รายการอ้างอิง

1. F. Wu, S. Li, and Y. Q. Zhang. A framework for efficient progressive fine granularity scalable video coding. IEEE Trans. Circuits and System for Video Technology Vol. 11 (March 2001): 332-344.
2. Marta Mrak, Nikola Sprljan, Sonja Grgic. Video coding schemes for transporting video over the internet. Faculty of Electrical Engineering and Computing. University of Zagreb. Technical Report (in Croatian) Zagreb. 2001. (Mimeographed)
3. Thomas Sikora. The MPEG-4 video standard verification model. IEEE Trans. Circuits and System for Video Technology Vol. 7 (February 1997): 19-31.
4. MPEG-4 video verification model version 18.0. ISO/IEC JTC1/SC29/WG11 N3908. Pisa. (January 2001).
5. Hayder M. Radha, Member. The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP. IEEE Trans. On Multimedia Vol.3. No.1 (March 2001): 53-68.
6. M. van der Schaar and H. Radha. Motion compensation fine-granularity-scalability for wireless multimedia. IEEE Workshop Multimedia Signal Processing Vol.10 (October 2001): 453-458.
7. Thomas Sikora. MPEG-4 very low bitrate video [Online]. Available from: http://wwwam.hhi.de/mpeg-video/papers/sikora/vlbv.htm [2002, November 15]
8. Thomas Sikora. MPEG-4 very low bit rate video. IEEE International Symposium on Circuits and Systems Vol.2 (June 1997): 1440 –1443.
9. Thomas Sikora. MPEG-4 video standard for multimedia applications [Online]. Available from: http://wwwam.hhi.de/mpeg-video/papers/sikora/ISCAS/IEEE-ISCAS.htm [2002, November 15]
10. Thomas Sikora. The MPEG-4 video image coding scheme [Online]. Available from: http://mpeg.telecomitalialab.com/standards/mpeg-4/mpeg-4.htm [2002, November 15]

11. Aggelos K. Katsaggelos, Lisimachos P. Kondi, Fablan W. Meier, Jorn Ostermann and Guido M. Schuster. MPEG-4 and rate-distortion-based shape-coding techniques. Proceedings of the IEEE Vol.86 No.6 (June 1998): 1126-1154.

12. S.Lawson and J. Zha. Image compression using wavelets and JPEG2000: a tutorial. Electronics & Communication Engineering Journal (June 2002): 112-121.

13. Information Technology-Coding of Audio-Visual Obects Part 2: Visual ISO/IED 14496-2:1999/ FDAM 4. ISO/IEC JTC1/SC29/WG11 N3904. January 2001.

14. Thomas Sikora. The MPEG-4 video image coding scheme [Online]. Available from: http://www.mpeg.org/MPEG/MSSG/tm5/Ch10/Ch10.html [2002, November 15]

15. W. Li. Overiew of fine granularity scalability in MPEG-4 video standard. IEEE Trans. Circuits and System for Video Technology Vol. 11 (March 2001): 301-317.

16. Hsiang-Chun Huang, Chung-Neng Wang, and Tihao Chiang. A robust fine granularity scalability using trellis-based predictive leak. IEEE Trans. Circuits and System for Video Technology Vol. 12 (June 2002): 372-385.

17. Martyn J. Riley, Iain E. G. Richardson. Digital video communications. Boston, London. Artech house. 1997.

18. Feng Wu, Shipeng Li, Xiaoyan Sun, Rong Yan, Ya-Qin Zhang. Comparisons between the one-loop and two-loop solutions for improving the coding eficiency of FGS. MPEG-4. 2001 Proceedings of Workshop and Exhibition (June 2001): 79-82.

19. M. van der Schaar and H. Radha. Adaptive motion-compensation fine-granular-scalability (AMC-FGS) for wireless video. IEEE Trans. on Circuit and Systems for Video Technology Vol.12 No.6 (June 2001): 360-371.

บทความทางวิชาการที่ได้รับการเผยแพร่

1. Robust Fine Granularity Scalability using Leaky Prediction for Low Bit-Rate Video Coding, ICT 2003, Assumption University, Bangkok, Thailand, pp. 132-136, 4-6 April 2003.

2. ได้รับการตอบรับงานประชุมวิชาการทางวิศวกรรมไฟฟ้า ครั้งที่ 26 (EECON-26) สำหรับบทความ "Improvement of Fine Granularity Scalability Using Adaptive Leaky Prediction for Low Bit-Rate Video Coding" กำหนดจัดการประชุม โดย ภาควิชาวิศวกรรมไฟฟ้า และ ภาควิชาเทคโนโลยีไฟฟ้าอุตสาหกรรม คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าพระนครเหนือ ในระหว่างวันที่ 6-7 พฤศจิกายน 2546

Robust Fine Granularity Scalability Using Leaky Prediction for Low Bit Rate Video Coding

S. Artnarongkorn, D. Tancharoen, S. Aramvith and S. Jitapunkul

Digital Signal Processing Research Laboratory, Department of Electrical Engineering
Chulalongkorn University, Bangkok, Thailand
E-Mail: sujanya@chula.com

Abstract

Transmission of video over bandwidth varying networks like the Internet requires a highly scalable solution capable of adapting to the network condition in real-time. To address this requirement, the MPEG-4 committee has approved the MPEG-4 fine granularity scalability (FGS) profile as a streaming video tool. This paper proposes a robust fine granularity scalability based on MPEG-4 video coding scheme by using leaky prediction to improve the temporal prediction. Leaky prediction is used to increase error robustness by trading off coding efficiency. The simulation results show that the proposed algorithm has 2.15 dB gain in PSNR over the original FGS for Carphone sequence.

Keywords: Fine Granularity Scalability (FGS), MPEG-4, Internet video streaming, leaky prediction

1. Introduction

Multimedia distribution over the Internet is becoming increasingly popular. However, the Internet was designed for computer data communication, and satisfy the necessary requirements for the effective delivery of multimedia streams, that poses significant challenges. Transmitting digital video over the Internet encounters two major problems: bandwidth variation due to heterogeneous access-technologies of the receivers (e.g., analog modem, cable modem, xDSL, etc.) or due to dynamic changes in network conditions (e.g., congestion events) and packet-losses or errors. It is very desirable to have a video-coding scheme that can adapt to the channel conditions [1][2].

Compared to video codecs for CD-ROM coded by MPEG-1 or TV broadcast coded by MPEG-2, codecs designed for the Internet require greater scalability, lower computational complexity, greater resiliency to network losses, and lower encode/decode latency for video conferencing. New algorithms specifically targeted at Internet video are being developed. Most recent efforts on video compression for streaming video have been focused on scalable video coding, which is included in MPEG-4 standard. The primary objectives of on-going research on scalable video coding are to achieve high compression efficiency at affordable cost and acceptable complexity. Scalable video coding is capable of coping with bandwidth variation [3].

Several scalable coding methods have been successfully proposed for video transmission through heterogonous networks. One of these techniques is the MPEG-4 Fine-Granular Scalability (FGS) scheme, that can adapt in real-time (i.e., at transmission time) to the bandwidth variation over heterogonous networks. Some key advantages of the MPEG-4 FGS framework are its packet-loss resilience and flexibility in supporting streaming application. Recently, techniques that include a certain amount of enhancement layer information into the prediction loop have been proposed to improve the coding efficiency of FGS [2][4]. However, to minimize the effect of drift introduced by prediction mismatch, higher computational effort is required at the decoder.

In this paper, we propose a novel scheme for improving motion-compensation in FGS coding by introducing leaky-prediction technique into prediction loop. The remainder of this paper is organized as follows. In Section 2, we describe the video coding scheme. Then the proposed scheme is described in Section 3. In Section 4, the simulation results of the improvements of video quality obtained by using the newly proposed scheme are presented. Finally, the conclusion is outlined in Section 5.

2. Video Coding

2.1. Layered video coding

A non-scalable video encoder generates one compressed bitstream. In contrast, a scalable video encoder compresses a raw video sequence into multiple layers. One of the compressed layers is the base layer, which can be independently decoded and provide coarse visual quality. Other compressed layers are enhancement layers, which can only be decoded together with the base layer and can improve visual quality. Therefore, the complete bitstream (i.e., the combination of all layers) provides the highest quality.

2.2. Fine granularity scalability

The basis idea of FGS is also to code a video sequence into a base layer and an enhancement layer. The FGS encoder and decoder are depicted in Fig. 1 and 2, respectively. The base layer uses nonscalable coding scheme to reach the lower bound of the bit-rate range and the difference between the original picture and the reconstructed picture is coded by using bit-plane coding of the DCT coefficients into the enhancement layer. The bitstream of the FGS enhancement layer may be truncated within any number of bits per picture after encoding completion. The decoder should be able to reconstruct an enhancement video from the base layer and the truncated enhancement layer bitstreams. At the enhancement layer, video quality is proportional to the number of bits decoded by the decoder for each picture.

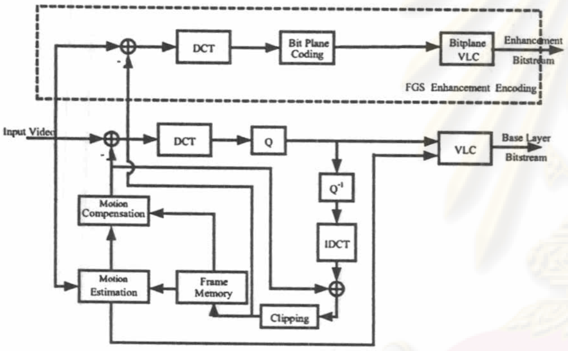


Figure 1. The FGS encoder

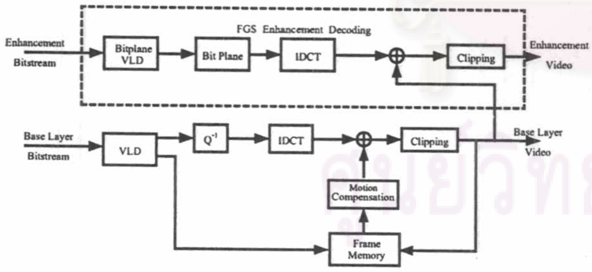


Figure 2. The FGS decoder

Bitplane coding of the DCT coefficients is used as the basic coding technique for FGS. The FGS enhancement layer encoder takes the original frame and reconstructed frame as input and produces a FGS enhancement bitstream. The difference between the original and reconstructed frames is transformed by the DCT to generate the DCT coefficients. After obtaining all DCT coefficients of a frame, bitplane shift operation can be performed. Then the maximum absolute value of the DCT coefficients is obtained and the maximum number of bitplanes for each DCT block is zigzag ordered into an array. A block bitplane is formed as an array of 64 bits and a block is taken one from each absolute value of the DCT coefficients at the same significant position. In order to cover a wide range of bit rate, there is a need to combine FGS with temporal scalability so that not only picture quality can be scalable but also temporal resolution (frame rate) is able to scale [5][6][7].

2.3. Comparison of scalable video coding techniques

The objective of video coding for Internet is to optimize the video quality over a given bit rate range. The bitstream should be partially decodable at any bit rate within the bit rate range to reconstruct a video signal with the optimized quality at that bit rate. The distortion-rate curve in Fig. 3 indicates the upper bound in quality for any coding technique at any given bit rate. Layered scalability techniques change the nonscalable single staircase curve to a curve with two stairs. The desired objective is to achieve the continuous curve that parallels the distortion rate curve with a single bitstream and it is done with FGS video coding technique.

The major difference between FGS and the layered scalable coding techniques is that, the FGS coding technique can code a video sequence into two layers, and the enhancement bitstream can also be truncated into any number of bits within each frame to provide partial enhancement proportional to the number of bits decoded for each frame. As a result, FGS provides the continuous scalability curve illustrated in Fig. 3 [7].

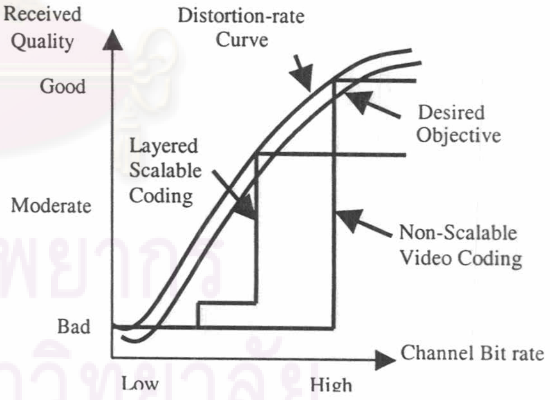


Figure 3. Video coding performance

3. Improved MPEG-4 FGS Codec Using Leaky Prediction

3.1. FGS Encoder

We propose a novel scheme for improving motion-compensation in FGS coding by introducing the notion of the extended base layer that includes an integrated base and enhancement layer. It is very clear that the reconstructed of references have higher quality than that of the original FGS scheme. As a result, when the network bandwidth is available enough to transmit the enhancement to the decoder, this scheme can provide higher quality. But if the decoder can not receive the enhancement layer bitstream used for reconstruction the references due to the limited network bandwidth, the difference between reference used in the encoder and the decoder will cause drifting error and error propagation. Then, the leaky prediction technique that introduced into the motion-compensation prediction loop for decreasing any error, is presented in Fig. 4.

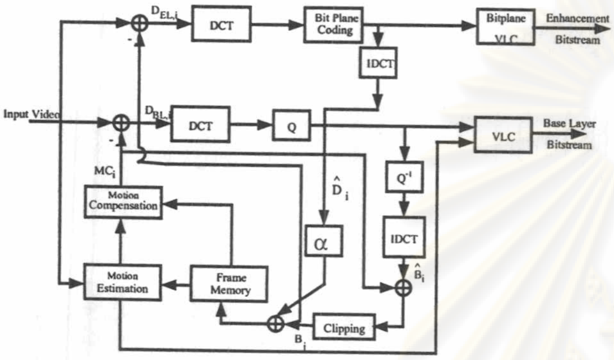


Figure 4. Structure of the proposed encoder

As illustrated in Fig. 4, we describe the technique form the base and enhancement layers. For the current frame, the original frame at time i is denoted F_i. At the base layer, the reconstructed frame of previous frame (at time i-1) is denoted B_{i-1} and residue error is denoted $D_{BL,i}$ (the subscript BL,i means the base layer at time i.). At the enhancement layer, the difference value of the previous frame between original and reconstructed frame is encoded and it is denoted $\hat{D}_{i-1}$. The leakage factor is denoted α. Then, the original frame at time i is shown in eq. (1)

$$F_i = (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + D_{BL,i} \tag{1}$$

By the subscript *mc* means that $(B_{i-1} + \alpha \hat{D}_{i-1})_{mc}$ is the motion compensation version of $(B_{i-1} + \alpha \hat{D}_{i-1})$. The coded version of the residue error as the base layer is denoted $\hat{B}_i$ and the quantization error is denoted $D_{EL,i}$. The relationship between $D_{BL,i}$, $\hat{B}_i$ and $D_{EL,i}$ is

$$D_{BL,i} = \hat{B}_i + D_{EL,i} \tag{2}$$

Substitute eq. (2) into eq. (1),

$$\begin{aligned} F_i &= (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + \hat{B}_i + D_{EL,i} \\ &= (B_{i-1})_{mc} + \hat{B}_i + (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \\ &= B_i + \hat{D}_i \end{aligned} \tag{3}$$

where

$$B_i = (B_{i-1})_{mc} + \hat{B}_i \tag{4}$$

$$\hat{D}_i = (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \tag{5}$$

The signal B_i and $\hat{D}_i$ will be used to predict next frame.

By expanding the recursive formula of $\hat{D}_i$ in eq. (5), we can obtain eq. (6)

$$\begin{aligned} \hat{D}_i &= (\alpha((\alpha \hat{D}_{i-2})_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \\ &= (\alpha((\alpha((\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2}))_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \end{aligned} \tag{6}$$

As demonstrated in eq. (6), it is obviously seen that any errors in the final residual $\hat{D}_i$ will be attenuated in this proposed framework.

3.2. FGS Decoder

The FGS decoder is shown in Fig. 5, the leaky prediction that was added at the enhancement layer with gain factor α (less than 1).

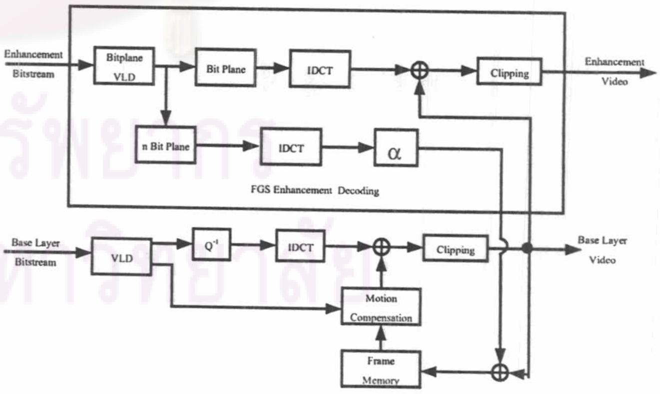


Figure 5. Structure of the proposed decoder

By assume that there is error at the enhancement layer for frame F_{i-2}, we denote the received enhancement layer bitstream $\bar{D}_{i-2}$ and the transmission error is denoted $\Delta\hat{D}_{i-2}$. Thus we have

$$D_{EL,i-2} = \bar{D}_{i-2} + \Delta\hat{D}_{i-2} \qquad (7)$$

and the reconstructed version of $\hat{D}_{i-2}$ is denoted $\tilde{D}_{i-2}$. Thus

$$\begin{aligned}\tilde{D}_{i-2} &= (\alpha\hat{D}_{i-3})_{mc} + \bar{D}_{i-2} \\ &= (\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2}. \qquad (8)\end{aligned}$$

Comparing eq. (6) and (8), the difference between $\hat{D}_{i-2}$ and $\tilde{D}_{i-2}$ is $\Delta\hat{D}_{i-2}$.

Now we back to the frame F_{i-1}. By assume that there is no error at the enhancement layer for frames F_{i-1} and F_i. Thus

$$\begin{aligned}\tilde{D}_{i-1} &= (\alpha\tilde{D}_{i-2})_{mc} + D_{EL,i-1} \\ &= (\alpha(\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2})_{mc} \\ &\quad + D_{EL,i-1}. \qquad (9)\end{aligned}$$

The difference between $\hat{D}_{i-1}$ and $\tilde{D}_{i-1}$ is $\alpha(\Delta\hat{D}_{i-2})$.

Now we move to the frame F_i and obtained

$$\begin{aligned}\tilde{D}_{i} &= (\alpha\tilde{D}_{i-1})_{mc} + D_{EL,i} \\ &= (\alpha((\alpha(\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2})_{mc} \\ &\quad + D_{EL,i-1}))_{mc} + D_{EL,i}. \qquad (10)\end{aligned}$$

The difference between $\hat{D}_i$ and $\tilde{D}_i$ is $\alpha^2(\Delta\hat{D}_{i-2})$.

3.3. Leaky prediction

Leaky prediction is a well-known technique to increase error robustness, which is a trade-off with coding efficiency [8]. In the proposed scheme, the first 3 bitplanes of the enhancement layer data is transformed back to the spatial domain using IDCT and is attenuated by a leakage factor α before added into the frame memory. Therefore, the drift or the difference between the encoder and decoder will be attenuated. If the leak factor is set to zero, the drift will be removed completely.

4. Simulation Results

The experimental results have been examined to show the performance of the proposed scheme. The MPEG-4 VM. 18.0 codec is used for the base layer [9]. From the fig. 6-9, the coding efficiency of the proposed scheme is compared with the original FGS coding. The carphone sequence in QCIF format is used for testing and every GOP has 20 frames of 1 I-frame and 19 P-frames. The bit rate of base layer based on the TM5 rate control is 32 kbps, and the frame rate is used at 30 frame/s. The leakage factor α was in the range of 0 and 1. If the leakage is more than 0.5, PSNR is mostly equal to the leakage at α=0.5 at high bit rate and PSNR is less than the leakage at α=0.5 at low bit rate. As the leakage is less than 0.5, PSNR is less than the leakage α=0.5 at high bit rate and we will obtain a little increase in PSNR if the leakage at α=0.5 at low bit rate. The best value of the leakage is 0.5. As shown in Fig. 6, the proposed scheme,

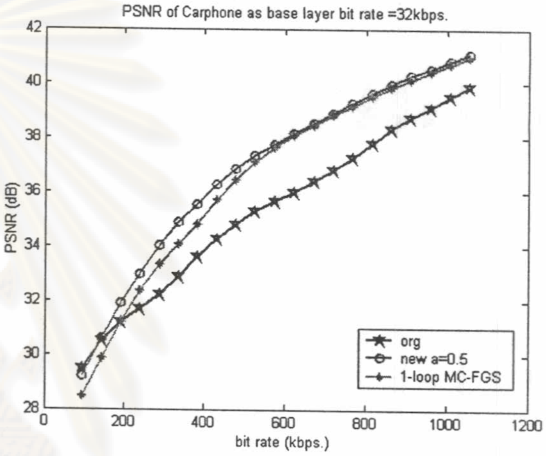


Figure 6. PSNR versus bit rate comparison between proposed scheme and original FGS of the Carphone sequence (The bit rate of base layer is 32kbps.).

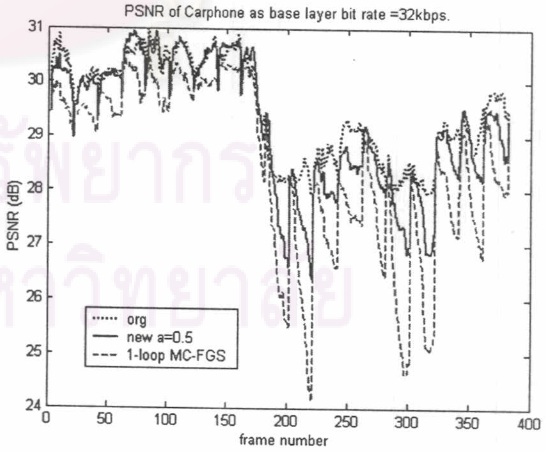


Figure 7. Comparison of the video quality in PSNR of the Carphone sequence between original FGS and proposed scheme (The bit rate of base layer is 32kbps and the total bandwidth is 96 kbps.).

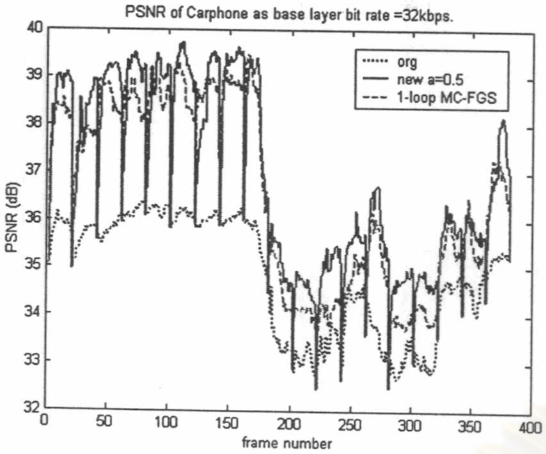
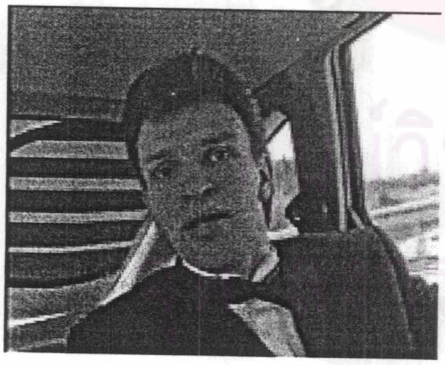


Figure 8. Comparison of video quality in PSNR of the Carphone sequence between original FGS and proposed scheme (The bit rate of base layer is 32kbps and the total bandwidth is 480 kbps.).

a. The original FGS coding



b. The proposed scheme

Figure 9. The 44th frame of Carphone sequence that was encoded by the original FGS and the proposed scheme. The bit rate of base layer is 32kbps and the total bandwidth is 480 kbps.

as set the leakage at $\alpha = 0.5$, has 2.15 dB gain in PSNR. Moreover the coding efficiency of all P-frames of the proposed method is higher than the original PSNR over the original FGS at medium and high bit rate and has 0.27 dB loss in PSNR at low bit rate. From Fig. 7, our results show that the coding efficiency of all I-frames have same quality for every coding method in this paper, and due to I-frame coding does not use temporal prediction. Then the error that is in previous P-frames of the proposed method is lower than the original FGS at high bit rate as depicted in Fig. 8. All frame, does not effect to I-frame (the I-frame insertion can also stop error propagation.).

From Fig. 9, comparison of the visual quality of the 44th frame of Carphone sequence between the original FGS coding and the proposed scheme. As the bit rate of base layer is 32kbps and the total bandwidth is 480 kbps. It is clearly observed around the face region that the visual quality of the area around the face region from the proposed scheme is better than that of original FGS.

5. Conclusions

In this paper, a novel FGS coding technique was presented to utilize the extended base-layer for prediction in order to improve coding efficiency. Error propagation due to the prediction mismatch or drift error is effectively controlled by leaky prediction in enhancement layers and I-frame insertion. Simulation results show that the proposed scheme, as set the leakage at $\alpha = 0.5$, has 2.15 dB gain in PSNR over the original FGS.

6. References

[1] Mihaela van der Schaar, "A hybrid temporal-SNR fine granularity scalability for Internet video," IEEE Trans. Circuits. Syst. Video Technol., vol. 11, pp. 318-331, March 2001.

[2] F. Wu, S. Li, and Y. Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," IEEE Trans. Circuits. Syst. Video Technol., vol. 11, pp. 332-344, March 2001.

[3] Marta Mrak, Nikola Sprljan, Sonja Grgic, "Video Coding Schemes for Transporting Video Over the Internet," Faculty of Electrical Engineering and Computing, University of Zagreb, Technical Report (in Croatian), Zagreb, 2001.

[4] M. van der Schaar and H. Radha, "Motion compensation fine-granularity-scalability for wireless multimedia," in Proc. IEEE Workshop Multimedia Signal Processing, vol.10, pp. 453-458, Oct. 2001.

[5] Hayder M. Radha, Member, "The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP," IEEE Trans. On Multimedia, vol.3, n.1, March 2001.

[6] Information Technology-Coding of Audio-Visual Obects Part 2: Visual ISO/IED 14496-2:1999/ FDAM 4, ISO/IEC JTC1/SC29/WG11 N3904, January 2001.

[7] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 301-317, March 2001.

[8] Hsiang-Chun Huang, Chung-Neng Wang, and Tihao Chiang, "A robust fine granularity scalability using trellis-based predictive leak," IEEE Trans. Circuits Syst. Video Technol., vol. 12, pp. 372-385, June 2002.

[9] MPEG-4 Video Verification Model version 18.0, ISO/IEC JTC1/SC29/WG11 N3908, Pisa, January 2001.

Improvement of Fine Granularity Scalability Using Adaptive Leaky Prediction for Low Bit Rate Video Coding

S. Artnarongkorn, S. Aramvith and S. Jitapunkul
Digital Signal Processing Research Laboratory, Department of Electrical Engineering,
Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand
Phone: (662) 2186915, Fax: (662) 2186912, E-Mail: sujanya@chula.com

Abstract

Transmission of video over bandwidth variation networks like the Internet requires a highly scalable solution capable of adapting to the network condition in real-time. To address this requirement, the MPEG-4 committee has approved the MPEG-4 fine granularity scalability (FGS) profile as a streaming video tool. In this paper, we propose an adaptive leakage factor algorithm of a leaky prediction to improve the motion compensation in low bit rate FGS coding. Simulation results show that the proposed algorithm achieve an increase of PSNR over the traditional FGS and one-loop motion compensation FGS.

Keywords: Fine Granularity Scalability (FGS), MPEG-4, Internet video streaming, adaptive leaky prediction

1. Introduction

Multimedia distribution over the Internet is becoming increasingly popular. However, the Internet is preliminary designed for computer data communication, and has yet to satisfy the necessary requirements for the effective delivery of multimedia streams. Transmitting digital video over the Internet encounters two major problems: bandwidth variation due to heterogeneous access-technologies of the receivers (e.g., analog modem, cable modem, xDSL, etc.) or due to dynamic changes in network conditions (e.g., congestion) and packet-losses or errors. This is very desirable to have a video-coding scheme that can be adapted to the varying channel conditions [1][2].

Comparing with video codecs for CD-ROM coded by MPEG-1 or TV broadcast coded by MPEG-2, codecs designed for the Internet require greater scalability, lower computational complexity, higher degree of error resilience to network losses, and lower codec latency for video conferencing. New algorithms specifically targeted at Internet video are being developed. Most recent efforts on video compression for streaming video have been focused on scalable video coding. The primary objectives of on-going research on scalable video coding are to achieve high compression efficiency at affordable cost and acceptable complexity. Scalable video coding is capable of coping with problems resulted from bandwidth variation [3].

Several scalable coding methods have been successfully proposed for video transmission through heterogonous networks. One of these techniques is the MPEG-4 Fine-Granular Scalability (FGS) scheme, that can adapt in real-time (i.e., at transmission time) to the bandwidth variation over heterogonous networks. The FGS encoder estimates the channel bandwidth before encoding and compresses the base layer into the rate lower than channel bandwidth. Therefore, this mechanism guarantees the quality of base layer bitstream. In the enhancement layer, the residual signal is compressed by bit-plane coding [4][5]. Some key advantages of the MPEG-4 FGS framework are its packet-loss resilience and flexibility in supporting streaming application. Recently, techniques that include a certain amount of enhancement layer information into the prediction loop have been proposed to improve the coding efficiency of FGS by introducing leaky prediction technique into prediction loop of a codec scheme [6]. However, the choice of leakage factor is varied according to characteristic of each video sequence and available channel bandwidth.

In this paper, we propose an adaptive leakage factor algorithm of a leaky prediction to improve the motion-compensation in FGS coding at bit rate lower than 256 kbps. The remainder of this paper is organized as follows. In Section 2, we describe the improved MPEG-4 FGS codec using leaky prediction, and our proposed adaptive leakage factor algorithm of leaky prediction in Section 3. Section 4 presents our simulation results under a comparison between original FGS and 1-loop MC-FGS coding scheme [7]. Finally, the conclusion is drawn in Section 5.

2. MPEG-4 FGS Codec Using Leaky Prediction

The basis idea of original FGS coding is also to code a video sequence into a base layer and an enhancement layer. The base layer uses nonscalable coding scheme to reach the lower bound of the bit-rate range. The difference between the original and the reconstructed picture is coded by using bit-plane coding of the DCT coefficients into the enhancement layer. The bitstream of the FGS enhancement layer may be truncated within any number of bits per picture after encoding completion. The decoder should be able to reconstruct an enhancement video from the base layer and the truncated enhancement layer bitstreams. At the enhancement layer, video quality is proportional to the number of bits decoded by the decoder for each picture.

Recently, techniques that include a certain amount of enhancement layer information into the prediction loop have been proposed to improve the coding efficiency of FGS [2][7]. However, to minimize the effect of drift introduced by prediction mismatch, the leaky prediction is used in the structure of FGS codec [6].

2.1 FGS Encoder Using Leaky Prediction

The FGS encoder using leaky prediction (FGS-LP) was proposed to improve motion-compensation in the original FGS coding scheme by introducing the notion of the extended base layer that includes an integrated base and enhancement layers. It is very clear that the reconstructed of reference frames have higher quality than that of the original FGS scheme. As a result, when the network bandwidth is available enough to transmit the enhancement layer bitstream to the decoder, this scheme can provide higher quality. But due to the limitation of network bandwidth, if the decoder can not receive the enhancement layer bitstream to reconstruct the reference frames, the difference between the reference frames used in the encoder and the decoder will cause drifting error and error propagation. Then, the leaky prediction technique that is introduced into the motion-compensation prediction loop to reduce such drift errors is presented in Fig. 1.

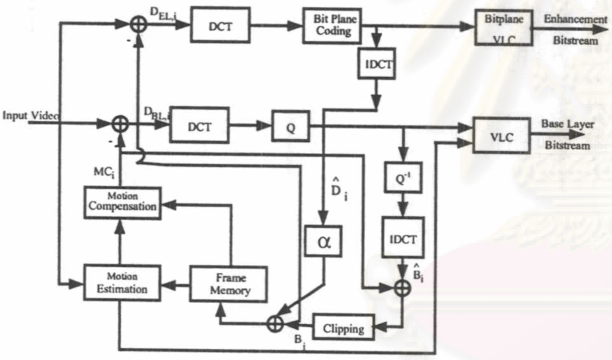


Figure 1. Structure of FGS encoder using the leaky prediction [6]

For the current frame, the original frame at time *i* is denoted F_i. At the base layer, the reconstructed frame of previous frame (at time *i-1*) is denoted B_{i-1} and the residue error of the base layer at time *i* is denoted $D_{BL,i}$. At the enhancement layer, the difference of the previous frame between original and reconstructed frames is encoded and denoted by $\hat{D}_{i-1}$. The leakage factor is denoted by α. Thus, the original frame at time *i* is shown in eq. (1).

$$F_i = (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + D_{BL,i} \qquad (1)$$

where $(B_{i-1} + \alpha \hat{D}_{i-1})_{mc}$ is the motion compensation version of $(B_{i-1} + \alpha \hat{D}_{i-1})$. The coded version of the residue error as the base layer is denoted $\hat{B}_i$ and the quantization error is denoted $D_{EL,i}$. The relationship between $D_{BL,i}$, $\hat{B}_i$ and $D_{EL,i}$ is shown in eq. (2).

$$D_{BL,i} = \hat{B}_i + D_{EL,i} \qquad (2)$$

By substituting eq. (2) into eq. (1), we obtain

$$\begin{aligned} F_i &= (B_{i-1} + \alpha \hat{D}_{i-1})_{mc} + \hat{B}_i + D_{EL,i} \\ &= (B_{i-1})_{mc} + \hat{B}_i + (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \\ &= B_i + \hat{D}_i \qquad (3) \end{aligned}$$

where

$$B_i = (B_{i-1})_{mc} + \hat{B}_i \qquad (4)$$

$$\hat{D}_i = (\alpha \hat{D}_{i-1})_{mc} + D_{EL,i} \qquad (5)$$

The signal B_i and $\hat{D}_i$ will be used to predict next frame.

By expanding the recursive formula of $\hat{D}_i$ in eq. (5), we can obtain eq. (6).

$$\begin{aligned} \hat{D}_i &= (\alpha((\alpha \hat{D}_{i-2})_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \\ &= (\alpha((\alpha((\alpha \hat{D}_{i-3})_{mc} + D_{EL,i-2}))_{mc} + D_{EL,i-1}))_{mc} + D_{EL,i} \qquad (6) \end{aligned}$$

As demonstrated in eq. (6), it is obvious that any drift errors in the final residual $\hat{D}_i$ will be attenuated in this proposed framework.

2.2 FGS Decoder Using Leaky Prediction

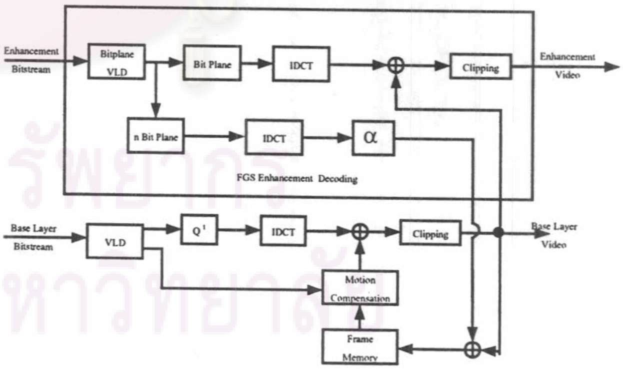


Figure 2. Structure of FGS decoder using the leaky prediction [6]

The FGS decoder using leaky prediction is shown in Fig. 2, the leaky prediction that was added at the enhancement layer with gain factor α (less than 1).

Assuming that there is an error at the enhancement layer for frame F_{i-2}, we denote the received enhancement layer bitstream, $\bar{D}_{i-2}$, and the transmission error, $\Delta\hat{D}_{i-2}$. Thus we have

$$D_{EL,i-2} = \bar{D}_{i-2} + \Delta\hat{D}_{i-2} \tag{7}$$

And the reconstructed version of $\hat{D}_{i-2}$ is denoted $\tilde{D}_{i-2}$. Thus,

$$\begin{aligned}\tilde{D}_{i-2} &= (\alpha\hat{D}_{i-3})_{mc} + \bar{D}_{i-2} \\ &= (\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2}.\end{aligned} \tag{8}$$

Comparing eq. (6) and (8), the difference between $\hat{D}_{i-2}$ and $\tilde{D}_{i-2}$ is $\Delta\hat{D}_{i-2}$.

Assuming that there is no error at the enhancement layer for frames F_{i-1} and F_i. Thus

$$\begin{aligned}\tilde{D}_{i-1} &= (\alpha\tilde{D}_{i-2})_{mc} + D_{EL,i-1} \\ &= (\alpha(\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2})_{mc} \\ &\quad + D_{EL,i-1}.\end{aligned} \tag{9}$$

The difference between $\hat{D}_{i-1}$ and $\tilde{D}_{i-1}$ is $\alpha(\Delta\hat{D}_{i-2})$.

Now we move to the frame F_i and obtain

$$\begin{aligned}\tilde{D}_{i} &= (\alpha\tilde{D}_{i-1})_{mc} + D_{EL,i} \\ &= (\alpha((\alpha(\alpha\hat{D}_{i-3})_{mc} + D_{EL,i-2} - \Delta\hat{D}_{i-2})_{mc} \\ &\quad + D_{EL,i-1}))_{mc} + D_{EL,i}.\end{aligned} \tag{10}$$

The difference between $\hat{D}_i$ and $\tilde{D}_i$ is $\alpha^2(\Delta\hat{D}_{i-2})$.

From the above derivations, it is obvious that the errors occurred in the decoded bitstream at the enhancement layer will be attenuated by a leakage factor of α for each iteration. After several iterations, the error will be attenuated to zero for α less than unity. Thus, the drift error is removed from the system.

3. Adaptive leakage factor algorithm of leaky prediction

Leaky prediction is a well-known technique to increase error robustness, which is a trade-off with coding efficiency [8]. In the proposed scheme, the first three bitplanes of the enhancement layer data is transformed back to the spatial domain using IDCT and is attenuated by a leakage factor α before adding into the frame memory. Therefore, the drift or the difference between the encoder and decoder will be attenuated. If the leakage factor is set to zero, the drift will be removed completely.

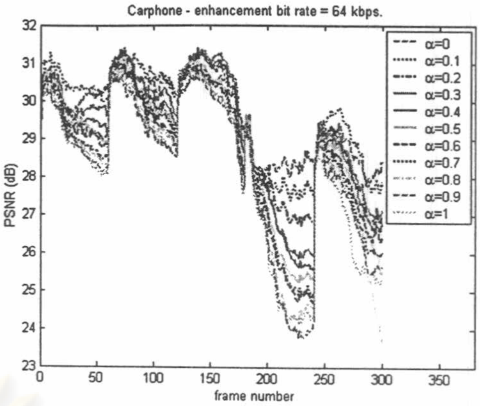


Figure 3. PSNR comparison of the video quality for the Carphone sequence with variable setting of the leakage factor. The bit rate of base layer is 32kbps. And the total bandwidth is 96 kbps.

However, the choice of leakage factor is important and it is shown to be varied among difference video frames and channel bandwidth. When we vary the leakage factor from zero to one in the Carphone sequence, as shown in Fig.3, we observe that the setting of leakage factor of each frame should be adapted according to spatial activity of that frame. However, the scheme proposed by [6] is still not dynamic and flexible to support characteristic variation from frame to frame and or from video sequence to another. Because its leakage factor is selected manually and fixed though each video sequence. Moreover, by varying the leakage factor from zero to one for several video sequences at different bit rates, the relationship between enhancement bandwidth and the leakage factor can be shown in eq. (11).

$$\alpha_{initial} = 0.1 \times \log_2\left(\frac{BW_{enh}}{64}\right) : 0 < BW_{enh} \le 256 \tag{11}$$

So, an adaptive leakage factor algorithm is proposed to identify a suitable value of leakage factor for each video sequence and channel bandwidth. The proposed algorithm is executed at the frame level. The processing steps are shown as follows.

Step 1. Determine average activity measure (avg_act) parameters [9]. It is the average value of 'Spatial Activity Measure (act)' in each frame. Compute a spatial activity measure for the macroblock j from the four luminance frame-organized sub-blocks (n=1,..,4) and the four luminance field-organized sub-blocks (n=5,..,8) using the original pixel values:

$$act_j = 1 + \min(vblk_1, vblk_2, \ldots, vblk_8) \tag{12}$$

where the variance of 8*8 block is defined as:

$$vblk_n = \frac{1}{64} \times \sum_{k=1}^{64} \left(P_k^n - P\_mean_n\right)^2 \qquad (13)$$

and

$$P\_mean_n = \frac{1}{64} \times \sum_{k=1}^{64} P_k^n \qquad (14)$$

and P_k^n is the sample value in the n-th original 8*8 block.

Step 2. Find Tot_avg_act parameter, which is the average value of avg_act from the first frame to the current frame encoded.

Step 3. Identify a suitable value of leakage factor. The decision can be defined as follows:

```
if((avg_act - Tot_avg_act) > Th)
    $\alpha = \alpha_{initial} - 0.1$;
else if((avg_act - Tot_avg_act) < -Th)
    $\alpha = \alpha_{initial} + 0.1$;
else
    $\alpha = \alpha_{initial}$ ;
```

4. Simulation Results

The experimental results have been implemented to show the performance of the proposed scheme. The MPEG-4 VM. 18.0 codec is used for the base layer [10]. From Figs. 4-6, the coding efficiency of the proposed scheme is compared with the original FGS coding. The Carphone and Foreman sequences in QCIF format (176x144 pixels/frame) are used for testing and every GOP has 60 frames which consist of 1 I-frame, and 19 P-frames. The bit rate of the base layer based on the TM5 rate control is 32 kbps, and the frame rate is set at 30 frames/s. The leakage factor α is adapted in the range of 0 and 1.

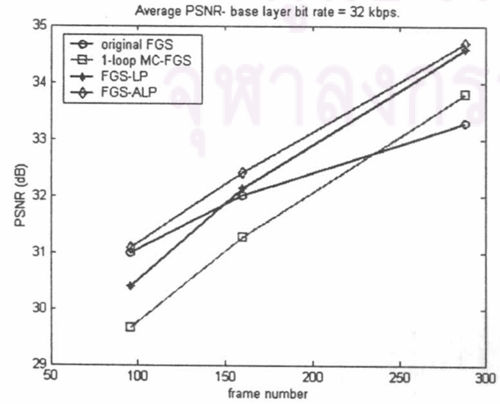


Figure 4. Average PSNR versus bit rate comparison between proposed algorithm and original FGS for several video sequences. The bit rate of base layer is 32kbps.

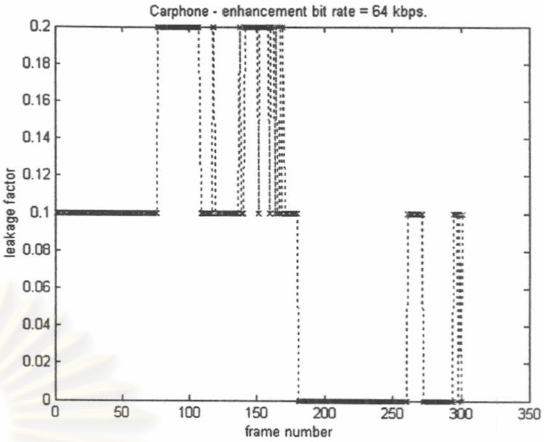


Figure 5. Variation of leakage factor (α) in Carphone sequence

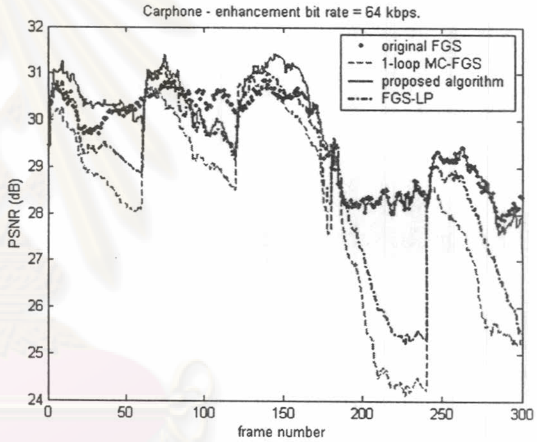


Figure 6. PSNR comparison of the video quality for the Carphone sequence between original FGS and proposed algorithm. The bit rate of base layer is 32kbps. And the total bandwidth is 96 kbps.

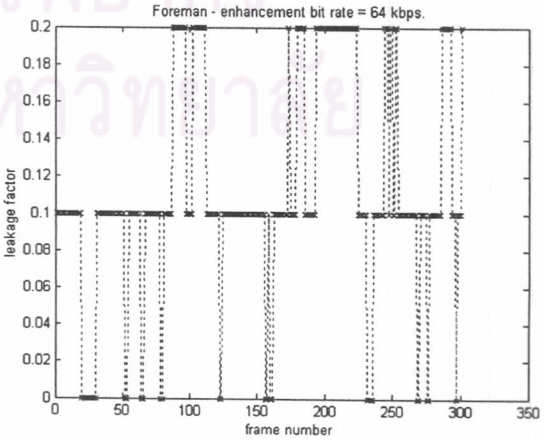


Figure 7. Variation of leakage factor (α) in Foreman sequence

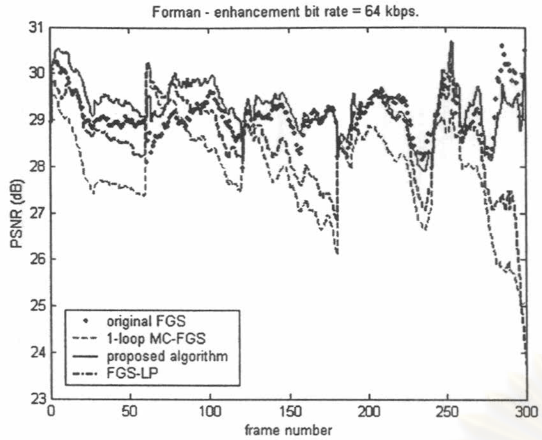


Figure 8. PSNR comparison of the video quality for the Foreman sequence between original FGS and proposed algorithm. The bit rate of base layer is 32kbps. And the total bandwidth is 96 kbps.

From Fig. 4, as comparing with the original FGS, the 1-loop MC-FGS and FGS using the leaky prediction (FGS-LP), our results show that this proposed algorithm has improved by about 1.3 dB, 1.4 dB, and 0.7 dB in PSNR, respectively. The leakage factor of each frame, which is identified by the proposed algorithm, is shown in Fig. 5 for Carphone sequence and Fig. 7 for Foreman sequence. Carphone and Foreman sequences are the standard sequences that are used for testing. For Fig. 6 and Fig. 8, our results show that this proposed algorithm can identify a suitable value of the leakage factor of each frame (or each video sequence).

5. Conclusions

In this paper, we propose an adaptive leakage factor algorithm of a leaky prediction in the MPEG-4 FGS codec to identify a suitable value of leakage factor to suit characteristic of each video sequence and channel bandwidth. As comparing with the original FGS, our simulation results show that our proposed algorithm has improved by about 1.3 dB in PSNR and up to 1.4 dB over the 1-loop MC-FGS. Moreover, average PSNR is up to 0.7 dB in the proposed algorithm compared with FGS using the fix leakage factor leaky prediction.

6. Acknowledgement

The authors would like to express the grateful thanks to grant from government research and development in cooperative project between Department of Electrical Engineering and Private Sector for Research for supporting this work and development under Chulalongkorn University.

7. References

[1] Mihaela van der Schaar, "A hybrid temporal-SNR fine granularity scalability for Internet video," IEEE Trans. Circuits. Syst. Video Technol., vol. 11, pp. 318-331, March 2001.

[2] F. Wu, S. Li, and Y. Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," IEEE Trans. Circuits. Syst. Video Technol., vol. 11, pp. 332-344, March 2001.

[3] Marta Mrak, Nikola Sprljan, Sonja Grgic, "Video Coding Schemes for Transporting Video Over the Internet," Faculty of Electrical Engineering and Computing, University of Zagreb, Technical Report (in Croatian), Zagreb, 2001.

[4] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," IEEE Trans. Circuits Syst. Video Technol., vol. 11, pp. 301-317, March 2001.

[5] Information Technology-Coding of Audio-Visual Objects Part 2: Visual ISO/IED 14496-2:1999/ FDAM 4, ISO/IEC JTC1/SC29/WG11 N3904, January 2001.

[6] S. Artnarongkorn, D. Tancharoen, S. Aramvith, and S. Jitapunkul, "Robust Fine Granularity Scalability using Leaky Prediction for Low Bit-Rate Video Coding," Proceeding of ICT 2003, Assumption University, pp. 132-136, April 2003.

[7] M. van der Schaar and H. Radha, "Motion compensation fine-granularity-scalability for wireless multimedia," in Proc. IEEE Workshop Multimedia Signal Processing, vol.10, pp. 453-458, Oct. 2001.

[8] Hsiang-Chun Huang, Chung-Neng Wang, and Tihao Chiang, "A robust fine granularity scalability using trellis-based predictive leak," IEEE Trans. Circuits Syst. Video Technol., vol. 12, pp. 372-385, June 2002.

[9] MSSG home page: http://www.mpeg.org/MPEG/MSSG/tm5/Ch10/Ch10.html

[10] MPEG-4 Video Verification Model version 18.0, ISO/IEC JTC1/SC29/WG11 N3908, Pisa, January 2001.

ประวัติผู้เขียนวิทยานิพนธ์

นางสาวสุจรรยา อาจณรงค์กร เกิดวันที่ 8 พฤศจิกายน พ.ศ. 2522 ที่จังหวัดอุตรดิตถ์ เข้าศึกษาในหลักสูตรวิศวกรรมศาสตรบัณฑิต คณะวิศวกกรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ในปีการศึกษา 2540 และเข้าศึกษาต่อในหลักสูตรวิศวกรรมศาสตรมหาบัณฑิต ที่ห้องปฏิบัติการวิจัยวิธีสัญญาณดิจิทัล ภาควิชาวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย ในปีการศึกษา 2544