ROBUST VIDEO TRANSMISSION OVER MIXED IP – WIRELESS CHANNELS USING MOTION-COMPENSATED OVERSAMPLED FILTERBANKS

J. C. Chiang, C. M. Lee

Department of Electrical Engineering
National Chung Cheng University, Taiwan

*M. Kieffer, P. Duhamel**

LSS – CNRS – Supélec – U. Paris-Sud, XI
Plateau de Moulon
F-91192 Gif-sur-Yvette, France

ABSTRACT

Robust video coding has attracted increasing attention during the past few years. This paper proposes a joint source channel coding scheme able to resist transmission errors over mixed internet-wireless channels. It involves motion-compensated oversampled filterbanks (OFBs). The redundancy introduced by the overcomplete representation in signals at the output of OFBs is employed for error correction of the motion compensated frames. The errors may be due to the wireless part of the channel (random noise), but also to the internet part (packet losses). The performance of the proposed approach is illustrated for compressed streams transmitted through a packet erasure channel with an averaged packet loss of 6.25% followed by a binary symmetric channel with a crossover probability of 10^{-2}.

1. INTRODUCTION

With the rapid advance of internet and wireless communication, robust transmission of multimedia signals over mixed internet and wireless channels becomes an important issue which has been actively addressed. As well known, during transmission of data over an heterogeneous network, there are various unavoidable impairments, such as packet losses and random bit errors. Usually, depending on the applications and the corresponding assumptions, only a single type of channel errors is considered. Common ways to protect data sent over a packet-erasure channel are multiple description coding (MDC) [1–4] and forward error correction (FEC), [5–7].

The idea of MDC is to generate several correlated descriptions for the same signal. The error resilience of MDC is due to the correlation existing between the various descriptions. Reconstruction of lost descriptions may be achieved using the received ones. FEC is another widely used technique in packet-switched networks, see, *e.g.*, [5, 7]. Such techniques first divide the input signal into several segments and protect every segment by a channel code with redundancy adapted to the channel conditions. Then the packets are built by combining elements of these segments. If some packets are lost, FEC ensures the possibility of recovering the original signal using the available symbols in the received segments. For channels introducing bit errors, standard channel codes may be used. In [6], a robust video coding for a binary symmetric channel with packet erasures has been proposed. A turbo code is used for protection against all channel errors.

*This work has been partly supported by the NEWCOM network of excellence. Thanks are also due to J. Woods, who kindly provided the initial $t + 2D$ coding program

MDC-based techniques are efficient to resist to packet losses, however, when corrupted packets are received, such techniques have difficulties in correcting errors. In FEC-based techniques, packet recovery and packet error correction may be performed. Nevertheless, when the channel improves they may be suboptimal, as redundancy introduced at a bit-level may not be exploited to remove some effect, *e.g.*, of the quantization noise. When the channel worsens, the correction capacity of FEC may be overflowed and the reconstructed signal may be very degraded.

This paper proposes a $t + 2D$ video coder based on oversampled filterbanks (OFBs) [8]. OFBs provide a redundant representation of the input signal. Previous works for robust image and video transmission using OFBs have been reported in [9] and [10] where redundancy is employed for error correction when transmission over binary symmetric channels is considered. This joint source-channel coding (JSCC) approach allows increased robustness against variations of the channel conditions, in contrast with tandem coding schemes. Here, the error correction ability of OFBs is extended to packet losses. The channel model considered is a packet-erasure followed by a binary symmetric channel (BSC) which stands, *e.g.*, wireless internet. Again, the redundancy between, but also inside the subbands at the output of the OFB is employed to resist both packet losses and transmission errors.

2. PROPOSED $T + 2D$ VIDEO CODING SCHEME

Figure 1 shows the proposed $t+2D$ video coding scheme based on OFBs. Two parts may be distinguished, namely motion-compensated (MC) temporal filtering (TF) and filtering of the MC frames. Unlike hybrid coding, where previously decoded frames are used as reference for the current frame, MCTF is performed along the motion trajectory, thus removing the dependency between consecutive frames.

The technique developed in [11] is employed for motion compensation. The obtained motion vectors (MV) are encoded by lossless DPCM and adaptive arithmetic code. Then, the MC frames are spatially filtered by an OFB, as in [10]. MDC is realized by data partitioning of the filtered subband signals. It is tuned for efficient error concealment in case of packet losses. Scalar quantization (SQ) and pyramid vector quantization (PVQ) are then performed. No entropy coding of the the MC frames is realized, in order to prevent propagation of channel errors in the bitstream. The bit assignment and the choice between SQ and PVQ are realized for every subband, based on the Shoham-Gersho algorithm [12].

Finally, the bitstream is generated by interleaving the output of quantized subband signals and coded MV. Headers and MV are

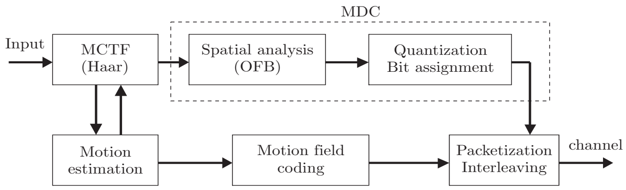

Fig. 1. Proposed MC-OFB based video coding scheme

particularly sensitive to transmission errors. This paper assumes that they are transmitted and received without any error (they are assumed to be strongly channel encoded). What will be demonstrated in this paper is the error correction and packet loss recovery ability of OFBs combined with MD. Only MC texture errors and losses are considered.

3. JOINT SOURCE-CHANNEL CODING TECHNIQUE

The robustness of the proposed video coding scheme is based on the redundancy introduced by the OFB and on the data partitioning technique. Before describing the error correction techniques, the considered channel model has to be described, see Figure 2. The first part of this model is an internet link where transmitted

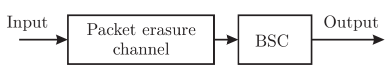

Fig. 2. Considered network model

data may encounter packet losses. The second part of the channel introducing random errors (wireless part) is modeled by a BSC. Thus, during transmission, some packets are lost and the remaining packets are received. However, the available packets may have been corrupted by the BSC noise.

First, the packet erasure channel model is described in Section 3.1. The way MDC is put at work in the proposed scheme to reestimate lost data is presented in Section 3.2. The Gaussian-Bernoulli-Gaussian (GBG) channel model is recalled in Section 3.3, which accounts for quantization noise, random bit errors, and poor re-estimation of lost data. Finally, Section 3.4 summarizes the error correction scheme.

3.1. Packet erasure channel model

The part of the channel introducing packet losses (internet part) is modeled by a 2-states Markov packet transmission model. This model is specified by two parameters p_{10}, and p_{01} where p_{10} denotes the transition probability of received state to lost state and p_{01} stands for lost to received. The averaged packet loss probability $p_{\text{pl}} = p_{10}/(p_{01}+p_{10})$ and the averaged packet loss burst length $L_{\text{pl}} = 1/p_{01}$ may also determine the network behavior. Next, we will present the protection technique for packet losses.

3.2. Multiple description coding for packet loss

Classical MDC algorithms first decorrelate the input signal by a transformation and then introduce some artificial redundancy in order to resist the transmission errors [13]. In the proposed scheme, OFBs are used to combine the signal decomposition and redundancy introduction in a single step. OFBs, as critically sampled filterbanks, perform a spectral signal decomposition. Due to oversampling, some redundancy is introduced in the spectral domain, which may be used for error correction [8]. Here, the spatial correlation that remains in each subband after OFB filtering is also used, as proposed in [14], in the context of image coding using wavelets. This can be accomplished by an efficient data partitioning after OFB filtering, to allow reconstruction of missing samples in subbands using their available neighbors.

A way to generate this kind of multiple descriptions is using polyphase downsampling [15] or domain-based partitioning [14]. In order to minimize the distortion under all possible transmission scenarios when packet losses occur, the domain-based partition has been considered here. In this partition problem, the objective function can be expressed as $P_{\text{opt}} = \arg \min_{|P|=N_p} d_{\min}(P)$, where P denotes the partition set, $P = \{P_0, P_1, ...P_{N_p-1}\}$, N_P is the given number of partitions, $|P|$ is the cardinal number of P and

$$d_{\min}(P) = \min d^i_{\min}, 0 \leq i \leq N_P - 1$$

with $d^i_{\min}$ representing the minimum distance between two samples in the $i-$th partition. The problem is equivalent to sphere packing in $\mathbb{Z}^2$. As well known, the hexagonal lattice solves this problem in $\mathbb{R}^2$ by relaxing some constraints. [4] presents the way of deciding which sublattice in $\mathbb{Z}^2$ is similar to the hexagonal lattice. The main purpose of such partitions is to make every sample in a partition surrounded by as many samples belonging to other partitions as possible. Thus, samples from a missing partition may be reestimated by their available neighbors. This approach is a way to improve error concealment, without increasing the bitrate, contrary to standard MDC techniques. For more details about the way the partitions are built, see [4, 14].

Here, every subbands are partitioned according to [14], see also Figure 3 where the number indicates which description the pixel belongs to. Contrary to the implementation in [1, 2], where each subband is a description, here, descriptions are created by grouping samples from several subbands, as done in the context of image coding using wavelets by [14]. Moreover, each network packet contains only one description.

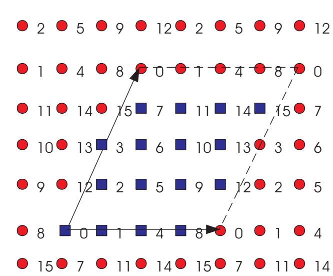

Fig. 3. An example of data partitioning with P=16

3.3. GBG channel model

In [8], it has been shown that OFBs may be seen as channel codes for real-valued data. Decoding with error correction using OFBs requires a model of the part of the communication scheme that lies between the output of the analysis OFB and the input of the synthesis OFB. This model is presented now.

In this paper, this *joint* channel model gathers the quantization error introduced by source coding, the errors remaining after channel decoding. Under these assumptions, for the received subband signal, the relation between $y(n)$ and $\widetilde{y}(n)$, the input and output of the joint channel can be written as $\widetilde{y}(n) = y(n) + a(n) + b(n)$, where $b(n)$ is some Gaussian noise (representing quantization errors) and $a(n)$ is an impulse noise (corresponding to uncorrected channel errors and badly reestimated samples). The Gaussian noise is assumed to be of zero mean and variance σ_g^2, while the impulse noise is modeled as a Bernoulli-Gaussian noise $a(n) = \xi(n)b'(n)$, where $\xi(n)$ stands for a Bernoulli process, an i.i.d. sequence of zeros and ones with prob$(\xi(n) = 1) = p$, and $b'(n)$ represents a Gaussian noise with zero mean and variance σ_i^2, such that $\sigma_i^2 \gg \sigma_g^2$. Such Gaussian-Bernoulli-Gaussian (GBG) channel is thus characterized by there 3 parameters σ_g^2, p and σ_i^2. These parameters may be estimated at encoder side. For example, σ_g^2 is the distortion introduced by the quantization. In absence of packet losses, a closed-form estimate for p and σ_i^2 may be obtained. When packet losses are encountered, p and σ_i^2 may be estimated at encoder side by simulating the samples reestimation in presence of packet losses.

3.4. Error correction scheme

The principle of the packet loss concealment and error correction scheme is as follows. First, samples from lost packets are reestimated from the samples belonging to received packets, as described in Section 3.2. The strategy used here is to evaluate the missing packet from its available nearest neighbors by averaging. Then, in order to correct transmission errors, due to the BSC part of the channel, but also samples that were badly reestimated, the error correction technique presented in [10] is put at work. It is briefly recalled in what follows.

Once missing samples have been estimated, all samples from all subbands are available. Nevertheless, for the synthesis OFB, they have passed through the GBG channel described in Section 3.3, and are thus corrupted by quantization and impulse noise. Using the impulse error detection and correction algorithm proposed in [8], the impulse noise may be partly removed. This scheme, implemented before the synthesis stage of the OFB, is shown in Figure 4.

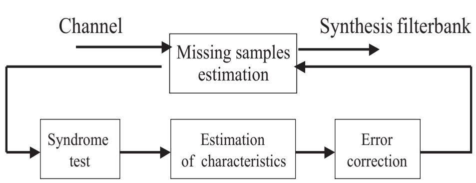

Fig. 4. Error correction scheme

The impulse error correction algorithm consists of two steps. First, an hypothesis test is used to determine whether impulse errors are present at a given time instant and optionally to estimate their number. For computational simplicity, this test is based on the norm of samples of the syndrome taken over a sliding window of length $N_V + 1$, where N_V is the order of the parity-check matrix associated to a given analysis OFB.

Once the time location is determined, it is possible to obtain maximum a posteriori estimates of the subband location and the amplitude of the impulse error. For more details, see [8]. Note that, this error correction technique can be executed iteratively. The missing packet can be re-evaluated by its neighbors after applying impulse correction technique to the impulse corrupted signals. Then the impulse correction technique is again applied based on the re-computed syndromes where the value of samples from to packets and from non-missing packets are both updated.

Parameters	Conditions	lost pack.	PSNR (Y)
$p_b = 0$, $p_{pl} = 0$, $L_{pl} = 0$	no error	0%	35.09 dB
$p_b = 0.01$, $p_{pl} = 0$, $L_{pl} = 0$	only BSC (with correction)	0%	30.75 dB
$p_b = 0.01$, $p_{pl} = 0.1$, $L_{pl} = 2$	losses and BSC (no correction)	6.25%	26.58 dB
$p_b = 0.01$, $p_{pl} = 0.1$, $L_{P_{pl}} = 2$	losses and BSC (correction)	6.25%	29.59 dB

Table 1. Simulation results.

4. SIMULATION RESULTS

The original $t + 2D$ compression algorithm was kindly provided by Dr. J. Woods. Our version with improved robustness is organized as follows: Each GOP (group of picture) contains 8 frames. A 3-level pyramid MCTF is realized with Haar filters. During this filtering, if the number of unconnected pixels between two consecutive frames is larger than a given threshold, MCTF is replaced by MC prediction. The obtained MV are encoded by lossless DPCM and adaptive arithmetic code before being strongly channel coded. Then, the MC frames are spatially filtered by an evenly-stacked DFT modulated OFB, with oversampling ratio $8/6$ and real prototype filter order of 47 [16]. All video components Y, U and V are filtered separately with the same OFB. Quantization involves SQ or PVQ, depending on the results of a bitrate optimization taking the channel effects into account. When PVQ is used, the vector components are taken from the same subband and the same description, at the largest possible distance. It should be noted that PVQ is used here to allow smaller bit rates than SQ; this is absolutely required for a good compression efficiency. Finally, data are sent over the channel described in Section 3.

The input sequence is `foreman.sif`, at 30 fps which corresponds to an original bitrate of 1100 kbps (including MV and MC frames), where the bitrate corresponding to the MC frames filtered by the OFB is 0.2 bpp (1000 kbps).

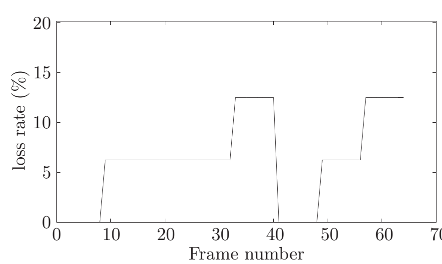

Fig. 5. Packet loss rate for the sequence

Here, the MV are assumed to be transmitted and received without any error while the MC frames may be corrupted by channel errors. The simulation parameters considered here are $p_{pl} = 0.1$, $L_{pl} = 3$ and $p_b = 0.01$, where p_b denotes the crossover probability of the BSC. The number of descriptions is 16 and every packet contains only a single description.

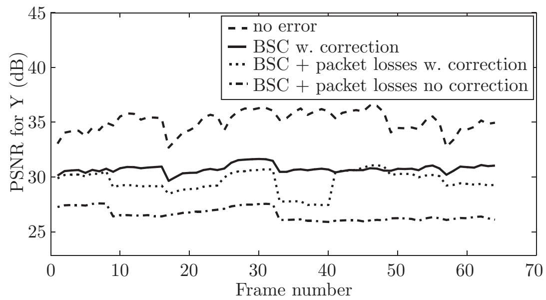

Fig. 6. Performance of the proposed scheme

Figure 5 presents the packet loss rate for the simulated sequence, with the corresponding PSNR performance shown in Figure 6. Table 1 summarizes these results: when tuned without error, the performance of our compression algorithm is sufficient for a practical use, despite the inherent bitrate increase due to the oversampling. Two simulation conditions are also provided. The first is for a BSC only, the second includes also packet losses. Even if the impulse correction is not performed, the scheme that has been tuned for robustness is quite robust to both transmission errors, since there is no error propagation, as the missing samples may be reconstructed from the received samples. Improved results may be obtained after impulse correction, as shown in the last line of Table 1. Moreover, Figure 6 shows the efficiency of the lost packet reestimation technique. Although the missing packets are evaluated using the impulse-corrupted received packets, most of the impulse errors have been eliminated after employing the impulse correction technique. Thus, when compared to the BSC case, for 6.25% of lost packets, only 1 to 2 dB are lost, while 2 to 3 dB are lost for 12.5% of missing packets. Figure 7 illustrates the performance of the proposed technique when both packet losses and bit errors occur.

Fig. 7. Decoded frames without (left) and with (right) correction of the impulse errors (both after estimation of lost samples to packet erasures)

5. CONCLUSIONS

In this paper, a $t + 2D$ video coder based on MC-OFBs has been proposed. The channel model considered in this work is a mixed internet and wireless channel. From simulations, it appears that the video signal can be well reconstructed using the structured redundancy introduced by the OFB for both packet losses and bit errors. In the future, we will extend this work to the MV part, and a more complete video coding scheme will be achieved.

6. REFERENCES

[1] V. K. Goyal, J. Kovacević, and M. Vetterli, "Multiple description transform coding : Robustness to erasures using tight frame expansion," *Proc. of IEEE Int. Symp. on Inofrmation Theory*, p. 408, August 1998.

[2] V. K Goyal, J. Kovacević, and J.A. Kelner, "Quantized frame expansions with erasures," *Journal of Appl. and Comput. Harmonic Analysis*, vol. 10, no. 3, pp. 203–233, 2001.

[3] P. G. Casazza and J. Kovacević, "Equal-norm tight frames with erasures," *Advances in Computational Mathematics, special issue on Frames*, 2002.

[4] I. Bajic and J. Woods, "Maximum minimal distance partitioning of the $\ell^2(\mathbb{Z})$ lattice," *IEEE Trans. on Information Theory*, vol. 49, no. 4, pp. 981–992, April 2003.

[5] J. Nonnenmacher, E. W. Biersack, and D. Towsley, "Parity-based loss recovery for reliable multicast transmission," *IEEE/ACM Transactions on Networking*, vol. 6, no. 4, pp. 349–361, 1998.

[6] B.A. Banister, B. Belzer, and T.R. Fischer, "Robust video transmission over binary symmetric channels with packet erasures," in *Proc. IEEE DCC*, 2002, pp. 162–171.

[7] J. Lacan and J. Fimes, "Systematic MDS erasure codes based on Vandermonde matrices," *IEEE Com. Letters*, vol. 8, no. 9, pp. 570–572, 2004.

[8] F. Labeau, J.C. Chiang, M. Kieffer, P. Duhamel, L. Vandendorpe, and B. Macq, "Oversampled filter banks as error correcting codes: theory and impulse correction," *IEEE Trans. on Signal Processing*, vol. 53, no. 12, pp. 4619–4630, 2005.

[9] J.C. Chiang, M. Kieffer, and P. Duhamel, "Robust image transmission using oversampled filterbank," in *Proceedings of EUSIPCO*, 2004, pp. 1325–1328.

[10] J. C. Chiang, M. Kieffer, and P. Duhamel, "Motion-compensated oversampled filterbanks for robust video coding," *Proceedings of ICASSP*, pp. 917–920, 2005.

[11] P. Chen, *Fully Scalable Subband/Wavelet Coding*, Ph.D. thesis, Rensselaer Polytechnic Institute, 2003, Rensselaer Polytechnic Institute.

[12] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, 1988.

[13] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.

[14] I.V. Bajic and J.W. Woods, "Domain-based multiple description coding of images and video," *IEEE Trans. on Image Processing*, vol. 12, no. 10, pp. 1211–1225, Oct. 2003.

[15] M. Caramma, M. Fumagalli, and R. Lancini, "Polyphase downsampling multiple description coding for IP transmission," *Proceedings of SPIE 2001-Visual Communication and Image Processing*, pp. 545–552, 2001.

[16] M. Harteneck, S. Weiss, and R. W. Stewart, "Design of near perfect reconstruction oversampled filter banks for subband adaptive filters," *IEEE Trans. on Circuit and Systems-II*, vol. 46, no. 8, pp. 1081–1085, August 1999.