QUESTION ITU-R 102-2/6

Methodologies for subjective assessment of audio and video quality

(1999-2011-2014)

The ITU Radiocommunication Assembly,

considering

*a)* that Recommendations ITU-R BS.1116, ITU-R BS.1283, ITU-R BS.1284, ITU-R BS.1285 and ITU-R BT.500, and Report ITU-R BT.1082, have established primary methods for the subjective quality assessment of audio (including multichannel presentation) or visual (including stereoscopic presentation) systems respectively;

*b)* that Recommendation ITU-R BS.1286 has established primary methods for the subjective quality assessment of audio in the presence of high quality television image;

*c)* that the perceptual interaction between the audio and visual modalities can affect their mutual qualities and the overall perceived quality;

*d)* that existing methods for the subjective assessment of audio quality are sometimes inadequate for audio systems with accompanying visual presentation;

*e)* that there are no generally applicable methods for the subjective assessment of visual quality with accompanying audio presentation;

*f)* that there are no known methods for the subjective assessment of both audio and visual presentation simultaneously;

*g)* that a wide range of multimedia systems, including digital multimedia video information systems (VIS) for collective, indoor and outdoor viewing, comprise audio-visual presentations. Such systems have a wide range of applicability in terms of:

– terminal types (standard and high definition television, computer terminals, (mobile-) multimedia terminals);

– applications (entertainment, education, information services);

– presentation quality (low, intermediate, high);

– presentation environments (domestic, office, outdoor, professional);

– delivery systems (internet, mobile networks, satellite, broadcast);

*h)* that the multiscreen technology is used in the broadcasting and multimedia information applications providing simultaneous presentation of several different images on the same screen;

*i)* that optical head-mounted displays (e.g. video glasses)[[1]](#footnote-1) have been implemented for the reception of TV broadcasting programmes and personal multimedia information,

decides that the following Questions should be studied:

1 What are the quality attributes for audio-visual perception?

2 How the context dependent quality balance between audio and visual presentation[[2]](#footnote-2)2 should be considered?

3 What are the subjective test methodologies[[3]](#footnote-3)3 required for different applications and quality levels for:

– audio-visual presentation?

– visual presentation in the presence of audio (audio presentation at a constant quality level)?

– audio presentation in the presence of visual (visual presentation at a constant quality level)?

4 How could such methodologies be used as criteria to identify quality attributes that are important for different application areas of audio-visual presentation, including VIS?

5 How could they be used to express quality requirements for audio and visual modalities for different application areas and to assess their optimization?

6What approaches could be used for image quality assessment when applied to multiscreen and optical head-mounted display (e.g. video glasses)?

further decides

1 that the results of the above studies should be included in (a) Recommendation(s);

2 that the above studies should be completed by 2015.

Category: S2

1. The personal displays utilizing optical glasses can be used with the PCs, smartphones and other devices. They can be used for the reception of TV broadcasting programmes and personal multimedia information at any time, at any place and in motion. [↑](#footnote-ref-1)
2. 2 Examples might include the importance of synchronization between audio and visual presentation for talking head applications, changing focus in sports transmissions (from fast moving objects, where video is most important, to the cheering crowd after certain event, where the audio catches the attraction). [↑](#footnote-ref-2)
3. 3 This should include, for example, the harmonization of grading scales employed in audio and visual testing at present (refer to present ITU-R BS and BT and ITU-T Recommendations),   
   test environments, viewing and listening distances, training procedures, etc. [↑](#footnote-ref-3)