14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

A PRACTICAL ALGORITHM FOR DISTRIBUTED SOURCE CODING BASED ON CONTINUOUS-VALUED SYNDROMES

Lorenzo Cappellari, and Gian Antonio Mian

Dept. of Information Engineering, University of Padova
via Gradenigo 6/B, 35131 Padova, Italy
email: lorenzo.cappellari@dei.unipd.it, mian@dei.unipd.it

ABSTRACT

Recent advances in video coding are based on the principles of distributed source coding to both relax the inherent complexity of classical encoding algorithms and offer robustness against transmission errors. Several practical frameworks for distributed source coding take advantage of channel coding principles to encode a suitably pre-quantized version of the source. In this paper, we present an alternative two-step approach for the problem of coding a source which is correlated with another one that is available only at the decoder. First, based on the correlation of the side information, a continuous-valued syndrome is computed. Then, according to the given rate-constraint, the syndrome is encoded and transmitted as in classical source coding. The great flexibility of the coding system is confirmed by the simulation results: for source coding of i.i.d. Gaussian sources, the performance is always within 4 dB from the Wyner-Ziv bound, for a wide range of side information correlations and target bit-rates.

1. INTRODUCTION

The video coding architectures have been driven predominantly by the "downlink" transmission model of TV broadcast. Consequently, in a traditional digital video encoding-decoding scheme, there is a computationally heavy encoder and a relatively light decoder. The high computational power required by the encoder is due to the fact that the best coding parameters must be found by exploiting the temporal and spatial correlation of the frames in the video sequence in an optimal manner, and it is dominated by the complexity of the motion compensated prediction task.

Nowadays, we are instead assisting to the emergence of applications which require an "uplink" transmission of digital video, such as for example video-phone calling and surveillance with low-power video-sensors. In this framework, we need low-complexity encoding algorithms (e.g. to prolong battery life), with high compression efficiency (due to the strict bandwith constraints) and robustness to channel losses.

A promising approach to simultaneously address these requirements is offered by distributed source coding. In this framework, several correlated sources are encoded independently but jointly decoded. In some case, it was shown that for a given distortion the rate needed by the "blind" encoders equals the rate that would be needed when using encoders that have perfect knowledge about the correlated sources. When thinking at the frames of a video sequence as the correlated sources, it should hence be possible to code them independently (i.e. without motion compensation) as *still* images while obtaining the same performance. Such principles are exploited in the PRISM coder [1], which is able to offer high compression ratios by transferring the heavy motion compensation task from the encoder to the decoder.

At the core of this video application, there is the problem of coding a random source (e.g. the current frame) which is correlated with another source (e.g. the previous frame) being known at the decoder but *not* at the encoder, on which is the focus of this work.

The rest of the paper is organized as follows. In Section 2, we review the current syndrome-based solutions to the source coding problem with side information. In Section 3, we present our proposed solution, which is based on the extension of the concept of syndrome to the continuous Euclidean space. We show the theoretical asymptotic optimality of this coding scheme in Section 4. The results of the experiments, with several numerical results in the Gaussian case, are presented in Section 5. Section 6 gives the final remarks on this work and concludes the paper.

2. DISTRIBUTED SOURCE CODING USING SYNDROMES

Consider the case where X and Y are two correlated and stationary memoryless sources, and we have to compress X, with Y (referred to as side information) being known at the decoder but *not* at the encoder (however, both the encoder and the decoder are assumed to have a perfect knowledge on the joint distribution of X and Y). If Y was known at both ends, then the encoder could exploit all the correlation between the two sources. Hence, the optimal solution would be represented by encoding the residual after *prediction* of X from Y only. In this case, the achievable distortion-rate function is the distortion-rate function $R_{X|Y}(D)$ of the random variable $X|Y$ [2].

If Y is known only at the decoder, then Wyner and Ziv have shown that the achievable distortion-rate function, denoted as $R^{WZ}_{X|Y}(D)$, equals $R_{X|Y}(D)$ only in particular cases, while in general $R^{WZ}_{X|Y}(D) \geq R_{X|Y}(D)$ [3]. Namely, the strict equality holds in the case of X and Y being jointly Gaussian [3], or in the case of X and Y being discrete, but with $D = 0$ [4]. Hence, at least in some cases, the limit $R_{X|Y}(D)$ is theoretically achievable. However, practical schemes towards this objective have only recently appeared.

As example, the DISCUS (*Distributed Source Coding Using Syndromes*) system [5] takes successfully advantage of channel coding principles to solve the source coding problem with side information at the decoder. In this approach, Y is seen as the "noisy" version of X produced by a virtual "correlation channel". Then, provided that an n-dimensional block of input data $\mathbf{X}$ is a codeword of some block code which is

14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

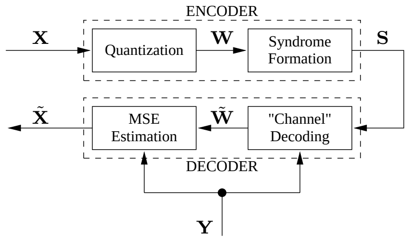

Figure 1: Wyner-Ziv encoder and decoder in the case of a discrete-valued syndrome (DISCUS system).

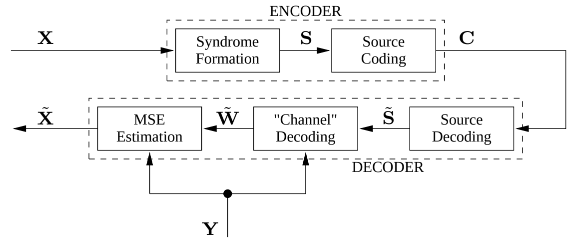

Figure 2: Wyner-Ziv encoder and decoder in the case of a continuous-valued syndrome.

good for channel coding, a "channel decoder" can recover the correct value of X with an arbitrary small probability of error. To make $\mathbf{X}$ appear like a codeword, two operations are needed at the encoder. As shown in Fig. 1, the first one is a "lossy" quantization (*scalar quantization*, SQ, or *trellis coded quantization*, TCQ [6]), which outputs a block of data $\mathbf{W}$ which can be put into one-to-one correspondence with the vectors in a subset of $\mathbb{Z}^n$. After partitioning the space $\mathbb{Z}^n$ into a *linear code* $\mathscr{C}_0$ and its *cosets* $\mathscr{C}_i, i = 1,2,\ldots,N-1$ (or into a *generalized coset code*, like for example the *trellis code* induced by TCQ, and its *label translates* [7]), the second operation "losslessly" codes the index i such that $\mathbf{W} \in \mathscr{C}_i$, which in the channel coding lingo is known as the *syndrome* of $\mathbf{W}$ w.r.t. the code $\mathscr{C}_0$. The code and the number of cosets N can be potentially chosen in a way such that the necessary rate to code i equals $R^{WZ}_{W|Y}(0)$ and $\mathbf{W}$ can be reconstructed with an arbitrary small probability of error from $\mathbf{Y}$.

At the decoder, after "channel decoding" of an n-dimensional block $\mathbf{Y}$ using the code $\mathscr{C}_i$ (note that for channel coding purposes each coset $\mathscr{C}_i$ is equivalent to the code $\mathscr{C}_0$), we obtain an almost exact approximation $\tilde{\mathbf{W}}$ of $\mathbf{W}$. Optimal linear MSE (*mean square error*) estimation leads then to the recovering of $\mathbf{X}$ using both $\mathbf{Y}$ and $\tilde{\mathbf{W}} \approx \mathbf{W} = \mathbf{X} - \mathbf{Q}$, where the *quantization noise* $\mathbf{Q}$ is assumed independent of $\mathbf{X}$.

3. CONTINUOUS-VALUED SYNDROMES

The great interplay between the quantizer and the syndrome former of Fig. 1 governs the final performance in terms of MSE, and makes somewhat difficult the ad-hoc design of the DISCUS system for specific "correlation channels" and transmission rates. For this reason, we propose an alternative approach which in a certain way swaps the two operations at the encoder and eventually offers simplified design rules.

As shown in Fig. 2, the first operation in the proposed encoder is the syndrome formation. In particular, we consider a trellis coded quantizer $Q_2[\cdot]$ based on the *geometrically uniform* partition $a\mathbb{Z}/4a\mathbb{Z}$, and the corresponding *unbounded* trellis code $\mathscr{C}$. The n-dimensional *Voronoi cells* of $\mathscr{C}$ are all similar each other and have the same *volume* V_2; hence they have the same *second moment per dimension* σ_2^2, the same *normalized volume* $V_{n2} \triangleq V_2^{2/n}$ and the same *normalized second moment* $G_2 \triangleq \sigma_2^2/V_{n2}$ [8]. Since it is known that, with a sufficiently high number of states in TCQ, G_2 becomes asymptotically very close to the normalized second moment of the sphere, it is reasonable to assume that the Voronoi cells of $\mathscr{C}$ become almost spheric (hence invariant under rotation). Consequently, the *translates* $\mathscr{C}_\mathbf{S} \triangleq \mathscr{C} + \mathbf{S}$ of the trellis code, for $\mathbf{S}$ belonging to the *basic* Voronoi cell of $\mathscr{C}$, asymptotically form a partition of $\mathbb{R}^n$. The value of $\mathbf{S}$ such that $\mathbf{X} \in \mathscr{C}_\mathbf{S}$, which simply satisfies

$$\mathbf{S} = \mathbf{X} - Q_2[\mathbf{X}] \quad , \tag{1}$$

is hence what we call a *continuous-valued* syndrome. In the case of $Y = X + N$ with $X \sim \mathscr{N}(0,\sigma_x^2)$, $N \sim \mathscr{N}(0,\sigma_n^2)$ and $N \perp X$, we choose a in a way such that σ_2^2 is proportional to $\sigma^2_{X|Y} = (1/\sigma_x^2 + 1/\sigma_n^2)^{-1}$, i.e. we set

$$V_{n2} = K\frac{\sigma^2_{X|Y}}{G_2} \quad , \tag{2}$$

where K is a proportionality *volumetric factor* and G_2 is set as the experimentally measured value of the normalized second moment relative to the used TCQ (which confirms the results in [6]).

The reason for this choice comes from the fact that the theoretical rate-distortion function in this setup equals [3]

$$D^{WZ}_{X|Y}(R) = D_{X|Y}(R) = \sigma^2_{X|Y}2^{-2R} \quad . \tag{3}$$

As we will show later, the distortion on the reconstructed source equals essentially the distortion that we commit on the syndrome. Hence, for a fixed rate, it is reasonable for the variance of the syndrome to be proportional to $\sigma^2_{X|Y}$ as in (2).

The syndrome $\mathbf{S}$ is then coded and decoded into $\tilde{\mathbf{S}}$ as in traditional source coding, according to the desired transmission rate. In particular, we again use a (*dithered*) trellis coded quantizer $Q_1[\cdot]$, obtaining

$$\tilde{\mathbf{S}} = Q_1[\mathbf{S} + \mathbf{Z}] - \mathbf{Z} \quad , \tag{4}$$

where $Z \sim \mathscr{N}(0,\sigma_z^2)$, $Z \perp S$, is the *dither*, known at the decoder. Denote with V_1, σ_1^2, V_{n1} and G_1 respectively the volume, the second moment, the normalized volume and the normalized second moment of $Q_1[\cdot]$. To satisfy the rate constraint R (in bit/sample), the volume ratio V_2/V_1 must equal 2^{nR}, and hence we set

$$V_{n1} = V_{n2} \cdot 2^{-2R} \quad ; \tag{5}$$

to guarantee that the quantization noise $\mathbf{Q} = \mathbf{S} - \tilde{\mathbf{S}}$ is independent of $\mathbf{S}$, we set $\sigma_z^2 = \sigma_1^2$.

At the decoder, a "channel decoder" then obtains an estimate of $\mathbf{X}$ from $\mathbf{Y}$ belonging to $\mathscr{C}_{\tilde{\mathbf{S}}}$. In the previously described setup, where N is Gaussian, the best MAP (*maximum a posteriori probability*) decoder is based on the minimum distance, and hence the "channel decoding" of $\mathbf{Y}$ is done through

$$\tilde{\mathbf{W}} = Q_2[\mathbf{Y} - \tilde{\mathbf{S}}] + \tilde{\mathbf{S}} \quad . \tag{6}$$

Then, assuming that with high probability $\tilde{\mathbf{W}} \approx \mathbf{X} - \mathbf{Q}$, where $Q \sim \mathcal{N}(0, \sigma_1^2)$, and $Q \perp X, N$, optimum linear MSE estimation reconstructs X through

$$\tilde{\mathbf{X}} = a_1 \mathbf{Y} + a_2 \tilde{\mathbf{W}} \quad , \tag{7}$$

where the optimum coefficients equal

$$a_1 = \frac{1}{1 + \frac{\sigma_n^2}{\sigma_x^2} + \frac{\sigma_n^2}{\sigma_1^2}} \quad a_2 = \frac{1}{1 + \frac{\sigma_1^2}{\sigma_x^2} + \frac{\sigma_1^2}{\sigma_n^2}} \quad . \tag{8}$$

4. THEORETICAL OPTIMALITY

Let the side information be $Y = X + N$ with X and N Gaussian as in the previous section. Hence [3]

$$R_{X|Y}^{WZ}(D) = \frac{1}{2}\log_2 \frac{\sigma_{X|Y}^2}{D} \quad 0 < D \le \sigma_{X|Y}^2 \quad . \tag{9}$$

If the quantizer of the DISCUS system was a lattice quantizer based on an *unbounded* lattice $\mathcal{L}_1$ and the code $\mathcal{C}_0$ was equal to a lattice $\mathcal{L}_2 \subseteq \mathcal{L}_1$ (without considering the mapping on $\mathbb{Z}^n$), then the asymptotic optimality of the system would be guaranteed, at least for $\sigma_n^2 \ll \sigma_x^2$, as shown in [9].

However, with similar arguments, *it is possible to prove the same result even in the presented case of continuous-valued syndromes*, where in general there is not any inclusion relation between the reconstruction points of the n-dimensional vector quantizers $Q_1[\cdot]$ and $Q_2[\cdot]$. In particular, assume that in correspondence of a target distortion D the following properties are satisfied (with $\varepsilon > 0$) for a sufficient large n:

1. the reconstruction points of the vector quantizer $Q_i[\cdot]$, $i = 1, 2$, form a geometrically uniform set [7] whose volume, second moment, normalized volume and normalized second moment are respectively V_i, σ_i^2, V_{ni} and G_i;
2. the normalized second moments satisfy $\log_2(2\pi e G_i) < \varepsilon$ (and obviously $G_i > 1/2\pi e$ [8]), $i = 1, 2$, (i.e. both quantizers have sphere-like Voronoi cells);
3. the second moment of $Q_1[\cdot]$ satisfies $\sigma_1^2 = (1/D + 1/\sigma_{X|Y}^2)^{-1}$ (i.e. $\sigma_1^2 \approx D$ for $D \ll \sigma_{X|Y}^2$);
4. the second moment of $Q_2[\cdot]$ satisfies $\sigma_2^2 \le \sigma_n^2 + \sigma_1^2 + \varepsilon$;
5. the quantizer $Q_2[\cdot]$ is such that $P[Q_2[c + \mathbf{N} + \mathbf{Z}^*] \ne c] < \varepsilon$, for each c s.t. $Q_2[c] = c$, where $Z^* \sim \mathcal{N}(0, \sigma_1^2)$, and $Z^* \perp N$.

These assumptions lead to the following. First of all, the rate of transmission satisfies

$$\begin{aligned} R &= \frac{1}{n}\log_2 \frac{V_2}{V_1} = \frac{1}{2}\log_2 \frac{G_1 \sigma_2^2}{G_2 \sigma_1^2} \le \frac{1}{2}\log_2 2\pi e G_1 \frac{\sigma_2^2}{\sigma_1^2} \le \\ &\le \frac{1}{2}\log_2\left(1 + \frac{\sigma_n^2}{\sigma_1^2} + \frac{\varepsilon}{\sigma_1^2}\right) + \frac{\varepsilon}{2} \le \\ &\le \frac{1}{2}\log_2 \frac{\sigma_n^2}{D}\left(\frac{D}{\sigma_n^2} + \frac{D}{\sigma_1^2} + \frac{D\varepsilon}{\sigma_1^2\sigma_n^2}\right) + \frac{\varepsilon}{2} \le \\ &\le \frac{1}{2}\log_2 \frac{\sigma_n^2}{D}\left(\frac{D}{\sigma_{X|Y}^2} + \frac{D}{\sigma_1^2} + \frac{D\varepsilon}{\sigma_1^2\sigma_n^2}\right) + \frac{\varepsilon}{2} = \\ &= \frac{1}{2}\log_2 \frac{\sigma_n^2}{D} + \mathcal{O}(\varepsilon) \quad , \end{aligned} \tag{10}$$

and hence goes to $R_{X|Y}^{WZ}(D)$ as $\varepsilon \to 0$ if $\sigma_n^2 \ll \sigma_x^2$. Then, with probability higher than $1 - \varepsilon$,

$$\begin{aligned} \tilde{\mathbf{W}} &= Q_2[(\mathbf{X} + \mathbf{N}) - (\mathbf{S} - \mathbf{Q})] + (\mathbf{S} - \mathbf{Q}) = \\ &= (\mathbf{X} - \mathbf{S}) + (\mathbf{S} - \mathbf{Q}) = \mathbf{X} - \mathbf{Q} \quad , \end{aligned} \tag{11}$$

where Q resembles asymptotically a Gaussian process with variance σ_1^2, uncorrelated with X and N (and it is reasonable to have a dither $Z \sim \mathcal{N}(0, \sigma_1^2)$ in quantizing S) [10]. Finally, with the coefficients of (8), it is easy to show that with the same probability

$$\sigma_e^2 = \frac{1}{n}E[\|\tilde{\mathbf{X}} - \mathbf{X}\|^2] = \frac{1}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_{X|Y}^2}} = D \quad , \tag{12}$$

and hence the Wyner-Ziv bound is reached as $\varepsilon \to 0$.

While properties 1, 2, and 3 are easily fulfilled by TCQ with sufficiently high n and number of states, properties 4 and 5 seem to be hard to be jointly satisfied. However, if $Q_2[\cdot]$ is asymptotically an optimal lattice quantizer, then the quantization noise resembles a Gaussian with variance σ_2^2 [10]. Hence, choosing $\sigma_2^2 \ge \sigma_n^2 + \sigma_1^2$, it is reasonable that property 5 holds while guaranteeing property 4.

5. EXPERIMENTAL RESULTS

The aim of the following preliminary experiments is to show both the flexibility and the optimality of the system proposed in Section 3. In all the experiments, the results refer to the Gaussian case previously analyzed, where $Y = X + N$. The Correlation-*SNR* (or C-SNR) is measured as the ratio σ_x^2/σ_n^2 and the data are averaged over 100 sequences whose length is $n = 10^4$.

First of all, we are interested to the right choice for the factor K to be used in (2). The solution is not straightforward. In fact, high values of K guarantee that $\tilde{\mathbf{W}} \approx \mathbf{X} - \mathbf{Q}$, but they imply very high values of σ_1^2, which is the variance of Q. In the opposite case, a low K value leads to a low variance of Q, but the probability that $\tilde{\mathbf{W}} \ne \mathbf{X} - \mathbf{Q}$ greatly increases.

Fig. 3 shows the performance of the system as a function of the C-SNR for various K. Both $Q_1[\cdot]$ and $Q_2[\cdot]$ are 8-state trellis coded quantizers. As expected, there is an optimum value of K. As example, a value $K = 2.2$ improves the performance w.r.t. to the case $K = 1.8$, but further increasing of K again leads to a poorer performance. When analyzing the performance of the system as a function of the rate, it can be however noted that higher values of K improve the system performance at medium bit-rates, as shown in Fig. 4. This effect is particularly evident when the C-SNR is low, as shown in Fig. 5.

When the best value is chosen for K, the experimental probability of $\tilde{\mathbf{W}}$ being different from $\mathbf{X} - \mathbf{Q}$ is about 10^{-3}, as obtained in the DISCUS system when it achieves the best performance w.r.t. the theoretical bound. However, in the DISCUS system the best performance is obtained for specific rates and correlations only, while in our scheme the measured performance becomes then about 4 dB (or equivalently $2/3$ of bit) far from the Wyner-Ziv bound, for any correlation and any rate.

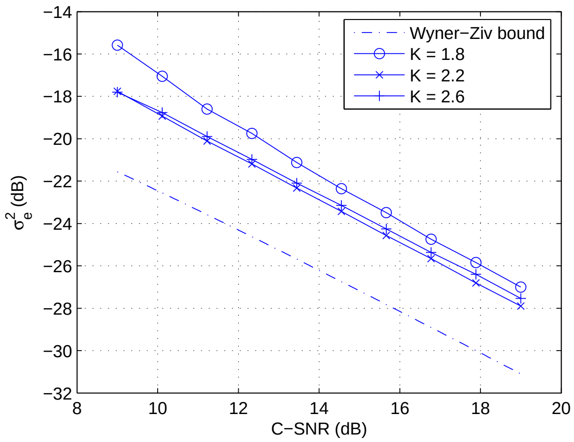

Figure 3: Average distortion at $R = 2$ bit/sample (with 8-state TCQ), for various values of the volumetric factor K.

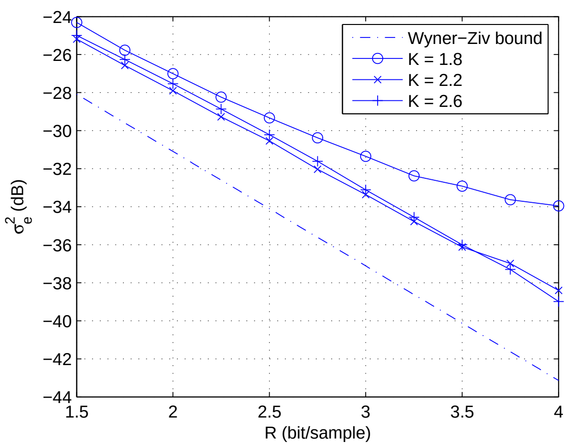

Figure 4: Average distortion at Correlation-$SNR = 19.0$ dB (with 8-state TCQ), for various values of the volumetric factor K.

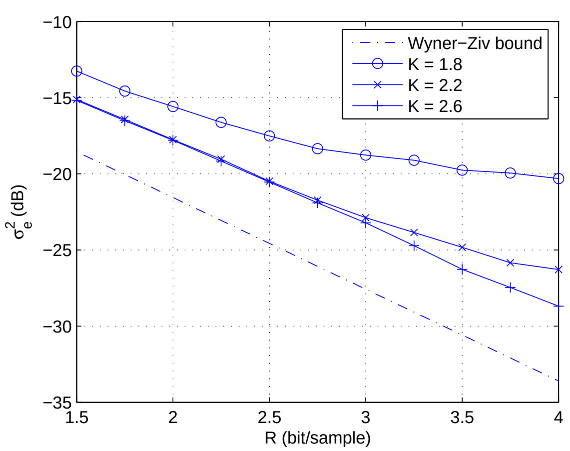

Figure 5: Average distortion at Correlation-$SNR = 9.0$ dB (with 8-state TCQ), for various values of the volumetric factor K.

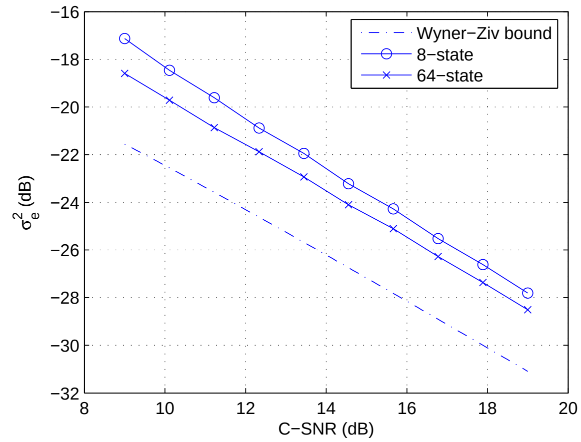

Figure 6: Average distortion at $R = 2$ bit/sample (with $K = 2$).

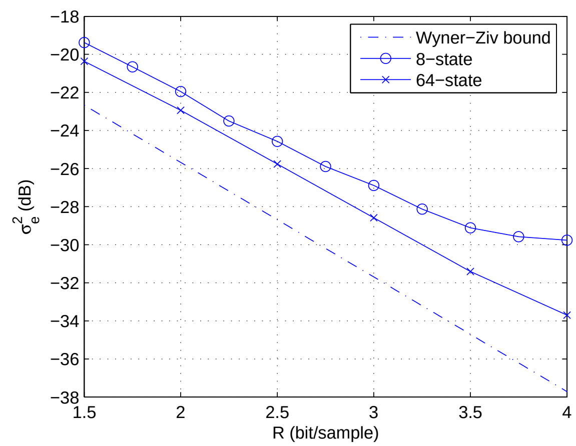

Figure 7: Average distortion at Correlation-$SNR = 13.4$ dB (with $K = 2$).

In the previous experiments, 8-state TCQ was used with the goal to maintain a very low computational complexity. However, the performance of such TCQ in terms of normalized second moment is quite far from the spheric-Voronoi cell limit (about 0.5 dB). Hence, we expect that further improvements can be achieved by increasing the number of states. As example, if both $Q_1[\cdot]$ and $Q_2[\cdot]$ are 64-state trellis coded quantizers, the performance already increases of about 1 dB w.r.t. the 8-state case. In particular, that improvement is more visible in the low correlation region, for a fixed value of the rate (as in Fig. 6), and in the high bit-rate region, for a fixed C-SNR (as in Fig. 7). It is worth noting that these results are obtained with a suboptimal value of K; it is however very probable that we reach similar results even with a more careful choice of K.

In designing the system, we simply use (5) to set the normalized volume of the *finer* quantizer $Q_1[\cdot]$ as a function of the target rate R. To show that, after syndrome quantization, the necessary rate is actually almost equal to the target R, in Fig. 8 we show the experimental first order entropy of the bit-stream output by TCQ. To compute this value, the experimental entropies of the *coded* and of the *uncoded* bits are summed together, and hence they are assumed to be inde-

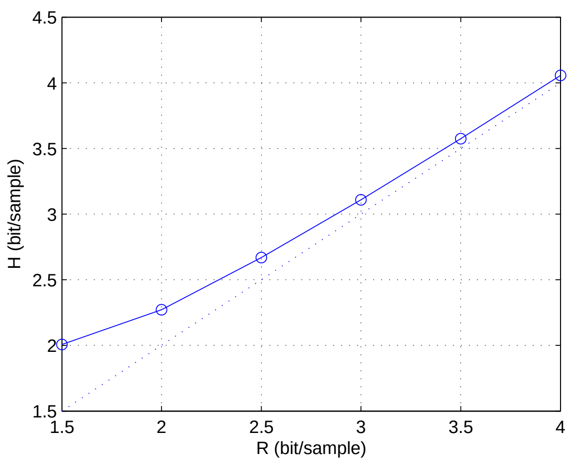

Figure 8: Estimated first-order entropy of the binary codeword **C** sent to the decoder.

pendent. At a first glance, it can be noted that at low bit-rates the entropy is actually greater than the target bit-rate R, but that at medium and high bit-rates essentially the two quantities are the same.

This phenomenon is explained by two reasons. First, at low bit-rates there are more side effects caused by the independent shapes of the Voronoi cells of the two quantizers, that lead to a final number of cells used to quantize the syndrome greater than the expected value V_2/V_1. Then, it is very likely that at low bit-rates the *coded* and the *uncoded* bits are not independent and hence that the measured entropy is overestimated. Consequently, it is likely that, as example, an ad-hoc context-based arithmetic coder leads to an average bit-rate which is eventually less than the estimated entropy, because it can better exploit the joint and temporal correlation of the *coded* and the *uncoded* bits.

6. CONCLUSION

In this paper, we presented a novel practical coding scheme for the source coding problem with side information at the decoder. The encoder consist of two cascaded blocks. The first one is a syndrome extractor and the second one is a classical source encoder.

In the first block, a continuous-valued syndrome is formed according to the known correlation between the variable to be coded X and the side information Y. We use TCQ based on the partition $a\mathbb{Z}/4a\mathbb{Z}$ to form this syndrome, and hence the complexity of the operation linearly increases with the number of samples and the number of states. Moreover, since the syndrome formation is simply driven by the value of a, it is straightforward to adapt the system to a changing X-Y correlation.

The second block allows for coding of the syndrome with the desired transmission rate R, that can be easily adapted to the changing transmission channel conditions. We again employ TCQ, but any classical source encoder can be used. Furthermore, there is the possibility to code the *same* syndrome at different rates, using embedded quantizers. Hence, the scheme is very suitable for *quality scalable* transmission.

The decoder consists of a "virtual" channel decoder and of a linear estimator (without considering the source decoder for the coded syndrome). The channel decoder is based on TCQ and hence its complexity equals the complexity of the syndrome extractor of the encoder. Both the channel decoder, driven by a, and the linear estimator, driven by the coefficients a_1 and a_2, can be easily adapted to the changing X-Y correlation.

The performance of the proposed coding system is shown for the jointly Gaussian case, using a wide range of transmission rates and correlations. In any case, the reconstruction error is within 4 dB from the theoretical bound, showing the increased flexibility w.r.t. the discrete-valued syndrome based system (DISCUS) [5]. Even if the DISCUS system can reach a performance as near to the Wyner-Ziv bound as 2.1 dB at rate 1 bit/sample or 3.2 dB at 2 bit/sample, that result holds *only* for a specific correlation. Moreover, this system is constrained to use integer bit-rates, and changing the rate implies to redesign the system.

As a final remark, while in DISCUS the encoder is tailored to the Gaussian distribution of X, in the proposed system the encoder is only based on the X-Y correlation, and it is likely to work even in the case of a general distribution. Consequently, its straightforward application to the distributed video coding framework seems to be very promising.

REFERENCES

[1] R. Puri and K. Ramchandran, "PRISM: a "reversed" multimedia coding paradigm," in *Proc. of IEEE Intl. Conf. on Image Processing*, 14-17 September 2003, vol. 1, pp. 617–620.

[2] T.M. Cover and J.A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, NY, USA, 1991.

[3] A.D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.

[4] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory*, vol. 19, no. 4, pp. 471–480, July 1973.

[5] S.S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): design and construction," *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.

[6] M.W. Marcellin and T.R. Fisher, "Trellis coded quantization of memoryless and gauss-markov sources," *IEEE Trans. Commun.*, vol. 38, no. 1, pp. 82–93, Jan. 1990.

[7] G.D. Forney, "Geometrically uniform codes," *IEEE Trans. Inform. Theory*, vol. 37, no. 5, pp. 1241–1260, Sept. 1991.

[8] J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, Springer-Verlag, New York, NY, USA, 1988.

[9] R. Zamir and S. Shamai, "Nested linear/lattice codes for Wyner-Ziv encoding," in *Proc. of IEEE Inform. Theory Workshop*, 22-26 June 1998, pp. 92–93.

[10] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Inform. Theory*, vol. 42, no. 4, pp. 1152–1159, July 1996.