The Quality Evaluation and QoS Requirement Analysis for MPEG-2 Video Service over the Internet

Jae-il Jung[1] and Yu-chan Kang[2]

[1] Division of Electrical and Computer Engineering Hanyang University 17, Haengdang-dong, Seongdong-gu, Seoul, 133-791, Korea
jijung@hanyang.ac.kr
[2] Personal Telecom, Seoul, Korea
yck73@perstel.co.kr

Abstract. The Next Generation Internet will provide large bandwidth and guarantee end-to-end user's Quality of Service (QoS) for high quality video service. In order to guarantee QoS for multimedia streaming services over the Internet, we analyze the effects of packet loss, delay, and delay variation on video quality and how user-perceived quality is related to the network performance parameters for MPEG-2 video service. We address the problem of QoS parameter translation between network level quality and user-oriented quality, and compute network performance objectives for guaranteeing video quality perceived by an end user.

1 Introduction

Many works considered the video QoS related issues. Some papers addressed the modeling of packet loss and delay on video quality with respect to MPEG system layer [1][2]. And others papers focused on network aspects, considering the buffering scheme, the relationship between layer's performance parameters, and the distributed systems [3][4][5]. To our knowledge, not much has been done in overall viewpoint on video QoS which includes the user opinion, application, and network.

We define the QoS translation as a mapping between network level QoS and user level QoS. The QoS translation process can be viewed as a mapping between objective QoS values and application specific subjective QoS values. The QoS evaluation system having this mapping process is divided into three parts: network level QoS, application level QoS, and user level QoS. The network level QoS monitors network performance and extracts network performance parameters: packet loss, delay, and delay variation (or jitter). These parameters reflect current network performance and cause application QoS degradation. The application level QoS analyzes the impact of network performance on a specific application service and the impairment factors to affect application quality. The application performance can be influenced by many factors, such as computing resources, but it is especially sensitive to network performance. Therefore, it is necessary to clarify how the application quality is related to network performance parameters and to extract a quantity of degraded quality from the application viewpoint. The user level QoS collects the end user's subjective opinions for the service and is evaluated by the statistics of opinion polls.

We evaluate the user-perceived QoS through the mapping process consisting of the above three QoS parts. Ultimately, we can apply the quality evaluation results to the network and service management for guaranteeing multimedia QoS.

In this paper, we propose a QoS evaluation methodology for MPEG-2 video service over the Internet in Section 2. First of all, it is essential to translate QoS values between network level QoS and application level QoS [4]. This method analyzes the effects of packet loss, delay and delay variation on video quality and employs the quantitative analysis discussed in Section 3. Therefore, the application level QoS is represented by quantitative QoS values, e.g., the number of impaired macroblocks in video data and the DIQ (Degraded Image Quality) [2]. We introduce an experimental system implementing quantitative analysis in Section 4. In Section 5, we analyze how the user-perceived quality is related to application level QoS for MPEG-2 video service. The subjective QoS assessment of the end user is accomplished by MOS (Mean Opinion Score) of ITU-R.BT 500-7 [6].

2 Models for Video Quality

The MOS (Mean Opinion Score) [6] is used for the perceptual user quality evaluation in Table 1, which is a general method used by many researchers when evaluating subjective quality. A 5-point quality grade is recommended for assessing video quality based on users' opinions. The quality tests typically require the viewer to watch a video sequence. The video sequence should be long enough for the viewer to experience the types of degradations common to video service.

The quality evaluation of the networked video service is intended to be made entirely on the basis of the picture quality because users are very sensitive to picture quality. The types of the impaired video quality are presented in Table 2.

There is a necessity for reliable and valid methods to measure subjective video qualities in the applications developed, and link them to the objective QoS factors that can be applied to network services.

2.1 Video Quality Impairment

Two bad things can happen to a packet delivered over the Internet: it may be lost, or it may be severely delayed. In reality, a delay longer than a certain length is considered to be a loss because video packets received after a long delay become useless. Delivery of video packets in real-time differs fundamentally from delivery of data packets in that video packets are time sensitive, and there is usually not enough time for redelivering lost video packets.

Table 1. ITU-R BT. 500-7 MOS (Mean Opinion Score)

MOS Grade	Service Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Table 2. Types of degraded video quality

Format	Status
Tiling & Pixelation	Formation of small blocks with distinct boundaries
Line blanking	Loss of video (black line)
Motion jerkiness	Irregular or unnatural motion observed
Frame freezing	Screen freezes
Error blocks	Small solid-color blocks appear (usually yellow or green)
Color cycling	Color stability lost

Two components of IP packet delay are the propagation delay and the queueing delay on the routers. The basis for designing common Internet "streaming" media applications is the intelligent use of receiver side (playout) buffering. The goal of receiver side buffering is to reduce packet losses due to delay by using the minimum possible buffer size. A too large playout buffer absorbs the delay jitter, but results in a longer start-up delay. Clearly, interactive applications are more constrained by the playout buffer size because their tolerance to delay jitter is much lower.

There are two key components in playout buffer size management. The first one is to determine an initial minimum buffer size (i.e., an initial waiting time). If the initial buffer is too small, the playback will have to pause; if it is too large, there will be an unnecessarily long initial wait. Since the delay jitter is variable, the buffer size should be able to adapt to its variations. The second component is to change the buffer size in real time with minimal effect on the playback quality.

2.2 Acceptable Delay Estimation

If every packet is delayed in a buffer such that it suffers cumulatively (in the network and buffer), a delay equal to the maximum network delay, the receiver will reproduce a jitter-free playback. The main problem of this approach is that it may be difficult to obtain an accurate estimate of maximum network delay over a stream lifetime, and such a value may be quite large in relation to the average delay [7][8].

The basic idea behind the perfect buffer size is to let each packet experience the same total amount of delay, which is equal to the maximum delay, from the retrieval time at the server side to the presentation time at the client side. Packets experiencing longer network delays stay in buffer for a shorter time, while those experiencing shorter delays stay in buffer for a longer time. Packets which have the maximum network delay (D_{max}) will be directly sent out from the buffer without staying. Packets which have the minimum network delay (D_{min}) will stay in the buffer for the longest time, which is equal to D_{max} - D_{min}.

For arbitrary packet i, $nd_i + bd_i$ is a constant TED, where TED is the total end-to-end delay, nd_i is the network delay suffered by i_{th} packet, and bd_i is the induced buffer delay which is a function of the buffer size [7][8].

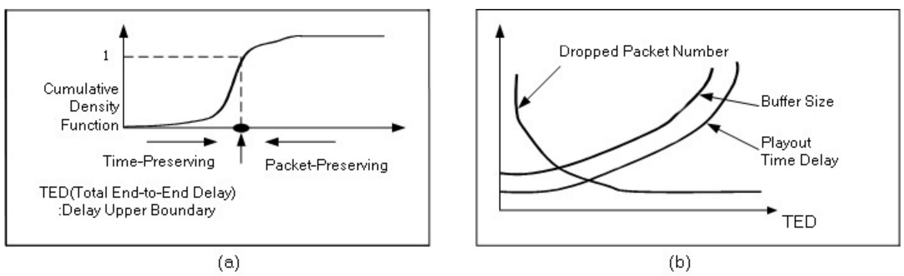

Fig. 1. Acceptable delay estimation for multimedia stream service: (a) a TED required to smooth network delay variation; (b) the buffer size must be adjusted to avoid lateness

In Fig.1, it is shown that a streaming application is bounded by an acceptable delay. In Fig.1(a), a networked application has a real-time characteristic by choosing an acceptable delay and this delay becomes a total end-to-end delay (TED) as an upper boundary of the total packet delay. Fig.1(b) shows a trade-off between the playout delay and the number of the dropped packets in the decoder. If the maximum network delay (D_{max}) and the minimum delay (D_{min}) are known in advance, then perfect synchroniza-tion can be achieved if the buffer size at the receiver side is configured as follows:

$$B(buffer\ size) = (D_{max} - D_{min})/R, \tag{1}$$

where R is the playout time interval [7][8].

3 Video Quality Evaluation

3.1 Quantitative Analysis

The number of impaired macroblocks is used as an objective QoS value. The calculation of impaired macroblocks is the mapping process from network level

QoS to MPEG-2 data level (application level QoS). In this paper, we assume that the size of the MPEG-2 TS (Transport Stream) packets is 188bytes [9].

Most video coding algorithms employ interframe and/or intraframe prediction, where decoding a picture may require availability of the previous pictures or previously transmitted parts of the current picture. For this reason, the impaired pictures affect the other pictures. For example, the impaired I picture affects all P pictures, and consecutively the impaired P picture affects the corresponding B pictures. In the case of a header fault of sequence, GOP or slice, the total data of the payload part below the header cannot be decoded [9]. Consequently, the number of the impaired macroblocks in the pictures should be examined. The total number of impaired macroblocks is defined as DIQ (Degraded Image Quality) [2].

$$DIQ = (MB(total\_loss)/MB(total)) \times 100\% \tag{2}$$

where $MB(total\_loss)$ is the total number of macroblacks destroyed due to packet impairments and $MB(total)$ is the total number of macroblocks.

We compute DIQ by the real-time network performance monitoring, which is explained in next section. The impaired macroblocks are analyzed by the received TS packets. First of all, we look at the impaired ones in the slice parts using TS information. The structure of macroblocks is characterized by the fixed number of macroblocks per picture, without regard to the type of picture. The macroblocks following the impaired macroblock are destroyed because of the absence of the header in the slice.

In the case of I picture impairment,

$$MB(P, propagation) = MB(I, lost) \times NUM(P), \tag{3}$$

$$MB(B, propagation) = 2 \times MB(I, lost) \times (NUM(P) + 1), \tag{4}$$

where $MB(P, propagation)$ and $MB(B, propagation)$ are the number of impaired macroblocks in P and B picture for the error propagation of the video reference coding, respectively. $MB(I, lost)$ is the number of impaired macroblocks in I picture. $NUM(p)$ is the number of P pictures in a GOP.

In the case of P picture impairment,

$$MB(P, propagation) = MB(P, lost) \times (NUM(P) - k), \tag{5}$$

$$MB(B, propagation) = 2 \times MB(P, lost) \times (NUM(P) - k + 2), \tag{6}$$

where k is the order of P pictures and is assumed to be ranged from 1 to 4. The total impairment is calculated as follows:

$$MB(total\_lost) = MB(lost) + MB(propagation), \tag{7}$$

where $MB(total\_lost)$ is the total number of impaired macroblocks.

The total number of impaired macroblocks is obtained in accordance with the position of the impaired macroblock [2]. This includes the effect of late packets due to network delay in the decoder side as well as lost packets in the network.

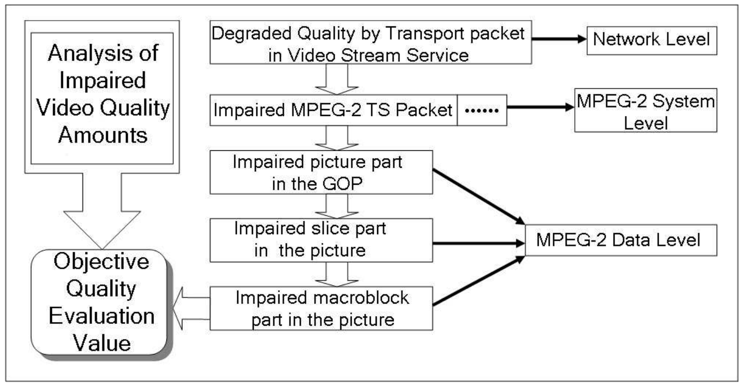

Fig. 2. Quantitative analysis for video quality impairment

4 Experimental System

In Fig.3, the outline of QoS evaluation system is presented. The server and client for MPEG-2 video service are implemented. The network parameter monitoring module is implemented both at the server and client. The video quality evaluation module is implemented only at the client. The total system is realized in software and constructed in PCs (Pentium-3 and MPEG-2 decoder card in client, and Windows 2000 in the server).

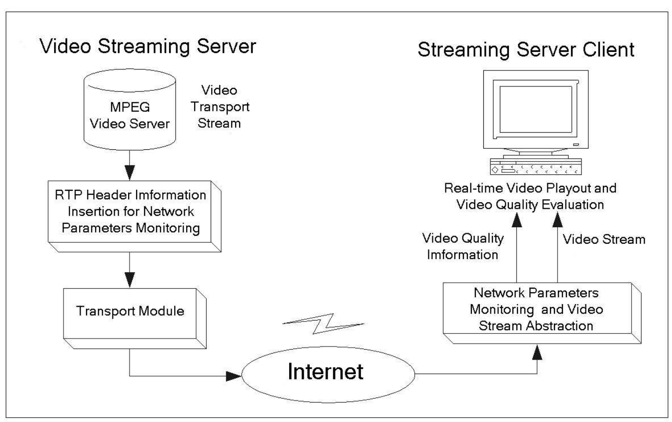

Fig. 3. Video QoS evaluation System

The server inserts RTP headers [10] for network parameter monitoring and transports TS packets to the client. The client consists of three modules: network parameter monitoring module, video playout module, and video quality evaluation module. The network parameter monitoring module measures packet loss, network delay and delay variations. It monitors the RTP header information of received packets. We use MPEG-2 decoder board for playback of a received video data in real-time. The video quality evaluation module analyzes the packets according to packet loss and delay. Packet lateness (see Section 2.2) and total packet loss (see Section 3) are recorded for quantitative analysis. We identifies the position of lost packets (including overly delayed packet) in MPEG2 layers and then calculates the total number of impaired macroblocks.

4.1 Network Parameter Monitoring

In Fig.4, the RTP-like header is inserted in UDP packets in order to measure the network parameters at the server side. This header is constructed by the packet sequence number and the timestamp indicating the sender transmission time. Packet loss is detected by monitoring the sequence number, and the packet delay and jitter are calculated by monitoring the timestamp. The packet header for a video delivery is described in Fig.4. Other anomalies such as out-of-order delivery of packets and packet replication can be detected by means of the timestamp and sequence numbers provided by the RTP header, and handling of these is quite straightforward. In summary, the network parameter monitoring module measures packet loss, network delay and delay variation.

4.2 Network Emulation

The network emulation module generates packet loss and delay in the network. Various network performance emulations are performed according to the designed loss and delay probability distribution. The server and client are connected through this module which helps in finding the network performance objectives and the qualitative analysis of the data.

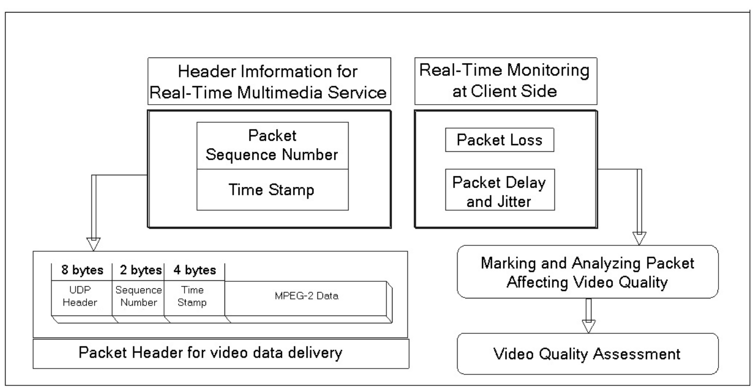

Fig. 4. Real-time network performance parameter monitoring

5 Qualitative Analysis and Experimental Results

In this section, we focus on the relationship between the objective quality and the subjective quality. The MOS grades are collected by human viewers for various videos having different network performance degradations. The end-user marks the quality of the specific video service into five grades. Results are stored and processed into the statistics as a standard input of the quality estimation. After the above subjective evaluation by statistics of user-oriented opinion polls, another analysis is performed, that is, a mapping between MOS grades and DIQ values. This relationship between MOS and DIQ is shown in Fig.5.

User level QoS can be expressed as application level QoS throughout this mapping. End-user's QoS can be expressed as the numerical DIQ value. The network and service provider can offer the appropriate service: the network provider can make a provision of network performance objectives required for the specific network application as a function of network management, and the service provider can collect numerical DIQ values as a function of service management.

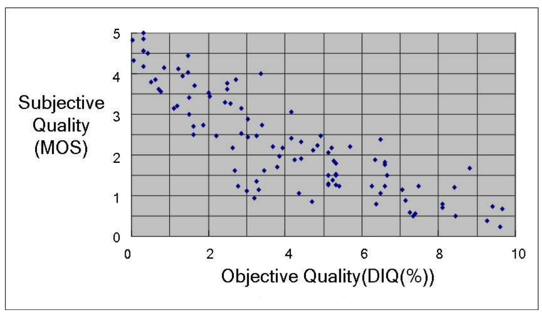

Fig. 5. Relationship between subjective and objective quality

Fig.6 shows experimental results in terms of the acceptable packet loss and late packet probability. The 4Mbps bitrate MPEG-2 video source is used and the total size of the video file is approximately 110Mbytes. The packet size is assumed to be a multiple of 188bytes, that is 4,512bytes (= 188bytes × 24). The size of the playout buffer is approximately of 60Kbytes, resulting in a buffer size (B) of 13 packets. The playout interval R is approximately 9ms (= 4,512bytes × 8bits / 4Mbps). From equation (1), we obtain an acceptable total end-to-end delay TED ($= D_{max}$) which is approximately 127ms ($= D_{min} + B \times R$). From these conditions, network parameters are generated so that packet loss varies from 0.01 to 0.5 and packet delay varies from 10ms to above 127ms having a typical delay distribution. We analyze the relationship between DIQ and network parameters. Subsequently, Fig.6.(a) and (b) show how the DIQ and MOS are related to the packet loss probability. This is used to find an acceptable probability of packet loss for the specific video service. In Fig.6(c) and (d), it is shown that how the

DIQ and MOS are related to the late packet probability. This is used to define an acceptable probability of late packets.

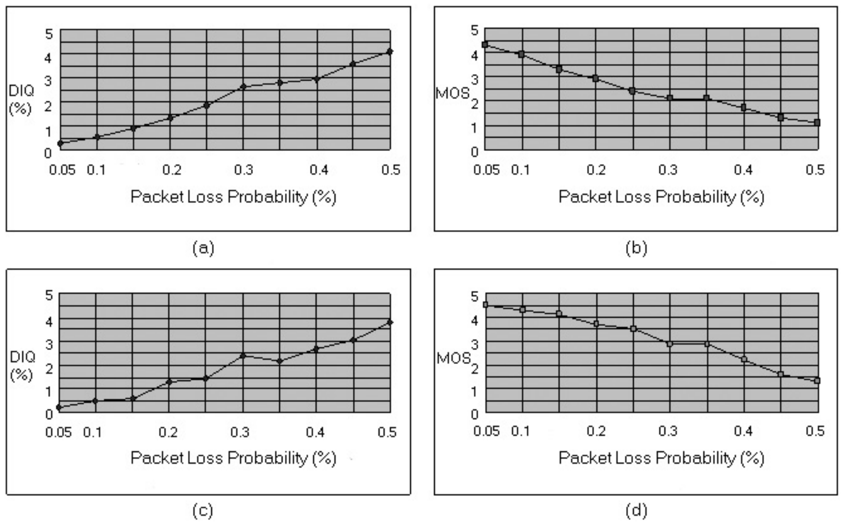

Fig. 6. Acceptable packet loss and late packet probability

If lost or late packet probability is below 0.01%, the user level QoS for high quality service for end-user is very satisfied. Above 0.1%, we can observe that the phenomenon of "motion jerkiness" occurs in the playback video. From 0.2 to 0.25%, "frame freezing" occurs. From 0.3 to 0.4%, the phenomenon of "tiling" appears. The effect of the packet loss is represented as "tiling" and the effect of the packet delay is represented as "frame freezing". Also, if the DIQ is above 3%, the video quality is very poor.

6 Conclusion

In order to evaluate the quality of video applications over the Internet, we proposed a mapping mechanism among the network performance parameters, the application QoS parameters, and the user-perceived QoS. The network performance objectives for MPEG-2 video service are evaluated. Packet loss probability should be below 0.01% and the late packet probability below 0.01%.

Our methodology for real-time QoS monitoring can be used as a part of the network diagnosis system for the improvement of service quality. Also, it can be used in the Service Level Agreement (SLA) monitoring and management system. The service provider must perform SLA monitoring to verify whether the offered service is meeting the QoS requirements specified in the SLA. From the user's viewpoint, he may desire to assess the service and to validate the QoS provided.

Acknowledgement

This work was supported by Center of Innovative Design Optimisation Technology (ERC of Korea Science and Engineering Foundation) and by Grid Middleware Center (Ministry of Information and Communications).

References

1. T. Zhang, U. Jennehag, Y. Xu, “Numerical modeling of transmission errors and video quality of MPEG-2”, Signal Processing: Image Communication, May 2001.
2. S. Jean, J. Park, H. Yoon, “QoS Parameter Translation for the MPEG Services between Layers in ATM Network”, IEEE SPIE vol.3231, 1997.
3. C. J. Sreenan, J. Chen, P. Agrawal, B. Narendran, “Delay reduction techniques for playout buffering”, IEEE Transactions on Multimedia, June 2000.
4. J.I.Jung, “Translation of user’s QoS requirements into ATM performance parameters in B-ISDN”, Computer Networks and ISDN System, October 1996.
5. A. Mehaoua, R. Boutaba, “The Impacts of Errors and Delays on the Performance of MPEG2 Video Communications”, ICASSP ’99, Mars 1999.
6. ITU-R. BT 500-7,“Methodology for The Subjective Assessment of The Quality of Televi-sion Pictures”, October 1995.
7. N. Shivakumar, C.J.Sreenan, B.Narendran, “The Concord Algorithm for Synchronization of Networked Multimedia Streams”, AT&T Bell Laboratories, May 1995
8. C.J.Sreenan, B.Narendran, P.Agrawal, “Internet stream synchronization using Concord”, AT&T Bell Laboratories, January 1996.
9. ISO/IEC 13818-3 International Standard MPEG-2, November 1994.
10. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, “A Transport Protocol for Real-Time Applications”,RFC 1889, January 1996.