Visual Comfort of Binocular and 3-D Displays

Frank L. Kooi
Alexander Toet
TNO Human Factors, Soesterberg, The Netherlands

INTRODUCTION

Binocular display systems are nowadays found throughout our society: in outdoor settings, office settings, and in the entertainment industry as well. The military deploy head-mounted displays in aircraft cockpits and various types of vehicles. 3-D displays are being used to gain a better understanding of the 3-D layout of complicated structures (e.g., molecules, seismic data, and anatomical data), virtual environments, and real environments (e.g., telerobotics). However, binocular display systems are still not widely used. A main reason is the viewing discomfort caused by binocular imperfections: differences between the left- and right-eye images. In stereo vision systems, eyestrain usually occurs to some extent, caused by inevitable left-eye/right-eye mismatches. Eliminating eyestrain in head-mounted displays greatly drives up the cost of the optical system. Being able to *predict* the level of visual discomfort from the specification of binocular viewing systems greatly helps the design and selection process. This article provides the basis for that.

OVERVIEW

It is well known that the human visual system is very sensitive to left-eye/right-eye differences. The visual discomfort experienced 1) when switching to a new pair of glasses, 2) while pressing lightly against one eye, or 3) when wearing multifocal lenses are examples of this sensitivity. Research on binocular discomfort typically addresses a particular type of device such as head mounted displays (HMDs)[1,2] or stereoscopic displays.[3,4] As a result, the impact on visual comfort of spatial distortions,[5] stereoscopic disparity,[6] luminance asymmetry,[7] and other factors[2,8] has been investigated separately. To our knowledge, no previous study has attempted to directly compare the relative impact of the various binocular factors that affect visual comfort. This article aims to provide such an overview.

First we give a short overview of the human binocular visual system followed by a brief overview of the major categories of binocular and stereoscopic display technologies. Then we present a study we performed to assess the relative impact of the various binocular factors that affect visual comfort. Finally, we will present the conclusions of this study pertaining to binocular viewing devices.

THE VISUAL SYSTEM: HOW DEPTH PERCEPTION WORKS

Binocular Vision: The Benefits of Having Two Eyes

Humans have two eyes rather than one for several (excellent) reasons. The second eye functions as a spare part, the combined field of view is larger, visual acuity is better, and depth perception is much better in specific circumstances. Binocular depth perception only functions properly when the left- and right-eye images match closely. If both images mismatch for some reason, the binocular advantage can easily turn into a binocular nightmare. Extreme discomfort can cause pain, which may lead to the active suppression of the image from one eye. This may eventually lead to the development of amblyopia.

Depth Cues

Our two eyes make 2-D images of the world, similar to photographs. The 3-D structure of the world around us therefore needs to be interpreted (mentally reconstructed) from the image pairs of the left and right eyes. This interpretation takes place in the visual part of the human brain. 3-D vision can therefore be disturbed both by a problem with the eyes (unsuitable input imagery) or by a problem in the brain (erroneous interpretation). The brain uses a number of ''tricks'' to make a 3-D interpretation of the 2-D input from the two eyes, commonly referred to as ''depth cues.'' The depth cues are essential for an understanding of 3-D technology. They can be divided into the four categories shown in Table 1. However, the transformation of the 2-D input into a 3-D interpretation is not always correct. Many visual illusions are indeed the result of an erroneous 3-D interpretation. Figure 1

DOI: 10.1081/E-EOE 120024169

Table 1 The four main categories of depth cues

	Depth cue	Parameter (unit)
1	Stereopsis	Binocular disparity (°)
2	Accommodation	Optical power (diopters)
3	(Motion) parallax	Relative position (°)
4	Pictorial depth cues	—

illustrates how the addition of some cast shadows and occlusion to a 2-D image creates a visual depth illusion.

Convergence

Stereopsis is the result of viewing with two eyes rather than with one.[9] The subtle differences in object position caused by the slightly different location of the left and right eyes are used by the visual system to see depth. The nearer an object is to the eyes, the more different its two images (retinal projections) are. From the difference between the two images, the brain can calculate the distance of each object in the scene. When viewing an object that is close by, the eyes turn inward to fuse the object. This process is called convergence. Turning the eyes outward is called divergence. The visual system is able to convert this stereoscopic disparity to a perception of depth.[9]

Accommodation

Accommodation is necessary because the eyes have a limited depth-of-focus. At any point in time, only one distance is truly seen as sharp; everything in front and everything behind is blurred to some extent. This helps us to focus visual attention on the object of interest because the rest of the world is (slightly) blurred (Fig. 2). Vergence is neurologically coupled to accommodation. When converging, the eyes also accommodate; when diverging, the eyes relax the accommodation. The reverse is also true: when the eyes accommodate, they also tend to converge. This coupling is very convenient because it helps to prevent diplopia (seeing double) and blur. Optometrists define the ''zone of clear binocular vision.'' A mismatch between the viewing distance and the binocular viewing angle (the required accommodation and the required vergence) leads to blur or double vision. The visual system cannot ''line up'' the eye vergence and the eye accommodation. Melzer[10] claims that this starts to happen when the two differ by more than 1/4 diopter (D). This is the difference between 80-cm and 100-m viewing distance. The coupling between accommodation and vergence therefore can be very tight, although our own experience shows that if allowed to adapt, subjects can cope with much larger differences.

(Motion) Parallax

Moving the head sideways or up/down has two effects:

1. It provides a depth percept during the motion from the optic flow.
2. It provides different points of view after the head motion is stopped.

The first (dynamic) effect is analogous to stereopsis. The second (static) effect is demonstrated in Fig. 3 where part of the background objects can only be seen from the right point of view. Moving the head sideways is a natural part of our behavior to obtain a better view and usually does not require a conscious effort.

Pictorial Depth Cues

Pictorial depth cues provide a depth perception when viewing the world with one eye closed and the other kept perfectly stationary. Pictorial depth cues are therefore also referred to as monocular depth cues. The depth that can be perceived in photographs is based exclusively on these cues. Examples of monocular depth cues are perspective, occlusion, and shading. The depth illusion shown in Fig. 1 includes all three of these cues. While pictures can convey a strong sense of depth, it is also easy to find scenes whose

Fig. 1 A depth illusion showing how easy it is to fool the visual system. Simply adding some cast shadows and occlusion to a 2-D image already does the job. (*View this art in color at www.dekker.com.*)

(a) (b)

Fig. 2 Illustration of the principle of accommodation. In an image taken with a small aperture (a), it is easier to ignore the (unsharp) background than in the image taken with a wider aperture (b). (*View this art in color at www.dekker.com.*)

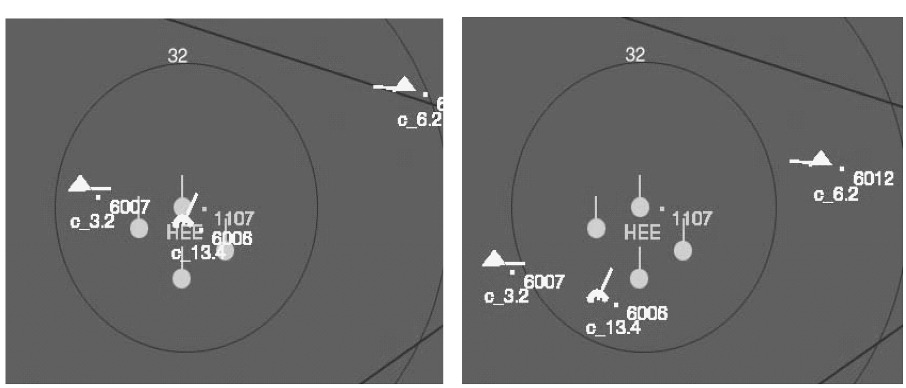
3-D layout cannot be perceived from a picture, whereas the 3-D structure immediately becomes apparent when stereo and/or parallax are available.

THE TECHNOLOGY OF STEREOSCOPIC DISPLAYS

Simply stated, 3-D technology adds the sense of depth by imitating one or more of the visual depth cues. Stereoscopic displays present each eye with its own image, mimicking the subtle differences in object position caused by the slightly different location of the left and right eyes. Here we briefly describe how the 3-D technologies achieve this result.

Stereoscopic Disparity

Stereoscopic disparity can be activated by presenting (slightly) different images to the left and right eyes. The most common stereoscopic display methods employ shutter glasses, polarized glasses, red/green glasses, and head-mounted displays. In most of these systems, the light coming from the display is split in two, either in time, in polarization, or in color. Head-mounted display systems generally use two displays, one in front of each eye (e.g., Fig. 4). These devices share the common disadvantage of constraining the user. For example, eye contact is disturbed, hampering communication with others. These devices are therefore not well suited 1) to work with for long periods of time or 2) for people who want to visually communicate with others.



(a) (b)

Fig. 3 Illustration of parallax. (a) The yellow foreground objects occlude the blue background objects. (b) A (head) movement to the right and upward makes all objects clearly visible, simulating the ''de-cluttering function'' of parallax. (*View this art in color at www.dekker.com.*)

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Fig. 4 Head mounted display. (*View this art in color at www.dekker.com.*)

Accommodation and Parallax

These 3-D glasses simulate the convergence depth cue but do not provide accommodation and parallax, which means that the depth percept is incomplete. Parallax can be added by tracking the head movements and adjusting the viewpoint accordingly. However, even with a fairly powerful computer, a time delay between head movement and image adjustment remains noticeable. Except for occasional prototype 3-D displays,[11,12] the accommodation cue can only be added by imaging the scene at physically different distances. Research[13] has shown that the depth percept immediately ''pops out'' when parallax and accommodation cues are present. If convergence (stereopsis) is the only depth cue, the depth percept requires some amount of time to build up. Figure 5 shows how large the perceptual time delay can be when objects are located in front of each other, i.e., in the situation in which depth is most important. Second, thanks to the parallax, occlusion of one object by another can easily be eliminated by moving the head sideways or vertically (Fig. 3). This is important if two objects are located in the same viewing direction but at different depth planes. Accommodation and parallax therefore have a large influence on the experienced quality of 3-D displays.

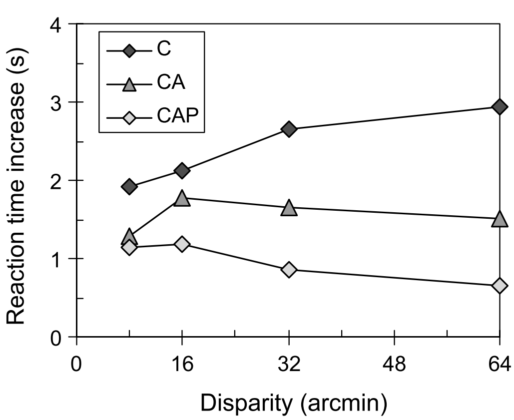


Fig. 5 Data showing how accommodation (A) and motion parallax (P) influence the ease of depth perception when the depth gradient is large. Shown is the extra time required to perceive the depth relationship of two adjacent dots when a distracting object is added at a different depth. The horizontal axis contains the amount of depth difference. The increase in reaction time caused by the distractor is 1–2 sec greater for the common type of 3-D displays (C: convergence cue only) than for transparent depth displays (CAP: accommodation and parallax as well as the convergence depth cue). These results imply that transparent depth displays are more natural to view than the standard 3-D displays described in ''The Visual System: How Depth Perception Works'' and particularly suited for cluttered 3-D imagery. (*View this art in color at www.dekker.com.*)

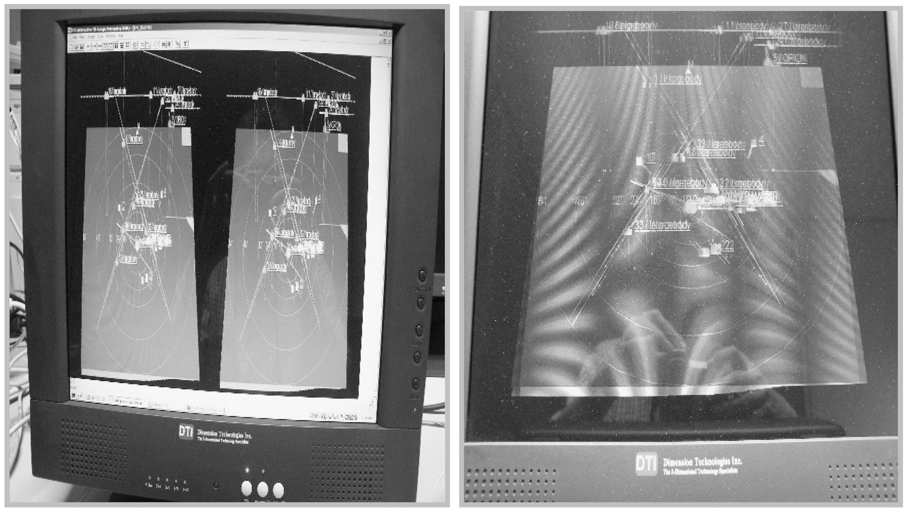
THE TECHNOLOGY OF AUTOSTEREOSCOPIC DISPLAYS

Autostereoscopic displays provide a 3-D percept similar to that produced by a stereoscopic display but do not require the use of stereoscopic glasses or other optical aids (Table 2). In the literature, 3-D displays are sometimes classified as stereoscopic, autostereoscopic, and true 3-D displays. The latter class includes what we term (motion) parallax, which is sometimes also called ''look around capability'': the ability to look at the object or scene from different viewpoints by moving sideways. The three major autostereoscopic techniques are lenticular (Fig. 6), holographic, and transparent.

Table 2 Optical aids that are required to view stereoscopic displays and their particular disadvantages

	Optical aid	Disadvantages
1	Shutter glasses	Low luminance, flicker in daylight environments
2	Polarized glasses	Need to keep the head straight up
3	Red/green glasses	No color vision, chromatic aberration, cross talk
4	HMDs	Image moves with the head, cables, or weight

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016



(a) (b)

Fig. 6 The individual left and right monocular images (a) and the stereoscopic presentation (b) of a lenticular autostereoscopic 3-D display. The stripes superimposed on the stereoscopic image occur because the camera is purposely placed in front of the correct viewing location. (*View this art in color at www.dekker.com.*)

Lenticular Screen Displays

Autostereoscopic displays based on a lenticular screen contain an optical layer glued to the screen that directs the light from each pixel to either the left or the right eye. To achieve this effect, the illumination coming from behind is limited to vertical lines that line up with the optical layers. This is typically called ''parallax illumination.'' To view the image correctly, the two eyes need to be at the right location in space, typically 60 cm from the screen.

Because the image needs to be split up between the left and right eyes, the stereoscopic resolution is lower than the monocular resolution. The light from each pixel either goes to the left eye or to the right eye. The luminance of a lenticular 3-D display tends to be lower than the luminance of a conventional display because the light needs to be funneled through the lenticular lenses. This makes the lenticular 3-D display more susceptible to annoying screen reflections than a standard 2-D display. Focusing the light precisely on the target eye is impossible because some of the light always leaks to the other (wrong) eye. Because of these limitations, lenticular 3-D displays are not well suited to view for long periods of time and an easy procedure to find the optimal head position is a must.

Holographic 3-D Displays

Holographic displays make use of coherent light to create an interference pattern that is identical to the light coming from real 3-D objects. The image contains parallax and accommodation besides stereo. Holography therefore provides, in principle, the most powerful tool to create 3-D displays. The technology is very complex, however, and dynamic 3-D holographic displays are something for the future.

Transparent Depth or Multilayer Displays

A relatively simple way to include accommodation and parallax in the depth percept is to optically superimpose two or more image slices representing the scene at different depth planes. Such a transparent display presents ''true depth'' in the sense that stereoscopic convergence, accommodation, and parallax are all present. The user does not pay a price in terms of viewing comfort as is the case with the other 3-D displays, especially the autostereoscopic ones. Our laboratory transparent display setup is shown in Fig. 7. So far, not much research has been performed on transparent displays. This is probably because the limited number of depth planes makes them unsuitable for the display of 3-D pictures and videos.

VISUAL DISCOMFORT RESULTING FROM BINOCULAR ASYMMETRIES

To assess their relative contribution, we measured the effect on visual comfort of each type of binocular asymmetry that may occur in practice, except for an

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

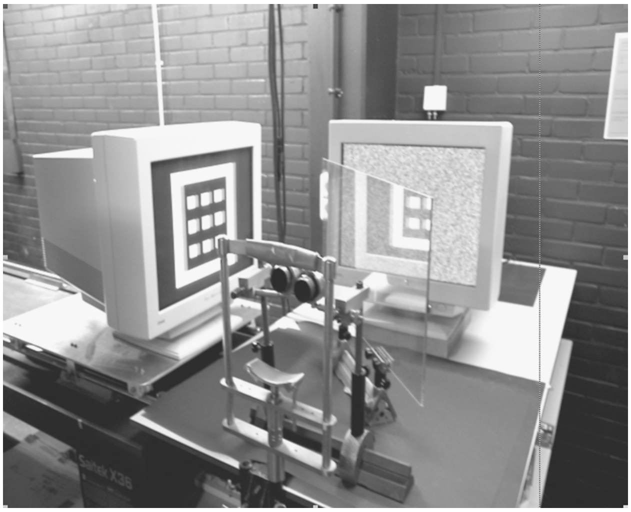


Fig. 7 Experimental transparent 3-D setup in which two images are combined with a half-silvered mirror. Because the light from the two displays adds up, we call it an additive transparent display. (*View this art in color at www.dekker.com.*)

optically induced accommodation difference. In the following, we briefly describe the experiment and the results. For details about the experimental setup and procedures, see Ref. [14].

Types of Binocular Asymmetries

Visual comfort of display systems can be seriously reduced by many factors including jitter, flickering, image motion, and poor resolution. In this study, we exclusively focus on visual discomfort that is caused by a difference between the left and right images of a stereo pair, i.e., *binocular asymmetry*. Three groups of binocular asymmetries may be encountered in practice, depending on the hardware of the viewing system. If *optics* are used, the geometry of the left and right images may differ by a shift, rotation, magnification, or by reduced resolution (image blur). If *filters* are used, the photometry of the left and right images may differ in their luminance, color, sharpness (blur), contrast, accommodation (because of chromatic aberration), or cross talk (the left-eye view leaks through to the right-eye view and vice versa). Third, if a *stereoscopic image* is presented, the stereoscopic disparity may cause an accommodation–convergence mismatch and a motion parallax–convergence mismatch unless the depth planes are optically located at the correct viewing distance[3,4] which is rarely the case.[13]

Another source of viewing discomfort associated with stereo display systems is cross talk. In a perfect stereoscopic display, the right-eye image should be seen by the right-eye only and should be completely invisible to the left-eye, and vice versa for the left-eye image. However, complete separation of the left- and right-eye images is often impossible in 3-D display systems. Interocular cross talk is typically in the range of 0.1–0.3% with polarization techniques and in the range of 4% to >10% with time-multiplexed 3-D displays.[15] Cross talk produces double contours (ghosting) and is a potential cause of eyestrain and headaches.[4,16] Pastoor[4] established visibility thresholds for cross talk using a system with perfect left/right image separation (a high-resolution mirror stereoscope; the experimental conditions were generated with image processing techniques). He found that visibility of cross talk increases (i.e., the threshold value is lowered) with increasing contrast and increasing binocular parallax (depth) of the stereoscopic image. To reproduce a reasonable depth range (up to 40 minarc) on a high-contrast display (100:1), cross talk should be as low as 0.3%.

A related issue of interest to optical designers are the tolerance limits for misalignment of imagery viewed against a background.[1] These are usually associated with see-through systems because these use a partially transparent mirror to superimpose the imagery on the world behind. Figure 8 shows an example of a see-through HMD. Two types of problems come into play specific to the see-through aspect: 1) a misalignment between the image and the real world; and 2) an accommodation difference between the image and the real world. The alignment of a see-through system needs to be extra accurate because any misalignment between the projected image and the real world is highly visible and cannot be adapted to. One image provides a reference to the other meaning that a misalignment causes one of the two images to be seen as double (diplopic).[1] The same argument holds for a convergence–accommodation mismatch in the displayed image. With the real world as reference, the

Fig. 8 The TopOwl see-through HMD, designed for helicopter piloting. (*View this art in color at www.dekker.com.*)

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

viewer does not get the chance to adapt to the mismatch. The high cost of a well-aligned optical system has spurred research into this topic.[1] Rash and McLean[17] recommend not to use symbology on a helicopter see-through head-mounted display to avoid these mismatches. A typical value for vertical misalignment is 5 arcmin. For a horizontal misalignment, typical values are 10 arcmin when the image is converged (requiring the eyes to turn in) and 4 arcmin when the image is diverged.

Experiment

Stimuli and setup

In this study, we used stereo images that represent a part of a typical office scene with two dominant depth planes (Fig. 9). The overall scene contains both highly structured and uniform areas.

A computer was used to present the stimuli and collect the observer responses. The left and right-eye images were displayed on a projection screen by two full color LCD projectors. Corresponding left- and right-eye images were projected simultaneously onto the projection screen. Both projectors were equipped with linear polarization filters. The filters were oriented such that the polarization of both images was orthogonal. The subjects viewed the images through two polarization filters: one in front of each eye. The polarization directions of these filters were matched to those of the image projections, such that each eye viewed the corresponding image from the stereo pair.

The reference, normal stereoscopic, and hyperstereoscopic images were also presented on an LCD monitor equipped with an autostereoscopic lenticular screen. The advantage of this type of display is the fact that an observer does not have to wear any optics to obtain a stereo percept because the technique to separate both views is integrated in the display. Lenticular screens inherently produce a large amount of cross talk.[2] Subjects that experience reduced viewing comfort because of cross talk on the projection screen should therefore also consistently give lower-quality ratings to the autostereoscopic screen.

All experiments were performed in a dimly lit room. Twenty-four subjects participated in the experiment.

Binocular image manipulations

The original image was transformed to create a set of image pairs corresponding to a range of well-defined binocular asymmetries (i.e., well-specified differences between the left- and right-eye images). The transformations that were applied include rotations, scaling, and deformation operations. Some of the binocular asymmetries involve a combination of two different types of these image transformations. Table 4 gives an overview of the 35 different image transformations employed in this study. In the following, we briefly describe these transformations in the order of their occurrence. The transformations were the following (the numbers in

Fig. 9 Two projectors (upper left) equipped with linear polarization filters project the corresponding left- and right-eye images simultaneously on the screen. The subject views the screen (170 × 128 cm) from 185-cm distance through two polarization filters such that the left and right eyes see a separately projected image. With this setup, nearly all binocular errors can be simulated. (*View this art in color at www.dekker.com.*)

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

parentheses refer to the corresponding transformations in Table 4):

A 1° image rotation was created in one of several, slightly different, ways:

1/2° excyclorotation of the two images (1).
1° incyclorotation of one image (i.e., the image to one eye) (2).
1° incyclorotation of one image with border (3).
1° excyclorotation of one image (4).
1/2° incyclorotation of the two images[a] (5).

The other distortions employed were:

1.5% and 2.5% magnification of one image (6, 9).
3% meridional magnification (astigmatic effect) in the horizontal and vertical directions (8, 10).
A trapezoidal distortion which vertically shrinks one side of the image by 1 PD (0.57°)[b] (7).

A binocular asymmetric shift was introduced as follows:

Horizontal by 2 and 3 PD (11, 13).
Vertical by 1 and 2 PD (12, 14).

Three levels of binocular disparity were presented (leading to stereo vision) corresponding to normal, 2×, and 4× hyperstereo (conditions 15, 16, and 17 in Table 4).

Cross talk results in a transparent overlay of the left-eye image over the right-eye image or vice versa and is perceived as ghost, shadow, or double contours. Even small amounts of cross talk can lead to headaches.[4] We deployed three levels of cross talk, corresponding to 5%, 15%, and 25% (conditions 18, 19, and 20 in Table 4). Because cross talk can only be perceived when the left and right images are different (i.e., when the contours in both images do not completely overlap), a small (1 PD) horizontal shift was added. To further investigate the effects of cross talk, the reference image (condition 33 in Table 4) and two stereodisparity levels (conditions 34 and 35 in Table 4) were also presented on the autostereoscopic display.

To investigate the interaction with other types of binocular differences, we combined the following three image manipulations:

1. Normal stereo and cross talk (21).
2. Blur and cross talk (22).
3. A vertical shift and stereo (23).

Filters are often used to separate the left and right images of a stereo pair for display. The use of filters can cause a difference in contrast (conditions 24 and 28 in Table 4), luminance (condition 25 in Table 4), sharpness, color, and/or local luminance. We simulate the latter three image changes by introducing blur (conditions 27 and 30 in Table 4), red/green color separation (condition 31 in Table 4), and by reducing the color depth of one image to 4 bits, respectively (condition 29 in Table 4). The reduction of color depth may lead to a visual impression of glossiness or sheen, which is usually associated with metallic surfaces, and which is sometimes called metallic luster.[c][18] As far as we know, there are no reports on the reduction of visual comfort because of luster.

For comparison, we also included the condition in which both images of the stereo pair are represented in grayscale only (black and white image representation: condition 26 in Table 4).

In the reference condition (condition 32 in Table 4), the left and right images are identical and should therefore receive a good score.

Procedure

Each stimulus presentation consisted of three temporal intervals:

1. A 3.5-sec presentation of the original (unmanipulated) reference image, followed by
2. A brief interval, followed by
3. A 5-sec presentation of a manipulated stereo image pair.

After viewing a stimulus presentation, the subject was asked to rate the experienced viewing comfort for the image presented in the last temporal interval relative to the reference image presented in the first temporal interval, according to a five-point category rating scale.[19] This scale ranges from 1 to 5, where 1 represents no perceptible reduction in viewing comfort compared with the reference stereo image and 5 represents an extreme reduction in viewing comfort relative to the reference stereo image (Table 3).

Each subject viewed all stimulus conditions. The order of the presentation was randomized and counterbalanced across subjects to reduce the possible effects of stimulus familiarity. Randomization of the presentation order also

[a]Excyclorotation is defined as a templeward turn of the upward extension of the vertical meridian (causing the top of the eye to turn out).
[b]1 PD = 1 prismatic diopter = 0.57° visual angle.

[c]Luster refers to 1) the appearance of two different surface colors viewed haploscopically and superimposed, the resulting percept being characteristically unstable and aptly described as one surface being seen through the other, and 2) a glossiness or sheen associated with metallic surfaces, sometimes called metallic luster.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table 3 Five-point category rating scale, ranging from 1 to 5, where 1 represents no perceptible reduction in viewing comfort compared with the reference stereo image and 5 represents an extreme reduction in viewing comfort relative to the reference stereo image

1	Equal viewing comfort
2	Slightly reduced viewing comfort
3	Reduced viewing comfort
4	Considerably reduced viewing comfort
5	Extremely reduced viewing comfort

served to eliminate the context effect and the after effect of the presentation. A high risk of subjective studies is that previous presentations will influence the judgments over the next presentations. The risk involves the possibility that subjects build up a maximal amount of discomfort (e.g., headaches) so that they are no longer able to make adequate judgments. To reduce this risk, the images were only briefly presented and the extent of the image manipulations was kept moderate. A large number of subjects (24) was used to reduce the above-mentioned effects as well. To obtain an estimate of training and context effects, all 24 subjects performed the complete experiment twice.

Subjects also had to compare images displayed on the autostereoscopic LCD display with the reference image displayed on the projection screen.

Halfway during the actual experiments, the optometric status of each subject was recorded including stereopsis, visual acuity, horizontal and vertical phorias at far and at near, eye dominance, and phoria.

Results

Table 4 summarizes the results of the experiment.

No significant difference was found between the rating scores resulting from the first and second run. This implies that the experiment was neither influenced by a strong context or learning effect. The consistency between the two repetitions (the first and second run), which is a good measure of the accuracy, is on average 0.5 points. Table 4 shows the mean rating score over the 24 subjects and the two repetitions. The quartiles give a good impression of the intersubject variability.

Table 4 shows that spatial distortions that mainly affect the edges of the images of a stereo pair turn out to cause very little discomfort. These distortions include all (1°) rotations, the trapezoid distortion, and the 1.5% magnification. This finding indicates that subjects tend to fixate the center of the screen. We expect that these distortions will cause more discomfort when the edges of the screen are fixated.

The results in Table 4 that correspond to shifts and meridional magnifications show that the human visual system is clearly more sensitive for stereo image pair mismatches in the vertical dimension than for mismatches in the horizontal direction.

Pastoor[4] proposes the rule of thumb that a 35-arcmin horizontal disparity is quite acceptable and 70-arcmin disparity is too much to be viewed comfortably. This rule agrees well with our findings for viewing stereo images (Table 4). The normal stereo image receives a good median score (2.0: corresponding to a slightly reduced viewing comfort) and an acceptable upper quartile score (2.5). The 4× hyperstereo image receives an acceptable median score (2.5) and a poor upper quartile score (3.8: corresponding to a considerably reduced viewing comfort). A significant part of the subjects are therefore troubled by the disparities encountered in the hyperstereo image. Notice that our test scene primarily contains objects at two depth planes separated by 14-arcmin disparity (the wall and the bookcases). The only significant structure in the scene that is at a different depth plane is the plant in the lower right corner, which is 31 arcmin in front of the wall. However, the subjects tended to fixate the central part of the scene and mostly ignored the plant. Therefore we may safely assume that the characteristic disparity of this scene in the 4× hyperstereo condition was about $4 \times 14 = 56$ arcmin. This is in the middle of the transition region between 35 and 70 arcmin as found by Pastoor.[4]

Both for the polarized screen and the autostereoscopic display, a fourfold increase in stereoscopic disparity results in significantly degraded rating scores. However, subjects tended to complain more about the hyperstereo condition when viewing the autostereoscopic display. This indicates that the cross talk and the luster combination present in the autostereoscopic display is equally important in causing eyestrain as the hyperstereo condition. These results suggest that lenticular screens should only be used to display stereo images with small amounts of disparity. Luster and cross talk will otherwise significantly degrade viewing comfort.

An important feature of the binocular visual system is its ability to effectively suppress blur from the defocused eye (interocular blur suppression: Ref. [20]). We therefore expected that the addition of image blur in the combined conditions would improve the rating scores. However, the results in Table 4 show that this is not the case. This indicates that the blur-suppression mechanism needs time to become effective.

The better rating of the ‘‘stereo-and-cross talk’’ condition than of the ‘‘1 PD-shift-and-cross talk’’ condition confirms that the discomfort as a result of cross talk scales with the amount of stereoscopic disparity.[6]

Of the filter manipulations, color separation, blur, and luster (resulting from the reduction of the color depth

Table 4 The rating scores collected in the experiment (1 = viewing comfort not reduced; 5 = extremely reduced)

No.	Image manipulation	LQ	Median	UQ
	Rotations (1°)			
1	Out, symmetric	1.0	1.3	1.5
2	In, one eye	1.0	1.3	1.7
3	In, including border, one eye	1.0	1.4	2.0
4	Out, one eye	1.0	1.4	1.9
5	In, symmetric	1.0	1.5	2.0
	Other distortions			
6	1.5% overall magnification	1.1	1.3	1.8
7	1 PD trapezoid	1.3	1.5	2.0
8	3% meridional horizontal	1.4	1.9	2.3
9	2.5% overall magnification	1.8	2.1	2.6
10	3% meridional vertical	1.8	2.4	3.0
	Shifts			
11	2 PD horizontal (converging)	1.5	1.6	2.6
12	1 PD vertical	2.1	2.8	3.3
13	3 PD horizontal (converging)	2.0	2.8	3.5
14	2 PD vertical	4.0	4.5	4.9
	Stereo images			
15	Normal stereo (IPD = 6 cm)	1.2	2.0	2.5
16	2 × hyperstereo (IPD = 12 cm)	1.5	1.9	2.6
17	4 × hyperstereo (IPD = 24 cm)	2.0	2.5	3.8
	Cross talk			
18	5% and 1 PD horizontal shift	1.5	2.0	2.7
19	15% and 1 PD horizontal shift	3.5	3.5	4.5
20	25% and 1 PD horizontal shift	4.4	4.8	5.0
	Combined manipulations			
21	Stereo and 5% cross talk	1.0	1.0	1.5
22	5% cross talk and blur	2.0	2.5	3.0
23	Stereo and 1 PD vertical shift	2.2	2.8	3.5
	Filter			
24	Higher contrast (+25%)	1.0	1.3	1.4
25	Overall luminance difference (−25%)	1.1	1.5	2.0
26	Black and white representation	1.0	1.5	3.0
27	Gaussian blur (sigma = 1 pixel)	1.7	2.2	2.7
28	Lowered contrast (−50%)	2.0	2.5	3.1
29	Quantization difference (4 bit)	2.0	2.8	3.3
30	Gaussian blur (sigma = 1.5 pixels)	2.3	3.0	3.5
31	Color asymmetry (red/green)	3.2	4.0	5.0
	Reference image			
32	Reference image	1.0	1.0	1.0
	Autostereoscopic display			
33	No stereo	1.0	1.3	1.8
34	Normal stereo	2.0	3.0	3.8
35	Hyperstereo	3.4	3.8	4.5

LQ = lower quartile, indicating that 25% of the subjects had a lower score. UQ = upper quartile indicating that 25% of the subjects had a higher score. The conditions are described in the text. 1 PD = 1 prismatic diopter which corresponds to the angle of 1 cm viewed from 1-m distance and which equals 0.57°.

of one image to 4 bits) have the most impact on viewing comfort.

The viewing comfort associated with the reference image is rated as ''not degraded,'' just as expected.

To provide more insight into the data, a part of the results is also shown graphically in Figs. 10 and 11. Figure 10 shows the median rating scores with upper and lower quartiles for the three levels of cross talk. The viewing comfort of images containing 5% cross talk is rated as ''a bit reduced'' (2), for 15% cross talk, it is considered as ''reduced a lot'' (3–4), and for 25% cross talk, it is experienced as ''extremely reduced'' (close to

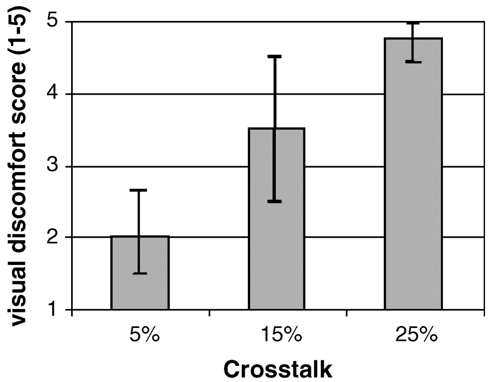


Fig. 10 Median rating scores with upper and lower quartiles for the three levels of cross talk. The visual comfort with 5% cross talk is rated as ''a bit reduced'' (2), 15% cross talk as ''reduced a lot'' (3–4), and 25% as ''extremely reduced'' (close to 5). (*View this art in color at www.dekker.com.*)

5). The three conditions differ statistically significant from each other ($P<0.05$). Figure 11 shows the median rating scores with upper and lower quartiles for four different horizontal and vertical binocular asymmetric shifts. This figure clearly shows that the human visual system is most sensitive (in terms of comfort) to vertical binocular asymmetries. Even a small vertical shift seriously degrades the visual comfort. However, horizontal shifts also degrade stereoscopic viewing comfort.

Optometry: between subject differences

We were particularly interested whether the level of experienced visual discomfort is related to the optometric status of the viewer. One can imagine, for example, that a subject with the tendency to squint could be bothered more by a spatial distortion. We therefore correlated the optometric data (visual acuity, stereopsis, phoria) with the rating scores. The results are listed in Table 5. Subjects with good vision are bothered more by an image misalignment than the subjects with poor vision. The reason is easy to understand: a person with limited binocular vision is less able to see the misalignment. We do not find an effect of the subject's eye quality on the image manipulations that are also noticeable with one eye. These include cross talk and the filter conditions. The virtual absence of a correlation with the visual acuity of the better eye (VA_{max}) indicates that the quality of binocular vision is determined by the poorer eye. This is confirmed by the high correlation between a subject's stereopsis score and the visual acuity of the poorer eye (VA_{min}) on the one hand and the absence of a correlation between stereopsis and VA_{max} on the other hand. These correlations are -0.96 and $+0.3$ respectively.

In summary, the quality of a person's binocular vision affects the binocular viewing comfort to a limited extent. Binocular misalignment and excessive stereoscopic disparity are less troublesome for people with reduced (binocular) vision.

Binocular image manipulation thresholds

The data in Table 4 can be used to estimate thresholds for each of the binocular image manipulations. As threshold, we choose to adopt a median score of 2 (viewing comfort is ''a bit reduced'' for 50% of the observers) or an upper quartile score of 3 (viewing comfort is ''reduced'' for 25% of the observers). The resulting thresholds are listed in Table 6. For some of the image manipulations, it is not possible to give accurate threshold estimates from our data. In that case, we list the highest value of the respective distortion for which the observers experienced no reduction in viewing comfort, and we indicate that the threshold is larger than the listed value.

For example, Table 4 shows that the observers experienced no reduction in viewing comfort for rotations up to 1°. This implies that the actual threshold is larger than 1° and designers should not be worried when image rotations up to 1° occur in practice.

From Table 6 we see that to *prevent reduced viewing comfort* with differentially distorted binocular image pairs,

An overall image magnification should be less than 2.5%.
The meridional vertical distortion should be less than 3%.
A horizontal shift should be less than about 2 and 3 PD.
A vertical shift should be less than 1 PD.
The disparity in a stereoscopic image should not exceed 30 arcmin.
For high disparity images, the cross talk should be less than 5%.

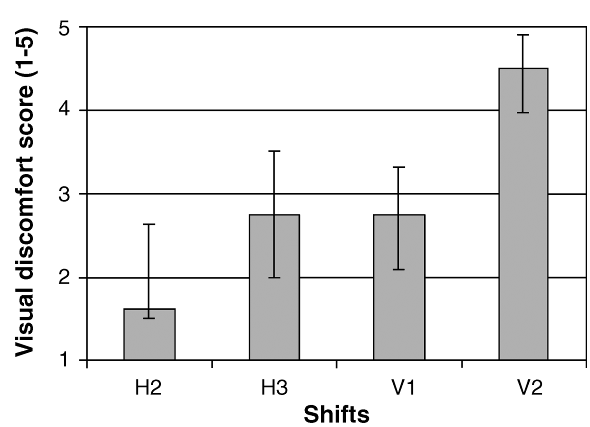


Fig. 11 Median rating scores with upper and lower quartiles for four binocular asymmetric shifts. H(V)1(2,3)=horizontal (vertical) convergent shift of 1 (2,3) prism diopters. Even small vertical shifts seriously degrade the visual comfort, but a horizontal shift also makes an impact. (*View this art in color at www.dekker.com.*)

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

Table 5 Correlations between three measures of eye quality (stereopsis, maximum, and minimum visus) and the viewing discomfort of a subset of the conditions (vertical)

	Stereopsis	Visus max	Visus min
4 × hyperstereo	−0.39 (p<0.063)	0.03	0.24
Cross talk	−0.01 (p<0.95)	0.17	−0.04
Horizontal shift	−0.33 (p<0.11)	0.19	0.23
Vertical shift	*−0.68 (p<0.0003)*	0.15	*0.59 (p<0.0025)*
Filters	0.11 (p<0.62)	0.17	−0.05
Autostereoscopic display	−0.20 (p<0.35)	0.11	0.04

Note that the correlations with stereopsis are negative because a high value indicates a poor score. The only binocular asymmetries that correlate significantly with the optometric status are the vertical misalignment and the 4 × hyperstereo (we calculated the correlations based on the logarithm of the stereopsis). Significant correlations are shown in italic. As expected, the visual acuity of the better eye does not correlate with the rating scores.

The contrast difference should not exceed 25%.
The Gaussian blur should have a standard deviation less than 2.3 arcmin (corresponding to 1 pixel on our display).
Color asymmetry should be less severe than complete red/green separation.
For 8-bit imagery, the quantization difference should be less than 4 bits.

Moreover, the *viewing comfort will not be reduced* for

A trapezoidal image distortion up to 1 PD.
A meridional horizontal distortion up to 3%.
Low disparity images if the cross talk is up to 5%.
Global luminance differences up to 25%.

A black and white image representation.

Table 6 Threshold values for each of the binocular image manipulations

Image manipulation	Threshold value
Rotations (1°)	>1°
Other distortions	
Overall magnification	2.5%
Trapezoid	>1 PD
Meridional horizontal	>3%
Meridional vertical	<3%
Shifts	
Horizontal	Between 2 and 3 PD
Vertical	<1 PD
Stereo images	Between 2 × and 4 × hyperstereo
Cross talk	
Low disparity	>5%
High disparity	5%
Filter	
Contrast difference	Between 25% and 50%
Luminance difference (−25%)	>25%
Black and white representation	OK
Gaussian blur	<1 pixel
Local luminance	Between 4- and 8-bit digitization
Color asymmetry	Less than red/green separation

The thresholds are derived from the data in Table 1. As threshold, we take a median score of 2 (viewing comfort is ''a bit reduced'' for 50% of the observers) or an upper quartile score[a] of 3 (viewing comfort is ''reduced'' for 25% of the observers). The conditions are described in the text.
[a]UQ = upper quartile indicating that 25% of the subjects had a higher score.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016

CONCLUSION

Visual discomfort resulting from binocular imperfections can be very severe, making it a factor that deserves serious consideration. The literature provides some data on visual discomfort of binocular displays, showing that the impact of binocular imperfections depends on many factors. The level of discomfort is determined by the degree and type of the imperfection, whether the system is see-through, and by the viewing duration. The impact of visual discomfort is also influenced by the circumstances: the ‘‘pain threshold’’ of a system used professionally is lower than for entertainment at home. Visual discomfort is relatively unimportant in the case of an advertising display near the entrance of a store meant to catch the attention of passersby because it is viewed for short periods, intermittent, and voluntary. In contrast, the Dutch KDC10 refueling operator station[21] is an example of a stereoscopic viewing system requiring excellent viewing comfort. It concerns paid work at odd hours, long viewing periods, and high responsibility.

Deriving engineering specifications for binocular viewing systems from the human factors data is not easy because many factors play a role. We advise to start by looking for human factors data that have been collected on hardware and conditions matching those of the system being designed. The references given here should provide a good starting point. A word of caution regarding 3-D displays is in order: no one has succeeded yet in building a 3-D viewing system that is free from eyestrain.

Binocular image imperfections can either have an optical origin (such as a shift, magnification, rotation, or distortion) or can result from an imperfect filter (resulting in a photometric asymmetry: luminance, color, contrast, cross talk). Stereoscopic disparity presented without the natural coupling to accommodation disparity and motion parallax can result in visual discomfort ranging from an odd feeling to the irresistible urge to close the eyes.

Nearly all binocular image asymmetries seriously reduce visual comfort if present in a large-enough amount. The data shown in Table 4 allow threshold value estimates for the onset of discomfort. Designers and users of binocular viewing systems can use these data to predict visual comfort from the specification of a binocular system.

The present study addresses static images that are presented for short amounts of time. This condition is representative for many actual viewing situations because people tend to regularly alternate fixation between a display and other parts of the visual world. A notable exception is formed by immersive HMDs that prevent this behavior by occluding everything except the display itself. Although display duration has a clear effect on stereoscopic vision in a performance-oriented context,[22–24] it has been shown before that brief image presentations suffice to obtain reliable appreciative judgments of static stereoscopic image material.[26,27]

REFERENCES

1. Kalich, M.E.; Rash, C.E.; van de Pol, C.; Rowe, T.L.; Lont, L.M.; Peterson, R.D. Biocular Image Misalignment Tolerance. In *Helmet- and Head-Mounted Displays VIII: Technologies and Applications*; Rash, C.E., Reese, C.E., Eds.; The International Society for Optical Engineering: Bellingham, WA, 2003; 284–295.
2. Velger, M. *Helmet-Mounted Displays and Sights*; Artech House: Norwood, MA, 1998.
3. Pastoor, S. Human factors of 3D displays. Displays **1993**, *14*.
4. Pastoor, S. Human Factors of 3D Imaging: Results of Recent Research at Heinrich-Hertz-Institut Berlin. In *Proceedings of the International Display Workshop ’95 (Asia Display ’95)*; Asian Technology Information Program: Tokyo, Japan, 1995; *3*, 66–72. for more information see: http://www.atip.org/fpd/src/conf/ad95/index.html.
5. Self, H.C. *Optical Tolerances for Alignment and Image Differences for Binocular Helmet-Mounted Displays (Report AAMRL-TR-86-019)*; Armstrong Aerospace Medical Research Laboratory: Wright-Patterson AFB, OH, 1986.
6. Pastoor, S. Human factors of 3D displays in advanced image communications. Displays **1993**, *14* (3), 150–157.
7. Beldie, I.P.; Kost, B. Luminance Asymmetry in Stereo TV Images. In *Stereoscopic Displays and Applications II*; Merritt, J.O., Fisher, S.S., Eds.; SPIE—The International Society for Optical Engineering: Bellingham, WA, 2003; 242–247.
8. Rash, C.E. *Helmet Mounted Displays: Design Issues for Rotary-Wing Aircraft (Report available on http://www.usaarl.army.mil/)*; U.S. Army Aeromedical Research Laboratory (USAARL): Fort Rucker, AL, 2003.
9. Howard, I.P.; Rogers, B.J. *Binocular Vision and Stereopsis*; Oxford University Press: Oxford, UK, 1995.
10. Melzer, J.E. *Head-Mounted Displays: Design and Applications (SPIE Short Course Notes SC-159)*; The International Society for Optical Engineering: Bellingham, WA, 2002.
11. Eagle, R.A.; Paige, E.; Sucharov, L.; Rogers, B.J. Accommodation cues reduce latencies for large-disparity detection. Perception **1999**; *28* (Supplement), 136c.
12. Gustafsson, T. *Mixed Reality i Militära Tillämpningar: En Förstudie med viss Fokus på Integration, Blickriktningssensor och Huvudburet Mikrodisplaysystem: Användarrapport (Mixed Reality in Military Applications: A Pilot Study with Certain Focus on Integration of Gaze-Tracker and Head Mounted Displays: User Report) (Report FOI-R-0587-SE)*; Totalförsvarets Forskningsinstitut (FOI): Linköping, Sweden, 2003.
13. Kooi, F.L.; Toet, A. Additive and Subtractive Transparent Depth Displays. In *Enhanced and Synthetic Vision 2003*; Verly, J.G., Ed.; The International Society for Optical Engineering: Bellingham, WA, USA, 2003; 58–65.

14. Kooi, F.L.; Toet, A. Visual comfort of binocular and 3D displays. Displays **2004**. Submitted.
15. Honda, T. Dynamic holographic 3D display using LCD. Asia Display **1995**, *95*, 777–780.
16. Yeh, Y.Y.; Silverstein, L.D. Limits of fusion and depth judgement in stereoscopic color displays. Hum. Factors **1990**, *32* (1), 45–60.
17. Rash, C.E.; McLean, W.E. Optical Performance. In *Helmet-Mounted Displays: Design Issues for Rotary-Wing Aircraft*; Rash, C.E., Eds.; U.S. Army Aeromedical Research Laboratory: Fort Rucker, AL, 1998; 101–167.
18. Cline, D.; Hofstetter, H.; Griffing, J.R. *Dictionary of Visual Science and Related Clinical Terms*; Butterworth-Heinemann: Oxford, UK, 1980.
19. ITU-R. *Methodology for the Subjective Assessment of the Quality of Television Pictures (Report BT.500-9)*; International Telecommunication Union: Geneva, Switzerland, 2003.
20. Schor, C.M.; Erickson, P. Ocular Dominance Accommodation, and the Interocular Suppression of Blur in Monovision. In *Presbyopia, Research: From Molecular Biology to Visual Adaption*; Obrecht, G., Stark, L., Eds.; Plenum Press: New York, USA, 1991; 273–288.
21. Kooi, F.L.; van Breda, L. *Quantifying the Image Quality of the KDC-10 Refuelling Vision System (Report TNO-TM 1997 A-052)*; TNO Human Factors Research Institute: Soesterberg, The Netherlands, 2003.
22. Tam, W.J.; Stelmach, L.B. Display duration and stereoscopic depth discrimination. Can. J. Exp. Psychol. **1998**, *52*, 56–61.
23. Patterson, R.; Cayko, R.; Short, L.; Flanagan, R.; Moe, L.; Taylor, E.; Day, P. Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. Percept. Psychophys. **1995**, *57*, 891–897.
24. Patterson, R.; Fox, R. The effect of testing method on stereoanomaly. Vis. Res. **1984**, *24*, 403–408.
25. IJsselsteijn, W.A.; de Ridder, H.; Vliegen, J. Effects of Stereoscopic Filming Parameters and Display Duration on the Subjective Assessment of Eye Strain. In *Stereoscopic Displays and Virtual Reality Systems VII*; Merritt, J.O., Benton, S.A., Woods, A.J., Bolas, M.T., Eds.; The International Society for Optical Engineering: Bellingham, WA, 2000; 12–22.
26. IJsselsteijn, W.A.; de Ridder, H.; Vliegen, J. Subjective evaluation of stereoscopic images: Effects of camera parameters and display duration. IEEE Trans. Circuits Syst. Video Technol. **2000**, *10* (2), 225–233.

Marcel Dekker, Inc.
270 Madison Avenue, New York, New York 10016