Efficient Compression of Video using Distributed Source Coding

Aparna P
Research Scholar
Dept. of Electronics and Communication Engg.
National Institute of Technology Karnataka, Surathkal,
Srinivasnagar, Mangalore 575025, INDIA
Email: aparnadinesh@rediffmail.com

Sumam David S. *SMIEEE*
Professor
Dept. of Electronics and Communication Engg.
National Institute of Technology Karnataka, Surathkal,
Srinivasnagar, Mangalore 575025, INDIA
Email: sumam@ieee.org

Abstract—**Current video standards based on the principle of predictive coding exploits temporal redundancy between the adjacent frames. Although such schemes promise good rate-distortion performance, encoder complexity is much higher. This paper looks into a new video paradigm based on principle of distributed source coding which supports a simple low complexity encoder by shifting the complexity to the decoder. Such a codec attempts to achieve the compression efficiency of interframe coding. This paper presents some simulation results of a video codec working on the principles of distributed source coding.**

Key Words: *Interframe coding, Intraframe coding*

I. INTRODUCTION

Tremendous development in advanced technologies like VLSI circuits and in computing capabilities have made many complex video applications like videophony, videoconferencing, HDTV, DVD etc, possible. With increasingly complex video services such as 3-D movies, 3-D games, high video quality as in HDTV, it is necessary to have advanced video and image compression techniques. Current video standards like ISO MPEG and ITU-H.26X schemes have made an effort in accomplishing the enhanced compression performance needs and providing a network friendly video representation addressing *conversational* applications (video telephony) and *non conversational* applications like storage, broadcast or streaming [1]. However with the advancement of technologies, demand for video applications is also increasing. Such conventional video standards does not match the needs of uplink friendly applications like mobile video cameras, wireless PC cameras, disposable video cameras, network camcorders, wireless video sensor networks etc. This is because conventional video codecs exploits the source statistics at the encoder, making the encoder complex and computationally intensive. Computational complexity of the video encoder is dominated by motion-compensated prediction operation required to strip the temporal redundancy existing between the adjacent video frames.

Based on the information theoretic bounds established in 1970's by Slepian-Wolf [2] for distributed lossless coding and by Wyner-Ziv [3] for lossy coding with decoder side information, it is seen that efficient compression can also be achieved by exploiting source statistics partially or wholly at the decoder. Video compression schemes that build upon these theorems are referred as distributed video coding and these suit well for uplink friendly video applications. Distributed coding exploits the source statistics only at the decoder, thus interchanging the traditional balance of complex encoder and simple decoder. Hence the encoder of such a video codec is very simple, at the expense of a more complex decoder. Such algorithms hold great promise for new generation mobile video cameras and wireless sensor networks [4], [5].

II. FUNDAMENTAL THEORIES OF DISTRIBUTED SOURCE CODING

A. Slepian-Wolf Theorem for Lossless Distributed coding

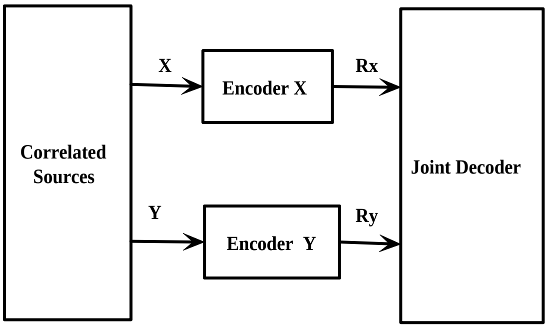

Fig. 1. Compression of source data by separate encoder but decoded by joint decoder

Consider two correlated information sequences X and Y as shown in Fig.1. Encoder of each source is constrained to operate without the knowledge of the other source while the decoder has access to both encoded binary message streams. The problem that Slepian-Wolf theorem [2] addresses is to determine the minimum number of bits per source character required for encoding the message stream in order to ensure accurate reconstruction at the decoder.

Considering separate encoder and decoder for X and Y, the rate required is $R_X \geq H(X)$ and $R_Y \geq H(Y)$ where $H(X)$ and $H(Y)$ represents the entropies of X and Y respectively. Slepian-Wolf [2] showed that good compression can be achieved with joint decoding but separate encoding.

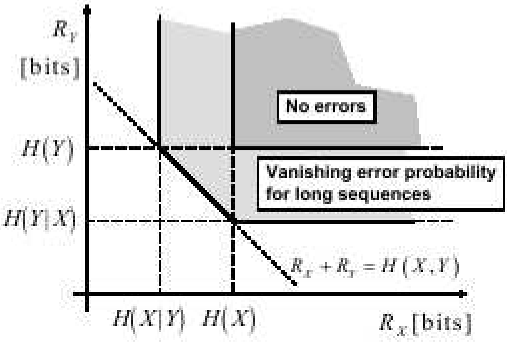

Fig. 2. Admissible Rate Region [4]

For doing this an admissible rate region is defined by the equations:

$$R_X + R_Y \geq H(X, Y) \tag{1}$$

$$R_X \geq H(X/Y), R_Y \geq H(Y) \tag{2}$$

$$R_X \geq H(X), R_Y \geq H(Y/X) \tag{3}$$

as shown in Fig.2.

Thus Slepian-Wolf showed that (1) is the necessary condition and (2) or (3) are the sufficient conditions required to encode the data in case of joint decoding.

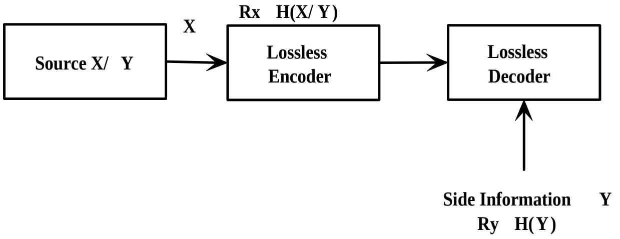

Fig. 3. Lossless Decoder with Side Information

With the above result as the base, we can consider the distributed coding with side information at the decoder as shown in the Fig.3. Let X be the source data that is statistically dependent on the side information Y. Side information Y is separately encoded at a rate $R_Y \geq H(Y)$ and is available only at the decoder. Thus as seen from Fig.2 X can be encoded at a rate $R_X \geq H(X/Y)$.

B. Wyner-Ziv Rate Distortion Theory

Aaron Wyner and Jacob Ziv [3], [6] extended Slepian-Wolf theorem and showed that conditional Rate-MSE distortion function for X is same whether the side information is available only at the decoder or both at encoder and decoder; where X and Y are statistically dependent Gaussian random processes. Let X and Y be the samples of two random sequences representing the source data and side information respectively. Encoder encodes X without access to side information Y. Decoder reconstructs $\hat{X}$ using Y as side information. Let $D = E[d(\hat{X}, X)]$ is the acceptable distortion. Let $R_{X/Y}(D)$ be the rate required for the case where side information is available at the encoder also and $R^{WZ}_{X/Y}(D)$ represent the Wyner-Ziv rate required when encoder does not have access to side information. Wyner-Ziv proved that Wyner-Ziv rate distortion function $R^{WZ}_{X/Y}(D)$ is the achievable lower bound for the bit rate for a distortion D

$$R^{WZ}_{X/Y}(D) - R_{X/Y}(D) \geq 0 \tag{4}$$

They also showed that for Gaussian memoryless sources

$$R^{WZ}_{X/Y}(D) - R_{X/Y}(D) = 0 \tag{5}$$

As a result source sequence X can be considered as the sum of arbitrarily distributed side information Y and independent Gaussian noise.

Based on these two fundamental theories a new video coding paradigm has been proposed which encodes each video frame separately but decodes jointly. This system makes use of previous frames (motion compensated or not) as their side information. Such an unconventional video coding system can have complexity and robustness of an intraframe coding system (Motion-JPEG) and at the same time can achieve compression efficiency near to the conventional interframe coding scheme.

III. PRINCIPLE

A source X is to be transmitted using least average number of bits. Statistically dependent side information Y is available only at the decoder. The encoder must therefore encode X in the absence of Y, whereas the decoder jointly decodes X using Y. Slepian and Wolf [2] and Wyner [6] recognized the use of channel codes for source coding with side information.

In this work distributed source encoder compresses X in to syndromes S with respect to a Channel code C [7]. Decoder on receiving the syndrome can identify the coset to which X belongs and using side information Y can reconstruct back X.

IV. DISTRIBUTED VIDEO CODEC

The video frames are divided into blocks of 8x8 and each block is processed one by one. The first frame is intracoded while the rest of the frames are classified as *Inter* or *Intra* or *Skip* blocks based on the Mean Square Error(MSE). If the MSE is very large between the frames then that block will be intracoded which means the temporal correlation or the correlation noise is less. If the MSE is very small then that block need not be coded as the temporal correlation between the blocks is very large and hence can be derived from the previous frame. If the MSE is in between the two limits of *Intra* and *Skip* then that block is classified as *Inter* [7].

Blocks that are classified as *Intra* are coded using conventional Intraframe coding. Block DCT (Discrete Cosine Transform) is applied to each Intra 8x8 block (or 16x16) and

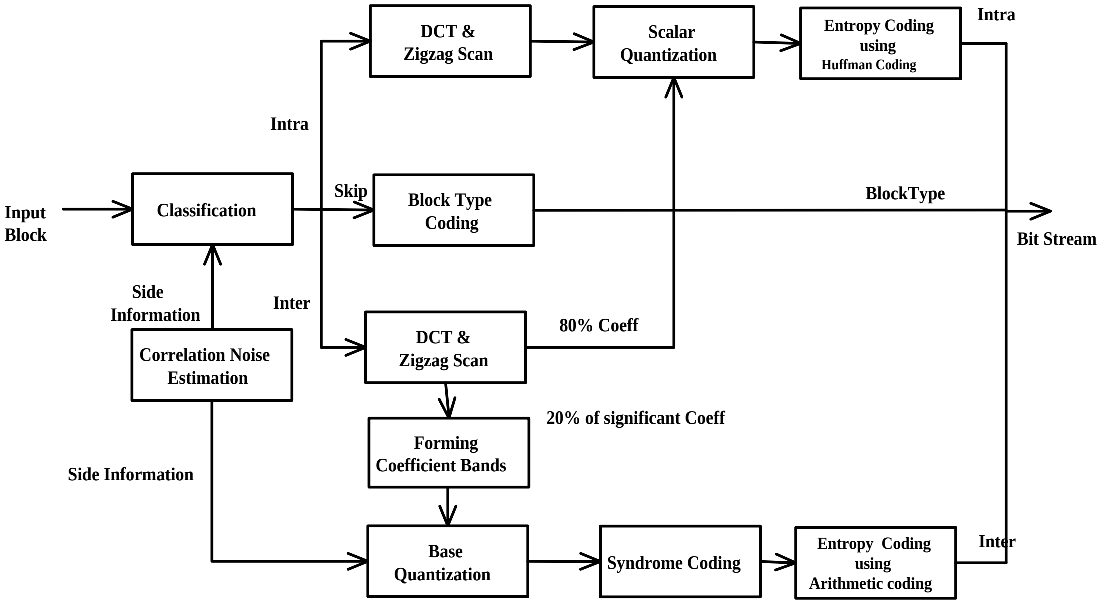

Fig. 4. Video Encoder

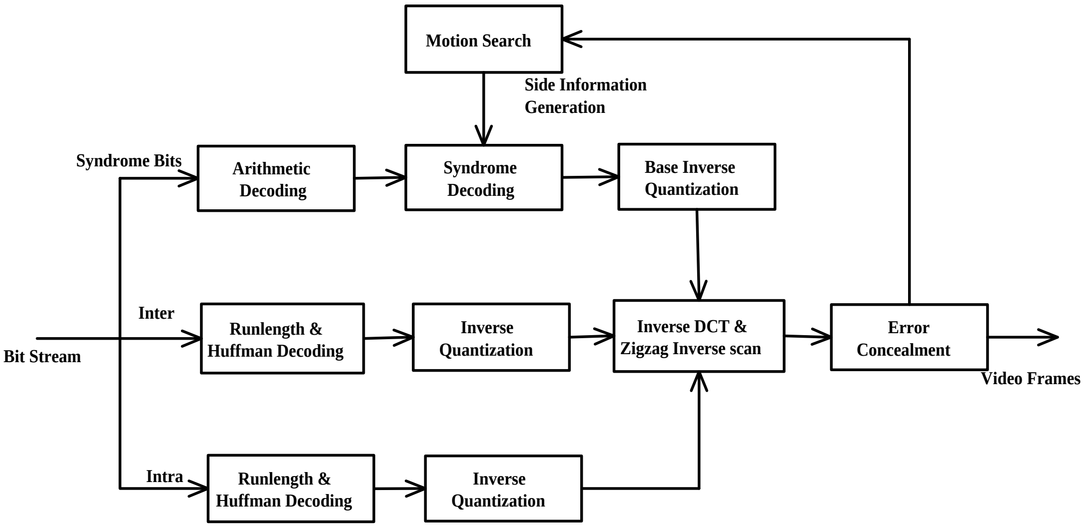

Fig. 5. Video Decoder

the DCT coefficients are quantized using a scalar quantizer. The quantized coefficients are zig-zag scanned and entropy coded using run length and Huffman coding. Blocks that are classified as *Inter* are syndrome coded. Block DCT is applied to the Inter blocks and coefficient bands are formed for each frame. Each coefficient band is uniform quantized with different bit allocations. Different parity check matrices H of simple block codes are used for each coefficient band to obtain syndrome bits. More number of bits are allocated for low frequency bands and few number of bits for high frequency bands. Syndrome bits are obtained by multiplying X with H. The syndrome bits obtained are then arithmetic coded to obtain better compression.

For the coefficients that have been intra-coded, decoding is done by entropy decoding followed by inverse quantization. Once all the coefficients are dequantized they are arranged into a 2-D block by inverse zigzag scan. Then the inverse transform is used to reconstruct the pixels. The syndrome bits are recovered by arithmetic decoding. The syndrome bits identifies the coset to which it belongs. Using the side information generated by the motion search module a sequence of quantized codewords are decoded within the coset identified by the transmitted syndrome. Motion search module will construct a side information frame by performing motion search between two previously decoded frames and extrapolating the estimated motion to current frame considering the pixels from previous frame. Then pixel values are reconstructed by dequantization and inverse transform.

V. IMPLEMENTATION

The encoder and decoder as shown in the block diagram Fig.4 and Fig.5 have been simulated. The video codec is simulated and tested with a object oriented approach using C++ in gcc. The program processes frames one by one and within each frame block wise processing is done. The input to the encoder is a QCIF (Quarter Common Intermediate Format) video file. Encoder allows the storage of one previous frame. Objective performance evaluation of the system is done by measuring the Compression Ratio(CR) (6), MSE and the Peak Signal to Noise Ratio(PSNR) (7) between the original and the reconstructed video file. The PSNR and CR for various video sequences is computed. These are compared with that of H.263+ Intra and H.263+ Predictive video codec [8].

$$CR = \frac{InputFileSize}{OutputFileSize} \tag{6}$$

$$PSNR = 10log_{10}\left[\frac{255^2}{\sum(X_{ij} - Y_{ij})^2}\right] \tag{7}$$

where X_{ij} is $(i,j)^{th}$ pixel of the original image and Y_{ij} is $(i,j)^{th}$ pixel of the reconstructed image.

VI. RESULTS

CR	Bit Rate (kbps)	VideoCodec	Luma PSNR(dB)
11.5917	98	H.263+	35.95
		DVC Implementation	33.13
		Intra (Motion-JPEG)	31.84
9.5360	119	H.263+	37.26
		DVC Implementation	35.28
		Intra (Motion-JPEG)	31.87
7.5728	150	H.263+	38.71
		DVC Implementation	32.87
		Intra (Motion-JPEG)	31.84

TABLE I
FILENAME :NEWS.QCIF,FRAME RATE=30 FPS

CR	BitRate(kbps)	VideoCodec	Luma PSNR(dB)
18.4786	61	H.263+	34.74
		DVC Implementation	32.14
		Intra (Motion-JPEG)	31.88
15.2088	75	H.263+	35.70
		DVC Implementation	33.65
		Intra (Motion-JPEG)	31.96
9.1135	125	H.263+	37.90
		DVC Implementation	33.23
		Intra (Motion-JPEG)	31.96

TABLE II
FILENAME :CONTAINER.QCIF, FRAME RATE=30 FPS

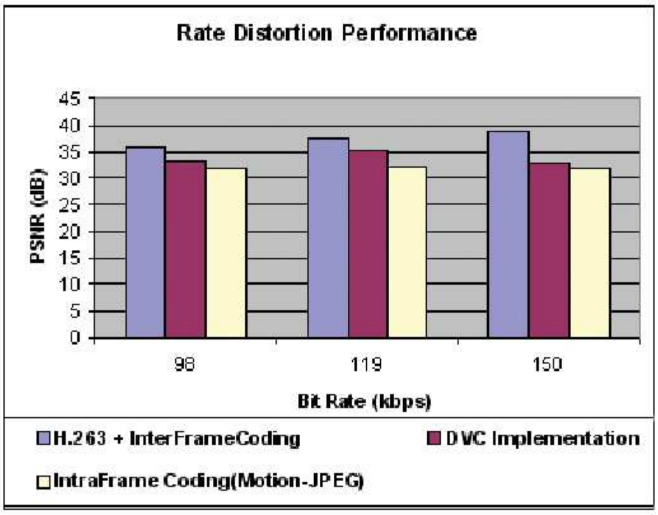

Fig. 6. Rate-Distortion performance for News.qcif

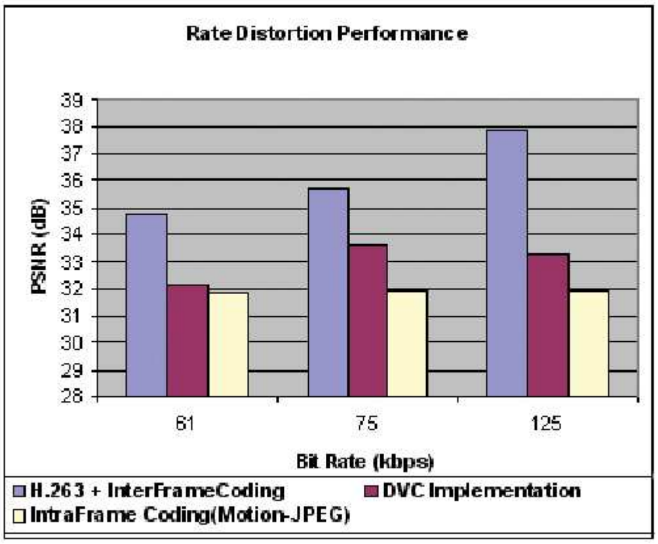

Fig. 7. Rate-Distortion performance for container.qcif

In this section some preliminary results are presented for two video sequences news.qcif and container.qcif in Fig.6 and Fig.7. The results tabulated are for a frame rate of 30 Frames Per Second (FPS)in Table 1 and 2. Constant bit rate (hence Compression ratio) is considered and corresponding PSNR for Luma is obtained for current implementation. Same CR and PSNR are considered for H.263+ Intra and H.263+ Predictive codecs. The current implementation outperforms Intraframe codec(Motion-JPEG), but is slightly inferior to H.263+ interframe(Predictive) codec [8].

VII. CONCLUSION

Distributed Video coding is a new coding paradigm that exploits the source statistics at the decoder thus making encoder simple. Video codec so developed introduces the concept of channel coding in to the problem of source coding with side information. Distributed codec is more robust due to the absence of prediction loop in the encoder. The quality of the reconstructed signal for the same CR can be improved by performing more complex motion search. However it is seen that the current implementation operates well in high quality (PSNR of order of 30dB) regime. The extension to lower bit rates with better quality using more robust channel codes and its real-time implementation using reconfigurable processors is currently in progress.

REFERENCES

[1] Iain E.G. Richardson, H.264 and MPEG-4 Video compression,Video coding for next generation.John Wiley 2003

[2] J. D. Slepian and J. K. Wolf, ”Noiseless coding of correlated information sources“,*IEEE Transactions on Information Theory*, vol. IT-19, pp. 471-480, July 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. IT-22, no. 1, pp. 1-10, Jan. 1976.

[4] Bernd Girod, Anne Margot Aaron,Shantanu Rane and David Rebollo-Monedero,“Distributed Video Coding”, *Proceedings of the IEEE*, Volume 93, Issue 1, pp 71 - 83 , Jan 2005

[5] P.Ishwar, Rohit.Puri, Abishek.Majumdar, K. Ramchandran,“Distributed Video Coding in Wireless Sensor Networks.”,*IEEE Signal Processing Magazine*, Vol 23,Issue 4, pp 94-106,July 2006.

[6] A. D. Wyner, ”Recent Results in the Shannon Theory“,*IEEE Transactions on Information Theory*, vol. 20, No. 1, pp. 2-10, Jan. 1974.

[7] R. Puri and K. Ramchandran, “PRISM: A new robust video coding architecture based on distributed compression principles”, *Proc. Allerton Conference on Communication, Control and Computing*, Allerton, IL,Oct. 2002.

[8] Guy Cote, Berna Erol, Michael Gallant, Faouzi Kossentini, “H.263+: Video Coding at Low Bit Rates”,*IEEE Transactions on circuits and systems for video Technology*, Vo 8, No 7,November 1998.

[9] S. S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): Design and construction”, *Proc. IEEE Data Compression Conference*, Snowbird, UT, pp. 158 -167, Mar. 1999.

[10] 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.