ETSI TR 126 905 V17.0.0 (2022-05)

TECHNICAL REPORT

Universal Mobile Telecommunications System (UMTS);
LTE;
5G;
Mobile stereoscopic 3D video
(3GPP TR 26.905 version 17.0.0 Release 17)

A GLOBAL INITIATIVE

Reference
RTR/TSGS-0426905vh00

Keywords
5G,LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Legal Notice

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Intellectual Property Rights 2
Legal Notice 2
Modal verbs terminology 2
Foreword 6
1 Scope 7
2 References 7
3 Definitions and abbreviations 8
3.1 Definitions 8
3.2 Abbreviations 9
4 General 10
4.1 Introduction 10
5 Technology description 10
5.1 Mobile 3D rendering technologies 10
5.1.1 Introduction 10
5.1.2 Glasses-free 3D video rendering technologies 10
5.1.1.1 Parallax barrier 10
5.1.1.2 Lenticular lens sheet 11
5.1.3 Glasses-based 3D video rendering technologies 12
5.1.3.1 Active-shutter glasses 12
5.1.3.2 Passive glasses 12
5.1.4 Potential impacts on a 3D service implementation 13
5.2 Stereoscopic 3D frame packing formats 13
5.2.1 Frame-compatible packing formats 13
5.2.2 Full resolution per view packing formats 14
5.3 Video codecs for stereoscopic 3D 15
5.3.1 H.264/AVC for frame compatible packing formats 15
5.3.2 H.264/AVC for temporal interleaving packing format 15
5.3.3 MVC (Multiview Video Coding) 15
5.3.4 Performance evaluation of the compression efficiency 16
5.3.4.1 Simulation setup 16
5.3.4.2 Simulation results 17
5.4 3D signalling 24
5.4.1 SIP/SDP codec and format signalling 24
5.4.2 File format signalling 25
5.4.2.1 Introduction 25
5.4.2.2 Frame compatible H.264/AVC 25
5.4.2.3 Temporally interleaved H.264/AVC 25
5.4.2.4 Multiview Video Coding MVC 26
5.4.2.5 Mixed 2D/3D video 26
5.4.2.6 MIME type signalling for 3D stereoscopic video files 26
5.4.3 Device capability exchange signalling of supported 3D video codecs and formats 26
5.4.4 Inclusion of 3D video information in the DASH MPD 27
6 Streaming use cases 27
6.1 PSS and MBMS-based 3D video services 27
6.1.1 Use case description 27
6.1.2 Working assumptions and operation points 27
6.1.3 Technical analysis 28
6.2 DASH-based streaming of 3D content 28
6.2.1 Use case description 28
6.2.2 Working assumptions and operation points 28
6.2.3 Evaluation of DASH-based streaming with HTTP-caching 28
6.2.3.1 Introduction 28

6.2.3.2 Coding of VoD content items ... 29
6.2.3.3 Simulation model ... 30
6.2.3.4 Simulation results ... 31
6.3 Common provisioning of 2D and 3D content for download and streaming ... 32
6.3.1 Use case description ... 32
6.3.2 Working assumptions and operation points ... 33
6.3.3 Technical analysis ... 34
6.4 3D Timed Text and Graphics ... 34
6.4.1 Use case description ... 34
6.4.2 Working assumptions and operation points ... 35
6.4.3 Possible solution ... 35
6.5 2D/3D mixed contents service ... 36
6.5.1 Use case description ... 36
6.5.2 Working assumptions and operation points ... 37
6.5.3 Technical analysis ... 37
6.6 Service provisioning based on depth range of the 3D content ... 37
6.6.1 Use case description ... 37
6.6.2 Working assumptions and operation points ... 37
6.6.3 Possible solution ... 37

7 Download use cases ... 38
7.1 Download of 3D video ... 38
7.1.1 Use case description ... 38
7.1.2 Working assumptions and operation points ... 38
7.1.3 Technical analysis ... 38
7.2 Progressive download of 3D video ... 38
7.2.1 Use case description ... 38
7.2.2 Working assumptions and operation points ... 39
7.2.3 Technical analysis ... 39
7.3 Correct rendering of downloaded 3D video ... 39
7.3.1 Use case description ... 39
7.3.2 Working assumptions and operation points ... 39
7.3.3 Technical analysis ... 39

8 Use cases for further study ... 39
8.1 Introduction ... 39
8.2 3D video delivering based on 2D video warehouse ... 39
8.2.1 Use case description ... 39
8.2.2 Working assumptions and operation points ... 40
8.3 3D video conversational services ... 40
8.3.1 Use case description ... 40
8.3.2 Working assumptions and operation points ... 41
8.4 Multiple-party 3D video conference ... 41
8.4.1 Use case description ... 41
8.4.2 Working assumptions and operation points ... 42
8.5 3D video call fall back to legacy phone ... 42
8.5.1 Use case description ... 42
8.5.2 Working assumptions and operation points ... 43
8.5.3 Gap analysis on supporting 3D video call fallback between 3D video phones ... 43
8.6 3D video call fall back between 3D capable phones ... 43
8.6.1 Use case description ... 43
8.6.2 Working assumptions and operation points ... 43
8.6.3 Gap analysis on supporting 3D video call fallback between 3D video phones ... 43
8.7 3D content in messaging ... 44
8.7.1 Use case description ... 44
8.7.2 Working assumptions and operation points ... 44
8.8 3D service in the converged environment ... 44
8.8.1 Use case description ... 44
8.8.2 Working assumptions and operation points ... 45
8.9 Bitrate adaptation ... 45
8.9.1 Introduction ... 45
8.9.2 Restricted access bandwidth ... 45

8.9.3 Rate adaptation in PSS and DASH ... 45
8.9.4 Rate adaptation in MTSI ... 46
8.9.5 Rate adaptation due to shared radio resources ... 46
8.10 View scalability for graceful degradation ... 46
8.10.1 Introduction ... 46
8.10.2 Graceful degradation in MBMS when entering bad reception conditions ... 46
8.10.3 Graceful degradation in MTSI ... 46
8.10.4 Combined support of heterogeneous devices and graceful degradation ... 46
9 Mobile 3D subjective tests ... 47
9.1 Introduction ... 47
9.2 Test description ... 47
9.2.1 Video sources ... 47
9.2.2 Content preparation ... 47
9.2.2.1 Frame rate evaluation ... 47
9.2.2.2 Resolution evaluation ... 47
9.2.3 Encoding profiles ... 47
9.2.4 Subjective test conditions ... 48
9.2.4.1 Methodology ... 48
9.2.4.2 Implementation ... 48
9.2.4.3 Observers ... 48
9.3 Test results ... 49
9.3.1 Frame rate evaluation ... 49
9.3.2 Resolution evaluation ... 50
9.4 Conclusion of the test ... 50
10 Content re-targeting ... 50
10.1 Introduction ... 50
10.2 Down-sampling/Up-sampling ... 51
10.3 Extraction of depth map ... 51
10.4 Occlusion handling ... 52
10.5 Depth adjustment ... 52
10.6 Creation of the second view ... 53
11 Conclusions ... 53
11.1 Introduction ... 53
11.2 Frame Compatible Format for Stereoscopic Video Coding ... 53
11.3 Stereoscopic Multi-view Video Coding ... 54
Annex A: Change history ... 56
History ... 57

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document provides a study of stereoscopic 3D video services over 3GPP networks and terminals. Technical definitions, use case descriptions, working assumptions, subjective tests results and technical studies are presented.

This document identifies the gaps within the Release 10 3GPP specifications in order to enable the implementation of the mobile 3D video use cases.

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 26.114: "IP multimedia subsystem (IMS); Multimedia telephony, Media handling and interaction".

[3] 3GPP TS 26.234: "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs".

[4] 3GPP TS 26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs".

[5] 3GPP TS 26.247: "Transparent end-to-end Packet-switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH)".

[6] 3GPP TS 26.140: "Multimedia Messaging Service (MMS); Media formats and codecs".

[7] 3GPP TS 26.245: "Transparent end-to-end Packet-switched Streaming Service (PSS); Timed Text Format".

[8] 3GPP TS 26.430: "Timed Graphics".

[9] 3GPP TR 26.904: "Improved Video Coding Support".

[10] IETF RFC 3261: "SIP: Session Initiation Protocol".

[11] IETF RFC 3264: "An Offer/Answer Model with the Session Description Protocol (SDP)".

[12] IETF draft draft-ietf-payload-rtp-mvc-01: "RTP Payload Format for MVC Video".

[13] IETF personal draft draft-greevenbosch-mmusic-sdp-3d-format-002: "Signal 3D format".

[14] ITU-R Recommendation BT 1788: "Methodology for the subjective assessment of video quality in multimedia applications".

[15] 3GPP TS 26.244: "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format".

[16] ISO/IEC 14496-15: 2010: "Information technology – Coding of audio-visual objects – Part 15: Advanced Video Coding (AVC) file format".

[17] ISO/IEC 14496-12:2008/Amd2 | 15444-12:2008/Amd2: " Part 12: ISO base media file format AMENDMENT 2: Support for sub-track selection & switching, post-decoder requirements, and color information ISO base media file format".

[18] ISO/IEC 14496-12:2008 | 15444-12:2008: "Information technology – Coding of audio-visual objects – Part 12: ISO base media file format" | "Information technology – JPEG 2000 image coding system – Part 12: ISO base media file format".

[19] JM H.264/AVC Reference Software, http://iphome.hhi.de/suehring/tml/download/.

[20] JMVC H.264/MVC Reference Software, Version 8.5, March 2011.

[21] Proceedings of the IEEE: "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard", A. Vetro, T. Wiegand, and G. Sullivan, ,, vol. 99, no. 4, p. 626642, 2011.

[22] Doc. JVT-AE022:"Coding performance of stereo high profile for movie sequences, London, U.K., Joint Video Team (JVT)", T. Chen, Y. Kashiwagi, C. S. Lim, and T. Nishi , Jul. 2009.

[23] Broadcasting, IEEE Transactions on: "Studies on the bit rate requirements for a HDTV format with 1920x1080 pixel resolution, progressive scanning at 50 Hz frame rate targeting large flat panel displays," H. Hoffmann, T. Itagaki, D. Wood, and A. Bock, , vol. 52, no. 4, pp. 420–434, 2006.

[24] 3GPP TS 26.237: "IP Multimedia Subsystem (IMS) based Packet Switch Streaming (PSS) and Multimedia Broadcast/Multicast Service (MBMS) User Service; Protocols ".

[25] IETF RFC 6381: "The 'Codecs' and 'Profiles' Parameters for ``Bucket´´ Media Types", Gellens R., Singer D. and Frojdh P., August 2011.

[26] ISO/IEC 23009-1: "Information technology -- Dynamic adaptive streaming over HTTP (DASH) -- Part 1: Media presentation description and segment formats.

[27] Open Mobile Alliance: "User Agent Profile Version 2.0", February 2006.

[28] 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON): ,"Adaptive parallax for 3D television", Ide, K. and Sikora, T., 2010, vol., no., pp.1-4, 7-9 June 2010.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in TR 21.905 [1].

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in TR 21.905 [1].

AVC	Advanced Video Coding
DASH	Dynamic Adaptive Streaming over HTTP
HDMI	High-Definition Multimedia Interface
HTTP	Hypertext Transfer Protocol
IMS	IP Multimedia Subsystem
IP	Internet Protocol
IPD	Interpupillary Distance
IR	Infrared
LTE	Long Term Evolution
MBMS	Multimedia Broadcast/Multicast Services
MPD	Media Presentation Description
MTSI	Multimedia Telephony Services for IMS
MVC	Multiview Video Coding
PSS	Packet Switched Streaming Service
RTP	Real Time Protocol
SDP	Session Description Protocol

4 General

4.1 Introduction

This Technical Report provides a study on mobile 3D stereoscopic video in 3GPP. Use cases and technical solutions are investigated regarding a variety of setups using 3GPP's streaming, multicast/broadcast, download and progressive download as well as conversational services. Clause 5 provides a definition of the stereoscopic 3D video technologies and terminology as well as a video codecs performance comparison. Clauses 6 and 7 focus on use cases for which the working assumptions and the operation points are defined before providing a technical analysis, whereas clause 8 provides a set of use cases in which further study is required so as to identify the gaps. Clause 9 introduces subjective tests conducted on a 3D capable mobile terminal and clause 10 presents a generic approach for 3D content adaptation depending on the client terminal charasteristics. The conclusion summarizes the recommended way forward for the introduction of 3D stereoscopic video support in 3GPP specifications.

5 Technology description

5.1 Mobile 3D rendering technologies

5.1.1 Introduction

Stereoscopy is the method of combining two plane pictures in order to produce a depth perception by the human brain. Each eye seeing a different angle of a scene, the human visual system - with subjective assessments - is able to interpret the depth information.

In the scope of the present document, this section provides some information on how the rendering technologies provide the depth perception. These technologies are split into two categories; the glasses based systems and the glasses free systems.

5.1.2 Glasses-free 3D video rendering technologies

5.1.1.1 Parallax barrier

The parallax barrier consists in a grid placed over the screen. When electrically activated, this barrier prevents the eyes of the user from viewing all the pixels of the display such as depicted in the figure 1. The resulting quality of experience is half the resolution per view compared with the 2D mode (i.e. when the barrier is switched off).

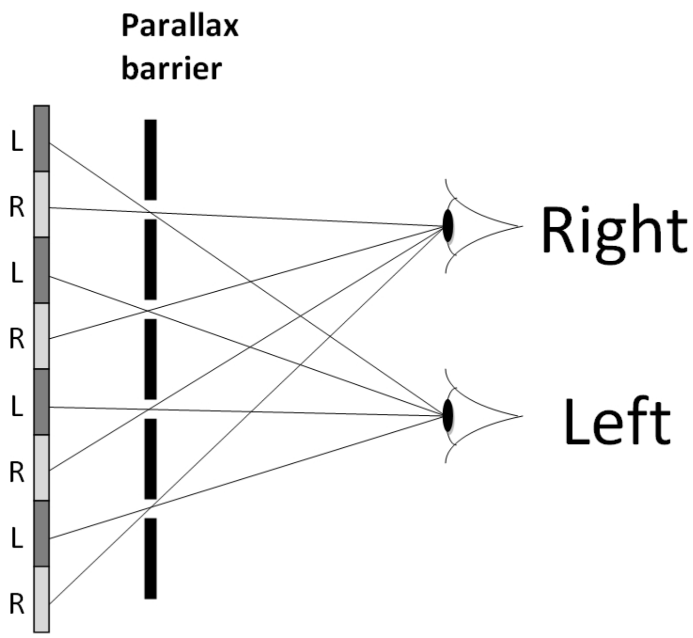

Figure 1: Parallax barrier

5.1.1.2 Lenticular lens sheet

This rendering technology is based on a lens sheet. It consists in a series of vertical hemi-cylindrical lenses placed so as to direct light in different viewing angles. When correctly placed, each eye can receive a different view from the other, as shown on the figure 2.

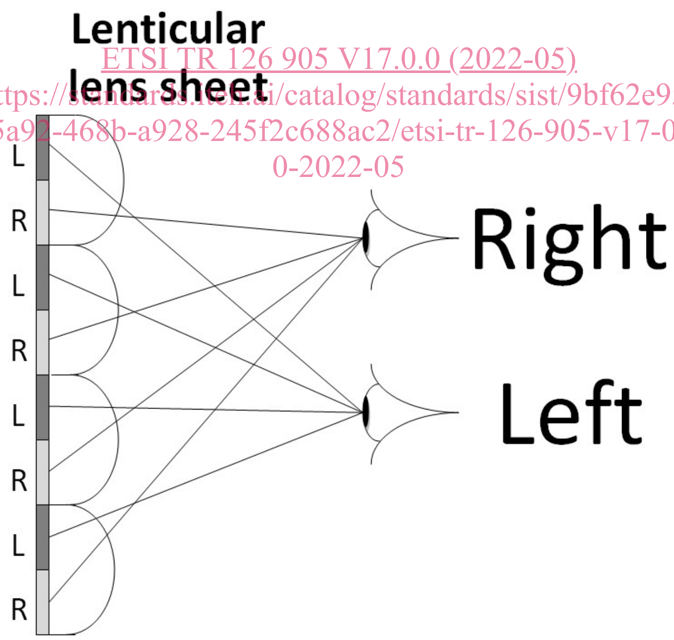

Figure 2: Lenticular lens sheet

5.1.3 Glasses-based 3D video rendering technologies

5.1.3.1 Active-shutter glasses

The active-shutter glasses are synchronized with the 3D display (potentially with IR signal transmitted from the glasses to the terminal) which displays alternatively the left and right views of a video. The figure 3 below illustrates such a case.

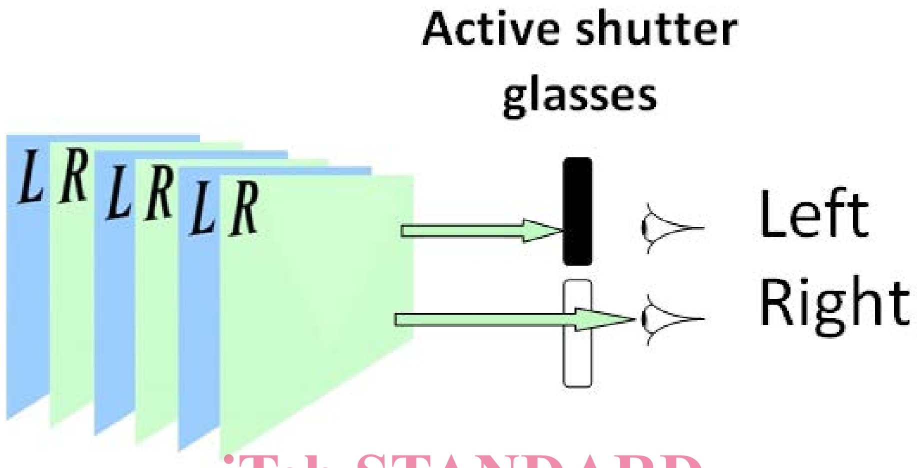

Figure 3: Active shutter glasses

5.1.3.2 Passive glasses

Passive glasses use a polarized filter placed on both the screen and the glasses. For example, the current 3D displays can interlace the left and right views in a single image on the screen whereas the filters on the glasses only allow the left eye to see the odd lines (in red on figure 4) and the right eye to see the even lines of the screen (in green on figure 4). In this case, image resolution is halved if compared to active systems but new systems such as active retarder will attempt to solve this problem.

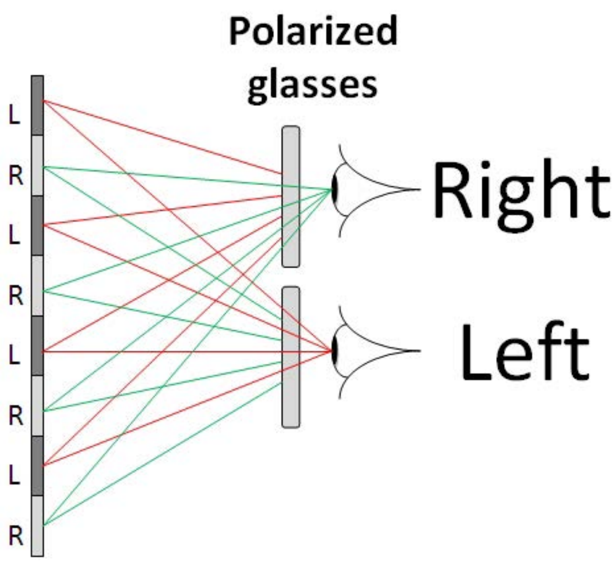

Figure 4: Passive polarized glasses

5.1.4 Potential impacts on a 3D service implementation

Given the fact that the rendering technologies offer different levels of quality of experience such as the resolution per view, the viewing angles… a service may benefit from adapting the provided 3D video format to the rendering technology in use. In this case appropriate signalling is necessary to either describe the different formats such that the client can select/request the format or the appropriate signalling of the rendering technology is important such that the server can select or annotate the appropriate format.

Depending on the service, these formats may have to be mapped to the different signalling frameworks in which the 3D video is offered, e.g. MPD in 3GP-DASH, SDP for MTSI and PSS, etc.

5.2 Stereoscopic 3D frame packing formats

5.2.1 Frame-compatible packing formats

The frame-compatible packing format consists in sub-sampling the two views which compose a stereoscopic 3D video and pack them together in order to produce a video signal compatible with a 2D frame infrastructure.

In a typical operation mode, the spatial resolution of the original frames of each view and the packaged single frame, have the same resolution. The spatial packing arrangement may use a side-by-side, top-bottom, interleaved, or checkerboard format as illustrated in figure 5 and the down-sampling process should be performed accordingly.

In most commercial deployments only side-by-side or top-bottom frame packing arrangements are applied.

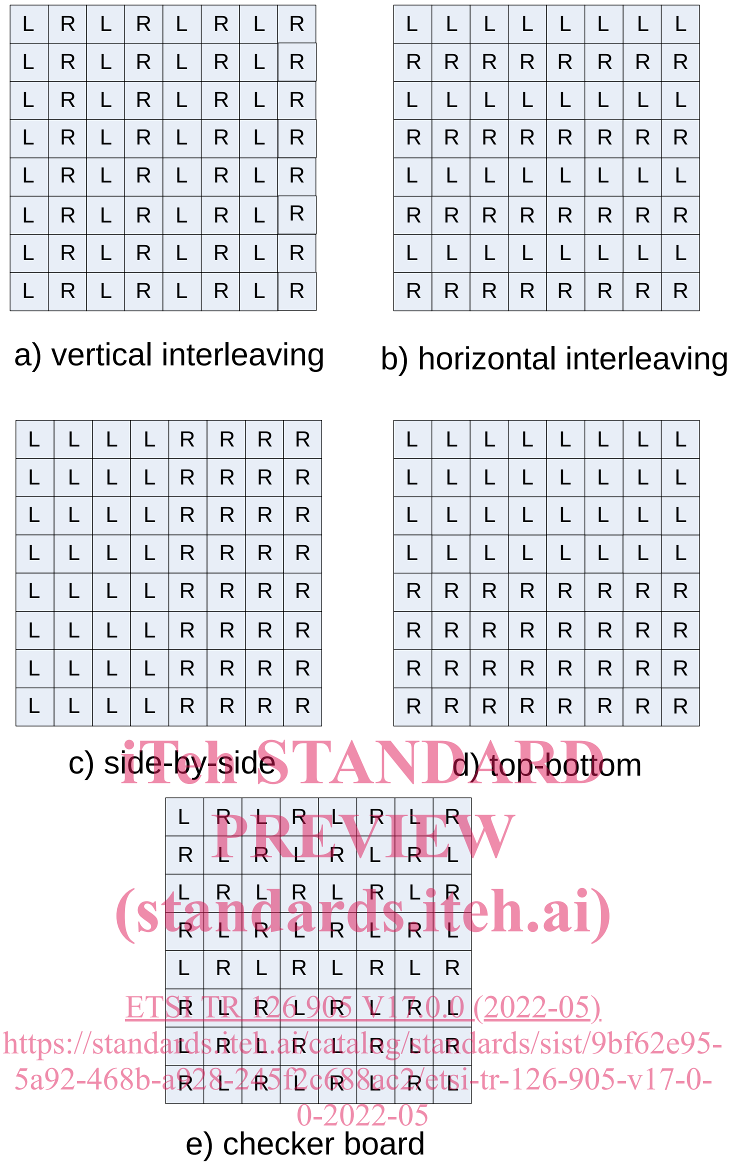

Figure 5: Spatial frame packing formats

5.2.2 Full resolution per view packing formats

In order to avoid the lack of definition introduced by the frame-compatible packing formats, it is possible to transmit both views at full resolution. In this case, the amount of data is twice as much as the frame compatible packing formats. Although the spatial packing format can be used in order to generate a twice bigger image, the most common format is the frame packing for which the left and right views are temporally interleaved such as shown on the figure 6 below.

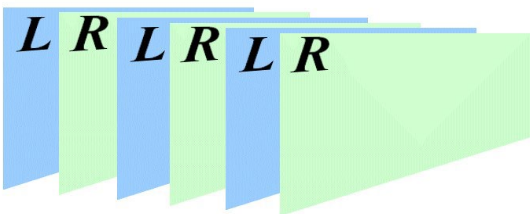

Figure 6: Temporal interleave packing format