

Video Quality Estimation for Packet Loss Based on No-Reference Method

Junghyun Han, Yo-han Kim, Jangkeun Jeong, and Jitae Shin

School of Information and Communication Engineering

Sungkyunkwan University, Suwon, 440-746, Korea

{kazikai, dos95, carrot252, jtshin}@skku.edu

Abstract— Objective video-quality assessment can be divided into Full Reference (FR), Reduced Reference (RR), and No Reference (NR) methods. Among them NR quality assessment method does not require original video sequence, so it is suitable for the real-time video streaming service. This paper presents an NR-based estimation of mean square error (MSE) for video quality degradation due to various pattern packet losses in packet video network environment. Our proposed algorithm predicts initial error due to packet loss in each frame by using motion information of received data. Proposed method shows high Pearson correlation (more than 95%) to match actual MSE.

Keywords— Packet Loss, Video Quality Estimation, No-Reference, Motion Vector, Mean Square Error (MSE)

I. INTRODUCTION

It is a rapid demand that the development of Internet and compression codec based on the process of real-time video streaming. Channel error can cause reconstructed frame at the decoder to be incorrect, which leads to significant error propagation to subsequent frames. It causes video quality degradation. For video quality assessment method, video quality metrics have been developed and standardized by International Telecommunication Union (ITU) and Video Quality Experts Group (VQEG) [1] [2].

Objective quality assessment can be divided into 3 categories according to the availability of the original video sequence: Full Reference (FR), Reduced Reference (RR) and No Reference (NR). FR metrics need the whole original video sequence as reference. RR metrics only need part of the information from the original video sequence. And finally, No Reference (NR) metrics need no information of original sequence and requires only distorted video information. Most widely used metrics for FR objective video quality are Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). They are simple and quick to calculate and provide a good method to evaluate the video quality [3].

For example, MSE is defined as

$$MSE_{kth} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (1)$$

where y_i is an original video sequence and $\hat{y}_i$ is the reconstructed video sequence. But FR-based methods are not practical for remote communication system, especially limited bandwidth environments due to require entire sender-side content. Then NR metrics are preferable to communication system. A conventional NR method is divided into no reference pixel (NR-P) method and no reference bitstream (NR-B) methods. Primarily, blockiness analysis and blurring analysis are examples of the NR-P method [4] [5]. And the NR-B method is a video quality assessment using only received bitstream through quantization parameter, error propagation model and error concealment method [6] [7]. But NR-based metrics have complex formula and lower Pearson correlation [8].

In this paper, we propose an NR-B based objective-quality metric that replaces the MSE metric and uses only received encoded streams to estimate reconstructed video quality. This paper investigates NR-B methods to estimate the MSE distortion under various packet loss patterns. Our applicable video limits to H.264/AVC video streaming system over packet video network. And we assume that each frame is coded into a single packet for low-bit wireless networks, and the loss of a packet corresponds to the loss of an entire frame.

The organization of the paper is as follows: In section 2, we describe the proposed algorithm for packet loss pattern. In section 3, we show the experimental results and section 4 presents conclusion.

II. PROPOSED ALGORITHM

In this section, objective NR video-quality metric for video quality degradation based packet loss is considered. Most video compression algorithms including MPEG-1/2/4 and H.263/H.264 are based on motion compensated prediction. In video encoder case, inter-frame prediction method is used.

If packet loss occurs, then error propagation is shown in Figure 1.

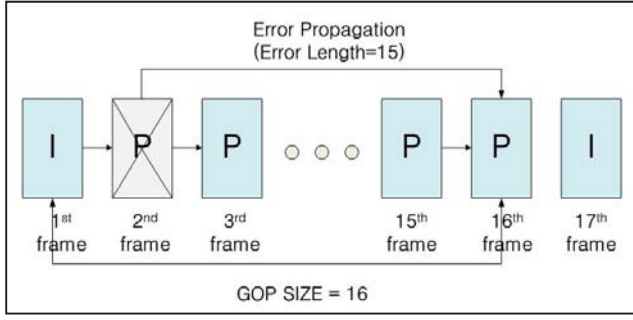


Figure 1. Error Propagation Caused by Single Packet Loss

If a single loss is occurred at k -th frame, the total MSE is given by

$$MSE_{kth}^{Single} = \sum_{i=0}^{El} (1 - \frac{i}{n}) \times MSE_{initial} \quad (2)$$

where MSE_{kth} is total distortion caused from the initial MSE of k -th frame's loss, $MSE_{initial}$. El is the number of frame between current loss location and next intra frame and n is number of frames in a group of picture (GOP). But, if the channel environment is bad, the multiple packet loss can occur in one GOP.

Therefore, we should also consider multiple packet loss case. So Eq.(2) can be changed into Eq.(3)

$$MSE_{kth}^{Multiple} = \sum_{j=0}^{Nl} \sum_{i=0}^{El} (1 - \frac{i}{n}) \times MSE_{initial}(j) \quad (3)$$

where Nl is the number of loss in a GOP. Eq. (3) may be used only in one GOP. So we need an extended equation to consider total video sequence. Let Ng be the number of GOP in a video sequence. We can get Eq.(4) for the initial MSE of k -th frame's loss

$$MSE_{kth}^{Total} = \sum_{k=0}^{Ng} \sum_{j=0}^{Nl} \sum_{i=0}^{El} (1 - \frac{i}{n}) \times MSE_{initial}^k(j) \quad (4)$$

Finally, the total distortion of a whole video sequence multiple frame loss is obtained roughly as Eq. (4). But, $MSE_{initial}$ value can be only calculated when the original video sequence is available. So we need a way to estimate $MSE_{initial}$ with only received video sequence. For the MSE estimation, we investigate the motion information metrics. Each frame in a video sequence has different motion vectors to reflect its motion activity information in each block's coordinates. If this motion vector value is bigger, it increases the initial value of MSE due to packet loss. In other words, fast motion video is more vulnerable to packet loss than slow motion video. An average value of motion vectors (MA) is a key parameter to represent the whole motion vectors. This MA is calculated as follows

$$MA = \frac{\sum_{j=0}^n \sum_{i=0}^m \sqrt{(mv_i^2 + mv_j^2)}}{W \times H} \quad (5)$$

where i, j are the coordinates of block in a frame. And the notation W is the width of the video sequence and H is the height of the video sequence.

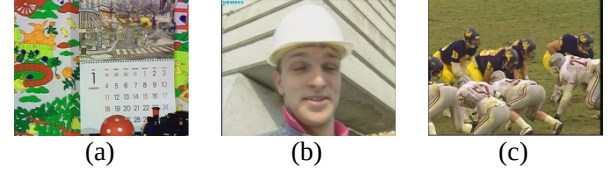


Figure 2. Test Sequences (a), (b), (c)

We use the three test sequences of Figure 2. Sequence (a) named 'Mobile' is a representative of small-motion videos, (b) 'Foreman' is for middle-motion videos, and (c) 'Soccer' is for large-motion videos. Shortly three sequences may represent various types of videos. From this assumption, next step is performed for an induction between MA and $MSE_{initial}$.

We performed several tests to find out the relationship between MA and $MSE_{initial}$. The sequence (QCIF, 176×144) (a), (b), and (c) are encoded by using H.264 encoder(JM 12.4 [9]) with 14 types of error as pattern presented in Table 1.

Table 1. Error pattern setting of Figure 2's test sequence

Error Pattern Setting							
Pattern Number	1	2	3	4	5	6	7
	8	9	10	11	12	13	14
Location of Error(P-frame)	1 st	2 nd	3 rd	4 th	5 th	6 th	7 th
	8 th	9 th	10 th	11 th	12 th	13 th	14 th

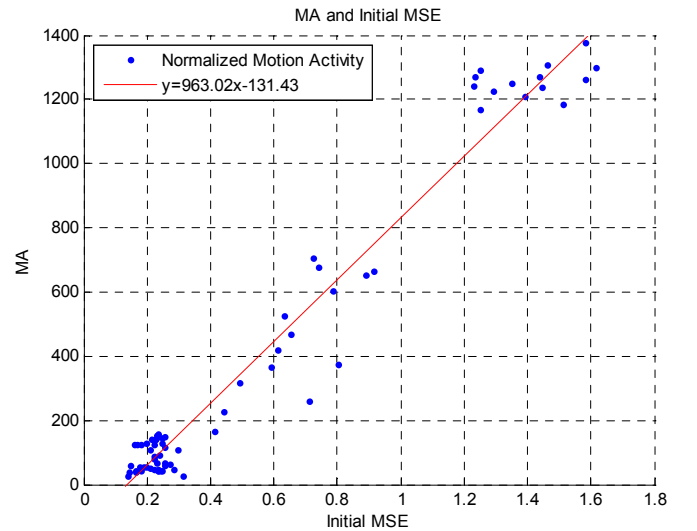


Figure 3. Relation between MA and Initial MSE

Figure 3 shows that $MSE_{initial}$ and MA are related in linear relationship. We can estimate the initial error from motion vector of video sequence. The *estimated* $MSE_{initial}$ is as follows

$$MSE_{initial} = \alpha \times MA + \beta \quad (6)$$

To find the linear parameter α and β , we conduct recursive analysis. So the linear parameter value is obtained as follow Table 2.

Table 2. Error pattern setting of experimental result

Parameter	Value
α	963.02
β	-131.43

In addition, Eq.(6) divided by the new constant to match the value of the estimated MSE and actual MSE. The new constant is γ . The value of γ is 3.98.

Therefore, the initial value of MSE is the following approximation using Table 2 and parameter γ .

$$MSE_{initial} \approx \frac{1}{3.98} (963.02 \times MA - 131.43) \quad (7)$$

Finally, we can obtain our estimation MSE by using Eq.(4) and Eq.(7). The estimated MSE represent as only motion vector, GOP size, and packet loss pattern information of received video sequence.

$$MSE_{est} = \sum_{k=0}^{Ng} \sum_{j=0}^{NI} \sum_{i=0}^{EI} (1 - \frac{i}{n}) \times (241.96 \times MA_j^k - 33.01) \quad (8)$$

III.EXPERIMENTAL RESULTS

In this paper, to evaluate performance of proposed algorithm, we compare the correlation between the actual MSE and estimated MSE value.

Table 3. Error pattern setting of experimental result

GOP Size in Sequence	Intra frame	1 (I-frame)
	Inter frame	15 (P-frame)
Number of GOP	15	
Number of Frame	240	
Error Pattern	Pattern Number	Location of Error
	1	1 st P-frame
	2	2 nd P-frame
	3	3 rd P-frame
	4	4 th P-frame
	5	5 th P-frame
	6	6 th P-frame

Single Packet Loss	7	7 th P-frame
	8	8 th P-frame
	9	9 th P-frame
	10	10 th P-frame
	11	11 th P-frame
	12	12 th P-frame
	13	13 th P-frame
	14	14 th P-frame
Multiple Packet Loss	15	1 st , 11 th P-frame
	16	3 rd , 8 th P-frame
	17	5 th , 6 th P-frame
	18	5 th , 6 th , 7 th P-frame

In order to test our proposed video quality method, we use the (a) “Crew (QCIF, 176 × 144)” and (b) “Soccer (QCIF, 176 × 144)” displayed in Figure 4 are encoded using H.264 encoder (JM 12.4 [9]) with encoding parameters, as presented in Table 3. The 18 types of packet loss are introduced by JVT loss simulator [10].



(a) (b)

Figure 4. Used sequences for evaluation

Figure 5 shows the performance comparison using test sequence (a) among actual measured MSE, MSE by using conventional method [11], and the estimated MSE by using Eq.(8). And also the case using test sequence (b) is presented in Figure 6. Pearson correlation (PC) of these value are presented in Table 4. The PC is an evaluation metric for objective quality assessment and is recommended by VQEG Hybrid / Bitstream Group test plan [8]. This PC is defined as Eq. (9).

$$PC = \frac{\sum (X_i - \bar{X}_i) \times (Y_i - \bar{Y}_i)}{\sqrt{\sum (X_i - \bar{X}_i)^2} \times \sqrt{\sum (Y_i - \bar{Y}_i)^2}} \quad (9)$$

The PC valued of 1 is in the case of an increasing linear relationship and the value of -1 is in the case of a decreasing linear relationship. If the variables are independent, the PC value is 0. In other words, the PC gives the information about correlation degree as well as correlation direciton. In general, the PC value above 0.80 are considered quite high correlation.

Table 4. Pearson Correlation

Video Sequence	Pearson Correlation (PC)
Crew	0.954
Soccer	0.983
Conventional [11]	0.732

The conventional method's PC is the correlation between the actual DMOS and PSNR [11].

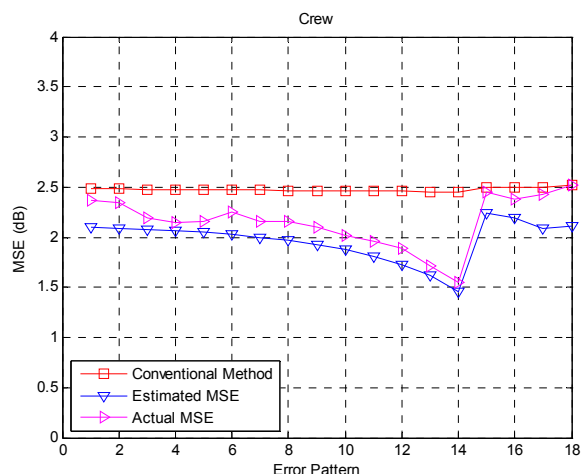


Figure 5. Proposed Estimated MSE and Actual MSE (Crew sequence)

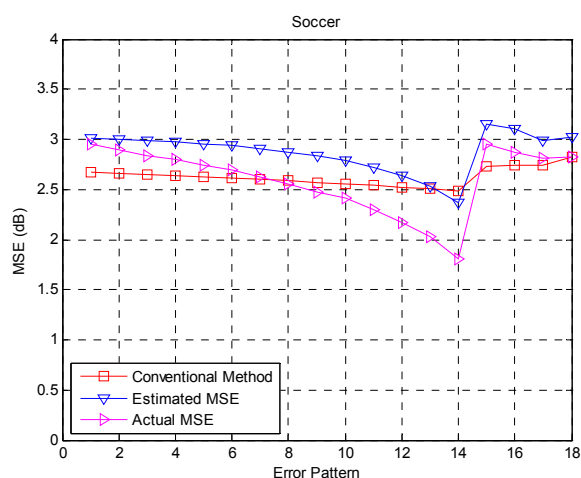


Figure 6. Proposed Estimated MSE and Actual MSE (Soccer sequence)

IV. CONCLUSIONS

In this paper, we propose an NR method for objective video-quality using motion vector considering packet loss. From the experimental result, we can obtain higher Pearson correlation than conventional method [11] in comparison with the actual MSE and its estimation value. All parameters can be extracted in decoding process. And it is a very simple and practical method. Its simplicity and easiness make this method suitable for real-time video streaming. For the future

work, we plan to combine the relationship video type and video quality by the packet loss. This metrics can more accurate measurement methods for video quality.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation (NRF) grant funded by the Korea government (MEST) (2009-0071289).

REFERENCES

- [1] Video Quality Experts Group. "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment," 2000.
- [2] Video Quality Experts Group, "Draft VQEG final report of FR-TV Phase II validation test," 2003.
- [3] B. Girod, What's wrong with mean-squared error. Digital Image and Human Vision, 1993
- [4] Zhou Wang Bovik, A.C. Evan, B.L. "Blind measurement of blocking artifacts in images", Image Processing, 2000. Proceedings. 2000 International Conference on Volume 3, pp. 10-13 Sept. 2000.
- [5] Marziliano, p.; Dufaux, F.; Winkler, S.; Ebragimi, T.; "A no-reference perceptual blur metric", Image Processing. 2002. Proceedings. 2002 International Conference on Volume 3, pp. 24-28 June 2002.
- [6] Tao Liu "Subjective Quality Evaluation of Decoded Video in the Presence of Packet Losses", ICASSP, 2007.
- [7] Yamada, Toru, Miyamoto, Yoshihiro, Serizawa, Masahiro "No-Reference Video Quality Estimation Based on Error-Concealment Effectiveness" Packet Video 2007, 2007.
- [8] VQEG "Hybrid / Bit stream Group test plan"
- [9] H.264 / AVC JM reference software home page, <http://iphome.hhi.de/suehring/tml/>
- [10] Yi Guo, Honuqiang Li "SVC/AVC loss simulator", JVT-Q.069.
- [11] Dukgu Sung, Seungseok Hong, Yohan Kim, Yonggyoo Kim, Taesung Park, Jitae Shin "No Reference Quality Assessment Over Packet Video Network" IWAIT 2009, 2009.