

STEREO-VISION FORMATS FOR VIDEO AND COMPUTER GRAPHICS

Lenny Lipton

StereoGraphics Corporation
2171 E. Francisco Blvd., San Rafael, CA 94901
phone: (415) 459-4500
fax: (415) 455-1867
email: lenny@crystaleye.com

ABSTRACT

There are several formats for time-shared stereoplex electronic displays. A stereo-vision format is the technique used for assigning pixels (or lines, or fields) for the left and right images, enabling them to be available at the display screen as an image with true binocular stereopsis.

These days most graphics workstations intrinsically output a high field rate, and don't require the above-and-below solution once used for workstations and now more commonly used on PCs.

Another approach uses spatial multiplexing of rows or columns, either for individual selection devices or autostereoscopic displays.

A new format, the White-Line-Code (WLC) system, was developed for PCs and offers a low cost but high resolution. This format doesn't care if the left and right fields are in interlace or progressive scan modes, and it doesn't care about field rate.

1. FORMAT DEFINITION

There are a number of ways to prepare multiplexed images for stereo-vision electronic displays, and there's more than one way to discuss and classify them. Nevertheless, it's my hope that this discussion, arbitrary by its nature, provides some insight into the structure of such displays.

First let's define what is meant by a stereo-vision format: *An electro-stereoscopic format is the method used for*

assigning pixels (or aggregates of pixels — lines or fields) to respective left and right images, thus making them available at the display screen, to the eyes of the observer, as an image with binocular stereopsis. Given this definition, one might simply apply the permutations that come to mind, based on a knowledge of electronic display structure, to come up with a variety of schemes such as pixel, line, or field sequential.

There are two major classifications I prefer: field sequential and pixel sequential. As it turns out, the most commercially important format is field sequential, but both field and pixel sequential are technically interesting approaches. We shall also see that the two approaches may be combined.

The major challenge facing the designer of an electronic stereoscopic display system is that of having the system interface seamlessly with the existing display infrastructure. The temporal multiplexing scheme involving encoding left and right views on alternate fields ("field" is used in the most general meaning and may in fact be a complete picture in the progressive scan mode) has proved to be the most agreeable in this regard, causing the fewest required hardware changes. This is an important consideration which has contributed to the commercial success of such products.

2. THE FIELD SEQUENTIAL FORMAT

The terms "alternate field," "time-multiplexed," "time-division multiplexed," or "interlaced" have all been used to describe what I have chosen to call the field sequential technique here. A cautionary note about the term "interlaced" is given at the conclusion of this section.

Field sequential products using eyewear as a selection device dominate the marketplace. The eyewear may be active, using liquid crystal (L.C.) shutters like StereoGraphics®, CrystalEyes® (professional applications) or SimulEyes® (consumer applications) products, or they may be passive polarizing spectacles as are used with StereoGraphics' Projector ZScreen®. A related technique employing the field sequential approach is embodied in products distributed by NuVision (formerly a Tektronix unit), in which a large L.C. modulator is fitted over a monitor screen and is used in conjunction with appropriate analyzing spectacles.

Whether active eyewear with L.C. shutters, or passive eyewear and a modulator are used, the user remains looking through a shutter. In the case of the on-screen modulator, one can consider the parts of the shutter to have been distributed between the modulator and the passive eyewear. But in either case the time-division multiplex technique is at work.

The most widespread method for viewing electronic stereo-vision images uses CrystalEyes on workstations for scientific visualization applications such as molecular modeling. CrystalEyes are active electronic eyewear which incorporate L.C. shutters whose occlusion rate and phase are controlled by an infra-red synchronization signal originating at or near the monitor. The information for the infra-red signal originates from the display system's video signal[1].

The term "interlace" is sometimes used to describe any time-multiplexed display, but the reader will understand that such use is incorrect when I explain interlaced stereoscopic display.

3. INTERLACED STEREO

The original and basic stereo-vision television format takes advantage of the odd-even interlaced structure of the medium to encode left and right images on alternate fields. It's a method that's still used today, and it has the virtue of using standard television sets or monitors, standard VCRs, and inexpensive demultiplexing equipment. In fact, the heart of the system is a simple field switch (off-the-shelf as a single chip) that shunts half the fields to one eye and half to the other.

Because of the low field rate (half of the 60 field per second North American video rate), the method results in flicker — which some people find more objectionable than others. One way to reduce the flicker when viewing such an image is to reduce its brightness by adding neutral density filters to the eyewear, and also by reducing room illumination. Another concern is that each eye sees half the number of lines (either all of the odds or all of the evens) which are normally available, so the image has half the resolution. Peculiarly enough, such images don't seem to have increased visibility of the raster structure, which I would expect to see. I've not run across this observation in the literature and consequently I haven't read an explanation for this puzzling phenomenon.

Another difficulty with this approach arises from the reduced temporal sampling rate (30 fields per second rather than 60, as is the case for standard video) which is, in addition, exacerbated by the fact that sampling of the field-switched cameras occurs out-of-phase. This can produce a temporal parallax artifact visible as a kind of jittery motion effect, especially noticeable for rapidly moving objects. I invented and patented a corrective technique involving capturing the images simultaneously, storing them, and then presenting them in the interlaced mode[2].

The interlace approach, in which odd and even fields are used for left and right information, has turned out to have an interesting application these days when used in conjunction with an HMD using L.C. displays, such as the consumer oriented products of Virtual IO. Because of the long persistence of L.C. displays, the lower than usual number of longer lasting fields will produce a more or less flicker-free image.

One source uses the *page flipping mode* for DOS based action games, typically running at 70 fields per second. Strictly speaking this is not an interlaced application, since page flipping is a variant of a VGA standard which produces a progressively scanned mode.

The other source is in fact truly interlaced: It is video, which (for North American NTSC) is 60 fields per second, or 30 lefts shuffled together with 30 rights, encoded on the odd and even fields. For that matter, since each eye — in an interlaced stereoscopic display — sees lines written in the same position, each eye sees, in effect, a progressively scanned image.

4. SEGMENT AND LINE SEQUENTIAL

The on-screen or monitor modulators using passive eyewear mentioned above use the Byatt[3] segmented shutter in which a number of horizontal segments of the L.C. modulator are activated in synch with the beam. As the beam reaches different raster locations, the Byatt modulator is switched or animated to follow its location. This approach might also be classified as a distinct variant, *segment sequential*, and is suggestive of the line sequential technique.

Although the line sequential approach has been discussed in the literature[4], there are no production shutters fast enough to realize the suggestion. The line sequential approach is interesting because, were it practical, it would be able to achieve a flicker-free result without having to change the standard television field rate. Left and right images would be presented on alternate lines, and switched to the eyes by appropriate shutters covering the eyes. Thus the present TV appliances in people's homes could serve as stereo-vision displays, with appropriate synch detection hardware and eyewear.

5. INTERDIGITATED IMAGES

Pixel sequential or interdigitated images have several advocates, for both stereoscopic and autostereoscopic applications.

To achieve their stereoscopic interdigitation method, VRex uses the interlaced format — but, interestingly, with a different selection technique. Unlike the time-division multiplex use of interlace for viewing through shuttering eyewear, or the HMD use of interlace for a flicker reduced display using a twin L.C.D. stereoscope, VRex uses interlace to interdigitate the left and right views. The VRex system uses an L.C.D. panel with a Parsell[5] Matrix they call micropol, made up of pixel- or line-wide strips of polarizing elements in juxtaposition with alternate rows of L.C. pixels. The L.C. panel, because of its fixed pixel location, guarantees good juxtaposition with the odd and even fields and the associated polarizing strips.

The long image lag of the L.C.D. has been used to good effect, and in this case suppresses flicker that might otherwise be seen in a display with short lived image

elements. The technique is used for both projection and direct viewing using a laptop modified with a Parsell Matrix.

Dimension Technologies and others use interdigitated images with vertical columns rather than horizontal rows. These columns, typically with left and right images positioned side-by-side in strips, are aligned with an appropriate selection device — in the case of several Japanese efforts, with an overcoated lenticular screen. Dimension Technologies uses an inverted raster barrier approach in which thin columns of rear illumination are created to direct the appropriate image stripe within a column to the appropriate eye.

6. ABOVE-AND-BELOW FORMAT

At StereoGraphics our concern has been to create stereo-vision formats and add-on selection devices which operate within the existing infrastructure of computer graphics and video systems, without modification to the infrastructure hardware or basic working procedures. The above-and-below method[6], which I invented, has survived for computer graphics on PCs, but has been eclipsed for workstations and video. **The method uses two subfields arranged above and below each other in a single standard field.** The images in these subfields are squeezed top to bottom by a factor of two[7].

At the standard 60 fields per second it takes half the duration of an entire field, or 1/120th second, to scan a subfield. When played back on a monitor operating at 120 fields per second, the subfields which had been juxtaposed spatially become juxtaposed temporally. Therefore each eye of the beholder, when wearing the proper shuttering eyewear, will see 60 fields of image per second, out of phase with the other 60 fields prepared for the other eye. Thus it is possible to see a flicker-free stereoscopic image, because each eye is seeing a pattern of images of 1/120th second followed by 1/120th second of darkness. When one eye is seeing an image, the other is not, and vice versa. (The field rate is typically 120, but anything somewhat slower or a great deal faster will work fine for many applications.)

Today there are many models of high-end graphics monitors which will run at field rates of 120 or higher. Providing a synchronization pulse is added between the

subfields in the subfield blanking area, such a monitor will properly display such images. The monitor can unsqueeze the image in the vertical so the picture has the normal proportions and aspect ratio.

When the term “interlace” is applied to such a display it is irrelevant, because the stereo format has nothing to do with interlace, unlike the odd-even field format explained above. The above-and-below format can work for interlaced or progressively scanned images and will preserve the interlaced characteristics of such displays.

StereoGraphics’ synch doubling emitter, the model EPC, for the above-and-below format is available for PCs. It adds the missing synchronization pulses to the vertical blanking for a proper video signal. The EPC unit also displays SimulEyes white-line-code formatted images. The shutters in the eyewear are triggered by the emitter’s infra-red signal.

If the image has high enough resolution to begin with, or — more to the point — enough raster lines, then the end result is pleasing. Below 300 to 350 lines per field the image starts to look coarse on a good sized monitor viewed from two feet, and that’s about the distance people sit from workstations or PC monitors — which sheds light on the basis for why this approach is obsolete for video. NTSC has 480 active video lines. It uses a two-fold interlace so each field has 240 lines. Using the subfield technique, the result is four 120-line fields for one complete stereoscopic image. The raster looks coarse, and there is a better approach for stereo video multiplexing as I shall shortly explain.

At the frequently used 1280x1024 resolution an above-and-below formatted image will wind up at about 1280x500 pixels per eye (some lines are lost to blanking). Even from the workstation viewing distance of two feet, most people would agree that this is good quality.

7. STEREO-READY COMPUTERS

These days most graphics computers, like those from SGI, Sun, DEC, IBM, and HP, use at least a double buffering technique to run their machines at a true 120 fields per second rate. Each field has a vertical blanking area associated with it which has a synchronization pulse. These computers are intrinsically outputting a high field rate and they don’t need the above-and-below solution

to make flicker-free stereo images, so they don’t need a synch doubling emitter to add the missing synch pulse. These computers are all outfitted with a jack that accepts the StereoGraphics workstation emitter, which watches for synch pulses and broadcasts the IR signal with each pulse. Most of these machines still offer a disproportionately higher pixel count in the horizontal compared with the vertical.

Some aficionados will insist that square pixels are de rigueur for high-end graphics, but the above-and-below format (or most stereo-ready graphics computers) produces oblong pixels — pixels which are longer in the horizontal than they are in the vertical. The popular 1280x1024 display produces a ratio of horizontal to vertical pixels of about 1.3:1, which is the aspect ratio of most display screens — so the result is square pixels. But in the above-and-below stereo-vision version for this resolution, the ratio of horizontal to vertical pixels for each eye is more like 2.6:1, and the result is a pixel which is longer than it is high by a factor of two.

SGI’s high-end machines, the Onyx® and Crimson® lines, may be configured to run square-pixel windowed stereo, with separate addressable buffers for left and right eye views, at 960x680 pixels per eye (at 108 fields per second). This does not require any more pixel memory than already exists to support the planar 1280x1024. Some SGI high-end computers have additional display RAM available, and they can support other square-pixel windowed stereo resolutions, though higher resolutions come at the expense of field rate. For example, such high-end high-display-memory SGI systems support 1024x768 pixels per eye, but only at 96 fields per second, which is high enough to have a flicker-free effect for all but very bright images in bright rooms.

For most applications, having pixels which are not perfectly square can result in a good looking picture. The higher the resolution of the image, or the more pixels available for image forming, the less of a concern is the shape of the pixels.

8. SIDE-BY-SIDE VIDEO

StereoGraphics developed the side-by-side technique[8] to address a significant problem of the above-and-below method as applied to video — not enough raster lines.

While the above-and-below solution is a good one for computer graphics applications because computer displays often output more raster lines than television, StereoGraphics' video products use a different technique to create stereo formats. First, for real-time viewing, this is what we do: The left and right images from the two video camera heads making up a stereoscopic camera are fed to our View/Record unit, and for viewing real-time they are stored and then played back at twice the rate at which they were read.

In addition, the fields are concatenated or shuffled to achieve the necessary left-right pattern. The result is an over-30-KHz or twice-normal-video-bandwidth signal which preserves the original image characteristics but, in addition, is stereoscopic. What I've described here is for real-time viewing using a graphics monitor with 120 fields per second capability, and it is the function of the View section of the View/Record box to produce such a signal.

Once it becomes necessary to interface with the existing television infrastructure a problem arises: NTSC recorders must be used if we are to have a generally useful system, and that means some effort must be expended to contain the image to within the NTSC specification of about a 15 KHz line rate. Thus, the Record section of the View/Record box serves to compress the left and right images so that they occupy the normal NTSC bandwidth. It does this by squeezing the images horizontally so that they occupy a single standard field, and the resultant signal is in fact an NTSC signal which may be recorded on an NTSC recorder. (We also make a PAL version.) When played back, the side-by-side analog video image is digitized by our Playback Controller, and formatted for stereo viewing. The result is an image which has characteristics which are similar to the real-time image described above.

9. DUAL-STREAM

Interest in artificial reality has produced a modern electro-stereoscopic format which is similar to the original stereoscopic format, but in a new guise. High-end products, like the Fakespace boom or Kaiser Electro-Optics HMD, use dual streams of images — one for each display. The observer looks into a Brewster stereoscope to see each display, which uses a high resolution

miniature monochrome CRT and a color shutter to produce field sequential color. Each display runs at 180 fields per second, to give each primary a chance to be displayed.

Typically these devices are used in conjunction with a Silicon Graphics Onyx-class workstation. The result is the best looking artificial reality display I've seen (rivaled only by the CAVE[9] which uses rear projection and CrystalEyes in a room environment). Thus the technique uses dual-stream for stereo-vision, and field sequential multiplexing for color — an impressive combination of technologies to achieve the desired result.

10. WHITE-LINE-CODE

The White-Line-Code (WLC) system is used for Pentium-class PCs and offers a high quality but low cost solution to the problem of stereo-vision imaging. For this format it is immaterial if the left and right fields are in interlace or progressive scan modes, and field rate is not directly of concern. Moreover, in the interlace mode either odd or even lines may be assigned to either perspective view. WLC was created to offer the most flexible possible stereo-vision system for the content providers, developers, and users[10]. The hardware components of the WLC allow for rapid user installation.

On the bottom of every field, for the last line of video, white lines are added to signify whether the field is a left or a right, as shown in the illustration. The last line of video was chosen because it is within the province of the software developer to add it. When our electronics see either a left or right white line it is prepared to shutter the eyewear once the vertical synch pulse is sensed.

The WLC is universal in the sense that it simply doesn't care about interlace or progressive scan or field rate or resolution. If the WLC is there, the eyewear will shutter in synchrony with the fields and you'll see a stereoscopic image. Our SimulEyes product is used with the WLC and PCs.

The two most popular modes of operation for WLC are (1) the page-swapping mode which, as mentioned above, is used most often for DOS action games running at either 70 or 80 fields per second (StereoGraphics has worked

out drivers for 80 fields per second or higher, which reduce flicker considerably), and (2) the multi-media or Internet mode which runs at 90 fields per second interlaced at 1024 by 768.

New video boards are coming along for the PC, and these use a technique which is similar to that used in stereo-ready workstations. They assign area in the video buffer for left or right images with a software “poke.” Most of these boards offer a 3.5mm jack that allows for the plug-in of SimulEyes OEM eyewear which includes on-board shutter driver capability.

Another significant beginning trend is that board and monitor manufacturers are enabling their products to run at higher field rates, say 120 fields per second, as is the case for workstations. When this happens the difference between PCs and workstations will vanish vis-à-vis stereo-vision displays.

11. ACKNOWLEDGMENT

Thanks are due to Bob Akka for sharing his knowledge about square pixel options.

12. REFERENCES

1. Lipton, Lenny and Marvin Ackerman, *Liquid Crystal Shutter System for Stereoscopic and Other Applications*, U.S. Patent No.4,967,268, Oct.30, 1990.

2. Lipton, Lenny, *Stereoscopic Television System with Field Storage for Sequential Display of Right and Left Images*, U.S. Patent No.4,562,463, Dec.31, 1985.

3. Byatt, Dennis W.G., *Stereoscopic Television System*, U.S. Patent No.4,281,341, Jul.28, 1981.

4. Nakagawa, Kenichi, Kojiro Tsubota, and Kunihiko Yamamoto, *Virtual Stereographic Display System*, U.S. Patent No.4,870,486, Sep.26, 1989.

5. Parsell, Richard K., *Method and Apparatus for Viewing Pictures in Stereoscopic Relief*, U.S. Patent No.2,218,875, Oct.22, 1940.

6. Lipton, Lenny, Michael R. Starks, James D. Stewart, and Lawrence D. Meyer, *Stereoscopic Television System*, U.S. Patent No.4,523,226, Jun.11, 1985.

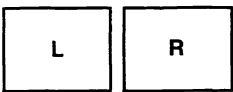
7. Lipton, Lenny, and Lhary Meyer, “A Flicker-Free Field-Sequential Stereoscopic Video System,” *SMPTE Journal*, November 1984, p.1047.

8. Lipton, Lenny, Lawrence D. Meyer, and Frank K. Kramer III, *Multiplexing Technique for Stereoscopic Video System*, U.S. Patent No.5,193,000, Mar.9, 1993.

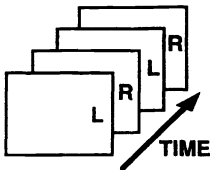
9. <http://evlweb.eecs.uic.edu/pape/CAVE/> (“The CAVE Virtual Reality System”).

10. Lipton, Lenny, Jeffrey J. Halnon, and Lawrence D. Meyer, *Universal Electronic Stereoscopic Display*, U.S. Patent No.5,572,250, Nov.5, 1996.

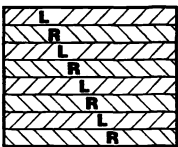
Stereo-vision Formats



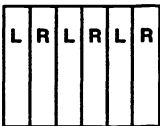
Stereopair. Independent left and right channels. Used in HMDs. Origin may be field sequential with field switch.



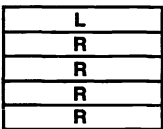
Field-sequential. Alternating left and right fields. Fields may be interlaced or progressively scanned.



Pixel sequential A. Rows of alternating left and right image stripes. Origin may be interlaced field sequential.



Pixel sequential B. Columns of alternating left and right image stripes. Usually an autostereoscopic format.



Segment sequential. Shown is a snapshot with top segment left and bottom segments right. Origin is field sequential.