J Vis (2010) 13:159–168
DOI 10.1007/s12650-009-0013-6

REGULAR PAPER

Huey-Min Sun · Yan-Kai Huang

The difference between perceived video quality and objective video quality

Received: 8 July 2009 / Accepted: 5 November 2009 / Published online: 16 December 2009

Abstract We select 93 video sequences encoded/decoded by Microsoft MPEG-4 software to classify them into six different content characteristics by the cluster analysis and the discriminant analysis in this study. We compare the peak signal noise ratio (PSNR) of objective video quality evaluation with the mean opinion score of subjective quality evaluation to understand their difference by varying different bit rate. We find that the acceptable satisfaction of user perceived quality for different motion and texture characteristics is significantly different in varying the bit rates. For the low motion and simple texture characteristic (Type 1) and the low motion and complex texture characteristic (Type 4), when the bit rate is 2,000 kbps, the tolerable discarded ratio is allowed to 44% and 49%, and the PSNR is 41.7 dB and 41.0 dB, subjects could perceive the acceptable satisfaction of video quality. For the middle motion and simple texture characteristic (Type 2) and the high motion and simple texture characteristic (Type 3), when the bit rate needs around 3,000 kbps, the tolerable discarded ratio must be controlled below 12% and 34%, and the PSNR is 45.2 dB and 41.7 dB, subjects could perceive the acceptable satisfaction of video quality. For the middle motion and complex texture characteristic (Type 5) and the high motion and complex texture characteristic (Type 6), when the bit rate is 6,000 kbps, the tolerable discarded ratio is allowed to 14%, and the PSNR is 43.2 dB and 42.9 dB, subjects could perceive the acceptable satisfaction of video quality.

Keywords Content characteristic · Bit rate · Subjective measurement · Objective measurement

1 Introduction

Most researchers study the video quality by mathematical formula to measure the difference between original image and processed image. However, different video content characteristics, such as low motion, high motion, simple texture, and complex texture, have different perception by varying different bit rates. Therefore, two methods can be employed to observe the change of perceived quality. One is the objective quality assessment computed by mathematical formula to quantify visual perception. The other is the subjective quality assessment measured by the rating-scale of subjects' perceived quality. The former usually uses the peak-signal-noise ratio (PSNR) to measure an individual image quality. The latter adopts the mean opinion score (MOS) to assess perceived quality. In this paper, we use both the PSNR and the MOS to analyze the differences among diversified video contents by varying bit rates.

For a constrained bandwidth network, MPEG-4 video coding stream with fine granularity scalability (FGS) can be flexibly dropped by very fine granularity to adapt to the available network bandwidth. However, what is the tolerable dropped ratio for the subjective perceived quality? Whether the tolerable dropped ratio for the different type of video sequences is significantly different.

H.-M. Sun (✉) · Y.-K. Huang
Department of Information Management, Chang Jung Christian University, Tainan 711, Taiwan, ROC
E-mail: prince@mail.cjcu.edu.tw

In our experiment, we take 15 video DVDs to edit 93 clips as our test databases. The clips are classified into six types by a cluster analysis method for the motion and texture characteristics. Our experimental process refers to (Hands 2004). At the same time, we implement the assessment interface for the DSCQE to quickly collect the responses of the questionnaires. We find that the perceived quality of different content characteristics is significantly different by varying the bit rates. We also examine the MOSs of 42 subjects by some statistic tools to understand the influence levels on different content characteristics which are defined as six types including the low motion and simple texture (Type 1), the middle motion and simple texture (Type 2), the high motion and simple texture (Type 3), the low motion and complex texture (Type 4), the middle motion and complex texture (Type 5), and the high motion and complex texture (Type 6). The results indicate that high motion characteristic clips varying the bit rates significantly affect the video quality on both subjective quality assessment and objective quality assessment. The affecting level of complex texture characteristic for subjective quality assessment is more than that for objective quality assessment.

The paper is organized as follows. Section 2 presents the related works. Section 3 states our research architecture. We make use of both the objective assessment method and the subjective assessment method to evaluate video quality in Sects. 4 and 5. Results of the experiment are shown in Sect. 6. The conclusions are drawn in Sect. 7.

2 Related works

Video contents are very diversifying, for example, sports footage, talk show, distance learning, and news. The contents are changed that affects the video quality. Apteker et al.(1995) explore the relationship between video acceptability and frame rate by different content. They show that users perceive a reduced frame rate for a continuous-media stream differently, depending on the content. The multimedia contents are categorized by three dimensions: (1) the temporal nature of the data, (2) the importance of the auditory, and (3) visual components to understanding the message. Schaar and Radha (2001) consider both the motion and the texture of the video sequences to analyze the temporal and the signal-to-noise scalabilities in MPEG-4 FGS. Yadavalli et al. (2003) consider only the motion characteristic to classify the video contents into low, medium, and high motions. Cuetos et al. (2003) apply the evaluation framework based on MPEG-4 FGS to investigate the rate-distortion optimized streaming at different video frame aggregation levels. The video quality is related to the motion and texture characteristic. Gulliver and Ghinea (2004) show that higher frame rates, although resulting in a better perceived level of quality and enjoyment, across different video contents, do not significantly increase the level of user information assimilation. Lu et al. (2005) introduce a very important mechanism of the human brain, visual attention, for visual sensitivity and visual quality evaluation. They propose perceptual quality significance map (PQSM) based upon the analysis of color contrast, texture contrast, and motion. Wang and Li (2007) propose to incorporate a recent model of human visual speed perception and model visual perception in an information communication framework. This could estimate both the motion information content and the perceptual uncertainty in video signals. Therefore, the factors of influence on video quality are bit rate, motion characteristic, and texture characteristic.

In the objective quality measurement studies, there are simply two major approaches. One is based on error sensitivity, such as the peak signal to noise ratio (PSNR) and the mean squared error (MSE). The other is based on perceived errors which computes the structural similarity index by the luminance, contrast, and structure comparison measures (Wang et al. 2004). Some researchers try to propose new objective quality metrics. However, the acquired process of related information in coding is too complicated. Aeluri et al. (2004) combine the four parameters: motion characteristic, encoder, frame rate, and bit rate based on the MSE to assess the video quality. The ANSI National Telecommunications and Information Administration (NTIA) General Model for measuring video quality adopts seven parameters including the loss of spatial information, the shift of edges from horizontal and vertical to diagonal, the shift of edges from diagonal to horizontal and vertical, the spread of the distribution of two-dimensional color samples, the quality improvement from edge sharpening or enhancements, the interactive effects from spatial and temporal, and the impairments from the extreme chroma. The seven parameters are based on four constructs including spatial alignment, processed valid region, gain and level offset, and temporal alignment (T1.801.03 2003; Pinson and Wolf 2004). Yao et al. (2003) employ the visual quality scores based on the combination of three objective factors: visually masked error, blurring distortion, and structural distortion.

The choice of the PSNR is motivated by video quality expert group (VQEG) (Rohaly 2000), which states that none of the objective measures performs better than the computationally very simple PSNR in predicting the scores assigned by humans. The VQEG is a designed and executed test program to compare subjective video quality evaluations to the predictions of a number of proposed objective measurement methods for video quality in different bit rates. However, using the method it is hard to understand the perceived quality of service for the different content characteristics. Zink et al. (2005) show that the PSNR is not an appropriate metric for variations in layer-encoded video. They conduct a subjective assessment on variations in layer-encoded video with the goal to assess the appropriateness of existing quality metrics. The quality of perception (QoP) and the user-level QoS are presented in (Ghinea and Thomas 2005; Ito and Tasaka 2005). QoP involves not only users' satisfaction but also their ability to perceive, synthesize, and analyze multimedia information. The authors examine the relationship between application-level QoS, users' understanding and perception on multimedia clips by empirical experiment. Ghinea and Thomas (2005) find that significant reductions in frame rate and color depth do not result in a significant QoP degradation. Nam et al. (2005) present visual content adaptation techniques considering the users visual perception characteristics. They address how the visual properties of image and video content are adapted according to two types of visual accessibility characteristics: color vision deficiency and low-vision impairment. Winkler and Mohandas (2008) review the evolution of video measurement from PSNR to hybrid metrics. They deem that PSNR is completely ignorant to things as complex as the interpretation of images by human visual system (HVS). Huynh-Thu and Ghanbari (2008) present that as long as the video content and the codec type are not changed, PSNR is a valid quality measure. However, when the content is changed, correlation between subjective quality and PSNR is highly reduced. Moreover, there are a number of objective video quality assessment algorithms that have been shown to perform much better than PSNR. Sheikh and Bovik (2006) propose a visual image information (VIF) by using natural scene statistics modeling in concert with an image degradation model and an HVS model. Chandler and Hemami (2007) propose the visual signal-to-noise ratio (VSNR) to operate the contrast thresholds for detection of distortions in the presence of natural images via wavelet-based models of visual masking and visual summation. Hence PSNR cannot be a reliable method for assessing the video quality across different video contents. Wang and Bovik (2009) review the reasons why we want to love or leave the venerable PSNR. This is true despite the fact that the PSNR exhibits weak performance. Yet, the PSNR has exhibited remarkable staying power. Therefore, this is necessary that the subjective quality measurement complements the objective quality measurement. The subjective quality evaluation has been defined as the double-stimulus impairment scale (DSIS), the double-stimulus continuous quality evaluation (DSCQE), the single-stimulus (SS), stimulus-comparison (SC), the single stimulus continuous quality evaluation (SSCQE), and the simultaneous double stimulus for continuous evaluation (SDSCE) methodologies by ITU-R Recommendation BT.500 (Alpert and Evain 1997; ITU-R Recommendation 2002).

In our experiment, the adopted DSCQE is suited to compare with the objective assessment for the PSNR. This is because our material must be with both test and reference sequences. In these subjective assessment methods, the mean opinion score (MOS) is used as the user-level QoS parameter or quality of perception. To measure MOS values, the rating-scale method (Hands 2004), where an experimental subject classifies objects into some categories, is used. Each category is assigned a score to represent from bad to excellent.

3 Research architecture and video classification

According to literature review, motion characteristic, texture characteristic, and bit rate can affect the satisfaction level of perceived quality. Our research architecture shown in Fig. 1 is based on the above concepts to illustrate the relationship between three independent variables and the dependent variable.

We use Microsoft MPEG-4 software encoder/decoder with FGS functionality[1] to encode the sequences with the CIF (352 × 288 pixels) format. The group of pictures (GoP) structure is set to *IBBPBBPBBPBB*. Every sequence has 300 frames, i.e. 10 s and is processed in the *YUV* format (*Y* is the luminance component; *U* and *V* are color components of a frame). The sampling rate of the base layer is 3, i.e., one frame per three frames is encoded for the base layer. The sampling rate of the enhancement layer is 1. In the encoding data, the number of bits for the motion characteristic of I frame are 0, while the number of bits for the texture

[1] Microsoft, ISO/IEC 14496 MPEG-4 Video Reference Software. Version: Microsoft-FDAM1-2.5-040207.

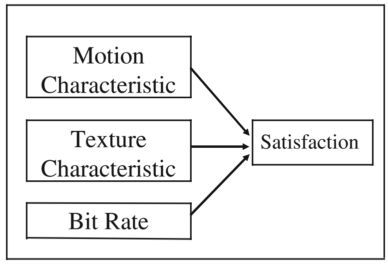

Fig. 1 Research architecture

characteristic of P and B frames are relatively very small. Therefore, we compute the average bits of all the *I* frames with the texture characteristic and the average bits of all the *P* and *B* frames with the motion characteristic for quantifying the characteristic of video content. We find out the centers of the motion and texture characteristics by the *K*-means of a cluster analysis method. Then, we adopt the classification function based on the discriminant analysis to classify all the sequences into six types with different motion and texture characteristics. In addition, we vary the bit rate to understand the influence on the different content characteristics.

In the clustering method, we employ the *K*-means to cluster all the clips into six groups. The six groups are defined as the six types of video content characteristics.

Table 1 illustrates the centers of six types. Considering the combination of motion and texture characteristics shown in Table 2, Type 1 stands for the low motion and simple texture clips, Type 2 for the middle motion and simple texture clips, Type 3 for the high motion and simple texture clips, Type 4 for the low motion and complex texture clips, Type 5 for the middle motion and complex texture clips, and Type 6 for the high motion and complex texture clips.

We find out the centers of the motion and texture characteristics by the *K*-means of a cluster analysis method. Then, we adopt the following classification function shown in the Eq. (1) based on the discriminant analysis to classify all the clips into six types.

$$d_i = \mu_i' \Sigma^{-1} x - \frac{1}{2} \mu_i' \Sigma^{-1} \mu_i + \ln p_i \tag{1}$$

where μ_i is the mean vector of the ith group, Σ is the variance–covariance matrix, and p_i is the ith group prior probability. According to the classification function, we make use of a statistical tool, called Statistica 6.0, to construct the discriminant function by the previous defined six groups as the following Eq. (2).

$$\begin{cases} d_1 = 0.0042x_1 + 0.0002x_2 - 18.6739 \\ d_2 = 0.0096x_1 + 0.0002x_2 - 37.8297 \\ d_3 = 0.0151x_1 + 0.0002x_2 - 59.0693 \\ d_4 = 0.0043x_1 + 0.0003x_2 - 43.3077 \\ d_5 = 0.0084x_1 + 0.0003x_2 - 50.1755 \\ d_6 = 0.0157x_1 + 0.0003x_2 - 82.2068 \end{cases} \tag{2}$$

where x_1 and x_2 stand for the motion and the texture characteristics, respectively. The discriminant function can test the hypothesis that the group means of a set of the independent variables x_1 and x_2 for

Table 1 The cluster analysis

Content characteristic	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
Motion (kb)	1,669	2,986	5,913	1,015	2,586	5,111
Texture (kb)	207,808	185,080	135,986	241,727	305,348	215,397

Table 2 The classification of all types by content characteristics

	Low motion	Middle motion	High motion
Simple texture	Type 1	Type 2	Type 3
Complex texture	Type 4	Type 5	Type 6

Fig. 2 The rating scale for both reference and test presentations

the six groups are equal. By averaging the discriminant scores for all the clips within a particular group, we arrive at the group average. The group average indicates the most typical location of any clip from a particular group.

We employ two video quality assessment methods to measure the effect of the video quality. One is the objective measurement, which computes the average PSNR of luminance component for assessing the spatial sharpness of an image. The PSNR is computed by the Eq. (3).

$$PSNR(f, f') = 10 \log\left(\frac{255^2}{MSE(f, f')}\right) \tag{3}$$

where f and f' are the original frame and the processed frame respectively, and $MSE(f, f')$ stands for the mean square error (MSE) between f and f'.

At the same time, we use variance to illustrate the change of frame sizes for each sequence. The variance of a sequence is stated as the following Eq. (4).

$$\frac{\sum_{i=1}^{N}(f_i - \mu)^2}{N - 1} \tag{4}$$

where N is the number of total frames, f_i is the ith frame size, and μ is the mean of all the frame sizes.

The other is the subjective measurement which performs the DSCQE test. Each trial consists of two presentations, one termed the “reference” (typically original source material) and the other termed the “test” (typically processed material). Both the source material and the processed material are identical in content. The processed material shows the material after alteration by varying the bit rate. Subjects provide quality ratings for both the reference and test presentations. Quality ratings are made using the rating scale shown in Fig. 2. The observers assess the quality twice for the reference and the test presentations, evaluating the fidelity of the video information by moving the slider of a voting function. The assessment scale is composed of 5 levels, which include bad, poor, fair, good, and excellent. We will analyze the mean opinion scores (MOSs) from the results of all the subjective assessments.

4 Objective assessment of video quality

We use the PSNR of objective quality assessment to measure the video quality by varying bit rates from 1,000 kbps to 8,000 kbps. The analysis results are shown in Table 3. We define ΔPSNR as the difference between the lowest and the largest PSNRs from bit rate 1,000 kbps to 8,000 ksps. The ΔPSNRs of Types 3, 5, and 6 are greater than that of Types 1, 2, and 4. This means that the video quality for the sequences with the high motion and complex texture characteristic has more change than that with the low motion and simple texture characteristic, especially on the bit rates from 4,000 kbps to 8,000 kbps. The required bit rate

Table 3 The PSNR analysis for six types

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
1,000 kbps	38.5	39.0	35.2	37.6	35.3	34.1
2,000 kbps	41.7	42.9	39.7	41.0	37.8	37.1
4,000 kbps	45.5	46.4	45.7	45.3	41.7	41.4
8,000 kbps	45.5	46.4	46.9	45.4	44.6	44.4
ΔPSNR	7	7.4	11.7	7.8	9.3	10.3

Table 4 The dropped ratio analysis for six types

	Type 1 (%)	Type 2 (%)	Type 3 (%)	Type 4 (%)	Type 5 (%)	Type 6 (%)
1,000 kbps	72	69	74	75	82	82
2,000 kbps	44	37	49	49	64	64
4,000 kbps	0	0	3	0	28	28
8,000 kbps	0	0	0	0	0	0

Table 5 The variance and the average frame size for six types

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
Variance (kb)	333	370	1,079	1,137	2,141	1,133
Average frame size (kb)	3,004	2,667	3,315	3,311	4,693	4,698

Table 6 The MOS analysis for six types

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
1,000 kbps	2.7	2.3	2.3	2.2	2.2	2.3
2,000 kbps	3.0	2.8	2.8	3.0	2.6	2.6
4,000 kbps	3.3	3.1	3.4	3.5	2.9	2.9
8,000 kbps	3.5	3.4	3.4	3.7	3.1	3.1
ΔMOS	0.8	1.1	1.1	1.5	0.9	0.8

without dropping any data for the sequences with high motion and complex texture characteristic needs around 8,000 kbps. When the bit rate is less than 1,000 kbps, the decoding with high dropped data ratio may result in failure. In our experiment, the available bit rates range from 1,000 kbps to 8,000 kbps for all our tested 93 sequences.

In the experiment, we discard the partial data when the bit rate is not sufficient for transmitting all the data. The simulation for each tested sequence adopts a real-time scheduling by the deadlines of all the frames. We show the dropped ratios varying the bit rate for each type in Table 4. The sequences of Types 1, 2, and 4 need only 4,000 kbps to transmit all frames, while that of Types 3, 5, and 6 need more than 4,000 kbps. This is because the sequences with the high motion and complex texture require more bandwidth resource to encode the frames. In addition, we analyze the variances, and the average frame sizes of six types shown in Table 5. The size of variance and the size of average frame are closely related to the motion and texture characteristic of sequence. The size of variance for Types 1 and 2 is far less than that of other types, while the size of average frame for Types 5 and 6 is far more than that of other types. The size of variance is small when the sequence with the low motion and simple texture characteristic, while the size of average frame is large when the sequence with the high motion and complex texture characteristic.

5 Subjective assessment of video quality

In the experimental results for subjective assessment, 42 viewers (the first test was given to 20 subjects and the second test was administered to 22 subjects 3 months after the first test) took part in the tests. Each video type is shown to 20 viewers in the first test, and 22 viewers in the second test. The ages of the test subjects are in the range of 19 to 26. Most of them are students majored in information management. They are

Table 7 The acceptable MOSs for each type

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
Acceptable bit rate (kbps)	2,000	3,333	2,667	2,000	6,000	6,000
Tolerable dropped ratio	44%	12%	34%	49%	14%	14%
PSNR	41.7	45.2	41.7	41.0	43.2	42.9

Table 8 Reliability analysis

	Type 1	Type 2	Type 3	Type 4	Type 5	Type 6
Cronbach's alpha	0.88	0.94	0.85	0.85	0.88	0.79

frequent computer and Internet users and are exposed to a variety of media contents each day. Each subject viewed the reference sequence and the test sequence and was allocated 24 combinations (6 types $\times$ 4 bit rates) for each type. Each subject spent about 15 min to complete a test. The MOS is used as the quality ratings.

The results of analysis are shown in Table 6. The highest MOS of perceived video quality has only 3.1 in the types 5 and 6. This means that the sequences with the high motion and complex texture characteristic have worse perceived quality than that with the low motion and simple texture characteristic. We compute the perceived quality difference, denoted as ΔMOS, by the difference between the lowest and the highest MOSs. The ΔMOS in the Types 2, 3, and 4 is relatively high, this implies the perceived quality for the sequence with the characteristic is relatively increasing fast when the bit rate is increasing. It means that the bandwidth resource should be poured into the sequence with the characteristic of Types 2, 3, and 4 to increase the perceived quality. However, the ΔMOS for Types 5 and 6 is relatively low, this means that high motion characteristic sequences varying the bit rates insignificantly affect the subjective quality assessment. Likewise, the ΔMOS for Type 1 is relatively low but the average MOS is relatively high, this means that the sequence with the low motion and simple texture characteristic has high perceived quality easily but it is slowly increasing.

Based on the rating scale of subjects, we define the middle level of five categories as the acceptable perceived quality score, i.e, $MOS = 3$. We compute the acceptable bit rate by using the linear interpolation method at the acceptable MOS from Table 6. Then, we compute the tolerable dropped ratio at the acceptable bit rate from Table 4. Therefore, the tolerable dropped ratio, the acceptable bit rate, and the PSNR for each type are shown in Table 7. We found that Types 1, 3, and 4 have low PSNR under the acceptable perceived quality score, and they have either simple texture or low motion characteristics. Whereas, Types 5 and 6 have high PSNR under the acceptable perceived quality score, and they have either complex texture or high motion characteristics. This implies that the sequences with either simple texture or low motion characteristics have acceptable perceived quality even if the PSNR is low.

Reliability is an assessment of the degree of consistency between multiple measurements of a variable. To reduce measurement errors, we adopt a test–retest method which measures the consistency between the responses for an individual in two points of time. The time difference between two tests is 3 months. The objective is to ensure that responses are too varied across time periods so that a measurement taken at any point of time is reliable. The first test consists of 22 subjects and the second test consists of 20 subjects. The reliability coefficient can assess the consistency of the entire scale with Cronbach's alpha. The generally agreed upon lower limit for Cronbach's alpha is 0.7 (Hair et al. 1998). The values of Cronbach's alpha in our reliability analysis is for six types. Table 8 shows that experimental reliability for each type is good.

6 Experimental results

In addition, we adopt a normalized scheme which defines the highest PSNR and MOS as 1 and the lowest PSNR and MOS as 0, respectively. This makes us understand the difference between the subjective assessment and the objective assessment. The MOSs of subjects are associated with the PSNRs varying bit rate from 1,000 kbps to 8,000 kbps for each type shown in Figs. 3, 4, 5, 6, 7, and 8.

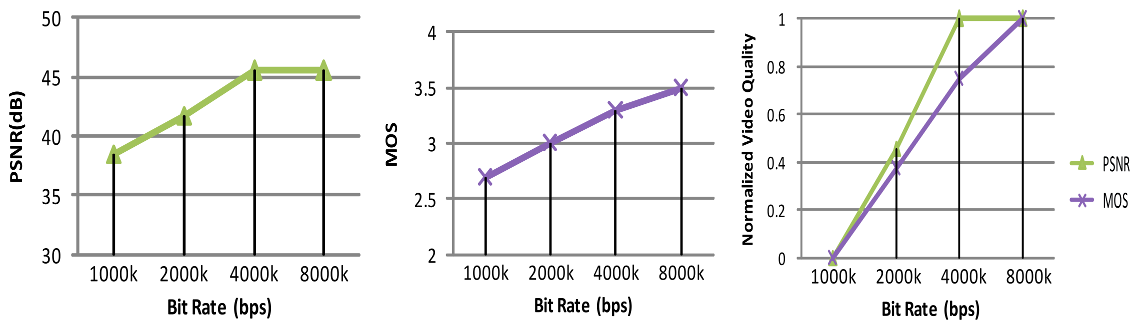

Fig. 3 The association between MOS and PSNR for Type 1

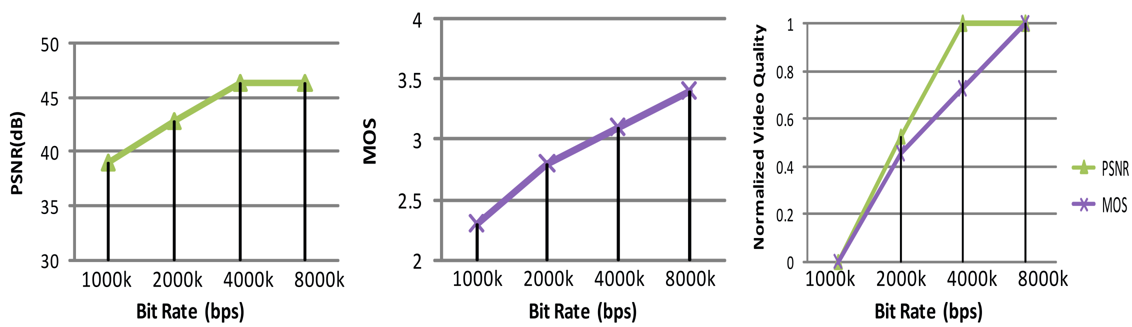

Fig. 4 The association between MOS and PSNR for Type 2

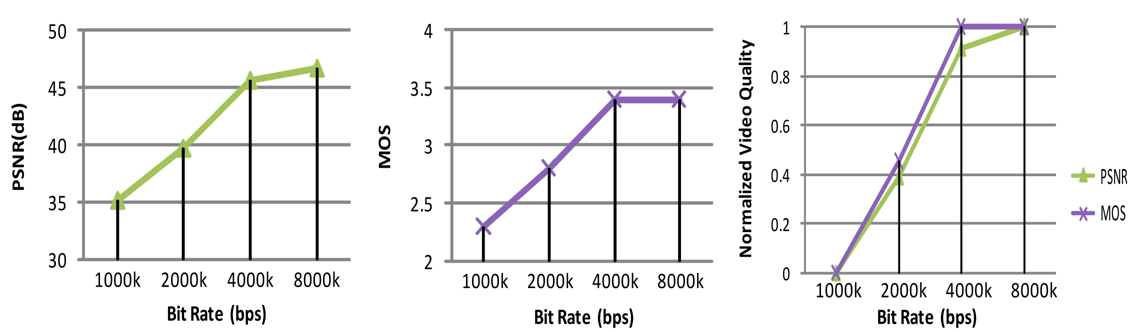

Fig. 5 The association between MOS and PSNR for Type 3

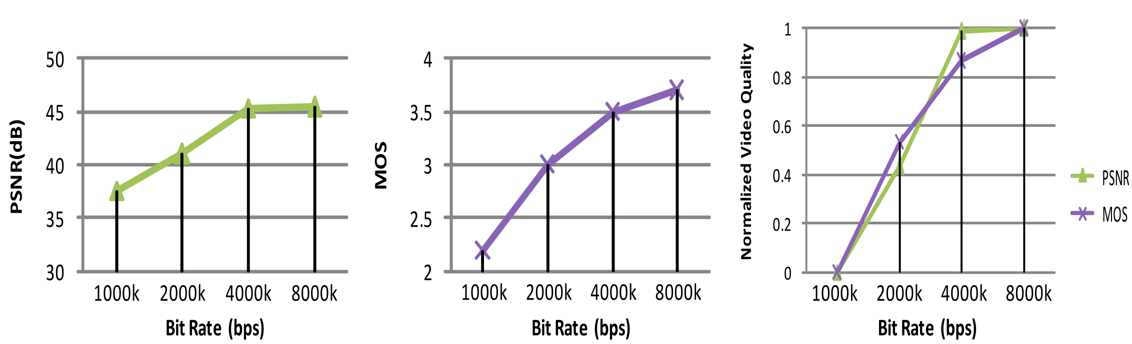

Fig. 6 The association between MOS and PSNR for Type 4

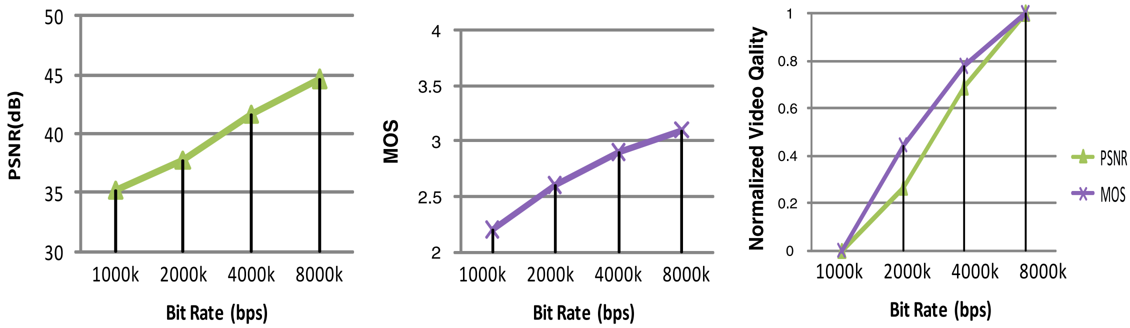

Fig. 7 The association between MOS and PSNR for Type 5

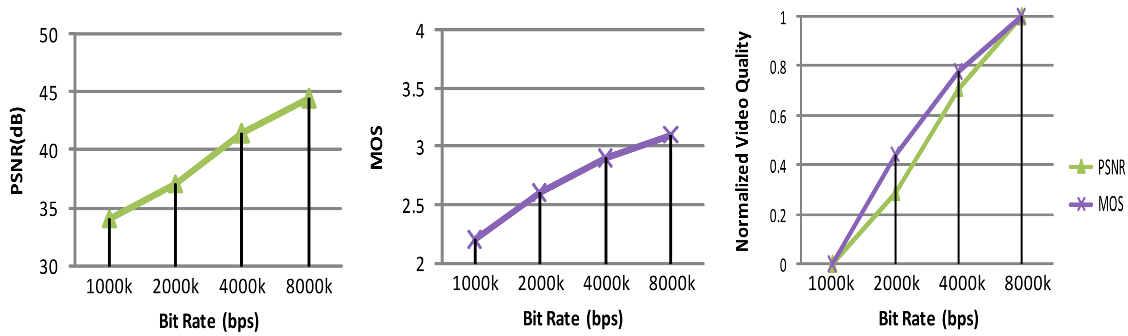

Fig. 8 The association between MOS and PSNR for Type 6

For Types 1, 2, and 4, we observe that the PSNRs are not changed from 4,000 kbps to 8,000 kbps. This is due the sequences with either low motion or simple texture characteristics need the less bandwidth resource for transmission. However, the MOSs of subjects are changed by increasing the bit rate for Types 1, 2, and 4. Figures 3 and 4 show the differences between PSNR and MOS at 4,000 kbps. We observe that the differences of normalized video quality about 0.3 (i.e., 30%) between the PSNR and the MOS at 4,000 kbps. In fact, the sequences at the bit rate 4,000 kbps and 8,000 kbps are identical for Types 1, 2, and 4, they do not discard any data. Obviously, the increasing bandwidth affects the perceived quality of the subjects. The subjects deem that more bandwidth resource has better video quality.

Figures 5 and 6 show the similar curves of normalized video quality between the PSNR and the MOS. The largest difference is less than 0.1 (i.e., 10%) at 4,000 kbps. Figure 5 shows that the MOSs are identical at 4,000 kbps and 8,000 kbps. We find that the highest MOSs with the higher motion characteristic for Types 2, 3, 5, and 6 are less than 3.5, while the highest MOSs with the low motion characteristic for Types 1 and 4 are greater than 3.5. This implies the sequences with high motion characteristic resulting in the low perceived quality even if the bit rate is greater than 4,000 kbps.

Figures 7 and 8 show the difference of normalized video quality about 0.15 (i.e., 15%) between the PSNR and the MOS at 2,000 kbps. The slope for the MOS is decreasing when the bit rate is increasing, while the slope for the PSNR is increasing. This means that the change of the subjective perceived quality is less than that of the objective video quality when the bit rate changes from 1,000 kbps to 8,000 kbps. That is the perceived video quality could not significantly increase but the PSNR could significantly increase when the bandwidth resource is poured.

7 Conclusions

According to our experimental results, different content characteristics and bit rates affect the quality of video. We have shown the difference between subjective quality assessment and objective quality assessment. All the tested clips are classified into six groups by the cluster analysis method. They are constructed by the discriminant function which can test the hypothesis that the group means of a set of independent variables motion (x_1) and texture (x_2) for six groups are equal. This function makes a new clip to be recognized easily by its content characteristics.

In addition, our experimental results indicate that the perceived acceptable quality of subjects for the content characteristics of different types is significantly different. The PSNR is no more change at the higher bit rate for the low motion or the simple texture types, such as Types 1, 2, and 4. However, the MOS is increasing at the higher bit rate for the types. We compare the MOSs of subjective quality assessment with the PSNRs of objective quality assessment and construct the relationship between both. In the analysis of comparing subjective perceived assessment with objective video quality, high motion characteristic clips varying the bit rates significantly affect the measurement of video quality. The affecting level on complex texture characteristic for the subjective quality assessment is more than that for the objective quality assessment, especially on the clips with low motion characteristic.

Acknowledgment This work was supported by the National Science Council of the R.O.C under Contract NSC 97-2221-E-309-010.

References

Aeluri PK, Bojan V, Richie S, Weeks A (2004) Objective quality analysis of MPEG-1, MPEG-2, and Windows Media video. In: The 6th IEEE southwest symposium on image analysis and interpretation, 28–30 March 2004, pp 221–225

Alpert Th, Evain J-P (1997) Subjective quality evaluation—the SSCQE and DSCQE methodologies, vol 271. ITU-R Recommendation BT.500–7, pp 12–20, Spring 1997

Apteker RT, Fisher JA, Kisimov JA, Neishlos H (1995) Video acceptability and frame rate. IEEE Multimed 2(3):32–40

Chandler DM, Hemami SS (2007) VSNR: a wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans Image Process 16(9):2284–2298

de Cuetos P, Reisslein M, Ross KW (2003) Evaluating the streaming of FGS-encoded video with rate-distortion traces. Institut Eurecom Technical Report RR-03-078, June

Ghinea G, Thomas JP (2005) Quality of perception: user quality of service in multimedia presentations. IEEE Trans Multimed 7(4):786–789

Gulliver SR, Ghinea G (2004) Changing frame rate, changing satisfaction? IEEE Int Conf Multimed Expo 1:177–180

Hair J, Anderson R, Tatham R, Black W (1998) Multivariate data analysis, 5th edn. Prentice Hall, Upper Saddle River

Hands DS (2004) A basic multimedia quality model. IEEE Trans Multimed 6(6):806–816

Huynh-Thu Q, Ghanbari M (2008) Scope of validity of PSNR in image/video quality assessment. Electron Lett 44(13):800–801

Ito Y, Tasaka S (2005) Quantitative assessment of user-level QoS and its mapping. IEEE Trans Multimed 7(3):572–584

ITU-R Recommendation BT.500-11 (2002) Methodology for the subjective assessment of the quality of television pictures. ITU, Geneva, Switzerland

Lu ZK, Lin W, Yang XK, Ong EP, Yao SS (2005) Modeling visual attention's modulatory after effects on visual sensitivity and quality evaluation. IEEE Trans Image Process 14:1928–1942

Nam J, Ro YM, Huh Y, Kim M (2005) Visual content adaptation according to user perception characteristics. IEEE Trans Multimed 7(3):435–445

Pinson MH, Wolf S (2004) A new standardized method for objectively measuring video quality. IEEE Trans Broadcast 50(3):312–322

Rohaly AM et al (2000) Video quality experts group: current results and future directions. In: Proceedings of SPIE visual communications and image processing, vol 4. Perth, Australia, June 2000, pp 742–753

Sheikh HR, Bovik AC (2006) Image information and visual quality. IEEE Trans Image Process 15(2):430–444

T1.801.03 (2003) American National Standard for Telecommunications—digital transport of one-way video signals—parameters for objective performance assessment. American National Standards Institute

van der Schaar M, Radha H (2001) Temporal-SNR rate-control for fine-granular scalability. Int Conf Image Process 2:1037–1040

Wang Z, Bovik AC (2009) Mean squared error: love it or leave it? IEEE Signal Process Mag 26(1):98–117

Wang Z, Li Q (2007) Video quality assessment using a statistical model of human visual speed perception. J Opt SocAm A 24(12):61–69

Wang Z, Lu L, Bovik AC (2004) Video quality assessment based on structural distortion measurement. Signal Process: Image Commun 19(2):1–9

Winkler S, Mohandas P (2008) The evolution of video quality measurement: from PSNR to hybrid metrics. IEEE Trans Broadcast 54(3):660–668

Yadavalli G, Masry M, Hemami SS (2003) Frame rate preferences in low bit rate video. Int Conf Image Process 1:441–444

Yao S, Lin W, Lu Z, Ong Z, Etoh M (2003) Objective quality assessment for compressed video. ISCAS '03. In: Proceedings of the 2003 international symposium on circuits and systems, vol 2. 25–28 May, pp 688–691

Zink M, Schmitt J, Steinmetz R (2005) Layer-encoded video in scalable adaptive streaming. IEEE Trans Multimed 7(1):75–84