OCTOBER 2010

MOTION IMAGING JOURNAL

INSIDE THIS ISSUE

A New Single Camera System for Stereoscopic Image Acquisition

By Zoran Perisic

Three key elements determine the ability of a stereoscopic system to capture a realistic three-dimensional (3D) image: wide angle of view, interocular distance, and convergence. Z3D achieves all three of these in a compact, format-independent unit that works with both film and digital cameras. A single camera and single lens system that includes an optical 3D viewfinder and a video 3D viewfinder which take the guesswork out of setting the convergence. It is compatible with existing 3D projection methods but can also be used for live 3D video projection.

3D IS EASY—GOOD 3D IS A BIT HARDER

We see the world about us from two viewpoints, separated by 2.5 in. on average. We get depth cues from the subtle disparity between the images produced on the retina of each eye from the corresponding viewpoint. So when our eyes are presented with two separate photographic images, we try to fuse them into one instinctively. Any amount of lateral displacement between the two viewpoints acts as a depth cue. Even when identical images are presented to each eye separately, we try to see depth in the scene by creating a pseudo three-dimensional (3D) illusion **(Fig. 1)**.

Figure 1. Frames from *Dracula Returns* Z3D demo; 35mm film. (Directed and Photographed by Z.P.)

In that respect, 3D is easy. However, it is much harder to achieve true stereoscopic reproduction with realistic depth perspective and comfortable viewing. Ideally, a 3D imaging system should be able to emulate the properties and functions of the human eyes in three specific areas: angle of view, interocular distance, and convergence.

When a scene is shot with an incorrect relationship among these three key elements, the resulting image does not appear realistic and is often uncomfortable to watch. Achieving the correct perspective is particularly important in a dramatic setting, in which the spatial relationship between the actors and the background needs to be maintained to enable smooth intercutting between one setup and another. Deliberate variations of these key elements can be used effectively in the right context to produce a desired effect, but the starting point should always be to reproduce true perspective.

Wide Angle of View

Ideally, the horizontal angle of view should correspond to the angle of view of the human eyes in the stereo zone—that is, where the monocular fields of the left and right eyes overlap (**Fig. 2**). However, it is advisable to take into account the size of the projection screen, as well as the size and shape of the presentation venue, when choosing a lens angle for a 3D shoot.

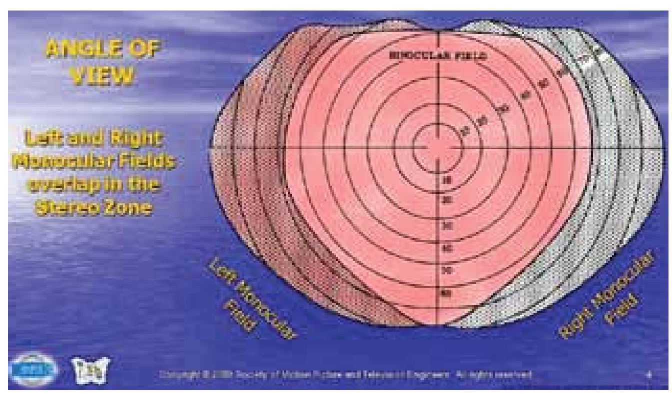

Figure 2. Binocular field and the Stereo Zone.

A narrow angle of view results in a 3D image with compressed depth, making the objects within a scene appear as flat cutouts, such as when viewed through binoculars. Extreme wide angle of view creates an egg-shaped distortion.

Interocular (Interpupillary) Distance

The average separation between the eyes in adults is 66mm. The distance between the two viewpoints of the stereoscopic imaging system should ideally match this interocular distance if it is to capture the correct relationship between the two images.

A wider-than-normal interocular distance results in a dwarfing effect: The objects appear smaller than in real life. A shorter-than-normal interocular distance makes the objects appear larger than life.

Convergence

Convergence is the most difficult aspect of 3D. One way of dealing with this is to ignore it and adopt the "parallel approach," where the optical axes of the two viewpoints run parallel to each other. The resultant images do not overlap across the full width of the frame, leaving an area outside the stereo zone, on each side of the frame, that is equal to the interocular distance used. This can be accommodated in post-production by cropping the sides of the frame; additional cropping at the top and bottom of the frame may be required to maintain the aspect ratio.

The two images produced by the parallel approach have a common center in the stereo zone. Axial lines traced from the original viewpoints to this new, common center show that they converge at an angle. Therefore, in practice, even the parallel approach results in convergence. This convergence is fixed but can be modified by repositioning the images relative to each other in post-production, with the inevitable loss of image area and the corresponding reduction in the effective angle of view.

The fixed convergence derived from the parallel approach produces satisfactory results and has been used by several 3D systems. However, a good case can be made for having a variable convergence (**Fig. 3**).

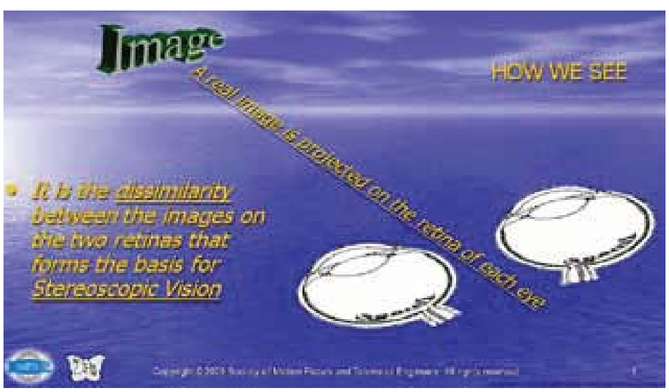

Figure 3. A real image of the scene is projected on the retina of each eye.

When we look at a scene, a real image of that scene is projected on the retina of each eye; those images are similar but not identical, because they are seen from two viewpoints. It is that subtle difference between the images on the corresponding area of the two retinas that forms the basis for stereoscopic vision.

Two important elements are associated with the retina, in addition to the blind spot. These are the macula, an area of high resolution between 2.5 and 3mm in diameter, and the fovea, an area of the highest visual acuity (resolution) that is only 1.5mm in diameter and is situated at the center of the macula.

The visual line of the eye is centered on the fovea and runs at an angle to the optical axis of the eye, diverging toward the nasal side (**Fig. 4**). Consequently, the left and right visual lines converge even when the eyes are facing forward. They converge even more when we are looking at something close. We see detail in the close to mid-distance range, and this is precisely where stereoscopic perception is most significant. It enables us to judge distance from an object. There is little or no perception of depth in the long distance beyond 25 ft.

Figure 4. Visual line of the eye is centered on the fovea and runs at an angle to the optical axis of the eye.

The eye muscles keep the visual lines of both eyes aligned on the center of what we are looking at. This gives us the ability to scan across a scene and zero in on a specific detail—a fixation point. We often think of this as zooming in on a detail, knowing full well that the eyes cannot zoom. While the visual lines of the eyes converge at the fixation point, the optical axes of the eyes inevitably converge at a point beyond the fixation point or may be parallel.

Convergence is the most difficult aspect of 3D. One way of dealing with this is to ignore it and adopt the "parallel approach," where the optical axes of the two viewpoints run parallel to each other.

THE "WINDOW" AND THE STEREO ZONE

The purpose of using convergence in 3D cinematography is not really to draw attention to a specific part of the frame (as is sometimes suggested) but to enable the eyes to zero in on any object within that frame with ease. By changing the convergence between the two viewpoints, we are effectively altering the shape of the stereo zone, which can make it easier to fuse the images and accommodate closer fixation points. When a stereoscopic image is correctly photographed and projected, the viewer should be able

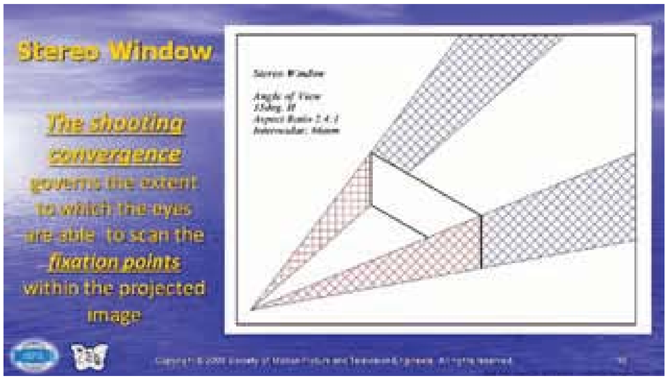

Figure 5. Parts of a scene occupy the stereo space in front of the "window" and the rest behind the window.

Figure 6. Z3D has a wide angle of view, correct interocular distance, and easy convergence control.

to "scan the projected mage" and zero in on any fixation point as when viewing a real scene. The big difference is that the viewer is seeing all this through a "window," with some parts of the scene occupying the stereo space in front of this window and the rest behind the window. The stereo window is not necessarily in the same vertical plane as the projection screen. However, it coincides with the convergence plane (**Fig. 5**).

The comfort factor in 3D viewing is determined by a fine balance between the angle at which the viewpoints of a stereo pair of images are recorded (the shooting convergence) and the extent to which the eyes are able to accommodate the scanning of the image by shifting from one fixation point to another within the projected image (the viewing convergence).

The convergence can also be altered in projection by increasing or decreasing the amount of overlap between the stereo pair of images on the screen with appropriate equipment. In practice, this is usually done in a digital intermediate (DI) suite. The final overlap position is locked when the material is transferred to a digital cinema package (DCP) for projection in RealD or similar systems.

The correct reproduction of stereoscopic depth is affected not only by the horizontal displacement of an image on the corresponding area of the retina but also by the size of that image on the retina. This is where the size of screen and the viewing distance play a major part. A well-executed and presented 3D scene places the viewer at the very spot where the picture was taken.

Z3D System

Z3D covers all of the preceding elements: wide angle of view, correct interocular distance, and easy convergence control (**Fig. 6**). It is a compact, format-independent system, employing only *one* camera and *one* lens. The camera can be either film or digital and either movie or still.

Left and right views of a scene are redirected toward each other by two front-surface mirrors that swivel in unison; two small, front-surface mirrors, placed at the intersection of these two beams, rotate each image through 90° as they redirect them toward the lens. The two views are recorded as two separate images occupying the top and bottom areas of the same frame in a toe-to-toe or head-to-head configuration.

An optical 3D viewfinder essentially repeats the imaging process in reverse—the left and right image on the ground glass of the camera is redirected to the corresponding left and right eyepiece. This allows the camera operator to see the image in 3D during photography and manipulate the convergence by simply rotating a knob at the side of the Z3D unit. The convergence can be preset or altered on the fly in much the same way as when pulling focus.

A 3D video viewfinder provides the same facility for simultaneous 3D viewing by other crew members, as well as for playback. It can also be used as the primary 3D viewfinder with digital cameras that do not have an optical viewfinder.

The composite frame containing both left and right images is displayed on a small monitor. This may be from a video-assist camera imaging the ground glass of the film camera or a direct video feed from a digital camera. An arrangement of mirrors enables each eyepiece to view only the corresponding left or right image. The viewer can also be placed over the flip-out monitor of a digital camcorder, enabling direct viewing in 3D.

SINGLE CAMERA ADVANTAGE

Most 3D systems available are rigs, with two cameras mounted side by side or at right angles to each other with a beam splitter mirror placed between the two lenses. The advantages of using a single camera system are self-evident (**Fig. 7**).

Z3D EVOLUTION

Z3D started out as an attachment supported on brackets in front of the lens. The intention was to enable the use of different lenses with the same Z3D unit. That way, only a small adjustment was required to accommodate the changes in the angle of view for different focal lengths. Prototype units were used in this way for a demo shoot with 35-mm film cameras using Nikon lenses.

Figure 7. The advantages of using a single camera system are self-evident. (Arri 2C Camera with Z3D unit and optical stereo viewfinder).

Figure 8. Z3D in-line configuration on a Mitchell camera. A relay lens transfers the aerial image formed by the Z3D unit to the camera.

After the demo shoot, it became apparent that the best solution was to integrate the lens and the Z3D unit into one self-contained optical unit that can be used on any camera in much the same way as a regular lens. This means having a separate unit for each focal length.

Other improvements followed, such as a custom-designed matte box. The demo shoot was hampered by the lack of a proper matte box. An integrated unit was later used for a test with an Arri D21 digital camera.

Configurations

There are two basic configurations of Z3D. One involves placing the integrated Z3D lens unit directly onto the camera, which is mounted vertically on the camera head. The optical viewfinder is situated behind the Z3D unit and in line with the direction of view. A good way of checking the 3D is to look through the 3D viewfinder and then look just above the unit at the scene in front, remembering to keep both eyes open.

An alternative method uses a relay lens to transfer the image from the Z3D unit to the camera, resulting in a more conventional *in-line configuration* (**Fig. 8**). This provides some advantages in that a lens of a symmetrical design can be used as the primary imaging lens. These lenses are particularly effective in capturing distortion-free perspective but cannot be used with conventional movie cameras because of their short back-focus distance.

Another advantage is that with lenses of symmetrical design it becomes easier to construct Z3D units with shorter interocular distances for use in specific applications where the object needs to appear larger than life. Units with longer-than-standard interocular distances are required when the objects needs to appear smaller than life. These are easier to construct.

The standard relay has a 1:1 image transfer ratio—but this approach also provides an option for magnification or reduction of the image size at the camera side to accommodate almost any format. Putting it simply, the images created by a Z3D unit can be transferred optically to the camera aperture using appropriate magnification for the required format. In this way, an integrated Z3D unit can be used on virtually any format, including IMAX.

The challenge is to construct the imaging lens as an integral part of the relay system.

FILM FORMATS

In the preferred configuration, the Z3D system splits the frame along the horizontal centerline, with left and right images lying along the split line. The left and right eye images are in either a head-to-head or a toe-to-toe relationship to each other, depending on whether the camera is pointing up or down relative to the direction of view. The deviding line is always parallel to the horizontal centerlines of the two images.

The horizontal angle of view is the determining factor in 3D. There is no 3D in the vertical. This is where a 2.4:1 aspect ratio has an advantage. It is in fact the scope format, and one eye can be transferred to 4 perforations with anamorphic squeeze for two-dimensional (2D) presentation.

On 35-mm film, with standard 4-perforation pull-down, each image occupies 2 perforations in height, as in the Techniscope format (**Fig. 9**). When the Z3D image is placed in the center of the full aperture, it covers an additional area on both sides of the image which can be used for altering the relationship between the left and the right eye images in post-production, if required, without sacrificing any of the frame area.

Figure 9. On standard 4 perf frame, each image occupies 2 perforations in height. Frames from *Dracula Returns* Z3D demo. (Directed and photographed by Z.P.)

Figure 10. 8-perforation VistaVision format. Mojave Desert and Las Vegas Strip at night.

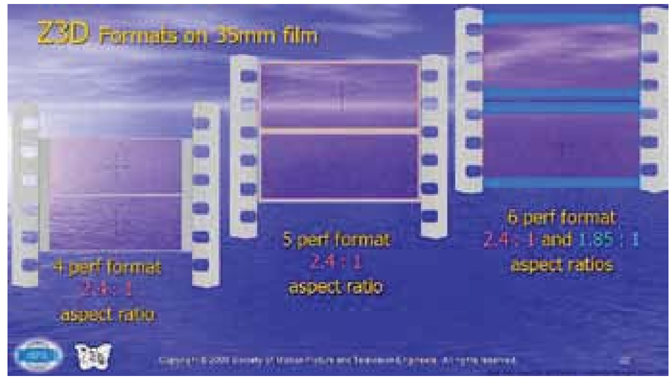

Figure 11. Possible film formats.

By splitting the frame along the horizontal centerline, the horizontal angle of view of the chosen aspect ratio is preserved—and with it the horizontal resolution of the image. Two images with a 2.4:1 aspect ratio can be accommodated on a 35mm full-aperture frame. On the other hand, an 8-perforation VistaVision frame can accommodate two full 4-perforation frames.

The image circle of a standard photographic lens covers up to 8-perforation VistaVision format because those lenses were designed for that format (**Fig. 10**).

Nonstandard 5 or 6-frame pull-downs offer interesting possibilities (**Fig. 11**).

DIGITAL FORMATS

With digital cameras, a squire sensor or one that matches the film format, like Arri D21, is most suitable. The height of the chip determines the effective width that will be used. Consequently, a camera with a 16:9 chip and 4K resolution can easily accommodate a pair of images in a 2.4:1 aspect ratio of 2048 × 858 pixels to produce a 2K DCP. (**Fig. 12**)

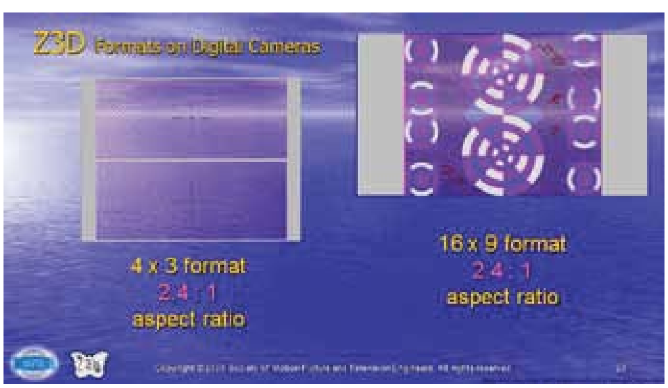

Figure 12. Possible digital formats.

The first step in post-production of Z3D-filmed material is to telecine the negative, full frame, (with time codes) for dailies and editing.

Half-Frame Split of the Digital Signal

Digital cameras scan a moving image *progressively* in much the same way as film cameras do, but the image is split horizontally into odd and even lines, which are stored sequentially as two interlaced components of the image—field A and field B.

However, if a moving image acquired by progressive scanning could be divided along the horizontal centerline so that all the lines in the top half of the frame form field A and all the lines in the bottom half of the frame form field B, the two fields could be stored *sequentially* as usual.

Because the Z3D system splits the frame in half along the horizontal line, it is desirable to store the progressively scanned digital images in this manner; effectively fields A and B become left and right eye images, respectively.

2-Chip Digital Camera

For maximum utilization of the 16 × 9 format, a 2-chip solution could be adopted. The image circle of a Z3D lens covers two sensors placed side by side—along the longer side of the rectangle. In this way, all advantages of Z3D image acquisition are preserved at the maximum resolution of the sensor. Data could be stored separately, such as when two digital cameras are used in a 2-camera rig.

Z3D Post-Production

The first step in post-production of Z3D-filmed material is to telecine the negative, full frame, (with time codes) for dailies and editing. Editing and post-production can be completed without separating the left and right eye images. Projection and viewing facilities enable the editor to see the image in 3D during the post-production. Visual effects can also be completed in this way. Essentially, the material can remain as one single stream during the entire post-production process.

Next, pull an edit list with time codes for scanning. When the picture is finalized, only the selected takes need be scanned at 4K and assembled for DI post-production.

The following FotoKem procedure is used for the production of a DCP from a Z3D full-frame format for the Z3D demo:

- Scan the 35 mm film negative full frame at 4K as per the edit list.
- Crop the sound-track area (or each side if shot on full-aperture center).
- In the DI suite, import full frame as one stream, split it along center line and cut off the bottom half. Import the full frame again as a second stream, rotate 180°, split the frame along center line and cut off the bottom half. The two halves of the original full frame now form two separate streams (left and right eye).
- Also in the DI suite, make color timing on one stream and simply apply the same setting it to the other stream as the images are identical in terms of color.
- Reduce to a 2048 × 858 format (per each eye) to produce a 2K DCP. (**Fig. 13**)

Figure 13. A Scaled up Z3D unit mounted on a digital projector provides for live 3D viewing. Handheld 3D viewer with monitor.

The following are Z3D presentation methods:

- Digital cinema projection in RealD and other theatrical 3D projection systems.
- Live or prerecorded digital video projection (polarized separation) using a scaled-up projection version of a Z3D unit in front of the projection lens.
- Film projection (polarized separation) using a scaled-up projection version of a Z3D unit in front of the projection lens.
- 3D viewing without glasses as a single-viewer autostereoscopic system.
- Handheld 3D viewer with a mini digital video player.
- 3D television.

CONCLUSION

A single camera alternative to the bulky dual camera rigs, Z3D is a compact, format-independent stereoscopic system for use with film or digital cameras. Designed to capture three-dimensional 3D images with a true perspective, it includes stereoscopic viewfinders (optical and/or video monitor based) that make it possible to adjust the convergence with confidence in the time honored approach of "what you see is what you got!"

BIBLIOGRAPHY

Cox, Arthur, *Photographic Optics*.

Lipton, Lenny, *Foundations of Stereoscopic Cinema*, Van Nostrand Reinhold: New York, 1982.

Optical Design, Military Standardization Handbook.

Presented at the SMPTE 2009 Annual Tech Conference & Expo, October 27-29, 2009.

Zoran Perisic won an Oscar and a BAFTA award for his work on *Superman—the Movie*. He is also the recipient of a Scientific and Technical Achievement Award from the Academy for the development of the Zoptic Front-Projection System—the principal method used to make Superman fly in *Superman 1,2, & 3*. He has worked as a visual effects supervisor or consultant on numerous productions and has directed feature films and TV programs. He is a member of the Academy, author of *Visual Effects Cinematography* and other books and has several patents to his credit, the latest of which is the Z3D System for Stereoscopic Cinematography.