

Low Complexity Lossless Video Compression

*Yao-Chun Fang, Chun-Yi Lee, Yin-Ming Wang, Chung-Neng Wang, Tihao Chiang**

National Chiao Tung University, Hsinchu, 30050, Taiwan

E-mail: tchiang@mail.nctu.edu.tw

ABSTRACT

We present a line-based adaptive lossless video compression (LALVC) algorithm for interactive multimedia applications that demand low complexity and low latency. Communications between high-resolution display and storage devices require high bandwidth for exchanging raw data. To reduce cost of video transmission without losing data accuracy, lossless video compression is necessary. Considering low complexity and low delay, LALVC adopts a simple and efficient architecture that adopts one-pass, raster-scan, transform-free coding process and a simple predictor. For low latency, zero-motion prediction and one-frame buffer are used to reduce temporal redundancy. In addition, to maximize the coding efficiency for both natural and computer-generated video sequences, LALVC adaptively selects the best coding mode for each line of a frame. The entropy coding of each line is based on Golomb code that can enhance coding efficiency with less computation load and is easy for hardware realization. The simulation results show that temporal preprocessing and line-based mode decision of LALVC can increase compression ratio with properly increased complexity as compared to that of JPEG-LS.

1. INTRODUCTION

Real-time interaction over universal multimedia access environment becomes more popular recently. The interactive applications include distance learning systems, tele-medicine diagnosis systems, on-line editing of shared documents, remote computer control, etc. In these applications, raw video data of VGA, SVGA, or higher resolutions are transmitted via the Internet or wireless channels. The transmission requires significant bandwidth to exchange video contents without loss. As video resolution increases, low bitrate and lossless video compression is strongly demanded to save the transmission bandwidth and time. For interaction with limited latency, it is challenging to retain a most compact and lossless representation of high-resolution video sequences with low complexity and low latency.

For lossless coding of still pictures, the existing algorithms covering CALIC [1] and JPEG-LS (LOCO-I) [2] have been shown to have low complexity. The existing

coding standards such as MPEG-2 for DVD can achieve the near perceptually lossless video fidelity but may still cause accuracy loss in reconstruction. As to lossless video coding, no existing standards can be found. Therefore, the basic idea is to propose a line-based adaptive lossless video compression (LALVC) approach with low complexity and low latency for real-time transport of raw video data.

Optimal lossless video coding algorithms need to remove both spatial and temporal redundancy simultaneously. For temporal redundancy reduction, a three-dimensional (3-D) context modeling approach was proposed in [6]. The 3-D method uses computationally expensive motion search and needs extra side information for transmitting motion vectors to the decoding process, which is not suitable for low-complexity application. In addition, block-based motion search is not essential for enhancing lossless video compression [7]. Thus, to remove temporal redundancy efficiently, LALVC adopts temporal pre-processing. The preprocessing adopts a near optimal mode decision for processing each line of pixels based on the residual energy of the handling line. The residual is found by subtracting the pixel values of the current frame and the previous frame under zero motion constraint.

To reduce the spatial redundancy with low complexity and low delay, LALVC uses a one-pass, raster-scan order, transform-free coding process and a simple predictor. The one-pass and raster-scan order process in CALIC and JPEG-LS (LOCO-I) have been proven to be simple and efficient for still picture coding. In addition, the process needs only a two-line buffer for the context model and inter-pixel prediction. Such time-consuming operations like DCT or wavelet transform are avoided in CALIC and JPEG-LS. For lossless compression, the transformation may not obtain a proper coding gain as the coefficients are encoded without quantization. Instead, these two algorithms de-correlated pixels via intra-picture prediction. Some algorithms [3]-[4] were modified to adapt for color pictures coding by adding simple predictors to process Red, Green and Blue color planes. For low complexity, we will not apply the inter-prediction across color components.

To select from various spatial and temporal techniques for optimizing coding gain of LALVC approach, 2-bit side information is used to present the best coding mode of each line. The experimental results show that LALVC can decrease the total bits to accurately represent the video sequences with properly increased complexity over that of JPEG-LS.

¹This work was supported by AU Optronics Corporation, Taiwan, R.O.C. (AUO Contract No. 920420)

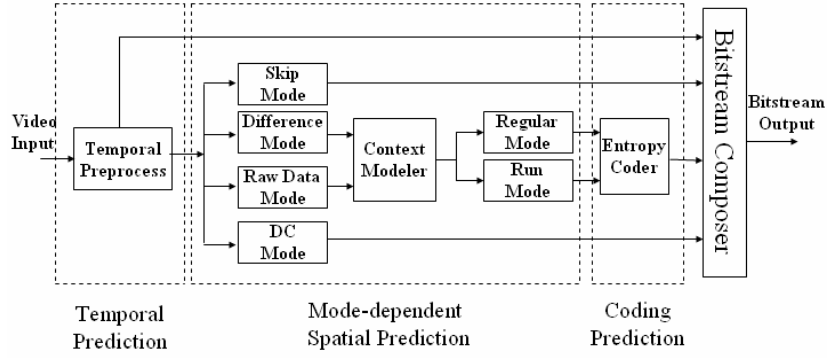


Figure 1. The block diagram of the LALVC encoder.

2. LINE-BASED ADAPTIVE LOSSLESS VIDEO COMPRESSION (LALVC)

LALVC algorithm contains three major parts. In Figure 1, the source video is first fed into a preprocessing unit for temporal redundancy removal. The output of the preprocessing process provides the best mode for coding each line. The mode information is then passed to the second part, which is referred to as a mode-dependent spatial prediction. Based on a simple spatial predictor, inter-pixel redundancy is removed and the residuals are encoded with an entropy coder using Golomb Code [2]. The decoder uses inverse operations to reconstruct the sequence in a manner of line-by-line and frame-by-frame.

2.1. Preprocessing for Temporal Redundancy Removal

LALVC encodes video sequences in a manner of pixel-by-pixel and then line-by-line in raster scan order. The line-based preprocessing algorithm employs some predictions to remove redundancy between temporally neighboring frames. The temporal prediction with zero motion vector, which has the lowest complexity, is applied to co-located lines of successive frames. The i -th and $(i-1)$ -th lines of the frame I_n and i -th line of the frame I_{n-1} are imported into the line comparator to calculate the difference of the two lines for advanced mode decision. The mode decision is based on two metrics including SAD (sum of absolute difference) and SAGD (sum of absolute gradient of difference).

$$SAD_{inter} = \sum_{j=1}^N |I_n(i, j) - I_{n-1}(i, j)| \quad (1)$$

$$SAGD_m = \sum_{j=2}^N |I_m^D(i, j) - I_m^D(i, j-1)|, m = \{intra, inter\} \quad (2)$$

, where difference signals $I_{intra}^D(i, j)$ and $I_{inter}^D(i, j)$ for the j -th pixel and i -th row within the n -th frame are derived by

$$I_{intra}^D(i, j) = I_n(i, j) - I_n(i-1, j), i, j = 1 \dots N \quad (3)$$

$$I_{inter}^D(i, j) = I_n(i, j) - I_{n-1}(i, j), i, j = 1 \dots N \quad (4)$$

With SAD_{inter} , $SAGD_{intra}$ and $SAGD_{inter}$, in Figure 2 we

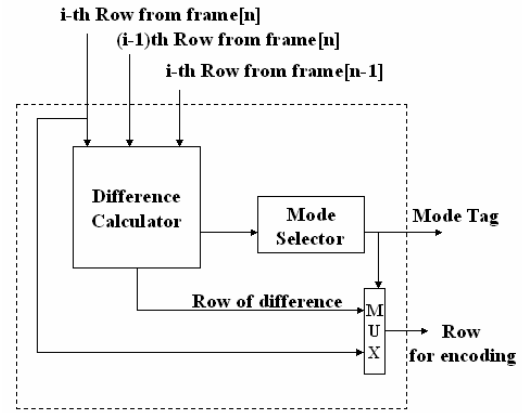


Figure 2. Line-based preprocessing.

```

if  $SAD_{inter} = 0$ 
  Enter Skip Mode and Skip the  $i$ -th line without coding
else if  $SAGD_{inter} = 0$ 
  Enter DC Mode and process the  $i$ -th line by one pixel value
else if  $SAGD_{inter} < SAGD_{intra}$ 
  Enter Difference Mode and take the difference row into coding process
else
  Enter Raw Mode and take the current row into coding process

```

Figure 3. Mode decision in the preprocessing.

show the line-based preprocessing with zero-motion prediction and mode decision algorithm in Figure 3. For optimal mode decision, switching of the raw data mode and difference mode is based on the observations on the absolute gradient difference (AGD) in Figure 5. The distributions of AGD are derived via an exhaustive search for the best coding mode of each line. When the raw data mode is preferred, the probability of small AGD(intra) is higher than that of small AGD(inter).

2.2. Line-Based Spatial Prediction

With the preprocessing, we choose the best coding mode for each line within the current frame. To further improve

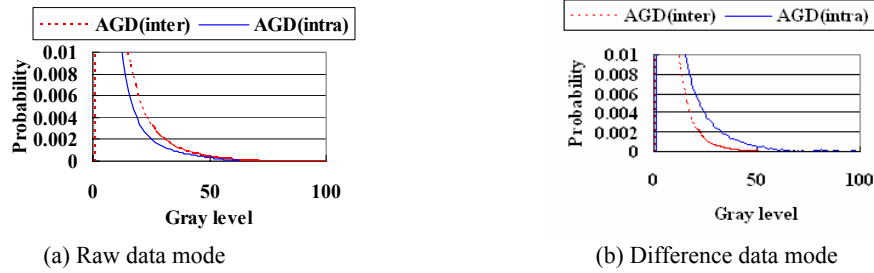


Figure 5. Probability distribution of AGD(inter) and AGD(intra) for Foreman sequence in CIF resolution.

the coding efficiency based on close relationship of spatially successive lines, a line-based intra-frame prediction method consisting of adaptive reference line and simplified context modeling is introduced.

2.2.1 Four line-based coding modes

Four modes including skip mode, difference mode, raw data mode and DC mode are used to categorize the output of the line comparator at the encoder. For decoding each line, 2-bit overhead is used to represent the coding mode.

A. Skip Mode (codeword ‘00’ in the syntax)

In the current frame, the line that is the same as the one in the previous frame is skipped. Side information is transmitted to the decoder to reconstruct the line by directly duplicating pixels from the previous frame.

B. Difference Mode (codeword ‘01’ in the syntax)

The difference mode is applied when $SAGD_{inter}$ is smaller than $SAGD_{intra}$ in magnitude. In this case, we take the residuals of zero motion prediction for coding. To save the sign bits of the prediction residuals, the same pixel value re-mapping approach adopted in JPEG-LS is used. If the image alphabet size is α ($= 2^8$ for a gray-level image), remapping the large prediction residual ($-\alpha \leq \varepsilon \leq \alpha$, ε is the current residual) into the small and unsigned one ($0 \leq \varepsilon \leq \alpha$) can decrease the value range without violating the data correction since the prediction value is known in the decoder [2].

C. Raw Data Mode (codeword ‘11’ in the syntax)

As $SAGD_{intra}$ is smaller than $SAGD_{inter}$, spatial coding algorithm is used. In this mode, only intra-frame prediction is taken on the original pixels.

D. DC Mode (codeword ‘10’ in the syntax)

When $I_{inter}^D(i, j)$ is identical for all pixels in a line, single value can represent DC offset. With the DC offset, the decoder can reconstruct the pixels of whole line.

2.2.2 Inter-Line Prediction

In LALVC, we use a simple template in Figure 4 to predict



Figure 4. The template of intra-frame prediction and context modeling.

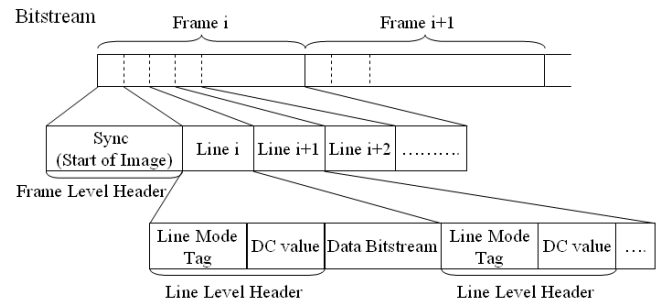


Figure 6. Syntax of CAVLC file format.

the pixel of the same frame with a fixed predictor.

$$\hat{x} = \begin{cases} \min(a, b) & \text{if } c \geq \max(a, b) \\ \max(a, b) & \text{if } c \leq \min(a, b) \\ a + b - c & \text{otherwise} \end{cases} \quad (5)$$

With this template, only two lines are buffered for advanced processing including the error bias cancellation in context modeling [2] and run mode extension. To simplify the context modeling, we use two gradients ($b-c$) and ($c-a$) for the error bias cancellation. To avoid performance degradation of inter-line prediction when the coding mode transition occurs, we adopt raw value reference lines and difference value reference lines for the raw data mode and the difference mode respectively. This adaptation is applied prior to the gradient computation in (5).

2.3. Syntax for Proposed Algorithm

We define the syntax for perfect LALVC decoding that supports the four line-based coding modes. In Figure 6, the bits for the frames are separated by the ‘Sync’ field. For line-level syntax, ‘Line Mode Tag’ is set in the preprocessing. In addition, the field ‘DC value’ can further enhance the coding efficiency when the brightness or color saturation variation of whole image pixels is adjusted. With the selected coding mode, the prediction residuals are generated and then passed through the entropy coder to form the data bitstream. With the syntax, the LALVC

decoding process reconstructs the sequence.

3. EXPERIMENT RESULTS

To evaluate the LALVC performance, we take both natural and computer-generated video sequences. For natural video sequences, we use test sequences for MPEG standards. The natural video sequences are in CIF (352x288) and YCbCr=4:2:0 formats. Each pixel is presented in 8 bits. Only Y component is used for simulation. The synthetic video sequences including scrolling web pages and adapting window size are captured from the computer screen. The synthetic video sequences are in CIF resolution and RGB format. The comparisons of compression ratio are shown in Table 1 for first 6 natural sequences and remaining 3 synthetic sequences. The performance of context modeling, intra predictor and entropy coder in LALVC is compressed with those within JPEG-LS, CALIC and WinRAR. WinRAR is commercial software for generic data compression. Prior to encoding, each input sequence is changed by zero motion prediction and sign remapping to form a new sequence with unsigned prediction residuals. Table 1 gives the comparisons based on the total amount of bits used to represent the new sequences. LALVC has the best performance in the natural sequences on the average. As to the computer-like sequences, LALVC can perform better than JPEG-LS and CALIC on the average. WinRAR is better than LALVC in the synthetic video sequences.

LALVC is low complexity, which is proper for software implementation and hardwired circuit design. For the preprocessing, if the frame resolution is $N*N$, LALVC needs $(7N^2-6N)$ adders to execute the mode decision and zero-motion prediction residual production. In addition, we have reduced the context model from three to two gradients to achieve identical performance with less memory requirement. The fixed predictor and entropy coder have been proven to have low complexity in JPEG-LS [2]. The execution time of LALVC algorithm is tested for P4 2.0 GHz Desktop and Windows XP. The results show that the encoding rate is about 55 frames per second (fps) for nature CIF sequences and near 133 fps for computer-generated CIF sequences. The averaged decoding rates of LALVC are 70 fps and 584 fps for the natural video and synthetic video bitstreams respectively.

4. CONCLUSION

In this paper, we have presented a low-complex and low-latency LALVC algorithm for real-time interaction applications for universal multimedia access environment. The LALVC consists of three parts covering the preprocessing, the mode-dependent spatial prediction and coding prediction. The preprocessing can efficiently remove the temporal redundancy via the zero-motion prediction and

Table 1 Compression ratios for various lossless coding methods of still images.

	WinRAR	JPEG-LS	CALIC	LALVC
Akiyo_Y	5.85	6.74	3.43	6.34
Bus_Y	1.29	1.44	1.09	1.64
Football_Y	1.67	1.75	1.28	1.99
Foreman_Y	1.80	1.84	1.26	1.96
Mobile_Y	1.32	1.36	1.07	1.40
Silent_Y	2.52	2.61	1.49	2.53
Browser_R	30.28	12.84	18.20	20.80
Browser_G	29.73	13.69	18.64	22.15
Browser_B	29.46	15.07	19.16	24.32

optimized mode decision that can select the best predictor to remove most of temporal prediction for various video sources. The reduced context modeling can reduce spatial redundancy. In addition, the syntax is simple and easy for the encoder and decoder implementation. The simulation results show that LALVC has significant compression ratios for both natural and synthetic sequences. The property of low complexity and low delay is suitable for real-time and low power applications, which makes LALVC useful for hardware realization such as FPGA implementation or ASIC design.

5. REFERENCES

- [1] X. Wu, and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Trans. on Communications*, vol. 45, no. 4, pp.437-444, Apr. 1997.
- [2] M.J. Weinberger, G. Seroussi and G. Sapiro, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *IEEE Trans. on Image Processing*, vol.9, no.8, pp.1309-1324, Aug. 2000.
- [3] R. Barequet and M. Feder, "SICLIC: a simple inter-color lossless image coder," *Proc. DCC '99*, 29-31 March 1999, pp.501-510.
- [4] Xiaolin Wu and N. Memon, "Context-based lossless interband compression-extending CALIC," *IEEE Trans. on Image Processing*, vol.9, no.6, pp.994-1001, June 2000.
- [5] E.S.G. Carotti, J.C. De Martin and A.R. Meo, "Backward-adaptive lossless compression of video sequences," *Proc. ICASSP'02*, vol.4, vol.4, 13-17 May, 2002, pp.3417-3420.
- [6] D. Brunello, G. Calvagno, G.A. Mian and R. Rinaldo, "Lossless compression of video using temporal information," *IEEE Trans. on Image Processing*, vol.12, no.2, pp. 132-139, Feb. 2003.
- [7] N.D. Memon and K. Sayood, "Lossless compression of video sequences," *IEEE Trans. on Communications*, vol. 44, no.10, pp.1340-1345, Oct. 1996.