Multiple Description Decoding of Overcomplete Expansions Using Projections onto Convex Sets

Philip A. Chou, Sanjeev Mehrotra, and Albert Wang

Microsoft Corporation, One Microsoft Way, Redmond, WA 98052-6399

pachou@microsoft.com, sanjeevm@microsoft.com, alwang@microsoft.com

Abstract

This paper presents a POCS-based algorithm for consistent reconstruction of a signal $x \in \mathcal{R}^K$ from any subset of quantized coefficients $y \in \mathcal{R}^N$ in an $N \times K$ overcomplete frame expansion $y = Fx$, $N = 2K$. By choosing the frame operator F to be the concatenation of two $K \times K$ invertible transforms, the projections may be computed in $\mathcal{R}^K$ using only the transforms and their inverses, rather than in the larger space $\mathcal{R}^N$ using the pseudo-inverse as proposed in earlier work. This enables practical reconstructions from overcomplete frame expansions based on wavelet, subband, or lapped transforms of an entire image, which has heretofore not been possible.

1 Introduction

Multiple description (MD) source coding is the problem of encoding a single source $\{X_i\}$ into N separate binary descriptions at rates $R_1, \ldots, R_N$ bits per symbol such that any subset S of the descriptions may be received and together decoded to an expected distortion D_S commensurate with the total bit rate of S. Early papers on multiple description coding are information theoretic and consider the problem of determining for $N = 2$ the set of rates and expected distortions $\{(R_1, R_2, D_1, D_2, D_{1,2})\}$ that are asymptotically achievable. More recent papers consider the problem of designing practical multiple description quantizers, and their use over erasure channels. Multiple description quantizers are a nice fit to erasure channels, because a multiple description decoder can reconstruct the source using however many descriptions it receives.

The papers on practical MD quantization have so far taken three distinctly different approaches. In the first approach, pioneered by Vaishampayan, MD scalar, vector, or trellis quantizers are designed to produce $N = 2$ descriptions, using a generalized Lloyd-like clustering algorithm that minimizes the Lagrangian of the rates and expected distortions $R_1, R_2, D_1, D_2, D_{1,2}$ [1, 2, 3, 4]. In the second approach, pioneered by Wang, Orchard, and Reibman, MD quantizers are constructed by separately describing (i.e., quantizing and coding) the N coefficients of an $N \times N$ block linear transform, which has been designed to introduce a controlled amount of correlation between the transform coefficients [5, 6, 7]. In the third approach, pioneered by Goyal, Kovačević, and Vetterli, MD quantizers are constructed by separately describing the N coefficients of an overcomplete $N \times K$ tight frame expansion [8, 7]. The present paper contributes to this last category. For completeness, it should be mentioned that a number of other papers take yet a fourth approach, in which the *natural* correlation between symbols is exploited for reconstruction. For example, odd pixels can be predicted from even pixels, and vice versa. This approach is similar to the second approach above, except that the transform is not actively designed. We consider this fourth approach to be more closely related to standard error resilience techniques.

In MD quantization using overcomplete (frame) expansions, an input signal $x \in \mathcal{R}^K$ is represented by a vector $y = Fx \in \mathcal{R}^N$, $N > K$. F is a so-called frame operator, whose N rows $\{\phi_i\}_{i=1}^N$ span $\mathcal{R}^K$. The coefficients of y are scalar quantized to obtain $\hat{y}$, and then are independently entropy coded and transmitted in (up to) N descriptions. The decoder receives descriptions of only $N' \leq N$ coefficients after potential erasures, and reconstructs the signal $\hat{x}$ from the received descriptions. Each received description is an encoding of the fact that some coefficient y_i lies in a particular quantization bin, say $[l_i, u_i)$.

Without loss of generality, assume that descriptions of the first N' coefficients are received, and that descriptions of the last $N'' = N - N'$ coefficients are erased. Let y' denote the first N' coefficients, and let y'' denote the last N'' coefficients, so that

$$y = \begin{bmatrix} y' \\ y'' \end{bmatrix} = \begin{bmatrix} F' \\ F'' \end{bmatrix} x = Fx,$$

where F' is $N' \times K$ and F'' is $N'' \times K$. A classical way for the decoder to reconstruct x from the received quantized coefficients $\hat{y}'$ is to use the linear reconstruction

$$\hat{x}_{lin} = (F')^{+}\hat{y}' \tag{1}$$

where $(F')^+$ is the pseudo-inverse of F'. The pseudo-inverse can be computed from the singular value decomposition $F' = U \cdot \mathrm{diag}(\sigma_1, \ldots, \sigma_{N'})V^t$ [9] as $(F')^+ = V \cdot \mathrm{diag}(\sigma_1^{-1}, \ldots, \sigma_{N'}^{-1}) \cdot U^t$. This is equal to $(F')^+ = ((F')^t(F'))^{-1}(F')^t$ when F' has full rank, i.e., when at least K descriptions are received (assuming any K rows of F are linearly independent). It can be shown that the reconstruction (1) has the property that

$$\hat{x}_{lin} = \arg\min_x ||\hat{y}' - F'x||^2. \tag{2}$$

That is, when $N' \geq K$, it chooses $\hat{x}_{lin} \in \mathcal{R}^K$ to be the coordinates of the point $F'\hat{x}_{lin}$ in the K-dimensional subspace spanned by the columns of F' which is closest to $\hat{y}' \in \mathcal{R}^K$, i.e., it projects $\hat{y}'$ onto the subspace $F'\mathcal{R}^{N'}$. Furthermore, when $N' < K$, i.e., when the x that minimizes (2) is not unique, then the reconstruction (1) chooses such an x with minimum norm.

The linear reconstruction (2) is not statistically optimal. The optimal reconstruction $\hat{x}_{opt}$, which minimizes the expected (squared error) distortion $E||X - \hat{X}_{opt}||^2$, is

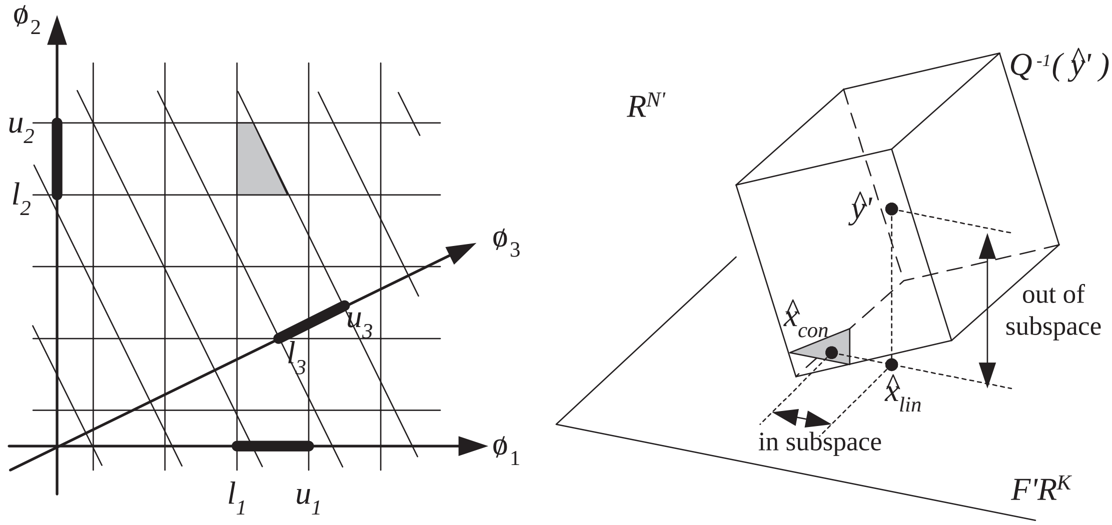

Figure 1: (a) Reconstruction in $\mathcal{R}^K$. (b) Reconstruction in $\mathcal{R}^{N'}$ (adapted from [10]).

the conditional mean of X given the descriptions $y_i \in [l_i, u_i)$ received. That is,

$$\hat{x}_{opt} = E[X|Q(F'X) = \hat{y}'] = E[X|l' \leq F'X < u'], \tag{3}$$

where the relations $l' \leq y'$ and $y' < u'$ are to be taken componentwise. Unfortunately, the conditional expected value of X given that it lies in the region $Q^{-1}(\hat{y}') = \{x : l' \leq F'x < u'\}$ is hard to compute. Figure 1a shows for $K = 2$ and $N' = 3$ a region $Q^{-1}(\hat{y}')$ for some $\hat{y}'$. Note that the regions $Q^{-1}(\hat{y}')$ for different $\hat{y}'$ are all dissimilar in general.

Although the optimal reconstruction (3) is difficult to compute, one thing is certain: x lies in $Q^{-1}(\hat{y}')$, and hence, since $Q^{-1}(\hat{y}')$ is convex, the optimal reconstruction (3) lies in $Q^{-1}(\hat{y}')$. Any reconstruction $\hat{x}$ which does not lie in $Q^{-1}(\hat{y}')$ is said to be inconsistent [10]. Figure 1b, adapted from [10], shows again for $K = 2$ and $N' = 3$ an inconsistent reconstruction from the linear projection (1). It is intuitive that consistent reconstructions have smaller expected squared error distortion than inconsistent reconstructions. In fact, Goyal, Vetterli, and Thao [10] show that while the expected distortion from linear reconstructions is asymptotically proportional to $1/N$, the expected distortion from consistent reconstructions is $O(1/N^2)$ in at least one DFT-based case, and they conjecture this to be true under very general conditions, when the frame is tight.

An algorithm for producing consistent reconstructions $\hat{x}_{con} \in Q^{-1}(\hat{y}')$ is the POCS (Projections Onto Convex Sets) algorithm [11]. In the POCS algorithm, an arbitrary initial point $p_t \in \mathcal{R}^n$ is alternately projected onto two closed convex sets $Q \subseteq \mathcal{R}^n$ and $P \subseteq \mathcal{R}^n$,

$$q_{t+1} = \arg\min_{q \in Q} ||p_t - q||^2 \tag{4}$$

$$p_{t+1} = \arg\min_{p \in P} ||q_{t+1} - p||^2 \tag{5}$$

until p_t and q_t converge to the intersection of P and Q, or if the intersection is empty, until p_t and q_t respectively converge to the sets $\{p \in P : ||p-q||^2 \leq ||p'-q||^2, \forall p' \in P, \forall q \in Q\}$ and $\{q \in Q : ||q-p||^2 \leq ||q'-p||^2, \forall q' \in Q, \forall p \in P\}$. For consistent reconstruction of $x \in \mathcal{R}^K$ from the quantized frame expansion $\hat{y}' = Q(F'x) \in \mathcal{R}^{N'}$, Goyal, Vetterli, and Thao [10] suggest alternately projecting the initial point $\hat{y}'$ onto the two convex sets $Q = F'\mathcal{R}^K \subseteq \mathcal{R}^{N'}$ and $P = Q^{-1}(\hat{y}') \subseteq \mathcal{R}^{N'}$, as shown in Figure 1b. The first projection (4) onto the linear subspace $F'\mathcal{R}^K$ can be accomplished, as usual, by the pseudo-inverse (1). The second projection (5) onto the quantization bin $Q^{-1}(\hat{y}')$ can be accomplished by component-wise clipping to the quantization bin. That is, the vector $\tilde{y}'$, $l' \leq \tilde{y}' \leq u'$, closest to an arbitrary vector $y' \in \mathcal{R}^{N'}$ is given component-wise by

$$\tilde{y}'_i = \text{clip}(y'_i; l_i, u_i) = \min\{\max\{y'_i, l_i\}, u_i\}. \tag{6}$$

Unfortunately, computation of the pseudo-inverse $(F')^+$ requires $O(K^2N')$ operations; the projections themselves require $O(KN')$ operations each.

To reduce the computational complexity of consistent reconstruction, Goyal, Vetterli, and Thao instead suggest finding $\hat{x} \in Q^{-1}(\hat{y}')$ by solving the linear program $\max c^t\hat{x}$ subject to

$$\begin{bmatrix} F' \\ -F' \end{bmatrix} \hat{x} \leq \begin{bmatrix} u \\ -l \end{bmatrix},$$

for an arbitrary objective functional c. They furthermore suggest that by varying c, it may be possible to find all the vertices of $Q^{-1}(\hat{y}')$, whereupon they can be averaged to approximate the region's centroid. (However, they do not appear to follow this latter suggestion.) Although this method avoids the high cost of computing the pseudo-inverse, the complexity of the simplex algorithm for solving the linear program is still $O(KN')$ operations per pivot. This complexity does not present a problem when K and N' are small. Goyal, Kovačević, and Vetterli [8, 7] apparently apply the algorithm for $K = 8$ and N' up to 10.

We are more interested in decoding multiple descriptions of overcomplete expansions based on overlapping functions, such as provided by wavelet, subband, or lapped transforms. In this case, the $N \times K$ frame operator F typically operates on an entire image at a time, for which $K = 512 \times 512 = 262144$ is common; N may be twice that. Clearly, in such cases it is not feasible to require $O(KN)$ operations for consistent reconstruction. For wavelet, subband, or lapped transforms, where $N = K$, F is sparse. In this case consistent reconstruction can be performed using $O(KL)$ operations, where L is the length of the support of the basis functions. Since L is usually on the order of a few hundred, such reconstruction is eminently feasible, and is used in all modern subband decoders. Both the pseudo-inverse and pivot operations destroy the sparsity of F. This paper presents an algorithm for consistent reconstruction from multiple descriptions of overcomplete expansions that preserves the efficiency of the sparse representation of F when the basis functions have finite support L. That is, the algorithm has complexity $O(KL)$. Our algorithm is based on POCS, but the projections are performed in the lower dimensional space $\mathcal{R}^K$, rather than in the space $\mathcal{R}^{N'}$.

2 Decoding Algorithm

Let F_1 and F_2 be two different invertible $K \times K$ transforms. For example, F_1 may be a wavelet transform (over an image $x \in \mathcal{R}^K$ suitably extended), and F_2 may be the identity transform, another wavelet transform, or the same wavelet transform over a 1-pixel shift of the image. Let $y_1 = F_1 x$ and $y_2 = F_2 x$ be the two corresponding sets of transform coefficients for $x \in \mathcal{R}^K$. Then

$$y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} x = Fx$$

defines an overcomplete $N \times K$ frame expansion of x with redundancy $N/K = 2$. The expansion is tight if both F_1 and F_2 are orthonormal[1]. We do not insist on orthonormality of F_1 and F_2, but it is best if F_1 and F_2 are orthonormally related, i.e., $F_{12} = F_2 F_1^{-1}$ and its inverse $F_{21} = F_1 F_2^{-1}$ are orthonormal, at least approximately, to ensure convergence of the POCS algorithm.

Let $\hat{y}_1$ and $\hat{y}_2$ be the quantized versions of y_1 and y_2, respectively, such that $\hat{y}_1$ and $\hat{y}_2$ lie (componentwise) between upper and lower quantization cell boundary vectors $l_1 \le \hat{y}_1 \le u_1$ and $l_2 \le \hat{y}_2 \le u_2$, respectively.

Let $R_1 \subseteq \{1, \ldots, K\}$ be the set of indices of the descriptions received by the decoder for $\hat{y}_1$, and let $R_2 \subseteq \{1, \ldots, K\}$ be the set of indices of the descriptions received by the decoder for $\hat{y}_2$. Descriptions not received by the decoder include those that have been erased as well as not sent at all.

A reconstruction $\hat{x}$ is consistent with the received descriptions if and only if it lies in the intersection of the following two closed convex sets:

$$\begin{aligned} P &= \{x : l_{1,i} \le (F_1 x)_i \le u_{1,i}, i \in R_1\} \\ Q &= \{x : l_{2,i} \le (F_2 x)_i \le u_{2,i}, i \in R_2\}. \end{aligned}$$

The following is our basic algorithm for finding a consistent reconstruction of x from the received descriptions.

1. *Initialization.* Start from an initial point in $F_1 P$: With $t = 0$, set

$$p_{1,i}(0) = \begin{cases} \hat{y}_{1,i} & \text{if } i \in R_1 \\ 0 & \text{if } i \notin R_1 \end{cases} .$$

2. *Transform $p_1(t)$ into the coordinate system of F_2.*

$$p_2(t) = F_2 F_1^{-1} p_1(t) = F_{12} p_1(t)$$

3. *Project $p_2(t)$ onto $F_2 Q$.*

$$q_{2,i}(t+1) = \begin{cases} \min\{\max\{p_{2,i}(t), l_{2,i}\}, u_{2,i}\} & \text{if } i \in R_2 \\ p_{2,i}(t) & \text{if } i \notin R_2 \end{cases}$$

[1]A frame expansion $y = Fx$ is tight if there exists a positive constant A such that for all x, $||Fx||^2 = A||x||^2$, which is a generalization of Parseval's relation. In the multiple description scenario, tightness of the original frame F is of little consequence, because the received frame F' will in general not be tight.

4. *Transform* $q_2(t+1)$ *into the coordinate system of* F_1.

$$q_1(t+1) = F_1 F_2^{-1} q_2(t+1) = F_{21} q_2(t+1)$$

5. *Project* $q_1(t+1)$ *onto* $F_1 P$.

$$p_{1,i}(t+1) = \begin{cases} \min\{\max\{q_{1,i}(t+1), l_{1,i}\}, u_{1,i}\} & \text{if } i \in R_1 \\ q_{1,i}(t+1) & \text{if } i \notin R_1 \end{cases}$$

6. *Check for convergence.* If $||p_1(t+1) - p_1(t)||^2 > \epsilon$, then set $t \leftarrow t+1$ and go to Step 2.

7. *Reconstruct* x.

$$\hat{x} = F_1^{-1} p_1(t+1)$$

A reasonable value for ϵ is K, so that the squared error per pixel is within one gray level.

There are two improvements that can be made to this basic algorithm. The first improvement that can be made is to reconstruct to a point in the interior of the quantization cell, rather than to a point on the boundary. Hopefully, such a reconstruction will be closer to the cell centroid. For this purpose, modify the algorithm to gradually shrink P and Q towards their centers (or approximate centroids) until there is no point of intersection. That is, run the basic algorithm to convergence, increase the lower limits, decrease the upper limits, and run the basic algorithm again to convergence. The limiting points for each run, say $p_1(\infty)$ and $q_1(\infty)$, will be equal until there is no point of intersection between P and Q, after which they will begin to diverge. Reconstruct $\hat{x}$ from the last limiting point $p_1(\infty)$ for which $p_1(\infty)$ approximately equals $q_1(\infty)$.

The second improvement that can be made, when the number of received descriptions $N' = |R_1| + |R_2|$ is less than K, is to reconstruct missing components to their conditional expected values given the received descriptions. The basic algorithm already does this if y_1 has a spherical density (e.g., if the components of y_1 are iid Gaussian). The reason is that if y_1 has a spherical density, then the subvector y_1'' consisting of the erased components $\{y_{1,i}\}_{i \notin R_1}$ has a spherical conditional density given the received components $\{y_{1,i}\}_{i \in R_1}$. Furthermore given the received components $\{y_{2,i}\}_{i \in R_2}$, y_1'' must lie in some $|R_2|$-dimensional linear variety. Thus the conditional density of y_1'' given all the received components is a spherical distribution in some linear variety with its mean at the point where the all-zero vector $y_1'' = 0$ projects onto the linear variety. Therefore, setting the missing components in the initial point $p_1(0)$ in Step 1 of the basic algorithm will result in their being replaced, after projection in Step 3, by their conditional means given the received descriptions. Although in most circumstances the density of y_1 is not spherical, it will be approximate spherical if F_1 is a decorrelating transform and x is preconditioned such that it has zero mean and the variance of each $y_{1,i}$ is constant. More precisely, if σ_i^2 is the variance of $(F_1 x)_i$, then replace x by $F_1^{-1}\text{diag}(\sigma_1^{-1}, \ldots, \sigma_K^{-1}) F_1(x - EX)$. The resulting vector y_1 will have approximately spherical density.

3 Experimental Results

Results are presented using the given algorithm to reconstruct a K dimensional vector x. The vector x is formed by taking the inverse DCT of a vector of transform coefficients y_1, which are in turn sampled from a Gaussian distribution with mean 0 and diagonal covariance with entries inversely proportional to frequency, i.e., $Y_{1,k} \sim N(0, \sigma^2/(1+ck))$, $k = 0, 1, \ldots, K-1$, where σ^2 and c are constants. The vector x is then transformed using a DCT for F_1 and the identity transform for F_2, yielding $y_1 = DCT(x)$ and $y_2 = x$. The transform coefficient vectors y_1 and y_2 are uniformly scalar quantized using a step size of Δ. In our experiment, $K = 512$, $\sigma^2 = 1.0$, $\Delta = 0.1$, and $c = 0.013$, so that the variance of the last coefficient is $\sigma^2/8$.

All of the quantized coefficients $\hat{y}_1$ are transmitted, along with 1/4 of the quantized coefficients $\hat{y}_2$ (randomly selected), for a redundancy of $N/K = 1.25$. Of these 640 transmitted coefficients, random subsets are received. The reconstruction algorithm is run for each subset received. The performance of the algorithm is measured by averaging the reconstruction error over all subsets having the same number of coefficients.

The performance of a comparable forward error correction (FEC) system is also obtained. A comparable FEC system is one in which the quantized coefficients from one of the transforms $\hat{y}_1$ are transmitted along with error correction information. The transmitted values are obtained by taking N linear combinations of the K quantized coefficients operating over a Galois field. The code is designed so that any K of the N basis vectors in the system are linearly independent. If each quantized coefficient is one packet, then the N transmitted packets are obtained by applying standard linear block codes of rate (N, K) to each bit of the quantized coefficients. The codes are systematic so that the actual coefficient values are transmitted in K of the N packets. If at least K of the packets are received, then all K coefficients of $\hat{y}_1$ can be recovered and the reconstruction is obtained by taking the inverse transform. If less than K of the packets are received, then only the received packets in the systematic portion are used, and the rest are set to 0.

A plot of the SNR vs. N' is shown in Figure 2. There is a slight gain with our system over the FEC system if all the coefficients are received. If fewer than K coefficients are received, then our system always outperforms the FEC system. However, there is substantial loss if exactly K coefficients are received. This is due to the fact that the partition induced by the received coefficients is usually not cubic since there is a mixture of coefficients from the two transforms. Also, the point found by the POCS algorithm is not necessarily the centroid of the cell. The exact value of the loss is somewhat arbitrary; performance of the FEC system when more than K coefficients are received can be made as high as desired by reducing Δ, whereas performance of even the optimal reconstruction when fewer than K coefficients are received reaches an upper bound independent of Δ.

Experimental results are also provided for the image Lena, using for F_1 a three-level dyadic separable 2D wavelet transform based on the 9/7 filter of [12], and using for F_2 the separable 2D DCT. The coefficients are uniformly scalar quantized to a stepsize of 16. All of the coefficients of $\hat{y}_1$ are transmitted, along with the coefficients

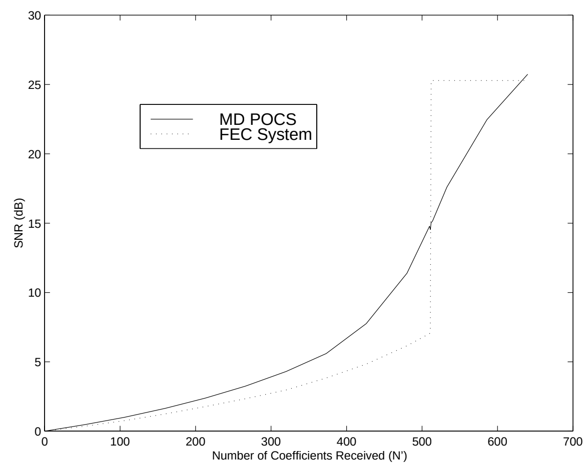

Figure 2: SNR Results using MD POCS and FEC System.

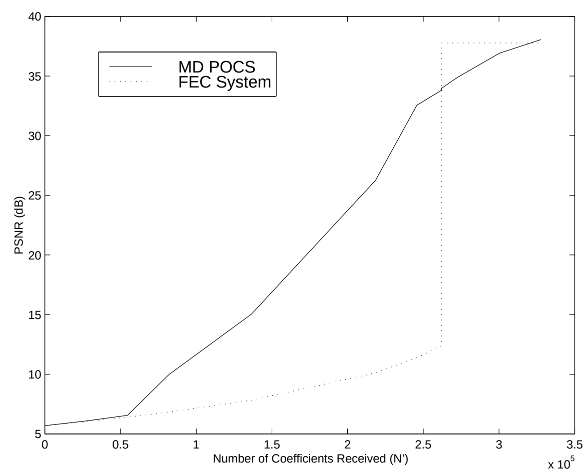

Figure 3: PSNR results for Lena comparing MD POCS and FEC system.

of $\hat{y}_2$ corresponding to the lower half horizontal frequencies and the lower half vertical frequencies, for a redundancy of $N/K = 1.25$. A plot of the PSNR vs. N' is show in Figure 3. A reconstruction is shown in Figure 4 when 1/8 of the coefficients are erased at random. The reconstruction PSNR is 36.58 dB, compared to 15.68 dB if the same 1/8 of the wavelet coefficients are erased without benefit of being reconstructed from the extra coefficients from F_2, as shown in Figure 5. However, this compares to 37.77 dB if the wavelet coefficients $y_1 = F_1 x$ are all received in their entirety.

4 Conclusion

In summary, we have developed a POCS-based algorithm for consistent reconstruction from multiple descriptions of overcomplete expansions. The algorithm operates in the data space $\mathcal{R}^K$ rather than in the expanded space $\mathcal{R}^N$, $N > K$. By constructing the frame from two complete transform bases, all projections can be expressed in terms of

Figure 4: Reconstruction from an overcomplete representation with redundancy 1.25, with 1/8 of the transmitted components missing.

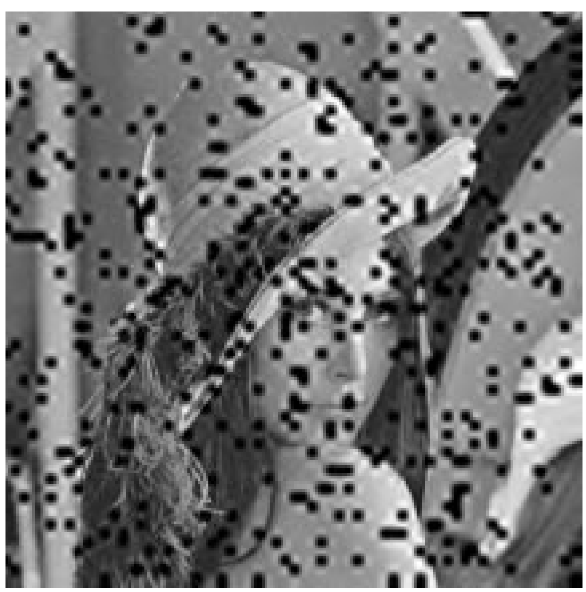

Figure 5: Reconstruction from a critically sampled representation with 1/8 of the transmitted components missing.

forward or inverse transforms. Since such transforms are usually efficient to compute, we can perform the reconstruction much faster than with previous methods. Indeed, our method enables overcomplete frame expansions of an entire image, which has heretofore not been possible.

References

[1] V. A. Vaishampayan. Vector quantizer design for diversity systems. In *Proc. CISS*, 1991.

[2] V. A. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Trans. Information Theory*, 39(3):821–834, May 1993.

[3] V. A. Vaishampayan and J. Domaszewicz. Design of entropy-constrained multiple description scalar quantizers. *IEEE Trans. Information Theory*, 40(1):245–250, January 1994.

[4] S. D. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt. Multiple description wavelet based image coding. In *Proc. Int'l Conf. Image Processing*, Chicago, IL, October 1998. IEEE.

[5] Y. Wang, M. T. Orchard, and A. R. Reibman. Multiple description image coding for noisy channels by pairing transform coefficients. In *Proc. Workshop on Multimedia Signal Processing*, pages 419–424. IEEE, Princeton, NJ, June 1997.

[6] V. K. Goyal and J. Kovačević. Optimal multiple description transform coding of Gaussian vectors. In *Proc. Data Compression Conference*, pages 388–397, Snowbird, UT, March 1998. IEEE Computer Society.

[7] V. K. Goyal, J. Kovačević, R. Arean, and M. Vetterli. Multiple description transform coding of images. In *Proc. Int'l Conf. Image Processing*, Chicago, IL, October 1998. IEEE.

[8] V. K. Goyal, J. Kovačević, and M. Vetterli. Multiple description transform coding: robustness to erasures using tight frame expansions. In *Proc. Int'l Symp. Information Theory*, page 408, Cambridge, MA, August 1998. IEEE.

[9] B. Noble and J. W. Daniel. *Applied Linear Algebra*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition, 1977.

[10] V. K. Goyal, M. Vetterli, and N. T. Thao. Quantized overcomplete expansions in $\mathcal{R}^n$: Analysis, synthesis, and algorithms. *IEEE Trans. Information Theory*, 44(1):16–31, January 1998.

[11] D. C. Youla. Mathematical theory of image restoration by the method of convex projections. In H. Stark, editor, *Image Recovery: Theory and Applications*. Academic Press, 1987.

[12] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Trans. Image Processing*, 1:205–221, April 1992.