TUPM 5.1

A User Adaptive Perceptual Rate Control Scheme for FGS Videos

Chia-Chiang Ho, *Student Member, IEEE*, Ja-Ling Wu, *Senior Member, IEEE*, and Sung-Wen Wang
Communications and Multimedia Lab., National Taiwan University, Taipei, Taiwan

***Abstract*—MPEG-4 fine-granularity-scalable video coding is novel for supporting continuous bitrate spectrum. In this paper, we propose a new video quality metric that considers the user's preference of spatial-temporal quality tradeoff. A practical rate control scheme for FGS videos is developed accordingly, to provide constant perceived quality over time.**

I. INTRODUCTION

Among several scalable coding schemes standardized in the new MPEG-4 standard, the fine-granularity-scalable (FGS) scheme receives great attention [1]. While being capable of providing both temporal and signal-to-noise ratio (SNR) scalabilities, FGS scheme is novel for providing the characteristic of continuous bitrate spectrum, using only one layer of enhancement bitstream. This is achieved through bit-plane coding, by which the generated enhancement layer bitstream can be truncated at any point. Nonetheless, when streaming FGS-coded bitstreams, a proper rate control scheme is required to achieve good balance between transmission rate and user perceived quality. In this paper, we propose a new video quality measurement that takes user preference into account. Then a new rate control scheme for FGS videos is proposed to provide constant perceived quality over time.

Although there are several researches concerning rate control of FGS videos [2]-[5], only few of them consider both temporal and SNR scalabilities together to optimize perceived quality. For one video segment, the scheme proposes in [4] determines either SNR or temporal quality should be maximized. However, such arbitration may be considered too rough. The scheme proposed in [5] depends on pre-determined subjective measurement to provide clues for achieving optimal SNR-temporal tradeoff. But we all knew that human interference is time consuming and inefficient. The contributions of our work are not only providing a practical solution but also taking user preference into account. By the aid of pre-stored metadata, the proposed rate control scheme can be performed with low computation burden, and can be easily applied to many applications based on real-time FGS video streaming. We believe that the efficiency and flexibility of the proposed scheme provide appealing add-on value for these applications.

This work was supported in part by the National Science Council, R.O.C., under the contract NSC91-2622-E-002-040.

II. THE PROPOSED RATE CONTROL SCHEME

A. User Adaptive Video Quality Index (UAVQI)

Peak signal-to-noise ratio (PSNR) is the most widely used image quality metric. However, it is widely criticized as well, for not correlating well with perceived quality measured by subjective quality assessments. Lots of researches were devoted to develop new objective measurement by taking characteristics of human visual system into account. However, most efforts are not so persuasive. For example, in 2000, the Video Quality Experts Group (VQEG) concluded that performances of those proposed perceptual quality measurements are statistically indistinguishable from that of PSNR [6]. While most previous researches were based on modeling of error sensitivity, Wang and Bovik took another point of view and proposed a new image quality measurement recently [7]. The measurement is called *Universal Image Quality Index* and is abbreviated as UIQI below. UIQI is based on structural distortions, including loss of correlation, mean distortion and variance distortion. The dynamic range of UIQI is [-1,1]. In our work, we biased the range to avoid negative value, that is, $Q = UIQI + 1$.

We knew that different individual might have different quality perception over the same media material. This motivates us to propose a new video quality measurement that takes explicit user preference into account. For a video segment $v=\{F_i \mid i=1,2,\ldots,N_v\}$ with N_v frames, we define the *User Adaptive Video Quality Index* (abbreviated as UAVQI) as

$$UAVQI(v) = \frac{1}{N_v}\sum_{i=1}^{N_v}\alpha(F_i)Q(F_i) \quad . \tag{1}$$

In which $Q(F_i)$ is the biased UIQI of the frame F_i. $\alpha(F_i)$ is the temporal quality decay factor, and is defined as

$$\alpha(F_i) = 1 + a\frac{f_i - f_{\max}}{f_{\max}}, \tag{2}$$

where a is the temporal quality decay constant (TQDC) which can be specified by the user, f_i is the instant frame rate of frame F_i, and $f_{\max}$ is the maximal supported frame rate. The valid range of TQDC is [0,1]. When a approaches zero, spatial quality is preferred. When a approaches one, temporal smoothness is preferred. The definition of UAVQI has taken both cases of dynamic and constant frame rates into consideration.

B. Temporal Scalability

Temporal scalability can be approximated by selective frame dropping of B-frames. In our experiments, we can

0-7803-7721-4/03 $17.00 © 2003 IEEE

support different frame rates by different frame dropping patterns.

C. Image Quality Scalability

Although the FGS enhancement layer bitstream can be truncated at any point, we restrict that the sent bitstream has been truncated only at the end of any coded bit-plane. This is because 1) there is no significant evidence that sending incomplete bit-plane do improve the overall perceived quality, and 2) we don't know if the decoder will discard the incomplete bit-plane or try to decode it.

D. Image Quality Scalability

A *transmission path* of one video segment v is an N_v-tuple $p=\{R_{F_1},\cdots,R_{F_{N_v}}\}$, where R_{F_i} is the bitcount allocated to frame F_i, including base layer and truncated enhancement layer bitstreams. The corresponding UIQI values of all frames are $\{Q_{F_1},\cdots,Q_{F_{N_v}}\}$. For one video segment v and a transmission path p, the UAVQI value is defined as:

$$UAVQI(v,p)=\sum_{i=1}^{N_v}\alpha(F_i)Q_{F_i}. \tag{3}$$

The problem to be solved is to choose a subset $v'=\{F_j \mid j=1,2,\cdots,N_{v'}\}$ and a transmission path p' on v', such that the corresponding UAVQI(v', p') is maximized and $\sum_{j=1}^{N_{v'}}\left|Q_{F_j}-Q_{F_{j-1}}\right|$ are minimized, under the constraint that $\sum_{j=1}^{N_{v'}}R_{F_j}\le R$. Here R is the bitrate constraint, which can be fixed or calculated based on available bandwidth.

E. Optimization by Newton's method

In our thought, the optimization problem is solved by 1) finding the solutions which minimize the quality variation, and then 2) finding the best one among those solutions obtained in 1), whose UAVQI value is maximal. Specifically, when v' is restricted to be of some constant frame rates, we can firstly figure out optimal transmission paths of different v', and then find the global optimal (v', p') pair among these transmission paths. The solution of the first part is found by using Newton's iteration method [2], based on the fact that the quality-rate curves are monotonous.

F. Miscellaneous considerations

A practical bitrate allocation algorithm for real-time FGS video transmission should be fast and efficient. Piecewise linear quality-rate curves for all frames, are off-lined calculated and stored. For efficiency, we restrict the number of iterations to be less than a pre-defined number. Another criterion to stop the iteration is when the error of rate allocation is smaller than a given threshold. The rate allocation error is compensated in the next video segment, so it won't accumulate in the long run. We expect that the user will prefer a constant frame rate for one video scene, so we use a scene cut detection algorithm to divide the input video into distinct video scenes. Then the rate allocation work is performed once for each video scene (segment) v in the runtime.

III. CONCLUSION

Our preliminary experiments gave evidences of the applicability, efficiency and user satisfaction of the proposed rate allocation scheme. A possible future work is to develop an automatic mechanism which determines the suitable value of a (defined in (2)) according to characteristic of input videos. Together with occasional user interference, we can then provide optimal perceived video quality subject to the viewpoint of video viewers.

REFERENCES

[1] W. Li, "Overview of the fine granularity scalability in MPEG-4 video standard," *IEEE Trans. Circuits Syst. for Video Technol.*, vol. 11, no. 3, pp. 301-317, Mar. 2001.

[2] X.-J. Zhao, Y.-W. He, S.-Q. Yang, Y.-Z. Zhong, "Rate Allocation of Equal Image Quality for MPEG-4 FGS Video Streaming", in *PV'2002*, Apr. 2002, Pittsburgh PA, USA.

[3] L. Zhao, J. Kim, and C.-C. J. Kuo, "Streaming MPEG-4 FGS Video with Constant-Quality Rate Control and Differentiated Forwarding", in *Proc. VCIP'02*, San Jose, CA, Jan. 2002.

[4] M. van der Schaar and H. Radha, "Temporal-SNR rate-control for Fine-Granular Scalability," in *Proc. IEEE ICIP'01*, vol. 2, pp. 1037-1040, Greece, Oct. 2001.

[5] R. K. Rajendran, M. van der Schaar, and S.-F. Chang, "FGS+: Optimizing the joint SNR-temporal video quality in MPEG-4 fine grained scalable coding," in *Proc. IEEE ISCAS'02*, vol. 1, pp. 445-448, Scottsdale, Arizona, May 2002.

[6] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," *http://www.vqeg.org*, Mar. 2000.

[7] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Processing Letters*, vol. 9 no. 3, pp. 81-84, Mar. 2002.

[8] Z. Wang, A. C. Bovik and L. Lu, "Why is image quality assessment so difficult," in *Proc. IEEE ICASSP'02*, vol. 4, pp. 3313-3316, Orlando, FL, May 2002.