SCTE® · ISBE®
STANDARDS

Digital Video Subcommittee

AMERICAN NATIONAL STANDARD

ANSI/SCTE 187-3 2019

Informative Guidance for Stereoscopic Video

NOTICE

TABLE OF CONTENTS

1.0 SCOPE 5

2.0 INFORMATIVE REFERENCES 5

3.0 COMPLIANCE NOTATION 6

4.0 DEFINITIONS AND ACRONYMS 7

5.0 INTRODUCTION 7

6.0 DOCUMENT STRUCTURE 8

7.0 GUIDANCE FOR HORIZONTAL IMAGE TRANSLATION (HIT) OF HIGH DEFINITION STEREOSCOPIC VIDEO PRODUCTION 8

8.0 INTRODUCTION TO HIT ADJUSTMENTS 9

9.0 RESIZING CONTENT FOR DIFFERENT SCREENS AND AUDIENCES 13

10.0 A BETTER APPROACH 20

11.0 CONCLUSION AND IMPLICATIONS: 20

LIST OF FIGURES

FIGURE 1 - FOCUS OF PART-1 AND PART-2 8

FIGURE 2 - THE EFFECTS OF THE STEREOGRAPHER'S PARALLAX BUDGET IN RELATIONSHIP TO THE SCREEN PLANE 10

FIGURE 3 - EXCESSIVE POSITIVE PARALLAX AND DIVERGENCE 11

FIGURE 4 - PARALLAX BUDGET SHIFTED FORWARD 11

FIGURE 5 - RED-CYAN ANAGLYPH STEREO IMAGES WITH SYMMETRIC HIT ADJUSTMENTS 12

FIGURE 6 - SYMMETRIC HIT ADJUSTMENTS WITH VIEW DISCREPANCIES 15

FIGURE 7 - SBS FORMAT AND TAB FORMAT AFTER SYMMETRIC HIT 16

FIGURE 8 - ASYMMETRIC HIT 16

FIGURE 9 - SBS AND TAB FORMAT OF ASYMMETRIC HIT 17

FIGURE 10 - HIT-ADJUSTED IMAGE AFTER CROPPING AND SIDE PANELS ADDED 17

FIGURE 11 - HIT-ADJUSTED IMAGE AFTER CROPPING AND PROCESSED AS SBS AND TAB 18

FIGURE 12 - STEPS FOR SUCCESSFUL HIT-ADJUSTED IMAGES WITHOUT VIEW DISCREPANCIES 19

FIGURE 13 - MAXIMUM HIT-ADJUSTMENT OF IMAGE USING 2K FORMAT AS SOURCE WITH 1920X1080 FINAL OUTPUT SIZE AFTER CROPPING 20

1.0 SCOPE

This document provides informative guidance for the construction or production of stereoscopic 3D programming material intended for transmission or distribution using the frame-compatible stereoscopic 3D format defined in part 1 [SCTE 187-1] and part 2 [SCTE 187-2] of this standard.

2.0 INFORMATIVE REFERENCES

The following documents may provide valuable information to the reader but are not required when complying with this standard.

2.1 SCTE References

[SCTE 187-1] SCTE 187-1 201x Stereoscopic 3D Formatting and Coding for Cable

[SCTE 187-2] SCTE 187-2 201x Stereoscopic 3D PSI Signaling

2.2 Standards from other Organizations

[ST-0428-1] SMPTE ST-0428-1-2006, "D-Cinema Distribution Master - Image Characteristics (2048 x 1080)"

2.3 Published Materials

[1] Lipton, L., "Stereoscopy glossary," Wordpress, URL: http://lennylipton.wordpress.com/2009/03/16/glossary/ (2009)

[2] McCarthy, S., "Glossary for video & perceptual quality of stereoscopic video," 3DatHome, URL: http://www.3dathome.org/files/ST1-01-01_Glossary.pdf (2010)

[3] Yamanoue, H., Okui, M. and Yuyama, I., "A study on the relationship between shooting conditions and cardboard effect of stereoscopic images," Circuits and Systems for Video Technology, IEEE Transactions on , vol.10, no.3, pp.411-416, (2000)

[4] Todd J. T., Tittle J. S. and Norman J. F., "Distortions of three-dimensional space in the perceptual analysis of motion and stereo, "Perception, 24(1), 75 – 86, (1995)

[5] Various authors, "Perspective distortion (photography),"Wikipedia, URL: http://en.wikipedia.org/wiki/Perspective_distortion_%28photography%29 (2010)

[6] Various authors, "Perspective distortion (graphical),"Wikipedia, URL: http://en.wikipedia.org/wiki/Perspective (graphical) (2010)

[7] Kurtz, H., "Orthostereoscopy, " J. Opt. Soc. Am. 27, 323–336 (1937).

[8] Harper, B. and Latto, R., "Cyclopean vision, size estimation, and presence in orthostereoscopic images," Presence: Teleoperators and Virtual Environments, Vol. 10, No. 3, Pages 312–330 (2001)

[9] Yamanoue, H., Nagayama, M., Bitou, M., and Tanada, J., "Orthostereoscopic conditions for 3D HDTV," Proc. SPIE 3295, 111 (1998)

[10] Masaoka, K., Hanazato, A., Emoto, M., Yamanoue, H., Nojiri, Y., and Okano, F., "Spatial distortion prediction system for stereoscopic images," J. Electron Imaging 15, 013002 (2006)

[11] Wattie, J., and Di Marzio, F., "Di Marzio formulae for stereo base, 3D base controversy further explained," URL: http://nzphoto.tripod.com/stereo/3dtake/fbmarzio.htm (2006)

[12] Wattie, J., "Bercovitz formulae for stereo base; 3D base controversy further explained," URL: http://nzphoto.tripod.com/stereo/3dtake/fbercowitz.htm (2009)

[13] Seigle, D. and Sanders, J., "3D's immersive experience at home: Why the home 3D experience will not rival that of the theater," In-Three, URL no longer available: http://www.practical-home-theater-guide.com/support-files/3dinthehome-v2.pdf (2009)

[14] Siegel, M., Tobinaga, Y., and Akiya, T., "Kinder gentler stereo," Proc. SPIE 3639, pp. 18–27 (1998).

[15] Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S., "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," Journal of Vision, 8(3):33, 1–30, URL: http://journalofvision.org/8/3/33/ (2008)

[16] Dodgson, N.A, "Variation and extrema of human interpupillary distance," Proc. SPIE Vol. 5291, 36–46, (2004)

[17] Mendiburu, B., [3D Movie Making: Stereoscopic Digital Cinema from Script to Screen], Focal Press, (2009)

[18] Gunnewiek, R. K., and Vandewalle, P., "How to display 3d content realistically", Proc. Fifth International Workshop on Video Processing and Quality Metrics for Consumer Electronics, (2010)

3.0 COMPLIANCE NOTATION

"SHALL"	This word or the adjective "REQUIRED" means that the item is an absolute requirement of this specification.
"SHALL NOT"	This phrase means that the item is an absolute prohibition of this specification.

"SHOULD"	This word or the adjective "RECOMMENDED" means that there may exist valid reasons in particular circumstances to ignore this item, but the full implications should be understood and the case carefully weighted before choosing a different course.
"SHOULD NOT"	This phrase means that there may exist valid reasons in particular circumstances when the listed behavior is acceptable or even useful, but the full implications should be understood and the case carefully weighed before implementing any behavior described with this label.
"MAY"	This word or the adjective "OPTIONAL" means that this item is truly optional. One vendor may choose to include the item because a particular marketplace requires it or because it enhances the product, for example; another vendor may omit the same item.

4.0 DEFINITIONS AND ACRONYMS

FC-S3D	Frame-Compatible Stereoscopic Three-Dimensional
HIT	Horizontal Image Translation
PSI	Program-Specific Information
PMT	Program Map Table
SbS	Side-by-Side
TaB	Top-and-Bottom
ZPS	Zero Parallax Setting

Frame-Compatible Stereoscopic Three-Dimensional (FC-S3D): Refers to video content composed of left and right stereoscopic image pairs assembled into single packed frames for delivery through legacy video distribution systems. The left and right image pair are typically subject to a filtering, decimation and formatting process to generate a packed frame that has the same pixel count as the original left or right frame. A reverse of this process is performed to reconstruct the full stereoscopic image prior to display. Examples of frame-compatible formats include top-bottom and side-by-side.

5.0 INTRODUCTION

This standard is part three of a three-part standard that describes the use of stereoscopic, three-dimensional (S3D) video programming using a frame-compatible delivery mechanism for cable systems in North America. In many ways the FC-S3D signals can be processed and handled in the same way as flat (2-dimensional) video programming, and hence it is described as frame-*compatible*. The purpose of this three-part standard is to define those parts that are necessarily different from conventional (2-dimensional) video programming.

6.0 DOCUMENT STRUCTURE

Part-1 of this standard (this document) [SCTE 187-1] defines the video formatting and constraints as well as specific 3D signaling that is part of the video user data bits as shown below in Figure 1. Part-2 of this standard SCTE 187-2] defines the program-specific information (PSI) requirements for signaling, which are carried in the program map table (PMT), which is also illustrated below in Figure 1. Finally this document, (part-3) is an informative document that provides tutorial and reference information about the implications of certain formatting as it applies to stereoscopic 3D production and content preparation.

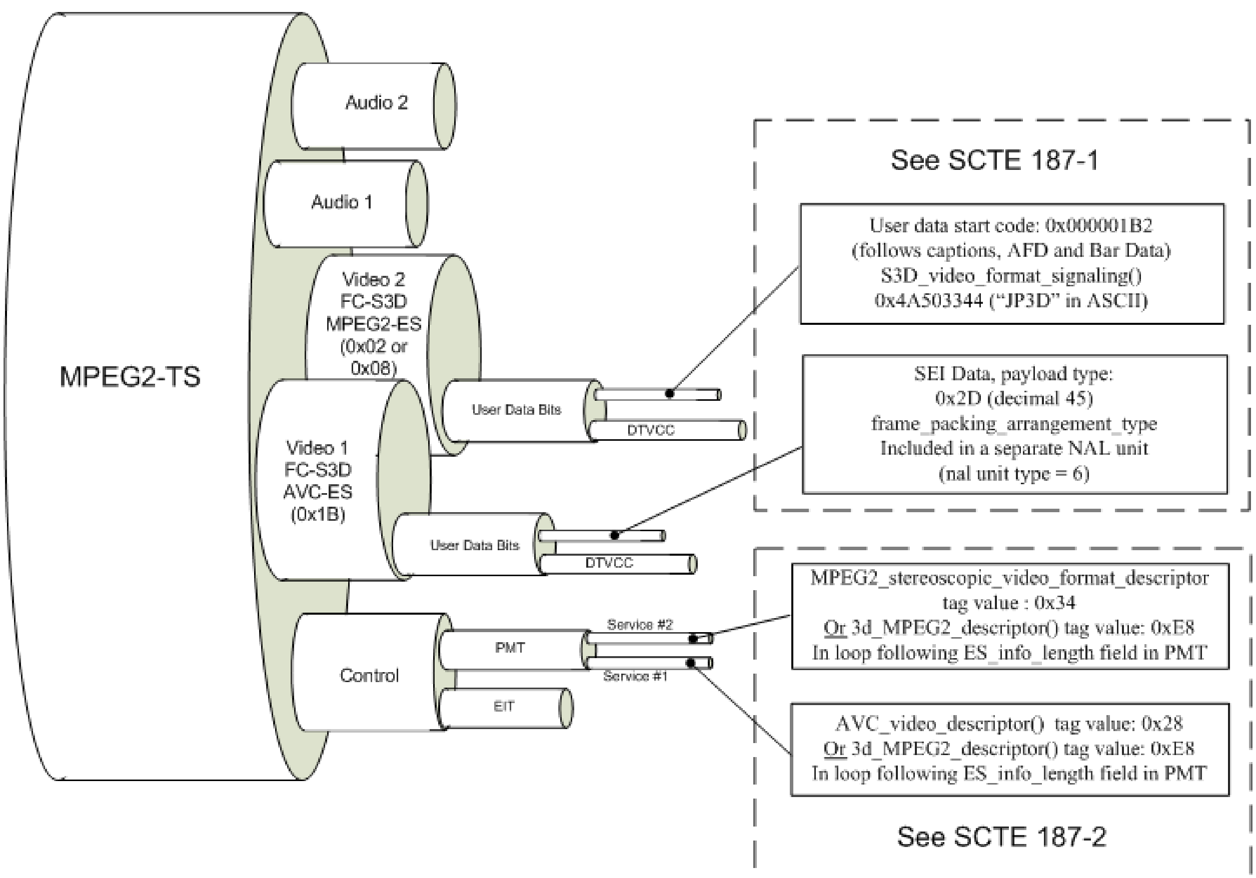

Figure 1 - Focus of Part-1 and Part-2

7.0 GUIDANCE FOR HORIZONTAL IMAGE TRANSLATION (HIT) OF HIGH DEFINITION STEREOSCOPIC VIDEO PRODUCTION

Horizontal image translation (HIT) is an electronic process for shifting the left-eye and right-eye images horizontally as a way to alter the stereoscopic characteristics and alignment of 3D content after signals have been captured by stereoscopic cameras [1]. When used cautiously and with full awareness of the impact on other interrelated aspects of the stereography, HIT is a valuable tool in the post production process as a means to modify stereoscopic content for more comfortable viewing. Most commonly it is used to alter the zero parallax setting (ZPS), to compensate for stereo window violations, or to compensate for excessive positive or negative parallax in the source material [2].

As more and more cinematic 3D content migrates to television distribution channels, the use of this tool will likely expand. Without proper attention to certain guidelines, the use of HIT can actually harm the 3D viewing experience. This paper provides guidance on the most effective use and describes some of the interrelationships and trade-offs. This paper recommends the adoption of the cinematic 2K video format as a 3D source master format for high-definition television distribution of stereoscopic 3D video programming.

8.0 INTRODUCTION TO HIT ADJUSTMENTS

There are at least three related variables that the stereographer is continuously trying to control when shooting for a given screen size and viewing distance. The stereographer uses (1) the focal length of the camera lens, (2) the distance from camera-to-subject, (3) the distance between the lenses, and sometimes (4) the toe-in or convergence angle of the cameras as his knobs to calibrate the stereo presentation for a single screen size and viewing distance. When horizontal image translation is used downstream in the post-production process, there is a risk of unintended consequences to the chosen balance of these interrelated parameters determined by the stereographer in the original production.

8.1 Interdependent Stereo Capture Parameters

The three primary stereographic factors impacted by these adjustments are listed below:

1. Geometric (or perspective) accuracy: The proper proportions between width and depth of objects are preserved avoiding depth distortion, elongation, flattening, or cardboard cut-out effects. [3] [4] [5] [6]

2. Orthostereoscopic accuracy: Proportions that best duplicate natural human vision are preserved avoiding hypostereo (giantism) or hyperstereo (Lilliputism). [7] [8] [9] [10] [11] [12]

3. Overall parallax: This includes the calibration and setting of the zero parallax setting (ZPS), avoidance of divergence at infinity, avoidance of window violations, minimizing the impact of ghosting, and control over negative parallax.

Of these parameters, the factors impacting items 1 and 2 are primarily determined by camera and lens settings and are not significantly impacted by subsequent HIT adjustments, but can also be impacted by changes to screen size and viewing distances. The references included above provides the reader with a sufficient understanding of the relationships of these first two parameters. The guidance this paper provides is primarily concerned with the impacts of HIT adjustments in item 3 above.

8.2 Parallax Budget

The calibration of the overall parallax budget and ZPS are by far the most critical of the three to get right. [13] [14]. Figure 2 below illustrates the concept of parallax budgets and its effect on perceived depth with reference to the viewing screen. The depth range of the original scene is represented by this model, but the perceived distance from the viewer can be altered directly through HIT adjustments.

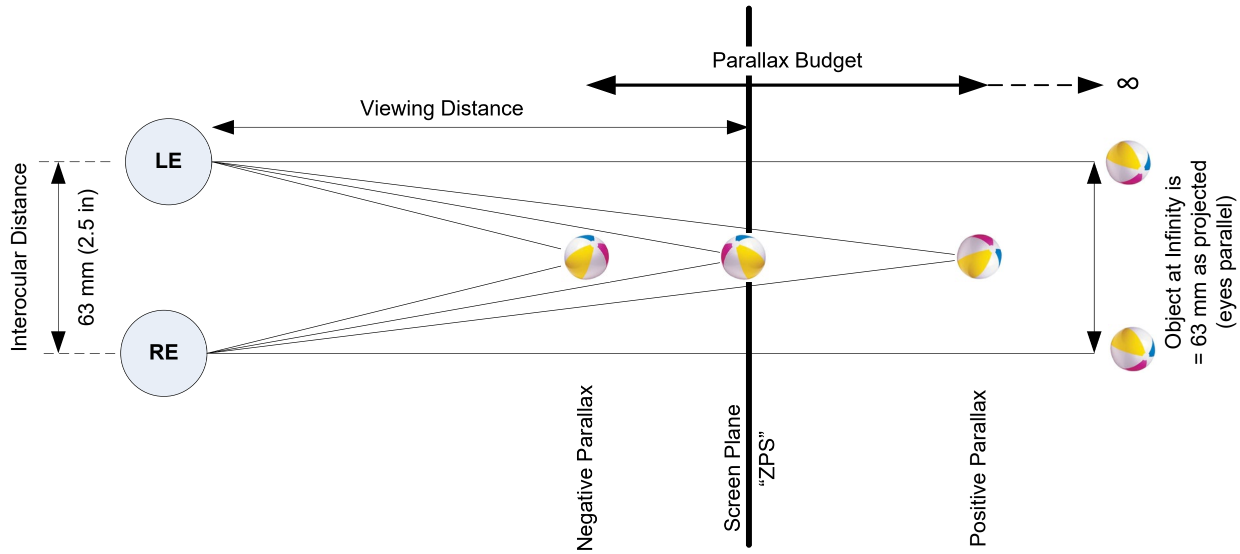

Figure 2 - The effects of the stereographer's parallax budget in relationship to the screen plane

Any failure or errors in the calibration here can compromise the viewer's ability to fuse, create excessive vergence-accommodation conflict [15], or force divergence of the eyes. Figure 3 illustrates an example with excessive positive parallax where objects behind the screen are projected with spacing greater than the distance between the eyes (typically 63 mm [16]). This kind of error can often be corrected using a negative HIT adjustment to pull the overall parallax budget forward enough to avoid the divergence.

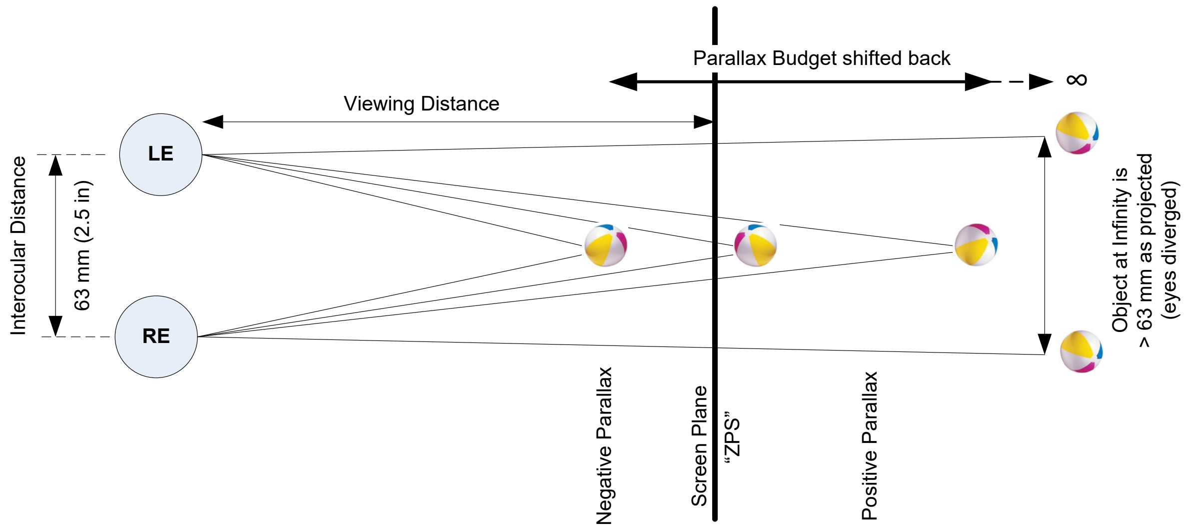

Figure 3 - Excessive positive parallax and divergence

On the other hand, the scene could contain a parallax budget calibrated too far in front of the screen plane as shown in Figure 4. In this case, a positive HIT adjustment would be necessary to push the scene back, but the amount of shift is limited by the maximum parallax of objects at infinity necessary to avoid divergence.

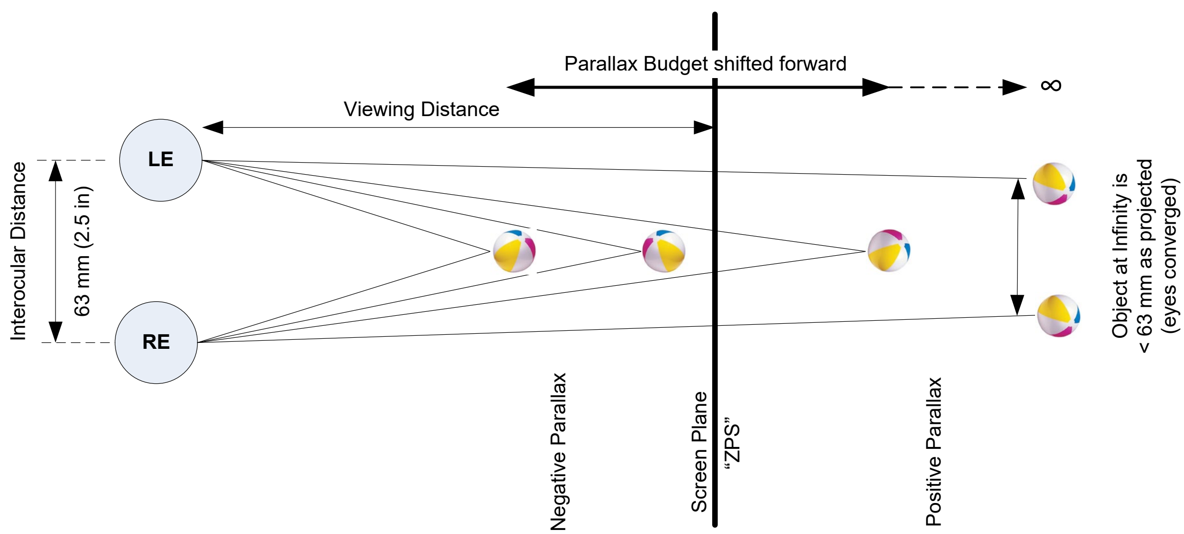

Figure 4 - Parallax budget shifted forward

To change the projected spacing between the two views of an element in a stereo image, an electronic shift is applied to the two frames in relation to each other horizontally to make the left/right spacing of the target object appear closer or further apart using horizontal image translation. An anaglyph example is shown below in Figure 5 where the left image has been shifted left and the right image shifted right, pushing the depth budget further back.

Unadjusted image with edge violations

HIT adjusted image with view discrepancies

HIT adjusted image with side panels

Figure 5 - Red-Cyan anaglyph stereo images with symmetric HIT adjustments

For example, pushing an object from audience space (negative parallax) back into the screen plane requires moving the left image left and/or the right image to the right (positive adjustment). If objects elsewhere in the same scene are already at or near infinity (63 mm), they will also be shifted and might require divergence of the eyes, which can be painful or destroy the viewer's ability to fuse. Conversely, for objects captured with excessive positive parallax (more than 63 mm), it will be necessary to pull them closer using a HIT adjustment to shift the left image right and/or the right image left to avoid the divergence. In this case any other objects elsewhere in the same scene at the screen plane will be moved into negative parallax, which could cause edge violations. [17]

9.0 RESIZING CONTENT FOR DIFFERENT SCREENS AND AUDIENCES

To understand the impact of shooting with one set of assumptions and viewing with another, an examination of the impact on all three of these areas is helpful. Even when the viewing distance and screen size proportions remain constant, there can still be some distortion. If 3D content is composed and captured for a 10 foot TV screen for viewing from 20 feet away, will the effect be different when viewed on a two-foot screen viewed from four feet away, or a 6-inch screen viewed from one foot away?

9.1 Protection from Divergence

Once sufficient protection against divergence with the largest screen size has been provided (maximum positive parallax limited), there is generally no problem with the risk of divergence on smaller screens. The geometric accuracy of objects generally appears uniform and natural across this range, provided the viewing distance remains proportional to screen size. However, this proportion does not often remain constant in the real world. As screens get smaller, the viewing distance as a proportion to screen size typically gets longer, creating a depth exaggeration. There will also be some miniaturization effect presented with smaller screens.

The bigger problems come when content developed for a very tall and very steep screen with close seating arrangements where the audience is viewing at a distance between 0.45x to 0.75x screen width from the screen (e.g., IMAX®), and it is moved to the small (television) screen optimized for typical viewing distances between 2x to 5x screen width. Many different shooting techniques have been used for the production of IMAX 3D content. Some of those techniques result in 3D formatted content with a more extensive depth range, predominantly negative parallax and very little positive parallax by proportion. Content typically projected in the original 3D IMAX (70mm) systems uses a maximum positive parallax set to no more than 0.3%[1] and maximum negative parallax of between 5% and 7%. This in-front-of-the-screen balance is okay when the edges of the screen are way off in your peripheral view (typical for IMAX viewing). However, when such content is converted for TV viewing and the viewing size is reduced down to where the edges are easily in view, these window violations are more problematic. To compensate, some stereo realignment or other techniques need to be applied. However, with this large depth range, pushing the ZPS back far enough to remove the edge violations can result in excessive positive parallax (divergence), which is best avoided.

To illustrate the interplay of these factors, consider a case where the potential divergence problem has been avoided by shooting the content with a smaller depth range in the first place. Now the problem will be the elongation effect (geometric distortion) as the viewing

[1] IMAX screens are calibrated for 2.5 inches at infinity and have a screen width of 70 feet (840 inches) = 2.5/840 = 0.0029 or 0.29%. Some IMAX screens are as wide as 110 feet and still maintain a calibration of 2.5 inches for infinity producing an even smaller percentage for positive parallax.

distance is increased in proportion to the screen height. If the content was shot for the cinema with the optimal viewing distance calibrated for less than 1x picture width and it is viewed on a TV at 3x to 5x picture width, it will have an exaggerated (elongated) depth distortion. This distortion will not hurt the viewer's eyes like the divergence error would, but it will not look natural.

The best way to minimize some of these problems, allowing for easier reapplication of cinema content for the small screen, is to use a combination of techniques. One approach would be to use some post-processing to realign the disparity between the left and right images. This HIT process is useful to push the ZPS back to avoid the strong edge violations or to pull the ZPS further forward to avoid divergence. Material that has been captured with a very wide interaxial spacing will have a correspondingly large depth range, especially when objects are proportionally too close to the cameras. In such cases, the HIT adjustment is desirable to push the ZPS back, but can easily result in excessive positive parallax (divergence). Likewise, material captured with cameras using a toe-in adjustment can more easily produce excessive positive parallax if the furthest object in the scene is not carefully controlled. In such cases, it is necessary to use the HIT adjustment to avoid divergence for TV viewing by limiting positive parallax to about 2% of screen width[2]. Another option for processing content with strong edge violations that is being repurposed for the smaller screen is to mask the offending edges with a dynamic floating window effect to minimize the conflicts.[18]

[2] The limit is calculated based on the largest potential screen size for home TV viewing, which can be as large as eight feet with home theater projection systems (2.5 inches / 96 inches = 2.6% so 2% is a relatively safe limit.)

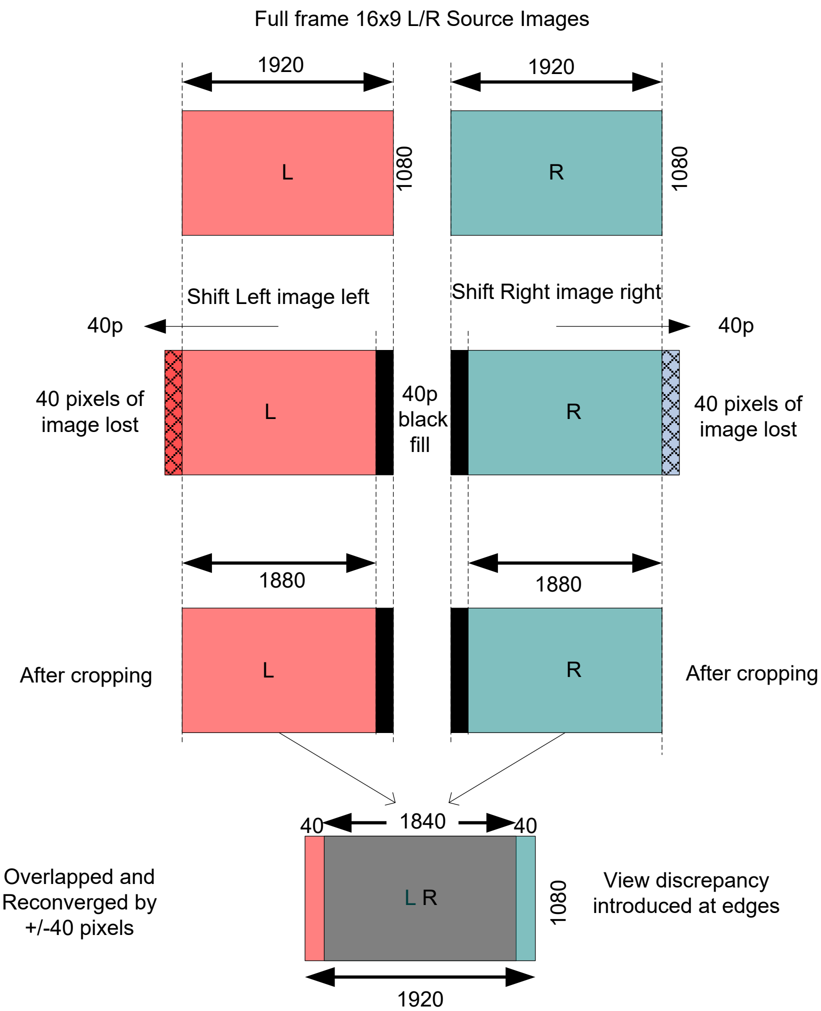

Figure 6 - Symmetric HIT adjustments with view discrepancies

HIT adjustment is illustrated in Figure 6 above with a full-frame left (red) and right (cyan) image as they would overlap when realigned or shifted by 40 pixels each. To move the ZPS further back or increase the separation at infinity, the left image is moved left and the right image is moved equally to right. Conversely, to move the ZPS further forward, reducing the separation at infinity requires moving the left image to the right and the right image equally to the left. Moving only one image is possible with many software tools, but should be avoided to prevent the resulting asymmetry of frame sizes that will result between the left and right images.

When one or both images are moved in a HIT adjustment without cropping and scaling the final scene, a border will be introduced into the stereo image visible in only one eye. This border area creates a view discrepancy on one or both edges of the stereo image as shown in Figure 6 above or Figure 7 below. The view discrepancy appears because some content remains visible for one eye, but is not visible for the other eye. This view discrepancy on the side of the screen could create false depth cues and can be distracting to some viewers.

When such HIT-processed content is panelized to the Side-by-side (SbS) or Top-and-bottom (TaB) frame-compatible formats, the image will include black bars at the center line in SbS or separate black bars on either side in TaB (as shown below in Figure 7) depending upon which way the images are moved. [SCTE 187-1] advises against such formatting in Section 8.5, however the alternative of no HIT could actually be a worse offense visually.

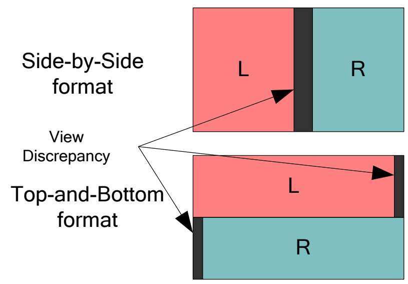

Figure 7 - SbS format and TaB format after symmetric HIT

Sometimes a HIT adjustment is applied that only moves one of the images horizontally. In this case, an asymmetry results since the viewable area of one image is larger than the viewable area of the second image. This frame size asymmetry produces a view discrepancy on only one side as the black bar on one frame lines up with a visible portion on the opposite frame. The illustration in Figure 8 below shows this condition. Some programmers, production facilities, and editing tools are using this technique as a short-cut, and it should be avoided.

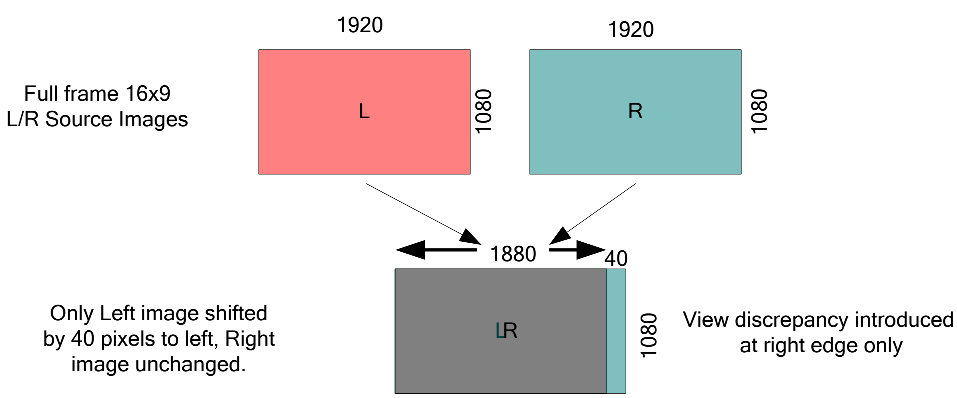

Figure 8 - Asymmetric HIT

The panelization of asymmetric HIT into SbS or TaB formats will be represented as shown below in Figure 9, where the bars are on one side of the image only. [SCTE 187-1] also advises against the use of such signals.

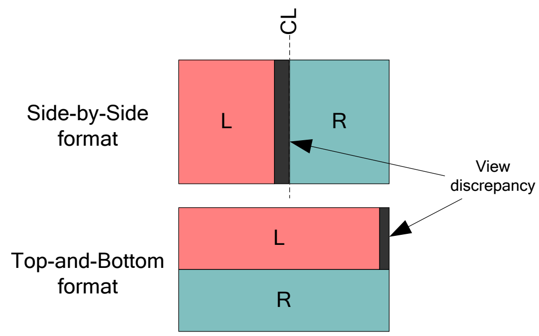

Figure 9 - SbS and TaB format of asymmetric HIT

9.2 Correcting the View Discrepancies

To avoid these unnecessary view discrepancies after any form of HIT adjustment, two correction methods are recommended. First, it is necessary to add black masks or borders to the opposite side of both images equal to the HIT adjustment shift. If the image was shifted 40 pixels, then 40 pixels of black mask is added. In the symmetric example of Figure 6 above, this leaves the usable portion of the image that remains visible in both eyes after the realignment equal to 1840 pixels, with 40 pixels of black mask on each side as shown in Figure 10.

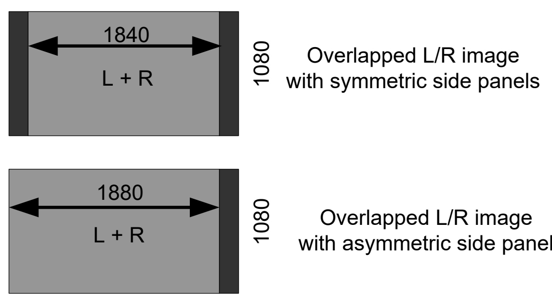

Figure 10 - HIT-adjusted image after cropping and side panels added

In the asymmetrical case of Figure 8, the usable image is 1880 pixels wide and 40 pixels of black mask should be added to the right side of both images. When these masked images are presented as SbS or TaB, they appear as shown in Figure 11 below (compare with Figure 7 and Figure 9).

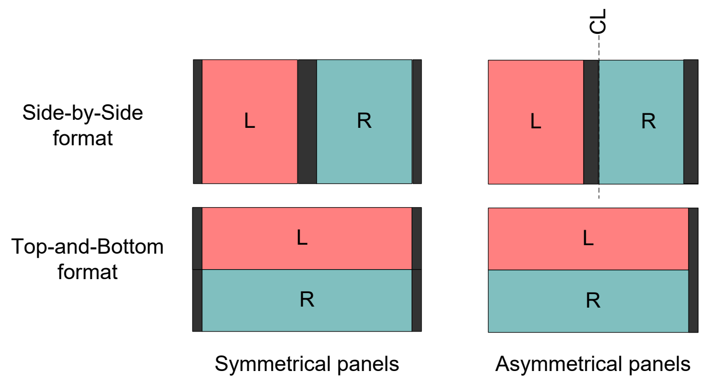

Figure 11 - HIT-adjusted image after cropping and processed as SbS and TaB

While this masking process effectively removes the view discrepancies, the remaining side panels (or pillar box) can also be problematic. These bars are visible on the TV screen and could possibly result in user modification of the aspect ratio when viewed. To avoid this, a second step is necessary. After the masking bars are added, it is helpful to rescale or resample the center-cut of the image so that it can be expanded back to the full-frame. This solution removes the need to transmit the side panels, but introduces a slight resolution loss. This is illustrated in this example where 1840 pixels are converted to 1920. It also creates an overscan in the vertical direction. This overscan is then cropped to restore the 1080 line image structure and maintain the correct aspect ratio. The top/bottom cropping can be asymmetrical as needed to preserve important details in the scene. Consequently, the content will lose some image content at the top and bottom of the frame. The full process is shown in Figure 12.

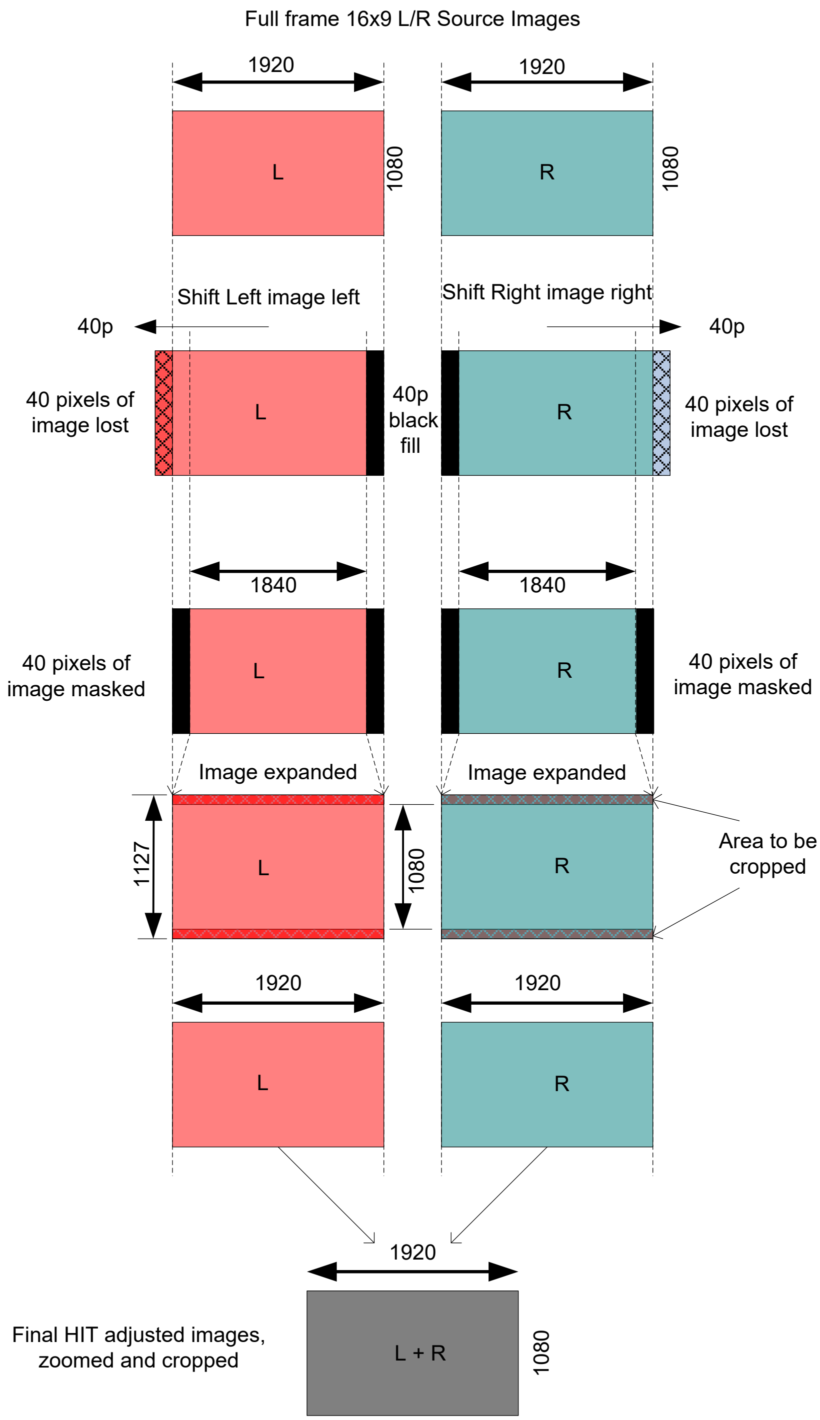

Figure 12 - Steps for successful HIT-adjusted images without view discrepancies

It should be noted that this HIT process is also quite useful as a simple technique for placing flat scenes shot in 2D behind the screen plane when they need to be edited into 3D programs. For example, when a multi-camera sporting event is captured, it is sometimes desirable to mix some 2D cameras angles into the production with 3D cameras. By applying the HIT adjustment on the same 2D frame used as both the left and right eyes, a simulated 3D image will result that appears in 3D space behind the screen plane that can easily be mixed with 3D shots dynamically. However, the same edge-clipping, and view discrepancies will occur when processing a 2D image in this way and the same corrections are necessary.

10.0 A BETTER APPROACH

None of these solutions is optimal. A better solution is to capture (2D and 3D) source material and graphics using the cinematic 2K format based on a frame size of 2048 pixels x 1080 pixels. In this way, an "adjustment margin" of up to 128 pixels[3] is preserved for subsequent stereo-window realignment that can be used without compromising the spatial resolution or cropping the image vertically. Even if no HIT is needed for distribution, the 1920 pixel portion can be center-cut by cropping the edges from the larger frame. An illustration of this is shown in the Figure 13. The standard that describes this video format is [ST-0428-1].

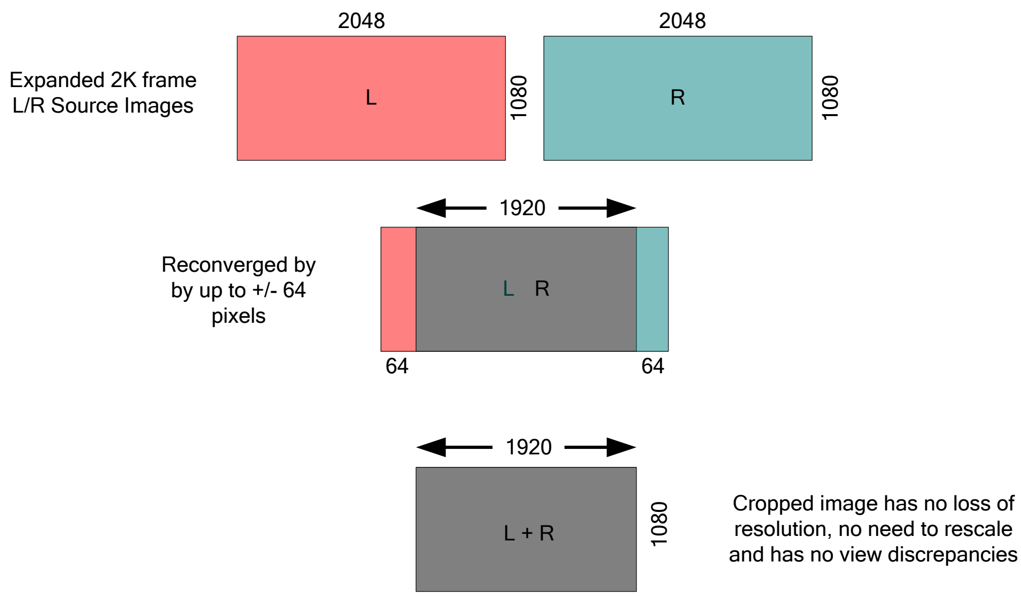

Figure 13 - Maximum HIT-adjustment of image using 2K format as source with 1920x1080 final output size after cropping

11.0 CONCLUSION AND IMPLICATIONS:

1. Horizontal Image Translation (HIT) is a valuable tool when it is necessary to alter the depth settings of stereoscopic material for a variety of reasons.

2. HIT adjustments should always be made symmetrically to both left and right frames.

---

[3] Applied as equal shifting in opposite directions for both left and right views, this leaves 64 pixels of movement for each eye or a total of 128 pixels of parallax displacement available.

3. It is necessary to have a full-frame left and right image to do the realignment.

4. If the original (full frame) images are the same frame size as the final product, HIT adjustment can introduce view discrepancies, side-panels, resolution loss, or rescaling errors.

5. To avoid these errors, it is best to capture any content to be used in a 3D broadcast at the cinematic 2K format (2048 x 1080).