**3GPP TSG SA WG4#129-e S4-241489**

**Online, 19th - 23rd August 2024**

|  |
| --- |
| *CR-Form-v12.2* |
| **PSEUDO CHANGE REQUEST** |
|  |
|  | **26.813** | **CR** | **<CR#>** | **rev** | **<Rev#>** | **Current version:** | **0.4.0** |  |
|  |
| *For* ***[HE](http://www.3gpp.org/3G_Specs/CRs.htm%22%20%5Cl%20%22_blank)******[LP](http://www.3gpp.org/3G_Specs/CRs.htm%22%20%5Cl%20%22_blank)*** *on using this form: comprehensive instructions can be found at* *<http://www.3gpp.org/Change-Requests>.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **X** | Radio Access Network |  | Core Network |  |

|  |
| --- |
|  |
| ***Title:***  | [FS\_AVATAR] 3D watermarking for mesh-based avatar protection  |
|  |  |
| ***Source to WG:*** | China Mobile Com. Corporation |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | FS\_AVATAR |  | ***Date:*** | 08-08-2024 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***Release:*** | Rel-18 |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:**Rel-8 (Release 8)**Rel-9 (Release 9)**Rel-10 (Release 10)**Rel-11 (Release 11)**…**Rel-16 (Release 16)**Rel-17 (Release 17)**Rel-18 (Release 18)**Rel-19 (Release 19)* |
|  |  |
| ***Reason for change:*** | Since designing and producing 3D mesh-based avatar representations is a time-consuming and labor-intensive process, protecting the copyright of mesh-based avatar representations has become a significant concern in the industry. According current TR 26.813 “UC4: Avatar Generation, Storage, and Access”:- the 5G system shall support mechanisms to identify an avatar and associate the avatar with a subscriber (i.e. the owner of the avatar). - the 5G system shall be able to identify the subscriber who has the right to use an avatar in mobile metaverse servicesDigital watermarking is a technology used in copyright protection of multimedia, such as images, videos, and 3D assets. It provides an efficient solution to the copyright protection for avatar representations. |
|  |  |
| Summary of change: | Provides an overview of 3D mesh watermarking technology, and how it can be used for mesh-based avatar protection.  |
|  |  |
| ***Consequences if not approved:*** | Lack of efficient methods for digital avatar representation copyright protections. |
|  ***c*** |  |
| ***Clauses affected:*** | 2, 9 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  |  |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  |  |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  |  |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

CHANGE #1

# 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 TR 22.856: “Feasibility Study on Localized Mobile Metaverse Services”

[3] ISO/IEC 23090-14:2023, Information Technology – Coded Representation of Immersive Media – Part 14: MPEG Scene Description.

[4] AFLW (Annotated Facial Landmarks in the Wild), accessible here: <https://paperswithcode.com/dataset/aflw>, January 2024.

[5] The LFPW (Labeled Face Parts in the Wild)，accessible here: <https://paperswithcode.com/dataset/lfpw>, July 2023

[6] WFLW (Wider Facial Landmark in the Wild), accessible here: <https://paperswithcode.com/dataset/wflw>， November 2023

[7] Wang, Feng, Hang Zhou, Han Fang, Xiaojuan Dong, Weiming Zhang, Xi Yang and Nenghai Yu. “Deep 3D mesh watermarking with self-adaptive robustness.” Cybersecurity 5 (2021): n. pag.

[8] Zhu, Xingyu, et al. "Rethinking Mesh Watermark: Towards Highly Robust and Adaptable Deep 3D Mesh Watermarking." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 7. 2024.

[9] Narendra, Modigari, et al. "Levenberg–Marquardt deep neural watermarking for 3D mesh using nearest centroid salient point learning." Scientific Reports 14.1 (2024): 6942.

[10] Zhu, Xingyu, et al. "Rethinking Mesh Watermark: Towards Highly Robust and Adaptable Deep 3D Mesh Watermarking." Proceedings of the AAAI Conference on Artificial Intelligence. Vol. 38. No. 7. 2024.

CHANGE #2

# 9 Security and Privacy

## 9.X 3D Watermarking for Mesh-based Avatar Protection

### 9.X.1 Introduction

Digital watermarking is a key technology for maintaining the integrity, ownership, and authenticity of 3D mesh-based avatars. The goal is to embed a message in 3D meshes that can withstand various attacks imperceptibly during the distribution and use of avatars, and reconstruct the message accurately from distorted/undistorted watermarked 3D meshes.

There are four requirements for 3D mesh watermarking:

- **Rubustness:** The watermark must be resilient against various attacks such as compression, scaling, and noise during transmission channels. The robustness of digital watermarking against a specific type of attack can be measured by the accuracy of watermark extracted from the attacked mesh.

- **Imperceptibility:** The watermark should not visibly change the appearance of the 3D mesh, ensuring that the visual quality of the avatars remains unchanged.

- **Efficiency:** The watermarking and extraction processes should be computationally efficient, ensuring that they do not introduce significant overhead or delay in the use and distribution of the avatars.

- **Capacity:** The ability to embed a larger amount of messages into the mesh.

### 9.X.2 3D Mesh Watermarking Methods

Existing 3D mesh watermarking methods can be categorized into traditional methods and deep learning-based methods. Traditional methods manually design watermark embedding and extraction algorithms in the spatial or transform domain to resist specific types of attacks. However, manually designing a watermarking algorithm for each specific scenario is both time-consuming and labor-intensive, and most of manually designed watermarking algorithms struggle to achieve universal robustness against various attack forms, such as clipping and smoothing.

Deep learning based-methods for 3D mesh watermarking achieved great progress in recent years. Deep learning-based watermarking techniques can optimize watermark embedding and decoding process by simulating various types of attacks, which effectively enhancing the universality of the watermark. It has been demonstrated that deep learning-based methods are effective and robust for copyright traceability of mesh-based digital avatar data files [7, 8. 9, 10]. Figure.9.X.2 shows the generic pipeline of deep learning-based 3D mesh watermarking.



**Figure.9.X.2 Generic Pipeline for Deep Learning-based 3D Mesh Watermarking**

The procedures of deep learning-based 3D mesh watermarking are summarized below:

S1. Acquire the original 3D mesh and embedding watermark;

S2. The pre-trained watermark processing system takes the original 3D mesh and embedding watermark as input, and use neural network model(s) simultaneously embed watermark into the geometric and texture parts of the 3D mesh.

S3. The watermarked 3D mesh may be attacked (e.g., transformations, random rotation, nosing, clipping, etc) during avatar transmission or implementation.

S4. The target watermark can be extracted from the attacked watermarked 3D mesh based on either the geometric part or texture part by neural network model(s). The target watermark should be consistent with the input watermark.

### 9.X.3 Mesh-based Avatar Protection

Mesh is a commonly used avatar representation format. Figure.9.X.3 provides the framework for implementing deep learning-based 3D mesh watermarking into mesh-based avatar protection without affecting the usability of digital avatar representations. .



**Figure.9.X.3 Deep Learning-based 3D Mesh Watermarking for Avatar Protection**

**Data Pre-processing Module:** Pre-process the 3D avatar file to obtain the original 3D mesh from the Avatar Storage entity.

**Watermark Embedding Module:** The geometric and texture parts of the 3D mesh are simultaneously embedded watermark to generated the watermarked 3D mesh . The geometric watermark embedding network extracts feature from geometric part of the 3D mesh, combining it with input watermark, and reconstructs the watermarked geometric part. The texture watermark embedding network follows the similar processes to reconstruct the watermarked textures.

**Attack Simulation Module:** During the training process, the attack simulation module simulates various attacks, including transformations, random rotation, nosing, clipping, etc, and applies them to the watermarked mesh during the end-to-end training of the framework, to enhance the robustness of the watermark against universal attacks.

**Watermark Extraction Module:** Features are extracted separately from the geometric structure and/or texture map of the attacked 3D mesh. The geometry and the texture features are independently fed into the watermark recovery network to obtain target watermark. This means that even if one part of the mesh's geometry or texture is intentionally removed, the complete watermark can still be extracted from the other part, thereby achieving the goal of copyright tracing.

3D watermarking is one of effective methods for embedding user identity information into 3D avatar assets without affecting their appearance or attributes. The latest algorithms are highly robust, meeting the latency requirements for real-time communication. An potential example to use it for user-authentication during the AR Call could be:

1. The application server generates a unique user-ID for each caller before the AR call setup, establishing a one-to-one mapping between the user-ID and the user.

2. The unique user-ID is embedded into the user's mesh-based avatar representation as a 3D watermark during the base avatar generation process.

3. During the call, the application server can extract the 3D watermark from the mesh-based avatar representation from either the geometric part or texture part, and verify the user’s identity by referencing the stored mapping in the database in application server.

END OF CHANGES