This item is the archived peer-reviewed author-version of:

Rate-distortion driven decoder-side bitplane mode decision for distributed video coding

Jürgen Slowack[a], Stefaan Mys[a], Jozef Škorupa[a], Nikos Deligiannis[b],
Peter Lambert[a], Adrian Munteanu[b], Rik Van de Walle[a]

[a] *Ghent University – IBBT, Department of Electronics and Information Systems (ELIS)
– Multimedia Lab, Gaston Crommenlaan 8 bus 201, B-9000 Ghent, Belgium*
[b] *Vrije Universiteit Brussel – IBBT, Electronics and Informatics Department (ETRO),
Pleinlaan 2, B-1050 Brussels, Belgium*

Abstract

Distributed video coding (DVC) features simple encoders but complex decoders, which lies in contrast to conventional video compression solutions such as H.264/AVC. This shift in complexity is realized by performing motion estimation at the decoder side instead of at the encoder, which brings a number of problems that need to be dealt with. One of these problems is that, while employing different coding modes yields significant coding gains in classical video compression systems, it is still difficult to fully exploit this in DVC without increasing the complexity at the encoder side. Therefore, in this paper, instead of using an encoder-side approach, techniques for decoder-side mode decision are proposed. A rate-distortion model is derived that takes into account the position of the side information in the quantization bin. This model is then used to perform mode decision at the coefficient level and bitplane level. Average rate gains of 13 to 28% over the state-of-the-art DISCOVER codec are reported, for a GOP of size four, for several test sequences.

Key words: Distributed Video Coding, mode decision, Wyner-Ziv coding.

1. Introduction

DVC is a new video coding approach where the complexity burden resides at the decoder instead of at the encoder. This shift in complexity compared to conventional solutions such as H.26X and MPEG-X enables new and inter-

esting applications including visual sensor networks, wireless video cameras and wireless video surveillance [1].

Most research in DVC extends or improves the well-known architecture developed at Stanford University [2, 3]. In this system, the frame sequence is split into key frames and so-called Wyner-Ziv (WZ) frames. Key frames are intra coded using conventional techniques such as H.263+ or H.264/AVC intra coding. For each WZ frame, the decoder generates a prediction using already decoded frames as references. This prediction Y – called the *side information* – is merely an estimation of the original frame available at the encoder. Hence, errors in Y are corrected by error correcting information sent by the encoder, using for example, turbo codes or LDPC codes.

Another pioneering architecture is PRISM [4], where a block-based approach is adopted. Blocks are either skipped, intra coded, or WZ coded using syndrome codes.

Many others have proposed techniques for improving rate-distortion performance of DVC systems, often based on Stanford's architecture. Since the quality of the side information has a significant impact on the overall performance, techniques for improved side information generation have been also assiduously investigated. For example, an interesting architecture featuring advanced side information generation is the DISCOVER codec [5], but many others have made contributions as well [6, 7, 8, 9]. A second important factor determining compression performance is the estimation of the conditional distribution $f_{X|Y}$. This distribution is transformed into soft information at the decoder, and used in high performance channel decoding (e.g. turbo or LDPC). Since X and Y are not simultaneously available at the encoder or decoder sides, $f_{X|Y}$ needs to be estimated. To solve this issue, decoder-side correlation estimation techniques have been proposed, by the authors of this paper [10, 11] as well as by other authors in the literature [12, 13, 14, 15].

Although the use of different coding modes has shown great advantages in conventional MPEG-like video compression systems, extending these techniques to DVC is not straightforward. The main problem here is that the bit rate and the distortion are determined by the quality of the side information Y at the decoder, compared to the original X at the encoder. Since X nor Y are both simultaneously available, it seems rather difficult to measure rate and distortion during the encoding or decoding process, and hence, approximations/estimations need to be used instead.

Most current techniques try to tackle this problem by generating an estimation Y' of Y at the encoder side, using low-complexity techniques. For

example, some techniques use a previously coded frame, while other techniques use fast motion estimation. A well-known example is the PRISM architecture, where the past neighboring frame is used, and thresholds on the difference between the current frame and the past neighboring frame define the mode for coding the current block. Mys et al. [16] use similar ideas in a Stanford-alike DVC architecture for the introduction of block-based skip. Error correcting information is still sent for an entire WZ frame, but the turbo decoding procedure is adjusted to take the information about skipped blocks into account. Liu et al. [17] propose an iterative method for deciding on a block-basis between intra mode and WZ mode, assuming Y' available. Both distribution parameters as well as the modes are determined through an iterative procedure. The models that are used are spatially stationary, having the same variance for all DCT coefficients at a given index and a given mode. Ascenso et al. [18] use fast motion estimation techniques to generate Y', and propose rate-distortion based mode decision for deciding between intra mode and WZ mode. Blocks classified as intra are more coarsely quantized than WZ blocks, and they are used for enhancing the side information at the decoder. Next, the enhanced side information is corrected using error correcting information calculated on the entire frame (intra + WZ blocks).

The disadvantage of these encoder-side techniques is that the compression performance depends on the quality of Y' at the encoder. This introduces a trade-off between rate-distortion performance and encoder complexity, since the best results will be obtained in case Y' coincides with Y.

As a solution, the decoder can be made responsible for performing mode decision. However, very few techniques have been proposed so far. In our previous work [19], techniques for skipping coefficient bands and bitplanes have been proposed using a threshold on the estimated distortion. Chien et al. [20] use rate-distortion based decoder-side mode decision, deciding between skipping or WZ coding of a given bitplane. This decision is based on a threshold. While their system outperforms the DISCOVER codec for sequences with low motion, objective results are inferior for sequences with more motion. In addition, only GOP's of size two are considered.

The idea of decoder-side mode decision is carried forward in this paper. Formulas for rate and distortion are developed that take into account the position of the realization of the side information in the quantization bin (Sect. 2). This lies in contrast to previous approaches (e.g. [18]). Based on these theoretical derivations, techniques for rate-distortion based decoder-side mode decision are developed. Two levels for mode decision are intro-

3

duced. Firstly, at the coefficient level, the decoder decides if the entire coefficient band should be skipped or not (Sect. 3.1). Secondly, at the bitplane level (Sect. 3.2), the decoder decides between bitplane skip, bitplane WZ coding and bitplane intra coding. These techniques have been implemented in our codec (Sect. 4), and significant improvements are reported for different test sequences (Sect. 5). Conclusions and future work are provided in Sect. 6, and Sect. 7, respectively.

2. Rate-distortion modeling

Most of the challenges in DVC are due to the inherent problem of not having the side information Y available at the encoder. One consequence is that, at the encoder, one has to choose a quantizer for coding X, without knowledge of Y. The latter is typically solved by quantizing and coding X independently from Y, in contrast to conventional systems where the residual $X - Y$ would be quantized and coded. However, not using the residual approach is less efficient, especially at reasonably low rates. We will illustrate this theoretically, and use these results to develop a method for decoder-side skipping of coefficients that would be too inefficiently coded.

Consider a uniform quantizer with step size Δ and quantization bins labeled q_i ($i \in \mathbb{Z}$). The low and high borders of each quantization bin are given by $q_i^L = \Delta(i - 0.5)$ and $q_i^H = \Delta(i + 0.5)$, respectively. As in most DVC systems, the correlation between X and Y is modeled by a Laplace distribution:

$$f_{X|Y}(x|y) = \frac{\alpha}{2} e^{-\alpha|x-y|}, \tag{1}$$

where α is the distribution scale parameter. Denote the quantization bin containing the realization of the side information as q_K. We introduce a parameter describing the position of the side information in q_K, through $y_N = \frac{y}{\Delta} - K$ taking values in $[-0.5; 0.5]$. As shown in appendix A, the minimum rate that is needed to communicate a quantized version of X from the encoder to the decoder is given by[1]:

$$
\begin{aligned}
H(Q(X)|Y_N = y_N) =& A \cdot \cosh(\alpha\Delta y_N) + B \cdot y_N \cdot \sinh(\alpha\Delta y_N) \\
& - (1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N)) \\
& \quad \cdot \log_2(1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N)),
\end{aligned}
$$

---

[1]Eq. 20 is repeated for convenience.

4

where,

$$A = e^{-\alpha\Delta/2} \left(\frac{\alpha\Delta}{\ln 2 \cdot (1 - e^{-\alpha\Delta})} - \log_2 \sinh(\alpha\Delta/2) \right),$$

$$B = -\frac{\alpha\Delta}{\ln 2} e^{-\alpha\Delta/2}. \tag{2}$$

Next, after the decoder receives the index of the bin containing the original X, a particular value in this bin needs to be chosen as a reconstruction X'. When using the reconstruction employed in the Stanford codec [3] and Mean Absolute Difference (MAD) as distortion measure (as in [18]), as shown in appendix B, we get the following expected distortion[2]:

$$E[|X - X'||Y_N = y_N] = -\cosh(\alpha\Delta y_N) \cdot \Delta e^{-\alpha\Delta/2} \cdot \left(\frac{e^{-\alpha\Delta}}{1 - e^{-\alpha\Delta}} + 0.5 \right)$$

$$+ \Delta y_N \sinh(\alpha\Delta y_N) \cdot e^{-\alpha\Delta/2} + \frac{1}{\alpha}. \tag{3}$$

Note that $H(Q(X)|Y_N = y_N)$ as well as $\alpha \cdot E[|X - X'||Y_N = y_N]$ are both functions of $\alpha\Delta$. Hence, the rate-distortion curves can be plotted for different positions y_N, by choosing any value for α and varying Δ. The result is depicted in Fig. 1. The accuracy of the rate-distortion model has been verified using a sample set of 100000 coded samples generated by a Laplacian source, with $\alpha = 1$. The rate-distortion curves generated experimentally and the ones generated using the proposed formulas for rate and distortion, coincide. The average absolute error is not higher than 0.002.

At the one end, when Δ is very small (i.e. for high rates), the rate and distortion become independent from the position of the side information. The curves coincide and no coding loss is observed. When Δ goes to infinity, the distortion approaches $\frac{1}{\alpha}$ while the entropy goes to zero, except for the case where $y_N = 0.5$, as explained further.

Consider the extreme case where the realization of the side information lies exactly on the bin border ($y_N = 0.5$). If Δ goes to infinity, this means that we end up with two quantization bins, each of infinite length, and having y exactly on the border between the two quantization bins. Due to the symmetry of the Laplace distribution, each of the quantization bins will have a probability of 50%. Therefore, the rate required to communicate the

---

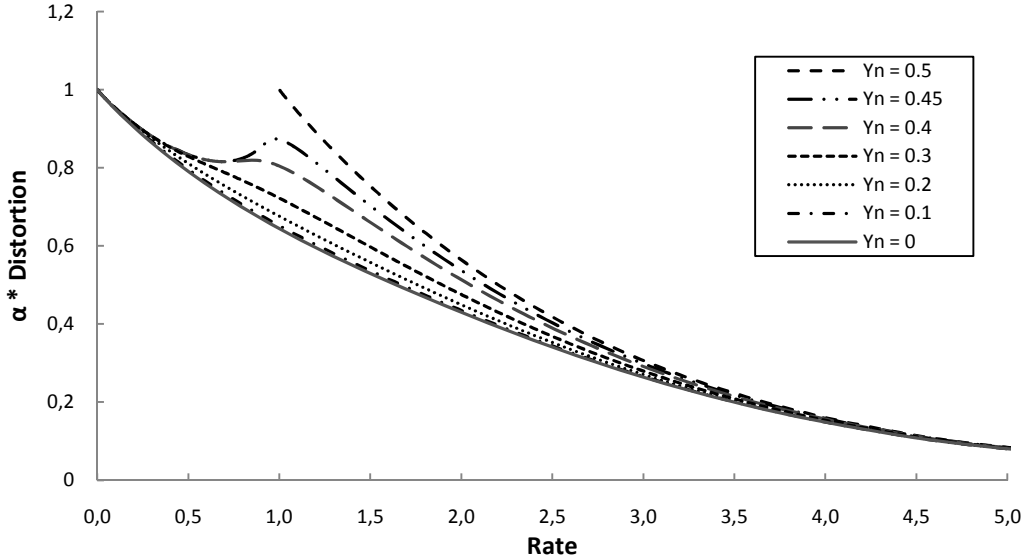[2]Eq. 29 is repeated for convenience.

Figure 1: Theoretical rate-distortion results for different positions of the side information in the quantization bin.

correct bin from the encoder to the decoder equals one (instead of zero as in all other cases depicted). At the decoder side, given the correct quantization bin, the reconstruction process will either select the side information y or the bin border (which is again y). Hence, this reconstruction is the same as in the case where the coefficient would have been skipped. As a consequence, the distortion is also the same in both cases, i.e. $\frac{1}{\alpha}$.

For other positions of the side information (e.g. $y_N = 0.45$), similar reasoning can be applied to explain the occurrence of distortion peaks around a bit rate of one.

To end the discussion of Fig. 1, we conclude that a coding loss is observed which grows as Y approaches the border of the quantization bin. This coding loss can be significant for relatively low rates. This is the case in most DVC systems where coefficients are quantized to a low number of bits (e.g. 2 to 7 bits [21]).

Now that we have quantified coding inefficiency, the question arises whether or not it is possible to avoid this loss. One solution is to perform residual coding by generating an estimation Y' of the side information both at the encoder as well as at the decoder. Next, the residual between X and Y'

6

is coded, as in [22]. An intuitive explanation for the gains achieved using this approach is that Y' is always in the center of the quantization bin. Hence, if there is strong correlation between Y' and Y, then Y has high probability to lie in the center of the bin as well. As such, there is high probability that the best coding performance is achieved. This introduces a trade-off between encoder complexity and rate-distortion performance, since the best results are obtained if $Y' = Y$.

A different strategy is to avoid inefficient coding by choosing between several coding modes. This technique is adopted in this paper. Since the mode decision is performed at the decoder-side, the complexity of the encoder is not increased. To develop this method, we first discuss some additional problems concerning rate and distortion in a practical DVC system.

2.1. Additional problems

The WZ rate in a practical DVC system is higher than the conditional entropy, due to a number of factors. A first problem is that the correlation between X and Y (i.e., the conditional distribution $f_{X|Y}$) needs to be estimated (e.g. by $f'_{X|Y}$), as discussed in the introduction of this paper. While techniques have been proposed in the literature, estimating the correlation is still a very difficult task due to the non-stationarity of the correlation noise $X - Y$ both in the temporal as well as in the spatial direction. This rate penalty can be significant in some cases, especially for higher rates where the mismatch between the bin probabilities increases.

A second issue is that – even if the conditional distribution $f_{X|Y}$ would be known – the conditional entropy can only be reached by a perfect conditional entropy coder, while in practice efficient but suboptimal solutions are used such as LDPC and turbo coding.

Finally, additional rate penalties are imposed by the feedback channel as well as the puncturing period (which determines the smallest chunk of bits that can be sent between encoder and decoder), but this overhead can usually be neglected.

Hence, to summarize we can say that estimating the bit rate using the conditional entropy (Eq. 2) will not always be entirely accurate. Our techniques compensate for this problem by the use of an intra mode for bitplane coding, and more strict criteria for coefficient band skipping and bitplane decoding, as explained further.

As far as distortion is concerned, the only assumption in Eq. 3 is that the correct quantization bin (i.e. the one containing the original) is suc-

cessfully communicated from the encoder to the decoder. This is usually a valid assumption, especially when using a CRC or hash code for additional security [8, 23, 24].

3. Proposed solution

The techniques proposed in this paper are entirely decoder-driven so that the complexity of the encoder remains fairly untouched.

To perform mode decision, typically, two approaches can be used. The first is to apply mode decision at the block level. Each WZ frame is partitioned into blocks and each block is classified into one of several classes (e.g. skip, WZ, intra). The advantage of this technique is that spatial variations can be taken into account. Regions well predicted by the decoder can be skipped, requiring no rate, while other regions can be WZ or intra coded. The second is to classify coefficient bands and bitplanes instead of blocks. The advantage of such a strategy is that high frequency information can often be discarded, as well as the most significant bitplane(s) of a coefficient band, which are often well predicted by the decoder. Since both techniques have their advantages, in this paper, we adopt a coefficient band/bitplane-level approach combined with a (block-like) coefficient classification method to differentiate spatially.

At coefficient level (Sect. 3.1) a conservative coefficient-band skip criterion is used, deciding between skipping the entire coefficient band or not. If the coefficient band can not be skipped, one or more bitplanes are decoded. At the bitplane level (Sect. 3.2), the decoder is granted the choice between three different coding modes for bitplane coding. A first option is again to skip the bitplane. Secondly, the bitplane can be coded in a WZ fashion, by using a turbo coding procedure with selective early stopping and rate request criteria. A third option is to use a conventional arithmetic coder. In that case, the bitplane is coded and decoded without using the side information or any other information, being hence referred to as a bitplane intra mode. While this mode has a theoretical minimal rate equal to the entropy of the bitplane (which can never be lower than the conditional entropy, i.e. given Y), the main idea is that this mode can be used in case the WZ coding mode is inefficient, for example, due to poor correlation noise estimation $f'_{X|Y}$ or poor side information.

3.1. Coefficient band coding

The theoretical results in Fig. 1 illustrate that the rate-distortion performance of quantizing and coding a coefficient depends on the position of the side information in the quantization bin. Some points on the rate-distortion curves seem much less interesting, in the sense that a reasonable amount of bits is spent while the decrease in distortion seems rather limited. In these cases, it could have been better to skip the coefficient and spend no rate at all. This trade-off can be expressed as a Lagrangian cost, where the cost to WZ code a coefficient X_i at index i in the coefficient band is defined as:

$$C_{WZ}^i = H(Q(X_i)|Y_i = y_i) + \lambda \cdot E[||X_i - X_i'|| \, |Y_i = y_i], \qquad (4)$$

while skipping the coefficient and using the side information as a reconstruction results in a cost of

$$C_{skip}^i = \lambda \cdot \frac{1}{\alpha}. \qquad (5)$$

The Lagrange parameter λ has been obtained through experiments on several sequences. First, the optimal Lagrange parameter has been determined offline, for each sequence and quantization level. Next, a curve has been fitted to these results, delivering the following formula for the Lagrange parameter:

$$\lambda = 7.6 e^{-0.1 \cdot IQP}, \qquad (6)$$

where IQP indicates the intra quantization parameter applied for the intra coded frames.

From a theoretical point of view, it seems advantageous to skip the entire coefficient band in case

$$\sum_i C_{skip}^i \leq \sum_i C_{WZ}^i. \qquad (7)$$

However, accounting for practical uncertainties, a more conservative coefficient band skipping method is proposed for use in practice, so that a coefficient band is skipped only if

$$C_{skip}^i \leq C_{WZ}^i, \forall i. \qquad (8)$$

If this criterion is fulfilled, all bitplanes are skipped and decoding proceeds with the next coefficient band. If this criterion is not fulfilled, one or more bitplanes are decoded as explained in the following section.

3.2. Bitplane coding

The second level, after the coefficient band level, is the bitplane level. In most systems, bitplanes are decoded so that high reliability is achieved for all bits. In our system, we only want to spend parity bits for the relevant coefficients, i.e., for which $C_{skip}^i > C_{WZ}^i$. The remaining coefficients are disregarded as much as possible, since it is better to skip them ($C_{skip}^i \leq C_{WZ}^i$).

For each coefficient (either relevant or non-relevant) the decoder keeps track of the quantization bins that can still contain the original coefficient X_i. Prior to decoding the first bitplane, this set of quantization bins S_i is initialized to contain all bins (e.g. 2^M bins when using a quantizer with 2^M levels). After each bitplane has been processed, the number of elements in S_i may be reduced, or not. This depends on the specific mode that has been selected for this bitplane, as explained further on. When all bitplanes have been processed, the coefficient is reconstructed by [25]:

$$x_i' = \frac{\sum_{q_k \in S_i} \int_{q_k^L}^{q_k^H} x_i \cdot f_{X_i|Y_i}'(x_i|y_i)dx_i}{\sum_{q_k \in S_i} \int_{q_k^L}^{q_k^H} f_{X_i|Y_i}'(x_i|y_i)dx_i}, \tag{9}$$

where q_k^L and q_k^H denote the low and high border of the quantization bin q_k, respectively. This reconstruction is more accurate than the one used for obtaining Eq. 3, so that the average distortion is slightly overestimated.

We will first describe how the decoder chooses between the different modes (Sect. 3.2.1), after which details are provided for each of the modes: bitplane skip mode (Sect. 3.2.2), bitplane WZ mode (Sect. 3.2.3), and bitplane intra mode (Sect. 3.2.4). For each of the modes, we will also indicate the effect it has on S_i.

3.2.1. Bitplane mode decision

The mode decision approach followed in this paper is to decode bitplanes so that the *relevant* coefficients have *high reliability* for all bits. As explained before, a coefficient is defined *relevant* for decoding in case $C_{skip}^i > C_{WZ}^i$. *High reliability* in this context means a probability higher or equal to the confidence threshold T (set to 90% in our experiments).

As such, the current bitplane is skipped if all relevant bits are already reliable.

If the bitplane is not skipped, a choice is made between intra coding and WZ coding. This choice is based on what was best for the co-located

bitplane in the previous frame, in hierarchical order[3]. However, only one coding mode has been used for coding the co-located bitplane in the previous frame. Hence, one needs to estimate the result of using other modes than the one used. This estimation is limited to the bit rate, since the expected distortion of the relevant coefficients is similar for each mode.

The following algorithm is used at the decoder for estimating the rate for coding bitplane k in F, denoted BP_k^F, given the decoded bitplane $BP_k'^F$:

- if BP_k^F was WZ coded, the intra rate is calculated by intra coding $BP_k'^F$,

- if BP_k^F was intra coded, the WZ rate is calculated by turbo coding and decoding of $BP_k'^F$.

The choice between the intra mode and WZ mode is then performed as follows. If the co-located bitplane in the previous frame, in hierarchical order, was skipped, then the correlation noise estimation is assumed to be accurate. In that case, the WZ mode is preferred. On the other hand, if the co-located bitplane was not skipped, the (possibly estimated) intra rate is compared to the (possibly estimated) WZ rate. If the intra rate is less than the WZ rate, the decoder selects the intra coding mode for coding the current bitplane, and vice versa.

3.2.2. Bitplane skip

Since no information is retrieved from the encoder, the set S_i of possible quantization bins containing the original value is not updated if the bitplane is skipped.

3.2.3. Bitplane WZ coding

If the bitplane needs to be WZ coded a turbo coding strategy is applied, similar to existing solutions in the literature. However, the criteria that are used at the decoder during this process are evaluated only on a subset of the bits in the bitplane, instead of all. This subset Ω contains the bits corresponding to relevant coefficients, as well as some additional elements.

---

[3]For example, consider a GOP of size four: $I_1 - WZ_2 - WZ_3 - WZ_4 - I_5 - WZ_6 - WZ_7 - WZ_8 - I_9$. Since a hierarchical GOP structure is used [26], the mode decision process for bitplane k in WZ_7 will use the results from bitplane k in WZ_3, while for WZ_8 it will use the results from WZ_6.

The latter are the bits corresponding to irrelevant coefficients that have high reliability before starting to decode the bitplane. This condition is used to compensate for inaccuracies in the correlation noise estimation. It has no effect if the correlation noise estimation is accurate.

Two criteria are used for defining when to stop the iterative turbo decoding process. Firstly, we use an early stopping criterion that is only evaluated for elements of Ω. If the probabilities of these elements remain fairly constant over four iterations (which is called *bit convergence*), the turbo decoding process is stopped. Secondly, if the early stopping criterion is not triggered, the turbo decoding process is stopped if a maximum of 20 decoding iterations has been reached.

If the turbo decoding process stops, the decoder determines if more bits need to be requested from the encoder through a selective rate request strategy. More specifically, additional parity bits are requested from the encoder if there are still bits in Ω for which convergence is not reached, or if they have a reliability less than T.

If no more bits need to be requested, the set of possible quantization bins S_i is updated for all coefficients in the coefficient band, in case both bit convergence as well sufficient reliability has been achieved. Otherwise, S_i is left untouched.

3.2.4. Bitplane intra coding

If the bitplane needs to be intra coded, a binary arithmetic entropy coder is used. This coder operates with an adaptive model that is initialized to a uniform distribution before each bitplane is coded. At the decoder, intra coded bitplanes can be decoded without using the side information. All decoded bits can be considered reliable, so that S_i is updated for each coefficient.

4. Test setup

The techniques proposed in this paper have been implemented in our codec, depicted in Fig. 2.

As in the Stanford architecture, the frame sequence is split into key frames I and WZ frames W. At the encoder, key frames are coded without using other frames as references, applying H.264/AVC intra coding techniques. The decoded intra frames I' are generated anyway during H.264/AVC intra coding, so they are used to calculate the quantization noise in I', for each

coefficient band. The result is sent to the decoder along with the coded key frame. WZ frames are partitioned into non-overlapping blocks of size 4-by-4. Each block is transformed using a DCT transformation, after which each of 16 coefficient bands is quantized and bitplanes are extracted. Next, depending on the information received from the decoder, all bitplanes in the coefficient band can be either (1) discarded, (2) only the current bitplane can be discarded, (3) the current bitplane can be intra coded, or (4) WZ coded. For the latter, parity bits are calculated by a turbo coder, and stored in a buffer. They are sent to the decoder in chunks, using a puncturing procedure.

At the decoder, the coded key frames are decoded to I'. For each WZ frame W, the decoder generates a prediction using already decoded frames, adopting the techniques used in the DISCOVER codec [5]. Next, information from the side information generation process as well as the amount of quantization noise in the key frames is used to estimate the correlation noise as $f'_{X|Y}$. More details about this correlation noise estimation can be found in our previous work [11]. Next, the side information Y is transformed, and mode decision is performed using the techniques discussed previously in this paper. The mode information is communicated to the encoder using the feedback channel, which is already available. If the WZ bitplane coding mode is selected, the turbo decoder uses the transformed side information for decoding. If the bitplane is intra decoded using the arithmetic decoder, the side information is not used. If all bitplanes are decoded, the coefficients are reconstructed, and the inverse DCT is applied to the results, delivering the decoded frame W'. This frame can then be used for side information generation of future frames to be decoded.

5. Results

Tests have been conducted on three sequences: "Mother and Daughter", "Table Tennis" and "Foreman". All sequences have CIF resolution, at a frame rate of 30 Hz. Only the luma component is coded, allowing for comparison with the DISCOVER codec [5].

First, we investigate the gains realized by introducing bitplane skip and bitplane intra mode, by comparing different configurations of our system (Sect. 5.1). Next, we compare decoder-side mode decision as proposed in this paper to a configuration with encoder-side mode decision (Sect. 5.2). Finally, the performance of our system is evaluated against comparable systems found in the literature (Sect. 5.3).
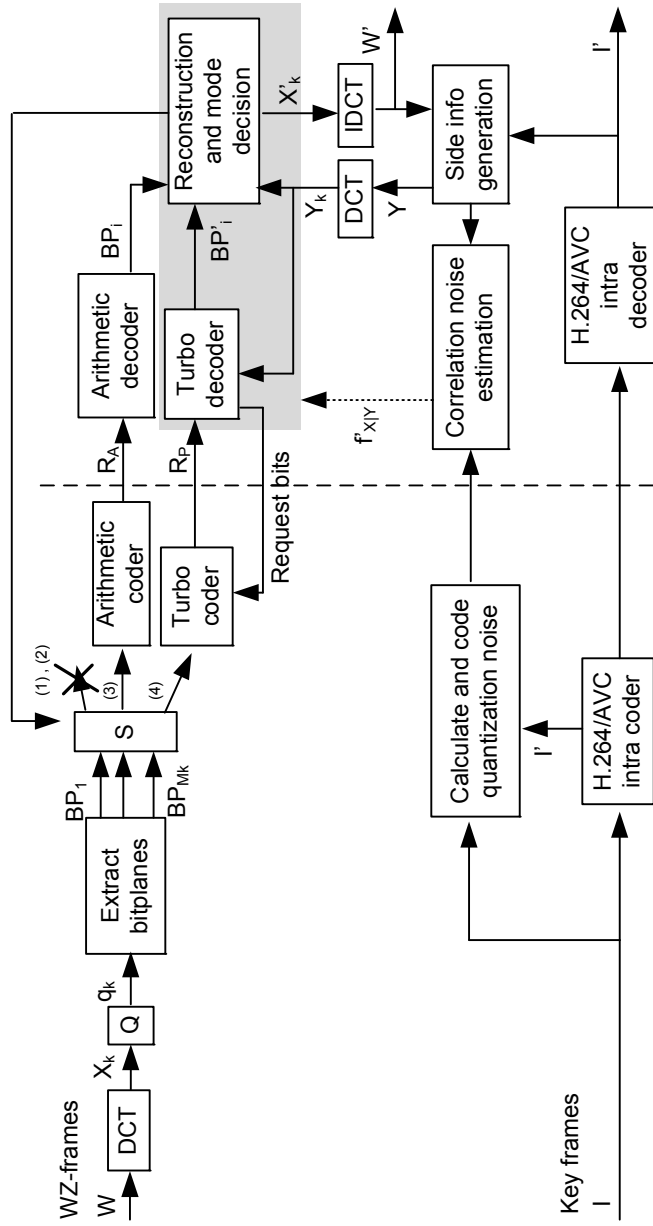
13

Figure 2: Codec architecture, featuring several modes for coding bitplanes.

5.1. Studying the gains realized by intra and skip

To illustrate the gains obtained by the different coding modes, results have been generated for a GOP of length four (i.e., $I - WZ - WZ - WZ - I - ...$), for a number of different configurations. A first configuration is the system proposed in one of our earlier papers [11]. This system has been taken as a basis in this paper. It supports bitplane WZ decoding and bitplane skip based on a threshold on the bit probabilities. A second configuration is obtained by using the system proposed in this paper, but excluding the intra mode. This system is labeled as *RD-based WZ + skip*. The third configuration, labeled *RD-based WZ + intra + skip*, supports all modes proposed in this paper.

Comparing rate-distortion performance of the different configurations (Fig. 3) indicates that most of the gain at low rates is due to the skip mode, while the gains at high rates are mostly due to the bitplane intra mode. This is rather expected. At low rates, for large quantization bins, the decoder is often able to predict to correct bin, so that the information can be skipped. At high rates, predicting the correct bin becomes difficult. In addition, inaccuracies in the correlation noise model make the WZ mode less efficient, especially at high rates where the accuracy of the estimation of the bin probabilities decreases. As a result, the intra mode is selected more often.

More details concerning the mode decision is provided in Tab. 1, for different sequences. For each sequence, results for the first eight coefficient bands (labeled CB0 to CB7) in raster scan are provided, from fine to coarse quantization (Q2 to Q5, respectively). Analyzing the occurrence of each mode, we can indeed see that the skip mode is selected more often at low rates, while the intra mode is more often selected at high rates. Across coefficient bands, we observe that high frequency information is often skipped. More skip is used for sequences that can be well predicted by the decoder, such as Mother and Daughter. On the other hand, for difficult sequences, more intra is used.

5.2. Comparing encoder and decoder-side mode decision

In this subsection, we compare the proposed system featuring decoder-side mode decision against a configuration applying encoder-side mode decision. In the latter configuration, at the encoder, the correlation between the original and the side information is estimated by generating an estimation of the side information using low-complexity techniques. More precisely, the side information for a particular WZ frame is estimated at the encoder by calculating the average between the past and future reference frames. These
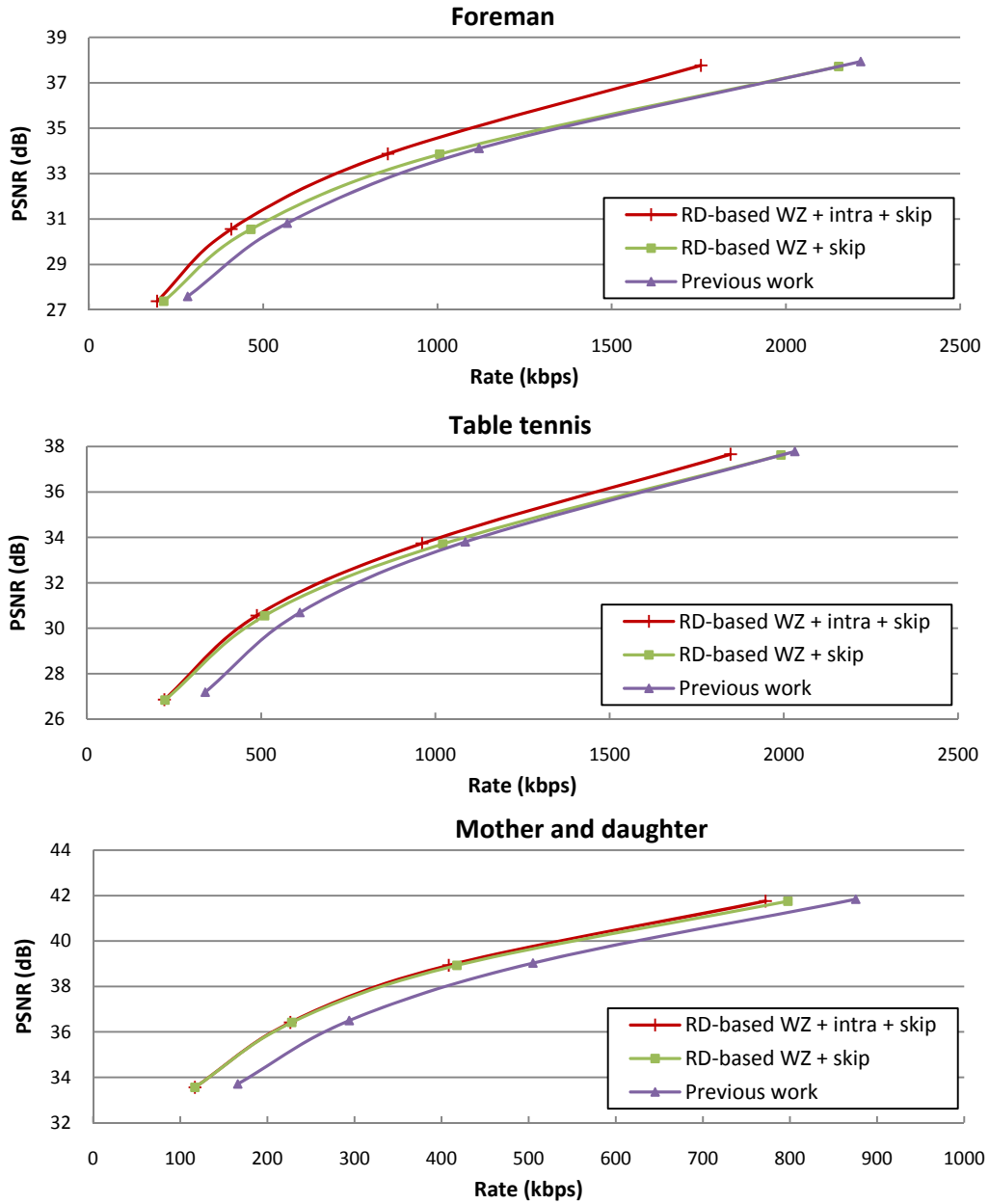
Figure 3: Comparing the system proposed in this paper to the same configuration applying only WZ and skip mode, and to our previous work [11].

Table 1: Average percentage of modes used for different coefficient bands (i.e., CB0 to CB7) and different quantization levels (i.e., Q2 to Q5).

FOREMAN

			CB0	CB1	CB2	CB3	CB4	CB5	CB6	CB7
	Q2	Skip	0%	6%	30%	43%	9%	28%	41%	53%
		WZ	68%	54%	27%	12%	43%	37%	29%	9%
		Intra	32%	40%	42%	45%	48%	35%	29%	38%
	Q3	Skip	0%	14%	46%	63%	16%	45%	65%	78%
		WZ	74%	56%	24%	7%	46%	33%	14%	6%
		Intra	26%	30%	30%	30%	38%	22%	20%	17%
	Q4	Skip	0%	41%	74%	84%	36%	68%	86%	92%
		WZ	82%	40%	7%	5%	38%	16%	4%	3%
		Intra	18%	19%	20%	12%	26%	17%	10%	5%
	Q5	Skip	5%	78%	93%	97%	71%	91%	100%	100%
		WZ	79%	10%	3%	1%	11%	4%	0%	0%
		Intra	16%	11%	4%	2%	18%	6%	0%	0%

TABLE TENNIS

			CB0	CB1	CB2	CB3	CB4	CB5	CB6	CB7
	Q2	Skip	0%	6%	19%	36%	3%	16%	28%	43%
		WZ	85%	69%	63%	50%	63%	58%	52%	40%
		Intra	15%	24%	18%	14%	35%	26%	20%	17%
	Q3	Skip	0%	8%	27%	44%	4%	22%	41%	59%
		WZ	90%	72%	59%	43%	63%	55%	44%	29%
		Intra	10%	20%	14%	13%	33%	23%	15%	12%
	Q4	Skip	0%	18%	47%	69%	10%	46%	72%	91%
		WZ	92%	64%	44%	20%	61%	40%	22%	6%
		Intra	8%	18%	9%	11%	30%	15%	6%	2%
	Q5	Skip	6 %	60%	96%	99%	49%	93%	99%	100%
		WZ	88%	35%	3%	1%	37%	5%	1%	0%
		Intra	7%	5%	1%	0%	15%	2%	0%	0%

MOTHER DAUGHTER

			CB0	CB1	CB2	CB3	CB4	CB5	CB6	CB7
	Q2	Skip	1%	32%	55%	74%	29%	52%	72%	87%
		WZ	97%	55%	33%	21%	67%	19%	9%	4%
		Intra	2%	13%	12%	5%	3%	29%	18%	8%
	Q3	Skip	4 %	50%	79%	93%	50%	77%	93%	98%
		WZ	96%	35%	14%	3%	49%	5%	3 %	1%
		Intra	0%	14%	8%	4%	1%	17%	4 %	1%
	Q4	Skip	22%	78%	97%	99%	81%	97%	100%	100%
		WZ	78%	13%	2%	1%	16%	2%	0%	0%
		Intra	0 %	9%	1%	0%	3%	1%	0%	0%
	Q5	Skip	64%	99%	100%	100%	100%	100%	100%	100%
		WZ	36%	1%	0%	0%	0%	0%	0%	0%
		Intra	0%	0%	0%	0%	0%	0%	0%	0%

reference frames are the original frames, i.e., without quantization noise, since the decoded WZ frames are not available at the encoder. Hence, to account for quantization noise, the interpolation between past and future references is intra coded and decoded, delivering the estimation Y' of the side information Y at the decoder. Next, the correlation between the original X and Y is estimated using X and Y', following the techniques described by Brites et al.[12] for the offline transform-domain correlation noise model at coefficient/frame level. Finally, given the correlation model, encoder-side mode decision is performed following the techniques proposed in this paper.

Both techniques are compared in Fig. 4, indicating better performance for the system applying decoder-side mode decision, compared to the encoder-side approach. This is due to a number of reasons. Firstly, the encoder uses an estimation of the side information instead of generating the actual side information, as computed by the decoder. While the encoder could generate the actual side information, this would make it significantly more complex, no longer fitting the typical DVC use case scenarios. Secondly, a different correlation model had to be used in an encoder-side mode decision approach, which could negatively impact performance. Hence, while these results give an indication about the performance of decoder-side mode decision compared to encoder-side mode decision (with simple side information estimation), a more thorough analysis is needed for pinpointing the major differences between the two approaches, and their advantages and disadvantages. This is part of future work.

5.3. Rate-distortion results compared to other systems

Our system is compared to a number of reference systems, namely DISCOVER, H.264/AVC intra coding, and low-complexity H.264/AVC inter coding (i.e. no motion estimation inter coding). Results for GOP sizes two (Fig. 5), four (Fig. 6), and eight (Fig. 7) are presented.

DISCOVER[4] is outperformed significantly. This is due to the improved correlation noise model proposed in our earlier work [11] as well as the addition of several coding modes as proposed in this paper. Average rate gains vary per sequence. The largest gains over DISCOVER are observed for sequences where the skip mode or intra mode is used a lot, as in Mother and Daughter, and Foreman, respectively (see Table 1). For example, for a GOP

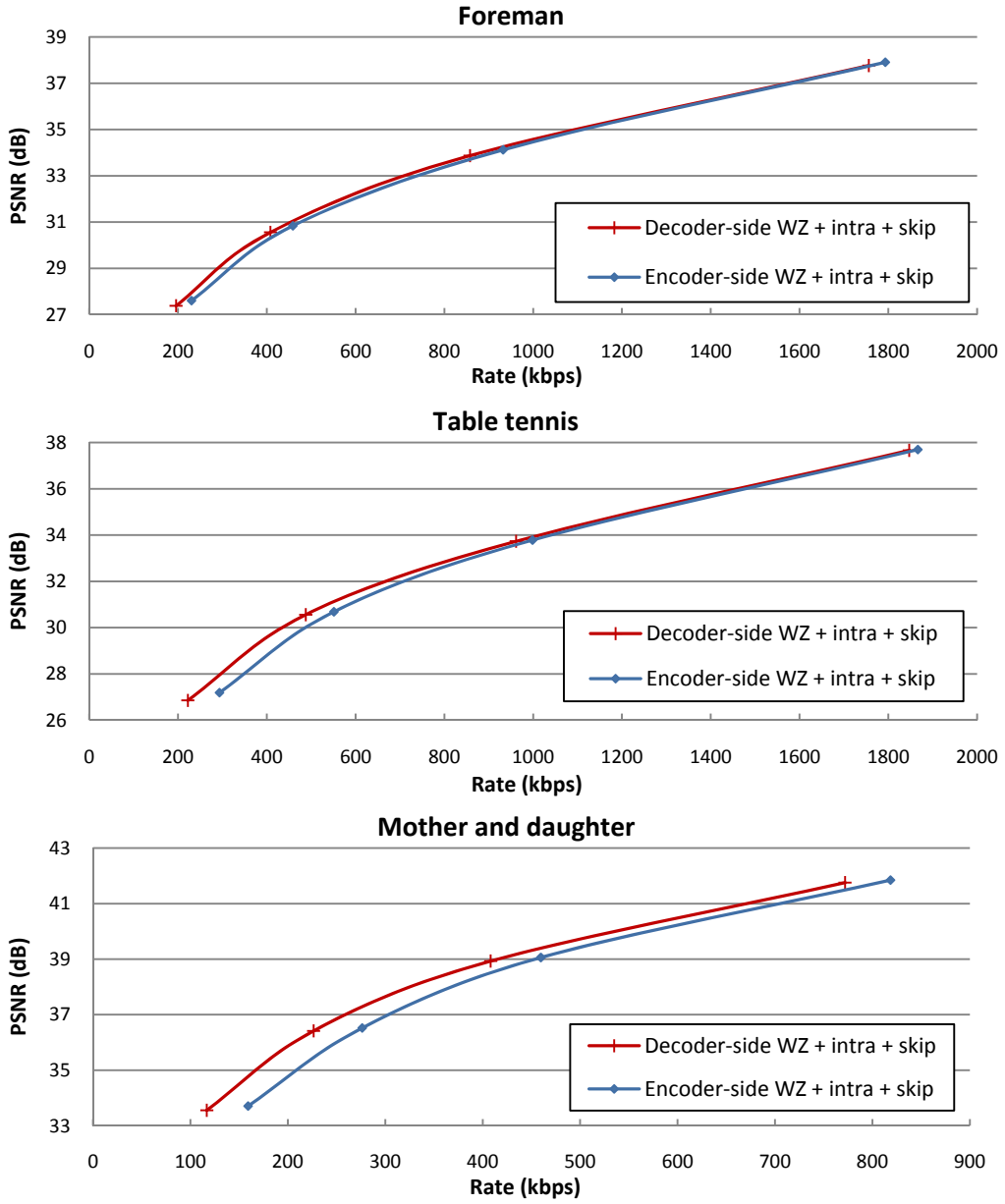[4]Executables of the DISCOVER codec are available online [27].

Figure 4: The proposed system performs better than a configuration where encoder-side mode decision is performed (with simple side information estimation).

of size four, average Bjøntegaard [28] rate improvements are achieved of 28% for Foreman, 13% for Table, and 24% for Mother and Daughter.

Unlike other DVC solutions, our system outperforms H.264/AVC intra coding consistently. On the other hand, compared to low complexity H.264/AVC inter coding, referred to as H.264/AVC no motion (i.e. no motion estimation inter coding), better performance is often achieved, even for long GOP sizes.

6. Conclusions

In this paper we have derived closed-form expressions for rate and distortion in DVC that take into account the position of the side information in the quantization bin. These formulas indicate that a significant coding penalty is paid if the side information lies close to the border of the quantization bin. A number of approaches can be used to compensate for this effect. The approach followed in this paper is to use a Lagrange cost function for mode decision, skipping cases of inefficient WZ coding.

Apart from the coding loss caused by inefficient quantization, a second issue addressed in this paper is the fact that the side information Y and correlation noise estimation $f_{X|Y}$ can be inaccurate. For these cases, the decoder had the possibility to select a bitplane intra coding mode, avoiding the use of unreliable Y and $f_{X|Y}$. The results indicate that this strategy improves significantly the coding performance. On the one hand, this illustrates that there is a need for further improving of side information generation and correlation noise estimation in DVC. On the other hand, this shows that adding additional coding modes can have a significant impact, improving the performance in DVC, without the need of performing motion estimation at the encoder side.

7. Future work

Modeling the correlation between the original and the side information is one of the major challenges in DVC, since this model has a significant impact on WZ decoding, coefficient reconstruction, and also, in our case, on mode decision. Inaccuracies in estimating the alpha parameter have a negative impact on the mode decision process. If alpha is overestimated (i.e. corresponding to a small variance), more information is skipped than necessary,
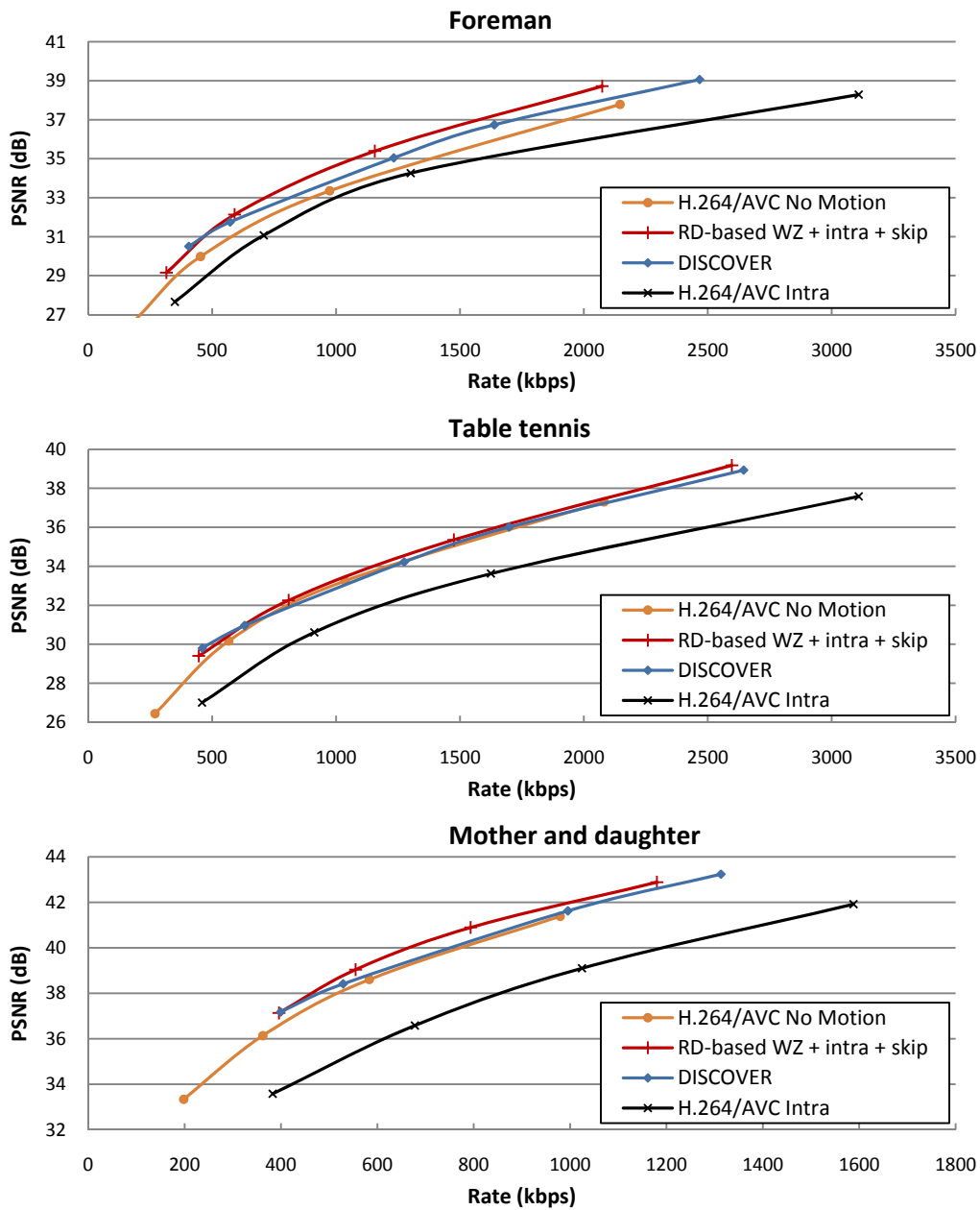
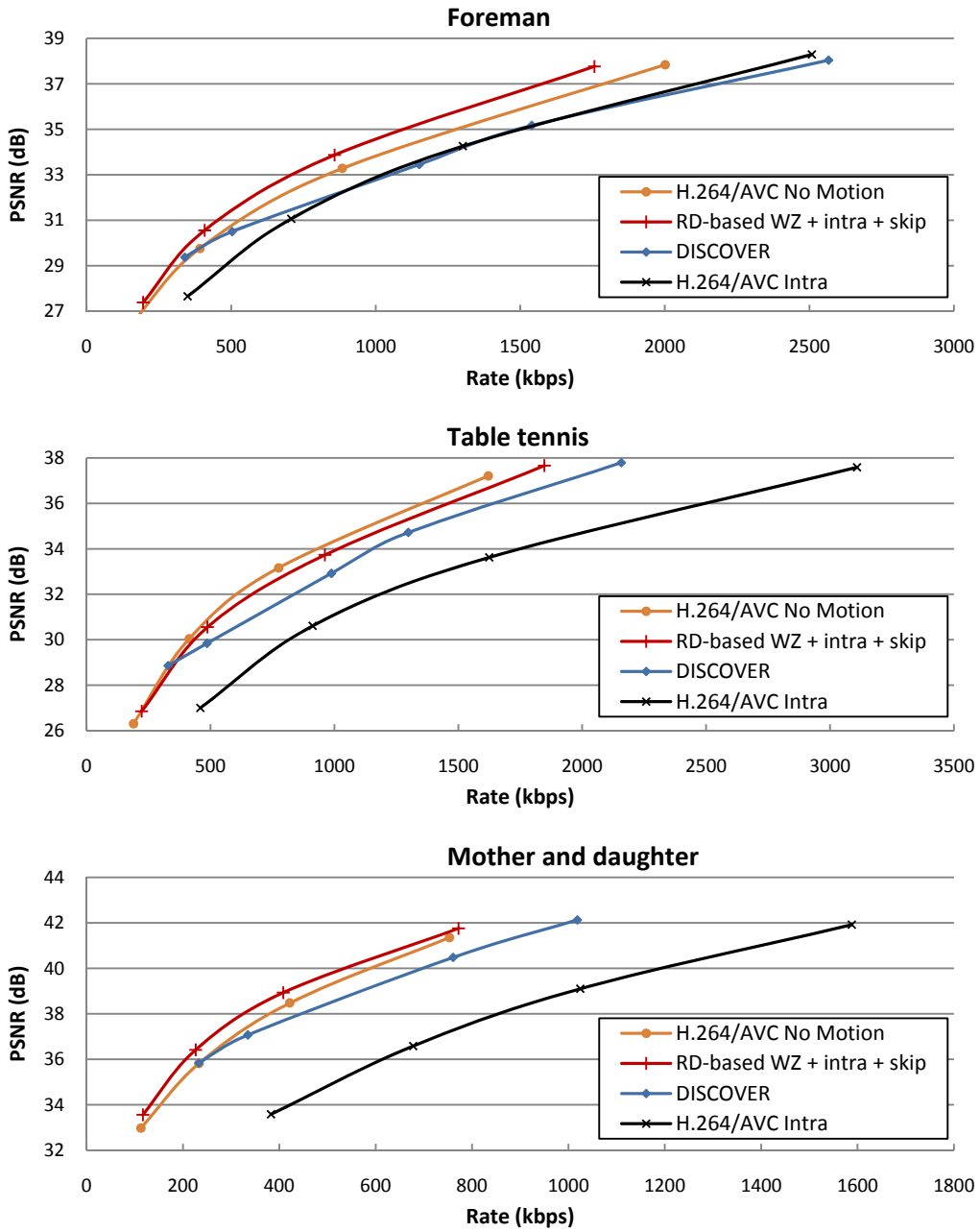Figure 5: Rate-distortion results for a GOP of size two.

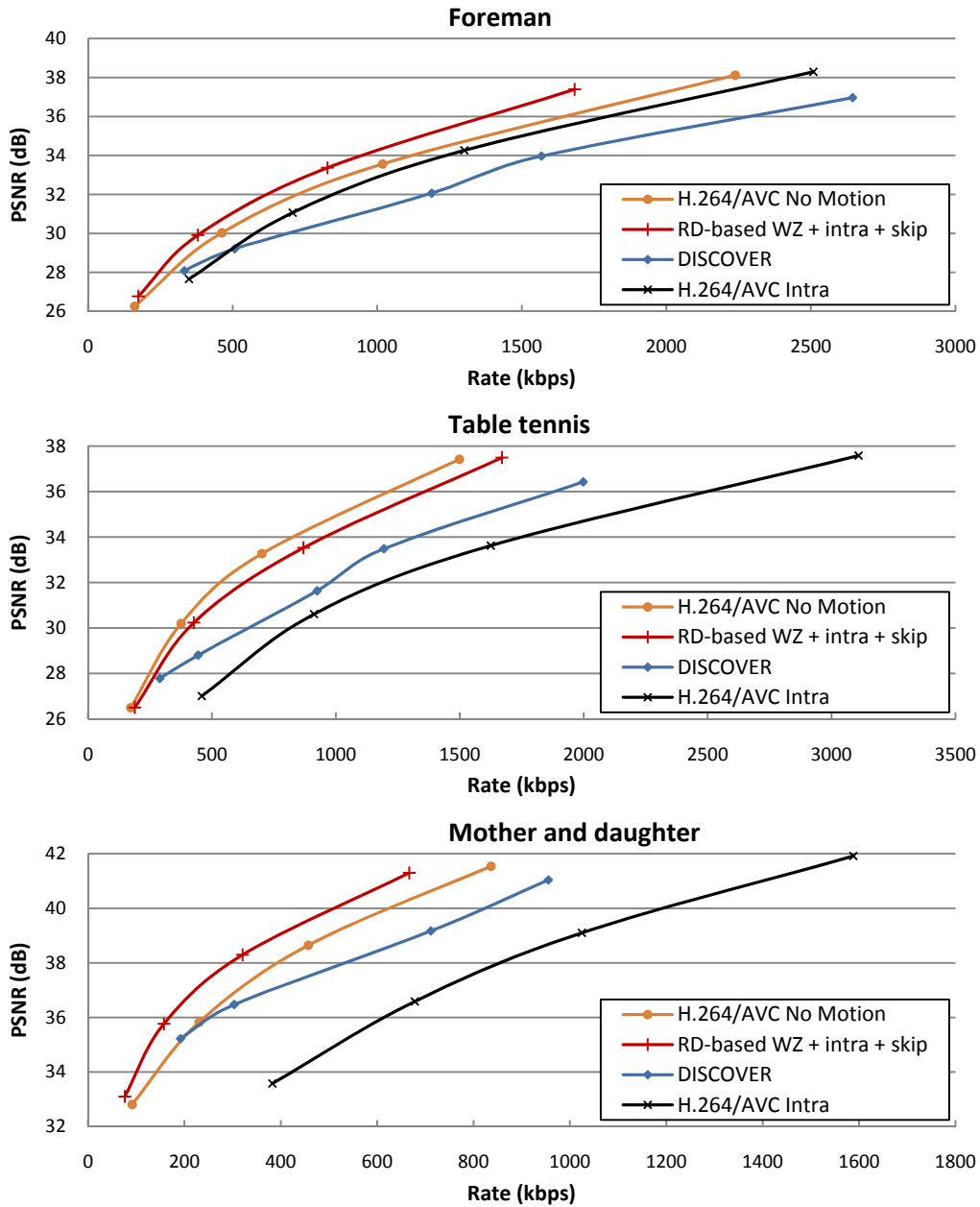Figure 6: Rate-distortion results for a GOP of size four.

Figure 7: Rate-distortion results for a GOP of size eight.

and vice versa. Also, inaccuracies in the correlation model make turbo decoding less efficient, so that more parity bits are requested from the encoder. If WZ decoding becomes too inefficient, intra coding provides a solution, as proposed in this paper, hereby ignoring the correlation model. Hence, due to the significant impact of the correlation model on the performance of a DVC system, we strongly believe that more research is necessary in this area in order to further improve the accuracy over that of existing correlation noise models.

New techniques can be investigated concerning the modes proposed in this paper. For example, the intra mode could be extended by using better adaptive models, as well as by using information from previously decoded bitplanes. In addition, research can be performed for defining new additional coding modes, as the introduction of the coding modes introduced in this paper showed significant gains.

References

[1] F. Pereira, L. Torres, C. Guillemot, T. Ebrahimi, R. Leonardi, and S. Klomp, "Distributed Video Coding: Selecting the most promising application scenarios," *Signal Processing : Image Communication*, pp. 339–352, 2008.

[2] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-domain Wyner-Ziv codec for video," in *Proc. SPIE Visual Communications and Image Processing*, vol. 5308, pp. 520–528, January 2004.

[3] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed Video Coding," in *Proc. IEEE, Special Issue on Video Coding and Delivery*, vol. 93, pp. 71–83, January 2005. Invited Paper.

[4] R. Puri and K. Ramchandran, "PRISM: A new robust video coding architecture based on distributed compression principles," in *Proc. Allerton Conference on Communication, Control and Computing*, October 2002.

[5] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "The DISCOVER codec: Architecture, techniques and evaluation," in *Proc. Picture Coding Symposium (PCS)*, November 2007.

[6] D. Kubasov and C. Guillemot, "Mesh-based motion-compensated inter-polation for side information extraction in Distributed Video Coding," in *Proc. IEEE International Conference on Image Processing (ICIP)*, October 2006.

[7] S. Ye, M. Ouaret, F. Dufaux, and T. Ebrahimi, "Improved side information generation for distributed video coding by exploiting spatial and temporal correlations," *EURASIP Journal on Image and Video Processing*, vol. 2009, 2009. Article ID 683510.

[8] R. Martins, C. Brites, J. Ascenso, and F. Pereira, "Refining side information for improved transform domain wyner-ziv video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, pp. 1327–1341, september 2009.

[9] X. Fan, O. C. Au, N. M. Cheung, Y. Chen, and J. Zhou, "Successive refinement based Wyner-Ziv video compression," *Signal Processing: Image Communcation*, 2009. doi:10.1016/j.image.2009.09.004.

[10] J. Škorupa, J. Slowack, S. Mys, P. Lambert, and R. Van de Walle, "Accurate correlation modeling for transform-domain Wyner-Ziv video coding," in *Proc. Pacific-Rim Conference on Multimedia (PCM)*, pp. 1–10, December 2008.

[11] J. Slowack, S. Mys, J. Škorupa, P. Lambert, C. Grecos, and R. Van de Walle, "Accounting for quantization noise in online correlation noise estimation for distributed video coding," in *Proc. Picture Coding Symposium (PCS)*, May 2009.

[12] C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, pp. 1177–1190, September 2008.

[13] X. Huang and S. Forchhammer, "Improved virtual channel noise model for transform domain Wyner-Ziv video coding," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 921–924, April 2009.

[14] X. Fan, O. C. Au, and N. M. Cheung, "Adaptive correlation estimation for general Wyner-Ziv video coding," in *IEEE International Conference on Image Processing (ICIP)*, November 2009.

[15] G. R. Esmaili and P. Cosman, "Correlation noise classification based on matching success for transform domain Wyner-Ziv video coding," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 801–804, April 2009.

[16] S. Mys, J. Slowack, J. Škorupa, P. Lambert, and R. Van de Walle, "Introducing skip mode in distributed video coding," *Signal Processing: Image Communication*, vol. 24, no. 3, pp. 200–213, 2009.

[17] L. Liu, D.-K. He, A. Jagmohan, L. Lu, and E. J. Delp, "A low-complexity iterative mode selection algorithm for Wyner-Ziv video compression," in *IEEE International Conference on Image Processing (ICIP)*, pp. 1136–1139, October 2008.

[18] J. Ascenso and F. Pereira, "Low complexity intra mode selection for efficient distributed video coding," in *Proc. International Conference on Multimedia and Expo (ICME)*, pp. 101–104, July 2009.

[19] J. Slowack, J. Škorupa, S. Mys, P. Lambert, C. Grecos, and R. Van de Walle, "Distributed video coding with decoder-driven skip," in *Proc. Mobimedia*, September 2009.

[20] W.-J. Chien and L. J. Karam, "BLAST-DVC: Bitplane selective distributed video coding," *Multimedia Tools and Applications*, July 2009. DOI: 10.1007/s11042-009-0314-8.

[21] C. Brites, J. Ascenso, J. Q. Pedro, and F. Pereira, "Evaluating a feedback channel based transform domain wyner-ziv video codec," *Signal Processing: Image Communication*, pp. 269–297, 2008.

[22] A. Aaron, D. Varodayan, and B. Girod, "Wyner-Ziv residual coding of video," in *Proc. Picture Coding Symposium (PCS)*, April 2006.

[23] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," in *Proc. IEEE International Conference on Image Processing (ICIP)*, pp. 3097–3100, October 2004.

[24] M. Tagliasacchi and S. Tubaro, "Hash-based motion modeling in Wyner-Ziv video coding," in *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, vol. 1, pp. 509–512, 2007.

[25] D. Kubasov, J. Nayak, and C. Guillemot, "Optimal reconstruction in Wyner-Ziv video coding with multiple side information," in *IEEE MultiMedia Signal Processing Workshop*, pp. 183–186, October 2007.

[26] A. Aaron, E. Setton, and B. Girod, "Towards practical Wyner-Ziv coding of video," in *Proc. IEEE International Conference on Image Processing (ICIP)*, pp. 869–872, September 2003.

[27] *DISCOVER-Distributed Coding for Video Services*, September 2005. Accessed: Dec. 1, 2009.

[28] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," tech. rep., VCEG, April 2002. Contribution VCEG-M33.

A. Coefficient rate calculation

Using the notations introduced in the beginning of Sect. 2 (concerning the uniform quantizer and Laplacian correlation model), the probability of q_i containing X given $Y = y$ is written as:

$$p_i = \int_{q_i^L}^{q_i^H} f_{X|Y}(x|y)dx. \tag{10}$$

It is quite straightforward to calculate p_i, by discriminating between different cases for i:

$$p_i = \begin{cases} \frac{e^{-\alpha(y-q_i^H)}-e^{-\alpha(y-q_i^L)}}{2}, & i < K \\ 1 - \frac{e^{-\alpha(y-q_i^L)}+e^{-\alpha(q_i^H-y)}}{2}, & i = K \\ \frac{e^{-\alpha(q_i^L-y)}-e^{-\alpha(q_i^H-y)}}{2}, & i > K. \end{cases} \tag{11}$$

This can be rewritten by introducing $y_N = \frac{y}{\Delta} - K$ and using the expressions for q_i^L and q_i^H:

$$p_i = \begin{cases} e^{\alpha\Delta(i-K)} \cdot \gamma_1, & i < K \\ 1 - e^{-\alpha\Delta/2} \cdot \cosh(\alpha\Delta y_N), & i = K \\ e^{-\alpha\Delta(i-K)} \cdot \gamma_2, & i > K \end{cases}$$

with

$$\begin{aligned} \gamma_1 &= \sinh(\alpha\Delta/2) \cdot e^{-\alpha\Delta y_N}, \\ \gamma_2 &= \sinh(\alpha\Delta/2) \cdot e^{\alpha\Delta y_N}. \end{aligned} \tag{12}$$

This result is used to derive a closed-form expression for the conditional entropy, given the position of the side information. With $j = i - K$, we get:

$$
\begin{aligned}
H(Q(X)|Y_N = y_N) &= - \sum_{i=-\infty}^{\infty} p_i \log_2 p_i, \\
&= - \sum_{i=-\infty}^{K-1} p_i \log_2 p_i - (p_K \log_2 p_K) - \sum_{i=K+1}^{\infty} p_i \log_2 p_i \\
&= - \sum_{j=-\infty}^{-1} p_{j+K} \log_2 p_{j+K} - (p_K \log_2 p_K) \\
&\quad - \sum_{j=1}^{\infty} p_{j+K} \log_2 p_{j+K}
\end{aligned}
\tag{13}
$$

The first term of this sum is calculated as:

$$
\begin{aligned}
\sum_{j=-1}^{-\infty} p_{j+K} \log_2 p_{j+K} &= \sum_{j=-1}^{-\infty} \gamma_1 e^{\alpha\Delta j} \log_2 \gamma_1 e^{\alpha\Delta j} \\
&= \sum_{j=1}^{\infty} \gamma_1 e^{-\alpha\Delta j} \log_2 \left(\gamma_1 e^{-\alpha\Delta j} \right) \tag{14} \\
&= \gamma_1 \sum_{j=1}^{\infty} e^{-\alpha\Delta j} \left(\frac{\ln \gamma_1}{\ln 2} + \frac{\ln e^{-\alpha\Delta j}}{\ln 2} \right) \\
&= \frac{\gamma_1}{\ln 2} \left(\ln \gamma_1 \sum_{j=1}^{\infty} e^{-\alpha\Delta j} - \alpha\Delta \sum_{j=1}^{\infty} j e^{-\alpha\Delta j} \right) \\
&= \frac{\gamma_1}{\ln 2} \left(\ln \gamma_1 \cdot \frac{e^{-\alpha\Delta}}{1 - e^{-\alpha\Delta}} - \alpha\Delta \cdot \frac{e^{-\alpha\Delta}}{(1 - e^{-\alpha\Delta})^2} \right) \tag{15}
\end{aligned}
$$

The third term is calculated in a similar way, recognizing that:

$$
\sum_{j=1}^{\infty} p_{j+K} \log_2 p_{j+K} = \sum_{j=1}^{\infty} \gamma_2 e^{-\alpha\Delta j} \log_2 \left(\gamma_2 e^{-\alpha\Delta j} \right), \tag{16}
$$

has a similar form as Eq. 14, so that the following expression is obtained:

$$
\begin{aligned}
H(Q(X)|Y_N = y_N) = & -\frac{e^{-\alpha\Delta}}{\ln 2 \cdot (1 - e^{-\alpha\Delta})} \cdot (\gamma_1 \ln \gamma_1 + \gamma_2 \ln \gamma_2) \\
& + \frac{\alpha\Delta e^{-\alpha\Delta}}{\ln 2 \cdot (1 - e^{-\alpha\Delta})^2} \cdot (\gamma_1 + \gamma_2) \\
& - (1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N)) \\
& \quad \cdot \log_2(1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N)). \qquad (17)
\end{aligned}
$$

By using the expressions for γ_1 and γ_2, the first term of this sum results in:

$$
\begin{aligned}
-\frac{e^{-\alpha\Delta/2}}{\ln 2} \cdot \alpha\Delta y_N \cdot \sinh(\alpha\Delta y_N) \\
- e^{-\alpha\Delta/2} \cdot \log_2 \sinh(\alpha\Delta/2) \cdot \cosh(\alpha\Delta y_N) \quad (18)
\end{aligned}
$$

while the second term simplifies to:

$$
\frac{\alpha\Delta}{\ln 2} \cdot \frac{e^{-\alpha\Delta/2}}{1 - e^{-\alpha\Delta}} \cdot \cosh(\alpha\Delta y_N), \qquad (19)
$$

which results in the following closed-form expression:

$$
\begin{aligned}
H(Q(X)|Y_N = y_N) = & A \cdot \cosh(\alpha\Delta y_N) + B \cdot y_N \cdot \sinh(\alpha\Delta y_N) \\
& - (1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N)) \\
& \quad \cdot \log_2(1 - e^{-\alpha\Delta/2} \cosh(\alpha\Delta y_N))
\end{aligned}
$$

with:

$$
\begin{aligned}
A = & e^{-\alpha\Delta/2} \left(\frac{\alpha\Delta}{\ln 2 \cdot (1 - e^{-\alpha\Delta})} - \log_2 \sinh(\alpha\Delta/2) \right), \\
B = & -\frac{\alpha\Delta}{\ln 2} e^{-\alpha\Delta/2}. \qquad (20)
\end{aligned}
$$

B. Coefficient distortion calculation

In this subsection, we develop an expression for the distortion using the mean absolute difference (MAD) as distortion metric. We assume that the output of the channel decoder is perfect, so that the decoder knows the bin

29

containing the original value. Reconstruction is performed as in the Stanford codec [3]:

$$x'_i = \begin{cases} q_i^H & i < K, \\ y & i = K, \\ q_i^L & i > K. \end{cases} \tag{21}$$

As such, average distortion given the specific realization of the side information, is calculated as follows:

$$
\begin{aligned}
E[|X - X'| \, | Y = y] &= \sum_{i=-\infty}^{\infty} \int_{q_i^L}^{q_i^H} f_{X|Y}(x|y)|x - x'_i| dx \\
&= \sum_{i=-\infty}^{K-1} \int_{q_i^L}^{q_i^H} \frac{\alpha}{2} e^{-\alpha(y-x)} \cdot (q_i^H - x) dx \\
&\quad + \int_{q_K^L}^{y} \frac{\alpha}{2} e^{-\alpha(y-x)} \cdot (y - x) dx \\
&\quad + \int_{y}^{q_K^H} \frac{\alpha}{2} e^{-\alpha(x-y)} \cdot (x - y) dx \\
&\quad + \sum_{i=K+1}^{\infty} \int_{q_i^L}^{q_i^H} \frac{\alpha}{2} e^{-\alpha(x-y)} \cdot (x - q_i^L) dx. \tag{22}
\end{aligned}
$$

Solving the integrals in the first term of this sum results in:

$$\sum_{i=-\infty}^{K-1} \int_{q_i^L}^{q_i^H} \frac{\alpha}{2} e^{-\alpha(y-x)} \cdot (q_i^H - x) dx = \frac{\alpha}{2} e^{-\alpha y} \cdot C \sum_{i=-\infty}^{K-1} e^{\alpha \Delta i}, \tag{23}$$

with

$$C = \frac{e^{\alpha \Delta/2}}{\alpha^2} - e^{-\alpha \Delta/2} \left(\frac{\Delta}{\alpha} + \frac{1}{\alpha^2} \right). \tag{24}$$

Using $y_N = \frac{y}{\Delta} - K$, and $j = i - K$, we can further simplify:

$$\sum_{i=-\infty}^{K-1} \int_{q_i^L}^{q_i^H} \frac{\alpha}{2} e^{-\alpha(y-x)} \cdot (q_i^H - x) dx = \frac{\alpha}{2} e^{-\alpha y} \cdot C \sum_{j=-\infty}^{-1} e^{\alpha \Delta (j+K)}$$

$$= \frac{\alpha}{2} e^{-\alpha \Delta y_N} \cdot C \sum_{j=1}^{\infty} e^{-\alpha \Delta j}$$

$$= e^{-\alpha \Delta y_N} \cdot \frac{e^{-\alpha \Delta/2} - (\alpha \Delta + 1) e^{-3\alpha \Delta/2}}{2\alpha(1 - e^{-\alpha \Delta})} \tag{25}$$

Similarly, the final term in Eq. 22 evaluates to:

$$e^{\alpha \Delta y_N} \cdot \frac{e^{-\alpha \Delta/2} - (\alpha \Delta + 1) e^{-3\alpha \Delta/2}}{2\alpha(1 - e^{-\alpha \Delta})}. \tag{26}$$

The second and third term are given by:

$$\frac{-2\alpha \Delta y_N - \alpha \Delta - 2}{4\alpha} \cdot e^{-\alpha \Delta(0.5 + y_N)} + \frac{1}{2\alpha}, \tag{27}$$

and

$$\frac{2\alpha \Delta y_N - \alpha \Delta - 2}{4\alpha} \cdot e^{-\alpha \Delta(0.5 - y_N)} + \frac{1}{2\alpha}, \tag{28}$$

respectively. Adding these results together, we get the following expression:

$$E[|X - X'||Y_N = y_N] = -\cosh(\alpha \Delta y_N) \cdot \Delta e^{-\alpha \Delta/2} \cdot \left(\frac{e^{-\alpha \Delta}}{1 - e^{-\alpha \Delta}} + 0.5 \right)$$

$$+ \Delta y_N \sinh(\alpha \Delta y_N) \cdot e^{-\alpha \Delta/2} + \frac{1}{\alpha}. \tag{29}$$