Published in the Proceedings - 2009 2nd IEEE International Conference on Computer Science and Information Technology, ICCSIT 2009
2009, Article number 5234835, Pages 130-135
DOI: http://dx.doi.org/10.1109/ICCSIT.2009.5234835

Recent Advances in Distributed Video Coding

Suvendu Rup, Ratnakar Dash, Niranjan Kumar Ray and Banshidhar Majhi
Department of Computer Science and Engineering
National Institute of Technology Rourkela
Rourkela–769 008, Orissa, India
{suvendurup, ratnakar.dash, rayniranjan, bmnitrourkela}@gmail.com

Abstract—**Distributed Video coding (DVC) is a new coding paradigm for video compression mainly based on two major information theory results, Slepian and Wolf's and Wyner-Ziv theorems. In some major ap- plications such as wireless low power surveillance and multimedia sensor networks, wireless PC cameras and mobile camera phones, it is really a challenge for traditional video coding architecture. For both encoder and decoder, it is necessary to have a low power consumption for some cases. So low power and low complexity encoder are essential to satisfy those requirements. This paper reviews some recent developments of dis- tributed video coding. Slepian and Wolf's coding is the lossless source coding and Wyner-Ziv coding is a lossy compression with receiver side information. It has been observed that in most of the traditional video coding algorithm such as MPEG,H.263+ or H.264 the encoder has com- plex computational operation rather than the decoder. The distributed video coding is a new kind of video coding which allows a low complexity video encoding where the major part of computational burden is shifted to decoder. DVC is mainly applicable to two areas that is low complex- ity video coding and robust video coding. The dierent techniques are analysed based on certain parameters such as compression rate, decoding and motion compensation.**

Index Terms—**Distributed Video Coding, Low complexity Video coding, Wyner-Ziv coding**

I. Introduction

Now a days, the digital video coding paradigm represented by the standardization of ITU-T or MPEG use discrete cosine transform (DCT) and interframe predictive coding [1], [2]. Recently, different video compression standards perform interframe predictive coding to exploit the similarities between successive frames. In order to explore those temporal correlations the encoder is five to ten times more complex than decoder, mainly due to motion estimation tasks. However, some applications may require both the encoder and decoder to be complex. In some cases, for example a mobile wireless camera uploading video to a fixed base station, we need a low complexity encoder and possibly a high expense of decoder because compression must be carried out at the camera where memory and computation are scarce. So we need an asymmetric video compression scheme as shown in figure 1 in which process is reversed that is low complexity encoder [3] high complexity decoder, where individual frames encoded independently (intraframe encoding) but decoded conditionally(interframe decoding). For most of the practical video coding there is a major problem asociated with Slepian and Wolf theorem [4] because it is a lossless coding. This comes from the fact that lossless coding achieve small compression factors since it does not eliminate the irrelevant video information. Later Wyner-Ziv analysed the problem for lossy compression case. According to Wyner-Ziv theorem [5], [6], [7] under certain conditions there is no coding efficiency loss. So together with Slepian-Wolf and Wyner-Ziv theorem it is possible to encode separately and decode jointly two statistical dependent variable. So this concept when applied to a video coding it is known as distributed video coding or Wyner-Ziv video coding [8], [9]. The information theory concept of DVC was established long years back but their practical implementation has been proposed only in recent years. The objective of the recent DVC schemes is to achieve efficient and low complexity encoder video compression and efficiently evaluate the use of channel codes [10] for DVC.

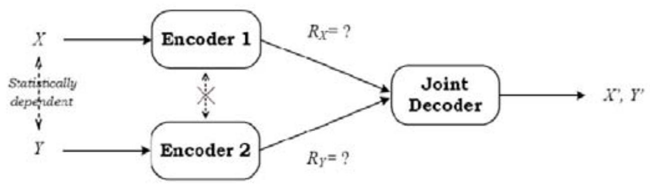

Fig. 1. Distributed compression of two statistically dependent discrete random sequences X, Y, independently and idenitically distributed (i.i.d)

This paper is organized as follows. Section 2 gives for understanding the foundation of DVC. In section 3 we review the recent advances in distributed video coding to till date. Section 4 illustrates results and discussion with some numerical performances of the existing methods and the conclusion is given in section 5.

II. Foundation of distributed video coding

A. Slepian-Wolf lossless coding theorem

Let X and Y are two statistically dependent independent identically distributed (i.i.d) sequences. These sequences are encoded independently with bit rates R_X and R_Y but jointly decoded . The expressions for rate combinations according to Slepian-Wolf theorem are as follows.

$$R_X \geq H\left(X|Y\right) \quad (1)$$

$$R_Y \geq H\left(Y|X\right) \quad (2)$$

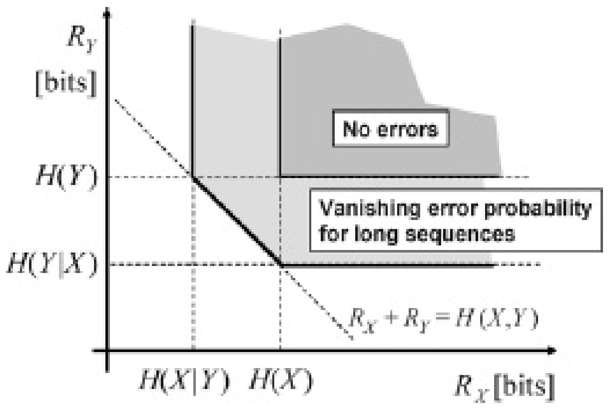

Fig. 2. Achievable rate region following the Slepian-Wolf theorem

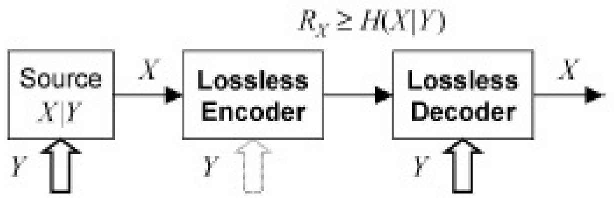

Fig. 3. Compression of a sequence of random symbols X using statistical related side information Y

$$R_X + R_Y \geq H\left(X, Y\right) \tag{3}$$

Where $H(X|Y)$ and $H(Y|X)$ are conditional entropy . The total sum rates $R_X + R_Y$ can achieve the joint entropy which is maximum. Another intersting feature of Slepian-Wolf coding is that it is well suited to channel coding. The achievable rate region is shown in figure. 2

B. Wyner-Ziv coding theorem

Wyner and Ziv extended the idea of Slepian and Wolf's theorem in 1976. They extended their work to establish information theoritic bounds for lossy compression. The theorem states that, if X and Y are two statistically dependent sequence and X is encoded independently without the access of side information Y (see figure. 3), the sequence X can be reconstructed with a distortion below D.

A distortion is written as $D = E[d\left(X, \hat{X}\right)]$ is acceptable as shown in figure. 4. When the encoder does not have any idea about Y, a rate loss $R^{WZ}_{X|Y}(D) - R_{X|Y}(D) >= 0$ is achieved. However, in case of mean squared error distortion and Gausian memory less source [5] a rateloss of $R^{WZ(D)}_{X|Y} - R_{X|Y}(D) = 0$ is achieved [11]. The basic architecture of Wyner-Ziv coding is shown in figure. 5.

III. Recent advances in distributed video coding

Recently most of the practical DVC solutions have been proposed by two groups Bernd Girod's group at stanford university and Ramchandran's group at the university of Berkley, carlifornia. This section briefly describes the advances of DVC and their performances.

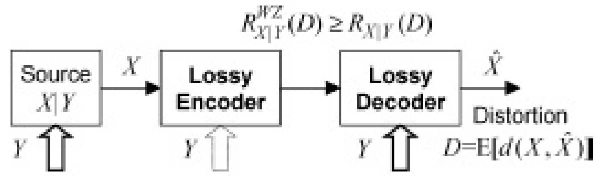

Fig. 4. Lossy Compression of a sequence X using statistically related side information Y

A. Stanford's low complexity video coding algorithm

Girod's group at stanford university have proposed several solutions related to distributed video coding. The following are the outline of those solutions.

Pixel- Domain coding solution:

Among the various solutions proposed by Girod's group the pixel domain coding solution is the most simplest among them (see figure. 6). In their proposed scheme [12], [13] the video frames are divided into key frames and Wyner-Ziv frame. The Wyner-Ziv frames are placed in between key frames which are encoded independently but decoded jointly. The scheme is simplest because neither DCT nor motion estimation and inverse discrete cosine transform are required . Every pixel in a Wyner-Ziv frame is uniformly quantized with 2^M intervals. The quantized indices is fed to Slepian-Wolf encoder with Rate Compatible Punctured Turbo (RCPT) code [14] . At the decoder side the side information $\hat{S}$ can be genereted by interpolation or extrapolation from previously key frames or from previously Wyner-Ziv frame .

At the decoder side, the decoder combines side information $\hat{S}$ and received parity bits to recover $\acute{q}$. After reconstruction of $\acute{q}$, the decoder reconstruct the $\acute{S}$ which can be written as $\acute{S} = E[S|\acute{q}, \hat{S}]$.

Transform domain coding solution:

Unlike the pixel domain coding, in this scheme neither motion estimation nor motion compensation are needed at the encoder. Here a blockwise DCT is performed in a Wyner-Ziv frame. The DCT coefficients are independently quantized and compressed by Slepian-Wolf turbo code [15], [16] . The side information can be generated from the previously reconstructed frame with or without motion compensation. This scheme has a higher encoder compexity than the pixel domain system.

Joint decoding and motion estimation solution:

In order to achieve high compression efficiency, motion estimation has to performed at the decoder side. In this scheme, the frames in a video are organised into group of picture(GOP)at the encoder end as shown in figure. 7. Each GOP is consist of key frames and Wyner-Ziv frames. For a given Wyner-Ziv frame a 4X4 discrete cosine transform is applied. The transformed coefficient band is then uniformly quantized, then bit plane extraction is performed. Each bit plane is then independently fed into turbo encoder. Some of the robust hash code word [17], [2] are send to the decoder for

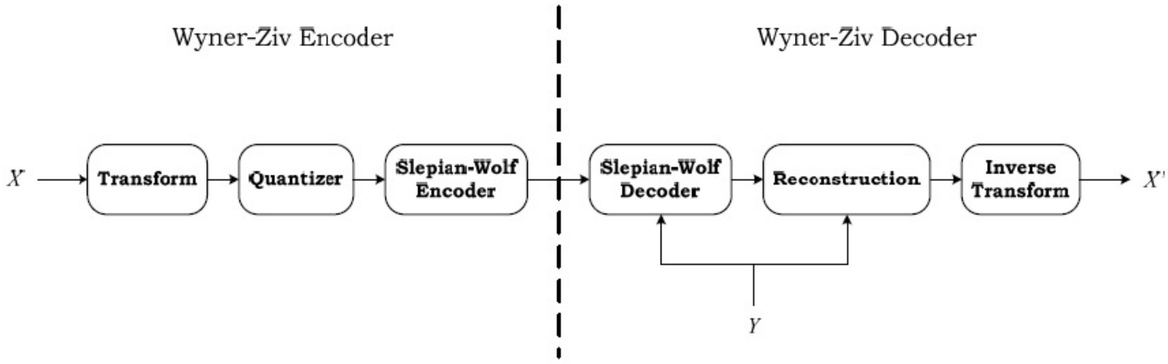

Fig. 5. Basic Architecture of Wyner-Ziv coding

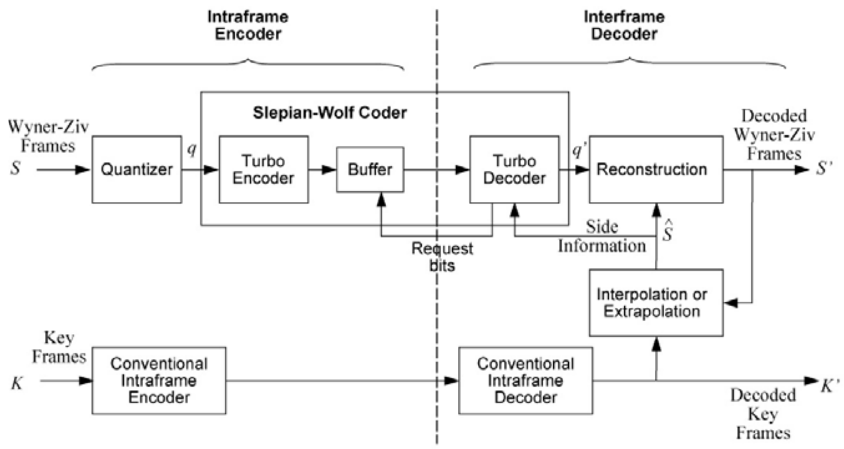

Fig. 6. Pixel domain encoding for low complexity encoder and decoder

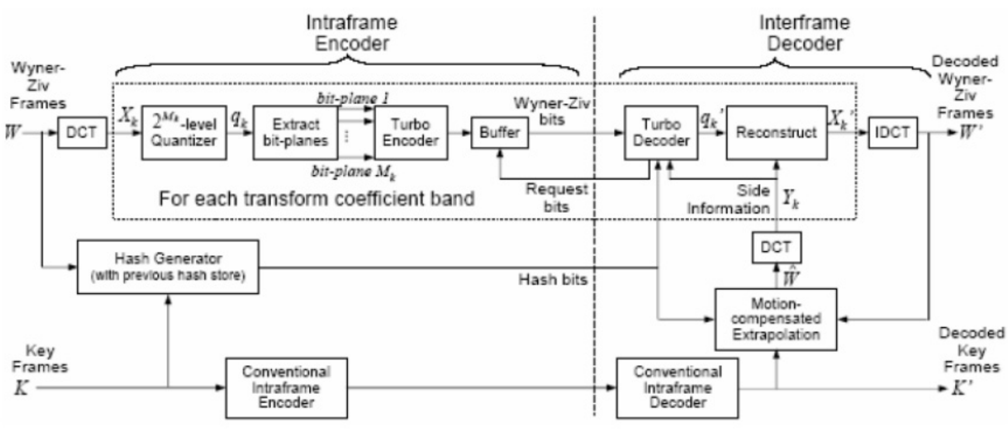

Fig. 7. Joint decoding and motion estimation for low complexity encoder and decoder

motion estimation. The key frame is encoded by conventional coding such as MPEG or H.26X. The encoded Wyner-Ziv bits are stored in buffer and only a minimum number of bits are sent on a request by the decoder.

The decoder receives three bit streams namely key frame bits, hash bits and Wyner-Ziv bits. The key frames are decoded using the conventional decoding and the hash bit is used for motion compensation. The Wyner-Ziv bit stream is decoded using turbo decoder. Then requantization followed by reconstruction of the Wyner-Ziv frame by motion compensation method. The reconstructed frame is then inversly transformed to obtain $\hat{W}$ which is the estimation of Wyner-Ziv frame.

B. Stanford robust video coding solution

In another video coding solution, a major application of systematic lossy source channel coding to error prone digital transmission has been proposed. It is also called systematic lossy forward error protection (SLEP) [18].

This scheme uses a hybrid video codec. At the encoder end, first bit stream i.e. key frame is encoded using traditional MPEG or H.26X and the second bit stream which is a Wyner-Ziv frame is encoded with coarse quantization. The quantized bits are applied to the Reed-Solomon encoder [19] and only the RS parity symbols are transmitted to the receiver.

At the decoder side, two bit streams are obtained and for the first bit stream, the traditional decoding is applied.When error occurs during the transmission, the Wyner-Ziv frame along with the side information from conventional decoder is used to reconstruct the sequence.

C. Berkeley robust video coding solution

Power efficient robust high compression syndrome based multimedia coding(PRISM):

This solution was proposed by Ramchandran's group at the

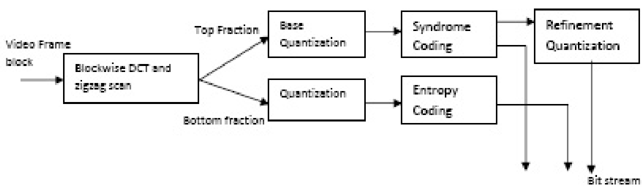

Fig. 8. Encoder architecture of PRISM

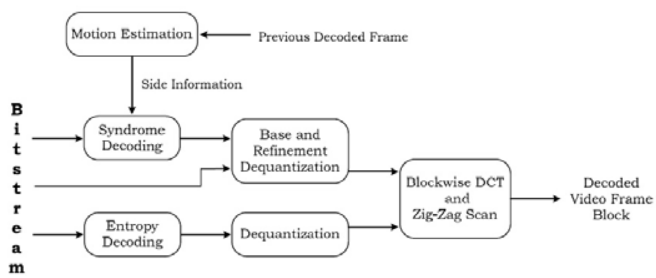

Fig. 9. Decoder architecture of PRISM

university of Barkley [20], [21] under the name of PRISM. It combines the features of intraframe coding with inter frame-coding compression efficiency.This architecture uses Wyner-Ziv coding but the generation of side information is different from other schemes. A new features of PRISM is the use of multiple candidates for side information.

At the encoder end, the video frame is first divided into 8X8 or 16X16 block.In the classification stage it is necessary to decide what kind of encoding is well suited for each block in the current frame. There are three classes of coding uesd at the encoder end such as no coding(skip class),traditional coding(intra coding class)and syndrome coding(syndrome coding class). Fig. 8 shows a PRISM encoder in which the syndrome coding [22] is the most essential part. For a input video frame block wise DCT is performed followed by zigzag scan. High frequency components are quanitzed coded with entropy encoder. The coarse quantization followed by syndrome coding is applied to the DC component.The syndrome encoded bits are further quantized and sent to the decoder. A cyclic redundancy check (CRC) of the base quantized transform coefficients is also computed and transmitted to help the decoder for motion estimation.

At the decoder end (see figure. 9) the frame blocks in skip class can be reconstructed by the colocated blocks in the previously reconstructed frame.The frame blocks in the intra coding class are reconstructed by the traditional decoder. Syndrome encoded blocks are decoded by performing motion estimation by using CRC bits. The previously decoded sequences and multiple candidates side information genaration are used for calculating motion estimation. The CRC is used as a reliable and unique signature for each block to identify the best candidate predictor.Then the bit streams are dequanitized and inversly transformed and scanned to reconsruct the video sequence.

IV. DISCOVER

Distributed coding for video services(DISCOVER) [23], [24], [2] is a new video coding scheme which has a strong potential of new applications, targeting new advances in coding efficiency, error resillence and scalability. The basic architecture of DISCOVER Codec is shown in figure. 10

At the encoder side the video is split into two parts. The first set of frames called key frame are encoded with conventional H.264/AVC encoder. The remaining frames known as Wyner-Ziv frames which undergoes a block based transform. Then the transformed coefficeints are quantized. A band arrangement of the coefficients are made in which each band contains the coefficient associated to same frequency with different blocks. Then the coefficients are ordered bit plane by bit plane and fed into a systematic channel encoder. The bits at the encoder output are stored in a buffer and transmitted to the decoder which can request more bits for better result using the feedback channel.

The decoder is more complex due to the exploitation of temporal correlation (Motion estimation). The conventional encoded key frames are first decoded by H.264/AVC decoder.

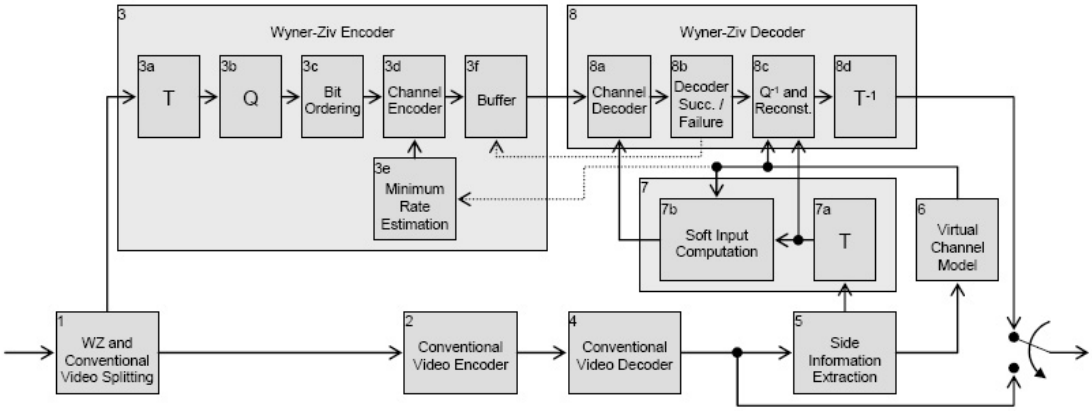

Fig. 10. Basic architecture of DISCOVER Codec

These frames are then used for construction of the side information. For a given Wyner-Ziv frame a motion compensated interpolation between the two closest references frame is performed. The difference between the original Wyner-Ziv frame and corresponding side information can be considered as correlation noise in a virtual channel. A laplacian model [8], [25] is used to obtain a good approximation of the residual. The side information is then transformed and an estimation of coefficients of Wyner-Ziv frame is obtained. Soft values for the information bits are computed from these coefficients, taking into account the stastical modelling of the virtual noise. These soft values are used for reconstruction at the decoder side. At the decoder end if the received parity bits are not enough then more parity bits are requested using feedback channel for better result. Then, reverse transform is applied to the transformed coefficeints. The Wyner-Ziv frame and H.264 frames are multiplexed to get the video sequence.

V. Results and discussion

It was observed that pixel domain Wyner-Ziv encoding performs 2-5dB better than traditional intraframe coding shown in figure. 11. However, there is still a significant gap toward H.263+ interframe coding. When MC-I (Motion Compensation interpolation) is used, the transform domain codec is only 0.5 dB better than pixel domain codec. With less reliable MC-E (Motion Compensation extrapolation), using a transform before encoding results in a 2 to 2.5 dB improvement which can be obserbved from the figure. 12 In joint decoding and motion estimation solution, QCIF video sequences at 10 frames per second with different GOP Length it performs a coding gain upto 9dB over traditional DCT based intrame coding. But for the same GOP size the rate distortion performance is 1 to 4 dB worse than that of H.263 + interframe coding. It is clearly shown in figure. 13 Results DISCOVER: The performance analysis of DISCOVER with other schemes like H.263+(Intra),H.264/AVC(Intra) and H.264/AVC(no motion) is described in below. QCIF video sequence with GOP length 2,DISCOVER performs 8dB gains than H.263 intra and 0.5 to 3 dB for H.264/AVC intra. So it is concluded that

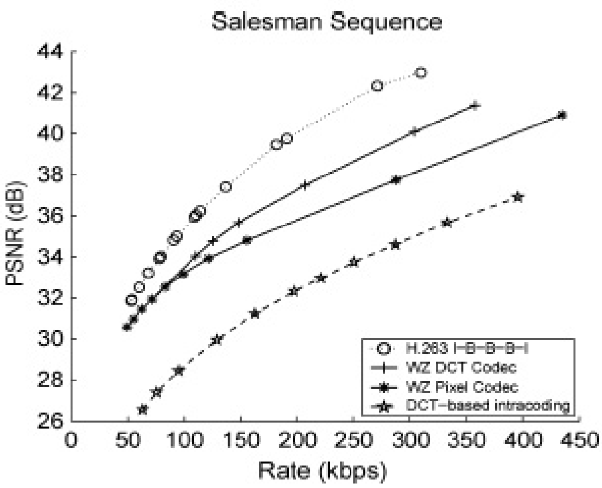

Fig. 11. Rate distortion performance of pixel domain Wyner-Ziv encoding

Fig. 12. (a)DCT-based intraframe coding (b) DCT-domain Wyner-Ziv coding, using MC-I (c) H.263+ interframe coding (B frames)

DISCOVER codec can exploit the temporal corelation in an efficient way while using a rather simpler encoder and still can be competative when compared to H.264/AVC intra coder. Some performance losses are observed when compared with H.264/AVC with no motion codec.

VI. Conclusions

In this paper, we have presented a review on Distributed Video coding based on Slepian and Wolf's and Wyner-Ziv

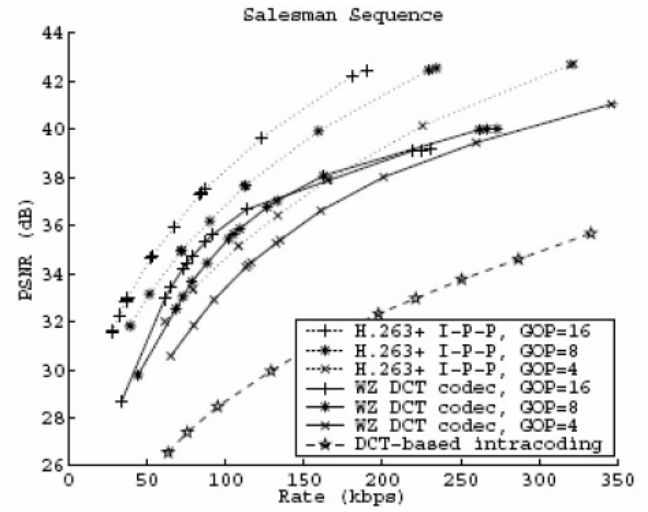

Fig. 13. Rate distortion performance of joint decoding and motion estimation encoding

theorem with side information at the decoder. This new video coding paradigm enables to explore the signal statistics, partially or totally, at the decoder. In other words, DVC enables to shift complexity from the encoder to the decoder.Distributed video coding algorithm will beat the traditional video coding algorithm in terms of low complexity encoding,robustness to errors and rate distortion performance. So DVC is an open area for emerging applications such as wireless video,sensor networks,disposable camera etc. Lots of future directions such as moving from lossless to lossy key frames, more accurate motion interpolation and extrapolation at the decoder, rate control at the encoder and channel codes can be exploited better.

REFERENCES

[1] C.Brites, "Advances in distributed video coding, instituto superior tecnico," MS Thesis, DEC 2005.

[2] N. Khan and N. Khalid, "Devlopements in distributed video coding," in *IWSSIP*.

[3] W.A.R.J.Weerakkody, W. Fernando, and A. Adikari, "Unidirectional distributed video coding for low cost video encoding," *IEEE Transaction on Consumer electronics–II*, vol. 53, no. 2, pp. 788–795, May 2007.

[4] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, pp. 471–480, July 1973.

[5] 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.

[6] A. Wyner, "On source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 21, no. 3, pp. 294–300, May 1975.

[7] A.Aaron, R.Zhang, and B.Girod, "Wyner-ziv coding of video," in *IEEE Data compression conference*, ser. IDCC-2003, 2003, pp. 93–102.

[8] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," in *Proc. IEEE*, vol. 93, no. 1, January 2005, p. 7183.

[9] R.Puri, A. Majumdar, P. Ishwar, and K. Ramchandran, "Distributed video coding in wireless sensor networks," IEEE signal Processing Magazine, pp. 94–106, July 2006.

[10] J. C. S. Pradhan and K. Ramchandran, "Duality between source coding and channel coding and its extension to the side information case," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1181–1203, May 2003.

[11] R.Zamir, "The rate loss in the Wyner Ziv problem," *IEEE Trans.Inf. Theory*, vol. 42, no. 6, p. 20732084, November 1996.

[12] A.Aaron, E.Setton, and B.Girod, "Toward practical Wyner-ziv coding of video," in *IEEE International Conference on Image processing*, ser. ICIP-2003, Barcelona, Spain, 2003.

[13] A.Aaron, R.Zhang, and B.Girod, "Wyner-Ziv coding for motion video," in *Proceedings of Asilomar Conf. Signals and Systems*, ser. ACSS-2002, vol. 4, May 2002, pp. 2002–2374.

[14] D. Rowitch and L. Milstein, "On the performance of hybrid fec/arq systems using rate compatible punctured turbo codes," *IEEE Trans. Commun.*, vol. 48, no. 6, p. 948959, June 2000.

[15] A.Aaron, S.Rane, E.Setton, and B.Girod, "Transform-domain wyner-ziv codec for video," in *SPIE Visual Communications and Image processing Conf.*, ser. VCIP-2004, San jose, CA, 2004.

[16] D.Rebollo-Monedero, A.Aaron, and B.Girod, "Transforms for high rate distributed source coding," in *AilmoreConf.Signals,Systems and Computers*, 2003.

[17] 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*, ser. ICIP-2004, Singapore, 2004.

[18] S.Rane, A.Aaron, and B.Girod, "Systematic lossy forward error protection for error resilient digital video broadcasting," in *SPIE Visual Communications and Image processing Conf.*, ser. VCIP-2004, San jose, 2004.

[19] A. Aaron, S. Rane, D. Rebollo-Monedero, and b. . I. I. C. I. P. y. . . s. I.-. B. Girod", title ="Systematic lossy forward error protection for video waveforms".

[20] R. Puri, A. Manjumdar, 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.

[21] R. Puri and K. Ramchandran, "PRISM: A reversed multimedia coding paradigm," in *IEEE International Conference on Image processing*, ser. ICIP-2003, Barcelona, Spain, 2003.

[22] S. Pradhan and K.Ramchandran, "Distributed source coding using syndrome(DISCUS)," in *IEEE Data Compression Conference*, ser. IDCC-1999, Snowbird,UT, 1999.

[23] X.Artigas, J.ascenso, M.Dalai, D.Kubasov, and M.quaret, "The discover codec: Architecture, techniques and evaluation," Picture Coding Symposium, 2007.

[24] "www.discoverdvc.org."

[25] C.Brites, J. Ascenso, and F. Pereira, "Studying temporal correlation noise modeling for pixel based wyner-ziv video coding," in *IEEE International Conference on Image Processing*, ser. ICIP-2006, Atlanta, 2006.

[26] A. Aaron and B. Girod, "Compression with side information using turbo codes," in *IEEE Data Compression Conference*, ser. IDCC-2002, 2002.