**3GPP TSG-SA WG1 Meeting #99e S1-222244r3**

**Electronic Meeting, 21st August – September 1st 2022** *(revision of S1-222244)*

Title: New use case on synchronized predictive avatars

Agenda Item: 7.4 - FS\_Metaverse

Source: Philips International B.V.

Contact: Oscar Garcia-Morchon <oscar.garcia-morchon@philips.com>

*Abstract: This contribution proposes a new use case for TR 22.856. All text is new for Change 1. Change 2 updates the reference section.*

**\* \* \* \* Start of Change 1 \* \* \* \***

5.X Synchronized predictive avatars

### 5.X.1 Description

In this first use case, three users are using the 5GS to join an immersive metaverse activity. The users Bob, Lukas, and Yong are located in the USA, Germany and China, respectively. Each of the users is served by a local metaverse edge computing server (MECS) hosted in the 5GS, each of the servers is located close to the user it is serving. When a user joins a metaverse activity, such as a joint game or teleconference, the avatar of the user is loaded in the metaverse edge computing servers of the other users. For instance, the metaverse edge computing server close to Bob hosts the avatars of Yong and Lukas.

The huge distance between the users, e.g., the distance between USA and China is around 11640 Km, determines minimum communication latency, e.g., 11640/c = 38 msec. This latency might also be higher due to causes such as hardware processing. This latency might also be variable due to multiple reasons, such as, e.g., congestion or delays introduced by (variable processing time of) hardware components such as sensors or rendering devices. Since this value maybe too high and variable for a truly immersive joint metaverse experience, each of the deployed avatars includes one or more predictive models of the person it represents and that allow rendering in the local edge server a synchronized predicted (current) avatar representation of the remote users. Similar techniques have been proposed for example in [15].

Figure 5.x.1-1 shows this exemplary scenario in which a MECS at location 3 (USA) runs the predictive models of remote users (Yong and Lukas) and takes as input the received sensed data from all users (Yong, Lukas, and Bob) as well as the current end-to-end communication parameters (e.g., latency) and generates a synchronized predicted (current) avatar representation of the users to be rendered in local rendering devices of Bob. A particular example of such scenario might be about gaming: Yong, Lukas, and Bob are playing baseball in an immersive Metaverse activity , and it is Yong’s turn to hit the ball that is going to be thrown by Lukas. If Yong hits the ball, then Bob can continue running since Yong and Bob are playing in the same team. In this example, the avatar predictive models of Lukas and Yong (deployed at the MECS close to Bob) will allow creating a combined synchronized prediction at Location 3 of Lukas throwing the ball and Yong reacting to the ball and hitting the ball so that Bob can start running without delays and can enjoy a great immersive metaverse experience.

This example aims at illustrating how predictive models can improve the user experience in a similar was as in [15]. Synchronized predictive avatars are however not limited to the gaming industry and can play a relevant role in other applications, e.g., immersive healthcare or teleconferencing use cases. This scenario involving synchronized predictive avatars assumes to require synchronization of user experiences to a single clock.



**Figure 5.x.1-1 Example of a joint metaverse experience with synchronized predicted avatar representation**

### 5.X.2 Pre-conditions

The following pre-conditions and assumptions apply to this use case:

1. Up to three different MNOs operate the 5GS providing access to metaverse services.

2. The users, Bob, Lukas, and Yong have subscribed to the metaverse services.

3. Each of the users, e.g., Bob, decide to join the immersive metaverse activity.

### 5.X.3 Service Flows

The following service flows need to be provided for each of the users:

1. Each of the users, e.g., Bob, decide to join the immersive metaverse activity and give consent to the deployment of their avatars.

2. Metaverse sensors at each user sample the current representation each of the users where sampling is done as required by the sensing modalities. The sampled representation of each of the users is distributed to the metaverse edge computing servers of the other users in the metaverse activity.

3. Each of the metaverse edge computing servers applies the incoming data stream representing each of the far located users to the corresponding avatar predictive models – taking into account the current communication parameters/performance, e.g., latency – to create a combined, synchronized, and current representation of the remote users that is provided as input to rendering devices in the local environment.

The service flows for the other users (i.e., Yong in China and Lukas in Germany) are the mirrored equivalent. For instance, even if not shown in Figure 5.x.1-1, the local metaverse edge computing server associated to Lukas will run the avatar predictive models of Yong and Bob and consume the data streams coming from those users.

### 5.X.4 Post-conditions

### The main post-condition is that each of the users enjoy an immersive metaverse activity.

### 5.X.5 Existing features partly or fully covering the use case functionality

### TS 22.261 includes in Clause 6.40.2 the following requirement related to AI/ML model transfer in 5GS: “Based on operator policy, 5G system shall be able to provide means to predict and expose predicted network condition changes (i.e. bitrate, latency, reliability) per UE, to an authorized third party.” This requirement is related to requirement [PR 5.X.6.2], but not exactly the same since the usage of predictive avatar models requires the knowledge of the end-to-end network conditions, in particular, latency.

### 5.X.6 Potential New Requirements needed to support the use case

[PR 5.X.6.1] the 5G system shall provide a means to synchronize the incoming data streams of multiple metaverse (sensor and rendering) devices associated to different users at different locations.

[PR 5.X.6.2] the 5G system shall provide a means to expose predicted network conditions, in particular, latency, between remote users.

[PR 5.X.6.3] The 5G system shall provide a means to support the distribution, configuration, and execution of a predictive model associated to a remote user in a local edge server.

**\* \* \* \* End of Change 1 \* \* \* \***

**\* \* \* \* Start of Change 2 \* \* \* \***

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 22.228: "Service requirements for the Internet Protocol (IP) Multimedia core network Subsystem (IMS)".

[3] 3GPP TS 22.173: "IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services".

[4] 3GPP TS 22.101: "Service principles".

[5] 3GPP TS 22.261: "Service requirements for the 5G system".

[6] Talaba D, Horvath I, Lee K H, "Special issue of Computer-Aided Design on virtual and augmented reality technologies in product design," Computer-Aided Design 42:361 - 363.

[7] Wang X, Tsai J J-H, "Collaborative Design in Virtual Environments," In: ISCA 48:1-2.

[8] Benford S, Greenhalgh C, Rodden T, "Collaborative virtual environments." Communications of the ACM 44: 79–85.

[9] M. Eid, J. Cha, and A. El Saddik, "Admux: An adaptive multiplexer for haptic-audio-visual data communication", IEEE Tran. Instrument. and Measurement, vol. 60, pp. 21–31, Jan 2011.

[10] K. Iwata, Y. Ishibashi, N. Fukushima, and S. Sugawara, "QoE assessment in haptic media, sound, and video transmission: Effect of playout buffering control", Comput. Entertain., vol. 8, pp. 12:1–12:14, Dec 2010.

[11] N. Suzuki and S. Katsura, "Evaluation of QoS in haptic communication based on bilateral control", in IEEE Int. Conf. on Mechatronics (ICM), Feb 2013, pp. 886–891.

[12] E. Isomura, S. Tasaka, and T. Nunome, "A multidimensional QoE monitoring system for audiovisual and haptic interactive IP communications", in IEEE Consumer Communications and Networking Conference (CCNC), Jan 2013, pp. 196–202.

[13] A. Hamam and A. El Saddik, "Toward a mathematical model for quality of experience evaluation of haptic applications", IEEE Tran. Instrument. and Measurement, vol. 62, pp. 3315–3322, Dec 2013.

[14] O. Holland et al., "The IEEE 1918.1 "Tactile Internet" Standards Working Group and its Standards," Proceedings of the IEEE, vol. 107, no. 2, Feb. 2019.

[15] Halbhuber, David & Henze, Niels & Schwind, Valentin. (2021). Increasing Player Performance and Game Experience in High Latency Systems. Proceedings of the ACM on Human-Computer Interaction. 5. 1-20. 10.1145/3474710.

…

[x] <doctype> <#>[ ([up to and including]{yyyy[-mm]|V<a[.b[.c]]>}[onwards])]: "<Title>".

**\* \* \* \* End of Change 2 \* \* \* \***