2—2

Adaptive Temporal Rate Control of Video Objects for Scalable Transmission

Hee Dong CHANG, Young Kwon LIM, Myoung Ho LEE, and Chieteuk AHN
Electronics and Telecommunications Research Institute,
{hdchang,yklim,mhlee,ahnc}@video.etri.re.kr

Abstract :The video transmission for real-time viewing over the Internet is a core operation for the multimedia services. However, its realization is very difficult because the Internet has two major problems, namely, very narrow endpoint-bandwidth and the network jitter. We already proposed a scalable video transmission method in [8] which used MPEG-4 video VM(Verification Model) 2.0[3] for very low bit rate coding and an adaptive temporal rate control of video objects to overcome the network jitter problem. In this paper, we present the improved adaptive temporal rate control scheme for the scalable transmission. Experimental results for three test video sequences show that the adaptive temporal rate control can transfer the video bitstream at source frame rate under variable network condition.

1. Introduction

The demand of video services such as VoD(Video on Demand) in the Internet is rapidly increasing in recent years despite no guarantee of QoS(Quality of Services) because the number of users is about 20-40 million and it will be doubled every year[1].

However, there are two major problems in the Internet for the video services. First, the network bandwidth is not sufficient to transfer video data in real-time. Second, the network jitters are unexpectedly variable according to network load so that fluctuating frame rate occurs at a client.

To cope with the low bandwidth, there are very low bit-rate coding methods such as DCT-based H.263[2], content-based MPEG-4[3], and region-based coding[4]. As to the network jitters, dynamic bit-rate control methods[5,6,7] have been proposed. These bit-rate control methods result in variable quality of image at a client since it uses different bit-rate according to the network condition.

The basic schemes for a scalable video transmission were presented in [8]. The schemes used MPEG-4 video VM(Verification Model) 2.0[3] for very low bit rate coding and an adaptive temporal rate control of video objects to overcome the network jitter problem.

The scalable video transmission keeps higher frame transfer rate (frames/second) than source frame rate under the variable network condition and keeps a constant quality of image.

In this paper, we propose the improved adaptive temporal rate control scheme for the scalable transmission.

The principle of the adaptive temporal rate control for the scalable transmission is described in Figure 1. In general video objects on frames have different degree of motion, for examples, the backboard screen

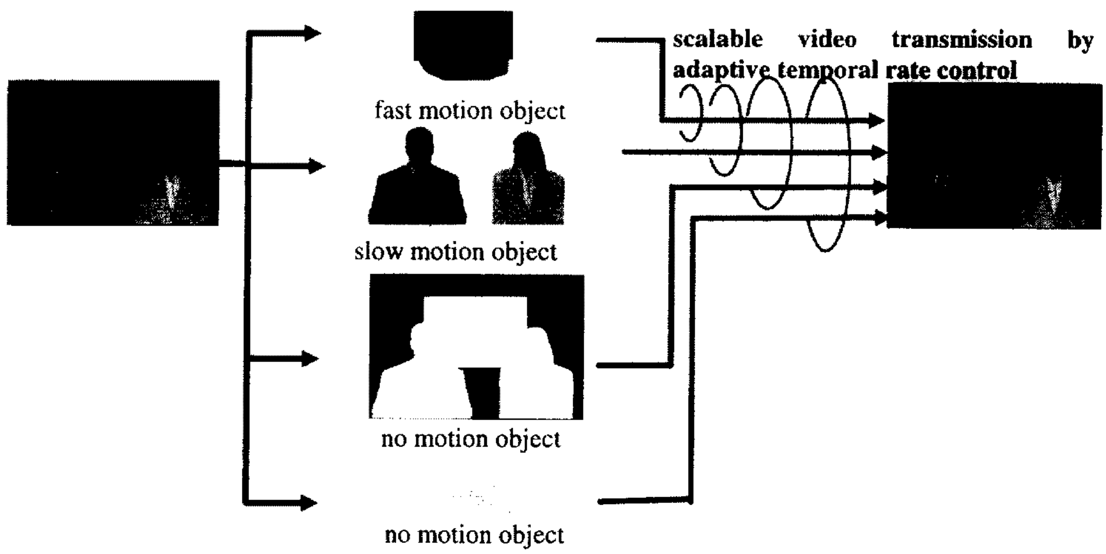


Figure 1: The principle of the adaptive temporal rate control for a scalable video transmission

has fast motion and the text 'MPEG-4 WORLD' and background have no motion in Figure1. The principle is to reassign temporal rates to video objects based on their motion degree so that the data size of video frames may be scaled down or reduced according to the network condition. Thus the objects with fast motion are transferred at source frame rate but the objects with no motion only once.

The overall system architecture for the scalable transmission is presented in Section 2, and the adaptive temporal rate control scheme in Section 3. The experimental results and concluding remarks are given in Section 3 and 4, respectively.

2. The Overall System Architecture

The overall system architecture for the scalable transmission method is shown in Figure 2. The VOP(Video Object Plane) Definition, Encoding, Decoding, Multiplex, Demultiplex, and Scene Composition form the modules of the MPEG-4 Video VM 2.0[3]; whereas the rest are the dedicated modules (rectangles shown in thick lines) for the temporal rate control.

The detailed explanation for the architecture is given in[8].

3. The Adaptive Temporal Rate Control Scheme

We use the following terminology to describe the adaptive temporal rate control scheme.

- Intra frame is that all its VOPs are encoded

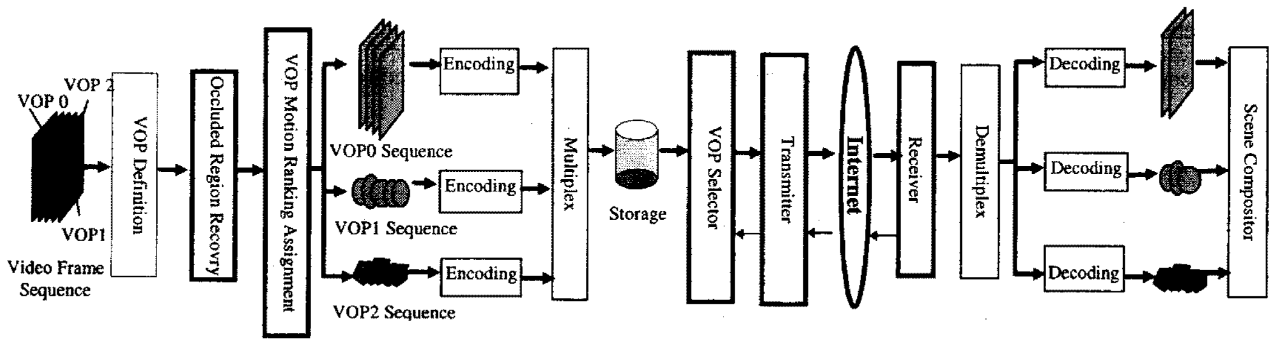


Figure 2: The overall system architecture

in intra mode.

- Inter frame has a VOP encoded in inter mode.
- GOV(Group of VOPs) is a group of all VOPs on video frames between an intra frame and the previous inter frame to the next intra frame.

A GOV is depicted in Figure 4-(a).

In fact, MPEG-4 video bitstream syntax provides the VOP-based random accessibility and editing. Thus we can access the starting points of a VOP, frame, and GOV in the bitstream; while the GOV is used as a processing unit for the temporal rate control.

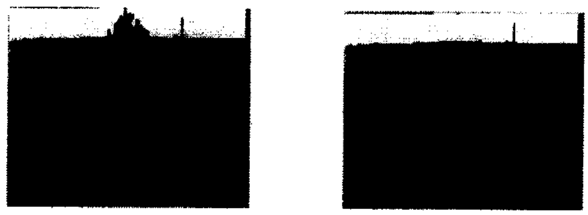
In this paper, we assume that the coding structure for each VOP id(identification) is I B B P B B P··· P B B P··· until the next intra frame; and the period of intra frame is within one second.

In the following, we describe the occluded region recovery and the adaptive temporal rate control of VOPs for the improved adaptive control scheme.

3.1 The Occluded Region Recovery

The recovery of the occluded region of VOP is needed for preventing cracks on the frame that is composed by VOPs on the current frame and VOPs on a past frame in the scene composition. Hence it is sufficient to recover the minimum region, able to prevent the cracks.

The recovery is performed automatically for stationary or translated VOPs, and performed manually for the others. The recovery for stationary or translated VOPs is made by union of the VOPs with the same id in the sense that each VOP is considered as a set of pixels. An example of the automatic region recovery is given in Figure 3. The recovery for the other VOPs is accomplished by editing of the VOP image using consecutive or entire VOPs.



(a) (b)

Figure 3: An example of the occluded region recovery for the land and sky: (a) before the recovery (b) after the recovery

3.2 The Adaptive Temporal Rate Control of VOPs

The adaptive temporal rate control is executed by the VOP selector in each GOV. It controls the temporal rate of each VOP adaptively based on the VOP motion characteristics, according to the network condition. An example of the temporal rate control is given in Figure 4. The control can be described in details as follows:

1) Select all VOPs on the intra frame except transferred stationary VOPs and those with the motion ranking 0(the fastest motion)[8] on the inter frames.
2) Transfer the selected VOPs using the Transmitter.
3) Get the current data transfer rate after the transferring.
4) Calculate the target data size of the GOV which is able to transfer the reduced bitstream at the frame rate of video source under the current data transfer rate:

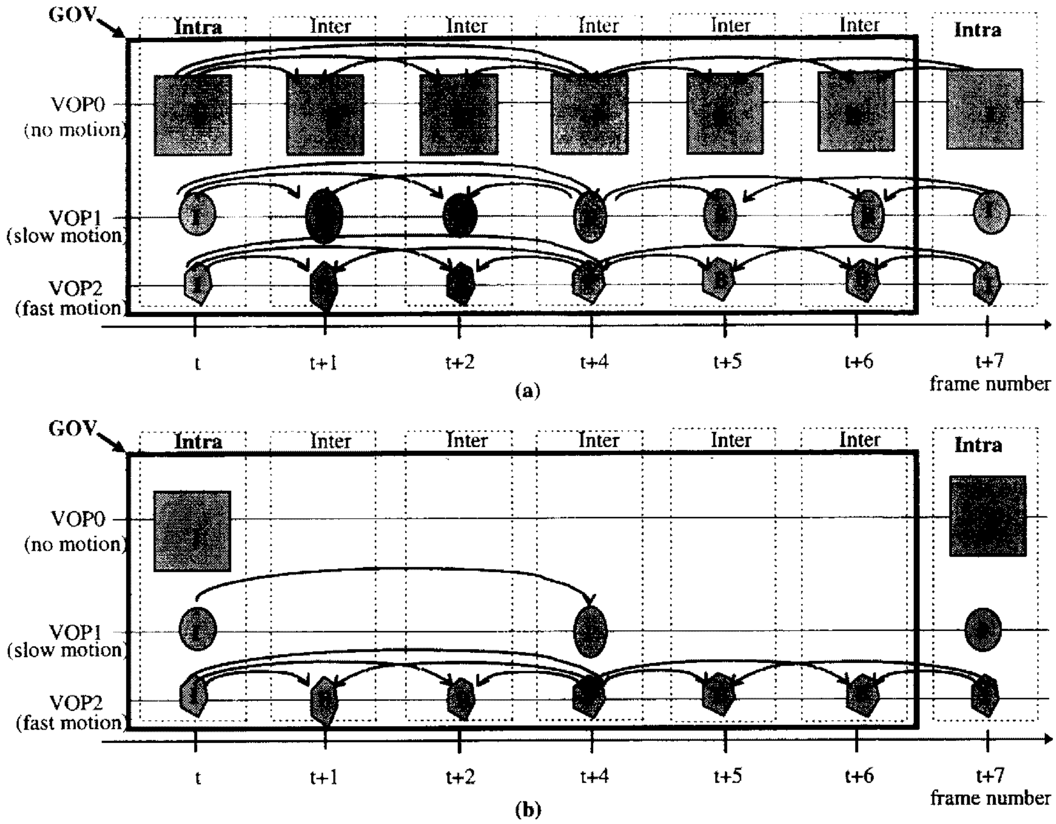


Figure 4: An example of temporal control of VOPs: (a) before the control (b) after the control

$$\textit{the target data size of the GOP} = \textit{frame number of the GOP} \frac{\textit{the current data transfer rate}}{\textit{the frame rate of video source}}$$

5) If the target data size of the GOP is greater than the total data size of the VOPs selected in 1), then VOPs with the nonzero motion ranking on the inter frames are selected until the difference of the target data size with the total data size of VOPs selected in 1) is filled according to the following order:

① in the motion ranking order

② in the encoding type order: P-VOP, B-VOP

③ in the frame number order

4 Experimental Results

We tested the improved adaptive temporal control scheme on two personal computers based on Microsoft Windows95 connected to 10 Mbps Ethernet LAN.

We assumed that the transmission channel bandwidth was the compression bit rate for each video bitstream and the network delay time was variable. To simulate the variable network delay on the assumed channel, we used the following delay time model:

Table 1. QCIF video sequences

class	sequence	# of VOPs	FPS	kbit/s
A	Akiyo	2	10	111
	Container	6	10	200
B	News	4	10	214

the delay time on the assumed network
= the delay time on the assumed channel
+ the additional delay time for variable network condition
= the delay time on the 10Mbps LAN + α
+ the additional delay time for variable network condition
= (the delay time on the 10Mbps LAN + α*)*

$$\left(1 + 2\frac{\text{random number}}{\text{maximum random number}}\right)$$

where α *= the delay time on the assumed channel*
- the delay time on the 10Mbps LAN

$$= \frac{\text{the transfered data size}}{\text{the assumed channel bandwidth}}$$

- the delay time on the 10Mbps LAN

Hence the delay time on the assumed network could be obtained using the delay time on the 10Mbps LAN, the transferred data size and the assumed channel bandwidth.

The QCIF video sequences for this test were listed in Table 1.

In encoding the test sequences, we chose 0.9 sec as the period of Intra frame and I B B P B B P··· as the encoding structure for each VOP in a GOP.

The experimental results included the comparison of the frame transfer rate (frames/sec) of the adaptive control scheme with that of the non-control scheme for each sequence, were shown in Figure 5~7.

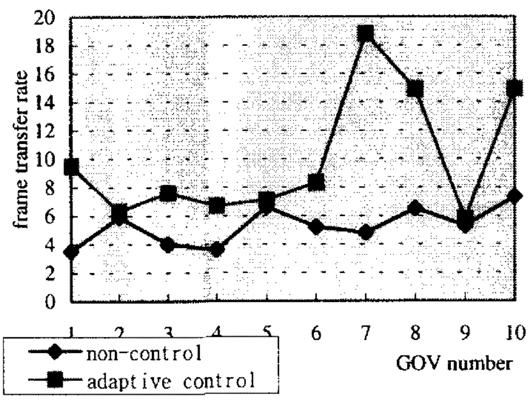


Figure 5: The comparison of the frame transfer rate of the adaptive control scheme with that of the non-control scheme for Akivo seauence

From Figure 5~7, the adaptive temporal control scheme maintains a higher frame transfer rate than that of the non-control scheme and able to transfer the reduced GOV at the frame rate of video source under variable network condition.

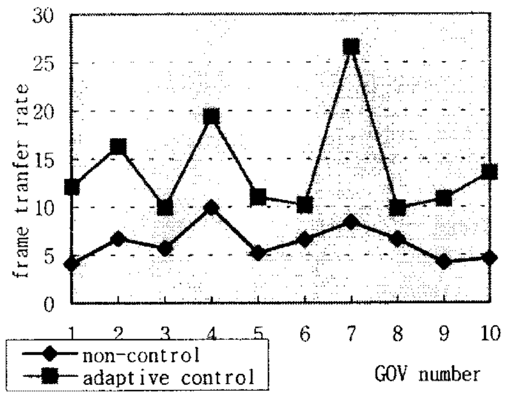


Figure 6: The comparison of the frame transfer rate of the the adaptive control scheme with that of non-control scheme for Container sequence

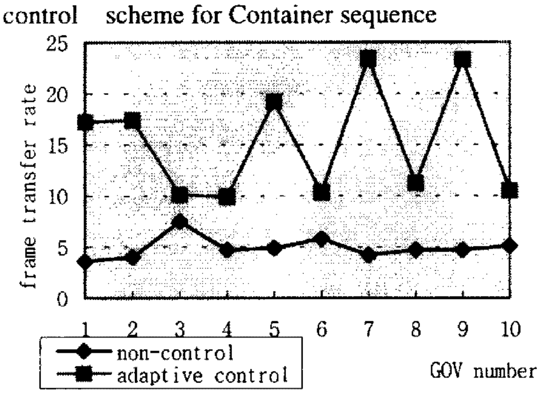


Figure 7: The comparison of the frame transfer rate of the the adaptive control scheme with that of non-control scheme for News sequence

4 Concluding Remarks

We propose the improved adaptive temporal rate control scheme of video objects for the scalable video transmission.

In order to transfer the video bitstream at source frame rate under the variable network condition, the adaptive control scheme scales down the data size of the encoded video frames by controlling the temporal rate of the content objects on the video frames according to the their motion characteristics.

By experiment results, the control method maintains higher frame transfer rate than source frame rate for more than two VOPs and constant quality of the image.

Further work is to develop buffer control and loss packet recovery methods.

References

[1] G. Bell and J. Gemmell, "On-Ramp Prospects for the Information Superhighway Dream," *Communications of the ACM*, vol.39, no.7, pp. 55-61, July 1996.

[2] ITU-T Rec. H.263: "Video Coding for Narrow Telecommunication Channels at < 64 kbit/s," March 1996.

[3] ISO/IEC JTC1/SC29/WG11 N1260: "Video Verification Model Version 2.0," March 1996.

[4] Y. Yokoyama, Y. Miyamoto, and M. Ohta, "Very Low Bit Rate Video Coding Using Arbitrarly Shaped Region-Based Motion Compensation," *IEEE Trans. on Circuits and System for Video Tech.*, vol. 5, no 6, pp. 500-507, Dec. 1995.

[5] S. Chakrabarti and R. Wang, "Adaptive Control for Packet Video," *Proc. the International Conference on Multimedia Computing and Systems*, pp. 56-62, May 1994.

[6] T. Turletti and C. Huitema, "Videoconferencing on the Internet," *IEEE/ACM Trans. on Networking*, vol. 4, no. 3, pp. 340-351, May 1994.

[7] H. Sun, W. Kwok, and J. W. Zdepski, "Architectures for MPEG Compressed Bitstream Scaling," *IEICE Trans. on Circuits and Systems for Video Tech.*, vol. 6, No. 2, pp. 191-199, April 1996.

[8] Hee Dong Chang, Young Kwon Lim, Myoung Ho Lee, and Chieteuk Ahn, "A Scalable Video Transmission Method over The Internet," *Proc. Third Joint Workshop on Multimedia Communications*, pp. 5-2-1 - 5-2-8, Oct. 1996.