QUESTION ITU-R 139/6

Methods for rendering of advanced audio formats

(2015)

The ITU Radiocommunication Assembly,

considering

*a)* that there is an increasing interest in the production of sound and television programmes in advanced sound systems, which can provide a listening experience that matches the enhanced viewing experience provided by image production in HDTV (see Recommendation ITU-R BT.709) and UHDTV (see Recommendation ITU-R BT.2020);

*b)* that Recommendation ITU-R BS.2051 specifies advanced sound systems that can provide an enhanced listening experience to a properly equipped radio or television audience;

*c)* that Recommendation ITU-R BS.1909 specifies as typical viewing environments theatre and large theatre environments as well as large-to-average size room environments, and mobile such as in-car or personal environments;

*d)* that consistency in sound production requires consistency in the sound reproduction system that is employed in the production environment and that this implies the need for consistency in the reproduction of the advanced sound system in the production chain;

*e)* that the rendering system that creates the loudspeaker signals from the advanced sound system signals is a critical component to provide the needed consistency in reproduction,

further considering

*a)* that a description of a baseline[[1]](#footnote-1) renderer should be complete and self-contained. Ideally it abstracts from implementation details and provides those by using a reference implementation;

*b)* that the description should clearly describe the operations and signal processing to be carried out, based on the incoming audio data, metadata and the local metadata which configure the rendering process and not contain any ambiguities. Extensions to the specification can allow points of enhancement but this is not part of the baseline renderer specification;

*c)* that if a file format does exist, this can be referred to in terms of parameters and storage, but in general the specification should not be linked to specific implementations of such parameters in aforementioned file format;

*d)* that a baseline rendering should be able to support all speaker setups as proposed in ITU‑R BS.2051,

decides that the following Questions should be studied

1 What are the requirements for a baseline renderer for use in the production of advanced sound programmes and quality evaluation?

2 What is the specification of a baseline renderer that is satisfactory for use in the production of advanced sound programmes and quality evaluation?

3 What signal processing and metadata inputs (environmental metadata, content related metadata) are required for satisfactory operation of a baseline renderer?

4 What algorithm should be used to derive the speaker signals based on all possible input formats (object-based, channel-based, scene-based and combinations thereof) according to Recommendation ITU-R BS.2051?

further decides

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

2 that the above studies should be completed by 2016.

Category: S1

1. A renderer converts a set of audio signals with associated metadata to a different configuration of audio signals and metadata, based on the provided content metadata, and local environmental metadata. A baseline renderer is an instance of a renderer which is used for quality evaluation purposes and in the programme production process. It being clearly defined does allow a comparison with other possible instances. It does not necessarily offer the best possible quality of the auditory scene and may not support all possible metadata, but can deliver a rendition which will preserve the artistic intent for a defined set of rendering conditions. [↑](#footnote-ref-1)