**Source: Orange**

**Title: On binaural rendering**

## Document for: Agreement

## Agenda Item: 7.5

1. **Introduction**

The IVAS design constraints [1] contain provisional text on binaural rendering, as follows:

|  |  |
| --- | --- |
| **…** | … |
| **Interface for binaural rendering** | The IVAS decoder/renderer shall provide an interface to provide [HRTF/BRIR] data for binaural rendering. The interface is [tbd].[Editor’s Note: There was some support for this interface to follow the SOFA SimpleFreeFieldHRIR convention - See AES69-2015].The IVAS decoder/renderer shall provide an API to provide [TBD scene displacement data].[The IVAS decoder/renderer shall support direct headphone presentation.] |
| **Control Data For Binaural Audio Rendering** | The IVAS decoder/renderer shall support the following control data for binaural audio rendering:[HRTF/BRIR] data for binaural rendering on command line interface. The format for [HRTF/BRIR] data is [tbd].[Editor’s Note: There was some support for this interface to follow the SOFA SimpleFreeFieldHRIR convention - See AES69-2015].[TBD scene displacement data]. |
| **…** | … |

The text is provisional (with brackets). Furthermore, the two boxes entitled ‘Interface for binaural rendering’ and ‘Control Data For Binaural Audio Rendering’ are obviously redundant and this could be confusing.

In this Tdoc we make some text proposals to progress the related IVAS requirements and address the issue of HRTF/BRIR filters to be used by default for testing.

1. **Proposed text updates**

The spirit of current IVAS-4 draft requirements on binaural rendering is to require that the IVAS decoder/renderer provides an external interface with the SOFA format and that IVAS natively supports binaural rendering. An initial proposal by Orange in [2] led to the inclusion of an Editor’s note FFS [3] on this aspect. This was later revisited in [4] with some discussion on the importance of individualized/personalized HRTFs and existing APIs for binaural rendering (SOFA, MPEG-H 3D Audio). Text updates were included in the box ‘Interface for binaural rendering’ in IVAS-4 [5]. Head-rotation aspects were included in IVAS-4 [7] based on proposal from Orange in [6], with some rewording of ‘rotation angles’ to ‘displacement data’. Direct headphone presentation was discussed in [8] and including provisionally in IVAS-4 [9|; this was revisited in [10] and further reflected in IVAS-4 [11]. The redundancy with two boxes ‘Interface for binaural rendering’ and ‘Control Data For Binaural Audio Rendering’ was introduced in [11] as a result of online editing.

Based on this SA4/IVAS history, we suggest text updates on several aspects:

* From early discussions there has been agreement on the idea of supporting an interface for binaural rendering, however it may not be ideal to require direct SOFA support for this interface. Instead, we propose to follow an approach inspired from MPEG-H 3D Audio where the interface shall be documented by IVAS codec candidate(s). Conversion tools from SOFA to the particular interface format are required. This will allow using equivalent binaural filters for reference and CuT conditions in IVAS testing, however the IVAS codec candidate(s) may benefit from a representation that is pre-processed offline and defined in terms of low-level parameters as in MPEG-H 3D Audio.
* An IVAS decoder/renderer with no predefined binaural filters is less usable, and in any case for testing purposes a default set of filters will be needed to process both reference and CuT conditions for experiments with binaural output and listening over headphones. It is unlikely that individual or personalised filters get used for selection subjective tests, so a set of generic filters is needed. We propose to define a default set of HRIR and BRIR filters for testing and to require that this default set is also natively used by IVAS when no filter set is provided through the binaural rendering interface; this will facilitate usage of the IVAS codec, designing test vectors, etc. See section 3 for a proposed set of HRIRs.
* The wording ‘scene displacement’ is general, and it may include aspects such as scene modifications (e.g. translations), we propose to be more specific and refer ‘head-tracking data’.

Proposed text updates are captured in the boxes updated below:

|  |  |
| --- | --- |
|  |  |
| **Control Data For Binaural Audio Rendering** | The IVAS decoder/renderer shall support the provision of HRIR / BRIR filter sets as control data for binaural audio rendering. The format of HRIR / BRIR filter sets shall be documented in IVAS candidate deliverables.Editor’s Note: potential requirements on conversion tools from HRIRs/BRIRs in SOFA format into the internal format are to be addressed in IVAS-6The IVAS decoder/renderer shall support a default HRIR / BRIR set for binaural rendering.Editor’s Note: default set to be specified in Annex of IVAS-4The IVAS decoder/renderer shall support head-tracking data as control data for binaural audio rendering. The format of the head-tracking data shall be either in quaternion or Euler angle domain [with other components TBD] every [5ms subframe / 20ms frame]Editor’s Note: head-tracking data format to be defined Annex of IVAS-4 |

1. **Proposed HRIR sets for default operation**

For IVAS testing it would be useful to define a set of default HRIRs (dry cases) and a set of default BRIRs (with reverberation). We address here the question of default HRIRs for binaural rendering.

There are many examples of HRTFs available in the field, see for instance the SOFA repository in [12]. Some HRIRs may have issues, such as insufficient spherical resolution, poor quality, licensing constraints, etc. One may consider defining minimum requirements on SOFA inputs to be acceptable for IVAS, however this topic deals with the usage of the IVAS codec and it could be resolved later.

In the present Tdoc we propose a set of HRIRs derived from measurements on one human subject in the context of the BiLi (BinauralListening) Project – see details in [13]. The raw HRIRs were post-processed including interpolation, symmetry enforcing, minimum phase+delay modeling, spectral smoothing to enable complexity reduction (due to impulse response shortening) and improve quality. This proposed set is referred to as the ‘Orange 53’ set.

We provide in attachment this proposed set of HRIRs in SOFA format (SimpleFreeFieldHRIR convention) together with related spherical grid. This set is defined at 48 kHz, and it defines HRIRs at 7658 positions (1-degree step in azimuth, azimuth-dependent grid in elevation). Note that this set of HRIRs is provided in 3GPP SA4 with the following license conditions:

(c) 2014 - 2022 Orange. All Rights Reserved Disclaimer :

This database called 'Orange HRIR BiLi database' is made available under Open Database License (ODbL) whose full text can be found at <http://opendatacommons.org/licenses/odbl/>.

Further, ORANGE reserves the right to revise this database and/or documentation and to make changes from time to time in the content hereof without obligation of ORANGE to notify any person of such revision or change. Use of Materials : ORANGE hereby grants users permission to use materials available therein for any purpose-educational, research or commercial. However, each reproduction of any part of the materials must include the copyright notice.

Other HRIR sets coming from the same BiLi project may also be provided (in SOFA format) if more examples are needed, for instance to consider HRIR personalization aspects. However, we suggest considering the attached proposal as a default HRIR set for the IVAS work item. It may be worth noting that France Radio has been using this HRIR set as a default set to binauralize audio content in their HyperRadio service [14].

Note that we do not propose default BRIRs set in this Tdoc, this is left open for discussion.

**References**

1. 3GPP Tdoc S4-220822, IVAS Design Constraints (IVAS-4), v0.5.0, Source: Editor (Huawei)
2. 3GPP Tdoc S4-170942, On IVAS design constraints and performance requirements, Source: Orange
3. 3GPP Tdoc S4-171036, IVAS Design Constraints (IVAS-4), v0.0.1, Source: Editor (Huawei)
4. 3GPP Tdoc S4-180480, API for binaural rendering (IVAS\_Codec), Source: Orange
5. 3GPP Tdoc S4-180605, IVAS Design Constraints (IVAS-4), v0.0.4, Source: Editor (Huawei)
6. 3GPP Tdoc S4-180823, On IVAS input-output formats and rendering, Source: Orange
7. 3GPP Tdoc S4-180926, IVAS Design Constraints (IVAS-4), v0.0.5, Source: Editor (Huawei)
8. 3GPP Tdoc S4-181127, On IVAS input and output formats, Source: Orange
9. 3GPP Tdoc S4-181218, IVAS Design Constraints (IVAS-4), v0.0.6, Source: Editor (Huawei)
10. 3GPP Tdoc S4-190136, On IVAS input-output formats, Source: Orange
11. 3GPP Tdoc S4-190248, IVAS Design Constraints (IVAS-4), v0.0.7, Source: Editor (Huawei)
12. SOFA repository, https://www.sofaconventions.org/mediawiki/index.php/Files
13. F. Rugeles Ospina, M. Emerit, J. Daniel, “A fast measurement of high spatial resolution head-related transfer functions for the BiLi Project,” in *Proc. ICSA,* Graz, Austria, 2015.
14. <https://hyperradio.radiofrance.com/>