SMPTE STANDARD

Stereoscopic 3D Frame Compatible Packing and Signaling for HDTV

Table of Contents

Page

3 Barker Avenue, White Plains, NY 10601
(914) 761-1100

Approved
July 29, 2013

Foreword

SMPTE (the Society of Motion Picture and Television Engineers) is an internationally-recognized standards developing organization. Headquartered and incorporated in the United States of America, SMPTE has members in over 80 countries on six continents. SMPTE's Engineering Documents, including Standards, Recommended Practices, and Engineering Guidelines, are prepared by SMPTE's Technology Committees. Participation in these Committees is open to all with a bona fide interest in their work. SMPTE cooperates closely with other standards-developing organizations, including ISO, IEC and ITU.

SMPTE Engineering Documents are drafted in accordance with the rules given in Part XIII of its Operations Manual.

SMPTE ST 2068 was prepared by Technology Committee 10E.

Intellectual Property

At the time of publication no notice had been received by SMPTE claiming patent rights essential to the implementation of this Standard. However, attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Specifically, one or more multiplexing methods signaled by this Standard may be the subject of patent rights. SMPTE shall not be held responsible for identifying any or all such patent rights.

Introduction

This section is entirely informative and does not form an integral part of this Engineering Document.

Stereoscopic 3D imaging systems deliver two images (left eye image and right eye image), that are arranged to be seen simultaneously or near simultaneously by the left and right eyes. Viewers then perceive increased depth in the picture, which becomes more like the natural binocular viewing experience.

One method of carrying stereoscopic signals in existing 2D HDTV systems is to multiplex two temporally or spatially sub-sampled images of a stereoscopic image pair into a single stream. This method of carriage is called the Frame Compatible packing method. This document provides a mechanism to signal the packing method used to create a Frame Compatible serial data stream. This signaling is based upon and is compatible with the H.264/MPEG-4 AVC Frame Packing Arrangement SEI message.

Annex B Frame Compatible Packing Methods (informative) provides illustrations of the frame packing methods covered by this standard.

Figure 1 illustrates the process of creating a frame compatible image stream from a Stereoscopic 3D image pair. The general order of operations that occur when two images of a stereoscopic image pair are packed into one frame compatible image are as follows:

(1) The input Stereoscopic serial data stream consisting of a stereoscopic image pair (Left Eye image and Right Eye image), as defined for example in SMPTE ST 292-2, Dual 1.5 Gb/s Serial Digital Interface for Stereoscopic Image Transport or SMPTE ST 425-4, Dual 3Gb/s Serial Digital Interface for Stereoscopic Image Transport, are first decoded into a Full resolution Left eye image (FrLe) and a Full resolution Right eye image (FrRe). The FrLe and FrRe images are then filtered prior to the application of temporal or spatial sub-sampling.

The sub-sampling process produces two half resolution sub-images — Half spatial resolution Left eye / Right eye (HrLe / HrRe), or Half Temporal resolution Left eye / Right eye (HTrLe / HTrRe). These sub images may be half the horizontal, vertical or temporal resolution of the originating full resolution stereoscopic image pair.

(2) The two HrLe and HrRe or HTrLe and HTrRe sub-images are then packed into an intermediate form in preparation for multiplexing into a single Frame Packed Image. The packing step may also format the sub-images by reversing the image orientation or pixel order. The two Frame Packed sub-images are then multiplexed together to form the frame packed image.

(3) The frame packed image (FpI) is then output as a frame compatible serial data stream in accordance with SMPTE ST 292-1, 1.5 Gb/s Serial Digital Interface or SMPTE ST 425-1, 3 Gb/s Serial Digital Interface.

(4) Information describing the sub-sampling, packing method and multiplexing (Packing Method Signaling), is appended to the frame compatible serial data stream to facilitate interchange of Frame Compatible signals.

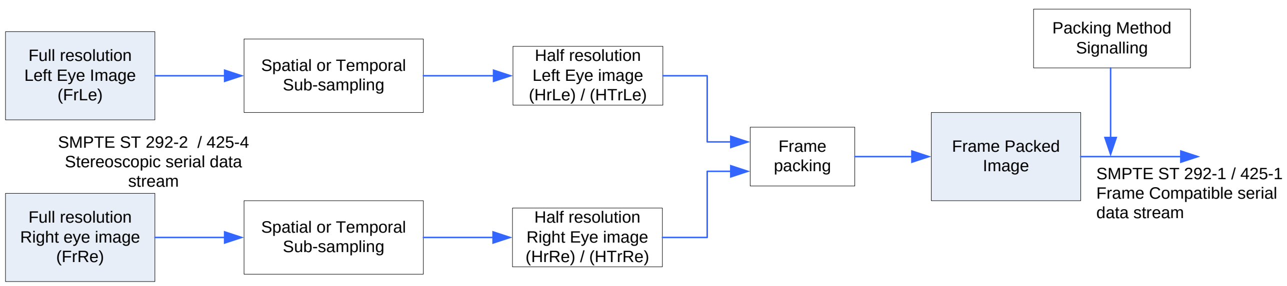

Figure 1 (Informative) – Overview of the frame packing operations

1 Scope

This standard defines a 3D Frame Compatible (3DFC) ancillary data packet that enumerates various frame packing methods in a manner compatible with Recommendation ITU-T H.264 (H.264) and ISO/IEC 14496-10 MPEG-4 AVC (MPEG-4 AVC), Frame Packing Arrangement SEI message.

This ancillary data packet provides information for the frame compatible transport of 10-bit stereoscopic image pairs (Left eye and Right eye images) — as defined in dual 1.5 Gb/s and 3 Gb/s serial digital interfaces for stereoscopic image transport — that are temporally or spatially sub-sampled and combined to form a frame compatible serial data stream in accordance with 1.5 Gb/s and 3 Gb/s serial digital interfaces.

The ancillary data packet is carried in the vertical ancillary (VANC) data space of the frame compatible serial data stream having the same pixel array structure as the originating stereoscopic image pair, and is used to signal the frame packing method utilized in order to facilitate interchange of Stereoscopic Frame Compatible signals.

The information contained in this ancillary data packet could then be used to set the H.264/MPEG-4 AVC Frame Packing Arrangement SEI message prior to delivery to a subsequent transmission delivery device.

The precise translation between the 3DFC ancillary data packet and the Frame Packing Arrangement SEI message is not within the scope of this standard. Nor does this standard address other additional H.264/MPEG-4 AVC SEI messages that may be required for transmission delivery of a 3D Frame Compatible signal.

Definition of the filtering mechanism is not within the scope of this standard since implementers can employ a variety of standard or proprietary filtering methods to improve image quality.

It is not necessary for implementations to include support for all frame packing arrangements defined in Table 4 to conform to this standard. Implementers should indicate supported frame packing arrangements in commercial publications.

2 Conformance Notation

Normative text is text that describes elements of the design that are indispensable or contains the conformance language keywords: "shall", "should", or "may". Informative text is text that is potentially helpful to the user, but not indispensable, and can be removed, changed, or added editorially without affecting interoperability. Informative text does not contain any conformance keywords.

All text in this document is, by default, normative, except: the Introduction, any section explicitly labeled as "Informative" or individual paragraphs that start with "Note:”

The keywords "shall" and "shall not" indicate requirements strictly to be followed in order to conform to the document and from which no deviation is permitted.

The keywords, "should" and "should not" indicate that, among several possibilities, one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required; or that (in the negative form) a certain possibility or course of action is deprecated but not prohibited.

The keywords "may" and "need not" indicate courses of action permissible within the limits of the document.

The keyword “reserved” indicates a provision that is not defined at this time, shall not be used, and may be defined in the future. The keyword “forbidden” indicates “reserved” and in addition indicates that the provision will never be defined in the future.

A conformant implementation according to this document is one that includes all mandatory provisions ("shall") and, if implemented, all recommended provisions ("should") as described. A conformant implementation need not implement optional provisions ("may") and need not implement them as described.

Unless otherwise specified, the order of precedence of the types of normative information in this document shall be as follows: Normative prose shall be the authoritative definition; Tables shall be next; followed by formal languages; then figures; and then any other language forms.

3 Normative References

The following standards contain provisions which, through reference in this text, constitute provisions of this standard. At the time of publication, the editions indicated were valid. All standards are subject to revision, and parties to agreements based on this recommended practice are encouraged to investigate the possibility of applying the most recent edition of the standards indicated below.

SMPTE RP 168:2009, Definition of Vertical Interval Switching Point for Synchronous Video Switching

SMPTE ST 291-1:2011, Ancillary Data Packet and Space Formatting

SMPTE ST 292-1:2012, 1.5 Gb/s Signal/Data Serial Interface

SMPTE ST 425-1:2011, Source Image Format and Ancillary Data Mapping for the 3 Gb/s Serial Interface

4 Definition of Terms

4.1
Frame Compatible
For the purposes of this standard, a frame compatible serial data stream means that the left-eye and right-eye images of a stereoscopic image pair are arranged in a spatial or temporal multiplex which results in a single image stream that can be treated like a conventional 2D serial data stream. The frame compatible serial data stream may be progressive, interlaced or progressive segmented frame in accordance with the format of the originating stereoscopic image pair.

5 3DFC Ancillary Data Packet — Specification and Carriage

5.1 Format of the Ancillary Data Packet

The 3DFC Ancillary Data Packet shall be a Type 2 ANC packet as defined in SMPTE ST 291-1. It consists of the ancillary data flag (ADF), the data ID (DID), the secondary data ID (SDID), the data count (DC), the user data words (UDW), and the checksum (CS).

The UDW consists of the data payload that describes the Frame Packing method employed for the frame compatible transport of the stereoscopic image pair.

The DID word shall be set to the value 41_h. The SDID word shall be set to the value of $0A_h$.

DC is a count of the number of words in the UDW and shall be set to a fixed value of 03_h.

5.2 Location and Timing of the Ancillary Data Packet

The 3DFC Ancillary Data Packet shall be located in the active line portion of the vertical ancillary space of the frame compatible serial data stream, in accordance with the image formats referenced by SMPTE ST 292-1, or SMPTE ST 425-1. Data may be located in any lines in the VANC area from the second line after the line specified for switching, as defined in SMPTE RP 168, to the last line before active video, inclusive.

The 3DFC Ancillary Data Packet shall describe the Frame Packing method employed for the current frame of the frame compatible serial data stream.

The packet shall be located in the luma (Y') data channel, for SMPTE ST 292-1 compliant frame compatible data streams.

The packet shall be mapped in accordance with the ancillary data mapping rules defined in SMPTE ST 425-1 for a frame compatible data stream compliant with SMPTE ST 425-1.

For interlaced video systems, the data packets shall be placed in the VANC area of the first field and in the VANC area of the second field. For progressive segmented frame systems, the data packets shall be placed in the VANC area of the first segment of the frame and may be in the VANC area of the second segment. For progressive formats, the ANC packet shall be placed in the VANC area of the frame.

Receiving equipment should identify the ANC packets on the basis of their ANC DID and SDID fields.

5.3 UDW Format

Three bytes of Frame Compatible signaling data shall be carried in the User Data Words of the 3DFC ancillary data packet as shown in Table 1.

The UDW shall be a sequence of three 10-bit words, and the Frame Compatible signaling data is transmitted in bits b7 through b0 of the 10-bit data word. Bit b8 is even parity for bits b7 through b0 of the 10-bit data word, and bit b9 equals the complement of bit b8.

Table 1 – ANC data payload

User Data Words		**User Data Word bits**									
UDW	**Function**	**b9**	**b8**	**b7**	**b6**	**b5**	**b4**	**b3**	**b2**	**b1**	**b0**
1	FC UDW 1	$\overline{b8}$	P	fpa3	fpa2	fpa1	fpa0	QSF	FPS	sv1	sv0
2	GRID Position	$\overline{b8}$	P	f1y1	f1y0	f1x1	f1x0	f0y1	f0y0	f0x1	f0x0
3	FC UDW 3	$\overline{b8}$	P	sf1	sf0	ci1	ci0	0	SP	CFR	FV

where:

sv1, sv0	is the Syntax Version code as defined in Table 2
FPS	is the Frame Packed Flag as defined in Table 3
QSF	is the Quincunx Sampling Flag as defined in Table 5
fpa3 - fpa0	is the Frame Packing Arrangement code as defined in Table 4
f0x1, f0x0	is the HrLe sample Position X as defined in Table 7
f0y1, f0y0	is the HrLe sample Position Y as defined in Table 8
f1x1, f1x0	is the HrRe sample Position X as defined in Table 9
f1y1, f1y0	is the HrRe sample Position Y as defined in Table 10
FV	is the Field Views Flag as defined in §5.5.1
CFR	is the Current Frame Reference Flag as defined in Table 14
SP	is the Sampling Phase flag as defined in Table 6
ci1, ci0	is the Content Interpretation Type as defined in Table 12
sf1, sf0	is the Spatial Flipping Type as defined in Table 13
P	is the even parity bit for b7 through b0

5.3.1 Syntax Version (sv1-sv0) – b1:b0 UDW 1

Two bits in the UDW1 byte are used to signal different versions of the mapping syntax for the ancillary data packet to decoding devices. Bit 0 shall be set to zero and bit 1 shall be set to zero for the mapping syntax described in this standard.

Table 2 – Syntax Version

sv1	sv0	Frame Reference
0	0	Syntax version 0 (Current)
0	1	Reserved
1	0	Reserved
1	1	Reserved

5.3.2 Frame Packed Signal (FPS) – b2 UDW 1

The Frame Packed Signal flag indicates whether the serial data stream containing the 3DFC ancillary data packet caries a conventional 2D image or a frame compatible image. The Frame Packed Signal values are shown in Table 3. When the Frame Packed Signal is set to zero, the rest of the values described in the following sections of this standard shall be set to zero.

Table 3 – Frame Packed Signal

Value	Frame Packed Signal
0	Image is not a frame packed stereoscopic image
1	Image is a frame packed stereoscopic image

5.3.3 Frame Packing Arrangement (fpa3-fpa0) – b7:b4 UDW 1

The Frame Packing Arrangement parameter is used to describe the overall frame packing method. Table 4 enumerates the frame packing methods supported by this standard.

Table 4 – Frame Packing Arrangement

Value	Short Name	Frame Packing Arrangement
0	Reserved	This value is reserved for future expansion.
1	Reserved	This value is reserved for future expansion.
2	Reserved	This value is reserved for future expansion..
3	Side-by-side	Each component plane of the frame compatible serial data stream contains a side-by-side packing arrangement of corresponding HrLe and HrRe images as illustrated in Annex B, Figure B.1 and Figure B.2
4	Top-and-bottom	Each component plane of the frame compatible serial data stream contains a top and-bottom packing arrangement of corresponding HrLe and HrRe images as illustrated in Annex B, Figure B.3
5	Temporal Interleave	The component planes of the frame compatible serial data stream contains a temporal interleaving of corresponding HTrLe and HTrRe images as illustrated in Annex B, Figure B.4
6-15	Reserved	These values are reserved for future expansion.

Refer to informative Annex B, Frame Compatible Packing Methods, for graphical representations of the frame packing arrangements enumerated in Table 4.

5.4 Sampling Method

5.4.1 Quincunx Sampling Flag (QSF) – b3 UDW 1

The Quincunx Sampling Flag indicates whether the sub-sampling is unidirectional or is bidirectional. The Quincunx Sampling Flag shall be set to zero when the pixels are sampled in either the horizontal or vertical direction only. The Quincunx Sampling Flag shall be set to one when the pixels are sampled in both the horizontal and vertical directions (i.e., in a quincunx pattern).

Quincunx sampling shall be applied only for the side-by-side packing arrangement.

Table 5 – Quincunx Sampling

Value	Sampling Structure
0	Pixels are not quincunx sampled
1	Pixels are quincunx sampled

5.4.2 Sampling Phase Flag (SP) – b2 UDW 3

The Sampling Phase Flag specifies whether there is constant phase between the sub-sampling sites and the original sampling grid.

Table 6 – Sampling Phase Flag

Value	Sampling Phase
0	Constant
1	Variable

For constant phase sub-sampling, where the sampling ratio is exactly 2:1 within a given sampling dimension, the grid position should be specified (see Section 5.4.3).

For variable phase sub-sampling, within a given sampling dimension, the first sample of the sub-sampled image is co-sited with the first pixel of the FrLe or FrRe image and the last sample of the sub-sampled image is co-sited with the last pixel of the FrLe or FrRe image. The remaining samples of the sub-sampled image are evenly spaced between the first and last pixels of the FrLe or FrRe image. In this case, the grid position parameters should all be set to their unspecified value.

5.4.3 Grid Position

This section describes the sampling position in spatial coordinates of the HrLe and HrRe image pixels relative to the original FrLe or FrRe image pixel locations where the FrLe or FrRe image is sub-sampled in an exact 2:1 ratio. The spatial coordinates refer to the location of the pixels in the original image prior to sub-sampling and image flipping (see Section 5.5.3 Spatial Flipping Flags). For the purposes of this document, the pixels in the image are considered to start at (0,0) in the upper left corner. The lower right corner is the furthest extent of the underlying image i.e. (1919,1079) for 1920 x 1080 images and (1279,719), for 1280 x 720 images

Figure 2 illustrates the potential sampling sites relative to the original sampling grid. The lower case e and o beneath the sample locations represent even and odd sampling respectively. There are three potential locations where samples may be drawn from. They are co-sited with even numbered pixels, co-sited with odd numbered pixels or they are drawn from a location mid-way between the even numbered and odd numbered pixels. Figure 2 is used to develop the grid positions for Sections 5.4.3.1 through 5.4.3.4.

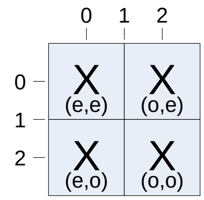

Figure 2 – Possible sampling locations

Color difference sampling follows the same pattern as luma sampling.

The spatial location reference information is not provided when the Quincunx Sampling Flag (QSF) is equal to 1 because the spatial alignment in this case is assumed to be such that the HrLe and HrRe images cover corresponding spatial areas with interleaved quincunx sampling patterns. In this case, the grid position parameters should all be set to their unspecified value.

5.4.3.1 HrLe image Horizontal Grid Position (f0x1- f0x0) – b1:b0 UDW 2

Table 7 contains the spatial location reference information for the horizontal position of the HrLe Image with respect to the upper left corner of the FrLe image.

Table 7 – HrLe Horizontal Grid Position value

Value	Spatial location of the HrLe samples relative to the FrLe image
0	Co-sited with even numbered samples
1	Between the even and odd numbered samples
2	Co-sited with odd numbered samples
3	Unspecified

Note: This parameter corresponds to the H.264 Frame Packing Arrangement SEI message syntax “frame0_grid_position_x”.

5.4.3.2 HrLe image Vertical Grid Position (f0y1- f0y0) – b3:b2 UDW 2

Table 8 contains the spatial location reference information for the vertical position of the HrLe Image with respect to the upper left corner of the FrLe image.

Table 8 – HrLe Vertical Grid Position value

Value	Spatial location of the HrLe samples relative to the FrLe image
0	Co-sited with even numbered samples
1	Between the even and odd numbered samples
2	Co-sited with odd numbered samples
3	Unspecified

Note: This parameter corresponds to the H.264 Frame Packing Arrangement SEI message syntax “frame0_grid_position_y”.

5.4.3.3 HrRe image Horizontal Grid Position (f1x1- f1x0) – b5:b4 UDW 2

Table 9 contains the spatial location reference information for the horizontal position of the HrRe Image with respect to the upper left corner of the FrRe image.

Table 9 – HrRe Horizontal Grid Position value

Value	Spatial location of the HrRe samples relative to the FrRe image
0	Co-sited with even numbered samples
1	Between the even and odd numbered samples
2	Co-sited with odd numbered samples
3	Unspecified

Note: This parameter corresponds to the H.264 Frame Packing Arrangement SEI message syntax “frame1_grid_position_x”.

5.4.3.4 HrRe image Vertical Grid Position (Y) (f1y1- f1y0) – b7:b6 UDW 2

Table 10 contains the spatial location reference information for the vertical or Y position of the HrRe Image with respect to the upper left corner of the FrRe image.

Table 10 – HrRe Grid Position Y value

Value	Spatial location of the HrRe samples relative to the FrRe image
0	Co-sited with even numbered samples
1	Between the even and odd numbered samples
2	Co-sited with odd numbered samples
3	Unspecified

Note: This parameter corresponds to the H.264 Frame Packing Arrangement SEI message syntax “frame1_grid_position_y”.

5.4.4 H.264 Frame Packing SEI message parameter translations (informative)

The H.264 Frame Packing SEI message signals the sample grid position in a manner different than this document. The grid positions are stored relative to the decoded and reconstituted image. Table 11 contains a translation from this document to the SEI message.

Table 11 – Mapping from the sampling grid positions to the Frame Packing SEI message

	Side-by-side		Top-and-bottom	
SMPTE	X Value	Y Value	X Value	Y Value
0	4	8	8	4
1	8	8	8	8
2	12	8	8	12
3	0	0	0	0

5.5 Pixel Packing Parameters

5.5.1 Field Views Flag (FV) – b0 UDW 3

This flag shall be set to zero for all applications compliant with this standard.

5.5.2 Content Interpretation Type (ci1-ci0) – b5:b4 UDW 3

The Content Interpretation Type parameter specifies the order in which the HrLe and HrRe images are packed, as shown in Figure 3 and Figure 4 for side-by-side orientation (packing arrangement 3). The same paradigm may be applied to other packing arrangements.

The order in which the HTrLe and HTrRe images are packed in the Temporal Interleave packing arrangement is specified in Section 5.5.4.

Table 12 – Content Interpretation Type

Value	
0	Unspecified relationship between the frame packed constituent frames
1	Indicates the top left pixel is derived from the HrLe image.
2	Indicates the top left pixel is derived from the HrRe image.

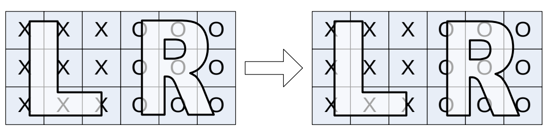

Figure 3 – Content Interpretation Type 1

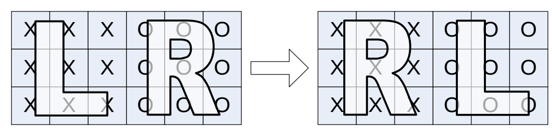

Figure 4 – Content Interpretation Type 2

5.5.3 Spatial Flipping Flags (sf1-sf0) - b7:b6 UDW 3

The Spatial Flipping flags specify whether the HrLe and HrRe images are flipped and if so, which of the two images are flipped. The sf1 flag indicates if one of the images is flipped, and the sf0 flag indicates which image is flipped. The spatial flipping flags are useful for packing arrangements 3 or 4 specified in Section 5.3.3. Figure 5 illustrates an example using side-by-side packing when neither view is flipped (sf1:sf0 = 00). Figure 6 illustrates an example when view flipping is enabled and the HrRe image is flipped (sf1:sf0 = 10).

Table 13 – Spatial Flipping Flags

sf1	sf0	**Spatial Orientation**
0	0	Images are not spatially flipped (Example 1)
0	1	Reserved
1	0	HrRe image is spatially flipped (Example 2)
1	1	HrLe image is spatially flipped

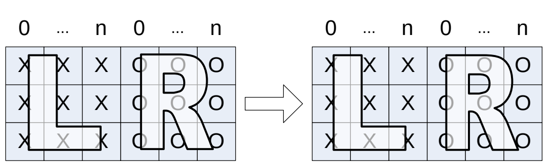

Figure 5 – Example 1

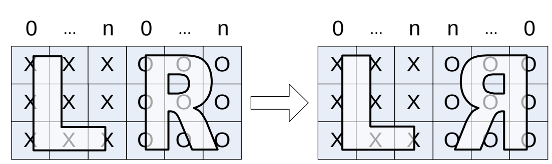

Figure 6 – Example 2

5.5.4 Current Frame Reference Flag (CFR) – b1 UDW 3

The Current Frame Reference flag specifies which HTrLe or HTrRe image is currently being presented in temporal Interleave mode (packing arrangement 5) as specified in Section 5.3.3.

This flag has no meaning for any of the other packing modes in which case it shall be set to 0.

Table 14 – Current Frame Reference

Value	Frame Reference
0	Current Frame is the HTrLe image
1	Current Frame is the HTrRe image

Annex A Bibliography (Informative)

Note: All references in this document to other SMPTE documents use the current numbering style (e.g. SMPTE ST 274:2008) although, during a transitional phase, the document as published (printed or PDF) may bear an older designation (such as SMPTE 274M-2008). Documents with the same root number (e.g. 274) and publication year (e.g. 2008) are functionally identical.

SMPTE ST 274:2008, Television — 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates

SMPTE ST 292-2:2011, Dual 1.5 Gb/s Serial Digital Interface for Stereoscopic Image Transport

SMPTE ST 296:2012, 1280 x 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure — Analog and Digital Representation and Analog Interface

SMPTE ST 424:2012, 3 Gb/s Signal/Data Serial Interface

SMPTE ST 425-2:2012, Source Image Format and Ancillary Data Mapping for Stereoscopic Image Formats on a Single-Link 3 Gb/s Serial Interface

SMPTE ST 425-4:2012, Dual 3 Gb/s Serial Digital Interface for Stereoscopic Image Transport

CableLabs OC-SP-CEP3.0-I02-110131, Content Encoding Profiles 3.0 Specification, 2011

Recommendation ITU-T Rec. H.262 | ISO/IEC 13818-2 (2000), Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Video, with Amendment 4 (2012)

Recommendation ITU-T Rec. H.264 | ISO/IEC 14496-10 (2010), Information Technology — Coding of Audio-Visual Objects — Part 10: Advanced Video Coding

Annex B 3D Frame Compatible Packing Methods — Graphical Representation (Informative)

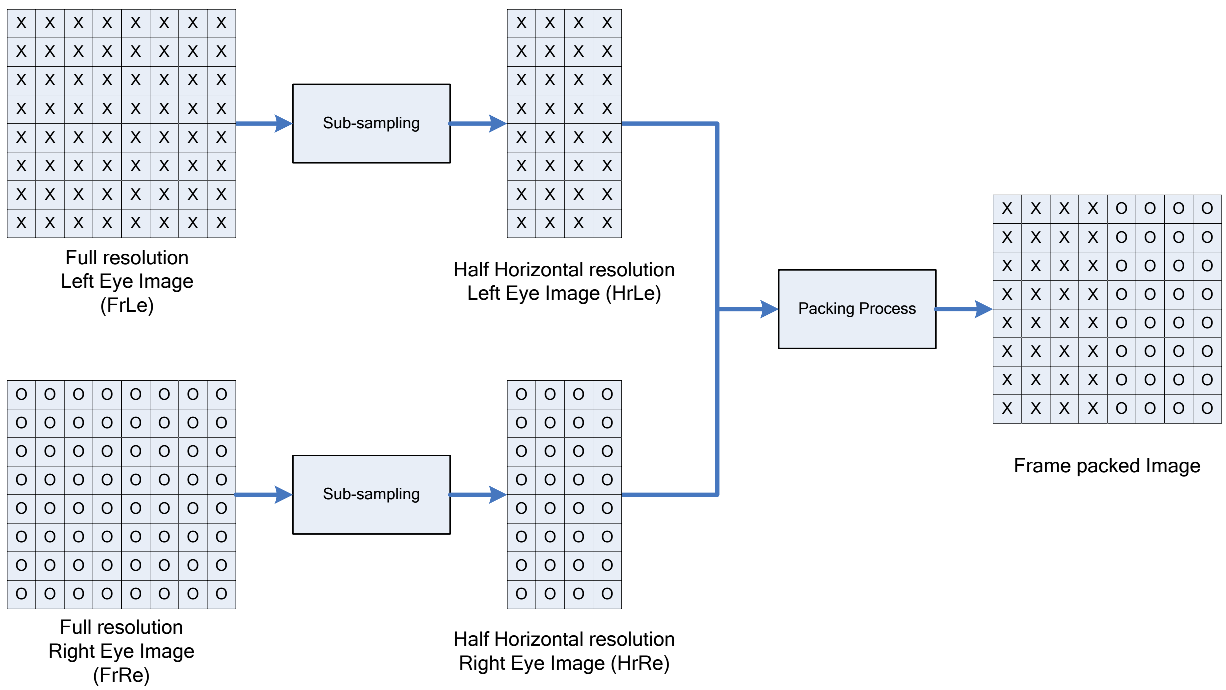

Figure B.1 – Side-by-Side — ½ resolution image is formed by horizontal sub-sampling

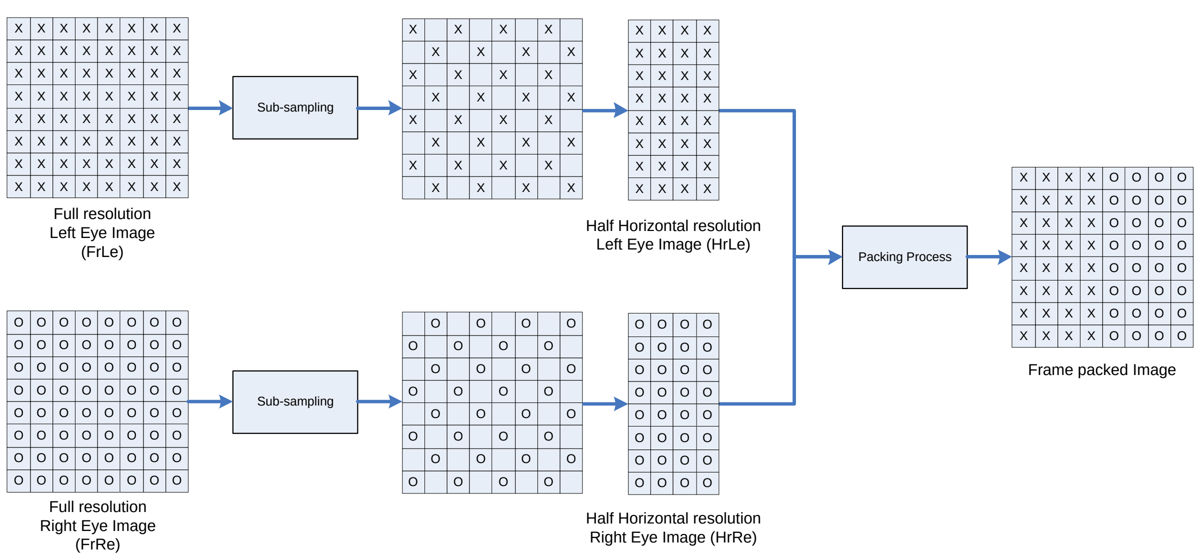

Figure B.2 – Side-by-Side — ½ resolution image is formed by Quincunx sub-sampling

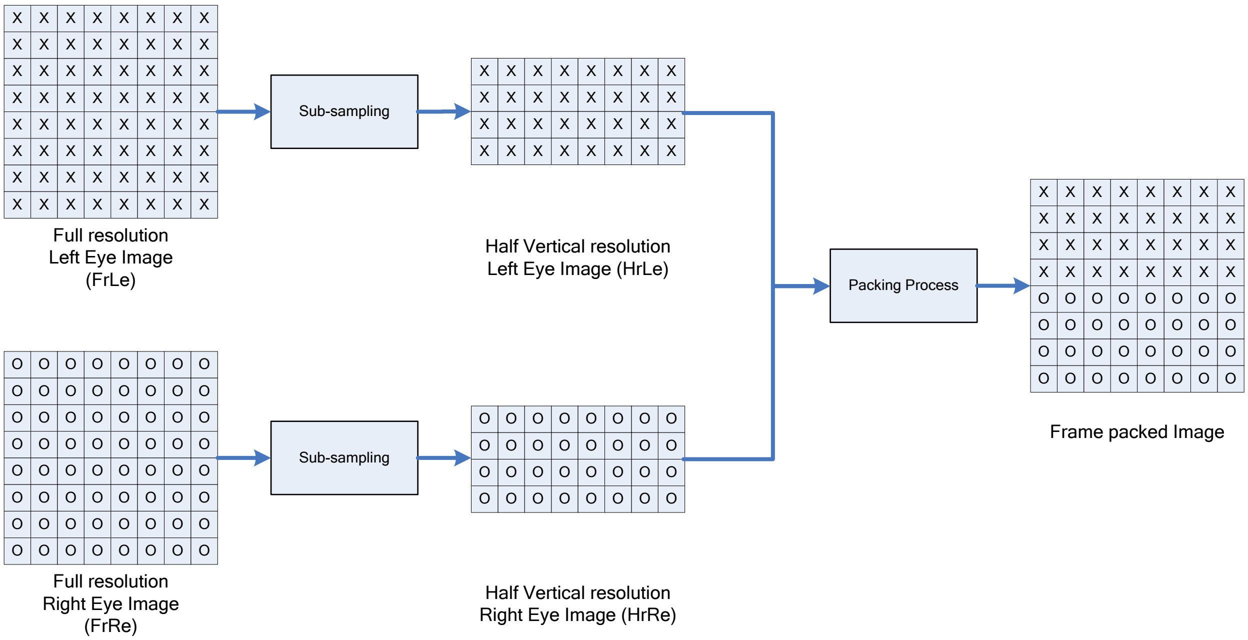

Figure B.3 – Top-and-Bottom — ½ resolution image is formed by vertical sub-sampling

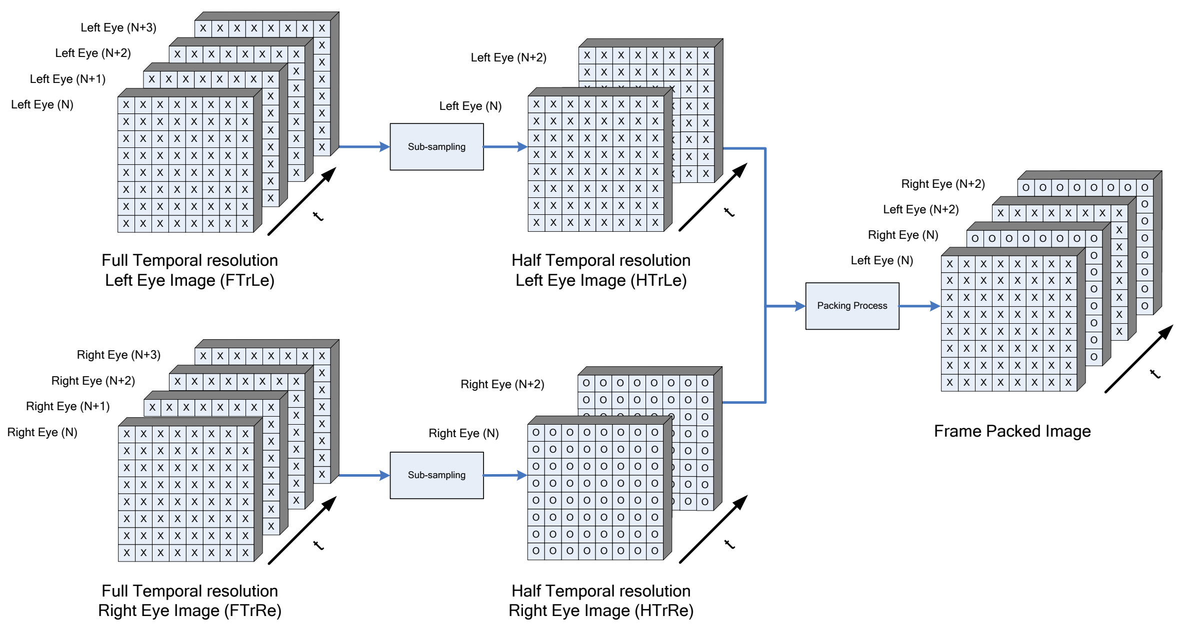

Figure B.4 – Temporal Interleave — ½ resolution image is formed by temporal sub-sampling