15

Smart SoftPhone Device for Networked Audio-Visual QoS/QoE Discovery & Measurement

Jinsul Kim
Digital Media Laboratory, Information and Communications University
Republic of Korea

1. Introduction

Today Multimedia over IP (MoIP) service is provided through the various access networks to Internet, allowing users to get service anytime and anywhere. To control many resources for QoS/QoE guaranteed services over converged networks, developing smart devices and applications applying pervasive network and computing are one of the hot research issues today. The ISO 8402 vocabulary defines quality as the totality of features and characteristics of a product or services that bear on its ability to satisfy stated and implied needs. Quality of Service (QoS) is the collective effect of service performance which determines the degree of satisfaction of the user of the service (from ITU-T E.8001). And then, Quality of Experience (QoE) is a term to allow for subjective as well as objective measures of QoS, performance and all aspects of the interaction (experience) with the service or product (from SLA management handbook). Both QoS and QoE are described very well with networked multimedia application services such as MoIP, IPTV, and mobile IPTV from most recently research (Kim et al., 2008). Due to the shared nature of current network structures, guaranteeing the QoS/QoE of Internet applications from an end-to-end is sometimes difficult and then it has been requested to develop smart devices which have multi-modal functionality for ubiquitous network and computing environment. There are two different aspects, i.e. 'network' and 'multimedia' that are both closely coupled in many critical issues such as QoS, QoE, etc., for MoIP services. An important problem is to provide realtime QoS/QoE-guaranteed multimedia services over packet-based networks. The problem is still unsolved because there are many parameters affecting quality between network and multimedia. In order to solve the problem, a study on QoS/QoE parameters discovery and measurement is necessary.

IP networks are on a steep slope of innovation that will make them the long-term carrier of all types of traffic, including multimedia services in the Next Generation Network (NGN) environment today. However, such networks are not designed for QoS/QoE guaranteed realtime multimedia communication because their variable characteristics (e.g. due to bandwidth, packet loss, delay, etc.) lead to a deterioration in voice/video quality. A major challenge in such networks is how to measure voice/video quality accurately and efficiently considering network resources that provide QoS/QoE-guaranteed services.

In this chapter, we design smart SoftPhone device for guaranteeing human perceived_QoS/QoE which can discover and measure various network parameters during realtime service through IP network. The smart SoftPhone for discovering and measuring of QoS/QoE-factors in realtime consists of four main blocks that is in order to control and measure various parameters independently based on SIP/UDP/RTP protocol during the end-to-end multimedia (voice and video) service. Also, we provide message report procedures and management schemes to guarantee QoS/QoE based on using smart SoftPhone device. In order to report quality parameters optimally during establishing call sessions for MoIP service, we design critical management module blocks for quality reporting. To sum up, for the performance evaluation of the smart SoftPhone with scientific exactitude of quality factors, we examine the proposed technique based on the realtime phone-call service through heterogeneous network. The experimental results confirm that the developed smart SoftPhone is very useful to quality-measuring for the quality guaranteed realtime MoIP service and then it could also be applied to improve quality as a packet compensation device. Finally, we propose QoS/QoE delivery and assessment methodology by model design and performance analysis in considering heterogeneous network and terminals.

The organization of this chapter is as follows. Section 2 describes previous approaches on the identification and characterization of MoIP services by using related works. In section 3, we design modules of smart SoftPhone for quality resource discovery and measurement. The message procedures are presented for call establishing and releasing. In section 4, we describe user, terminal, and network-aware QoS/QoE supported methods with personal mobile broadcasting services. In section 5, RTCP-XR based block types are introduced to monitoring and managing media quality for MoIP services. Section 6 and section 7 present measurement methods with performance evaluations for voice and video quality. Finally, section 8 concludes the chapter.

2. Previous Works

There have been many related research and development efforts in the field of QoS management and measurement for the past decade. Also, today multimedia quality management aspects of QoE have become an important issue with the development of realtime applications such as IP-phones, TV-conferencing and video streaming over IP networks. Specifically, when voice/video data is mixed with various application data, there are worries that there will be a critical degradation in voice/video quality. For the measurement of network parameters, many useful management schemes have been proposed in this research area (Imai et al., 2006). Managing and Controlling of QoS/QoE-factors in realtime is required importantly for stable MoIP service. An important factor for MoIP-quality control technique involves the rate control, which is based largely on network impairments such as jitter, delay, packet loss rate, etc due to the network congestions (Eejaie et al., 1999) (Beritelli et al., 2002). In order to support application services based on the NGN, an end-to-end QoS monitoring tool is developed with qualified performance analysis (Kim et al., 2006).

The different parts of multimedia have different perceptual importance and each part of multimedia does not contribute equally to the overall media quality. Voice/video packets that are perceptually more important are marked, i.e. acquire priority in our approach. If there is any congestion, the packets are less likely to be dropped than the packets that are of less perceptual importance. The QoS schemes which are based on the priority marking are open loop ones and do not make use of changes in the network (Cole & Rosenbluth, 2001).

Currently, most interactive multimedia applications use the realtime transport protocol (RTP) for data transmission with realtime constraints. RTP runs on top of existing transport protocols, typically UDP, and provides realtime applications with end-to-end delivery services such as payload type identification and delivery monitoring. RTP provides transport of data with a notion of time to enable the receiver to reconstruct the timing information of the sender. Besides, RTP messages contain a message sequence number to allow applications to detect packet loss, packet duplication, or packet reordering. RTP is extended by the RTP control protocol (RTCP) that exchanges member information in an on-going session. RTCP monitors the data delivery and provides the users with some statistical functionality. The receivers can use RTCP as a feedback mechanism to notify the sender about the quality of an on-going session. The original RTCP provides overall feedback on the quality of end-to-end networks (Schulzrinne et al., 2005). However, the standard RTCP packet type is defined for quality control in realtime without bidirectional quality reporting and managing procedures in detail through IP networks. The RTP Control Protocol-Extended Reports (RTCP-XR) are management protocol which defines a set of metrics that contains information for assessing the media quality by the IETF (Friedman et al., 2003). The RTCP-XR reports the packet loss rate, the packet discard rate and the distribution of lost/discarded packets. The loss/discard distribution describes calls in terms of bursts (periods during which the loss/discard rate is high enough to cause noticeable quality degradation) and gaps (periods during which lost or discarded packets occur infrequently and hence quality is generally acceptable). To guarantee quality, the RTCP-XR can report the quality directly in terms of the estimated ***R***-factor or the mean opinion score (MOS). The ***R***-factor is a conversational-quality metric in the range of 0 to 100. And the both MOS-LQ and MOS-CQ are in the range of 1 to 5. The RTCP-XR can be implemented as software is integrated into IP phones and gateways inexpensively. Then the messages containing key call-quality-related metrics are exchanged periodically through SoftPhones. However, the RTCP-XR is adequate to monitor the QoS-factor on end-to-end MoIP networks because it doesn't have media quality monitoring functionality. To solve this problem, we propose upgrading some components in the RTCP-XR scheme. Because current IP networks are not designed to support the QoS, quality measurement becomes more important and urgent for more reliable higher quality multi-media services over IP networks. We explore impact of individual packet loss, delay, and jitter on the perceptual media quality on the smart SoftPhone as one of MoIP systems. The evaluation of MoIP service quality is carried out by firstly encoding the input media pre-modified with given network parameter values, and then decoded to generate degraded output signals.

In this paper, we implement smart SoftPhone device for guaranteeing QoS/QoE which can discover and measure various network parameters such as jitter, delay, and packet loss rate, etc., and then propose an end-to-end quality management scheme with the realtime message report procedures to manage the QoS/QoE-factors. The newly presented QoS-factor transmission mechanism for QoE related QoS-factors managing over IP networks is assessed with the performance analysis in the realtime transmission of QoS parameters through various IP networks.

3. Smart SoftPhone Module Design for Discovering and Measuring

In this section, we clarify and design each functionality blocks which are carried on smart SoftPhone for discovering and measuring call-quality over IP network in realtime. There are ten different technical functionality, i.e. 'SIP Stack Module', 'Codec Module', 'RTP Module', 'RTCP-XR Module', 'Transport Module', 'Measurement Module', 'UA Communication Module', 'User Communication Module', 'User Interface Module' and 'Control Module' (Kim et al., 2007).

3.1 Modeling of Smart SoftPhone Functionality

In order to discover and measure quality status, we design 11 critical modules for User Agents (UA) as illustrated in Fig. 1. It comprises in four main blocks and each module is defined as follows:

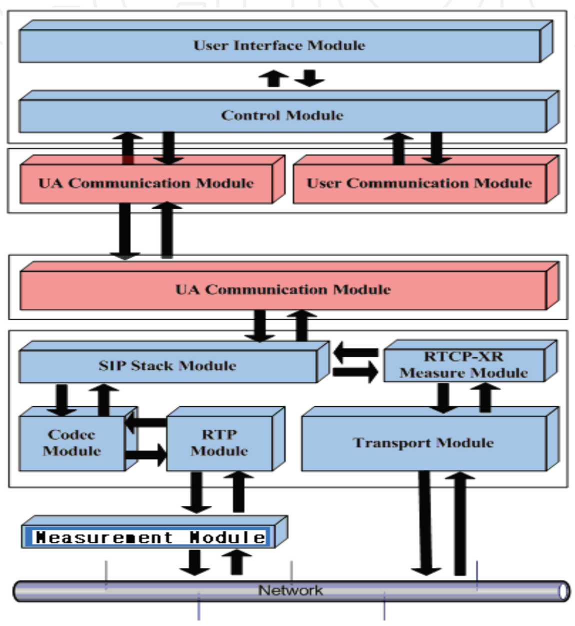


Figure 1. Main blocks and modules of smart SoftPhone functionality

* SIP Stack Module -
  - Analysis of every sending/receiving messages and creation response messages
  - Sending to transport module after adding suitable parameter and header for sending message
  - Analysis of parameter and header in receiving message from transport module
  - Management and application of SoftPhone information, channel information, codec information, etc.
  - Notify codec module of sender's codec information from SDP of receiving message and negotiate with receiver's codec
  - Save up session and codec information
* Codec Module –
  - Providing the encoding and decoding function about two different voice/video codecs
  - Processing of codec (encoding/decoding) and rate value based on SDP information of sender/receiver from SIP stack module
* RTP Module –
  - Sending created data from codec module to receiver SoftPhone through RTP protocol

* RTCP-XR Measure Module –
  - Formation of quality parameters for monitoring and sending/receiving information of quality parameters to SIP stack/transport modules
* Transport Module -
  - Address messages from SIP stack module to network
  - Address receiving message from network to SIP stack module
* Measurement Module –
  - Measure voice/video quality by using packet and rate which is received from RTP module and network
* UA Communication Module -
  - For requesting call connection, interchange of information to SIP stack module and establish SIP session connection
  - Address information to control module in order to show information of SIP message to user
* User Communication Module -
  - Sending and receiving of input information through UDP protocol
* User Interface Module -
  - User give any command by using GUI and sending information to control module
* Control Module -
  - Management of UA communication, user communication, and user information modules
  - Management of various optional information module

3.2 Blocks and Modules for Call Session Control & Quality Management

In this work, we propose realtime message report procedures and management scheme between MoIP-Quality Management (QM) server and smart SoftPhones. The proposed method for the realtime message reporting and management consists of four main processing blocks, as illustrated in Fig. 2. These four different processing modules implement call session module, UDP communication module, quality report message management module and quality measurement/computation/processing modules. All of the call session messages are addressed to quality report message management module by UDP communication. After call-setup is completed, QoS-factors is measuring followed by computation of each quality parameters base on the message processing. Followed by each session establish and release, quality report messages are also recorded in database management module immediately.

An endpoint of SIP based Softswitch (SSW) is known as smart SoftPhone (UA). That is, SIP client loosely denotes SIP end points where UAs run, such as SoftPhones. SSW is intermediated network elements between the end points and engages in the routing of SIP messages from a UA to other UA based on a logical SIP address. SSW also performs functions of authentication, authorization, and signaling compression. A logical SIP URI address consists of a domain and identifies a UA ID number. The UAs belonging to a particular domain register their locations with the SIP registrar of that domain by means of a REGISTER message. Fig. 3 shows SIP based SSW connection between UA#1-SoftPhone and UA#2-SoftPhone.

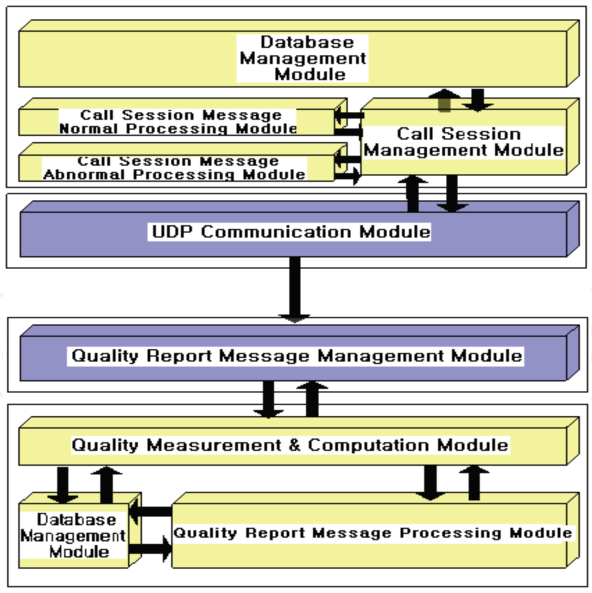


Figure 2. Main blocks and modules for call session control & quality management

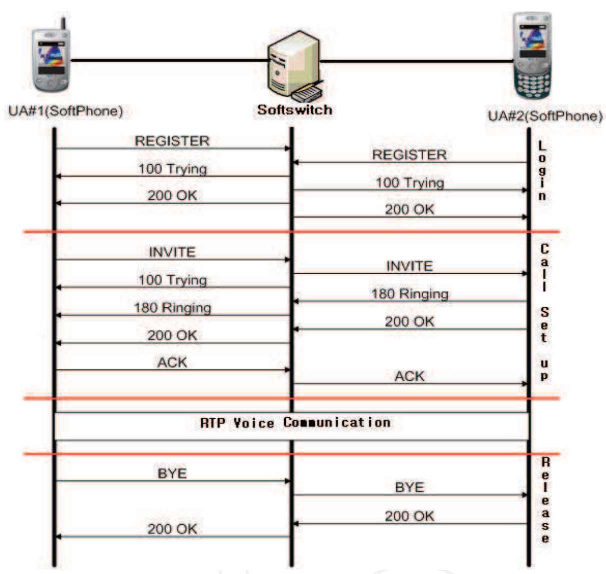


Figure 3. Procedures of call establish/release between Softswitch and smart SoftPhone

4. User, Terminal, and Network-Aware QoS/QoE Guaranteed MoIP Services

For the techniques of new challenging over application source levels we focus on the relationship of user terminals and media streaming sources to support user perceived quality of experience (QoE) based on high quality of service between heterogeneous terminals and heterogeneous networks. These techniques depend on the characteristics of media processing and terminal capabilities such as LCD panel size, resolution, etc., on the heterogeneous network environment.

4.1 IP-based Personal Mobile Broadcasting Services

IP-based communications over a wide area have become more and more popular because of the emerging wireless IP networks and services. However, multimedia transmission and streaming may suffer from an unreliable Internet connection and heterogeneous bandwidth to the different receivers. The multimedia streaming service, which is aware of the network resource, is still a critical topic for the user perceived QoS/QoE guaranteed service. For example, if users wish to call and watch callee on the media phone, they will require resource to support QoS/QoE. The bandwidth allocation in the distribution network will be very different for these two users in order to ensure that both users get the same QoE. The bit rate required for the delivery of content at a fixed quality varies. Therefore, the priority of any individual media stream must correspondingly be allowed to vary both over time and from one stream to another.

Also, the personal mobile broadcasting service as one of MoIP services considers that the QoE-guaranteed media contents are transferred seamlessly between heterogeneous devices based on each user profile. The user currently has various handheld devices. It is always possible to buy an additional new device and use more than one at the same time. In this case, in order to maintain a high QoE-guaranteed media service for specific devices that a user owns, all of the terminal capability information is associated with each user subscription profile on the home subscription server (HSS) system. The HSS function is defined as one of the subdivided functions of IP multimedia subsystem (IMS) service network that is contained in the initial filter criteria.

There is a need to be able to coordinate the access to the supporting terminal capability and user profile information so that they can receive their interesting context service from the originating device to the target client device. This service involves seamlessly transferring QoE-guaranteed video and displaying it between different devices based on user profiles. In order to display the proper scene, HSS, application server (AS), and SSW systems are composed to provide video streams seamlessly for the heterogeneous devices environment. These systems consider both the terminal capability and the user profile for personal mbile broadcasting service as shown in Fig. 4. The HSS system controls and matches all of the profile information in terms of service providers, users and devices. Call session control function (CSCF) can either play the role of a proxy (P)-, interrogating (I)-, or serving (S)-CSCF for seamless session controls.

The personal mobile broadcasting service is more suited for transferring realtime sessions. It is basically to support capturing the session control information from the originating terminal device and transferring them to the target terminal device. This is done by a session control function that allows a user to have heterogeneous mobile devices. For the scenarios to provide a personal mobile broadcasting service, the provider would first find an available network resource for streaming (e.g., bandwidth, multicast address). Then, it would give this information to a content providing end-user that is controlled by the HSS. This example is shown in Fig. 4 as explicated at a football stadium. Second, the content providing end-user sends an extracting video stream by first considering the LCD panel sizes of heterogeneous devices. It also considers an actual broadcasting video stream with multicast or multiple unicasts by using the information in the AS. Third, the receiving client in the mobile environment may be able to select a specific content. This will be based on user profile and terminal capability with logical source information provided by content search results. In this process, service control functions may participate in session routing information gathered by the SSW. This message contains actual content address and session information for receiving it. Fourth, receiving client devices in the mobile environment can request a content delivery function to join the session. Receiving client devices obtain the

content from the content delivery function which is designed and located in the AS. The providing functions in the HSS, SSW and AS control all of the service providers, end-users and terminal capability, together.
The contents provider provides a video stream on many heterogeneous devices such as a cellular phone, PDA, computer, HDTV (IPTV), etc. These devices have various LCD panel sizes and different resolutions from small to big considering heterogeneous networks (e.g., WLAN, Wibro, CDMA). The viewer can feel very uncomfortable if the multimedia contents just transfers from a widescreen sized LCD panel to a small sized LCD panel without considering the resolution and aspect ratios. The user cannot recognize what the scene describes on the device in the mobile service environment. Quality degradation due to down sampling, up-sampling, en(de)coding, etc in the delivery channel can happen for personal mobile broadcasting service.

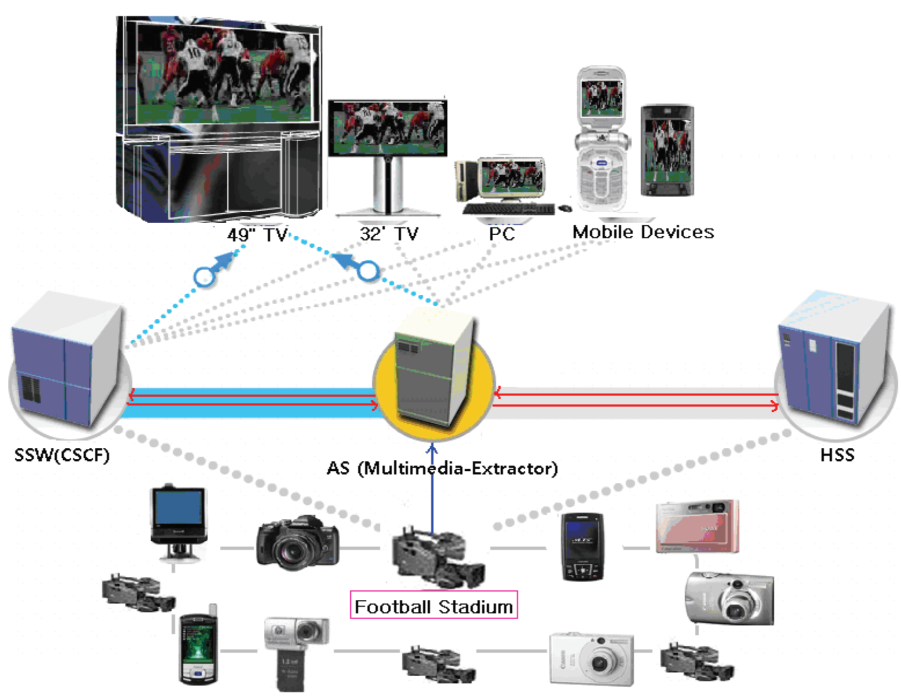


Figure 4. Configuration of personal mobile broadcasting service with considering terminals capability and user profiles

The term resolution is often used as a pixel count and as the spatial dimension in a digital imaging which is captured and displayed on device. The resolution is defined by three cases: Low Resolution (LR), High Resolution (HR), and Super Resolution (SR). Consider the following case: a low resolution image captured by a mobile phone has a resolution of 128 * 128 and we would like to display it on a higher resolution screen of 1024 * 768. Then, SR processing techniques are needed so that the blurring effect can be reduced to improve user perceived QoE. In this case, some more complicated processing techniques are required to convert the LR image to a higher one. The aspect ratio of an image is defined as its width divided by height. If an image is displayed on a device with an aspect ratio different from that of an image, the modification is required, and it is still an interesting issue for frame rate conversion (Telkap, 1995). The resolutions of commonly used displays and several commonly used aspect ratios for various applications are shown in Table 1 (Lee et al., 2007).

	Resolution	Aspect ratio
SDTV	640 * 480	4:3
HDTV	1920*1080	16:9
Computer-VGA	640*480	4:3
Cellular phone	128*128	1:1
PDA	320*200/480*320	5:4

Table 1. Frame resolution and aspect ratio comparison of heterogeneous devices

5. Quality Management for QoS/QoE Guaranteed MoIP Services

The RTP protocol is used for transmitting realtime data information and the RTCP, for sending the control information. The main function of the RTCP is to provide a detailed representation of the voice packets exchanged during an RTP session. Its structure includes the sender report (SR) and the receiver report (RR) transmitted periodically to all participants in the session. It aims at providing a feedback on the quality of the transmission (e.g., delay, jitter, packet loss, etc.), where transmitters send "sender reports" and receivers send "receiver reports" using the RTCP-XR. While the SR includes transmission and reception information for active senders in the session, the RR would also contain the reception information for non-active senders. The MoIP-QM server receives the QoS-factor followed by messages procedure and reports QoS information for the monitoring QoS-factor every 2 seconds.

5.1 RTCP-XR based Quality Monitoring and Management for MoIP Services

For the management of bidirectionial quality resources, we have developed a RTCP-based packet structure to provide an end-to-end transmission controlling method that can report delay, jitter, and packet losses in a timely manner. Our packet structure is similar to the RTCP extended report, which is primarily defined to provide more detailed statistics, particularly for multicast applications. In our case, the RTCP-XR scheme is specifically designed to report delay, jitter, and packet losses for every frame of voice signal. Also, the loss and the discard rates are designed to be calculated for each session at the end-receiver in order to measure realtime values. The original RTCP-XR packet type defined that can be used for speech quality monitoring. However, it is not familiar with realtime speech quality reporting for conversational speech through IP networks. Thus, we modified the RTCP-XR packet scheme in order for reporting and monitoring of bidirectional quality because MoIP communication services such as Internet phone, cellular phone, etc must be recognized as dialogical speech.

For the delay as one of the significant QoS-factor, BT-1 is formatted with both sub-block 1 and sub-block 2. SSRC_1 and SRRC_2 are for the sender and the receiver numbers which are defined randomly. The DLRR in sub-block-1 reports one way delay between the sender and the receiver. The DLRR in sub-block-2 reports the round trip delay which is measured using one way delay information from DLLR in sub-block-1. Fig. 5 shows the message format-II of BT-1 for delay monitoring. In Fig. 6, the information of the jitter and the packets is controlled in BT-2. At first, in order to manage the jitter, we categorize it into three levels of min_jitter, max_jitter, and mean-jitter. Those are reported as the cumulative effects of jitter values obtained through the jitter buffer in our cases. Second, to get the realtime communication, speech quality monitoring field of packet count information of the RTP/RTCP is included in the report block. The Tx/Rx RTP Packets format in BT-2 is designed for monitoring sender/receiver RTP. The Rx RTCP Packets are the RTCP-XR packets which are received at the MoIP-QM server, and the Tx RTCP Packets are the RTCP XR packets which are sent from the SoftPhone to the MoIP-QM server. The final result value of the cumulative packet loss is also included in the BT-2 scheme. The format of the BT-3 scheme is similar to the standard RTCP-XR. However, the loss and the discard rates are computed as soon as the call session is established by using the cumulative packets which is controlled on BT-2. The accuracy of the realtime counter for each packet is really critical point. Specially, for time synchronization, we set our current execution time of SoftPhone by using the NTP (National Time Protocol). That is, time of the host system is designed to be synchronized to the national standard time. Finally, by applying the QoS-factor obtained from speech quality measurement in UA, the MOS and MOS-CQ are defined separately Other factors in BT-3 are added to the standard format scheme as shown in Fig. 7.

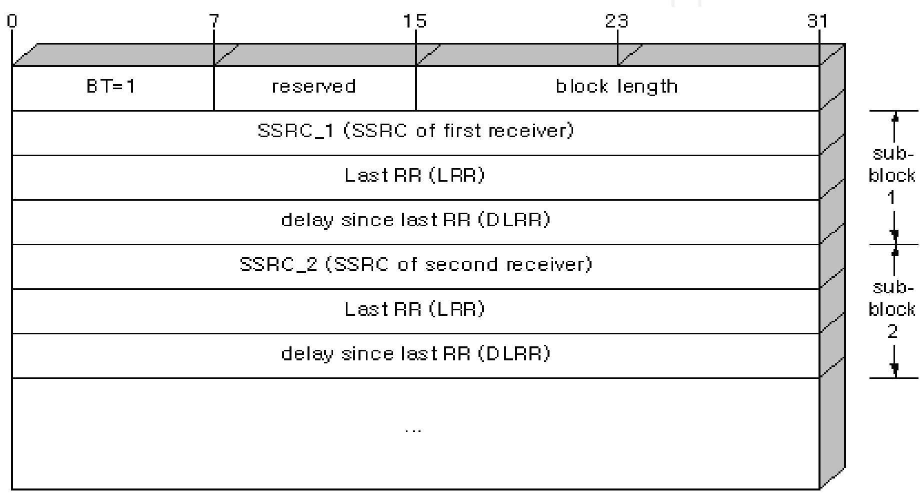


Figure 5. BT=1 delay monitoring for transmission control

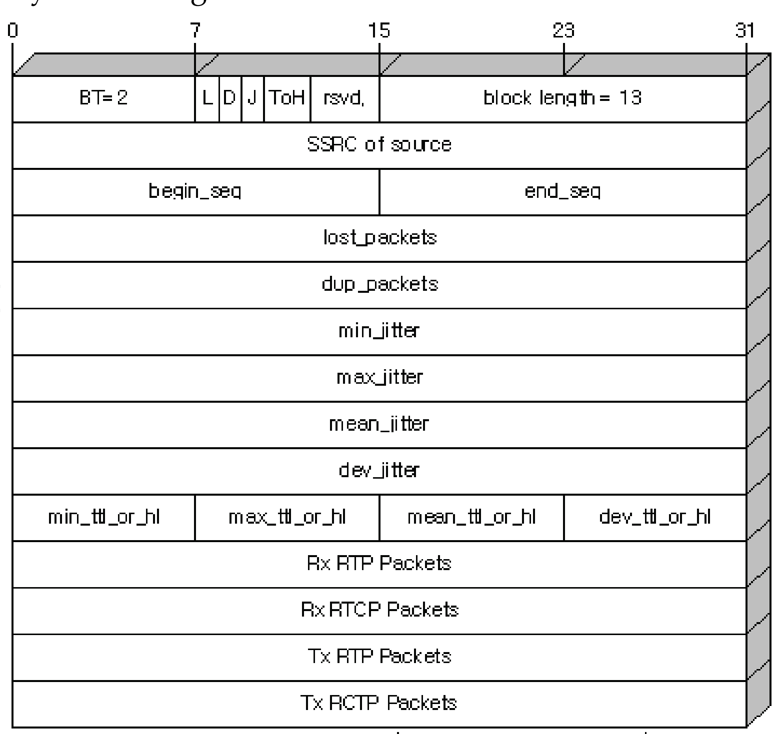


Figure 6. BT=2 jitter, packets for transmission control

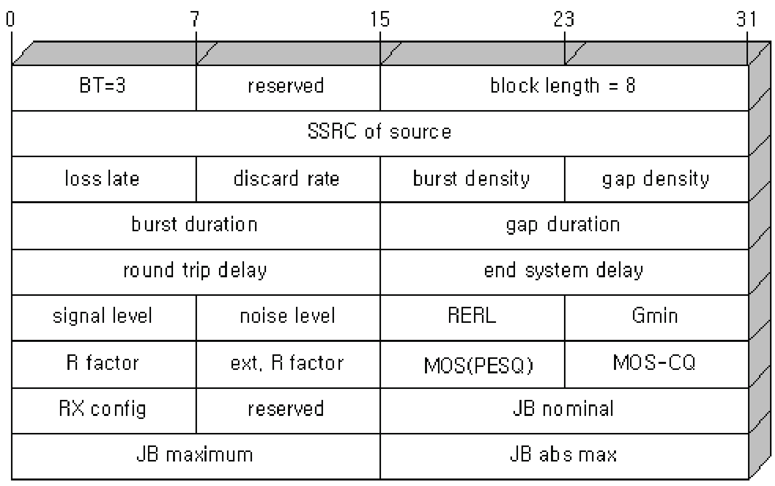


Figure 7. BT=3 loss & discard late, MOS (PESQ), & MOS-CQ for transmission control

6. Voice Quality for QoS/QoE Guaranteed MoIP Services

We propose method of the realtime message report procedures and management schemes for the quality guaranteed MoIP services above. The QoS/QoE-factors control mechanism is experimented with applying Packet Loss Concealment (PLC) algorithm under different packet loss simulating conditions using G.711 and G.729A codecs. When packet losses occur over IP networks, the PLC algorithms employed in speech codecs reconstruct lost speech frames based on the previously received speech information. The PLC algorithm in G.711 Appendix I repeatedly inserts pitch period which is detected from the previous speech in history butter which is called the pitch period replication method. The PLC algorithm employed in G.729A estimates an excitation signal and synthesis filter parameters from last frame which is good condition. To prove the efficiency of the proposed message procedures and management schemes for QoS/QoE-factor control, we evaluate packet loss rate with G.711 and G.729A codecs and then the management scheme is proved as following improvement of results in this chapter.

6.1 Voice Quality Measurement for MoIP Services

Because present IP networks are not designed to support the QoS, the quality measurement becomes more important and urgent for more reliable higher quality multimedia services over IP networks. We explore the impact of the individual packet loss, the delay, and the jitter on the perceptual speech quality in MoIP systems. The MoIP service quality evaluation is carried out by firstly encoding the input speech pre-modified with given network parameter values and then decoded to generate degraded output speech signals. In order to obtain an end-to-end (E2E) MOS between the caller-UA and the callee-UA, we apply the PESQ and the E-Model method. In detail, to obtain the ***R*** factors for E2E measurement over the IP network we need to get ***Id, Ie, Is*** and ***Ij***. Here, ***Ij*** is newly defined as in equation (1) to represent the E2E jitter parameter.

$$R\text{-}factor = R0 - Is - Id - Ij - Ie + A \quad (1)$$

The ITU-T Recommendation provides most of the values and methods to get parameter values except ***Ie*** for codecs, ***Id*** and ***Ij***. First, we obtain ***Ie*** value after the PESQ algorithm applied. Second, we apply the PESQ values to ***Ie*** value of R-factor. We measure the E2E ***Id*** and ***Ij*** from our current network environment. By combining ***Ie, Id*** and ***Ij,*** the final ***R-***factor could be computed for the E2E QoS performance results. Finally, obtained ***R***-factor is reconverted to MOS by using equation (2), which is redefined by the ITU-T SG12.

$$\begin{cases} R \leq 6.5: \ MOS = 1 \\ 6.5 \leq R \leq 100: \ MOS = 1 - \frac{7}{1000}R + \frac{7}{6250}R^2 - \frac{7}{1000000}R^3 \\ R \geq 100: \ MOS = 4.5 \end{cases} \quad (2)$$

6.2 Experimental Environment and Performance Evaluation

To model various packet loss environments, we design burst and random packet losses with 0%, 3%, 5%, and 10% loss rates, and it is considered that a packet contains 1 speech frame, 2 speech frames, or 3 speech frames.

Loss type \ PLC		No PLC			Applied PLC		
		PESQ	R	MOS	PESQ	R	MOS
Random	0%	4.2	93	4.4	4.2	93	4.4
	3%	3.5	71	3.6	3.8	78	3.9
	5%	3.0	59	3.0	3.4	68	3.5
	10%	2.5	48	2.5	3.0	58	3.0
Burst	0%	4.2	93	4.4	4.2	93	4.4
	3%	3.2	64	3.3	3.2	65	3.4
	5%	3.0	57	3.0	3.3	64	3.3
	10%	2.7	50	2.6	2.8	53	2.7

Table 2. Result for realtime environment with G.711

Loss type \ PLC		No PLC			Applied PLC		
		PESQ	R	MOS	PESQ	R	MOS
Random	0%	3.5	72	3.7	3.5	72	3.7
	3%	3.2	63	3.4	3.3	68	3.5
	5%	3.1	62	3.2	3.2	65	3.4
	10%	2.9	57	2.9	3.0	58	3.0
Burst	0%	3.5	72	3.7	3.5	72	3.7
	3%	3.2	64	3.3	3.3	67	3.5
	5%	3.1	61	3.2	3.2	64	3.2
	10%	2.9	56	2.9	3.0	57	2.9

Table 3. Result for realtime environment with G.729A

For the performance evaluation of PLC algorithm based on proposed management scheme, we use the systemic evaluation method called PESQ, defined by ITU-T Recommendation P.862 for objective assessment of quality. After comparing an original signal with a degraded one, the output of PESQ provides a score from -0.5 to 4.5 as a MOS-like score. The reference speech for the real-time environment simulation is the decoded speech without any packet loss, respectively. In the real-time environment, after getting the measured value from PESQ evaluation method, the value is applied to the E-Model evaluation method and then, finally, the MOS is acquired.

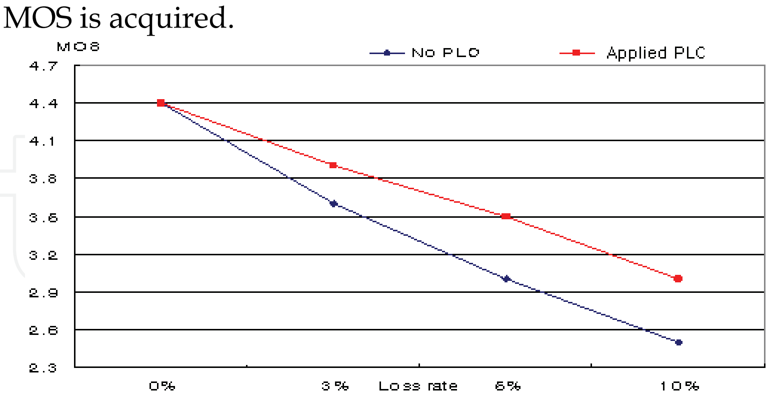


Figure 8. MOS result for random losses with G.711

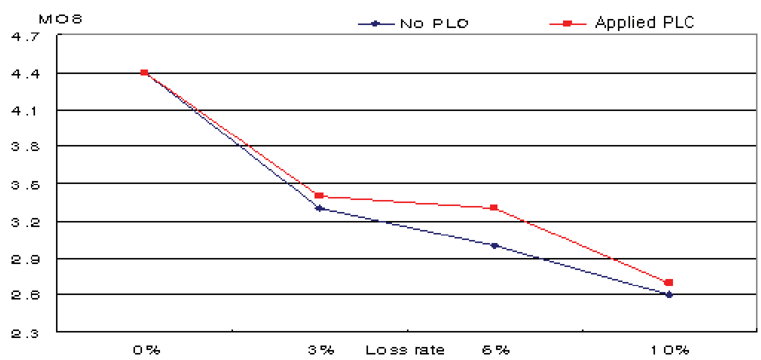


Figure 9. MOS result for burst losses with G.711

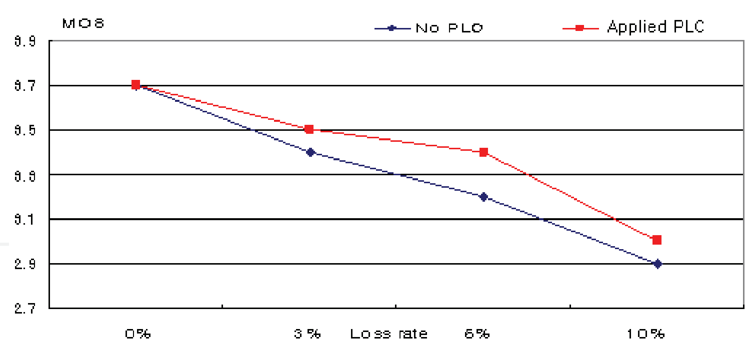


Figure 10. MOS result for random losses with G.729A

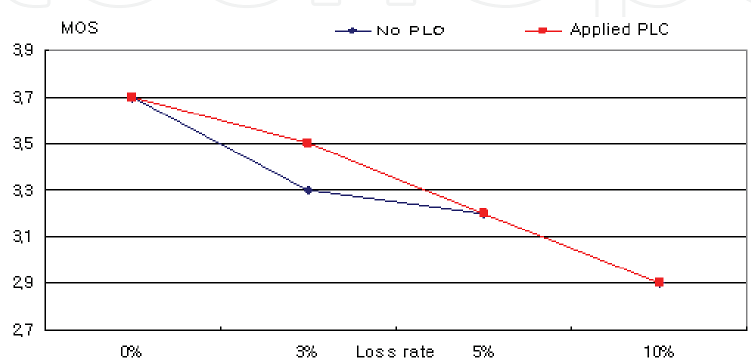


Figure 11. MOS result for burst losses with G.729A

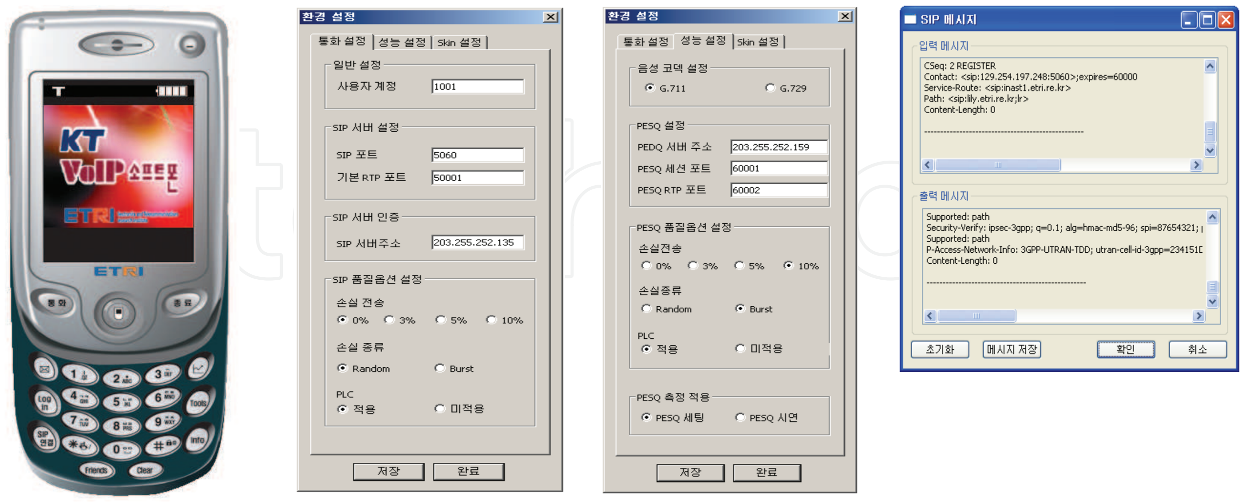
We use 200 Korean dialogue speech utterances from 2 male and 2 female speakers as test data. The duration of each utterance is about 10 seconds. The simulation result shows that the QoS-factor transmission control mechanisms with applying PLC algorithms achieve the distinguished result for the realtime speech quality monitoring than other MoIP systems which use the standard scheme without applying PLC algorithm during conversation through IP-based network environment. In the following result tables, 'fpp' refers to frame per packet. The performance of the PLC algorithm in G.711 and G.729A is compared to that of the no-PLC algorithm employed in G.711 and G.729A. In Table 2, the results of PLC performance evaluation for G.711 in realtime environment are summarized by the measurement methods of PESQ, the R-factor, and the MOS. The applied PLC achieves the PESQ gains between 0.1 and 0.3 for packet loss. The corresponding gains for the R-factors and the MOS scores are also achieved by these PESQ gains. High gains are achieved at random losses with high loss rates as shown in Table 2 (in Fig. 8 and 9). In Table 3, the results of PLC performance evaluation for G.729A in real-time environment are also summarized by the measurement methods of PESQ, the R-factor, and the MOS. The PLC algorithm achieves the PESQ gain of 0.1 for all types of loss. The corresponding gains for the R-factors and the MOS scores are achieved by these PESQ gains. Finally, even though the MOS is not highly improved for burst losses, the MOS gain of 0.1 is generally achieved as shown in Table 3 (in Fig. 10 and 11).

6.3 Implementation of SoftPhone and MoIP-QM Server

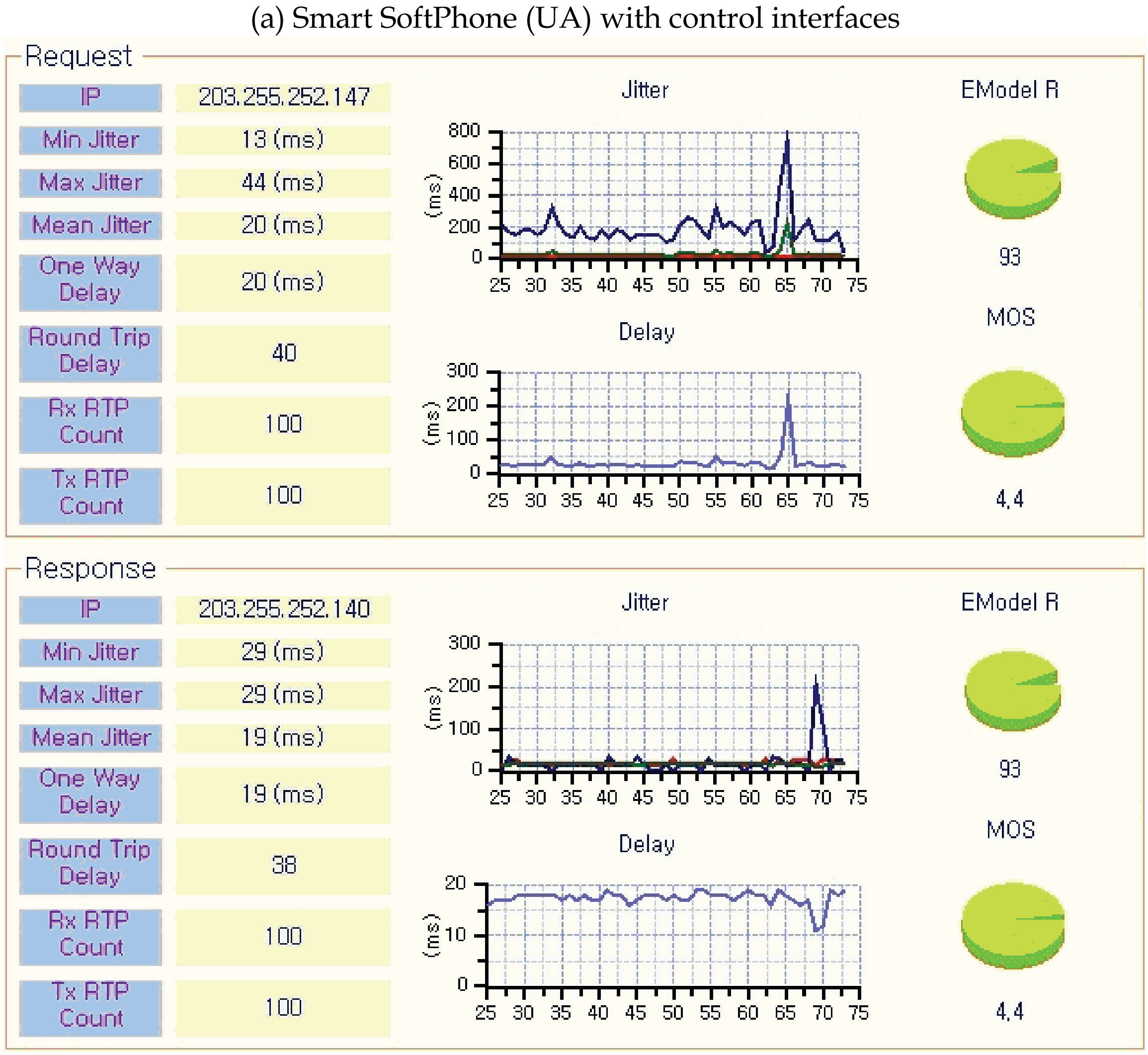
The main part of the smart SoftPhone is implemented in MFC and C# in .Net environment for delivering and measuring of QoS/QoE-factors with various parameters transmission control in realtime. As illustrated in Fig. 12 (a), the realtime MoIP packets delivering and measuring function on the smart SoftPhone is designed as various GUI function to manage an end-to-end media quality over IP network. Smart SoftPhone has functionality for media (voice/video) communication with other devices through SIP stack. The SIP stack implement to connect between network and any user's GUI of smart SoftPhone. We use vocal 1.5.0 server in order for role of SSW, which is registration, establish, and release from smart SoftPhone. The user can call session establish and release by using GUI on smart SoftPhone which display the SIP message for sending and receiving to user. By the SIP message, the user checks current situation in realtime between caller and callee. Also, it measures voice quality by using objective measurement method which is called PESQ as standard from ITU-T and has reporting function network parameters which follows by RTCP-XR formation especially to monitoring QoS/QoE-factors. While the sessions establish and release by phone, three different RTCP-based packet structures BTs are modeled by realtime information such as caller ID, callee ID, delay, jitter, packet loss, packet discard, etc., for each frame of speech signal. For speech signal measurement, our smart SoftPhone designs to control of encoding/decoding voice packets in G.711 and G.729A as shown in Fig. 12 (a). One important objective point with the implementation is to explore the measured values of PESQ, ***R-factor*** and MOS with time synchronization of each session at the end-sender and end-receiver by using NTP.

MoIP-QM server is implemented in Delphi and C# in .Net environment followed by the structure of BT=1, BT=2, and BT=3 for monitoring of QoS/QoE-factor transmission control in realtime. While the sessions establish and release by SoftPhone, three different RTCP-based packet structures BTs are modeled by realtime information.

Two main parts of MoIP-QM server, the request and response are designed based on message procedures while call session is opened, and then jitter, delay, etc., per packet are reported and described visually on MoIP-QM server as shown in Fig. 12 (b).



(a) Smart SoftPhone (UA) with control interfaces



(b) MoIP-QM server with QoS/QoE-factors measurement

Figure 12. Implementation of SoftPhone and MoIP-QM server

7. Video Quality Measurement for QoS/QoE Guaranteed MoIP Services

Video quality for MoIP services can be affected by variety of factors such as video coders, transmission type, bandwidth limitation etc. We need to measure video quality in a fundamental requirement in modern communications systems for technical, legal and commercial reasons. Video quality measurement can be carried out using either objective or subjective methods of video quality.

7.1 Video Quality Indicators Extraction and Measurement

MoIP service can be defined as a kind of convergence service composed of broadcasting and telecommunication sectors. In this sense, various multimedia can be provided through IP networks with interactivity. Since MoIP services are very sensitive to the network degradation such as packet loss, out of order, and jitter, the quality of service cannot be guaranteed. With considering various effects both network and video levels for MoIP service, several distortions including blurring, block distortion, color error, jerkiness, edge busyness, etc. aspects of user perceived QoE, occur during transmitting and encoding/decoding processes times. At the measuring points, block distortion, blurring effects, and color error mainly happen on the video source. From the transport area, we can assess at the points where there are areas before/after the IP network and before/after the access network. Color error, jerkiness, edge busyness, etc., in the source is affected by packet loss, delay, jitter, etc.

Degradation measures, which can give the numerical information of video quality, play an important role. Most researchers have used many forms of quantitative quality metrics such as the mean squared error (MSE) and peak-to-noise ratio (PSNR) as Full Reference (FR) based objective video quality measure method. The most common objective criterion is the mean square error (MSE). The MSE of original and processed image refer to (3).

$$MSE = \frac{1}{M.N}\sum_{x=1}^{M}\sum_{y=1}^{N}| f(x,y) - \tilde{f}(x,y)|^2 \quad (3)$$

where $f(x,y)$ is the original image, $\tilde{f}(x,y)$ is the processed image, $\boldsymbol{M}$ and $\boldsymbol{N}$ are the height and width of the images. Peak Signal-to-Noise Ratio (PSNR) is another widely used way to measure all image quality evaluation. PSNR is a MSE derived objective quality measure. PSNR is defined in (4) where peak signal strength is assumed as 255.

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \quad (4)$$

However, since pixel values of original and degraded videos are used in the full-reference model in MSE and PSNR, computational burden is very large.

Also, although MSE and PSNR metrics are simple and widely used metric results are poorly correlated with subjective rating since they do not model the human visual system. The example for comparing MSE and PSNR values is shown in Fig. 13. People feel uncomfortable in spite of the same PSNR among the image.

Figure 13. Correlation between MSE/PSNR results and subjective rating

It is necessary to develop new objective metrics in order to reflect network based QoS-status aware end-user perceived QoE indicators for accurate prediction and measurement in considering correlation with subjective measurement.

Error	Model	Method	Cause
Blurring	RR	Comparing edge standard deviation between OS and PS	Encoding/ Decoding
Block distortion	RR	Estimation using edge of vertical/horizontal direction and weight of around pixels	Encoding/ Decoding/ Network Transport
Color error	RR	Comparing hue and saturation between OS and PS	Encoding/ Decoding/ Network Transport
Edge busyness	RR	Comparing edge average values between OS and PS	Network Transport
Jerkiness	NR	Freezing frame extraction using edge difference between frames	Network Transport

Table 4. Video quality indicators

The current issue in the area is to measure in realtime with face value which service providers really want the greatest accuracy. Thus, our focus is prediction and measurement quality of the distorted video contents frames with considering user perceived QoE, basically, and then the several additional proposed methods are useful for applying to estimation of networked-QoS aware video-QoE indicators base on reduced-reference (RR) for blurring, block distortion, color error, and edge busyness/no-reference (NR) for jerkiness video measurement method. In Table 4, we briefly present the cause of errors and key of measuring method.

In the RR model, extracted features of the original and degraded videos are used instead of all pixel values. Perceptual video quality is computed by using these features. Finally, no-reference models use only the degraded video sequence without using the original video sequence. Although the NR model is very fast, accuracy for measuring degradation cannot be guaranteed. The perceptual objective video quality models are shown in Fig. 14.

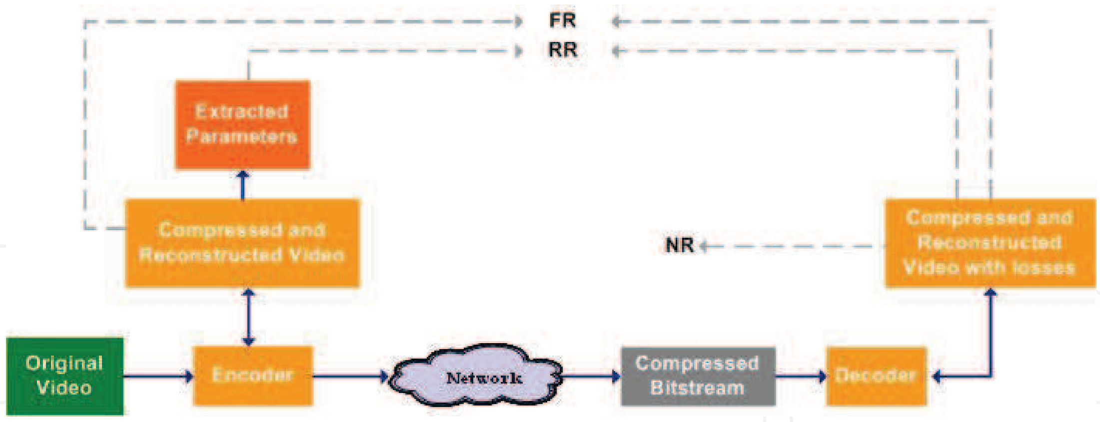


Figure 14. Perceptual video quality measurement model

7.2 Hybrid Objective Metrics for Video Quality Measurement

Differences in quality of video are due to loss compression/decompression as well as transmission errors, which lead to artifacts in the received viewing contents. The amount of artifacts and visibility of these distortions strongly depend on the video contents. There are two types of quality to measure and to verify digital video quality which is delivered to the end user to identify content quality degradation: objective and subjective quality. Both of these are to develop Video Quality Metrics (VQM) which is intended to provide calculated values that are strongly correlated with a viewers' assessment. In this research, we mention above five indicators which include blurring, block distortion, edge busyness, color error, and jerkiness according to the whole transmission process which can produce artifacts to digital video QoE. We design hybrid VQM model which is defined in (5).

$$VQM = a \times E_{edge} + b \times E_{block} + c \times E_{blur} + d \times E_{color} - E_{jerky} + C \quad (5)$$

We use the multiple linear regression analysis. If we suppose as $y = a + bx$, it can be described by the results of QoE indicators and subjective MOS as in $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$, and then to get linear coefficient a, b, the procedure is as follows,

$$\Pi = \sum_{i=1}^{n} [y_i - f(x_i)]^2 = \sum_{i=1}^{n} [y_i - (a + bx_i)]^2 \quad (6)$$

By differentiation of a, b, as follows, and then define by x, y

$$\begin{aligned} \frac{\partial \Pi}{\partial a} &= 2\sum_{i=1}^{n} [y_i - (a + bx_i)] = 0 \\ \frac{\partial \Pi}{\partial b} &= 2\sum_{i=1}^{n} x_i [y_i - (a + bx_i)] = 0 \end{aligned} \quad (7)$$

$$\sum_{i=1}^{n} y_i = a\sum_{i=1}^{n} 1 + b\sum_{i=1}^{n} x_i$$
$$\sum_{i=1}^{n} x_i y_i = a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 \qquad (8)$$

Finally, linear coefficient a, and b are as follows, and then from the equation, coefficients a, b, c, d, C for the hybrid VQM is a = -17.809, b = -3.352, c = 5.340, d = 32.191, C = 4.424 from the equation (5).

$$a = \frac{(\sum_{i=1}^{n} y_i)(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} x_i y_i)}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}, \quad b = \frac{n\sum_{i=1}^{n} x_i y_i - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \qquad (9)$$

Finally, we design the hybrid VQM model as follows,

$$VQM = -17.809E_{edge} - 3.352E_{block} + 5.340E_{blur} + 32.191E_{color} - E_{jerky} + 4.424 \qquad (10)$$

7.3 Heterogeneous Networks and Terminals-Aware for MoIP Services

The content providing end-user sends an extracting video stream by first considering the LCD panel sizes of heterogeneous devices. The personal mobile broadcasting contents provider provides a video stream on many heterogeneous handheld devices such as a cellular phone, PDA, computer, etc. These devices have various LCD panel sizes and different resolutions from small to big considering heterogeneous networks (e.g., WLAN, WIMAX, (W)CDMA). The viewer can feel very uncomfortable if the multimedia contents just transfer from a widescreen sized LCD panel to a small sized LCD panel without considering the resolution and aspect ratios. The user cannot recognize what the scene describes on the device in the personal mobile broadcasting service environment. Quality degradation due to down sampling, up-sampling, en(de)coding, etc in the delivery channel can happen for the personal mobile broadcasting service.

Table 5 shows the results packet loss rates with considering LCD panel size of heterogeneous terminals in their bandwidth limitation when the personal mobile broadcasting services deliver video content to their various target terminals through heterogeneous networks which has different LCD size. We consider VGA (resolution: 649*480, 150kbps (video), 192kbps (audio)), CIF(resolution: 356*288, 75kbps (video), 192kbps (audio)), QVGA(resolution: 320*240, 63kbps (video), 192kbps (audio)), QCIF(resolution: 178*144, 41kbps (video), 192kbps (audio)), Cellular Phone Size (resolution: 128*128, 34kbps (video), 192kbps (audio)) with same commercial content for heterogeneous handheld devices.

Handover (HO) Time		No HO	30.0s	60.5s	90.5s
Analysis Section		0.0s ~ 10.0s	25.5s ~ 35.5s	55.5s ~ 65.5s	85.5s ~ 95.5s
Source	HO delay	LAN(100M)	WLAN(11M)	WiMAX(5M)	WCDMA(384Kb)
VGA (Computer, SDTV)	0.0ms	0/1731	42/803	90/1325	271/1457
	0.3ms	0/1045	103/1077	58/1513	276/792
	0.7ms	0/953	89/960	108/1012	1002/1680
CIF (PDA-I)	0.0ms	0/635	72/791	7/703	87/579
	0.3ms	0/642	27/518	102/770	194/794
	0.7ms	0/913	148/1062	49/520	94/559
QVGA (PDA-II)	0.0ms	0/667	30/601	7/945	78/569
	0.3ms	0/518	42/720	48/457	82/711
	0.7ms	0/509	35/431	57/847	159/699
QCIF (Cellular Phone-I)	0.0ms	0/538	19/398	1/431	11/500
	0.3ms	0/430	22/390	34/394	31/336
	0.7ms	0/391	54/442	36/446	43/454
Cellular Phone	0.0ms	0/343	19/338	0/327	4/386
	0.3ms	0/478	20/355	30/322	39/353
	0.7ms	0/385	35/386	30/308	33/316

Table 5. HO in heterogeneous networks/terminals for the personal mobile broadcasting service

The display of a specific target scene considering the context-aware viewer's visual sight is one of the important facts in providing QoE-guaranteed viewer centric mobile broadcasting service (Kim et al, 2008). When the original contents used for a big LCD panel are transferred to a small LCD panel, the video sequence captured for normal viewing on a standard IPTV may have an adverse effect. The viewer trying to view the image on the smaller display may have uncomfortable experiences. In order to provide the QoE guaranteed service to satisfy the visual perception of the viewer, the specific context based extracting methodology should be applied to the contents on devices with considering LCD panel size of the targeted device, together.

8. Conclusion

The demand on the guaranteed QoS/QoE of the flexible audio-visual content through the heterogeneous networks and the display heterogeneous terminals will increase as much as the MoIP service has been developed rapidly in residential and business communication markets.

- In this chapter, several related active research issues of Mobile IPTV service are highlighted and some new research directions have been pointed out.
- Provide critical message procedures applying RTCP-XR based packet structures BTs to manage call session with quality factors such as jitter, delay, loss, etc.
- Design management module for call session and for quality reporting using SoftPhone.
- Present QoS/QoE-factors transmission control mechanism
- Assess voice quality with a performance analysis of the PLC algorithms
- Derive video quality guaranteed technologies that enable end-to-end personal mobile broadcasting service in a heterogeneous environment.
- Ability to correlate the impact of networking resource, terminal capability, and user profile at each of the media stream applications.

Finally, our proposed methods of transmission procedure and management scheme using SoftPhone are very useful to manage QoS/QoE audio-visual quality through IP network. Also, hybrid objective metric is very useful for user perceived QoE-aware video quality measurement.

9. References

Beritelli, F.; Ruggeri, G. & Schembra, G. (2002). TCP-Friendly Transmission of Voice over IP, *Proceedings of IEEE International Conference on Communications,* pp. 1204-1208, 0-7803-7400-2, April 2002, New York, USA

Cole, R. G. & Rosenbluth, J. H. (2001). Voice over IP Performance Monitoring, *ACM SIGCOMM Computer Communications Review*, Vol. 31, No.2, April 2001, pp. 9-24, 0146-4833

Eejaie, R.; Handley, M. & Estrin, D. (1999). RAP: An End-to-end Rate-based Congestion Control Mechanism for Realtime Streams in the Internet, *Proceedings of IEEE INFOCOM,* pp. 21-25, 0-7803-5417-6, March 1999, New York, USA

Friedman, T.; Caceres, R. & Clark, A. (2003). RTP Control Protocol Extended Reports, *IETF RFC* 3611, Nov. 2003

Imai, S.; Yamada, A.; Ueno, H.; Nakamichi, K. & Chugo, A. (2006). Voice Quality Management for IP Networks based on Automatic Change Detection of Monitoring Data, *Proceedings of APNOMS, LNCS* (Lecture Notes in Computer Science) Springer-Verlag, Berlin Heidelberg, Vol. 4238, pp. 27-29, September 2006, 0302-9743

ITU-T Recommendation G.711 Appendix I. (1999). A High Quality Low-Complexity Algorithm for Packet Loss Concealment with G.711, *ITU-T,* Sept. 1999

ITU-T Recommendation G.729 Annex A. (1996). Coding of Speech at 8 kbit/s using Conjugate-Structure Algebraic-Code-Excited-Linear-Prediction (CS-ACELP), *ITU-T,* Nov. 1996

Kim, C. C.; Shin, S. C.; Ha, S. Y.; Han S. Y. & Kim, Y. J. (2006). End-to-End QoS Monitoring Tool Development and Performance Analysis for NGN, *Proceedings of APNOMS,* LNCS (Lecture Notes in Computer Science), Springer-Verlag, Berlin Heidelberg, Vol. 4238, pp. 332-341, September 2006, 0302-9743

Kim, J. S.; Hahn, M. S. & Lee, H. W. (2007). Smart SoftPhone Device for the Network Quality Parameters Discovery and Measurement, *HCI International 2007,* LNCS (Lecture Notes in Computer Science), Springer-Verlag, Berlin Heidelberg, Vol. 4555, pp. 898-907, January 2007, 0302-9743

Kim, J. S.; Um, T. W.; Won, R.; Lee, B.S. & Hahn, M. S. (2008). Heterogeneous Networks and Terminal-Aware QoS/QoE-Guaranteed Mobile IPTV services, *IEEE Comm. Maga.*, Vol.46, No.5, May 2008, pp.110-117, 0163-6804

Lee, M. S.; Shen, M. Y.; Kuo, C. C. & Yoneyama, A. (2007). Techniques for Flexible Image/Video Resolution Conversion with Heterogeneous Terminals, *IEEE Comm. Maga.*, Vol. 45, No. 1, Jan. 2007, pp. 61-67, 0163-6804

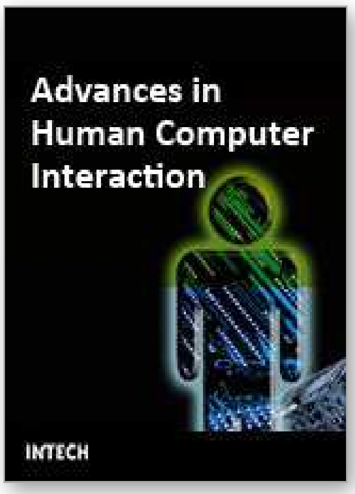
Park, C. S.; Wang, T. S. & Ko, S. J. (2007). Error Concealment using Inter-Layer Correlation for Scalable Video Coding, *ETRI Journal,* Vol. 29, No. 3, June 2007, pp. 390-392, 1225-6463

Recommendation J.149. (2004). Method for Specifying Accuracy and Cross-Calibration of Video Quality Metrics (VQM), *ITU-T,* March, 2004

Schulzrinne, H.; Casner, S.; Frederick, R. & Jacobson, V. (2005). RTP: A Transport Protocol for Real-Time Application, *IETF RFC 3550,* July 2005

Schwarz, H.; Marpe, D. & Wiegand, T. (2005). Comparison of MCTF and Closed-Loop Hierarchical B Pictures, *ITU-T & ISO/IEC JTC1, JVT-P059,* July 2005

Telkap, A. M. (1995). *Digital Video Processing,* Prentice Hall, Englewwod Cliffs, NJ, 1995



Advances in Human Computer Interaction

Edited by Shane Pinder

ISBN 978-953-7619-15-2

Hard cover, 600 pages

Publisher InTech

Published online 01, October, 2008

Published in print edition October, 2008

In these 34 chapters, we survey the broad disciplines that loosely inhabit the study and practice of human-computer interaction. Our authors are passionate advocates of innovative applications, novel approaches, and modern advances in this exciting and developing field. It is our wish that the reader consider not only what our authors have written and the experimentation they have described, but also the examples they have set.

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821