3D Display size matters: Compensating for the perceptual effects of S3D display scaling

Karim Benzeroual
Centre for Vision Research
York University
4700 Keele St, Toronto, ON
Canada
karim@cse.yorku.ca

Robert S. Allison
Centre for Vision Research
York University
4700 Keele St, Toronto, ON
Canada
allison@cse.york.ca

Laurie M. Wilcox
Centre for Vision Research
York University
4700 Keele St, Toronto, ON
Canada
lwilcox@yorku.ca

Abstract

Over the last decade, advances in technology have made stereoscopic 3D (S3D) displays widely available with an ever-expanding variety of technologies, dimensions, resolution, optimal viewing angle and image quality. Of these, one of the most variable and unpredictable factors influencing the observer's S3D experience is the display size, which ranges from S3D mobile devices to large-format 3D movie theatres. This variety poses a challenge to 3D content makers who wish to preserve the three dimensional artistic context and avoid distortions and artefacts related to scaling. This paper will review the primary human factors issues related to S3D image scaling and the techniques and algorithms used to scale content.

1. Introduction

In recent years the consumer electronics market has been flooded with a variety of S3D products, which rely on a variety of display and image segregation technologies. For each display system, the ideal viewing conditions (eg. viewing angle) can be defined in order to obtain the desired 3D experience. SMPTE and THX [1, 2] have provided specific standards and guidelines for the ideal viewing angle for theatre and television. However, screen dimension[1] is an uncontrolled variable since the same content could be displayed on a mobile autostereoscopic device, 3D monitor, HD 3DTV or in a 3D movie theatre. Adapting a S3D film to a variety of screen sizes is necessary for most, if not all, popular movies if the distributors are to maximize their exposure and therefore earnings. However, unlike 2D film the S3D scaling process is complicated by a variety of computational and perceptual issues that can significantly impact the audience experience.

As outlined below, the existing approaches to scaling S3D content for a variety of delivery form factors can be divided into two main categories: those applied during acquisition and those applied during postproduction or display. The most common strategy is some combination of pre and post-production approaches. However, inevitably some degree of perceptual and geometric distortion will remain. A better understanding of these distortions and their perceptual consequences will provide S3D content creators with insight and context for using sophisticated scaling approaches based on both acquisition and post-production techniques. This paper will review the principal issues related to S3D content scaling, some of the technical solutions available to content makers/ providers and the perceptual consequences for audiences.

2. Stereoscopic Geometry

As was shown by Spottiswood in the early 1950's [3], displaying stereoscopic 3D content at different sizes may dramatically influence the audience's S3D experience. Given the interdependence of acquisition and display parameters; most filmmakers, while trying to protect for different screen dimensions will have a target viewing condition when they begin filming.

Figures 1 and 2 depict stereoscopic viewing and acquisition geometry, respectively [3] [4]. In figure 2, the observer fixates point *F* (actor at the screen plane), and an object is positioned beyond the screen at location *P* (house). Angles θ and ω are the binocular parallax of *P* and *F*, respectively. The absolute (binocular) retinal disparity is the angular difference in position of the object on the two retinas and can be computed as the binocular parallax minus the convergence angle. Relative binocular disparity η is the retinal disparity of one object relative to another:

$$\eta = \theta - \omega$$

[1] The range of viewing distances typically used are correlated with the size of the display, with audiences moving closer as screens get smaller. If field of view is constant it is often the distance that is more important. Since they normally co-vary here we will focus on screen size and related disparity scaling issues, but will point out the role of viewing distance in particular when it is warranted.

978-1-4673-1612-5/12/$31.00 ©2012 IEEE

If V is large compared to D and IO, then, η varies with the inverse of distance squared:

$$\eta \approx \frac{IO \cdot D}{V^2}$$

Images of objects maintain constant relative positions on the retina as the eye move [5, 6], making relative retinal disparity a reliable sources of depth information, to which people are remarkably sensitive [7].

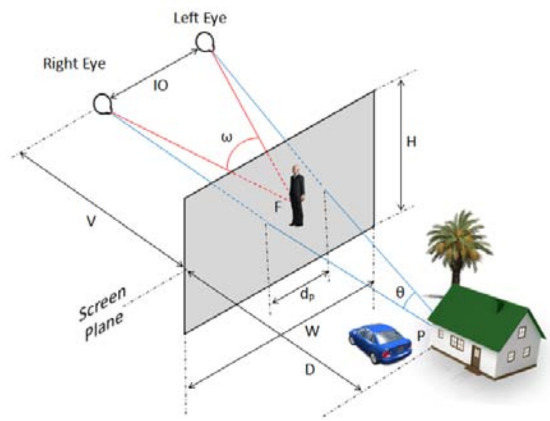

Figure 1: Here we illustrate the geometry of stereopsis with V: viewing distance; D: relative distance between an object P and the fixation point F; IO: interocular distance; ω: angle subtended by the eyes at the fixation point (screen plane); θ: angle between eyes and the object; d_p: is the on-screen disparity of the point P.

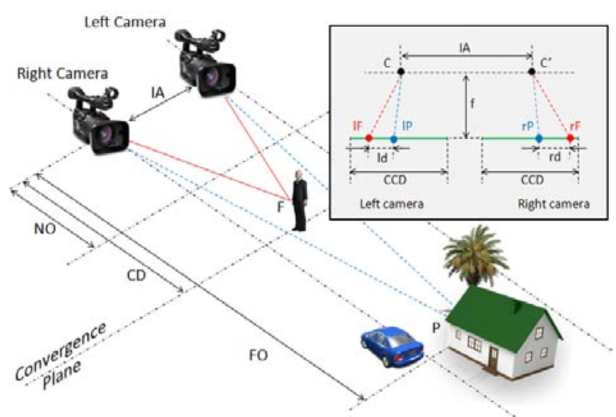

Figure 2: Here we illustrate the geometry of S3D acquisition, where, IA: Interaxial Distance; CD: Convergence Distance; NO: Nearest object; F and P: points in the scene; FO: Furthest Object in the scene; f: Focal length; C and C': cameras' optical centers; CCD: sensors' width; lF and rF: left and right projections of the point F on the left and right camera sensors, respectively; lP and rP: left and right projections of the point P on the left and right camera sensors; rd and ld: the right and left relative disparity in pixels between the points P and F, respectively.

In Figure 2, the cameras are parallel and shooting a scene which contains a car, actor, house and tree positioned at different distances from the cameras. One parameter that stereographers typically control is the maximum range of binocular parallax and hence the range of relative disparity presented to the viewer. For example, Bercovitz [8] proposed a formula to calculate the optimal IA based on object distances in the scene:

$$IA = d * \frac{FO * NO}{FO - NO} * (1/f - 1/a)$$

Where: "d" is a parameter specifying the maximum disparity range (upper bound) in the images formed on the sensor plane. This is the difference in image disparity between the closest and furthest points in the scene, d= rd – ld (see Fig. 2) and "a" is the focal distance.

3. Technical Solutions

As mentioned above, S3D content may be adapted for different screen sizes during content creation, or in postproduction. In this section we review some of these approaches.

3.1. During content creation

3.1.1 Multiple-films

The most obvious way to ensure that S3D content is optimized for a given display is to create multiple films shot with different camera parameters to target a variety of screen sizes. This strategy involves capturing the same mise-en-scène using different rig configurations where each configuration is optimized for specific viewing conditions. This approach is time consuming, complicated in terms of staff management, very costly and, since it relies on multiple takes, and can lead to variability in quality or content between the versions of the film. It is rarely used in live action filmmaking but is more feasible for animation as scenes can be re-rendered with different stereo camera rigs more reliably than in live action. Nevertheless, even for animation it is uncommon as it increases the required computational, storage and post-production resources.

3.1.2 "Protecting" for different screen sizes

Given limited funding, most productions create one film to be viewed on all devices. To ensure that their content is suitable for a large range of displays, the stereographer and director of photography may 'protect' for other screen sizes (typically smaller than the target) by choosing settings that work well on the target and at least adequately on the other screen sizes. This ensures the stereoscopic experience on these other devices is, if not optimal, at least comfortable. This compromise is normally then locked in, although "repurposing" techniques could be applied during postproduction (see section 3.2) to adapt the film.

3.1.3 Multi-View Systems

In multi-view systems a scene is shot with multiple S3D cameras simultaneously making it more time efficient than sequential re-filming. Two variants of this approach have been suggested for S3D film production.

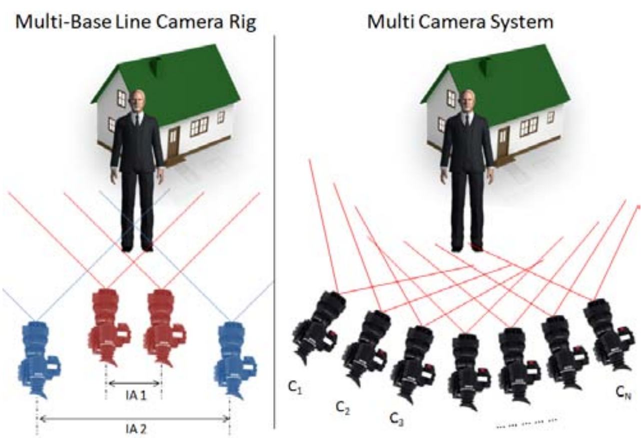

Figure 3: Multi camera S3D capture.

The first is a multibaseline stereo system [9], where two groups of two cameras with different IAs are embedded in one system (see fig.3 left). This approach may be a suitable compromise for live action film makers because it allows them to target different screen sizes without resorting to costly postproduction treatments. In terms of efficiency, this offers a "half way solution" since only two viewing conditions can be perfectly matched. Generally, filmmakers will configure their system for the average large and small display scale in order to minimize (but not avoid) perceptual artefacts.

The second variant is a multi camera configuration in which N cameras are used to capture the scene (see Fig.3 right) simultaneously. In this case, the N captures can be used for a multi-view autostereoscopic displays or as camera pairs which provide different IAs [10] [11]. This approach is well-suited to multi-view autostereoscopic displays and can serve as input for view interpolation algorithms (see section 3.2.2).

3.2. Post-production techniques for virtual rig control

In the standard S3D post-production process it is common to adjust the location of the scene relative to the screen plane by manipulating the zero parallax setting (ZPS). A more complex process is required to alter the IA, which is typically applied to address content rescaling [12]. These techniques allow film makers to optimize their footage for specific viewing conditions after the content is captured, and so provide much needed flexibility for viewing platforms. In this section we will describe the main techniques used to adapt the effective acquisition geometry during the postproduction stage.

3.2.1 Zero Parallax Setting Correction

The ZPS defines the position of the convergence plane in the scene (see Fig.2); it can be changed by shifting the left and the right images horizontally in opposite directions. The ZPS adjustment is often used for artistic purposes to control how much of the scene is positioned in front of or behind the screen plane. It is also used to avoid or minimize frame violation artefacts, and to reposition the subject of interest relative to the screen plane. ZPS adjustments are also performed in postproduction to adapt content for different screen sizes. For instance, to remove divergent parallax introduced in far objects (see section 4.1.2) when content is moved to a larger screen.

3.2.2 View Interpolation

Assume that we have a S3D acquisition system composed of two parallel cameras placed at positions P_0 and P_1, and oriented toward a point in space. View interpolation [13] [14] [15] [16] is a technique that generates a new virtual camera view positioned between P_0 and P_1. This technique makes it possible to change the effective IA during postproduction. Such algorithms are suitable for decreasing the IA and can be useful for adapting content that has been targeted for medium to large displays, to make it suitable for hand-held devices. Specifically such algorithms may prove useful for reducing the appearance of cardboard cut-out effects (see section 4.1.3).

3.2.3 View Synthesis & Free view point

View Synthesis [17] [18] [19] [20] [16] is a technique that offers more freedom than view interpolation as a new view can be generated from a virtual camera positioned at any location in the scene. This technique requires a complete geometric model of the scene, typically constructed from multiple camera views (from 10 to 100 or more). Because view synthesis provides 'virtual' views from any point in the scene, it can be used to increase the IA (unlike the interpolation technique). This is a potentially powerful solution to display scaling issues since the acquisition geometry could be completely redefined in the post-production process. While promising, the use of image synthesis algorithms to repurpose S3D film faces serious challenges. For instance, view synthesis requires depth information for all points in a scene from any arbitrary position. A spherical camera array arrangement might appear to resolve this issue, but poses practical issues, such as camera coverage and visibility. Deep concavities for example can cause difficulties as their interior can only be viewed from a narrow set of vantage points and insufficient density or errors in depth estimates can be a problem with stereo vision approaches.

Therefore, instead of using stereoscopic camera rigs, other acquisition systems may be used to reconstruct a dynamic scene in 3D, for example Gurram et. al [21] fused imagery data from a LIDAR laser scanner and a passive camera, Yang et. al [22] designed a real-time light field camera which allows the navigation in the scene through different virtual view points. However, these systems also typically have visibility or model completeness issues. An intermediate solution is to use disparity remapping techniques [23] [24]. Such approaches use stereo-matching algorithms to identify corresponding points in the two eye's views, and create an intermediate disparity. While easier to implement from a practical perspective than view synthesis or interpolation techniques, the stereo-matching processes are error prone (requiring monitoring and correction) and time consuming.

3.2.4 2D plus Depth

Using the 2D plus depth technique [25] [16] [26], stereoscopic content is represented using a 2D image sequence plus a depth map sequence, rather than separate left and right stereo image sequences. The depth map is a grayscale image associated with each frame and defines the depth information for each pixel in the frame. Such a representation can be extracted from stereovision algorithms from traditional S3D content or from a single camera view combined with an associated range image (say from a Lidar scanner if the scene is static). To display the content, the process is reversed and new views must be synthesized. The quality of the reconstructed S3D content is highly dependent on the precision and resolution of the depth map.

The main advantage of the 2D plus depth technique is that the representation is agnostic about the intended display device and multi-view synthesis for multi-view autostereoscopic displays is treated in the same way as two view stereo synthesis. An additional benefit is that the views are synthesized and so can be tailored to the size and viewing distance of the display or for user preferences and tolerances. The main disadvantage of this technique is that depths are only assigned to pixels and thus parts of the image that are visible in the camera view.

3.2.5 2D to 3D conversion

2D to 3D conversion techniques are a set of algorithms that allow the generation of a second camera view from 2D content. This is similar to 2D plus depth but the image is segmented into regions and disparity in the regions is generated from the image, not the scene, based on human or automated disparity assignment. Sisi et. al [27] described the most popular 2D to 3D conversion algorithms based on techniques such as: Motion Parallax [28], Focus [29], Still image analysis [30], Depth map generation [31], Vanishing line detection [32], Motion detection [33], Line tracing method [34], and Color analysis (real time application) [35].

It is rare that a film production will have access to extensive depth maps or multi-camera imagery. Therefore, 2D to 3D conversion techniques are currently the most common approach to modifying S3D content. The biggest challenge facing these efforts is the need to adequately reproduce geometrically consistent monocular occlusions [36] (see section 4.2). Some of the conversion process can be achieved using computational tools. However, quality conversion requires specialized rotoscoping, compositing and in-painting techniques that are typically performed by hand, on a frame-by-frame basis. As outlined in section 4.2 the human visual system is very sensitive to inappropriate occlusion information.

4. Perceptual Consequences

4.1. Effects of Disparity Range

Many of the unwanted perceptual effects of display scaling are due to the disparity range of the scaled content being inappropriate for the intended viewing distance. This can produce compressed or exaggerated depth and scale, and/or oculomotor and fusion issues.

In some films, S3D effects are used for dramatic impact. The goals of increasing the immersion and the spectacular aspects of S3D can challenge oculomotor and perceptual mechanisms that maintain binocular visual comfort. These challenges can be exaggerated when the content is scaled for a different display size (or more precisely when viewing distance is changed to be appropriate for these scaled displays). Perceptual, psychological and physiological side effects may arise when the viewing conditions deviate excessively from natural human binocular vision. Two of the most important factors to consider are vergence-accommodation mismatch and the motor fusion range (including divergence limits).

4.1.1 Vergence & Accommodation

Under natural binocular viewing conditions, vergence and accommodation represent two of the main oculomotor mechanisms that allow the viewer to maintain sharp focus on objects at different distances. In horizontal vergence the two eyes move in opposite directions to align the high acuity fovea of both eyes on an object of interest by crossing (convergence) or uncrossing (divergence). However, pointing the fovea at a target does not guarantee clear vision if the image is not well-focussed. Focus or accommodation is achieved by intraocular muscles that change the shape of the lens to position the image of an object on the retina. Under normal viewing conditions these two processes are tightly coupled, so that as an observer converges on an object, the accommodative system maintains its focus. These links are not just a habitual correlation between the two responses but are 'wired in' through neural cross-links between the two systems.

In most 3D display systems, the normal vergence/accommodation relationship is not preserved since focus should always be at the screen plane while vergence should follow the disparity of the target. This is a particularly acute conflict at close viewing distances (less than 2m) and can cause visual fatigue [37] when sustained. As small displays are usually viewed at short distances, depth range in S3D content should be reduced for handheld devices and other small form factors.

4.1.2 Fusion Limits & Divergence

If the retinal disparity of an object is too large it fails to fuse into a single binocular object and may be seen as a pair of monocular objects (diplopic)[2]. The upper limit of binocular fusion using an HD 3DTV is approximately 2° to 3° positive or negative parallax [38]. As a rule of thumb, conservative content makers use a smaller range such as 1° [12] to prevent discomfort and visual fatigue.

To see how scaling can introduce divergence, suppose that the image of an object displayed on a small screen has positive parallax equal to the observer's IO. When the object is properly fixated the observer's eyes will be parallel, consistent with the object being placed at infinity. If that same content is displayed on a much larger screen, the magnification will cause the parallax to exceed the observer's IO and the eyes will need to diverge beyond parallel to fixate it. Divergence beyond parallel is an unnatural eye movement (requiring fixation 'beyond' infinity) and observers have only limited ability to diverge. Sustained divergence may cause discomfort and visual fatigue and is usually avoided [12]. If the disparity range is not excessive, this issue can be met with in the postproduction stage by adjusting the ZPS (3.2.1).

[2] Diplopia is a natural feature of normal binocular vision and thus should not pose a problem. However in S3D displays the normal relationship between fusion range and depth of field and other factors is disturbed.

4.1.3 Distortions of Size and Depth

Scaling S3D content simultaneously affects size and depth percepts but in different ways, resulting in a number of inter-related visual distortions. The size/depth interaction occurs partly because, when the screen size and/or the viewing distance are changed, the desired IA should be recalculated to match the changes in the viewing conditions (but is not).

It is known that the perceived size of an object can vary depending upon the IA, with objects appearing smaller when large IAs are used. This has been attributed to convergence micropsia [39] but size distortions can also be caused by failures of size constancy that are not related to convergence, when the visual system misestimates the distance to an object.

The visual system interprets cues to distance and size in a flexible manner to maintain object form or to resolve paradoxes. With a large IA, particularly when combined with long focal length lenses, miniaturization may occur. This is a phenomenon where objects appear tiny, but appropriately proportioned (also known as the puppet theatre effect). Note that when the IA is increased, the image size of the object does not change, but the additional disparity will distort the apparent shape of the object. If we assume that the distance to a display or to the content presented in the scene is somewhat ambiguous, then the visual system compensates for the disparity-induced shape distortion by assuming it is closer to the observer and therefore smaller.

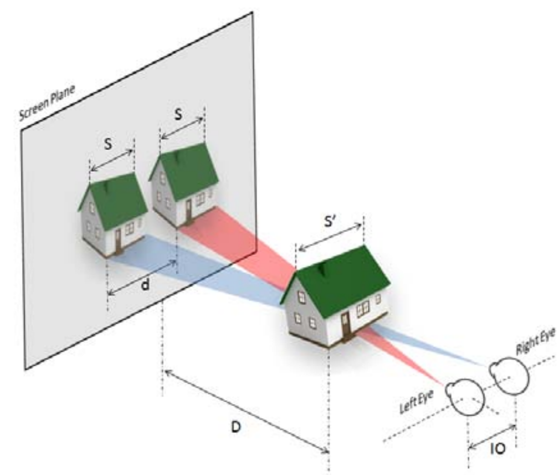

Figure 4: Here we illustrate the phenomena where the visual system scales the apparent size (S') based on depth (D) from disparity (d) and size from visual angle (S).

In effect, the visual system 'chooses' an apparent distance where the ratio of depth from disparity and size from visual angle is normalized, leading to a perception of a proportionate but scaled object (Fig. 4). An analogous effect can be produced by excessively small IA causing the scene to appear enormous in scale (gigantism). In both cases, the normalization depends on an internal model or bias for the shape of the object or scene.

The cardboard cut-out effect is another well-known phenomenon in S3D filmmaking and photography that is tied to rescaling content. It is related to the puppet theater effect described above but without the normalization process that preserves proportions.

Phenomenologically, the cardboard cut-out effect is the perception of objects being flattened, and lying on a series of layers in a scene at different depths rather than having a natural appearance of volume. A mismatch between perceived size and perceived depth is the basis of this effect with cut-out percepts occurring when the depth is small compared to the perceived size. As outlined by Howard and Rogers[40], if the viewing distance is underestimated then the perceived depth from disparity (which increases proportionally to the square of the viewing distance (V^2)) will be small compared with the apparent size from perspective (which is proportional to the viewing distance *(V)),* resulting in object flattening. This explanation does not explain why the spaces *between* objects seem less compressed. Rogers has highlighted the role of cue conflict in these percepts [41].

In S3D displays these images are usually presented on screen planes that are much nearer than the distance of the original scene. This combination of factors results in severely foreshortened objects and the compression of depth across the scene into separate layers. Thus, content intended for theatrical release may exhibit cardboard cutout phenomena when displayed on a smaller display at a nearer distance.

4.1.4 Scaling techniques and disparity range effects

Control of effective IA through any of the methods in section 3 can limit disparity to a fusible range and prevent divergent disparity in the background. While shifting the ZPS does not change the total amount of binocular parallax, ZPS control can readjust the disparity range relative to the screen and avoid divergent disparity in the large format variant of a film.

Techniques that use a fixed number of views are necessarily constrained in adjusting for distortions and disparity range. For example, film makers often deliberately select larger IAs when they know that their film will primarily be shown on a large screen. If that content is then shown on a small display close to the observer, the resultant scaling artefacts may degrade the depth percepts and introduce unwanted distortions. The creation of a single S3D film and protecting it for multiple screen sizes implies a compromise between control of disparity range and stereoscopic distortion and the range of screen sizes that can be protected.

Similarly, the success of multi-camera/multi-film approaches depends on how well the set-up matches targeted screen sizes. In the extreme, the multi-film approach could eliminate the perceptual consequences of scaling, by providing a camera pair for each possible display. In practice these must be approximated or view interpolation used. Thus, geometric depth distortions resulting from non-optimal viewing arrangement will remain. Fortunately for the S3D industry, audiences are generally able to tolerate modest distortion for both 2D and S3D films.

Techniques that create a specific IA by generating a new view can ideally match a desired stereoscopic viewing geometry. However, the matching and reconstruction processes are error prone and the quality of the reconstruction depends on the segmentation, disparity assignment and occlusion handling processes.

4.2. Monocular Half-Occlusions

Many of the computational algorithms described in section 3 require the generation of new camera views to adapt the acquisition parameters for specific display sizes. One of the developers' main challenges is to detect and manage monocular half-occlusions.

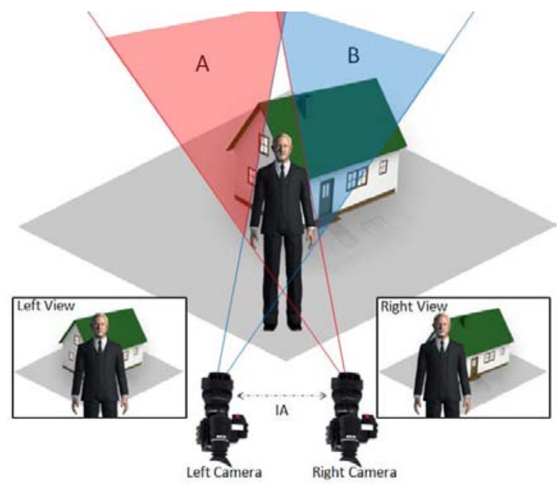

Figure 5: Monocular occlusions, the central area (Actor) is visible to both eyes; areas A and B are visible to the right and left eyes only, respectively.

In S3D filming, as in natural vision, there are regions of a scene that are visible to one eye but not the other. Such monocularly occluded regions occur at object boundaries, where the object occludes part of the background or another object from one eye's view. Such occluded regions also occur within an object where the observer can see more of the right side of the object in the right eye, and vice versa as shown in Figure 5. When an object is occluded by another object in both eye's views it is a binocular occlusion and invisible. However, such objects may become binocularly or monocularly visible when the IA is changed. To further complicate matters, movement of objects or of the cameras will dynamically introduce and obscure additional monocular regions.

Generally, the monocular occlusions contained in stereoscopic content shot with a dual camera rig are fully consistent with the binocular disparity in the scene. Thus, techniques that use pairs of real cameras with the desired IA such as capturing multiple versions of the film, protecting for screen size and multi-view acquisition techniques, normally acquire the appropriate monocular and binocular occlusion information. As outlined by Wilcox et. al [42], monocular occlusions only become problematic when they are accidentally added, distorted or simulated for scaling purposes.

Distorted monocular occlusions can occur even when viewing images captured with a binocular camera rig. Depending on the acquisition and viewing conditions, S3D content may include monocular regions at locations that are not consistent with the scene geometry or scale. Wilcox & Allison [42] suggest that the phenomenon may be particularly apparent in scaled shots containing self occlusion. Self occlusion refers to the fact that parts of an object can be hidden by other parts of the object in one eye because of the angle of view as described above. An example of occlusion distortion occurs when an object, such as a cube, whose width is less that the observers' IO, is displayed directly in front of the viewer. When viewed at the targeted display size, the monocular occlusions will be consistent with the viewing geometry and the cube will not be distorted. However, if this S3D image is presented on a much larger display, the width of the object may exceed the IO. In this case the observer should not be able to see the monocular regions, but does. The visual system copes with the discrepancy either by distorting the object's shape or by alternately suppressing parts of one eye's image, a phenomena known as binocular rivalry [43].

Techniques that scale IA by generating new views (section 3.2 must pay close attention to monocular occlusion phenomena. This is because regions that were previously occluded may be visible in the new viewpoints and this content must be reconstructed. If this content is not available in other views or in a 3D model then it must be synthesized (in-painted). While this is an issue for all methods in this class of techniques it is most pronounced for 2D plus depth and 2D-3D conversion. The main disadvantage of the 2D plus depth and the 2D-3D conversion techniques is that depths are only assigned to pixels and thus parts of the image that are visible from the camera viewpoint. The description is therefore incomplete and phenomena such as monocular occlusion (see section 4.2) and transparencies are not represented. When a new view is synthesized some previously occluded features may be uncovered that are not in the image or the depth map producing 'holes' in the synthesized image. Including occlusion information as well as image and depth data can help, but there is a tradeoff between the completeness of such information and the compactness of the representation [24, 27].

5. Conclusion

With the emergence of new widely accessible S3D display technologies, content scaling has become an increasingly important issue. Displaying unmodified S3D content on vastly different screen sizes, for instance from mobile devices to 3D movie theatres, can negatively impact the S3D experience. Perhaps more importantly, some of the scaling artefacts may produce discomfort and fatigue. In this paper, we reviewed the human factors issues that may occur when S3D content is not scaled appropriately for the viewing device. As outlined here, the ideal is to match the shooting variables to specific viewing conditions. This strategy is not viable (for economic and artistic reasons), therefore a range of strategies are applied during content production, during post-production or at the display itself. The combination of good pre-production planning (eg. 'protecting' content for target screen sizes) and careful application of post-production techniques can go a long way towards ensuring the suitability of the resulting content for different display sizes. This is particularly true if the known sensitivities of the human visual system to particular artefacts are taken into account.

6. Acknowledgments

This work was supported by Ontario Centre of Excellence (OCE) and Ontario Media Development Corporation (OMDC) within the 3D Film Innovation Consortium (3DFLIC).

7. References

1. *SMPTE*. Available from: www.smpte.org.
2. *THX*. Available from: www.thx.com.
3. Spottiswoode, R., N. Spottiswoode, and C. Smith, *Basic principles of the three-dimensional film.* SMPTE Journal, 1952. **59**: p. 249-286.
4. Woods, A., T. Docherty, and R. Koch, *Image Distortions in Stereoscopic Video Systems.* SPIE, 1993. **1915**: p. 36-48.
5. Westheimer, G. and S.P. McKee, *Steroscopic acuity for moving retinal images.* J Opt Soc Am, 1978. **68**(4): p. 450-455.
6. Steinman, R.M., W.B. Cushman, and A.J. Martins, *The precision of gaze.* HUM NEUROBIOL, 1982. **1**: p. 97-109.
7. Howard, H.J., *A test for the judgment of distance.* AM J PSYCHOL, 1919. **2**: p. 656-75.
8. Bercovitz, J., *Image-side perspective and stereoscopy.* SPIE, 1998. **3295**: p. 288-298.
9. Kang, S.B., et al., *A multibaseline stereo system with active illumination and real-time image acquisition.* Proceedings of the Fifth International Conference on Computer Vision (ICCV '95), 1995: p. 88-93.

10. Collins, R.T., O. Amid, and T. Kanade, *An active camera system for acquiring multi-view video* Proceedings of the 2002 International Conference on Image Processing (ICIP '02), 2002: p. 517-520.
11. *3D Omnicam*. Available from: http://www.hhi.fraunhofer.de/.
12. Devernay, F. and P. Beardsley, *Stereoscopic Cinema.* Image and Geometry Processing for 3-D Cinematography, 2010. **5**: p. 11-51.
13. Zitnick, C.L., et al., *High-quality video view interpolation using a layered representation.* ACM T GRAPHIC, 2004. **23**(3): p. 600-608.
14. Chen, S.E. and L. Williams, *View interpolation for image synthesis.* COMP GRAPH, 1993: p. 279-288.
15. Doyen, D., L. Blondé, and T. Borel, *The use of a dense disparity map to enhance Quality Of Experience in stereoscopic 3D content.* SMPTE 3D, 2011.
16. Kauff, P., et al., *Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability.* Signal Processing: Image Communication, 2007. **22**(2): p. 217-234.
17. Zinger, S., L. Do, and P.H.N.d. With, *Recent developments in free-viewpoint interpolation for 3DTV.* 3D Research, 2012. **3**(1): p. 1-6.
18. Didyk, P., et al., *Adaptive Image-space Stereo View Synthesis.* Vision, Modeling, and Visualization, 2010: p. 299-306.
19. Scharstein, D., *Stereo Vision for View Synthesis.* Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'96), 1996: p. 852-858.
20. Rhee, S.-M., J. Choi, and S.-M. Choi, *Accurate Stereo View Synthesis for an Autostereoscopic 3D Display* IEEE ICCE, 2010: p. 375 - 376
21. Gurram, P., et al., *3D Scene Reconstruction Through a Fusion of Passive Video and Lidar Imagery.* Proceedings of the Applied Imagery Pattern Recognition Workshop, 2007: p. 133-138.
22. Yang, J.C., et al., *A real-time distributed light field camera.* EGRW '02 Proceedings of the 13th Eurographics workshop on Rendering, 2002: p. 77-85.
23. Lang, M., et al., *Nonlinear Disparity Mapping for Stereoscopic 3D.* COMP GRAPH, 2010. **29**(4): p. 75:1 - 75:10.
24. Devernay, F. and S. Duchéne, *New View Synthesis for Stereo Cinema by Hybrid Disparity Remapping.* IEEE IMAGE PROC, 2010: p. 5-8.
25. Vetro, A., S. Yea, and A. Smolic, *Toward a 3D video format for auto-stereoscopic displays.* P SOC PHOTO-OPT INS, 2008. **7073**: p. 70730F.1-70730F.10.
26. Coll, B., et al., *3D TV at home: Status, challenges and solutions for delivering a high quality experience.* Int. Workshop Video Processing and Quality Metrics for Consumer Electronics, 2010.
27. Fei, W. and L. Wei, *The overview of 2D to 3D conversion system.* CAIDCD, 2010. **2**: p. 1388 - 1392.
28. Matsumoto, Y., et al., *Conversion system of monocular image sequence to stereo using motion parallax.* SPIE, 1997. **3012**: p. 108-115.
29. Chen, W.-Y. and L.-G. Chen, *Fast Algorithm and Hardware Architecture for Stereo Image Synthesis Systems.* National Taiwan University, 2006. **1**.
30. Harman, P., et al., *Rapid 2d to 3d conversion.* P SOC PHOTO-OPT INS, 2002. **4660**: p. 78-86.
31. Battiato, S., et al., *Depth map generation by image classification.* P SOC PHOTO-OPT INS, 2002. **5302**: p. 95-104.
32. Tsai, Y.-M., Y.-L. Chang, and L.-G. Chen, *Block-based vanishing line and vanishing point detection for 3d scene reconstruction.* ISP ACS, 2006: p. 586-589.
33. Choi, C.-H., B.-H. Kwon, and M.-R. Choi, *A real-time field sequential stereoscopic image converter.* IEEE T CONSUM ELECTR, 2004. **50**(3): p. 903-910.
34. Jung, Y.J., et al., *A novel 2D-to-3D conversion technique based on relative height depth cue.* Proc. SPIE, 2009. **7237**: p. 72371U:1 - 72371U:8.
35. Ideses, I. and L. Yaroslavsky, *A method for generating 3D video from a single video stream.* VMV 2002 2002: p. 435-438.
36. Tran, L.C., et al., *On Consistent Inter-view Synthesis For Autostereoscopic Displays.* 3D Research, 2012. **3**(1): p. 1-10.
37. Hoffman, D.M., et al., *Vergence–accommodation conflicts hinder visual performance and cause visual fatigue.* J VISION, 2008. **8**(3): p. 1-30.
38. Yano, S., M. Emoto, and T. Mitsuhashi, *Two factors in visual fatigue caused by stereoscopic HDTV images.* Displays, 2004. **25**: p. 141-150.
39. Ono, H., P. Muter, and L. Mitson, *Size-distance paradox with accommodative micropsia.* ATTEN PERCEPT PSYCHO, 1974. **15**(2): p. 301-307.
40. Howard, I.P. and B.J. Rogers, *Seeing in Depth* 2002, Thornhill, Ont.: I. Porteous.
41. Rogers, B.J., *Charles Wheatstone and the cardboard cut-out phenomenon.* Perception, 2002. **31**: p. ECVP Abstract Supplement.
42. Wilcox, L.M., I. Tsirilin, and R.S. Allison, *Sensitivity to monocular occlusions in stereoscopic imagery: Implications for S3D content creation, distribution and exhibition.* SMPTE 3D Conference, 2010.
43. Ooi, T.L. and Z.J. He, *Binocular rivalry and surface-boundary processing.* Perception, 2006. **35**(5): p. 581-603.