

Distributed Video Coding: Principles and Evaluation of Wavelet-Based Schemes

Riccardo Bernardini, Roberto Rinaldo and Pamela Zontone
University of Udine
Italy

1. Introduction

The current scenario considered for video coding and transmission, as assumed by the MPEG standards, uses complex encoding algorithms, including motion compensation and multi-hypothesis rate-distortion optimized procedures, to achieve high compression efficiency. This is the right solution when the encoding is carried out by high performance devices, while the receivers (e.g., handheld devices or mobile phones) should be kept as simple and cheap as possible. This scenario is rapidly evolving into a new one, where users have the interest to produce and transmit video and multimedia, possibly using their mobile and battery operated lightweight devices. Of course, this calls for new multimedia coding paradigms, where the *encoder* is as simple as possible to reduce the computational power requested for compression and radio transmission. Similar constraints should be considered in monitoring or surveillance applications, where a number of video sensors, with limited computational power, cooperate to send video information to a receiving station.

Distributed Source Coding (DSC) refers to the compression of two or more correlated sources that do not communicate with each other (hence the term distributed coding). These sources send their information to a central decoder that performs joint decoding. In this situation, the challenging problem is to achieve the same efficiency (the joint entropy of correlated sources) while not requiring sources to communicate with each other.

The Slepian-Wolf Theorem is a celebrated result of information theory which assures that this unexpected result is indeed achievable. In other words, it is possible to code two correlated sources (X, Y) one independently of the other, and achieve the same performance obtainable when a coder can exploit knowledge of both. For instance, in a conventional video coder, two consecutive frames can be compressed by computing the motion compensated difference between the first and the second frame, then transmitting the first frame in intra-mode and the inter-mode coded difference frame. The Slepian-Wolf result assures that, in principle, we can achieve the same coding efficiency by coding the two frames independently, i.e., without performing a costly motion compensation operation.

1.1 Chapter organization and contribution

The following chapter will be subdivided into the following sections.

1. An overview of the Slepian-Wolf and Wyner-Ziv theorems, with a short summary of essential Information theoretic concepts;

2. a review of the state-of-the-art in Distributed Source Coding and Distributed Video Coding (DVC), with references to the latest advancements;
3. a presentation of some original work by the authors on the design and evaluation of wavelet-based video coding;
4. an overview of applications of the DVC paradigm to other aspects besides video coding. In particular, we will present results relative to robust transmission of video using an auxiliary DVC stream;
5. conclusions, advantages and limitations of the DVC paradigm.

In particular, we will introduce an original distributed video coder based on processing the wavelet transform with a modulo-reduction function. The reduced wavelet coefficients are compressed with a wavelet coder. At the receiver side, the statistical properties between similar frames are used to recover the compressed frame. A second contribution is the analysis and the comparison of DVC schemes in two different scenarios: in the first scenario the information frames are separated from the other frames, and they are compressed following the original framework considered for Wyner-Ziv coding. In the second scenario, all the frames are available at the encoder making this an interesting proposal for the design of a low-complexity video coder, with no motion compensation, where the information frames are coded using DSC techniques. The whole set of experiments show that the proposed schemes - that do not use any feedback channel - have good performance when compared to similar asymmetric video compression schemes considered in the literature. Finally, we will consider an original error-resilient scheme that employs distributed video coding tools. A bitstream, produced by any standard motion-compensated predictive codec, is sent over an error-prone channel. A Wyner-Ziv encoded auxiliary bitstream is sent as redundant information to serve as a forward error correction code. We propose the use of an extended version of the Recursive Optimal per-Pixel Estimate (ROPE) algorithm to establish how many parity bits should be sent to the decoder in order to correct the decoded and concealed frames. At the decoder side, error concealed reconstructed frames and parity bits are used by the Wyner-Ziv decoder, and each corrected frame is used as a reference by future frames, thus reducing drift. Tests with video sequences and realistic loss patterns are reported. Experimental results show that the proposed scheme performs well when compared to other schemes that use Forward Error Correcting (FEC) codes or the H.264 standard intra-macroblock refresh procedure.

2. Foundation of Distributed Source Coding

In this section we will introduce in detail two major results provided by Information Theory that prove that, under the DSC paradigm, it is still possible to achieve or to approach, in total generality, the optimal performance of a joint coder: the Slepian-Wolf theorem and the Wyner-Ziv theorem. We first introduce the main ideas behind distributed source coding by means of examples.

2.1 A glimpse at Distributed Source Coding

Consider the case of N distributed sensors that communicate with a data collection center using a radio link. Data recorded by each sensor, at each time instant, can be modelled as a set of X_1, \dots, X_N , random variables, and it is reasonable to expect that there is a strong correlation between data at each sensor. In the case of two sensors X and Y , suppose we can model Y as a random variable with equiprobable integer values in the set $\{0, \dots, 7\}$ while $d = X - Y$,

the difference between the values recorded by the two sensors, can be modelled as a variable assuming values in $\{0, \dots, 3\}$ with equal probability. Assuming d and Y independent, X would have a probability mass function as depicted in Fig. 1,

$$p_X(k) = \sum_{h=0}^7 Pr[d = k - h, Y = h] = \sum_{h=0}^7 p_d(k - h) p_Y(h).$$

We know from Shannon’s theory (6) that the entropy of X

$$H(X) \triangleq - \sum_k p_X(k) \log_2(p_X(k))$$

represents the minimum average number of bits per source symbol necessary to represent the source. In this case, $H(X)$ is obviously greater than the entropy $H(Y) = 3$ bit of Y (the eight equiprobable values of Y can be coded most efficiently with 3 bits). A simple calculation would show that $H(X) \simeq 3.328$ bit.

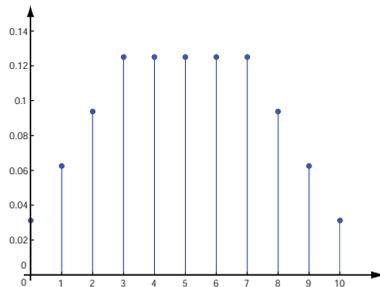


Fig. 1. The probability mass function of X .

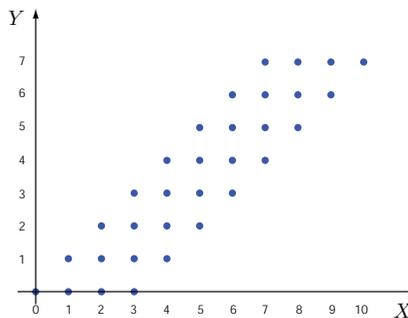


Fig. 2. Support of the probability mass function of (X, Y) .

The pair (X, Y) assumes 32 equiprobable values in the set represented in Fig. 2, with associated 5 bit entropy. As it is well known, supposing pairs (X, Y) are generated independently, $H(X, Y) = 5$ bit represents a lower bound for the number of bits (per source pair) with which we can represent sequences of pairs generated by the two sources. If a coder has access both

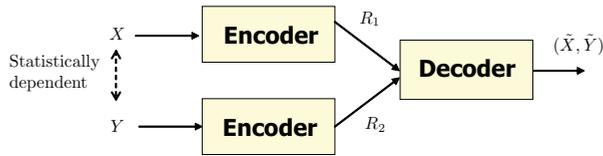


Fig. 3. Distributed Source Coding. Correlated variables X and Y are compressed independently and sent to a joint decoder.

to X and Y , the way to code (X, Y) is obvious: one can use 3 bits to represent the 8 values of Y , and other 2 bits to represent $d = X - Y$.

How can we proceed if the coders for X e Y cannot communicate, as in Fig. 3? Are we forced to use an efficient coding procedure for X and Y , separately, and use $R_1 = H(X)$ and $R_2 = H(Y)$ bit per symbol, respectively? Note that this procedure is inefficient to code the pair, since $H(X, Y) \leq H(X) + H(Y)$.

Fig. 4 illustrates a coding procedure which allows optimal coding even if the two coders do not communicate with each other and act separately. In particular, one can use a unique code (in the figure, a different shape symbol) for the subsets $\{0, 4, 8\}$, $\{1, 5, 9\}$, $\{2, 6, 10\}$, $\{3, 7\}$, of the X alphabet. For each value of X , its coder transmits to the decoder the index (the shape symbol) of the subset to which the value of X belongs. With four subsets, 2 bits are sufficient. Observe from the figure that, for each value of Y , we have 4 possible values of X , each one belonging to a *different* subset. Thus, if we know at the decoder the value of Y (represented with $R_2=3$ bit), and the index of the subset to which X belongs (2 bit), it is possible to uniquely decode the *pair* (X, Y) . Note that the two coders act independently, and each of them transmits the index corresponding to the actual value of Y , or the subset index for the actual value of X in each experiment. In particular, we have $R_1 = H(X|Y) = 2$ bit and $R_2 = H(Y) = 3$ bit.

Fig. 5 shows a different example, where the pair (X, Y) assumes 16 equiprobable values (the joint entropy is therefore $H(X, Y) = 4$ bit), and one uses a distributed coding procedure that associates a different code symbol J to the subsets $\{-1, -5\}$, $\{-3, -7\}$, $\{1, 5\}$, $\{3, 7\}$ of Y , using two bits, and one symbol I for the subsets $\{-7, 1\}$, $\{-5, 3\}$, $\{-3, 5\}$, $\{-1, 7\}$ of X values, using other 2 bits. From the pair of code symbols (I, J) , one can uniquely identify (X, Y) , but the two coders can act independently. Note that in this case we have $R_1 > H(X|Y) = 1$ bit, $R_2 > H(Y|X) = 1$ bit, and $R_1 + R_2 = H(X, Y) = 4$ bit.

In these two examples, one necessary requirement is that each pair (X, Y) can be uniquely identified by (i, j) , where $i \in I$ and $j \in J$ identify the labels for subsets of X and Y values, respectively. Note that, in order for the procedure to work, the total number of label pairs $|I||J|$ must be at least as large as the number of (X, Y) pairs with non-zero probability. Moreover, for each value of X (respectively, Y), there must be a sufficient number of labels J (respectively, I) to uniquely identify the non-zero probability corresponding pairs (X, Y) . In addition, the key point is to associate a code symbol to a subset (or *bin*) of values of one variable that are sufficiently far apart, so that its exact value can be discriminated once the value (or set of possible values) of the other variable is also known.

The preceding considerations can be justified by a general result of Information Theory derived by Slepian and Wolf in 1973 (7), which we describe below.

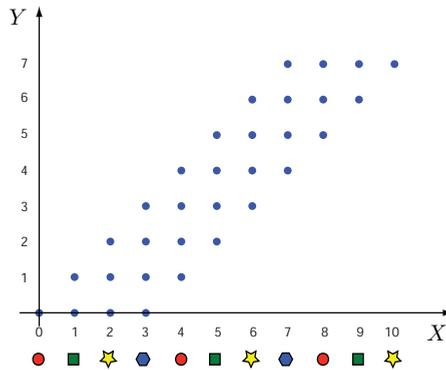


Fig. 4. A “distributed” code for (X, Y) .

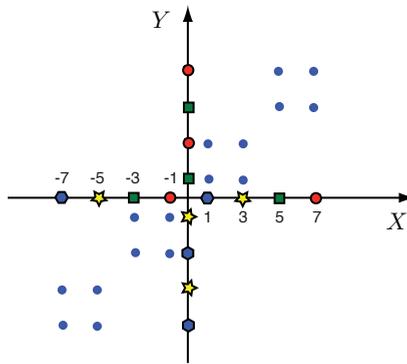


Fig. 5. A “distributed” code for (X, Y) .

2.2 The Slepian-Wolf theorem

Let $(X_i, Y_i)_{i=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d) drawings of a pair of correlated discrete random variables X and Y . For lossless reconstruction, a rate given by the joint entropy $H(X, Y)$ is sufficient if we perform joint coding. The Slepian-Wolf theorem refers to the case of X and Y separately encoded but jointly decoded, i.e., the encoder of each source is constrained to operate without knowledge of the other source, while the decoder has available both encoded message streams (see Fig. 3). It appears that the rate at which we can code the two sources in this case is $H(X) + H(Y)$, which is greater than $H(X, Y)$ if X and Y are not independent.

Let X take values in the set $\mathcal{A}_X = \{1, 2, \dots, A_X\}$ and Y in the set $\mathcal{A}_Y = \{1, 2, \dots, A_Y\}$. Denote their joint probability distribution by

$$p_{X,Y}(x, y) = P(X = x, Y = y) \quad x \in \mathcal{A}_X, y \in \mathcal{A}_Y.$$

Next, let $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ be a sequence of n independent realizations of the pair of random variables (X, Y) . Denote by \mathbf{X}^n the block sequence of n -characters X_1, X_2, \dots, X_n produced by the source X , and by \mathbf{Y}^n the block sequence Y_1, Y_2, \dots, Y_n produced by the other

source. The probability distribution for this correlated pair of vectors is

$$p_{\mathbf{X}^n, \mathbf{Y}^n}(\mathbf{x}^n, \mathbf{y}^n) = P(\mathbf{X}^n = \mathbf{x}^n, \mathbf{Y}^n = \mathbf{y}^n) = \prod_{i=1}^n p_{X,Y}(x_i, y_i) \quad (1)$$

$$\mathbf{x}^n = (x_1, x_2, \dots, x_n) \in \mathcal{A}_X^n$$

$$\mathbf{y}^n = (y_1, y_2, \dots, y_n) \in \mathcal{A}_Y^n$$

where \mathcal{A}_X^n is the set of \mathcal{A}_X^n distinct n -vectors whose components are in \mathcal{A}_X and \mathcal{A}_Y^n is defined analogously.

The first encoder (see Fig. 3) maps the input \mathbf{X}^n to the index $I = C_1(\mathbf{X}^n)$, where $I \in \mathcal{M}_X = \{1, 2, \dots, M_X\}$; similarly, the other encoder maps the input \mathbf{Y}^n to the index $J = C_2(\mathbf{Y}^n)$, where $J \in \mathcal{M}_Y = \{1, 2, \dots, M_Y\}$. I and J are called *encoded- X message number* and *encoded- Y message number*, respectively (7). At the decoder side, the joint decoder is a function $g : \mathcal{M}_X \times \mathcal{M}_Y \rightarrow \mathcal{A}_X^n \times \mathcal{A}_Y^n$ such that

$$g(C_1(\mathbf{X}^n), C_2(\mathbf{Y}^n)) = (\tilde{\mathbf{X}}^n, \tilde{\mathbf{Y}}^n).$$

Let P_e^n be the probability of decoding error, i.e., $P_e^n = P[(\mathbf{X}^n, \mathbf{Y}^n) \neq (\tilde{\mathbf{X}}^n, \tilde{\mathbf{Y}}^n)]$.

Associated with these encoders and the joint decoder are rates $R_1 = (1/n)\log M_X$ and $R_2 = (1/n)\log M_Y$. We think of the two encoders as producing the integers I and J after n correlated source pairs (X, Y) have been generated. R_1 units of information per source character are sufficient to transmit I to the joint decoder and R_2 units are sufficient to transmit J . The decoder then produces the estimates $\tilde{\mathbf{X}}^n$ and $\tilde{\mathbf{Y}}^n$ of the input sequences \mathbf{X}^n and \mathbf{Y}^n .

The pair of rates R_1 and R_2 is said to be an *admissible rate point* (7) if for every $\epsilon > 0$ there exist for some $n = n(\epsilon)$ encoders and decoders (considering the case with two decoders as well) with $M_X = \lfloor \exp(nR_1) \rfloor$ and $M_Y = \lfloor \exp(nR_2) \rfloor$ such that $P_e^n < \epsilon$. Here the symbol $\lfloor \cdot \rfloor$ denotes the largest integer not greater than the argument of the function. In other words, the pair of rates R_1 and R_2 is an admissible rate point if it is possible to construct a sequence of codes with rate R_1 for \mathbf{X}^n and rate R_2 for \mathbf{Y}^n , such that $P_e^n \rightarrow 0$ with $n \rightarrow \infty$.

The *achievable rate region* is the closure of the set of admissible rate points.

The **Slepian-Wolf theorem** says that if R_1 is the rate corresponding to the coding of X and R_2 to the coding of Y (see Fig. 3), the *achievable rate region* of DSC is given by:

$$R_1 \geq H(X|Y), \quad (2)$$

$$R_2 \geq H(Y|X), \quad (3)$$

$$R_1 + R_2 \geq H(X, Y). \quad (4)$$

Fig. 6 shows the achievable region for the Slepian-Wolf theorem. The Slepian-Wolf theorem suggests, therefore, that it is possible to compress statistically dependent signals, in a distributed scenario, to the same rate as with a system where the signals are compressed jointly.

The proof of the Slepian-Wolf theorem uses, as it is common in Information Theory, the concepts of typical set and of random coding. We give here the main ideas, while a complete development can be found, for instance, in (6). As a matter of fact, it can be shown, using the Law of Large Numbers, that, for large n , there are basically $2^{nH(X)}$ highly probable (*typical*) \mathbf{X}^n sequences, while the other possible source sequences are generated with vanishing

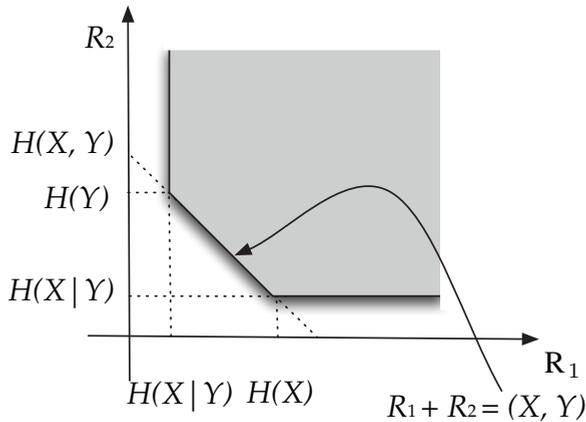


Fig. 6. Achievable rate region for two sources in Distributed Source Coding.

probability. Along the same lines, there are $2^{nH(Y)}$ sequences \mathbf{Y}^n , and only $2^{nH(X,Y)}$ jointly typical sequences $(\mathbf{X}^n, \mathbf{Y}^n)$. Furthermore, for each typical $\mathbf{X}^n = \mathbf{x}^n$, there are about $2^{nH(Y|X)}$ jointly typical \mathbf{Y}^n and $2^{nH(X|Y)}$ jointly typical sequences for each $\mathbf{Y}^n = \mathbf{y}^n$. The requirements of the Theorem become therefore obvious: the number 2^{nR_1} of I labels and 2^{nR_2} of J labels have to be large enough so that $(i, j), i \in I, j \in J$, can identify uniquely each typical $(\mathbf{X}^n, \mathbf{Y}^n)$ (this is ensured by Eq. (4)). Moreover, since the codes are assigned by separate coders, for each typical value of \mathbf{X}^n (respectively, \mathbf{Y}^n), there must be at least $2^{nH(Y|X)}$ possible labels (respectively, $2^{nH(X|Y)}$). This is indeed ensured by conditions in Eq. (2) and Eq. (3). Note that, in the examples of the previous section, we had similar requirements to code each random pair. In addition, the conditions of the Theorem assure that, as n increases, there are exponentially many more labels than sequences in the typical sets, so that choosing the codes at random will provide unique labelling with high probability.

In Slepian-Wolf coding schemes, it is practical to design codes to approach one corner point of the achievable region (Fig. 6), e.g., with $R_1 = H(X|Y)$ and $R_2 = H(Y)$. This problem is known as the problem of source coding X with *side information* Y at the decoder, and it can be referred to as an *asymmetric scenario*. In this asymmetric context, therefore, the aim is to code X at a rate that approaches $H(X|Y)$ based on the correlation model between X and Y and not using the specific Y at the encoder.

2.3 Practical schemes for Distributed Coding

Wyner first realized the close connection between DSC and channel coding (9), suggesting the use of linear channel codes as a constructive approach for Slepian-Wolf coding (10). The basic idea in Wyner's work was to partition the space of all possible source outcomes into disjoint bins that are the cosets of some good linear channel code for the specific correlation model between X and Y .

If we take two binary symmetric sources and we consider an (n, k) binary linear block code, there are 2^{n-k} different syndromes, each one indexing a different bin of 2^k binary words of length n . Each bin is a coset code of the block code, which means that the distance properties of the original code are preserved in each bin. In particular, the Hamming distance between any

two source binary words in the coset (i.e., the number of symbols in which the two codewords differ) is at least equal to the *minimum distance* of the linear code, a characteristic that has to be established at design time. As a matter of fact, a linear code with minimum distance $d = 2t + 1$ can successfully correct any error vector \mathbf{e} with t symbol errors in the received noisy codeword.

As we will see below, a practical scheme for distributed coding consists in sending to the receiver, for a sequence of n input bits, the corresponding $(n - k)$ syndrome bits, thus achieving a compression ratio $\frac{n}{n-k}$. This approach was only recently used for practical Slepian-Wolf code schemes based on conventional channel codes. If the correlation model between X and Y can be seen as a binary channel, this syndrome concept can be extended to all binary linear codes such as Turbo and LDPC codes.

In a typical transmission system, given an (n, k) systematic linear channel code with the $(n - k) \times n$ parity-check matrix \mathbf{H} , and using this channel code for error correction, the length- k input message is transformed into a length- n message \mathbf{X} by appending $n - k$ parity bits. The codeword \mathbf{X} has now length n and $n - k$ syndrome bits are computed as $\mathbf{s} = \mathbf{X}\mathbf{H}^T$.

We transmit the codeword \mathbf{X} and we receive a vector $\mathbf{Y} = \mathbf{X} + \mathbf{e}$, where \mathbf{e} is the error vector which indicates the positions where the received vector \mathbf{Y} differs from the transmitted one \mathbf{X} . As it is well known, knowledge of the syndrome \mathbf{s} allows to determine the minimum weight $\mathbf{e} = g(\mathbf{s})$ such that $\mathbf{Y} = \mathbf{X} + \mathbf{e}$. At the receiver side, if $g(\mathbf{Y}\mathbf{H}^T) = g((\mathbf{X} + \mathbf{e})\mathbf{H}^T) = g(\mathbf{e}\mathbf{H}^T)$ is the decoding function based on the syndrome, we can write therefore $\mathbf{e} = g(\mathbf{e}\mathbf{H}^T)$ with probability close to 1 and recover from this the original codeword \mathbf{X} .

We see now how a similar procedure can be used to code \mathbf{X} and recover it from the side-information \mathbf{Y} in a distributed coding scenario. A length- n vector \mathbf{X} of source symbols is compressed as the $n - k$ bit syndrome $\mathbf{s} = \mathbf{X}\mathbf{H}^T$ of a linear code. The syndrome is sent to the receiver, where the side information \mathbf{Y} is available. Suppose the correlation model implies that, to each n -length source binary word \mathbf{X} , corresponds the side-information vector $\mathbf{Y} = \mathbf{X} + \mathbf{e}$, where \mathbf{e} is an error vector that can be corrected by the code with probability close to 1. Then, it is possible to reconstruct \mathbf{X} with the knowledge of the syndrome \mathbf{s} and \mathbf{Y} . In fact, if $g(\cdot)$ denotes the decoding function based on the syndrome, we can calculate the difference $\mathbf{Y}\mathbf{H}^T - \mathbf{s} = (\mathbf{Y}\mathbf{H}^T - \mathbf{X}\mathbf{H}^T) = (\mathbf{Y} - \mathbf{X})\mathbf{H}^T = \mathbf{e}\mathbf{H}^T$, derive $\mathbf{e} = g(\mathbf{e}\mathbf{H}^T)$ and finally determine $\mathbf{X} = \mathbf{Y} - \mathbf{e}$.

In summary, the source messages can be partitioned by means of a linear channel code, in such a way that all the messages with the same syndrome are assigned to the same coset. The messages in the coset are sufficiently far apart, since they are separated, at least, by the minimum distance of the code. The receiver identifies the coset from knowledge of the syndrome. Furthermore, using the side-information, it can discriminate the actual source message, as soon as the differences between the side-information and the source message can be corrected by the code. An alternative partition can be obtained by assigning to the same coset all the messages that generate the same parity bits. This last approach is known to be suboptimal (11) since there is no guarantee that these cosets have the good geometrical properties of the syndrome-based cosets in terms of minimum distance of the elements in each coset.

A practical correlation model that is often assumed between binary X and Y is the binary symmetric model where the correlation between X and Y is modeled by a binary symmetric channel (BSC) with cross-over probability p . We know that for this channel $H(X|Y) = H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)$. Although this model looks simple, the Slepian-Wolf coding problem is not trivial.

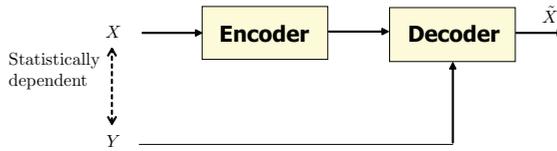


Fig. 7. Asymmetric scenario: source coding of X with *side information* Y .

2.4 Wyner-Ziv theorem

The Slepian-Wolf theorem is focused on the case of lossless compression of two correlated sources. The counterpart of this theorem for lossy source coding is the Wyner and Ziv’s theorem on source coding with side information (8). The theorem considers the problem of how many bits are needed to encode X under the constraint that the average distortion between X and the coded version \tilde{X} does not exceed a given distortion level, assuming the side information available at the decoder but not at the encoder (see Fig. 7). In detail, let $(X_i, Y_i)_{i=1}^{\infty}$ be a sequence of independent and identically distributed (i.i.d) drawings of a pair of correlated discrete random variables X and Y . Let X take values in the set $\mathcal{A}_X = \{1, 2, \dots, A_X\}$. Denote by \mathbf{X}^n the blocks of n -characters X_1, X_2, \dots, X_n that are coded into a binary stream of rate R , which can in turn be decoded as a sequence $\tilde{\mathbf{X}}^n$. The average distortion level is $1/n \sum_{i=1}^n E[d(X_i, \tilde{X}_i)]$, where $d(x, \tilde{x}) \geq 0, x \in \mathcal{A}_X$, is a pre-assigned distortion measure.

Let $R^*(D)$ be the infimum of rates R such that communication is possible at an average distortion level not exceeding $D + \epsilon$ (with $\epsilon > 0$ arbitrarily small and with a suitably large n) when only the decoder has access to the side information \mathbf{Y}^n ; let $R_{X|Y}(D)$ be the rate-distortion function which results when the encoder as well as the decoder has access to the side information. In (8) it is shown that when $D > 0$ then

$$R^*(D) > R_{X|Y}(D).$$

Therefore, knowledge of the side information at the encoder allows the transmission of \mathbf{X}^n at a given distortion level using a smaller transmission rate.

With this theorem, we can notice that a Wyner-Ziv scheme suffers some rate loss when compared to lossy coding of X when the side information Y is available at both the encoder and the decoder. One exception is when X and Y are jointly gaussian and the MSE (Mean Squared Error) distortion measure is used. There is no rate loss with Wyner-Ziv coding in this case, which is of special interest in practice; in fact, as a first approximation, many images and video sources can be modeled as jointly gaussian, and so may be the case for measured values in sensor networks applications.

Finally, it is easy to show that, in case of discrete variables and zero distortion, we obtain the Slepian-Wolf theorem:

$$R^*(0) = R_{X|Y}(0) = H(X|Y).$$

3. State-of-the-art in DSC and DVC

Recently, several schemes based on the Slepian-Wolf (and its continuous variable counterpart – Wyner-Ziv) theorem have been proposed for distributed video coding (DVC). In general, the current implementations consider X as a noisy version of Y . Typically, X and Y are constructed

as the bitplanes of some representation of the source (in the pixel, or transform domain) so that efficient binary codes (like Turbo or LDPC codes) can be used. In particular,

1. one coder transmits information (with a standard coding scheme), from which the decoder can calculate the side-information Y ;
2. an independent coder protects X by means of an error correction code;
3. the independent coder transmits the code syndrome or parity bits to represent X ;
4. the receiver, by exploiting the protection properties of the error correction code, recovers X from its "noisy" version Y and the code syndrome or parity bits (see Fig. 8).

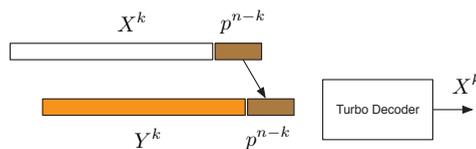


Fig. 8. Distributed coding using the parity bits of an efficient Turbo binary code.

This paradigm has been implemented in some practical schemes presented in the literature. In (12), Pradhan and Ramchandran presented a syndrome-based framework that employs trellis-coded quantization and trellis channel codes, successively extended to a video coding system called PRISM (13). The idea is to consider every video frame as a different source; DSC allows the encoding of the frames without performing motion estimation at the encoder, with performance similar to a standard video coder that exploits the temporal correlation between consecutive frames. Hence, this scheme requires a light encoder and a complex decoder.

Other works are based on channel codes such as Turbo Codes and LDPC codes (14; 15). In particular, in (14) Aaron et al. apply a Wyner-Ziv coding to the pixel values of a video sequence. The reference scheme of (14), with two separate coders and a joint decoder, assumes that the video sequence is divided into Key frames (i.e., the even frames of the sequence), and Wyner-Ziv (WZ) frames (the odd frames). One coder codes the Key frames without knowledge of the WZ frames and sends them to the decoder. The decoder computes a prediction of the WZ frames that will be used as side information in the distributed coding paradigm. Such an approach is extended to the transform domain in (15). The DCT transform enables the coder to exploit the statistical dependencies within a frame, and so better rate-distortion performance can be achieved. In general, LDPC codes show some performance advantage with respect to Turbo codes. More recently, in (24) a probability updating technique (PUT) to enhance the Turbo coding performance in the context of Wyner-Ziv video coding has been presented.

The algorithms to generate the side information at the decoder influence significantly the rate-distortion performance of the Wyner-Ziv video coding schemes. The techniques described in (28; 29) were selected for the DISCOVER mono-view codec (21). The architecture of this codec is based on the scheme proposed in (15) but many improvements have been added in order to enhance the performance of the basic building blocks. However, as in the original scheme, a feedback channel is still used to request more parity bits until the decoder reconstruction is successful. An improved side information generation method using field coding has been also proposed in (23). WZ frames are divided into the top and bottom fields

as the field coding of a conventional video codec. Top fields are coded with the generation method presented in (28) and bottom fields are reconstructed using the information of the already decoded top fields. Hash-based motion estimation approaches have been presented in (26; 27). In these schemes additional bits are sent by the WZ encoder to aid the decoder in estimating the motion and generate the side information.

Other possible schemes have been presented in the literature. In (16) the pixels of a frame are divided into two sub frames: the key sub frame, consisting of the odd vertical pixel lines, is conventionally encoded and it is used at the decoder to compute the side information that will be used to reconstruct the Wyner-Ziv sub frame (the even vertical pixel lines of the original frame). In (17) Tagliasacchi et al. propose another WZ sub frame coding. The WZ frames are split in two parts: the first part is decoded using the side information only (obtained from the Key frames). The second part is instead decoded using the side information and the previously decoded WZ sub frame.

Wavelet based coding schemes have the potential advantage to naturally allow multiresolution and embedded coding. A wavelet domain DVC scheme has been proposed in (30). A pair of lattice vector quantizers (LVQ) is used to subtract the dependence between wavelets coefficients. The Authors extend the motion compensation refinement concept of pixel domain to wavelet domain and propose a new search strategy for vector reconstruction. In (31), a wavelet domain DVC scheme based on the zero-tree entropy (ZTE) coding is then presented. The wavelet coefficients are quantized using scalar quantization and reorganized in terms of the zero-tree structure. Only the significant coefficients are encoded with a Turbo coder and the punctured parity bit are transmitted. In (32), the Authors exploit the multiresolution properties of the wavelet decomposition to refine motion estimation at the receiver, in order to improve the quality of the side information.

In (18) a scalable video coding scheme is proposed, which performs the DSC between the base and the enhancement layer. In (19), instead, the DSC principles are applied to hyperspectral images. A technique for Wyner-Ziv coding on multispectral images based on a set theory is investigated in (20). Recent advances in multi-view distributed video coding have been also reported in (25).

4. Wavelet-based video coding schemes

In this section we present and compare different Distributed Video Coding (DVC) schemes based on the use of the wavelet transform, which naturally allows for spatial and other forms of scalability. The results presented here summarize the content of (1–4).

The video frames are separated into Key frames, i.e., the ones that are coded using standard techniques and sent to the receiver, and Wyner-Ziv (WZ) frames, which are coded using the distributed coding paradigm. For the results we present below, the Key frames are the even frames of the sequence, while the WZ frames are the odd ones, as in (14).

Two scenarios have been considered (see Fig. 10). In the first, the WZ frames are encoded independently of the Key frames, and the Key frames are encoded and decoded using a conventional intraframe codec. This is the original framework considered for Wyner-Ziv coding, e.g., in (14; 15). In the second scenario, all frames (Key frames and WZ frames) are available at the encoder. This scenario is interesting for the design of a low-complexity video coder, with no motion compensation, and where half of the frames (the WZ frames) are coded using distributed source coding techniques. This framework is considered, for example, in (33).

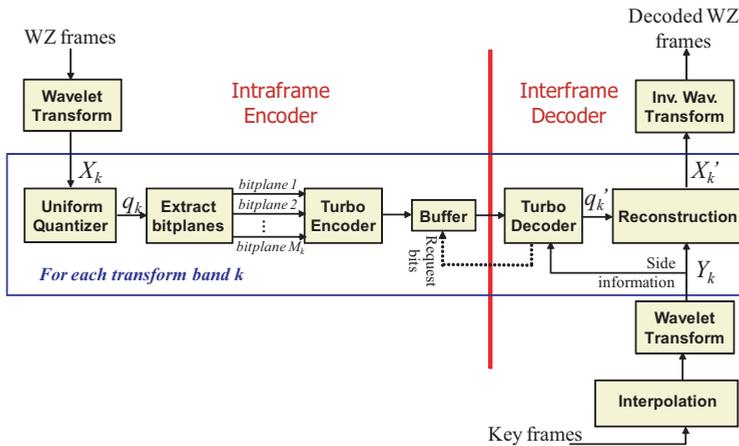


Fig. 9. The reference wavelet based DVC scheme.

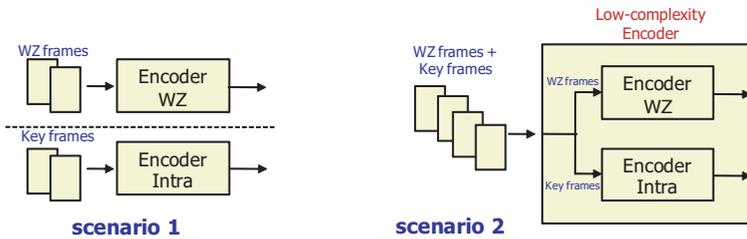


Fig. 10. The considered scenarios.

The reference scheme, which we describe below, is an extension of the one considered in (15), and operates in the wavelet domain, according to the scheme of Fig. 9.

4.1 Wyner-ziv wavelet domain scheme

This scheme operates on the Wavelet Transform of the WZ frames. A three level, ten band wavelet transform is considered for QCIF sequences (see Fig. 11.a). At the encoder, the wavelet transform coefficients are grouped together to form coefficient subbands. Each subband is then quantized using a midtreand uniform quantizer where the quantization step is set to be equal for all the subbands (this is the optimal solution for orthogonal transforms). Bits are assigned according to a modified sign/module labeling procedure (4). For each subband, the bitplanes are then independently coded using a *Rate Compatible Punctured Turbo* (RCPT) coder. Using different puncturing schemes, it is possible to send incremental subsets of parity bits, thus allowing to vary the protection offered by the coder to the bitstream.

At the decoder, the side information is generated from the Key frames using temporal interpolation based on Motion Compensated (MC) interpolation with symmetric motion vectors (34). The purpose of this procedure is to reconstruct at the receiver a good approximation of each WZ frame, which will be used by the decoder as side information. The parity bits sent by the encoder are used to recover the bitplanes of the wavelet transform of the WZ frames from those of the side-information.

This scheme uses a feedback channel, and to allow the decoder to request additional parity bits until correct decoding is possible, we consider the transmission of a 16 bit CRC code for each bitplane. If the transmitted CRC does not match with the decoded bitplane, the decoder requests additional parity bits from the encoder buffer until the reconstructed bitplane matches the CRC and the decoding is declared to be successful.

As in (15), the iterative turbo decoder uses information about already decoded bitplanes to improve a-priori knowledge while decoding the next bitplane. Moreover, since the WZ frames are typically quantized more coarsely than the Key frames, the decoder implements a Maximum Likelihood reconstruction strategy, where the WZ wavelet coefficient is reconstructed as the value in the quantization interval, determined on the basis of the WZ decoded bitplanes, which is closest to the value of the side-information. The scheme considered in this section has performance similar to the one of (15), with the possible advantage that the use of the wavelet transform naturally allows for various forms of scalability.

4.2 Hybrid wavelet domain Wyner-ziv scheme with rate estimation

One of the drawbacks of the scheme described above, is that it requires a feedback channel to request additional parity bits, until successfully decoding is achieved (with a residual decoding error probability if the CRC fails). The use of the feedback channel may not be possible in certain applications, e.g., interactive applications, live streaming or multicast transmission, because of the excessive delay that is introduced by the procedure.

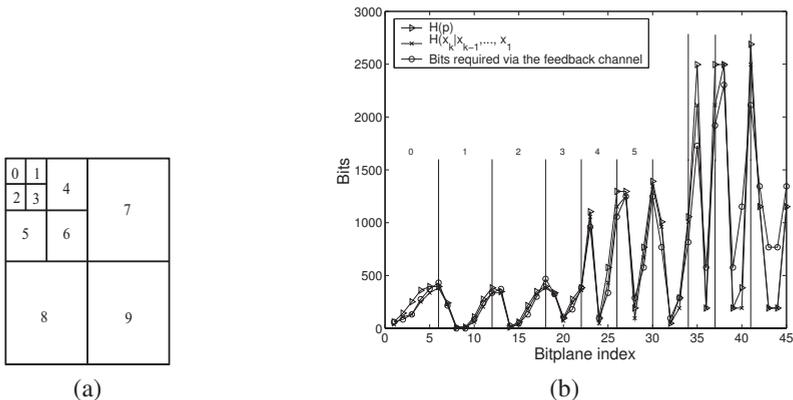


Fig. 11. (a) The 3-level wavelet transform; (b) bitrate prediction using the statistical models.

Here we present a scheme that does not use a feedback channel, and includes a procedure to estimate the required bitrate for WZ frames at the encoder. Since the decoder cannot make requests to the WZ encoder, it is necessary that the latter estimates the required parity bits for each wavelet coefficient bitplane.

In particular, we propose that the WZ wavelet coefficients in each subband are related to those of the corresponding side information according to the model $X = Y + e$, where Y and e are independent random variables, and e has a Laplacian distribution, i.e., e has a probability density function

$$f_e(a) = \frac{\alpha}{2} e^{-\alpha|a|}. \tag{5}$$

Let us denote with x_k the k -th bitplane of X , with x_1 being the most significant bit. We show in (4) that, as suggested by the Slepian-Wolf theorem, the conditional entropy

$$H(x_k|x_{k-1}, \dots, x_1, Y) \quad (6)$$

provides a good estimate of the required WZ bitrate for the bitplanes of coefficients belonging to the lower resolution subbands (in particular, subbands 0-6 in Fig. 11.a). Note that $H(x_k|x_{k-1}, \dots, x_1, Y)$ can be computed at the encoder using the model $X = Y + e$, by estimating α in Eq. (5) based on an approximation \tilde{Y} of the side-information that will be constructed at the receiver.

In particular, in the first scenario (see Fig. 10), the Key frames are not available at the encoder. Therefore, we compute the average of the WZ frames closest to the current frame, and approximate the side information as the wavelet coefficients \tilde{Y} of this average. In the second scenario, \tilde{Y} is the wavelet coefficient of the average of the Key frames closest to the current frame. The two scenarios differ only for the side information \tilde{Y} which is constructed at the transmitter for rate estimation.

We show in (4) that the entropy $H(p)$ corresponding to the bitplane crossover probability $p = P[x_k \neq y_k]$ (1; 35; 36) also provides an acceptable estimate of the required bitrate, with $H(p)$ assuming a more conservative larger value. Note that, if one assumes the binary symmetric channel model $x_k = y_k + q_k$, where q_k is independent on y_k , $P[q_k = 1] = p$, and the sum is modulo 2, we have $H(p) = H(x_k|y_k)$. This is consistent with Eq. (6), where dependence from WZ and side information bitplanes, other than the current bitplane, is neglected. Entropy $H(x_k|y_k)$ or probability p can be computed from x_k , which is known at the encoder, and y_k , calculated from an approximation \tilde{Y} of the side information.

For high resolution subbands (subbands 7-9 in Fig. 11.a), the models tend to underestimate the required bitrate thus leading to incorrect decoding. Therefore, a hybrid procedure where the quantized high resolution subbands are entropy coded using low-complexity intra-coding procedures (37) is proposed. For the lower resolution subband, $H(p)$ of the bitplane crossover probability $p = P[x_k \neq y_k]$ (1; 35; 36) is used as the estimate. As an example, Fig. 11.b shows the required bits for each bitplane of all wavelet subbands for one frame of the QCIF sequence *Teeny*, quantized with a quantization step $\Delta = 32$. The vertical lines and the index from 0 to 9 separate the bitplanes of different subbands. In the figure $H(p)$, the entropy and the bitrate actually requested via the feedback channel are shown.

4.3 DVC via modulo reduction

In (1; 4) we propose an alternative procedure for DVC that does not use Turbo codes and does not require feedback from the receiver. As seen in Fig. 12, it comprises three steps: 1) reduction modulo M of the unquantized original wavelet coefficient X to obtain the *reduced variable* $\bar{X} = \Phi_M(X) \triangleq X \bmod M$ (see Fig. 13); 2) lossy coding of \bar{X} . The reduced coefficients can be compressed by means of an efficient wavelet coder. In our implementation we use the low complexity coder presented in (37), but other choices are possible; 3) at the receiver, maximum likelihood (ML) decoding of X from quantized \bar{X} and side information Y . As before, the side information Y is generated by temporal interpolation based on Motion Compensated (MC) interpolation with symmetric motion vectors (34). In (4) it is discussed how to choose M to guarantee the recovery of X , after detailing the reconstruction procedure.

The idea behind this scheme can be understood with the help of Fig. 13. The original coefficient X is reduced modulo M to obtain \bar{X} , thus producing values in $[-M/2, M/2]$.

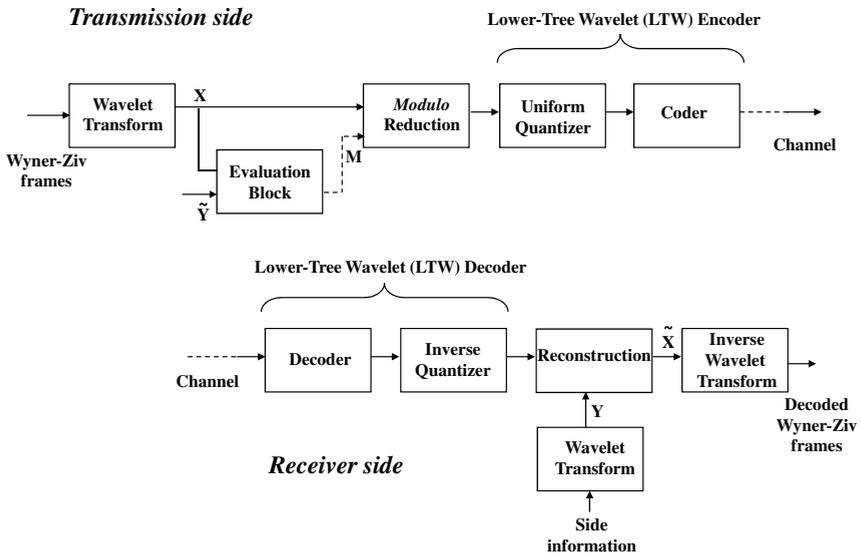


Fig. 12. The proposed scheme with the Evaluation block.

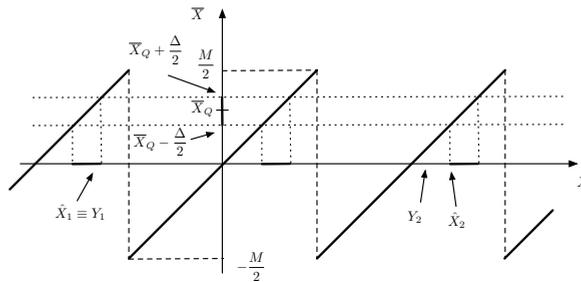


Fig. 13. Construction of the reduced variable \bar{X} , and examples of the reconstruction rule at the receiver.

The reduced range values \bar{X} can therefore be quantized and coded more efficiently than the original coefficients. Knowledge of the quantized \bar{X} value \bar{X}_Q at the receiver, allows to conclude that the original X belongs to the set $I_Q + nM$, $n \in Z$, given by the translations of the correct quantization interval for X . Finally, one selects n so that the reconstructed X belongs to the set $I_Q + nM$ and is closest to the side-information Y at the receiver.

It is worth giving the rationale behind this scheme by comparing it with a syndrome-based Wyner-Ziv scheme. In a WZ scheme, instead of transmitting each bitplane of X , we transmit a *syndrome* which allows the receiver to deduce that the encoded binary word belongs to a *coset*; similarly, in the proposed scheme, from the knowledge of \bar{X} one can deduce that X belongs to the *coset* $\Phi_M^{-1}(\bar{X}) = \{\bar{X} + nM; n \in Z\}$ (see Fig. 13; we neglect here the effect of quantization). The reduced value \bar{X} can be interpreted as an *analog syndrome* of X . At the receiver, ML reconstruction estimates X by choosing the element of $\Phi_M^{-1}(\bar{X})$ that is closest to the side information Y . Disregarding quantization, it is clear that no error occurs if $|X - Y| < M/2$.

In the usual syndrome-based Wyner-Ziv paradigm, the number of bits of the syndrome must be large enough to allow for the correction of all the “flipped” bits in the bitplanes of X and of the side information. If the syndrome length is not sufficient, X is recovered with an error; similarly, in the proposed scheme, the value of M is chosen large enough to grant for the reconstruction, and if the value of M is underestimated, errors will occur. The major difference between this scheme and a classical WZ scheme is that having an *analog syndrome* allows us to move the quantizer *after* the syndrome computation and use any lossy scheme to encode the reduced values.

4.4 Experimental results for wavelet-based DVC

To have an idea of the performance which can be obtained with DVC schemes, we report here some experiments with the wavelet-based schemes considered above. Further experiments and details can be found in (4).

We consider 299 frames of the QCIF *Foreman* sequence, and 73 frames of the QCIF *Teeny* sequence, coded at 30 frames/s. Only the performance relative to the luminance component of the WZ frames (i.e., the even frames) is considered. The Key frames (i.e., odd frames) are compressed at the encoder with the H.264/AVC standard coder. We set a quantization parameter QP in order to have an average PSNR, for the Key frames, of about 33 dB.

The Turbo code is a Rate Compatible Turbo Punctured (RCPT) code with a puncturing period equal to 33 (15). The Wavelet transform is computed by using the well known 9/7 biorthogonal Daubechies filters, using a three level pyramid. As mentioned before, the difference between the two considered scenarios determines how the approximation \tilde{Y} is calculated at the transmitter. To this respect, we recall that motion compensation is used at the receiver only. We consider the results relative to the use of the reference scheme presented in Section 4.1 (WD WZ), the scheme with rate estimation described in Section 4.2 (WD WZ RE), and the scheme using modulo reduction of Section 4.3 (MR). We also report the results relative to a simple scheme where the WZ frames are intra-coded, but the actual reconstruction is computed as the X value that belongs to the coded quantization interval and is closest to the side-information Y at the receiver (MLJD in the figures). We consider also a scheme where the WZ schemes are intra-coded (IC), and, for scenario 2, a scheme where the frame difference between consecutive frames is intra-coded. Finally, we also report the performance

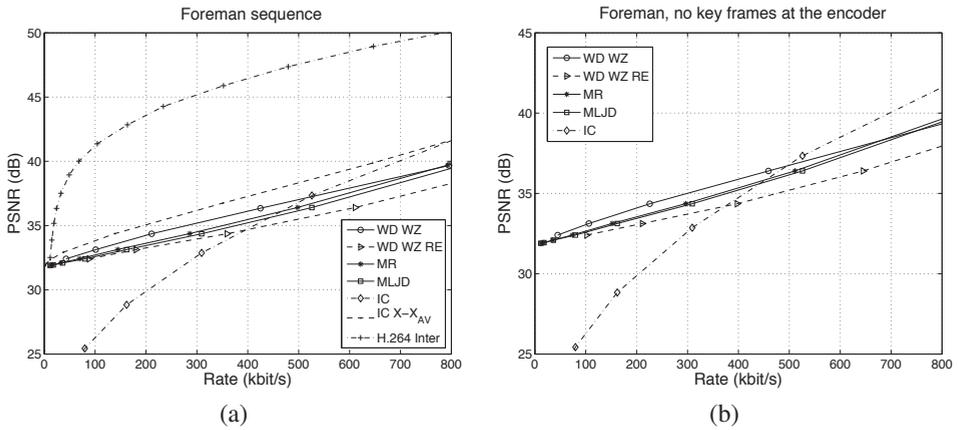


Fig. 14. (a) Rate-PSNR performance for the *Foreman* sequence (scenario 2), the Key frames are compressed using a QP = 35. (b) Rate-PSNR performance for the *Foreman* sequence (scenario 1), the Key frames are not available at the encoder and they are compressed using a QP = 35.

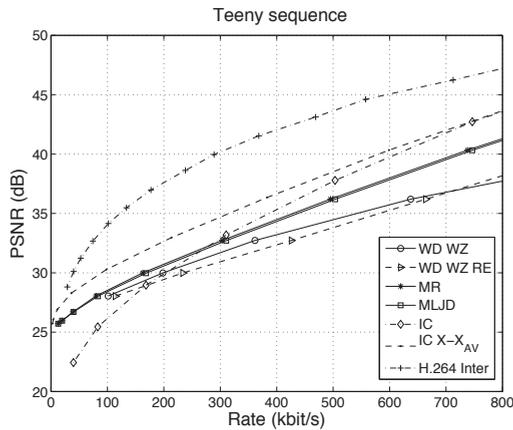


Fig. 15. Rate-PSNR performance for the *Teeny* sequence (scenario 2), the Key frames are compressed using a QP = 5.

of a standard H.264 video coder with inter-frame coding. In this case, the WZ frames are encoded as B frames (predicted from the previous and next frame with motion compensation). As we can see from the figures, for scenario 2, intra coding of the difference $X - X_{AV}$ with joint decoding performs much better than the other schemes. As mentioned, the intra coder can be implemented in this case with low complexity (37), with a clear performance advantage with respect to the DVC schemes considered in this paper and in related papers in the literature. However, note that this scheme can not be used in scenario 1. Among the other schemes, the WZ Wavelet Domain scheme with feedback from the receiver has the best performance at some bit-rates, while we notice some performance loss when the rate is estimated at the

encoder. The modulo reduction scheme has comparable or better performance, with a slight advantage (around 0.3 dB) over the MLJD scheme. In Fig. 15, one can notice that, for the high motion video sequence *Teeny*, the performance of the DVC schemes based on channel codes degrades. In all cases, the performance loss with respect to H.264 in inter-mode is significant.

5. Robust transmission of video using an auxiliary DVC stream

As another application of the Distributed Coding paradigm, we summarize in this section the results presented in (2). In particular, we consider the problem of protecting a video stream from data losses, that may be caused by transmission errors. Error protection is achieved by producing an auxiliary redundant stream encoded according to the Wyner-Ziv (WZ) video coding paradigm. This auxiliary scheme can protect a primary stream encoded with any motion-compensated predictive codec.

Along similar lines as those described in the previous sections, the proposed scheme works in the transform domain, and protects the most significant bitplanes of the Discrete Cosine Transform (DCT) coefficients. It uses LDPC codes to compute the syndrome bits of the auxiliary stream.

At the receiver side, the primary stream is decoded and motion-compensated error concealment is applied, in order to do partial recovery of the transmission errors. The concealed reconstructed frame is used as side information by the Wyner-Ziv decoder, which performs LDPC decoding based on the received syndrome bits. The prior information that can be obtained at the decoder, based on the observed error pattern, can be also used to efficiently help LDPC decoding.

One key point of the proposed procedure is that, in order to allocate the appropriate number of syndrome bits, one has to define an appropriate model relating X and Y and, in particular, one has to estimate the variance of their difference, as it was done, in a different context, in the procedure described in Section 4.2. To this purpose, a modified version of the ROPE algorithm (Recursive Optimal per-Pixel Estimate of end-to-end distortion) (38), that works in the DCT domain, is introduced. The proposed EDDD algorithm (Expected Distortion of Decoded DCT coefficients) provides an estimate of the channel induced distortion for each frame and DCT subband. This information is then used to determine the model parameters and estimate the number of syndrome bits to be produced by the Wyner-Ziv encoder.

The proposed scheme was compared with one where Forward Error Correction (FEC) codes are used. The FEC scheme adopts (N, K) Reed-Solomon channel codes. Moreover, the scheme was also compared to the use of the intra-macroblock refresh procedure, which is a non-normative tool in the standard H.264/AVC which increases the robustness to transmission errors (39). Experimental results (see Fig. 16) show that the proposed scheme has comparable or better performance, especially at high packet loss probability, than a scheme using FEC codes. One possible advantage of the proposed solution, is that it naturally allows for rate adaptivity and unequal error protection (UEP) achieved at the frame, DCT band and bitplane granularity.

In addition, the proposed scheme outperforms the intra-macroblock refresh procedure. Note that the latter requires to be applied either at encoding time, or to transcode a pre-encoded bitstream to perform mode switching. Conversely, in the proposed scheme, one can deal with a pre-encoded sequence and simply add Wyner-Ziv bits for protection, maintaining the original bitstream unaltered.

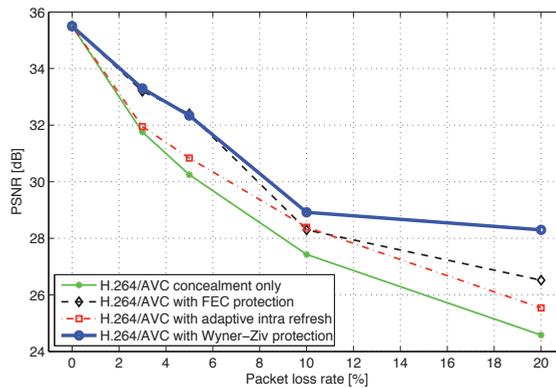


Fig. 16. PSNR vs. PLR for the *Foreman* sequence.

6. Conclusions

In this chapter, we presented the main concepts relative to Distributed Source Coding (DSC), and presented its application to video coding (DVC) and to error protection for video transmission. Although distributed coding is a well known result in Information Theory, its practical implementation, in particular for video coding, is rather recent. DVC is particularly attractive because it can simplify the video compression algorithm, which, as seen, becomes in principle a channel coding procedure. This allows to shift the complexity from the encoder to the decoder, which now has to compute the side-information, typically using a costly motion compensation procedure. Moreover, since decoding exploits a statistical, rather than deterministic, dependence between the source and the side information, it is possible that the decoding process is tolerant to errors and more robust than in a conventional decoder. This makes DVC an interesting option for emerging applications where geographically separated sources capture correlated video.

Experiments show, however, that some conventional techniques (e.g., intra coding with joint decoding and intra coding of the difference between the current frame and the one obtained by averaging the closest Key frames), which do not or partially use the distributed coding paradigm, can have comparable or better performance than the considered DVC schemes, at least for some sequences and bit-rates. In addition, an H.264 interframe coding has significantly better performance than the considered DVC schemes. However, DVC can have a role in some applications, especially when a good quality side information can be constructed at the decoder.

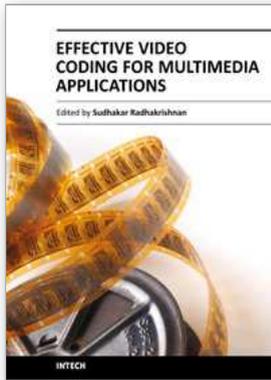
7. References

- [1] R. Bernardini, R. Rinaldo, and P. Zontone, "Wavelet domain distributed coding for video", *Proc. International Conference on Image Processing 2006*, Atlanta, USA, October 2006, pp. 245-248.
- [2] R. Bernardini, M. Fumagalli, M. Naccari, R. Rinaldo, M. Tagliasacchi, S. Tubaro and P. Zontone, "Error Concealment Using a DVC Approach for Video Streaming Applications", *Proc. Eusipco 2007*, Poznań, Poland, pp. 668-672, Sept. 2007.

- [3] Riccardo Bernardini, Roberto Rinaldo, Pamela Zontone, Andrea Vitali, "Performance Evaluation of Distributed Video Coding Schemes", *Proc. IEEE International Conference on Signal Processing 2008*, Las Vegas, Nevada, USA, pp. 709-712, Apr. 2008.
- [4] R. Bernardini, R. Rinaldo, A. Vitali, P. Zontone (2009), "Performance evaluation of wavelet-based distributed video coding schemes," *SIGNAL, IMAGE AND VIDEO PROCESSING*, ISSN: 1863-1703, doi: 10.1007/s11760-009-0141-4.
- [5] R. Bernardini, M. Naccari, R. Rinaldo, M. Tagliasacchi, S. Tubaro and P. Zontone, "Rate allocation for robust video streaming based on distributed video coding," *Signal Processing: Image Communication*, Volume 23, Issue 5, June 2008, pp. 391-403.
- [6] T. Cover and J. Thomas, "Elements of Information Theory", New York, Wiley, pp. 158-167, 1991.
- [7] D. Slepian and J. K. Wolf, "Noiseless Coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. 19, n. 4, pp. 471-480, 1973.
- [8] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, n. 1, pp. 1-10, 1976.
- [9] Zixiang Xiong, Angelos D. Liveris, and Samuel Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine*, pp. 80-94, 2004.
- [10] A. Wyner, "Recent results in the Shannon Theory," *IEEE Transactions on Information Theory*, vol. 20, n. 1, pp. 2-10, 1974.
- [11] A. D. Liveris, Z. Xiong, and C. N. Georghiades, "Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes," *Proceedings of the IEEE Data Compression Conference*, 2003.
- [12] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (DISCUS): Design and construction," *IEEE Transactions on Information Theory*, vol. 49, pp. 626-643, Mar. 2003.
- [13] R. Puri, A. Majumbar, P. Ishwar, and K. Ramchandran, "Distributed video coding in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, pp. 94-106, July 2006.
- [14] A. Aaron, R. Zhang, and B. Girod, "Wyner-ziv coding of motion video," in *Proc. Asilomar Conference on Signals and Systems*, (Pacific Grove, California), Nov. 2002.
- [15] A. Aaron, S. Rane, E. Setton, and B. Girod, "Transform-domain wyner-ziv codec for video," in *Visual Communications and Image Processing*, San Jose, CA, Jan. 2004.
- [16] A. Adikari, W. Fernando, H. K. Arachchi, and W. Weerakkody, "Wyner-ziv coding with temporal and spatial correlations for motion video," in *IEEE Electrical and Computer Engineering, Canadian Conference*, May 2006.
- [17] M. Tagliasacchi, A. Trapanese, S. Tubaro, J. Ascenso, C. Brites, and F. Pereira, "Exploitation spatial redundancy in pixel domain wyner-ziv video coding," in *IEEE International Conference on Image Processing*, Atlanta, GA, October 2006.
- [18] H. Wang and A. Ortega, "Scalable predictive coding by nested quantization with layered side information," in *Proc. of IEEE International Conference on Image Processing*, 2004.
- [19] M. Barni, D. Papini, A. Abrardo, and E. Magli, "Distributed source coding of hyperspectral images," in *Proc. of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2005.
- [20] X. Li, "Distributed coding of multispectral images: a set theoretic approach," in *Proc. of IEEE International Conference on Image Processing*, 2004.

- [21] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "The discover codec: Architecture, techniques and evaluation," in *Picture Coding Symposium (PCS)*, Lisboa, Portugal, Nov. 2007.
- [22] W. Liu, L. Dong, and W. Zeng, "Motion refinement based progressive sideinformation estimation for wynerziv video coding," *to appear in IEEE Trans. Circuits and Systems for Video Technologies*, 2010.
- [23] C-H Han, Y-I Jeon, S-W Lee and H-S Kang, "Improved Side Information Generation Using Field Coding for Wyner-Ziv Codec", *2nd International Congress on Image and Signal Processing*, Tianjin, October 2009.
- [24] C. Brites and F. Pereira, "Probability Updating for Decoder and Encoder Rate Control Turbo based Wyner-Ziv Video Coding", *Proc. of IEEE International Conf. on Image Processing*, Hong Kong, China, September 2010.
- [25] F. Dufaux, M. Ouaret and T. Ebrahimi, "Recent advances in multi-view distributed video coding", *Proc. of SPIE Mobile Multimedia/Image Processing for Military and Security applications*, 2007.
- [26] A. Aaron, S. Rane, and B. Girod, "Wyner-ziv Video Coding with Hash-based Motion Compensation at the Receiver", *IEEE International Conference on Image Processing*, Singapore, October 2004.
- [27] J. Ascenso and F. Pereira, "Adaptive hash-based side information exploitation for efficient Wyner-Ziv video coding", *IEEE International Conference on Image Processing*, Saint Antonio, USA, September 2007.
- [28] J. Ascenso and C. Brites and F. Pereira, "Improving Frame Interpolation With Spatial Motion Smoothing for Pixel Domain Distributed Video Coding", *5th EURASIP Conference on Speech and Image Processing*, July 2005.
- [29] J. Ascenso, C. Brites and F. Pereira, "Content Adaptive Wyner-Ziv Video Coding Driven by Motion Activity," *IEEE International Conference on Image Processing*, Atlanta, USA, October 2006.
- [30] A. Wang, Y. Zhao, and L. Wei, "Wavelet-Domain Distributed Video Coding with Motion-Compensated refinement", *Proc. ICIP 2006*, Atlanta USA, Oct. 2006, pp. 241-244.
- [31] X. Guo, Y. Lu, F. Wu, and W. Gao, "Distributed Video Coding using Wavelet", *Proc. ISCAS 2006*, Island of Kos, Greece, May 2006, pp. 5427-5430.
- [32] W. Liu, L. Dong and W. Zeng, "Estimating side-information for Wyner-Ziv video coding using resolution-progressive decoding and extensive motion exploration", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, April 2009, pp. 721-724.
- [33] M. Tagliasacchi, A. Trapanese, S. Tubaro, J. Ascenso, C. Brites, and F. Pereira, "Intra Mode Decision Based on Spatio-temporal Cues in Pixel Domain Wyner-Ziv Video Coding", *IEEE International Conference on Acoustic, Speech and Signal Processing*, Toulouse, May 2006.
- [34] D. Alfonso, D. Bagni, D. Moglia, "Bi-directionally motion-compensated frame-rate up-conversion for H.264/AVC decoder", *ELMAR Symposium*, June 2005, Zadar, Croatia.
- [35] D. Kubasov, K. Lajnef and C. Guillemot, "A Hybrid Encoder/Decoder Rate Control for Wyner-Ziv Video Coding with a Feedback Channel, *Proc. of MMSP, IEEE International Workshop on Multimedia Signal Processing*, Chania, Crete, Greece, October, 2007, pp. 251-254.

- [36] C. Guillemot, F. Pereira, L. Torres, T. Ebrahimi, R. Leonardi, and J. Ostermann, "Distributed Monoview and Multiview Video Coding", *Signal Processing Magazine, IEEE*, Sept. 2007, Vol. 24, Issue 5, pp. 67-76.
- [37] J. Oliver, M.P. Malumbres, "Low-Complexity multiresolution image compression using wavelet lower trees", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16, Nov. 2006, pp. 1437-1444.
- [38] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience", *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 966-976, June 2000.
- [39] ITU-T, *Information Technology - Coding of audio-visual objects - Part 10: advanced video coding*, 2003, Final Draft International Standard, ISO-IEC FDIS 14 496-10.



Effective Video Coding for Multimedia Applications

Edited by Dr Sudhakar Radhakrishnan

ISBN 978-953-307-177-0

Hard cover, 292 pages

Publisher InTech

Published online 26, April, 2011

Published in print edition April, 2011

Information has become one of the most valuable assets in the modern era. Within the last 5-10 years, the demand for multimedia applications has increased enormously. Like many other recent developments, the materialization of image and video encoding is due to the contribution from major areas like good network access, good amount of fast processors e.t.c. Many standardization procedures were carried out for the development of image and video coding. The advancement of computer storage technology continues at a rapid pace as a means of reducing storage requirements of an image and video as most situation warrants. Thus, the science of digital video compression/coding has emerged. This storage capacity seems to be more impressive when it is realized that the intent is to deliver very high quality video to the end user with as few visible artifacts as possible. Current methods of video compression such as Moving Pictures Experts Group (MPEG) standard provide good performance in terms of retaining video quality while reducing the storage requirements. Many books are available for video coding fundamentals. This book is the research outcome of various Researchers and Professors who have contributed a might in this field. This book suits researchers doing their research in the area of video coding. The understanding of fundamentals of video coding is essential for the reader before reading this book. The book revolves around three different challenges namely (i) Coding strategies (coding efficiency and computational complexity), (ii) Video compression and (iii) Error resilience. The complete efficient video system depends upon source coding, proper inter and intra frame coding, emerging newer transform, quantization techniques and proper error concealment. The book gives the solution of all the challenges and is available in different sections.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Riccardo Bernardini, Roberto Rinaldo and Pamela Zontone (2011). Distributed Video Coding: Principles and Evaluation of Wavelet-Based Schemes, Effective Video Coding for Multimedia Applications, Dr Sudhakar Radhakrishnan (Ed.), ISBN: 978-953-307-177-0, InTech, Available from:
<http://www.intechopen.com/books/effective-video-coding-for-multimedia-applications/distributed-video-coding-principles-and-evaluation-of-wavelet-based-schemes>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai

Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.