3GPP TSG SA WG4 Meeting #122S4-230146

20th – 24th February 2023

**Source:** Samsung Electronics Co., Ltd.

**Title: [MeCAR] On pose information**

**Document for** Discussion and Agreement

# Introduction

For the split rendering service, a MeCAR device may send an estimated pose for the estimated target display time to the split rendering server. The device also receives a rendered media frame with the pose used to generate the frame from the server. The current MeCAR PD v0.4 describes user pose as a pose and a timestamp, but it is unclear whether the time can be interpreted as the pose acquisition time or the target display time.

To make an accurate estimation of the target display time, the device should refer to the statistical history of the delays previously occurred and estimated. The estimated delay is the gap between the time when an estimation was made (T1 in the figure) and the time estimated as the display time (T2.estimated). The actual delay which occurs is the gap between the time when an estimation was made (T1) and the time when a photon was displayed (T2.actual).

In case the device sends more than one pose at a time, a group of poses can be considered as a sequence of pose and metadata pairs with several timestamps. The time when the group of poses is sent by the device (T1') may help the split render function in the server to identify the version of this pose group in the case where the device intends to overwrite, or resend certain estimated poses updated using the more recent and accurate estimation parameters. The server may replace poses for the same target display time (poses with a new T1') if the poses have not been rendered yet.

In the case where there are poses stacked in the pose buffer of split render function, for example with a granularity finer than the device's supported frame rate, the split render function should select the pose closest to the target display time, according to the previous-render-to-photon delay. The previous-render-to-photon delay from the most recent frame information may help the server to make this selection.

# Proposed changes

<Change 1>

6.X Pose information

The MeCAR device sends a group of pose information to the server's split render function to generate rendered media frames based on the poses. Each pose is associated with time metadata, such as the time when the pose estimation was made (T1), the estimated target display time of the content (T2.estimated), and the time the group of poses was sent (T1').

The gap between the actual-target-display-time (T2.actual) and the pose estimate time (T1) is the pose-to-render-to-photon delay, which allows the MeCAR device to know the amount of processing time as well as the connection delay required for a loop of split rendering. The next round of pose estimation should refer to the pose-to-render-to-photon delay for the estimation of a new T2.estimated.

The split render function in the server may refer to T1', which is the time when the group of poses is sent from device, if multiple pairs of pose and metadata for the same target display time are received from the device. The T1’ information may be used to manage poses by the server, such as allowing the MeCAR device to update former estimations by resubmitting a new pose with the same estimated-target-display-time.

The split render function in the server sends rendered media frames and associated metadata. To allow the MeCAR device to know which version and pair of pose and estimated-target-display-time was selected and rendered among the pose pairs in the pose buffer of server, the associated metadata may include the pose used for the rendered frame, as well as corresponding time information, such as T1, T2.estimated, and the time when the rendering started (T3) to the MeCAR device in order to measure the render-to-photon delay.

In the case where there are poses stacked in the server’s pose buffer, for example with a granularity finer than the device's supported frame rate, the split render function should select the pose closest to the display time, according to the previous-render-to-photon delay. The previous-render-to-photon delay from the most recent frame information may help the server to make this selection.

Note: In addition to the timestamp differences between the estimated and actual displayed poses, the accuracy of the estimated pose with regards to the real pose is also relevant. Metrics related to such pose estimation accuracy is FFS.

The figure below illustrates the time flows such as pose-to-render-to-photon delay Hand render-to-photon delay occurred between the MeCAR device and split rendering function in the server.



* Pose and timestamp information from the device
	+ Estimated pose
	+ Estimated-at-time (T1)
	+ Estimated-target-display-time (T2.estimated)
	+ Sent-at-time (T1')
	+ Previous-render-to-photon-time (T2.actual-T3)
* Pose and timestamp information associated with rendered media frame from the server
	+ Pose used for rendering
	+ Estimated-at-time (T1)
	+ Estimated-target-display-time (T2.estimated)
	+ Start-to-render-at-time (T3)

Pose includes location and direction information.

Estimated-at-time (T1) is the actual time of when the pose estimation was made.

Estimated-target-display-time (T2.estimated) is the estimated target display time for the media frame which is rendered, or will be rendered, using to this pose.

Sent-at-time (T1') is the actual time when a pose or a group of poses is sent from the device to the server.

Previous-render-to-photon-time (T2.actual.previous-T3.previous) is the render-to-photon delay for the most recent frame.

Start-to-render-at-time (T3) is the actual time when the renderer in the split rendering server starts to render the associated media frame.

<End of change 1>

# Proposal

It is proposed to add the proposed change in clause 2 to section 6.X of MeCAR PD v0.4.