**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11/N17886**

**July 2018, Ljubljana, SI**

**Title: MPEG-I Phase 1 Use Cases (v1.5)**

**Source: WG11**

**Status: Approved, Public**

**Editor: Mary-Luc Champel (mary-luc.champel@technicolor.com)**

**Rob Koenen (rob.koenen@tno.nl)**

**Madhukar Budagavi (m.budagavi@samsung.com)**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Contents**

[1 Introduction 2](#_Toc488422708)

[2 Use Cases 2](#_Toc488422709)

# Introduction

This document gathers use cases for immersive media. They are used to derive requirements for immersive applications and services.

Use cases presented here are targeting the first two phases of MPEG-I Phase 1a and Phase 1b. A separate document, N17683, documents the use cases used to derive ruirements for MPEG-I Phase 2.

# Use Cases

|  |
| --- |
| Multiple users in VR environment |
| Several users can navigate simultaneously within a shared VR environment. User interactions with VR objects and possibly between VR users are shared in real-time among all users. |
| **Overlap with other use cases** |
| user embodiment in 360 video  user interactions with VR environment objects |
| **Target Phase (1a, 1b, 2) :** Phase 1b, Phase 2 |
| **Required  features** |
| Interactions with VR objects and between users  Synchronization of interactions among users  Rendering of other users in the virtual environment;  synchronization of audio and video of users and the scene (this may duplicate what goes above) - not that there is a trade-off between synchronization and immediate interaction |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| Detection and rendering of interactions between users within the VR environment  Synchronization among users of interactions with VR objects |

|  |
| --- |
| 360 3DoF+ video with depth |
| While still sitting, the user can experience full 3DoF VR experience (the three head rotation degrees of freedom are fully supported) with the addition of constrained degree of freedom on translation axis (heave, sway and surge). Such 3DoF+ experience allows the user to benefit from a true 3D experience thanks to the support of motion parallax. |
| **Interested People:** |
| **Overlap with other use cases** |
| Overlap with : 360 3DoF video with depth. The present use case adds the capability to have constrained movement along heave, sway and surge axis. The constraints reproduce the actual physical constraints experienced by a sitting user. |
| **Target Phase (1a, 1b, 2) :** Phase 1b |
| **Required  features** |
| 3DoF+ navigation  Depth information  Motion parallax support |
| **Identified Gaps and Optimization Potentials** |
| Quality of motion parallax perception  Experience shall not rely on heave , sway and surge sensors but benefit from them when present.  Composition of video objects with different depths. |
| **Potential Requirements and Specifications** |
| Depth rendering and composition with support of motion parallax |

|  |
| --- |
| High Quality 360 3DoF+ video with depth and OTT delivery |
| Users can experience the highest quality VR supported by its HMD. |
| **Interested People:** |
| **Overlap with other use cases** |
| 360 3DoF+ video with depth  user embodiment in 360 video  user interactions with VR environment objects |
| **Target Phase (1a, 1b, 2):** Phase 1b |
| **Required  features** |
| High Quality VR builds on 3 quality pillars : audio, video and immersiveness.   * Visual quality is such that #pixels per degree of the content is at least above what current HMDs can support (40). In other words, visible viewport resolution is equal or higher to HMD display resolution. * Audio quality is high enough to support 3DoF+. * Immersiveness quality is obtained thanks to high video framerate (90fps), very low motion-to-photon latency (20ms), motion parallax support, user embodiment and interactions.   Despite its limited and not guaranteed bandwidth, OTT delivery shall not negatively impact the high quality of the VR experience. More precisely delivery shall be done without sacrificing on visual quality, motion-to-photon latency or framerate. |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| 3DoF+ navigation  Depth information  Motion parallax support  User shall be present in the VR environment through some form of user-embodiment.  Composition of user-embodiment with VR environment shall be as realistic as possible (proper lighting information, presence of realistic shadows…).  User actual movements shall be monitored and rendered in real-time within VR environment.  Interactive objects within VR environments  Detection & rendering of user interactions with such objects  Delivery and interaction means that can ensure low enough motion-to-photon latency  High Quality VR adaptive streaming |

|  |
| --- |
| Distributing 360 A/V content library in 3GPP |
| Introduction: content library and target devices A Service provider has access to a library of 360 A/V content. The library is a mixture of content formats from user generated content, documentaries, promotional videos, as well as highlights of sports events. The latter content is replaced and updated daily. The content enables to change the field-of-view based on user interaction. The Service provider wants to create a portal to distribute the content to Head Mounted Displays that are powered by mobile devices. The service provider wants to target two types of applications:   * View on a screen with the field-of-view for the content adjusted by manual interaction (e.g. mouse input or finger swipe) * View in a HMD with head motion tracking. The Service provider expects different types of consumption and rendering devices with different capabilities in terms of decoding and rendering.   The Service provider has access to the original footage of the content and may encode and transcode it to appropriate formats. The footage includes different types of VR content, including   * For video:   + Monoscopic video, i.e. a spherical video without real depth perception.   + Stereoscopic video, i.e., a spherical video using a separate input for each eye.   + Different spatial and temporal resolutions, e.g. up to 16K and 240fps with 10 bit and 12bit as well as different colour spaces and transfer characteristics. * For audio:   + Channel-based audio   + Object-based audio   + Scene-based audio   + Mixtures of the above   The service provider wants to reach as many devices as possible and wants to minimize the amount of different formats that need to be produced while ensuring that the content is presented in highest quality on the different devices. Specifically, the content provider wants to target existing and emerging mobile phones, for example powered by Snapdragon 821 and 835.  The service provider also wants to ensure that the original content provides consistent information that can be used by the encoding and decoding to maintain the artistic intent of the content producer. Downloading content The service provider wants to enable that a certain amount of the content can be downloaded to devices through HTTP and is played back on the device after downloading. The service provider wants to ensure that a device downloads only content that it can decode and render while providing the best user experience for the device capabilities. Non real-time broadcast distribution of VR content Many of the devices support broadcast file delivery services. The service provider agrees with a mobile network operator that certain content is pre-cached on mobile devices for offline consumption using an MBMS file delivery services. The service provider wants to ensure that MBMS delivered content can be consumed by as many devices as possible. Streaming content For certain contents, the service provider wants to ensure that content is rendered instantaneously after selection, so a DASH-based streaming is considered. The service provider wants to ensure that a device accesses only content that it can decode and render while providing the best user experience for the device's capabilities. The service provider also wants to ensure that the available bandwidth for the user is used such that the rendered content for the user is shown in the highest quality possible. |
| **Overlap with other use cases** |
| Overlaps with OMAF objectives |
| **Target Phase (1a, 1b, 2) :** Phase 1a, Phase 1b |
| **Required features** |
| * Consistent formats of the original footage including metadata * Audio and video encoding tools for 360 video and 3D audio that do not exceed capabilities of existing and emerging devices * Encapsulation tools to annotate and distribute content through ROI streaming and for broadcast * Reference decoder to rendering interfaces in order to enable rendering of 360 video and 3D audio |
| **Identified Gaps and Optimization Potentials** |
| * Consistent metadata for encoding * Specification for collection of audio and video encoding tools for 360 video and 3D audio that do not exceed capabilities of existing and emerging devices * Encapsulation tools to annotate and distribute content through ROI streaming and for broadcast * Reference decoder to rendering interfaces in order to enable rendering of 360 video and 3D audio |
| **Potential Requirements and Specifications** |
| tbd |

|  |
| --- |
| Social TV and VR |
| Alex and his friends have signed up to a service provider for a new live football experience. Alex watches a live football match on his HMD. His friends do the same, each at different places. All of them wear a headset with microphone to communicate. They connect socially to at least communicate verbally with each other. Everyone sees the football match in such a way that they can comment and discuss the actions of the match live as if they would sit in front of the same TV. They experience at least the sound communication as if the friends are sitting next to each other at different fixed positioned seats.  In an extended version each of them are placed at a virtual stadium seat and watch the match while being able to follow the action by head movement. |
| **Overlap with other use cases** |
| OMAF, MORE |
| **Target Phase (1a, 1b, 2) :** Phase 1b, 2 |
| **Required features** |
|  |
| **Identified Gaps and Optimization Potentials** |
| * Mixing of 3D scene sound with non-diegetic content |
| **Potential Requirements and Specifications** |
| Tbd |

|  |
| --- |
| Multi-viewpoint 3DoF with monoscopic/stereoscopic 360 video content |
| This use case extends from 3.1 and 3.2 in that the viewer is presented with a multi-viewpoint 3DoF with either monoscopic or stereoscopic 360 video content experience. Compared to 3.1 and 3.2, the viewer is given the additional opportunity to change his/her 360 degree viewpoint, amongst a set of multiple predefined (fixed) viewpoints within the content. Changing between the fixed viewpoints can be either controlled manually by the viewer (using some kind of controller), or automatically through specified metadata (e.g. a director’s cut or guided viewing). The corresponding audio for the current viewpoint location will also be delivered and rendered accordingly. View synthesis or interpolation is not used in this use case. |
| **Overlap with other use cases** |
| Same as 360 Movie, but with positional meta-data either transmitted or via local interface |
| **Target Phase (1a, 1b, 2) :** Phase 1a,1b |
| **Required  features** |
| Muti-viewpoint |
| **Identified Gaps and Optimization Potentials** |
| **Potential Requirements and Specifications** |
| OMAF version 1 |

|  |
| --- |
| Single viewpoint 3DoF+ with full 3D 360 video (limited 6DoF) |
| Single viewpoint 3DoF+ with full 3D 360 video is an enhanced experience of the fixed single viewpoint use case for Phase 1a (section 3.1). This use case differs from 3.1 in that it provides a realistic, natural full 3D 360 video experience, where the rendered content provides a natural 3D representation depending on all head rotation orientations (yaw, pitch, roll) of the viewer. As a 3DoF+ experience, this single viewpoint content also provides a limited amount of motion parallax, enough to give the viewer a sense of natural depth whilst changing his/her view within the scene, as well as the capability for small head translational movements. Small head translational movements are defined as movements which can be achieved whilst the viewer is in a seated position, without the use of the lower body. In this manner, this content is comparable to a restricted 6DoF experience, where translational movements are limited to small head movements. Audio is also rendered accordingly based on the user’s head position and orientation, including possible sonic occlusion, if perceptually relevant.  Whilst most content for this use case will be professional content which is post-produced, future consumer capture devices may support full 3D capture, introducing user generated content which also falls into this use case category.  Example user case content scenarios  Full 3D 3DoF virtual rollercoaster:  Omnidirectional real-scene content captured on real rollercoasters with motion parallax for full 3D 360 video.  User generated content:  Future generations of consumer targeted omnidirectional camera systems may support the acquisition of single viewpoint 3DoF+ full 3D content. The live capture and delivery of this user generated content can deliver a live immersive experience. |
| **Overlap with other use cases** |
| 3D Audio with 3 DoF. |
| **Target Phase (1a, 1b, 2) :** Phase 1b |
| **Required  features** |
| 3DoF+ navigation with associated interface to renderer Audio rendering responsive to small head movements |
| **Identified Gaps and Optimization Potentials** |
| Composition of audio objects with change in position and possible occlusion due to head movement. |
| **Potential Requirements and Specifications** |
| Audio metadata shall describe the orientation and directivity and potentially the behavior of a sound source when it is occluded, if perceptually relevant. |

|  |
| --- |
| Multi-viewpoint 3DoF+ with full 3D 360 video |
| Description  This use case is a combination of the two use cases from 3.3 and 4.1. The combination of the two use cases results in a multi-viewpoint experience where each viewpoint presents a full 3D 360 video with limited motion parallax and capabilities for small head translational movements. The corresponding audio for the current viewpoint location will also be delivered and rendered accordingly.  Example user case content scenarios  Multi-viewpoint concert experience:  Full 3D 3DoF 360 content captured from multiple viewpoints dotted around the concert venue allows viewers to experience the concert from multiple locations in the concert space, especially the acoustic differences in audio perspective from each individual viewpoint.  Multi-viewpoint sports experience:  Although similar to a multi-viewpoint concert experience, a multi-viewpoint sports experience does not necessarily require multi-viewpoints to exist within the same space, on the condition that multi-viewpoint videos are aligned temporally. One example is a Formula One race event, where multi-viewpoints may include: stadium viewpoints, racer vehicle viewpoints and pit-stop viewpoints. |
| **Overlap with other use cases** |
|  |
| **Target Phase (1a, 1b, 2) :** Phase 1b |
| **Required  features** |
| Multi-viewpoint, Full 3D |
| **Identified Gaps and Optimization Potentials** |
| **Potential Requirements and Specifications** |
| OMAF version 2 |

|  |
| --- |
| Playing a game in Social VR |
| Alice and her friend Bob play a game together using social VR.   * Alice and Bob are at different physical locations   + Both wear a head-mounted display (HMD) that has a microphone   + Both have game consoles and connected controllers to engage in the game   + Both have 3D audio rendering, either at the HMD or with a fixed speaker set-up   + Each has a camera that captures his/her activity, and a solution to remove background, e.g.     - green screen a.k.a. chroma key, or     - depth camera, or     - stereo camera * The HMD places the two players in the same virtual environment   + The virtual environment is available in the form of 360o photo(s) or 360o video(s), or as a rendered model   + The virtual environment has 3D audio   + The virtual environment features a shared screen that is used for the game * Everything is orchestrated and aligned to create a consistent social experience   + The game screen is aligned in the virtual environment onto the virtual screen   + Alice and Bob are placed and oriented consistently in the virtual environment     - E.g. sitting in front of the shared screen and next to each other   + The cameras are placed to achieve consistent eye contact     - The location of the camera in Alice's physical set-up corresponds with Bob's location in the virtual environment     - Ditto vice versa   + The 3D audio objects are placed consistently with the virtual environment     - The game audio is placed e.g. at the shared screen (or surround, if required by the game)     - The audio of Alice is placed at the location of Alice in Bob's view.     - Ditto vice versa   Note that this use case has several features in common with use case 2.5 Social TV and VR  Note that several features from other use cases could be mixed in, e.g.   * 2.3 User Embodiment in 360 video 🡪 users sees own gestures towards the other person * 2.1 Multiple users in VR environment 🡪 extend Social VR to more than two users * … several more |
| **Overlap with other use cases** |
| MPEG-V, OMAF, MORE, use case 2.5 Social TV and VR |
| **Target Phase (1a, 1b, 2):** Phase 1b |
| **Required features** |
| * Composition of the virtual environment * Spatial alignment in the virtual environment * Spatial alignment in the physical environment * Temporal alignment of the virtual environment * Messaging and control signaling |
| **Identified Gaps and Optimization Potentials** |
| * T.b.d. |
| **Potential Requirements and Specifications** |
| See m40263 and m41999  Potentially standardisable aspects for these requirements are   * Metadata that provides yaw/pitch/roll coordinates that provide the position/orientation of a camera (or microphone) with respect to the perceived VR environment * Metadata that provides the type and other details of the camera (or microphone) * Metadata to identify and distinguish camera’s (or microphones) in case there are multiple * Metadata to signal details of the visual (and/or auditive/haptic) indication of the camera/microphone position/orientation in the VR environment * Containers (e.g. ISOBMFF) to carry the above metadata to the other user(s).   Such metadata could be specified in the MPEG-V activity, in close coordination with MPEG-OMAF.  See m40263 and [m42000](http://wg11.sc29.org/doc_end_user/current_document.php?id=60906&id_meeting=173)  Potentially standardisable aspects for these requirements are   * Metadata that conveys the angle of view of the camera * Metadata that conveys the distance between the camera and the video-captured person * Metadata that conveys the position of the video-captured person in the video frame * Containers (e.g. ISOBMFF) to carry the above metadata to remote users.   Such metadata could be specified in the MPEG-MORE activity, in close coordination with MPEG-V and MPEG-OMAF. |

|  |
| --- |
| A still subject looks around in a room with up-close views |
| A single person stands still in a center of a room of the museum and looks around, with the capability to look closely. |
| **Overlap with other use cases** |
|  |
| **Target Phase (1a, 1b, 2) :** Phase 1b |
| **Required  features** |
| The media content consists of multiple nested spheres with the same center, and the subject is at the center of the sphere, looking from inside out, with the 3 rotational DoFs. The subject is capable of view changing from sphere to sphere. |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| Media Format shall support   * the spherical geometry type * content with nested spherical geometry types * switching among content with the nested spherical geometry types   Presentation Format shall support   * spherical viewports with translations and orientations in 3/2/1 dimensions, moving along the z-axis * viewport dependent presentation that can be adapted to different network conditions and device capabilities and configurations   Orchastration Format shall support   * orchestartion of the nested spherical content along the z-axis |

|  |
| --- |
| Single user in VR environment, 6 DoF |
| A single user can navigate within a VR environment.  It may be that there is no or simple interaction, i.e. interaction that has been predetermined by the content creator. Interaction may be responsive to user coordinates, orientation, or user/object proximity, e.g. if I am near an object and look at it, it begins to “talk.”  User interactions with VR objects are rendered in real-time.  It may be that only audio components are presented in VR environment, e.g. AR augmented with audio (e.g. without augmented visuals) may be used for AR audio advertisement. |
| **Overlap with other use cases** |
| **Target Phase (1a, 1b, 2) :** Phase 1b, Phase 2 |
| **Required  features** |
| 6DoF navigation with associated interface to renderer  Audio rendering responsive to user movements  Navigation within “background” audio  Navigation within and around VR audio objects  Interactions with VR objects |
| **Identified Gaps and Optimization Potentials** |
| **Potential Requirements and Specifications** |
| Correct rendering of background and object audio from user perspective for all head orientations and user positions in VR world. |

|  |
| --- |
| Multiple users in VR environment, 6 DoF |
| Several users can navigate simultaneously within a shared VR environment. User interactions with VR objects and between VR users are shared in real-time among all users. |
| **Overlap with other use cases** |
| Single user in VR environment, 6 DoF |
| **Target Phase (1a, 1b, 2) :** Phase 1b, Phase 2 |
| **Required  features** |
| Interactions with VR objects  Rendering of other users in the virtual environment, including possible speech or audio from other users.  Interaction between users, including voice communications. Voice communications between users must be low-latency.  Synchronization of interactions among users, e.g. lip movement and lip synch of speech.  Synchronization of audio and video of users and the scene. |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| Detection and rendering of interactions between users within the VR environment  Audio generated by one user and rendered for other must be sufficiently low-latency so as to support conversational communication  Synchronization of interactions among users, e.g. lip movement and lip synch of speech.  Synchronization among users of interactions with VR objects |

|  |
| --- |
| Heterogeneous interactions |
| Several users can interact with each others with different capabilities (e.g. device capabilities, network capabilities) in a shared VR environment:   * Users with high bandwidth can be provided high-resolution omnidirectional view, those users can be represented as some sort of user-embodiment in shared VR environment, user's movements, voice can be captured and rendered in real-time within VR environment, and can also get other user-embodiment's movements in view. * Users with low bandwidth can be provided omidirectional view with low resolutions according to their bandwidth. those users can interact with other users in the same way with high bandwidth users. * For those users whose devices with limited capabilities (e.g. device without motion tracking, limited bandwidth) can only support omnidirectional view by manual interaction (e.g. TV remote controller input), they can choose a certain viewports (among a set of multiple predefined viewports) to get into a shared VR environment, and zoom in/out the interested field of view. If device capabilities allowed, user interaction shall also be done in the share VR environment. |
| **Overlap with other use cases** |
| Multiple users in VR environment  User embodiment in 360 video  User interactions with VR environment objects |
| **Target Phase (1a, 1b, 2) :** Phase 1b, Phase 2., |
| **Required  features** |
| Dynamic resolutions according to bandwidth  Multiple viewports can be pre-defined before set up a shared VR environment  User Interactions with multiple devices with different capabilities, especially capabilities limited device. |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| Detection and rendering of interactions between users within the VR environment  Synchronization among users of interactions with VR objects  Zooming in/out of region of interest by user interactions  Heterogeneous devices supported in same VR environment |

|  |
| --- |
| Low-latency Viewport Switching by Multi-sources |
| In 360 video, viewport is switched in different ways, e.g. user head motion, recommended viewport of director’s cut and friend’s viewing. High-quality immersive experience is guaranteed only with low-latency in viewport switching. There are three cases.   1. Bob watches a 360 video and selects his viewport by himself by head motion or manual input based on his view interest. The player with strong capabilities collects his selection/preference and moving track, and then, predicts his next viewport which Bob will switch into in a high probability. When he really switch his viewport, the player fast switches into next viewport based on prediction. 2. During the viewing process, if Bob don’t want to miss some splendid views that are recommended by director, he can select “director recommendation mode” to playback. When Bob is going to watch a scene with recommended viewport, his viewport can be fast and smoothly switched into recommended viewport without degraded experience by pre-fetching the recommended VR content. 3. Bob watches a football game with his friend Jack in VR environment, when Jack points out match highlight, Bob will align his viewport with Jack’s by acquiring Jack’s viewport and fast switching into. Moreover, if Bob selects “friend’s viewport mode”, he can follow Jack’s viewport and viewport switching all the time. |
| **Overlap with other use cases** |
| Multi-viewpoint 3DoF with monoscopic/stereoscopic 360 video content |
| **Target Phase (1a, 1b, 2):** Phase 1b. |
| **Required  features** |
| Viewport prediction based on user behavior, director’ cut or other viewer’s recommendation  VR content pre-fetching with viewport prediction or recommendation |
| **Identified Gaps and Optimization Potentials** |
|  |
| **Potential Requirements and Specifications** |
| The specification shall support low latency viewport switching by pre-fetching VR content based on viewport prediction or recommendation. |

|  |
| --- |
| VR Conferencing |
| User A is in a VR environment, experiencing a VR conference with remote users B, C and D. Users B, C and D are depicted in the virtual environment of user A, e.g. as video placed in that VR environment.  During the VR conference, user A directs its attention and view to different remote users at different times. At every moment, it is clear for each of remote users whether user A is looking at them or not. |
| **Overlap with other use cases** |
| MPEG-V, MORE, OMAF, use case 2.1 Multiple users in VR environment, use case 2.5 Social TV and VR, |
| **Target Phase (1a, 1b, 2):** Phase 1b |
| **Required features** |
| * Features that help a remote user determine whether it is being looked at by user A or not, when user A is looking around in user A’s VR environment. |
| **Identified Gaps and Optimization Potentials** |
| * T.b.d. |
| **Potential Requirements and Specifications** |
| See [m41998](http://wg11.sc29.org/doc_end_user/current_document.php?id=60904&id_meeting=173)  Potentially standardisable aspects for these requirements are   * Metadata that provides the position/orientations of remote users in the VR environment of user A * Metadata that provides the direction of view of user A in its VR environment * Metadata that indicates which remote user is being looked at by user A * Metadata that indicates whether a video stream corresponds to user A looking straight into the camera (main camera) or not (side camera) * Containers (e.g. ISOBMFF) to carry the above metadata to remote users.   Such metadata could be specified in the MPEG-MORE activity, in close coordination with MPEG-V and MPEG-OMAF. |

|  |
| --- |
| VR Magnifier |
| A user is experiencing a 360 video. The director wants to focus the user’s attention to details of a specific object in the 360 video. The director introduces a VR magnifier that magnifies in the direction of the object. The user looks at the object, increases the magnification even more and moves the direction of the VR magnifier. The quality of the 360 video remains constant, as a higher-resolution tile is provided for the VR magnified area. |
| **Overlap with other use cases** |
|  |
| **Target Phase (1a, 1b, 2):** Phase 1b |
| **Required features** |
| * Features that enable the configuration and control of a VR magnifier |
| **Identified Gaps and Optimization Potentials** |
| * T.b.d. |
| **Potential Requirements and Specifications** |
| See m42001  Potentially standardisable aspects for these requirements are   * Metadata about the VR magnifier algorithm * Metadata with parameters of the VR magnifier algorithm, e.g. the magnitude of the central magnification, and the boundary of the magnified area * Metadata about the location (jaw, pitch) of the VR magnifier * Metadata about the symbolisation of the VR magnifier, e.g. a drawn ring * Metadata about the allowed user controls of the VR magnifier, varying between none, some parameters and all. * Containers (e.g. ISOBMFF) to carry the above metadata to the other user(s).   Such metadata could be specified in the MPEG MORE activity, in close coordination with MPEG-OMAF. |