NO REFERENCE QUALITY ASSESSMENT OVER PACKET VIDEO NETWORK

Dukgu Sung, Seungseok Hong*, Yohan Kim*, Yonggyoo Kim**, Taesung Park**, Jitae Shin**

*School of Information and Communication Engineering
Sungkyunkwan University
Suwon, Korea

**Telecommunication R&D Center
Samsung Electronics
Suwon, Korea

E-mail: {dksung, sshong80, dos95, jtshin}@skku.edu
{taesung, yonggyoo}@samsung.com

ABSTRACT

This paper presents NR (No Reference) Quality assessment method for IPTV or mobile IPTV. Because No Reference quality assessment method does not access the original signal so it is suitable for the real-time streaming service. Our proposed method use decoding parameters, such as quantization parameter, motion vector, and packet loss as a major network parameter. To evaluate performance of the proposed algorithm, we carried out subjective test of video quality with the ITU-T P.910 ACR (Absolute Category Rating) method and obtained the mean opinion score (MOS) value for QVGA 180 video sequence coded by H.264/AVC encoder. Experimental results show the proposed quality metric has a high correlation (84%) to subjective quality.

Keywords: Video quality, No Reference, MOS,DMOS

1. INTRODUCTION

Objective video quality assessment means to compute automatically quality scores well correlated with that given by human observers. Such metrics can provide Quality of Experience (QoE) for IPTV or mobile IPTV. According to ITU-T SG12, QoE is the overall acceptability of an application or service as perceived subjectively by the end-user. Service provider can monitor of end-user's perspective QoE and can control the service parameter value for guaranteed service quality. So, objective video quality assessment is very important technology in QoE.
The objective video quality assessment metrics can divided into Full Reference (FR), Reduced Reference (RR) and No Reference (NR). The FR metric requires both the original video and the distorted video. And The RR metric require information of original video sequence and distorted video. The NR metric requires only distorted video. FR method and RR method have high correlation compared with NR method. But FR and RR is not suitable for real time streaming service. Because FR and RR need a additional data. Additional data is overhead in transmission. On the other hand, NR does not need additional data. So, we investigate more accurate NR method that is suitable for quality assessment in real time streaming service.
Our proposed method uses the decoding parameters of quantization parameter (QP), motion vector (MV) and the error-length caused by frame loss. To measure the video quality, we use the Difference Mean Opinion Score (DMOS). The DMOS score will be produced for each Processed Video Sequence (PVS) by subtracting the score from that of the hidden reference score for the Source Reference Channel or Circuit (SRC) used to produce the PVS. To make a better quality assessment algorithm we find the relationship between the video quality and related parameters. And make the general quality assessment model. And we find the optimal weight factor by linear regression analysis. Finally, we propose the No-Reference quality assessment method. To evaluate performance of proposed algorithm, we carried out subject video quality test with the Absolute Category Rating (ACR) method and obtained MOS. The testers are chosen from Sungkyunkwan University undergraduate students majoring in information and communication engineering. Totally 20 students participated in the subject test.
This paper is organized as follows. Section 2 presents the subjective video quality test. In Section 3, data analysis from subjective test is deduced and results in proposed algorithm. Section 4, shows the experiment result. and our conclusions in Section 5.

2. SUBJECTIVE TEST

1.1 Test method

To obtain MOS Score, we conduct subjective test. Our subjective test method follows the ACR in ITU-T P.910 [1]. The method specifies that after each presentation the subjects are asked to evaluate the quality of the sequence shown. The time pattern for the stimulus presentation can be illustrated by Fig.1. The voting time should be less than or equal to 10s. And the following five-level scale for rating overall quality should be used as the scores of five though one standing for.

Excellent, Good, Fair, Poor, and Bad, respectively.

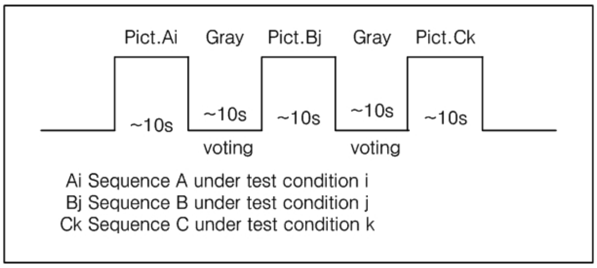

Fig. 1: The Scheme of ACR method

(a) (b) (c)

Fig. 2: Test Sequences (a) foreman, (b) football, (c) mobile

1.2 Test materials

We use the three test sequence. Foreman, Football and Mobile test sequences which are shown in Fig.2. We use the 6 original test sequences that have a length of 8 seconds a resolution of QVGA, and a frame rate of 15 / 30fps.
And we made 18 sequences having different quality coded by H.264/AVC encoder (JM 12.4 [2]) from the original 6 sequences. We use the quantization parameters of 28 / 32 / 36, H.264/AVC baseline profile, a GOP structure of 15, and an error concealment of a frame copy method. Then, we made that 10 loss patterns, using the SVC/AVC loss simulator (JVT Q.069 [3]). So we obtained 6 original test sequences and 180 distorted sequences. Total 20 undergraduate students participated in our subjective test and assessing the 6 original sequences and 180 distorted sequences.

3. DATA ANALYSYS

3.1 General Quality Assessment Model

We obtain the MOS score from the subjective test. But our method needs DMOS(Difference Mean Opinion Score). From VQEG Hybrid / Bitstream Group test plan [4]. DMOS value is calculated using the following formula.

$$DMOS = 5 + MOS(PVS) - MOS(SRC) \qquad (1)$$

PVS is the received video sequence, and SRC is the original video sequence. So, we proposed the formula of DMOS estimation.

$$DMOS_{est} = 5 - D \qquad (2)$$

In eq. (1), DMOS is processed for distorted sequence by subtracting original sequence. so, we define the D is the total distortion. The total distortion D is divided by D_e (encoding distortion), D_t(frame loss distortion), and D_a (motion activity distortion). Encoding distortion is occurred by quantization in encoding process. In quantization process, original video sequences lose the information. Frame loss distortion is occurred by packet loss in transmission. The Frame loss causes distortion not only current frame but also next frame until the end of GOP. So, we estimate the error propagation. The video sequence that has high motion activity has lower video quality. All the three parameters are distortion parameters. So, we make a DMOS estimation model as follows.

$$DMOS_{est} = 5 - D_e - D_t - D_a \qquad (3)$$

3.2 Encoding Distortion Model

The encoding distortion is occurred by quantization process. So, the quantization parameter (QP) is the key parameter of the encoding distortion. Following eq. (4) is the model of the encoding distortion.

$$D_e = w1 * QP \qquad (4)$$

The QP denotes the quantization parameter. And w1 denotes the weighting factor of the encoding distortion.

3.3 Frame Loss Distortion Model

The packet loss makes frame loss distortion. The frame loss is propagated in GOP structure. Because H.264/ AVC encoder uses inter prediction algorithm. All the P-frame (predictable frame) depends on previously frame. To measuring of error propagation, we need error propagation model. But, traditional error propagation model [5] is very complex. So, we use the error length (EL) model [6]. The error length is defined as the number of frame between the lost frame and the end of the GOP. For example, when the second frame loss, the error length is 14 in 15 GOP structure. Fig.3 shows that the error length of single frame loss.
When a multiple error occurs, we summate the each error length such as Fig.4.
Following eq. (5) is the model of the frame loss distortion

$$D_t = w2 * EL \qquad (5)$$

The frame loss distortion D_t presents EL and w2. The w2 is the frame loss weighting factor.

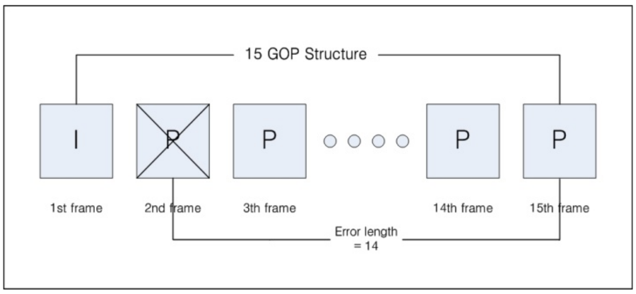

Fig. 3: The error length of the single frame loss

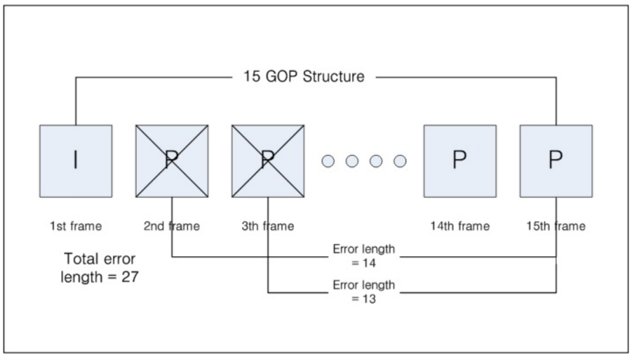

Fig. 4: The error length of the multiple frame losses

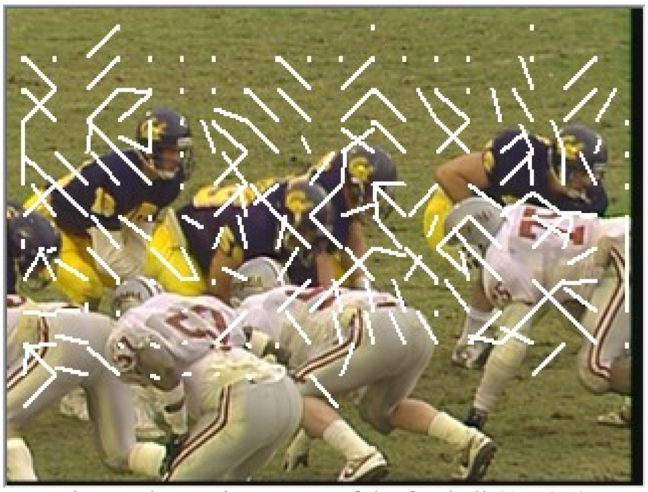
Fig. 5: The motion vector of the football (QVGA)

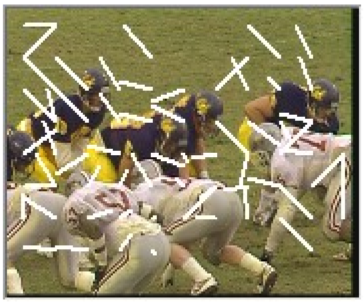
Fig. 6: the motion vector of the football (QCIF)

3.4 Motion Activity Distortion Model

We focus on the motion activities. The video sequences, have a big motion vector, are measured low quality than small motion vector video sequences. So we define the motion activity. The eq. (6) shows the motion activity

$$MA = \frac{\sum |mv|}{w * h} \tag{6}$$

The notation w is the width of the video sequence and h is the height of the video sequence. We calculated motion activity from absolute value of the 1 frame's motion vector and divided by resolution of the video sequence. Because, same video sequence which have the different resolution can have different motion vector. Figure 5 and 6 show that different motion vector. The eq. (7) shows that the motion activity distortion.

$$D_a = w3 * MA \tag{7}$$

The motion activity distortion is calculated by motion activity (MA) and weighting factor w3.

3.5 DMOS Estimation Model

Finally, we find all the distortion. We make equation (8) using the equation (4), (5), and (7) shows that the DMOS estimation equation.

$$DMOS_{est} = 5 - w1 * QP - w2 * EL - w3 * MA \tag{8}$$

Now, we have to find the weighting factor w1, w2, and w3. So, we perform the linear regression analysis.
The optimal weighting factor is 0.02, 0.05, and 0.7 for w1, w2, and w3, respectively.
So, our DMOS estimation equation is :

$$DMOS_{est} = 5 - 0.02 * QP - 0.05 * EL - 0.7 * MA \tag{9}$$

4. EXPERIMENTAL RESULTS

4.1 Experiment Method

From the subjective test, we obtained the 6 original MOS data and 180 distorted MOS data. And using the eq. (1), we obtained the 180 DMOS data. Each distortion parameter is extracted by H.264/AVC decoder (JM 12.4). And using eq. (9), we calculate the estimated DMOS values.

4.2 Pearson Correlation

To evaluate performance of proposed algorithm, we calculate the correlation between the actual DMOS and estimated DMOS value. And, calculate the correlation between the actual DMOS and PSNR. We use the Pearson correlation. The Pearson correlation is the evaluation metric for objective quality assessment. Pearson correlation recommended by VQEG Hybrid / Bitstream Group test plan [3]. Pearson correlation R is shown as the eq. (10).

$$R = \frac{\sum_{i=1}^{N} (X_i - \overline{X}) * (Y_i - \overline{Y})}{\sqrt{\sum (X_i - \overline{X})^2} * \sqrt{\sum (Y_i - \overline{Y})^2}} \tag{10}$$

Xi denotes the subjective score DMOS and Yi the estimated DMOS value. N represents the total number of video samples considered in the analysis.
The Pearson correlation is 1 in the case of an increasing linear relationship, −1 in the case of a decreasing linear relationship, If the variables are independent then the Pearson correlation is 0.

4.3 Experimental Results

Table.1 shows that the proposed method has higher correlation than PSNR. The proposed method has 0.84 and the PSNR has 0.73. So, proposed method is improved

Table. 1: The correlation of quality assessment method

	PSNR	Proposed Method
Pearson Correlation	0.732	0.842

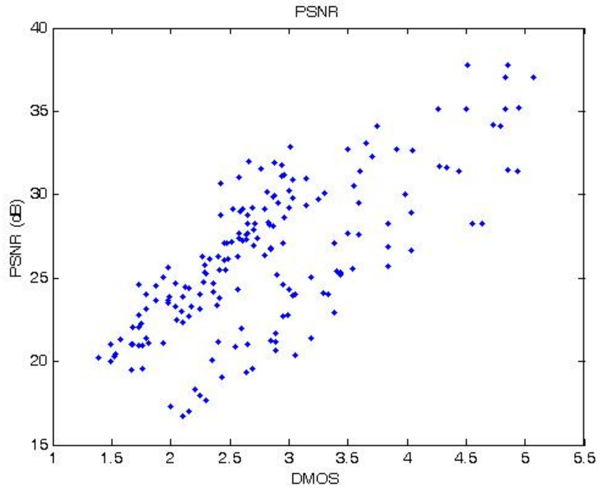

Fig. 7 : The graph of the PSNR

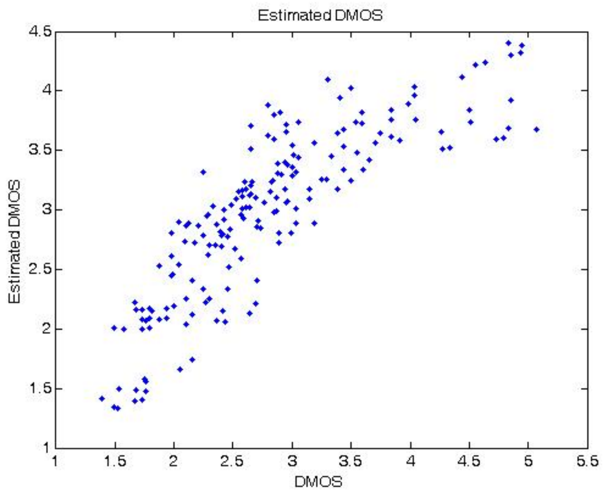

Fig. 8 : The graph of the proposed method

about 15%. Fig.6 shows that the relationship between the PSNR and actual DMOS score. All points of the Fig.7 are widely spread in the graph. This means PSNR is not suitable objective quality assessment method. But Fig.8 shows that the proposed method is more linear than PSNR method.

5. CONCLUSIONS

In this paper, we examined the quantization parameter, motion vector, and frame loss on the perceptual quality of end-user's perspective. and proposed a new objective quality assessment method. From the experimental result, proposed method's accuracy exceeds about 15% compared with PSNR method. All parameters can be extracted in decoding process. And it is very simple and practical method. So it is suitable for real-time streaming service. For example mobile IPTV or Video telephone.

6. ACKNOWLEDGMENT

This work was supported by the Korea Science and Engineering Foundation (KOSEF) grant funded by the Korea government (MEST) (R01-2006-000-10556-0).

7. REFERENCES

[1] ITU-T P.910 "Subjective Video Quality Assessment Methods For Multimedia Applications"
[2] H.264 / AVC JM reference software home page , http://iphome.hhi.de/suehring/tml/
[3] Yi Guo, Houqiang Li "SVC/AVC loss simulator", JVT-Q.069
[4] VQEG "Hybrid / Bitstream Group test plan"
[5] Y.J. Liang, J. G. Apostolopoulos, and Bernd Girod "Analysis Of Packet Loss For Compressed Video : Does burst-length matter?" ICASSP, 2003
[6] Tao Liu, Yao Wang "Subjective quality evaluation of decoded video in the presence of packet losses", IEEE, ICASSP, 2007