International Telecommunication Union

ITU-R
Radiocommunication Sector of ITU

Recommendation ITU-R BT.2133-0
(10/2019)

Transport of advanced immersive audio-visual content in IP-based broadcasting systems

BT Series

Broadcasting service (television)

International Telecommunication Union

Foreword

The role of the Radiocommunication Sector is to ensure the rational, equitable, efficient and economical use of the radio-frequency spectrum by all radiocommunication services, including satellite services, and carry out studies without limit of frequency range on the basis of which Recommendations are adopted.

The regulatory and policy functions of the Radiocommunication Sector are performed by World and Regional Radiocommunication Conferences and Radiocommunication Assemblies supported by Study Groups.

Policy on Intellectual Property Right (IPR)

ITU-R policy on IPR is described in the Common Patent Policy for ITU-T/ITU-R/ISO/IEC referenced in Resolution ITU-R 1. Forms to be used for the submission of patent statements and licensing declarations by patent holders are available from http://www.itu.int/ITU-R/go/patents/en where the Guidelines for Implementation of the Common Patent Policy for ITU-T/ITU-R/ISO/IEC and the ITU-R patent information database can also be found.

Series of ITU-R Recommendations

(Also available online at http://www.itu.int/publ/R-REC/en)

Series	Title
BO	Satellite delivery
BR	Recording for production, archival and play-out; film for television
BS	Broadcasting service (sound)
BT	**Broadcasting service (television)**
F	Fixed service
M	Mobile, radiodetermination, amateur and related satellite services
P	Radiowave propagation
RA	Radio astronomy
RS	Remote sensing systems
S	Fixed-satellite service
SA	Space applications and meteorology
SF	Frequency sharing and coordination between fixed-satellite and fixed service systems
SM	Spectrum management
SNG	Satellite news gathering
TF	Time signals and frequency standards emissions
V	Vocabulary and related subjects

Note: *This ITU-R Recommendation was approved in English under the procedure detailed in Resolution ITU-R 1.*

Electronic Publication
Geneva, 2019

© ITU 2019

RECOMMENDATION ITU-R BT.2133-0

Transport of advanced immersive audio-visual content in IP-based broadcasting systems

(2019)

Scope

This Recommendation provides guidance on using ISO/IEC 23090-2 (Omnidirectional media format, also known as OMAF) to transport advanced immersive audio-visual (AIAV) content in IP-based broadcasting systems.

Keywords

Immersive media, virtual reality, 360° video, transport, IP-based broadcasting

The ITU Radiocommunication Assembly,

considering

a) that Virtual Reality, 360° video, and other immersive media technologies have caught the attention of the content providers, audiences, and the associated consumer technology vendors;

b) that Recommendation ITU-R BT.2123 specifies parameter values of AIAV systems including 360° video for production and international programme exchange in broadcasting;

c) that ISO/IEC 23090-2 "Omnidirectional media format (OMAF)" specifies encapsulation, signalling, and streaming of omnidirectional media for MPEG Media Transport (MMT) as specified in ISO/IEC 23008-1 and for MPEG-dynamic adaptive streaming over HTTP (DASH) as specified in ISO/IEC 23009-1;

d) that IP-based broadcasting systems can be extended to support OMAF for transporting AIAV content;

e) that guidance on using OMAF in IP-based broadcasting systems will facilitate broadcasters in the development and deployment of the systems,

considering further

a) that Recommendation ITU-R BT.1869 specifies multiplexing schemes for IP packets in broadcasting systems;

b) that Recommendation ITU-R BT.2074 defines the service configuration, media transport protocol, and signalling information required for IP-based broadcasting systems using MMT,

recommends

that when AIAV content is to be transported in an IP-based broadcasting system using the Omnidirectional media format (OMAF) standard specified in ISO/IEC 23090-2, the guidance in Annex 1 should be referred to.

NOTE 1 – Annex 2 shows an overview of the OMAF standard specified in ISO/IEC 23090-2 for information.

NOTE 2 – Annex 3 presents possible layer structures for AIAV content in IP-based broadcasting systems for information.

Annex 1

Guidance on using ISO/IEC 23090-2 for IP-based broadcasting systems

Presentation profiles of OMAF define requirements and constraints for tracks or items of any number of media types. The OMAF viewport-independent baseline presentation profile should be used. This profile fulfils the basic requirements to support 3D audio and omnidirectional and 3D video. Both monoscopic and stereoscopic video are supported.

A video track should conform to the HEVC-based viewport-independent OMAF video profile, which supports both monoscopic and stereoscopic spherical video up to 360°. The bitstream of the track complies with the HEVC main 10 profile, Main tier, Level 5.1.

An audio track should conform to the OMAF 3D audio baseline profile, which supports channels, objects, and higher-order ambisonics (HOA) as 3D audio by using MPEG-H 3D Audio codec. MPEG-H 3D Audio specifies a normative interface for the user's viewing orientation and permits low-complexity, low-latency rendering of the audio scene to any user orientation. The bitstream of the track complies with the MPEG-H 3D Audio Low Complexity (LC) Profile, Levels 1, 2, and 3.

A timed text may be transported for providing subtitles and closed captions for omnidirectional video, which should conform to either W3C TTML profiles for Internet Media Subtitles and Captions 1.1 (IMSC1.1) or the Web Video Text Tracks format (WebVTT). The timed text may be either fixed positioned, in which the timed text does not move as the user's viewing orientation moves, or always-visible, in which the timed text is always visible to the user irrespective of the user's viewing orientation.

An initial viewing orientation may be transported as one of the timed metadata. It indicates an initial viewing orientation to enable the content author to indicate that a particular viewing orientation is recommended for users.

Annex 2
(informative)

Overview of OMAF standard specified in ISO/IEC 23090-2

ISO/IEC 23090-2 (Omnidirectional media format, OMAF) specifies a media format that enables omnidirectional media applications, focusing on 360° video, images, and audio, as well as associated timed text. It also specifies various technologies such as region-wise packing methods that may be used for improving the picture quality on the basis of information obtained from head/eye tracking at the user's device. The first version of OMAF supports 3DoF[1] VR/360° content. Figure 1 shows a typical content flow process for an omnidirectional media application with projected video.

---

[1]   Three Degrees of Freedom: Three rotational and un-limited movements around the X, Y, and Z axes (respectively pitch, yaw, and roll). A typical use case is a user sitting in a chair freely looking around in any direction.

FIGURE 1

Content flow process for omnidirectional media with projected video

(This figure is taken from ISO/IEC 23090-2.)
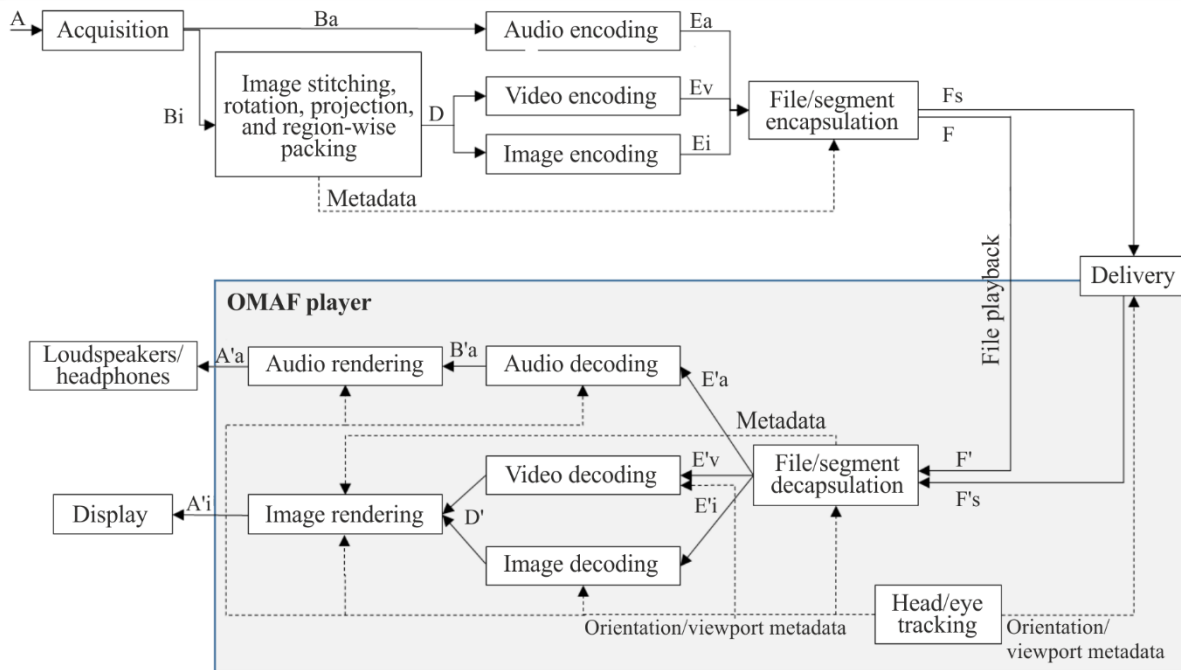
A: Real-world audio-visual scene
B: Multiple-sensors-captured video or audio
D/D': Projected/packed video
E/E': Coded video or audio bitstream
F/F': ISO Base Media File Format

BT.2133-01

The file/segment encapsulation process in Fig. 1 includes encapsulating metadata into the file or the segments. Examples of metadata are projection and region-wise packing information assisting in rendering the decoded packed pictures.

On the basis of the aforementioned flow process, OMAF has the following specifications:

– a coordinate system and the equations for the equirectangular and cubemap omnidirectional projections formats, the conversion from the local coordinate axes to the global coordinate axes, and the rectangular region-wise packing;

– extensions to the ISO Base Media File Format for omnidirectional media as well as for timed metadata for sphere regions;

– extensions to MMT and MPEG-DASH for omnidirectional media;

– media profiles such as the HEVC-based viewport-independent video profile, 3D audio baseline profile, HEVC image profile, and IMSC1.1 timed text profile;

– presentation profiles based on some of the OMAF media profiles such as the viewport-independent baseline presentation profile and viewport-dependent baseline presentation profile.
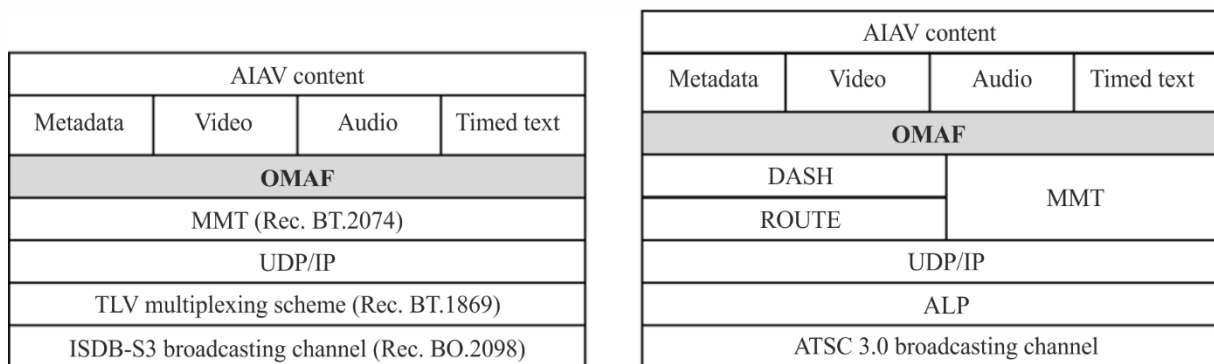
Annex 3
(informative)

Possible layer structures for transport of AIAV content in broadcasting systems

The OMAF specification stipulates delivery of OMAF content in MMT or MPEG-DASH. AIAV content encapsulated in OMAF can be delivered in IP-based broadcasting systems that adopt MMT or MPEG-DASH. Figure 2 shows possible layer structures for AIAV content in ISDB-S3 and ATSC 3.0. In these cases, 360° video with omnidirectional views is delivered to huge numbers of users simultaneously through unidirectional broadcasting channels.

FIGURE 2

Possible layer structure for transport of AIAV content in broadcasting systems

AIAV content			
Metadata	Video	Audio	Timed text
OMAF			
MMT (Rec. BT.2074)			
UDP/IP			
TLV multiplexing scheme (Rec. BT.1869)			
ISDB-S3 broadcasting channel (Rec. BO.2098)			

AIAV content			
Metadata	Video	Audio	Timed text
OMAF			
DASH	MMT		
ROUTE			
UDP/IP			
ALP			
ATSC 3.0 broadcasting channel			

* ROUTE: Real-Time Object Delivery over Unidirectional Transport
* ALP: ATSC 3.0 Link layer Protocol

BT.2133-02