Digital Photography and Ortho-stereo

S. Spicer, June 2008 steve.spicer@gmail.com

With digital photography now firmly entrenched in the world of stereo, many people may not have yet seen digital ortho-stereo photography. This article describes what ortho-stereo imaging is, how to achieve it, and some possible visual display systems. Ortho-stereo is the most realistic way to view stereo images, as the original scene is reproduced at life-size scale and perspective.

Ortho-stereo

Three requirements must to be met to achieve ortho-stereo reproduction of images:- (1) correct perspective; (2) correct camera-lens spacing; and (3) image resolution on the retina that meets or exceed the resolving power of the human eye. [Item (3) this is not formally required for ortho-stereo, but is desirable to achieve as it further adds to the realism of the reproduce image.]

1. Perspective

To obtain correct perspective of the original scene, an image that is identical to the image created by the original scene (for the field of view of the photograph) has to be reproduced on the retinas of the person viewing the photograph. This requires each object in the photo to subtend the same solid-angle as the objects in the original scene did. For example – imagine a scene in which there is a football on the ground nearby, and some goal-posts in the distance. This football creates an image on the retina that is perhaps 0.5mm in diameter, while the images of the goal posts are 1mm apart. Now, if a photo is taken of the same scene, the original perspective is reproduced when the photo also creates an image on the retinas with the same dimensions – 0.5mm for the football, and 1mm for the goal post separation. To meet this requirement, the focal length of the camera lens, the size of the digital CCD, the size of the print (or computer monitor) and the viewing distance must be correctly matched.

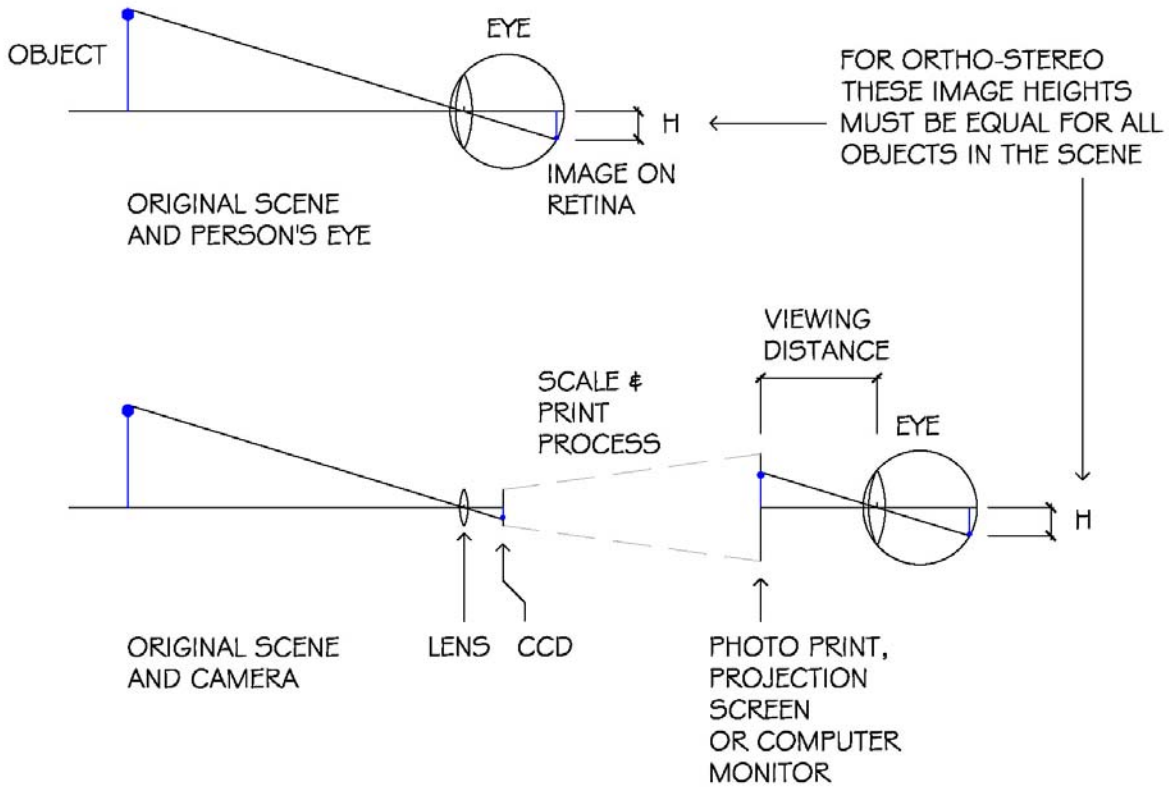

Fig 1 – Perspective requirements for ortho-stereo

2. Camera Separation

Ortho-stereo requires that the camera lens spacing be equal to the eye-separation of the person who will view the stereo photo. This means spacing the lenses by 65mm to 70mm. Even so, some people will have eye-separation outside these limits.

3. Resolution

The *resolving power* of the human eye is generally taken as 1 minute-of -arc (1/60'th of a degree) for high-contrast, sharply defined objects (see Appendix for types of visual acuity). This corresponds to the size of the horizontal and vertical strokes that comprise the letter "E" on the bottom line of the Snellen eye-test chart. For lower contrast objects, the eye's resolving power is reduced accordingly. However, there are two other types of resolution, which are called *stereo acuity* and *vernier acuity* – both are about 6 times greater than the resolving power (i.e. about 0.17 minutes-of-arc). At a minimum an ortho-stereo display should achieve 1 minute-of-arc resolution (i.e. the solid-angle of one pixel is 1 minute-of-arc); ideally it would achieve full stereo resolution.

Worked Example

Camera Parameters

For a digital twin-camera stereo rig using Canon A720 cameras, the closest spacing that can be achieved is 71mm. This is the same spacing as that used by the Stereo Realist film camera. The CCD (the Charge Coupled Device that takes the place of film) in the camera has a diagonal dimension of 7.60mm (based on Canon published data for 35mm equivalent focal length). The aspect ratio is the same as that of a television screen, being 4 units wide and 3 units high. This means the width of the CCD is 6.08mm and the height is 4.56mm.

Now, suppose we want to display images from this stereo camera on two computer monitors that are 1920 pixels wide and 1200 pixels high each. The diagonal dimension of the monitor is 24 inches, and the aspect ratio is 16:10 (the most common for current computer monitors).

Because the display shape is a wider format than the camera, let's make a decision to crop the image from the camera at the top and bottom, so it can fit to the computer monitor display. We can now re-work the image size from the camera, which will now be cropped to 16:10 format, with width 6.08mm as before, but the height is reduced to 3.8mm. Therefore the diagonal is 7.17mm.

It is convention that the "standard lens" for a given format has a focal length equal to the diagonal dimension of the format. So let's assume that our Canon twin camera rig has its zooms set to 7.17mm. (The nearest available focal length provided by the A720 is 7mm. The value 7.17mm has been retained for clarity).

The above calculations are shown in the table below from an Excel spreadsheet. The numbers in blue are the only numbers that you need to supply as "inputs" to the spreadsheet. All the other numbers are automatically calculated. This spreadsheet can be downloaded from
http://www.44bx.com/stereo/orthostereo.xls

Display & Viewing

Now, on the display side, we have a pair of monitors to display the image pair. We could use these in the following ways:

- side-by-side, with four mirrors for viewing
- facing each other, with two mirrors for viewing (after Wheatstone)
- arranged at right-angles, with half-silvered mirrors and polarising glasses

To view the images in ortho-stereo, the distance from the screen to the viewer's eyes has to be calculated. Using the spreadsheet, and entering values for: the screen-diagonal size; the number of

pixels; and the "ortho factor" (equal to 1 for true ortho-stereo) …the viewing distance is then calculated and found to be 610mm.

Also calculated is the image resolution at the retina, which for this case is found to be 1.52 minutes-of-arc. This is the angle subtended at the retina for one pixel of the display, and represents the maximum resolution of the display and viewing system. This angle is somewhat larger that the 1 minute-of-arc resolving power of the eye. Therefore, the individual pixels that make up the image may be visible, and the image could look somewhat soft (not sharp). [When displaying just a white image, the matrix or raster of pixels that comprise an LCD display is generally discernable for people with good vision at a distance of 500 to 600mm. When viewing stereo images, this raster will have a location in space.]

Obtaining a resolution of 1 minute-of -arc will cost more money, and require two displays that are about 2920 pixels wide and 1825 pixels high. The viewing distance will depend on the physical size of the screen, as well as other factors already discussed.

Orthoscopic Stereo Camera & Display
S.Spicer, June 2008
INPUTS are in **BLUE**

Camera			Display		
Lens Focal Length*	**7.17**	mm	Ortho Factor	**1**	
35mm-film equivalent focal length	43	mm	Display Size Diagonal	**24**	inch
* "standard focal length" = size of the masked-image diagonal			Ortho-stereo Viewing Distance	610	mm
CCD size, diagonal	**7.60**	mm			
Aspect ratio width	**4**		Display Pixels, width	**1920**	
Aspect ratio height	**3**		Display Pixels, height	1200	
Width/Height Ratio	1.33		Width/Height Ratio	**1.60**	
Native Megapixels	**8.0**	Mpixels	Displayed Megapixels	2.30	Mpixels
CCD Width	6.08	mm	Display Width	20.4	inch
	3.27	Mpixels		516.9	mm
CCD Height	4.56	mm	Display Height	12.7	inch
	2.45	Mpixels		323.1	mm
Horizontal Field of View	46.0	degrees	Horizontal Field of View	46.0	degrees

Mask the Camera Image (top & bottom) to Match the Shape of the Display			Image Resolution at Eye's Retina		
Aspect Ratio - Masked to Match Display	1.60		Pixel Horizontal Angle	0.00044	radians
Megapixels - Masked to match Display	6.7	Mpixels		0.02531	deg
Masked-Image Width, at CCD	6.08	mm		1.52	minutes of arc
	3.27	Mpixels			
Masked Image Height, at CCD	3.80	mm	Human Eye Resolving Power (for reference only)		
	2.04	Mpixels	Resolving Power	1	minutes of arc
Masked-Image Diagonal	7.17	mm (= "std. lens")	Vernier/Stereo Acuity	0.17	minutes of arc

Ortho Factor of 0.8

With the Stereo Realist hand viewer for film, the lens focal length is about 43mm, while the Stereo Realist camera has lenses of 35mm. This gives an approximation that is about 80% ortho-stereo. If we are willing to accept this compromise, so the display system mimics the stereo image produced by a hand-viewer, then the viewing parameters for our 24" computer monitors work out as follows:

- viewing distance: 762mm
- image resolution at the retina: 1.2 minutes of arc

Alternatively, if we decide that we want resolution to match the eye (1 minute-of-arc) then we must upgrade the computer monitors to be 2360 x 1475 pixels (maintaining the ortho factor at 0.8).

Orthoscopic Stereo Camera & Display
S.Spicer, June 2008
INPUTS are in **BLUE**

Camera			Display		
Lens Focal Length*	**7.17**	mm	Ortho Factor	**0.8**	
35mm-film equivalent focal length	43	mm	Display Size Diagonal	**24**	inch
* "standard focal length" = size of the masked-image diagonal			Ortho-stereo Viewing Distance	762	mm
CCD size, diagonal	**7.60**	mm			
Aspect ratio width	**4**		Display Pixels, width	**2360**	
Aspect ratio height	**3**		Display Pixels, height	1475	
Width/Height Ratio	1.33		Width/Height Ratio	**1.60**	
Native Megapixels	**8.0**	Mpixels	Displayed Megapixels	3.48	Mpixels
CCD Width	6.08	mm	Display Width	20.4	inch
	3.27	Mpixels		516.9	mm
CCD Height	4.56	mm	Display Height	12.7	inch
	2.45	Mpixels		323.1	mm
Horizontal Field of View	46.0	degrees	Horizontal Field of View	37.5	degrees

Mask the Camera Image (top & bottom) to Match the Shape of the Display		
Aspect Ratio - Masked to Match Display	1.60	
Megapixels - Masked to match Display	6.7	Mpixels
Masked-Image Width, at CCD	6.08	mm
	3.27	Mpixels
Masked Image Height, at CCD	3.80	mm
	2.04	Mpixels
Masked-Image Diagonal	7.17	mm (= "std. lens")

Image Resolution at Eye's Retina		
Pixel Horizontal Angle	0.00029	radians
	0.01647	deg
	0.99	minutes of arc

Human Eye Resolving Power (for reference only)		
Resolving Power	1	minutes of arc
Vernier/Stereo Acuity	0.17	minutes of arc

Screen Layouts

Two possible screen layouts are shown below (Fig 2 & 3). The pros and cons are as follows:

	Advantages	Disadvantages
Twin-Mirror System	- Brighter display (no light-loss) - Two side-by-side monitors are useful for normal computer work	- Some chance of stray reflections - Only one person can view - requires two first-surface mirrors for each monitor - Is not easily moved - More effort required to allow use of monitor for normal computer use
Polarised System	- Can sit closer (will provide ortho-stereo for short focal length camera lenses, which need closer viewing distances) - More than one person can view it, but only one person can be in the ortho-stereo position - Makes use of inherent 45-degree polarisation in many LCD screens (no separate polariser required)	- About 50% of light is lost in half-silvered mirror - More chances of stray reflections - Readily adapts to use for normal computer functions (e.g. turn one monitor off, or hinge the mirror).

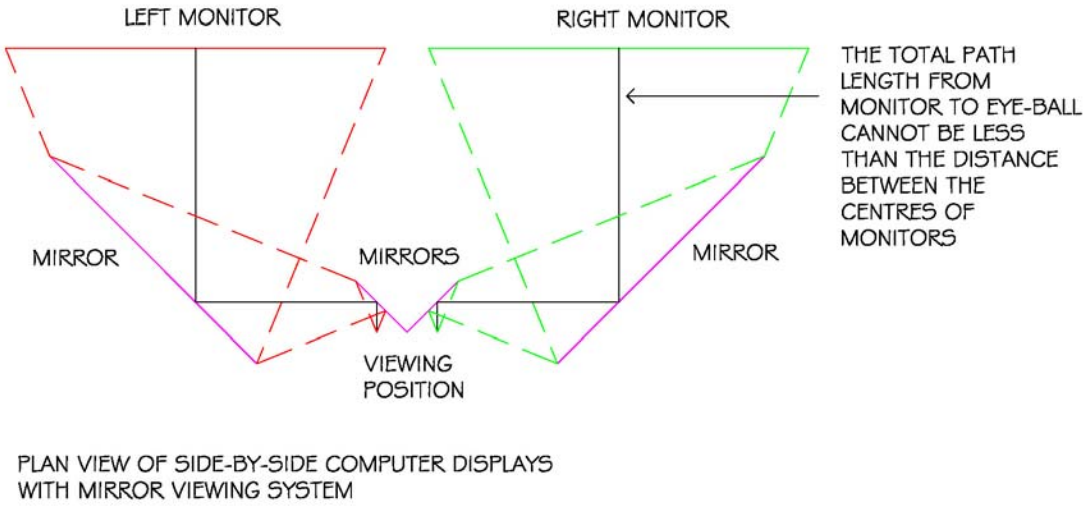

Fig 2 – Possible layout for side-by-side computer monitors and first-surface mirror system.

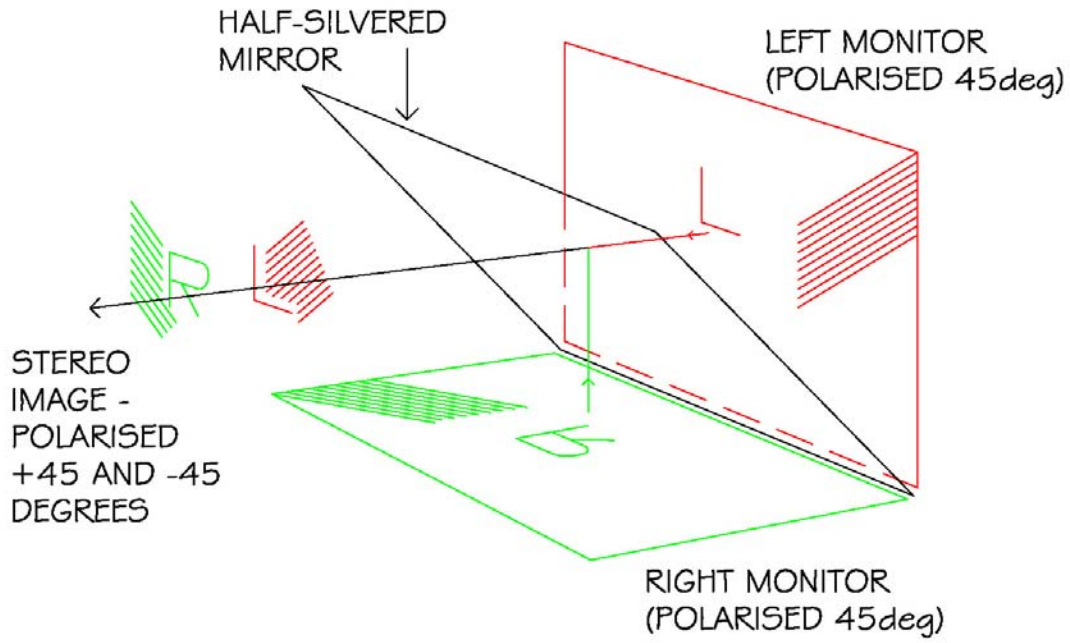

Fig 3 – Possible layout for polarised computer monitors. Note that the polarisation on each individual screen is the same orientation, but the mirror system results in orthogonal polarisation when viewed.

Summary

A pair of 24" computer monitors with about 2.3 megapixels (1920 x 1200) can be used for ortho-stereo viewing, but the image resolution at the eye will less than that of normal good vision (by a factor of about 1.5) and the pixel structure may be visible to people with normal vision. The viewing distance should be about 0.6m.

To replicate the image size of hand-viewers used for stereo images on film, the requirements are relaxed a little, as these viewers provide an approximation of about 80% ortho-stereo. The same monitors as above will yield about 1.2 minutes-of-arc resolution, viewed from about 0.76m.

To achieve ortho-stereo viewing with 1-minute-of-arc resolution (normal vision), approximately 5.3 megapixels are required in each monitor (2920 x 1825pixels), assuming a "standard lens" is used on the camera. Each time the camera lens focal length is halved, four times as many pixels are required in the monitors if resolution is to remain unchanged, and ortho-stereo viewing is to be maintained.

For viewing system to provide resolution that matches the vernier and stereo-acuity of the eye (0.17 minutes-of-arc) then about 34 times as many pixels are required in the display, compared to that required for a 1 minute-of-arc system (i.e. 184 megapixels). A camera with accordingly high lens and CCD resolution will also be required.

Attachments

1. Spreadsheet

D:\miscellaneous\
stereo\ortho stereo\c

Appendix

1. Types of Visual Acuity [extract from Reference 1]

Type of Visual Acuity	Value (minutes-of-arc)	Description
Point Acuity	1	The ability to resolve two distinct point targets.
Grating Acuity	1 to 2	The ability to distinguish a pattern of bright and dark bars from a uniform grey patch.
Letter Acuity	5	The ability to resolve a letter. The Snellen eye chart is a standard way of measuring this ability. 20/20 vision means that a 5-minute letter target can be seen 90% of the time.
Stereo Acuity	0.17	The ability to resolve objects in depth. The acuity is measured as the difference between two angles for a just-detectable depth difference.
Vernier Acuity	0.17	The ability to see if two line segments are collinear.

References

1. "A Survey of Display Device Properties and Visual Acuity for Visualization", Technical Report TR-2005-32, *Amit P. Sawant, Christopher G. Healey*, Knowledge Discovery Lab, Department of Computer Science, North Carolina State University, Raleigh, NC 27695-8207, http://www.csc.ncsu.edu/faculty/healey/download/cstr.05a.pdf
2. "Sensor Sizes", *Vincent Bockaert*, Digital Photography Review, http://www.dpreview.com/learn/?/Glossary/Camera_System/sensor_sizes_01.htm
3. Canon A720IS Digital Camera Specifications