ROBUST IMAGE TRANSMISSION USING OVERSAMPLED FILTERBANKS

J. C. Chiang, M. Kieffer and P. Duhamel

LSS – CNRS – Supélec – Université Paris-Sud
Plateau de Moulon, 91192 Gif-sur-Yvette, France

ABSTRACT

It is well known that the redundancy in oversampled filterbanks can be used for obtaining efficient coding schemes. In this paper, we emphasize the fact that they can also be used for resisting transmission errors. Here, a robust image coding system is established based on an oversampled filterbank. Then, some techniques are derived for detecting and correcting the impulse noise resulting from transmission errors. Finally, the full scheme is described, and its performances are evaluated by simulation on a binary symmetric channel.

1. INTRODUCTION

In conventional signal transmission schemes, source and channel coding are designed separately, according to Shannon separation principle [11], see Figure 1-a). This result states that the end-to-end performance of a transmission scheme can be optimized by separately optimizing the source encoder-decoder pair and the channel encoder-decoder pair. However, this holds only in the limit of infinite channel code block length. No design algorithm is provided for good channel codes with finite block length. In addition, Shannon theory does not address the design of good source codes when the probability of channel error is nonzero, which is unavoidable for finite-length channel codes.

As a consequence, for practical systems, the channel decoder is not always able to remove all errors introduced by the channel. Then the residual errors will impact seriously the reconstructed signal. For systems in which the delay or complexity is constrained, alternative approaches based on joint source and channel coding (JSCC) have thus been proposed, see, *e.g.*, [16].

A possible technique for JSCC is to introduce some structured redundancy *before* source coding, see Figure 1-b). This redundancy is obtained by imposing some specific property to the signal to be compressed and transmitted. If the source coder preserves part of this property in the reconstructed signal, then when some errors remain uncorrected after channel decoding, they may be detected and corrected using the introduced redundancy before source coding.

Previous works have introduced structured redundancy to images using BCH codes on the reals [5] or frames of $\mathbb{R}^n$ or $\mathbb{C}^n$ [10]. Error-correction schemes were proposed. However, the redundancy is introduced after transformation of the image, which may impair the coding efficiency.

In this paper, we propose to combine image transformation and redundancy introduction in a single step using oversampled filterbanks (OFBs), see Figure 1-b). In OFBs, the output signal is an overcomplete representation of the input signal. This redundancy has been put at work in [1] to correct part of the noise introduced in the subbands by a Gaussian Bernoulli-Gaussian (GBG) channel, *i.e.*, a channel introducing a noise consisting of a mixture of Gaussian background noise and Bernoulli-Gaussian impulse noise. OFB used in such a way can thus be seen as channel coders placed before the source coder. Note that several results have been obtained for correcting errors in an erasure channel setting, see, *e.g.*, [8]. However, the erasure channel is somewhat different from the GBG channel since external elements (such as the internet channel) provide the information on which samples (or subbands) are in error. In contrast, our techniques can be used in a wireless channel, since they allow error localization.

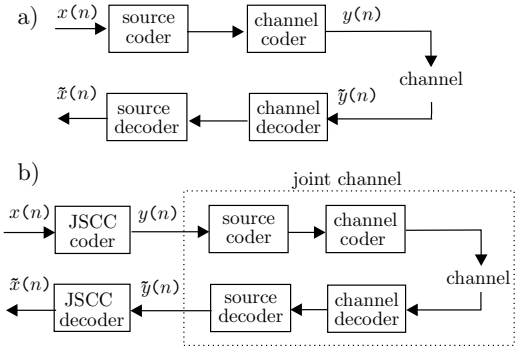

Figure 1: Separate and joint source-channel image transmission schemes

In Section 2, the proposed communication scheme will be detailed. The OFB will be implemented by a discrete Fourier transform (DFT) modulated OFB and the source coding scheme is based on scalar and pyramid vector quantization of the subbands generated by the OFB. When the transmission channel is binary symmetric, the concatenation of quantization and transmission errors can be efficiently modeled as a GBG channel, described in Section 2.3. In Section 3, the way OFBs can be used to correct errors will be recalled. An example comparing a classical separate source-channel communication scheme and the proposed scheme is presented in Section 4.

2. JOINT SOURCE-CHANNEL CODING SCHEME

The joint source-channel communications scheme is described on Figure 1-b). The OFB will be implemented using a discrete Fourier transform (DFT) modulated OFB, see Section 2.1. Source coding is then implemented by nonuniform scalar or pyramid vector quantization, see Section 2.2. The *joint channel* is defined as part of the communication scheme between the input of the source coder and the output of the source decoder and represented by a dashed box on Figure 1-b). Using the elements described before, it will be shown in Section 2.3, that the joint channel can accurately be

This work is part of VIP RNRT project

described by a GBG channel model, when the conventional channel is modeled, *e.g.*, by a binary symmetric channel.

2.1 DFT modulated oversampled filterbanks

Modulated filterbanks [13] have the advantage over standard filterbanks that all subband filters are derived from a single lowpass prototype filter $h(n)$, hence only a single filter design is required. When considering robust image transmission, the first idea would be to employ modulated OFBs with real outputs, such as cosine modulated OFBs. However, as stated in [14], it is impossible to synthetize such type of perfect reconstruction modulated OFB with a small noninteger oversampling ratio.

This is why DFT modulated OFB have been chosen. All L subband filters $h_k(n)$ are derived from the lowpass prototype $h(n)$ as $h_k(n) = h(n)w_L^{-(k+1/2)n}$ for odd-stacked filters and as $h_k(n) = h(n)w_L^{-kn}$ for even-stacked ones, where $k = 0,1,...,L-1$ and $w_L = \exp(-j2\pi/L)$, see [2]. Based on FFT, DFT modulated OFBs allow an efficient implementation. Moreover, when the input and the prototype filter are real, half of the subband coefficients can be obtained from the other half. Consider for example the odd-stacked case, for subbands k and $L-k-1$,

$$\begin{aligned} h_k(n) &= h(n)w_L^{-(k+1/2)n} = h(n).e^{\frac{2\pi}{L}j\left(k+\frac{1}{2}\right)n} \\ h_{L-k-1}(n) &= h(n)w_L^{-(L-k-1/2)n} = h(n).e^{-\frac{2\pi}{L}j\left(k+\frac{1}{2}\right)n} \end{aligned}$$

are complex conjugate. Similar results can be obtained for even-stacked DFT modulated OFB.

2.2 Source coding

In separate source and channel coding, efficient compression is usually realized by quantization followed by entropy coding realized using variable length codes (VLC). The main disadvantage of VLC is that errors not corrected by the channel decoder can desynchronize the VLC decoder and errors may propagate, perturbing the reconstruction of the original signal.

In order to avoid these problems no entropy coding is realized in the proposed scheme. Nonuniform scalar quantization is employed for subbands transmitted at high bitrates. For subbands that are transmitted at low bitrates, pyramid vector quantization (PVQ) [3] is applied to obtain sufficient performances. This type of quantizers belongs to the family of product code vector quantizers. For a given vector of dimension D, its norm is quantized first. After normalization, it is assigned to a point lying on a hyperpyramid and the binary codeword corresponding to this point is then transmitted. The performances are quite satisfying compared to nonuniform scalar quantization, when D is large enough. Moreover, there is no need to transmit any dictionary, as the association of a codeword to a point on the pyramid is realized in a algorithmic way, see [3] and [7] for more details.

2.3 Joint channel model

In separate source and channel coding, the channel is the part of the communication scheme that is between the output of the channel coder and the input of the channel decoder. Similarly, the joint channel is situated between the output of the analysis OFB and the input of the synthesis OFB. The joint channel model gathers the quantization error introduced by source coding and the errors remaining after channel decoding. In [6], such joint channel has been modeled as a memoryless communication channel corrupted by the sum of Gaussian plus Bernoulli Gaussian noises (*GBG channel*). Under these assumptions, for every subband of the OFB, the relation between $y(n)$ and $\widetilde{y}(n)$, the input and output of the joint channel can be written as

$$\widetilde{y}(n) = y(n) + a(n) + b(n), \tag{1}$$

where $b(n)$ is some gaussian noise (quantization errors) and $a(n)$ is an impulse noise (uncorrected channel errors). The gaussian noise has zero mean and variance σ_g^2, while the impulse noise is modeled as Bernoulli gaussian $a(n) = \xi(n)b'(n)$, where $\xi(n)$ stands for a Bernoulli process, an i.i.d. sequence of zeros and ones with $\text{prob}(\xi(n)=1) = p$, and $b'(n)$ represents a gaussian noise with zero mean and variance σ_i^2, such that $\sigma_i^2 \gg \sigma_g^2$. The probability density function (pdf) of the channel noise $c(n) = a(n) + b(n)$ can be expressed as

$$p(c) = (1-p)\,G\left(c,0,\sigma_g^2\right) + pG\left(c,0,\left(\sigma_g^2+\sigma_i^2\right)\right), \tag{2}$$

with $G(c,m,\sigma^2)$ denoting a gaussian pdf having mean m and variance σ^2. A GBG channel is thus characterized by 3 parameters σ_g^2, p and σ_i^2.

If the L quantized subbands are sent into a binary symmetric channel with crossover probability p_{B}, in fact, L GBG channel models have to be built, depending on the number of bits per sample b assigned to the subband and the characteristics of the quantization that has been used. The parameter σ_g^2 corresponds to quantization noise only. The sample error probability p and the variance of the impulse noise σ_i^2 depend on b and on the quantizer outputs when scalar quantization is considered. For pyramid vector quantization, the dimension D of the quantized vectors has also an impact on p and σ_i^2, see [7] and [4] for more details.

3. OVERSAMPLED FILTERBANKS SEEN AS CHANNEL CODES

The following results have been taken from [1] and [9]. The polyphase representation $\mathbf{E}(z)$ provides a convenient description of the relation between the polyphase components $\mathbf{X}(z)$ and $\mathbf{Y}(z)$ of the input and of the output of an L subbands OFB with decimation factor $M < L$. The Smith form decomposition

$$\mathbf{E}(z) = \mathbf{U}(z)\begin{pmatrix}\mathbf{\Lambda}(z)\\ \mathbf{0}\end{pmatrix}\mathbf{W}(z) = \mathbf{U}(z)\begin{pmatrix}\mathbf{\Lambda}(z)\mathbf{W}(z)\\ \mathbf{0}\end{pmatrix}, \tag{3}$$

where $\mathbf{U}(z)$ and $\mathbf{W}(z)$ are unimodular matrices of sizes $L\times L$ and $M\times M$, respectively and $\mathbf{\Lambda}(z)$ is a diagonal matrix of size $M\times M$, evidences the redundancy introduced by an OFB to a input signal. Since unimodular, their inverses exist. After passing $\mathbf{X}(z)$ through the filter $\mathbf{\Lambda}(z)\mathbf{W}(z)$, the obtained signal is padded with $L-M$ zeros before passing through $\mathbf{U}(z)$.

3.1 Parity-check matrix and syndrome

A syndrome is easily defined by partitioning $\mathbf{U}^{-1}(z)$ as

$$\mathbf{U}^{-1}(z) = \begin{pmatrix}\mathbf{V}^0(z)\\ \mathbf{V}(z)\end{pmatrix}, \tag{4}$$

where $\mathbf{V}^0(z)$ is of size $M \times L$ and $\mathbf{V}(z)$ of size $(L-M) \times L$. From (3) and (4), one obtains in the noiseless case

$$\mathbf{V}(z)\mathbf{Y}(z) = \mathbf{V}(z)\mathbf{E}(z)\mathbf{X}(z) = \mathbf{0}. \quad (5)$$

$\mathbf{V}(z)$ plays thus the role of a *parity-check matrix* for the OFB seen as a channel code. Now, the output $\mathbf{Y}(z)$ of the OFB passes through a GBG channel, as described in Section 2.3, to get $\widetilde{\mathbf{Y}}(z)$. After multiplying $\widetilde{\mathbf{Y}}(z)$ by the parity-check matrix one gets

$$\begin{aligned} \mathbf{V}(z)\widetilde{\mathbf{Y}}(z) &= \mathbf{V}(z)(\mathbf{Y}(z)+\mathbf{A}(z)+\mathbf{B}(z)) \quad (6) \\ &= \mathbf{V}(z)\mathbf{A}(z)+\mathbf{V}(z)\mathbf{B}(z), \quad (7) \end{aligned}$$

where $\mathbf{A}(z)$ and $\mathbf{B}(z)$ are the polyphase representations of impulse and gaussian noises respectively. $\mathbf{V}(z)\widetilde{\mathbf{Y}}(z)$ is thus the *syndrome* associated to $\widetilde{\mathbf{Y}}(z)$. This syndrome may be used to realize impulse error correction, even in the presence of background Gaussian noise as shown below.

3.2 Impulse correction scheme

The main difficulty with such scheme comes from the fact that if an impulse occurs at time ℓ, in subband k, all subbands of the syndrome may be affected from time ℓ to time $\ell+N_V$, where N_V is the order of the entries of $\mathbf{V}(z)$. Moreover, as no simple finite state machine can be considered to represent the input-output behavior of the OFB, no simple Viterbi-like decoding technique can be put at work..

Figure 2 summarizes the impulse error detection and correction algorithm proposed in [1]. This scheme is implemented before the synthesis stage of the OFB.

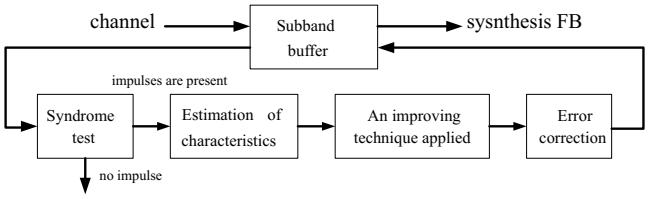

Figure 2: Error correction scheme

The algorithms consists of two steps. First, an hypotheses test is used to determine whether impulse errors are present at a given time instant ℓ and optionally to estimate their number. For computational simplicity, this test is based on the norm of samples of the syndrome (5) taken over a sliding window of length $2N_V+1$. Once ℓ is determined, it is possible to obtain maximum *a posteriori* estimates of the subband k where the impulse error occurs and its amplitude a. For more details, see [9].

4. SIMULATION RESULTS

Here, a paraunitary even-stacked DFT modulated OFB with oversampling ratio $8/6$ is considered. Because of the effect of quantization, the perfect reconstruction constraints play a less important role than when there is no quantization noise. Therefore, the criterion for the prototype filter design is based on a best trade-off between stopband attenuation maximization and reconstruction error minimization [15]. Evenly-stacked DFT modulated OFBs have been preferred to odd-stacked DFT modulated OFBs as, for the same prototype, they give in general a better coding gain. Note that this coding gain must be expressed in terms of the real-valued filters (the independent ones in the outputs of the complex modulated ones) rather than in terms of the complex filter outputs since only the independent real and imaginary parts are quantized and transmitted.

In this example, $L=8$, $M=6$ and the prototype filter length is 48. $\mathbf{E}(z)$ has order 7 and results in a parity-check matrix with order 21. This parity-check matrix has been obtained according to the technique described in [9]. For a given image, the subband samples are obtained by row filtering followed by column filtering to obtain 64 complex subbands. Among them, 4 subbands are real. To reconstruct the image, these 4 real subbands, the real part of 30 subbands and the imaginary part of 30 subbands have to be transmitted. Globally, the corresponding code has a rate $36/64 \approx 1/2$.

As the variances of the samples in differing subband are quite distinct, the entries of the vectors for the PVQ are not taken from a combination of several subbands, but only from a single one. Optimization of the bit-assignment and of the choice of the quantizer (nonuniform SQ or dimension 64 PVQ) for all subbands is performed by the Shoam-Gersho algorithm [12]. At 0.38 bpp, the optimization tells that the first subband coefficients are quantized by a SQ and the remaining subband coefficients by a dimension 64 PVQ.

All quantized subband coefficients are sent over a binary symmetric channel with crossover probability $p_{\mathrm{B}} = 5 \times 10^{-3}$. No channel code has been considered here. The impulse error localization and correction algorithm is applied after inverse quantization and before the synthesis stage of the OFB. First experiments have shown that the algorithm of Section 3.2 does not perform well in the subbands where a dimension 64 PVQ is used. This is mainly due to the fact that a single bit error results in many sample errors in the subband, due to the high dimension of the vector quantization. In order to reduce the sample error-rate p in some subbands and to preserve a low bit-rate, dimension 4 and 16 PVQ has been used for some subbands. The lowest dimension has been assigned to subbands that are perceptually more important. The tuning has been performed experimentally in order to get the best *SNR* in the subband with a low value of p. The resulting bitrate is then 0.41 bpp, and the obtained $PSNR = 31.46$ dB (with no noise) for the reconstructed image when the input image is Lena.

S.band	1	2_{R}	2_{I}	9_{R}	9_{I}
Quant	SQ	4D PVQ	4D PVQ	16D PVQ	16D PVQ
b	5	2.25	2.25	0.875	1.125
p	0.025	0.015	0.015	0.02	0.02
INR	27.8	16.7	16.3	13	14
SNR_0	37.1	11.4	10.6	5.9	7.3
SNR_1	24.2	9.1	8.6	4.7	5.5
SNR_2	34.6	9.8	9.5	4.9	5.9

Table 1: Simulation results

Table 1 shows some results and the corresponding performance of impulse noise correction. In this table, m_{R} and m_{I} denote the real and imaginary part of subband m. For a given subband, b is the assigned number of bits per sample, p the sample-error-rate and INR the impulse to quantization noise ratio. SNR_i indicates the signal to noise ratio between the input and output of the joint channel (SNR_0: without channel error, SNR_1: with channel error, SNR_2: after applying the error correction techniques). Figure 3 illustrates the re-

constructed image without correction. Its $PSNR = 28.7$ dB. Figure 4 shows the one obtained after correction and a corresponding $PSNR = 30.75$ dB. If we compare visually both figures, it is observed that almost all the big spots have disappeared after correction. In terms of *PSNR*, since the *PSNR* without transmission error was 31.46 dB, it is seen that the correction procedure did a very good job in correcting the errors. Globally small spots can still be observed in the corrected image. This can be interpreted as follows : Since the joint channel parameters depend on the dimension of the PVQ, at constant b, the apparent numbers for the sample error-rate p on a PVQ is always larger than for a SQ. The same kind of variation about the *INR*. $INR = 27.8$ dB for the first subband, and as a result, the impulse noise is easily detected and corrected. On the other hand, for the remaining subbands, they have $INR \leq 17$dB. It turns out that with such *INR* value, our algorithm has difficulties for efficiently detecting the impulses. As a result, only the largest ones will be corrected.

Figure 3: The reconstructed image without correction

These performances have to be compared with a classical separate source and channel coding scheme. The source coder is a JPEG image encoder (providing at 0.20 bpp a $PSNR = 30.4$ dB, quality factor 10) followed by a rate $1/2$ convolutionnal code. The bitrate on the channel is then 0.40 bpp. Sent over the same channel as before, the mean performances are $PSNR = 30.4$ dB.

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we proposed an image coding scheme based on OFBs and shows the ability of the OFBs to correct the transmission error. It was shown that, even in a somewhat simple framework involving fixed code-length only, the scheme was able to provide good robustness to error transmission, even for realistic values such as 5×10^{-3}. In the future, we will pay attention to develop an OFB with better frequency localization and the technique with improved impulse noise correction capacities.

REFERENCES

[1] J. C. Chiang, M. Kieffer, and P. Duhamel. Oversampled filterbanks seen as channel codes : impulse noise correction. In *Proceedings of ICASSP*, 2003.

[2] R. E. Crochiere and L. R. Rabiner. *Multirate Digital Signal Processing*. Prentice-Hall, 1983.

[3] T. R. Fischer. A pyramid vector quantization. *IEEE Trans. Inform. Theory*, 32:568–583, 1986.

Figure 4: The reconstructed image with correction

[4] A. Gabay. *Codage conjoint source canal, application à la transmission d'images par satellite*. PhD thesis, ENST, Paris, 2001.

[5] A. Gabay, O. Rioul, and P. Duhamel. Joint source-channel coding using structured oversampled filters banks applied to image transmission. In *Proc. ICASSP*, 2001.

[6] M. Ghosh. Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems. *IEEE Trans. on Communication*, 44(2):145–147, 1996.

[7] A. C. Hung, E. K. Tsern, and T. H. Meng. Error-resilient pyramid vector quantization for image compression. *IEEE Trans. Image Processing*, 7(10):1373–1386, 1998.

[8] J. Kovacević, P. L. Dragotti, and V. K. Goyal. Filter bank frame expansions with erasures. *IEEE Trans. Information Theory*, 48(6), 2002.

[9] F. Labeau, J. C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Marq. Oversampled filter banks as error correction codes: Theory and impulse noise correction. *submitted to IEEE Trans. on Signal Processing*, 2004.

[10] G. Rath and C. Guillemot. Characterization of a class of error-correcting frames and their application to image transmission. In *Proceedings of PCS*, St Malo, 2003.

[11] C. E. Shannon. A mathematical thoery of communication. *Bell Syst. Tech. J.*, 1948.

[12] Y. Shoam and A. Gersho. Efficient bit allocation for an arbitrary set of quantizers. *IEEE Trans. Acoustics, Speech, Signal Processing*, 36(9):1445–1453, 1988.

[13] P. P. Vaidyanathan. *Multirate Systems and Filterbanks*. Prentice-Hall, Englewood-Cliffs, NJ, 1993.

[14] R. F. von Borries, R. L. de Queiroz, and C. S. Burrus. On filter banks with rational oversampling. In *Proc. ICASSP*, 2001.

[15] S. Weiß, M. Harteneck, and R. Stewart. Design and efficient implementation of oversampled GDFT filter banks for subband adaptive filtering. *Digest IEE Colloquium on Digital Filters: an Enabling Technology*, (252):12/1–12/8, 1998.

[16] S. B. Zahir Azami, P. Duhamel, and O. Rioul. Joint source channel coding : Panorama of methods. *Proceedings of CNES workshop on Data Compression*, Nov. 1996.