $\|e_x\|^2$. Let,

$$\hat{x} = Bx_{\mathcal{M}} \approx x \tag{2.4}$$

where B is a N x K matrix whose columns are the K chosen vectors. Using the pseudo-inverse [Brogan 1985],

$$x_{\mathcal{M}} = (B^T B)^{-1} B^T x \tag{2.5}$$

and combining (2.4) and (2.5),

$$\begin{aligned}\hat{x} &= Bx_{\mathcal{m}} \\ &= B(B^T B)^{-1} B^T x\end{aligned} \tag{2.6}$$

By the orthogonal projection theorem, x may be represented as the sum of two signals,

$$x = P_{\mathcal{m}} x + P_{\mathcal{m}^\perp} x \tag{2.7}$$

and,

$$\hat{x} = P_{\mathcal{m}} x \tag{2.8}$$

where $P_{\mathcal{m}}$ is the orthogonal projection of x onto $\mathcal{m}$ and $P_{\mathcal{m}^\perp}$ is the orthogonal projection of x onto $\mathcal{m}^\perp$. By comparing the equations in (2.6) and (2.8), it can be seen that,

$$P_{\mathcal{m}} = B(B^T B)^{-1} B^T \tag{2.9}$$

We can approximate x better by choosing the optimum subset of K vectors from L > N vectors than by choosing the K vectors from N vectors [Safar 1988]. Thus, given x, we would like to find the optimum set of K vectors. Ideally, the K vectors are chosen such that e_x is minimized but this requires a full search of all $\binom{K}{L}$ permutations, which is not practical due to time constraints. Safar [Safar 1988] has proposed a sub-optimal method of selecting the K vectors called Recursive Residual Projection (RRP). In this procedure, K vectors are still chosen to form $\mathcal{m}$, however, the choice of the K vectors is sub-optimal so that,

$$\hat{x}_{RRP} = P_{\mathcal{m}}^{\circ} x + \Delta x \tag{2.10}$$

where $P_{\mathcal{m}}^{\circ}$ is the projection of x onto the optimal set of K vectors and Δx can be viewed as noise introduced into the process to speed calculation of $\hat{x}_{RRP}$.

2.2 Reconstruction of the MBR Signal

The signal recovery problem may be described using the discrete linear model [Andrews 1977, Pratt 1978],

$$y = Dx + n \tag{2.11}$$

where x is the unknown original signal, n represents unknown additive noise, and y is the observed signal. The distortion operator D, is a N x N distortion matrix representing, for example, distortion by a linear system, or a nonlinear band-limited system, and may be known or partially known. Let,

$$D_{MBR} = B(B^T B)^{-1} B^T \tag{2.12}$$

where D_{MBR} represents the distortion of x defined in (2.6). Let D_{RRP} represent the distortion defined in (2.10). The objective of signal recovery is to find x.

If we assume that there is no additive noise and D is nonsingular then (2.11) is solved by,

$$x = D^{-1} y \tag{2.13}$$

There are several problems that make this approach unattractive. First of all, finding the inverse is an ill-posed problem in which small perturbations in the input lead to large perturbations in the output. Thus, if the assumption about additive noise is wrong and noise is present, the computed solution may be nowhere near the original signal due to numerical instabilities. To see this more clearly, assume that (2.13) is computed when noise is present,

$$\begin{aligned} x' &= D^{-1} y \\ &= D^{-1}(Dx + n) \\ &= x + D^{-1} n \end{aligned} \tag{2.14}$$

and the second term may actually dominate the solution. The problem is worsened if D is nearly nonsingular so that D^{-1} has large entries, exacerbating the effect of the noise.

If D is singular then this corresponds to a linear dependence between at least two of the rows or columns and formally the solution does not exist. When D is singular, and applied to x, certain components of the signal are irretrievably lost, which means that instead of a unique solution, there is an infinite number of solutions. A pseudo-inverse may be used to determine the "best" solution from the infinite number, however, the pseudo-inverse problem is still ill-posed [Stark 1987]. The MBR distortion operator D_{MBR} is singular, because B is not full rank.

The above discussion has assumed that D is known exactly. While D_{RRP} may be applied exactly to any signal, and D_{RRP} can be found in closed form, the solution in (2.13) cannot be used because D_{RRP}^{-1} does not exist. However, as mentioned earlier, D_{RRP} is an approximation to D_{MBR}. In the case that D is only known approximately, there are fewer alternatives. Certainly, computing the inverse of the approximation is even more numerically unstable and many optimization techniques, both iterative and non-iterative, rely on exact knowledge of D.

Typically, there are two broad classes of algorithms, iterative and non-iterative. The non-iterative algorithms are usually based on a least squares algorithm with the more sophisticated algorithms incorporating knowledge of noise statistics and other *a priori* knowledge. Iterative techniques also incorporate *a priori* knowledge.

We wish to apply a number of constraints to aid in the reconstruction and allow D to remain only approximately known. For example, if it is an image, each of the pixels has positive brightness; negative brightness is not possible. The distorted signal, y, may contain pixel values that are negative and an improved signal can be obtained if this additional information or constraint is incorporated into the reconstruction process. Also, if the image is quantized the final image should be quantized. The method of Projection

Onto Convex Sets satisfies the above two requirements. It allows any number of nonlinear constraints to be subsumed automatically.

A mapping is used that has one or more fixed points that are known to be either the original or an acceptable approximation to the original. The mapping usually consists of constraints that the original signal is known to satisfy and is iteratively applied until a criterion for convergence is satisfied. Iterative techniques are also ill-posed but if the types of *a priori* information are limited, the iterative process can be proven to converge.

2.3 Definitions

To be useful, iterative algorithms must be proven to converge. Convergence and uniqueness proofs for iterative algorithms are usually based on set theory. Therefore, before examining iterative algorithms, several definitions will be introduced. Griffel and Brogan [Griffel 1981, Brogan 1985] provide good introductions to set theory.

2.3.1 Properties of Sets

First, let us consider S, a subset of $\mathcal{H}$. Let $f, g \in S$. The subset S is **bounded** if there is a β such that,

$$\|f\| \leq \beta \tag{2.15}$$

for all $f \in S$. The subset S is said to be **linear** if,

$$af + bg \in S \tag{2.16}$$

for all scalars a, b. These two properties of sets are mutually exclusive. A bounded set cannot be linear and a linear set cannot be bounded. A distance, d, between two elements, $f, g \in S$, may be defined as,

$$d = \|f - g\| \tag{2.17}$$

With the distance between two elements defined, we can define convergence of a sequence of elements. For a sequence $\{f_n\}$ of elements, $f_1, f_2, ..., f_\infty \in \mathcal{H}$, we say $\{f_n\}$ converges to f_∞, written $f_n \to f_\infty$, if

$$\|f_n - f_\infty\| \to 0 \text{ as } n \to \infty \tag{2.18}$$

This is known as **strong convergence** and states that a sequence of elements converges strongly if the distance between successive elements tends to zero. Another type of convergence occurs only in infinite dimensional spaces. Let $\mathcal{X}$ be an infinite dimensional Hilbert space. If $f_1, f_2, ..., f_\infty \in \mathcal{X}$, we say $\{f_n\}$ converges weakly to f_∞, written $f_n \to f_\infty$ weakly, if

$$(f_n, g) \to (f_\infty, g) \tag{2.19}$$

for all $g \in \mathcal{X}$. Strong convergence implies **weak convergence**. Weak convergence implies strong convergence for a closed and finite dimensional space.

Two classifications of sets are open and closed. However, before these types of sets can be defined a limit point must be defined. The element f is a **limit point** of S if there is a sequence $\{f_n\}$ of elements of S such that $f_n \to f$ and for each n, $f_n \neq f$. This says that a sequence converges to a limit point but never actually reaches the limit point. The subset S is **open** if for each $f \in S$ there is a $\partial > 0$ such that $g \in S$ whenever $\|f - g\| < \partial$. A set is **closed** if it contains all of its limit points. Open and closed are not mutually exclusive as the names may suggest. Thus, a set may be either closed or open, or neither open nor closed. A common example of an open set is the interval (0,1) on the real number line. The set is not closed because the sequence $\left\{\frac{1}{2}, \frac{1}{3}, \frac{1}{4} ...\right\}$ clearly converges to the limit point zero which is not in the set. The set may be made closed by including

the points 0 and 1. Thus, the set [0,1] is closed. Likewise, the set (0,1] is neither open nor closed. A **subspace** is a linear and closed subset.

A set may be defined as either convex or nonconvex. The subset S is called **convex** if for any $f, g \in S$, $\mu f + (1-\mu)g \in S$ for all $0 \le \mu \le 1$. Geometrically, this means that a line segment joining any two points in the set must be entirely inside the set. A sphere is the archetypal example of a convex set. A set is **nonconvex** if it is not convex. All subspaces are convex.

Lastly, an angle may be defined between signals f and g as

$$cos(\theta) = \frac{(f,g)}{\|f\| \, \|g\|} \tag{2.20}$$

The angle between two orthogonal vectors is 90°. The angle is not defined if f or g are zero vectors, because the norm of a zero vector is zero.

2.3.2 Projection Operators

Closed and convex sets have an important role in this thesis. Let C be a closed and convex subset of $\mathcal{H}$. Given $f \in \mathcal{H}$ and $g \in C$ we wish to find P_C such that $P_C : H \to C$. The, in general, nonlinear projection operator $P_C : \mathcal{H} \to C$ has the property,

$$\inf_{x \in C} \|f - x\| = \|f - g\| \tag{2.21}$$

The operation in equation (2.21) assigns to every element f, in $\mathcal{H}$, the closest vector g in C, which then has the property of being unique. The property in (2.21) may also be expressed,

$$\|f - P_C f\| = \min_{x \in C} \|f - x\| \tag{2.22}$$

and may be solved using the method of Lagrange Multipliers. All points g, in C, are fixed points of P_C, i.e. $P_C g = g$. If C is a subspace then x is called the orthogonal projection of f onto C.

2.3.3 Mappings

Consider a mapping $T: S \rightarrow S$. An iterative equation may be written,

$$f_{n+1} = Tf_n \tag{2.23}$$

To be useful, an iterative equation must be proven to converge to a solution. Generally a solution is obtained only after an infinite number of iterations, however, acceptable convergence is usually obtained after a finite number of iterations. If $f = Tf$ then f is called a **fixed point** of the mapping T.

Two important classes of mappings are contraction and nonexpansive mappings. The mapping T is a **contraction mapping** if $\|Tf - Tg\| \leq a\|f - g\|$ where $0 < a < 1$. In other words, Tf and Tg are closer together than f and g by some definite distance. The contraction mapping theorem states [Youla 1982] that if $T: S \rightarrow S$ is a contraction mapping then there is exactly one $f \in S$ such that $f = Tf$. Furthermore, for any $f_0 \in \mathcal{H}$ the sequence $\{f_n\}$ defined by (2.23) converges to f. This theorem is useful in that it gives a prescribed method for finding the desired point as well as proving that a solution exists and that it is unique.

A less restrictive mapping is the nonexpansive mapping. The mapping $T: S \rightarrow S$ is **nonexpansive** if $\|Tf - Tg\| \leq \|f - g\|$. The corresponding iterative equation in (2.23) is not guaranteed to converge, because a nonexpansive mapping does not require a definite reduction in distance by a definite amount. However, when reconstructing an image, it is beneficial to know that application of a nonexpansive mapping will not worsen an image.

Nonexpansive mappings have at least one fixed point, and usually an infinite number. The projection operator P_C defined in Section 2.3.2 is a nonexpansive mapping [Youla 1982].

Let T be a product of two or more individual projections. The mapping T will be a nonexpansive mapping if all T_i are nonexpansive mappings. Furthermore, if at least one of the T_i is a contraction then T is also a contraction,

$$\begin{aligned} \|T_2T_1x - T_2T_1y\| &\le a\|T_1x - T_1y\| \\ &\le a\|x - y\| \end{aligned} \tag{2.24}$$

where T_2 is a contraction mapping, $0 < a < 1$, and T_1 is a nonexpansive mapping.

2.4 Constrained Iterative Restoration

The idea behind iterative restoration algorithms is to find at least a nonexpansive mapping $T: S \rightarrow S$ that has fixed points that are acceptable approximations to x. If T is a contraction then there is only one fixed point which should be x, the original signal, and equation (2.23) can be used to find this fixed point.

To find a mapping that has fixed points that are feasible solutions, T is usually derived from constraints that the original signal is known to satisfy. Schafer et al. [Schafer 1981] propose the identity,

$$x = Cx + \lambda(y - DCx) \tag{2.25}$$

as a general form to be used for iterative restoration. This equation requires knowledge of the distortion operator D, and the constraint operator C, both of which may be nonlinear. The parameter λ, is usually a constant but it may also be a function of independent variables, or a function of x, and is useful to help speed convergence [Schafer 1981]. The constraint operator C, may be derived from J unrelated constraints, $C = C_1C_2...C_J$.

Equation (2.25) can be made iterative,

$$\begin{aligned} x_{k+1} &= Cx_k + \lambda(y - DCx_k) \\ &= Cx_k + \lambda y - \lambda DCx_k \\ &= \lambda y + (I - \lambda D)Cx_k \\ &= \lambda y + Gx_k \\ &= Tx_k \end{aligned} \tag{2.26}$$

where,

$$G = (I - \lambda D)C \tag{2.27}$$

In order for G to be useful to signal reconstruction it must be either a contraction or a nonexpansive operator. From the contraction mapping theorem, the iterative equation in (2.26) is guaranteed to converge to a unique point if T is a contraction mapping, or because,

$$\left\| Tx_i - Tx_j \right\| = \left\| Gx_i - Gx_j \right\| \tag{2.28}$$

it is sufficient for G to be a contraction mapping. The initialization vector x_0, may be any signal in $\mathcal{H}$, but it is usually chosen to be λy. Whether G is a contraction mapping depends on the transformations D and C and the parameter λ which may be used to force G to be a contraction mapping. According to inequality (2.24), if $(I - \lambda D)$ and C are at least nonexpansive and one or both are contractions then G is a contraction. Likewise, given that all the C_i's are at least nonexpansive, if any of the C_i's are contractions then C is a contraction and if all the C_i's are nonexpansive then C is nonexpansive.

It is possible to prove that G is a contraction by directly applying the definition of a contraction mapping. By doing this the bounds of the parameter λ can be determined such that a contraction is formed. However, this becomes increasingly difficult as the number of constraints increases because each constraint is nonlinear.

If G is a contraction mapping then, by the contraction theorem, the iterative equation in (2.26) has one and only one fixed point, which is the original signal. This can

be seen by replacing x_k with x. Thus, if T is a contraction, there is only one fixed point and it is the original signal. On the other hand, if G is non-expansive then the iterative equation in (2.26) typically has an infinite number of fixed points. Each of these fixed points may or may not be a fixed point of C. However, if the fixed point of T is also a fixed point of C, then,

$$x_k = x_k + \lambda(y - Dx_k) \tag{2.29}$$

which implies that,

$$y = Dx_k \tag{2.30}$$

Thus, for the non-expansive case it is a simple matter to check if the fixed point of T is also a fixed point of C. If so, then the fixed point of T is a valid solution in that it satisfies all constraints and produces the observed signal upon distortion. However, the fixed point may not be the original signal [Trussell 1983].

Normally, two successive iterations are compared to one another to determine if the algorithm has converged. If the convergence is slow then this criterion may not be appropriate. Trussell proposes redefining convergence, especially when the observed signal is noisy [Trussell 1983]. The term $\{y - Dx_k\}$ is called the residual and is the difference between the original distorted observed image and the distorted restored image. The residual signal is zero when the approximation x_k produces the observed signal y. However, if y is the distorted signal plus additive noise then the residual signal produces the additive noise. Trussell uses the residual signal to redefine convergence by forcing the residual to be a likely realization of the noise. When $\|y - Dx_k\|^2$ is approximately equal to the noise power the algorithm is stopped.

If G is a contraction mapping, the iterative equation in (2.26) requires,

$$\lim_{k \to \infty} Dx_k = y \tag{2.31}$$

and so D is required to implement the algorithm. The next section introduces an algorithm, called Projection Onto Convex Sets (POCS), that can be considered a generalization of Schafer's algorithm in that they both iteratively enforce constraints. However, the distortion operator D is not required and the limit in (2.31) is not enforced in POCS. A constraint is defined in Section 3.1 that can enforce the limit in (2.31) by defining the constraint $C = \{f : Df = y\}$.

2.5 Projection Onto Convex Sets

Youla [Youla 1982] derived a signal reconstruction algorithm that has two significant benefits over other signal reconstruction algorithms. The method of Projection Onto Convex Sets (POCS) allows the inclusion of an arbitrary number of nonlinear constraints and guarantees at least weak convergence in an infinite dimensional space and strong convergence in a finite dimensional space. The only restriction is that the sets formed by the constraints must be closed and convex.

Consider J constraints each forming a closed and convex set of signals in $\mathcal{H}$. Let C_i represent the i^{th} set. The intersection of these J constraints is,

$$C_0 = \bigcap_{i=1}^{J} C_i \tag{2.32}$$

and every signal in C_0 satisfies all J constraints. In general, there will be many signals in C_0, including x.

All signals in C_0 are considered to be close approximations to x. Given y, we would like to find a signal belonging to C_0. If $P_0 : \mathcal{H} \rightarrow C_0$ were known then the problem would be solved by,

$$x = P_0^{-1} y \tag{2.33}$$

however, P_0 is a combination of J nonlinear constraints and in general the inverse will be difficult to find, even if it exists. Therefore, iterative techniques are used to find a signal in C_0.

Considering an iterative approach, we would like to find a mapping $T:\mathcal{H} \to \mathcal{H}$ that has all its fixed points in C_0. Let P_i represent the, in general, nonlinear projection operator $P_i:\mathcal{H} \to C_i$. If f satisfies the i^{th} constraint, then f is a fixed point of P_i and if f does not satisfy the i^{th} constraint, it can be made to do so via the projection operator P_i. A composite projection operator is defined,

$$P_C = P_1 P_2 ... P_J \tag{2.34}$$

In general, P_C is not P_0. As mentioned in Section 2.3.3, P_C is a nonexpansive mapping. Youla shows that all fixed points of P_C are points in C_0. It can easily be seen that each point in C_0 satisfies every C_i and is a fixed point of every P_i. Thus,

$$\begin{aligned} P_C x &= P_1 P_2 ... P_J x \\ &= P_1 P_2 ... P_{J-1} x \\ &= x \end{aligned} \tag{2.35}$$

and every signal in C_0 is a fixed point of P_C, but the converse, that all fixed points of P_C are members of C_0 is not obvious.

The fact that the mapping P_C, is nonexpansive and contains fixed points is not enough to guarantee that the iterative equation,

$$x_k = P_C^k x_0, \quad k = 1 \to \infty \tag{2.36}$$

converges to a limit point. However, Youla does show that the iterative equation in (2.36) is guaranteed to converge weakly in infinite dimensional spaces and strongly in finite dimensional spaces [Youla 1982]. Note that each x_k, $k > 0$ belongs to C_1 because P_1 is the last projection that operates on the signal.

In general, the final solution will not be the projection of the initialization signal onto C_0. The restored signal lies on the boundary of C_0 just as if the initialization signal were projected onto C_0. However, unlike the projection onto C_0, the restored signal is generally not the closest signal in C_0 to the initialization vector. The final solution also depends on the ordering of the projections in the composite projection P_C in (2.34).

The mapping P_C may be difficult to find for many nonlinear constraints. Youla [Youla 1982] shows that the projection operators may be applied successively,

$$\begin{aligned} x_{n+1} &= P_{\alpha(n)} x_n \\ \alpha(n) &= 1 + n \bmod J \qquad n \geq 0 \end{aligned} \tag{2.37}$$

and $\{x_n\}$ converges to the same limit as $\{x_k\}$ in (2.36). Thus, any signal in $\mathcal{H}$ can be projected iteratively from one set to another until an acceptable criterion for convergence has been satisfied. Consider Figure 2, which simulates the algorithm for J = 2 and N = 2. C_0 is the intersection of two closed convex sets C_1 and C_2. A point in C_0 can be found by iteratively projecting an arbitrary point from one closed convex set to another.

Youla [Youla 1982] points out that slow convergence may occur when the iteration is near a point of tangency of two or more sets. To speed convergence, each projection operator may be relaxed,

$$T = T_1 T_2 \ldots T_J \tag{2.38}$$

where,

$$T_i = I + \lambda_i (P_i - I) \qquad 0 < \lambda_i < 2 \tag{2.39}$$

where I is the identity matrix. Youla shows that the relaxed projection operator in (2.39) has the same fixed points as P_i and the mapping in (2.38) can be used in place of that in

(2.37). Levi and Stark [Levi 1984] show that λ may be varied from one iteration to the next and discuss optimization of λ for $J = 2$.

The POCS algorithm is robust in that even if noise is introduced into the observed signal, a feasible solution is always found. A different final solution may be obtained, but a signal in C_0 will always be found in the limit. POCS also allows an arbitrary number of nonlinear closed and convex sets to be included in the reconstruction process. While many constraints form closed and convex sets, some do not. For example, the set of signals consisting of quantized samples is not convex. If the constraints are not closed and convex, they may cause divergence when the constraints are included in POCS. However, as demonstrated in Section 4.4, they may still be used successfully. Even with this limitation, the robustness of the algorithm and the ease of incorporating many constraints, make the POCS algorithm applicable to many signal reconstruction problems.

2.5.1 Softened Constraint

A constraint may be a function of the originally observed signal y. If the observed signal is corrupted by additive noise then the projection $P:\mathcal{H} \rightarrow C$ now maps $P:\mathcal{H} \rightarrow C'$ and the projection is no longer correct. This can be seen by considering the fact that every fixed point of P is an element of C, so that if and only if $x = Px$ then $x \in C$. If P is a function of y and y is noisy then P changes to P' and $C' = \{x : x = P'x\}$. Essentially, the constraint set C which is enforced by P has moved and C is not being enforced anymore. Furthermore, if $C' \cap C_i$, $i = 1,2,...,J$ is empty then constraint incompatibility results and the algorithm cannot converge to a solution because no solution exists.

To prevent constraint incompatibility from occurring, the constraint may be modified to include the original set plus possible realizations of the additive noise [Beex 1983, 1984]. In this case, $P:\mathcal{H} \rightarrow C_{Soft}$ and $C' \subset C_{Soft}$ which prevents constraint incompatibility. Thus, something must be known about the noise in order to soften C.

Geometrically, the size of the constraint set is increased so that C_0 is ensured to be nonempty. An example of a softened constraint is given in Section 2.5.2.

The process of softening constraints can be compared to redefining convergence introduced in Section 2.4. The process of softening constraints is applied to only the constraints that are known to be dependent on the observed signal. Other constraints that are known to be correct are not affected. Redefining convergence is like softening the solution set rather than a particular constraint. The final solution does not have to satisfy constraints used in the algorithm. Although the iteration will move toward solutions that satisfy constraints the final solution must only be a signal that, when distorted, produces a realization of the additive noise. If it is important that the reconstructed signal satisfy the defined constraints, the individual constraints should be softened.

2.5.2 Prototype Signal-Based Constraint

The reconstructed signal may be known to be similar to a known signal. This knowledge may be used to constrain the reconstructed signal. Sezan et al. [Sezan 1990] define a prototype image-based constraint to be of the form,

$$C = \left\{ f : \left\| f_p - f \right\|^2 \le \delta \right\} \tag{2.40}$$

where f_p is the prototype signal, and δ is an upper bound on the variation of the signal from the prototype. The signal f_p is found by applying an operator O to the observed image. A confidence measure δ is defined, that is the expected variation of the known image from the prototype image. Thus, the constraint defined in (2.40) becomes,

$$C = \left\{ f : \left\| Oy - f \right\|^2 \le c\Psi \right\} \tag{2.41}$$

where,

$$\Psi = E\left\{\|Oy - x\|^2\right\} \tag{2.42}$$

and $c \geq 1$ and is a confidence measure of whether $y \in C$. The operator O defines the constraint and Ψ must be known *a priori* unless certain simplifying assumptions are made about the expected value in (2.42).

The constraint in (2.40) is a softened constraint. To see this, let

$$\begin{aligned} C_{Hard} &= \left\{f : f = f_p\right\} \\ &= \left\{f : \|f_p - f\|^2 = 0\right\} \end{aligned} \tag{2.43}$$

The hard constraint C_{Hard} forces the iterated signal to be equal to the prototype signal. If we assume that differences between the observed signal and the prototype signal are the result of noise, we can soften C,

$$C_{Soft} = \left\{f : \|f_p - f\|^2 \leq \delta\right\} \tag{2.44}$$

The soft constraint C_{Soft} includes C_{Hard} and allows for the original signal plus realizations of noise with noise power δ and is identical to the constraint in (2.40).

Let $e_p = f_p - f$ where f_p and f are defined previously and e_p is then the error between a signal and its prototype. The constraint in (2.40) is then equivalent to placing a bound on the variation of e_p. The soft constraint C_{Soft}, is not limited to placing a bound on the variation of e_p; it can be any closed convex constraint that restricts e_p somehow. If a constraint is closed and convex, it can be applied to f as well as to $f - f_p$.

In this thesis, the observed image is used as a prototype, so that $f_p = y$ and $e_p = y - x$. Thus, the prototype signal is a distorted version of the original signal. Let $e_{p,k} = y - x_k$. The computation of $e_{p,k}$ requires that y be stored during the entire reconstruction. Knowledge about e_p is transmitted and used during the reconstruction process.

2.5.3 Global versus Block Constraint

During distortion the signal is split into blocks. A constraint may be applied to either the entire signal or to each block of the signal. For the same constraint, block constraints require more *a priori* information than global constraints because there is usually more than one block in a signal. Block constraints require the same *a priori* information to be transmitted for each block as that transmitted for a global constraint. On the other hand, a block constraint may be more effective than a similar global constraint so that the fidelity of the final reconstruction may well determine a combination of global and block constraints to be most effective.

2.6 Image Fidelity Measures

It is important to have an objective measure of image distortion in order to compare the effectiveness of different reconstruction procedures. A common measure is the mean square error defined by,

$$e_{ms}^2 = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} E(x_{i,j} - \hat{x}_{i,j})^2 \tag{2.45}$$

where the x_{ij} represent the NxM original image and the $\hat{x}_{ij}$ represent the distorted image or reconstruction. The mean square error in (2.45) can be approximated by,

$$\begin{aligned} e_{ms}^2 &\cong \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{i,j} - \hat{x}_{i,j})^2 \\ &\equiv MS(x - \hat{x}) \end{aligned} \tag{2.46}$$

Other error measures may be defined such as absolute error or maximum error. However, it is known that these error measures do not reflect subjective evaluation [Mannos 1974]. We would like to evaluate distortion to reflect the subjective evaluation by humans.

Mannos and Sakrison [Mannos 1974] note that although the visual system is not linear, it may be treated as such after an initial nonlinear transformation. They suggest the initial transformation,

$$f(x) = x^{1/3} \tag{2.47}$$

followed by a linear spatially invariant operation,

$$A(f_r) \approx 2.6\left[0.0192 + 0.114 f_r\right] exp\left[-\left(0.114 f_r\right)^{1.1}\right] \tag{2.48}$$

where

$$f_r = \sqrt{f_x^2 + f_y^2} \tag{2.49}$$

The parameters in (2.47) and (2.48) were chosen to reflect subjective evaluations of various images. Thus, a subjective evaluation may be obtained after applying (2.47) and (2.48) to the image.

The above human visual system error measure will not be implemented in our simulations. Instead, mean square error is used as an initial indicator of the success of a particular algorithm. Images are then included for the reader's subjective evaluation.

2.7 Error Signal Characteristics

The reconstruction process is meant to produce the original signal,

$$x = P_{\mathcal{M}} x + P_{\mathcal{M}^\perp} x \tag{2.50}$$

from the observed signal,

$$y = P_{\mathcal{M}} x \tag{2.51}$$

Thus, the error signal,

$$e_x = P_{\mathcal{M}^\perp} x \tag{2.52}$$

is orthogonal to the observed signal. In Section 3.1, a constraint will be defined that forces $P_{\mathcal{M}} x_k = y$. Furthermore, let,

$$e_{x,k} = P_{\mathcal{M}^\perp} x_k \tag{2.53}$$

Thus, to reconstruct x, $e_{x,k} = e_x$, $k \to \infty$. It is therefore reasonable to include constraints on e_x, and so we examine a few of its properties.

If the basis vectors of $\mathcal{M}$ are formed from the discrete cosine and Haar transforms, the probability density function of e_x is modeled well as Laplacian [Khanna 1990]. The Laplacian probability density function can be specified with one parameter α as follows,

$$p_{e_x}(z) = \frac{\alpha}{2} e^{-\alpha|z|} \tag{2.54}$$

and has variance,

$$\sigma_{e_x}^2 = \frac{2}{\alpha^2} \tag{2.55}$$

In general, the signal $P_{\mathcal{M}^\perp} x_k$, will not have a Laplacian density function. Indeed, the first iteration has all elements zero in $P_{\mathcal{M}^\perp} x_k$. Therefore, it is desired to modify this component of x_k to have a Laplacian density function defined by,

$$\alpha = \sqrt{\frac{2}{\sigma_{e_x}^2}} \tag{2.56}$$

where $\sigma_{e_x}^2$ is known *a priori*.

The RRP procedure always removes the DC component from the image block x [Khanna 1990]. Thus, e_x must have a mean of zero. This fact could be used as a constraint but is actually included in another more powerful constraint which is introduced in Section 3.1. Also, the probability distribution function is double sided and centered at zero and so has zero mean as the probability distribution indicates.

We will see later, in Section 3.1, that D_{MBR} always removes energy from x. For now we can see from (2.50) that,

$$\|x\|^2 = \|P_m x\|^2 + \|P_{m^\perp} x\|^2 \tag{2.57}$$

Thus it should be clear that the reconstruction process must add energy back into y in order to obtain x. Typically, the iterated signal will have less energy than the original. However, simulations show that this is not always the case.

2.8 Quantization of *a priori* Information

Let C be defined such that,

$$C = \{f : F(f) \geq k\} \tag{2.58}$$

where $F(f)$ is a linear or nonlinear function of f. It is beneficial to quantize $k \geq 0$ such that the required *a priori* information of C is decreased to a sufficient level to make reconstruction from *a priori* information advantageous. We may assume k to be quantized to $Q[k]$ leading to

$$\hat{C} = \{f : F(f) \geq Q[k]\} \tag{2.59}$$

Let k be a continuous variable with probability density function $p_K(k)$. The continuous variable k is to be represented by a set of Q reconstruction levels such that,

$$Q[k] = r_i\left[u_i - u_{i+1}\right] \qquad \mathrm{i} = 1,\ldots,Q \tag{2.60}$$

where r_i is the i^{th} reconstruction level, $d_i \le r_i \le d_{i+1}$, and d_i is the i^{th} decision level. The objective is to find d_i and r_i for all i that minimizes $\|k - Q[k]\|^2$. The solution cannot be found in closed form and a numerical approximation procedure can be found in Shanmugan [Shanmugan 1988].

The quantized variable $Q[k]$ is a noisy version of k so that the constraint $\hat{C}$ is moved, just as a constraint that is dependent on the noisy observed signal is moved. As mentioned earlier, the noisy information may cause constraint incompatibility. To eliminate the possibility of incompatibility, the quantized constraints can be softened to include expected noise from the quantization process.

The constraints can be softened for the reconstruction process (on the receiving end) because the noise from the quantization process is known. Furthermore, the quantization process itself can be modified to produce softened constraints by restricting the selection of the reconstruction levels. This is a suboptimal quantization process and corresponds to modifying the noise so that constraint incompatibility will not occur. The noise is used to soften the constraint so that $\hat{C} \subset C$. This will cause a degradation in the effectiveness of the constraint depending on the degree of quantization.

The particular modification to the quantization process depends on the constraint whose *a priori* information is being quantized. For the constraint defined in (2.58) the appropriate method is to choose each reconstruction level such that $r_i = d_i$ so that $Q[k] \le k$ and the constraint is softened.

The reconstruction and decision levels on the transmitting and receiving ends can be computed using the same algorithm, based on a known probability density function so that the reconstruction and decision levels do not have to be transmitted. Transmission of a parametric representation for the probability density function is all that is required. Recall that for the Laplacian, only a single parameter is needed. Even if the quantization

process is modified as described above it is possible for both ends to do this modification because it is only dependent on the particular constraint not on the quantized variable k. The levels depend on $p_K(k)$ which varies from one constraint to another.

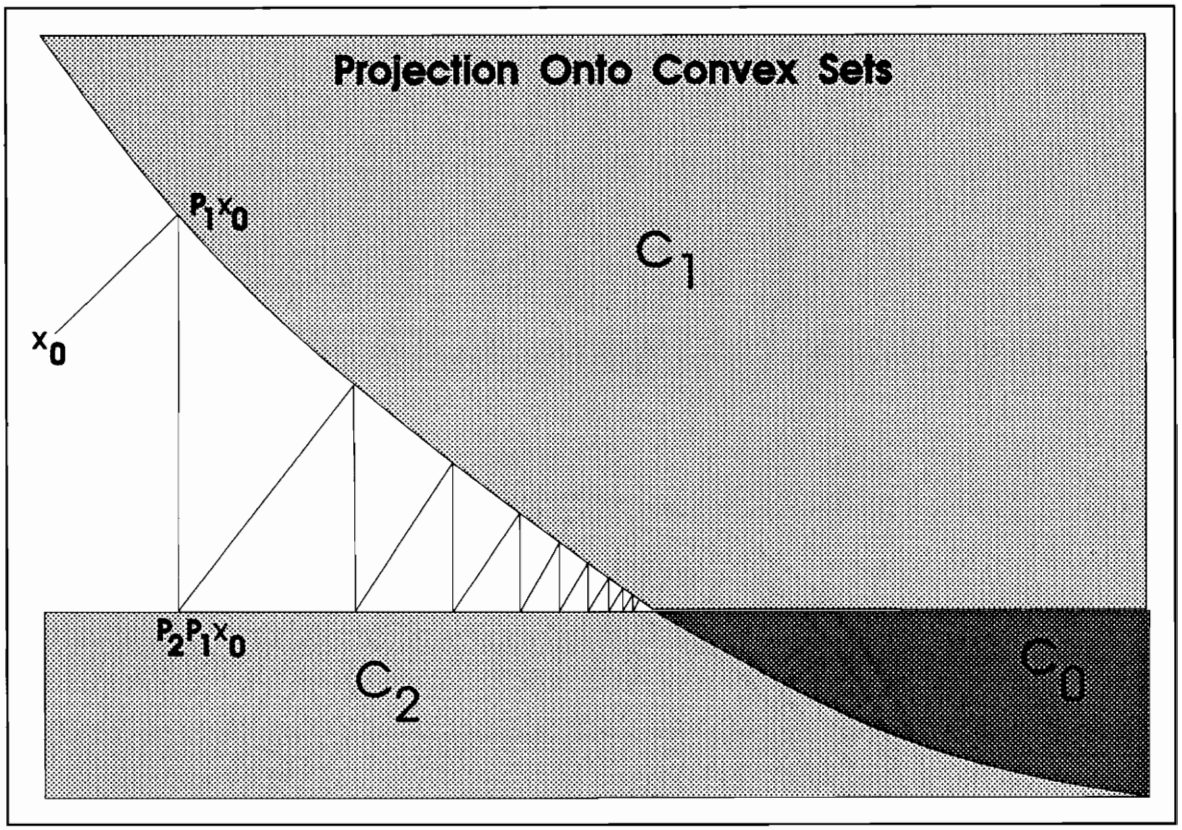


Figure 2: Projection Onto Convex Sets

3.0 Constraint Definitions and Projection Operators

Chapter 3 lists constraints which can be used with the POCS algorithm and are useful to the reconstruction of an MBR distorted signal. Our goal is to define constraints that use little *a priori* information yet severely limit the size of C_0, the intersection of all constraint sets. The smaller C_0 is the more likely we are to get a good approximation to the original signal. Ideally, we would like only one signal in C_0. In this case the reconstruction would be exact. However, this uniqueness is difficult and costly in terms of *a priori* knowledge.

The POCS algorithm is guaranteed to converge only if all of the constraints forming the composite projection operator are closed and convex. If one of the constraints is nonconvex, the algorithm is not guaranteed to converge. However, using a nonconvex constraint may significantly increase the convergence rate when convergence does occur. Convex constraints in the composite projection can improve the likelihood that the POCS algorithm will converge when the composite projection contains a nonconvex constraint. All but one of the constraints introduced in Chapter 3 are closed and convex.

The information requirements of each constraint are discussed. Each constraint requires *a priori* knowledge about the original signal which may be transmitted along with the observed signal as side information. However, some *a priori* knowledge is understood by both transmitter and receiver and does not have to be transmitted. For example, knowledge that the original signal is known to always originate from an image can be used to form a constraint that always forces all elements to be positive.

In this case, no *a priori* information needs to be transmitted because the signal is known to *always* contain positive intensity values. For purposes of computing the maximum information requirements, the intensity values of the original signal x are assumed to be in the range [0,255].

A projection operator is defined for each constraint. Also, some constraints may be applied to both the prototype signal and the original signal. If the constraint can only be applied to one or the other, it is indicated.

3.1 MBR Constraint

As alluded to in Section 2.4, the distortion operator D, may be included as a constraint as follows,

$$C = \{f : y = Df\} \tag{3.1}$$

where y is the observed signal. Thus, when a signal that belongs to C is distorted by the distortion operator D, that operation produces the observed signal. If D is singular there are an infinite number of signals that, when distorted, produce y. Let the distortion operator be D_{MBR} and compare the distortion operator in (2.12) to the projection in (2.9). Then,

$$C_{MBR} = \{f : y = P_{\mathcal{M}} f\} \tag{3.2}$$

Assuming no noise, the observed signal belongs to C_{MBR}. However, in general x_k does not belong to C_{MBR} and so the constraint in (3.2) is used.

The MBR projection [Safar 1988] is given by,

$$\begin{aligned} P_{MBR} g &= y + P_{\mathcal{M}^\perp} g \\ &= y + [I - P_{\mathcal{M}}] g \\ &= y + g - P_{\mathcal{M}} g \end{aligned} \tag{3.3}$$

where $P_{\mathcal{M}}g$ can be computed using the definition of $P_{\mathcal{M}}$ in (2.9). Also, Safar [Safar 1988] discusses a more efficient method for computing $P_{\mathcal{M}}g$. The MBR projection produces two terms, the observed signal y and the portion of g that is orthogonal to M, $P_{\mathcal{M}^\perp}g$, which is the part that must be reconstructed so that $P_{\mathcal{M}^\perp}g \approx P_{\mathcal{M}^\perp}x$. When P_{MBR} is used in the composite projection operator, it forces $P_{\mathcal{M}}x_k = P_{\mathcal{M}}x$, but the latter does not affect $P_{\mathcal{M}^\perp}x_k$ because, the projection P_{MBR}, adds $P_{\mathcal{M}^\perp}x_k$ to $y = P_{\mathcal{M}}x$. The projection operator in (3.3) requires that y be stored in memory during reconstruction.

If y is noisy, C_{MBR} may be softened to include the signal plus possible realizations of the added noise [Beex 1983, 1984]. The softened constraint may also be used when the Recursive Residual Projection (RRP) procedure is used to approximate the MBR projection, because the RRP procedure does not produce an exact projection and therefore may be regarded as a noisy version of the MBR projection.

The observed signal y is a linear combination of K vectors chosen from L, NxN unitary transforms. Thus, the observed signal may be written,

$$y = \sum_{i=1}^{L} y_{T_i} \tag{3.4}$$

where y_{T_i} is the linear combination of the J≤K basis vectors from the i^{th} transform. Also, let $f_{\mathcal{M}} = P_{\mathcal{M}}f$ where $f \in \tilde{C}_{MBR}$, i.e. f satisfies the softened MBR constraint. The projection of the reconstructed signal onto $\mathcal{M}$ can be expressed similarly,

$$f_{\mathcal{M}} = \sum_{i=1}^{L} f_{\mathcal{M}_{T_i}} \tag{3.5}$$

where $f_{\mathcal{M}_{T_i}}$ is the linear combination of the J≤K basis vectors from the i^{th} transform. The expansions in (3.4) and (3.5) allow us to place bounds on the additive noise corresponding to each participating unitary transform. Assuming the additive noise to be uniformly distributed, the soft constraint $\tilde{C}_{MBR}$ may be defined as follows,

$$\tilde{C}_{MBR} = \left\{ f : \left\| y_{T_i} - f_{m_{T_i}} \right\| \le e_i \quad \forall\, i \in [1, L] \right\} \tag{3.6}$$

where e_i is the positive square root of the constant noise power corresponding to the i^{th} participating transform.

Let $g_m = P_m g$ so that the projected signal g_m, can also be expanded similar to f_m in (3.5). Safar [Safar 1988] shows the soft projection operator $\tilde{P}_{MBR}$ to be,

$$\tilde{P}_{MBR} g = g + \sum_{i=1}^{L} s_i (y_{T_i} - g_{m_{T_i}}) \tag{3.7}$$

and,

$$s_i = \begin{cases} 0 & \left\| y_{T_i} - g_{m_{T_i}} \right\| < e_i \\ 1 - \dfrac{e_i}{\left\| y_{T_i} - g_{m_{T_i}} \right\|} & \left\| y_{T_i} - g_{m_{T_i}} \right\| \ge e_i \end{cases} \tag{3.8}$$

If $e_i = 0$ then (3.6) is equivalent to (3.2) which can be seen as follows,

$$\begin{aligned} \left\| y_{T_i} - f_{m_{T_i}} \right\| &= 0 \\ y_{T_i} &= f_{m_{T_i}} \end{aligned} \tag{3.9}$$

so that,

$$\begin{aligned} \sum_{i=1}^{L} y_{T_i} &= \sum_{j=1}^{L} f_{m_{T_j}} \\ y &= f_m \\ &= P_m f \end{aligned} \tag{3.10}$$

and (3.8) reduces to $s_i = 1$ so that (3.7) is equivalent to (3.3),

$$\begin{aligned}\tilde{P}_{MBR}g &= g + \sum_{i=1}^{L} y_{T_i} - \sum_{j=1}^{L} g_{\mathfrak{m}_{T_j}} \\ &= g + y - g_{\mathfrak{m}} \\ &= g + y - P_{\mathfrak{m}}g \\ &= P_{MBR}g\end{aligned} \tag{3.11}$$

Thus, the hard MBR constraint C_{MBR} is a special case of the soft MBR constraint $\tilde{C}_{MBR}$ where $e_i = 0$, i.e. no additive noise exists.

The MBR projection operator defined in (3.3) requires the observed signal which is a linear combination of K < N basis vectors. Thus, we must know which K vectors from a possible L basis vectors, as discussed in Section 2.1, along with their corresponding magnitudes in order to obtain the observed signal and to compute $P_{\mathfrak{m}}g$ in (3.3).

3.2 Positivity Constraint

Let the positivity constraint C_P be defined by [Youla 1982],

$$C_P = \{f : f_i \geq K \quad \forall i \in [0, N-1]\} \tag{3.12}$$

where K represents the minimum intensity value that may occur in the original signal x. The constraint C_P is used for signals derived from digitized images because the intensity values have a minimum defined for the receiver and the transmitter, and it typically equals zero. For example, a digitized image may have intensity values zero and greater whereas others may have intensity values one and greater.

The positivity projection P_P is defined as follows,

$$P_P g_i = \begin{cases} g_i & g_i \geq K \\ K & g_i < K \end{cases} \tag{3.13}$$

The constraint C_P can be enforced without transmission of any *a priori* information because K is known by both the transmitter and receiver. The positivity constraint C_P cannot be used to constrain the error signal $e_{x,k}$ because the lower bound of e_x is not constant from one signal to another.

3.3 Extremum Bound Constraint

Let the extremum bound constraint C_{EB} be defined by,

$$C_{EB} = \{f : b \le f_i \le a \quad \forall i \in [0, N-1]\} \tag{3.14}$$

where,

$$a = \max(e_x) \tag{3.15}$$

and

$$b = \min(e_x) \tag{3.16}$$

The extremum bound projection P_{EB} is performed element-wise,

$$P_{EB}g_i = \begin{cases} a & g_i > a \\ g_i & a \le g_i \le b \\ b & g_i < b \end{cases} \tag{3.17}$$

The two parameters, a and b, are required for the extremum bound projection P_{EB} and may be quantized. The error signal e_x has no bounds because the probability density function of e_x is Laplacian. However, the error signal is unlikely to have a magnitude greater than the maximum intensity value in the image. Therefore, we assume that the error signal is in the range [-256, 255] and 9 bits per N pixels are required to transmit each of the parameters.

3.4 Maximum Energy Constraint

Let the maximum energy constraint C_E be defined by,

$$C_E = \left\{ f : \sum_{i=1}^{N} |f_i|^2 \le \rho^2 \right\} \tag{3.18}$$

where,

$$\rho^2 = \sum_{i=1}^{N} |e_x^2| \tag{3.19}$$

The maximum energy projection P_E is,

$$P_E g = \begin{cases} \sqrt{\frac{\rho^2}{\|g\|^2}}\, g & \|g\|^2 > \rho^2 \\ g & \|g\|^2 \le \rho^2 \end{cases} \tag{3.20}$$

One parameter, ρ^2, must be transmitted with the signal. If we assume that each element of the error signal is in the range [-256, 255], then the square of an element of the error signal requires 16 bits. The parameter ρ^2 requires $16 + \log_2 N$ bits per N pixels.

3.5 Zero Crossing Constraint

The zero crossing constraint C_Z was introduced by Safar [Safar 1988]. It is desired to define the set of all discrete signals that have zero crossings at specified locations in the signal. A zero crossing consists of a change in sign between consecutive elements either from positive to negative or from negative to positive. Thus, a zero crossing consists of two elements of opposite sign. Let f_i be the first

element and f_{i+1} be the second where i represents the location of the first element. Then the location i zero crossing constraint $C_Z(i)$ is defined by,

$$C_Z(i) = \{f : sign(f_i) = s_i ; sign(f_{i+1}) = -s_i\} \qquad i \in [1, N-1] \tag{3.21}$$

where,

$$sign(f_i) = \begin{cases} -1 & f_i < 0 \\ 1 & f_i \geq 0 \end{cases} \tag{3.22}$$

and

$$s_i = sign(e_x)_i \tag{3.23}$$

where i is the location at which a zero crossing is to be enforced. The parameter s_i is known *a priori* and is based on the error signal because x is nonnegative and therefore does not have zero crossings. The projection for the zero crossing constraint $C_Z(i)$ is,

$$P_Z(i)g = \begin{bmatrix} g_1 \\ \vdots \\ s_i|g_i| \\ -s_i|g_{i+1}| \\ \vdots \\ g_N \end{bmatrix} \tag{3.24}$$

Each zero crossing constraint $C_Z(i)$ requires the location of an element i in the signal and the sign s_i of that element to be transmitted. The location requires $\log_2 N$ bits and the sign of the element requires 1 bit, resulting in a transmission requirement of $1 + \log_2 N$ bits per N pixels.

Let L_Z be a set of j locations where zero crossings are known to exist and let C_{Z_j} be the intersection of multiple $C_Z(i)$ constraints where each $C_Z(i)$ constraint

enforces a zero crossing at a different location $i \in L_Z$. The constraint C_{Z_j}, enforces a zero crossing at each location in L_Z,

$$C_{Z_j} = \bigcap_{i \in L_Z} C_Z(i) \qquad 1 \le j \le N-1 \tag{3.25}$$

so that C_{Z_j} enforces j zero crossings in the signal. The projection operator for C_{Z_j} is,

$$P_{Z_j} g = \prod_{i \in L_Z} P_Z(i) g \tag{3.26}$$

The signal $P_{Z_j} g$ is the same regardless of the ordering of the j individual zero crossing projection operators, $P_Z(i)$.

A signal may have up to N-1 zero crossings. If some subset of the total number of zero crossings is constrained, it is beneficial to constrain the zero crossings that are more important to reconstructing the signal. The most important zeros are the zero crossings that add new information to the reconstruction. However, it is difficult to determine analytically which zero crossings are more important than others because it will depend on a number of nonlinear constraints that are involved in the reconstruction. Two criteria for choosing a subset of zero crossing locations are defined,

Criterion 1. The j zero crossings per signal that have the largest distance between the intensity values of the two elements forming the zero crossing.

Criterion 2. The j zero crossings per signal that have the smallest distance between intensity values of the two elements forming the zero crossing.

The criteria are compared in terms of mean square error in Chapter 4.

The constraint C_{z_j}, does not prevent zero crossings from appearing elsewhere in the signal where there were none in the original signal. It simply enforces some (or all) zero crossings where they are known to exist according to *a priori* knowledge.

If C_{z_j} enforces zero crossings at two or more consecutive locations, some of the *a priori* information used to enforce C_{z_j} is redundant. This can be seen by examining the definition of C_{z_j} in (3.25) for the case of $j=2$. The *a priori* information enforces the sign at locations i and $i+1$ and the next zero crossing enforces the sign at locations $i+1$ and $i+2$. Thus, information about the sign at location $i+1$ is redundant and it would be beneficial to transmit sign information about another location, possibly $i+3$ instead of $i+1$, or to pick another zero crossing altogether.

The zero crossing constraint, C_{z_j} enforces j zero crossings so that $j(1+\log_2 N)$ bits per N pixels are required for transmission.

3.6 Sign Constraint

The zero crossing constraint $C_z(i)$, enforces the sign of the two consecutive elements that define a zero crossing. If it is desired to constrain all zero crossings it may be beneficial to transmit the sign of every element, instead of a location plus a sign, which allows every zero crossing to be enforced without transmitting the location of every zero crossing. Also, the sign constraint allows only those zero crossings that are known to have existed in the original signal.

One other drawback of the zero crossing constraint is that C_z requires location information. The locations of zero crossings in the signal cannot be predicted in advance. However, the sign of a specific element may be enforced without transmitting the location of that element. For example, we may enforce the sign of the third element in the signal. The definition of the constraint contains the location that is to be enforced.

Following this reasoning, let S_j be the set of the first j locations in g, i.e. $S_j = [1, j]$. Let C_{S_j} be defined by,

$$C_{S_j} = \left\{ f : sign(f_i) = s_i \quad \forall i \in S_j \right\} \tag{3.27}$$

If $j = N$ then the sign of every element is constrained. Let $P_S = P_{S_N}$.

The projection P_{S_j} onto the sign constraint C_{S_j} is implemented element wise as follows,

$$P_{S_j} g_i = \begin{cases} s_i |g_i| & i \in S_j \\ g_i & i \notin S_j \end{cases} \tag{3.28}$$

The projection in (3.28) requires one bit for each element g_i for which $i \in S_j$, to indicate the value of s_i. Thus, the number of bits required to implement the sign constraint C_{S_j} is the number of locations in S_j, or j bits per N pixels.

3.7 Spike Constraint

A location in a signal in which the magnitude increases significantly and then decreases down close to where the signal started is referred to as a spike. A spike is a fairly vague term and the definition below is motivated by a desire to improve on the zero-crossing constraint. Thus this is not the only way to define a spike constraint.

A spike is defined over three consecutive elements denoted $i-1, i, i+1$ where $2 \le i \le N-1$. Let i denote the location of a spike in f,

$$C_{Spk} = \left\{ f : sign(f_{i-1}) = sign(f_{i+1}) \neq sign(f_i) \quad and \quad |f_i| \ge K \right\} \tag{3.29}$$

Thus a spike is equivalent to two consecutive zero crossings with the magnitude of the middle element, f_i, constrained above a constant. In fact this definition may allow

features to be enforced that do not look like spikes. This definition is also motivated by an extension to the zero crossing constraint. The spike projection operator is,

$$P_{Spk}g_i = \begin{cases} -s_{i-1}|g_{i-1}| \\ s_i|\chi| \\ -s_{i+1}|g_{i+1}| \\ g_i \qquad\qquad \textit{otherwise} \end{cases} \tag{3.30}$$

where,

$$\chi = \begin{cases} K & |g_i| < K \\ g_i & |g_i| \geq K \end{cases} \tag{3.31}$$

The intersection of multiple spike constraints C_{Spk} where each C_{Spk} constraint enforces a spike at a different location is the set of all signals that have spikes at the specified locations. Let

$$C_{Spk_j} = \bigcap_{1}^{j} C_{Spk} \qquad 1 \leq j \leq N-1 \tag{3.32}$$

so that C_{Spk_j} enforces j spikes in the signal. The spike constraint projection operator P_{Spk_j} is,

$$P_{Spk_j} = \prod_{1}^{j} P_{Spk} \tag{3.33}$$

As with the zero crossing constraint C_Z, it is difficult to determine which spikes it is best to enforce. However, it is reasonable to assume that spikes which constrain f_i to a larger K will lower mean square error more than spikes which constrain f_i to a relatively small constant K. Therefore, spikes are selected according to the rank order of the magnitude of the middle element, with the highest magnitude selected first.

The parameter K varies for each spike that is enforced over a range of [0, 255] and so the location for each spike (at $\log_2 N$ bits/spike) and the parameter K (at 8 bits/spike) must be transmitted for each spike that is to be enforced in a given block. Thus, a total of $j(8+\log_2 N)$ bits per N pixels are required for transmission.

3.8 Minimum Absolute Deviation Constraint

Let the minimum absolute deviation constraint C_{MA} be defined by,

$$C_{MA} = \{f : |(P_S f)_i| \geq a \quad \forall i \in [1, N]\} \tag{3.34}$$

where,

$$a = \min |e_x| \tag{3.35}$$

and is known *a priori*.

The C_{MA} constraint must enforce the sign of each element because e_x contains both positive and negative elements. Consider the set,

$$C = \{f : |f_i| \geq a \quad i \in [1, N]\} \tag{3.36}$$

The constraint C is not convex. To see this, consider two signals, $f, g \in C$, each with elements $f_i \geq a$ and $g_i \leq -a$. The constraint C is convex if $f' = \mu f + (1-\mu)g$ is also an element of C for all $0 \leq \mu \leq 1$. However, $f' \notin C$ because the element $f_i' = 0$ if $\mu = \frac{g_i}{g_i - f_i}$. The parameter μ is always positive because both the numerator and denominator are negative. Also, the parameter μ is always less than 1 because $|g_i| \leq |g_i - f_i|$. Therefore, $0 \leq \mu \leq 1$ and C is not convex. Constraining the sign using the sign projection operator P_S in (3.34) prevents this case from occurring.

The minimum absolute deviation projection P_{MA} is defined by,

$$P_{MA}g_i = \begin{cases} s_i a & |(P_S g)_i| < a \\ g_i & otherwise \end{cases} \tag{3.37}$$

The constraint C_{MA}, enforces the sign plus a minimum magnitude a, for every element in the signal, so C_{MA} requires a relatively large amount of *a priori* information. The sign of every element s_i and the parameter a must be transmitted to implement the minimum absolute deviation projection operator P_{MA}. Thus, the requirement for transmission is 1 bit per pixel plus 8 bits per N pixels or $8+N$ bits per N pixels.

The information requirements of P_{MA} are relatively large. Another disadvantage of C_{MA} is that edges are missed entirely. The smallest magnitude is enforced element-wise. A signal with an edge will have both small and large magnitudes. Therefore, no information about the high valued elements is enforced. The high valued elements contribute more to the mean square error than the low valued elements.

The advantage of the minimum absolute deviation constraint C_{MA} is discussed when the histogram constraint is introduced in Section 3.12.

3.9 Minimum Increase Constraint

The minimum increase constraint C_{MI} attempts to reconstruct edge content by detecting the location of sudden increases in the signal and enforcing the increases in the reconstructed signal. The increase is forced to be larger than some minimum value, K. Let the minimum increase constraint C_{MI} be defined by,

$$C_{MI} = \{f : f_{i+1} - f_i \geq K\} \tag{3.38}$$

where the parameter K, is a constant. Let,

$$d_i = (e_x)_{i+1} - (e_x)_i \qquad i = 1,\ldots,N-1 \tag{3.39}$$

and I_{MI} be defined by,

$$I_{MI} = \{i : d_i \geq K\} \tag{3.40}$$

as the set of indices at which the increase exceeds the constant K. The minimum increase projection P_{MI}, is then defined as,

$$P_{MI} f_i = \begin{cases} f_i & i \notin I_{MI} \text{ or, } i \in I_{MI} \text{ and } f_{i+1} - f_i \geq K \\ f_i = \dfrac{f_{i+1} + f_i}{2} - \dfrac{K}{2} & i \in I_{MI} \text{ and } f_{i+1} - f_i < K \\ f_{i+1} = \dfrac{f_{i+1} + f_i}{2} + \dfrac{K}{2} & \end{cases} \tag{3.41}$$

The set I_{MI} is *a priori* information that must be transmitted, at a cost of $\log_2 N$ bits per index.

The C_{MI} constraint may be applied to blocks of a signal or to the entire signal. In either case, K may be adjusted to control the amount of *a priori* information that must be transmitted. However, if the constraint is applied to the entire signal, finding the value of K that produces the desired amount of *a priori* information is time consuming because the histogram of d_i must be computed.

A faster method is to apply C_{MI} to each block and let $K = \max_i(d_i)$ so that I_{MI} contains one location. The parameter K must then be transmitted along with the location of the maximum increase, because K varies for every signal. Although this method is faster, it may not be as effective because a signal block where the signal is almost constant uses the same amount of *a priori* information as a block for which the signal varies widely.

If the elements of e_x are in the range [-255, 256] then according to the equation in (3.39), d_i and thus K may range over [0, 512]. However, it is unlikely in practice that this will occur and the parameter K is represented over a range of [0, 255] and

thus requires 8 bits per N pixels. Thus a total of $\log_2 N$ bits to indicate a location plus 8 bits to indicate the increase are required, or a total of $8+\log_2 N$ bits per N pixels are required for transmission.

3.10 Minimum Decrease Constraint

The minimum decrease constraint C_{MD} is similar to the minimum increase constraint C_{MI}, except now decreases in the signal are enforced instead of increases. The definitions are similar to those for C_{MI} because f_i is replaced with f_{i+1} and f_{i+1} is replaced with f_i. Let the minimum decrease constraint C_{MD} be defined by,

$$C_{MD} = \{f : f_{i+1} - f_i \le K'\} \tag{3.42}$$

where K' is a constant less than zero. Let

$$d_i = (e_x)_i - (e_x)_{i+1} \qquad i = 1,\ldots,N-1 \tag{3.43}$$

then I_{MD} is defined by,

$$I_{MD} = \{i : d_i \le K'\} \tag{3.44}$$

as the set of indices at which the decrease exceeds the constant K'. The minimum decrease projection P_{MD} is then defined to be,

$$P_{MD} f_i = \begin{cases} f_i & i \notin I_{MD} \ \text{or} \ f_{i+1} - f_i \le K' \\ f_i = \dfrac{f_{i+1} + f_i}{2} - \dfrac{K'}{2} & i \in I_{MD} \ \text{and} \ f_{i+1} - f_i > K' \\ f_{i+1} = \dfrac{f_{i+1} + f_i}{2} + \dfrac{K'}{2} & \end{cases} \tag{3.45}$$

As in the minimum increase constraint C_{MI}, the parameter K' used in the C_{MD} constraint may be adjusted to control the amount of *a priori* information that must be transmitted.

The definitions are the same as those for the minimum increase constraint C_{MI} if all inequalities are flipped, i.e. greater than symbols are replaced with less than symbols and vice versa, and the constant K is made negative.

The transmission requirements for the minimum decrease constraint C_{MD} are the same as for the minimum increase constraint C_{MI} so that a total of $8+\log_2 N$ bits per N pixels are required for transmission.

3.11 Mean Constraint

Let the mean constraint C_M be defined by,

$$C_M = \left\{ f : \frac{1}{N}\sum_i f_i = \beta \right\} \tag{3.46}$$

where β is a scalar and,

$$\beta = \frac{1}{N}\sum_i (e_x)_i \tag{3.47}$$

and is transmitted with the signal. The mean projection P_M can be implemented by,

$$P_M f_j = f_j + \frac{1}{N}\left(\beta - \sum_i f_i \right) \tag{3.48}$$

The mean constraint C_M may be useful to the reconstruction process because it can reduce or increase the energy in the signal. However, $C_M \subset C_{MBR}$ and because P_{MBR} requires only the observed signal y, the mean constraint C_M is of little use.

The mean of the error signal may range over [-256, 255] so that the mean constraint C_M requires 9 bits per N pixels.

3.12 Histogram Constraint

Histogram modification is typically used as an image enhancement technique. The intensity levels of an image may be spread over a small region of the dynamic range of the image. The result is an image in which much of the detail is hidden or imperceptible. Histogram modification attempts to rescale the intensity levels evenly over the entire dynamic range, thus making details more visible. This process enhances the contrast in the image and is called *histogram equalization.*

Frequently, a distorted signal has a probability density function different from that for the original signal, and a more faithful reconstruction of the image would have a density function similar to the original density function. Furthermore, the histogram may possibly be specified with only a few parameters, making it especially easy to define the desired density function. Finally, Yenping et al. [Wang 1988] have shown that a histogram constraint can accelerate convergence in the initial iterations of an iterative reconstruction procedure by forcing the density function to fit the *a priori* density function. The process of transforming a signal with an arbitrary histogram to a signal with a specified histogram is called *histogram specification.*

Let the Laplacian histogram constraint C_H be defined as follows,

$$C_H = \left\{ f : p_f(f) = \frac{\alpha}{2} e^{-\alpha|f|} \right\} \tag{3.49}$$

where $p_f(f)$ is the probability density function of f and the Laplacian probability density function parameter α is defined in equation (2.56). The parameter α is transmitted with the signal. Thus, C_H is the set of all signals that have a Laplacian

density function. The Laplacian histogram constraint C_H is convex if and only if, $f_1, f_2 \in C_H$ and $f_3 = \mu f_1 + (1-\mu) f_2$ implies that $f_3 \in C_H$. The characteristic function of a Laplacian density function is,

$$\Psi_L(\omega) = \frac{\alpha^2}{\alpha^2 + \omega^2} \tag{3.50}$$

Let Ψ_i be the characteristic function of f_i,

$$\begin{aligned}
\Psi_3(\omega) &= E\left\{e^{j\omega f_3}\right\} \\
&= E\left\{e^{j\omega[\mu f_1 + (1-\mu) f_2]}\right\} \\
&= E\left\{e^{j\omega\mu f_1}\right\} E\left\{e^{j\omega(1-\mu) f_2}\right\} \\
&= \Psi_1(\omega\mu)\Psi_2(\omega - \omega\mu) \\
&= \frac{\alpha^2}{\alpha^2 + (\omega\mu)^2} \frac{\alpha^2}{\alpha^2 + (\omega - \omega\mu)^2} \\
&= \frac{\alpha^4}{\alpha^4 + \alpha^2\omega^2 - 2\alpha^2\omega\mu + 2\alpha^2\omega^2\mu^2 + \omega^4\mu^2 - 2\omega^3\mu^3 + \omega^4\mu^4}
\end{aligned} \tag{3.51}$$

The last equality in (3.51) cannot be reduced to the form of the characteristic function in (3.50) so that f_3 does not have a Laplacian density function; i.e. $f_3 \notin C_H$ and therefore the set of all continuous signals that have a Laplacian density function is not convex.

It is desired to find the transformation T, such that,

$$f = T(g) \tag{3.52}$$

where g is a signal with arbitrary probability density function $p_g(g)$, and f is a signal that satisfies the Laplacian histogram constraint C_H. The transformation T will not

have the properties of a projection operator because C_H is not convex. Thus, f will not be the closest signal in C_H to g and f may also not be unique.

For simplicity, the continuous case will be examined first [Woods 1981]. The results for the continuous case will then be translated to the discrete case. The correct transformation can be found by [Papoulis 1965],

$$p_f(f) = \left[p_g(g) \frac{dg}{df} \right]_{g=T^{-1}(f)} \tag{3.53}$$

The ordering of the pixel values must be preserved. That is, elements that have intensity values less than other pixels before the transformation should have intensity values less than those same pixels after the transformation. To preserve the order, T must be increasing.

If,

$$G(g) = \int_{-\infty}^{g} p_g(s) ds \tag{3.54}$$

then, from (3.53), $p_g(g) = 1;\ 0 \le g \le 1$, which is the uniform probability density function. Note that $G(g)$ in (3.54) is the cumulative distribution function of $p_g(g)$ as shown in Figure 3. Regardless of what the original density function of g is, the density function of f after the transformation in (3.54) will be the uniform probability density.

Now suppose the original density function is uniform and the desired density function is Laplacian. Figure 3 shows that the inverse transformation would be appropriate,

$$f = F^{-1}(y) \tag{3.55}$$

where F is the cumulative distribution function of f,

$$f = F_f^{-1}(y) \tag{3.56}$$

and y is the original g transformed by (3.56) to a uniform density function,

$$f = F_f^{-1}(G(g)) \tag{3.57}$$

where g is the original signal with arbitrary density function and f is a signal with the desired density function.

The transformation in (3.57) is used to convert from one continuous density function to another continuous density function. However, most applications are discrete. The transformation could be applied to every discrete point in the signal, however, this can be time-consuming even for moderately sized images. Instead, the points are quantized and then transformed.

The continuous density function can be approximated [Hummel 1975, 1977] by a piecewise constant density as follows,

$$p_g(g) = \frac{p_g(g_{Q_j})}{g_{j+1} - g_j} \qquad g_{Q_j} \in [g_j, g_{j+1}) \tag{3.58}$$

where g_{Q_i} is the point to which all intensity levels in the j^{th} bin are assigned and,

$$p_g(g_{Q_j}) = \int_{g_j}^{g_{j+1}} p_g(z)dz \tag{3.59}$$

Both the decision levels g_i, and reconstruction levels g_{Q_i}, can be computed as described in Section 2.8 where we were concerned with quantization. The objective is the same; find decision and reconstruction levels that approximate the signal g by minimizing the mean square error between g and $Q[g]$. As noted in Section 2.8 the optimal decision and reconstruction levels cannot be found in closed form. However,

given the decision levels, the optimum reconstruction level g_{Q_i} is the centroid or mean of the data in the interval $[g_j, g_{j+1})$.

The distribution function can then be approximated by the following piecewise linear function,

$$G_g(g_{Q_j}) = \frac{g_{Q_j} - g_j}{g_{j+1} - g_j} p_g(g_{Q_j}) + \sum_{i=0}^{j-1} p_g(g_{Q_i}) \tag{3.60}$$

The first term in (3.60) can be approximated, yielding

$$G_g(g_{Q_j}) = \frac{1}{2} p_g(g_{Q_j}) + \sum_{i=0}^{j-1} p_g(g_{Q_i}) \tag{3.61}$$

without significant quantization error. The transformation is then,

$$T(g_{Q_j}) = F_f^{-1}(G_g(g_{Q_j})) \tag{3.62}$$

The process of histogram specification from discrete to discrete signal is outlined in Table 1.

For the Laplacian density function in (2.54), the cumulative distribution function $F_f(f)$ is found,

$$\begin{aligned} F_f(f) &= \int_{-\infty}^{f} \frac{\alpha}{2} e^{-\alpha|s|} ds \\ &= \begin{cases} \frac{1}{2} e^{\alpha f} & f < 0 \\ 1 - \frac{1}{2} e^{-\alpha f} & f \geq 0 \end{cases} \end{aligned} \tag{3.63}$$

Finding the inverse transformation for $F_f(y)$ in (3.63), yields

$$f = \begin{cases} \frac{1}{\alpha}\ln(2F_f(f)) & f \leq .5 \\ -\frac{1}{\alpha}\ln(2-2F_f(f)) & f > .5 \end{cases} \tag{3.64}$$

The cumulative distribution is found at each quantized point using (3.63) and the new transformed quantization points are found using (3.64).

An inherent limitation of the histogram specification technique is that if there is a discontinuity in the original cumulative distribution function then the rescaled signal will also exhibit a discontinuous cumulative distribution function at the corresponding transformed point. In other words, jumps in the arbitrary density function translate into jumps in the desired density function even when the desired density function is not expected to contain any jumps.

A signal that is to be reconstructed using the Laplacian histogram constraint may have a discontinuous cumulative distribution function at zero, which would correspond to a jump in the distribution function at zero. Indeed, the error signal at iteration zero $e_{p,0}$, contains all zeros. The probability mass at zero is removed from the calculation of the distribution function and the rest of the points are transformed as discussed above and in Table 1. The probability mass at zero is not modified. Thus, the net effect is a transformation that produces an approximately symmetrical Laplacian probability density function instead of a nonsymmetrical density function that is produced if the probability mass at zero is transformed also.

The Laplacian histogram constraint C_H may cause divergence as the mean square error decreases because C_H is not convex. A tolerance is built in to remove the constraint from the iterative process before divergence occurs. The function of the histogram constraint is to add energy into the image, which for a zero mean signal will increase the variance. Thus, when the variance approaches the *a priori* variance the

histogram constraint is no longer needed and is likely to cause divergence. A tolerance that works well in practice is,

$$P_H g = \begin{cases} g & \sigma^2_{e_{p,k}} \geq 0.8\sigma^2_{e_x} \\ T(g) & \textit{otherwise} \end{cases} \tag{3.65}$$

where the iterated error signal $e_{p,k}$ is defined in Section 2.5.2.

It is important to remember that C_H is statistical in nature. For a typical image, the density function of the error signal is not exactly Laplacian, which is another cause of divergence in the reconstruction algorithm. For example, the histogram specification process attempts to make the probability density function symmetrical because the specified density function is symmetrical. In general, this will be inconsistent with the data, because there is no reason to assume that a single realization of the error signal has a perfectly symmetrical density function.

Furthermore, the projection in (3.65) contains a tolerance that is defined using the variance of the global error signal e_x. The signal g must be large enough to make this statistically-based tolerance meaningful by closely approximating the Laplacian probability density function. The tolerance will not work well for short blocks because the variance of an individual block may be below $0.8\sigma^2_{e_x}$ or above $\sigma^2_{e_x}$. If the Laplacian histogram constraint is applied to individual blocks a better tolerance may be obtained by defining a variance for each block.

The histogram constraint is advantageous because it has a global effect on the signal and it has the ability to add large amounts of energy into the signal in the early stages of the iterative reconstruction algorithm with little *a priori* knowledge. However, because the histogram is a statistical characterization, and because the constraint is not convex, the constraint may cause divergence as the error between the

original and its reconstruction decreases for any particular signal block. Therefore a tolerance is built into the constraint based on the variance of the reconstructed image.

The histogram projection P_H requires only the transmission of the variance of the error signal $\sigma_{e_x}^2$. The mean of the error signal is zero, as discussed in Section 2.7, so that the variance is the sum of the squares of every element in the image. Thus, the variance does not have a bound and may be quite large. Therefore, the variance is mapped into a finite range by the arctangent function as follows,

$$\theta = \tan^{-1}(\sigma_{e_x}^2) \tag{3.66}$$

and the parameter θ is transmitted using 16 bits to represent the range $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

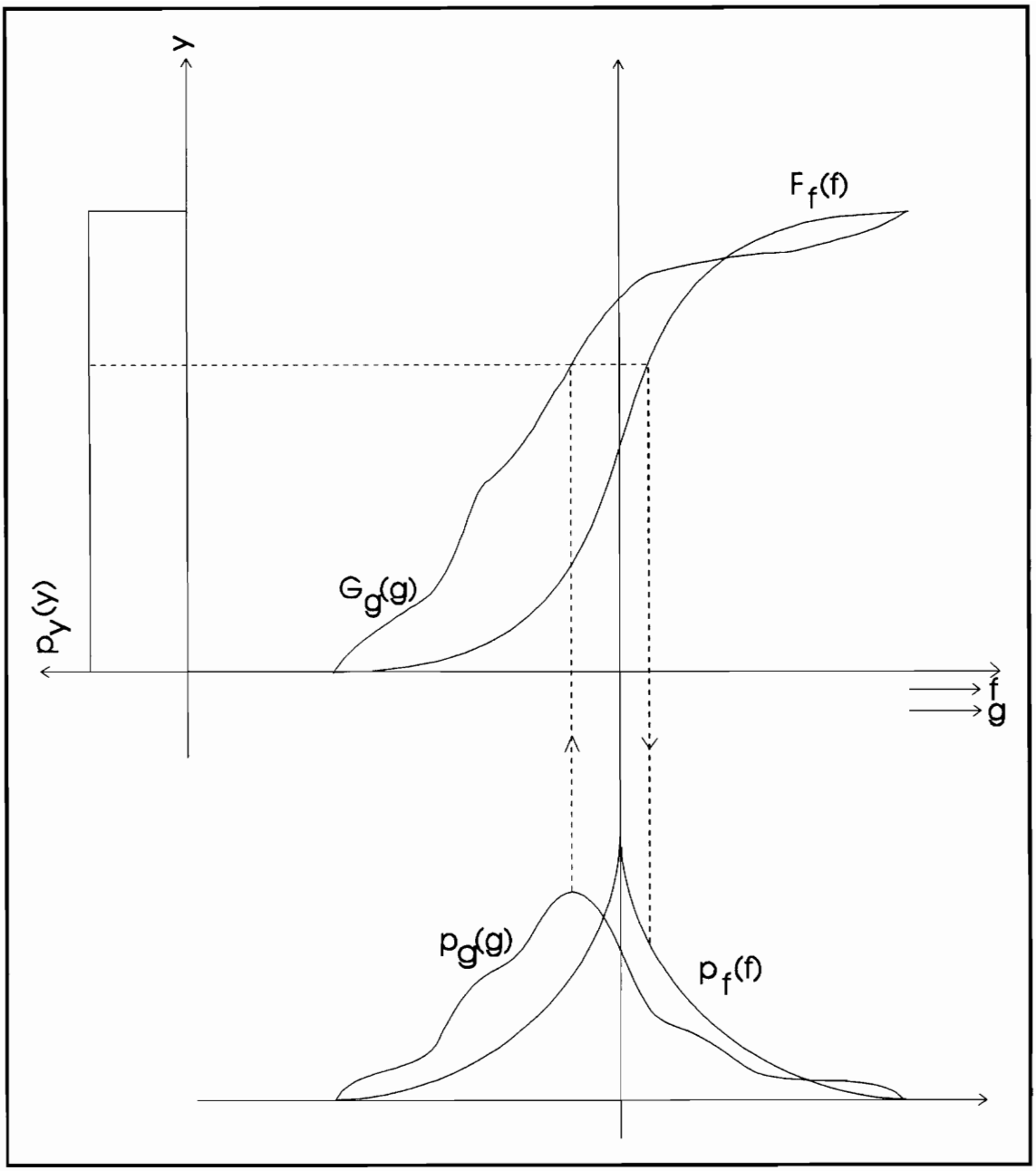


Figure 3: Histogram specification

Table 1: Histogram specification

1. For a signal g, with an arbitrary density function $p_g(g)$, calculate bins and a quantization point g_{Q_i}, for each bin. Steps 2 - 4 are applied to each bin.

2. Compute the cumulative distribution function $G_g(g_{Q_i})$ using (3.61) and normalize to 1 by dividing by the total number of elements in g.

3. Pass $G_g(g_{Q_i})$ through the precalculated inverse of the cumulative distribution function of f as in (3.62), to yield $T(g_{Q_i})$.

4. Move all points in the i^{th} bin of g to $T(g_{Q_i})$.

4.0 Simulation Results Using the Different Constraints

In this chapter, the constraints defined in Chapter 3 are compared to one another in terms of both the resulting mean square error and the image quality. It is not feasible to obtain analytical results to make these comparisons because the constraints are in general nonlinear. Furthermore, Projection Onto Convex Sets requires iteration between two or more constraints. Therefore, the effectiveness of one constraint cannot be compared to the effectiveness of another constraint directly, instead, a set of constraints must be compared to another set of constraints. Thus instead of attempting to produce analytical results we set out to obtain performance results from simulations and attempt to generalize these results to applicability to other images.

Section 4.1 compares the performance of sets of constraints in terms of ensuing mean square error and rate of convergence. In Section 4.2 the effect of different initialization vectors is discussed. The nonconvex histogram constraint is considered in Section 4.3. Finally, it is known that mean square error does not give a good indication of subjective picture quality, and therefore the last section investigates subjective image quality to a human observer.

The simulations in this chapter use a test image. A subsection of an image is converted to a one-dimensional signal by extracting successive lines of the image. The image used is a 64 x 64 subsection containing the Mandrill's left eye [Matlab]. The signal is divided into blocks of 16 pixels and the RRP procedure is simulated by selecting two MBR coefficients to approximate the signal. The compression is simulated in that no actual quantization is performed on the signal. The unitary

transforms used in the RRP procedure are the DCT and Haar transforms. The signal is distorted by approximating the MBR representation of the signal and then the POCS algorithm of Section 2.5 is used to reconstruct the signal using a specific set of constraints.

The constraints in the next section follow the order in which they were introduced in Chapter 3. Figures 4-15 and 18,19 are graphs that plot MSE (%) where,

$$MSE(\%) = \frac{MS(x - x_k)}{MS(x)} 100 \tag{4.1}$$

versus number of POCS iterations. Mean square (MS) is defined in (2.46). All reconstructions are calculated for $k = 0,...,15$ iterations where the zeroth iteration is the MSE (%) of the observed image, y.

4.1 Closed and Convex Sets with MBR and Positivity Constraints

The positivity constraint C_P requires no information to be transmitted to calculate its projection. The original signal is known to be derived from a digitized image, therefore all elements of the signal are positive and C_P can always be used to limit the number of feasible realizations. On the other hand, the distorted signal is *a priori* information that must be transmitted to implement the MBR constraint C_{MBR}. However, the type of distortion, namely D_{MBR}, is known without it having to be transmitted. Therefore, C_{MBR} and C_P may be used in every reconstruction.

If additional constraints are included in the composite projection operator, the resulting MSE will be equal to or lower than from the POCS performance resulting with the composite projection operator $P = P_{MBR} P_P$, so the latter can be used as a standard to gauge the effect of additional constraints.

The composite projection operator $P = P_{MBR}P_P$, is used in POCS and the resulting MSE performance is plotted in Figure 4. The standard composite projection, $P_{MBR}P_P$, is not very effective in reconstructing the signal, mainly because C_{MBR} does not affect $e_{x,k}$ as noted in Section 3.1. Thus, any reduction in mean square error (MSE) is via C_P which forces the relatively few negative elements in the image to be positive.

The distortion operator implemented in this simulation is D_{RRP}. This distortion operator is known to produce a noisy version of the signal that D_{MBR} would have produced. Therefore, instead of using the C_{MBR} constraint, $\tilde{C}_{MBR}$ should be used. However, at no time during these simulations did C_{MBR} cause constraint incompatibility and divergence of the POCS iteration.

4.1.1 Zero Crossing Constraint

Figures 5 and 6 each are based on POCS using $P_1 = P_{MBR}P_{Z_1}P_P$, $P_2 = P_{MBR}P_{Z_2}P_P$, and $P_3 = P_{MBR}P_{Z_3}P_P$. The composite projections in Figure 5 use criterion 1 as defined in Section 3.5 for the selection of the zero crossings and the composite projections in Figure 6 use criterion 2. Both criteria produce similar results. Criterion 1 produces better results than criterion 2, however, in Section 4.5, we will see that this is not always the case.

Each zero crossing that is enforced requires 5 bits of *a priori* information to be transmitted with the signal. Each of the P_{Z_i} are enforced per block where N=16. The constraint, C_{Z_1} enforces one zero crossing per block and requires 0.3125 bit/pixel for the projection P_{Z_1}. Likewise, P_{Z_2} and P_{Z_3} require 0.6250 bit/pixel and 0.9375 bit/pixel respectively. Enforcing over 3 zero crossings is disadvantageous because C_S enforces all zero crossings while requiring less information.

Criterion 1 not only leads to lower MSE but also enforces more signs in the error signal. As mentioned in Section 3.5, if two consecutive zero crossings are constrained, the sign of the common element is transmitted twice. Table 1 lists the

number of actual signs that are constrained using the composite projection operators above.

Table 2: Number of signs constrained by zero-crossing constraint

Composite projection	Number of signs constrained
$P = P_{MBR} P_{Z_1} P_P$ criterion 1	512
$P = P_{MBR} P_{Z_2} P_P$ criterion 1	962
$P = P_{MBR} P_{Z_3} P_P$ criterion 1	1376
$P = P_{MBR} P_{Z_1} P_P$ criterion 2	512
$P = P_{MBR} P_{Z_2} P_P$ criterion 2	943
$P = P_{MBR} P_{Z_3} P_P$ criterion 2	1348

For both $P_2 = P_{MBR} P_{Z_2} P_P$, and $P_3 = P_{MBR} P_{Z_3} P_P$, criterion 1 enforced slightly more signs than criterion 2, with the number of signs approximately 90% of the maximum that could have been enforced had the constraint excluded consecutive zero crossings.

4.1.2 Sign Constraint

Figure 7 shows the performance of POCS using $P_1 = P_{MBR} P_{S_8} P_P$, $P_2 = P_{MBR} P_{S_{12}} P_P$, and $P_3 = P_{MBR} P_S P_P$. Each of the P_{S_i} in the projections are enforced per block. POCS using the composite projection $P = P_{MBR} P_{S_8} P_P$ performs slightly better than when using $P = P_{MBR} P_P$ alone and the final MSE decreases as the number of enforced signs increases. However, the MSE does not decrease smoothly as the number of signs increases, which can be seen by comparing the results using $P = P_{MBR} P_{S_{12}} P_P$ and $P = P_{MBR} P_S P_P$. Each sign that is enforced requires 1 bit so that P_{S_8}, $P_{S_{12}}$, and P_S require 8, 12, and 16 bits/block or for a block length of 16, 0.5, 0.75, and 1 bits/pixel, respectively.

To compare POCS performance using C_S to C_Z, examine Figures 5 and 8. Each projection in Figure 8 uses the same amount of *a priori* information as the corresponding projection in Figure 5. However, for each projection in Figure 8 POCS performs better than with the corresponding projection in Figure 5 in terms of MSE. This better performance occurs for several reasons. The zero-crossing constraint C_{Z_1}, enforces 2 signs; C_{Z_2} enforces 3 or 4; and C_{Z_3} enforces 4 to 6 signs. The actual number of signs that C_{Z_2} and C_{Z_3} enforce depends on the signal and the selection criterion. The sign constraints C_{S_5}, $C_{S_{10}}$, and $C_{S_{15}}$ always enforce 5, 10, and 15 signs, respectively. Because the zero crossing constraints always enforce fewer signs than the sign constraint for the same amount of *a priori* information, the only way in which C_Z would be better than C_S is if the locations of zero crossings were more important to signal reconstruction than just sign information. The zero crossings may provide more information about the signal than just the signs but require so much more a priori information that the gain is turned into a loss of effectiveness.

4.1.3 Spike Constraint

POCS performance using composite projection operators containing C_{Spk} is shown in Figures 9 and 10. Figure 9 compares results for $P_1 = P_{MBR} P_{Spk_1} P_P$ and $P_2 = P_{MBR} P_{Spk_2} P_P$. The constraint C_{Spk} is applied to every block in the signal. Not every block has one and two spikes as defined in Section 3.7. In this simulation, approximately 85% of the blocks had a first spike and 62% had a second spike.

Comparing Figures 7 and 9, POCS using C_{Spk_1} performs much better than using C_{S_8} in terms of mean square error while using the same amount of *a priori* information. The spike constraint C_{Spk_1} places the magnitude of one element directly into the signal. This can also be seen by the difference in MSE from the zeroth to the first iteration. After the first iteration, the reconstructions converge more slowly because the energy is already in the reconstruction and is typically not removed by projections onto other

constraints. Also, the constraint C_{Spk_2} performs only slightly better than C_{Spk_1} because the secondary spikes are of less importance than the primary ones.

The spike constraint C_{Spk} can be approximated by quantizing the magnitude of the middle element. Figure 10 shows the effect of quantization of the magnitude using uniform quantization. Sixteen quantization intervals are centered at zero so that there are eight intervals above zero and eight below zero. Only the size of the interval needs to be transmitted. The continuous C_{Spk} produces MSE only slightly below the quantized C_{Spk}.

The magnitude of the middle element is quantized to 4 bits so that each spike that is enforced requires 4 bits for location and 4 bits for the magnitude, or 8 bits total. Thus, C_{Spk_1} and C_{Spk_2} require 8 and 16 bits per block or .5 and 1 bit/pixel respectively.

4.1.4 Minimum Absolute Deviation Constraint

A composite projection using C_{MA} is shown in Figure 11. POCS using the composite projection $P_1 = P_{MBR}P_{MA}P_P$ performs better than with the projection operators considered above. However, $P_1 = P_{MBR}P_{MA}P_P$ also requires much more *a priori* information than previous projection operators, well over 1 bit/pixel. The usefulness of C_{MA} is shown later when the nonconvex constraint, C_H, is considered in Section 4.2.

4.1.5 Minimum Increment and Minimum Decrement Constraints

POCS performance using composite projections using C_{MI} and C_{MD} is shown in Figure 12. For this simulation, $P_2 = P_{MBR}P_{MD}P_P$ produced a somewhat better result than $P_1 = P_{MBR}P_{MI}P_P$ in terms of MSE. When used together, in $P_3 = P_{MBR}P_{MD}P_{MI}P_P$, the two projections do very well in reducing the MSE compared to previously defined composite projections used in the POCS algorithm.

The constraints C_{MI} and C_{MD}, in P_1 and P_2 above, are implemented as block constraints so that the increase and decrease constants are required for each block. The constants are quantized to 4 bits so that P_{MI} and P_{MD} each require 4 bits for locations plus 4 bits for magnitude, or 8 bits total. For a block length N = 16, the projections, P_{MI} and P_{MD} each require 0.5 bit/pixel.

The C_{MI} and C_{MD} constraints may also be global constraints in which case the constant is transmitted once for the entire image. A global constraint constant K was chosen for the projection of P_{MI} such that approximately 256 locations were enforced. This is the same number as are enforced by the block C_{MI} constraint. Figure 13 compares global versus block constraints in which $P_1 = P_{MBR}P_{MI}P_P$ and $P_2 = P_{MBR}P_{MI}P_P$ where the P_{MI} in P_1 is a global constraint. Use of the global projection P_{MI} decreases MSE a little bit more than use of the block projection P_{MI}.

Even though use of the global C_{MI} constraint in POCS produces slightly lower MSE, it still may not be beneficial. The histogram of the increases must be computed in order to determine what the constant K must be to transmit a certain amount of *a priori* information. Furthermore, the coding of the locations in a global constraint C_{MI} takes more *a priori* information than the coding of the block C_{MI} constraint.

4.1.6 Extremum Bound and Maximum Energy Constraints

Use in the POCS algorithm of both $P = P_{MBR}P_{EB}P_P$ and $P = P_{MBR}P_EP_P$ does not help in the reconstruction of y because both C_{EB} and C_E are not used. For a constraint to be useful, the constraint must constrain the signal in some way. The signal x_k always belongs to both C_{EB} and C_E and thus the corresponding projections do not help in composite projections. However, C_{EB} and C_E will be useful in later sections where more than just closed and convex sets are considered.

4.1.7 Summary

Table 2 summarizes the MSE produced by the POCS algorithm using each of the composite projections considered in the previous sections of this chapter. Small MSE decreases are produced by the use of P_P, P_{MBR}, P_S, and P_{Z_j}. The first thirteen composite projections in Table 2 use combinations of these projection operators and decrease MSE between 5 and 8% because C_P, C_{MBR}, C_S, and C_{Z_j} do not introduce much energy into the reconstruction. Larger MSE decreases are produced by using projection operators that directly constrain the magnitude of elements in a block; P_{Spk}, P_{MI}, P_{MD}, and P_{MA}. The last eight composite projections in Table 2 use C_{Spk}, C_{MI}, C_{MD} and C_{MA} and produce MSE between 14 and 40%. Use of the minimum decrease, and minimum increase constraints produces lower MSE than the use of C_{Spk} and requires less *a priori* information than C_{MA}.

Table 3: Comparison of projections containing convex projections.

Composite Projection	MSE Decrease (%)	Bit/Pixel
$P = P_{MBR} P_P$	5.012	0
$P = P_{MBR} P_{Z_1} P_P$ (criterion 1)	5.744	0.3125
$P = P_{MBR} P_{Z_2} P_P$ (criterion 1)	6.254	0.6250
$P = P_{MBR} P_{Z_3} P_P$ (criterion 1)	6.388	0.9375
$P = P_{MBR} P_{Z_1} P_P$ (criterion 2)	5.258	0.3125
$P = P_{MBR} P_{Z_2} P_P$ (criterion 2)	5.798	0.6250
$P = P_{MBR} P_{Z_3} P_P$ (criterion 2)	6.092	0.9375
$P = P_{MBR} P_{S_8} P_P$	6.525	0.5
$P = P_{MBR} P_{S_{12}} P_P$	6.748	0.75
$P = P_{MBR} P_S P_P$	7.423	1
$P = P_{MBR} P_{S_5} P_P$	6.284	0.3125
$P = P_{MBR} P_{S_{10}} P_P$	6.539	0.6250
$P = P_{MBR} P_{S_{15}} P_P$	7.25	0.9375
$P = P_{MBR} P_{Spk_1} P_P$ (discrete)	15.27	0.5
$P = P_{MBR} P_{Spk_2} P_P$	18.51	1
$P = P_{MBR} P_{Spk_1} P_P$ (continuous)	14.72	-
$P = P_{MBR} P_{MA} P_P$	22.66	-
$P = P_{MBR} P_{MI} P_P$ (block)	22.74	0.5
$P = P_{MBR} P_{MD} P_P$	25.67	0.5
$P = P_{MBR} P_{MD} P_{MI} P_P$	39.15	1
$P = P_{MBR} P_{MI} P_P$ (global)	24.1	-

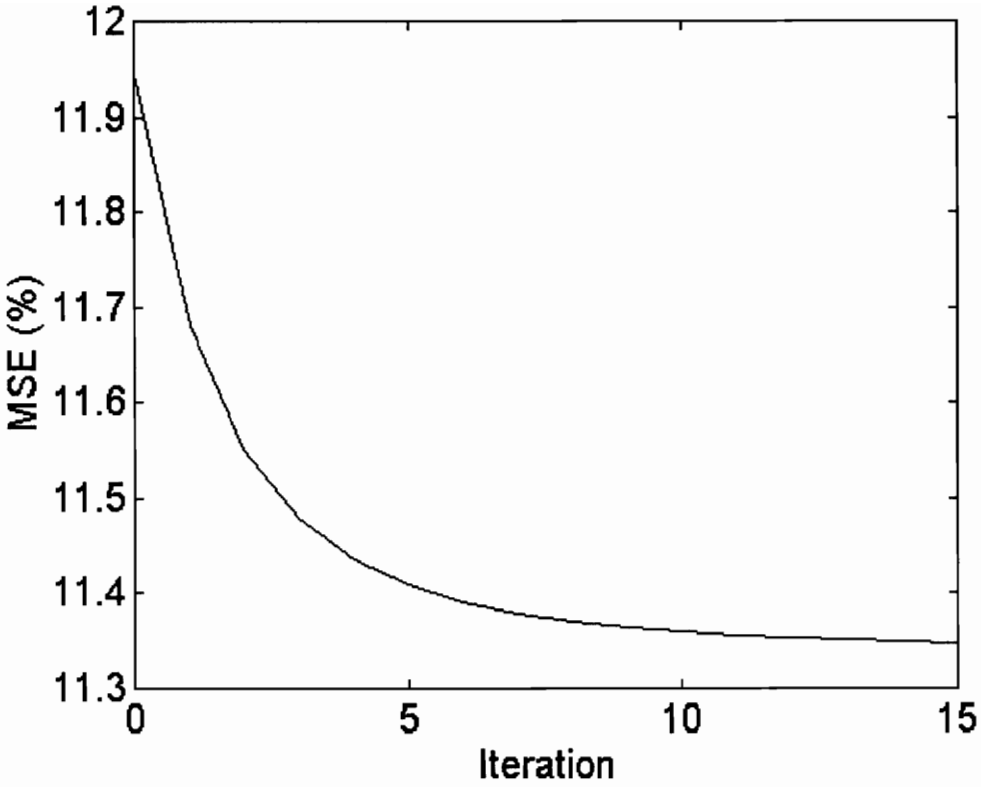


Figure 4: MSE of POCS reconstruction using, $P = P_{MBR} P_P$

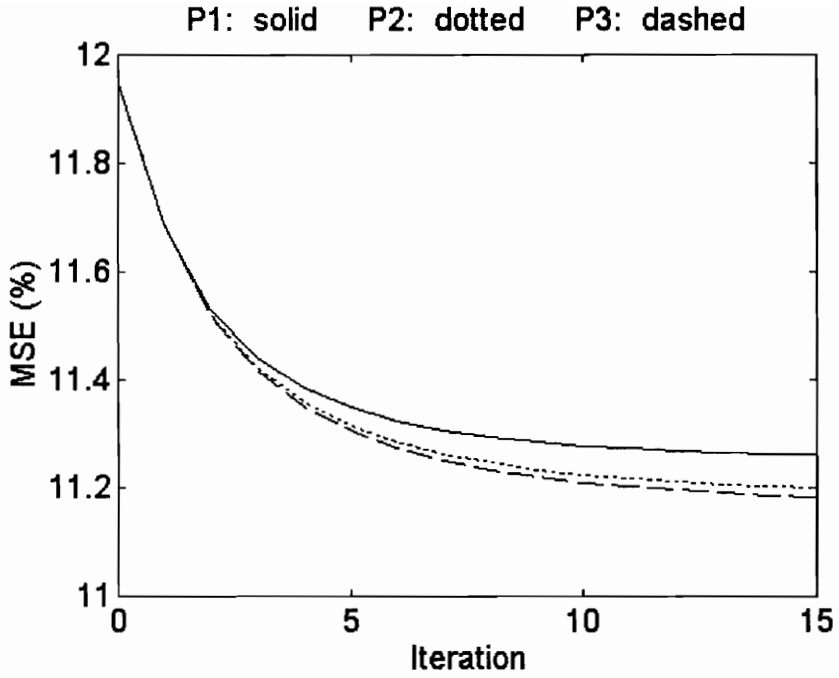


Figure 5: MSE of POCS reconstruction using zero crossing constraint, $P_1 = P_{MBR}P_{Z_1}P_P$, $P_2 = P_{MBR}P_{Z_2}P_P$, $P_3 = P_{MBR}P_{Z_3}P_P$, criterion 1

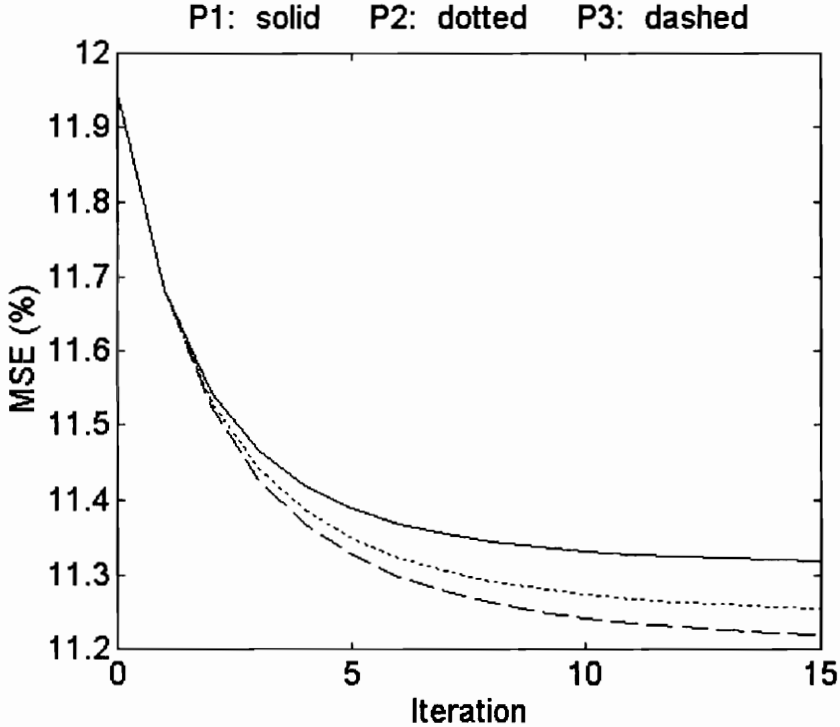


Figure 6: MSE of POCS reconstruction using, $P_1 = P_{MBR}P_{Z_1}P_P$, $P_2 = P_{MBR}P_{Z_2}P_P$, $P_3 = P_{MBR}P_{Z_3}P_P$, criterion 2

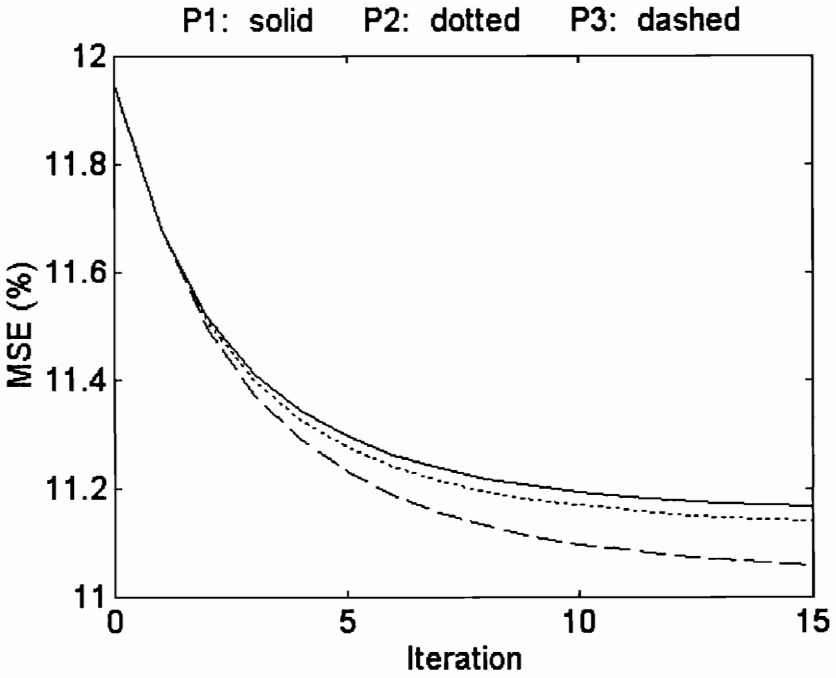


Figure 7: MSE of POCS reconstruction using, $P_1 = P_{MBR}P_{S_8}P_P$, $P_2 = P_{MBR}P_{S_{12}}P_P$, $P_3 = P_{MBR}P_SP_P$

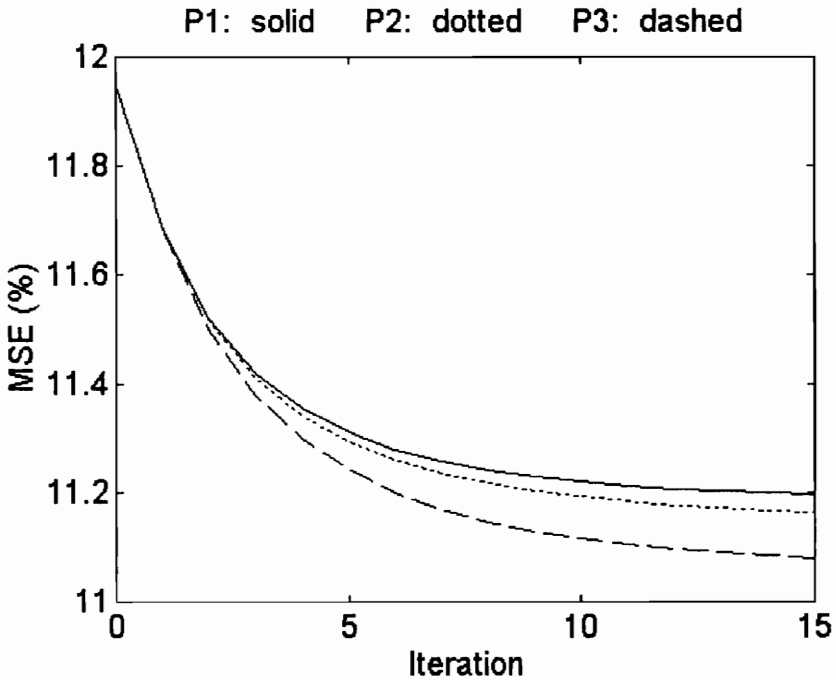


Figure 8: MSE of POCS reconstruction using, $P_1 = P_{MBR}P_{S_5}P_P$, $P_2 = P_{MBR}P_{S_{10}}P_P$, $P_3 = P_{MBR}P_{S_{15}}P_P$

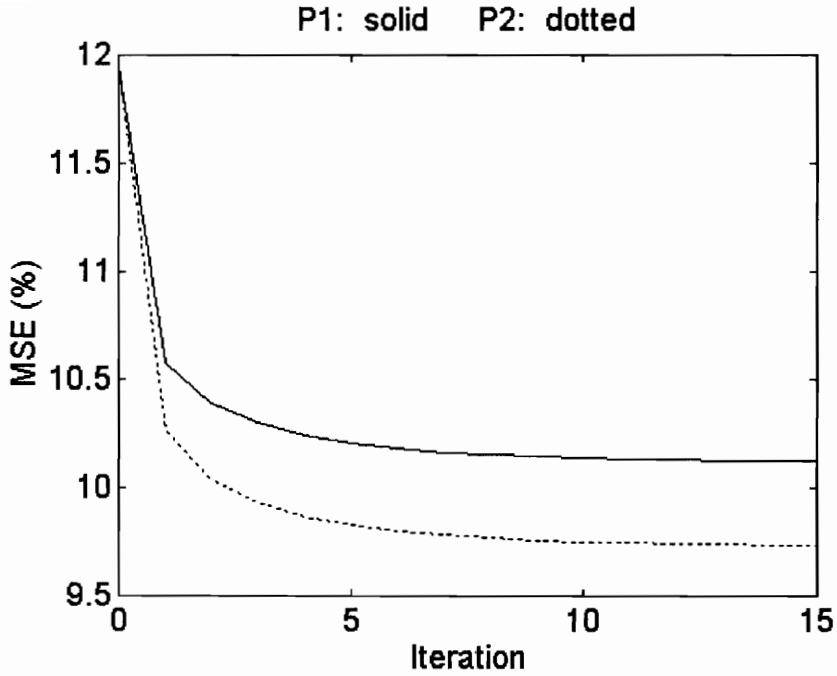


Figure 9: MSE of POCS reconstruction using,
$P_1 = P_{MBR} P_{Spk_1} P_P$, $P_2 = P_{MBR} P_{Spk_2} P_P$

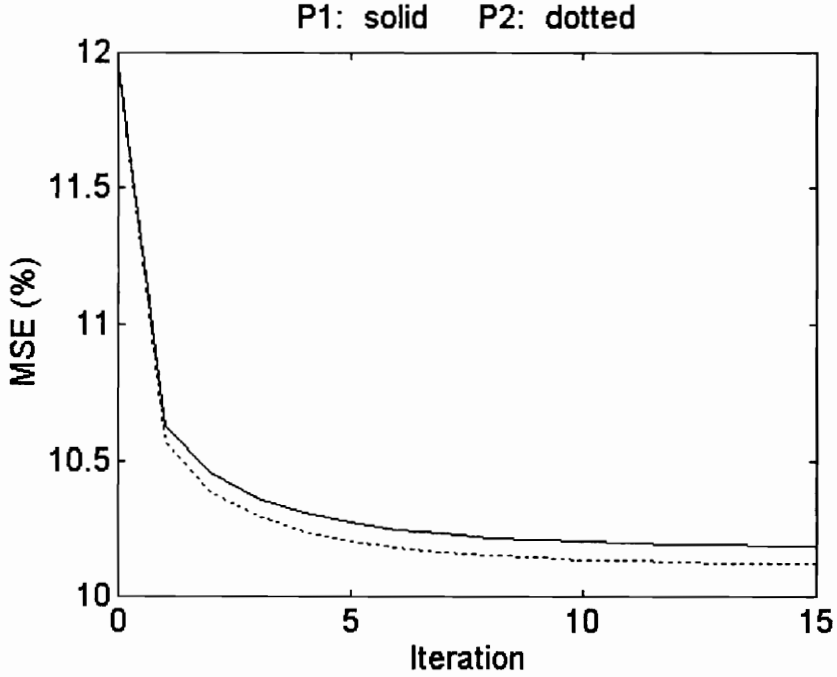


Figure 10: Comparison of continuous and quantized spike constraint,
$P_1 = P_{MBR} P_{Spk_1} P_P$ (quantized), $P_2 = P_{MBR} P_{Spk_1} P_P$ (continuous)

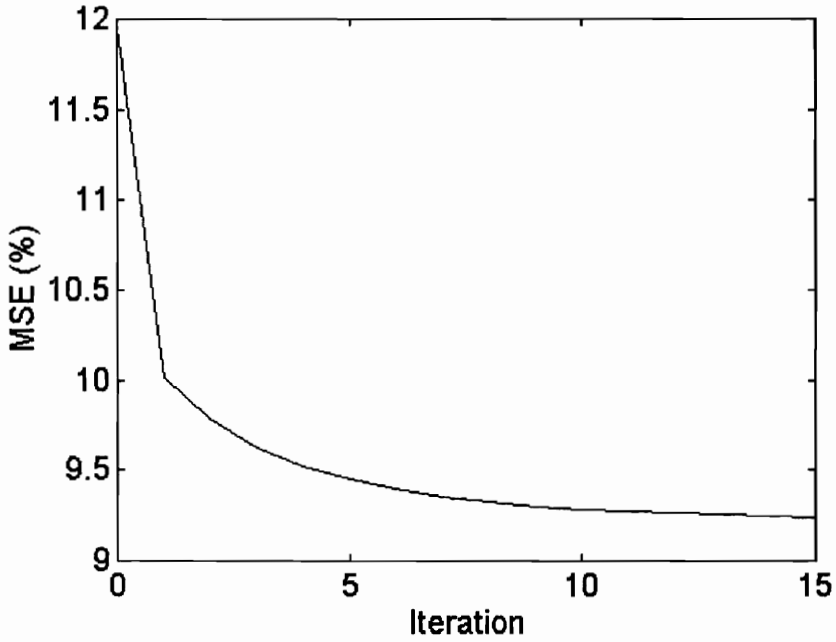


Figure 11: MSE of POCS reconstruction using, $P = P_{MBR} P_{MA} P_P$

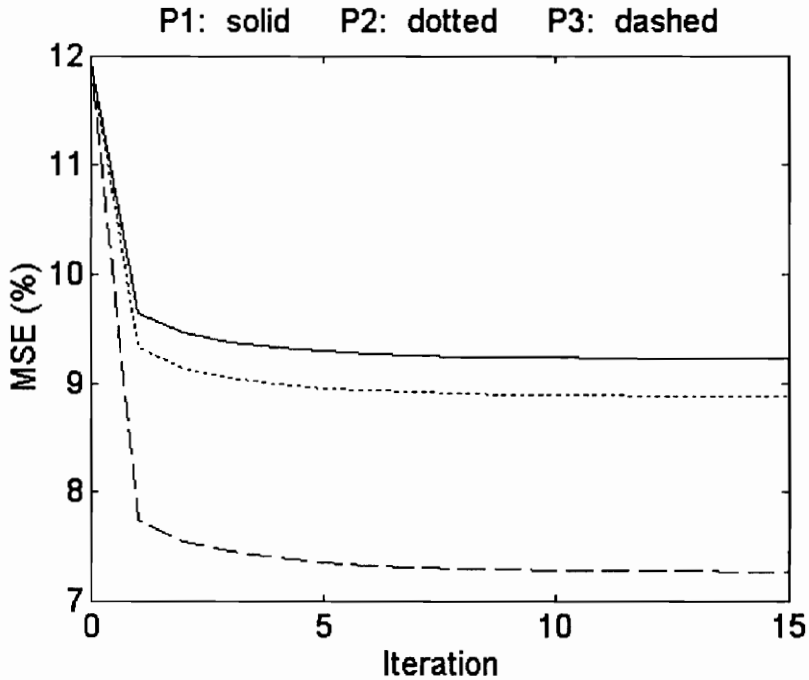


Figure 12: MSE of POCS reconstruction using, $P_1 = P_{MBR} P_{MI} P_P$, $P_2 = P_{MBR} P_{MD} P_P$, $P_3 = P_{MBR} P_{MD} P_{MI} P_P$

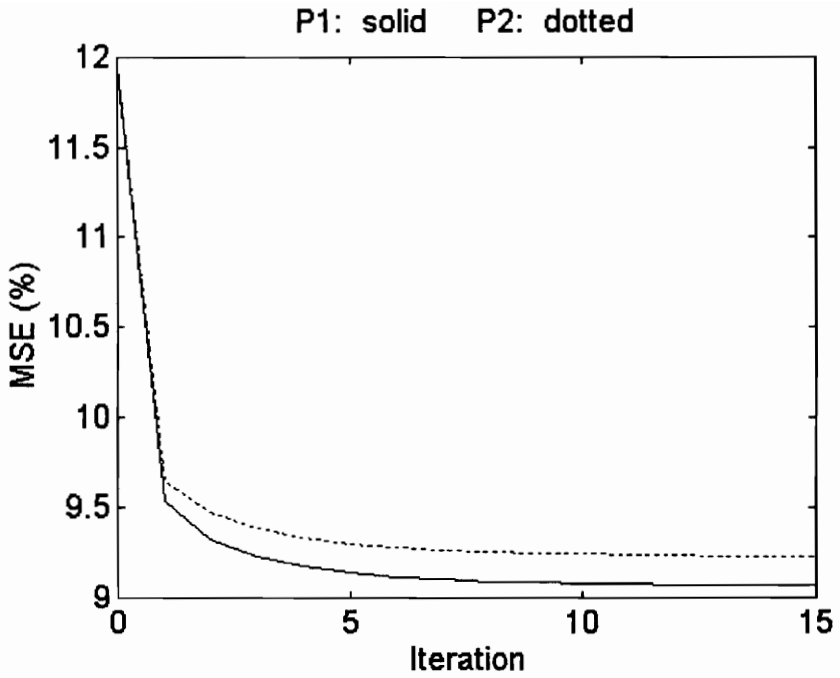


Figure 13: Comparison of MSE of block and global minimum increase constraint, $P_1 = P_{MBR} P_{MI} P_P$ (global), $P_2 = P_{MBR} P_{MI} P_P$ (block)

4.2 Initialization Vector

The initialization vector that has been used in the previous simulations is the observed signal,

$$\begin{aligned} x_0 &= y \\ &= P_{\mathcal{M}} x \end{aligned} \tag{4.2}$$

assuming no noise. With this initialization vector the error signal $e_{x,0}$, defined in (2.53) will be zero. If all the sets are closed and convex, POCS iterates to a solution. Just enough energy is introduced into the signal to cause convergence and once the signal satisfies all constraints the algorithm has converged and the signal is no longer modified. The solution will be equal to or close to the minimum norm solution.

However, there is no reason to assume that the minimum norm solution will be close to the original signal.

A reasonable effort should be made to find constraints that introduce energy back into the signal because the initialization vector is known to start with less energy than the original signal. It is difficult however, to find constraints that do not require substantial *a priori* knowledge of how much and where to add the energy back into the signal. If the initialization vector were to begin with $P_{\mathcal{M}^\perp} x_0 \neq 0$, earlier constraints that were ineffective could be used to remove the unwanted energy and to keep the desired energy that was inadvertently added to the signal.

Thus, we would like to find an initialization vector that, after reconstruction, produces an image that, after reconstruction, is closer to the original image than to the observed image. It is not important that the initialization vector itself be closer to the original image than to the observed image. The only objective is to find an initialization vector that in the end consistently produces a better reconstruction than the observed image would have produced.

It is difficult to determine without simulation which initialization vectors are advantageous because the composite projection is nonlinear. Two initialization vectors are introduced in the next two sections that produce better images after reconstruction.

4.2.1 Observed Signal Plus White Noise

Consider the following initialization vector,

$$I_n = y + n \tag{4.3}$$

where y is the observed signal and n is white noise of zero mean and variance, σ_n^2. In general, this initialization vector contains a component in $\mathcal{M}$ which is removed by C_{MBR}

during the first iteration. The projection $P_{\mathcal{M}^\perp} x_0$ is nonzero and constraints can be used to remove the unwanted portions of $P_{\mathcal{M}^\perp} x_0$. In the simulations below, $\sigma_n^2 = \sigma_{e_x}^2$.

The additive noise introduces energy into the signal. The reconstruction should remove any unwanted energy from the signal and keep the desired energy that corresponds to a feasible signal. The constraints involved in this reconstruction must be efficient at removing unwanted energy to prevent the MSE from increasing from the zeroth to the first iteration. It is important to remember that the algorithm is still guaranteed to converge to a solution. However, the constraints may not remove enough unwanted energy so that the MSE of the reconstruction is larger than the MSE of the observed signal.

The standard composite projection, consisting of the two constraints C_{MBR} and C_P, does not remove enough unwanted energy to decrease the mean square error as shown in Figure 14. Through simulations it was found that using the sign constraint C_S improved the mean square error when I_n is used for initialization. During the first iteration C_S removed enough unwanted energy from the signal that the MSE decreased. Figure 15 shows the error curves for 100 independent realizations of additive noise.

Use of the maximum energy constraint C_E, and the extremum bound constraint C_{EB} in POCS is effective in lowering the MSE even further. However, they require more *a priori* information; when C_S is used the extra bit rate exceeds 1 bit/pixel if used with C_E or C_{EB}. Use of another composite projection, $P = P_{MBR} P_{S_8} P_P$, yielded a reduced MSE for the simulation image when the initialization vector in (4.3) is used. However, the reconstruction did not converge for every test image.

It is important to examine the effect that the use of the initialization vector I_n has on individual blocks. The mean square error of the entire image, such as for example in Figure 15, does not show the change in MSE for individual blocks. Furthermore, the first iteration is our only concern, because the POCS algorithm is guaranteed to reduce MSE after the reconstruction process has begun. Figure 16 shows

the minimum and maximum MSE after the first iteration for each block, relative to the MSE at iteration zero, for the POCS reconstruction using $P_{MBR}P_SP_P$ over 100 realizations of I_n. Thus, MSE above the zero reference line corresponds to an increase in MSE, i.e. where the additive noise worsened the signal reconstruction. MSE below the zero reference line corresponds to a decrease in error, i.e. where the added noise helped the reconstruction.

In Figure 16, the order of the blocks is modified so that correlated blocks are next to one another in the plot. All of the previous simulations have produced the original signal by taking blocks from the first row of the image, then the second row, and so on. The modified original signal is produced by taking the same blocks but in a different order. The blocks are still sub-sections of a row and are transmitted down the first column, then down the next column, and so on. This produces a signal in which more of the pixels in a block are closer to more of the pixels in the adjacent block so that the plot in Figure 16 is smoother. Also, for any given MSE value in Figure 16, it is easy to find the corresponding block in the image.

4.2.2 Observed Signal Plus a Constant

Another ad hoc initialization vector is defined as follows,

$$I_K = y + K \tag{4.4}$$

where K is a vector with length N and,

$$K = -\left[\sigma_{e_x}, \sigma_{e_x}, \ldots, \sigma_{e_x}\right] \tag{4.5}$$

With all elements equal to the standard deviation of the error signal, it is ensured that a constant is added to the signal that is not too large or too small relative to the other intensity values.

Comparing the result obtained with P_2 in Figure 15 with that for P_3 in Figure 17, use of the initialization vector I_K in POCS produces considerably lower MSE and increased speed of convergence than use of the initialization vector, I_n. Furthermore, I_K requires only C_{MBR} and C_P constraints to lower the mean square error as shown in Figure 17. The variance of the signal must be transmitted for both initialization vectors. A reconstruction that uses I_n requires other constraints, such as P_S, to lower MSE, whereas using I_K requires no extra constraints to ensure that the MSE decreases.

The initialization vector I_K does not necessarily help to lower MSE as a comparison of Figure 18 to Figure 12 illustrates. The reconstruction for which the MSE is given in Figure 18 uses I_K but the performance is similar to that for the reconstruction which is initialized with the observed signal. Typically however, the use of I_K lowered the MSE of a reconstruction significantly.

Figure 19 shows the MSE after the first iteration for each block, relative to the MSE at iteration zero, for the standard projection using the initialization vector I_K. The figure shows an MSE increase for some particular blocks but generally the MSE decreases. The places in the plot where the MSE did not change correspond to where the initialization signal was added, left unmodified by the constraints and then removed by C_{MBR}, because C_{MBR} removes any bias from the error signal to maintain a zero mean.

Figure 20 shows the results obtained using $P = P_{MBR}P_SP_P$ and the initialization vector I_K. Inclusion of the sign constraint P_S in the POCS reconstruction prevents more MSE block errors from increasing as can be seen by comparing Figure 19 and Figure 20. Using $P = P_{MBR}P_SP_P$ and the initialization vector I_K to reconstruct the test image, all block errors decrease during the first iteration except for block 102 which is slightly above zero. (Other test images contain more than one block for which the MSE increases.) This is because use of the sign constraint P_S removes more unwanted energy from the initialization vector than use of the standard projection. Comparing

POCS reconstructions using $P = P_{MBR}P_SP_P$ for different initialization vectors, I_K yields lower MSE than the minimum MSE produced by I_n for approximately 70% of the blocks.

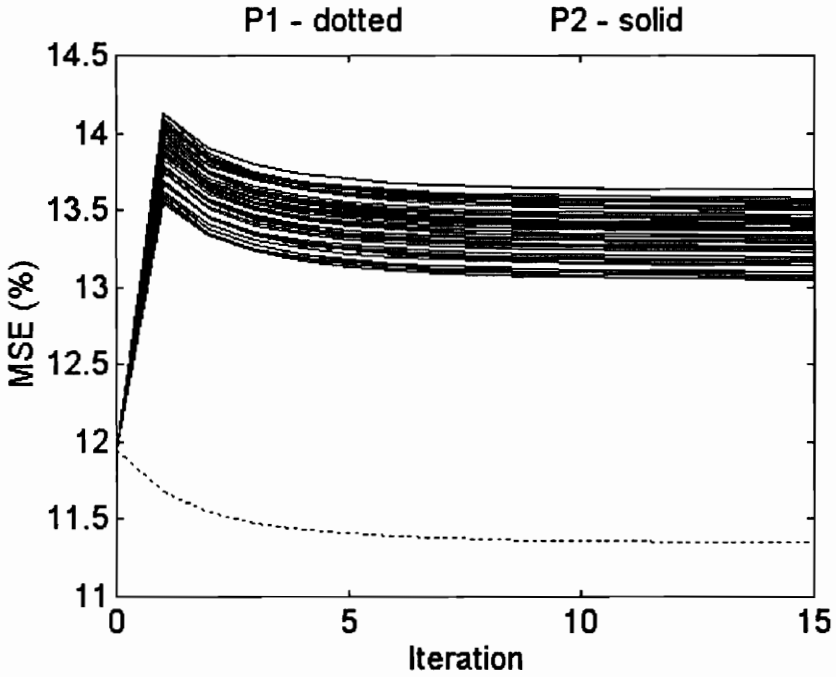


Figure 14: MSE of reconstruction (50 realizations) using, $P_1 = P_{MBR}P_P$, ($P_2 = P_{MBR}P_P$ with I_n)

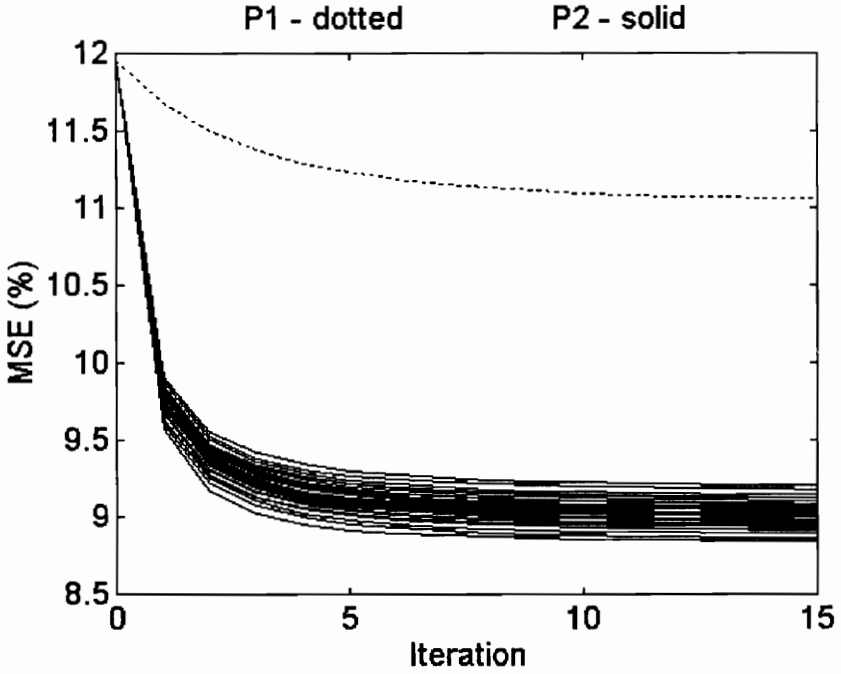


Figure 15: MSE of reconstruction (50 realizations) using, $P_1 = P_{MBR}P_SP_P$, ($P_2 = P_{MBR}P_SP_P$ with I_n)

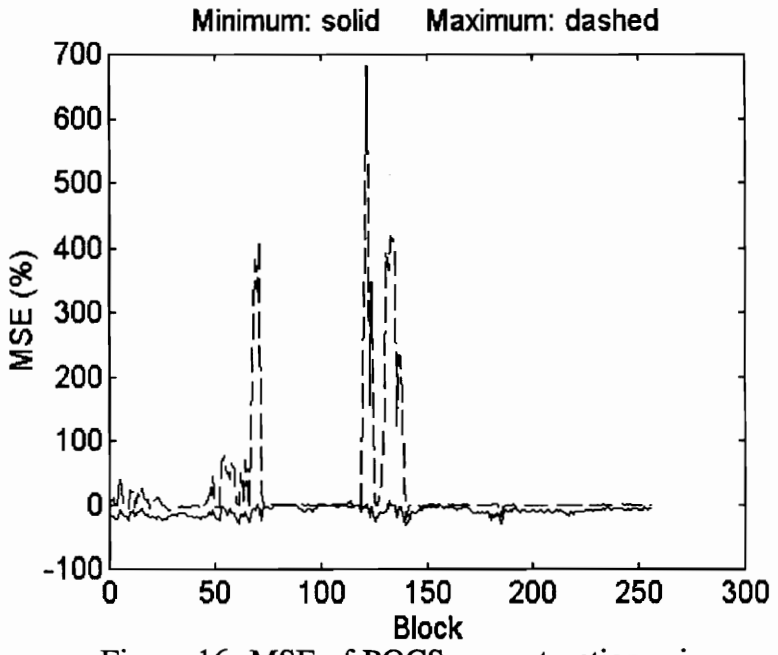


Figure 16: MSE of POCS reconstruction using, $P = P_{MBR}P_SP_P$ with I_n

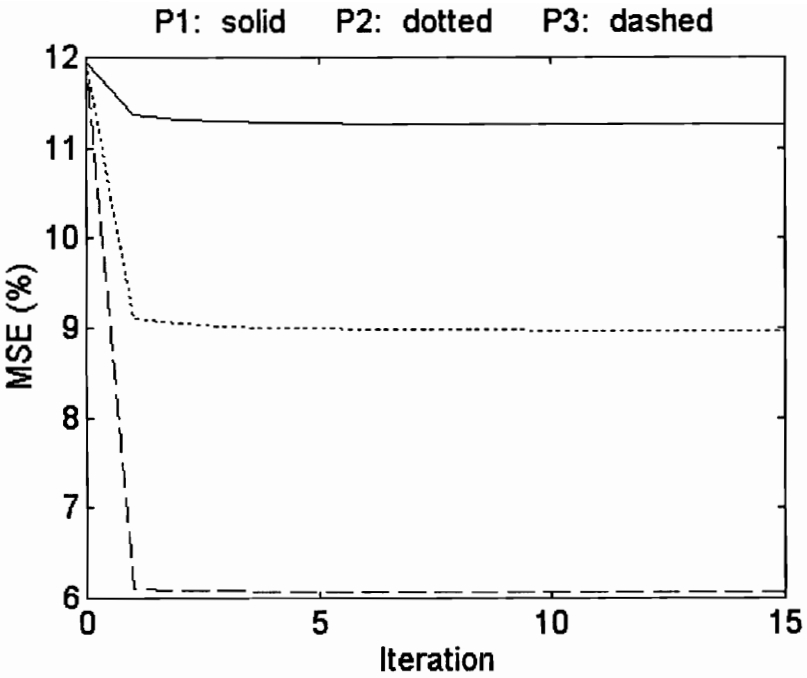


Figure 17: MSE of POCS reconstruction using, $P_1 = P_{MBR}P_P$, $P_2 = P_{MBR}P_{S_8}P_P$, $P_3 = P_{MBR}P_SP_P$ with I_K

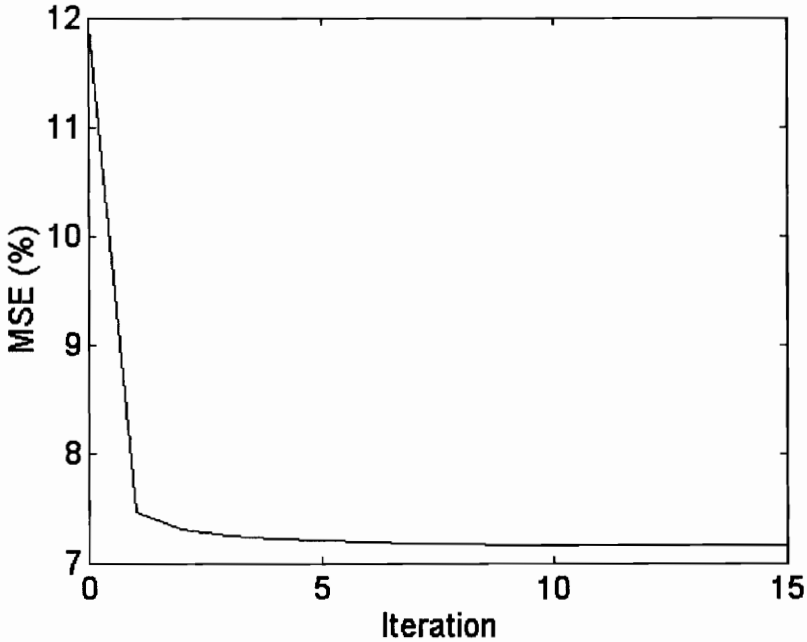


Figure 18: MSE of POCS reconstruction using, $P = P_{MBR}P_{MD}P_{MI}P_P$ with I_K

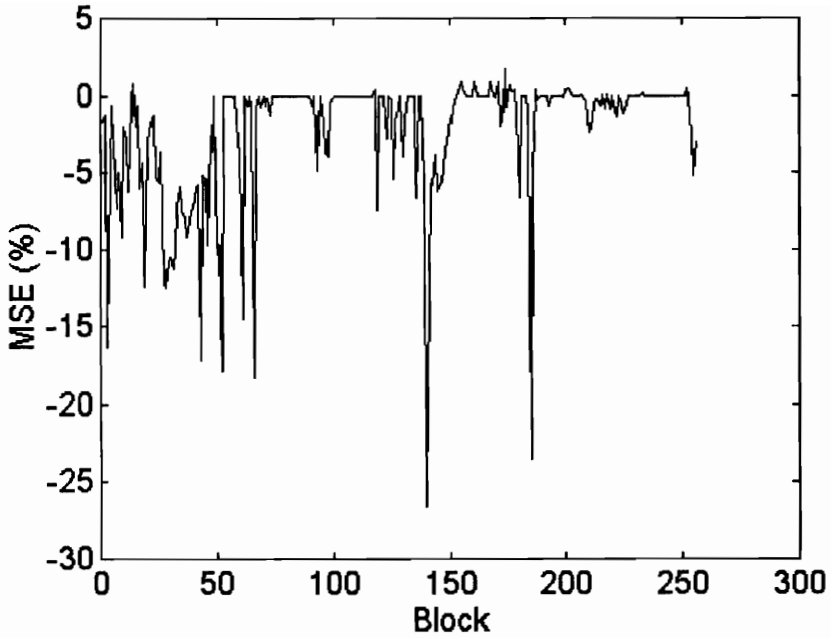


Figure 19: Difference in MSE between zeroth and first iteration, $P = P_{MBR} P_P$ with I_K

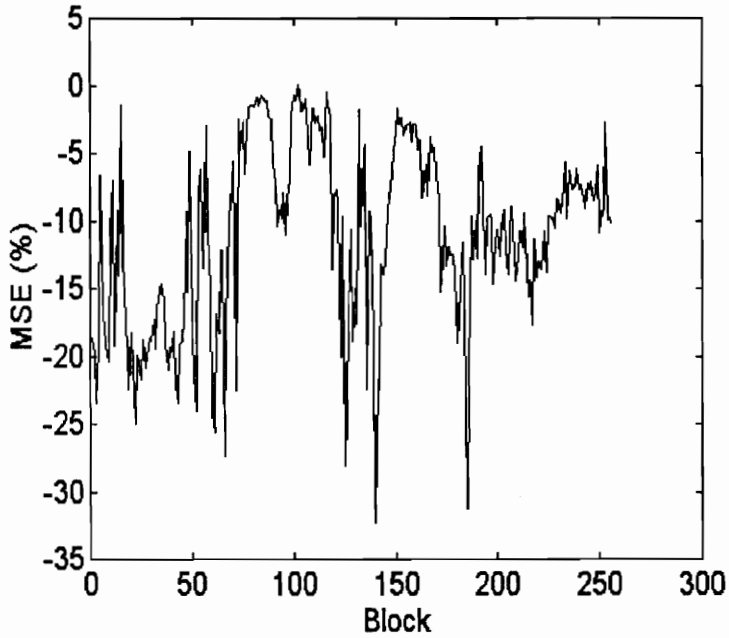


Figure 20: Difference in MSE between zeroth and first iteration, $P = P_{MBR} P_S P_P$ with I_K

4.3 Histogram Constraint

POCS is not guaranteed to converge to a solution when the histogram constraint is used in the composite projection because C_H is not a convex set. Other constraints must be included in the composite projection to limit the divergence caused by C_H and to help guarantee convergence. By giving up the guarantee of convergence it is hoped that the nonconvex constraint will provide much better reconstruction than can be obtained using just closed and convex sets.

Simulations show that the composite projection $P_H P_{MBR} P_P$ causes divergence. Other closed and convex sets will be required in the composite projection to force the iterations to converge to a feasible solution, i.e. one that satisfies all convex constraints and has an approximately Laplacian error signal.

4.3.1 Cause of Divergence

The ordering of the intensity values is of utmost importance to C_H. This ordering is not modified when P_H is applied to the signal. It is the actual intensity values that are modified so that the histogram of $e_{x,k}$ has the correct probability density function. Now, suppose that an intensity value is higher than in the original signal. The usual effect of P_H is to spread the intensity values outward, away from zero, so that when C_H is imposed that particular intensity value increases further away from its correct location instead of toward it. Likewise, intensity values that are lower in the original signal increase only slightly in intensity. However, this does not cause divergence. It is the first case that, on a larger scale, causes divergence.

Thus, to increase the chances of convergence, the ordering of the intensity values should be as accurate as possible before C_H is imposed for the first time. The objective then becomes to find out which constraints impose the correct order on the error signal

$e_{x,k}$. By the term "correct order" we mean that the set of all pixels with magnitude greater than the magnitude of the i^{th} pixel e_i in $e_{x,k}$ is equal to the set of all pixels with magnitude greater than e_i in the original error signal e_x. While it is unlikely that one or more constraints will impose the correct ordering, some constraints will do better than others.

4.3.2 Convex Constraints that Prevent Divergence

A POCS reconstruction using the projection $P_H P_{MBR} P_P$ diverges. To see why, consider Figure 21 which shows a histogram of the error signal $e_{x,k}$ after the first iteration of the POCS reconstruction using the projection $P_{MBR} P_P$. The peak of the histogram is at zero and approximately 70% of the elements of $e_{x,k}$ are identically equal to zero. The histogram constraint C_H does not affect zero elements. Futhermore, the elements that are not identically zero may not necessarily be the pixels that have the highest intensity values in the original error signal e_x, i.e. intensity values that should be rescaled to larger magnitudes. Thus, divergence will likely occur.

If x_k is projected onto the minimum absolute deviation constraint C_{MA} before C_H, then C_{MA} imposes an ordering of the error signal intensity values that will prevent divergence when C_H is imposed. Figure 22 shows the histogram after the first iteration of $P = P_{MBR} P_{MA} P_P$. Approximately 1% of the elements of $e_{x,k}$ are identically equal to zero. The variance of the error signal is still low, however, the order of the intensity values now reflects the error signal e_x more accurately. Figure 23 shows a histogram of the error signal after P_H has been applied in addition. The density function now has a Laplacian shape as well as a much higher variance. The composite projection $P = P_H P_{MBR} P_{MA} P_P$ decreases the MSE by over 50% as shown in Figure 24. Unfortunately, C_{MA} requires over 1 bit/pixel for the projection. The extremum bound constraint C_{EB} and the energy constraint C_E are useful in lowering the MSE even further, but it is undesirable to increase the already high transmission rate imposed by C_{MA}.

No other combination of convex constraints was useful in preventing divergence when C_H is used. Therefore, it is surprising that using C_S and C_H in a POCS reconstruction causes convergence when I_k is used as the initialization vector. Figure 25 shows the MSE resulting from projection $P = P_H P_{MBR} P_S P_P$ used with I_k, which produced the lowest MSE observed for any combination of constraints at slightly above 1 bit/pixel. Table 3 shows the MSE achieved by composite projections containing P_{S_8} and $P_{S_{12}}$ when combined with P_H.

The energy constraint C_E and the extremum bound constraint C_{EB} help to lower the MSE, however, they do not significantly improve the convergence properties of a composite projection that contains C_H. The MSE of the POCS reconstruction using $P = P_H P_{MBR} P_{S_8} P_{EB} P_P$ and I_k is shown in Figure 26. This projection requires the same amount of *a priori* information as the projection $P = P_H P_{MBR} P_S P_P$ but the POCS reconstruction using the projection produces a higher MSE.

The significant lowering of MSE by $P = P_H P_{MBR} P_S P_P$, using the intialization vector I_k, is not typical. In other cases, C_H was shown to increase MSE as well as significantly increasing computation time. A good example is the POCS reconstruction using $P = P_H P_{MBR} P_{S_8} P_{MI} P_P$ and the initialization vector I_k shown in Table 3 which produces an MSE decrease of 35.62% as opposed to $P = P_{MBR} P_{S_8} P_{MI} P_P$ using I_k which produces an MSE decrease of 41.11% shown in Table 3. Thus, the histogram constraint C_H must be used with caution. The simulations in Table 3 show that close to 1 bit/pixel is required to force convergence when C_H is used and still the composite projection may produce a further decrease in MSE by simply removing C_H.

Table 4: MSE decrease and bit/pixel required for projections containing C_H

Composite Projection/Initializaton Vector	MSE Decrease (%)	Bit/Pixel
$P = P_{MBR} P_S P_P,\ I_n$	24.66 (average)	1.004
$P = P_{MBR} P_P,\ I_k$	5.822	0.004
$P = P_{MBR} P_{S_8} P_P,\ I_k$	24.94	0.5004
$P = P_{MBR} P_S P_P,\ I_k$	49.24	1.004
$P = P_{MBR} P_{MD} P_{MI} P_P,\ I_k$	40.07	1.004
$P = P_{MBR} P_{S_8} P_{MI} P_P,\ I_k$	41.11	1.004
$P = P_H P_{MBR} P_{MA} P_P$	52.77	1.254
$P = P_H P_{MBR} P_S P_P,\ I_k$	61.33	1.004
$P = P_H P_{MBR} P_{S_8} P_P,\ I_k$	13.8	0.5004
$P = P_H P_{MBR} P_{S_{12}} P_P,\ I_k$	35.33	0.7504
$P = P_H P_{MBR} P_{S_8} P_{MI} P_P,\ I_k$	35.62	1.004
$P = P_H P_{MBR} P_S P_{EB} P_P,\ I_k$	66.38	1.504
$P = P_H P_{MBR} P_{S_8} P_{EB} P_P,\ I_k$	20.83	1.004

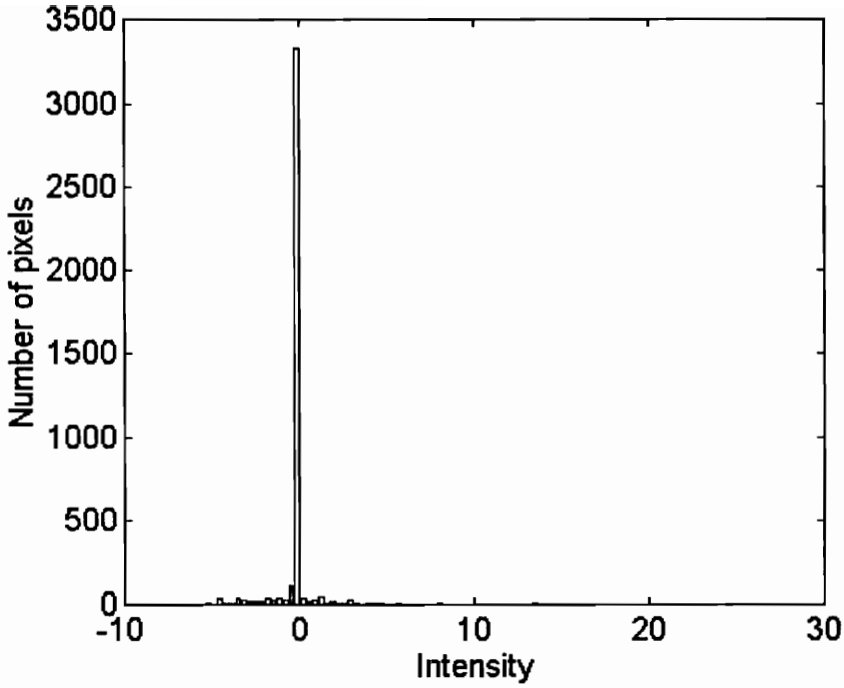


Figure 21: Histogram of $e_{x,k}$ after the first iteration using, $P = P_{MBR} P_P$

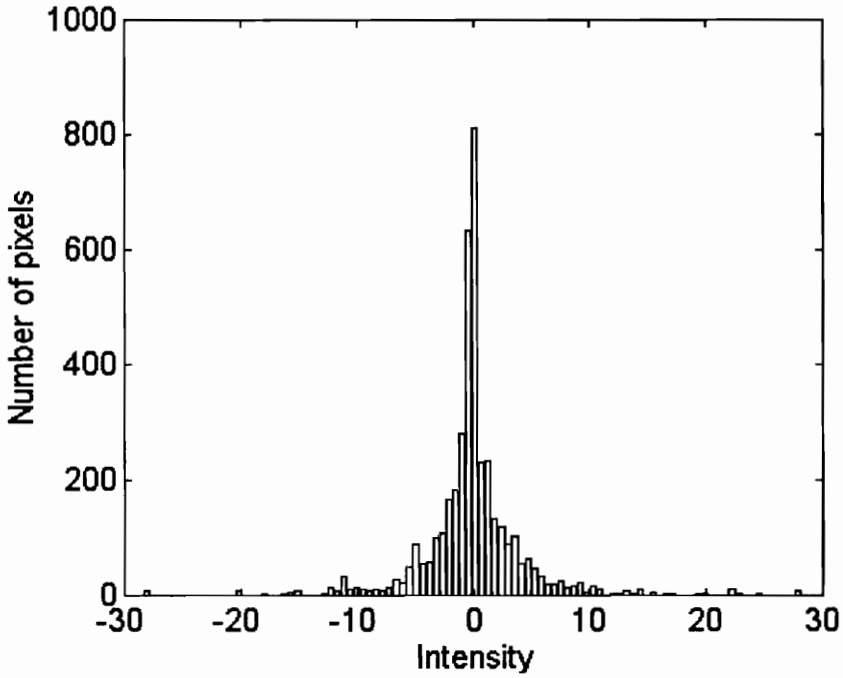


Figure 22: Histogram of $e_{x,k}$ after the first iteration using, $P = P_{MBR} P_{MA} P_P$

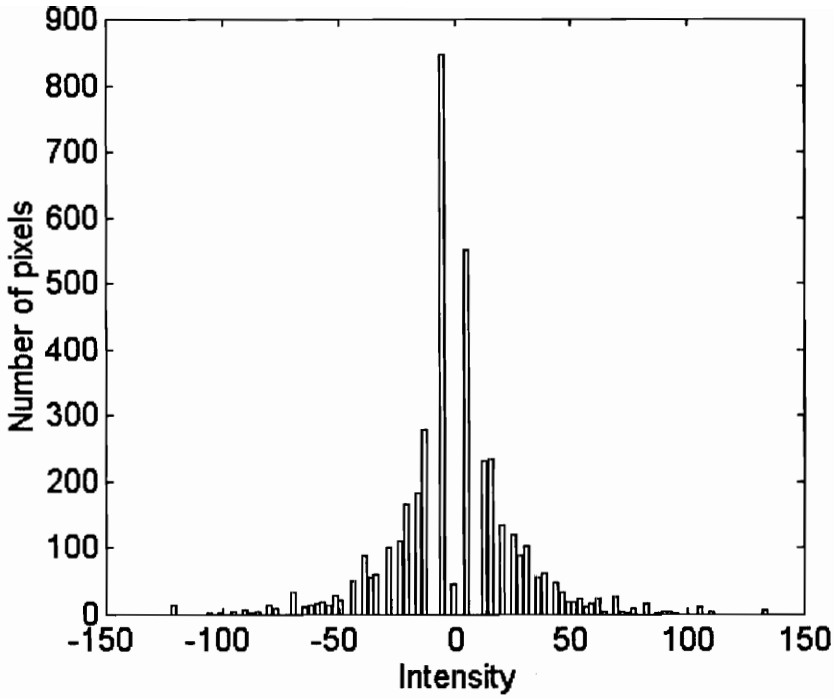


Figure 23: Histogram of $e_{x,k}$ after the first iteration using, $P = P_H P_{MBR} P_{MA} P_P$

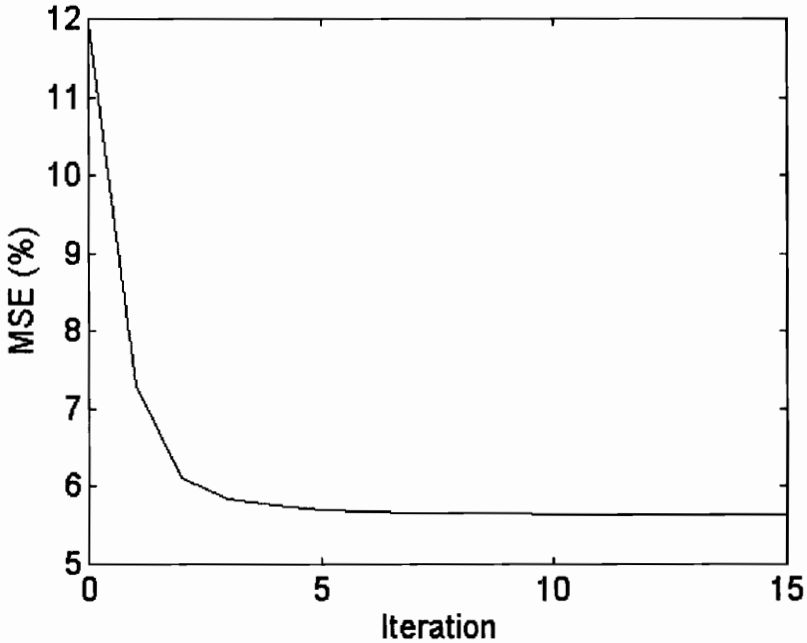


Figure 24: MSE of POCS reconstruction using, $P = P_H P_{MBR} P_{MA} P_P$

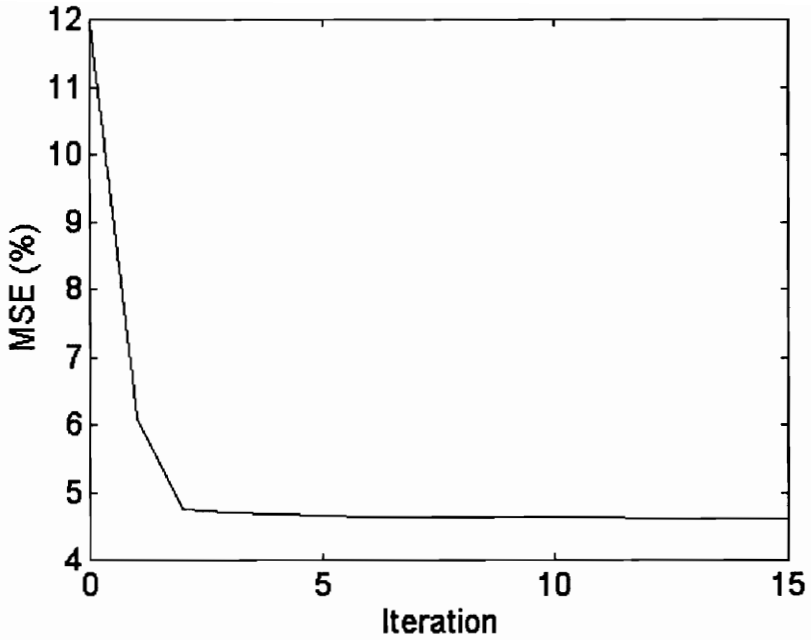


Figure 25: MSE of POCS reconstruction using, $P = P_H P_{MBR} P_S P_P$ and I_k

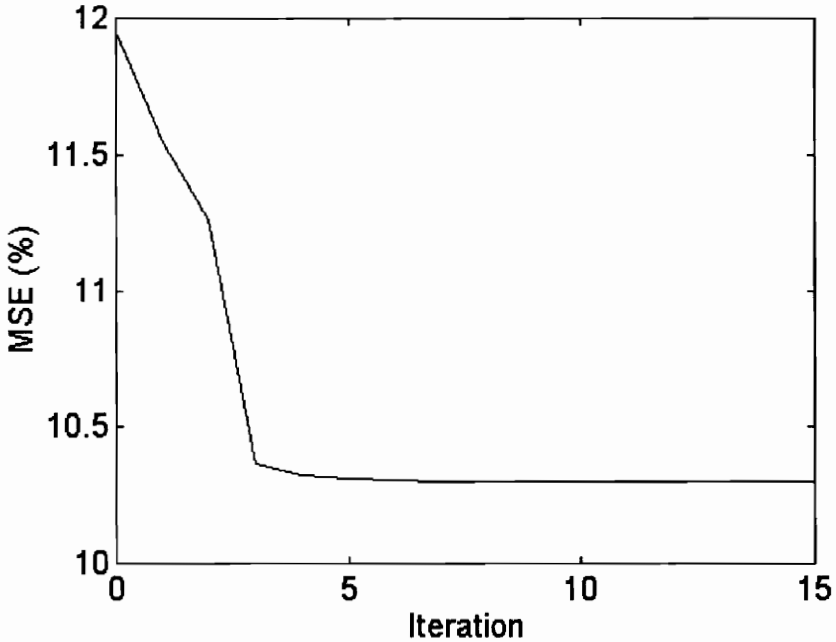


Figure 26: MSE of POCS reconstruction using, $P = P_H P_{MBR} P_{S_8} P_{EB} P_P$ and I_k

4.4 Image Quality

Just as important as reduction in mean square error is a subjective improvement in image quality. The end user will care little that the mean square error is reduced by 50% if details of the image are difficult to discern. Each constraint described above imposes distinct characteristics on the image so it is important to consider image quality obtained by using particular constraints.

The test image and distortion are the same as described earlier. Figures 27 and 28 are POCS reconstructions using the composite projection listed under each plot. If an initialization vector other than the observed image is used, that initialization vector is also listed under each plot. Each reconstruction is stopped after the fifteenth iteration. The same color map, which is a smooth transition from black to white, is shown above each plot.

The distorted image is shown in the lower middle. The distorted image contains a noticeable number of negative intensity pixels. Plotting images, Matlab represents negative intensity values as maximum intensity values. As a result both negative and relatively high intensity values are displayed using white. A good amount of blockiness shows up as well. The intensity values of the undistorted image range in integers from 1 to 220. Also, the original image is shown in the upper middle for comparison purposes.

The composite projections are divided into two groups. Figure 27 shows the first group of reconstructions that require 1 bit/pixel of *a priori* information or less. The second group shown in Figure 28 requires more than 1 bit/pixel of *a priori* information.

4.4.1 1 bit/pixel or Less Required for a Composite Projection

The minimal reconstruction in Figure 27 (a) produces a reduction in MSE of approximately 5%. It contains a bit of black spots that have replaced the white artifacts caused by negative pixel intensity. The positivity constraint C_P forces these white artifacts to an intensity value of 1 (in Matlab) and the constraints used in the reconstruction are not enough to move the low intensity values away from the minimum so that they remain at an intensity value close to 1. Little of the blockiness has been removed and a reduction of 5.0204% in MSE is barely noticeable except for the removal of the white artifacts. However, no *a priori* information needs to be transmitted for a POCS reconstruction using $P_{MBR}P_P$ so it may be performed for every distorted image.

The black spots are reduced a little by repeating the previous reconstruction with a different initialization vector, the result of which is shown in Figure 27 (b). The black spots still remain but have become smaller. The reduction in MSE is 5.8245% and there is little noticeable difference between Figures 27 (a) and (b). The variance of e_x must be transmitted for this reconstruction which is quantized to 16 bits as described in Section 3.12. The variance is transmitted once for the entire image so that the transmission requirement is 16 bits per 4096 pixels.

The sign constraint C_S helps little in the reconstruction shown in Figure 27 (c) even though C_S uses 1 bit/pixel in its projection. As mentioned earlier, energy must be added to y to obtain x. Using C_S does not add energy; instead it can only remove energy. The signal is not modified unless the sign is incorrect, in which case that element is set to zero. Both the MBR constraint C_{MBR} and the positivity constraint C_P can add energy into the signal x_k. However, little energy is introduced because the observed signal nearly satisfies both constraints. The positivity constraint C_P is the only constraint that actually introduces some energy into the signal because 158 of 4096 intensity values which are negative are set to 1 by C_P (in Matlab). Indeed, the sets formed by C_{MBR} and

C_S actually contain y. Thus $P = P_{MBR}P_S$ will not improve y at all, because y satisfies both constraints already. The slight decrease in error is because only small amounts of energy are being introduced into the signal initially by C_P and later by C_P and C_{MBR}.

As defined in Section 3.2, the C_P constraint enforces a minimum K, in the signal x_k. Normally the constant K is defined to be zero and C_P only removes energy. However, Matlab defines the minimum possible intensity value to be one. Thus, the constant K equals one and energy is introduced into the signal if an element of x_k has an absolute value less than one.

The composite projection shown in Figure 27 (d) improves image quality by enforcing spikes in the signal. Much of the reduction in mean square error is due to intensity values being directly replaced in the signal. Some of the blockiness seen in earlier results seems to have been removed and the application of the spike constraint gives the appearance of more detail in the reconstruction, especially in the upper third of the image. Not all blocks contained a spike and still fewer contained two spikes. Thresholds may be set so that only information about those blocks that have spikes is transmitted.

The reconstruction in Figure 27 (e) shows the effect of using the minimum increase constraint C_{MI} in the reconstruction. Edges are improved around the eye, however, they are too thin to improve subjective image quality. Still, the eye is more clearly defined. Much of the decrease in mean square error is due to intensity values being directly replaced. Some blockiness still remains in the image.

The reconstructions in Figures 27 (f) and (g) show the image quality produced by the sign constraints C_{S_8} and $C_{S_{12}}$, respectively, with the standard deviation of e_x added to the observed image y. The mean square error significantly decreased compared to the same POCS reconstruction using the intialization vector $x_0 = y$. Comparing the reconstructions shown in Figures 27 (c) and (f), the image quality in Figure 27 (f) is also much better than that in Figure 27 (c) with approximately half the *a priori* information.

Much of the blockiness has been removed with these two reconstructions. However, the edges are still only slightly improved.

The reconstruction in Figure 27 (h) has good edges due to the minimum increase constraint C_{MI} and the minimum decrease constraint C_{MD}. These two constraints together use 1 bit/pixel and significantly help in reconstructing the edges around the eye. Some blockiness is still present however.

4.4.2 More than 1 bit/pixel Required for the Composite Projection

All the images in this section require over 1 bit/pixel of *a priori* information so it is expected that they will do better in terms of MSE and image quality. Typically, if the initialization vector is the same, this is the case, but not always, as shown in the reconstruction in Figure 28 (a) which gives the effect of the minimum absolute deviation constraint C_{MA}. Relative to the MSE of the previous POCS reconstructions, the MSE of the POCS reconstruction using $P_{MBR}P_{MA}P_P$ is high given the amount of *a priori* information that must be transmitted and this is reflected in the image quality. Blockiness is evident and there is little edge reconstruction.

The reconstruction shown in Figure 28 (b) shows the effect of the initialization vector I_n with a composite projection operator containing C_S. The image quality is poor in places due to the initialization vector, however, the blockiness has been reduced substantially. For the most part the image quality is acceptable except in low intensity regions where the added white noise has a more visible effect. The *a priori* information requires approximately 1.004 bit/pixel; 1 bit/pixel to implement the sign constraint C_S plus 16 bits to transmit the variance of the error signal, $\sigma^2_{e_x}$. Possibly the extremum bound constraint C_{EB} would remove this noisy effect, however, C_S is required because I_n is used. Therefore, the transmission requirements will never go below 1 bit/pixel because C_S is required.

The composite projection $P_{MBR}P_{MD}P_{MI}P_P$ with the initialization vector I_k, produces the result shown in Figure 28 (c), which can be compared to the result in Figure 27 (h), where the observed image y was used for the initialization. Both the MSE decrease and the image quality of the reconstruction shown in Figure 28 (c) are similar to those for the reconstruction in Figure 27 (b). However, the dark spots are now smaller in size. Blockiness is somewhat visible even with a relatively large decrease in mean square error.

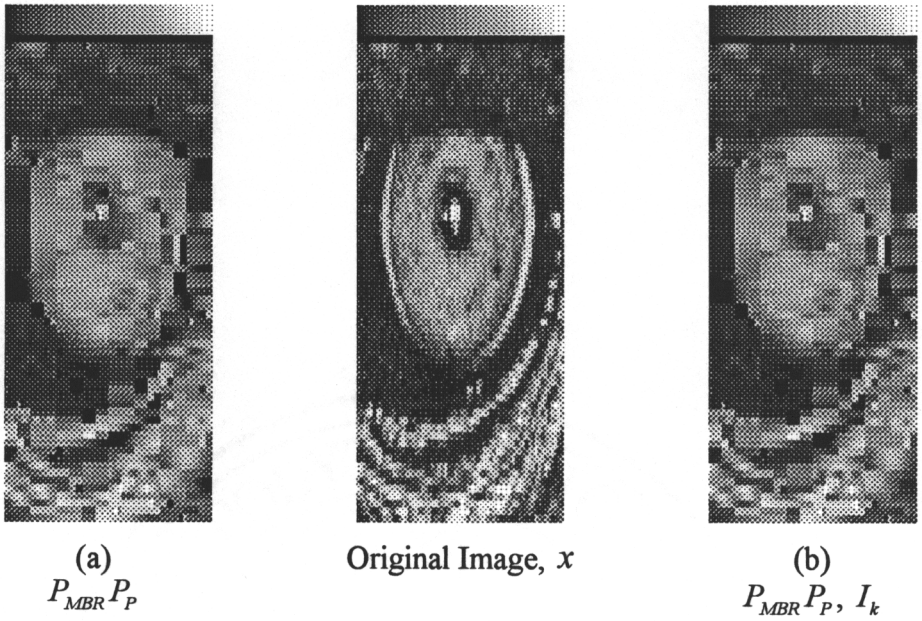
The composite projection shown in Figure 27 (c) is used again for the reconstruction in Figure 28 (d), now with the intialization vector I_k instead of the observed image y. The initialization vector I_k improves the image quality and lowers MSE significantly. Most of the extraneous black spots have been removed. The upper third of the image is still a bit noisy; this is caused by not having a constraint to remove the excess energy introduced by the initialization vector I_k. Even edge content has increased a bit.

The MSE decrease obtained by $P = P_{MBR}P_{MA}P_P$ is almost doubled by the reconstruction represented in Figure 28 (e), which uses $P = P_H P_{MBR}P_{MA}P_P$. The number and size of the black spots over the low intensity regions of the image have increased. However, the reconstruction is improved by the Laplacian histogram constraint as can be seen by comparing Figure 28 (e) with Figure 28 (a). The blockiness has been removed and edges are clearer.

The reconstruction producing Figure 28 (f) uses the composite projection $P = P_H P_{MBR}P_S P_P$, initialized by I_k, which is C_H plus the projection that produced the reconstruction in Figure 28 (d). This combination of constraints performed best in lowering MSE with an *a priori* information requirement slightly over 1 bit/pixel. The reconstruction in Figure 28 (f) also has good image quality. The edges around the eye are clearer than in other reconstructions except those using C_{MI} and C_{MD}.

The reconstruction shown in Figure 28 (g) uses the composite projection $P = P_{MBR}P_{S_8}P_{MI}P_P$ with initialization vector I_k. The image quality is relatively good in that

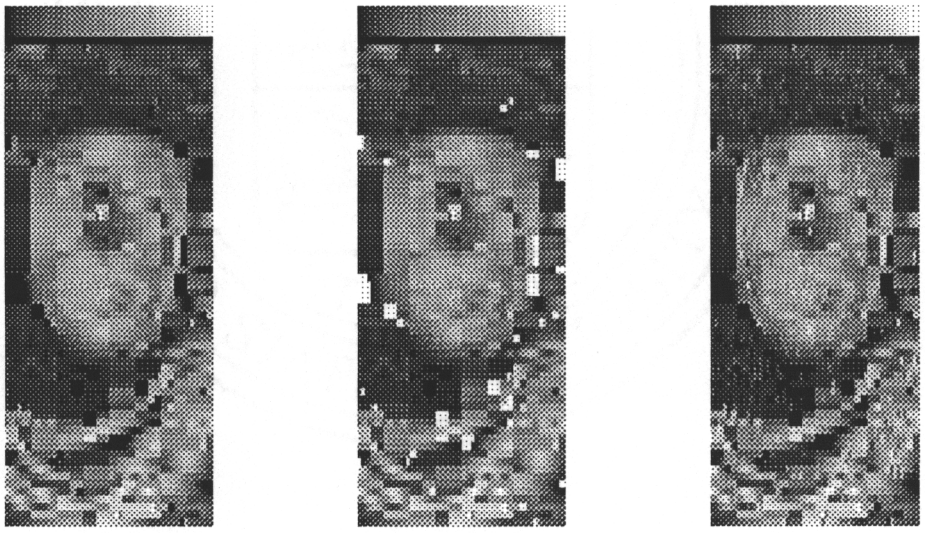
the blockiness is removed and there are edges around the eye. Many of the black spots have also been removed.



(a)
$P_{MBR}P_P$

Original Image, x

(b)
$P_{MBR}P_P$, I_k

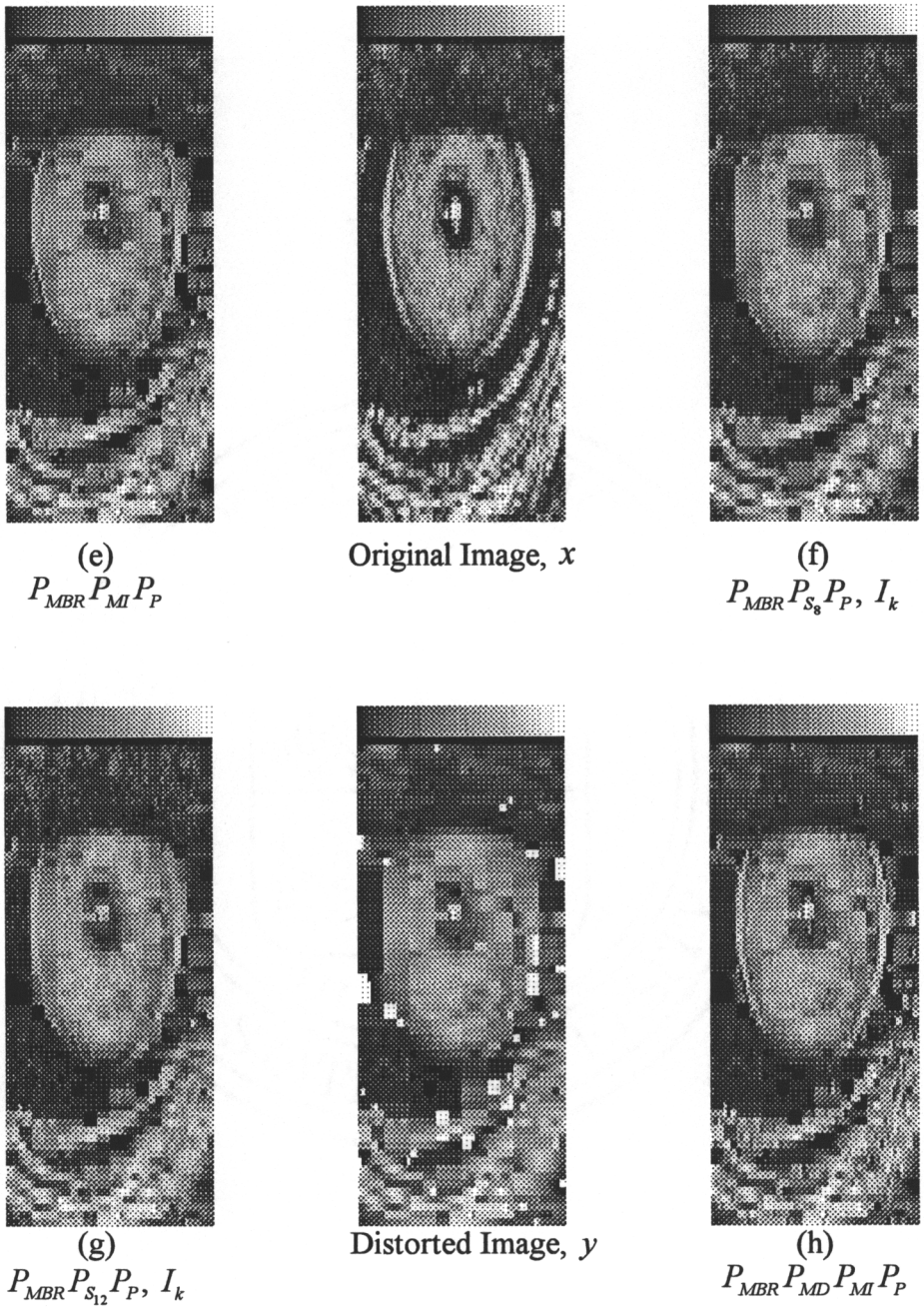


(c)
$P_{MBR}P_SP_P$

Distorted Image, y

(d)
$P_{MBR}P_{Spk_2}P_P$

Figure 27: POCS reconstructions of Mandrill's eye at 1 bit/pixel or less



(e)
$P_{MBR}P_{MI}P_P$

Original Image, x

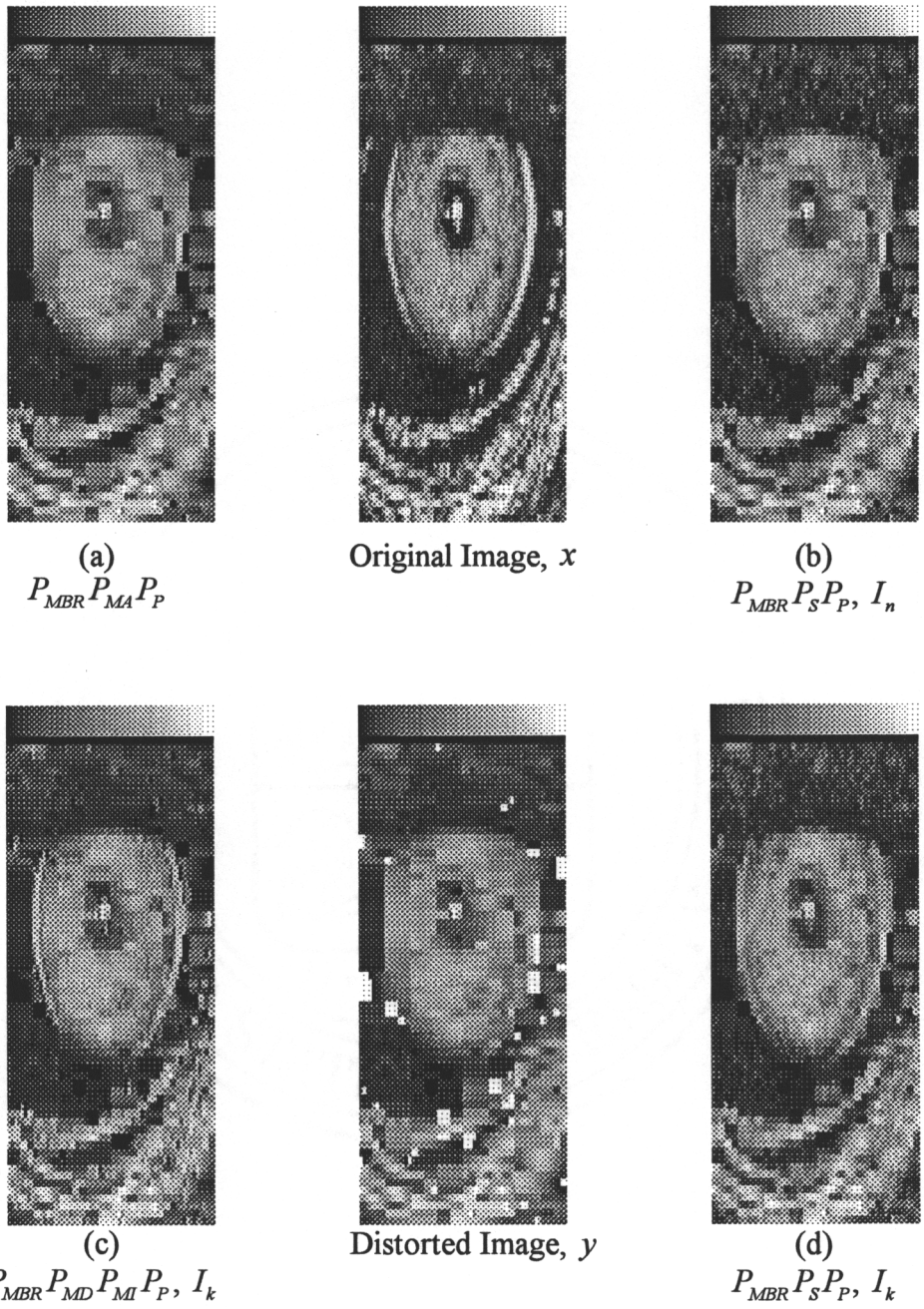
(f)
$P_{MBR}P_{S_8}P_P$, I_k

(g)
$P_{MBR}P_{S_{12}}P_P$, I_k

Distorted Image, y

(h)
$P_{MBR}P_{MD}P_{MI}P_P$

Figure 27: POCS reconstructions of Mandrill's eye at 1 bit/pixel or less (cont'd)



(a)
$P_{MBR}P_{MA}P_P$

Original Image, x

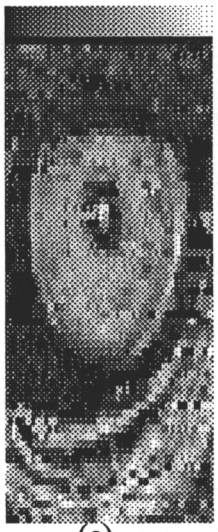
(b)
$P_{MBR}P_SP_P$, I_n

(c)
$P_{MBR}P_{MD}P_{MI}P_P$, I_k

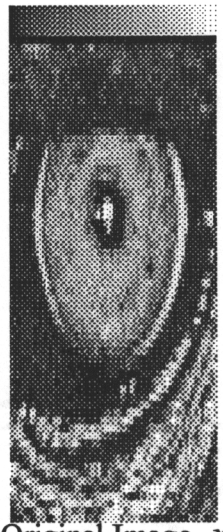
Distorted Image, y

(d)
$P_{MBR}P_SP_P$, I_k

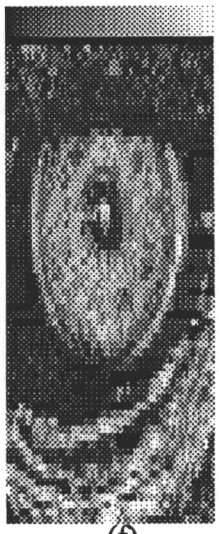
Figure 28: POCS reconstruction of Mandrill's eye at more than 1 bit/pixel



(e)
$P_H P_{MBR} P_{MA} P_P$

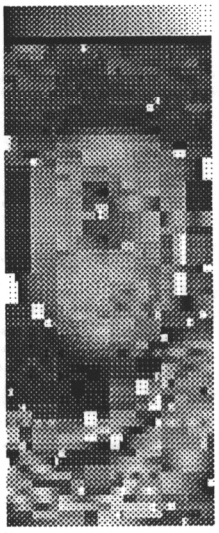


Original Image, x



(f)
$P_H P_{MBR} P_S P_P$, I_k

(g)
$P_{MBR} P_{S_8} P_{MI} P_P$, I_k



Distorted Image, y

Figure 28: POCS reconstruction of Mandrill's eye at more than 1 bit/pixel (cont'd)

4.5 Other Images

The simulations thus far have dealt with only one image. The performance for this single image may or may not be representative of the performance for other images. To determine if the above results are typical the distortion and reconstruction process is carried out for two other images. Image 2 is a section of another image taken from Matlab's library titled "Clown" with intensity values that range from 1 to 81. Image 3 is a section of another image taken from Matlab's library, titled "Gatlin," with intensity values that range from 1 to 62.

The results of the POCS reconstructions of Image 2 and Image 3 are summarized together with the previous results for Image 1. The results for Image 2 and 3 are similar to those for Image 1 except that the MSE (%) as defined in (4.1) is consistently lower for Images 2 and 3 than that for Image 1. The only exceptions to this are the last two projections in Table 5, in which the MSE decrease for Image 3 is higher than that for Image 1. The reason for this difference is that the distortion is not as pronounced in Images 2 and 3 as in Image 1. The MSE of the distorted image 1 is 11.95% whereas for Images 2 and 3 the MSE of the distorted image is 2.81% and 1.37% respectively.

Table 5: % MSE decrease of reconstructions using convex constraints

Composite Projection	Image 1	Image 2	Image 3
$P = P_{MBR} P_P$	5.012	0.6735	0.8995
$P = P_{MBR} P_{Z_1} P_P$ criterion 1	5.744	0.6777	0.904
$P = P_{MBR} P_{Z_2} P_P$ criterion 1	6.254	0.9018	1.072
$P = P_{MBR} P_{Z_3} P_P$ criterion 1	6.388	0.916	1.086
$P = P_{MBR} P_{Z_1} P_P$ criterion 2	5.258	0.9098	1.078
$P = P_{MBR} P_{Z_2} P_P$ criterion 2	5.798	0.9584	1.084
$P = P_{MBR} P_{Z_3} P_P$ criterion 2	6.092	0.9637	1.086
$P = P_{MBR} P_{S_8} P_P$	6.525	0.6945	0.9286
$P = P_{MBR} P_{S_{12}} P_P$	6.748	1.006	0.9827
$P = P_{MBR} P_S P_P$	7.423	1.136	1.509
$P = P_{MBR} P_{S_5} P_P$	6.284	0.6905	0.9269
$P = P_{MBR} P_{S_{10}} P_P$	6.539	0.6945	0.9286
$P = P_{MBR} P_{S_{15}} P_P$	7.25	1.125	1.509
$P = P_{MBR} P_{Spk_1} P_P$	15.27	9.691	3.283
$P = P_{MBR} P_{Spk_2} P_P$	18.51	12.75	3.63
$P = P_{MBR} \tilde{P}_{Spk_1} P_P$	14.72	9.015	2.774
$P = P_{MBR} P_{MA} P_P$	22.66	19.53	16.99
$P = P_{MBR} P_{MI1} P_P$	22.74	17.62	10.61
$P = P_{MBR} P_{MD1} P_P$	25.67	18.54	9.062
$P = P_{MBR} P_{MD1} P_{MI1} P_P$	39.15	30.77	16.51
$P = P_{MBR} P_{MI} P_P$	24.1	20.84	15.55

Table 6: % MSE decrease of reconstructions using C_H, I_k, or I_n

Composite Projection	Image 1	Image 2	Image 3
$P = P_{MBR}P_SP_P,\ I_n$	24.66	21.64	18.8
$P = P_{MBR}P_P,\ I_k$	5.822	1.298	2.862
$P = P_{MBR}P_{S_8}P_P,\ I_k$	24.94	23.09	17.41
$P = P_{MBR}P_SP_P,\ I_k$	49.24	48.04	41.56
$P = P_{MBR}P_{MD}P_{MI}P_P,\ I_k$	40.07	31.33	18.52
$P = P_{MBR}P_{S_8}P_{MI}P_P,\ I_k$	41.11	37.48	26.95
$P = P_HP_{MBR}P_{MA}P_P$	52.77	32.97	42.58
$P = P_HP_{MBR}P_SP_P,\ I_k$	61.33	46.64	39.56
$P = P_HP_{MBR}P_{S_8}P_P,\ I_k$	13.8	5.383	Divergence
$P = P_HP_{MBR}P_{S_{12}}P_P,\ I_k$	35.33	21.49	2.095
$P = P_HP_{MBR}P_{S_8}P_{MI1}P_P,\ I_k$	35.62	23.95	7.062
$P = P_HP_{MBR}P_SP_{EB}P_P,\ I_k$	66.38	53.35	67.52
$P = P_HP_{MBR}P_{S_8}P_{EB}P_P,\ I_k$	20.83	21.32	24.48

5.0 Conclusion

The method of Projection Onto Convex Sets (POCS) incorporates many constraints and guarantees strong convergence if all the constraints are closed and convex. Also, the method of POCS reconstructs a signal entirely from constraints and so does not require knowledge of the distortion operator. For these reasons, POCS is well suited to the problem of reconstructing a signal given partial information. Constraints that do not form closed and convex sets may still be useful if they significantly reduce the MSE of the final reconstruction.

For tractability, the problem is narrowed to distortion caused by approximating the Multiple Bases Representation of the signal and considering signals that are generated from images. This thesis has examined the question of which constraints are best in lowering mean square error and producing good quality images while requiring little *a priori* information.

Simulations of POCS reconstructions of a test image are used to compare composite projections by examining the final MSE and image quality. As always there is a trade-off to be considered; lower final MSE and better image quality generally require higher transmission rates.

5.1 Convex Constraints

Knowledge that the signal is derived from an image and that the type of distortion is D_{MBR} allow us to partially reconstruct the signal without transmitting any additional *a*

priori information. The positivity constraint C_P and the MBR constraint C_{MBR} use this knowledge to decrease the resulting MSE in the test image by approximately 5%.

The original image x always contains more energy than the observed image y because the distortion operator D_{MBR} always removes energy from x. Thus, constraints that increase the energy in x_k are much more effective in decreasing MSE. When the observed image is used to initialize the iterative algorithm, C_Z, C_S, C_E, C_{MBR}, C_{EB}, and C_P have little effect on the signal because they introduce a relatively small amount of energy into the reconstruction. The zero crossing constraint C_Z, and the sign constraint C_S can only remove energy. Indeed, the observed signal already satisfies C_Z, C_S, C_E, C_{MBR} and C_{EB} and the function of these constraints is to remove unwanted portions of the signal. Therefore, if any reconstruction is to occur, at least one other constraint, i.e. C_{Spk}, C_{MI}, C_{MD}, C_{MA}, or C_P must be used.

The spike constraint C_{Spk}, the minimum increase constraint C_{MI}, and the minimum decrease constraint C_{MD} significantly lower the resulting MSE of the POCS reconstruction. Each of these constraints introduces energy directly into the signal at specific locations so they tend to improve the image at discrete points, which may make the image appear grainy. A smoothing constraint may improve the image quality by spreading the energy introduced by these constraints into adjacent pixels. Also, blockiness is reduced noticeably by C_{Spk}, C_{MI}, C_{MD}.

5.2 Initialization Vector

A change in the initialization vector was demonstrated to have a large effect on both the mean square error and the image quality for most, but not all, composite projections. Two vectors were considered, the observed signal plus white noise I_n, and the observed signal plus a constant I_k. The latter was determined to be superior to the first for two reasons. First, I_k did not require other constraints to prevent divergence,

while I_n did require additional constraints. Secondly, I_k produced better reconstructions in terms of both MSE and image quality. The initialization vector I_n produced images that were noisy in the low intensity regions of the image.

If the initialization vector is $x_0 \neq y$, x_0 does not satisfy C_Z, C_S, C_E, C_{MBR} and C_{EB} so that combinations of these constraints may form more useful composite projections. Here, energy is added more or less randomly and the constraints are used to remove the undesirable energy that has been introduced. Thus, the better each of these constraints is at removing the unwanted noise the better the constraint performs. Of these constraints C_S was found to be the most effective in lowering MSE. Moreover, C_S can be easily modified to require varying amounts of *a priori* information.

5.3 Laplacian Histogram Constraint

The only nonconvex constraint considered was the Laplacian histogram constraint C_H. A good initial start of the algorithm was shown to be necessary to prevent divergence. To accomplish this, other convex constraints in the composite projection were implemented first.

The minimum absolute deviation constraint C_{MA} is the only constraint that was able to force convergence when the POCS reconstruction used C_H. The reason for this is that C_{MA} provides a good initial start for C_H. However, C_{MA} requires a significant amount of *a priori* information to be transmitted.

The sign constraint C_S also forced C_H to converge if I_k was used. While C_{S_8} forced convergence in the test image, it is not reliable because using C_{S_8} on other images, divergence was found to occur. Although C_E and C_{EB} lowered MSE further, they could not force convergence without C_S.

References

Andrews, H. C. and Hunt, B. R., *Digital Image Restoration*, Englewood Cliffs, NJ: Prentice Hall, Inc., 1977.

Beex, A. A., "Iterative Reconstruction of Space-Limited Scenes from Noisy Frequency-Domain Measurements," ICASSP pp. 147-150, Boston, MA, April 14-16, 1983.

Beex, A. A., "Soft Constraint Iterative Reconstruction from Noisy Projections," ICASSP pp. 12A.6.1-4, San Diego, CA, March 19-21, 1984.

Brogan, W. L., *Modern Control Theory*, Englewood Cliffs, NJ: Prentice Hall, 1985.

Gerchberg, R. W. and Saxton, W. O., "A Practical algorithm for the Determination of Phase from Image and Diffraction Plane Pictures," Optik, Vol. 35, No. 2, pp. 237-246, 1972.

Gerchberg, R. W., "Super-Resolution through Error Energy Reduction," Optica Acta, Vol. 21, No. 9, pp. 709-720 1974.

Griffel, D. H., *Applied Functional Analysis*, New York, NY: John Wiley & Sons, 1981.

Hummel, R. A., "Histogram Modification Techniques," Comput. Graph. and Image Proc., Vol. 4, No. 3, pp. 209-224, 1975.

Hummel, R. A., "Histogram Enhancement by Histogram Transformation," Comput. Graph. and Image Proc., Vol. 6, No. 2, pp. 184-195, April 1977.

Khanna, R., *Image Data Compression Using Multiple Bases Representation*, M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, September 1990.

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.

Mannos, J. L. and Sakrison, D. J., "The Effects of a Visual Fidelity Criterion on the Encoding of Images," IEEE Trans. Inf. Theory, Vol. 20, No 4, pp. 525-536, July 1974.

Matlab. The MathWorks Inc., Natick, MA.

Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, New York, NY: McGraw-Hill 1965.

Papoulis, A., "A New Algorithm in Spectral Analysis and Bandlimited Extrapolation," IEEE Trans. Circuits Syst., Vol. 22, pp. 735-742, 1975.

Pratt, W. K., *Digital Image Processing*, New York: John Wiley & Sons, 1978.

Safar, F. G., *Signal Compression and Reconstruction Using Multiple Bases Representation*, M.S. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061, June 1988.

Schafer, R. W., Russell, M. M., and Richards, M. A., "Constrained Iterative Restoration Algorithms," Proc. IEEE, Vol. 69, No. 4, pp. 432-450, April 1981.

Sezan, M. I. and Stark, H., "Image Restoration by the Method of Convex Projections: Part 2 - Applications and Numerical Results," IEEE Trans. Med. Imaging, Vol. 1, pp. 95-101, October 1982.

Sezan, M. I. and Trussell, H. J., "Prototype Image-Based Constraints for Set-Theoretic Image Restortation," ICASSP Vol. 4, pp. 1897-1900, 1990.

Shanmugan, K. S. and Breipohl, A. M., *Random Signals - Detection, Estimation and Data Analysis*, New York, NY: John Wiley & Sons 1988.

Stark, H., editor, *Image Recovery: Theory and Application*, Orlando, FL.: Academic Press, Inc., 1987.

Trussell, H. J., "Convergence Criteria for Iterative Restoration Methods," IEEE Trans. ASSP, Vol. 31, No. 1, pp 129 - 136, Febuary 1983.

Trussell, H. J. and Civanlar, M. R., "The Initial Estimate in Constrained Iterative Restoration," ICASSP Vol. 2, pp. 643-646, 1983.

Trussell, H. J. and Civanlar, M. R., "The Feasible Solution in Signal Restoration," IEEE Trans. ASSP, Vol. 32, No. 2, pp. 201-212, April 1984.

Wang, Y., Liu, Y., and Han, L., "Image Reconstructions from Limited Data with Histogram Constraint," ICASSP Vol. 2, pp. 1292-1295, 1988.

Woods, R. E. and Gonzalez, R. C., "Real-Time Digital Image Enhancement," Proc. IEEE, Vol. 69, No. 5, pp. 643-654, May 1981.

Youla, D. C., "Generalized Image Restoration by the Method of Alternating Orthogonal Projections," IEEE Trans. Circuits Syst., Vol. 25, No. 9 pp. 694-702, September 1978.

Youla, D. C. and Webb, H., "Image Restoration by the Method of Convex Projections: Part 1 - Theory," IEEE Trans. Med. Imaging, Vol. 1, No. 2, pp. 81-94, October 1982.

Vita

Phillip Moose grew up near Natural Bridge, Virginia. He was born on May 1, 1968 five minutes after his identical twin John, in Lexington, Virginia. He went to Natural Bridge High School where he graduated in the top tenth of his class in 1986.

He then began his undergraduate degree in Electrical Engineering at Virginia Tech and worked at General Electric in Salem as a cooperative education student.

After graduating in 1991 he began graduate work at Virginia Tech while supporting himself by working part-time at Tele-Works Inc. in Blacksburg. There he programmed voice response systems on the OS/2 operating system until the time he quit in August, 1993. At which time he began work at Image Processing Technologies. He graduated in May 1994.