Efficient Hash-Driven Wyner-Ziv Video Coding for Visual Sensors

Nikos Deligiannis,[a,c] Marc Jacobs,[a,c] Frederik Verbist,[a,c] Jürgen Slowack,[b,c] Joeri Barbarien,[a,c]
Rik Van de Walle,[b,c] Peter Schelkens,[a,c] and Adrian Munteanu[a,c]

[a]Department of Electronics and Informatics (ETRO), Vrije Universiteit Brussel (VUB), Pleinlaan 2, 1050 Brussels, Belgium.
[b]ELIS Department, Multimedia Lab, Ghent University, Gaston Crommenlaan 8, 9050 Ghent, Belgium.
[c]Interdisciplinary Institute for BroadBand Technology (IBBT), Gaston Crommenlaan 8 (b102), 9050 Ghent, Belgium.

Abstract—Recent advances in wireless visual sensor technology, have been calling for innovative architectures realizing efficient video coding under stringent processing and energy restrictions. Driven by profound findings in network information theory, Wyner-Ziv video coding constitutes a suitable paradigm for video sensor networks. This work presents a novel hash-driven Wyner-Ziv video coding architecture for visual sensors, which coarsely encodes a low resolution version of each Wyner-Ziv frame to facilitate accurate motion-compensated prediction at the decoder. The proposed method for side-information generation comprises hash-based multi-hypothesis pixel-based prediction. Once critical Wyner-Ziv information is decoded, the derived dense motion field is further enhanced. Experimental validation illustrates that the proposed hash-driven codec achieves significant compression gains with respect to state-of-the-art Wyner-Ziv video coding, even under demanding conditions.

Keywords-Wyner-Ziv coding; hash-driven distributed video coding; low-cost encoding; visual sensors

I. INTRODUCTION

Recently, wireless multimedia sensor technology has undergone a quiet revolution, significantly influencing a mixture of domains associated to modern lifestyle. Unlike current display and acquisition devices, sensor networks are task-oriented and application-specific, e.g., detection of biological agents, environmental measurements, video surveillance. This explosive growth of diverse visual sensor applications is calling for multimedia communication technology capable of dealing with the associated harsh operating conditions. In particular, multimedia content requires high data rates which consume elevated bandwidth and energy resources. Hence, high compression performance is a basic requirement in modern video communications. Moreover, the restricted energy supply of wireless sensors plays a key role in all aspects of a wireless sensor network, from the node's processing power, to the volume of information it can transmit and receive. What is more, establishing a reliable transmission over error-prone wireless channels demands coding architectures delivering advanced error resilience.

To meet these strict requirements, new coding paradigms need to be considered. An emerging and promising technology for scalable, error resilient, low complexity encoding is distributed source coding (DSC). The latter stems from information-theoretic findings established by Slepian and Wolf [1] and by Wyner and Ziv [2], according to which efficient compression can be achieved by leveraging knowledge of the source statistics at the decoder only. Conventional video coding principles, used in the widely spread ISO MPEG standards, or the ITU-T recommendations, H.263 [3] and H.264/AVC [4] (also known as MPEG-4 part 10), are not suitable to offer high-performance low-cost video encoding in wireless visual sensor networks [5], [6]. In these traditional architectures, the encoder bears the brunt of the computational burden, and its complexity is mainly dominated by the motion estimation operation. Conversely, inspired by DSC principles, distributed video coding (DVC) frameworks [7], [8], [9], employ simple yet efficient encoders, combined with more complex decoders performing joint decoding. Furthermore, as Slepian-Wolf coding can be realized by channel codes, communication error robustness is an inherent feature of DVC schemes. Due to their attractive characteristics, DVC systems have been identified as a promising solution for distributed networks of video sensors, and are expected to have a major impact on the future design of protocols for wireless multimedia sensors [5], [6].

Despite all the research efforts dedicated to DVC, current systems still fail to meet the compression efficiency of their predictive video coding counterparts. One of the main reasons for this deficit is attributed to the shortcoming of Wyner-Ziv video decoders to generate efficient motion-compensated prediction [10], [11]. Predictive video encoders exploit the available temporal correlation in the video sequence by generating high quality prediction using motion estimation and compensation. In contrast, most common DVC schemes employ motion compensated interpolation (MCI) to generate side-information at the decoder. However, MCI assumes a linear motion model which falls short in case of irregular motion [10]. To alleviate this problem, joint decoding and side-information refinement [12], [13], [14], [15], or hash-based motion estimation [16], [17], [18], [19] have been proposed.

Along these lines, a favorable feature of predictive video coding systems is their capacity to exploit correlations in long group of pictures (GOP). However, the performance of state-of-the-art DVC architectures [9] deteriorates significantly when

the considered GOP size increases; this being the reason why the most commonly used GOP in the DVC literature is 2 [9].

This work presents a novel hash-driven DVC architecture for visual sensors, which builds on several contributions. Unlike MCI-based codecs, the proposed system encodes a robust hash to facilitate efficient motion-compensated prediction at the decoder. In contrast to other hash-based Wyner-Ziv video codecs in the literature, the proposed system does not perform block-based mode decision [16], [17], temporal correlation exploitation [18], or encoder-side motion estimation [20] in order to code the hash data. In addition, at the basis of the proposed side-information generation method, overlapped block motion estimation and compensation [19] enables accurate multi-hypothesis pixel-based prediction at the decoder. After decoding the DC coefficient band, the proposed codec refines the generated side-information and uses this result to decode the remaining DCT bands. It is notable that, in contrast to existing hash-based DVC codecs, e.g., [17], the proposed system is liberated from a prediction loop at the encoder, thus retaining the communication error robustness traits of Wyner-Ziv coding.

Experimental evaluation shows that the proposed hash-based technique enables accurate side-information creation even under demanding conditions, that is, irregular motion content and long GOPs. Results illustrate that the proposed codec achieves remarkable compression gains compared to the state-of-the-art DISCOVER [9] codec of up to 40.85% Bjøntegaard rate reduction [21]. Moreover, the proposed codec retains the principle characteristics of Wyner-Ziv coding, i.e., low-cost encoding, error robustness and temporally scalable coding, thus constituting a competitive scheme for visual sensors.

II. DISTRIBUTED SOURCE AND VIDEO CODING

A. Distributed Source Coding Prologue

The Slepian-Wolf [1], and Wyner-Ziv [2] theorems are the foundation stones of DSC frameworks.

1) Slepian-Wolf Coding: Consider the compression of two correlated, discrete, identically and independently distributed (i.i.d.), random sources X and Y. In traditional predictive coding, where both the encoder and the decoder access the dependencies between the sources, the lowest achievable bound for lossless compression is the joint entropy $H(X|Y)$. Surprisingly, the Slepian-Worf theorem [1] proves that, even when the sources are separately encoded and jointly decoded, the achievable lowest rate point for decoding, with an arbitrarily small probability of error, is still given by $H(X|Y)$.

2) Wyner-Ziv Coding: Wyner and Ziv [2] studied the case of distributed lossy coding with side-information at the decoder. Let X and Y be two statistically dependent i.i.d. random sequences, where X is independently encoded and jointly decoded using Y as side-information, to a reconstructed sequence $\hat{X}$ yielding an expected distortion $D = E\left[d(x,\hat{x})\right]$. The Wyner-Ziv theorem states that a rate increase is incurred due to unaware source encoding of $R_{X|Y}^*(D) - R_{X|Y}(D) \geq 0$,

where, $R_{X|Y}^*(D)$, $R_{X|Y}(D)$ denote the Wyner-Ziv and predictive coding rate-distortion (RD) function, respectively. However, Wyner and Ziv demonstrated that this performance loss is eliminated for Gaussian memoryless sources and a mean-square distortion metric. This result has been generalized for sources defined by the sum of arbitrarily distributed side-information and Gaussian noise independent of the side-information [22].

B. Distributed Video Coding Architectures

Distributed video compression has attracted significant attention lately, also due to its application in wireless sensor networks. Several practical systems have been proposed. Initially, Puri *et al.* presented PRISM [8], a codec which carries out block classification at the encoder, and employs Bose-Chaudhuri-Hocquenghem (BCH) and cyclic redundancy check (CRC) codes to exploit temporal correlations at the decoder. Moreover, layered Wyner-Ziv coding was shown to facilitate scalable video coding [23]. In an alternative architecture [7], intra frame coding was employed to code the key frames, based on which side-information is generated by the decoder. Modeling the correlation between the source and the side-information as a virtual channel, powerful channel codes were used to encode the Wyner-Ziv frames. Rate control was performed at the decoder using a probabilistic model, requesting additional data from the encoder via a feedback channel. This architecture was afterwards improved by the VISNET research framework funded by the EU, leading to the state-of-the-art DISCOVER codec [9].

III. THE PROPOSED WYNER-ZIV VIDEO CODEC

This section presents the architecture of the proposed hash-driven transform-domain Wyner-Ziv video codec. The block diagram of the codec is given in Figure 1. The frames of the video sequence are split into key and Wyner-Ziv (WZ) frames. The key frames, K, are intra frame coded using H.264/AVC, while the WZ frames, W, are encoded in two parts, namely, a hash and a WZ bit-stream.

The hash bit-stream is composed of a low resolution, coarsely coded description of each WZ frame. In specific, to create the hash, each WZ frame is first down-scaled with a factor of 2. Then, the down-scaled WZ frame is intra coded with H.264/AVC at a lower quality compared to the quality of the key frames.

Together with the hash data, a Wyner-Ziv bit-stream is formed for each WZ frame. In contrast to other hash-driven DVC schemes, e.g., [17], the Wyner-Ziv encoder is chosen to encode the original WZ frame – rather than its difference with the hash – in order to enable robustness against communication errors [7]. Following the principles of the architecture in [7], the WZ pixel values are transformed using the integer discrete cosine transform (DCT) [4]. Subsequently, the obtained DCT coefficients are grouped together into DCT bands b which are quantized with 2^{M_b} levels. Predefined quantization matrices (QM), as in [9], are employed. The quantized symbols are then converted to binary codewords and fed to an LDPC Accumulate (LDPCA) codec [24].
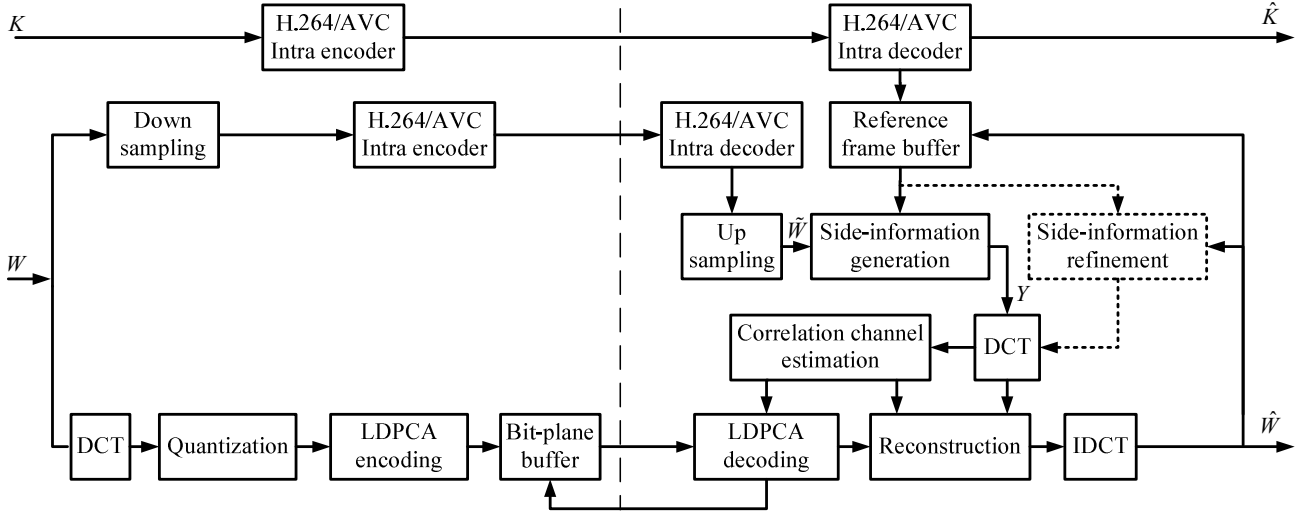
Figure 1. Schematic description of the proposed hash-driven Wyner-Ziv video codec.

At the decoder, the key frames are H.264/AVC Intra decoded and stored in the reference frame buffer. Also, the H.264/AVC intra decoded and up-scaled [25] hash data is used to produce side-information. A detailed presentation of the proposed side-information generation framework is given in Section IV. The resulting side-information is first transformed and converted to soft information in order to decode the bit-planes of the DC coefficient band of the WZ frame. After LDPCA decoding, the derived bit-planes are grouped into quantization indices and minimum mean square error (MMSE) reconstruction [5] is carried out to obtain the decoded DC coefficients. Thereafter, the transformed side-information is updated with the decoded DC band coefficients and the inverse DCT is performed. Next, using the updated WZ frame and reference frames from the buffer, a side-information refinement approach improves the derived motion field. This yields improved side-information, which is used to LDPCA decode and MMSE reconstruct the remaining DCT coefficient bands of the WZ frame. Finally, after decoding all the DCT coefficient bands, the inverse DCT is performed and the decoded WZ frame is displayed and buffered.

IV. HASH-DRIVEN SIDE-INFORMATION GENERATION

This section describes the proposed framework to perform enhanced motion-compensated prediction in the presented hash-driven Wyner-Ziv video codec. Utilizing the decoded hash $\tilde{W}$ and two reference frames, R^n, $n = \{0,1\}$, the proposed codec supports a hierarchical bidirectional motion prediction structure. This structure is commonly utilized by DVC systems, [7], [9], [12], [17], enabling temporal scalability, a feature of great importance in heterogeneous sensor networks in which resources and demands are greatly varying.

A. Hash-based Motion Estimation

Using the decoded and up-scaled hash data and the reference frames, the proposed Wyner-Ziv video decoder performs overlapped block motion estimation (OBME) and compensation to generate side-information. This approach is

driven by our previous work in [19] with key modifications as a result of our novel DVC scheme.

First, the decoded and up-scaled hash data $\tilde{W}$ is divided into overlapping spatial blocks of 16×16 pixels with an overlapping step of 4 pixels. For each overlapping block the best matching block, within a search range of 15 pixels, is found in each of the reference frames R^n. The derived motion vectors minimize the sum of absolute differences (SAD), that is

$$SAD^n = \sum_{i=0}^{15} \sum_{j=0}^{15} \left| \tilde{W}\left(i + p_x, j + p_y\right) - R^n\left(i + p_x + v_x^n, j + p_y + v_y^n\right) \right|$$

(1)

where, $\mathbf{p} = \left(p_x, p_y\right)$ denotes top-left coordinates of a block in $\tilde{W}$, and $\mathbf{v}^n = \left(v_x^n, v_y^n\right)$ is the candidate motion vector per reference frame R^n. Next, per overlapping block $\tilde{W}_{\mathbf{p}^k} = \tilde{W}\left(i + p_x^k, j + p_y^k\right)$, $0 \le i, j < 15$, the algorithm retains only the motion vector $\mathbf{v}_k = \left\{ \mathbf{v}_k^{n=0} \text{ or } \mathbf{v}_k^{n=1} \right\}$, which provides the minimum SAD value. The derived motion field identifies the best temporal predictor $R_{\mathbf{p}^k + \mathbf{v}_k}$, per side-information overlapping block $Y_{\mathbf{p}^k}$. Since the motion estimated blocks are overlapping, each pixel $Y\left(i', j'\right)$ in Y belongs to several overlapping blocks $Y_{\mathbf{p}^k}$. For each overlapping block, OBME has identified the best temporal predictor $R_{\mathbf{p}^k + \mathbf{v}_k}$ in one of the reference frames. Therefore, each pixel in the side-information frame has a number of pixel predictors $R_{\mathbf{p}^k + \mathbf{v}_k}\left(i^k, j^k\right)$ from the two reference frames. During compensation, the estimated value of a pixel $Y\left(i', j'\right)$ in Y, is determined as the mean value of the predictor pixel values determined by OBME, namely,

$$Y\left(i', j'\right) = \frac{1}{K_{\mathbf{p}^k}} \sum_{k=0}^{K_{\mathbf{p}^k}-1} R_{\mathbf{p}^k + \mathbf{v}_k}\left(i^k, j^k\right),$$

(2)

where, $K_{\mathbf{p}^k}$ is the number of overlapping blocks $Y_{\mathbf{p}^k}$ to which the $Y\left(i', j'\right)$ pixel belongs. We notice that the presented algorithm features overlapping of the motion-estimated blocks, thereby diminishing blocking artifacts.
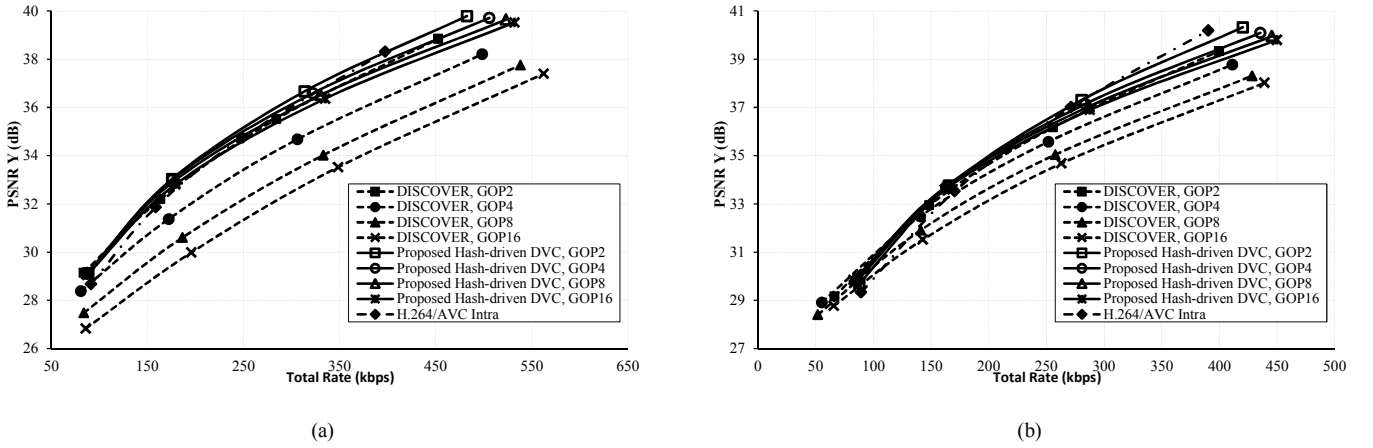
Figure 2. The compression performance of the proposed hash-driven Wyner-Ziv video codec against DISCOVER and H.264/AVC Intra for (a) Foreman, QCIF, 15Hz, and (b) Carphone, QCIF, 15Hz.

TABLE I. BJØNTEGAARD GAINS OF THE PROPOSED HASH-DRIVEN WYNER-ZIV CODEC AGAINST DISCOVER

Sequence	GOP 2		GOP 4		GOP 8		GOP 16	
	ΔR (%)	ΔPSNR (dB)	ΔR (%)	ΔPSNR (dB)	ΔR (%)	ΔPSNR (dB)	ΔR (%)	ΔPSNR (dB)
Foreman	-5.37	0.35	-21.57	1.38	-34.43	2.38	**-40.85**	**3.01**
Carphone	-2.26	0.20	-6.64	0.47	-16.90	1.06	**-22.69**	**1.40**
Ice	-6.22	0.48	-18.56	1.44	-28.19	2.37	**-32.85**	**2.94**

B. Side-information Refinement

After decoding the DC coefficient band, the WZ decoder has an improved approximation of the WZ frame at its disposal. This stems from the fact that the DC coefficient compacts most of the block energy, since a 4×4 transform is used. Therefore, the decoder can exploit this additional information to significantly improve the quality of the already obtained motion field. In contrast to other works in the literature, e.g., [13], [14], in the proposed algorithm side-information refinement is executed in the spatial domain in order to avoid having to construct a complex overcomplete DCT representation. Therefore, after DC coefficient decoding, the initial DCT transformed side-information is updated with the decoded DC band coefficients and an inverse DCT is performed, yielding the partially decoded WZ frame. The proposed side-information refinement technique uses the available partially decoded WZ frame and the reference frames to perform motion estimation as explained in Section IV.A.

V. EXPERIMENTAL RESULTS

This section assesses the performance of the presented Wyner-Ziv video codec featuring the proposed hash-driven methodology to generate side-information at the decoder. The efficiency of the proposed codec has been compared to that of the DISCOVER codec [9], and the DVC scheme of [17], which deliver state-of-the-art distributed video compression. Moreover, the H.264/AVC Intra frame codec [4] has been used as the lowest-complexity reference for H.264/AVC.

Concerning test conditions, compression results are provided for all frames of Foreman, Carphone and Ice

sequences, at QCIF resolution and a frame rate of 15Hz. Four GOP sizes are considered, that is, GOP of size 2, 4, 8, and 16. To retain a constant quality along the decoded sequence the quantization parameters (QPs) of the H.264/AVC intra coded frames and the QMs of the Wyner-Ziv frames are selected as in [9]. The QP for hash coding is empirically set to 42.

The proposed codec is compared against DISCOVER [9], which generates side-information with an advanced hierarchical MCI framework [26]. The experimental results for Foreman and Carphone are depicted in Figure 2. The average RD performances of the evaluated codecs for all sequences are also listed in Table I. For convenience, the results are depicted in terms of the Bjøntegaard (BD) metric [21]. The results in Figure 2 and Table I show that the proposed codec outperforms DISCOVER, at all GOPs and both sequences. Additionally, it is important to highlight that, while DISCOVER's performance drops dramatically when the GOP size increases, the proposed system suffers only a fairly small compression loss. As a consequence, the longer the GOP size, the higher the compression gain brought by the proposed codec, which can be up to 22.69%, 32.85% and 40.85% BD rate reduction in Carphone, Ice and Foreman GOP16, respectively. The dramatic performance loss of DISCOVER as the GOP size increases is due to the failure of the linear motion assumption at the basis of the MCI technique. To alleviate this problem, the hash-based DVC scheme of [17], performs block-based decision at the encoder in order to identify the blocks for which MCI is predicted to fail. These blocks are low quality intra coded and are employed to assist MCI at the decoder. Compared to the system of [17], the proposed codec brings significant BD rate improvements of 4.28%, 17.38%, and 28.28% in Foreman GOP 2, 4, and 8, respectively.
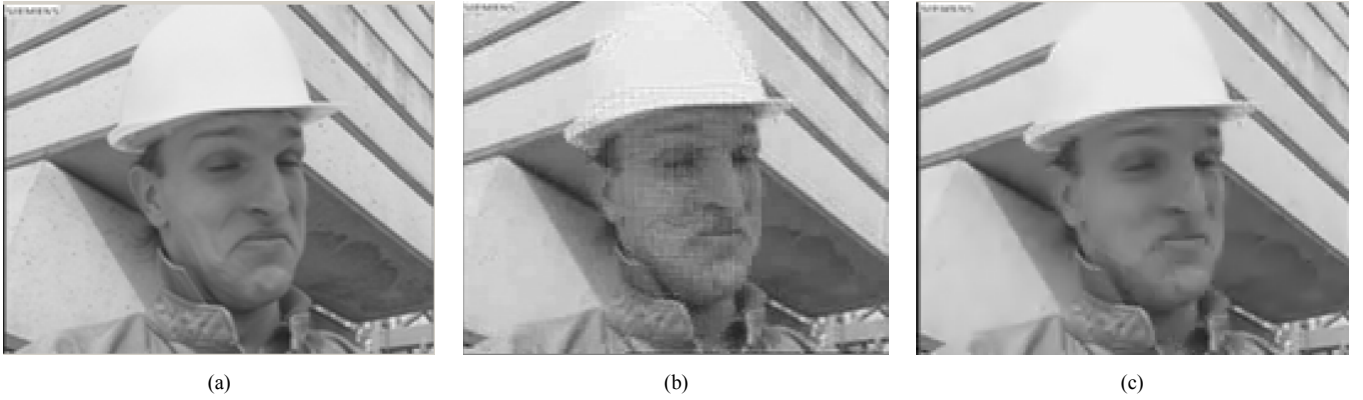
(a)   (b)   (c)

Figure 3. Visual comparison for Foreman, QCIF, 15Hz, GOP16 (a) original, (b) coded with DISCOVER (348.3kbps, 33.5dB), (c) coded with the proposed hash-driven Wyner-Ziv video codec (334.8kbps, 36.4dB). The parameters of the DISCOVER codec follow the test conditions of the DISCOVER project website [9].
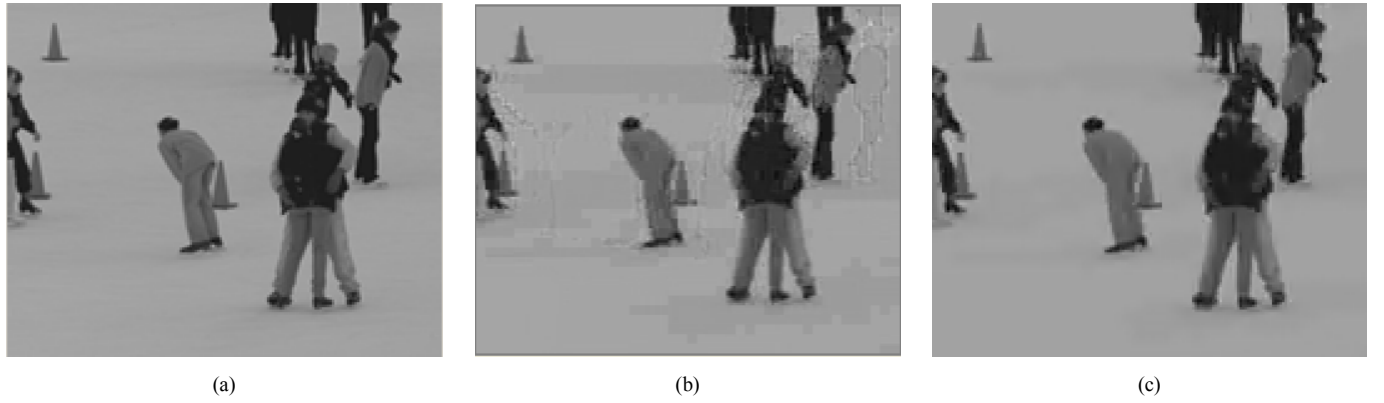

(a)   (b)   (c)

Figure 4. Visual comparison for Ice, QCIF, 15Hz, GOP16: (a) original, (b) coded with DISCOVER (185.7kbps, 31.9dB), (c) coded with the proposed hash-driven Wyner-Ziv video codec (176.9kbps, 34.8dB). The parameters of the DISCOVER codec follow the test conditions of the DISCOVER project website [9].

What is more, in contrast to DISCOVER and the scheme of [17], the proposed codec can support low-delay coding without significant loss in performance. Initial results have shown that, when the proposed codec deploys unidirectional prediction, a compression loss of 4.3% in rate (Foreman QCIF, 15Hz, GOP16) appears with respect to bidirectional prediction.

The subjective performance of the proposed system, with respect to DISCOVER, for Foreman and Ice is illustrated in Figure 3 and Figure 4, respectively. From the depicted sequence snapshots it is clear that the proposed codec delivers significantly better visual quality than DISCOVER.

Triggered by low power processing limitations imposed by sensor nodes, low-cost encoding is a key aspect of distributed video compression. In general, the encoding complexity of a Wyner-Ziv video codec is dominated by H.264/AVC intra frame coding. Essentially, the complexity associated to Wyner-Ziv encoding is negligible in relation to key frames' encoding complexity [27]. Hence, the lower the number of key frames in the encoded video sequence, i.e., the longer the GOP, the higher the gain in complexity offered by DVC with respect to H.264/AVC intra frame coding. Execution time measurements under controlled conditions, as defined by the VISNET I project [9], [27], have shown that the proposed codec brings an average encoding time reduction of approximately 30%, 50%,

60%, 65% for a GOP of size 2, 4, 8, and 16, respectively, compared to H.264/AVC Intra.

In contrast to pure Wyner-Ziv codecs, e.g., DISCOVER [9], the proposed codec exhibits higher encoding complexity due to hash formation and coding. However, this complexity overhead is kept quite low, given that the hash data consists of one fourth of the original frame size, which is coarsely coded with H.264/AVC intra frame coding. It also worth mentioning that compared to DISCOVER, for a given distortion level, the proposed codec offers a vast reduction of the encoding rate (up to 40.85%). This significant rate reduction brings an important decrease in power consumed by the transmission part of a visual sensor node. Given the scale between the processing requirements of multimedia coding and transmission, the proposed codec can reduce the overall power consumption of a visual sensor compared to DISCOVER. One concludes that the proposed hash-driven Wyner-Ziv video codec can be a competitive solution for wireless multimedia sensor nodes.

VI. CONCLUSIONS

Driven by the strict prerequisites of visual sensor technology, this paper has presented a novel Wyner-Ziv video codec. The proposed codec features an advanced hash-driven motion-compensated prediction framework, which enables accurate side-information creation at the decoder. Concrete

experimental assessment has demonstrated the superior compression performance of the proposed codec against state-of-the-art Wyner-Ziv video coding systems. Coding gains up to 40.85% in rate reduction with respect to DISCOVER have been observed, with a small increase in encoding complexity. Additionally, the proposed codec offers inherent communication error robustness and temporal scalability, thereby being a strong candidate for wireless visual sensors.

REFERENCES

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

[2] 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, January 1976.

[3] G. Cote, B. Erol, M. Gallant, and F. Kossentini, "H.263+: video coding at low bit rates," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 8, no. 7, pp. 849-866, November 1998.

[4] 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, July 2003.

[5] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Processing Magazine,* vol. 21, pp. 80-94, September 2004.

[6] I. Akyildiz, T. Melodia, and K. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks,* vol. 51, no. 4, pp. 921-960, March 2007.

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

[8] 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, October 2007.

[9] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Quaret, "The DISCOVER codec: Architecture, techniques and evaluation," in *Picture Coding Symposium, PCS 2007*, Lisboa, Portugal, November 2007 [Online]. Available: www.discoverdvc.org.

[10] 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, January 2007.

[11] G. Petrazzuoli, M. Cagnazzo, and B. Pesquet-Popescu, "High order motion interpolation for side information improvement in DVC," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX, March 2010.

[12] R. Martins, C. Brites, J. Ascenso, and F. Pereira, "Refining side information for improved transform domain Wyner-Ziv video coding," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 19, no. 9, pp. 1327-1341, Sep. 2009.

[13] A. B. B. Adikari, W. A. C. Fernando, W. A. R. J. Weerakkody, and H. K. Arachchi, "Sequential motion estimation using luminance and chrominance information for distributed video coding of Wyner-Ziv frames," *IEE Electronic Letters,* vol. 42, no. 7, pp. 398-399, March 2006.

[14] D. Varodayan, D. Chen, M. Flieri, and B. Girod, "Wyner-Ziv coding of video with unsupervised motion vector learning," *Signal Processing: Image Communication,* vol. 23, no. 5, pp. 369-378, June 2008.

[15] 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, pp. 1-15, 2009.

[16] A. Aaron, S. Rane, and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," in *IEEE International Conference on Image Processing, ICIP*, Singapore, October 2004.

[17] J. Ascenso, C. Brites, and F. Pereira, "A flexible side information generation framework for distributed video coding," *Springer, Multimedia Tools and Applications, Special Issue on Distributed Video Coding,* July 2009, DOI 10.1007/s11042-009-0316-6

[18] E. Martinian, A. Vetro, J. S. Yedidia, J. Ascenso, A. Khisti, and D. Malioutov, "Hybrid distributed video coding using sca codes," in *IEEE Multimedia Signal Processing Workshop*, Victoria, Canada, October 2006.

[19] N. Deligiannis, A. Munteanu, T. Clerckx, J. Cornelis, and P. Schelkens, "Overlapped block motion estimation and probabilistic compensation with application in distributed video coding," *IEEE Signal Processing Letters,* vol. 16, no. 9, pp. 743-746, September 2009.

[20] B. Macchiavello, D. Mukherjee, and R. L. de Queiroz, "Iterative side-information generation in a mixed resolution Wyner-Ziv framework," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 19, no. 10, pp. 1409-1423, October 2009.

[21] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves," ITU-T Video Coding Experts Group (VCEG),, Austin, TX, Document VCEG-M33, April 2001 2001.

[22] S. Pradhan, j. Chou, and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Transactions on Information Theory,* vol. 49, no. 5, pp. 1181-1203, May 2003.

[23] Q. Xu and Z. Xiong, "Layered Wyner-Ziv video coding," *IEEE Transactions on Image Processing,* vol. 15, no. 12, pp. 3791-3803, December 2006.

[24] D. Varodayan, A. Aaron, and B. Girod, "Rate-adaptive codes for distributed source coding," *Signal Processing, Elsevier,* vol. 86, pp. 3123–3130, 2006.

[25] C. A. Segall and G. J. Sullivan, "Spatial scalability within the H.264/AVC scalable video coding extension," *IEEE Transactions on Circuits and Systems for Video Technology,* vol. 17, no. 9, pp. 1121-1135, September 2007.

[26] J. Ascenso, C. Brites, and F. Pereira, "Content adaptive Wyner-Ziv video coding driven by motion activity," in *IEEE International Conference on Image Processing*, Atlanta, GA, Oct. 2006.

[27] 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 *IEEE Pacific Rim Symposium on Image Video and Technology*, Santiago, Chile, December 2007.