Entering the Third Dimension - Viewing in Stereo

Cues, tricks, and how Human Vision works

By Chris Russ

Abstract: There are many cues for viewing the world in 3-D. Only one of those is stereo vision. This paper examines how and what people see in a 3-D world with some guidelines on generating such images.

What do people see in a 3-D world?

It turns out that we use a number of methods of identifying depth and distance. Stereo vision is not the complete answer. Rather, a number of visual cues convince us that we are seeing familiar 3-D objects and our brains fill in the rest.

Let us examine some of the cues that provide the impression that there is actual depth in images.

Visual Cues

- **Shading & Highlights**: People assume that the light source is above the object so that highlights are on top and dark shades are below.
- **Occlusion**: When an object is partially occluded, it appears to be larger and farther away. There is a famous optical illusion where the moon or sun is at the horizon and appears much larger. The optical size change due to the atmosphere is *very* minor.
- **Shadows**: If the position of the light source is known (or assumed) a shadow can provide significant depth information, especially if the shadow is cast on a surface or other object that is "known."
- **Bigger means Closer**: If there are two objects that are supposed to be the same size, the closer one is assumed to be bigger and visa-versa.
- **Hazy with Distance**: Objects that are farther away can appear to be "hazy."
- **Motion Flow, Perspective**: This is an entirely different concept. Basically objects that are getting closer move away from the vanishing point and objects that are going away converge on the vanishing point. The edges of these objects also have the same perspective effects. These perspective effects (parallel lines converge, closer objects are bigger, etc.) provide enough clues to drive an automobile with vision in only one eye.
- **Focusing Distance**: Although this varies greatly from person to person, it is possible to tell how far an object is away due to how hard it is to focus. This can be overridden by other clues (especially stereo images), but is another cue for depth. Objects that are not in the focal plane appear out-of-focus, making it easier to tell what objects are at the same distance. When focused on a nearby object, far away objects appear blurry (and easier to ignore).
- **Stereoscopy**: Finally, this is a powerful method for determining distance or enhancing an estimate in the near working distance. The vergence angle needed to "fuse" an object allows the brain to make an approximation of distance to the object.

Effective working Distance

Generally, the area that is covered by stereo vision is between eight inches and fifteen feet. Closer than that, it is virtually impossible to focus (either from eyestrain or lens limitations) and farther than that there is little useful stereo information because the eyes are so close together. This range is governed by the distance between the eyes. If they were farther apart, a greater working distance would be possible.

During World War I large binoculars with a two to three foot spacing were used (on tripods, of course) to range enemy positions in the trenches. By increasing the baseline between the lenses the effective working distance for stereo vision was over 300 feet. This made it possible for a human to judge distance for artillery and damage assessment more accurately.

Hand-Eye coordination

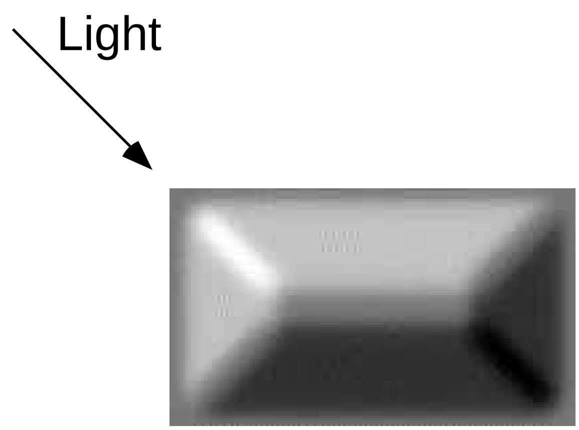
There are many advantages to stereo vision for tool users like us. Try to eat with chopsticks with one eye covered. It is possible, but a great deal more difficult for people that have stereo vision. The interesting thing is that it is possible to function with only one eye. Just more difficult. (The other main advantage to having two eyes is redundancy.)

Generating Images

From a computer graphics perspective, there are other reasons to understand 3-D vision. For instance, the better the representation of a scene, the more likely the observer will believe that it is real. While it is possible to suspend disbelief with relatively primitive graphics, most computer graphics people are professional skeptics and look for the places in images where the effect falls apart. Sometimes this is because edges do not look right or the physics model is not exactly correct. Regardless of this, a professional observer can tell a lot by looking at the images.

Shading & Highlights

People assume that the side of an object that light strikes is going to be brighter. As a result, as the angles fall away from normal to the light source, the surfaces become darker. Both the light source highlights and the shading from the angular changes tell a lot about the shape of the surface. In addition, if the material is recognized (wood, plastic, metal, skin, etc.) the shading is expected to conform to the real world. (One of the major problems with rendering surfaces is that the BRDF: *Bidirectional Reflectance Distribution Function* of the surfaces is poorly simulated by a specular / diffuse model. Also, surface roughness must be approximated and this can be difficult with simple bump maps.)



Shadows

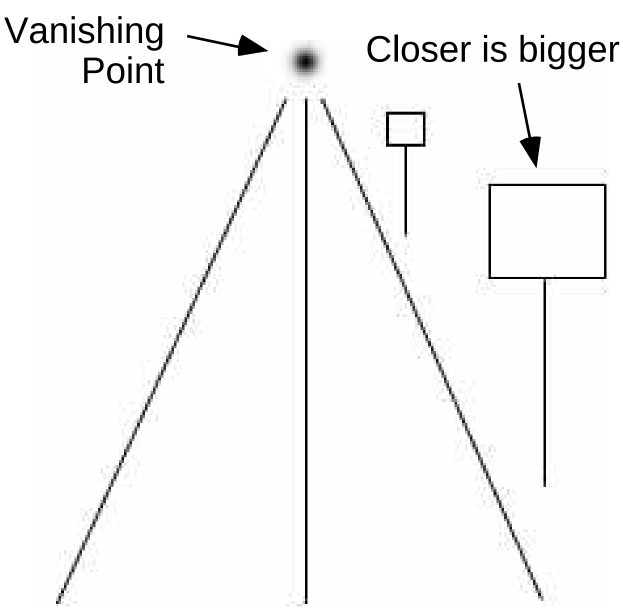
Shadows are a powerful cue, especially when the light source position is known. In the Macintosh user interface, the light source is assumed to be the upper left corner, approximately 45° out from the plane of the page -- or over the left shoulder of the user.

These icons show cast shadows on an assumed planar surface. The icons achieve depth and the “desktop” is made painted in the viewer's brain.

Icon Shadows

Perspective

Objects that are farther away are smaller and if the viewer is in motion tend to move toward the vanishing point. Objects that are moving toward the viewer come out of the vanishing point and get larger. The notion of "motion flow" and the attempt to measure it was a mainstay of computer vision in the 1980's.



Haze

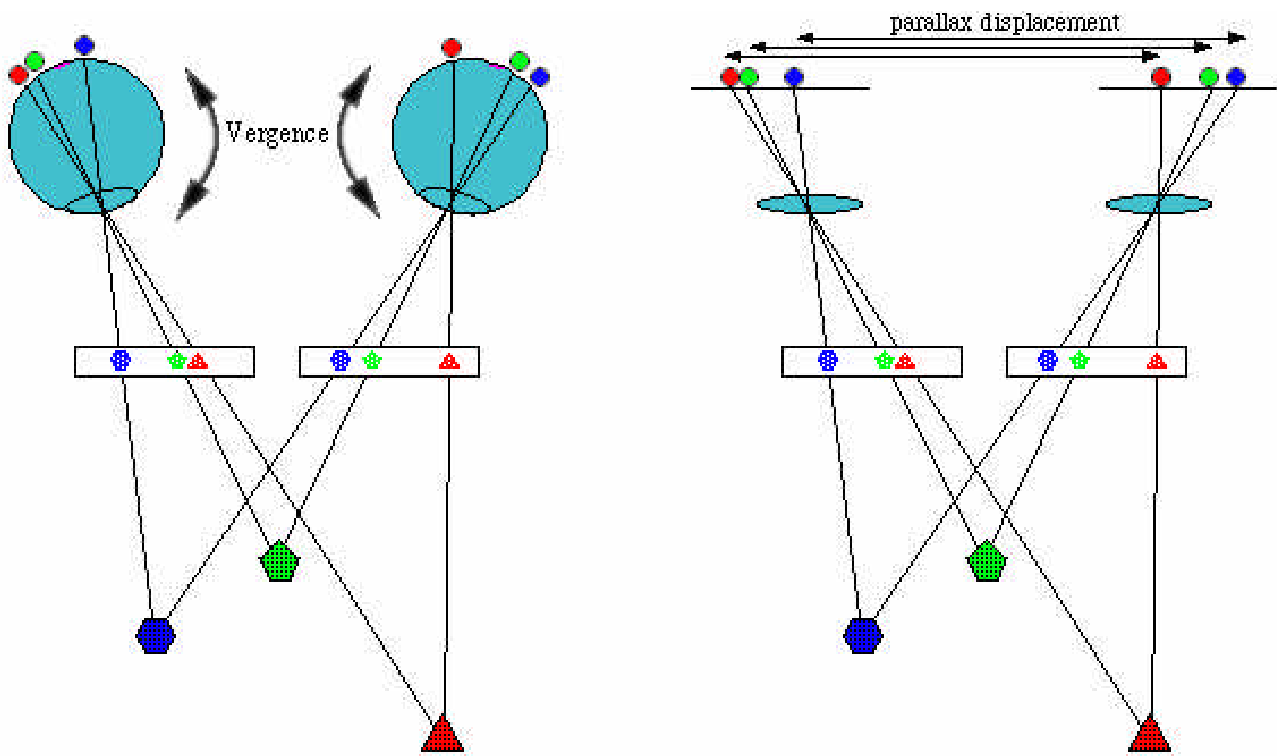
Objects that are a long distance away appear hazy, as in this photograph. This is especially important in scenes that show a large range of distances (near vs. far) and have particulate material or water droplets in the air, or are showing illuminated beams of light (Rayleigh scatter).

(Photograph taken at 800 yards.)

Parallax Diagram

In the following example, the **vergence** angle that the eyes have to turn in order to "fuse" an object is a strong indicator of distance. By increasing the distance between the projections for the two views, the apparent distance is increased. This distance is called **parallax** and there are some real-world limits on the amount of parallax that people can see.

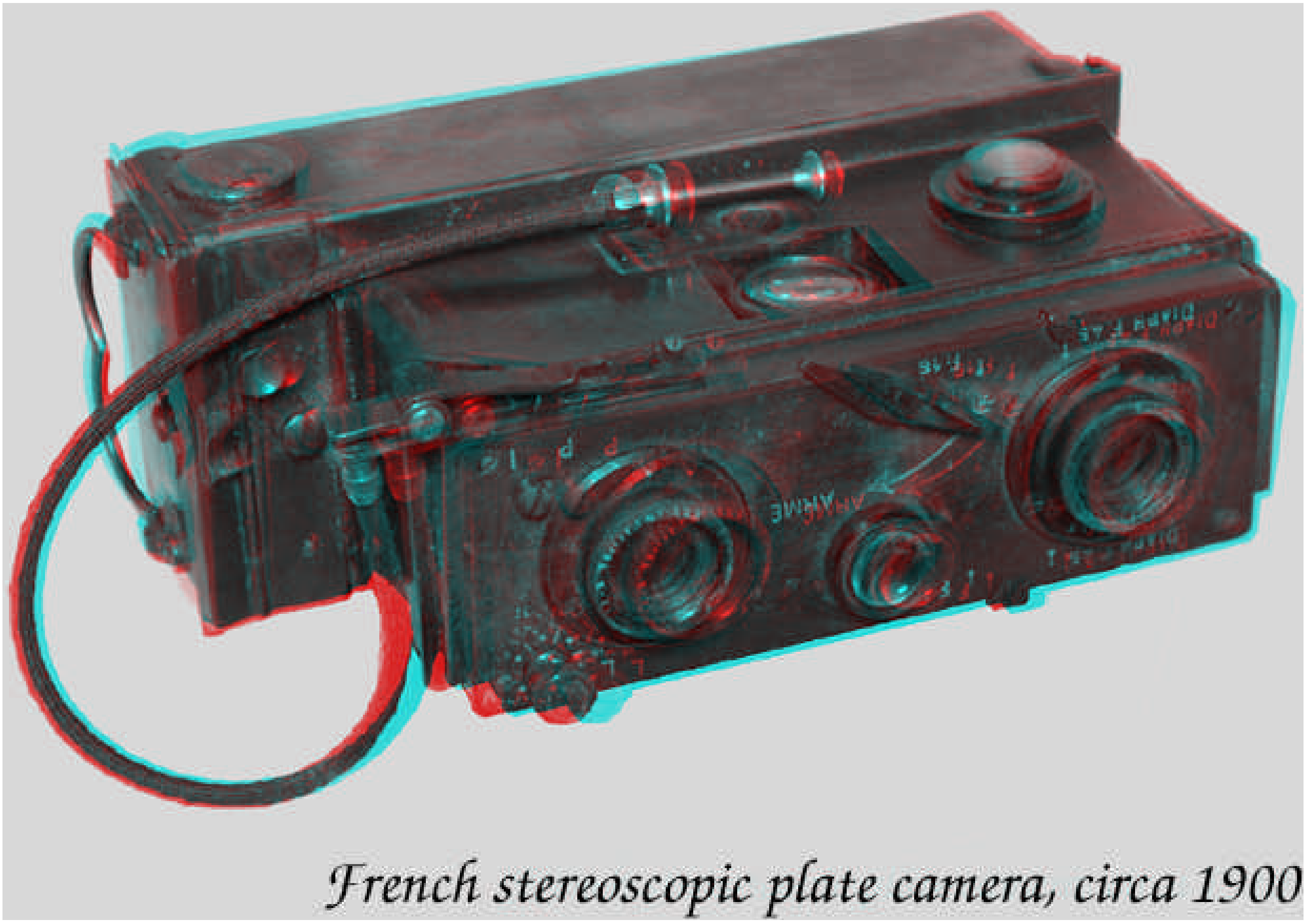
Positive parallax causes the objects to appear to be farther away than the projection plane. Negative parallax causes the objects to appear closer than the projection plane. The real physical limit is the distance between the viewer's eyes. (Although some people can handle more parallax than that it is generally not a good idea.) Practical limits are +2" for positive parallax (apparently infinity) and -1" (half the distance between the projection plane and the viewer) for negative parallax.



Measuring distance from disparity

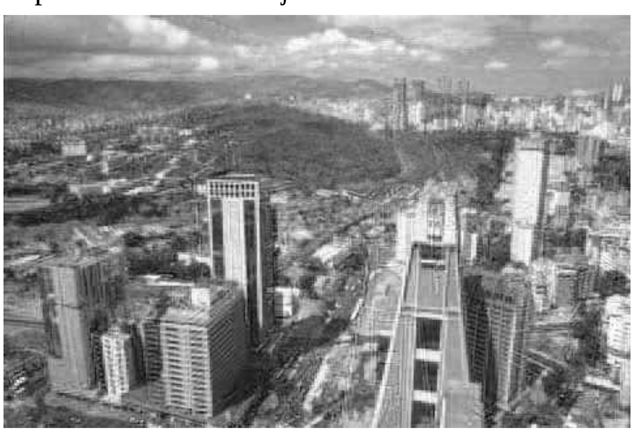
Stereoscopy

In the next figure, this stereo camera (from ~1900 A.D.) was used to take plates for projection. Notice the center view-finder, the shutter cable and the independent optics for left and right.



French stereoscopic plate camera, circa 1900

With practice, people can view stereo pairs side-by-side by diverging their eyes (as if focussing at infinity). It requires practice and it is possible to have objects that are more than normal eye-spacing apart, but it can be painful if that distance is exceeded by much. There is an alternative method for fusing by crossing eyes, but it is not used as much.



Left eye *Stereo pair of Caracas, Venezuela* *Right eye*

Red-Cyan Anaglyphs

Once the stereo images are captured, the problem of how to view them comes up. A simple method, called Anaglyph Stereo, projects two images, one in red and one in blue, so that a simple pair of glasses can be used to separate the images for each eye.

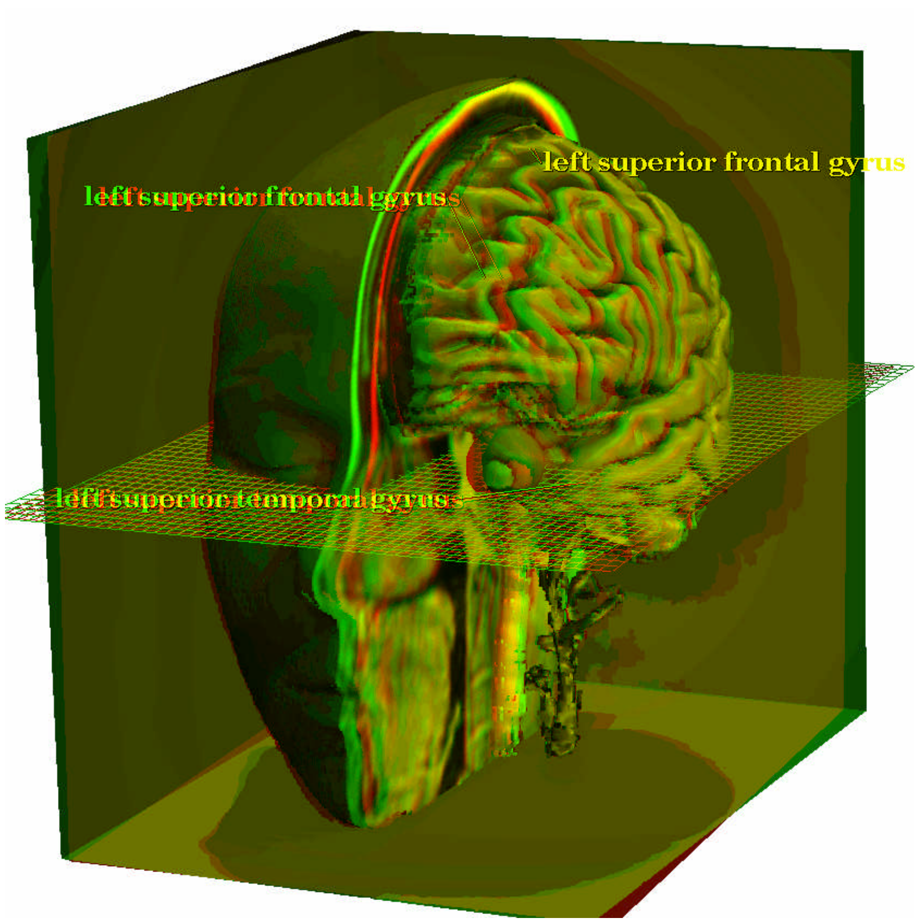
This is the most common form of 3-D viewing and gets re-discovered every ten years or so.

(A number of monster movies in the 1950's and 1960's were made in anaglyph stereo.)

One of the side effects is that it is difficult to view color in images. It can be done by using red and blue for the stereo effect and painting details in green in areas that are otherwise uniform.

Red/Blue Polarized Glasses

Note that the standard for anaglyph stereo is that the l**eft eye is red**.



Red-Greed Anaglyph of section of human head ©1994 University of Hamburg, Germany

Anaglyph Stereo Picture of Wrigley Field

Polarized light

Polarized glasses at 45° and 135°

Most two-lens 3-D projection systems are polarized systems, while most print methods are Red / Blue or Red / Green anaglyph. (Except for "Magic Eye" autostereograms.) The polarizers are fixed at angles of 45° and 135° (at right angles to each other) were established in the 1940's. It is important to use a silver (mirror) screen instead of a white screen since a diffuse surface will depolarize the light and make stereo viewing impossible.

Shuttered Glasses

Some video monitors can project at 120 frames per second (FPS). By synchronizing a shuttered (LCD) pair of glasses with an IR transmitter, these glasses allow each eye to see only one version of the image, each at 60 FPS.

Goggles with Alternating Images

Another method is to simply project a different image to each eye. The brute force method is to use a pair of virtual reality goggles.

i-Glasses™

The problem with LCD's is getting glass with enough pixels to make a high-resolution image. This particular brand encodes S-Video so that the even lines go to one eye and the odd lines go to the other. This is rapidly becoming a standard in the 3-D videotaping community. Also, there are lenses that can be placed on video cameras to record stereo home video.

Autostereoscopic Methods

The Holy Grail of display technology is a way of displaying 3-D images without any special hardware at all -- no glasses, goggles, helmets, etc. Some have shown more success than others. Here are a few examples of current technology:

HOLOGRAM

Holography works by transmitting a coherent light source through an interference pattern. The most recent breakthrough is color holography where a "white" laser is used on 3 different interference patterns to generate a color hologram. Also, there are outfits that can create custom holographic images from a 3-D data set.

Hologram of a Ford P2000 Prodigy Sedan

The two special components to the Ford hologram are 1) that it is in color and 2) that it is 10 feet x 4 feet and created in tiles. (Image created for Ford by Zebra Imaging - Austin, TX.)

radargifts.com - Color Hologram Key-chain

At the other end of the price spectrum are color holographic ornaments and key-chains, again using "white" lasers.

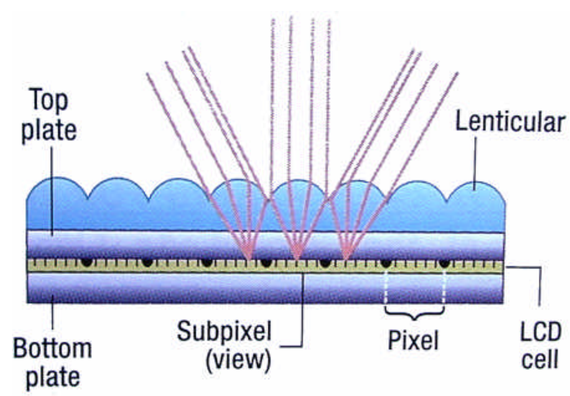
At this time there is no way to generate a holographic interference pattern in real time, nor is there a way to display it with current display technologies (although it may be possible with e-Paper in the future). There is real hope for this technology. (Note: It will be very silicon intensive when it is done. The coding will be impressive and probably be a gigantic Fourier Transform.)

VARIFOCAL MIRROR

This is a cute technology for viewing a 3-D volume. It works by projecting a set of ~30 images on a concave mirror that moves up and down once a second. In effect, a volume is "painted" in your brain. Unfortunately it is very difficult to view a volume. Humans are very good at viewing surfaces and this technology has not been used much.

CYLINDRICAL LENS DISPLAY

In this display, a sheet of cylindrical lenses is placed on a fairly high resolution LCD panel. The requirement is that each cylindrical lens covers a number of pixels horizontally (and thus it is important that the pixels be small). By changing which pixel in a set of pixels is turned on, the angle that the light is projected is different. Thus, it is possible to project different images to each of the viewer's eyes, or as the viewer moves, a new image will become available.



Subpixels project at different angles

Rules for Anaglyph Images

Back to the easiest kind of stereo images to generate: Anaglyph.

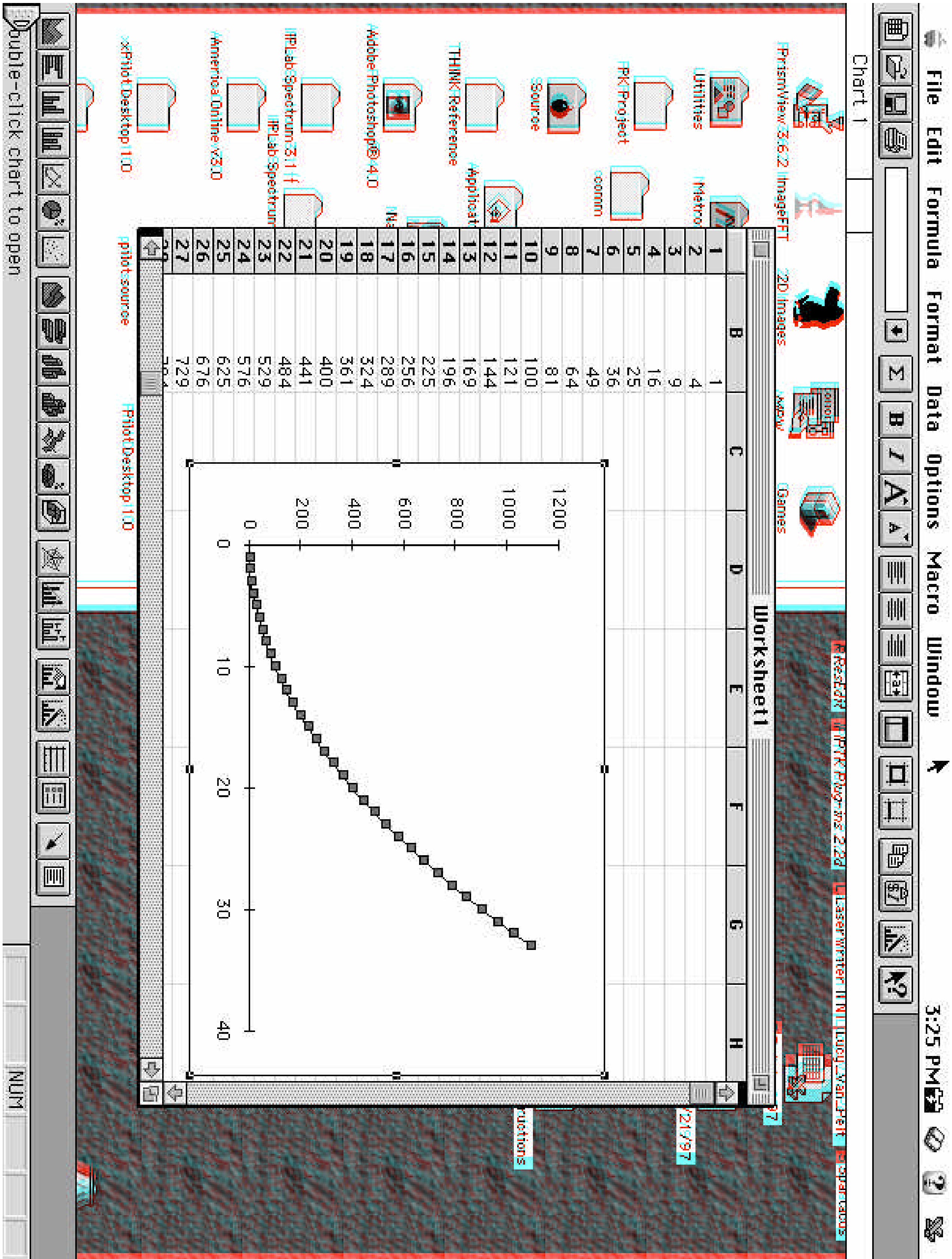
Here are some rules for making practical Anaglyph images:

- Planes perpendicular to viewer are easiest to fuse (and generate paired views)
- Planes and simple objects at other angles are preferred.
- Good fusion requires texture or edges. The viewer's brain fills in the rest. Surprisingly little of the scene is needed for fusion. (Random Dot Stereogram images are a good example of this.)
- Once the object is "grokked" you're anchored to it (and don't need to fuse it again). It is possible to examine other objects in the scene without having to re-"fuse" the scene.
  (Warning: If the complexity is too great, it is very possible to completely loose the scene, or not be possible to fuse it at all.)
- Rapidly moving objects with big parallax changes are hard to track especially if they don't agree with other cues. Disney and others provide effects where objects "leap" out of the screen at the audience. Many people cannot track the changes that fast.
- Keep the functional parallax to a reasonable working distance to minimize eyestrain. It is hard to track things two feet away and 8 feet away at the same time.
- Color blind people can still use anaglyph stereo glasses. They can *see* red, just not tell that it *is* red.

Stereo Desktop Metaphor (MacOS 8)

The next figure shows what a 3-D desktop would look like. The following things could be done to make a real implementation of a stereo desktop

- Project Windows in flat planes. (The window bars could be 3-D "textured" very easily. Hidden windows could be projected off to the side or at different angles.)
- Windows in front are "In Front." This means that windows that are in back get pushed farther away as more things get pulled in front.
- There should be a shadow under mouse and windows. This cue gives an impressive 3-D effect.
- Millions -> 256 grays, 256 colors -> 16 grays (You don't get color!)
- Alerts can "Pop Out" using negative parallax.
- Keep the main window in the plane so there is no parallax -- this allow the computer to be used even if the user is not wearing glasses *and* provides minimum eye-strain since the focal distance and the fusion distance are the same.
- The desktop will "fuse" much better if it is textured. In fact, a 3-D desktop image would look very cool.
- Icons and controls can be drawn in 3-D directly. By supplying a change to PlotIcon functions, it would be possible to make a depth resource for how to render 3-D icons. CDEFs and WDEFs could be very interesting.
- Practical limits for parallax are 2" for positive (infinity) and -1" (half the distance between the projection plane and the viewer) for negative parallax. Because the viewer is looking into a screen that is often two feet away, this allows a very versatile range.



Stereo Desktop Snapshot

Conclusion

We are eagerly awaiting better technology for 3-D viewing. However, using the myriad tools that we have now, it is very possible to build extensions to our existing desktop metaphor and create new, entirely 3-D, metaphors for operating computers.

Bibliography

Wilson, Andrew "Vision Systems Design, May 2000," pp.45-52

Russ, John C. "The Image Processing Handbook, Third Edition," 1998, CRC Press.

URLs for stereo

http://stereoscopy.com/reel3d/

http://www.studio3d.com/pages2/store_glasses.html

http://stereoscopy.com/stereosociety/

http://stereoscopy.com/3d-info/index.html

http://axon.physik.uni-bremen.de/research/stereo/

URLs for Holography

http://www.radar-eoir.com/

http://www.zebraimaging.com/

http://www.thecarconnection.com/news/nf/ford_hologram.asp