Joint Source-Channel coding of scalable video with partially coded index assignment using Reed-Muller codes

Georgia Feideropoulou *, James E. Fowler†, Béatrice Pesquet-Popescu*, and Jean-Claude Belfiore*
*ENST Paris
46 rue Barrault, 75634 Paris, France
{feiderop}@tsi.enst.fr
† Department of Electrical & Computer Engineering
GeoResources Institute, Mississippi State ERC
Mississippi State University, USA

***Abstract*— Joint source-channel coding of scalable video using motion-compensated temporal filtering is considered. The proposed coding scheme consists of a structured vector quantizer based on lattice constellations and a linear index assignment which minimizes simultaneously the channel and source distortions. Both uncoded linear index assignment as well as partially coded linear index assignment via Reed-Muller codes are considered. The proposed system is compared to an unstructured quantizer with minimax index assignment. Simulation results indicate that, for a Gaussian channel, the structured-codebook scheme is very robust, maintaining near-noiseless performance even when the channel is very noisy. Additionally, the proposed structured-quantizer scheme outperforms its unstructured counterpart when channel noise levels are high.**

This work was funded in part by the French Centre National de la Recherche Scientique and by the US National Science Foundation under Grant No. CCR-0310864.

I. Introduction

Joint source-channel coding is recent methodology for the coding of video for situations in which Shannon's separability theorem is not applicable due to high complexity required of both the source and channel coders. Typically, joint source-channel coding has focused on either optimizing channel coding with respect to the source (*source-optimized channel coding*), or optimizing source coding with respect to the channel (*channel-optimized source coding*). An important aspect of joint source-channel coding is the index-assignment problem: the assigning of indices, i.e., codewords, to codevectors so as to reduce the effect of channel errors. Index assignment of codevectors plays an important role in determining the overall performance of quantizers [1], [2].

In this paper, we consider the transmission over noisy channels of scalable video coded using motion-compensated temporal filtering (MCTF). Specifically, we propose a joint source-channel coding scheme based on a structured vector quantization (VQ) of spatio-temporal wavelet coefficients from a $t + 2D$ MCTF decomposition coupled with a linear index assignment mapping VQ codevectors to channel codewords. In [3], it was shown that linear index assignment minimizes the channel distortion for binary symmetric channels. In [4], linear transforms constructed by "maximum component-diversity" lattice constellations are used to build structured VQ codebooks which are proved to minimize simultaneously the source and channel distortion for Gaussian sources. We extended this previous work to the coding of video wherein the source distribution is not Gaussian. Recently [5], we developed a bit-allocation algorithm which takes into consideration a nonnegativity constraint, and we evaluated the performance of our coding scheme on a Gaussian channel using uncoded index assignment.

In this paper, we ameliorate the performance of the coding scheme from [5] using partially coded index assignment based on Reed-Muller codes. We emphasize that the Reed-Muller codes are not appended to the output of our quantizer; rather, the codevectors are mapped directly to codewords which belong to the set of the binary vectors of a Reed-Muller code. We compare our coding scheme, both with and without partially coded index assignment, to a coding scheme consisting of an unstructured codebook produced by VQ training using the generalized Lloyd algorithm (GLA) [6] combined with careful index assignment based on a minimax error criterion [8].

The paper is organized as follows. In the next section, we review the structure of both the coded and uncoded variants of our joint source-channel coding scheme with linear index assignment. In Sec. III, we extend the previous structure to the coding of video, and, in Sec. IV, we present some simulation results. Finally, Sec. V concludes the paper.

II. Joint source-channel coding with uncoded and coded linear index assignment

Let a d-dimensional vector $\boldsymbol{x}$ be the input of a vector quantizer, producing a n-bit binary codeword, which is the index of the vector used for signal reconstruction at the receiver. The source codebook can be viewed as as function of $(b_1...b_n) \in \{+1, -1\}^n$ representing the index assignment. Under the assumption of a maxentropic quantizer, the total distortion D is

$$D = D_s + D_c,$$

where D_s is the distortion due to the quantization, and D_c is the distortion dependent on the index assignment. In [3], it is proved that, for the binary symmetric channel, D_c is minimized by linear labeling. Moreover, in [4], a linear labeling that minimizes the source distortion D_s is constructed. This linear labeling is fully described by a $d \times n$ matrix $\boldsymbol{G}_{d,n}$, where $d \leq n$. For the case of a stationary, memoryless, zero-mean Gaussian source with variance one and a maxentropic source coding, the linear labeling represented by the matrix $\boldsymbol{G}_{d,n}$ necessarily transforms an identically distributed random variable into a random variable (the source codebook) which mimics the source distribution.

Let $\boldsymbol{M}_n$ be an $n \times n$ generator matrix of a "maximum-component-diversity" lattice constellation as described in [7]. Its construction is based on number-field theory, and it is expressed by the standard embeddings in $\boldsymbol{R}^n$ of the ideal ring of the totally real subfield of cyclotomic fields. The matrix $\boldsymbol{G}_{d,n}$ can be constructed as any combination of d rows of the matrix $\boldsymbol{M}_n$. Similar properties are achieved when selecting columns of the matrix $\boldsymbol{M}_n$, and we shall denote by $\boldsymbol{G}'_{n,r}$ the $n \times r$ matrices constructed this way, where $r \leq n$. Further detail on the properties and construction of $\boldsymbol{M}_n$, $\boldsymbol{G}_{d,n}$, and $\boldsymbol{G}'_{n,r}$ is given in [10]; however, we note here that, due to the properties of the matrix $\boldsymbol{M}_n$, the mapping

$$\boldsymbol{b} \in BPSK_n \rightarrow (\boldsymbol{G}_{d,n} \cdot \boldsymbol{b}) \tag{1}$$

is linear, and permits the construction of an asymptotically Gaussian source dictionary.

In the coded case, if C is a (n,k) linear code, then $\boldsymbol{b} = (b_1, \cdots, b_n)$ ranges in $m(\boldsymbol{c})$ where $\boldsymbol{c} = (c_1, \cdots, c_n)$ is one of the 2^k possible binary codewords belonging to C. The function $m(\cdot)$ maps $\boldsymbol{c} = (c_1, \cdots, c_n)$ onto $m(\boldsymbol{c}) = (m(c_1), \cdots, m(c_n))$ with $m(0) = 1$ and $m(1) = -1$. The new codebook with codevectors $\boldsymbol{y}$ is given by $\boldsymbol{y} = \boldsymbol{G}_{d,n}\boldsymbol{b}$, where $\boldsymbol{b}$ now ranges in $m(\boldsymbol{c})$.

III. Coding of Video Sequences

In MCTF decomposition of video, the distribution of spatio-temporal wavelet coefficients is not Gaussian, and a direct application of the structured VQ scheme described above is not appropriate. Thus, in order to take into account a non-Gaussian source distribution, we classify the spatio-temporal coefficients in each subband into two classes, and we adapt the quantizer to each class. In the detail frames, this classification of vectors of coefficients is based on a stochastic model of the spatio-temporal dependencies between coefficients which we introduced in [11]. The vectors of coefficients in the approximation frames are classified according to their norm. We find the source codebook for each class by minimizing

$$\min_{\boldsymbol{b}} \mathrm{E}||\boldsymbol{x} - \beta \boldsymbol{G}_{d,n}\boldsymbol{b}||^2, \tag{2}$$

where $\boldsymbol{b} = (b_1, ..., b_n)^t$, in the uncoded case, ranges in the set of 2^n possible codewords of a $BPSK_n$, and, in the coded case, in the set of 2^k possible codewords of a Reed-Muller (n', k), with $n' = n$, applying the mapping function $m(\cdot)$. $\boldsymbol{G}_{d,n}$ is the matrix obtained as explained above, and β is a parameter which scales the lattice constellation to the source dynamics. In order to find the parameter β as well as the codebook with vectors $\boldsymbol{y} = \boldsymbol{G}_{d,n}\boldsymbol{b}$, an iterative optimization algorithm is applied. A similar optimization is applied when using the matrix $\boldsymbol{G}'_{n,r}$ for VQ; in this case, a Reed-Muller (r,k) is used for coded index assignment.

Recall that a Reed-Muller $RM\{l,m\}$ of order l, is a linear code of block length $n' = 2^m$, dimension k,

$$k = \sum_{i=0}^{l} \binom{m}{i}, \tag{3}$$

and minimum Hamming distance $d_{min} = 2^{m-l}$, where m is a positive integer, and $0 \leq l \leq m$. The error correction capability of an $RM\{l,m\}$ is $\max(0, 2^{m-l-1} - 1)$ errors. An $RM\{l,m\}$ is the set of all binary vectors of length n' associated with the Boolean polynomials of degree at most l.

In previous work [5], we presented an iterative bit-allocation algorithm taking into account a nonnegativity constraint which was proved close to the optimal distribution of the available bits among the spatio-temporal subbands. In the same work, we evaluated the performance of the above coding scheme on a Gaussian channel with uncoded index assignment. However, even if the entire scheme was proved to be efficient in the presence of noise, we noticed that, especially in the case of very low channel SNR, increasing the bitrate did not lead to a significant improvement of the PSNR of the reconstructed video sequence, which was not the case when the channel is noiseless. In order to examine the previous observation, we calculated the end-to-end distortion D in both cases. Obviously, in the noiseless case, the end-to-end distortion is the source distortion D_s, while in the presence of noise, it is $D = D_s + D_c$ as explained in Sec. II. We remarked that for some subbands, especially for the subbands with high energy, D was much higher than D_s, meaning that the channel distortion D_c was dominant for those high-energy subbands.

This observation motivates us in the present work to apply a coded index assignment. We choose to employ Reed-Muller codes due to their symmetry, their widespread use in lattice construction, and their error-correcting capability. In creating the coded index assignment, rather than cascade a Reed-Muller code, we simply restrict the mapping space to be the space of the binary vectors belonging to this code. In this way, we decrease the end-to-end distortion for noisy-channel transmission without changing the source-coding rate of the uncoded case as dictated by the bit-allocation algorithm, and, consequently, without changing the total bitrate. Based on this constraint, the Reed-Muller code should be chosen with consideration of the trade-off between the following three conditions: (1) the error correction capability of the code, (2) the blocklength n' of the code should be equal to n or r (of the matrices $\boldsymbol{G}_{d,n}$ or $\boldsymbol{G}'_{n,r}$, respectively) in order not to increase the bitrate, and (3) the dimension of the code k should be close to n or r (of $\boldsymbol{G}_{d,n}$ or $\boldsymbol{G}'_{n,r}$, respectively) so that the number of 2^k possible codewords is close to the 2^n or 2^r possible codewords of the uncoded case, in order not to increase the source distortion.

Encoding then involves the following steps:

- In each subband, we calculate D_s (in a noiseless environment) and the end-to-end distortion D for the given noisy channel as would be obtained with an uncoded index assignment.
- If the difference between D and D_s is high (which means that D_c is significant), we restrict the index assignment to codewords of $RM\{r,m\}$, taking into consideration the three conditions presented above.
- Otherwise, we apply the mapping to uncoded codewords.

At the decoder side, when we receive coded codewords, we apply soft-decision decoding via the Viterbi algorithm with the BCJR trellis [9]; otherwise, when we receive uncoded codewords, we apply hard-decision decoding.

IV. Simulation Results

We consider a temporal Haar decomposition applied on GOFs of 16 frames, with 4 temporal and 2 spatial resolution levels. The spatial multiresolution analysis uses biorthogonal $9/7$ filters.

The bit-allocation algorithm presented in [5] dictates the size of $\boldsymbol{G}_{d,n}$ or $\boldsymbol{G}'_{n,r}$. The choices for $\boldsymbol{G}_{d,n}$ or $\boldsymbol{G}'_{n,r}$ are, however, limited by computational complexity and the dependences between the spatio-temporal coefficients. It is known that the spatio-temporal coefficients exhibit strong correlation with their spatial or spatio-temporal neighbors. Thus, in order to capture these relationships, the dimensions d in $\boldsymbol{G}_{d,n}$ or n in $\boldsymbol{G}'_{n,r}$ should be a power of 4. However, in order to attain high or low coding rates, we allow d or r, respectively, to be equal to 2. In addition, in order to keep the complexity low, we limit the dimensions n in $\boldsymbol{G}_{d,n}$ and r in $\boldsymbol{G}'_{n,r}$ to be no greater than 16. For our tests, we consider CIF (352×288) test sequences at 30 fps.

In following the coding procedure presented in Sec. III, we noticed that partially coded index assignment is necessary for the subbands with high energy. Most of these subbands are part of the temporal approximation frames at the coarsest spatial-resolution level. However, as the bitrate increases, we remark that the low-frequency subbands of some of the temporal detail frames, especially at the coarser temporal-resolution levels, need to be protected.

In our coding scheme, the subbands with partially coded index assignment are quantized according to the bit allocation algorithm with $\boldsymbol{G}_{d,16}$ or $\boldsymbol{G}'_{n,16}$, and thus $RM\{2,4\}$ with $n' = 16$ and $k = 11$ proved to be the most flexible compromise between the three conditions discussed previously in Sec. III. However, compared to the uncoded case, we expect the source distortion to increase, as now the dimension of the mapping space has decreased to 2^{11} possible codewords, instead of 2^{16} as in the uncoded case. On the contrary, the end-to-end distortion of the entire scheme in a noisy environment will decrease significantly compared to the uncoded case, and the total bitrate remains the same.

For comparison, we consider an alternative joint source-channel coding scheme—straightforward unstructured VQ of spatio-temporal MCTF coefficients coupled with careful index assignment. In this approach, a good, albeit suboptimal, minimax index assignment based on a polynomial-time algorithm, the Minimax Cover Algorithm (MCA) [8], is used. In adopting a minimax error criterion, a channel code is designed against the worst-case performance, without sacrificing the average performance [8]. We applied the minimax algorithm to an unstructured codebook generated by GLA [6] trained on our video sequences. The total bitrates (in Kbs) for this unstructured scheme are the same as the ones produced by other, structured coding schemes. The size of the unstructured codebook for each subband is as prescribed by the bit-allocation algorithm of [5].

In Tables I and II, we present the results of the three coding schemes. In these tables, JSC refers to our first coding scheme with linear uncoded index assignment [5], JSC+RM is our second coding scheme using partially coded index assignment with $RM\{2,4\}$ as proposed here, and finally, MCA is the coding scheme using an unstructured codebook produced by GLA coupled with minimax index assignment.

“hall-monitor” 1660 Kbs			
	noiseless	SNR=4.33	SNR=6.75
JSC	35.59	27.23	33.47
JSC+RM	N/A	33.45	34.18
MCA	40.31	24.15	33.13
“hall-monitor” 1113 Kbs			
	noiseless	SNR=4.33	SNR= 6.75
JSC	34.38	26.96	32.68
JSC+RM	N/A	32.08	33.36
MCA	38.03	23.90	32.54
“hall-monitor” 565 Kbs			
	noiseless	SNR=4.33	SNR=6.75
JSC	31.78	26.41	30.81
JSC+RM	N/A	30.29	31.15
MCA	35.36	23.78	31.61
“hall-monitor” 317 Kbs			
	noiseless	SNR= 4.33	SNR=6.75
JSC	27.88	23.91	27.22
JSC+RM	N/A	25.02	27.22
MCA	32.81	23.54	30.31

TABLE I

AVERAGE PSNR (IN DB) OF THE RECONSTRUCTED SEQUENCE “HALL-MONITOR” AT DIFFERENT BITRATES AND DIFFERENT CHANNEL SNR. NO CODED IA IS APPLIED IN THE NOISELESS CASE.

We remark that the choice of a partially coded index assignment via $RM\{2,4\}$ is justified by the considerable increase in PSNR of the reconstructed sequence as well as by the robustness of the new coding scheme in the presence of noise as compared to the uncoded case. As expected, due to the unstructured nature of the codebook generated by GLA, in the noiseless case, the PSNR of the MCA scheme is quite high as compared to that of the structured-codebook schemes. On the other hand, as the channel noise increases, the PSNR for the unstructured MCA scheme drops dramatically despite the index assignment provided by the minimax algorithm. Meanwhile, one can clearly observe that both our coding schemes, especially JSC+RM, remain quite close to

"foreman" 1530 Kbs			
	noiseless	SNR=4.33	SNR=6.75
JSC	34.51	26.49	32.63
JSC+RM	N/A	32.03	32.85
MCA	38.72	25.62	34.06
"foreman" 1104 Kbs			
	noiseless	SNR=4.33	SNR= 6.75
JSC	33.30	26.27	31.72
JSC+RM	N/A	31.33	32.00
MCA	37.44	25.59	33.16
"foreman"524 Kbs			
	noiseless	SNR=4.33	SNR=6.75
JSC	30.78	25.66	29.86
JSC+RM	N/A	29.51	30.01
MCA	34.80	25.22	32.08
"foreman" 376 Kbs			
	noiseless	SNR= 4.33	SNR=6.75
JSC	30.07	25.53	29.27
JSC+RM	N/A	29.05	29.40
MCA	33.64	24.99	31.38

TABLE II

AVERAGE PSNR (IN DB) OF THE RECONSTRUCTED SEQUENCE "FOREMAN" AT DIFFERENT BITRATES AND DIFFERENT CHANNEL SNR. NO CODED IA IS APPLIED IN THE NOISELESS CASE.

their noiseless performance, even for very noisy channels (i.e., low SNR). Indeed, JSC+RM using partially coded index assignment significantly outperforms both the uncoded JSC approach as well as the unstructured MCA technique when the channel SNR is very low.

V. CONCLUSION

This paper presents an approach to joint source-channel coding of scalable $t + 2D$ video. The coding scheme is based on a structured vector quantizer constructed by lattice constellations and a linear index assignment. We made the distinction among two cases of interest: uncoded index assignment, and partially coded index assignment via Reed-Muller codes. Simulation results revealed a high improvement in performance for the latter scheme, especially at low channel SNR. Additionally, both of these coders were compared to an unstructured quantizer using careful index assignment. The coding schemes using a structured quantizer proved to be much more robust than the coder based on an unstructured quantizer, the proposed structured coders obtaining between 1- and 5-dB greater average PSNR over very noisy channels.

REFERENCES

[1] A. Mehes and K. Zeger, "Binary Lattice Vector Quantization with Linear Block Codes and Affine Index Assignments," IEEE Trans. Inform. Theory, vol. 44, no. 1, pp. 79-94, Jan. 1998.

[2] R. Hagen and P. Hedelin, "Robust Vector Quantization by a Linear Mapping of a Block Code," IEEE Trans. Inform. Theory, vol. 45, no. 1, pp. 200-218, Jan. 1999.

[3] P. Knagenhjelm and E. Agrell, "The Hadamard Transform - A Tool for Index Assignment," IEEE Trans. Inform. Theory, vol. IT 42, pp 1139-1151, July 1996.

[4] J. C. Belfiore, X. Giraud, and J. Rodriguez-Guisantes, "Optimal Linear Labeling for the Minimization of both Source and Channel distortion," ISIT 2000, Sorrento, Italy, June 25-30, 2000.

[5] G. Feideropoulou, B. Pesquet-Popescu, and J. C. Belfiore, "Bit Allocation Algorithm for Joint Source-Channel Coding of $t + 2D$ Video Sequences," accepted for publication in IEEE ICASSP'05, Philadelphia, USA, March 2005.

[6] Y. Linde, A.Buzo, and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun., vol. COM-28, pp. 84-95, Jan. 1980.

[7] X. Giraud, E. Boutillon, and J.-C. Belfiore, "Algebraic Tools to Build Modulation Schemes for Fading Channels," IEEE Trans. Inform. Theory, vol. IT 43, pp.938-952, May 1997.

[8] L. C. Potter and D.-M. Chiang, "Minimax Nonredundant Channel Coding," IEEE Trans. Commun. vol. 43, no.4, pp. 804-811, Apr. 1995.

[9] R. J. McEliece, "On the BCJR Trellis for Linear Block Codes," IEEE Trans. Inform. Theory, vol. 42, no. 4, pp. 1072-1091, Jul. 1996.

[10] G. Feideropoulou, B. Pesquet-Popescu, and J. C. Belfiore, "Joint Source-Channel Coding of Scalable Video," in IEEE GLOBECOM'04, Dallas, Texas, Nov. 2004.

[11] G. Feideropoulou, B. Pesquet-Popescu, and J. C. Belfiore and J. Rodriguez, "Non-linear modeling of wavelet coefficients for a video sequence," Proc. IEEE Int. Workshop NSIP03, Grado, Italy, June 8-11, 2003.