3GPP SA4 WG4#130 S4-241959

Orlando, USA, 18-22 November 2024 Revision of S4aA240050

**Source:** **Xiaomi Technology,** **Beijing Bytedance Technology Co., Ltd, Huawei Technologies Co Ltd, vivo Mobile Communication Co., Ltd, Nokia**

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

**Document for: Discussion**

**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 |
|  | Normative – Stage 1 |
|  | Normative – Stage 2 |
| X | 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 can 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 |
| 1050113 | Terminal Audio quality performance and Test methods for Immersive Audio Services, Phase 2  (ATIAS\_Ph2) | The output can be used in this WI |

# 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 life. They serve not only as a means of real-time communication, but also as important devices for recording live events, creating contents, and enjoying multimedia entertainment. These diverse experiences continue to increase the level of expectations for multimedia services. The real-time voice and audio communication services have evolved from narrow band voice to super wide band or even full band audio. Given the increasing expectations, mono voice and audio services are increasingly unable to meet people's audio service experience requirements, and spatial audio is expected to be the next step up in voice and audio service quality.

IVAS codec was standardized in August 2023. IVAS supports several immersive encoder input formats and provides attractive features and performance for multiple use cases. For smartphone devices to generate and transmit immersive voice and audio content, the first step is to capture sound field signals and then transform them into appropriate formats data that can be fed 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. This may lead to some inconsistencies in level of immersion or variations in the performance of the audio service. Specifically, there may be a risk that such inconsistencies might reduce consumers interest and negatively affect their expectations for immersive audio services.

Based on the study in TR 26.933, while immersive audio capture is feasible, its widespread adoption in current products is still limited. This could potentially slow down the deployment speed of a fully end-to-end IVAS ecosystem, and provision of immersive audio capture guidelines and/or example solutions can therefore help with more widespread IVAS adoption.

Standardized minimum performance requirements/objectives for immersive audio capture solutions can ensure consumers experience consistent performance of audio quality and immersion across different devices. This consistency boosts consumers’ interest and satisfaction in immersive voice and audio services in general. The consistency can facilitate the development of new applications and technologies that use immersive audio with guaranteed minimum performance, and evaluating the performance in terms of minimum requirements/objectives will ensure opportunities for differentiation and further improvements of the immersive audio capture algorithms and solutions in the future. Provision of a clear set of guidelines to follow can potentially lead to a greater variety of applications and technologies that use immersive audio. This allows consumers to seamlessly switch between different devices and applications without experiencing unexpected inconsistencies in audio experience, benefiting both consumers and developers.

Deployment of smartphone devices with immersive audio capture capabilities meeting or exceeding minimum performance requirements/objectives can ensure high-performance immersive audio data generation. This together with use of the IVAS standard that is already available may further accelerate the maturation speed of the end-to-end immersive audio solutions in the market.

# 4 Objective

The overall objective of this work item is to define minimum performance criteria and requirements/objectives for immersive audio capture capabilities for smartphones and develop example solutions that meet the minimum performance requirements/objectives. To enable the execution of the work item and evaluation of solutions for immersive audio capture in a transparent manner, a representative set of target devices will be defined, and sufficient test signals are generated based on the devices. So, the following objectives should be achieved within the work item in two consecutive phases.

* Phase 1:
  + Definition of a set of target devices or target device types that allow the development of immersive audio capture example solutions and evaluation of their performance. The target devices or target device types are characterized by various properties, including at least the number of microphones and the overall microphone configuration model (detailing, e.g., the microphone placement and the relative distances of the microphones).
  + Definition of minimum performance requirement/objective criteria for raw microphone signal performance and characteristics including, e.g., expected compensation of the microphone signals: e.g., directional response, SNR, frequency response.
  + Definition of data requirements for signals based on the set of target devices or target device types and collection of a common database of raw/compensated microphone signals for development of example solutions for converting raw/compensated microphone signals into at least one IVAS encoder input format.
* Phase 2:
  + Development of immersive audio capture example solutions for acoustic processing and converting raw/compensated microphone signals into at least one IVAS encoder input format.
  + Evaluation of available immersive audio capture example solutions against the minimum performance requirements/objectives.
  + Specification of selected example solutions meeting the minimum performance requirements/objectives.

# 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 |
| TS | 26.xxx | xxx | SA#1xx  (xxx) | 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 |
| vivo Mobile Communication Co., Ltd |
| Nokia |
|  |