A Novel Content-Based Video Streaming Algorithm for Fine Granular Scalable Coding[1]

Li Zhao, Qi Wang, Yuwen He, Shiqiang Yang, and Yuzhuo Zhong

Department of Computer Science and Technology, Tsinghua University,
Beijing 10084, China
Li.Zhao@media.cs.tsinghua.edu.cn

Abstract. Most of current video streaming technologies, which protect the video stream by rate allocation or CRC means according to the bandwidth of network and the video quality user expected, is unrelated with the content of video. This paper proposes a new content-based video streaming method, which is based on video content analysis. The method discusses the rate distortion function of Fine Granular Scalable (FGS) enhancement layer bitstream based on the video content, and allocates the suitable bits for video units by the importance of video content. The experimental results indicate that our content-based rate allocation method can protect the transmission of video efficiently, improve the subjective quality of video and use the network bandwidth in maximum utility.

1 Introduction

Video streaming application over the Internet booms rapidly in recent years. Because of the lack of Quality-of-Service (QoS) guarantees over the current Internet, the media streaming data undergoes the dynamic variations of bandwidth, delay and packet loss. Due to the bandwidth of the channel between the encoder and decoder is not a fixed number but changes constantly along with time.

To resolve the issues of video streaming over Internet mentioned above, there are two general types of approaches have been proposed. The one is Network-Centric, which requests the router/switcher of the Internet to provide the guarantees of the bandwidth, delay, jitter, packet loss rate for video streaming applications. The other is End System-based, which does not rely on the implement of network at present or in the future and be achieved easily. In this paper, we focus on the second (End System-based) approach.

From the End System-based perspective, there are some conventional methods proposed to solve the issue mentioned above: (1) Transcode approach is to encode the video with the highest quality and store the bitstream at the server. Then the encoded bitstream is transcoded into different bitrate for different users [3]; (2) Simulcast approach is to generate multiple bitstreams of multiple bitrates at the encoder time. (3) Scalable encoding or adaptive encoding methods [5] [6]. (4) Rate-shaping approach is a technique to adapt the rate of compressed video bit streams to the target rate constraint[1]. Almost all of the current video streaming technologies, which

---

protect the video stream by rate allocation or CRC means[2] according to the bandwidth of network and the video quality user expected, is unrelated with the content of video.

In this paper, we propose a content-based video streaming algorithm for FGS coding, which combines the video streaming with video content analysis technique. This approach allocates the bits by video content based on video content analysis result and can protect the important video content by selectively and actively discarding some bits of unimportant video data while the bandwidth descends sharply.

The rest of this paper is organized as follows. In section 2 we focus on our content-based video streaming algorithm. Section 3 proposes our experimental results. Concluding remarks are in section 4.

2   Algorithms Description

The rate control is a fundamental technique in the encoding process, which is based on the rate-distortion theory[7][8]. The objective of rate control is to make encoded bitstream meet the bandwidth of channel. The rate allocation for the FGS enhancement is different from the traditional rate control in the encoding process for that it is independent of the FGS encoding progress and is executed by the feedback information from the channel before transmission.

In this paper, we propose a content-based video rate allocation algorithm for video streaming over Internet. In our method the video content is analyzed at first, including video structure analysis, motion information extraction, video objective retrieval/identification etc. Then allocate the bits for FGS enhancement based on the result of video content analysis. In FGS coding, the base layer restructures the low quality VOP and the enhancement layer, improves the quality of the VOP. In our rate allocation algorithm, we assume that the decoder can receive all base layer bits. We focus on how to allocate the enhancement bitstream by video content.

Segmented video clips are used to represent the video sequence: $\{\text{Seg}(1), \text{Seg}(2), \ldots \text{Seg}(N-1), \text{Seg}(N)\}$, here N denotes the total number of segmented video clips.

We propose a video content-weighted rate-distortion (R-D) function:

$$D(R) = w * \sigma^2 * e^{-aR} \tag{1}$$

Here, $D(R)$ denotes the mean square value of the error between decoded video clip and original video clip; w denotes the weight coefficient of the video clip, which is determined by the result of video content analysis; ? is a constant number; $?^2$ is the variance of the encoding signal; R is the bits number used to encode the video clip.

If R_i denotes the allocated bits number for encode video clip $Seg(i)$ and D_i denotes the mean square error of $Seg(i)$, the essence of the problem of rate allocation is a global optimization issue and can be described by the following formula:

$$f(R_1, R_2 .. R_3) = D_1(R_1) + D_2(R_2) + \cdots D_N(R_N) = \sum_{i=1}^{N} w_i * \sigma_i^2 * e^{-aR_i} \tag{2}$$

Subject to $R_1 + R_2 + \cdots R_N = S$ (const), S is the given total number of bits used for the sequence decoding.

Where $f(R_1, R._2..R_3)$ denotes the total degree of distortion of all video clips.

In order to minimize total distortion $f(R_1, R._2..R_3)$, we can apply Lagrange multiplier method to solve it. Firstly, we define the following function:

$$g(R_1, R_2..R_N) = f(R_1, R_2..R_N) - \lambda(S - (R_1 + R_2 + \cdots + R_N)) \tag{3}$$

Then, we obtain some equations:

$$\frac{dg}{dR_1} = \frac{dg}{dR_2} = \cdots = \frac{dg}{dR_N} = \frac{dg}{d\lambda} = 0 \tag{4}$$

We solve above equations:

$$R_i = \frac{C}{N} + \frac{2}{aN} \sum_{i \neq j} In \frac{\sigma_i}{\sigma_j} + \frac{1}{aN} \sum_{i \neq j} In \frac{w_i}{w_j} \tag{5}$$

Here, $i = 1,2,\cdots,N$.

However calculating $\frac{1}{a} In \frac{\sigma_i}{\sigma_j}$ is a not easy thing during the rate allocation

progress. So a simple method for estimating $\frac{1}{a} In \frac{\sigma_i}{\sigma_j}$ must be found. Similarly, we

think the distortions approximately equal when the all enhancement layers are fully decoded, i.e.

$$\sigma_i^2 e^{-a\hat{R}_i} = \sigma_j^2 e^{-a\hat{R}_j} \tag{6}$$

Now we can get:

$$\frac{1}{a} In \frac{\sigma_i}{\sigma_j} = \frac{\hat{R}_i - \hat{R}_j}{2} \tag{7}$$

Here, $\hat{R}_i$ and $\hat{R}_j$ denote the used bits number to fully encode the enhancement layer of video clip $Seg(i)$ and $Seg(j)$. In addition, this information can be generated during encoding.

Bring equation (7) into equation (5), the bit allocation can be simply represented as:

$$R_i = \frac{C}{N} + \frac{1}{N} \sum_{i \neq j} (\hat{R}_i - \hat{R}_j) + \frac{1}{aN} \sum_{i \neq j} In \frac{w_i}{w_j} \tag{8}$$

Here $i = 1,2,\cdots,N$.

3   Experimental Result

To evaluate the performance of the content-based video streaming algorithm we proposed in this paper, we use the *Stefan* sequence that is the standard test sequence of MPEG-4 as our experimental sequence. The parameters are set as follows: The sequence format is SIF (352*240). The encoding frame rate is 30f/s. The quantization scheme of the base layer is the same as in H.263. The bit rate of base layer is 320Kbits/s.

   In detail, in our experiment, we adopt motion activity to describe the importance of video content: we assume that the greater the motion activity is, the higher the importance of video content is. Here global motion estimation (GME) method is adopted to measure the motion activity. Detailed algorithm of robust GME can be referred in[4]. Then, according to the motion activity, the sequences are segmented into several segments with different importance degree. The segment result is given in Fig.1.
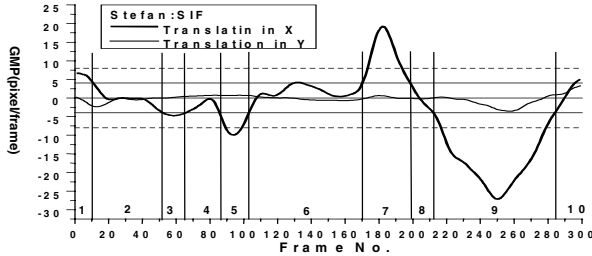

Fig. 1. The motion activity segmentation result according to global motion estimation method.

   In our experiment, the enhancement layer bits are allocated with two methods: One is to allocate the bits to each frame averagely. The other is to allocate the bits with our algorithm. Fig.2 shows the experimental result under the bitrates of enhancement layer at 320Kbits/s. Here the horizontal axis denotes the frame number and vertical axis denotes the SNR.
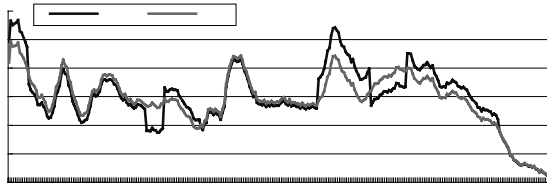

Fig. 2. The curves of PSNR versus frame number of "Stefan" (320Kbits/s)

   The experimental results prove that the proposed bit allocation scheme can improve the performance of FGS to 0.5~1.0 dB at the important parts. For example, the part of No. 175~200 of sequence "Stefan" has very high motion activity (See fig. 1). We can think this part is more important. From fig.2, we can see that the PSNR of our method at this part is 1.0 dB higher than contrasting method's.

We can draw the conclusion that the content-based rate allocation scheme proposed for video streaming can protect important video content while transmitting. In other words, this method can improve the video quality at the receiver and increase the utilization ratio of the channel.

4   Conclusions

In this paper, we propose a new content-based video rate allocation algorithm for video streaming over Internet, which is based on video content analysis and fine granular scalable video coding techniques. This scheme can stream video by content, so that it can protect the important video content while transmitting under the conditions of limited bandwidth of channel and improve the utilization of the channel. The experimental results show that the method proposed can protect video content efficiently and improve the subjective quality of important video segments.

In the future, we will discuss the content-based video streaming in big picture, including rate control, error resilience and error concealment etc research areas, and propose a whole architecture of content-based video streaming/delivery.

References

[1]   Jacobs, S., Eleftheriadis, A.: Real-Time Video on the Web Using Dynamic Rate Shaping, Image Processing, 1997. Proceedings., International Conference on Volume: 3 , 1997 , Page(s): 14 -17 vol.3.
[2]   Tan, W.T., Zakhor, A.: Real-Time Internet Video Using Error Resilient Scalable Compression and TCP-Friendly Transport Protocol, Multimedia, IEEE Transactions on Volume: 1 2 , June 1999 , Page(s): 172 -186.
[3]   Wang, L.M., Luthra, A., Eifrig, B.: Adaptive Rate Control for Mpeg Transcoder, Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on Volume: 4 , 1999 , Page(s): 266 -270 vol.4.
[4]   He, Y.W., Qi, W., Yang, S.Q., Zhong, Y.Z.: Feature-based Fast and Robust Global Motion Estimation Technique for Sprite Coding, *ISO/IEC JTC1/SC29/WG11, MPEG00/M6226*, July 2000.
[5]   Yu, Y., Chen, C.W.: SNR Scalable Transcoding for Video over Wireless Channels, Wireless Communications and Networking Conference, 2000. WCNC. 2000 IEEE Volume: 3 , 2000 , Page(s): 1398 –1402
[6]   Lee, B.R., Park, K.K.: H.263-Based SNR Scalable Video Codec, Consumer Electronics, 1997. Digest of Technical Papers. ICCE., International Conference on , Page(s): 254 –255
[7]   Chiang, T., Zhang, Y.Q.: A new rate control scheme using quadratic rate distortion model, IEEE Transactions on CSVT, 7(1): 246-250, Feb, 1997.
[8]   Mallat, S., Falzon, F.: Analysis of low bit rate image transform coding, IEEE Transactions on SP, 46(4): 1027-1042, Apr 1998.