International Journal of Principles and Applications of Information Science and Technology
December 2007, Vol.1, No.1

Performance Assessment of Fine Granularity Scalability in MPEG-4 Standard[1]

Namyun Kim, Kitae Hwang

Hansung University
Samseon-dong, Seongbuk-gu, Seoul, Korea
{nykim, calafk}@hansung.ac.kr

Abstract. This paper discusses a performance assessment of fine granularity scalability (FGS) in MPEG-4 video streaming. The performance issues used in this paper are encoding time, decoding time, data bit-rate, and PSNR. The measurements were carried out through real experiments by varying quantization parameter and the number of bit planes. The experiment results show that a streaming system can adapt video streams without sacrificing visual quality and timeliness by determining the appropriate quantization parameter and the number of bit planes.

Key words : MPEG-4 FGS, Video Streaming, Heterogeneous Environment, Performance Assessment.

1 Introduction

There have been increasing demands for video streaming applications over the Internet [1][2]. Since the current Internet provides best-effort services, video applications suffer an end-to-end delay and frame loss due to bandwidth constraints of communication channels. In addition, it is not easy for video applications to support several kinds of video devices with diverse display capabilities. Therefore, techniques for video scalability are needed to adapt the streamed-video quality according to changes of the network condition and the video device capability.

In MPEG-2 and MPEG-4, several scalability techniques such as SNR scalability, temporal scalability, and spatial scalability have been introduced [2]. In such a scalable coding technique, a video sequence is coded into a base layer (BL) and an enhancement layer (EL). The base layer stream can be independently decoded and provide coarse visual quality whereas the enhancement layer stream can be decoded together with the base stream and can provide better visual quality. Thus, the encoder can compress the video signal into a bit

1 This research was financially supported by Hansung University in the year of 2007.

rate that is less than the current channel condition (BL only or BL+EL).

As a new scalable coding scheme to extend the existing scalability, the MPEG-4 fine granularity scalability (FGS) [3] has been proposed as a part of MPEG-4 standards. In contrast to the conventional scalable video coding which requires the reception of full EL, the EL consists of a few bit planes. In FGS coding, the EL bit stream can be truncated into any number of bits within each frame to provide partial enhancement. Therefore, the FGS video encoder is capable of achieving continuous rate control through a way of dropping some bit planes to meet seamless video display. On receiver side, the received EL (full or some part of EL) can be successfully decoded and the better visual quality can be achieved.

To fully take advantage of FGS video coding technique for streaming video applications, many performance issues have to be dealt with, which are addressed in this paper. For example, how the encoding time and decoding time of the MPEG-4 FGS video stream vary with the number of bit planes. However, they have not been addressed in the past researches [4][5]. The encoding and decoding time are deeply related with quantization process and creation of bit planes of EL, and have great impacts on accessibility of client devices to seamless video streaming.

In this paper, we focused on live video streaming using MPEG4-FGS and identified performance issues to be considered; encoding time, decoding time, data bit-rate, and PSNR. The measurements were carried out through real experiments by varying quantization parameter and the number of bit planes. The objective of these experiments is to get the impact of FGS bit plane and quantization parameter on four performance issues. This result can be used in dynamic adaptation streaming system.

The remainder of the paper is organized as follows. Section 2 discusses MEPG-4 FGS coding method. Section 3 describes the performance issues to enhance adaptation to network condition and device capability. Section 4 discusses the assessment of performance issues for different quantization parameters and the number of bit planes. Finally, we conclude in section 5.

2 MPEG-4 FGS Coding Method

As shown in Fig.1, the basic MPEG-4 FGS structure consists of two separate streams of BL and EL. In MPEG-4, the VOP (Video Object Plane) is the instance of a video object at a given time and corresponds to a frame of video sequences in MPEG-1 or MPEG-2 [3].

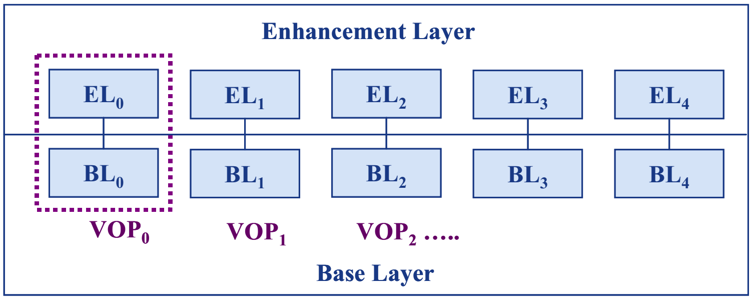

Fig. 1. A Basic MPEG-4 FGS Structure

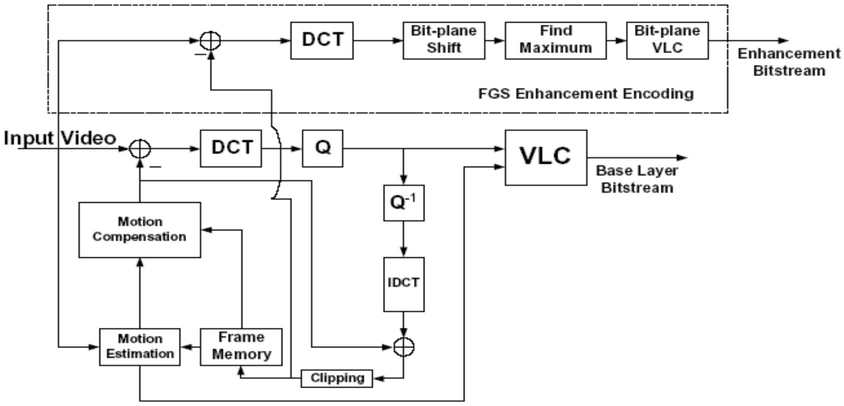

Fig. 2. MPEG-4 FGS Encoder Structure

Fig. 2 shows a typical structure of the MPEG-4 FGS encoder [6]. The BL is encoded by the MPEG-4 encoder for a low bit rate. In the encoding process of the BL, all values in the DCT matrix are divided by a quantization parameter (QP) during the quantization (Q) step, where the stream size of the BL is determined. That is, as QP value gets larger, the stream size gets less and video quality gets worse. However, it consumes less network bandwidth as a trade off.

The difference between the original picture and the reconstructed picture is coded as an EL stream using bit plane (BP) coding of DCT coefficients. The BP coding method considers each DCT coefficient as a binary number of several bits instead of a decimal integer. For each 8*8 DCT block, 64 DCT coefficients are zigzag ordered into an array. A bit plane of the block is defined as an array of 64 bits, taken one bit from each absolute value of the DCT coefficients at the same significant position. For each bit plane, (RUN, End-of-Plane) symbols are formed and coded to produce the output bit stream. As a result, in the BP coding method, an EL is coded into multiple layers (EL1, EL2, EL3 …), where EL1 is a layer composed of the most significant bit (MSB) of the entire EL. The layer EL1 is known to make the most important impact on the video quality, so it is usually sent to the receiver with the highest priority. And then the EL2, EL3, and so on are sent in sequence, if needed. Note that EL2 includes MSB and MSB-1 bit stream. Thus, the sender can adjust the number of the bit planes depending on network bandwidth or client device capability, and the receiver can decide how many EL bit planes it will decode to play the video seamlessly according to its capability.

3 Performance Issues for Adaptation

The existing researches have proposed the scalable video coding using MPEG-4[2]. In live video steaming, traditional video encoders typically rely on altering the QP value of the encoder to achieve rate adaptation. So the sender can dynamically adjust the target bit rate by changing the QP value. If the available network bandwidth is detected to be low, the QP value can be changed higher and thus the bit rate of the stream is smaller. Thus, the receiver can display the video stream continuously under high network traffic.

In MPEG-4 FGS, the number of bit planes can be controlled by the sender in order to adapt varying network bandwidth [4][5]. That is, the sender can send base layer with EL1 or other enhancement layer according to network condition. However, it was not well-known how data rate, video quality and encoding/decoding time in MPEG-4 FGS are affected by the number of bit planes. For the adaptive algorithm to be practical in real world, interdependency between the above issues and the number of bit planes should be analyzed.

In this paper, we identified four performance issues and conducted experiments on the impact of the number of bit planes: *encoding time, decoding time, bit rate, and PSNR.*

4 Measurement of Performance Issues

4.1 Test System

For the measurement of MPEG-4 FGS performance issues, a test system illustrated in Fig. 3 has been built. For setting up the test system, we utilized the MPEG4-FGS codec and JRTP library as shown in Table 1. We have used Akiyo QCIF (176x144) frames as the video sequence. We used a YUV file recorded previously as a video source. The FGS server creates separately two video streams such as base bit stream and enhancement bit stream by varying QP values and the number of bit planes. The server sends two bit streams to the FGS client through RTP (Real-time Transfer Protocol) [7]. The FGS client receives both bit streams, decodes them by each VOP unit into YUV video frames, and displays them.

Table 1. Experiment environments

Parameter	Contents
MPEG-4 FGS Codec	Microsoft MPEG-4 FGS V2.4[8]
RTP Library	JRPT Open Source[9]
Video Sequence	Akiyo QCIF(176x144) 100 frame

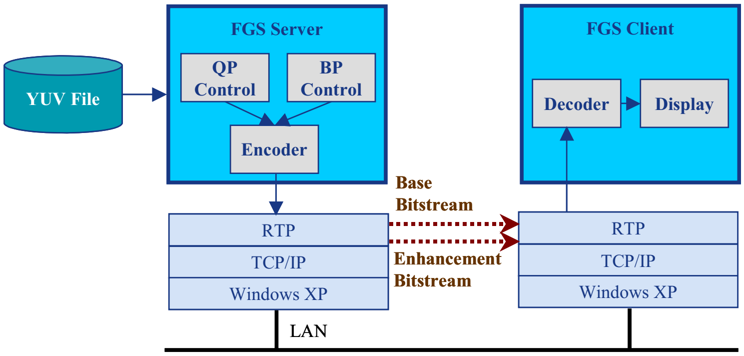

Fig. 3. MPEG-4 FGS Test System

4.2 Encoding Time

Fig. 4 shows the average encoding time for 100 frames according to different QP values and the number of BPs. In the legends of all following graphs from Fig. 4 to Fig. 7, BL means the base layer only and BL+ELN means both the base layer and N (number) of bit planes in the enhancement layer, respectively. In general, for smaller QP values, MPEG-4 FGS codec generates less bit planes. As shown in Fig. 4, the encoding time of all video streams is less than 33 msec.

The purpose of this measurement is to examine the change of encoding time by varying QP values and the number of BPs. As you can imagine from the encoding process shown in Fig. 2, the QP value makes an impact on the VLC(Variable Length Coding) process. Huffman coding included in the VLC process depends on the data size generated by the quantization process. In general, the big QP removes more values in high frequency area and produces more zero values, which reduces encoding time by Huffman coding. However, this trend gets weaker after a certain QP value which might be 6 in this experiment. This is why high frequency area is slowly exhausted as the QP becomes greater over a certain point. On the other hand, encoding an enhancement stream causes an additional delay by about 35% compared to a base layer stream.

4.3 Decoding Time

Fig. 5 shows the average decoding time of a frame (VOP) according to different QP values and the number of bit planes. We measured the decoding time as the elapsed time from the instant of a frame reception to playout time. The decoding time is related with device capability.

As you can see, decoding time is much larger than encoding time. This is because decoding phase of Huffman coding takes a long time. The decoding time of BL and more than the second bit plane takes longer than 33 msec.

If video servers would like to support video streaming for devices with low CPU capability, they have to consider the number of bit planes.

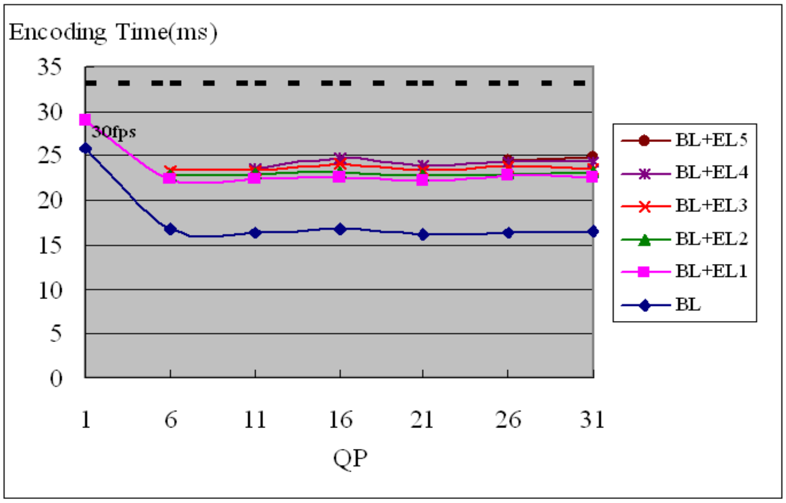

Fig. 4. Average Encoding Time for QP Value

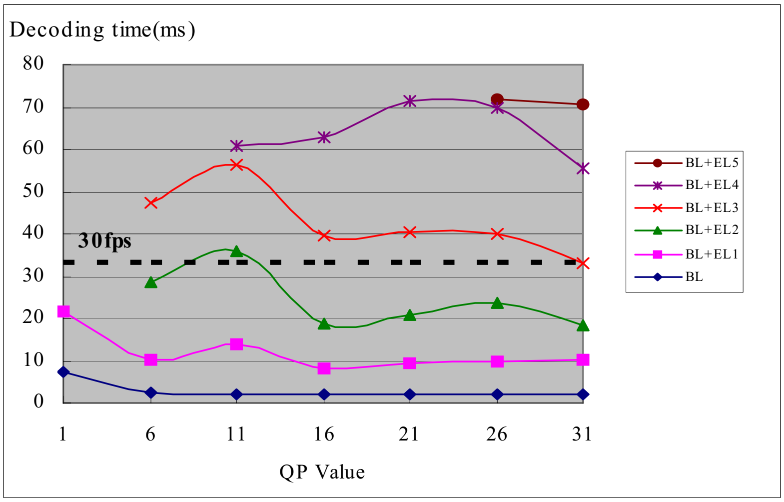

Fig. 5. Average Decoding Time for QP Value

4.4 Bit Rate

Bit rate of the video frame is one of the most important issues in adaptation. Fig. 6 shows the average bit rate. For the fixed QP value, the bit rate increases rapidly as the number of BP increases. In both cases of the BL and the BL+EL1, bit rates have a trivial difference. This is because EL1 bit stream is much smaller than other bit planes. The server should determine the number of BPs so that target bit rate is less than the available network bandwidth.

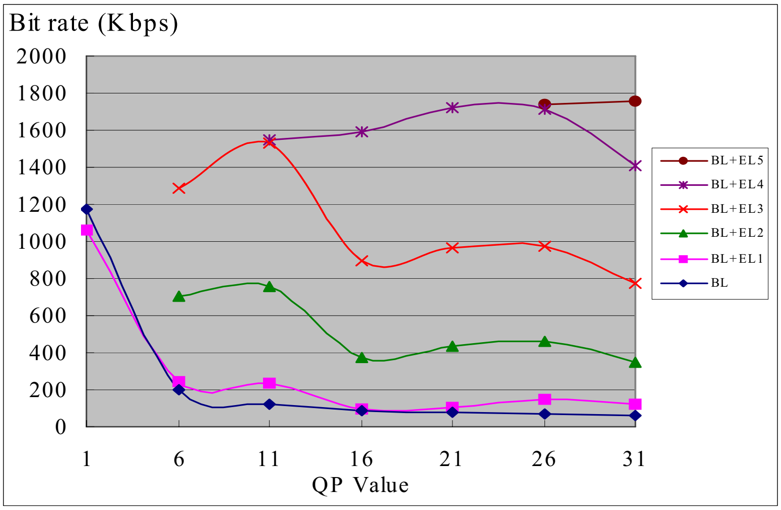

Fig. 6. Average Bit Rates for QP Value

4.5 PSNR (Peak Signal to Noise Rate)

We measured the average of PSNR values to examine the relationship between video quality and parameters such as QP and the number of bit planes. The PSNR is often used as a measure of the difference between the original video sequence and the received video sequence, i.e., video quality in the receiver side.

The experiment result in Fig. 7 shows that the quality is better as the number of bit planes increases for fixed QP value and the quality is worse as the QP value increases for a specific bit plane. Fig. 8 shows pictures of PSNR 30dB and 46dB when QP value is 31.

PSNR can be converted to MOS (Mean Opinion Score) [10], which is an evaluation value for the human quality impression and is given on a scale from 5(best) to 1(worst). More than 37db in PSNR value converts to 5 in MOS and the quality of video image is evaluated to be excellent.

As shown in Fig. 6, when more than the second bit plane of EL is transferred, the bit rate increases rapidly and consumes larger network bandwidth. If bigger values than 37dB PSNR assume to be the same quality (MOS=5), the server doesn't have to send the third bit plane of EL, which results in less consumption of the network bandwidth and therefore provides seamless video streaming under network congestion.

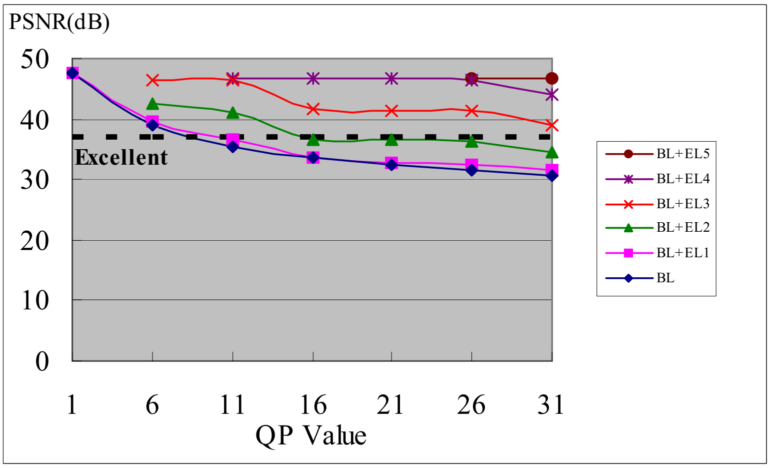

Fig. 7. Average PSNR for QP Value

Fig. 8. Video quality for different number of BP

5 Conclusion and future work

In this paper, we built a test model using MPEG-4 FGS and analyzed the performance issues for video streaming through real measurement. The objective of these experiments is to get the impact of FGS bit plane on four performance issues such as encoding times, decoding times, bit rate and video quality. This result can be used in dynamical adaptation streaming algorithm.

Experiment results show that selecting the appropriate number of bit plane can reduce bit rate of a video stream without sacrificing video quality. This information can be used for rate control under network congestion. In addition, we found that decoding time should be considered for client devices with low capabilities.

In the future, we need to experiment with diverse video samples and analyze in more details.

References

1. D.Wu: Streaming Video over the Internet, Approaches and Directions. IEEE Transactions on Circuits and Systems for Video Technology (2001).
2. D. Wu, et al: On End-to-End Architecture Transporting for MPEG-4 Video Streaming over the Internet. IEEE Transaction on Circuits and Systems for Video Technology (2000).
3. ISO/IEC JTC1/SC29/WG11: Coding of audio-visual objects, Visual ISO/IEC 14496-2 / Amd 2, Streaming video objects (2002).
4. Q. Liu and J. N. Hwang : Scalable Video Transmission System using Bandwidth Inference in Congestion Control. IEEE Int. Conf. on Circuits and Systems (2004).
5. C. Huang and L. Guan: A New Rate Adaptation Framework for MPEG-4 FGS Video over IP. IEEE Canadian Conf. on Electrical and Computer Engineering (2004).
6. Weiping Li: Overview of Fine Granularity Scalability in MPEG-4 Video Standard. IEEE Trans. on Circuits and Systems for Video Technology (2001).
7. Schulzrinne, Casner, Frederick, and Jacobson: Real-Time Transport Protocol. RFC 1889 (1996).
8. MPEG-4 FGS: ISO/IEC part 5, Reference Software, microsoft-v2.5-040207-NTU.
9. JRTPLIB, http://research.edm.uhasselt.be/jori/page/index.php.
10. J. Klaue, et al: Eval - A Framework for Video Transmission and Quality Evaluation. Proc. of the Int. Conf. on Modeling Techniques and Tools for Computer Performance Evaluation, (2003).