How to Shoot Good 3D with The Stereocam 3D Camcorder Adapter


By

How to Shoot Good 3D With The Stereocam 3D Camcorder Adapter

Introduction

The Stereocam 3D Camcorder Adapter is a "break-through" product. As an attachment to camcorders, not only does it provide an easy, low-cost, way to record "true", stereoscopic, three-dimensional (3D) video, but also the user has access to all the facilities built into his camera: sound, zoom, auto focus, auto exposure, special effects, etc. All these features, including digital video, can be used with **the Stereocam**, but cannot be found in other 3D camera systems, even very expensive ones. Present-day camcorders are an amazing product, and they will only get better with the "digital revolution". **The Stereocam** gives us the opportunity to take advantage of these developments for making 3D videos. In addition, **the Stereocam** provides another important feature, one that can only be found on very expensive 3D camera systems: dynamic convergence adjustment. You will learn in this "app note" that this is a critical adjustment for good 3D video. We will also discuss other basic principles of 3D videography, such as the "Stereo Window" and composition, how to do "out-the-window 3D", and some "more technical" topics (vertical convergence, vignetting, near-field effects, and motion effects). Let's start with a review of how the **Stereocam** works.

How Does the Stereocam 3D Adapter Work?

Look into a **Stereocam** from the "big" end – where light enters the adapter. You will see two liquid crystal shutters (LCSs), one straight-on, one through the mirror. Think of these as two view ports, one for the left-eye view and the other, the right-eye view. They are positioned 2.2 inches (56 mm) apart, slightly less than the average human interpupillary distance (IPD) of 2.6 inches (66 mm). This distance forms the "stereo base" of our 3D system. Now look through the "small" end – where **the Stereocam** is connected to the camera. You will see two images, from these two viewports, superimposed on each other. If you now turn the Convergence Adjustment Knob (see the diagram on the Quick Start Card of your **Stereocam**), the two views move horizontally relative to each other. You can make the two views line up for distant objects or for near objects, but never for both at the same time. This is called *parallax*.
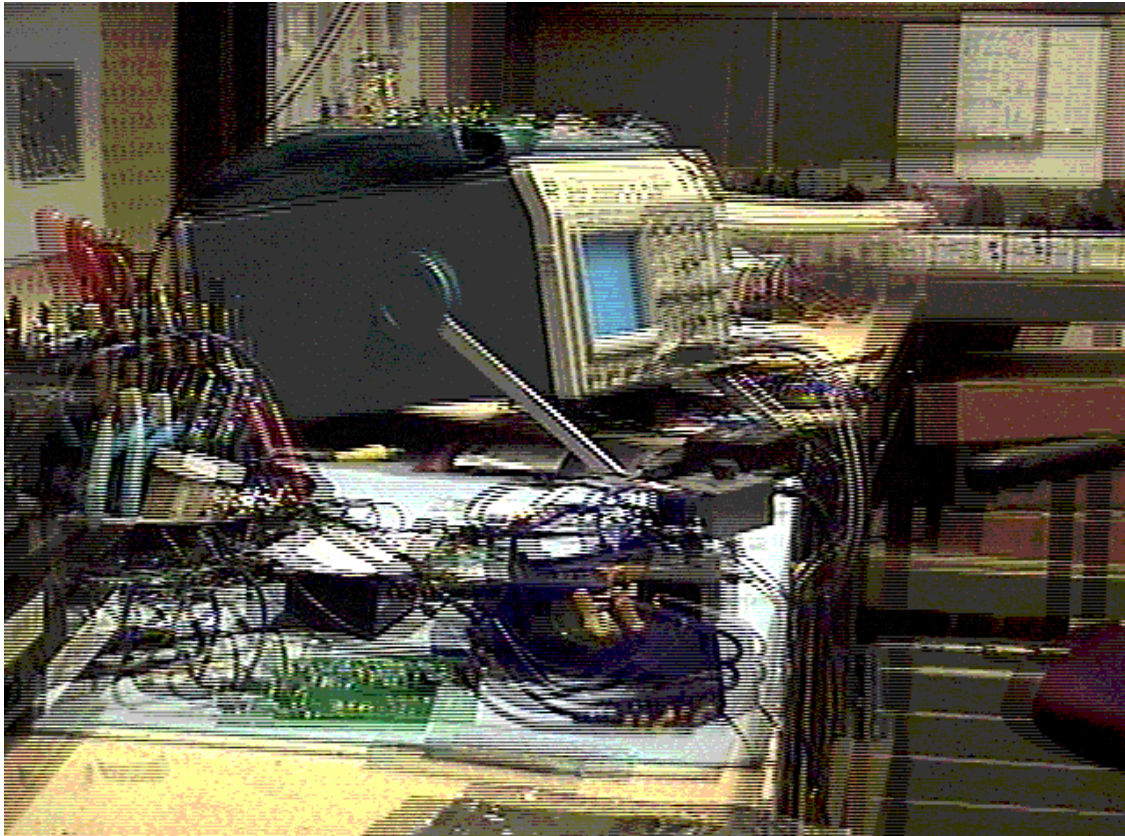
[parallax illustration]

After you mount the **Stereocam** on your camcorder, you will see the same double image view in the camera's viewfinder. The place where objects line up is called the *convergence point* or *convergence distance*. The distance between object points that do not line up is called *disparity*. Note that it is easy to monitor the convergence point and the disparity in the viewfinder. These important concepts are discussed fully below. The next step in the 3D recording process is to store these two images separately. We do this on a single camera using the "time-sequential" or "time-multiplexed" method. The LCSs act as light valves; one is open when the other is closed, and vice-versa. At any one instant of time, only one image, left or right is recorded by the camera. The switching is driven by the video signal (via the Video Signal Cable, page 6) so that the left image is recorded during the "odd" video field and the right one, during the "even" field. (Thus, this format is also called "field-sequential".)

[blowup of parallax illustration to show black lines]

The Stereocam creates and stores two views that are very similar to what we see with our eyes. When we look at a natural scene, our eyes fixate on objects in the scene. That is, they converge on some object point and adjust their focus to that distance (accommodate). We continually do this as we look at different things. The two views are different; they have parallax because of the IPD between our eyes. This parallax causes disparity in the two images on our retinas, just the right amount for our brain to "fuse" the two images into a three-dimensional, solid view of the world. This process is called "stereopsis". ("Stereo" comes from the Greek word for "solid".)

How Do We View the 3D Images?

Now that we have recorded the left-eye and right-eye views, they need to be presented to our left and right eyes separately. This is not easy, but there are various ways to do this. They range from personal head-mounted devices to monitors and large-screen projectors. The most direct method is for each eye to look directly at its own display, for example, in a head-mounted display. A second way is to use passive polarization. In this method, the two images are projected using dual projectors fitted with crossed polarizers; the viewer wears polarized glasses. A third method is called "auto stereoscopic" displays. This means that the viewer does not need to wear anything; optical devices are used to ensure that the left eye sees only the left image, and the right eye, the right image. (Stereo 3D displays will be summarized in other App Notes.) The most common method for 3D viewing is a CRT monitor or TV set that displays the left/right images time-sequentially. The two images are presented one after the other, and our eyes are shuttered in some way so that the left eye sees only the left images, and the right, only the right. If the sequence runs fast enough, each eye sees its own continuous image. The shuttering is done by either "active" glasses worn by the user (each eye is "opened" sequentially) or by a

switching polarizer screen on the monitor (in which case the viewer wears "passive" glasses). As you may have guessed, time-sequential 3D displays are very compatible with time-sequential 3D video recording, but other types of 3D displays can also be used with this format. For the present discussion, we do not particularly care what type of method is used except that both images, left and right, are presented to the viewer on the same image plane. This alone will dictate certain 3D imaging principles that we should keep in mind.

Accommodation, Convergence, and Disparity

Although the objective is to make sure that each eye sees the same thing it would see in nature, *no display device*, whether 2D or 3D, duplicates the way our eyes really work. In a 2D display, both eyes are looking at the same, single, image instead of the two parallax views. In addition, in most images, the whole scene is in focus at the same time. This is not the way our eyes work in nature, but it has to be this way so that we can look wherever we want on the display surface. In reality, only a very small, central, part of our field of view (the fovea) is in sharp focus, and then only at the fixation distance. Our eyes continually change focus, or accommodate, as we look at near and far objects. However, when viewing a (flat) image, all the objects are in focus at the same time. In a sense, the first time man viewed a painted image, he had to learn a new way of seeing, especially images which try to portray depth. In stereoscopic 3D displays, our eyes are now each given their proper parallax view, but they still must deal with the fact that both images are, in reality, *flat*! The two images are superimposed on some plane at a fixed distance from the viewer, and this is where he or she must focus to see the images clearly. As in real nature, our eyes roam around the scene on the monitor and fixate on certain objects or object points. Now, however, our eyes are converging at one distance and focusing at another. There is a "mismatch" between ocular convergence and accommodation.

For example, suppose that the Stereocam is converged at 10 ft (i.e., the *camera* convergence distance), and a near object, "A", is 5 ft away and a far object, "B", is at 15 ft. Objects at the convergence distance do not have any disparity and appear exactly overlaid on the screen. In the 3D space surrounding the monitor, they appear to reside *on* the screen surface. Object A, which appears in front of the screen, is said to have *negative* disparity, and object B, behind the screen, has *positive* disparity. In order to view object A, our eyes converge to a point that is in front of the screen; for object B, the convergence point is behind the screen. As in real nature, our eyes converge on the various objects in the scene, but they remain focused on the monitor screen. Thus we are learning a new way of "seeing" when we view stereo pairs of images. Some people may find it uncomfortable at first, but soon get used to it; most viewers have no problems at all. Usually difficulties in seeing 3D on stereoscopic displays are due to other factors, such as "ghosting" (i.e., each eye sees part of the other image, also called "crosstalk") or imbalances in image quality (e.g., brightness and color). When the two images match well and are seen distinctly - and separately - by the two eyes, it becomes easy to fuse objects, even if there is a large amount of horizontal disparity.

Guidelines for Adjusting Convergence and Disparity

There are no hard and fast rules about how much disparity to allow, and, in any case, it is very hard to monitor it precisely as you are shooting a 3D video. Different objects – at different depths – will have differing amounts of disparity. In addition, the size of the screen and the distance the viewers are from the screen affect how easy it is to fuse a 3D scene. As you increase screen size, disparity increases, but, as you increase the viewing distance, the mismatch between accommodation and ocular convergence decreases.
In general, the best rule of thumb is to keep the maximum disparity in a scene to a minimum. With the convergence control, **the Stereocam** allows you to adjust the disparity for what is best for a particular scene. As mentioned above, disparity and convergence is easily monitored in the viewfinder of your camera as you shoot 3D with a **Stereocam**
.

When first learning how to shoot 3D videos, here are some guidelines to follow:

1. **Set the convergence for about 5 or 6 feet, via the Convergence Adjustment Knob and keep all objects in the scene at a farther distance. This will place the closest objects at the surface of the monitor and the more distant objects "inside" the monitor. (See The Stereo Window, below). In the beginning, do not change convergence as you are shooting.**

2. **Do not attempt close-up shots, meaning closer than 5 or 6 feet. Objects closer than this may have a lot of disparity (and other problems see *Near-Field Effects* below).**

3. **Do not attempt high zoom shots. For a fixed setting of convergence, disparity increases as you zoom. Use a wide-angle lens setting when you are beginning and then move to medium zoom shots later on.**

Later, as you get more practiced, try converging on your object of interest, especially if you zoom in. This will be "bring" the object to the surface of the monitor, and it will be easier to see. You will find that this "dynamic" convergence control gives you another way to direct your viewers' attention to what you want them to look at. Take care whenever you do this that objects closer to you than the main object will be in front of the screen and may be hard to see clearly. This starts us thinking about the "Stereo Window".

The Stereo Window

In 3D photography or videography, the display device changes from a display *surface* to a display *volume*; the photograph or monitor becomes a window or opening into a 3D space, a scaled representation of the true physical space. The display volume extends both into the display and in front of it. But the space in front of the monitor must be used judiciously, for two reasons. First, consider two objects that are the same distance from the convergence point; say object "A" is in front and "B" is behind. In the display space, objects at the convergence point will appear to be at the surface of the monitor, and objects A and B will appear equidistant in front of and behind the screen. However, because of triangulation of the light rays, the disparity on the screen for A will be greater than that for B. Thus it will be more difficult to fuse object A and see it clearly.
A second concern about the space in front of the screen is that a 3D object that appears here is perceived to be in front the actual physical object that is providing the window, e.g., the edge of the opening. In order to see and understand this object, each eye must see

their respective view of the object, clearly and completely. When recording, it is easy to "cut off" near objects in one of the views. If part of the object is missing in one eye, it becomes very difficult to recognize and fuse that object. (It is OK to cut off objects that are behind the window. Since the object is imagined to be behind the edge, it is our normal expectation not to see all of it.)

This is why accomplished stereoscopic photo/videographers use the concept of the "Stereo Window". The edge of the display surface is considered to be a window or view port into the 3D display space; viewing the 3D scene is like looking through the window at the outside world. The position of the window in that space is defined by the camera's convergence distance. Thus, the next guideline for the 3D videographer is

4. **In general, compose your scene and adjust the convergence so that everything is "behind" or "inside" the Stereo Window. That is, all objects in the scene are behind the convergence point. Objects can be at the Window only if they are not cut off by an edge.**

"Out-the-Window" 3D

But, you might ask, how do I make things "pop out" of the screen?! Yes, objects in front of the convergence point will appear in front of the display surface, but *they will be very hard to see if they are cut off by the edge of the screen*. In addition, the impression that objects are coming out of the window is much more effective if some objects in the scene are *not* out-the-window. We sense the third dimension by comparing objects at different depths. If all the objects are at the same depth, e.g., in front of the monitor, the feeling of depth will not be so strong.

These ideas lead to the next two guidelines:

5. **Whenever you bring objects in front of the screen, for example, for a "pop-out" effect, make sure that they stay completely *within* the Stereo Window! That is, they are not cut off in either the left or right view.**

6. **When bringing objects in front of the Stereo Window, leave other things behind the Window.**

Here is a very effective test shot that will help you learn about the Stereo Window and how to make things pop out of it: Place someone about 5 or 6 feet in front of your camera (with the **Stereocam** attachment on it, of course!). Center your shot on his/her face and upper body and use moderate zoom and adjust convergence so that your subject fills the screen and is "at the Window" (i.e., zero disparity). Now ask the person to raise an arm and point it at the camera. When you view this on your 3D monitor, his or her arm will come eerily out of the monitor. Next, ask the person to move his/her arm to the left or right, enough so that it moves out of the camera view. You will see how the out-the-window effect is lost as soon as the arm is cut off in one of the views.

Good 3D Composition

All of the rules of good composition of 2D photo/videography go over to 3D, things like framing, using diagonals, dividing the view by thirds, leading the viewer's eyes to the

subject matter, slow pans, moving the camera with moving objects, zooming for head shots, etc. What additional things should we do now that we can record depth?

In many shots with **the Stereocam**, you will find that 3D just "comes naturally"; after all, *all* images have depth (at least, images of the real world!). Having the third dimension is just something that *should* be there. But there are times when you can use depth in a more deliberate way, to create an impression or to emphasize your subject material. One way, discussed above, is to bring things out of the window, but there are other techniques that are of more general usage. For example,

7. **Make sure that in most shots or scenes there are objects positioned at various distances from the camera: foreground, mid-distance, and background.**

8. **Frame a distant scene with near objects. (A "useless" 3D shot is a distant mountain range. While it may be a good 2D subject, objects that are very far away, even with zoom, will not have a stereoscopic effect. Put something in the foreground, and, all of a sudden, you have a very dramatic depth effect.)**

9. **Use a change in depth to lead the viewer's eyes from foreground to background. A very effective 3D shot is a path or road that goes off into the distance.**

"More Technical" Topics

As you get more proficient with **the Stereocam**, you may become interested in the "more technical" aspects and adjustments of the device and its operation.

Following is a discussion on a few of these topics.

Vertical Convergence and its Adjustment

In normal vision when we fixate on an object point, it is natural for our eyes to converge horizontally. Even though limited differential vertical movement is possible (up to several degrees depending on the individual), it is not easy for our eyes to move differentially in the vertical. Thus, one goal of any stereoscopic system is to eliminate, or, at least, minimize vertical disparity.

All 3D camera systems have vertical disparity. This is due to the fact that the two lines of sight (that correspond to the two viewpoints) are converged. A rectangular shape in one view will not be rectangular in the other view (this is called "keystone distortion"), and some object points will be displaced vertically in the two images (for example, the corners of a box). Though theoretically possible, it would be very difficult to build a 3D camera that eliminates vertical disparity at all object points in the 3D scene. So why do 3D camera systems work? Because the amount of vertical disparity is small enough for our eyes to be able to adjust.

Even though **Stereocam** is not meant to be a precision product, during the manufacturing process we use a tight specification on the allowed vertical misalignment: 4 minutes of arc. This number cannot be directly related to the vertical disparity that might appear on the viewing device. As discussed above, disparity is affected by the amount of zoom, size

of the screen, etc. When using high zoom, vertical disparity will be much more noticeable, even though the unit is within the manufacturing specification. Camcorders are also not precision devices. The relative positioning of their optical system to their housing can vary, even within the same model and manufacturer. This can also cause vertical misalignment. Another factor is that your **Stereocam** might become vertically misaligned (i.e., out of specification) through shipping, handling, bumping, etc. Therefore,

10. **Check the vertical disparity when you first install your Stereocam and before each taping session. Adjust it if necessary. (I.e., keep the Adjustment Screwdriver handy! CAUTION DO NOT REMOVE THE SCREW ALL OF THE WAY! This can cause the screw to be removed completely. It is recommended to tighten the screw all of the way and then slightly back it off as desired)**

Depending on the maximum zoom of your camera, 4 arc-minutes can appear as 1/2 inch disparity on a 21 inch diagonal TV set. It is rare that you will be shooting 3D at full zoom (for a start, you will need a very steady tripod). Since most shots will be at medium to wide zoom, it is perfectly acceptable to set vertical convergence using medium zoom. You will find the adjustment to be much easier to do. A good rule of thumb then is

11. **Adjust vertical convergence using the highest zoom that you intend to be shooting with.**

If you do shoot at full zoom, you will also notice some hysteresis in the Convergence Knob. Again, this is due to the sensitivity at high zoom and the fact that **Stereocam** is not a precision product. You will find, however, that it is not difficult to adjust the knob in such a way that you can still eliminate vertical disparity.

Do not forget, as you examine and adjust the vertical disparity in a scene, that it will never be eliminated everywhere in the field of view. Optimize it near the center of the camera's view. The more the two views are converged, that is, for closer convergence settings, the more vertical disparity varies across the scene.

Thus, the "worst-case scenario" for vertical disparity is near convergence at high zoom.

Vignetting

In optics, vignetting ("vin-YET-ting") is when the field of view is cut off on one or more edges because of the physical construction of the lens or housing. **The Stereocam** does cause this at wide zoom on some cameras. The reason is that **the Stereocam** was designed to support the widest possible range of camcorders, and they can vary substantially in their dimensions, especially in the placement of their lens system with respect to their housing. To eliminate the effect entirely would have meant a bigger, heavier unit and a more expensive product.

We feel that the vignetting seen in this version (which will be eliminated in higher-end versions currently on our drawing boards) is not a serious limitation because, where it occurs, it is only at the widest zoom settings. The solution is for the user to "watch for it" and to not use those widest settings. In other words:

12. **When using the widest zoom settings, look for vignetting in the viewfinder and increase zoom accordingly.**

Cameras vary in how well the viewfinder gives you "what you see is what you get". You will have to learn how to judge the degree of vignetting on your camera by doing test shots and playing them back on your display system. Displays also vary in how much they cut off the edges of the full image. Televisions sets are normally set to "overshoot", but computer monitors are not. All these factors affect whether your audience will see vignetting in the final production. You may find that you see vignetting in the viewfinder, but it is not present on the display used for your presentation.

Near-Field Effects

The optical design of **the Stereocam** does not allow for near-field 3D recording. There are several reasons.

One is that the two viewports (see **How Does the Stereocam 3D Camcorder Adapter Work?** above) are not actually positioned on the same depth plane. After going through the mirror, the side view is set back about 2 inches from the direct view. (You may have noticed this as you looked into the adapter from the "big" end.) This results in different object sizes in the two views ("size distortion"), which causes additional vertical disparity throughout the scene. But the effect is only noticeable for near objects. Our tests have shown that it is not serious for objects beyond 3 feet.

A second reason is that the convergence angle for near objects becomes relatively large. Therefore the keystone distortion is greater, and there is greater vertical disparity. Another consideration is that, unless you take special care, close-up views do not result in good 3D, even if vertical disparity is eliminated. Often the depth dimension will simply look "wrong", and objects will appear elongated.
Thus, our guideline for near-field shooting is simply "don't do it", or
13. **Do not attempt to shoot objects that are closer than about 3 feet to the camera.**
Of course, this is a subjective result; it really depends on the amount of vertical disparity or distortion your audience can tolerate. If you are careful about area of interest and convergence, it is possible to shoot selected near subjects.

Motion Effects

As discussed above, **the Stereocam** uses the field-sequential 3D video format. Each eye is updated at a rate of 30 times per second. One result is that motion is not as "smooth" as compared to the 60 Hz refresh of 2D video; however, it is generally not objectionable or even noticeable. Indeed, many computer-based video systems use the same update rate.
A second effect is a possible mismatch of the left and right-eye views. The two images are actually captured at different instants of time, separated by 1/60th of a second. (It is important to recognize that this is not a "delay". Each channel is a faithful temporal record of the scene.) If you were to examine a single image pair, you would see that a moving object appears at different places in the two views.

However, a video recording is a continual record of images of the moving object, and, in particular, the left and right views are two continual streams of images. After the

shuttering process, each eye sees its own sequence of images, and our brain interprets this as two continuous streams, not as individual stereo pairs. It senses the *motions* of the object in the two views, not its exact positions, and this is what it uses to recreate the sense of the third dimension.

Now, there can be rare instances when your brain does not get an accurate record of an object's motion and this can lead to strange 3D effects. For example, suppose you are panning along with a rapidly moving object and miscellaneous stationary foreground objects "flash" across the screen. They may be moving so fast relative to the camera that both left and right images are not even recorded. However, in normal 3D recording, these events will be transitory and unimportant.

Things to try

Finally, the best teacher is "experience". As you use your **Stereocam**, the suggestions given above will seem like second nature. To get you started here are some ideas for scenes that you might shoot that will "show off" the 3D effect:
- A head/shoulder shot of a person, with his/her arm towards the camera.
- A mountain range in the distance framed by near trees or bushes.
- A road or path leading away from the camera, surrounded by trees.
- A field of flowers, shot from low level.
- A soccer game, shot from somewhat above eye level and medium zoom.
- A child on a swing, coming out of the screen.
- A sunset at the beach, looking along the breaking waves.
- A tree or a bush or a flower, from some distance away with sufficient zoom to fill the screen with the object.

Have fun!