International Journal of Advanced Smart Convergence Vol.4 No.1 18-30 (2015)
http://dx.doi.org/10.7236/IJASC.2015.4.1.18

IJASC 15-1-3

Proposed Assessment for Quality of Experience of Live IPTV in Home Environments

Jongpil Jeong[1], Jae-Young Choi[1*]

[1, 1*]*SungKyunkwan University, South Korea*
{jpjeong, jychoi1001}@skku.edu

Abstract

As the speed of networks that subscribers can use has greatly increased, demand for high-quality broadcast content, such as from Internet Protocol Television (IPTV) and Video on Demand (VoD), is likewise increasing. Therefore, while broadcasters are increasing content and channels, they are striving to improve consumer quality of experience (QoE) to differentiate themselves from competitors, including by producing higher physical-quality content. Recently, subjective measurement methods have been internationally standardized as the most reliable approach for measuring and evaluating IPTV QoE. However, a majority of these methods are performed in experimental environments and are based on the extremely brief viewing period of approximately ten seconds using original reference videos. It is actually difficult to apply standard evaluation methods based on a ten-second viewing interval to assess real broadcast watching of IPTV or other services that involve a longer time (i.e., more than thirty minutes). In this paper, we therefore propose a method that accommodates actual viewing environments. Using the mean opinion score, we experimentally analyze the effects of evaluation interval changes under actual conditions in which IPTV service is provided. In addition, we propose improvements by applying the results into actual live broadcast IPTV service and by analyzing consumer service QoE.

Keywords: *Internet Protocol Television (IPTV), Quality of Experience (QoE), Subjective Video Quality Assessment, Mean Opinion Score (MOS), Absolute Category Rating (ACR).*

1. Introduction

Supply of video services that require high quality distribution, such as Internet Protocol Television (IPTV) and Video on Demand (VoD), in IP-based network environments is greatly increasing. Accordingly, because the Internet employs a packet-based 'best effort' data transfer method, broadcasting quality of service (QoS) support and management have become very important. In particular, unlike data service, if quality is deteriorated, such as by discontinuation of audio or lagging of video, in the broadcast service—including in IPTV, in which audio and video are streamed in real time—consumers strongly react, which marks the

Manuscript Received: Jan. 10, 2015 / Revised: Feb. 5, 2015 / Accepted: Mar. 13, 2015
Corresponding Author: jychoi1001@skku.edu
Tel: +82-31-299-4607, Fax: +82-31-290-5819
Sungkyunkwan University

decrease in overall QoS. Therefore, interest in quality of experience (QoE), which represents a measure of quality that people experience, is greatly increasing [1]. QoE is affected by various factors, such as video digitalization, and encoding and decoding in the head-end system, including packet loss and delays in networks and terminals (STB, etc.). The quality of each variable affecting QoE can be measured by objective methods; however, it is not easy to objectively measure and evaluate quality without using original video at the current level. In the current method, video quality is evaluated by the mean opinion score (MOS) to quantify the QoE when consumers watch a broadcast service. Methods to objectively measure QoE for video are classified by whether or not the video is compared to the original video. These methods include full reference (FR), reduced reference (RR), and no reference (NR) approaches [7]. However, these methods have some limitations in operational or practical aspects when evaluating picture quality [8]. ITU-R BT.500 [2] and ITU-T P.910 [3] present various methods to evaluate the subjective quality of TV picture quality. In addition, many studies on effects of quality degradation factors, such as loss and delay occurring during data transmission of IPTV, have been conducted [4-6]. These evaluation methods are largely divided into relative evaluation, requiring the original video, and absolute evaluation, which does not require it. When considering evaluation of picture quality in IPTV viewing in a home environment, relative evaluation methods [2]—such as the double-stimulus continuous quality-scale method (DSCQS), double-stimulus impairment scale (DSIS), and subjective assessment method for video quality evaluation (SAMVIQ)—cannot be applied because they require the original video. However, because the absolute category rating (ACR) [3] is the standard method for video evaluation without the original video, it is appropriate to use it for assessing video QoE in home IPTV service. Nevertheless, because it is a method for evaluating picture quality within ten seconds of viewing, its application is limited in the actual environment where consumers watch IPTV in very frequent intervals. Moreover, because this method is designed to provide the score in situations where viewers do not watch the video for ten seconds immediately after viewing it for ten seconds—that is, the next video is paused—it is not suitable for environments where broadcasts are viewed in real time.

To address the above issues, we experimentally analyze the effects of change in the evaluation cycle (interval) on MOS in environments in which IPTV service is provided. In this paper, we additionally propose an evaluation method that is suitable for an actual viewing environment. Further, we apply this novel evaluation method to an actual live broadcast IPTV service to measure and analyze consumer QoE for service, and service improvement areas are derived.

The rest of this paper is organized as follows. In Section 2, the ACR standard methodology is described. In Section 3, our tests of an evaluation cycle appropriate for an actual home IPTV evaluation environment are outlined. Our measurements of the IPTV picture quality QoE level in an actual environment by using both the previously verified method and proposed tests are described in Section 4. In Section 5, our conclusions are presented

2. Existing methods of measuring video quality

The Subjective picture quality evaluation methods use guidelines to evaluate the quality of video that viewers experience according to international standards P.910, P.911, and ITU-R BT.500. Among them, ACR [3] is the standard method for evaluating picture quality of only video and without the original source. This method is used to evaluate picture quality within ten seconds of starting to watch the video. It is being used to develop models for evaluating multimedia picture quality based on VQEG and ITU. Because the video is shown only once, many videos can be evaluated in a short period of time.

One example of the ACR evaluation procedure is shown in Figure 1. Ai, Bj, and Ck—which are created by applying the conditions of i, j, and k on videos A, B, and C—are presented once within ten seconds. Evaluators assess the video quality using a five-point scale: 1 = Bad, 2 = Poor, 3 = Fair, 4 = Good, and 5 = Excellent. The assessment is performed within a ten-second interval. The length of time each video is shown can be decreased or increased depending on the content type (drama, sports, and so on). However, because the ACR is used to evaluate video quality within ten seconds, it can be a burden on evaluators. The given video is played for ten seconds and then paused (yielding a gray picture) for another ten seconds while it is being evaluated. Therefore, the video used for evaluation purposes should be separately created or manipulated to be unsuitable for real-time broadcast evaluation.

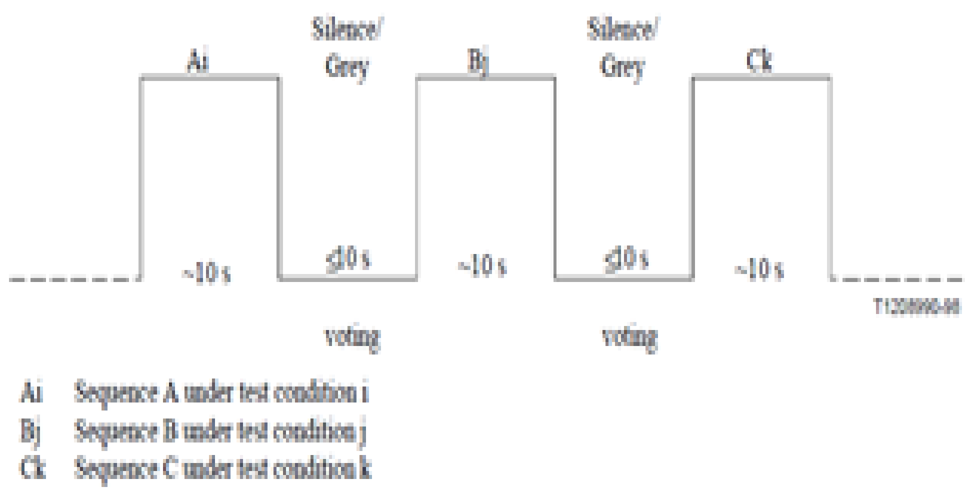


Figure 1. ACR evaluation procedures

Considering the above, an ACR evaluation method that is suitable for most evaluation approaches should be developed to evaluate QoE for videos that consumers view in real time in an actual IPTV watching environment. In addition, improvements to the existing methods, including the evaluation cycle, that are regarded as having the greatest disadvantages are needed

3. Proposed evaluation cycle for live broadcast IPTV QoE

3.1. Experimental Method for Evaluation Cycle Selection

In our study, an experimental environment was prepared, and varying content with degradation was produced, to establish evaluation conditions that are applicable to a real-time IPTV viewing environment. Experiments with ACR and to measure QoE for videos with cycles of 2, 5, 10, and 30 minutes were carried out. The experimental environments complied with ITU-R BT.500-11 recommendations. Environmental factors, such as TV types, screen brightness, viewing distance and angle, TV location height, distance between evaluator and TV, and so on, were addressed with consideration of an actual environment for watching TV. The surrounding brightness was based on the illumination of 200, and evaluators viewed the

TV with fluorescent lighting turned on. The details are shown in Table 1.

In addition, with consideration of the evaluator 30° viewing angle recommended by international standards, four evaluators were assigned to one TV. To avoid evaluation method errors, basic methods, such as evaluation scale, watching time, and elapsed time, were reviewed and rehearsed more than ten times. To represent a variety of video conditions, such as background screen, subject movement, and screen changes, three representative video content types—sports, drama, and documentary—were selected, as shown in Table 2. With consideration of a real-time viewing environment, the lengths of the documentary, drama, and sports experimental videos were 30, 60, and 90 minutes, respectively.

For the ACR measurements, all videos were separated by an interval of 10 seconds and a blank image was added to each of them for the 10 second evaluation period. On the contrary, in the IPTV experiments proposed for analysis with cycles of 2, 5, 10, and 30 minutes, the video continued to play during the evaluation without the addition of a blank image. The sound of an alarm at the end of each viewing window marked the start of its respective evaluation period.

Table 1. Experimental Environments

Classification	Content	Remarks
TV type	Liquid crystal display (LCD)	Set as default mode
TV size	32 in	Reflected based on preferred viewing distance of ITU-R BT.500-11 and size of screen
Viewing height	TV located 0.5 m above floor	
Viewing distance (between evaluator and TV)	2.7m	
Surrounding environment during viewing	Fluorescent lighting on	Surrounding brightness of evaluation environment (based on illumination of 200)
Number of evaluators per TV	Maximum 4 (viewing angle is within 30°)	Viewing angle of 30° complies with ITU-R BT. 500-11 recommendation

The measurement scale was the same five-point rating system described earlier. The results of these measurements were compared with ACR and MOS values at 2, 5, 10, and 30 minute cycles. For the comparison, the number of MOS values measured during one evaluation cycle (2, 5, 10, and 30 minutes) was 1. However, in the case of the ACR, because every 10-second interval within the cycle of each evaluation was measured, and several values were derived, it was difficult to directly compare them to each other. Therefore, in this study, the average evaluation value for 10 seconds measured within the cycle was used as the ACR result value.

Table 2. Experimental Video Specifications

Content Type	Description of Features	
	Video	Features
Sports	Soccer, Basketball, Music, Dance	□ Sudden movement frequently occurs due to movement of objects (players) on screen □ Angle of screen is constantly moving
Drama	Drama, News	□ Static background and the least angle movement □ Movement occurs in local area (gesture and face) through action or conversation (actor, anchor)
Documentary	Documentary, Landscape	□ Movement area is the least attributed to wide-angle shooting □ Occurrence of smooth movement in entire screen due to natural camera movement

However, most ACR result values consisted of values for evaluating the original video in which degradation was not applied when the average was calculated. Therefore, the deteriorated result value of the degraded video was not reflected. To address this problem, only the ACR values of a 10-second cycle in which degradation was applied were extracted, and the average of these values was used. Figure 2 shows the configuration in which the experimental video was produced and the experiment was carried out.

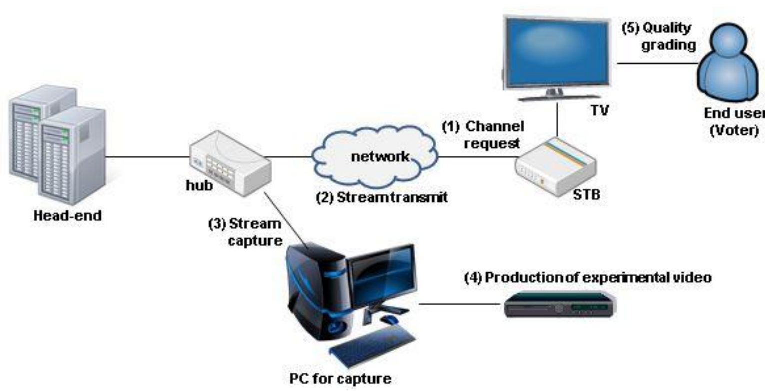


Figure 2. Experimental scenario and topology

Figure 3 illustrates this process with a 2-minute evaluation cycle. In the case of a 60 minute drama, there were 30 intervals with a cycle of 2 minutes, and a total of 12 intervals of 10 seconds during each 2 minute cycle. As an example, the following situation is outlined. Among 12 intervals of 10 seconds in the first cycle of 2 minutes, degradation occurred in three intervals: the third, sixth, and ninth 10-second intervals.

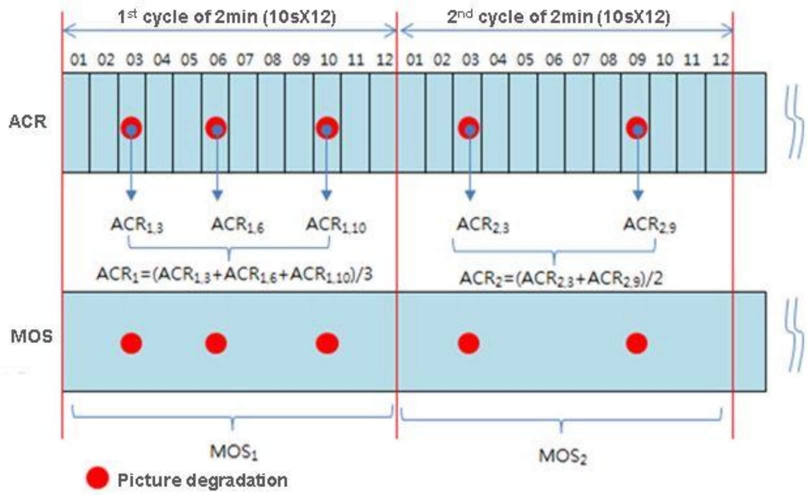


Figure 3. Application of video deterioration and ACR and MOS deterioration measurements

Degradation occurred in two 10-second intervals (the third and ninth) within the second cycle of 2 minutes. The MOS value in the first cycle of 2 minutes was MOS1, which was measured by the evaluators once after watching the video for 2 minutes. The ACR measurement value for comparison was the arithmetic average of ACR1,3, ACR1,9, and ACR1,10, which were measured by ACR in the third, sixth, and ninth 10-second intervals where deterioration occurred. Similarly, the ACR and MOS values in the second intervals of 2 minutes were similarly measured, and measurements were similarly performed in the other cycles.

3.2. Choice of Evaluation Cycle

Values measured by ACR methods were obtained according to the test methods described in the previous section. These methods measured video quality in 10-second cycles for three content types (documentary, drama and sports), and the MOS quality values were measured with video evaluation cycles of 2, 5, 10, and 30 minutes.

To obtain our experimental results, the video evaluation cycles appropriate for an actual IPTV viewing environment were selected. The averages for each content type and for entire videos were then calculated for analyzing the correlation of values measured by ACR methods and video evaluation cycles. This method was used to prove the validity of the selected evaluation cycle. The correlation between the following two variables X and Y was defined as follows:

$$corr(X,Y) = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \sigma_Y} \tag{1}$$

Here, μ_X and μ_Y represent the mean of X and Y, respectively. σ_X and σ_Y represent the standard deviation of X and Y, respectively. The results of our experiments confirmed that a variety of videos with cycles of 10 minutes showed the greatest overall correlation, as shown in Figure 4.

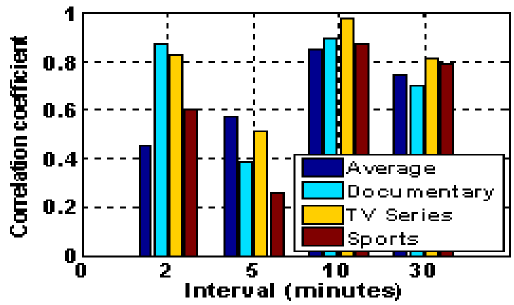
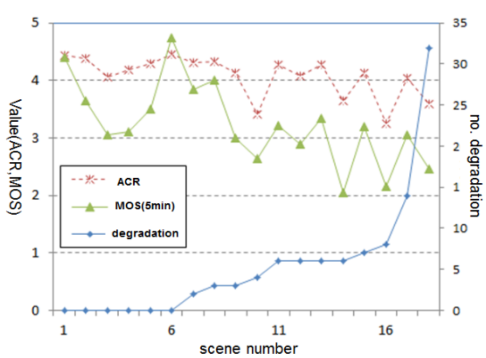


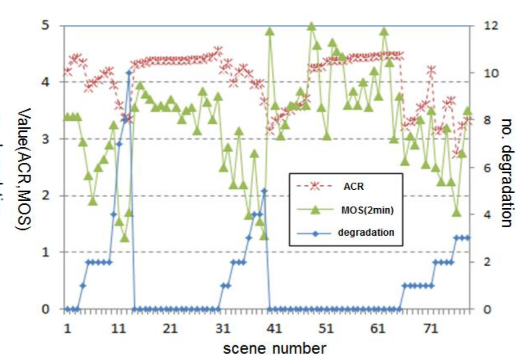
Figure 4. Comparison of correlations with ACR based on evaluation cycle

However, the correlation values based on the evaluation cycle appear somewhat differently depending on genre. Dramas showed the greatest value of the correlation between the ACR and MOS when the cycle was 10 minutes; however, this value was relatively less affected compared to other content types. Although sports videos showed a very low correlation in cycles of 2 and 5 minutes, they demonstrated the largest correlation in 10-minute cycles, and the second largest correlation in 30-minute cycles.

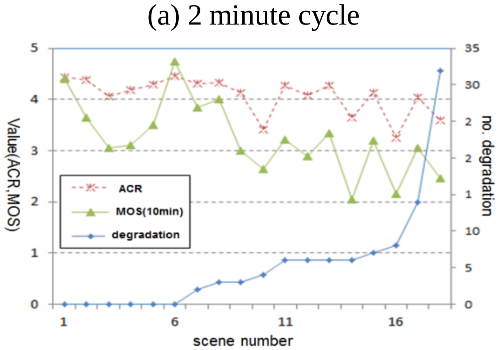
We believe that the evaluation results were greatly varied depending on the extent of evaluation cycle movement; in sports videos, the evaluation cycle became shorter because there were many variations of scenes. If the time became longer, it was less affected by such movements. Documentaries showed a large correlation in cycles of 2 and 10 minutes, and a small correlation in 5-minute cycles. This content type did not demonstrate a large correlation in cycles of 30 minutes or irregular pattern. This may be attributed to documentaries typically having complex characteristics, such as in nature videos and in videos that are artificially manufactured and combined.



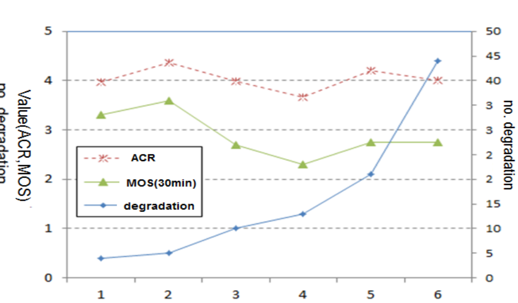
(a) 2 minute cycle



(b) 5 minute cycle



(c) 10 minute cycle



(d) 30 minute cycle

Figure 5. MOS changes by ACR values and evaluation cycles based on degree of deterioration

Moreover, in situations of less or little deterioration, the difference in ACR and MOS result patterns was not large, as shown in Figure 5. However, as the degree of degradation became greater, or the evaluation cycle became shorter, QoE values in patterns different from the results obtained in ACR were shown. Values measured in ACR on the graph, and values measured in the proposed method, did not produce the same pattern; that is, the MOS value of the proposed method was lower than the average of 0.5. These values were different because the factors used to recognize degradation of QoE in actual home environments—including differing variables in all IPTV sectors, such as head-end, network, and STB services that were provided—were more numerous than those in experimental environments of ACR, which had many experimental restrictions as a precondition. Therefore, it was shown that IPTV users in a home environment had the higher possibility of feeling uncomfortable with picture quality when viewing a live broadcast.

This experiment compared the evaluation cycle of 10 seconds, which was the standard method of ACR, with the changed evaluation cycle. It showed that the correlation between the measured MOS values was the largest in the evaluation method using a 10-minute cycle. The evaluator discomfort measurement and occurrence of experimental errors can be reduced by replacing the evaluation cycle that was found to have the greatest issues (when ACR was applied in an actual home environment) with a practical cycle that considers and applies evaluation in home environments. In addition, even if supplemented cycles were applied, the evaluation values of the entire MOS showed the same result value pattern obtained in a laboratory environment of the existing ACR. Based on this information, the high reliability of our experiment was confirmed. Therefore, an efficient and practical assessment was made in a real home environment by adhering to an actual evaluation cycle using evaluation techniques with existing ACR standards (e.g., absolute rating evaluation).

4. Measurement and analysis of live broadcast IPTV QoE

4.1. Measurement Environment and Conditions

The measurement cycle was the greatest disadvantage of the ACR, which is a representative QoE measurement method mainly used in existing laboratory environments. For our study, it was improved to be suitable for an actual home TV viewing environment. The MOS of evaluators was measured with the target IPTV service being provided as a live broadcast. As evaluators watched the IPTV broadcast content (news, drama, and music channels were selected depending on the individual) scheduled in advance in real time, their QoE of the picture was measured with MOS. Because entire sectors in which services were provided, such as head-end, network, and STB, contained original quality content at the time it was produced, the entire measurement sector represented the evaluators' QoE.

Evaluation procedures were carried out as follows. Picture QoE measurement software, with which viewers can rate MOS values in real time on their portable notebook or smartphone, was installed. The evaluation was performed according to instructions of the measurement program. If evaluators clicked the installed measurement software prior to evaluation, the popup window where their basic information (gender, region, channel, and so on) is entered was opened and the alarm signaled the start of the evaluation. If users began the evaluation, the measurement software marked the evaluation time as a cycle of 10 minutes. Users could rate MOS values for their experience level about videos they watched within 20 seconds. Total evaluation time was 30 to 60 minutes based on the time needed to play the TV content. In addition, to increase the accuracy of evaluation results, users could not change the channel after the evaluation began.

Once they gave their ratings with MOS values, the result would be transmitted in real time to a remote collection server. Measurement values which did not comply with the above procedures were excluded from the final data, which increased the accuracy of measurement results.

The evaluation was conducted with consumers of IPTV service from three major providers in South Korea. Evaluation content (channels) were divided by diverse content characteristics into drama, documentary (news), and music broadcasts, which were fully played to verify the quality level of the entire service. In addition, the number of population samples with the proper size was ensured by applying generalized statistical sampling methods to verify the statistical reliability of the measurement results. In our study, the probability sampling method was selected because it was generally utilized. In addition, it had the advantage of high representativeness if more than a certain sample size was secured. Moreover, to estimate the proper number of size necessary for evaluation, the confidence level and margin of error were set in advance. We applied the confidence of 95% and margin of error of ±3%, which was superior to the confidence of 95% and margin of error of ±5% that is typically used. Finally, determined sample size was classified into population mean estimation and ratio allocation methods of probability sampling. However, population mean estimation was used for the IPTV service quality because information on mean and standard deviation of population was ensured through prior investigation. Therefore, the number of samples necessary for the actual test was calculated by using the mean and standard deviation, which was the result of a pre-test with 30 subscribers and the following (2). The result was approximately 2,000 samples.

$$sample(n) = \frac{N(z\sigma)^2}{Nd^2+(z\sigma)^2} \approx \frac{(z\sigma)^2}{d^2} \tag{2}$$

In the above equation, N, $z(=z_{\alpha/2})$, σ and d represented the population size, the level of confidence, the standard deviation, and the margin of error, respectively. $z(=z_{\alpha/2})$ applied the constant of 1.96 for a confidence level of 95%.

4.2. Measurement Results and Analysis

An actual measurement was carried out as a sample survey over three months. The survey targeted subscribers of nationally distributed live broadcast IPTV. The samples included a higher proportion of subscribers in urban areas with consideration of service subscription rate. A total of 2,021 subscribers were selected and three content types (drama, news, and music channels) were shown to each person. A total of 5,951 types of contents were evaluated. Measurements of subscriber voting were aligned to QoE according to service provider and region. The measurements were analyzed to verify the practicality of the proposed method. Detailed standards were proposed to utilize the results of analysis as valuable data for future IPTV service performance improvements.

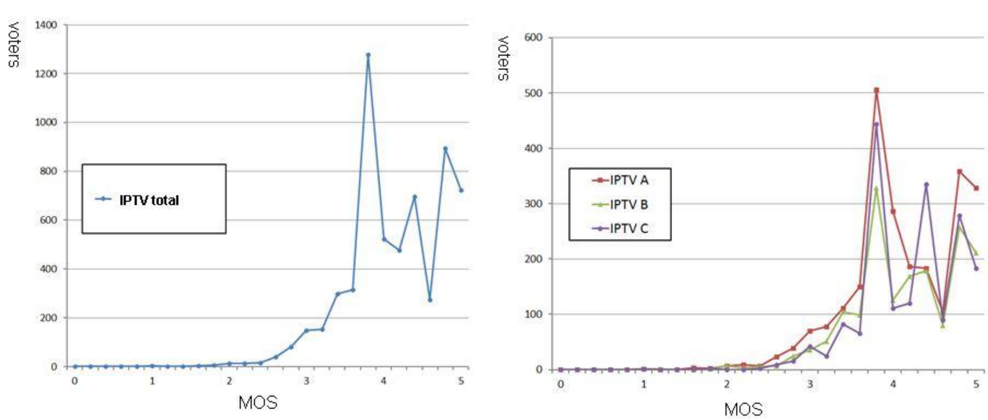
Table 3 shows the subscriber-rated level of QoE for live broadcast IPTV service. As a result, the MOS value for the full-service content was an average of 4.1, which represented “good” (good = 4.0 or higher, as determined by the international standards). As mentioned previously, international standards BT.500 of ITU-R, P.910 of ITU-T, and others have classified into five levels the user quality experience when viewing videos. Based on the quality level, ACRs were evaluated as: MOS of 1 is “bad” (very annoying), MOS of 2 is “poor” (annoying), MOS of 3 is “fair” (slightly annoying), MOS of 4 is “good” (perceptible but not annoying), and MOS of 5 is “excellent” (imperceptible). When the result of this evaluation was interpreted by applying the above international standards, a mean of MOS of 4.1 implied that “Most users were satisfied with quality when they watched live broadcast IPTV service in a home environment. In addition, some rarely found the degradation in videos that they were watching, but they did not feel uncomfortable watching TV."

Table 3. Status and Results of QoE for IPTV Service Measurement

Service Provider	No. House-hold	No. Contents	MOS Mean	Standard Deviation	Ratio of MOS 4.0 or Lower
IPTV A	832	2,453	4.1	0.7	41.0%
IPTV B	576	1,691	4.1	0.6	39.6%
IPTV C	613	1,807	4.2	0.6	38.1%
Total	2,021	5,951	4.1	0.6	39.0%

However, when results of detailed measurements were compared with numbers of the entire measurement, the numbers of MOS values of 4.0 or lower accounted for 39%, which was a very large proportion. As shown in Figure 6(a), MOS values were widely distributed from 1 to 5, and many of them were distributed between 3.5 and 4.0. In addition, it was found that the number of samples of MOS of 3.0 or lower, which represented significant discomfort on quality, was not small. This result showed that users felt uncomfortable watching live broadcast IPTV service at the level of 40% or less, or some felt greatly dissatisfied about participating in the entire quality measurement.

When examining the results in more detail, current IPTV service typically adopted international standards for infrastructure and service configuration. However, the differences in quality level may appear depending on differences in technical and support policy, such as network configuration and QoS application methods by service providers. Therefore, the level of quality of each service from different providers was analyzed, and the result showed a slight difference, as depicted in Figure 6(b). As shown in Table 3, the numbers of household members and content items measured in each service were somewhat different; however, they were accepted for statistical comparison. The average MOS values showed a slight difference with 4.0 and 4.1. In addition, there was a slight difference in proportion of MOS values of 4.0 or lower, such as 39.1% and 41%, which confirmed that the level of quality was similar in each service.



(a) Total service (b) Services by communications provider

Figure 6. MOS distribution diagram

4.3. Discussion

As noted above, it was confirmed that measurement methods for user QoE with 10-minute cycles proposed in this paper had a significant advantage for measuring and analyzing details of QoE levels for live broadcasting of video content in an actual home or office TV viewing environment. The proposed methods overcame the largest disadvantage of picture quality evaluation within a laboratory environment, by which current measurement methods, including ACR, have been affected. In addition, the proposed methods were tested and applied in an actual environment. The results indicated that the level of user QoE for live broadcast IPTV service was overall satisfactory according to quality levels recommended by the international standards. However, if measured and analyzed in more detail by this evaluation approach, the proportion of subscribers who may potentially have complaints or doubts about quality accounted for approximately 40%, which suggested that improvements are required. In addition, there were some regional deviations in subscribers' perceived quality level. As a result, it was found that investment and efforts by service providers are required to provide excellent IPTV service, which would be distributed more stably and evenly in the future.

As shown in the measurement results, unlike factors such as delay and packet loss that are mechanically measured and have proportionally affected QoE, the quality that subscribers perceived was satisfactory. Although they experienced the quality at the same level, these subscribers might perceive quality differently based on personal preferences and standards. Therefore, we recommended consideration of setting in advance the compliance rate (%) into the standards set and managing it. We suggested this approach as a means to measure and manage the stability of service with standards to classify existing user QoE levels into the five MOS levels. Some international standards (Y.1541, TS120-6, and others) have even recommended compliance with the recommended standards by applying the average. The average should be compiled in advance when managing network packet loss and failure, and the standards at the level of 95 to 99% should be compared to the full service. For example, this paper recommends that we should be able to consider the step-by-step application of the average MOS 4.0 and satisfactory ratio of service standard (MOS 4.0) compared to the number of samples up to 90~100% on the basis of existing international standards and results of measurement and analysis, as shown in Table 4. MOS 4.0 complied with the "satisfactory" level specified in the international standards, including ACR. Some countries, including Switzerland, have applied the standards of MOS 4.0 for distributing IPTV service. In addition, the ratio to satisfy the service standard in 95% was referred to recommendations applied in the international standards. Many countries, including the US and Japan, have applied the success rate of 95% as the standard of good performance when the number was assigned to specific services, such as VoIP.

Table 4. Management Standard for IPTV User QoE

Measurement Item	Service Standard	Description
MOS	4.0 or higher	Level of user satisfaction based on the International Standard
Service Standard Satisfactory Ratio	90%~99%	Ratio to satisfy the service standard compared to entire users. More than 95% compared to number of entire quality measurements

If the live broadcast IPTV services currently measured by applying the above standards were analyzed, it would meet the overall average MOS 4.0. Therefore, it seemed sufficient to provide the service. However, because there was a high proportion of MOS 4.0 or lower, the clear standards could be set in which the standard satisfaction of service was improved up to 90%~100%, step-by-step, in terms of improvements on user service satisfaction and vitalization of service. Likewise, IPTV service providers simultaneously manage both entire MOS values and the proportion of user dissatisfaction compared to the entire MOS value. Therefore, the goal of service quality improvements can be quantitatively set and satisfaction of users can be increased.

5. Conclusion

To compensate for the evaluation cycle that is regarded as the drawback of the ACR standard, an evaluation cycle applicable to an actual evaluation environment was derived and verified. To this end, broadcast content was classified into three types—documentary, drama, and sports—and the correlation between the ACR of the evaluation method with a 10-second cycle and MOS values measured with varying evaluation cycles was obtained. The findings showed that the correlation with the measurement result from the ACR method, which was standard in its cycle of 10 minutes, was the highest. This result implied that the result had the same reliability as measurements with a 10-second cycle, even if it was completed with a 10-minute cycle and users comfortably measured it. Finally, it was found that the drawback of the ACR could be complemented and utilized to measure QoE for video in real time.

In addition, the ACR standard method that complemented the evaluation cycle was applied, and the QoE of the picture of commercial IPTV service was measured. The result of the analysis showed that continuous efforts to improve service quality of providers were required because the overall quality level was “good;” however, the proportion of users who felt or might feel uncomfortable in quality was greater. Therefore, in this paper we proposed a management standard to satisfy 95% or higher of service viewers, as long as the entire quality level satisfied MOS 4.0 as the standard of reference upon proving and managing IPTV service by providers.

With the dissemination of IPTV service, the interest in QoE representing a scale of quality perceived by consumers is growing. However, an appropriate method to measure video QoE in an IPTV environment has not yet been presented. Results of this study are expected to be used as important data with regard to selection of evaluation cycles and quality improvement among factors for measuring QoE for various broadcast services. In addition, further studies will be needed on evaluation cycles that utilize more content types, and on verification of accuracy of results, for applying them as more reliable evaluation methods.

ACKNOWLEDGEMENTS

This research was supported in part by contribution of Korea Communications Commission and Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (No.2010-0020737).

References

[1] A. Takahashi, D. Hands, V. Barriac, "Standardization Activities in the ITU for a QoE Assessment of IPTV," *IEEE Communications Magazine*, Vo. 46, No. 2, pp.78-84, 2008.

[2] ITU-R Rec. BT.500-11, Methodology for the Subjective Assessment of the Quality of Television Pictures, *ITU-R*, 2002.

[3] ITU-T Rec. P.910, Subjective video quality assessment methods for multimedia applications, *ITU-T*, 1999

[4] R. Shumeli, O. Hadar, R. Huber, M. Maltz, M. Huber, "Effects of an Encoding Scheme on Perceived Video Quality Transmitted Over Lossy Internet Protocol Networks," *IEEE Tr. Broadcasting*, Vol. 54, No. 3, pp. 628-640, 2008.

[5] H. Kim, S. Choi, "A Study on a QoS/QoE Correlation Model for QoE Evaluation on IPTV Service," *IEEE ICACT'2010*, 2010

[6] Q. Dai, R. Lehnert, "Impact of Packet Loss on the Perceived Video Quality," *IEEE INTERNET'2010*, 2010

[7] ITU-T J.143 User Requirements for Objective Perceptual Video Quality Measurements in Digital Cable Television Series J: Transmission of Television, Sound Programme and Other Multimedia Signals Measurement of the Quality of Service, 2000

[8] K. Yamagishi, T. Hayashi, "Parametric Packet-Layer Model for Monitoring Video Quality of IPTV Services," *IEEE ICC'2008*, 2008