2011 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) December 7-9, 2011

Distributed Source Coding : Theorem and Its Application To Video Coding

Takayuki Nakachi
NTT Network Innovation Laboratories
Nippon Telegraph and Telephone Corp. 239-0847 JAPAN
E-mail: nakachi.takayuki@lab.ntt.co.jp

***Abstract*—Distributed Source Coding (DSC) implements inter source redundancy by encoding correlated information sources, without letting the sources communicate with each other. The DSC theorem states that the optimal rate achieved by joint encoding and decoding can be reached by separate encoding and joint decoding. Video coding based on the DCS theorem is Distributed Video Coding (DVC). DVC is attracting attention as a new paradigm for video compression. While conventional video coding schemes such as H.264/AVC reduce inter-frame redundancy at the encoder, DVC reduces inter-frame redundancy at the decoder by using distributed channel coding. The architecture suits applications that require low-complexity encoders. Examples of such applications include wireless video surveillance and mobile camera phones. In this paper, after describing the DSC theorem, we offer a survey of recent trends and problems in DVC and then introduce our contributions.**

I. Introduction

A new coding paradigm, referred to as Distributed Source Coding (DSC), has emerged based on two Information Theory theorems from the seventies: Slepian-Wolf [1] and Wyner-Ziv [2]. The Slepian-Wolf theorem states that for lossless coding of two or more correlated sources, the optimal rate achieved when performing joint encoding and decoding can be reached by separate encoding and joint decoding. The Wyner-Ziv theorem shows that this still holds for lossy coding.

Distributed Video Coding (DVC) applies the DSC paradigm to video coding. It encodes frames independently, and decodes them jointly, thus shifting the computational burden to the decoder side. This architecture suits applications that require low-complexity encoders. Examples of such applications include wireless video surveillance, sensor networks, wireless PC cameras and mobile camera phones. While the conventional MPEG and ITU-T standards such H.264/AVC [3] imply complex encoders and lightweight decoders, the conventional codec is well-suited for broadcasting-like applications, where a single sender is transmitting data to many receivers.

The known Wyner-Ziv coding techniques include practical schemes with low-complexity encoders [4]-[17]. The DVC paradigm offers a number of major differentiations compared to conventional coding. It is based on a statistical framework with the built-in joint source-channel coding structure. The tasks of motion estimation (ME) and interpolation are shifted to the decoder side.

Current topics of interest in the field of DVC are 1) coding efficiency improvement, 2) decoder complexity reduction, 3) robust transmission 4) multi-view video coding and 5) application beyond coding, etc. In particular, coding efficiency improvement and decoder complexity reduction are challenging problems. According to the Slepian-Wolf theorem and the Wyner-Ziv theorem, DVC can be achieved without a coding performance penalty relative to conventional coding. However, coding efficiency remains a challenging issue for DVC despite the considerable improvements achieved over the last few years. Regarding decoder complexity, a DVC decoder can be an order of magnitude more complex than a conventional video encoder such as H.264/AVC. This is because it (DVC decoder) must not only generate side-information, but also use a belief propagation algorithm to realize channel coding.

In this paper, after describing the DSC theorem, we describe the fundamental framework and prominent features of DVC. We then offer a survey of recent trends and problems and detail our contributions. This paper is organized as follows. Section II reviews the Slepian-Wolf theorem and the Wyner-Ziv theorem. In Sec. III, we briefly describe the fundamental framework and prominent features of DVC. We then survey the recent trends and problems of DVC in Sec IV and introduce our contributions in Sec V. Finally, a conclusion is drawn.

II. Distributed Source Coding

We review the Slepian-Wolf theorem [1] and the Wyner-Ziv theorem [2].

A. Slepian-Wolf theorem

Slepian-Wolf considered the distributed lossless compression of two statistically dependent sources X and Y (Fig. 1(a)). According to Ref. [1], given the distributed source coding problem for i.i.d. sources (X, Y), the achievable rate region is given by

$$R_X \geq H(X|Y) \quad (1)$$
$$R_Y \geq H(X|Y) \quad (2)$$
$$R_X + R_Y \geq H(X, Y). \quad (3)$$

In this case, the rate region is as illustrated in Fig. 2; the lower bound is called the "Slepian-Wolf limit". That is, the Slepian-Wolf theorem shows that correlated sources, (X, Y), can be compressed while requiring less rate mutual information, $I(X; Y)$, without observation of each other.

978-1-4577-2166-3/11/$26.00 ©2011 IEEE

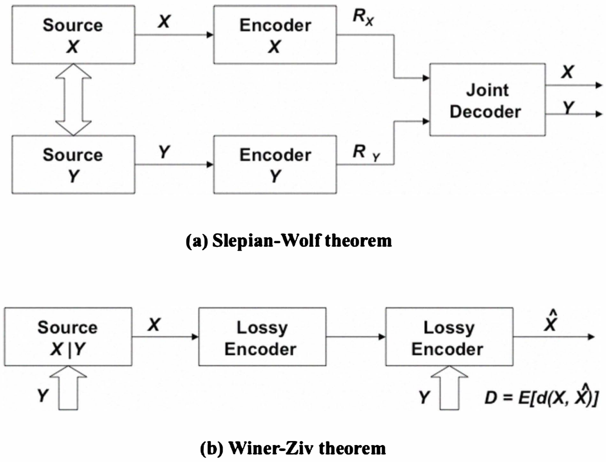


Fig. 1. Distributed compression of two statistically dependent random processes X and Y.

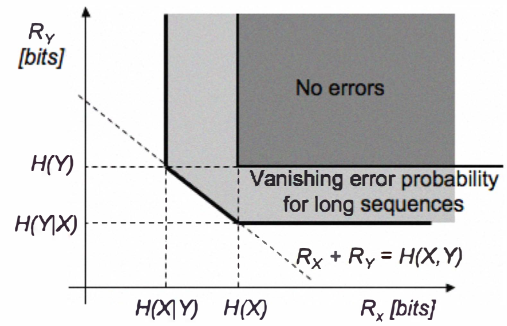


Fig. 2. Slepian-Wolf theorem: achievable rate region for the distributed compression of two statistically dependent i.i.d. sources X and Y.

B. Wyner-Ziv theorem

The Wyner-Ziv theorem expands the Slepian-Wolf theorem to cover lossy compression. The Wyner-Ziv theorem establishes the rate region for the distributed lossy compression of two statistically dependent sources X and Y (Fig. 1(b)). In [2], Wyner and Ziv introduce the rate distortion function with side information (called Wyner-Ziv rate-distortion function) $R_{X|Y}^{WZ}(D)$ which is based on the simple distortion function ($D = E[d(X, \hat{X})]$). The Wyner-Ziv rate-distortion region is given by

$$R_{X|Y}(D) - R_{X|Y}(D)^{WZ} \geq 0. \qquad (4)$$

In DVC systems, the rate region and rate-distortion region match if the side information is available at the encoder. Studies have examined how closely we can approach the Slepian-Wolf limit and the Wyner-Ziv limit with Turbo and LDPC codes.

III. Fundamental Distributed Video Coding (DVC) Framework

This section briefly describes a fundamental framework and prominent features of DVC. DVC applies the DSC paradigm to video coding. Early DVC solutions were proposed by two groups. One is Girod's group at Stanford university [4]-[6], and the other is Ramchandran's group at the University of Berkley [7].

A. Stanford Architecture [4]-[6],

The Stanford WZ video coding architecture was first proposed in 2002 for the pixel domain [4] and later extended to the transform domain [5] [6], where DCT coefficients are WZ coded. The pixel domain DVC architecture is shown in Fig. 3. The video frames are divided into key frames and Wyner-Ziv frames. Key frames are encoded by conventional intra-frame codec such as JPEG. The Wyner-Ziv frames are placed in between key frames, which are encoded independently but decoded jointly. Every pixel in a Wyner-Ziv frame is uniformly quantized with the interval of 2^M. The quantized indices are fed to the Slepian-Wolf encoder with Rate Compatible Punctured Turbo (RCPT) codec.

At the decoder side, the side information $\hat{S}$ can be generated by interpolation or extrapolation of decoded key frames. The side information $\hat{S}$ is used in the turbo decoder, along with the parity bits of the Wyner-Ziv frames requested via a feedback channel, in order to reconstruct the bitplanes, and subsequently the decoded video sequence. In [18], rate-compatible Low-Density Parity-Check Accumulate (LDPCA) codes, which better approach the communication channels capacity, replace the Turbo codes. The decoder combines the side information $\hat{S}$ and received parity bits to recover q'. After reconstruction of q', the decoder reconstructs S' which can be written as $S' = E[S|q', \hat{S}]$.

In the transform domain Wyner-Ziv codec, the RD performance of this scheme is improved by using a transform coding tool to exploit the spatial correlation between neighboring sample values and to compact the block energy into as few transform coefficients as possible.

B. PRISM Architecture [7]

The PRISM (Power-efficient, Robust, hIgh compression Syndrome-based Multimedia coding) architecture is shown in Fig. 4. A new feature of PRISM is the use of multiple candidates for side information. At the encoder, each frame is split into 8×8 blocks which are DCT transformed. This information is used to classify blocks into three encoding classes. The first class corresponds to blocks with very low correlation; they are encoded using traditional coding. The second class consists of blocks that have very high correlation and are merely signaled as skipped. Finally, the remaining blocks, in the third class, are encoded based on syndrome encode principles. The syndrome bits are computed from the least significant bits of the transform coefficients. The lower part of the least significant bit planes is entropy coded. The upper part of the least significant bit planes is coded using a coset channel code.

At the decoder, the skip-class frame blocks can be reconstructed by the colocated blocks in the previously reconstructed frame. The frame blocks in the intra coding class are

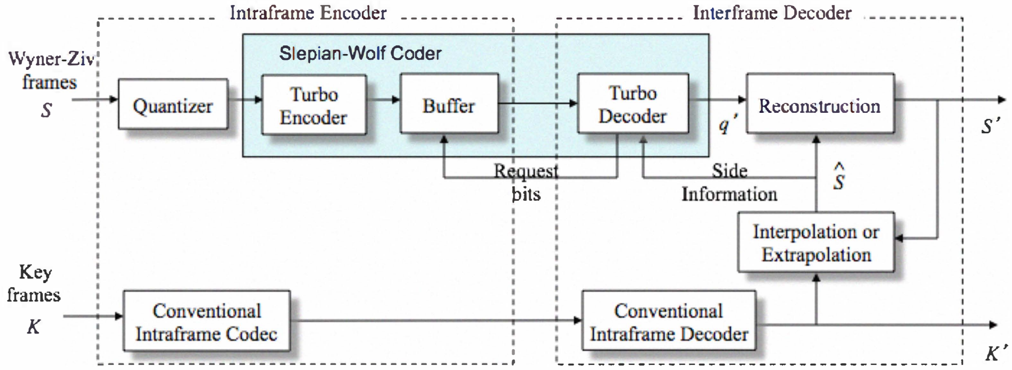


Fig. 3. Stanford Pixel Domain Distributed Video Coding Architecture [4].

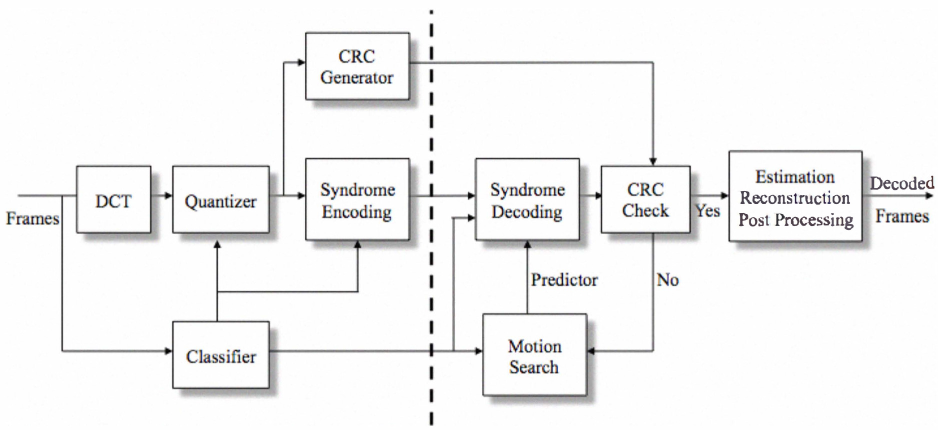


Fig. 4. PRISM Architecture [7].

reconstructed by the traditional decoder. Syndrome encoded blocks are decoded by performing motion estimation by using CRC bits. The previously decoded sequences and multiple candidate's side information are used for motion estimation. The CRC is used as a reliable and unique signature for each block to identify the best candidate predictor. The bit streams are the dequanitized and inversely transformed and scanned to reconstruct the video sequence.

C. Prominent Features of DVC

The DVC paradigm offers a number of major differences from conventional coding. First, it is based on a statistical framework with the built-in joint source-channel coding structure. As it does not rely on joint encoding, content analysis can be performed at the decoder side.

Due to its intrinsic joint source-channel coding framework, DVC is robust to channel errors. DVC provides codec-independent scalability because it does not rely on a prediction loop. Furthermore, DVC does not need a temporal prediction loop, unlike 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 very low complexity encoding.

DVC offers a number of potential advantages which make it well-suited for the aforementioned emerging upstream applications. DVC is well-suited for multi-view coding, i.e. camera sensor networks, by exploiting the correlation between views without requiring inter-camera communications, an important architectural advantage. Furthermore, the DSC principles are useful beyond coding applications. For instance, DSC can be used for data authentication, tampering localization, and secure biometrics.

IV. Recent Trends and Problems

In this section, we review current topics of interest in DVC.

A. Coding Efficiency

According to the Slepian-Wolf theorem and the Wyner-Ziv theorem, DVC 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 the considerable improvements made over the last few years.

Although first described in 2007, the codec developed by the European project DISCOVER [11] remains one of the best performing DVC schemes. DVC codecs now consistently outperform H.264/AVC Intra coding, except for scenes with complex motion. In some cases, e.g. video sequences with

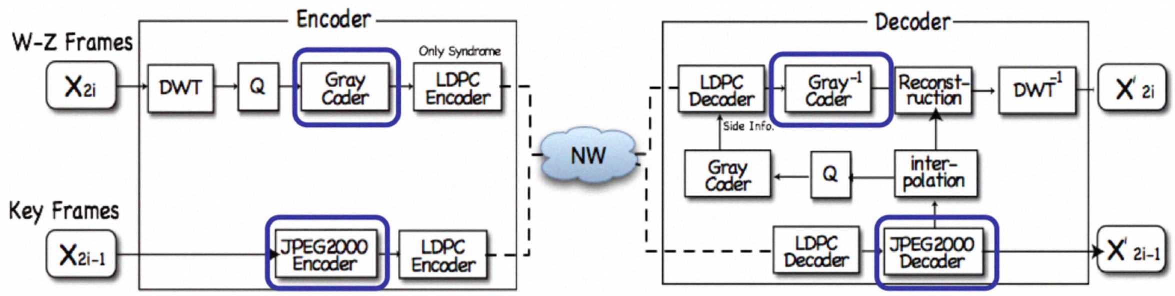


Fig. 5. Wavelet Domain Distributed Video Coding [13].

simple motion structure, DVC can even top H.264/AVC No Motion. Nevertheless, their performance remains generally significantly lower than a full-fledge H.264/AVC codec.

B. Computational Complexity

A DVC decoder can be an order of magnitude more complex than a conventional video encoder such as H.264/AVC. This is because the DVC decoder must generate side-information at the decoder and use a belief propagation algorithm to realize channel coding. If the accuracy of the side information is very high, the time taken for channel decoding can be decreased. In general, however, several iterations are required to converge on a solution.

In [11], it is shown that the DVC decoder is several orders of magnitude more complex in terms of software execution time compared to a conventional H.264/AVC Intra-frame decoder, and about 10-20 times more complex than an H.264/AVC Intra-frame encoder. More research is needed to achieve desirable performance. Optimized decoder implementations on multi-core processors and FPGA's should be specifically considered.

C. Robust Transmission

Distributed source 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. 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 a noise margin 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, e.g., due to packet losses impacting the reference frame.

Distributed video coding is applied to error resilient MPEG-2 video broadcasting in [19], 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; the coarsely quantized video bitstream, obtained from a conventional hybrid video coder, is subjected to Reed-Solomon coding, and only the parity symbols are transmitted.

D. Other Current Topics

With its ability to exploit inter-camera correlation at the decoder side, without communication between cameras, DVC is also well-suited for multi-view video coding [20] where it offers a noteworthy architectural advantage. It is attractive for a number of applications such as stereoscopic video, free viewpoint television, multi-view 3D television, and camera networks for surveillance and monitoring. With the emergence of heterogeneous multimedia networks and the greater variety of client terminals, scalable coding is becoming an attractive feature. WZ enhancement layers can be built upon conventional codec or DVC base layers.

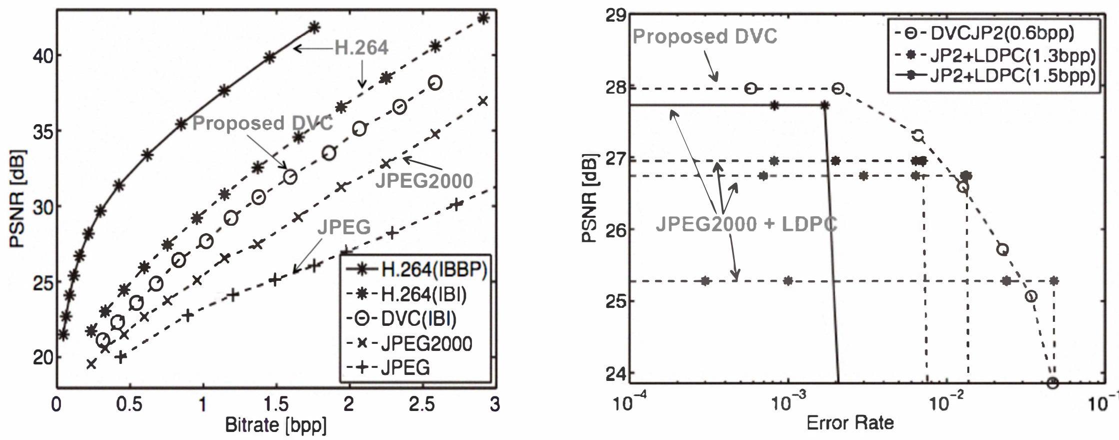
Outside the coding scenario, DSC has also found applications for other domains. For instance, DSC can be used for sensor networks [21], image authentication [22] and secure biometrics [23]. The application of DSC to other domains beyond coding is still a relatively new topic of research. It is not unexpected that further exploration will lead to significant results and opportunities for successful applications. A more detailed survey of DVC studies can be found in [12].

V. Our Contributions

We have studied DVC from several points of view [13]-[17]. This section introduces our contributions as they apply to 1) coding efficiency improvement, 2) decoder complexity reduction, and 3) a new layered coding framework.

A. Coding Efficiency Improvement

1) Wavelet Domain DVC [13]: We proposed a wavelet-based DVC scheme that utilizes the scalability of JPEG 2000. The wavelet domain DVC architecture is shown in Fig. 5. The proposed wavelet-based DVC scheme generates side information without requiring the decompression of all bit streams. This makes decoding far more efficient; the scheme supports scalability with regards to resolution and quality. In addition, we proposed two methods to increase the coding gain of the new DVC scheme. One method is the introduction of Gray code, and the other involves optimum bit allocation.



(a) Rate-Distortion (b) Error Resiliency

Fig. 6. Rate-Distortion Performance and Error Resiliency of the Proposed Wavelet DVC Scheme [13]

We showed that these methods increase overall coding performance by 4.7 [dB]; the proposed scheme offers 1 to 3 [dB] better compression performance than JPEG 2000 as shown in Fig. 6. Moreover, we investigated the scalability and error resiliency of the wavelet-based DVC (right side of Fig. 6). Tests confirmed that the wavelet-based DVC has good scalability and strongly resists bit errors while maintaining high compression performance.

2) An Intra/Inter Mode Decision Scheme based on Wavelet Domain DVC [14] : Rate-distortion theory suggests that generating accurate side information can significantly benefit overall performance. In order to get better side information, average interpolation is used in [13] . However, the correlation noise statistics describing the relationship between the Wyner-Ziv frames and the side information is not always spatially stationary.

In order to overcome this problem, we proposed an intra/inter frame mode decision scheme at the decoder for wavelet domain distributed video coding. In our scheme, the mode decision is successively performed by using not only the past or future key frames but also the decoded lower subband of the current Wyner-Ziv frame. Due to the iterated updates, the mode decision of higher subbands becomes more accurate and higher coding efficiency can be attained. By comparison, mode decision in the conventional intra/inter frame mode decision scheme [10] uses only the past or future frames without any information from the current frame. PSNR Simulations showed that the new scheme gives a coding gain about 1[dB] over the wavelet-based DVC without mode decision.

B. Decoder Complexity Reduction [15]

A DVC decoder can be an order of magnitude more complex than a conventional video encoder such as H.264/AVC [3] . Its use of the belief propagation algorithm means that DVC systems suffer from excessive decoding times. Our solution is to propose a parallelized DVC scheme that treats each bitplane independently. The simple application of parallelization would cause low encoding efficiency since parallelized systems can't use additional side information for the decoding of subsequent bitplanes. There, we proposed an information index assignment method to estimate bit probability as accurately as possible. For the parallelized DVC scheme we investigated the relationship between the bit probabilities of each pixel. Moreover, we proposed a parity rate estimation method for the parallelized DVC scheme that uses the derived bit probability. Therefore, the proposed parallelized DVC system can perform rate control at the encoder in a highly efficient manner.

Simulation results show that the proposed method can reduce the decoding time by up to 30-35 [%] with only slight parallelization loss. Moreover, we proposed a parity rate estimation method for the parallelized DVC system. Using this estimation method, the parallelized DVC can perform rate control at the encoder side with quite accurate bitrates.

C. A New Layered Coding Framework [16]

We proposed a new scheme of layered video coding for real-time streaming and lossless archiving. The layered video coding proposal provides delay scalability which is not offered by conventional scalable coding methods. Motion JPEG2000 is used as base layer coding for real-time streaming and a Slepian-Wolf coder is used as an enhancement layer coder to enhance lossless coding performance.

The design concept of using the Slepian-Wolf technique is depicted in Fig. 8. The computational burden is shifted to the network side; this enhances the coding efficiency since it allows long delays. By using this scheme, the base layer video can be transmitted with low-delay for real-time streaming and the enhancement layer signal can be efficiently compressed by using inter-frame redundancy.

VI. Conclusion

In this paper, we have presented a review of distributed source coding (DSC). We described the fundamental framework and prominent features of distributed video coding

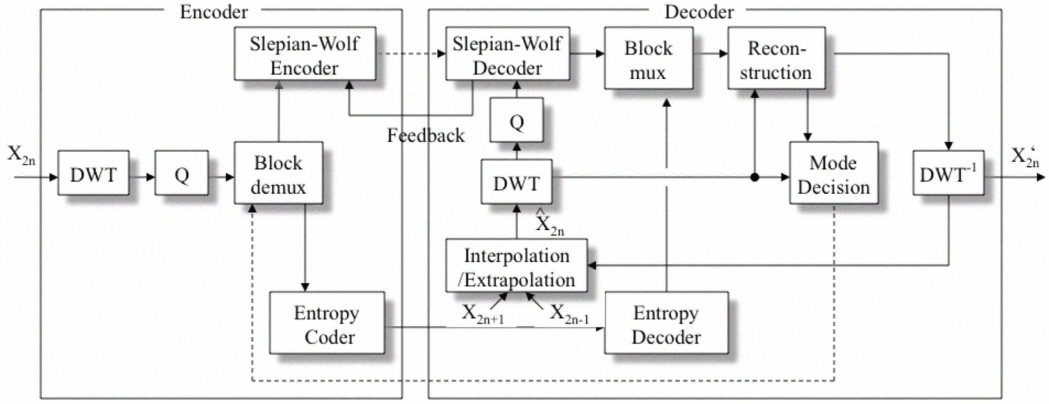


Fig. 7. Intra/inter frame mode decision scheme in wavelet domain distributed video coding [14]

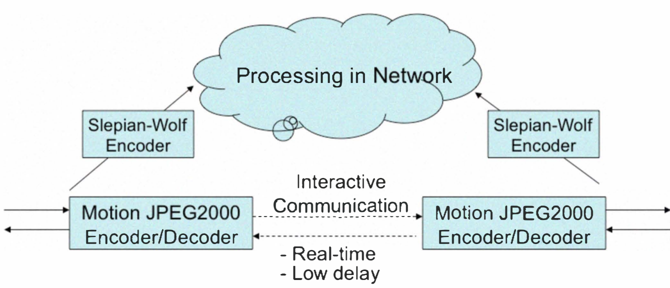


Fig. 8. 1) Real-time interactive communication using JPEG2000 and 2) residual signal coding using Slepian-Wolf encoder for lossless archiving [16].

(DVC), a key DSC application. A survey was provided of recent trends and problems and our contribution were introduced. This new video coding paradigm (DVC) shifts shift complexity from the encoder to the decoder. The DVC algorithm can best the traditional video coding algorithm in terms of complexity encoding, robustness to errors. Moreover, DVC will open new areas in terms of emerging applications such as multi-view video coding, sensor networks, image authentication and secure biometrics, etc.

References

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

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

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

[4] A. Aaron, R. Zhang, and B. Girod, "Wyner-Ziv coding of motion video," *Proc. Asilomar Conference on Signals and Systems*, Pacific Grove, CA, Nov. 2002.

[5] A. Aaron, S. Rane, E. Setton and B. Girod, "Transform-domain Wyner-Ziv codec for video," *VCIP-2004*, San Jose, CA, Jan. 2004.

[6] A. Aaron, S. Rane and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," *IEEE ICIP-2004*, Singapore, Oct. 2004.

[7] R.Puri and K.Ramchandran, "PRISM:a new robust video architecture based on distributed compression principles," *Allerton Conf. Communicationm Control, and Computing*, Allerton, IL, 2002

[8] J. Ascenso, C. Brites, and F. Pereira, "Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding," *5th EURASIP*, Slovak Republic, July 2005.

[9] J. Ascenso, C. Brites, C, and F. Pereira, "Motion compensated refinement for low complexity pixel based distributed video coding," *IEEE AVSS 2005*, pp. 593 - 598, Sept. 2005.

[10] M. Tagliasacchi, A. Trapanese, S. Tubaro, J. Ascenso, C. Brites, F. Pereira, "Intra mode decision based on spatio-temporal cues in pixel domain wyner-ziv video coding," *IEEE ICASSP2006*, vol. 2, pp. II57-II60, May 2006

[11] Discover Project, http://www.discoverdvc.org/

[12] F. Dufaux, W. Gao, S. Tubaro, A. Vetro, "Distributed video coding: trends and perspectives," *EURASIP Journal on Image and Video Processing 2010*, R2010-026.

[13] Y. Tonomura, T. Nakachi and T. Fujii, "Distributed video coding using JPEG 2000 coding scheme," *IEICE Trans. on Fundamentals,* E90-A, pp. 581-589, March 2007.

[14] T. Nakachi, Y. Tonomura and T. Fujii, "Intra/inter frame mode decision scheme in wavelet domain distributed video coding, *IEEE ISIVC2008*, paper no. 171, 2008.

[15] Y. Tonomura, T. Nakachi and T. Fujii and H. Kiya, "Parallel processing of distributed video coding to reduce decoding time," IEICE Trans. on Fundamentals, vol. E92-A, no. 10, pp. 2463-2470, Oct. 2009.

[16] T. Nakachi, Y. Tonomura and T. Fujii, "Delay constrained layered video coding for real-time streaming and lossless archive," *2011 IEEE DSP/SPE Workshop,* pp. 134-139, Jan. 2011.

[17] R. Miyazawa , S. Kudoh , T. Nakachi and N. Hamada, "Coding efficiency improvement on distributed video coding utilizing super resolution approach," *IEEE ISPACS2011.*

[18] D. Varodayan, A. Aaron and B. Girod, "Rate-adaptive codes for distributed source coding," *EURASIP Signal Processing Journal, Special Section on Distributed Source Coding*, vol. 86, no. 11, pp. 3123-3130, November 2006.

[19] S. Rane, A. Aaron, B. Girod, "Systematic lossy forward error protection for error resilient digital video broadcasting—a Wyner-Ziv coding approach," *IEEE ICIP2004*, October 2004.

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

[21] S. S. Pradhan, J. Kusuma, K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Process. Mag.* 19(2), 51-60, 2002.

[22] Y.-C. Lin, D. Varodayan, and B. Girod, "Image authentication based on distributed source coding", *IEEE ICIP2007*, TX, October 2007.

[23] S.C. Draper, A. Khisti, E. Martinian, A. Vetro, J.S. Yedidia, "Using distributed source coding to secure fingerprint biometrics", *IEEE CASSP2007*, Honolulu, HI, April 2007.