

FILTER BANKS FOR MULTIPLE DESCRIPTION CODING

Pier Luigi Dragotti

Electrical and Electronic Engineering
Imperial College,
Exhibition Road, London SW7-2BT, UK
e-mail: p.dragotti@imperial.ac.uk

ABSTRACT

In this paper we review some of our recent results on the design of critically sampled and oversampled filter banks for multiple description coding. For the case of critically sampled filter banks, we show that optimal filters are obtained by allocating the redundancy over frequency with a reverse 'water-filling' strategy. Then we present families of oversampled filter banks that attain optimal performance as well.

1. INTRODUCTION

Multiple Description (MD) coding is a source coding technique for information transmission over unreliable networks. A MD coder represents an information source using multiple bit streams (descriptions). Each individual description provides an approximation to the original message, and multiple descriptions can refine each other to produce a better approximation than that attainable by any single one alone. The simplest formulation of the MD problem involves only two descriptions. This is the so called case of two channels and three receivers. The source generates two descriptions at rates R_1 and R_2 . If both descriptions are received then the decoder can reconstruct the source at some small distortion value D_0 (the *central* distortion), but if either one is lost, the decoder can still reconstruct the source at some higher distortion D_1 or D_2 (the *side* distortions).

Early papers on MD coding are information theoretic in nature and try to find the set of achievable values for the quintuple $(R_1, R_2, D_0, D_1, D_2)$ [6, 15, 1, 28, 27, 13, 25, 26, 5]. The MD problem can be generalized to more than two channels and more than three receivers. The natural extension is to M channels and $2^M - 1$ receivers. The situation of three channel and seven receivers was studied by Zhang and Berger [27]. While an achievable region for the M -channel case has been found recently [22].

Several efforts have also been made to design practical MD coding systems. These systems can be divided into two main families which follows two different philosophies. One focuses on the problem of designing particular quantizers that can meet the MD constraint, while the other family uses ordinary quantizers and get the MD property from the choice of a particular transform.

In [19], a design procedure for the construction of fixed-rate scalar quantizers was presented. In [20], that design procedure was extended to the entropy-constrained case. Various constructions of MD vector quantizers have been proposed [4, 7, 11, 21]

This paper includes research conducted jointly with Martin Vetterli, Sergio Servetto, Jelena Kovačević and Vivek Goyal [5, 12].

and the MD lattice quantizers of [21] closes the gap between the performance of the entropy constrained MD scalar quantizer and the MD rate-distortion bound.

A rather different approach pioneered by Wang et al. [24] and then extended by Goyal and Kovačević [8] consists of applying a suitable blockwise transform to the input vector before coding to obtain the MD property. This approach is usually called MD Transform Coding. The basic idea is to decorrelate the vector components and then to introduce again correlation between coefficients, but in a known and controlled manner, so that erased coefficients can be statistically estimated from those received.

Finally, techniques based on overcomplete frame expansions have been proposed in [2, 9, 14].

For an excellent review on MD coding refer to [10].

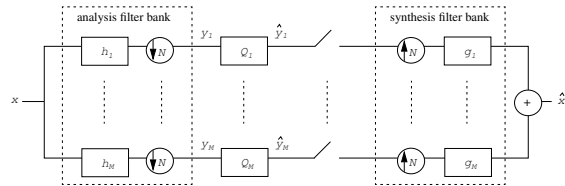


Fig. 1. Abstraction of a lossy network with a frame expansion implemented by an oversampled filter bank.

In this paper, we consider the communication model depicted in Figure 1. An input sequence $x[n]$ is fed through an M -channel filter bank followed by downsampling by N . The M output sequences are then separately scalar quantized with uniform scalar quantizers and sent over M different channels. Each channel either works perfectly or not at all. The decoder receives only $M - e$ of the quantized output sequences, where e is the number of erasures during the transmission. The reconstruction process is linear. We wish to find properties of the filter banks that minimize the mean square error (MSE) between the input and the reconstructed sequences. The case $M = N$ (critically sampled filter banks) is treated in the next section. We show that in this case optimal filter banks form a biorthogonal basis. In Section 3 and Section 4, we consider the case when $M > N$ and present families of oversampled filter banks that achieve optimal performance. Finally, we conclude in Section 5 by showing a simple application example.

2. CRITICALLY SAMPLED FILTER BANKS

In this section, we consider the classical two-channel filter bank scheme, that is, we assume $M = N = 2$. Moreover, we assume the input sequence $x[n]$ to be a stationary Gaussian random process with known statistics. The two output sequences are independently coded at rates R_1 and R_2 and sent over two different erasure channels. Now, assume that R bits per sample are sufficient to achieve the central distortion D_0 . We call redundancy $\rho = R_1 + R_2 - R$ the difference between these two cases. This redundancy represents the price we have to pay to reduce the effect of losses. Our target is to find filter banks which minimizes the side distortions for a given fixed redundancy ρ .

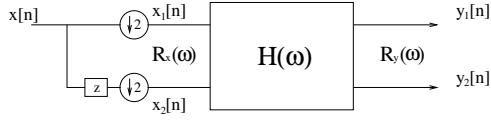


Fig. 2. The polyphase representation of the analysis stage

For convenience we will formulate our problem in the polyphase domain [18, 23]. In this case the analysis stage can be equivalently represented by the block scheme shown in Fig 2. The input-output relation can be expressed in matrix notation introducing the analysis polyphase matrix $H(\omega)$:

$$\begin{pmatrix} Y_1(\omega) \\ Y_2(\omega) \end{pmatrix} = \begin{pmatrix} H_{11}(\omega) & H_{12}(\omega) \\ H_{21}(\omega) & H_{22}(\omega) \end{pmatrix} \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \end{pmatrix}. \quad (1)$$

Call $R_x(\omega)$ the 2×2 polyphase power spectral density (p.s.d.) matrix of the input process. Likewise $R_y(\omega)$ is the p.s.d. matrix of the outputs. The system response has the following form:

$$R_y(\omega) = H(\omega)R_x(\omega)H^+(\omega), \quad (2)$$

where $H^+(\omega)$ denotes the Hermitian transpose of $H(\omega)$. The synthesis part of the system can be analyzed in a similar fashion.

As a first step we decompose the matrix $H(\omega)$ into the product of two matrices $M(\omega)$ and $T(\omega)$

$$H(\omega) = T(\omega)M(\omega). \quad (3)$$

$M(\omega)$ is a unitary decorrelating matrix that diagonalizes the input covariance matrix $R_x(\omega)$. Thus: $R_x(\omega) = M(\omega)\Lambda(\omega)M^*(\omega)$ where $\Lambda(\omega)$ is a diagonal matrix which contains the spectral eigenvalues of $R_x(\omega)$:

$$\Lambda(\omega) = \begin{bmatrix} \lambda_1^2(\omega) & 0 \\ 0 & \lambda_2^2(\omega) \end{bmatrix}. \quad (4)$$

For a stationary input process, the decorrelating matrix can be found analytically and has the following form [17]:

$$M(\omega) = \frac{\sqrt{2}}{2} \begin{bmatrix} e^{j\omega/2} & 1 \\ -1 & e^{-j\omega/2} \end{bmatrix}; \quad (5)$$

the filter bank related to $M(\omega)$ is usually called principal component filter bank [17]. Now, this factorization does not reduce the generality of the solution, since $M(\omega)$ is a unitary invertible

matrix independent of ρ and we are considering square error distortions. So it is enough to solve the simpler problem of optimally designing the matrix $T(\omega)$ for the two input sequences with p.s.d. matrix $\Lambda(\omega)$. Then the final solution will be represented by the product between this matrix and the decorrelating matrix $M(\omega)$. From now on we assume that the two sequences $(x_1[n], x_2[n])$ have already been decorrelated and are represented by the diagonal p.s.d. matrix $\Lambda(\omega)$. Notice that these two sequences are still a realization of a stationary Gaussian process.

We can now state the following theorem:

Theorem 1 Assume that $\rho \gg 0$ and that the two p.s.d. $\lambda_1^2(\omega), \lambda_2^2(\omega)$ of the two decorrelated input sequences $x_1[n], x_2[n]$ are such that $\delta_1 > \Delta_2$ where δ_1 is the essential infimum of $\lambda_1^2(\omega)$ and Δ_2 is the essential supremum of $\lambda_2^2(\omega)$. Then the optimal analysis filters for MD Coding of $x_1[n]$ and $x_2[n]$ are represented by the following polyphase matrix:

$$T(\omega) = \begin{bmatrix} a(\omega) & \frac{1}{2a(\omega)} \\ -a(\omega) & \frac{1}{2a(\omega)} \end{bmatrix},$$

where:

$$a(\omega) = \sqrt{\frac{\lambda_2(\omega)}{2\lambda_1(\omega)(2^{2\rho(\omega)} - \sqrt{2^{4\rho(\omega)} - 1})}}$$

and:

$$\rho(\omega) = \rho + \frac{1}{4} \log(\lambda_1^2(\omega) - \lambda_2^2(\omega)) - \frac{1}{8\pi} \int_{-\pi}^{\pi} \log(\lambda_1^2(\omega) - \lambda_2^2(\omega)) d\omega.$$

Even though it is not possible to extend this result to the case of more than two channels, it is possible to find some approximate solutions to this problem.

3. OVERSAMPLED FILTER BANKS

We now focus on the case $M > N$ and we assume that there are no more than $M - N$ erasures. In this case we do not make any assumptions on the input source. However, we use a statistical model for the quantization error; the reconstruction then depends only on the characteristics of the filter bank.

We will use notation: TF for a tight frame, UF for a uniform frame and UTF for a uniform tight frame.

Call $H_i(\omega) = [H_{i1}(\omega), H_{i2}(\omega), \dots, H_{iN}(\omega)]^*$ the polyphase representation of the i th analysis filter and call $\mathbf{H}(\omega)$ the corresponding $M \times N$ polyphase analysis matrix, which is a matrix whose i th row equals $H_i^*(\omega)$. Many properties can be stated easily in terms of this matrix. In particular, we can say that a filter bank implements a frame decomposition in $l_2(\mathbb{Z})$ if and only if its polyphase analysis matrix is of full rank on the unit circle [3] and a filter bank implements a tight frame expansion in $l_2(\mathbb{Z})$ if and only if $\mathbf{H}^*(\omega)\mathbf{H}(\omega) = AI_N$. [3]. The pseudo-inverse is defined as in the finite-dimensional case:

$$\mathbf{H}^\dagger(\omega) = (\mathbf{H}^*(\omega)\mathbf{H}(\omega))^{-1}\mathbf{H}^*(\omega). \quad (6)$$

A frame implemented with filter banks is uniform if: $\|h_i[n]\| = 1$, $i = 1, \dots, M$ or, using Parseval's relation, if:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{j=1}^N |H_{ij}(\omega)|^2 d\omega = 1, \quad i = 1, \dots, M.$$

If we call $\lambda_i(\omega)$ the spectral eigenvalues of $\mathbf{H}^*(\omega)\mathbf{H}(\omega)$, then:

1. the integral sum of the spectral eigenvalues of $\mathbf{H}^*(\omega)\mathbf{H}(\omega)$ equals the sum of the filters' norms:

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^N \lambda_i(\omega) d\omega = \sum_{i=1}^M \|h_i[n]\|^2;$$

2. for a UF, the integral sum of the eigenvalues equals M ;
3. for a TF, $\mathbf{H}^*(\omega)\mathbf{H}(\omega)$ has eigenvalues constant over the unit circle and equal to A with multiplicity N : $\lambda_i(\omega) = A$, $i = 1, \dots, N$;
4. for a UTF, $\mathbf{H}^*(\omega)\mathbf{H}(\omega)$ has eigenvalues constant over the unit circle and equal to $\frac{M}{N}$ with multiplicity N .

We now introduce a new definition:

Definition 1 (Strongly uniform frame) A frame expansion in $l_2(\mathbb{Z})$ implemented by an $M \times N$ polyphase matrix $\mathbf{H}(\omega)$ is strongly uniform if:

$$\sum_{j=1}^N |H_{ij}(\omega)|^2 = 1, \quad i = 1, \dots, M.$$

3.1. Examples of Strongly Uniform Frames

It will be shown in next sections that strongly uniform tight frames constitute an important class of frames. We propose the following factorization to design polyphase matrices corresponding to strongly uniform tight frames:

$$\mathbf{H}(\omega) = F\mathbf{U}(\omega), \quad (7)$$

where F is a uniform tight frame in \mathbb{C}^N and $\mathbf{U}(\omega)$ is an $N \times N$ paraunitary matrix. It is easy to see that such a polyphase matrix corresponds to a strongly uniform tight frame.

Although we cannot claim that our factorization includes all possible strongly uniform tight frames, we can state the following:

Theorem 2 Define an equivalence relation by bundling a frame (implemented with an FIR oversampled filter bank) with all frames that result from rigid rotations of its elements as well as negation or shift of some individual ones (i.e. $h_i[n] \rightarrow -h_i[n-k]$ $k \in \mathbb{Z}$). When $M = N + 1$, there is a single equivalence class for all strongly uniform tight frames.

4. INTRODUCING ERASURES

Here we consider the effect of erasures on the structure of the frame and on the MSE. We denote by E the index set of erasures and by $\mathbf{H}_E(\omega)$ the polyphase matrix after $e = |E|$ erasures. $\mathbf{H}_E(\omega)$ is an $(M - e) \times N$ matrix obtained by deleting the E -numbered rows from the $M \times N$ polyphase matrix $\mathbf{H}(\omega)$. The first question to be answered is under which conditions $\mathbf{H}_E(\omega)$ still represents a frame. We then study the effect of erasures on the MSE.

Effect of Erasures on the Structure of a Frame

Our aim is to use the pseudo-inverse operator to reconstruct after e erasures. The pseudo-inverse matrix is defined only if the matrix $\mathbf{H}_E(\omega)$ is still a frame, that is, if and only if it is still of full rank on the unit circle. This leads to the following definition:

Definition 2 An oversampled filter bank which implements a frame expansion represented by a polyphase matrix $\mathbf{H}(\omega)$ is said to be robust to $e = |E|$ erasures if and only if for any index set E of erasures, $\mathbf{H}_E(\omega)$ is of full rank on the unit circle.

Let us consider first the case where there is only one erasure.

Theorem 3 An oversampled filter bank which implements a uniform tight frame is robust to one erasure if and only if

$$\sum_{j=1}^N |H_{ij}(\omega)|^2 < \frac{M}{N} \text{ for all } i = 1, \dots, M, \text{ for all } \omega.$$

Recall that by definition a strongly uniform frame is such that: $\sum_{j=1}^N |H_{ij}(\omega)|^2 = 1$, $i = 1, \dots, M$, for all ω . Thus, as a consequence of the previous theorem we can state that any oversampled filter bank which implements a strongly uniform tight frame is robust to one erasure. The result of Theorem 3 does not reveal anything about the existence of filter banks which are robust to more than one erasure.

In [8], it has been shown that a complex harmonic frame in \mathbb{C}^N or a real harmonic frame in \mathbb{R}^N is robust to e erasures ($e \leq M - N$). The following theorem guarantees the existence of at least one family of strongly uniform tight frames in $l_2(\mathbb{Z})$ which are robust to e erasures ($e \leq M - N$):

Theorem 4 Consider an oversampled filter bank with polyphase analysis matrix $\mathbf{G}(\omega) = F\mathbf{U}(\omega)$, where F is a complex harmonic frame in \mathbb{C}^N or a real harmonic frame in \mathbb{R}^N and $\mathbf{U}(\omega)$ is an $N \times N$ polyphase matrix nonsingular on the unit circle ($\det(\mathbf{U}(\omega)) \neq 0$). This filter bank is robust to e erasures ($e \leq M - N$).

If $\mathbf{U}(\omega)$ is a paraunitary matrix, the resulting oversampled filter bank $\mathbf{G}(\omega) = F\mathbf{U}(\omega)$ represents a strongly uniform tight frame robust to e erasures ($e \leq M - N$).

Effect of Erasures on the MSE

In the previous section, it has been shown that it is possible to design oversampled filter banks which are robust up to $M - N$ erasures. We assume such filter banks for the rest of the paper.

Now, we want to compute the effect of the erasures on the MSE. Consider first a strongly uniform frame and $e = 1$. It follows

Theorem 5 Consider encoding with a strongly uniform frame and decoding with linear reconstruction. The MSE averaged over all possible erasures of one channel is minimum if and only if the original frame is tight. Moreover a tight frame minimizes the maximum distortion caused by one erasure. The MSE is given by:

$$MSE_1 = \left(1 + \frac{1}{M - N}\right) MSE_0.$$

It is not possible to extend the result of this theorem to the case of more than one erasure. However, it is possible to compute the MSE with $e > 1$ when the original frame is strongly uniform and tight:

$$MSE_E = \left(1 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{i=1}^e \frac{\mu_i(\omega)}{M - N\mu_i(\omega)} d\omega\right) MSE_0, \quad (8)$$

where $\mu_i(\omega)$ are the spectral eigenvalues of $\mathbf{T}^*(\omega)\mathbf{T}(\omega)$ and $\mathbf{T}(\omega)$ is the $N \times e$ polyphase matrix with columns $\{H_i(\omega)\}_{i \in E}$.

Since the spectral sum of the e eigenvalues of $\mathbf{T}(\omega)$ is constrained to be a constant, the minimum in (8) occurs when all the eigenvalues are equal and constant, which is true when $\mathbf{T}(\omega)$ is tight.



Fig. 3. Multiple description coding of images. (a) Original image. (b) Nine descriptions of the image. (c) Reconstructed image with nine descriptions (PSNR=27.9dB, 0.2bps). (d) Reconstructed image with four descriptions (PSNR=27.2dB, 0.2bps.)

5. A SIMPLE APPLICATION EXAMPLE

To conclude the paper, we show how these families of filter banks can be used for MD coding of images. In Figure 3 we show a simple example of the use of oversampled filter banks. In this case, we have a three channel filter bank with downsampling by two (i.e., $M = 3$ and $N = 2$). The image is filtered along columns and rows, and this leads to nine descriptions shown in Figure 3(b). Each single description is compressed with SPIHT [16] and sent over a different erasure channel. If all the descriptions get to the receiver the quality of the reconstructed image is PSNR=27.9dB (Figure 3(c)), but we still get a good reconstruction quality with only four descriptions (PSNR=27.2dB, Figure 3(d)).

6. REFERENCES

- [1] R. Ahlswede. The rate distortion region for multiple descriptions without excess rate. *IEEE Trans. Information Theory*, 31(6):721–726, November 1985.
- [2] P.A. Chou, S. Mehrotra, and A. Wang. Multiple description decoding of overcomplete expansions using projection onto convex sets. In *Data Compression Conference*, Snowbird, Utah, March 1999.
- [3] Z. Cvetković and M. Vetterli. Oversampled filter banks. *IEEE Trans. Sig. Proc.*, 46(5):1245–1255, May 1998.
- [4] S.N. Diggavi, N.J.A. Sloane, and V.A. Vaishampayan. Design of asymmetric multiple description lattice vector quantizers. In *Data Compression Conference*, Snowbird, Utah, March 2000.
- [5] P.L. Dragotti, S. Servetto, and M. Vetterli. Optimal filter banks for multiple description coding: Analysis and synthesis. *IEEE Trans. Information Theory*, 48(7):2036–2052, July 2002.
- [6] A. A. El Gamal and T.M. Cover. Achievable rates for multiple descriptions. *IEEE Trans. Information Theory*, 28(6):851–857, November 1982.
- [7] M. Fleming and M. Effros. Generalized multiple description vector quantization. In *Data Compression Conference*, Snowbird, Utah, March 1999.
- [8] V. Goyal and J. Kovačević. Generalized multiple description coding with correlating transforms. *IEEE Trans. Information Theory*, 47:2199–2224, September 2001.
- [9] V. Goyal, J. Kovačević, and J.A. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10(3):203–233, May 2001.
- [10] V.K. Goyal. Multiple description coding: Compression meets the network. *IEEE Sig. Processing Mag.*, 18:74–93, September 2001.
- [11] J.A. Kelner, V. Goyal, and J. Kovačević. Multiple description lattice vector quantization: Variations and extensions. In *Data Compression Conference*, Snowbird, Utah, March 2000.
- [12] J. Kovačević, P.L. Dragotti, and V. Goyal. Filter bank frame expansions with erasures. *IEEE Trans. Information Theory*, 48(6):1439–1450, June 2002. Special issue in Honour of Aaron D. Wyner.
- [13] T. Linder, R. Zamir, and K. Zeger. The multiple description rate region for high resolution source coding. In *Data Compression Conference*, Snowbird, Utah, March 1998.
- [14] S. Mehrotra and P.A. Chou. On optimal frame expansions for multiple description quantization. In *International Symposium on Information Theory (ISIT)*, Sorrento, Italy, June 2000.
- [15] L. Ozarow. On a source-coding problem with two channels and three receivers. *Bell Syst. Tech. J.*, 59(10):1909–1921, 1980.
- [16] A. Said and W.A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Trans. Circuits and Systems for Video Technology*, 6(3):243–249, June 1996.
- [17] M.K. Tsatsanis and G. B. Giannakis. Principal component filter banks for optimal multiresolution analysis. *IEEE Trans. Sig. Proc.*, 43(8):1766–1777, August 1995.
- [18] P.P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall, Englewood Cliffs, NJ, 1993.
- [19] V.A. Vaishampayan. Design of multiple description scalar quantizers. *IEEE Trans. Information Theory*, 39(3):821–834, May 1993.
- [20] V.A. Vaishampayan and J. Domaszewicz. Design of entropy-constrained multiple-description scalar quantizers. *IEEE Trans. Information Theory*, 40(1):245–250, January 1994.
- [21] V.A. Vaishampayan, N.J.A. Sloane, and S.D. Servetto. Multiple description vector quantization with lattice codebooks: design and analysis. *IEEE Trans. Information Theory*, 47(5):1718–1734, July 2001.
- [22] R. Venkataramani, G. Kramer, and V.K. Goyal. Bounds on the achievable region for certain multiple description problems. In *International Symposium on Information Theory (ISIT)*, Washington, DC, June 2001.
- [23] M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice-Hall, Englewood Cliffs, NJ, 1995.
- [24] Y. Wang, M. Orchard, V.A. Vaishampayan, and A.R. Reibman. Multiple description coding using pairwise correlation transforms. *IEEE Trans. Image Processing*, 10(3):351–366, March 2001.
- [25] R. Zamir. Gaussian codes and Shannon bounds for multiple descriptions. *IEEE Trans. Information Theory*, 45(7):2629–2635, November 1999.
- [26] R. Zamir. Shannon-type bounds for multiple descriptions of a stationary source. In *Journal of Combinatorics, Information and System Sciences*, December 2000.
- [27] Z. Zhang and T. Berger. New results in binary multiple descriptions. *IEEE Trans. Information Theory*, 33(4):502–521, July 1987.
- [28] Z. Zhang and T. Berger. Multiple description source coding with no excess marginal rate. *IEEE Trans. Information Theory*, 41(2):349–357, March 1995.