Dolby® Open Specification for Frame-Compatible 3D Systems

Issue 2

Dolby Laboratories, Inc.

Corporate Headquarters

Dolby Laboratories, Inc.
100 Potrero Avenue
San Francisco, CA 94103-4813 USA
Telephone 415-558-0200
Fax 415-863-1373
www.dolby.com

European Headquarters

Dolby Laboratories, Inc.
Wootton Bassett
Wiltshire SN4 8QJ England
Telephone 44-1793-842100
Fax 44-1793-842101

Dolby Laboratories, Inc. does not require consideration for the use of the information regarding base layer contained in this document. If Dolby Laboratories, Inc. develops an enhancement layer technology (see section 5 of this document), then a license to use such technology will be required.

TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, IN NO EVENT SHALL DOLBY LABORATORIES, INC., ITS SUBSIDIARIES, OR ITS AGENTS BE LIABLE FOR ANY SPECIAL, INCIDENTAL, CONSEQUENTIAL, DIRECT, OR INDIRECT DAMAGES WHATSOEVER (INCLUDING WITHOUT LIMITATION LOSS OF SYSTEM AVAILABILITY, DAMAGES FOR LOSS OF BUSINESS PROFITS OR REVENUES, BUSINESS INTERRUPTION, LOSS OF DATA, COST OF DEFENSE, OR OTHER PECUNIARY LOSS) ARISING OUT OF THE USE, INABILITY TO USE, OR RELIANCE ON THE INFORMATION CONTAINED IN THIS DOCUMENT HOWEVER CAUSED AND WHETHER BASED IN CONTRACT, TORT (INCLUDING NEGLIGENCE) OR ANY OTHER THEORY OF LIABILITY, EVEN IF DOLBY LABORATORIES, INC., ITS SUBSIDIARIES, OR ITS AGENTS HAVE BEEN ADVISED, EITHER ORALLY OR IN WRITING, OF THE POSSIBILITY OF SUCH DAMAGES.

Issue 2
S10/22833/23688

Table of Contents

Dolby Open Specification for Frame-Compatible 3D Systems

Annex 1 Technical Specifications for the Dolby Open Specification for Frame-Compatible 3D Systems

Dolby Open Specification for Frame-Compatible 3D Systems

Dolby Laboratories has conducted a series of experiments into frame-compatible[1] systems to characterize the advantages and disadvantages of the various methods for both decimation and packing.

This document outlines the results and proposes that a specific method be adopted to optimize the performance of frame-compatible systems for both 3D and 2D content.

Optimizing for 2D content packed as frame-compatible 3D content has value, because some content, including interstitial material such as commercials, will be available in 2D format only. Current specifications refer only to static 3D operation and do not consider dynamic 2D to 3D transitions (or vice versa).This is appropriate, as it takes existing 3D televisions one to three seconds to change between 2D and 3D display modes.

Moreover, in the same way that color (or lack thereof) was used to great creative effect in films like *The Wizard of Oz* and *Schindler's List,* a combination of 2D and 3D scenes can be used to similar effect, as witnessed in the recent trailer for *Tron: Legacy.*

As the broadcast industry has already reached consensus on the following issues, these will not be elaborated upon further in this document:

- Quincunx decimation requires a significantly higher compression bit rate than either column or line decimation systems, and is therefore not efficient for emission systems.
- Column interleave decimation and side-by-side (SbS) packing is required for interlaced formats; otherwise, vertical resolution becomes too limited for the picture to be considered HDTV.

[1] Walt Husak, Dolby Laboratories, Inc., Dolby® 3D white paper

1 Dolby 3DTV Frame-Compatible Tests

Dolby's experiments using the ITU-T H.264 or equivalent ISO/IEC MPEG-4 Advanced Video Coding (AVC[2]) video coding standard have revealed no significant difference in encoding efficiency between SbS and top-and-bottom (TaB). TaB packing is referred to referred to as over/under (O/U) in Figure 1, which shows a typical plot for live action programming.

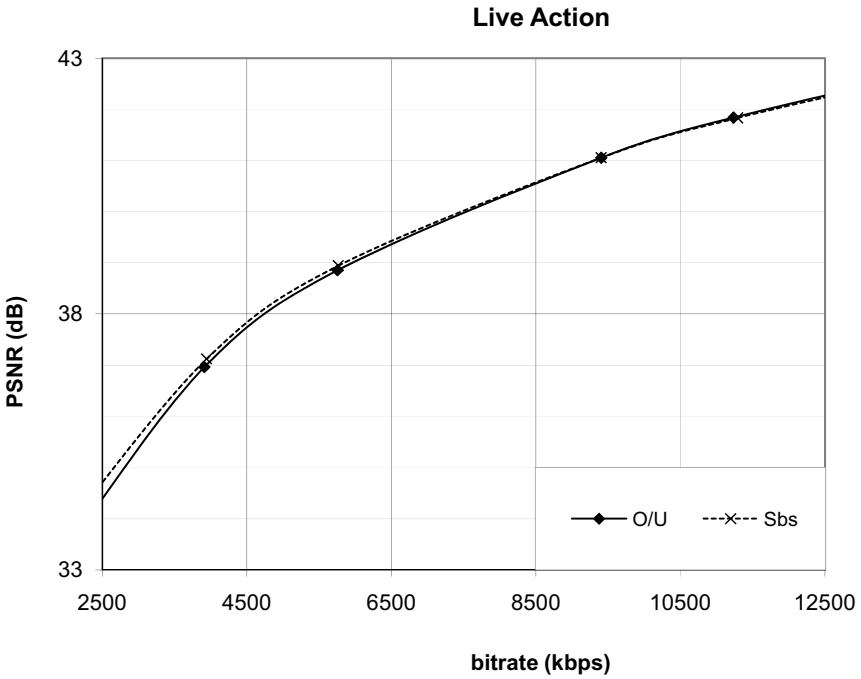

Figure 1 Relative Coding Efficiencies of Side by Side and Top and Bottom (Labeled O/U in This Diagram)

Dolby's internal tests with compressed progressive content at transmission bit rates resulted in no overall preference for SbS or TaB; the choice between the two is dictated largely by the content type (high horizontal or vertical resolution).

More recently, the Entertainment Technology Center (ETC), in collaboration with ESPN and Warner Bros. Entertainment, demonstrated the differences between full resolution 3D and half resolution in both SbS and TaB formats for progressive content during the 2010 NAB Show. Although these demonstrations were not conducted in laboratory conditions, there was general consensus that the full-resolution images were perceptibly sharper than the half resolution, even with the limited content available. There was no consensus regarding the half-resolution demos; the preference for SbS or TaB was dependent on content and individual preference.

Given that support of interlaced formats is mandatory for broadcasters, and in the interests of simplifying the decoding and formatting requirements, we suggest that all content, interlaced and progressive, should be encoded in the SbS format.

Unfortunately, the most common interface currently used to transport 3D images—3D HDMI™ (version1.4a)—only defines SbS progressive formats as optional, rather than

[2] ISO/IEC 14496-10:2005, *Information Technology—Coding of Audio-Visual Objects—Part 10: Advanced Video Coding*

mandatory, although such formats are supported in the DVB BlueBook document A151[3] and used by operators in North America. Support for progressive TaB systems has therefore been included in this issue of the specification.

2 Side-by-Side Decimation

Having made the initial choice to use an SbS frame-compatible structure, Dolby Laboratories then compared the performance of the different SbS decimation options. In Figure 2 and Figure 3, columns are counted left to right starting from zero (0, 1, 2, 3, and so on). Thus, the left-most column is considered an even column.

We investigated the two classes of decimation available for the left- and right-eye images:

- Common decimation for left and right images (even, even or odd, odd)
- Complementary decimation for left and right images (even, odd or odd, even)

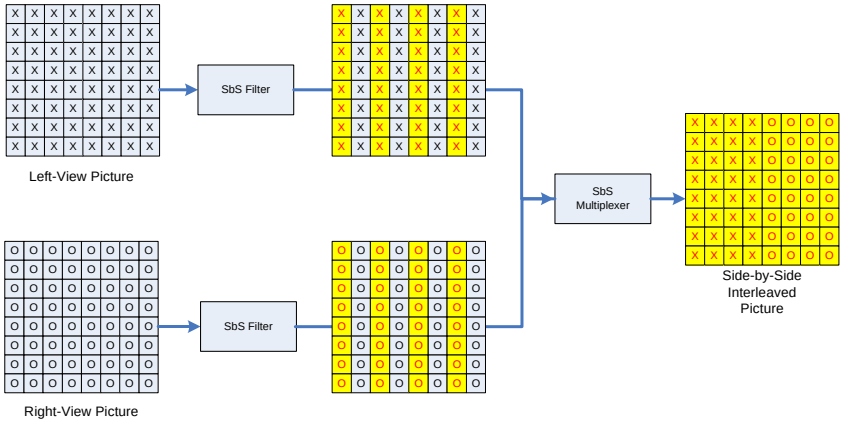

Figure 2 Common SbS Decimation Method (Even, Even)

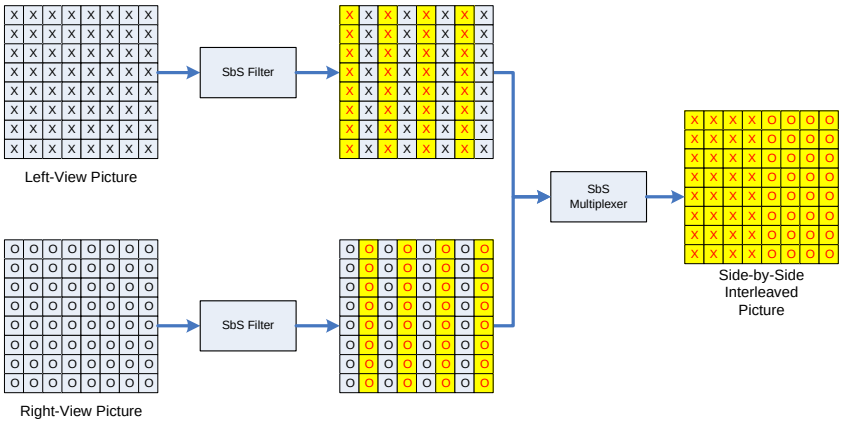

Figure 3 Complementary SbS Decimation (Even, Odd)

[3] DVB BlueBook A151, *DVB Commercial Requirements for DVB 3D-TV*, July 2010

In the compressed MPEG-4 AVC/H.264 domain, signaling to define the exact decimation method used has already been standardized in the supplemental enhancement information (SEI) messaging.[4]

The key parameters evaluated were:

- 3D half-resolution performance in SbS
- 2D images encoded as SbS
- Efficiency of enhancement layer coding needed to re-create full-resolution images

3 Top and Bottom Decimation

The same choices of common or complementary sampling are also available for systems employing TaB packing.

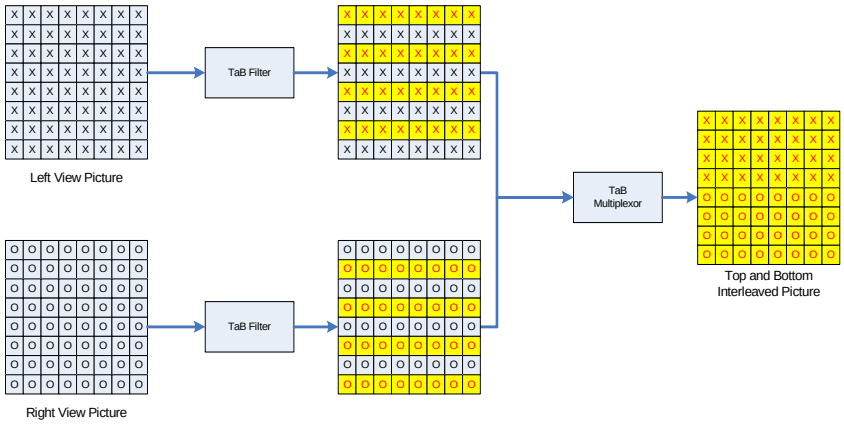

Figure 4 Complementary TaB Decimation (Even, Odd)

Similar benefits can be accrued for the TaB system by using complementary sampling techniques, but this can be used only for progressive sources.

4 3DTV Half-Resolution Performance

To ensure acceptable picture quality of the half-resolution images, some baseband filtering may be applied to reduce alias artifacts generated by decimation. Although this will inevitably reduce the overall image resolution, these filters typically operate above the Nyquist limit for optimum subjective performance (balancing resolution and aliasing).

In the case of SbS or TaB common decimation, the coincident odd (for even, even decimation systems) pixels from both left and right images must be interpolated in the display.

For complementary decimation systems, a coded pixel is always available in either the left-eye or right-eye view. At screen depth, where vergence and accommodation match, the fusion of the left- and right-eye images produces a full-resolution image (within the limitations of the baseband filtering employed).

[4] *Text of ISO/IEC 14496-10:2009/FDAM 1 Constrained Baseline-Profile, Stereo High-Profile, and Frame-Packing Arrangement SEI Message*, GJ Sullivan, et al, Doc. N10707, London, UK, July 2009

In 2D content production, it is normal to use focus to draw the attention of the viewer to a particular object. In 3D production, this is also normally the point of vergence. Placing the object where vergence and accommodation of the eye coincide helps to avoid asthenopia, or visual fatigue[5].

Complementary decimation techniques provide the highest resolution at this point. Experiments are currently underway to confirm that visual acuity is at a maximum when vergence and accommodation stimuli are coincident.

5 2D Compatible Images Coded Side by Side

In the case of 2D encoded into an SbS format (to avoid display mode switching times), near full-resolution 2D signals can be displayed when using the complementary decimation process.

True full-resolution 2D can be displayed if signaling is provided to the receiver to indicate 2D SbS content. In this case, all baseband filtering could be removed for 2D content before encoding. Full-resolution 2D content can then be displayed in 3D display modes by inhibiting the interpolation process in the 3D monitor.

Existing 3D displays will be unable to take advantage of this signaling; consequently, some minimal baseband filtering may be desirable.

If signaling is not provided, 3D display manufacturers could check whether left- and right-eye images for the entire 3D image are essentially the same (that is, 2D), and suitably modify their upconversion algorithm.

6 Efficiency of Enhancement Layer Encoding

Dolby Laboratories' internal tests on enhancement layer efficiency have clearly demonstrated that, for a base layer using the principle of complementary decimation, there is a significant overall gain in coding (enhancement) efficiency for 2D content, and for those portions of 3D content at the screen plane.

7 The Dolby Open Specification

Dolby's specification incorporates the following key elements:

1. Side-by-side decimation and packing should be adopted as the method of implementing frame-compatible 3D systems, for both progressive and interlaced systems.
2. A complementary left-eye/right-eye decimation structure should be adopted, as this provides higher resolution at the screen plane in the case of 3D images. It is also capable of providing full-resolution 2D images in the 3D SbS format without the need for receiver mode switching.
3. If only one format is used for decimation and packing, it is not necessary to provide detailed signaling for these parameters down to the display.

---

[5] "Accommodative Responses to Stereoscopic Three-Dimensional Display," Inoue, Ohzu, *Applied Optics,* Vol. 36, Issue 19, July 1997 /
"Visual Discomfort and Visual Fatigue of Stereoscopic Displays: A Review," Lambooij, IJsselsteijn, Fortuin & Heynderickx, *Journal of Imaging Science & Technology,* Vol. 53, Issue 3, May/June, 2009

The detailed technical specification is included in Annex 1, beginning on page 7.

Annex I Technical Specifications for the Dolby Open Specification for Frame-Compatible 3D Systems

1 Introduction

To provide a single 3D (stereoscopic) frame-compatible base layer solution for both progressive and interlaced formats, the left and right views are packed together using an SbS arrangement, as shown in Figure 1.

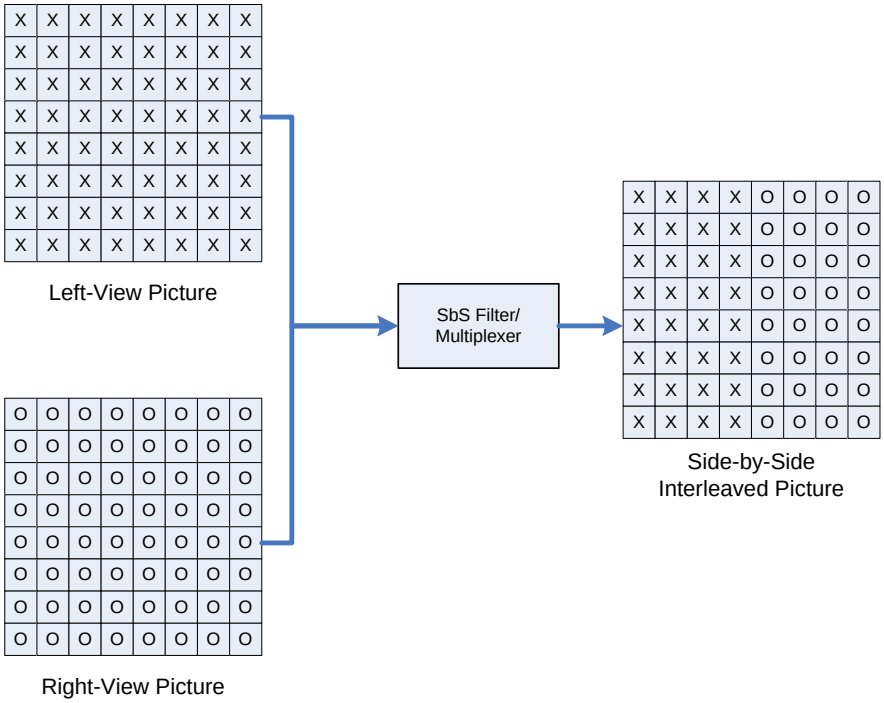

Figure 1 Side-by-Side Interleaved Arrangement for the Delivery of Stereoscopic Material

If only progressive material is being considered, the alternative TaB packing, shown in Figure 2 can also be used.

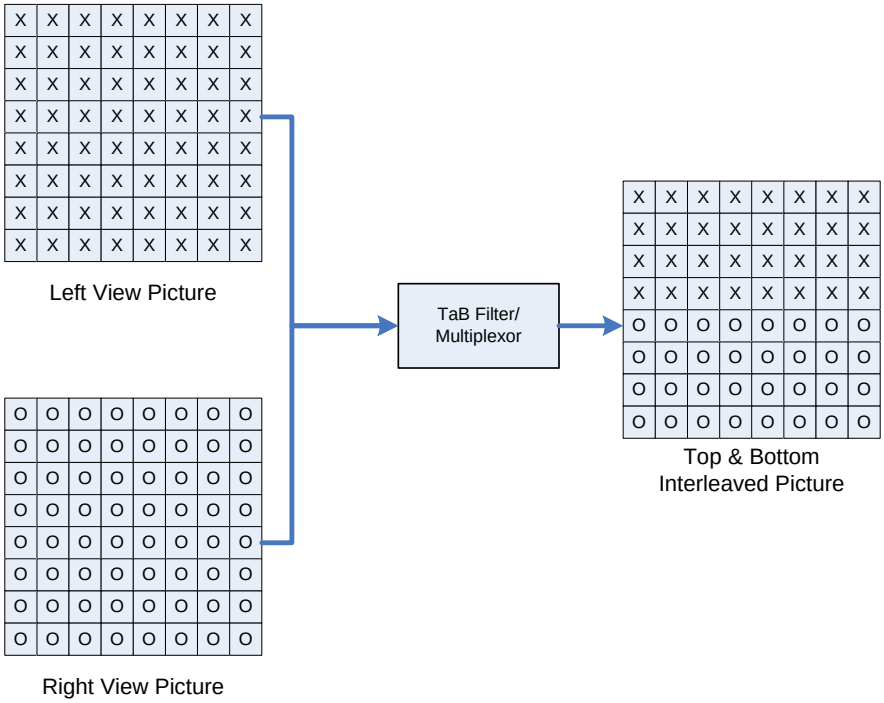

Figure 2 Top and Bottom Interleaved Arrangement for the Delivery of Progressive Stereoscopic Material

2 Dolby Frame Compatible Format Specifications

The specifics of the Dolby compatible SbS 3D video format are shown in the following table.

Packing Arrangement	Side by side	Top and bottom
Sampling	Horizontal (Columns)	Vertical (Rows)
Sub Frame Arrangement	Left first	Left top
Complementary Sampling	Left subframe position (x,y): (0,0) Right subframe position (x,y): (1,0)	Left subframe position (x,y): (0,0) Right subframe position (x,y): (0,1)

The last parameter specifies that a different sampling process is performed between the left and right views. (Essentially, the right view is sampled with a delay of 1 pixel either horizontally or vertically.) This provides considerable performance benefits, especially when delivering 2D content using the same packing arrangement. In this scenario, an intelligent decoding system could reconstruct the full 2D view without any loss in resolution because, excluding information lost due to compression, no information was removed due to the packing process[1]. This process is shown in Figure 3 for SbS interleaving.

[1] It is assumed that no filtering is also applied onto the image.

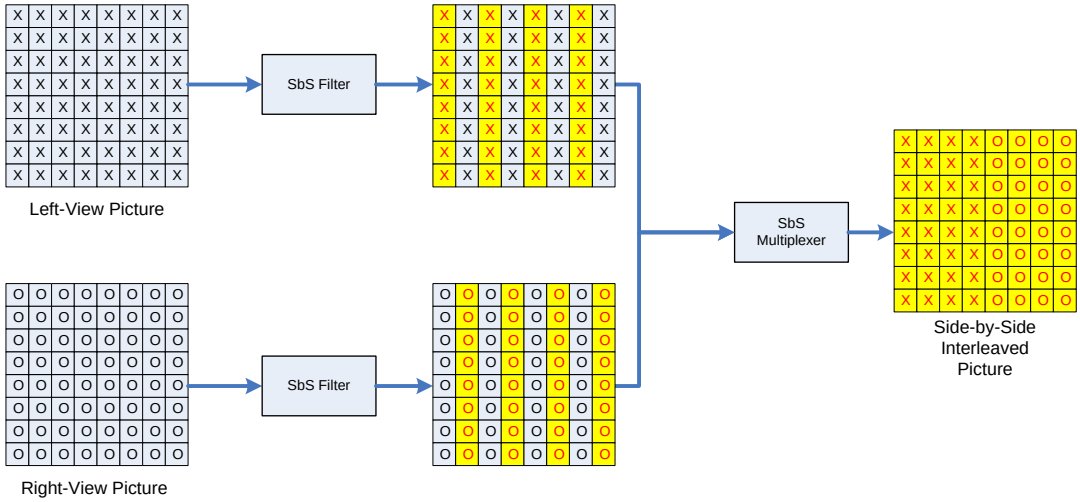

Figure 3 Side-by-Side Interleaving with Delayed Right View Sampling

An example process of deinterleaving and reconstruction the full-resolution images from an SbS 3D video format is presented in Figure 4.

This document does not specify the upconversion process that should be used at the decoder/display. One system may, for example, elect to use a simple horizontal filter and interpolation method, while another may consider the use of a more complex, edge-adaptive interpolation system for upconverting the two frame views. Information from both the subframes could also be utilized during interpolation, which may be particularly useful when upconverting 2D images packed in this arrangement.

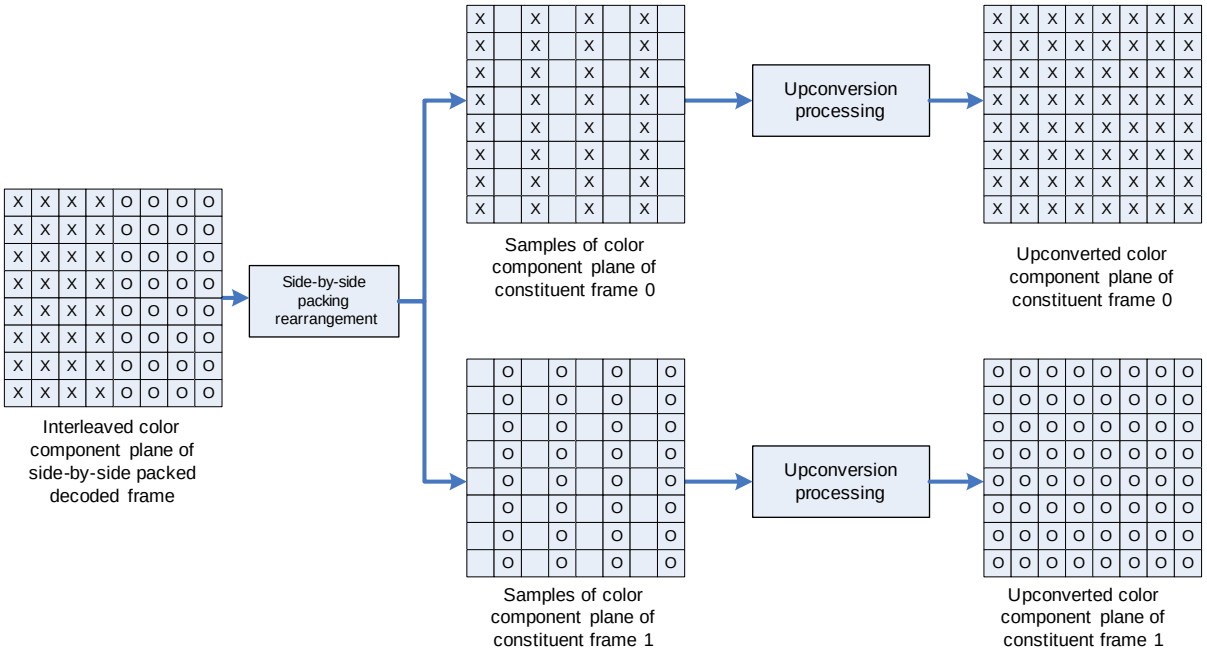

Figure 4 Example Side-by-Side Deinterleaving Process

Similar techniques can also be employed for TaB systems for progressive images.

3 Dolby Frame-Compatible Format Recommendations

Filtering before horizontal or vertical sampling is not part of the specification, as this will be determined by the target application, device, and bit rate.

However, two specific categories of filters should be made available:

- Filtering for 3D frame-compatible content
- Filtering for 2D frame-compatible content

3.1 Filtering for 3D Content Multiplexed into SbS

A variety of filters may be used to downsample the images, including linear filters, adaptive filters, frequency domain filters and motion-compensated temporal filters.

The reference linear filter design specified in the following table provides a good trade-off in terms of sharpness, and compression ratio.

Video Component		
Frame-Compatible Content	**Luma**	**Chroma**
SbS	[–5 19 29 –68 –47 305 558 305 –47 -68 29 19 –5]	[–20 –13 84 154 84 –13 –20]
TaB	[48 34 –93 –36 301 516 301 –36 –93 34 48]	[–20 –13 84 154 84 –13 –20]

3.2 Filtering for 2D Content Packed as 3D Frame-Compatible Content

By using complementary sampling, it is possible to reconstruct a full-resolution 2D image, excluding the effects of the compression system, if no filtering is applied. Consequently, where the source material is 2D, all filtering can be removed, or appropriately modified given the characteristics of the signal, before the SBS multiplexer (see Figure 3).

4 SEI Messaging for the Dolby Frame-Compatible System

The mandatory parameters for the specification of the Dolby compatible SbS 3D video format using the frame packing arrangement SEI message are listed in Table 1.

Table 1 Mandatory Parameters

Parameter	SbS Value	TaB Value
`frame_packing_arrangement_type`	3	4
`quincunx_sampling_flag`	0	0
`content_interpretation_type`	1	1
`spatial_flipping_flag`	0	0
`frame0_flipped_flag`	0	0
`frame0_grid_position_x`	4*	8
`frame0_grid_position_y`	8	4
`frame1_grid_position_x`	12	8
`frame1_grid_position_y`	8	12

* Redefined after Issue 1 of this document. The values given correspond to those defined in the *Defect Report on Rec. ITU-T H.264 (2010-03)* document. This can be downloaded from http://wftp3.itu.int. Check for updates to this document before implementing this specification.

5 References

[1] *Advanced Video Coding for Generic Audio-Visual Services*, available for download at http://www.itu.int

[2] ISO/IEC JTC 1/SC 29/WG11, *Text of ISO/IEC 14496 10 Advanced Video Coding* (Third Edition), Sullivan, Wiegand, Marpe, and Luthra, Doc. N6540, July 2004

[3] ISO/IEC JTC 1/SC 29/WG11, *ISO/IEC 14496-10:200X/FPDAM 1,* GJ Sullivan, AM Tourapis, Yamakage, and Lim, April 2009