International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1011

(06/2010)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Multimedia Quality of Service and performance – Generic and user-related aspects

Reference guide to quality of experience assessment methodologies

Recommendation ITU-T G.1011

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T G.1011

Reference guide to quality of experience assessment methodologies

Summary

Recommendation ITU-T G.1011 provides a reference guide to existing standards for quality of experience (QoE) assessment methodologies. This Recommendation specifies QoE assessment approaches and classifications of different applications, and identifies a taxonomy of QoE assessment standards with different technical categorizations.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T G.1011	2010-06-29	12

Keywords

Quality of experience (QoE).

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

Recommendation ITU-T G.1011

Reference guide to quality of experience assessment methodologies

1 Scope

Quality of experience (QoE) is a widely used term in the telecommunication industry. However, there is no single document available which describes methods for assessing QoE. This Recommendation provides a reference guide to existing standardized methods in ITU for QoE assessment. The overall approaches for the different assessment methodologies typically used are described, and a taxonomy of QoE assessment standards is defined.

For different applications, this Recommendation identifies suitable subjective assessment methodologies, and where applicable, objective assessment methods which can be used to estimate subjective opinion, and gives guidance on their usage and limitations.

The intended users of this Recommendation include carriers, service providers and equipment manufacturers.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T G.107] Recommendation ITU-T G.107 (2002), *The E-model, a computational model for use in transmission planning*.

[ITU-T G.1030] Recommendation ITU-T G.1030 (2005), *Estimating end-to-end performance in IP networks for data applications*.

[ITU-T G.1070] Recommendation ITU-T G.1070 (2007), *Opinion model for video-telephony applications*.

[ITU-T J.140] Recommendation ITU-T J.140 (1998), *Subjective picture quality assessment for digital cable television systems*.

[ITU-T J.144] Recommendation ITU-T J.144 (2004), *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*.

[ITU-T J.246] Recommendation ITU-T J.246 (2008), *Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference*.

[ITU-T J.247] Recommendation ITU-T J.247 (2008), *Objective perceptual multimedia video quality measurement in the presence of a full reference*.

[ITU-T J.249] Recommendation ITU-T J.249 (2010), *Perceptual video quality measurement techniques for digital cable television in the presence of a reduced reference*.

[ITU-T P.10 Amd.2] Recommendation ITU-T P.10/G.100 Amendment 2 (2008), *New definitions for inclusion in Recommendation ITU-T P.10/G.100*.

[ITU-T P.561] Recommendation ITU-T P.561 (2002), *In-service non-intrusive measurement device – Voice service measurements*.

[ITU-T P.562] Recommendation ITU-T P.562 (2004), *Analysis and interpretation of INMD voice-service measurements*.

[ITU-T P.563] Recommendation ITU-T P.563 (2004), *Single-ended method for objective speech quality assessment in narrow-band telephony applications*.

[ITU-T P.564] Recommendation ITU-T P.564 (2007), *Conformance testing for voice over IP transmission quality assessment models*.

[ITU-T P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.

[ITU-T P.805] Recommendation ITU-T P.805 (2007), *Subjective evaluation of conversational quality*.

[ITU-T P.830] Recommendation ITU-T P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.

[ITU-T P.835] Recommendation ITU-T P.835 (2003), *Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm*.

[ITU-T P.862] Recommendation ITU-T P.862 (2001), *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*.

[ITU-T P.862.1] Recommendation ITU-T P.862.1 (2003), *Mapping function for transforming P.862 raw result scores to MOS-LQO*.

[ITU-T P.862.2] Recommendation ITU-T P.862.2 (2007), *Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs*.

[ITU-T P.910] Recommendation ITU-T P.910 (1999), *Subjective video quality assessment methods for multimedia applications*.

[ITU-T P.911] Recommendation ITU-T P.911 (1998), *Subjective audiovisual quality assessment methods for multimedia applications*.

[ITU-T P.920] Recommendation ITU-T P.920 (2000), *Interactive test methods for audiovisual communications*.

[ITU-R BS.1116-1] Recommendation ITU-R BS.1116-1 (1997), *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*.

[ITU-R BS.1285] Recommendation ITU-R BS.1285 (1997), *Pre-selection methods for the subjective assessment of small impairments in audio systems*.

[ITU-R BS.1387] Recommendation ITU-R BS.1387 (2001), *Method for objective measurements of perceived audio quality*.

[ITU-R BS.1534-1] Recommendation ITU-R BS.1534-1 (2003), *Method for the subjective assessment of intermediate quality levels of coding systems*.

[ITU-R BT.500-12] Recommendation ITU-R BT.500-12 (2009), *Methodology for the subjective assessment of the quality of television pictures*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 quality of experience (QoE) [ITU-T P.10 Amd.2]: The overall acceptability of an application or service, as perceived subjectively by the end-user.

NOTE 1 – Quality of experience includes the complete end-to-end system effects (client, terminal, network, services infrastructure, etc.).

NOTE 2 – Overall acceptability may be influenced by user expectations and context.

4 Abbreviations

This Recommendation uses the following abbreviations:

CIF	Common Intermediate Format
FR	Full Reference
HDTV	High Definition Television
IPTV	Internet Protocol Television
ISDN	Integrated Services Digital Network
MOS	Mean Opinion Score
$MOS_{A,V,MM}$	Mean Opinion Score for Audio, Video or Multimedia
NB	Narrowband
NR	No reference
PSTN	Public Switched Telephone Network
QCIF	Quarter Common Intermediate Format
RR	Reduced Reference
SDTV	Standard Definition Television
SWB	SuperWideband
VGA	Video Graphics Array
VoIP	Voice over Internet Protocol
WB	Wideband

5 Conventions

This Recommendation uses the following conventions:

LAB: laboratory testing;

MON: monitoring;

PLN: planning.

6 General approach to QoE assessment

In principle, assessment of QoE must be performed using subjective tests with metrics such as mean opinion score (MOS). However, it is also possible and sometimes more convenient to estimate QoE based on objective testing and associated quality estimation models. Through different quality estimation models, it is possible to measure or calculate the objective parameters affecting QoE, in order to evaluate QoE. Subjective testing needs more resources and efforts to be carried out because it requires human subjects. On the other hand, objective measurement and automatic calculation using appropriate quality estimation models is generally much faster and cheaper, but the accuracy of the final evaluation depends on the accuracy of these models.

In general, if QoE is evaluated objectively, it may be tested in intrusive mode, non-intrusive mode, or planning mode. "Intrusive mode" means that the quality assessment system requires that a signal be injected into the system under test in order to generate a degraded output signal. This implies that the channel must be taken out of service for normal traffic. Conversely, for "non-intrusive mode", the quality assessment system can be used whilst live traffic is carried by the channel, without the need for any active test signals. "Planning mode" is not used in a real-time environment, but as a tool for the design of systems, and hence does not require any real-time inputs.

Irrespective of which mode it is in, the most important thing is to identify the key objective parameters that affect QoE and then relate these parameters to QoE measurement indices through specific quality estimation models.

Since quality of experience, i.e., perceived quality, is internal to the user, perceptual quality tests are ultimately the only means of validly and reliably assessing quality (see Figure 6-1 for a graphical explanation of the terms "valid" and "reliable").

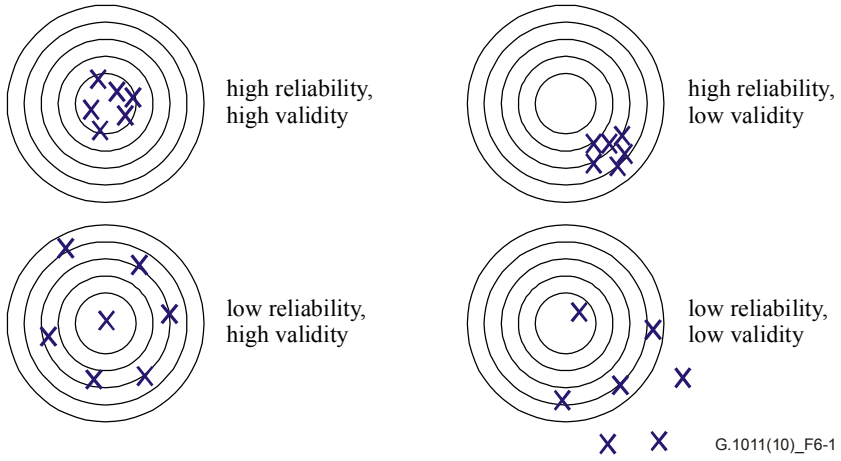

Figure 6-1 – Validity and reliability

NOTE – Validity describes how well a method actually measures what it is intended to measure. Reliability describes the accuracy of a method in terms of the scattering of results, e.g., when test assessment is repeated ([b-Guilford, 1954], pp. 349-357). Nowadays, validity is often replaced with the term accuracy, and, reliability is replaced with the term precision.

The intrinsic variation that subjective test methods may be related with (e.g., from subject to subject or test laboratory to test laboratory) can be reduced based on a proper method design, yielding reliable and valid test results. There are a number of ITU Recommendations on how to properly design subjective tests, as summarized in clause 9.

Although objective assessment methods will lead to a certain level of reliability (i.e., the same quality estimate for a given input whenever it is applied to that input), their validity and

reliability may be limited. Such model limitations are typically reflected in the corresponding Recommendations (see clause 9 for a summary of objective models related ITU Recommendations). Within the range of system configurations, the objective methods that have been designed for, and proven to provide, valid and reliable quality predictions, they can be used as a replacement for or a complement to subjective tests.

All methods can be classified according to how well diagnostic information can be obtained:

- Glass-box approaches enable insight into the internal functionality of the system. Such insight, in turn, may give access to diagnostic information. Hence, the source of a certain quality level may be identified and the relation between system configuration and quality can be observed. For example, parameter-based models are typically glass-box models, since they build on a 'diagnostic' description of the system.
- Black-box approaches do not enable insight into system internals. Hence, with black-box approaches direct indication of the source of low quality is not available. Certain signal-based models are typical examples. For black-box methods, the system under test is described only by its input and output. Based on additional processing of the results, diagnostic information may be inferred.

Current objective quality assessment techniques cannot necessarily be classified in a strict manner according to this black-box/glass-box dichotomy. Instead, different methods enable more or less diagnostic information to be obtained. Sometimes the terms 'white box' and 'greybox' are used in the literature in these cases ([b-Möller, 2005], pp. 86-89).

6.1 Subjective assessment

According to [b-Jekosch, 2005], a (speech) quality test is defined as:

"[a] routine procedure for examining one or more empirically restrictive quality features of perceived speech with the aim of making a quantitative statement on these features" ([b-Jekosch, 2005], p. 91).

Using this definition, subjective test methods can be broadly categorized as utilitarian and analytic test methods on the basis of whether the totality or a certain subset of perceived quality features is to be investigated (analytical), or whether single individual features or integral quality are to be measured (utilitarian).

Utilitarian test methods employ a unidimensional quality rating scale, so that a direct comparison between the quality of different systems can be made. In turn, analytical methods are related with a multidimensional analysis of test stimuli: They aim at identifying and/or quantifying some or all perceptual features underlying quality.

In order to capture the entire range of underlying auditory features, and to ensure a common understanding of these features, the reliability and validity of analytical methods can be significantly improved when trained listeners are used (e.g., [b-Mattila, 2001]; [b-Möller, 2000], pp. 105-120). In turn, tests of the utilitarian type are typically carried out with naive subjects. These should ideally represent the user group at which the system or service under investigation is aimed (e.g., age range, social and cultural background, etc.).

It should be noted that this utilitarian-analytical categorization of auditory test methods can be considered as being of utilitarian type itself, as it employs a unidimensional view on test methods. A multidimensional, that is, analytical view on speech quality assessment methods was provided by [b-Jekosch, 2005] (pp. 105-111).

Utilitarian methods have three main goals: "To be reasonably efficient in test administration and data analysis, to measure speech quality on a unidimensional scale, and to have good reliability in the test method" ([b-Quackenbush, 1988], pp. 15-16).

All test methods can be classified according to the categorization generally applicable to psychometric methods (after [b-Möller, 2000], pp. 48-49):

– The applied scaling method and scale level, with the scale levels being (after [b-Stevens, 1951], p. 25):
  1) Nominal (identifying equality).
  2) Ordinal (identifying greater or less).
  3) Interval (identifying equality of intervals or differences).
  4) Ratio (identifying equality of ratios).
– The presentation method used, such as:
  • Adjustment, i.e., the stimulus is changed to meet a certain description or judgment.
  • Constancy, i.e., one or several stimuli are presented per trial and a description or judgment is given (single stimulus, paired comparison, etc.).
– The test 'modality':
  • Passive, e.g., listening or viewing.
  • Active, e.g., talking or moving.
  • Interactive, e.g., in a conversation.
– The instance at which quality is being judged:
  • Judgment after stimulus/stimuli presentation.
  • Continuous judgment during stimulus/stimuli presentation (e.g., in case of time-varying degradations).

6.2 Objective assessment

In order to reduce the necessity for time-consuming and costly perception tests to measure the quality of networks and systems, much effort has been spent on the development of alternative, instrumental, i.e., 'objective' methods. It has to be noted, however, that the speech quality of today's telecommunication systems cannot readily be determined solely on the basis of basic signal measures like the signal-to-noise ratio. Although quality may be correlated with different instrumentally measurable signal characteristics, it is typically not possible to establish a simple relationship between these instrumentally measurable magnitudes and the quality perceived by a user of the system.

To account for the more complex interdependencies, different types of quality estimation and prediction models have been developed. Each model has its proper domain of application, and range of system or service conditions it has been designed for. Consequently, there is no universal quality model that can be applied in all circumstances. In order to better determine what model to choose for a given application, the models can be categorized according to different criteria (adapted from [b-Möller, 2002]):

– What the application is aimed at (network planning, codec optimization, service monitoring, etc.).
– The predicted quality features (integral quality – audio, video, audiovisual, speech, etc.), intelligibility, conversational aspects, etc.
– The network components and configuration under consideration (entire connection end-to-end, codecs, etc.).
– The respective model input parameters (measured parameters like noise levels, information extracted from protocol headers like loss rates, entire signals, etc.).

– The way in which the input information is obtained:
  - By using passive measurements, which do not interfere with the running system or service, but involve an observation of the system or service during its normal operation.
  - By using active measurements, which involve dedicated measurement signals or data being processed by the system, and evaluating the transmitted result.
  - By estimation, e.g., during the network or service planning phase.
– The level of interactivity of the service aspects that can be assessed with the model, related with:
  - a passive perception (listening, viewing, etc.);
  - an active usage (talking, moving);
  - an interactive usage (e.g., for a conversation).
– The extent to which psychophysical knowledge or rather purely empirical data have been incorporated.

7 Target services

This Recommendation gives guidelines for QoE assessment of various telecommunication services mainly utilizing audio and visual media.

7.1 Audio

– Conversational voice and voice messaging:

  Speech communication services such as mobile telephony and VoIP, as well as conventional PSTN and ISDN services, are important targets of this Recommendation. The speech bandwidth can be either narrowband (300-3400 Hz) or wideband (100-7000 Hz).

– Streaming/on-demand audio:

  Audio delivery services provide audio content through telecommunication networks. They can be either download based or streaming based. Content is usually, but not limited to, musical material. Audio bandwidth can be from narrowband to full band (20-20000 Hz), although usually either superwideband (50-14000 Hz) or full band (20-20000 Hz) is used.

7.2 Video

– Videophone/videoconferencing:

  Audiovisual communication refers to two-way interactive communication using speech and video signals, such as video telephony and videoconferencing. The bandwidth of speech signals can be either narrowband (300-3400 Hz) or wideband (100-7000 Hz). Video resolution can range between QCIF and HDTV.

– Streaming/on-demand video:

  Video delivery services provide video content (normally with audio signals) through telecommunication networks. They can be either download based or streaming based. Typical content includes movies, documentaries, sports, news, and advertisement, for example. Audio bandwidth can be from narrowband (300-3400 Hz) to full band (20-20000 Hz), although usually either superwideband (50-14000 Hz) or full band (20-20000 Hz) is used. Video resolution can be between QCIF and HDTV.

7.3 Data

– Data communication services targeted in this Recommendation are primarily web browsing.

8 Measurement approaches for different scenarios

Basically, there are five types of objective quality assessment methodologies. Depending on the application, these can be divided into media-layer models, packet-layer models, bitstream models, hybrid models and planning models.

Objective quality assessment models can be used for several purposes:

– Planning: "Planning" refers to estimating the perceived quality of services of networks/systems before they are implemented. Since it is not used in a real-time environment, no real-time inputs are required to the objective model.

– Lab-testing: "Lab-testing" refers to estimating the perceived quality of services of networks/systems in the laboratory while the equipment is being developed.

– Monitoring: "Monitoring" refers to estimating the perceived quality of services of networks/systems that are operational. Necessary information is collected from the network and analysed to reflect the degradation of the quality experienced by users.

8.1 Media-layer models

Media-layer models (Figure 8-1) use the actual media signals (audio/video) as their input, and may take into account codec compression and channel characteristics. They use complex perceptually-based psycho-physical models to estimate QoE, by comparing (FR/RR) the output (degraded) signal to the input (clean) signal or just analysing the output (degraded) signal (NR).

The main applications of FR models are QoE assessment in the laboratory, for example, codec comparison/optimization, since it uses not only the degraded (received) signal but also the original (transmitted) signal to estimate QoE. Conversely, RR/NR models can be applied to QoE monitoring at the mid-point or end-points of the IPTV network.

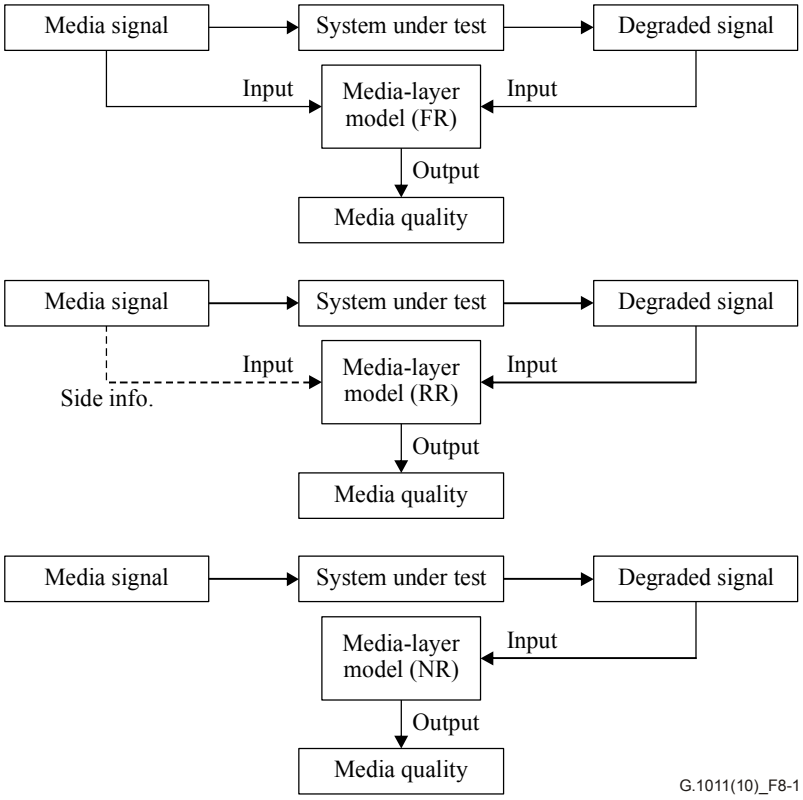

Figure 8-1 – Media-layer models

8.2 Packet-layer models

Packet-layer models (Figure 8-2) utilize only packet header information for QoE prediction. Because they do not parse the packet payload information, it is difficult to incorporate in such models aspects of QoE which are related to the media content, although they have a very light measurement of computational efficiency. These models are primary utilized as network probes at the mid-point or end-points of the network.

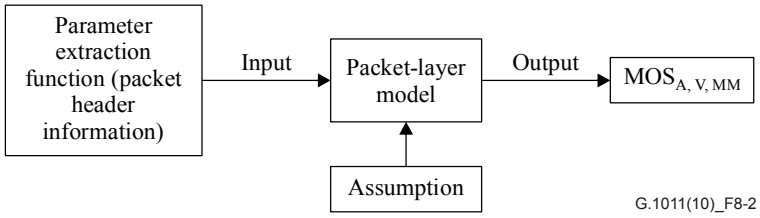

Figure 8-2 – Packet-layer model

8.3 Bitstream-layer models

A bitstream model (Figure 8-3) takes not only the encoded bitstream information, but also the packet header information as its input. Therefore, such a model can be viewed as a combination model between media-layer models and packet-layer models.

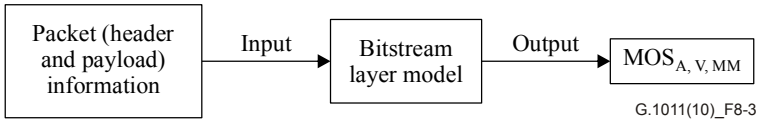

Figure 8-3 – Bitstream-layer model

Since the bitstream-layer model uses only the received packet information (degraded signal), it can be applied to QoE monitoring at the mid-point or end-points of the IPTV network.

8.4 Hybrid models

A hybrid model (Figure 8-4), just as its name implies, is the combination of the previously mentioned models. It employs as much information as possible to predict QoE. An example of these models is shown below:

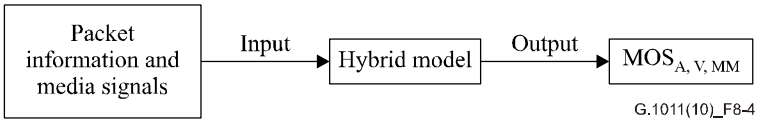

Figure 8-4 – Example of a hybrid model

8.5 Planning models

The input for planning models (Figure 8-5) includes the quality planning parameters of networks or terminals. It usually requires prior knowledge about the system under test. Such models can be applied to network planning and terminal/application design.

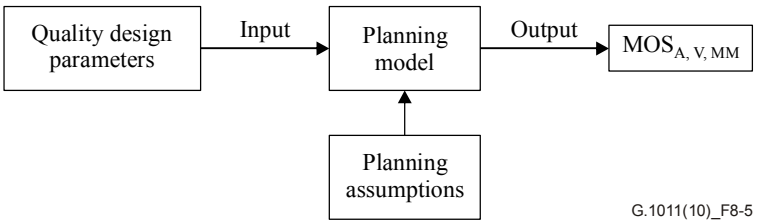

Figure 8-5 – Planning model

Standard examples of such models are [ITU-T G.107] for speech and [ITU-T G.1070] for videophone.

9 Taxonomy

Existing QoE assessment standards are categorized in Table 9-1. Appendix I provides information on planned standards.

Table 9-1 – Current ITU Recommendations for QoE assessment

Application	Media	Conversational (CONV)/Non-conversational (NONCONV)	Subjective test methodology	Objective test methodology		
				Model	FR/RR/NR	Primary usage
Telephony	Speech	NONCONV	[ITU-T P.800] [ITU-T P.830] [ITU-T P.835]	[ITU-T P.862] + [ITU-T P.862.1] (NB) [ITU-T P.862.2] (WB)	FR	LAB, MON
				[ITU-T P.563] (NB) [ITU-T P.564] (NB/WB)	NR	MON
		CONV	[ITU-T P.800] [ITU-T P.805]	[ITU-T G.107] (NB)	NR	PLN
				[ITU-T P.561] + [ITU-T P.562] (NB/WB)	NR	MON
Video-telephony	Multimedia (Note)	CONV	[ITU-T P.920]	[ITU-T G.1070] (NB/WB)	NR	PLN
Video-streaming (Mobile TV/IPTV)	Video	NONCONV	[ITU-T P.910] [ITU-T J.140] [ITU-R BT.500-12]	[ITU-T J.144] (SDTV) [ITU-T J.247] (QCIF, CIF, VGA)	FR	LAB, MON
				[ITU-T J.249] (SDTV) [ITU-T J.246] (QCIF, CIF, VGA)	RR	MON
	Audio	NONCONV	[ITU-T P.830] [ITU-R BS.1116-1] [ITU-R BS.1285] [ITU-R BS.1534-1]	[ITU-R BS.1387]	FR/RR	MON/PLN
	Multimedia	NONCONV	[ITU-T P.911]			
Web-browsing	Data			[ITU-T G.1030]	NR	PLN
NOTE – For individual media (i.e., speech and video), the Recommendations used in telephony and video-streaming applications are applicable.						

Appendix I

(This appendix does not form an integral part of this Recommendation)

In order to give the users of this Recommendation an overall view of existing and ongoing projects in ITU, information on existing and planned QoE assessment standards is also provided in Table I.1.

Table I.1 – Existing Recommendations and current work in ITU on QoE assessment

Application	Media	Conversational (CONV)/Non-conversational (NONCONV)	Subjective test methodology	Objective test methodology		
				Model	FR/RR/NR	Primary usage
Telephony	Speech	NONCONV	[ITU-T P.800] [ITU-T P.830] [ITU-T P.835]	[ITU-T P.862] + [ITU-T P.862.1] (NB) [ITU-T P.862.2] (WB) (Notes 2 and 3)	FR	LAB, MON
				[ITU-T P.563] (NB) [ITU-T P.564] (NB/WB)	NR	MON
		CONV	[ITU-T P.800] [ITU-T P.805]	[ITU-T G.107] (NB)	NR	PLN
				[ITU-T P.561] + [ITU-T P.562] (NB/WB) (Note 4)	NR	MON
Video-telephony	Multimedia (Note 1)	CONV	[ITU-T P.920]	[ITU-T G.1070] (NB/WB)	NR	PLN
Video-streaming (Mobile TV/IPTV)	Video	NONCONV	[ITU-T P.910] [ITU-T J.140] [ITU-R BT.500-12]	[ITU-T J.144] (SDTV) [ITU-T J.247] (QCIF, CIF, VGA) (Note 5)	FR	LAB, MON
				[ITU-T J.249] (SDTV) [ITU-T J.246] (QCIF, CIF, VGA)	RR	MON
				(Note 6)	NR	MON
	Audio	NONCONV	[ITU-T P.830] [ITU-R BS.1116-1] [ITU-R BS.1285] [ITU-R BS.1534-1]	[ITU-R BS.1387]	FR/RR	MON/PLN
	Multimedia	NONCONV	[ITU-T P.911]	(Notes 7 and 8)	NR	MON
				(Note 9)	NR	PLN

Table I.1 – Existing Recommendations and current work in ITU on QoE assessment

Application	Media	Conversational (CONV)/Non-conversational (NONCONV)	Subjective test methodology	Objective test methodology		
				Model	FR/RR/NR	Primary usage
Web-browsing	Data			[ITU-T G.1030]	NR	PLN

NOTE 1 – For individual media (i.e., speech and video), the Recommendations used in telephony and video-streaming applications are applicable.
NOTE 2 – Objective listening quality assessment model (NB/WB/SWB).
NOTE 3 – Objective noise reduction assessment model (NB/WB/SWB).
NOTE 4 – Objective conversational voice quality assessment model.
NOTE 5 – Objective perceptual multimedia video quality measurement of HDTV (HD).
NOTE 6 – Hybrid perceptual/bit-stream models for objective video quality measurements.
NOTE 7 – Non-intrusive parametric model for the assessment of performance of multimedia streaming.
NOTE 8 – Non-intrusive bit-stream model for the assessment of performance of multimedia streaming.
NOTE 9 – Opinion model for video and audio streaming application.

Bibliography

[b-ITU-T P.862.3] Recommendation ITU-T P.862.3 (2007), *Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2*.

[b-DSL Forum] DSL Forum TR-126 (2006), *Triple-play Services Quality of Experience (QoE) Requirements*.

[b-Guilford, 1954] Guilford, J.P. (1954), *Psychometric Methods. McGraw Hill, New York, NY*.

[b-Möller, 2005] Möller, S. (2005), *Quality of Telephone-based Spoken Dialogue Systems, Springer, Boston MA*.

[b-Jekosch, 2005] Jekosch, U. (2005), *Voice and Speech Quality Perception – Assessment and Evaluation, Springer, Berlin*.

[b-Mattila, 2001] Mattila V.V., (2001), *Perceptual Analysis of Speech Quality in Mobile Communications, Vol. 340. Doctoral dissertation thesis, Tampere University of Technology, Tampere*.

[b-Möller, 2000] Möller, S. (2000), *Assessment and Prediction of Speech Quality in telecommunications. Kluwer Academic Publ., Boston, MA*.

[b-Quackenbush, 1988] Quackenbush, S.R. (1988), *Barnwell III, T.P., and Clements, M.A., Objective Measures of Speech Quality. Prentice Hall, Englewood Cliffs, NJ*.

[b-Stevens, 1951] Stevens S.S. (1951), *Handbook of Experimental Psychology. Wiley, New York*.

[b-Möller, 2002] Möller, S., Raake, A. (2002), *Telephone Speech Quality Prediction: Towards Network Planning and Monitoring Models*.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2010