Review Article

Distributed Video Coding: Trends and Perspectives

Frederic Dufaux,[1] Wen Gao,[2] Stefano Tubaro,[3] and Anthony Vetro[4]

[1] *Multimedia Signal Processing Group, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland*
[2] *School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China*
[3] *Dipartimento di Elettronica e Informazione, Politecnico di Milano, 20133 Milano, Italy*
[4] *Mitsubishi Electric Research Laboratories, Cambridge, MA 02139, USA*

Correspondence should be addressed to Frederic Dufaux, frederic.dufaux@epfl.ch

This paper surveys recent trends and perspectives in distributed video coding. More specifically, the status and potential benefits of distributed video coding in terms of coding efficiency, complexity, error resilience, and scalability are reviewed. Multiview video and applications beyond coding are also considered. In addition, recent contributions in these areas, more thoroughly explored in the papers of the present special issue, are also described.

1. Introduction

Tremendous advances in computer and communication technologies have led to a proliferation of digital media content and the successful deployment of new products and services. However, digital video is still demanding in terms of processing power and bandwidth. Therefore, this digital revolution has only been possible thanks to the rapid and remarkable progress in video coding technologies. Additionally, standardization efforts in MPEG and ITU-T have played a key role in order to ensure the interoperability and durability of video systems as well as to achieve economy of scale.

For the last two decades, most developments have been based on the two principles of predictive and transform coding. The resulting motion-compensated block-based Discrete Cosine Transform (DCT) hybrid design has been adopted by all MPEG and ITU-T video coding standards to this day. This pathway has culminated with the state-of-the-art H.264/Advanced Video Coding (AVC) standard [1]. H.264/AVC relies on an extensive analysis at the encoder in order to better represent the video signal and thus to achieve a more efficient coding. Among many innovations, it features a 4×4 transform which allows a better representation of the video signals thanks to localized adaptation. It also supports spatial intraprediction on top of inter prediction. Enhanced inter prediction features include the use of multiple reference frames, variable block-size motion compensation, and quarter-pixel precision.

The above design, which implies complex encoders and lightweight decoders, is well suited for broadcasting-like applications, where a single sender is transmitting data to many receivers. In contrast to this downstream model, a growing number of emerging applications, such as low-power sensor networks, wireless video surveillance cameras, and mobile communication devices, are rather relying on an upstream model. In this case, many clients, often mobile, low-power, and with limited computing resources, are transmitting data to a central server. In the context of this upstream model, it is usually advantageous to have lightweight encoding with high compression efficiency and resilience to transmission errors. Thanks to the improved performance and reducing cost of cameras, another trend is towards multiview systems where a dense network of cameras captures many correlated views of the same scene.

More recently, a new coding paradigm, referred to as Distributed Source Coding (DSC), has emerged based on two Information Theory theorems from the seventies: Slepian-Wolf (SW) [2] and Wyner-Ziv (WZ) [3]. Basically, the SW theorem states that for lossless coding of two or

more correlated sources, the optimal rate achieved when performing joint encoding and decoding (i.e., conventional predictive coding) can theoretically be reached by doing separate encoding and joint decoding (i.e., distributed coding). The WZ theorem shows that this result still holds for lossy coding under the assumptions that the sources are jointly Gaussian and a Mean Square Error (MSE) distortion measure is used. Distributed Video Coding (DVC) applies this paradigm to video coding. In particular, DVC relies on a new statistical framework, instead of the deterministic approach of conventional coding techniques such as MPEG and ITU-T schemes. By exploiting this result, the first practical DVC schemes have been proposed in [4, 5]. Following these seminal works, DVC has raised a lot of interests in the last few years, as evidenced by the very large amount of publications on this topic in major conferences and journals. Recent overviews are presented in [6, 7].

DVC offers a number of potential advantages which make it well suited for the aforementioned emerging upstream applications. First, it allows for a flexible partitioning of the complexity between the encoder and decoder. Furthermore, due to its intrinsic joint source-channel coding framework, DVC is robust to channel errors. Because it does not rely on a prediction loop, DVC provides codec independent scalability. Finally, DVC is well suited for multiview coding by exploiting correlation between views without requiring communications between the cameras, which may be an important architectural advantage. However, in this case, an important issue is how to generate the joint statistical model describing the multiple views.

In this paper, we offer a survey of recent trends and perspectives in distributed video coding. More specifically, we address some open issues such as coding efficiency, complexity, error resilience, scalability, multiview coding, and applications beyond coding. In addition, we also introduce recent contributions in these areas provided by the papers of this special issue.

2. Background

The foundations of DVC are traced back to the seventies. The SW theorem [2] establishes some lower bounds on the achievable rates for the lossless coding of two or more correlated sources. More specifically, let us consider two statistically dependent random signals X and Y. In conventional coding, the two signals are jointly encoded and it is well known that the lower bound for the rate is given by the joint entropy $H(X, Y)$. Conversely, with distributed coding, these two signals are independently encoded but jointly decoded. In this case, the SW theorem proves that the minimum rate is still $H(X, Y)$ with a residual error probability which tends towards 0 for long sequences. Figure 1 illustrates the achievable rate region. In other words, SW coding allows the same coding efficiency to be asymptotically attained. However, in practice, finite block lengths have to be used. In this case, SW coding entails a coding efficiency loss compared to lossless source coding, and the loss can be sizeable depending on the block length and the source statistics [8].

$$R_x \geq H(X|Y)$$
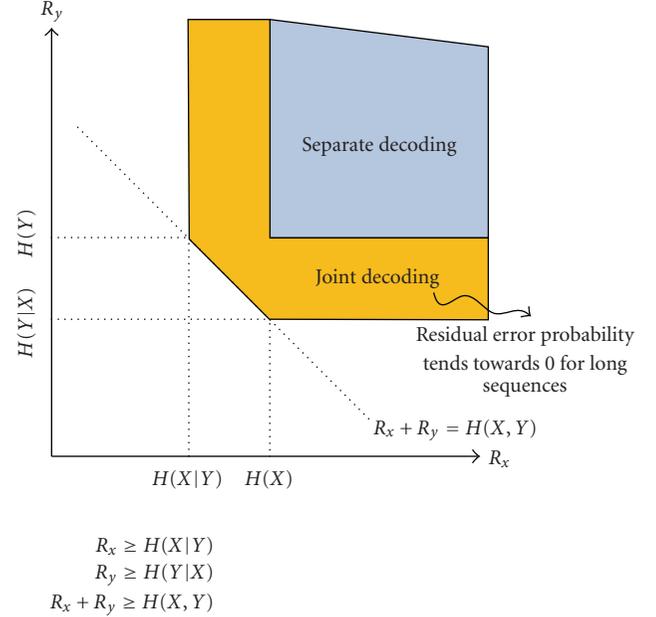$$R_y \geq H(Y|X)$$
$$R_x + R_y \geq H(X, Y)$$

FIGURE 1: Achievable rates by distributed coding of two statistically dependent random signals.

Subsequently, Wyner and Ziv (WZ) extended the Slepian-Wolf theorem by characterizing the achievable rate-distortion region for lossy coding with Side Information (SI). More specifically, WZ showed that there is no rate loss with respect to joint encoding and decoding of the two sources, under the assumptions that the sources are jointly Gaussian and an MSE distortion measure is used [3]. This result has been shown to remain valid as long as the innovation between X and Y is Gaussian [9].

2.1. PRISM Architecture. PRISM (Power-efficient, Robust, hIgh compression Syndrome-based Multimedia coding) is one of the early practical implementations of DVC [4, 10]. This architecture is shown in Figure 2. For a more detailed description of PRISM, the reader is referred to [10]. More specifically, each frame is split into 8×8 blocks which are DCT transformed. Concurrently, a zero-motion block difference is used to estimate their temporal correlation level. This information is used to classify blocks into 16 encoding classes. One class corresponds to blocks with very low correlation which are encoded using conventional Intra-coding. Another class is made of blocks which have very high correlation and are merely signaled as skipped. Finally, the remaining blocks are encoded based on distributed coding principles. More precisely, syndrome bits are computed from the least significant bits of the transform coefficients, where the number of least significant bits depends on the estimated correlation level. The lower part of the least significant bit planes is entropy coded with a (*run, depth, path, last*) 4-tuple alphabet. The upper part of the least significant bit planes is coded using a coset channel code. For this purpose, a BCH code is used, as it performs well even with small block-lengths. Conversely, the most significant bits are
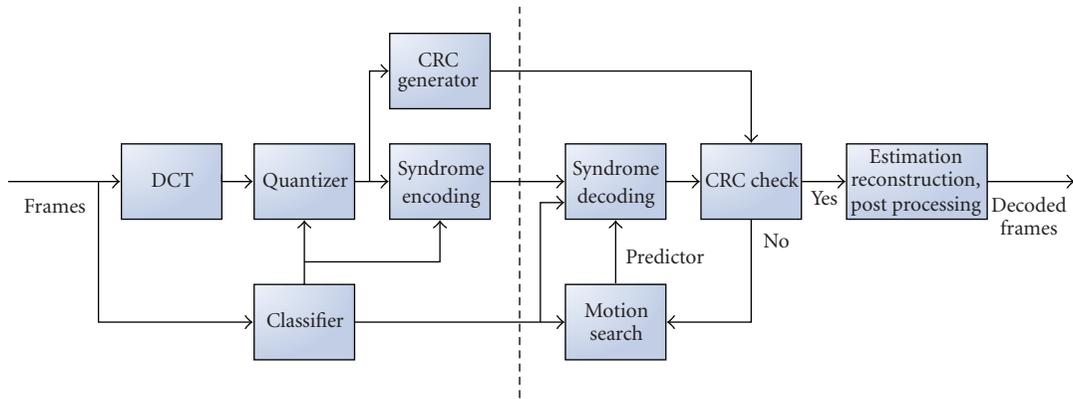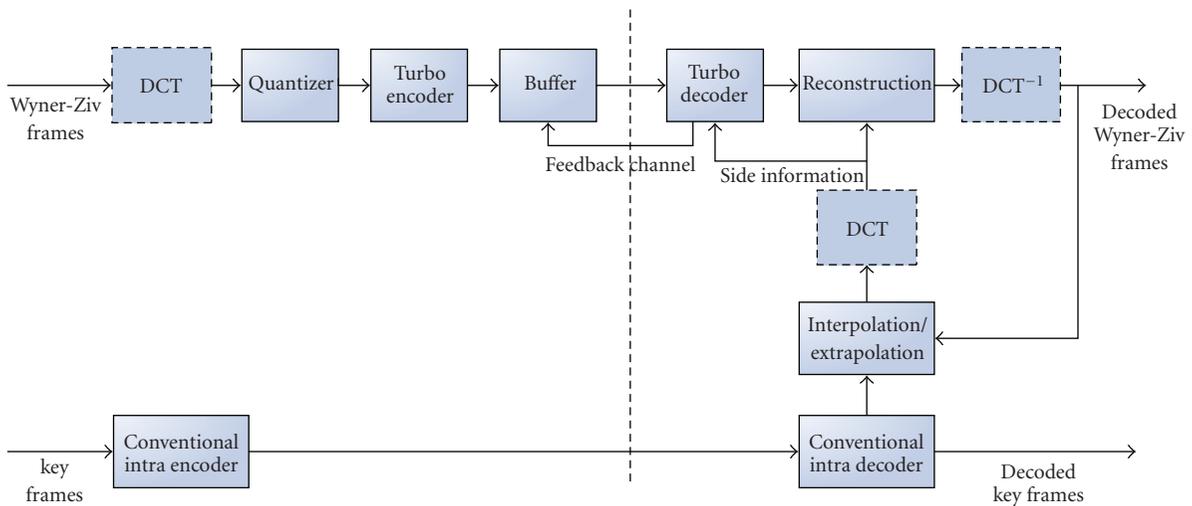
Figure 2: PRISM architecture.

Figure 3: Stanford pixel-domain and transform-domain DVC architecture.

assumed to be inferred from the block predictor or SI. In parallel, a 16-bit Cyclic Redundancy Check (CRC) is also computed. At the decoder, the syndrome bits are then used to correct predictors, which are generated using different motion vectors. The CRC is used to confirm whether the decoding is successful.

2.2. Stanford Architecture. Proposed at the same time as PRISM, another early DVC architecture has been introduced in [5, 11]. A block diagram of this architecture is illustrated in Figure 3, whereas a more detailed description is given in [11]. The video sequence is first divided into Group Of Pictures (GOPs). The first frame of each GOP, also referred to as key frame, is encoded using a conventional intraframe coding technique such as H.264/AVC in intraframe mode [1]. The remaining frames in a GOP are encoded using distributed coding principles and are referred to as WZ frames. In a pixel-domain WZ version, the WZ frames first undergo quantization. Alternatively, in a transform-domain version [12], a DCT transform is applied prior to quantization. The quantized values are then split into bitplanes which go

through a Turbo encoder. At the decoder, SI approximating the WZ frames is generated by motion-compensated interpolation or extrapolation of previously decoded frames. The SI is used in the turbo decoder, along with the parity bits of the WZ frames requested via a feedback channel, in order to reconstruct the bitplanes, and subsequently the decoded video sequence. In [13], rate-compatible Low-Density Parity-Check Accumulate (LDPCA) codes, which better approach the communication channels capacity, replace the Turbo codes.

2.3. Comparison. The two above architectures differ in a number of fundamental ways, as we will discuss hereafter. A more comprehensive analysis is also given in [14].

The block-based nature of PRISM allows for a better local adaptation of the coding mode in order to cope with the nonstationary statistics typical of video data. By performing simple interframe prediction for block classification based on correlation at the encoder, the WZ coding mode is only used when appropriate, namely, when the correlation is sufficient. However, this block partitioning implies a short block-length

which is a limiting factor for efficient channel coding. For this reason, a BCH code is used in PRISM. In contrast, in the frame-based Stanford approach, a frame is WZ encoded in its whole. Nevertheless, this enables the successful usage of more sophisticated channel codes, such as Turbo or LDPC codes.

The way motion estimation is performed constitutes another important fundamental distinction. In the Stanford architecture, motion estimation is performed prior to WZ decoding, using only information directly available at the decoder. Conversely, in PRISM, motion vectors are estimated during the WZ decoding process. In addition, this process is helped by the transmitted CRC check. Hence, it leads to better performance and robustness to transmission errors.

In the Stanford approach, rate control is performed at the decoder side and a feedback channel is needed. Hence, the SW rate can be better matched to the realization of the source and SI. However, the technique is limited to real-time scenarios without too stringent delay constraints. As in PRISM rate control is carried out at the encoder, the latter does not have this restriction. However, in this codec, the SW rate has to be determined based on a priori classification at the encoder, which may result in decreased performance.

Note that some of these shortcomings have been addressed in subsequent research works. For instance, the Stanford architecture has been augmented with hash codes transmitted to enhance motion compensation in [15], a block-based Intracoding mode in [16], and an encoder-driven rate control in order to eliminate the feedback channel in [17].

2.4. State-of-the-Art Performance. The codec developed by the European project DISCOVER, presented in [18], is one of the best performing DVC schemes reported in the literature to date. A thorough performance benchmark of this codec is publicly available in [19]. The DISCOVER codec is based on the Stanford architecture [5, 11] and brings several improvements. It uses the same 4×4 DCT-like transform as in H.264/AVC. Notably, SI is obtained by motion compensated interpolation with motion vectors smoothing resulting in enhanced performance. Moreover, the issue of online parameter estimation is tackled, including rate estimation, virtual channel model and soft input calculation, and decoder success/failure.

In [19], the coding efficiency of the DISCOVER DVC scheme is compared to two variants of H.264/AVC with low encoding complexity: H.264/AVC Intra (i.e., all the frames are Intra coded) and H.264/AVC No Motion (i.e., interframe coding with zero motion vectors). It can be observed that DVC consistently matches or outperforms H.264/AVC Intra, except for scenes with complex motion (e.g., the test sequence "Soccer"). For scenes with low motion (e.g., the test sequence "Hall Monitor"), the gain can reach up to 3 dB.

More recently, the performance of the DVC codec developed by the European project VISNET II has been thoroughly assessed [20]. This codec is also based on the Stanford architecture [5, 11]. It makes use of some of the same techniques as in the DISCOVER codec and includes a number of enhancements including better SI generation, an iterative reconstruction process, and a deblocking filter. In [20], it is shown that the VISNET II DVC codec consistently outperforms the DISCOVER scheme. For low-motion scenes, gains up to 5 dB are reported over H.264/AVC Intra. On the other hand, when compared to H.264/AVC No Motion, the performance of the VISNET II DVC codec typically remains significantly lower. However, DVC shows strong performance for scenes with simple and regular global motion (e.g., "Coastguard"), where it outperforms H.264/AVC No Motion.

In terms of complexity, [19] shows that the DVC encoding complexity, expressed in terms of software execution time, is significantly lower than for H.264/AVC Intra and H.264/AVC No Motion.

3. Current Topics of Interest

The DVC paradigm offers a number of major differentiations when compared to conventional coding. First, it is based on a statistical framework. As it does not rely on joint encoding, the content analysis can be performed at the decoder side. In particular, DVC does not need a temporal prediction loop characteristic of past MPEG and ITU-T schemes. As a consequence, the computational complexity can be flexibly distributed between the encoder and the decoder, and in particular, it allows encoding with very low complexity. According to information theory, this can be achieved without loss of coding performance compared to conventional coding, in an asymptotical sense and for long sequences. However, coding efficiency remains a challenging issue for DVC despite considerable improvements over the last few years.

Most of the literature on distributed video coding has addressed the problem of light encoding complexity, by shifting the computationally intensive task of motion estimation from the encoder to the decoder. Given its properties, DVC also offers other advantages and functionalities. The absence of the prediction loop prevents drifts in the presence of transmission errors. Along with the built-in joint source-channel coding structure, it implies that DVC has improved error resilience. Moreover, given the absence of the prediction loop, DVC is also enabling codec independent scalability. Namely, a DVC enhancement layer can be used to augment a base layer which becomes the SI. DVC is also well suited for camera sensor networks, where the correlation across multiple views can be exploited at the decoder, without communications between the cameras. Finally, the DSC principles have been useful beyond coding applications. For instance, DSC can be used for data authentication, tampering localization, and secure biometrics.

In the following sections, we address each of these topics and review some recent results as well as the contributions of the papers in this special issue.

3.1. Coding Efficiency. To be competitive with conventional schemes in terms of coding efficiency has proved very

challenging. Therefore, significant efforts have focused on further improving the compression performance in DVC. As reported in Section 2.4, the best DVC codecs now consistently outperform H.264/AVC Intracoding, except for scenes with complex motion. In some cases, for example, video sequences with simple motion structure, DVC can even top H.264/AVC No Motion. Nevertheless, the performance remains generally significantly lower than a full-fledge H.264/AVC codec.

Very different tools and approaches have been proposed over the years to increase the performance of DVC.

The compression efficiency of DVC depends strongly on the correlation between the SI and the actual WZ frame. The SI is commonly generated by linear interpolation of the motion field between successive previously decoded frames. While the linear motion assumption holds for sequences with simple motion, the coding performance drops for more complex sequences. In [21, 22], spatial smoothing and refinement of the motion vectors is carried out. By removing some discontinuities and outliers in the motion field, it leads to better prediction. In the same way, in [23], two SIs are generated by extrapolation of the previous and next key frames, respectively, using forward and backward motion vectors. Then, the decoding process makes use of both SI concurrently. Subpixel accuracy, similar to the method in H.264/AVC, is proposed in [24] in order to further improve motion estimation for SI generation.

Another approach to improve coding efficiency is to rely on iterative SI generation and decoding. In [25], motion vectors are refined based on bitplane decoding of the reconstructed WZ frame as well as previously decoded key frames. It also allows for different interpolation modes. However, only minor performance improvements are reported. The approach in [26] shares some similarities. A partially decoded WZ frame is first reconstructed. The latter is then exploited for iteratively enhancing motion-compensated temporal interpolation and SI generation. An iterative method by way of multiple SI with motion refinement is introduced in [27]. The turbo decoder selects for each block which SI stream to use, based on the error probability. Finally, exploiting both spatial and temporal correlations in the sequence, a partially decoded WZ frame is exploited to improve the performance of the whole SI generation in [28]. In addition, an enhanced motion compensated temporal frame interpolation is proposed.

A different alternative is for the encoder to transmit auxiliary information about the WZ frames in order to assist the SI generation in the decoder. For instance, CRCs are transmitted in [4, 10], whereas hash codes are used in [15, 29]. At the decoder, multiple predictors are used, and the CRC or hash is exploited to verify successful decoding. In [30], 3D model-based frame interpolation is used for SI. For this purpose, feature points are extracted from the WZ frames at the encoder and transmitted as supplemental information. The decoder makes use of these feature points to correct misalignments in the 3D model. By taking into account geometric constraints, this method leads to an improved SI, especially for static scenes with moving camera.

Another important factor impacting the performance of DVC is the estimation of the correlation model between SI and WZ frames. In some earlier DVC schemes [5], a Laplacian model is computed offline, under the unrealistic assumption that original frames are available at the decoder. In [31], a method is proposed for online estimation at the decoder of the correlation model. Another technique, proposed in [32], consists in computing the parameters of the correlation model at the encoder by approximating the SI.

For the blocks of the frame where the SI fails to provide a good predictor, in other words for the regions where the correlation between SI and WZ frame is low, it is advantageous to encode them in Intramode. In [16], a block-based coding mode selection is introduced based on the estimation of SI at the encoder side. Namely, blocks with weak correlation estimation are Intracoded. This method shares some similarities with the mode selection previously described for PRISM [4, 10].

The reconstruction module also plays an important role in determining the quality of the decoded video. In the Stanford architecture [5, 11], the reconstructed pixel is simply calculated from the corresponding side information and boundaries of the quantization interval. Another approach is proposed in [33], which takes advantage of the average statistical distribution of transform coefficients. In [34], the reconstructed value is instead computed as the expectation of the source coefficient given the quantization interval and the side information value, showing improved performance. A novel algorithm is introduced in [35], which exploits the statistical noise distribution of the DVC-decoded output.

Note that closing the performance gap with conventional coding is not simply a question of finding new and improved DVC techniques. Indeed, as stated in Section 2, some theoretical hurdles exist. First, the Slepian-Wolf theorem states that SW coding can achieve the same coding performance asymptotically. In practice, using finite block lengths results in a performance loss which can be sizeable [8]. Then, the Wyner-Ziv theorem holds for Gaussian sources, although video data statistics is known to be non-Gaussian.

The performance of decoder side motion interpolation is also theoretically analyzed in [36, 37]. In [36], it is shown that the accuracy of the interpolation depends strongly on the temporal coherence of the motion field as well as the distance between successive key frames. A model, based on a state-space model and Kalman filtering, demonstrates that DVC with motion interpolation at the decoder cannot reach the performance of conventional predictive coding. A method to optimize the GOP size is also proposed. In [37], a model is proposed to study the performance of DVC. It is theoretically shown that conventional motion-compensated predictive interframe coding outperforms DVC by 6 dB or more. Subpixel and multireference motion search methods are also examined.

In this special issue, three contributions address different means to improve coding efficiency. In [38], Wu et al. address the shortcoming of the common motion-compensated temporal interpolation which assumes that

the motion remains translational and constant between key frames. In this paper, a spatial-aided Wyner-Ziv video coding is proposed. More specifically, auxiliary information is encoded with DPCM at the encoder and transmitted along with WZ bitstream. At the decoder, SI is generated by spatial-aided motion-compensated extrapolation exploiting this auxiliary information. It is shown that the proposed scheme achieves better rate distortion performance than conventional motion-compensated extrapolation-based WZ coding without auxiliary information. It is also demonstrated that the scheme efficiently improves WZ coding performance for low-delay applications.

Sofke et al. [39] consider the problem that current WZ coding schemes do not allow controlling the target quality in an efficient way. Indeed, this may represent a major limitation for some applications. An efficient quality control algorithm is introduced in order to maintain uniform quality through time. It is achieved by dynamically adapting the quantization parameters depending on the desired target quality without any a priori knowledge about the sequence characteristics.

Finally, the contribution [40] by Ye et al. proposes a new SI generation and iterative reconstruction scheme. An initial SI is first estimated using common motion-compensated interpolation, and a partially decoded WZ frame is obtained. Next, the latter is used to generate an improved SI, featuring motion vector refinement and smoothing, a new matching criterion, and several compensation modes. Finally, the reconstruction step is carried out again to get the decoded WZ frame. The same idea is also applied to a new hybrid spatial and temporal error concealment scheme for WZ frames. It is shown that the proposed scheme outperforms a state-of-the-art DVC codec.

3.2. Complexity.

Among the claimed benefits of DVC, low-complexity encoding is often the most widely cited advantage. Relative to conventional coding schemes that employ motion estimation at the encoder, DVC provides a framework that eliminates this high computational burden altogether as well as the corresponding memory to store reference frames. Encoding complexity was evaluated in [19, 41]. Not surprisingly, it showed that DVC encoding complexity (DISCOVER codec based on the Stanford architecture) was indeed providing a substantial speed-up when compared to conventional H.264/AVC Intra and H.264/AVC No Motion in terms of software execution time.

Not only does the DVC decoder need to generate side information, which is often done using computationally intense motion estimation techniques, but it also incurs the complexity of a typical channel decoding process. When the quality of the side information is very good, the time for channel decoding could be lower. But in general, several iterations are required to converge to a solution. In [19, 41], it is shown that the DVC decoder is several orders of magnitude more complex in term of software execution time compared to that of a conventional H.264/AVC Intraframe decoder and about 10–20 times more complex than an H.264/AVC Intraframe encoder.

Clearly, this issue has to be addressed for DVC to be used in any practical setting. In [42], a hybrid encoder-decoder rate control is proposed with the goal to reduce decoding complexity while having a negligible impact on encoding complexity and coding performance. Decoding execution time reductions of up to 70% are reported.

While the signal processing community had devoted little research effort to reduce the decoder complexity of DVC, there is substantial work on fast and parallel implementations of various channel decoding algorithms, including turbo decoding and belief propagation (BP). For instance, it has been shown that parallelization of the message-passing algorithm used in belief propagation can result in speed-ups of approximately 13.5 on a multicore processor relative to single processor implementations [43]. There also exists decoding methods that use information from earlier-decoded nodes to update the latter-decoded nodes in the *same* iteration, for example, Shuffled BP [44, 45]. It should also be possible to reduce complexity of the decoding process by changing the complexity of operations at the variable nodes, for example, replacing complex trigonometric functions by simple majority voting. These and other innovations should help to alleviate some of the complexity issues for DVC decoding. Certainly, more research is needed to achieve desirable performance. Optimized decoder implementations on multicore processors and FPGAs should specifically be considered.

3.3. Robust Transmission.

Distributed video coding principles have been extensively applied in the field of robust video transmission over unreliable channels. One of the earliest examples is given by the PRISM coding framework [4, 10, 46], which simultaneously achieves light encoding complexity and robustness to channel losses. In PRISM, each block is encoded without the deterministic knowledge of its motion-compensated predictor, which is made available at the decoder side only. If the predictor obtained at the decoder is within the noise margin for the number of encoded cosets, the block is successfully decoded. The underlying idea is that, by adjusting the number of cosets based on the expected correlation channel, decoding is successfully achieved even if the motion compensated predictor is noisy, for example, due to packet losses affecting the reference frame.

These results were extended to a fully scalable video coding scheme in [47, 48], which is shown to be robust to losses that affect both the enhancement and the base layers. This is due to the fact that the correlation channel that characterizes the dependency between different scalability layers is captured at the encoder in a statistical, rather than deterministic, way.

Despite PRISM, most of the distributed video coding schemes that focus on error resilience try to increase the robustness of standard encoded video by adding redundant information encoded according to distributed video coding principles. One of the first works along this direction is presented in [49], where auxiliary data is encoded only for some frames, denoted as "peg" frames, in order to stop drift propagation at the decoder. The idea is to achieve the

robustness of intrarefresh frames, without the rate overhead due to intraframe coding.

In [50], a layered WZ video coding framework similar to Fine Granularity Scalability (FGS) coding is proposed, in the sense that it considers the standard coded video as the base layer and generates an embedded bitstream as the enhancement layer. However, the key difference with respect to FGS is that, instead of coding the difference between the original video and the base layer reconstruction, the enhancement layer is "blindly" generated, without knowing the base layer. Although the encoder does not know the exact realization of the reconstructed frame, it can try to characterize the effect of channel errors (i.e., packet losses) in statistical terms, in order to perform optimal bit allocation. This idea has been pursued, for example, in [51] where a PRISM-like auxiliary stream is encoded for Forward Error Protection (FEP), and rate-allocation is performed at the encoder by exploiting the information provided by the Recursive Optimal Per-pixel Estimate (ROPE) algorithm.

Distributed video coding has been applied to error resilient MPEG-2 video broadcasting in [52], where a systematic lossy source channel coding framework is proposed, referred to as Systematic Lossy Error Protection (SLEP). An MPEG-2 video bitstream is transmitted over an error-prone channel without error protection. In addition, a supplementary bitstream is generated using distributed video coding tools, which consists of a coarsely quantized video bitstream obtained using a conventional hybrid video coder, applying Reed–Solomon codes, and transmitting only the parity symbols. In the event of channel errors, the decoder decodes these parity symbols using the error-prone conventionally decoded MPEG-2 video sequence as side information. The SLEP scheme has also been extended to the H.264/AVC video coding standard [53]. Based on the SLEP framework, the scheme proposed in [53] performs Unequal Error Protection (UEP) assigning different amounts of parity bits between motion information and transform coefficients. This approach shares some similarities with the one presented in [54] where a more sophisticated rate allocation algorithm, based on the estimated induced channel distortion, is proposed.

To date, the robustness to transmission errors has proved to be one of the most promising directions for DVC in order to bring this technology to a viable and competitive level in the market place.

In this special issue, two papers propose the use of DVC for robust video transmission. In particular, the contribution by Tonoli et al. [55] evaluates and compares the error resilience performance of two distributed video coding architectures: the DISCOVER codec [18] which is based on the Stanford architecture [5, 11], and a codec based on the PRISM architecture [4, 10]. In particular, a rate-distortion analysis of the impact of transmission errors has been carried out. Moreover, a performance comparison with H.264/AVC, both without error protection and with a simple FEP, is also reported. It is shown that the codecs behavior strongly depends on the content. More specifically, PRISM performs better on low-motion sequences, whereas DISCOVER is more efficient otherwise.

In [56] Liang et al. propose three schemes based on Wyner-Ziv coding for unequal error protection. They apply different levels of protection to motion information and transform coefficients in an H.264/AVC stream, and they are shown to provide with better error resilience in the presence of packet loss when compared to equal error protection.

3.4. Scalability. With the emergence of heterogeneous multimedia networks and the variety of client terminals, scalable coding is becoming an attractive feature. With a scalable representation, the video content is encoded once but can be decoded at different spatial and temporal resolutions or quality levels, depending on the network conditions and the capabilities of the terminal. Due to the absence of a closed-loop in its design, DVC supports codec-independent scalability. Namely, WZ enhancement layers can be built upon conventional or DVC base layers which are used as SI.

In [47], a scalable version of PRISM [4, 10] is presented. Namely, an H.264/AVC base layer is augmented with a PRISM enhancement layer, leading to a spatiotemporal scalable video codec. It is shown that the scalable version of PRISM outperforms the nonscalable one as well as H.263+ Intra. However, the performance remains lower when compared to motion compensated H.263+.

In [57], the problem of scalable predictive video coding is posed as a variant of the WZ side information problem. This approach relaxes the conventional constraint that both the encoder and decoder employ the very same prediction loops, hence enabling a more flexible prediction across layers and preventing the occurrence of prediction drift. It is shown that the proposed scheme outperforms a simple scalable codec based on conventional coding.

A framework for efficient and low-complexity scalable coding based on distributed video coding is introduced in [32]. Using an MPEG-4 base layer, a multilayer WZ prediction is introduced which results in improved temporal prediction compared to MPEG-4 FGS [58]. Significant coding gain is achieved over MPEG-4 FGS for sequences with high temporal correlation.

Finally, [59] proposes DVC-based scalable video coding schemes supporting temporal, spatial, and quality scalability. Temporal scalability is realized by using a hierarchical motion-compensated interpolation and SI generation. Conversely, a combination of spatial down- and upsampling filters along with WZ coding is used for spatial scalability. The codec independence is illustrated by using both H.264/AVC Intra and JPEG 2000 [60] base layers, with the same enhancement WZ layer.

While the variety of scalability offered by DVC is intriguing, a strong case remains to be made where its specificities play a critical role in enabling new applications.

In this special issue, two contributions address the use of DVC for scalable coding. In the first one [61] by Macchiavello et al. the rate-distortion performance of different SI estimators is compared for temporal and spatial scalable WZ coding schemes. In the case of temporal scalability, a new algorithm is proposed to generate SI using a linear motion model. For spatial scalability, a superresolution method is introduced for

upsampling. The performance of the scalable WZ codec is assessed using H.264/AVC as reference.

In the second contribution [62] Devaux and De Vleeschouwer propose a highly scalable video coding scheme based on WZ, supporting fine-grained scalability in terms of resolution, quality, and spatial access as well as temporal access to individual frames. JPEG 2000 is used to encode Intrainformation, whereas blocks changing between frames are refreshed using WZ coding. Due to the fact that parity bits aim at correcting stochastic errors, the proposed approach is able to handle a loss of synchronization between the encoder and decoder. This property is important for content adaptation due to fluctuating network conditions.

3.5. Multiview. With its ability to exploit intercamera correlation at the decoder side, without communication between cameras, DVC is also well suited for multiview video coding where it could offer a noteworthy architectural advantage. Moreover, multiview coding is gathering a lot of interests lately, as it is attractive for a number of applications such as stereoscopic video, free viewpoint television, multiview 3D television, or camera networks for surveillance and monitoring.

When compared to monoview, the main difference in multiview DVC is that the SI can be computed not only from previously decoded frames in the same view but also from frames in other views. Another important matter concerns the generation of the joint statistical model describing the multiple views.

Disparity Compensation View Prediction (DCVP) [63] is a straightforward extension of motion compensated temporal interpolation, where the prediction is carried out by motion compensation of the frames in other views using disparity vectors. Multiview Motion Estimation (MVME) [64] estimates motion vectors in the side views and then applies them to the view to be WZ encoded. For this purpose, disparity vectors between views have also to be estimated. A homography model, estimated by global motion estimation, is rather used in [65] for interview prediction, showing significant improvement in the SI quality. Another approach is View Synthesis Prediction (VSP) [66]. Pixels from one view are projected to the 3D world coordinates using intrinsic and extrinsic camera parameters and then are used to predict another view. The drawback of this approach is that it requires depth information and the quality of the prediction depends on the accuracy of the camera calibration as well as the depth estimation. Finally, View Morphing (VM) [67], which is commonly used to create a synthesized image for a virtual camera positioned between two real cameras using principles of projective geometry, can also be applied to estimate SI from side views.

When the SI can be generated either from the view to be WZ encoded, using motion compensated temporal interpolation, or from side views, using one of the method previously described, the next issue is how to combine these different predictions. For fusion at the decoder side, the challenge lies in the difficulty of determining the best predictor. In [68], a technique is proposed to fuse intraview

temporal and interview homography side information. It exploits the previous and next key frames to choose the best predictor on a pixel basis. It is shown that the proposed approach outperforms monoview DVC for video sequences containing significant motion. Two fusion techniques are introduced in [69]. They rely on a binary mask to estimate the reliability of each prediction. The latter is computed on the side views and projected on the view to be WZ encoded. However, depth information is required for intercamera disparity estimation. The technique in [70] combines a discrete wavelet transform and turbo codes. Fusion is performed between intraview temporal and interview homography side information, based on the amplitude of motion vectors. It is shown that this fusion technique surpasses interview temporal side information. Moreover, the resulting multiview DVC scheme significantly outperforms H.263+ Intracoding. The method in [71] follows a similar approach but relies on the H.264/AVC mode decision applied on blocks in the side views. Experimental results confirm that this method achieves notably better performance than H.263+ Intracoding and is close to Intercoding efficiency for sequences with complex motion. Taking a different approach, in [63] a binary mask is computed at the encoder and then transmitted to the decoder in order to help the fusion process. Results show that the approach improves coding efficiency when compared to monoview DVC. Finally, video sensors to encode multiview video are described in [72]. The scheme exploits both interview correlation by disparity compensation from other views as well as temporal correlation by motion compensated lifted wavelet transform. The proposed scheme leads to a bit rate reduction by performing joint decoding when compared to separate decoding. Note that in all the above techniques, the cameras do not need to communicate. In particular, the joint statistical model is still derived at the decoder.

Two papers address multiview DVC coding *in this special issue*. In the first one [73], Taguchi and Naemura present a multiview DVC system which combines decoding and rendering to synthesize a virtual view while avoiding full reconstruction. More specifically, disparity compensation and geometric estimation are performed jointly. The coding efficiency of the system is evaluated, along with the decoding and rendering complexity.

The paper by Ouaret et al. [74] explores and compares different intercamera prediction techniques for SI. The assessment is done in terms of prediction quality, complexity, and coding performance. In addition, a new technique, referred to as Iterative Multiview Side Information, is proposed, using an iterative reconstruction process. Coding efficiency is compared to H.264/AVC, H.264/AVC No Motion and H.264/AVC Intra.

3.6. Applications beyond Coding. The DSC paradigm has been widely applied to realize image and video coding systems that shift a significant part of the computational load from the transmitter to the receiver side or allow a joint decoding of images taken by different cameras without any

need of information exchange among the coders. Outside the coding scenario, DSC has also found applications for some other domains.

For example, watermarks are normally used for media authentication, but one serious limitation of watermarks is lack of backward compatibility. More specifically, unless the watermark is added to the original media, it is not possible to authenticate it. In [75], an application of the DSC concepts to media hashing is proposed. This method provides a Slepian-Wolf encoded quantized image projection as an authentication data which can be successfully decoded only by using an authentic image as side information. DSC helps in achieving false acceptance rates close to zero for very small authentication data size. This scheme has been extended for tampering localization in [76].

Systems presented in [75, 76] can do successful image authentication for JPEG compressed images but are not able to work correctly if the transmission channel applies any linear transformation on the image such as contrast and brightness adjustment in addition to JPEG compression. Some improvements are presented in [77]. In [78], a more sophisticated system for image tampering detection is presented. It combines DVC and Compressive Sensing concepts to realize a system that is able to detect practically any type of image modification and is also robust to geometrical manipulation (cropping, rotation, change of scale, etc.).

In [79, 80], distributed source coding techniques are used for designing a secure biometric system for fingerprints. This system uses a statistical model of relationship between the enrollment biometric and the noisy biometric measurement taken during authentication.

In [81], a Wyner-Ziv coding technique is applied for multiple bit rate video streaming, which allows the server to dynamically change the transmitted stream according to available bandwidth. More specifically, in the proposed scheme, a switching stream is coded using Wyner-Ziv coding. At the decoder side, the switch-to frame is reconstructed by taking the switch-from frame as side information.

The application of DSC to other domains beyond coding is still a relatively new topic of research. It is not unexpected that further explorations will lead to significant results and opportunities for successful applications.

In this special issue, the paper by Valenzise et al. [82] deals with the application of DSC to audio tampering detection. More specifically, the proposed scheme requires that the audio content provider produces a small hash signature by computing a limited number of random projections of a perceptual, time-frequency representation of the original audio stream; the audio hash is given by the syndrome bits of an LDPC code applied to the projections. At the user side, the hash is decoded using distributed source coding tools, provided that the distortion introduced by tampering is not too high. If the tampering is sparsifiable or compressible in some orthonormal basis or redundant dictionary (e.g., DCT or wavelet), it is possible to identify the time-frequency position of the attack.

4. Perspectives

Based on the above considerations, in this section we offer some thoughts about the most important technical benefits provided by the DVC paradigm and the most promising perspectives and applications.

DVC has brought to the forefront a new coding paradigm, breaking the stronghold of motion-compensated DCT-based hybrid coding such as MPEG and ITU-T standards, and shedding a new light on the field of video coding by opening new research directions.

From a theoretical perspective, the Slepian-Wolf and Wyner-Ziv theorems state that DVC can potentially reach the same performance as conventional coding. However, as discussed in Section 2.4, in practice, this has only been achieved when the additional constraint of low complexity encoding is taken into account. In this case, state-of-the-art DVC schemes nowadays consistently outperform H.264/AVC Intracoding, while encoding is significantly simpler. Additionally, for sequences with simple motion, DVC matches and even in some cases surpasses H.264/AVC No Motion coding. However, the complexity advantage provided by DVC may be very transient, as with Moore's law, computing power increases exponentially and makes cost-effective within a couple of years the implementation that is not manageable today. As a counter argument to this, the time to have a solution with competitive cost relative to alternatives could be more than a couple years and this typically depends on the volumes that are sold and level of customization. Simply stated, we cannot always expect a state-of-the-art coding solution with a certain cost to be the best available option for all systems, especially those with high-resolution video specifications and nontypical configurations. It is also worth noting that there are applications that cannot tolerate high complexity coding solutions and are typically limited to intraframe coding due to platform and power consumption constraints; space and airborne systems are among the class of applications that fall into this category. For these reasons, it is possible that DVC can occupy certain niche applications provided that coding efficiency and complexity are at competitive and satisfactory levels.

Another domain where DVC has been shown to be appealing is for video transmission over error-prone network channels. This follows from the statistical framework on which DVC relies, and especially the absence of prediction loop in the codec. Moreover, as the field of DVC coding is still relatively young and the subject of intensive research, it is not unreasonable to expect further significant performance improvements in the near future.

The codec-independent scalability property of DVC is interesting and may bring an additional helpful feature in some applications. However, it is unlikely to be a differentiator by itself. Indeed, scalability is most often a secondary goal, surpassed by more critically important features such as coding efficiency or complexity. Moreover, the codec-independent flavor brought by DVC has not found its killer application yet.

Multiview coding is another domain where DVC shows promises. On top of the above benefits for monoview, DVC allows for an architecture where cameras do not need to communicate, while still enabling the exploitation of interview correlation during joint decoding. This may prove a significant advantage from a system implementation stand-point, avoiding complex and power consuming networking. However, multiview DVC coding systems reported to date still reveal a significant rate-distortion performance gap when compared to independent H.264/AVC coding for each camera. Note that the latter has to be preferred as a point of reference instead of Multiview Video Coding (MVC), as MVC requires communication between the cameras. Moreover, the amount of interview correlation, usually significantly lower than intraview temporal correlation, depends strongly on the geometry of the cameras and the scene.

Taking a very different path, it has been proposed in [83] to combine conventional and distributed coding into a single framework in order to move ahead towards the next rate-distortion performance level. Indeed, the significant coding gains of MPEG and ITU-T schemes over the years have mainly been the result of more complex analysis at the encoder. However, these gains have been harder to achieve lately and performance tends to saturate. The question remains whether more advanced analysis at the decoder, borrowing from distributed coding principles, could be the next avenue for further advances. In particular, this new framework could prove appealing for the up-and-coming standardization efforts on High-performance Video Coding (HVC) in MPEG and Next Generation Video Coding (NGVC) in ITU-T, which aim at a new generation of video compression technology.

Finally, while most of the initial interest in distributed source coding principles has been towards video coding, it is becoming clear that these ideas are also helpful for a variety of other applications beyond coding, including media authentication, secure biometrics, and tampering detection.

Based on the above considerations, DVC is most suited for applications which require low complexity and/or low power consumption at the encoder and video trans-mission over noisy channels, with content characterized by low-motion activity. Under the combination of these conditions, DVC may be competitive in terms of rate-distortion performance when compared to conventional coding approaches.

Following a detailed analysis, 11 promising applica-tion scenarios for DVC have been identified in [84]: wireless video cameras, wireless low-power surveillance, mobile document scanner, video conferencing with mobile devices, mobile video mail, disposable video cameras, visual sensor networks, networked camcorders, distributed video streaming, multiview video entertainment, and wireless capsule endoscopy. This inventory represents a mixture of applications covering a wide range of constraints offering different opportunities, and challenges, for DVC. Only time will tell which ones of those applications will span out and successfully deploy DVC-based solutions in the market place.

5. Conclusions

This paper briefly reviewed some of the most timely trends and perspectives for the use of DVC in coding applications and beyond. The following papers in this special issue further explore selected topics of interest addressing open issues in coding efficiency, error resilience, multiview coding, scalability, and applications beyond coding. This survey provides with a snapshot of significant research activities in the field of DVC but is by no means exhaustive. It is foreseen that this relatively new topic will remain a dynamic area of research in the coming years, which will bring further significant developments and progresses.

Acknowledgments

References

[1] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576, 2003.

[2] D. Slepian and J. K. Wolf, "Noiseless coding of correlated infor-mation sources," *IEEE Transactions on Information Theory*, vol. 19, no. 4, pp. 471–480, 1973.

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

[4] R. Puri and K. Ramchandran, "PRISM: a new robust video coding architecture based on distributed compression princi-ples," in *Proceedings of Allerton Conference on Communication, Control and Computing*, Allerton, Ill, USA, October 2002.

[5] A. Aaron, R. U. I. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," in *Proceedings of the 36th Asilomar Conference on Signals Systems and Computers*, pp. 240–244, Pacific Grove, Calif, USA, November 2002.

[6] C. Guillemot, F. Pereira, L. Torres, T. Ebrahimi, R. Leonardi, and J. Ostermann, "Distributed monoview and multiview video coding," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 67–76, 2007.

[7] P. L. Dragotti and M. Gastpar, *Distributed Source Coding: Theory, Algorithms and Applications*, Academic Press, New York, NY, USA, 2009.

[8] D. A. K. He, L. A. Lastras-Montano, and E. N. H. Yang, "A lower bound for variable rate slepian-wolf coding," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '06)*, pp. 341–345, Seattle, Wash, USA, July 2006.

[9] S. S. Pradhan, J. I. M. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Transactions on Informa-tion Theory*, vol. 49, no. 5, pp. 1181–1203, 2003.

[10] R. Puri, A. Majumdar, and K. Ramchandran, "PRISM: a video coding paradigm with motion estimation at the decoder," *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2436–2448, 2007.

[11] B. Girod, A. M. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71–83, 2005.

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

[13] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive distributed source coding using low-density parity-check codes," in *Proceedings of the 39th Asilomar Conference on Signals, Systems and Computers*, pp. 1203–1207, Pacific Grove, Calif, USA, November 2005.

[14] F. Pereira, C. Brites, J. Ascenso, and M. Tagliasacchi, "Wyner-Ziv video coding: a review of the early architectures and further developments," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '08)*, pp. 625–628, Hannover, Germany, June 2008.

[15] A. Aaron, S. Rane, and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," in *Proceedings of International Conference on Image Processing (ICIP '04)*, pp. 3097–3100, Singapore, October 2004.

[16] 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," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 57–60, Toulouse, France, May 2006.

[17] C. Brites and F. Pereira, "Encoder rate control for transform domain Wyner-Ziv video coding," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 2, pp. 5–8, San Antonio, Tex, USA, September 2007.

[18] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "The discover codec: architecture, techniques and evaluation," in *Proceedings of Picture Coding Symposium (PCS '07)*, Lisboa, Portugal, November 2007.

[19] Discover DVC Final Results, http://www.img.lx.it.pt/~discover/home.html.

[20] J. Ascenso and F. Pereira, "Integrated software tools for distributed video coding," VISNET II Deliverable D1.2.3, June 2009, http://ltswww.epfl.ch/~dufaux/visnet2/dels/d0072.pdf.

[21] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," in *Proceedings of the 5th EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services*, Smolenice, Slovak Republic, June-July 2005.

[22] C. Brites, J. Ascenso, and F. Pereira, "Improving transform domain Wyner-Ziv video coding performance," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '06)*, vol. 2, pp. 525–528, Toulouse, France, May 2006.

[23] K. Misra, S. Karande, and H. Radha, "Multi-hypothesis based distributed video coding using LDPC codes," in *Proceedings of Allerton Conference on Commun, Control and Computing*, Allerton, Ill, USA, September 2005.

[24] L. Wei, Y. Zhao, and A. Wang, "Improved side-information in distributed video coding," in *Proceedings of International Conference on Innovative Computing, Information and Control*, Beijing, China, August-September 2006.

[25] J. Ascenso, C. Brites, and F. Pereira, "Motion compensated refinement for low complexity pixel based distributed video coding," in *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS '05)*, pp. 593–598, Como, Italy, September 2005.

[26] X. Artigas and L. Torres, "Iterative generation of motion-compensated side information for distributed video coding," in *Proceedings of IEEE International Conference on Image Processing (ICIP '05)*, pp. 833–836, Genova, Italy, September 2005.

[27] W. A. R. J. Weerakkody, W. A. C. Fernando, J. L. Martínez, P. Cuenca, and F. Quiles, "An iterative refinement technique for side information generation in DVC," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '07)*, pp. 164–167, Beijing, China, July 2007.

[28] S. Ye, M. Ouaret, F. Dufaux, and T. Ebrahimi, "Improved side information generation with iterative decoding and frame interpolation for distributed video coding," in *Proceedings of IEEE International Conference on Image Processing (ICIP '08)*, San Diego, Calif, USA, October 2008.

[29] E. Martinian, A. Vetro, J. S. Yedidia, J. Ascenso, A. Khisti, and D. Malioutov, "Hybrid distributed video coding using SCA codes," in *Proceedings of the 8th IEEE Workshop on Multimedia Signal Processing (MMSP '06)*, pp. 258–261, Victoria, Canada, October 2006.

[30] M. Maitre, C. Guillemot, and L. Morin, "3-D model-based frame interpolation for distributed video coding of static scenes," *IEEE Transactions on Image Processing*, vol. 16, no. 5, pp. 1246–1257, 2007.

[31] C. Brites and F. Pereira, "Correlation noise modeling for efficient pixel and transform domain Wyner-Ziv video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 9, pp. 1177–1190, 2008.

[32] H. Wang, N. M. Cheung, and A. Ortega, "A framework for adaptive scalable video coding using Wyner-Ziv techniques," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 60971, 18 pages, 2006.

[33] Y. Vatis, S. Klomp, and J. Ostermann, "Enhanced reconstruction of the quantised transform coefficients for Wyner-Ziv coding," in *Proceedings of IEEE International Conference on Multimedia and Expo (ICME '07)*, pp. 172–175, Beijing, China, July 2007.

[34] D. Kubasov, J. Nayak, and C. Guillemot, "Optimal reconstruction in Wyner-Ziv video coding with multiple side information," in *Proceedings of the 9th IEEE International Workshop on Multimedia Signal Processing (MMSP '07)*, pp. 183–186, Crete, Greece, October 2007.

[35] W. A. R. J. Weerakkody, W. A. C. Fernando, and A. M. Kondoz, "An enhanced reconstruction algorithm for unidirectional distributed video coding," in *Proceedings of the 12th IEEE International Symposium on Consumer Electronics (ISCE '08)*, Algarve, Portugal, April 2008.

[36] M. Tagliasacchi, L. Frigerio, and S. Tubaro, "Rate-distortion analysis of motion-compensated interpolation at the decoder in distributed video coding," *IEEE Signal Processing Letters*, vol. 14, no. 9, pp. 625–628, 2007.

[37] Z. Li, L. Liu, and E. J. Delp, "Rate distortion analysis of motion side estimation in Wyner-Ziv video coding," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 98–113, 2007.

[38] B. Wu, X. Ji, D. Zhao, and W. Gao, "Spatial-aided low-delay Wyner-Ziv video coding," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 109057, 11 pages, 2009.

[39] S. Sofke, F. Pereira, and E. Müller, "Dynamic quality control for transform domain Wyner-Ziv video coding," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 978581, 15 pages, 2009.

[40] 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, Article ID 683510, 15 pages, 2009.

[41] F. Pereira, J. Ascenso, and C. Brites, "Studying the GOP size impact on the performance of a feedback channel-based Wyner-Ziv video codec," in *Proceedings of IEEE Pacific-Rim Symposium on Image and Video Technology*, Santiago, Chile, December 2007.

[42] J. Areia, J. Ascenso, C. Brites, and F. Pereira, "Low complexity hybrid rate control for lower complexity Wyner-Ziv video decoding," in *Proceedings of European Conference on Signal Processing (EUSIPCO '08)*, Lausanne, Switzerland, August 2008.

[43] C.-H. Lai, K.-Y. Hsieh, S.-H. Lai, and J.-K. Lee, "Parallelization of belief propagation method on embedded multicore processors for stereo vision," in *Proceedings of the 6th IEEE Workshop on Embedded System for Real-Time Multimedia (ESTIMedia '08)*, pp. 39–44, Atlanta, Ga, USA, October 2008.

[44] J. Zhang and M. Fossorier, "Shuffled belief propagation decoding," in *Proceedings of the 36th Annual Asilomar Conference on Signals Systems and Computers*, pp. 8–15, Pacific Grove, Calif, USA, November 2002.

[45] J. Zhang, Y. Wang, M. Fossorier, and J. S. Yedidia, "Replica shuffled belief propagation decoding of LDPC codes," in *Proceedings of the 39th Conference on Information Sciences and Systems (CISS '05)*, The Johns Hopkins University, Baltimore, Md, USA, March 2005.

[46] A. Majumdar, J. I. M. Chou, and K. Ramchandran, "Robust distributed video compression based on multilevel coset codes," in *Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*, pp. 845–849, Pacific Grove, Calif, USA, November 2003.

[47] M. Tagliasacchi, A. Majumdar, and K. Ramchandran, "A distributed-source-coding based robust spatio-temporal scalable video codec," in *Proceedings of Picture Coding Symposium (PCS '04)*, pp. 435–440, San Francisco, Calif, USA, December 2004.

[48] M. Tagliasacchi, A. Majumdar, K. Ramchandran, and S. Tubaro, "Robust wireless video multicast based on a distributed source coding approach," *Signal Processing*, vol. 86, no. 11, pp. 3196–3211, 2006.

[49] A. Sehgal, A. Jagmohan, and N. Ahuja, "Wyner-Ziv coding of video: an error-resilient compression framework," *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 249–258, 2004.

[50] Q. Xu, V. Stanković, and Z. Xiong, "Layered Wyner-Ziv video coding for transmission over unreliable channels," *Signal Processing*, vol. 86, no. 11, pp. 3212–3225, 2006.

[51] M. Fumagalli, M. Tagliasacchi, and S. Tubaro, "Drift reduction in predictive video transmission using a distributed source coded side-channel," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '06)*, Toulouse, France, May 2006.

[52] S. Rane, A. Aaron, and B. Girod, "Systematic lossy forward error protection for error-resilient digital video broadcasting—a Wyner-Ziv coding approach," in *Proceedings of International Conference on Image Processing (ICIP '04)*, pp. 3101–3104, Singapore, October 2004.

[53] L. Liang, P. Salama, and E. J. Delp, "Adaptive unequal error protection based on Wyner–Ziv coding," in *Proceedings of Picture Coding Symposium (PCS '07)*, Lisbon, Portugal, November 2007.

[54] 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*, vol. 23, no. 5, pp. 391–403, 2008.

[55] C. Tonoli, P. Migliorati, and R. Leonardi, "Error resilience in current distributed video coding architectures," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 946585, 18 pages, 2009.

[56] L. Liang, P. Salama, and E. J. Delp, "Unequal error protection techniques based on Wyner-Ziv coding," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 474689, 13 pages, 2009.

[57] A. Sehgal, A. Jagmohan, and N. Ahuja, "Scalable video coding using Wyner-Ziv codes," in *Proceedings of Picture Coding Symposium (PCS '04)*, pp. 441–446, San Francisco, Calif, USA, December 2004.

[58] T. Ebrahimi and F. Pereira, *The MPEG-4 Book*, Prentice Hall, Englewood Cliffs, NJ, USA, 2002.

[59] M. Ouaret, F. Dufaux, and T. Ebrahimi, "Codec-independent scalable distributed video coding," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 3, pp. 9–12, San Antonio, Tex, USA, October 2007.

[60] A. Skodras, C. Christopoulos, and T. Ebrahimi, "The JPEG 2000 still image compression standard," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 36–58, 2001.

[61] B. Macchiavello, F. Brandi, E. Peixoto, R. L. de Queiroz, and D. Mukherjee, "Side-information generation for temporally and spatially scalable Wyner-Ziv codecs," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 171257, 11 pages, 2009.

[62] F.-O. Devaux and C. De Vleeschouwer, "Parity bit replenishment for JPEG 2000-based video streaming," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 683820, 18 pages, 2009.

[63] M. Ouaret, F. Dufaux, and T. Ebrahimi, "Multiview distributed video coding with encoder driven fusion," in *Proceedings of European Conference on Signal Processing (EUSIPCO '07)*, Poznan, Poland, September 2007.

[64] X. Artigas, F. Tarres, and L. Torres, "Comparison of different side information generation methods for multiview distributed video coding," in *Proceedings of International Conference on Signal Processing and Multimedia Applications (SIGMAP '07)*, Barcelona, Spain, July 2007.

[65] F. Dufaux, M. Ouaret, and T. Ebrahimi, "Recent advances in multi-view distributed video coding," in *Mobile Multimedia/Image Processing for Military and Security Applications*, vol. 6579 of *Proceedings of SPIE*, Orlando, Fla, USA, April 2007.

[66] E. Martinian, A. Behrens, X. I. N. Jun, and A. Vetro, "View synthesis for multiview video compression," in *Proceedings of the 25th Picture Coding Symposium (PCS '06)*, Beijing, China, April 2006.

[67] S. M. Seitz and C. R. Dyer, "View morphing," in *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*, pp. 21–42, New Orleans, La, USA, August 1996.

[68] M. Ouaret, F. Dufaux, and T. Ebrahimi, "Fusion-based multiview distributed video coding," in *Proceedings of the 4th ACM International Workshop on Video Surveillance and Sensor Networks (VSSN '06)*, pp. 139–144, Santa Barbara, Calif, USA, October 2006.

[69] X. Artigas, E. Angeli, and L. Torres, "Side information generation for multiview distributed video coding using a fusion approach," in *Proceedings of the 7th Nordic Signal Processing Symposium (NORSIG '06)*, pp. 250–253, Reykjavik, Iceland, June 2006.

[70] X. Guo, Y. Lu, F. Wu, W. Gao, and S. Li, "Distributed multiview video coding," in *Visual Communications and Image Processing*, vol. 6077 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2006.

[71] X. U. N. Guo, Y. A. N. Lu, F. Wu, D. Zhao, and W. E. N. Gao, "Wyner-Ziv-based multiview video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 6, pp. 713–724, 2008.

[72] M. Flierl and B. Girod, "Coding of multi-view image sequences with video sensors," in *Proceedings of IEEE International Conference on Image Processing (ICIP '06)*, Atlanta, Ga, USA, October 2006.

[73] Y. Taguchi and T. Naemura, "Rendering-oriented decoding for a distributed multiview coding system using a coset code," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 251081, 12 pages, 2009.

[74] M. Ouaret, F. Dufaux, and T. Ebrahimi, "Iterative multiview side information for enhanced reconstruction in distributed video coding," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 591915, 17 pages, 2009.

[75] Y. A. O. C. Lin, D. Varodayan, and B. Girod, "Image authentication based on distributed source coding," in *Proceedings of the 14th IEEE International Conference on Image Processing (ICIP '07)*, vol. 3, pp. 5–8, San Antonio, Tex, USA, October 2007.

[76] Y. A. O. C. Lin, D. Varodayan, and B. Girod, "Image authentication and tampering localization using distributed source coding," in *Proceedings of the 9th IEEE International Workshop on Multimedia Signal Processing (MMSP '07)*, pp. 393–396, Chania, Greece, October 2007.

[77] N. Khanna, A. Roca, G. T.-C. Chiu, J. P. Allebach, and E. J. Delp, "Improvements on image authentication and recovery using distributed source coding," in *Media Forensics and Security*, vol. 7254 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2009.

[78] M. Tagliasacchi, G. Valenzise, and S. Tubaro, "Localization of sparse image tampering via random projections," in *Proceedings of International Conference on Image Processing (ICIP '08)*, pp. 2092–2095, San Diego, Calif, USA, October 2008.

[79] S. C. Draper, A. Khisti, E. Martinian, A. Vetro, and J. S. Yedidia, "Using distributed source coding to secure fingerprint biometrics," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, vol. 2, pp. 129–132, Honolulu, Hawaii, USA, April 2007.

[80] Y. Sutcu, S. Rane, J. S. Yedidia, S. C. Draper, and A. Vetro, "Feature transformation of biometric templates for secure biometric systems based on error correcting codes," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop (CVPR '08)*, Anchorage, Alaska, USA, June 2008.

[81] M. E. I. Guo, Y. A. N. Lu, F. Wu, D. Zhao, and W. E. N. Gao, "Wyner-Ziv switching scheme for multiple bit-rate video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 5, pp. 569–581, 2008.

[82] G. Valenzise, G. Prandi, M. Tagliasacchi, and A. Sarti, "Identification of sparse audio tampering using distributed source coding and compressive sensing techniques," *EURASIP Journal on Image and Video Processing*, vol. 2009, Article ID 158982, 12 pages, 2009.

[83] F. Pereira, "Video compression: still evolution or time for revolution?" in *Proceedings of the 10th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '09)*, London, UK, May 2009.

[84] 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*, vol. 23, no. 5, pp. 339–352, 2008.