Kinder Gentler Stereo

Mel Siegel[a], Yoshikazu Tobinaga[b], Takeo Akiya[c]

[a]The Robotics Institute
School of Computer Science, Carnegie Mellon University, Pittsburgh PA 15213 USA
[b]Kansai Research Institute (KRI), Kyoto, Japan
[c]Electronics Intelligence Incorporated, Tokyo, Japan

ABSTRACT

Not only binocular perspective disparity, but also many secondary binocular and monocular sensory phenomena, contribute to the human sensation of depth. Binocular perspective disparity is notable as the strongest depth perception factor. However means for creating it artificially from flat image pairs are notorious for inducing physical and mental stresses, e.g., "virtual reality sickness". Aiming to deliver a less stressful "kinder gentler stereo (KGS)", we systematically examine the secondary phenomena and their synergistic combination with each other and with binocular perspective disparity. By KGS we mean a stereo capture, rendering, and display paradigm without cue conflicts, without eyewear, without viewing zones, with negligible "lock-in" time to perceive the image in depth, and with a normal appearance for stereo-deficient viewers. To achieve KGS we employ optical and digital image processing steps that introduce distortions contrary to strict "geometrical correctness" of binocular perspective but which nevertheless result in increased stereoscopic viewing comfort. We particularly exploit the lower limits of interocular separation ("microstereopsis"), showing that unexpectedly small disparities stimulate accurate and pleasant depth sensations. Under these circumstances crosstalk is perceived as depth-of-focus (blur) rather than as ghosting (doubling). This suggests the possibility of radically new approaches to stereoview multiplexing that enable zoneless autostereoscopic display.

KEYWORDS: stereoscopic displays, 3D displays, 3D-TV, virtual reality sickness, simulator sickness, microstereopsis

1. INTRODUCTION

How geometrically to construct a 3D-stereoscopic camera and display system that reproduces exactly the retinal images of the original scene has been well known for a very long time[1,2]. Almost as old and well known are the many heuristic rules for deviating from this perfect geometry for the sake of mitigating the negative side effects of perfection[2,3]. How can perfection have negative side effects? Because the human eye-brain system uses many cues at many cognitive levels to understand the three dimensional structure of the scene it observes, and even a geometrically perfect 3D-stereoscopic image pair fails to stimulate all of these correctly. The perceptual conflict between the cues that are stimulated correctly and the cues that are not stimulated correctly causes physiological and psychological stresses that have recently come to be known as "virtual reality sickness", "simulator sickness", etc.[4,5]. One well known conflict, for example, is between convergence and accommodation: the eyes converge to a location in front of or behind the screen, as is appropriate for the rendered disparity, but they focus on the screen *per se*. The discrepancy between the state of two sets of muscles whose states are normally tightly coupled causes discomfort. Despite the availability *and routine application* of mitigating heuristics, we often notice that workers at computer monitors equipped with 3D-stereoscopic capability nevertheless spend most of their time looking at the screen either with the stereo turned off (e.g, watching anamorphic images in the above-below or side-by-side format) or with the stereo turned on but not using their eyewear (e.g., watching overlaid left and right "ghost" images) in preference to viewing in stereo; only when they really *must* use stereo depth-disambiguation to understand what they are looking at on the screen do they then briefly turn on the stereo or put on the eyewear.

These observations stimulate us to seek a "kinder gentler stereo" paradigm, a natural and unobtrusive approach to 3D-stereoscopic display of images, video, graphics, and animation that is as free of physical and mental stress is naked eye viewing of the real world. We seek an approach to stereo image capture (and, for computer graphics, stereo image generation) without cue conflicts, without eyewear, without viewing zones, and with negligible "lock-in" time to perceive the image in depth. Another desirable goal is that the display should appear normal to the several percent of otherwise normally sighted individuals who cannot fuse stereo pairs.

In Section 2 we review the many monocular and anomalous binocular cues that can contribute to depth perception. In Section 3 we introduce our experiments with *microstereopsis*, a limiting case of binocular image capture and display that has many of the desirable kinder gentler stereo characteristics we seek. In Section 4 we describe in how microstereopsis enables a radically new approach to engineering zoneless autostereoscopic display systems. In Section 5 we summarize our conclusions and indicate the future work that is needed to fully realize the kinder gentler stereo goal in a practical commercial system.

2. MONOCULAR AND ANOMALOUS BINOCULAR DEPTH PERCEPTION

We begin by examining the known range of phenomena that contribute to the perception of depth in (generally still) flat pictures. While high-level cognitive phenomena (e.g., an anomalously tiny person "next to" an anomalously enormous person is probably actually two people of approximately normal size, one very far away and one very close) are familiar, the existence of five to ten lower-level depth perception-stimulating phenomena is less well known. It seems that, as a general rule, the brain is inclined to infer "depth from distortion".

2.1. A historical survey of "the illusion of depth in single pictures"

This section takes its title from the 1925 article by Ames discussed in the first of the following subsections. As we have noted, it is well known that many cognitive-level phenomena contribute to the perception of depth in visual observation of the 3D world: size, interposition, shading, motion, etc., are all well known high level cues to understanding the depth structure of a 3D scene that augment (or can substitute for) binocular perspective disparity. It is however not so generally well known that there is about as long a list of phenomena that stimulate the perception of depth *when in fact only a single perspective is available*. Some of these phenomena require viewing the picture binocularly through some apparatus or in a particular way. The unifying feature of these seems to be that *some kind* of disparity is created between the left eye's and the right eye's rendering of a single picture; we speculate that the brain's way of resolving any such disparity is to perceive it in terms of depth. Others of these are monocular; in fact, the first one in the first section below is the long-known "Claparade effect": looking at a picture with only one eye makes it seem more three dimensional than looking at it with both eyes. Several of the authors reviewed below plausibly hypothesize that the monocular effects are attributable to the "turning off" of the binocular recognition that the picture is flat, allowing the brain to create the depth dimension based entirely on scene content understanding. Note that in some cases the literature cited does not unambiguously tell whether the phenomena described is monocular, binocular, or both.

2.1.1. Ames (1925)

The first comprehensive discussion we have found on "the illusion of depth in single pictures" is by Ames in 1925[6] in an article with this title. Ames says his work is stimulated by the then-recent work (in German) of Strieff[7], which discusses Claparade's "paradox monokulare Stereoskopie", the greater perception of depth that is observed when a picture is observed with one eye than with two. Ames then goes on to describe several ways to increase the illusion of depth in a single picture. Below we summarize his list, with some comments.

1. Looking at a picture with only one eye. This is the Claparade effect. [monocular]
2. Looking at a picture through an Iconoscope. The Iconoscope is a viewing device that optically reduces perspective parallax. [binocular]
3. Viewing a picture from a greater distance. It is observed that there is a greater illusion of depth in large pictures observed from a large distance (as a movie screen) than in small pictures observed from a small distance. [binocular or monocular?]
4. Changing the convergence of the eyes from that normally required by the distance from which the picture is viewed. Prisms can be used to alter the convergence *either inward or outward*. [binocular]
5. Looking at a picture through a small hole 2 mm or more in diameter held close to the eye. This may be done monocularly or, using both eyes, with the aperture in front of only one eye; the monocular case gives a stronger illusion. [monocular or binocular]
6. Changing the accommodation of the eyes from that normally required by the distance from which the picture is viewed. Lenses can be used to alter the accommodation *either inward or outward*. [binocular?]
7. Looking at a picture binocularly, with one eye receiving a sharp image and the other a blurred one[A]. [binocular]
8. Looking at the reflection of a picture in a mirror. This presumably depends on some imperfection of the mirror, e.g., visible dirt on its surface, as an ideal mirror performs *exactly* a unit magnification projective transformation. [binocular?]
9. Looking at a picture with abnormal rotation of the visual images about the axes of vision. [binocular?]

The common element among all of these phenomena is that viewing conditions that *reduce* the perception of depth in solid (3D) scenes are observed paradoxically to *increase* the perception of depth in flat (2D) scenes[B]. Ames proposed explanation

A This is the basis of "mixed resolution coding" for compressed storage and transmission of 3D-stereoscopic imagery.

B Ames notes, however, that Stief does not consider the Claparade effect (item 1 above) to be paradoxical.

(expressed in modern language) is that interfering with normal binocular stereopsis (with which the picture would be measurably flat) frees the brain to model the 3D world using only the high level cues contained in the image.

2.1.2. Judge (1926)

In his 1926 book[8], Judge discusses “Pseudo-Stereoscopic Effects”. In modern usage “pseudo-stereoscopic” or “pseudoscopic” generally refers to a true stereoscopic pair that has been reversed left for right. However Judge uses the term to refer to various effects that contribute to the perception of depth in single images. At the outset he cautions that “... there are devices available which certainly enable single flat prints to show a partial relief effect ... on the other hand, the usual result obtained resembles that of a series of flat objects, which might be compared to photographs pasted on thin sheets, cut around in silhouette, and placed in different planes, somewhat like stage scenery”.

He notes that a good photographic print, viewed through a large convex lens or by reflection in a concave mirror, will show “a certain amount of relief of the stage scenery type”. He discusses several special lenses (the Zeiss “Verant” is often mentioned by Judge and others of the same period) and commercial devices that exploit this observation. Some use two eyes and one photograph, some one eye and one photograph, and some two eyes separately viewing identical copies of one photograph. He speculates that the perception of depth is due to disparity between the two retinal images caused by viewing-optics aberrations. He also notes that viewing a photograph through a tube that blocks out all light except that coming directly from the photograph enhances the perception of depth in the flat image.

2.1.3. Schlosberg (1941)

In 1941 Schlosberg[9] notes that “in the period around 1910, when interest in stereoscopy was high, it was well known that the ‘plastic’ effect could be obtained almost as well by viewing a single picture through a lens as by the use of disparate pictures in the binocular stereoscope”. Schlosberg also lists several ways in which “‘flatness’ cues can be eliminated, permitting the depth cues to yield a plastic effect”:

1. Looking at a picture from a distance. [binocular or monocular?]
2. Monocular viewing, including three improvements (a) looking through a tube, (b) looking through a lens, and (c) looking at a picture monocularly in a mirror. [monocular]
3. Partial binocular vision, including (a) blurring the image in one eye [see footnote A above], and (b) prisms to displace or rotate the image in one eye. [binocular]
4. Full binocular vision, either through one large lens for both eyes, or through separate lenses, especially of the Zeiss “Verant” design. [binocular]
5. The Iconoscope. [binocular]

These methods are clearly quite similar to those discussed by Ames, but the explanations are more “psychological” and, coming 16 years later, they are expressed in somewhat more modern terms. Schlosberg concludes, similarly to our paraphrasing of Ames conclusion, “monocular plastic depth is due to the release of certain monocular factors from overpowering cues, largely binocular, that show the picture to be flat”.

2.1.4. Motokawa et al (1956)

In 1956 Motokawa et al[10] report on neurological measurements relating to observers’s understanding of monocularly viewed “eclipsed” (occluded) shapes and perspective drawings.

2.1.5. Valyus (1962)

In his comprehensive 1962 book Valyus[11] briefly discusses several “monocular stereoscopic effects”. He defines the term via the remark “... all secondary factors which make for an increased plasticity of the visual image independently of binocular vision may sometimes be united under the common term ‘monocular stereo effect’”. He states “the monocular stereoscopic effect arises as a result of features of perspective”, and notes the contributions of familiarity, shading, aerial perspective, and motion parallax due to intentional head movements.

2.1.6. Schwartz (1971)

In 1971 Schwartz[12] primarily reviews the earlier literature, including the work of Ames and Schlosberg mentioned above. Schwartz points out that in the absence of binocular perspective disparity, *any* kind of disparity between the two retinal images of a single picture will stimulate the illusion of depth in a picture with enough high level cues. Methods include:

1. Luminance disparity, as in the Pulfrich effect[C], and geometric disparity, such as will occur when two pictures from the same perspective but taken at different times, viewed with a stereoscope, reveal scene elements that changed between the

two shots as "floating in space in front or behind" the otherwise flat picture. [binocular]

2. Color, via chromatic aberration, which requires different accommodation for objects at the same depth but with different color. [monocular and binocular]
3. Large magnifier (slide viewer) via luminance disparity and geometrical disparity. Luminance disparity occurs because the light diffuser does not produce a Lambertian illumination pattern. Geometrical disparity occurs because the single large lens system (through which both eyes look) has significant aberrations which produce different distortions at each eye.
4. Stereopsis enhancement factors include sharpness of focus, contrast, vertical contours, and high retinal illumination.
5. Familiarity: people perceive more depth in single pictures of familiar objects and scenes than in pictures of unfamiliar objects and scenes. Related cues include interposition, size of known objects, association with other objects, perspective of parallel lines, vertical position in field, foreshortening due to perspective, distribution of light and shadow, and aerial perspective. [These are of course well known factors, which we repeat here only for completeness.]

2.1.7. Okoshi (1976)

In his comprehensive 1976 book Okoshi[13] briefly identifies accommodation as a 'monocular depth cue'; he says it "is effective only when it is combined with other binocular cues, and for a viewing distance less than two meters".

2.1.8. Aloimonos (1987)

In a series of three technical reports ending in 1987 Aloimonos[14,15,16] proposes a method that can recover absolute depth by a moving observer (following a known path) "without using any kind of correspondence" because "optical flow may be computed with a variational method that does not require correspondence"[D]. He derives the expression

$$Z = \frac{f_x(xW - U) + f_y(yW - V)}{f_x(Axy - B(x^2 + 1) + Cy) + f_y(A(y^2 + 1) - Bxy - Cx) + f_t}$$

where (with all distances relative to the camera focal length) Z is depth from the camera lens, (x,y) are the image coordinates of the point (X,Y,Z) under observation, (A,B,C) is the vector whose cross product with (X,Y,Z) gives its rotational velocity, (U,V,W) is its translational velocity, and (f_x, f_y, f_t) are the spatiotemporal derivatives of the image intensity function $f(x,y,t)$, such that $f_x u + f_y v + f t = 0$ (approximately).

Having derived this relationship, Aloimonos considers first the case of a vibrating camera, then (in the paper's conclusion) the possibility of ocular tremor and optical flow being a source of monocular depth perception in humans. His conclusion is worth quoting in its entirety:

> *We have demonstrated in a theoretical way how a moving observer can recover depth. Our theory, which is correspondenceless, can be applied to the case of a vibrating camera to recover depth.*
>
> *One might imagine that the human eye could take advantage of ocular tremor, in combination with vernier acuity for edge displacements, to achieve "monocular stereopsis". However, the amplitude of the tremor and the focal length of the eye are such that, even if measurement of the tremor were available to the brain, depths could be discriminated only out to rather short distances.*

Unfortunately Aloimonos does not disclose his assumptions about parameter values or his actual calculations in detail, only his conclusion, and he did not respond to several attempts we made to obtain these details via personal correspondence.

2.1.9. Section Summary

We have examined the literature of "the illusion of depth in single pictures" and identified five to ten phenomena, other than binocular perspective disparity conveyed by image pairs taken from separate left and right eye perspectives, that contribute to the perception of depth in pictures. Most of these are described for still pictures, but there is every reason to expect they will be at least as effective with moving pictures. However the Pulfrich effect, described by Schwarz [Section 2.1.6] as a luminance

C The Pulfrich effect may not be an entirely appropriate example, since it is a dynamic phenomena (it works only for movies and animations) whose origin is attributed to a very low level characteristic of the human eye-brain system, i.e., a reciprocal relationship between retinal illumination and transmission delay of the corresponding signal.

D We think this is at best a formal argument, as practically the optical flow field calculation requires correspondence. However this issue does not affect the part of the discussion that is of primary interest to us.

disparity effect, does require motion. Aloimonos [Section 2.1.8] describes an algorithm for extracting the three dimensional structure of a scene from a moving picture; he suggests the possibility that small vibrational motions of the camera, analogous to eye tremor, might be adequate, but he concludes that realistic numerical values for the parameters argue against the existence of this kind of "monocular depth perception" in humans.

At the risk of oversimplifying, we can summarize very briefly:

1. Depth is perceived in a single picture that is viewed monocularly with optics (including field stops) that remove the "flatness cues" that tell the brain that the viewed object (picture) is actually two dimensional.
2. Depth is perceived in a single (sometimes duplicated) picture that is viewed binocularly with optics and illumination that introduce anomalous left eye/right eye disparities.

In both cases we should presume that the "illusion of depth" requires a corresponding "knowledge of depth" that is derived from high-level understanding of the picture's content; in other words, these illusions allow us to see in depth when we don't need stereo to understand the depth content of the picture we are seeing!

2.2. Perspective distortion as a depth illusion stimulating modality

A phenomena that is not mentioned explicitly in the literature that we have encountered but that we would include from our own observations is *perspective distortion*, the distortion of the relative size of closer and farther objects (and parts of objects) when the scene is captured from a longer distance by a longer focal length lens vs. from a shorter distance by a shorter focal length lens. [Valyus (Section 2.1.5) and Schwartz (Section 2.1.6) both mention the contribution of "perspective" to depth perception, but it is unclear whether they are referring specifically to perspective distortion or rather to the better known (and apparently different) perception of, e.g., parallel lines appearing to meet at the horizon.] Perspective distortion is illustrated in Figure 1, showing a scene photographed by telephoto, normal, and wide angle lenses at different camera-to-subject distances:,

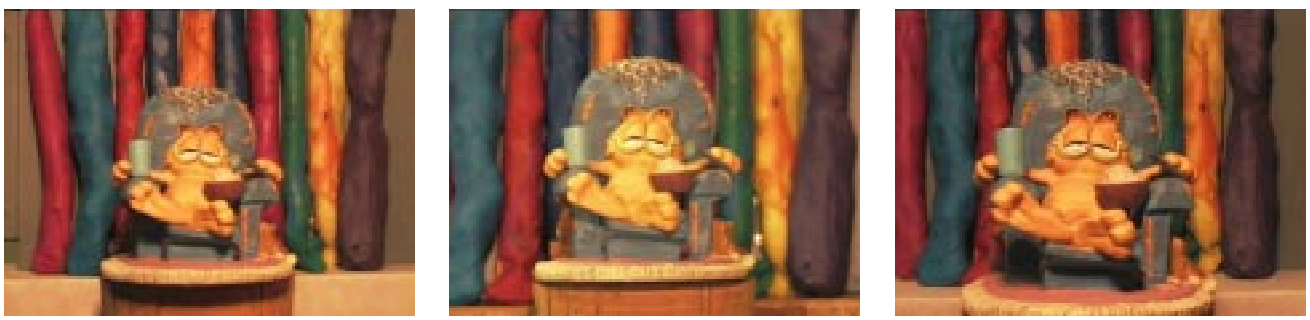

Figure 1. Pictures of the same scene with telephoto (left), normal (middle), and wide-angle (right) lenses at corresponding long, intermediate, and short distances from camera to scene. Notice "flatness" at left, "depth" at right.

Notice the "flatness" of the left image vs. the "depth" of the right image due to the exaggerated largeness of the nearer features (like the feet) in the right image. Perspective distortion is often exploited artistically by photographers and cinematographers; for example, portrait photographers typically use a longer-than-normal focal length lens to make the subject's nose look smaller and to make it seem to protrude less from her face.

2.3. Motion parallax ("depth-from-motion")

By far the most important depth cues in the real live word, after binocular stereopsis, are the "motion parallax" cues related to the relative motion and the internal dynamics of the objects in the scene, plus the motion of the viewer (or camera) with respect to the scene. In the artificial worlds of cinema and video, excepting only the relatively few 3D-stereoscopic productions that

Figure 2. Motion parallax ("depth-from-motion") illustrated by an animated GIF file that can be viewed at http://www.cs.cmu.edu/afs/cs/user/mws/ftp/spie99/Figure_2.gif.

have been made, binocular perspective disparity is absent and motion parallax becomes probably the most important cue. Even complex wireframe drawings, which in a still picture look like a confused jumble of intersecting lines in a plane, acquire depth and solidness the instant they are put into rotational or translational motion. What we might call "artificial motion", e.g., vertical oscillations of the camera at several cycles per second, have on several occasions been used as to create an illusion of "3D-TV" within the unaugmented monoscopic TV infrastructure[E]. Figure 2 is an animated GIF file that illustrates the power of motion to stimulate 3D-ness; the animation can be viewed by opening http://www.cs.cmu.edu/afs/cs/user/mws/ftp/spie99/Figure_2.gif with a web browser.

3. MICROSTEREOPSIS

So it seems clear that if a scene contains enough familiar detail that its depth content can be deduced by high-level reasoning (vs. triangulation based on the binocular perspective disparity between corresponding points in a stereopair), then under appropriate viewing conditions even a flat picture of this scene will stimulate the illusion of depth. This observation leads to the following hypotheses:

1. If a scene contains enough familiar detail that its depth content can be deduced by high-level reasoning, then under "ordinary" viewing conditions depth perception can be stimulated by binocular perspective disparity that is smaller than the disparity demanded by "geometrical correctness" in the sense of Reference 1.
2. Smaller-than-"correct" disparities stimulate smaller-than-"typical" measures of physical and mental discomfort.
3. Disparity reduction by left/right shifting of the members of a stereopair (equivalent to converging the cameras inward from their "correct" parallel axes geometry and rectifying the resulting keystone distortion) is effective.
4. Disparity reduction by reducing the interocular separation to a value smaller than the human interocular separation (which is the required camera separation for geometrical correctness) is also effective.
5. Left/right shift and interocular separation reduction are complementary: they are especially effective when used together.
6. The combination of reduced interocular separation and other depth perception stimulating factors, especially motion parallax, are also complementary.

We have tested these hypotheses in a series of informal experiments in which the authors, their colleagues, visitors to their labs, prospective supporters of ongoing research, etc., have been shown various implementations and their perceptions queried and noted. Formal human factors experiments with the necessary controls, quantitative data recording and statistical analysis, etc., have not yet been conducted. However we believe with high confidence that the informal results reported *and demonstrated* in the following subsections will be fully validated by future measurements made under a strict protocol.

These preliminary experiments demonstrate that disparities that are so small that they are perceived (when overlaid on one screen and viewed without stereo-viewing eyewear) not as ghosting but instead as blurring. This occurs at interocular separations that are only a few percent of the normal human interocular separation. If left-right shifting is used to make the disparity zero in the vicinity of the "center-of-interest" of the scene, then the blurring occurs in the foreground and background regions, and it is interpreted as the blurring expected because of the lens's limited depth-of-focus rather than as a fault. These pictures look "normal" when viewed without stereo-viewing eyewear; when stereo-viewing eyewear is put on, the depth appears immediately and naturally, with no (informally) observable "lock-in" time to fuse the left and right images.

We call depth perception stimulated by anomalously small disparities "microstereopsis"[F].

E People who have seen this effect often report a sensation of 3D-ness followed by symptoms of motion sickness.

F The term "microsteroscopy" has also been used for photography through a stereo microscope[17], but confusion seems unlikely.

3.1. "Center-of-Interest Correction"

Figure 3, Figure 4, and Figure 5 illustrate the effectiveness of "center-of-interest correction" in reducing the difficulty of viewing this scene stereoscopically. Figure 3 and Figure 4 are stereo pairs taken under conditions that were held as identical as possible; the only difference is that in Figure 4 the images have been shifted left/right to bring corresponding points in the vicinity of the center-of-interest into coincidence. These pairs can be viewed stereoscopically by "free viewing". It should be apparent that the exact circumstances that make it somewhat *easier* to free view Figure 3 than Figure 4 (Figure 3 is viewed with the eyes converged, Figure 4 with the eyes parallel) make it substantially *harder* to view Figure 3 than Figure 4 when these images are viewed overlaid (e.g., temporally multiplexed) at the greater distance of a workstation CRT or projected onto a distant screen. Figure 5 provides gray level representations of the disparities in Figure 3 and Figure 4.

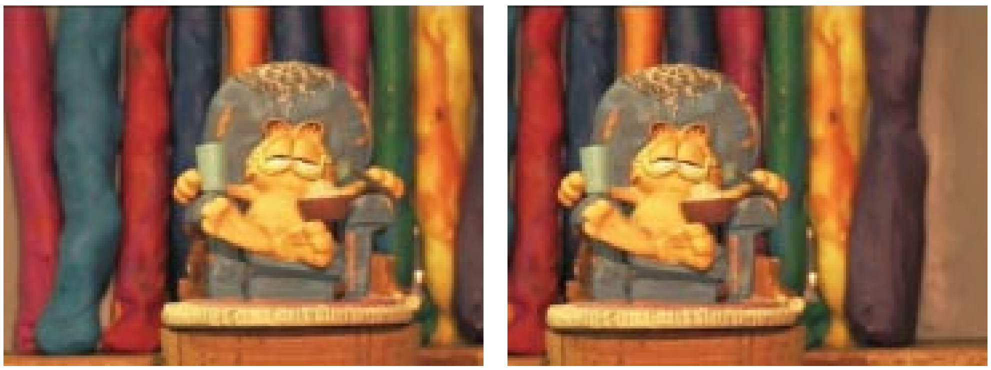

Figure 3. Left/right stereo pair, 1 m camera-to-scene, 50 mm lens, parallel axes, 40 mm interocular separation. Can be comfortably fused by "free viewing" with 65 mm on-page interocular separation, but disparities are uncomfortably large on a CRT or a distant screen.

Figure 4. Left/right stereo pair, 1 m camera-to-subject, 50 mm lens, parallel axes, 40 mm interocular separation, center-of-interest compensated. In contrast to Figure 3, this pair can be fused only with difficulty on-page, but very comfortably on a CRT, or projected on a screen.

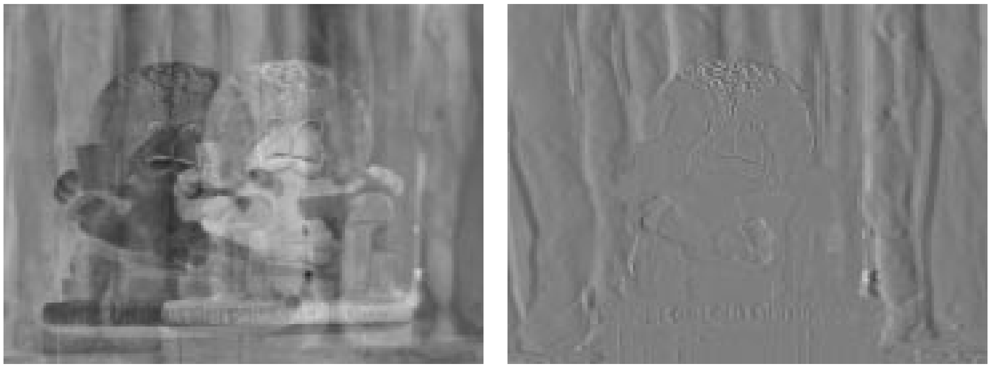

Figure 5. Gray level representations of disparities between the image pairs of Figure 3 (left above) and Figure 4 (right above). When corresponding pixels are identical this representation is midrange gray; negative (left-right) differences are darker, positive differences lighter.

3.2. Very small interocular separation

Figure 6 and Figure 7 illustrate the effect of using a very small interocular separation, only a few percent of the normal 65 mm human interocular separation. Depth perception is enhanced by using a wide angle lens and taking advantage of the resulting perspective distortion, but binocular stereopsis is also adequately stimulated when the pictures are taken with telephoto lenses. Figure 6 can be free viewed reasonably comfortably, and the pair is very easy to view multiplexed on a CRT or a screen. The gray level Figure 7 shows how small the left/right disparity is, even in comparison to the already small disparity in Figure 4

Figure 6. Microstereopsis illustrated by interocular separation ~3% of normal. The effect is visible with the telephoto lens used in the previous figures, but is enhanced here by also using a 12.5 mm wide angle lens and a correspondingly closer camera-to-subject distance.

Figure 7. Gray level representation of the difference between left and right views in Figure 6.

4. ZONELESS AUTOSTEREOSCOPIC DISPLAYS

The bane of conventional 3D-display systems is crosstalk, the failure of the multiplexing system to block all of the right image content from the viewer's left eye and all of the left image content from the viewer's right eye. When disparities assume their typical values, crosstalk is perceived as ghosting: a (hopefully) dim "wrong eye" image is superimposed on a (hopefully) bright "correct eye" image. Since viewers find ghosts annoying, one of the major challenges of 3D-display system design is to minimizing ghosting. The conventional way to achieve this has been to concentrate effort on multiplexing hardware that minimizes crosstalk.

As illustrated below by Figure 8, with microstereopsis, crosstalk is perceived not as ghosting but as blur, and in fact, as a normally acceptable kind of blur, the blur associated with the lens's finite depth-of-focus. This suggests that microstereoscopic imagery could be displayed on hardware that permits more crosstalk than is acceptable with conventional stereoscopic imagery. This idea has been tested by viewing temporally-multiplexed conventional and microstereoscopic imagery through LCD shutter hardware that was modified to allow controllable degradation of its on/off transmission ratio. Although to date the experiment has been conducted only informally with a few colleagues as subjects, the hypothesis appears to be correct.

Figure 8. (left) Crosstalk perceived as "ghosting" illustrated by averaging the two views in Figure 4. (right) Crosstalk perceived as "blurring" illustrated by averaging the two views in Figure 6. Most viewers find ghosting objectionable, but consider blurring a normal phenomena.

Freeing the 3D-display designer of the overriding requirement to achieve low crosstalk presents the opportunity to propose radically new multiplexing methods. To offer but one simple example, a uniform angular shading gradient that resulted in a slight left eye/right eye luminosity differential might be sufficient to convey enough binocular perspective disparity information to stimulate stereopsis. This display would be zoneless (correctly ordering the perspectives from any viewing angle) and autostereoscopic (requiring no eyewear, headtracking, etc.).

We will discuss implementation details and design principle alternatives in future publications.

5. CONCLUSIONS

In pursuit of "kinder gentler stereo", a depth-displaying paradigm that would overcome the negative aspects of conventional left/right multiplexed binocular image pair display systems, we have surveyed the literature that describes the perception of depth by means other than binocular perspective disparity. We have found and summarized references to five to ten apparently low-level effects in addition to the well known high-level effects of size, interposition, etc. This survey leads to the conclusion that the brain is able to create the perception of depth if the scene is sufficiently complex that knowledge of depth can be inferred from high-level understanding of the scene's content. This suggests that, in conventional left/right multiplexed binocular image pair display systems, disparities that are much smaller than conventional might adequately stimulate depth perception but not "virtual reality sickness", etc. Our (still informal) experiments support the hypothesis that "microstereopsis" does provide a low stress alternative to 3D-displays that rely on left/right disparities that correspond to the normal human interocular separation. Synergistic combination of microstereopsis with motion parallax, perspective distortion, and other "single image" depth perception stimulating effects provides the sought after "kinder gentler stereo". We suggest, and demonstrate in a preliminary experiment, that in a 3D-display system based on microstereopsis, crosstalk between the multiplexed left and right eye channels is negligibly objectionable, in contrast with its being extremely objectionable in conventional 3D-display systems. This suggest several possibilities for engineering zoneless autostereoscopic 3D-displays.

ACKNOWLEDGEMENTS

We thank: H. Samura and Y. Nakazono of Kansai Research Institute (KRI) for supporting the study of monocular depth perception (MDP) which stimulated the ideas for kinder gentler stereo (KGS) and the exploitation of microstereopsis that are described in this paper; S. Nagata of NHK and Intervision for numerous valuable discussions and suggestions, and for sharing with us his human factors data on motion parallax perception; Jeff Halnon of Stereographics Corporation for teaching us how to modify the Stereographics SimulEyes LCD driver circuit to controllably degrade its on/off transmission ratio; and Alan Guisewite for data collection and other assistance in the laboratory.

REFERENCES

1. V. S. Grinberg, G. W. Podnar, M. W. Siegel, “Geometry of Binocular Imaging”, SPIE/IS&T Stereoscopic Displays and Applications V, San Jose, February 1994, pp 56-65.
2. A. C. Hardy and F. H. Perrin, “The principles of optics”, New York, London, McGraw-Hill, 1932 (1st ed.).
3. L. Lipton, “The CrystalEyes Handbook”, Stereographics Corporation, 1991. Also available on web page links at http://www.stereographics.com/documents.html.
4. P. A. Howarth, Visual Ergonomics Research Group, http://info.lboro.ac.uk/departments/hu/groups/viserg/viserg1.htm.
5. E. M. Kolasinsky, Simulator Sickness in Virtual Environments, http://www.cyberedge.com/4a7a.html.
6. A. Ames, Jr., “The Illusion of Depth from Single Pictures”, J. Opt. Soc. Am., V.10, pp.137-148 (1925).
7. Strieff, “Die binokulare Verflachung von Bildern, ein vielseitig bedeutsames Sehproblem”, Klinische Monatsblatter fur Augenheilkunde, V.70, p.1, (1923).
8. A. W. Judge, “Stereoscopic Photography: Its Application to Science, Industry, and Education”, American Photographic Publishing Co, Boston, 1926, pp. 137-139.
9. H. Schlosberg, “Stereoscopic Depth from Single Pictures”, Amer. J. Psychol., V.54, p.601-605 (1941).
10. K. Motokawa, D. Nakagawa, and T. Kohata, “Monocular stereoscopic vision and gradients of retinal induction”, Journal of Comparative and Physiological Psychology, V.49 p.4, 1956.
11. N. A. Valyus, “Stereoscopy”, The Focal Press, London, 1966. [English translation of 1962 edition in Russian.]
12. A. H. Schwartz, “Stereoscopic Perception with Single Pictures”, Optical Spectra, September 1971, p.25-27.
13. T. Okoshi, “Three-Dimensional Imaging Techniques”, Academic Press, New York, 1976 [a translation and extension of the Japanese “Sanjigen-Gazo Kogaku”, Sangyo-Tosho, Tokyo, 1972], p.29.
14. J. Aloimonos, “One Eye Suffices: A Computational Model of Monocular Robot Depth Perception”, University of Rochester Department of Computer Science TR 160, December 1984.
15. J. Aloimonos and P. Chou, “Detection of Surface Orientation and Motion from Texture: 1. The Case of Planes”, University of Rochester Department of Computer Science TR 161, January 1985.
16. J. Aloimonos and A. Rozenfeld, “On Monocular Stereopsis: A Note”, University of Maryland Department of Computer Science Technical Report CAR-TR-218, CS-TR-1698, February 1987. [This report repeats most of the previous two, and also contains some new material of particular interest to us.]
17. J. G. Ferwerda, “The world of 3-D: A practical guide to stereo photography”, 3-D Book Productions, Harry zur Kleinsmiede, The Netherlands, 1987 (2nd edition).