SNR Scalability Based on Wavelet

Yen-Yu Chen
Department of Information Managemnet
Chungchou Institute of Technology
miscyy@tcts1.seed.net.tw

Ying-Wen Chang
Department of Information Managemnet
Chungchou Institute of Technology
ywchang@dragon.ccut.edu.tw

Abstract

Scalable video coding methods and flexible streaming approaches are required for adapting changing network conditions in real time. This work presents scalable video coding based on a wavelet scheme. With respect to encoding, a scalable video encoder encodes predictive residual coefficients into several layers of a bit-stream. The bit-stream of the enhancement layer is truncated at various bit-rates, according to the bandwidth of the network. The residual image is reordered to form wavelet blocks. Wavelet blocks are classified into three types, and the most significant coefficients in each type of block are preserved. After the coefficients are preserved, significant coefficients are encoded using the concept of the significant link. With respect to decoding, the bit-stream of the enhancement layer is combined with the bit-stream of the base layer to reconstruct a better quality video. Simulation shows that the proposed method outperforms MPEG-4 FGS by around 0.5 dB in terms of PSNR.

Key words: scalable video coding, MPEG4, FGS

1. Introduction

The real-time transmission of live video or stored video dominates real-time multimedia. This investigation is concerned only with stored video. In an Internet streaming video system [1-6], the video server operates between the encoder and the channel. The encoder is no longer aware of the channel capacity and does not know at which bit rate the video quality is optimal. Also, the decoder may not be able to decode all of the bit-stream received from the channel sufficiently quality enough to reconstruct the video signal. The bitstream is partially decoded at any bit rate within a range, to reconstruct a video signal with the optimal quality at that bit rate.

Several streaming solutions have used variations of scalable (layered) video coding methods [7-10] and typically have been complemented by packet loss recovery and/or error resilience mechanisms to the relatively high packet-loss rate normally encountered on the Internet. In MPEG-2 [11-16] and MPEG-4 [17-18], many layered techniques [19-21], namely SNR scalability, temporal scalability, and spatial scalability, are included. These technologies code a video sequence in a base layer and an enhancement layer. The enhancement layer bitstream must be both completely received and decoded to ensure that it can enhance the quality of the video.

2. Proposed Algorithm

Figure 1 illustrates the main procedures. First, the positions of significant coefficients must be determined. Secondly, the most important coefficients are preserved and the less important coefficients are eliminated. Finally, the coefficients are encoded as a bit-stream by the zero-tree encoding method. The bit-stream is delivered to the decoder over best-effort network. The following subsections describe each process.

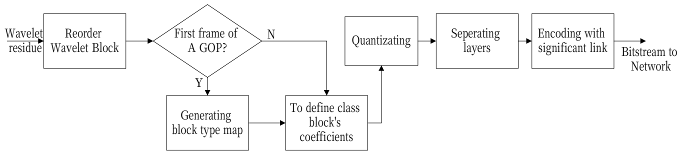


Figure 1 Overview of the algorithm

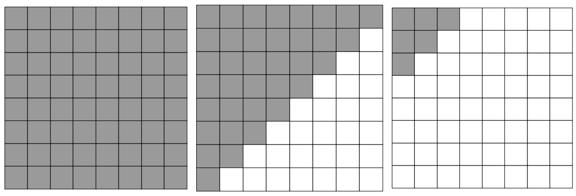
2.1 Reorganized and Classification wavelet block

Quadtree decomposition is a simple technique for representing images at various resolutions. The decomposition clarifies the reorganization of coefficients of each wavelet tree into a wavelet block. The lower band coefficients are centralized in the upper left of each wavelet block

For three-level DWT images, all individual 8×8 wavelet blocks are stored for further processing. First, the square sum of each wavelet block in the first frame of a GOP is computed. Then, the wavelet blocks are classified into three types by the square sum: type 1 blocks have larger square sums than type 2, and type 2 have larger square sums than type 3. The square sum *T* is given by

$$T = \sum_{j=1}^{N}\sum_{i=1}^{N} block^2(i, j) \quad \ldots\ldots(1)$$

Figure 2 depicts the method of preservation. Of type 1, all coefficients are preserved, while of types 2 and 3, one half and one-tenth of coefficients are preserved, respectively.



(a) Type 1 (b) Type 2 (c) Type 3

Figure 2 Coefficient preservation by type

2.2 Quantization

The low band coefficients contribute greatly to the PSNR of the video sequence, and are preserved using a lower quantization factor, Q_1. Simulations results indicate that more of the larger absolute coefficients are clustered in HL and LH bands than in the HH band. The HH band coefficients are quantized by a larger factor, Q_3, to preserve the edge information. A smaller factor, Q_2, is applied in other places in the wavelet block to reduce the amount of transmitted information at the proper bit-rate. Another reason is that if a wavelet coefficient at a coarse scale is significant with respect to a given threshold *T*, then all wavelet coefficients with the same orientation at the same spatial location, on finer wavelet scales, are also likely to be significant with respect to *T* in the natural image but not in the wavelet residue image.

2.3 Separating Layers

The advantages of layer coding are such that the residue coefficient is converted into several sub-streams. The coefficients are divided by a power of 2. The data each layer are 0, 1 or –1, expressed in binary using two bits. The method efficiently reduces the data transmission overhead. A simple example is presented below.

Wavelet coefficients	10	0	-6	0	3	2	0	8	-5	1	-5	3	-9
L1 (2^3)	1	0	0	0	0	0	0	1	0	0	0	0	-1
L2 (2^2)	0	0	-1	0	0	0	0	0	-1	0	-1	0	0
L3 (2^1)	1	0	-1	0	1	1	0	0	0	0	0	1	0
L4 (2^0)	0	0	0	0	1	0	0	0	-1	1	-1	1	-1

After the data of each layer are separated, the encoding method is applied to generate the transmission bit-streams.

2.4 Significant Link

The concept of significance-link is adopted and four symbols are redefined to encode the shape of a cluster: LINK, POS, NEG and ZERO. POS/NEG represents the positive/negative significant coefficient. ZERO represents an insignificant coefficient. LINK indicates the presence of a significance-link. Two lists of coefficients are maintained by the algorithm: LSC (list of significant coefficients) and LCC (list of child clusters). $c_n[x,y]$ represents the coefficient at position [x,y] in frame n.

The steps of a t-level DWT residue image encoding algorithm are described as follows:

Main_Enc( )

Step1 ∶for the subband LL_t, each position (x,y) of LL_t .

if c[x,y] is a significant coefficient and has not been encoded, then output POS/NEG and put (x,y) to LSC ;

Otherwise output ZERO.

Step2 ∶ for the subband LH_t, for each position (x,y) belong to LH_t, //LH_t,HL_t,HH_t

Step2.1 ∶ if c[x,y] is a significant coefficient and has not been encoded, then encode (x,y) output POS/NEG and put (x,y) to LSC

Step2.2 ∶ output encode (x,y) ZERO

Step2.3 ∶ if any one child of (x,y) is a significant coefficient and has not been encoded, then output LINK and put children of (x,y) to LCC;

Step 2.4 ∶call ENC_LCC( )

. Step 3 ∶ encode the output symbol as bit-streams.

END_enc( )

ENC_LCC( )

while LCC is not empty, remove (x,y) from LCC.

Step 1 ∶ for △x=0,1, △y=0,1,

Step1.1 ∶ if c(x+△x, ,y +△y) is a significant coefficient and has not been encoded, then output POS/NEG and put (x,y) to LSC, go to step 1.2Otherwise output ZERO.

Step1.2 ∶if any one child of (x+△x, ,y +△y) is a significant coefficient and has not been encoded then output LINK put children of (x+△x, ,y +△y) to LCC;

Step 2 ∶call ENC_LCC( )

END ENC_LCC( )

As in most image compression algorithms, the final step includes entropy coding, which employs adaptive arithmetic coding. The decoding algorithm is straightforward and can obtained by simply reversing the encoding process.

3. Simulation Results

Some test sequences are used in an experiment to establish the performance of the proposed. Two 352×288 sequences, “Foreman”, and “Coastguard” are used, and a total 300 frames in YUV color format 4∶2∶0 are employed. In the coding scheme, the video sequences are encoded in two parts the base layer and the enhancement layer. An MPEG-4 encoder encodes these sequences in the base layer, such that the base layer bit-streams can be transmitted over a band-limited channel. The proposed method is used to code the enhancement layer and improve visual performance. The resulting enhancement layer bit-streams can be transmitted at any bit-rate.

The reconstructed frames from the base layer bit-streams encoded by MPEG-4 at 128K bit/s with TM5 rate control and only the first frames, are I-frames while the others are P-frames. The average PSNR are 29.8597 dB and 26.7939 dB for Foreman and Coastguard, respectively. Notably the base layer bit-streams yield the minimum quality of bandwidth adaptation. In the sequence “Foreman”, the camera is first set on a talking man and then moved to a building. These moving frames suffer from serious blocking artifacts. In the sequence “Coastguard”, the camera moves with the boat such that the boat is always in the center of the frames. However, regions of ripple around the boat and bushes around the coast are blurred because of compression.

Table 1. Simulation results of proposed method and FGS.

Enh. Layer Bit-Rate	Foreman		Coastguard	
	Proposed	FGS	Proposed	FGS
64K	31.92	31.17	27.89	27.75
128K	32.19	31.89	28.98	28.52
192K	32.69	32.51	29.56	29.27
256K	33.21	33.14	29.80	29.79

For comparison, Table 1 lists the simulated results of the proposed method and MPEG4 FGS. By observation, the proposed method always outperforms FGS. Figure 3 list the reconstructed PNSR of Foreman by adding enhancement layer at (a) 64 k bit/s (b) 128k bit/s (c) 192 k bit/s (d) 256 k bit/s. Figure 4 list the reconstructed PNSR of Coastguard by adding enhancement layer at (a) 64 k bit/s (b) 128k bit/s (c) 192 k bit/s (d) 256 k bit/s. In regions of motion, motion vector is normally large and motion compensation may not suffice to stand for these regions after DCT and quantization. In the base layer, the blocking effect is very serious in these regions at a low bit rate. After additional information, an enhancement layer, is added, these blocking effects are efficiently removed and PSNRs are thus improved.

(a)

(b)

(c)

(d)

Figure 3 Reconstructed image in Foreman (frame 255) obtained by adding enhancement layer at (a) 64 k bit/s (b) 128k bit/s (c) 192 k bit/s (d) 256 k bit/s

(a)

(b)

(c)

(d)

Figure 4 Reconstructed image in Coastguard (frame 69) obtained by adding enhancement layer at (a) 64 k bit/s (b) 128k bit/s (c) 192 k bit/s (d) 256 k bit/s

4. Conclusion

This work presents a scalability algorithm based on DWT. It use significant link to preserve the relationship of significant coefficients. The more important information can be put in the front of the bistream. The bistream would be truncated to adapt to the network bandwidth varying. Vital coefficients are always received by decoder end and the better quality of video sequence can be reconstructed and play. In a simulation, the PSNR of the reconstructed video with the enhancement layers is enhanced to 2.5 dB over that of the reconstructed video with only the base layer bit-stream. Moreover, the proposed method outperforms MPEG-4 FGS by around 0.5 dB in terms of PSNR.

5. Acknowledgement

This research is supported partially by the National Science Council of Taiwan under the contract number of NSC 92-2416-H-235-001

References

[1] Dapeng Wu, Tiwei Thomas Hou, Wenwu Zhu, Ya-Qin Zhang, and Jon M. Peha, "Streaming Viedo Over The Internet: Approaches and Directions," *IEEE Trans. on Circuit and Systems for Video Technology*, vol. 11, no.3, pp. 282-300 March 2001.

[2] H.Radha,Y.Chen,"Fine-Granular-Scalable Video for Packet Networks," *Packet Video'99*, Columbia University, NY, April 1999.

[3] H. Radha, Y. Chen, K. Parthasarathy, R. Cohen, "Scalable Internet Video Using MPEG-4," *Signal Processing: Image Communication*, pp. 95-126, September 1999.

[4] S.-J. Choi, J. W. Woods, "Motion-Compensated 3-D Subband Coding of Video " *IEEE Trans. on Image Process*, pp. 155-167, February 1999.

[5] S. McCanne, V. Jackobson, and M. Vetterli, "Receiver-driven Layered Multi-cast," *Proc. SIGCOMM'96*, Standford, CA, pp. 117-130, Aug. 1996.

[6] S. McCanne, V. Jackobson, and M. Vetterli, "Low-Complexity Video Coding for Receiver-Driven Layered Multicast," *IEEE JSAC*, vol.16, no.6, Aug. 1997,pp. 983-1001.

[7] W.Tan and A.Zakhor, "Real-Time Internet Video Using Error Resilient Scalable Compression and TCP-Friendly Transprot Procotol," *IEEE Trans. on Multimedia*, vol.1, no.2, pp.172-186, June 1999.

[8] A. Eleftheriadis and D. Anastassiou, "Meeting Arbitrary QoS Constraints Using Dynamic Rate Shaping of Coded Digital Video," *in Proc. 5th Int. Workshop Network and Operating System Support for Digital Audio and Video (NOSSDAV'95)*, pp.95-106, Apr. 1995.

[9] X. Wang, and H. Schulzrinne, "Comparison of Adaptive Internet Multimedia Applications," *IEICE Trans. Communication.*, vol.E82-B, no. 6, pp. 806-818, June 1999.

[10] Q. Zhang, G. Wang, W. Zhu, and Y. Q. Zhang, "Robust Scalable Video Streaming Over Internet with Network-adaptive Congestion Control and Unequal Loss Protection," in *Proc. Packet Video Workshop*, Kyongju, Korea, Apr. 2001.

[11] T.Sikora, "MPEG Digital Video-Coding Standards," *IEEE Signal Processing Magazine*, September 1997.

[12] V.Bhaskaran and K.Konstantinides, “Image and Video Compression Standards: Algorithm and Architectures,” *Kluwer Academic Publishers*, 1995.

[13] ISO/IEC 13818-2 “Generic Coding of Moving Pictures and Associated Audio Information: Video,” May 1994.

[14] J. Black, W. Pennebaker, C. Fogg, and D. LeGall “MPEG Video Compression Standard, Digital Multimedia Standards Series,” *Chapman and Hall*, 1996.

[15] K. Rao and J. Hwang, “Techniques and Standards for Image, Video and Audio Coding ,” *Prentice Hall*, 1996.

[16] B. Haskell, A. Puri, and A. Netravali “Digital Video: An Introduction to MPEG-2,” *Chapman and Hall*, 1997.

[17] ISO/IEC/JTC1/SC29/WG11 N1642, “MPEG-4 Video Verification Model Version 7.0,” April 1997.

[18] ISO/IEC JTC1/SC29/WG11 N3908 “January 2001/Pisa MPEG-4 Video Verification Model version 18.0”

[19] Weiping Li, “Overview of Fine Granularity Scalablility in MPEG-4 Video Standard,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11 no. 3 pp. 301-310 March 2001.

[20] Hayder M. Radha, Mihaela van der Schaar, and Yingwei Chen,”The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP,” *IEEE Trans. on Multimedia* vol. 3 no.1 pp. 53-68, March 2001.

[21] Feng Wu, Shipeng Li, Ya-Qin Zhang, “DCT-Prediction Based Progressive Fine Granularity Scalable Coding,” *IEEE Image Processing Proc. International Conference*, vol. 3 pp.556-559, 2000.