ISSN 2229-5518

Enhancement in lossless video sequnces compression using adaptive pixel based prediction method

Saraswati Jadhav, **Kishori Deganokar**

Abstract—We investigate lossless compression schemes for video sequences of adaptive pixel based prediction method.The proposed technique exploits spatial, temporal correlation and spectral redundancy in backward adaptive method. The proposed coder gives the relational between the operational domain (Wavelet or Spatial) and amount of temporal, spatial redundancy in video sequences.No need to send motion vector as side information like block motion compensation method. It sends very low side information like operational domain and search range.Due to less side information; encoder increases the compression efficiency.This algorithm use simple lossless arithmetic entropy coding. It also shows relation between the size of search window and compression ratio.The experimental result shows the improvement in compression efficiency.

Index Terms—. Adaptive pixel based prediction; arithmetic coding; Integer Wavelet transforms; spatial prediction; temporal prediction

1. Introduction

There are two types of compression, lossy compression and lossless compression. Lossless compression of a video sequence [1]-[9] may be required for medical and scientific applications, and also for studio quality archival purposes.

Memon and Sayood investigate [1] lossless compression scheme for video sequences. This hybrids video compression scheme makes use of both spectral and temporal correlations.A sequence of video images has spatial, temporal, and spectral correlations. Hence, for the specific case of video images, the selection of the predictor can be based either on information from spectrally or temporally adjacent pixels, or both. The lossless JPEG predictors are quite effective in removing spatial correlations present in individual frames.

Zhang [7] suggested the method for lossless video compression using combination of temporal and spatial adaptive prediction. This video encoder uses movement estimation to remove the temporal redundancy and uses GAP gradient-adjusted predictor to remove the spatial redundancy, and then an adaptive method is used to combine the two predictions.This schemes also use context -based coding.

CALIC [12] (context based adaptive lossless image compression) uses both context and prediction of pixel values. Brunello [5] consider the problem of lossless compression of video by taking into account temporal information.Video lossless compression is an interesting possibility in the line of production and contribution. Brunello propose a compression technique which is based on motion compensation, optimal three-dimensional (3-D) linear prediction. This algorithm, motion vectors are sending at the output of encoder. Due to this compression efficiency is decreases.

Kyeong [2] proposed context based predictive coding of video, where the interframe or intraframe coding mode is adaptively selected on a pixel basis. We perform the coding mode selection using only the previously reconstructed samples which are also available at the decoder, so that any overhead information on the coding mode selection does not need to be transmitted to the decoder.

Park [9] describes an adaptive lossless compression algorithm for color video sequences utilizing backward adaptive temporal prediction and an integer wavelet transform. Park exploits two redundancies in color video sequences specifically spatial and temporal redundancies. Park shows that an adaptive scheme exploiting the two redundancies has better compression performance than lossless compression of individual image frames.

In this algorithm, we take the temporal predictor in [9] as reference.We proposes an enhanced pixel predictor which exploits the motion information among adjacent frames using extremely low side information. It gives the relationship between the operational domain and the amount of the temporal and spatial redundancies of the sequence to be encoded. The computational complexity entropy coder is reducing. As opposed to [9], the proposed predictor uses a backward adaptive prediction mode selector instead of using extra bits to indicate the prediction mode. Meanwhile, to maximize the prediction performance, the proposed scheme adaptively selects the predictor from all candidates based on previous prediction accuracy.

- *Saraswati Jadhav is PG Research Scholar & Lecturer inVishwakarama Instistiute of Technology, Pune-411037, India. E-mail: saraswatijadhav@gmail.com*
- *Kishori Degaonkar is Assistant Profressor in Vishwakarama Instistiute of Technology, Pune-411037, India.*

ISSN 2229-5518

2. Lossless video sequnces compression using adaptive prediction method algorithm:

We propose this algorithm in which we use pixel based compensation algorithm.This algorithm motion vectors are not send as side information.So size of side information is decrease .This method encoder uses fixed size side information and it is independent of video frame size.So compression efficiency of video frame is increases.As shown in figure, the proposed scheme is divided into following parts: preprocessing, adaptive symbol prediction, adaptive prediction mode selection, and arithmetic coding of the prediction residuals.

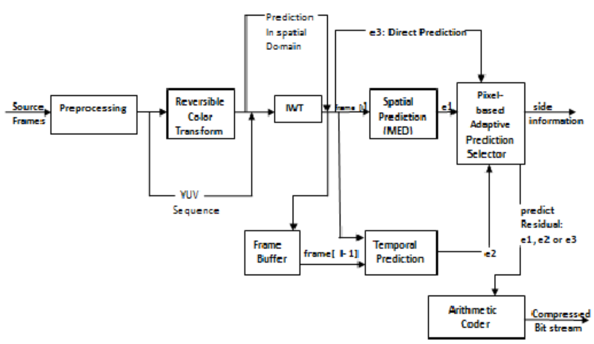

Fig. 1 Block diagram of lossless video sequnces compression using adaptive prediction method

Preprocessing select the operational domain of video sequnces as wavelet or spatial depending on it's temporal or spatial redundancy. Adaptive symbol prediction method determines the spatial, temporal and direct prediction residual. In adaptive prediction mode selection step, predictor is adaptively selected from prediction residuals of the past neighboring pixels.Simple arithmetic coding encode minimum prediction residual.

We divide preprocessing of video sequnces in two steps. First step of preprocessing, we determine the temporal redundancy of video sequences. The temporal redundancy of video sequences is calculated by interframe correlation coefficients of video sequences. Consider Pi(x, y) is the pixel to be encoded that is located at (x, y) in frame [i], then inter frame correlation coefficients frame[i] and frame [i+1] estimated by (1), where Pi is the average of pixels in frame[i] and Pi+1 is the average of pixels in frame [i+1].We calculate the average of inter frame correlation coefficients of all video test sequences. According to result of test sequences, we decide the threshold value of video test sequnces as 0.9 .If average of interframe correlation coefficients higher than predefined threshold value then sequences is low motion video sequences. In this case motion compensation of video sequences, we select the wavelet domain.Wavelet transform [19]-[21] gives energy compensation property.If average of inter frame correlation coefficients smaller than predefined threshold value then sequences is high motion video.

$$\rho i,i+1=\frac{\sum_x\sum_y\left(pi(x,y)-\overline{pi}\right)\left(pi+1(x,y)-\overline{pi+1}\right)}{\sqrt{\left(\sum_x\sum_y\left(pi(x,y)-\overline{pi}\right)^2\right)\left(\sum_x\sum_y\left(pi+1(x,y)-\overline{pi+1}\right)^2\right)}} \quad (1)$$

Wavelet transform is inefficient for motion compensation of high motion video due to shift variant property of transform. So high motion video is operated in spatial domain. Fig. 2 shows the interframe correlation coefficient of video sequnces of 0^{th} to 39^{th} frame.

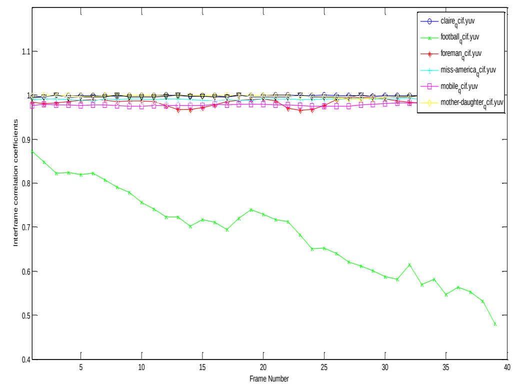

Fig. 2 Interframe correlation coefficients of video test sequnces

Our experiment, we consider all video test sequnces as YUV and size of each video is 176X144 (Quadratic Common Interchange format-qcif).

Second step of preprocessing, we estimate spatial redundancy of low motion video sequences. Depending on spatial redundancy of sequences, we select suitable IWT for sequences. We estimate spatial redundancy of sequences such as correlation coefficients, as high frequency coefficients magnitude of discrete cosine transform (DCT)[15] .The sum of the amplitudes of the four highest frequency coefficients are shown in fig(3) as follows.This sum indicates that intra frame correlation coefficients between the video frames.We estimate the average of intra frame correlation coefficient of 0^{th} to 39^{th} frames.

The small value of intra frame correlation coefficient indicates that the adjacent pixels are varying greatly within a frame. The small value of intra frame correlation coefficients video frames operated in 5/3 integer wavelet transform [13]-[14] (5/3 IWT). 5/3 transform performs well for both lossy and lossless compression. The high value of intra frame correlation coefficients video frames implies that the adjacent pixels are very close to each other and S transform is used to estimate the spatial redundancy of these video frames. S transform is integer version of Haar transform which has lowest computational complexity among all transforms.

ISSN 2229-5518

The forward S transform equations are given below:

$$d[n] = x[2n+1] - x[2n]$$
$$s[n] = x[2n] + \lfloor d[n]/2 \rfloor \qquad (2)$$

The forward 5/3 transform equations are given below:

$$d[n] = x[2n+1] - \lfloor (x[2n] + x[2n+2])/2 \rfloor$$
$$s[n] = x[2n] + \lfloor (d[n-1] + d[n])/4 + 1/2 \rfloor \qquad (3)$$

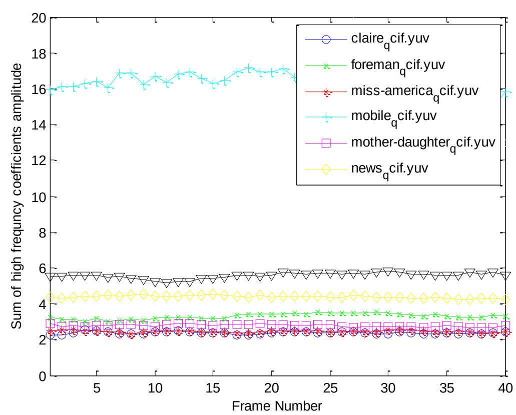

Fig. 3 Sum of high frequency coefficients amplitude of video test sequnces

Table I shows data bit rate of RGB video test sequnces after compression.

Table I

Sequnces	Bit Rate IWT5/3 Wavelet	Bit Rate IWTS Wavelet	Bit Rate SPATIAL
miss -america	10.82633	7.81894	8.53718
foreman	11.97155	9.64035	10.03511
carphone	10.83419	8.76167	9.18563
mobile	11.78995	12.28147	12.71177
football	11.09516	11.50565	10.98012

To take into account spatial redundancy, prediction based on the values of pixels belonging to the same frame of pixel to be encoded is computed. Median edge detector predictor, primitive edge detection of horizontal and vertical edges is achieved by the examining the neighboring pixels of current pixel Pi(x,y).The spatial predictor [16] estimate the pixel to be coded based on the three already coded neighboring pixels.
The spatially predicted value is:

$$\hat{pi}^{s}(x,y) = \begin{cases} \min(A,B) & if C \ge \max(A,B) \\ \max(A,B) & if C \le \min(A,B) \\ A+B-C, & otherwise \end{cases} \qquad (4)$$

The spatially predicted value is:

$$e_1 = pi(x,y) - \hat{pi}^{s}(x,y) \qquad (5)$$

Temporal Prediction:

Temporal Prediction reduces the redundancy in frames. To reduce the temporal redundancy, we present an adaptive pixel based predictor. The proposed temporal predictor does not require transmission of motion vectors. So high amount of side information is reducing.

Assume that Pi(x, y) is the symbol to be encoded in frame[i]. The proposed temporal predictor aims to find best matched symbol in references frame [i-1].The best matched symbol in references frame[i-1] denoted as temporal predictor Pi^T(x,y).The predictor find the motion activity of target window Pi(x,y) in frame[i] and references frame [i-1] within search range W*H as shown in fig. The target window from current frame formed by upper left neighboring symbols of Pi(x,y).The temporal predictor searches best match of target window using cumulative absolute differences between target window frame[i] and frame[i-1].The minimum cumulative absolute differences within search range ,where

$$CAD(T\omega) = \sum_{(m,n)\in T\omega} \left| p_i(x,y) - p_{i-1}(x+m, y+n) \right| \qquad (6)$$

Where Tw denotes target window Pi(x,y) and Pi-1(x,y) denotes the symbol values of the current frame[i] and frame[i-1] respectively. Motion vectors (m,n) determined for the region

$$(mo, no) = \arg\min_{(m,n)} CAD(T_\omega) \qquad (7)$$

(m_0, n_0) is motion displacement of target window.Then temporal predictor Pi(x, y) obtained by

$$Pi^{T}(x,y) = P_{i-1}(x+mo, y+no) \qquad (8)$$

Then temporal prediction residual calculated by

$$e_2 = Pi(x,y) - P_i^{T}(x,y) \qquad (9)$$

Test result shows that this temporal predictor gives better performance. But computational complexity of temporal predictor is very high due to search and match in target window of frame[i] and frame[i-1].This temporal predictor method is very time consuming .It gives undue computational burden in real application.

But compression efficiency is first objective of this algorithm. So computational complexity of algorithm is ignored. It is very difficult to find out parameter W, H, M, N, and L of target window and search range because we know that video is non stationary sequences and charterstics of different video sequences always vary greatly from each other. So proposed algorithm W, H, M, N, and L are adjustable to improve compression performances.

Direct mode:

We know the energy compaction property of IWT. After application of IWT, we check high frequency sub band (LH, HL, and HH) wavelet coefficient of transform. If this high frequency sub band amplitude is less than temporal predictor residual or spatial predictor residual, then wavelet coefficients

ISSN 2229-5518

are encoded and transmitted directly denoted as e3.

Adaptive backward prediction mode selection:

The proposed algorithm select predictor among three residuals spatial predictor, temporal predictor and direct mode (e1, e2, e3) based on previous prediction accuracy.

The adaptive prediction selection is based on sum of amplitudes of prediction residuals of past neighboring pixels. The template of backward adaptive prediction mode selection of symbol Pi(x, y) as shown below:

$$\begin{aligned} e^{(S)} = & |e_1(x-1, y-2)| + |e_1(x, y-2)| + |e_1(x+1, y-2)| \\ & + |e_1(x-2, y-1)| + |e_1(x-1, y-1)| + |e_1(x, y-1)| \\ & + |e_1(x+1, y-1)| + |e_1(x-2, y)| + |e_1(x-1, y)| \end{aligned} \quad (10)$$

e(s) represents the sum of amplitude of spatial prediction residuals of past neighboring symbol.

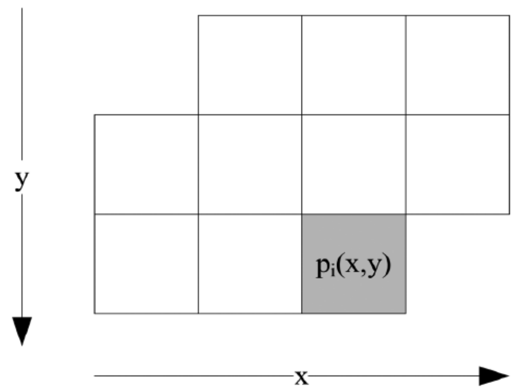

Fig. 2. Template of backward adaptive prediction mode selector of symbol

$$\begin{aligned} e^{(T)} = & |e_2(x-1, y-2)| + |e_2(x, y-2)| + |e_2(x+1, y-2)| \\ & + |e_2(x-2, y-1)| + |e_2(x-1, y-1)| + |e_2(x, y-1)| \\ & + |e_2(x+1, y-1)| + |e_2(x-2, y)| + |e_2(x-1, y)| \end{aligned} \quad (11)$$

e(T) represents the sum of amplitude of temporal prediction residuals of past neighboring symbols

$$\begin{aligned} e^{(D)} = & |e_3(x-1, y-2)| + |e_3(x, y-2)| + |e_3(x+1, y-2)| \\ & + |e_3(x-2, y-1)| + |e_3(x-1, y-1)| + |e_3(x, y-1)| \\ & + |e_3(x+1, y-1)| + |e_3(x-2, y)| + |e_3(x-1, y)| \end{aligned} \quad (12)$$

e(D) represents the sum of amplitude of direct prediction residuals of past neighboring symbols.

The final prediction mode is indicated by the mode with minimum value among e(s), e (T), e (D) which is

$$mode = \arg\left(\min_{S,T,N}\left\{e^{(S)}, e^{(T)}, e^{(D)}\right\}\right) \quad (13)$$

If e(s) is smallest value in {e(s), e(T),e(D)} } for symbol Pi(x, y),then spatial Prediction is selected as the prediction mode and prediction residual is e2.The selection and calculation of the prediction mode only uses past Information. So it has advantage that no need to transmission of extra side information.

This way reduce the size of compressed data by removing extra bits used to represent the prediction mode, but also gain high prediction efficiency by Adaptively selecting the predictor that performs best for neighboring pixels.

Side information at the output of encoder we send as follows:

1. Video sequences RGB or YUV
2. Operational domain: SIWT, 5/3IWT, Spatial transform.
3. Frame width and height.
4. Target window size: W, H
5. Search Range: M, N, L

In proposed algorithm, size of side information is independent of the size of video sequnces.

Arithmetic coding:

The prediction residuals are encoded using arithmetic entropy coding. Arithmetic coding is form of variable length encoding. Encoding and decoding of prediction residuals involves arithmetic operation. So it is called arithmetic coding.When prediction residual symbols are converted to arithmetic coding ,frequently used symbols are stored with fewer bits and not so frequently occurring symbols will be stored with more bits, resulting fewer bits used in total. Arithmetic coding encodes the entire message into a single number, a fraction n where (0.0 ≤ n < 1.0). The next step is to encode this fraction using a fixed-point binary number of sufficient precision

3 RESULTS

We tested proposed algorithm using RGB and YUV color video sequences. We consider lossless image coding JPEG-LS [11] and CALIC [12] to compare proposed algorithm result.

TABLE I
DATA BIT RATE FOR YUV SEQUNCES

Sr. No.	Name of video sequences	Data bit rate	JPEG_LS	Compression gain over JPEG_LS(%)
1	Foreman	4.67456	5.869	20.30
2	News	2.83702	5.063	43.96
3	Mother-daughter	3.41643	5.253	34.96
4	Salesman	3.70754	5.864	36.77

Data bit rate also depends on the serach range and target window size of video frame as shown in Table III-A and Table III-B. We utilize search range and target window size in temporal prediction of spatial domain. If size of search

ISSN 2229-5518

window and target window size of frame is less then data bit rate is high. If size of search window and target window size of frame is high then data bit rate is low.

TABLE II
DATA BIT RATE FOR YUV SEQUNCES

Sr. No.	Name of video sequences	Data bit rate	CALIC	Compression gain over CALIC (%)
1	Foreman	4.67456	5.679	17.68
2	News	2.83702	5.108	44.45
3	Mother-daughter	3.41643	5.176	33.39
4	Salesman	3.70754	5.779	35.84

TABLE III-A
DATA BIT RATE FOR YUV SEQUNCES AND WINDOW SIZE

Sr. No.	Name of video sequences	Data bit rate	Target Window size	Search Range
1	foreman	4.67456	(4,4,4)	(12,12)
2	news	2.83702	(4,4,4)	(12,12)
3	salesman	3.70754	(8,8,4)	(15,15)
4	tennis	5.00043	(4,4,4)	(12,12)
5	mother	3.416437	(10,10,5)	(17,17)

TABLE III-B
DATA BIT RATE FOR YUV SEQUNCES AND WINDOW SIZE

Sr. No.	Name of video sequences	Data bit rate	Target Window size	Search Range
1	foreman	4.61657	(8,8,4)	(15,15)
2	news	2.76298	(8,8,4)	(15,15)
3	salesman	3.78352	(4,4,4)	(12,12)
4	tennis	5.04881	(8,8,4)	(15,15)
5	mother	3.51491	(4,4,4)	(12,12)

4 Conclusion

We presented new scheme lossless adaptive prediction for video coding of video sequences. This algorithm exploits spatial, temporal redudancy of video frames.We select operational domain spatial or wavelet for video sequences depending on interframe correlation coefficients. Also this algorithm adaptively select the best predictor out of spatial, temporal, direct depending on minimum residual.In this algorithm, side information like motion vectors of video frames are not send as output of encoder .So data bit rate of video sequnces is decrease.Like this,We enhance data bit rate of video sequences. Side information of video sequences is fixed, it is independent of size of video.The fixed side information gives the information of W, H, M, N, L of video frame in temporal prediction and operational domain (spatial, wavelet).Table shows that slow motion video average data bit rate is low and vice versa.Table shows the relation between target window size and data bit rate of video sequnces.In temporal prediction, target window size of video frame is inversely propentional to data bit rate of video frame. Computional complexity of this algorithm is moderate.

5 References

[1]N. D. Memon and K. Sayood, "Lossless compression of video sequences," IEEE Trans. Commun. vol. 44, no. 10, pp. 13401345, Oct. 1996.

[2] K. H. Yang and A. F. Faryar, "A context-based predictive coder for lossless and near-lossless compression of video," Int. Conf. Image Processing,Sep. 2000, vol. 1, pp. 144147.

[3] E. S. G. Carotti, J. C. De Martin, and A. R. Meo, "Low-complexity lossless video coding via spatio-temporal prediction," Int. Conf. Image Processing,Sep. 2003, vol. 2, pp. 197200.

[4] -, "Backward-adaptive lossless compression of video sequences," IEEE Int. Conf. Audio, Speech, Signal Processing,Apr. 2002, pp. 34173420.

[5] D. Brunello, G. Calvagno, G. A. Mian, and R. Rinaldo, "Lossless compression of video using temporal information," IEEE Trans. Image Process.,vol. 12, no. 2, pp. 132139, Feb.
2003.

[6] B. Martions and S. Forchgammer, "Lossless compression of video using motion compensation," IEEE DCC, 1998, pp. 560589.

[7] Z. Ming-Feng, H. Jia, and Z. Li-Ming, "Lossless video compression using comnination of temporal and spatial prediction," IEEE. Int. Conf. Neural Netwoks Signal Processing, Dec.
2003, pp. 11931196.

[8] Y. Gong, S. Pullalarevu, and S. Sheikh, "A wavelet-based lossless video coding scheme,"Int. Conf. Signal Processing, 2004, pp. 11231126.

[9] S.-G. Park, E. J. Delp, and H. Yu, "Adaptive lossless video compression using an integer wavelet transform," Image Processing, 2004, pp. 22512254.

[10] S.-G. Park, "Adaptive lossless video compression," Ph.D. dissertation School Elect. Comput. Eng., Purdue Univ., West Lafayette, Dec. 2003

[11] -, "JPEG-LS, Lossless and Near-Lossless Coding of Continuous Tone Still Images,"ISO/IEC JTC1/SC 29/WG,1, Jul. 1997.

[12] N. Memon and X. Wu, "Context-based, adaptive, lossless image coding," IEEE Trans. Commun.,vol. 45, no. 4, pp. 437444, Apr. 1997.

[13] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, "Wavelet transforms that map integers to integers," Appl. Comput. Harmon. Anal., vol. 5, no. 3, pp. 332369, Jul. 1998.

[14] -, "Lossless image compression using integer to integer wavelet transforms," IEEE Int. Conf. Image Processing, Oct. 1997, vol. 1, pp. 596599.

[15] K. Sayood, Introduction to Data Compression, 2nd ed. San Mateo, CA: Morgan Kaufmann, 2000.

[16] D. S. Taubman and M.W. Marcellin, JPEG2000 Image Compression Fundamentals, Standards and Pratice, Norwell, MA: Kluwer, 2002.

[17] M. G. Bulmer, "Principles of Statics," New York: Dover, 1979.

[18] Ze-Nian li.Mark S.Drew, fundametals of Multimedia, 2n ed,Prentice-Hall of2006

[19] Refacel C.Gonzalez and Richard E.Woods, Digital Image Processing,

ISSN 2229-5518

3rd ed., Pearson Prentice Hall of India Pvt Ltd, 2006.
[20] Iain E.Richardson, The.H.264 Advance Video Compression Standard, 2nd ed., 2010 John Wiley and Son Ltd, United Kingdam, 2003.
[21] Peter D. Symes, "fundamental compression techniques and an overview of the JPEG and MPEG compression systems," Video Compression, 3rd ed., McGraw-Hill, California, 1998.