Journal of Theoretical and Applied Information Technology
31st March 2013. Vol. 49 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

THE OPTIMAL PARTITION FOR SINGLE VIDEO SEQUENCE IN STREAMING MEDIA SYSTEM

[1]MANDE XIE

[1]School of Computer & Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang, 310018

E-mail: [1]xiemd@zjgsu.edu.cn

ABSTRACT

The fine grained scalable (FGS) coding is one of the most popular coding technologies in network video stream. In FGS stream, the base layer rate is an important parameter and should generally make a compromise between the number of beneficial clients and the efficiency of video coding. To maximize all clients' perceived quality, the optimal partition algorithm for single FGS-coded video sequence is proposed. By analyzing the algorithms oriented FGS-coded stream proposed by Hsu et al, this paper first puts forward a method to build the R-Q function curve and proposes an algorithm to cut search space of the base layer rate. Second, according to the optimization problem of single sequence, the paper proposes the improved FGSOPT_N algorithm. At last, some illustration examples are given to demonstrate and compare the effect of algorithms.

Keywords: *Fine Grained Scalable, Greedy Algorithm, Streaming Media System, Base Layer Rate*

1. INTRODUCTION

In recent years, the number of broadband users rapidly increases. As the main application of value added services of broadband, the streaming media service also has obtained fast development. Among the network services, the online film and TV watching and download services account for a considerable proportion. The market scale of online video service reaches to hundreds of millions yuan. Therefore, lots of big network operators, even the telecom operators are making efforts in the streaming media system construction. Meanwhile, in order to improve the clients' perception experience, the scalable coding is extensively being used. Due to the inherent low fault tolerance and limited rate expansion ability of multi-layers and multiple descriptions coding, FGS has become the most popular coding technology in network video, which has been adopted by MPEG4. The FGS encodes the stream sequence into a base layer and an enhancement layer, in which the former can be decoded separately to provide the basic video quality, the later will overlay on the base layer after being decoded, thus to provide a better video quality. In a streaming media system, the base layer rate is very important for content providers. If the base layer rate is high, this may increase perceived quality for some clients, which could allow the provider to charge higher service rates. However, the coding efficiency will be reduced. On the other hand, a higher base layer rate may disqualify other clients from receiving the complete base layer stream. Hence, this paper puts forward the optimal partition algorithm for multiple fine grained scalable video to maximize the clients' perception experience.

2. RELATED WORKS

In present, studies on the basic theories of streaming media system are mountainous, among which the streaming media playback and transmission technology are more mature. The hot topics focus on how to improve the quality of service in streaming media system, and the FGS coding method occupy wide attention. The biggest difference between the FGS and ordinary scalable coding is that the enhancement layer of the FGS adopts a Bit-Plane coding technology, which makes the enhancement layer bit stream of each frame truncated at any place, and the reconstructed video quality proportional to the number of bits received in decoding. The base layer coding of FGS is the same as that of common non-scalable coding. But as for the enhancement layer coding, the DCT residual will be got when the original DCT coefficients minus the reconstructed one after the inverse quantization of the base layer. After the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

zigzag scan of residuals, the plane coding will be adopted. According to the bit matrix of the coding sequence, that is, the first coding will be the top bits. The benefit of Bit-Plane coding is to code based on the priority of each factor, thus it ensures the video decoder to get the important video information in streaming truncation, and provides a fine scalability.

Literature [1] and [2] study the coding efficiency gap of FGS video decoder in the MPEG-4. The author surveys the relationship between FGS video coding efficiency gap and time domain correlation. They find that the correlation coefficients between enhancement layer frame and the motion compensated reference frames are good indications of the FGS coding efficiency. Literature [3], [4] and [5] study the coding efficiency of scalability coders through the layer overload. The layer overload refers to the bit rate that can not improve the video quality. However, no papers establish proper, accurate coding efficiency gap function, which cannot be helpful to propose the accurate algorithm. Through experiments, Hsu et al [7] shows that the appropriate base layer rate of FGS can improve clients' perception experience. But it doesn't put forward a systematic method for base layer rate choice. Radha et al [6] consider the optimization of a single stream that has multiple layers with different granularities, which can also be used in stream systems that have multiple non-scalable version of the stream to calculate the best rate of each version. Radulovic et al [5] compare the performance of scalable and non-scalable stream. To maximize the average video perception experience, the author formalizes a dynamic planning to calculate the rate of each layer, and selects the square root rate-distortion model to estimate coding efficiency of each layer. In literature [4], Liu et al study the broadcasting multilayer streaming video problem in the wireless celluar system in which the number of channel and users' capacity distribution are given. Yang et al [8] research the multi-cast stream system. The algorithm divides the receivers into several groups to maximize the system's overall utility. This system can be encoded as several cumulative layers as well as multiple versions that have different rates, but it doesn't take account of the layer overload. In a word, the existing literatures have not yet provided a systematic method to optimize the structure and maximize the average quality of all clients. Hence, Hsu et al. [9] [10] formalize the selection problem of the optimal base rate of single sequence, and put forward a optimized FGSOPT algorithm to maximize the average video perception experience of given clients distribution. On this basis, they further expand the method into multi-sequence model, and put forward a branch bound algorithm MFGSOPT and greedy algorithm MFGS. However, the algorithm FGSAG (FGS algorithm group) proposed by Hsu et al need to introduce the gap function, and in most cases, it can't find the optimal solution. This paper introduces a new optimization strategy to increase the probability of getting the optimal solution without the introduction of differences function.

3. OVERVIEW OF THE ALGORITHM FOR SINGLE VIDEO SEQUENCE

3.1 Modeling and Attribute Analysis of the Quality Gap Function

In order to define and formal the problem, Hsu et al first give the following definition:

Definition 1: Quality gap function $\Delta(r_b)$: The Quality gap function $\Delta(r_b)$ is the quality difference between a non-scalable stream and a fine-grained scalable stream coded with base layer rate r_b, when both streams are decoded at the same bit rate.

The quality gap function is the discrete function of r_b . They model the function through experiments, and draw the curve of the function. By the observation, the features of quality gap function are summarized as follows:

(1) FGS streams have lower coding efficiency than a non-scalable stream.

(2) Higher base layer rates lead to smaller quality gaps. This fact indicates that the quality gap $\Delta(r_b)$ is a non-increasing function of the base layer r_b .

(3) Sequences with different characteristics lead to different quality gap $\Delta(r_b)$.

3.2 The Algorithm for Single Video Sequence

The optimizing problem for the single video sequence can be formulated into a search problem which is to search the best base layer rate for the given video sequence, thus to improve the perceived quality for all clients. Since the clients are heterogeneous, firstly, the cluster analysis should be done. All the clients who have the same bandwidth b_c belong to the same class, and C refers to the final number of client class. Suppose the fraction of clients in each class $c(1 \le c \le C)$ is given by a probability mass function f_c, where $\sum_{c=1}^{C} f_c = 1$. Without loss of

Journal of Theoretical and Applied Information Technology
31st March 2013. Vol. 49 No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

generality, we can also suppose $b_1 < b_2 < \cdots < b_C$, $b_C \le r_{\max}$, where $r_{\max}$ refers to the most likely corresponding video sequence bit rate. Accordingly, the optimization problem of single video sequence that aims to the maximization of average perceived quality can be formalized into:

$$\max_{r_b} \sum_{c=1}^{C} q(b_c) f_c \,, \quad \text{Where } r_b \in [0, r_{\max}] \qquad (1)$$

Where $q(b_c)$ is the quality (measured as PSNR in dB) achieved by clients in class c.

The most direct solution is to search the optimal r_b by enumerating all possible values r_b in the interval. To reduce search space, Hsu et al. introduce and prove the theorem 1.

Theorem 1: An optimal solution r_b^* for the base layer rate that maximizes the average perceived quality for all users can be found at one of the rates b_c , where $1 \le c \le C$.

Based on the theorem 1, Hsu et al. design a FGSOPT algorithm, through the enumeration method to find the r_b^* value.

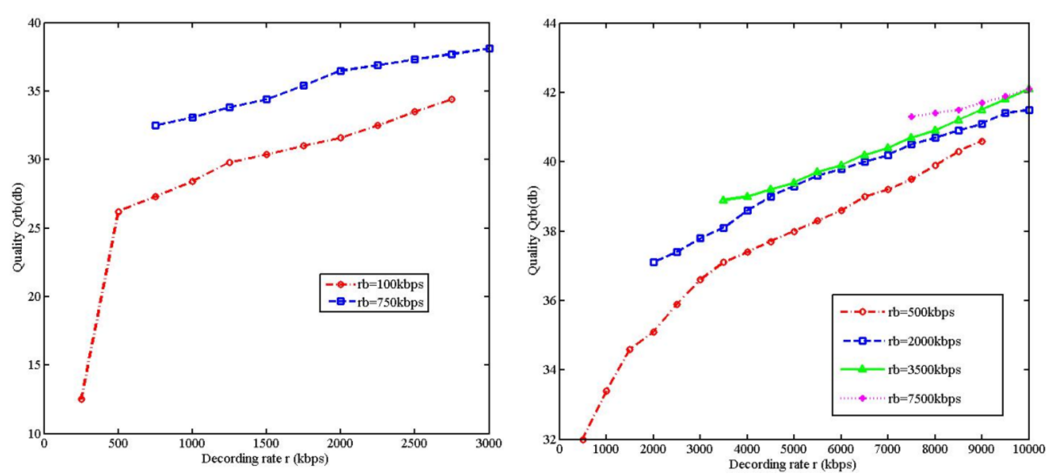
4. THE IMPROVED FGSOPT ALGORITHM

4.1 Building the R-Q Function Curve

In FGSOPT [9] [10], the quality gap function is introduced to compute the video quality under all kinds of base layers rate. From the view of building the R-Q function curve, the following steps are constantly repeated.

1) Encode the sequence with the non-scalable coder at rate $r_i, (1 \le i \le N)$, where N is the number of sample point in the curve.
2) Decoding the non-scalable sequence at decoding rate $r_i, (1 \le i \le N)$.
3) Calculate the video quality Q_{non} by comparing the original video with the reconstructed video.
4) Encode the sequence with FGS coder at base layer rate r_b (This step may not be repeated).
5) Decoding the FGS sequence at decoding rate $r_i, (1 \le i \le N)$.
6) Calculate the video quality Q_{FGS} by comparison the original video with the reconstructed video.
7) Compute the $\Delta(r_b) = Q_{non} - Q_{FGS}$.

As a result, tremendous computational overhead are required. Therefore, video quality gap function introduces a larger amount of calculation, which is unnecessary. It can be solved by building a quality table of a certain base layer rate directly. The method is to code sequence with FGS coder at a base layer rate r_b , then, decode the video at a variety of decoding rate, and calculate the video quality. Based on quality table, the function between bandwidth and quality at the base layer r_b can be obtained through the curve fitting method. First of all, the definition is given:



(a) Mobile (b) Crew

Figure.1 The video quality function curve

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Definition 2: video quality function $q_{r_b}(r_c)$: For a fine-grained scalable stream coded with base layer rate r_b, $q_{r_b}(r_c)$ is the video quality achieved by decoding it at the rate r_c, which is measured as PSNR in dB.

If the r_b is fixed, $q_{r_b}(r_c)$ is a discrete function related to rate r_c. The experiments similar to those in [9][10] will be used to model the function. We use the reference software of the Joint Scalable Video Model (JSVM)[11], which contains several executables programs: H264AVCEncoderLibTest, BitStreamExtractor, H264AVCDecoderLibTest, and FixedQPEncoder. Five diverse video sequences: City, Mobile, Harbour, Soccer, and Crew are chosen. The first four sequences are in CIF format with 30 frames per second, and the last one is in 4CIF format which is a high-resolution. We set r_b at a specific value and encode the whole stream. Then we compute the quality that would be perceived by various clients decoding the stream at different rates r_c. Assume clients lie in the range between 250 kbps and 3000 kbps with a step of 250 kbps for CIF sequences, and between 500 kbps and 10000 kbps with a step of 500 kbps for 4CIF sequences. The quality is determined by decoding the stream and computing the peak signal to noise ratio (PSNR) in dB.

According to the Rate -Quality (R-Q) point calculated above, the $q_{r_b}(r_c)$ curve can be drawn through the curve fitting way. Change the base layer rate r_b, and repeat the above experiment for different r_b, different curve can be drawn. In Figure 1, (a) and (b) show the quality function curve of Mobile and Crew video, respectively. In the program realization, when the decode rate r_c is given, the quality value Q can be achieved by looking up the table. Through the observation of lots of quality function curves, the following conclusions can be made:

When the receiving rate r_c is fixed, the quality function $q_{r_b}(r_c)$ is non-decreasing function of the base layer r_b. That is to say, according to the clients with the same receiving rate r_c, the stream with the higher base layer rates r_b will generate the better video quality.

4.1 The improved FGSOPT Algorithm

As for the optimization problem of single video sequence, Hsu et al. introduce and prove theorem 1. Furthermore, they design a FGSOPT algorithm. In fact, without the introduction of quality difference gap, theorem 1 is still valid.

Without the introduction of $\Delta(r_b)$, Theorem 1 can be proved as follows.

We can re-write the quality of the FGS stream $q(b_c)$ for clients in class c as:

$$q(b_c) = \begin{cases} 0 & b_c < r_b \\ q_{r_b}(b_c) & b_c \ge r_b \end{cases} \tag{2}$$

As shown in Eq. (2), the quality for clients in any class c is zero if these clients do not have enough bandwidth to receive the complete base layer, i.e., if $b_c \ge r_b$. The whole search range $[0, r_{\max}]$ can be divided into non-overlapping intervals $(b_{c-1}, b_c]$, where $c = 1,2,...,C$ and $b_0 = 0$. Now, assume that the optimal base layer rate r_b occurs in an arbitrary interval $(b_{z-1}, b_z]$. Because all clients with $b_c \le b_{z-1}$ receive quality of zero, the maximization problem is converted into:

$$\max_{r_b} \sum_{c=z}^{C} q_{r_b}(b_c) f_c, \quad where \;\; r_b \in (b_{z-1}, b_z]$$

(3)

Recall that in Section 3, we argued that when the receiving rate r_c is fixed, the quality function $q_{r_b}(r_c)$ is an increasing function of the base layer r_b . Since $q_{r_b}(r_c)$ is increasing in the interval $(b_{c-1}, b_c]$, no point in that interval could make the quality better than $r_b = b_z$. Thus, an optimal solution for r_b occurs at b_z.

Based on the theorem 1, the FGSOPT algorithm, called FGSOPT_N, is redesigned without the introduction of quality difference gap. Figure 2 shows the algorithm.

In Figure 2, $q_i (i = 1,2,...,C)$ indicates the sum of the video quality in all clients when the basic layer rate is r_{b_i}.

5. CONCLUSIONS

The FGS is one of the most popular coding technologies in network video and is adopted by

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MPEG4. In a streaming media system, the base layer rate is an important parameter when the content providers encode the video by FGS. If the base layer rate is high, content providers may redu-

FGSOPT_N
{
1) $q_1 = q_2 = ...q_C = 0$;
2) $q_{\max} = q_C = q_{b_C}(b_C)$;
3) $r_b^* = b_C$;
4) for $c = C-1$ to 1
{
5) $d = q_{b_c}(b_c)\, f_c + \sum_{i=c+1}^{C} (q_{b_{c+1}}(b_i) - q_{b_{c+1}}(b_i))\, f_i$;
6) $q_c = q_{c+1} + d$;
7) if ($q_c > q_{\max}$)
{
8) $q_{\max} = q_c$;
9) $r_b^* = b_c$;
}
}
10) return r_b^* ;
}

Figure. 2 The improved FGSOPT_N algorithm

ce the coding efficiency gap and achieve high quality and increase perceived quality for some clients. However, a high base layer rate may disqualify other clients from receiving the complete base layer stream. Hence, to maximize all clients' perceived quality, the optimal partition algorithm for multiple FGS-coded video is proposed. By analyzing the algorithms proposed by Hsu et al, this paper first puts forward a building method of R-Q function curve, and proposes a theorem to cut the search space of best base layer rate. As for the optimizing problem of single video sequence, the paper puts forward the improved FGSOPT_N algorithm based on the analysis of the defects of FGSOPT. Compared with FGSOPT proposed by HSU et al, the new algorithms provided by this paper can improve the probability to find the best base layer rate and reduce the runtime

6. ACKNOWLEDGMENTS

This work was partially supported by Grant No. 61100214 and No. 61100057 from the National Natural Science Foundation of China and Zhejiang Science and Technology Program No. 2011C14024 and No. 2010C13005 and Grant No. Z201122764 from Educational Committee of Zhejiang Province.

REFERENCES:

[1] H. Radha, M. van der Schaar, and Y. Chen. The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP. IEEE Transactions on Multimedia, 3(1):53–68, March 2001.

[2] M. van der Schaar and H. Radha. Adaptive motion-compensation fine-granular- scalability (AMC-FGS) for wireless video.IEEE Transactions on Circuits and Systems for Video Technology, 12(6):32–51, June 2002.

[3] De Cuetos, D. Saparilla, and K. Ross. Adaptive streaming of stored video in a TCP-friendly context: Multiple versionsor multiple layers? In Proc. of International Packet Video Workshop (PV'01), Kyongju, Korea, Acpril 2001.

[4] J. Liu, B. Li, Y. Hou, and I. Chlamtac. Dynamic layering and bandwidth allocation for multi-session video broadcasting with general utility functions. In Proc. of IEEE INFOCOM'03, pages 630–640, San Francisco, CA, March 2003.

[5] Radulovic, P. Frossard, and O. Verscheure. Adaptive video streaming in lossy networks: versions or layers? In Proc. Of IEEE International Conference on Multimedia and Expo (ICME'04), pages 1915–1918, Taipei, Taiwan, June 2004.

[6] H. Radha, M. van der Schaar, and S. Karande. Scalable video transcaling for the wireless internet. EURASIP Journal onApplied Signal Processing, 24(2):265–279, February 2004.

[7] C. Hsu and M. Hefeeda. Structuring multi-layer scalable streams to maximize client-perceived quality. In Proc. of the IEEE International Workshop on Quality of Service (IWQoS'07), pages 182–187, Evanston, IL, June 2007.

[8] Y. Yang, M. Kim, and S. Lam. Optimal partitioning of multicast receivers. In Proc. of IEEE International Conference on Network Protocols (ICNP'00), pages:129–140, Osaka, Japan, November 2000

[9] C. Hsu and M. Hefeeda, Optimal Partitioning of Fine-Grained Scalable Video Streams, In Proc. of ACM International Workshop on Network and Operating Systems Support for Digital Audio & Video (NOSSDAV'07), pp. 63--68, Urbana-Champion, IL, June 2007

[10] C. Hsu and M. Hefeeda, Partitioning of Multiple Fine-Grained Scalable Video Sequences Concurrently Streamed to Heterogeneous Clients, IEEE Transactions on Multimedia, 10(3), pp. 457--469, April 2008.

[11] Joint Video Team. Joint scalable video model reference software. JSVM 8.0, February 2007.