저속 영상부호화를 위한 최적 프레임 율과 공간 양자화 결정

정회원 부 소 영*, 이 병 욱*

An Optimal Selection of Frame Skip and Spatial Quantization for Low Bit Rate Video Coding

So-Young Bu*, Byung-Uk Lee* ***Regular Members***

요 약

본 논문에서는 낮은 전송률에서 공간적 시간적 해상도를 tradeoff하는 새로운 비디오 코딩 알고리듬을 제안한다. 먼저 영상의 움직임을 측정하여 최적의 frame rate를 선택한다. DCT quantization parameter (QP)는 앞에서 결정된 frame rate와 주어진 bit rate에 따라 결정한다. 마지막으로, 본 논문은 skip된 비디오 영상에 대해 단순하고 효과적인 distortion 측정방법을 제안한다. 실험결과, 제안한 알고리듬이 H.263 TMN5 보다 영상 품질이 2 dB 정도 개선되었다.

Key Words : rate control, frame rate, skipped frame, video coding, spatio-temporal optimization

ABSTRACT

We present a new video coding technique to tradeoff frame rate and picture quality for low bit rate video coding. We show a model equation for selecting the optimal frame rate from the motion content of the source video. We can determine DCT quantization parameter (QP) using the frame rate and bit rate. For objective video quality measurement we propose a simple and effective error measure for skipped frames. The proposed method enhances the video quality up to 2 dB over the H.263 TMN5 encoder.

I. Introduction

Video compression is widely employed for video communication over the internet or wireless channels. Video compression standards, such as MPEG-4 and H. 263 achieve acceptable video quality over limited bandwidth of mobile communication [1][2]. The compression standards guarantee decoder compatibility. However it is an open issue how to achieve an optimal encoding: we need to decide the number of frames to be coded and the quantization parameter for a frame. High quality encoding is based on rate-distortion models [3] to search for optimum encoding parameters for minimum distortion at a given bit budget. Most of works have been limited to spatial optimization assuming that every source frame is encoded; however it is unavoidable to skip frames from encoding at very low bit rate under 50 Kbps. Therefore we must choose between encoding more frames with low image quality or sending less frames with better video quality. We present an intuitive measure for objective video quality measurement, and a model

* 이화여자대학교 정보통신학과 영상신호처리 연구실 (bulee@ewha.ac.kr)
논문번호 : 030146-0401, 접수일자 : 2003년 4월 3일

for tradeoff between the optimal number of skipped frames and spatial quantization parameter depending on the motion content of the source video.

Ⅱ. Video Compression and Video Quality

Video signal has a lot of temporal and spatial redundancy. Temporal redundancy comes from high similarity between adjacent frames while spatial redundancy arises from the high correlation among the neighboring pixels. The video compression standards eliminate the redundancy without noticeable degradation in image quality. Video compression reduces the cost of video delivery over wired and wireless networks.

There are two types of data channels, i.e. constant bit rate (CBR) and variable bit rate (VBR) channels. Mobile channels can have time varying channel capacity depending on the number of users and signal level. Therefore a video encoder for such a channel must adapt the bit rate depending on the channel condition [1]. The proposed algorithm can be applied to both CBR and VBR rate control.

We propose an encoding algorithm achieving spatio-temporal optimization: we first determine optimal frame rate depending on motion content of the source video and then decide quantization parameter from the given bit rate. The basic idea is that we reduce frame rate for motionless video so that we can spend more bits per frame, and vice versa. The block diagram of encoding is shown in Fig. 1.

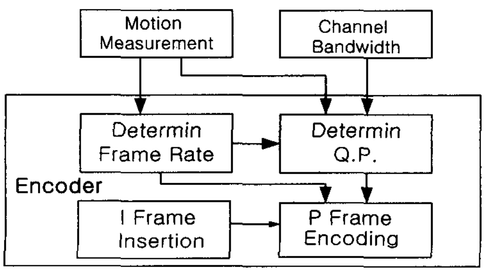

Fig. 1. Block diagram of the proposed spatio-temporal encoding algorithm

We need to define an objective measure for video quality for spatio-temporal optimization. The most popular measure is PSNR (peak signal to noise ratio). It compares the original image with the coded image. For video encoding which allows frame skip, we need to define an error measure for skipped frames. If we measure the error of coded video ignoring the skipped frame, we will have better PSNR when we send few frames with high quality, which is not acceptable especially for moving scenes. Vetro et. al. proposed a statistical model based on optical flow from computer vision. It is based on image spatial gradient and variance of the motion vectors [4]. Liu and Kuo proposed a method of affine warping [5]. They assume that a skipped frame is interpolated by affine transformation of the previous coded frame. However finding affine parameters is time consuming and of no practical value, since decoders do not employ affine transform for temporal interpolation. The affine parameters are not transmitted to a decoder. In most cases decoders hold the previous frame until the next coded frame is decoded.

We propose a simple and effective error measure. Suppose a frame N is encoded and then frames N+1 through N+4 are skipped. Since frame N+1 is not coded, we may want to compare it with the previous coded frame N, which is reasonable since most decoders repeat the previous frame for a skipped image. However if there is a scene change at a skipped frame, it will result in huge error. For the remedy we compare the original frame with adjacent coded frames N and N+5 as shown in Fig. 2, and then elect the minimum of the two. Equations (1) through (5) show the definition of the error measure. However there is a minor shortcoming: timing jitter of scene change is ignored. A scene change in the source video may appear to be delayed if it happens during one of the skipped frames. Since it is very rare to drop consecutive frames, the delay of scene change is usually limited to one frame period.

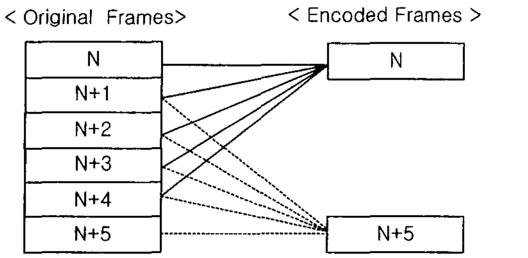

Fig. 2. Proposed video error measure

The following Eqn. (1) shows the difference of the original frame n to the previous coded frame and Eqn. (2) shows the error from a coded frame in the nearest future. The PSNR in Eqn. (5) depends on the mean square error of coded frames as well as skipped ones.

$$MSE_1 = \frac{1}{P} \sum_{i=0}^{P-1} (OF_n(i) - EFP_n(i))^2 \quad (1)$$

$$MSE_2 = \frac{1}{P} \sum_{i=0}^{P-1} (OF_n(i) - EFF_n(i))^2 \quad (2)$$

$$MSE_n = \min(MSE_1, MSE_2) \quad (3)$$

$$D = \frac{1}{N} \sum_{n=0}^{N-1} MSE_n \quad (4)$$

$$PSNR = 10 \log_{10}(255^2 / D) \quad (5)$$

P: Picture Size, OF: Original Frame
EFP: Encoded Frame of the nearest past
EFF: Encoded Frame in the nearest future
D: Average mean square error (MSE) for N frames

Ⅲ. Optimal Frame Skip and Quantization Parameter

Image quality depends on bit rate and image complexity. Fig. 3 shows the dependency of the image quality on bit rate and image complexity [6]. We obtain higher picture quality for high bit rates and simple scenes. We need to allocate more bits for a complex content video to achieve the same level of quality as a low complexity video. The motion content in the source video can be an indicator for video complexity.

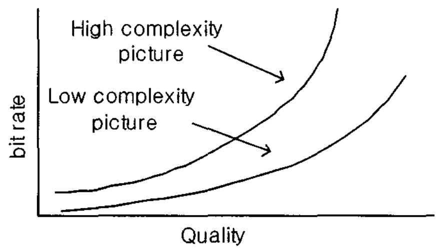

Fig.3. Relationship of bit rate and video quality

1. Motion Measure

It is intuitive to send more frames with low video quality if the video has many changes, while it is better to send a high quality image and skip the same images for a static sequence. The proposed method determines the frame rate from motion measure of the video. We measure the mean square of the difference of the previous and the current frame as shown in Eqn. (6), and use the mean of the difference over the N frames as in Eqn. (7). This measure determines the coded frame rate very accurately as will be shown in the following sections.

$$Diff_n = \frac{1}{P} \sum_{i=0}^{P-1} (OF_n(i) - OF_{n+1}(i))^2 \quad (6)$$

$$Motion = \frac{1}{N} \sum_{n=0}^{N-1} Diff_n \quad (7)$$

P: Picture Size
Motion: average of N frame difference
Diff: frame difference between the two adjacent frames.

2. Optimal Frame Rate

We model the relationship of the optimum frame rate and the proposed motion measure for a given bit rate. It is intuitive that higher frame rate is desirable for rapidly changing scenes. We measure motion content for a window of 100 frames. Fig. 4. shows the variation of PSNR depending on the coded frame rate for Container sequence. It shows that 4.3 Hz is the optimal frame rate at bit rate of 26 Kbps. Fig. 5 shows the same experiment on Mother & Daughter test sequence and we observe that the optimal frame rate is 10 Hz. The observations are in agreement with our intuition since the Mother & Daughter sequence has more motion contents as shown in Table 1.

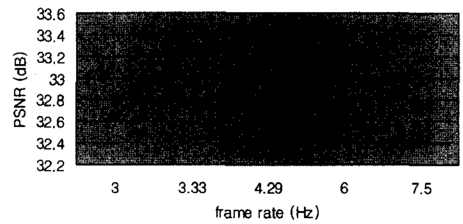

Fig. 4. PSNR of Container sequence for various frame rates. The bit rate is 26 Kbps.

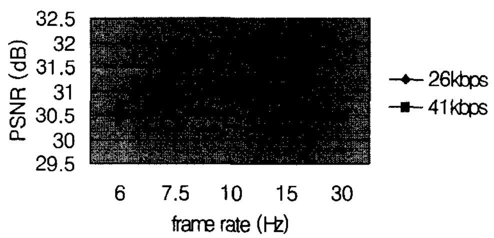

Fig. 5. PSNR of Mother & Daughter sequence at 26 Kbps and 41 Kbps

As exemplified in Fig. 5, the optimum frame rate is not sensitive to bit rate. We have done extensive experiments in the range of 20 ~ 60 Kbps, and concluded that the effect of bit rate can be ignored in the range.

The desired frame rate is implemented by skipping frames during the encoding process. The source frame rate is 30 Hz for NTSC television, and encoder skips a frame after every S frames. Then the coded frame rate is given as

$$Coded\ frame\ rate = 30\frac{S}{S+1} \quad (8)$$

Table 1. Optimal frame skip for sample video

	Container	Salesman	Mother& Daughter	Carphone
Motion Measure	277	1179	2457	6005
Optimal frame skip	6	2	2	1

Table 1 shows experimental results. We found that the optimal frame skip and the motion measure is inversely proportional from experiments and model the relationship as shown in the following equation,

$$\text{frame skip} = \text{round}(1390/M + 1) \quad (9)$$

where, M is the motion measure defined in Eqn.(7). The model equation matches the experimental data.

3. Quantization Parameter for DCT

DCT Quantization Parameter (QP) in H. 263 or MPEG-1, 2, and 4 has a range from 1 to 31. As QP increases it reduces the size of video data. At the same time it increases quantization error, thus resulting in degradation of image quality. As the quantization parameter increases, the corresponding image quality degrades. The image quality is the best when the QP is 1.

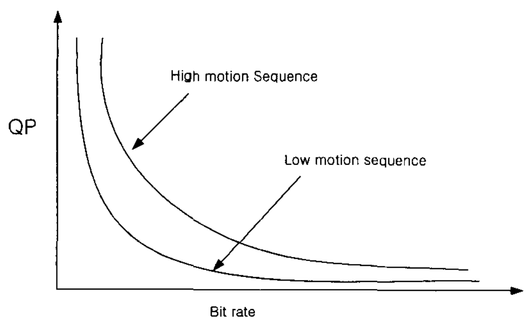

Fig. 6. The relationship of bit rate and Quantization Parameter

The optimal frame rate was determined by the motion measure M, and the QP can be decided by the available bandwidth of the transmission channel. Fig. 6 shows our model for QP and bit rate for H.263 TMN 5 encoder.

We employ test sequences such as Container, Salesman, Mother&Daughter, Carphone to derive the model parameters. Table 2. shows a results of experiments showing the quantization parameter for a given bit rate on H.263 encoder. We model the inversely proportional relationship of QP and bit rate of Fig. 6 as

$$QP=round(c/bitrate + d), \quad (10)$$

Table. 2. Optimal QP from experiments

Bit rate / Motion Measure	20 kbps	30 kbps	40 kbps	50 kbps	60 kbps
277	9	6	5	4	3
1179	13	8	7	6	5
2457	14	11	8	7	6
6005	23	17	12	10	9

where the parameters c and d depends on the motion contents as shown in the following equations. It is obtained from experiments as following equations.

$$c = 32.8 \ln^2 M - 387.3 \ln M + 1315.7 \quad (11)$$

$$d = 0.408 \ln M - 1.83 \quad (12)$$

We can verify that the model equations follow the data faithfully.

Table. 3. QP from the model equations

Bit rate / Motion Measure	20 kbps	30 kbps	40 kbps	50 kbps	60 kbps
277	9	6	5	4	3
1179	12	8	6	5	5
2457	16	11	9	7	6
6005	23	16	12	10	9

Ⅳ. Experimental Results

We analyze motion measure for 100 frames and determine the optimal frame rate using the proposed model Eqn. (9). Initial quantization parameter is determined by the proposed model Eqn. (10). The bit rate is confined to 20 ~ 60 Kbps which is typical for wireless channels. The actual bit rate may differ from the target bit rate.

We compare the PSNR of H.263 TMN 5 rate control algorithm and the proposed algorithm for 200 frames. Table 3 shows the performance of the proposed algorithm for sample sequences employed in the modeling. We observe that the proposed algorithm shows up to 2 dB improvements over the current method. Table 4 shows the same experiment applied to new test sequences that have not been employed during the modeling. We still notice that the improvement shows similar trend. It assures that the proposed algorithm is effective for the wide range of video sequences.

Table 4. Performance of proposed algorithm for test sequences employed for the modeling

Target bit rate(kbps)		20	30	40	50	60
Container Motion measure: 277(0-99) 271(100-199)	actual bit rate(kbps)	19.81	30.90	39.91	50.16	74.64
	PSNR(proposed)	32.37	33.83	34.65	35.30	36.61
	PSNR(TMN5)	30.28	32.29	33.65	34.68	36.55
Salesman Motion measure: 1179(0-99) 984(100-199)	actual bit rate(kbps)	22.06	33.56	48.58	63.41	71.52
	PSNR(proposed)	30.95	31.91	33.30	34.29	34.70
	PSNR(TMN5)	30.50	31.67	33.13	34.17	34.62
Mother& Daughter Motion measure: 2457(0-99) 494(100-199)	actual bit rate(kbps)	19.53	29.29	42.38	57.65	63.26
	PSNR(proposed)	31.35	32.42	33.27	34.03	34.30
	PSNR(TMN5)	30.48	31.69	32.76	33.72	33.90
Carphone Motion measure: 6005(0-99) 10529(100-199)	actual bit rate(kbps)	23.71	29.83	39.21	49.71	59.68
	PSNR(proposed)	28.45	29.12	29.77	30.25	30.59
	PSNR(TMN5)	28.07	28.34	29.00	29.25	29.47

Table 5. Performance of proposed algorithm for test sequences not employed during the modeling

Target bit rate(kbps)		20	30	40	50	60
News Motion measure: 3446(0-99) 5809(100-199)	actual bit rate(kbps)	24.15	33.49	44.93	49.71	59.68
	PSNR(proposed)	29.50	30.38	31.09	30.25	30.59
	PSNR(TMN5)	28.49	29.34	30.27	29.25	29.47
Silent Motion measure: 4175(0-99) 4293(100-199)	actual bit rate(kbps)	21.66	29.50	42.78	54.98	73.46
	PSNR(proposed)	29.44	30.15	32.90	33.60	34.40
	PSNR(TMN5)	28.86	29.75	30.95	32.14	33.18
Foreman Motion measure: 12211(0-99) 11636 (100-199)	actual bit rate(kbps)	30.87	36.35	45.93	55.37	66.68
	PSNR(proposed)	26.93	27.52	28.07	28.43	28.75
	PSNR(TMN5)	24.59	25.22	26.09	26.32	26.87

V. Conclusion

We propose a simple and effective algorithm for optimal encoding of video at low bit rate, which selects the optimum frame rate from a motion measure. We define a motion measure from mean square of adjacent frame differences. Then we present an experimental equation that derives the optimum frame rate from the motion measure. We observe that the frame rate is not sensitive to bit rate in 20 ~ 60 Kbps range. Optimum quantization parameter is determined from target bit rate and the motion measure, and experimental equations are also presented. A new PSNR measure for skipped frames played an essential role in comparing various encoding parameters. The measure is fast and effective.

Future work includes improvements in motion measure, optimal size of window for motion measure and buffer control.

References

[1] Wang Yao, Ostermann Jorn and Zhang Ya-Qin, Video Processing and Communications, Prentice Hall, 2001.

[2] Berna Erol and Michael Gallant, "H.263+; Video Coding at Low Bit Rates", IEEE transactions on circuits and systems for video technology, Vol.8, No.7, November 1998.

[3] N.S. Jayant, and Peter Noll, Digital Coding of Waveforms, Prentice-Hall, 1984.

[4] Anthony Vetro, Yao Wang, and Huifang Sun, "Rate-Distortion Optimized Video Coding Considering Frameskip," International Conference on Image Processing, Oct. 2001, Thessaloniki, Greece.

[5] Shan Liu and C.-C. Jay Kuo, "Complexity Reduction of Joint Temporal-spatial Bit Allocation Using R-D Models for Video Streaming," International Conference on Image Processing, Sep. 2002, Rochester, N.Y.

[6] Anthony G. Nguyen and Jeng-Neng Hwang, "Scene Context Dependent Rate Control", Proceedings of Ninth ACM International Conference on Multimedia, 2001, pp. 309-318, Ottawa, Canada

부 소 영(So-Young Bu) 정회원

2000년 2월 : 이화 여자 대학교 전자공학과 졸업

2002년 8월 : 이화 여자 대학교 정보통신학과 석사

2003년 9월~현재 : 컴퓨터 그래픽스/가상현실 연구소 연구원

<관심분야> 영상처리, 영상통신

이 병 욱(Byung-Uk Lee) 정회원

이화여자대학교 정보통신학과 교수

한국통신학회논문지 제 22권 6 호 참조