Question 19/12 – Objective and subjective methods for evaluating perceptual audiovisual quality in multimedia and television services

(Continuation of Question 19/12)

1   Motivation

In digital transmission systems, the perceptual quality of the audiovisual signal is influenced by a number of interacting factors, such as source coding and compression, bit rate (fixed or variable), delay, bandwidth, synchronization between the media, transmission impairments, and many others. New services that use IP, wireless, mobile, NGN, etc. are providing ubiquitous access for multimedia services. Audiovisual multimedia cover multichannel audio, television, and 3D video applications including interactive ones, in addition to other applications such as videoconferencing, personal computer desktop conferencing, interactive educational and training services, groupware, interactive gaming, and videotelephony. This Question focuses on perceptual impacts of compression, transmission, and decompression on audiovisual quality of these multimedia services and applications.

The effect of the source and display is particularly important and necessary for the case of 3DTV and high-dynamic range (HDR) displays, as both these technologies are not mature and still introduce quality problems. Display technologies are evolving from 2D to 3D, high-definition to ultra-high definition, low dynamic range to wide-gamut and high-dynamic range displays. In particular, HDR images are currently typically displayed on low-dynamic range (LDR) displays because of the limited availability of HDR displays. In order to visualize HDR images on LDR displays, tone mapping is necessary and this creates information loss that can deteriorate the quality and details of the HDR image. Recently, HDR displays have appeared on the market but they use internal processing that can affect the video quality. 3DTVs exhibit crosstalk to various degrees and can impact negatively the viewing experience. For these new technologies, the quality impact of the display and transmission (or camera, production and transmission) cannot always be separated. Although bandwidths available in cable transmission are well suited for ultra-high definition television (UHDTV), maintaining adequate video quality still represents a challenge. ITU-R has recommended methods for the subjective assessment of picture quality (e.g. BT.500-13, BT.1788, BT.2021). There is a need to confirm that those subjective assessment methods and set-up requirements (including selection of the display, settings/calibration of the display, viewing distance, angle, luminance levels etc.) are equally applicable to the case of next-generation visual media, such as television transmission on digital or mixed analogue-digital chains, 3D, HDR and UHDTV images.

Concerning the measurement of the overall quality of experience (QoE), it includes not only a single impairment of each mono-media but also inter-media relation and response time of user operation. There is a need to identify the group of parameters that can provide objective measurement of the overall QoE and continuous in-service monitoring and control of it along the transmission chain.

In order to develop the two-way measurement techniques required for conversational applications, a basis in one-way audio and video quality evaluation must first be defined and validated. Considering the spread of broadband connections to business and the home, the bandwidths will support both low resolution, e.g. quarter video graphics array (QVGA), and standard, high and ultra-high definition imagery. As an example, audio multimedia applications currently range from audio for narrow-band applications, e.g. video telephony, to the enhanced audio contained in 7.1 surround sound systems for interactive gaming. In the

future, HDR, 3D programmes and 3D games are expected to become more widely available. Objective and subjective methods for assessing the perceptual quality of these media services are needed, particularly those relating to transmission.

Objective methods: Current objective quality measuring techniques for audiovisual applications do not correlate to the user opinion on the perceived audiovisual quality with the desirable accuracy. It is therefore necessary to identify objective techniques for measuring the various individual and combined effects of factors such as digital compression, transmission, storage, and others on the perceived quality of audiovisual systems. It is also important to verify that these techniques are meaningful by correlating proposed objective tests with corresponding subjective test data.

Subjective methods: There is a need to continue to develop new subjective methods to address new audiovisual services. The perceived quality depends on the kind of application and on the tasks the applications are used for. For example, in a free conversation through a videophone or videoconferencing application, the perceived quality may primarily depend on delay, lip-synchronization and audio quality, while in a mainly one-way application like remote-teaching the perceived quality could be primarily related to the quality of graph and low motion picture sequences.

These studies include the maintenance of and enhancements to existing Recommendations, and the development of new Recommendations as needed.

Much of the work on this Question (and its predecessors) was and will be done in cooperation with the video quality experts group (VQEG).

2   Question

Study items to be considered include, but are not limited to:

–   Interaction of media: What subjective and objective measurement methods should be used to evaluate end-to-end quality of each medium (e.g. video, audio, television, 3D video) and the interactions between the media, with particular attention to the audiovisual quality assessment of systems used for videoconferencing/videotelephony and other interactive multimedia services? What are the quality levels that can be defined by objective or subjective methods in different applications (or tasks) taking into account the interactions between media?

–   Transmission errors: What objective methods could be used for in-service measurement and monitoring of transmission systems for such multimedia services in the presence of transmission errors? What new subjective measurement methods should be used for the evaluation of transmission quality of real time audiovisual services by expert observers resulting in the identification of specific flaws in the transmission equipment or environment? What procedures should be used, and which dimensions, transforms, and partial or differential signals should be viewed by experts to evaluate specific impairments of real time audiovisual services? What objective and subjective methods can be used to evaluate audiovisual signals with time-varying quality?

–   Impairment characterizations: Among the most significant factors (e.g. spatial resolution, temporal resolution, colour fidelity, audio and visual artefacts, media synchronization, delay, cross-talk etc.) affecting the overall quality of multimedia services, what objective and subjective methods assess the extent of or can differentiate between these factors? How can the mutual interaction between these factors be objectively and subjectively measured with respect to their influence on

overall audiovisual quality? For what applications can the assessment methods be shown to be useful and robust over a range of conditions? What kind of artificial impairment generator would be useful for subjective or objective methods?

– Evaluation of specific services: What assessment methods (objective and subjective) can be used to characterize the quality effects of multipoint distribution for interactive communication and other new audiovisual services such as remote monitoring, interactive gaming, and mobile audiovisual communication?

– Test methodologies: What subjective methods and assessment tools are required to fully describe perceived visual or audiovisual impairments in terms of measurable system parameters? What kind of references should be used in subjective tests? What methods can be used to measure the video quality of 3D video? What new subjective methods are needed when analysing new applications and usage scenarios? What kind of service or application design is needed to minimize visual fatigue in 3D video applications? What methods can be used to measure the visual fatigue level introduced into a 3D video signal by the source content (e.g. amount of motion, depth of field), compression and transmission?

– Combination of test results: In some cases it may be useful to combine objective measures (e.g. video measures, audio measures, media synchronization) to provide a single figure of merit. In this regard, which objective measures and/or techniques should be combined, and in what manner, so that the figure of merit correlates satisfactorily with subjective test results?

– Test sequences: While the library of test sequences has increased greatly recently (e.g. www.cdvl.org), there is still a need for more test sequences, especially those with audio included and 3D. Which audiovisual test material (e.g. audiovisual test sequences, 3D video) can be standardized for subjective and objective evaluations? In addition to the definitions of SI and TI in P.910, which criteria (objective and/or subjective) should be used to characterize and classify multimedia test material?

– Validation and applicability of objective methods: There are three basic methodologies of objective picture quality measurement. Full-reference (FR) uses the full bandwidth video input. Reduced-reference (RR) uses lower bandwidth features extracted from the video input. No reference (NR) has no information about the video input. What objective methodology should be used for different multimedia applications? What subjective methods should be used to validate each of the three basic objective methodologies? How can hybrid perceptual/bitstream (hybrid) methodologies use information about the encoded bit-stream to supplement FR, RR or NR methodologies?

– What enhancements to existing Recommendations are required to provide energy savings directly or indirectly in information and communication technologies (ICTs) or in other industries? What enhancements to developed or new Recommendations are required to provide such energy savings?

– What are the quality requirements for transmission of UHDTV?

– Are the current methods recommended for subjective assessment of digital picture quality also applicable to scenarios where the display is not transparent, such as in 3DTV or HDR images? Are the current quality assessment methods applicable to ultra-high definition television?

– How should the impairment introduced by the display be taken into account in evaluation of the viewing experience?

–   How should the impairments introduced by the transmission chain be taken into account, such as those introduced by digital or mixed analogue-digital television transmission chains?

–   How should the impairment introduced by the (stereo-) camera be taken into account in evaluation of the viewing experience?

–   What objective methodology can be used to jointly analyse the perceptual quality of the entire stream, including the quality of both the camera and the display?

–   How should the objective measurement of impairments introduced by digital or mixed analogue-digital transmission networks be carried out?

–   Which network parameters should be used to provide objective measurement of the overall QoE and should be the basis for continuous in-service monitoring along the transmission chain both for digital and for mixed analogue-digital television transmission?

–   What perceptual image/video quality assessment methods can be used to determine which tone-mapping operator maintains best the visual information of an HDR image or produces the highest-quality LDR image? What perceptual image/video quality assessment methods can be used to assess the quality of HDR content?

–   What methods can be used to measure the visual fatigue in 3D video from the video capture, rendering and display?

3   Tasks

Tasks include, but are not limited to:

–   Quality assessment in multimedia services requires on the one hand the continuous updating of Recommendations under the responsibility of Study Group 12 and also the definition of new task oriented/application-dependent evaluation and subjective methods for the combined evaluation of audio and video signals.

–   A new Recommendation utilizing expert viewers is expected. Three Recommendations defining objective methods for assessing audiovisual quality in multimedia services are expected to be approved.

–   Initial work on quality assessment of interactive gaming applications will result in a new Recommendation.

–   Maintenance and revision of Recommendations on 3D subjective methods.

–   It is anticipated that new Recommendations will address: methods to characterize and select appropriately 3D displays for subjective evaluation of 3D picture quality; methods for HDR and UHDTV quality evaluation and methods to assess/characterize the impact of non-transparent displays on viewing experience.

An up-to-date status of work under this Question is contained in the SG12 work programme http://www.itu.int/ITU-T/workprog/wp_search.aspx?q=19/12,

4   Relationships

WSIS Action Lines

–   C2

Sustainable Development Goals

–   9

Recommendations

–　　　P- and J-series

Questions

–　　　14/12

Study Groups

–　　　ITU-T SG9, SG13, SG15, SG16

–　　　ITU-R SG6

Other bodies

–　　　ITU IRG-AVQA, VQEG, IETF and regional standardization bodies (e.g. ATIS)