3GPP TSG|WG-4 Meeting #129-e S4-241627

Online, 19-23 August 2024 Revision of S4-241567

**Source:** **Xiaomi Technology,** **Beijing Bytedance Technology Co., Ltd, Huawei Technologies Co Ltd**

**Title:** **Draft WID on Diverse audio CApturing system for Smartphone devics (Da****CAS)**

**Document for: Approval**

**Agenda Item: 17**

3GPP™ Work Item Description

Information on Work Items can be found at <http://www.3gpp.org/Work-Items>
See also the [3GPP Working Procedures](http://www.3gpp.org/specifications-groups/working-procedures), article 39 and the TSG Working Methods in [3GPP TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm)

Title: Draft WID on Diverse audio CApturing system for Smartphone devices

Acronym: DaCAS

Unique identifier:

Potential target Release: Rel-19

# 1 Impacts

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Affects: | UICC apps | ME | AN | CN | Others (specify) |
| Yes |  | X |  | X |  |
| No | X |  | X |  | X |
| Don't know |  |  |  |  |  |

# 2 Classification of the Work Item and linked work items

## 2.1 Primary classification

### This work item is a …

|  |  |
| --- | --- |
|  | Study  |
| X | Normative – Stage 1 |
|  | Normative – Stage 2 |
|  | Normative – Stage 3 |
|  | Normative – Other\* |

**\* Other = e.g. testing**

## 2.2 Parent Work Item

For a brand-new topic, use “N/A” in the table below. Otherwise indicate the parent Work Item.

|  |
| --- |
| Parent Work / Study Items  |
| Acronym | Working Group | Unique ID | Title (as in 3GPP Work Plan) |
| N/A |  |  |  |

### 2.3 Other related Work Items and dependencies

|  |
| --- |
| Other related Work /Study Items (if any) |
| Unique ID | Title | Nature of relationship |
| 770024 | EVS Codec Extension for Immersive Voice and Audio Services | IVAS codec can be used to encode the output audio signals of this WI |
| 830005 | Terminal Audio quality performance and Test methods for Immersive Audio Services | The output will be used in this WI |
| 980008 | Study on Diverse audio Capturing system for End-user Devices | TR document of FS\_DACED are background of this WI |
| 1040021 | EVS Codec Extension for Immersive Voice and Audio Services, Phase 2 (IVAS\_Codec\_Ph2) | Codec can be used to encode the output format audio from this WI |
|  | Terminal Audio quality performance and Test methods for Immersive Audio Services, Phase 2(ATIAS\_Ph2) |  |

# 3 Justification

Following the development of the wireless communication technologies and the increasing diversification of people's lifestyles. Smartphones have become an indispensable part of people’s daily lives, not only as a means of real-time communication, but also as an important device for recording lives, creating contents and enjoying entertainments. On audio service aspect, the real-time communication evolves from narrow band voice to super wide band or even full band audio, mono audio services are increasingly unable to meet people's audio service experience requirements and spatial audio are expected to be the next step-up audio service.

IVAS codec has been standardized in August 2023, it is a single codec with attractive features and performance, it supports several encoder input formats. For the immersive audio content that generated in smartphone devices, the first step is to capture sound field signals and then transform them into appropriate formats data that can be feed into IVAS encoder. Due to their various form factors as described in TR 26.933, the smartphone devices with different immersive audio capture solutions may provide significant different performance of immersive audio signals, therefore this may lead to inconsistencies of the performance and style of the audio service got from different smartphone devices, specifically, the inconsistencies might reduce consumer interests and expectations for immersive audio services.

Based on the conclusion of TR 26.933, although immersive audio capture is indeed feasible, its widespread adoption in current products remains limited, which could potentially slow down the deployment speed of a fully end-to-end IVAS ecosystem, this situation highlights the need for a normative solution.

Standardizing immersive audio capture can ensure consumers experience a consistent performance of audio quality and immersion across different smartphone models and brands. This consistency enhances consumer satisfaction and loyalty. The consistency is especially important for the immersive audio experience in XR scenarios; Secondly, it would facilitate the development of new applications and technologies that utilize immersive audio. There will be a clear set of guidelines to follow which would potentially lead to a greater variety of applications and technologies that utilize immersive audio and allow consumers to seamlessly switch between different devices and applications without experiencing unexpected inconsistencies in audio experience, benefiting both consumers and developers; Thirdly, it could be used for setting requirements and evaluating performance to determine test methods and quality metrics.

The goal is to output normative immersive audio capture capabilities for smartphone devices, the consistencies and high-performance immersive audio data generated from sending side may accelerate the maturation speed of the end-to-end immersive audio solution together with the IVAS codec.

The selected normative solution(s) will introduce (a) mathematic algorithm with corresponding requirements, enables the transformation of compensated microphone signals into expected formats including IVAS input formats.

-The mathematic algorithm is designed to be flexible, catering to a broad spectrum of devices, from current market products to those anticipated in future landscape.

-The requirements incorporate a flexible microphone configuration model and signal characteristics such as directional response, SNR, and frequency response. This requirement will be helpful in the development of immersive audio devices, providing a roadmap for design microphone configuration and tuning microphone signal to meet the requirements of solution.

# 4 Objective

The overall objective of this work item is to develop an immersive audio capabilities solution for smartphones, which can by be used for conversational and non-conversational use case where immersive audio generated and consumed by smartphones. The following objectives should be achieved within the work item:

* immersive audio capture and test technologies for relevant smartphone devices,
	+ The selected solutions are expected to set requirements for compensated microphone signals including flexible microphone configuration model and characteristic (directional response, SNR, and frequency response.
	+ The selected solutions will define mathematic algorithms which convert microphone signals into expected formats which include IVAS input formats.
	+ The selected solutions have expected performance for the converted signals, refer to the requirements in TS 26.131 and TS 26.261.

*The developments under this work item should lead to a set of new specifications. Following 3GPP practice, floating-point xxxx code and associated test vectors should also be part of this set of specifications.*

# 5 Expected Output and Time scale

|  |
| --- |
| New specifications {One line per specification. Create/delete lines as needed} |
| Type  | TS/TR number | Title | For info at TSG#  | For approval at TSG# | Rapporteur |
| TR | 26.xxx | xxx | SA#1xx(D) | SA#1xx(xxx) | xxx |
|  |  |  |  |  |  |

|  |
| --- |
| Impacted existing TS/TR {One line per specification. Create/delete lines as needed} |
| TS/TR No. | Description of change  | Target completion plenary# | Remarks |
| N/A |  |  |  |
|  |  |  |  |

# 6 Work item Rapporteur(s)

xxxxx

# 7 Work item leadership

SA4

# 8 Aspects that involve other WGs

*SA4 will coordinate this work with relevant WGs if necessary.*

# 9 Supporting Individual Members

|  |
| --- |
| Supporting IM name |
| *Xiaomi Technology* |
| Beijing Bytedance Technology Co., Ltd |
| Huawei Technologies Co Ltd |
|  |
|  |
|  |