Multilevel Coset coding of video with Golay codes

Aparna.P , Dr Sumam David
Department of Electronics and Communication Engineering
National Institute of Technology Karnataka, Surathkal
Email: p.aparnadinesh@gmail.com, sumam@ieee.org

***Abstract* -This paper presents a video coding method based on the principle of distributed source coding. This work aims in shifting the encoder complexity to the decoder to support uplink friendly video applications, simultaneously achieving the rate-distortion performance of the conventional predictive coding system. In this work concept of syndrome coding with Golay codes is adopted for compression. The simulation results presented in this paper reveals the superior performance of this distributed video coder over the Intraframe coders and predictive coders for video data with less correlation between frames.**

Keywords: predictive coding, coset, uplink, distributed video coding (DVC)

I. INTRODUCTION

Extensive developments in VLSI technology and computing capabilities have made a number of advanced video applications like video conferencing, video telephony, DVD, video-on demand etc possible. With increased technology power more complex video applications like HDTV, 3-D movies, 3-D games have come in to existence demanding higher compression performance. Current popular video standards like ISO MPEG and ITU-H.26x have been successful in delivering the required compression and quality standards. These standards have an interframe coding model that exploits source statistics at the encoder resulting in a very high compression performance at an acceptable quality level. The interframe coding model uses motion estimation and compensation algorithm at the encoder that removes the temporal redundancy between frames. The motion estimation and compensation process amounts to 80% of the encoder complexity and computational resources resulting in a bulky encoder with higher power consumption. This feature makes it inappropriate for uplink friendly applications like mobile video cameras, wireless video sensor networks, wireless surveillance etc. These wireless video applications demand a simple encoder since power, size and the computational resources are of primary concern in the wireless scenario. Distributed Video Coding is a new coding paradigm that attempts to fulfill the requirement of wireless- video applications. In this scheme the complex motion search algorithm at the encoder is eliminated, but can be incorporated at the decoder. Distributed video coder thus exploits the source statistics at the decoder alone, interchanging the traditional balance of complex encoder and simple decoder. A video codec working on this principle is thus very promising for uplink friendly video applications. In such a coding system the encoder encodes each video frame separately with respect to the correlation statistics between itself and the side information. The decoder decodes the frames jointly using the side information available only at the decoder. This video paradigm is as opposed to the conventional coding system where the side information is available both at the encoder and decoder. Alternatively, we have the intraframe coding scheme (Motion- Jpeg) with a low complexity encoder but a poor rate-distortion performance as the temporal redundancy across the frames are not taken care. However it provides robustness against channel errors [1] as every frame is coded independent of others. All these factors lead to the development of a new coding paradigm called the Distributed video coding that needs to incorporate within itself the merits of interframe coding scheme and intraframe coding scheme. Such an architecture promises higher compression efficiency, robustness to wireless channel errors and at the same time distribution of complexity between encoder and decoder. In this paper, we present an approach used for video coding based on the principle of multilevel coset coding using Golay codes and compare it with H.263+ Intra coding [6], H.264 Intra coding [7], H.263+ Inter coding [6]. We also compare the complexity and performance of this work with [8] which presents a syndrome coding approach based on LDPC codes.

II. PRINCIPLE

Distributed Video coding concept is based on the information theoretic bounds established in 1970s by Slepian-Wolf [2] for distributed lossless coding and by Wyner-Ziv [3] for lossy coding with decoder side information. Practical distributed video coders specifically work on the principle of Wyner-Ziv coding considering a distortion measure. In conventional coding system, the source is coded with respect to the side information that is available both at the encoder and decoder. In distributed video coding system side information is not available at the encoder, but the encoder encodes each frame separately with respect to the correlation statistics between itself and the side information. The decoder has access to the side information and hence decodes the frames jointly using the side information. In distributed video coding environment side information Y is treated as the noisy version of the source X i.e $X = Y + N$. Statistically dependent side information Y, is available only at the decoder and if X is the source that is to be transmitted using least average number of bits. The encoder must therefore encode X in the absence of Y, whereas the decoder jointly decodes X using Y. In this context

978-1-61284-690-3/11/ $26.00 ©2011 IEEE

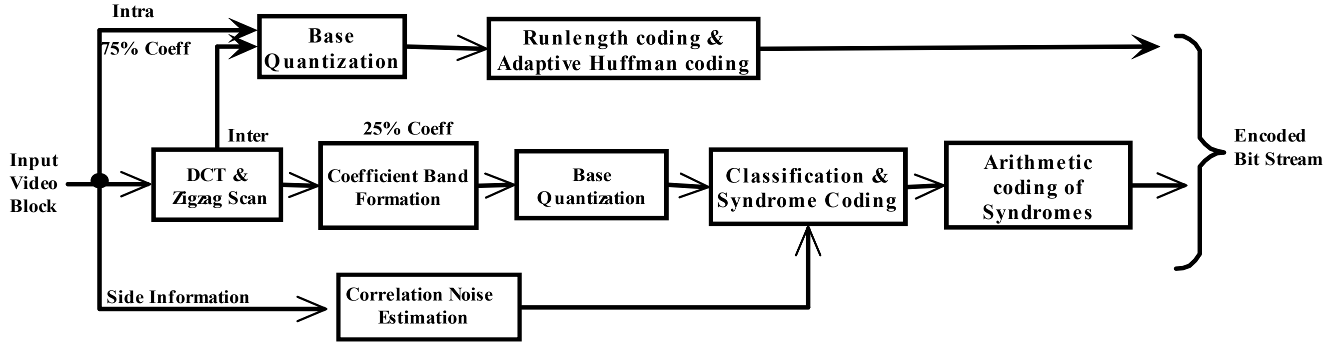


Fig.1 Video Encoder Block Diagram of the Proposed Method

X is compressed to syndrome S of a channel code [5]. These syndromes identify the coset to which X belongs to. The receiver on receiving the syndrome S identifies the code word from the corresponding coset that is close to the side information Y. This corresponds to coset channel coding. This is done by using a parity check matrix H of a (n, k) linear channel code. Compression is achieved by generating syndrome bits of length $(n - k)$ for each n bits of data.

III. PROPOSED SCHEME

Encoder: The encoder block diagram of the current implementation is as shown in the Fig.1. The first frame is coded as the intra frame. Each of the consecutive frames are intercoded. An intra frame is introduced after an interval of p frames. Each 8 × 8 block of the frame is transformed using Discrete Cosine Transform. The transformed coefficients are zig zag scanned so that they are arranged in the order of their importance. For the frames to be coded as intra, these coefficients are scalar quantized and entropy coded using Huffman and run length coding. For the frames to be intercoded, these transformed coefficients are formed into coefficient bands such that each coefficient at the same spatial location within a block is combined together in one coefficient band. The first coefficient band corresponds to all DC coefficients and hence is very significant. Around 1/4 of the coefficients in a block are chosen for intercoding and hence number of bands is limited to 16 in a block of size 8 × 8. The remaining 3/4th of the coefficients are less important and hence can be quantized and entropy coded like intra blocks. These coefficients are insignificant and hence contribute less to the compression performance. Each of the coefficient bands is assigned different number of bits for quantization. More number of bits are assigned to higher bands and less number of bits to lower bands. Proper decoding of the syndrome coded bits requires that the band step size is greater than the correlation noise. Hence a different set of bit allocation is pre-defined. One of these data sets is chosen based on the input data. The bit allocation data set is chosen based on the average of the correlation noise between each band and the corresponding coefficient band of the side information, such that the above criterion is fulfilled. Each of these coefficients bands are uniformly quantized with reference to the bits allocated for each band. Syndrome generation for these bits is based on the principle of multilevel coset code proposed in [5]. Syndrome for each quantized coefficient is generated based on the correlation noise between the source and the side information. In this case syndrome corresponds to the bits that cannot be inferred from the best side information. The Fig.2 shows the quantization bin, where X_i represents the source, X_{qi} quantized value of the source, Y_i represents side information, $\triangle$ represents the stepsize of the corresponding band. The N_i represents the correlation noise, given by $N_i = X_{qi} - Y_i$. In the presence of side information Y_i the number of least significant bits that needs to be communicated to the encoder is given by

$$l_i = 2 + \left\lfloor \log_2\left(\frac{|(N_i)|}{\Delta}\right) \right\rfloor ; N_i > \Delta \qquad (1)$$

$$l_i = 0; \text{else} \qquad (2)$$

The syndrome bits to be communicated to the decoder can be obtained by:

$$S_i = X_{qi} \,\&\, (2^{li} - 1) \qquad (3)$$

where & represents bitwise AND operation. The number of least significant bits l_i , for each coefficient should also be sent along with these bits. Hence for each coefficient, the value of l_i and l_i number of least significant bits are mapped to a unique symbol :

$$S_{ti} = 2^{li} + S_i \qquad (4)$$

where + denotes bitwise OR operation. Transmitting syndrome bits is equivalent to dividing the quantization lattice into sub lattices as shown in the Fig.2 up to the level specified by l_i . Thus the number of least significant bits that needs to be communicated to decoder is dependent on the correlation noise between the source X and the side information Y. Bit planes marked in gray in Fig.3 are transmitted to the decoder and the bit planes in white can be inferred from the side information. More the correlation noise, more number of bits are to be transmitted to the decoder. This corresponds to Wyner-Ziv coding where in the least significant bits that cannot be obtained from the side information is transmitted to

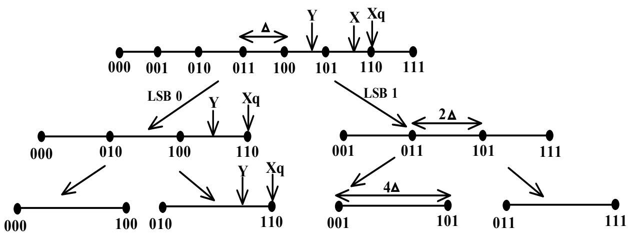


Figure 2. The quantization bin [5]

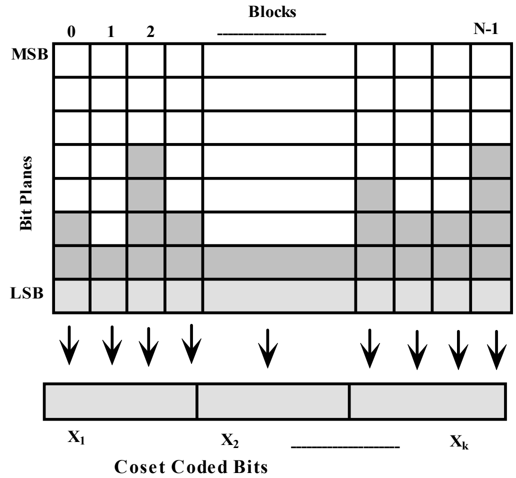


Figure 3. Bitplanes of a single coefficient Band

the decoder. The unique symbol S_{ti} obtained from l_i and S_i is then arithmetic coded and transmitted to the decoder.

Coset Channel Coding using Golay code : In order to reduce the bitrates further, some of the bitplanes can be coset channel coded. These bits as shown in Fig.3 cannot be generated by side information and hence are transmitted to the decoder. However this may incur higher bitrate. In order to further reduce the bitrate, some of the least significant bitplanes are further compressed based on the principle of distributed source coding [5]. Distributed video coding in [8] uses LDPC codes for coset channel coding. This gives only a marginal gain in the bitrate with reasonable quality, but at the cost of very high complexity at the decoder. It also needs a very long block length *n* to give a good performance. In this method (23,12,7) Golay code is used where the block length *n* is 23 and the message bit length *k* is 12 and the hamming distance d_{min} between the code words is 7. In this method the parity check matrix *H* of a (23,12,7) Golay code is used to generate the syndrome bits S_i from the input bits X_i . One or two least significant bitplanes of the syndrome bitplanes can be considered for coset channel coding which are formed into a block of *n* bits each as shown in Fig.3. Each of the *n* bits of *X*, say b_{xi} block is transformed into $(n - k)$ syndrome bits Z_i by using the parity check matrix *H* according to

$$Z_i = Hb_{xi} \quad (5)$$

Thus *n* data bits are compressed to $(n - k)$ syndrome bits giving a compression rate of $(n / n - k)$. Two methods can be adopted for side information generation. If a less complex decoder is desirable we consider the co- located block of the reference frame as side information. The other method of generating the side information is by performing the motion search for the candidate predictor *Y*. This is done by transmitting some cyclic Redundancy check (CRC) bits generated on the quantized codewords which helps the decoder to find the best side information.

2) Decoder: The block diagram of the decoder is shown in the Fig.4. The frames that are intracoded are passed through an entropy decoder, then dequantized and inverse transformed to get back the intra coded frame. The coefficients that are syndrome coded are first passed through the arithmetic decoder to decode the unique symbol S_{ti} . From this symbol S_{ti} the number of least significant bits l_i and the syndrome bits Z_i are obtained. At the decoder only the syndrome bits Z_i and the side information bits Y_i are available. As long as the hamming distance between X_i and Y_i is less than d_{min} of the Golay code, the data b_{xi} can be recovered. The syndrome bits Z_i indicates the coset to which b_{xi} belongs to. The coset leader of that coset is chosen as say A_i . We then find Y_i' such that $Y_i' = Y_i \oplus A_i$, where $\oplus$ is bitwise EXOR operation. We then find b_{xi}' a codeword corresponding to the Golay code closest to Y_i'. The required data bits X_i is then obtained from b_{xi}' as $b_{xi} = b_{xi} \oplus A_i$. These decoded bits b_{xi} are combined with rest of the multilevel coset coded bits to reconstruct back the current frame. Side information generation unit provide the best side information Y_i for the current frame. Based on the side information Y_i the rest of the MSBs are retrieved and combined together with the LSBs to form the current frame coefficient band for *X*. The coefficient bands are uniform dequantized based on the bit allocation set chosen. Rest of the coefficient that are intracoded are further combined with syndrome decoded coefficients and then block inverse DCT is applied to get back the original frame. In case of complex decoder side information is generated by performing the motion search and quantizing the candidate block and computing the CRC. If the CRC matches, that block is considered as side information. By using a *(n, k)* linear channel code, the encoding rate achieved is $(n - k)/n$. In the work [8] the syndrome coding of the inter coefficients was done using 3/4 and 1/2 rate LDPC code. It is observed that for a 1/2 rate LDPC code, the correlation of the sources with the side information should be very high, which other wise would result in high distortion. On the other hand the use of 3/4 rate LDPC encoder results in less distortion, but the compression achieved is quite low. The other issue with the LDPC code is the parity check matrix *H* which needs to be generated in real-time or stored at the encoder. This would increase the complexity of the encoder as storing increases the resource requirement and power consumption. Also LDPC codes require long block lengths and high decoding complexity. As the block length of the data to be coset channel coded is small, LDPC code seems to be unsuitable for the current implementation. Hence in this work Golay codes have been considered for coset channel coding. The encoder of this

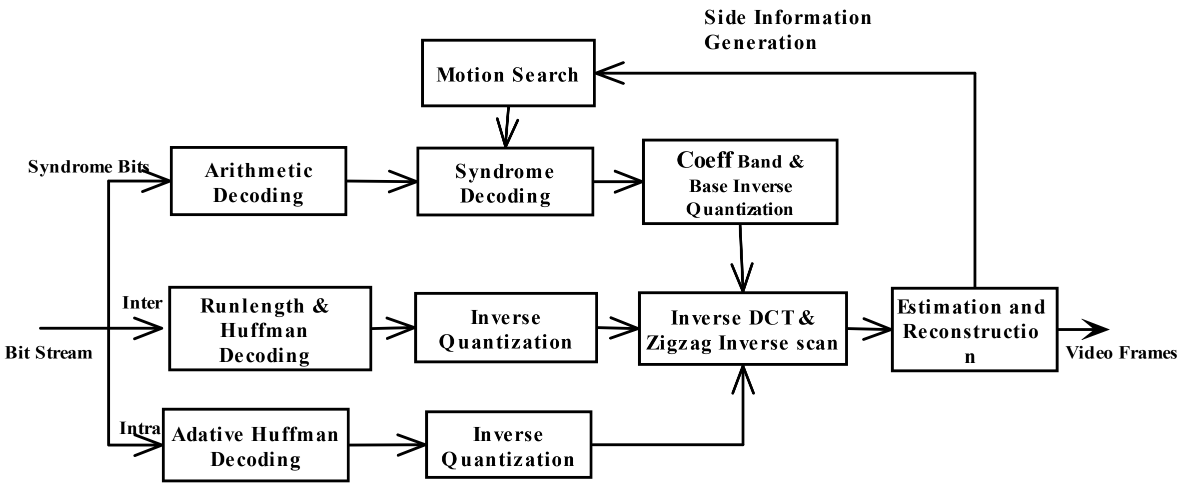
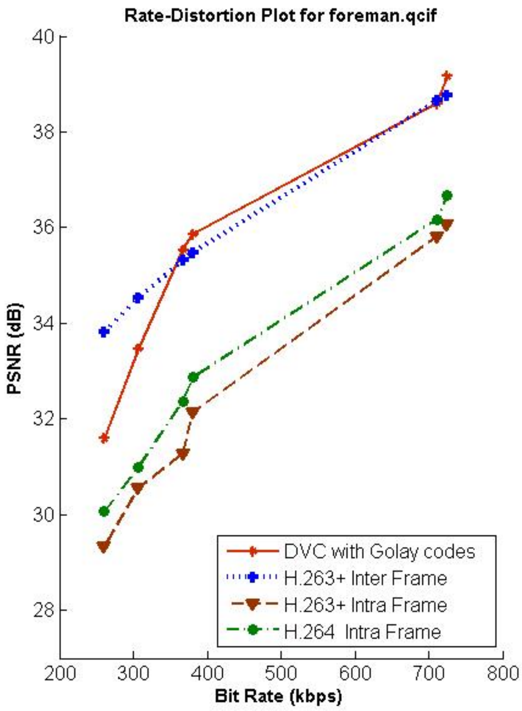


Figure 4. Video Decoder of the proposed scheme



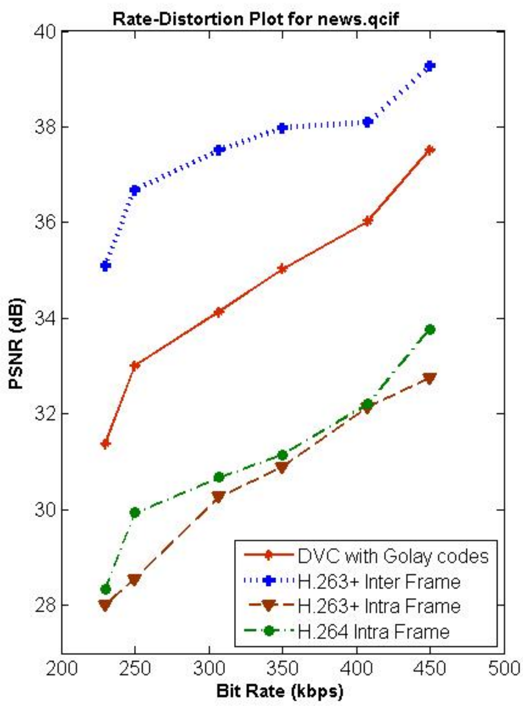


Fig. 5. Comparison of rate-distortion performance of various methods for a frame rate of 30 fps

method is simple satisfying the main objective of this work. Also quality of the reconstructed sequence is good with Golay codes for a 1/2 rate encoder.

IV. SIMULATION RESULTS

Video Codec is designed to support an application to wireless network of video camera equipped with cell phones. Encoder allows the storage of one previous frame. Objective performance evaluation of the system is done by comparing the bit rate and the Peak Signal to Noise Ratio (PSNR) between the original and the reconstructed video sequence. The method discussed above has been simulated and the results are presented in this paper. Two test video QCIF sequences with a resolution of 176 × 144 are considered for evaluating the rate distortion. These video files are considered based on their motion content. The frames of foreman.qcif have less correlation when compared to news.qcif file. As seen in the Fig.5, the Luma PSNR is plotted against different bitrates for different methods. The performance of distributed video codec discussed in this paper is compared with H.263+ Intra Frame coder [6], H.264 Intra Frame coder [7], H.263+ In- ter Frame coder [6] for foreman.qcif and news.qcif at a frame rate of 30 fps. From the simulation results we can see that the DVC implementation performs considerably well for video sequences with higher motion content i.e foreman.qcif. For video sequences with less motion content (i.e more correlation between adjacent frames) H.263+ Inter coder performs better than DVC implementation. Also an improvement of at least 3dB is seen in the current implementations when compared to the syndrome coding with LDPC codes [8]. However the results of DVC implementations are consistently better than H.263+ Intracoder and H.264 Intra coder.

V. CONCLUSION

Coset channel coding using Golay Code has been introduced in this method. Use of simple block codes will improve the rate-distortion performance further without increasing the encoder complexity. Use of LDPC codes for coset channel coding is eliminated as it increases decoder complexity and the requirement of resources at the encoder without contributing much to the rate distortion performance. Also a proper statistical model for classification will drastically improve the performance of the coder. The main aim of this work is to reduce encoder complexity making it pertinent to uplink friendly architecture which seems to be satisfied. It is also observed that the current implementation operates well in high quality (PSNR of order of 30dB) regime. The extension to lower bit rates without any compromise in the quality so that it is comparable with the conventional codecs can be further considered without increasing the complexity.

REFERENCES

[1] 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

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

[5] 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.

[6] G.Cote, B.Erol, M.Gallant, and F.Kosssentini, "H.263+:Video coding at low bitrates ", IEEE Transactions. Circuits Sys. Video Technology, Vol 8, No 7, pp 849-866, November 1998.

[7] T. Wiegand, G.J.Sullivan, G.Bjontegaard and A.Luthra, "Overview of the H. 264/AVC video coding standard ", IEEE Transactions on circuits and systems for video technology, Vol 13, No 7, pp 560-576, 2003.

[8] Aparna P, Sivaprakash Reddy, Sumam David, "Distributed video coding using LDPC codes for wireless video ", Journal on Wireless Sensor Networks, Scientific Research Publishing, Inc. USA, pp 334-339, Nov 2009,

[9] Rohit Puri, Abhik Majumdar and Kannan Ramachandran, "PRISM: A video Coding Paradigm with Motion Estimation at the Decoder", IEEE Transactions On Image Processing, Vol 16, No 10, October 2007.

[10] 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.