3D-TV Terrestrial Broadcasting, Part 5 – Service Compatible 3D-TV using Main and Mobile Hybrid Delivery

Doc. A/104 Part 5
28 October 2015

Advanced Television Systems Committee
1776 K Street, N.W.
Washington, D.C. 20006
202-872-9160

The Advanced Television Systems Committee, Inc., is an international, non-profit organization developing voluntary standards for digital television. The ATSC member organizations represent the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

Specifically, ATSC is working to coordinate television standards among different communications media focusing on digital television, interactive systems, and broadband multimedia communications. ATSC is also developing digital television implementation strategies and presenting educational seminars on the ATSC standards.

ATSC was formed in 1982 by the member organizations of the Joint Committee on InterSociety Coordination (JCIC): the Electronic Industries Association (EIA), the Institute of Electrical and Electronic Engineers (IEEE), the National Association of Broadcasters (NAB), the National Cable and Telecommunications Association (NCTA), and the Society of Motion Picture and Television Engineers (SMPTE). Currently, there are approximately 120 members representing the broadcast, broadcast equipment, motion picture, consumer electronics, computer, cable, satellite, and semiconductor industries.

ATSC Digital TV Standards include digital high definition television (HDTV), standard definition television (SDTV), data broadcasting, multichannel surround-sound audio, and satellite direct-to-home broadcasting.

Note: The user's attention is called to the possibility that compliance with this standard may require use of an invention covered by patent rights. By publication of this standard, no position is taken with respect to the validity of this claim or of any patent rights in connection therewith. One or more patent holders have, however, filed a statement regarding the terms on which such patent holder(s) may be willing to grant a license under these rights to individuals or entities desiring to obtain such a license. Details may be obtained from the ATSC Secretary and the patent holder.

Revision History

Version	Date
Candidate Standard approved	1 January 2014
A/104 Part 5:2014 Standard approved	29 August 2014
Revision 1 approved	28 October 2015

Table of Contents

1. SCOPE 5
  1.1 Documentation Structure 5
  1.2 Introduction and Background 5
  1.3 Organization 5
2. REFERENCES 5
  2.1 Normative References 6
  2.2 Informative References 6
3. DEFINITION OF TERMS 6
  3.1 Compliance Notation 6
  3.2 Treatment of Syntactic Elements 7
    3.2.1 Reserved Elements 7
  3.3 Acronyms and Abbreviation 7
  3.4 Terms 7
4. SERVICE COMPATIBLE 3D-TV USING MAIN AND MOBILE HYBRID DELIVERY 8
  4.1 Overall Description of SC-MMH 8
  4.2 Encoding and Decoding 8
  4.3 Video Format 8
  4.4 Service Multiplex and Transport 9
  4.5 Audio 10
  4.6 Closed Captioning for SC-MMH 10
  4.7 Signaling for SC-MMH 10
    4.7.1 PSI 10
    4.7.2 PSIP 12
    4.7.3 Signaling at 2D/3D Boundaries 12
ANNEX A: ASPECT RATIO ADJUSTMENT OF THE ATSC MOBILE DTV VIDEOS 14
ANNEX B: VIDEO ENHANCEMENT INFORMATION 15
  B.1 System Overview of SC-MMH with Video Enhancement Information 15
  B.2 References 16
  B.3 Definition of Terms 16
  B.3.1 Acronyms and Abbreviation 16
  B.3.2 Terms 16
  B.4 Bit Stream Syntax and Semantics 16
  B.4.1 VEI Bit Stream 16
  B.4.2 VEI Frame 16
ANNEX C: DECODING PROCESS OF VIDEO ENHANCEMENT INFORMATION 21
  C1. Decoding Process 21
  C.1.1 Decoding VEI_frame_header and Pre-processing 22

Index of Tables and Figures

Table 4.1 List and Location of SC-MMH Descriptors 10
Table 4.2 MDTV Hybrid Stereoscopic Service Descriptor 11
Table 4.3 Stereoscopic Service Types 11
Table 4.4 Example TVCT Composition 12
Table A.1 Examples of Aspect Ratio Adjustment before 3D Presentation 14
Table B.1 VEI frame Syntax 17
Table B.2 VEI_frame_header Syntax 17
Table B.3 VEI_frame_type 17
Table B.4 LPU_level 18
Table B.5 SPU_level 18
Table B.6 VEI_frame_payload Syntax 19
Table B.7 Signed Exp-Golomb Code Table 19
Table B.8 Record Mode of PU Except Level ‘0’ 20
Table B.9 Trailing Bits Syntax 20

Figure 1.1 Concept of SC-MMH. 5
Figure 4.1 Overview of SC-MMH system. 8
Figure B.1 Image Quality Enhancement of SC-MMH System with VEI. 15
Figure C.1 Decoding process of VEI for a frame. 21
Figure C.2 The prediction directions of IVV. 23

3D-TV Terrestrial Broadcasting, Part 5 – Service Compatible 3D-TV using Main and Mobile Hybrid Delivery

1. SCOPE

This document provides detailed specification of the parameters of the Service Compatible 3D-TV using Main and Mobile Hybrid Delivery (SC-MMH) which is one particular case of the Service Compatible 3D-TV system using Real-time Delivery (SCRT). This specification includes the stereoscopic video formats, the service multiplex and transport layer characteristics, and normative/informative specifications.

1.1 Documentation Structure

This document provides a general overview, technical description of SC-MMH system and a list of reference documents.

1.2 Introduction and Background

3D-TV broadcasting service using SC-MMH specification consists of Stereoscopic 3D video, audio and ancillary data. In SC-MMH, The Stereoscopic 3D video has left view and right view, where two views can be used as 2D videos for the ATSC main 2D-TV service and ATSC Mobile DTV service, respectively. The Stereoscopic 3D video is transmitted as two independent video elementary streams, where one of them is compatible with the legacy 2D TV service based on ATSC A/53 and the other video is compatible with ATSC Mobile DTV service based on ATSC A/153. In this manner, SC-MMH fully guarantees backward-compatibility with the existing fixed and mobile 2D services. Figure 1.1 illustrates the concept of SC-MMH.

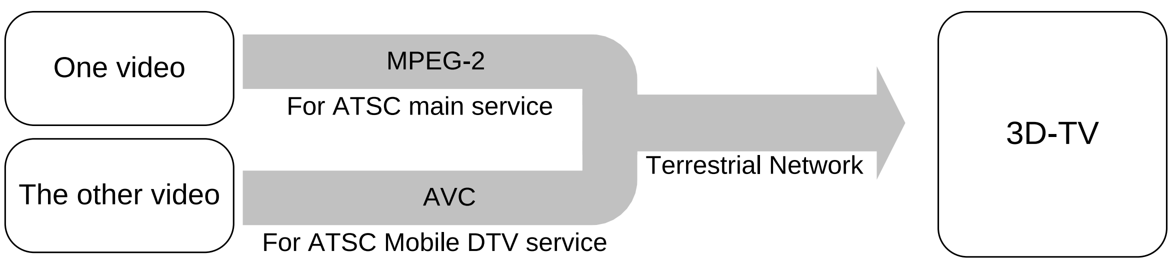

Figure 1.1 Concept of SC-MMH.

1.3 Organization

This document is organized as follows:

- Section 1 – Scope of this document and a general introduction.
- Section 2 – List of references and applicable documents.
- Section 3 – Definition of terms, acronyms, and abbreviations for this document.
- Section 4 – Description of SC-MMH

2. REFERENCES

All referenced documents are subject to revision. Users of this Standard are cautioned that newer editions might or might not be compatible.

2.1 Normative References

The following documents, in whole or in part, as referenced in this document, contain specific provisions that are to be followed strictly in order to implement a provision of this Standard.

[1] IEEE/ASTM: "Use of the International Systems of Units (SI): The Modern Metric System," Doc. SI 10-2002, Institute of Electrical and Electronics Engineers, New York, N.Y.

[2] ATSC: "ATSC Mobile/Handheld Digital Television Standard, Part 1 – ATSC Mobile Digital Television System," Doc. A/153 Part 1:2011, Advanced Television Systems Committee, Washington, D.C., 1 June 2011.

[3] ATSC: "ATSC Digital Television Standard, Part 4 – MPEG-2 Video System Characteristics," Doc. A/53, Part 4:2009, Advanced Television Systems Committee, Washington, D.C., 7 August 2009.

[4] ATSC: "ATSC Mobile/Handheld Digital Television Standard, Part 7 – AVC and SVC Video System Characteristics," Doc. A/153 Part 7:2012, Advanced Television Systems Committee, Washington, D.C., 4 July 2012.

[5] ATSC: "ATSC Digital Television Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics," Doc. A/53, Part 3:2013, Advanced Television Systems Committee, Washington, D.C., 7 August 2013.

[6] ATSC: "ATSC Mobile/Handheld Digital Television Standard, Part 3 – Service Multiplex and Transport Subsystem Characteristics," Doc. A/153 Part 3:2013, Advanced Television Systems Committee, Washington, D.C., 29 October 2013.

[7] CEA: "Digital Television Closed Captioning: 3D Extensions," Doc. CEA-708.1, Consumer Electronics Association, Arlington, VA, June 2012.

[8] ITU: "Information technology – Generic coding of moving pictures and associated audio information: Systems," ITU-T Recommendation H.222.0 (2012) | ISO/IEC 13818-1:2012, International Telecommunications Union, Geneva.

[9] ATSC: "Parameterized Services Standard," Doc. A/71:2012, Advanced Television Systems Committee, Washington, D.C., 3 December 2012.

[10] ATSC: "ATSC 3D Digital Television Standard, Part 1 – System Definition," Doc. A/104, Part 1:2014, Advanced Television Systems Committee, Washington, D.C., 4 August 2014.

2.2 Informative References

The following documents contain information that may be helpful in applying this Standard.

[11] K. H. Jung, M. S. Bang, S. H. Kim, H. G. Choo, and D. W. Kang: "Quality Enhancement for Hybrid 3DTV with Mixed Resolution Using Conditional Replenishment Algorithm," ETRI Journal, October. 2014.

3. DEFINITION OF TERMS

With respect to definition of terms, abbreviations, and units, the practice of the Institute of Electrical and Electronics Engineers (IEEE) as outlined in the Institute's published standards [1] shall be used. Where an abbreviation is not covered by IEEE practice or industry practice differs from IEEE practice, the abbreviation in question will be described in Section 3.3 of this document.

3.1 Compliance Notation

This section defines compliance terms for use by this document:

shall – This word indicates specific provisions that are to be followed strictly (no deviation is permitted).

shall not – This phrase indicates specific provisions that are absolutely prohibited.

should – This word indicates that a certain course of action is preferred but not necessarily required.

should not – This phrase means a certain possibility or course of action is undesirable but not prohibited.

3.2 Treatment of Syntactic Elements

This document contains symbolic references to syntactic elements used in the audio, video, and transport coding subsystems. These references are typographically distinguished by the use of a different font (e.g., restricted), may contain the underscore character (e.g., sequence_end_code) and may consist of character strings that are not English words (e.g., dynrng).

3.2.1 Reserved Elements

One or more reserved bits, symbols, fields, or ranges of values (i.e., elements) may be present in this document. These are used primarily to enable adding new values to a syntactical structure without altering its syntax or causing a problem with backwards compatibility, but they also can be used for other reasons.

The ATSC default value for reserved bits is '1.' There is no default value for other reserved elements. Use of reserved elements except as defined in ATSC Standards or by an industry standards setting body is not permitted. See individual element semantics for mandatory settings and any additional use constraints. As currently-reserved elements may be assigned values and meanings in future versions of this Standard, receiving devices built to this version are expected to ignore all values appearing in currently-reserved elements to avoid possible future failure to function as intended.

3.3 Acronyms and Abbreviation

The following acronyms and abbreviations are used within this document.

ATSC – Advanced Television Systems Committee

MMH – Main and Mobile Hybrid Delivery

SCRT – Service Compatible 3D-TV using Real-time Delivery

SC-MMH – Service Compatible 3D-TV using Main and Mobile Hybrid Delivery

The "%" modulus operator shall be as defined in ISO/IEC 13818-1 [8] Section 2.2.1.

3.4 Terms

The following terms are used within this document.

left view – Video provided for the left eye.

reserved – An element that is set aside for use by a future Standard.

right view – Video provided for the right eye.

Service compatible – 3D-TV broadcasting service composed of two or more compressed video images, where at least one of them is the legacy 2D-TV image having the same resolution as the production resolution.

Stereoscopic 3D video – Video composed of a left view and a right view.

VEI – Video Enhancement Information

4. SERVICE COMPATIBLE 3D-TV USING MAIN AND MOBILE HYBRID DELIVERY

4.1 Overall Description of SC-MMH

SC-MMH provides stereoscopic 3D services and fully guarantees backward-compatibility with the existing fixed/mobile 2D services based on ATSC A/153 standard [2]. As illustrated in Figure 4.1, in SC-MMH, one video is transmitted via ATSC main service and the other video is transmitted via ATSC Mobile DTV service within the same RF channel. Signaling information for 3D-TV service is transmitted via ATSC main service. In addition, Video Enhancement Information (VEI) is optionally transmitted via ATSC main service for enhancement of stereoscopic 3D image quality.

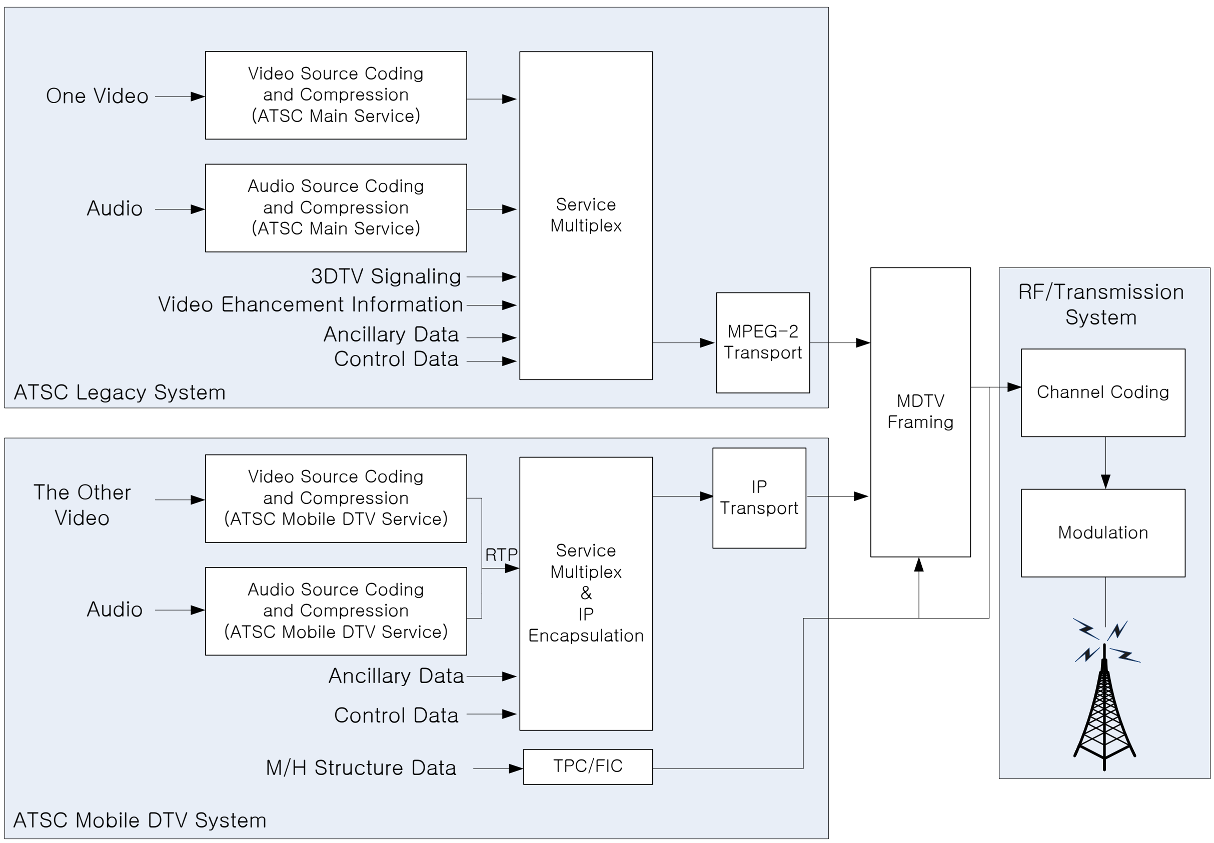

Figure 4.1 Overview of SC-MMH system.

4.2 Encoding and Decoding

Video coding of the view on ATSC main service shall comply with A/53 Part 4 [3] and video coding of the view on ATSC Mobile DTV service shall comply with A/153 Part 7 [4].

4.3 Video Format

Video formats of the view on ATSC main service shall comply with A/53 Part 4 [3] and video formats of the view on ATSC Mobile DTV service shall comply with A/153 Part 7 [4] with the following additional constraints:

- Both views shall have the same aspect ratio of the active area.

- The vertical resolution of the ATSC Mobile DTV pictures shall be greater than or equal to 0.4 × the vertical resolution of the ATSC main DTV pictures, when stereoscopic_service_type of MDTV_hybrid_stereoscopic_service_descriptor() is set to '001'.
- VEI shall be transmitted and stereoscopic_service_type of MDTV_hybrid_stereoscopic_service_descriptor() shall be set to '010' when the vertical resolution of the ATSC Mobile DTV pictures are less than 0.4 x the vertical resolution of the ATSC main DTV pictures.
- MPEG-2 video shall not use bar data.
- The frame rate of the ATSC Mobile DTV service shall be that of the ATSC main DTV service divided by an integer, including one.

When aspect ratios of L/R pictures are not exactly identical, the input pictures on the mobile broadcast side shall be letterboxed (or pillarboxed) before AVC compression to make the both views have the same aspect ratio of the active area, and the existence of the bars embedded in the transmitted pictures shall be indicated by Active Format Description (AFD) and optionally Bar Data information as specified in A/153 Part 7[4]. [1]

When the frame rate of the ATSC Mobile DTV service is that of the ATSC main DTV service divided by an integer, other than one, the frames of ATSC main service to be matched to the counterparts are selected by the timestamps in the same synchronization manner as described in Section 4.7.1.1.

In order to facilitate fast channel change or quick random access, it is recommended that the video structure of the both views be aligned together. In other words, when the video on the ATSC main service is an I picture of MPEG-2 video, it is recommended that the corresponding video on the ATSC MDTV service be an AVC RAP (Random Access Point) picture.

4.4 Service Multiplex and Transport

Service multiplex and transport of the ATSC main service and the MDTV service shall comply with A/53 Part 3 [5] and A/153 Part 3 [6], respectively.

When stereoscopic_srevice_type of MDTV_hybrid_stereoscopic_service_descriptor() is set to '010', VEI stream shall be encapsulated by MPEG-2 PES packets and transmitted in the form of TS stream over ATSC main service. stream_type value of TS stream shall be 0x92, and stream_id of PES shall be 0xBD (private_stream_1).

For streams of stream_type 0x92, PES syntax and semantics shall conform to the requirements of ATSC A/53 Part 3[6]. Each PES packet shall begin with a VEI frame, which shall be aligned with the PES packet header. The first byte of a PES packet payload shall be the first byte of a VEI frame. Each PES header shall contain a PTS. Additionally, it shall contain a DTS as appropriate. The PES packet shall not contain more than one VEI frame, and shall be void of VEI frame data only when transmitted in conjunction with the discontinuity_indicator to signal that the continuity_counter may be discontinuous.

Within the PES packet header, the following restrictions apply:

- The PES_packet_length shall be coded as 0x0000
- ata_alignment_indicator shall be coded as '1'

[1] The examples of aspect ratio adjustment processing and the corresponding behaviors of the receivers are described in Table A.1.

Note that although PTS values of VEI frames are assigned in the same manner as the ATSC main service, the VEI frame data is used for the frame on the ATSC Mobile DTV service, of which the counterpart frame on the ATSC main service has the same PTS value as the VEI frame. Representation of VEI frame data and its decoding process are described in Annex B.

4.5 Audio

The receiver can use either the AC-3 audio stream of ATSC main service or the HE AAC audio stream of ATSC Mobile DTV service. Note that surround sound can only be offered in the AC-3 audio stream.

4.6 Closed Captioning for SC-MMH

Closed captioning data is transported in the video of ATSC main service in compliance with ATSC A/53 Part 4 [3]. Closed captioning commands to support z-axis placement of caption windows (e.g., disparity data) shall be formatted in accordance with CEA-708.1 [7] and carried in the cc_data() specified in Section 6.2.3.1 of A/53 Part 4 [3].

4.7 Signaling for SC-MMH

4.7.1 PSI

The following descriptor shall be used for signaling of information on SC-MMH. Table 4.1 indicates the core descriptor, its tag, and its allowed locations in the PSI/PSIP tables. When used, the descriptor shall be in each indicated location (shown with an “M”). The descriptor also may be present in a second location (shown with an “O”).

Table 4.1 List and Location of SC-MMH Descriptors

Descriptor Name	Descriptor Tag	Terrestrial Broadcast					
		PMT	MGT	TVCT	EIT	DCCT	DCCSCT
MDTV hybrid stereoscopic service descriptor	0xEC	M			O		

4.7.1.1 MDTV Hybrid Stereoscopic Service Descriptor

MDTV_hybrid_stereoscopic_service_descriptor() carries information on SC-MMH. This descriptor includes stereoscopic service type information, a service identifier of ATSC Mobile DTV service, synchronization information, and a flag that indicates which stream is displayed for the left eye. When used, MDTV_hybrid_stereoscopic_service_descriptor() shall be present in the loop following program_info_length field in PMT [8]. MDTV_hybrid_stereoscopic_service_descriptor() may be used in the descriptor loop of the 3D event in the EIT in order to indicate the future event is in 3D.

Table 4.2 MDTV Hybrid Stereoscopic Service Descriptor

Syntax	No. of Bits	Format
MDTV_hybrid_stereoscopic_service_descriptor() {		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
reserved	5	'11111'
stereoscopic_service_type	3	bslbf
If (stereoscopic_service_type == '001' | '010') {		
MDTV_service_id	16	uimsbf
reserved	6	'111111'
leftview_flag	1	bslbf
offset_sign_bit	1	bslbf
timestamp_offset	32	uimsbf
}		
}		

descriptor_tag – This 8-bit field shall be 0xEC.

descriptor_length – This 8-bit field specifies the length (in bytes) immediately following this field up to the end of this descriptor.

stereoscopic_service_type – This 3-bit field specifies the type of the current service. This field shall be set to "000" when 2D DTV service is provided. This field shall be set to "001" when SC-MMH service is provided. This field shall be set to "010" when SC-MMH service is provided with VEI.

Table 4.3 Stereoscopic Service Types

Values	Description
'000'	2D DTV
'001'	SC-MMH
'010'	SC-MMH with VEI
'011' ~ '111'	Reserved

MDTV_service_id – This 16-bit field specifies ID of Mobile DTV Service that carries one view images.

leftview_flag – This 1-bit field specifies which stream is to be displayed for the left eye. A value of '1' indicates that the view on ATSC main service is the left view and the view on ATSC mobile DTV service is the right view. A value of '0' indicates that the view on ATSC main service is the right view and the view on ATSC mobile DTV service is the left view.

offset_sign_bit – This 1-bit field is the arithmetic sign for the timestamp_offset. A value of '0' indicates that PTS % 2^{32} is greater than RTP timestamp in calculation of timestamp_offset. A value of '1' indicates that PTS % 2^{32} is less than RTP timestamp in calculation of timestamp_offset. When timestamp_offset is '0', this field may be ignored.

timestamp_offset – This 32-bit field specifies the absolute value of the difference between PTS % 2^{32} and RTP timestamp. In offset calculation, MSB of PTS is ignored by applying modular arithmetic. Two timestamps (ATSC main service: PTS, ATSC Mobile DTV service: RTP timestamp) corresponding to the access units to be presented at the same time in 3D services are used to obtain the offset.

4.7.2 PSIP

4.7.2.1. Virtual Channel Signaling

A virtual channel that carries an SC-MMH service shall be identified by service_type set to 0x09 (Extended Parameterized Service) in the TVCT. In addition, the following descriptor shall be present in the descriptor loop following the descriptors_length field of the terrestrial_virtual_channel_table_section() or cable_virtual_channel_table_section():

- Parameterized Service Descriptor (PSD) as specified in A/71[9] and A/104 Part 1[10].

This placement is shown as an example in Table 4.4.

Table 4.4 Example TVCT Composition

```
TVCT
…
for (i<num_channels_in_section) {
    …
    major_channel_number = 0x003
    minor_channel_number = 0x002
    …
    program_number = 0x0002
    …
    service_type = 0x09 (extended parameterized service)
    …
    parameterized_service_descriptor()
    …
}
```

The parameterized_service_descriptor() with application_tag = 0x01 shall provide information about the type of 3D service carried. This information can facilitate the behaviours of the 3D-TV receivers in displaying Stereoscopic 3D video.

4.7.3 Signaling at 2D/3D Boundaries

4.7.3.1 PMT Signaling

In SC-MMH, PMT is basically used to signal service type and video stream signaling is optionally used in addition. Unless video stream signaling of 4.7.2.2 is used, PMT should be signaled timely as follows to prevent the receivers from combining two different images as 3D.

In the case of a 2D-to-3D switch, the associated PMT signaling shall not precede the actual start of the 3D service. On the other hand, in the case of a 3D-to-2D switch, the associated PMT signaling shall not take place after the actual end of the 3D service. Here, 3D service may include zero-disparity content that consists of two identical images as described in Section 4.7.3.3.

In addition to PMT signaling, video stream signaling may be used for frame-accurate signaling as described in Section 4.7.3.2.

4.7.3.2 Video Stream Signaling

Frame_packing_arrangement_data in extensions_and_user_data(2) of MPEG-2 video may be used in addition to PMT signaling in order to provide frame-accurate signaling of the mode transition from 3D to 2D or 2D to 3D. If frame_packing_arrangement_data is used for frame-accurate signaling, frame_packing_arrangement_data with arrangement_type value ‘0001000’ (2D video) shall be included in the MPEG-2 video stream during the 2D portion of the service, and frame_packing_arrangement_data shall not be included in the MPEG-2 video stream during the 3D portion of the service. When PMT signaling indicates that the current program is 2D, the existence or absence of frame_packing_arrangement_data shall be ignored.

4.7.3.3 3D with Zero-disparity

A subset of 3D content in which the left and right images are identical may be used for the 2D portion such that the video, transport, and signaling parameters match the adjacent 3D contents.

Annex A: Aspect Ratio Adjustment of the ATSC Mobile DTV Videos

The actual aspect ratios of the videos over the ATSC main service and the ATSC Mobile DTV service may be slightly different as specified in A/153 Part 7[4]. Accordingly, the input video on the mobile broadcast side may need adjustment processing before compression for L/R matching. To resolve this issue, input pictures, which are not exactly 16:9 aspect ratio, shall be letterboxed (or pillarboxed) before AVC compression. In this case, the existence of the bars embedded in the transmitted pictures shall be indicated by Active Format Description (AFD) and optionally Bar Data information as specified in A/153 Part 7[4]. At the receiver side, to create 3D images by combining the L/R images which have slightly different aspect ratios, the images may need adjustment processing before presentation. In this annex, some examples of aspect ratio adjustment processing and the corresponding behaviors of the receivers are described in Table A.1.

> Note: It is assumed that the display size is the same as the size of the video over ATSC main service.

Table A.1 Examples of Aspect Ratio Adjustment before 3D Presentation

Video size (MDTV)	Mobile DTV Video Processing Before Compression	Video Size (main)	Aspect Ratio Adjustment Processing Before 3D Presentation
416 x 240	Scale to 416 x 234 Add 2 and 4 line bars to top and bottom, respectively	1920 x 1080	**Images over Mobile DTV:** Extract the 416 x 234 active portion (from line 3 to line 236) Scale to 1920 x 1080
		1280 x 720	**Images over Mobile DTV:** Extract the 416 x 234 active portion (from line 3 to line 236) Scale to 1280 x 720
		704 x 480	**Images over Mobile DTV:** Extract the 416 x 234 active portion (from line 3 to line 236) Scale to 704 x 480
640 x 368	Scale to 640 x 360 Add 4 line bars to each top and bottom	1920 x 1080	**Images over Mobile DTV:** Extract the centre 640 x 360 portion Scale to 1920 x 1080
		1280 x 720	**Images over Mobile DTV:** Extract the centre 640 x 360 portion Scale to 1280 x 720
		704 x 480	**Images over Mobile DTV:** Extract the centre 640 x 360 portion Scale to 704 x 480
832 x 480	Scale to 832 x 468 Add 6 line bars to each top and bottom	1920 x 1080	**Images over Mobile DTV:** Extract the centre 832 x 468 portion Scale to 1920 x 1080
		1280 x 720	**Images over Mobile DTV:** Extract the centre 832 x 468 portion Scale to 1280 x 720
		704 x 480	**Images over Mobile DTV:** Extract the centre 832 x 468 portion Scale to 704 x 480

Annex B: Video Enhancement Information

B.1 SYSTEM OVERVIEW OF SC-MMH WITH VIDEO ENHANCEMENT INFORMATION

In SC-MMH System, supplemental information called Video Enhancement Information (VEI) may be utilized to enhance the 3D visual quality when the spatial resolutions of the first view and second view are quite different. VEI can help the viewers who may be interfered due to the quality gap when the difference of resolutions between two views is too large. VEI may be used to enhance the enlarged image of the 2nd view (ATSC Mobile DTV) with lower resolution. When VEI is used, VEI may be generated at 3D encoding step and then transmitted to the 3D-TV receivers. A SC-MMH receiver supporting VEI decodes VEI and thereby generates the 3D images with better quality. When VEI is not used, the intact images from ATSC mobile service is used for 3DTV service without this enhancement process.

VEI is used to enhance the image quality of the enlarged 2nd view video as shown in Figure B.1.

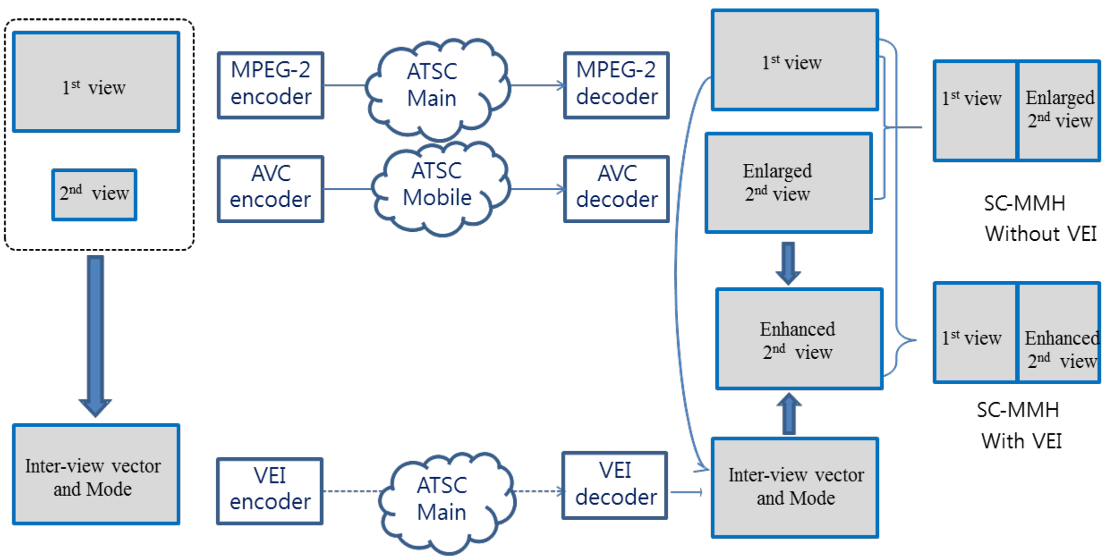

Figure B.1 Image Quality Enhancement of SC-MMH System with VEI.

VEI shall consist of two kind of information; one is the mode and the other is the inter-view vector. The mode represents which substitute is selected from the inter-view vector compensated 1st view and the enlarged 2nd view.

The inter-view vector (IVV) represents the correlations between two stereoscopic images. The IVV at (x, y) pixel position shall consist of horizontal and vertical components as below:

$$IVV(x, y) = (h\text{-}ivv, v\text{-}ivv)$$

Since there is a spatial correlation between IVVs and modes, the pixels with the same IVVs and modes can be merged into a single PU (Processing Unit). And the optimized size of PUs depends on the spatial characteristics of images. Hence, a frame shall be partitioned into a quad-tree structure with variable-sized PUs.

There are also temporal correlations between IVVs of consecutive frames. The VEI frames that use the temporal correlation shall be denoted as 'Inter-type' and the other frames shall be denoted as 'Intra-type'.

B.2 REFERENCES

There are no references in this Annex.

B.3 DEFINITION OF TERMS

For the purpose of this Annex, where an abbreviation or a term is not covered in Section 3 of the main document, the abbreviation and the term in question will be described in this section.

B.3.1 Acronyms and Abbreviation

For the purpose of this Annex, the following acronyms and abbreviations are applied in addition to the definitions in Section 3.

IVV – Inter-View Vector

PU – Processing Unit

LPU – Largest Processing Unit

SPU – Smallest Processing Unit

B.3.2 Terms

For the purpose of this Annex, the following definitions are applied in addition to the definitions in Section 3 of the main document. These definitions are either not present in Section 3 or replace definitions in Section 3.

1st view Frame – A frame received from the ATSC Main service.

2nd view Frame – A frame received from the ATSC Mobile DTV service.

Enlarged 2nd view – An enlarged frame of the 2nd view for 3D service.

Enhanced 2nd view – An enhanced frame of the enlarged 2nd view by using VEI.

PU – A basic processing unit of VEI, a square block of luma samples of an enhanced 2nd view frame. PUs shall be partitioned with Quad-tree Structure.

level – A value indicates the size of PU (the size of the PU = $2^{level} \times 2^{level}$).

quad-tree – A tree in which a parent node can be split into four child nodes, each of which may become parent node for another split into four child nodes.

syntax element – An element of data represented in the bit stream.

tree – A tree is a finite set of nodes with a unique root node.

B.4 BIT STREAM SYNTAX AND SEMANTICS

B.4.1 VEI Bit Stream

The VEI bit stream is made up of a sequence of VEI frames.

B.4.2 VEI Frame

Each VEI frame shall consist of a VEI_frame_header and VEI_frame_payload. See Table B.1.

Table B.1 VEI frame Syntax

Syntax	No. of Bits	Format
VEI frame(){		
VEI_frame_header ()	88	
VEI_frame_payload (VEI_type, W, H, LPU_level, SPU_level)	var	
}		

B.4.2.1 VEI_frame_header Syntax

The VEI_frame_header syntax is given in Table B.2.

Table B.2 VEI_frame_header Syntax

Syntax	No. of Bits	Format
VEI_frame_header ( ) {		
VEI_frame_start_code	32	bslbf
Presentatin_time_stamp	32	bslbf
VEI_payload_length	16	uimsbf
VEI_frame_type	1	bslbf
LPU_level	3	bslbf
SPU_level	3	bslbf
reserved	1	'1'
}		

VEI_frame_start_code – The field shall be 0x000001B2 (user_data_start_code). This field identifies the beginning of a VEI frame.

Presentation_time_stamp – This field shall specify the presentation time stamp of the 1st view image.

VEI_payload_length – This field shall specify the byte length of VEI_frame_payload.

VEI_frame_type – This field shall indicate whether the IVVs of the previous frame are used or not. When set to "1", the IVVs of the previous frame shall be utilized for decoding the IVV of the current frame. When set to "0", the previous information shall be discarded. When VEI_frame_type is equal to 0 (Intra-type), each PU in a VEI frame shall be one of the three modes: 'Main-view mode' (M mode), which uses the inter-view compensation PU with IVV from the frame of ATSC main service; 'MH-view mode' (H mode), which uses the collocated PU of enlarged frame of ATSC Mobile DTV service; and 'Not Available mode' (NA mode), which indicates this PU is not defined at the current level, but will be defined at the following levels.

When VEI_frame_type is equal to 1 (Inter-type), 'Skip mode' (S mode), which uses the IVV of collocated PU in the previous VEI frame, shall be used additionally. See Table B.3.

Table B.3 VEI_frame_type

Values	Description
'0'	Intra-type
'1'	Inter-type

LPU_level[2] – This field specifies the largest size of the PU (LPU).

Table B.4 LPU_level

Values	Description	LPU_level
‘000’	Sample resolution of LPU is a 8x8	3
‘001’	Sample resolution of LPU is a 16x16	4
‘010’	Sample resolution of LPU is a 32x32	5
‘011’	Sample resolution of LPU is a 64x64	6
‘100’	Sample resolution of LPU is a 128x128	7
‘101’ ~ ‘111’	Reserved	

SPU_level – This field specifies the smallest size of the PU (SPU).

Table B.5 SPU_level

Values	Description	SPU_level
‘000’	Sample resolution of SPU is a 1x1	0
‘001’	Sample resolution of SPU is a 2x2	1
‘010’	Sample resolution of SPU is a 4x4	2
‘011’	Sample resolution of SPU is a 8x8	3
‘100’	Sample resolution of SPU is a 16x16	4
‘101’ ~ ‘111’	Reserved	

B.4.2.2 VEI_frame_payload Syntax

The VEI_frame_payload syntax is given in Table B.6.

[2] PU with Quad-tree Structure: The hierarchical depth of quad-tree shall be specified by the LPU_level and SPU_level values, where LPU_level and SPU_level specifies the sizes of the largest processing unit (LPU) and smallest processing unit (SPU), respectively. The size of PU shall increase four times per level increase.

Table B.6 VEI_frame_payload Syntax

Syntax	No. of Bits	Format
VEI_frame_payload (VEI_type, W, H, LPU_level, SPU_level) {		
for(level = LPU_level ; level >= SPU_level ; level--) {		
PUsize = 2^{level}		
for(i= 0; i < H; i += PUsize)		
for(j= 0; j < W; j += PUsize){		
if (decoded_flag(i,j) == 1) continue		
If (level == 0) **pel_diff**	var	vlclbf
else {		
record_mode	var	vlclbf
if ((VEI_frame_type == ‘0’&&		
((record_mode == ‘10’ && level != SPU_level)||		
(record_mode == ‘1’ && level == SPU_level)))		
|| (VEI_frame_type == ‘1’&& record_mode == ‘11’)) {		
h-ivv_diff	var	vlclbf
v-ivv_diff	var	vlclbf
}		
}		
}		
trailing_bits( )	var	vlclbf
}		

pel_diff – When the level is equal to 0 and the size of PU is 1x1, DPCM is used instead of using record mode and accompanying IVV. This field specifies the difference value between the current pixel and the reference pixel. This field shall be used only at the level 0. This field should be coded by signed Exp-Golomb coding. The bitstrings and codeNum is show in Table B.7.

Table B.7 Signed Exp-Golomb Code Table

Codewords	Description
‘1’	0
‘011’, ‘010’	+1, -1
‘00110’, ‘00111’, ‘00100’, ‘00101’	+2,+3,-3,-2
‘0001100’ ~ ‘0001111’, ‘0001000’ ~ ‘0001011’	+4~+7, -7~-4
‘000011000’ ~ ‘000011111’, ‘000010000’ ~ ‘000010111’	+8~+15, -15~-8
‘00000110000’ ~ ‘00000111111’, ‘00000100000’ ~ ‘00000101111’	+16~+31, -31~-16
‘0000001100000’ ~ ‘0000001111111’, ‘0000001000000’ ~ ‘0000001011111’	+32~+63, -63~-32
…	…

record_mode – This field indicates the mode for the current PU. The semantics of this field for each VEI_frame_type and the current level of PU is as shown in Table B.8.

Table B.8 Record Mode of PU Except Level '0'

VEI_frame_type	Level	Values	Mode
'0' (Intra-type)	Other Levels	'0'	H Mode
		'10'	M Mode
		'11'	NA Mode
	SPU Level	'0'	H Mode
		'1'	M Mode
'1' (Inter-type)	Other Levels	'00'	S Mode
		'10'	H Mode
		'11'	M Mode
		'01'	NA Mode
	SPU Level	'0'	S Mode
		'10'	H Mode
		'11'	M Mode

h-ivv_diff – This field specifies the difference value for horizontal component of IVV for the pixels in the current PU. This field should be coded by signed Exp-Golomb coding and the values should be represented in Table B.7.

v-ivv_diff – This field specifies the difference value for vertical component of IVV for the pixels in the current PU. This field should be coded by signed Exp-Golomb coding and the values should be represented in Table B.7.

trailing_bits() – This bits may present as following case:

i) The first (most significant, left-most) bits of the final byte in VEI_frame_payload contains the remaining bits of the final byte, (if any),

ii) The next bit shall consist of a single stop_one_bit equal to 1, and

iii) When the stop_one_bit is not the last bit of a byte-aligned byte, one or more alignment_zero_bit shall be present to result in byte alignment.

Table B.9 Trailing Bits Syntax

trailing_bits() {	Descriptor
stop_one_bit /* equal to 1 */	'1'
while(!byte_aligned())	
alignment_zero_bits /* equal to 0*/	'0'
}	

Annex C: Decoding Process of Video Enhancement Information

C1. DECODING PROCESS

In VEI decoding process, each VEI frame shall be sequentially decoded. In each frame, VEI frame header shall be decoded and then its payload shall be decoded depicted as shown in Figure C.1.

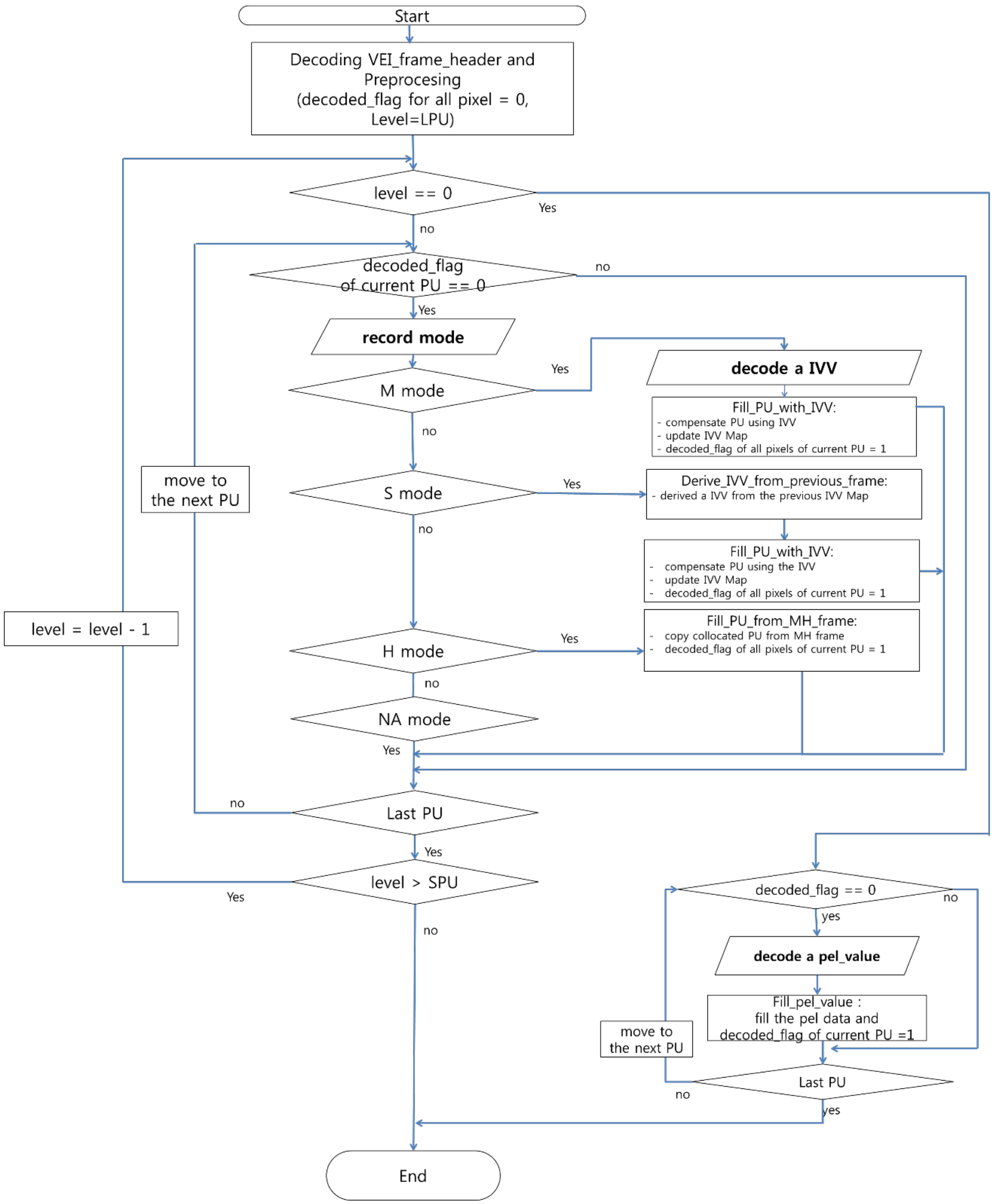

Figure C.1 Decoding process of VEI for a frame.

The inputs of decoding process of VEI shall be a bit stream of VEI, a frame of the 1st view and a frame of the 2nd view. The outputs of decoding process of VEI shall be the enhanced 2nd view and IVVs.

VEI for each frame shall consist of PUs structured by quad-tree. The hierarchical depth of quad-tree shall be specified by the LPU_level and SPU_level values. Here, each PU shall be decoded in the order from left-top to right-bottom. Once all PUs at the given level are decoded, the scanning of PUs for the next lower level shall be started. The PUs, already decoded at the higher level, shall be skipped.

C.1.1 Decoding VEI_frame_header and Pre-processing

The VEI_frame_header shall be decoded to get the parameters which are used for encoding VEI. Pre-processing shall be performed for decoding process of VEI for a frame as below.

- Enlarging a frame of the 2nd view received from the ATSC mobile DTV (Enlarged 2nd view Frame)
- Resetting 'decoded_flag' in Figure C.1 to 0 for all pixels
- Setting the level value to LPU_level

For a Level:

The size of the PU at a level shall be determined by the following:

$$\text{Size\_of\_PU} = 2^{\text{level}} \times 2^{\text{level}}$$

In each level, the PU shall be processed from left-top to right-bottom.

For Each PU:

If the level is equal to 0, pel_diff should be decoded according to Table B.7.

- The pel_diff should be a value after decoding the signed exp-Golomb code.
- Then, the following Fill_pel_value process shall be invoked:
  - i) The current pixel value shall be derived as follows:

$$\text{Pel\_value}(x, y) = \text{pel\_diff} + \text{reference pel value}$$

  - ii) The reference pel value shall be derived as follows:
    If the neighboring left pixel has been encoded at the level 0, the reference pel value shall be the left pixel.
    Otherwise, if the neighboring upper pixel has been encoded at the level 0, the reference pel value shall be the upper pixel.
    Otherwise, the reference pel value shall be the fixed value of 128.
  - iii) decoded_flag shall be set to "1" for the current PU.

Otherwise, record_mode shall be firstly decoded according to Table B.8.

- If record_mode is set to M mode:
  - i) h_ivv_dff and v_ivv_diff should be decoded according to Table B.8.
  - ii) The horizontal and vertical components of an IVV shall be derived as follows:

$$\text{h-ivv} = \text{h\_ivv\_diff} + \text{h\_ivv\_predictor}$$

v-ivv = v_ivv_diff + v_ivv_predictor

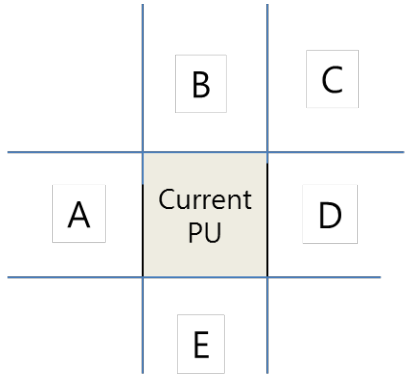

Figure C.2 The prediction directions of IVV.

iii) The IVV predictor is derived as follows:
If the left neighboring PU (A in Figure C.2) is decoded and having an IVV, the IVV predictor shall be the IVV of A.
Otherwise, if the upper neighboring PU (B) is decoded and having an IVV, the IVV predictor shall be the IVV of B.
Otherwise, if the upper-right neighboring PU (C) is decoded and having an IVV, the IVV predictor shall be the IVV of C.
Otherwise, if the right neighboring PU (D) is decoded and having an IVV, the IVV predictor shall be the IVV of D.
Otherwise, if the below neighboring PU (E) is decoded and having an IVV, the IVV predictor shall be the IVV of E.

iv) The following Fill_PU_with_IVV process shall be invoked:
All pixels of the current PU shall be the value compensated from the frame of the 1st view and the IVV, which shall be derived as follows:

$$P(x, y) = P_{1^{st}\ view}(x+h\text{-}ivv,\ y+v\text{-}ivv)$$

The IVV has half pixel accuracy and the bilinear interpolation scheme should be used to calculate sub-pixel positions. Update all pixels of current PU in the Interview vector Map as follows:

$$IVV(x, y) = (h\text{-}ivv,\ v\text{-}ivv)$$

decoded_flag shall be set to “1” for the current PU.

- Otherwise, if record_mode is set to S mode:
  i) The horizontal and vertical components of the IVV shall be replicated from those of collocated PU in the IVV map of the previous frame.
  ii) The Fill_PU_with_IVV process shall be invoked.
  iii) decoded_flag shall be set to “1” for the current PU.
- Otherwise, if record_mode is set to H mode:
  i) The following Fill_PU_from_MH_frame process shall be invoked:

All pixels of current PU shall be filled with the pixels of collocated PU in the enlarged 2nd view.

ii) decoded_flag' shall be set to "1" for the current PU.

- Otherwise, if record_mode is set to NA mode:

  Moving to the next PU.

Test of Completion:

If the current level is not equal to the SPU level, the next level shall be processed from the step two.

Otherwise, the decoding process for the current frame shall be finished by skipping trailing_bits.

End of Document