Stephan Reichelt, Ralf Häussler, Gerald Fütterer, and Norbert Leister, *"Depth cues in human visual perception and their realization in 3D displays,"* Three-Dimensional Imaging, Visualization, and Display 2010, Bahram Javidi and Jung-Young Son, Editors, Proc. SPIE 7690, 76900B (2010), DOI:10.1117/12.850094

http://dx.doi.org/10.1117/12.850094

Depth cues in human visual perception and their realization in 3D displays

Stephan Reichelt, Ralf Häussler, Gerald Fütterer, and Norbert Leister
SeeReal Technologies GmbH, Blasewitzer Str. 43, 01307 Dresden, Germany

ABSTRACT

Over the last decade, various technologies for visualizing three-dimensional (3D) scenes on displays have been technologically demonstrated and refined, among them such of stereoscopic, multi-view, integral imaging, volumetric, or holographic type. Most of the current approaches utilize the conventional stereoscopic principle. But they all lack of their inherent conflict between vergence and accommodation since scene depth cannot be physically realized but only feigned by displaying two views of different perspective on a flat screen and delivering them to the corresponding left and right eye. This mismatch requires the viewer to override the physiologically coupled oculomotor processes of vergence and eye focus that may cause visual discomfort and fatigue.

This paper discusses the depth cues in the human visual perception for both image quality and visual comfort of direct-view 3D displays. We concentrate our analysis especially on near-range depth cues, compare visual performance and depth-range capabilities of stereoscopic and holographic displays, and evaluate potential depth limitations of 3D displays from a physiological point of view.

Keywords: 3D Displays, Human factors, Depth cues, Accommodation, Vergence, Binocular vision, Stereoscopic displays, Holographic displays

1. INTRODUCTION

When considering that we live today in a communication society where information exchange widely relies on visual representation, it is amazing that most of the screens we are using plenty of hours per day for work or entertainment get along with a flat 2D image. Generally, complex data can be interpreted more effectively when displayed in three dimensions. In information display industry, three-dimensional (3D) imaging, display, and visualization are therefore considered to be one of the key technology developments that will enter our daily life in the near future. Natural perception of depth as in daily life, however, remains still a challenging task in display technology as much as for content creation. But why is that? What causes us to perceive 3D images and movies as being comfortable or not? How do we sense that something is different or missing? How do we know in which aspects our visual system is more or less tolerant toward impairments in 3D imaging? As answers to these questions are crucial for successful product development, launch and consumer acceptance of 3D displays it is all the more important to address human factor issues at the best. This involves display performance, eye visual acuity, and visual perception.

The purpose of this paper is to review current 3D display technologies, and especially how they provide crucial depth cues. In particular, we discuss motion parallax and consistent accommodation/convergence cues, which become essential for 3D desktop displays intended for longer use at short viewing distances.

2. DEPTH PERCEPTION AND VISUAL CUES

The human visual system relies on a large number of cues for estimating distance, depth, and shape of any objects located in the three-dimensional space of the surrounding. Fig. 1 gives an overview of the various cues for depth perception. For the sake of completeness, this section briefly recaps the different cues and their meanings. For optimum visual comfort all depth cues delivered by a 3D display have to be both mutually linked and consistent with natural viewing. In our discussion, however, we concentrate on the interaction of accommodation and convergence because providing well-matched focus and disparity cues is still an unsolved issue for most of the known 3D display systems.

Further author information: (Send correspondence to Stephan Reichelt)
E-mail: sre@seereal.com, Phone: +49 (0)351 450 3240, Web: `http://www.seereal.com`

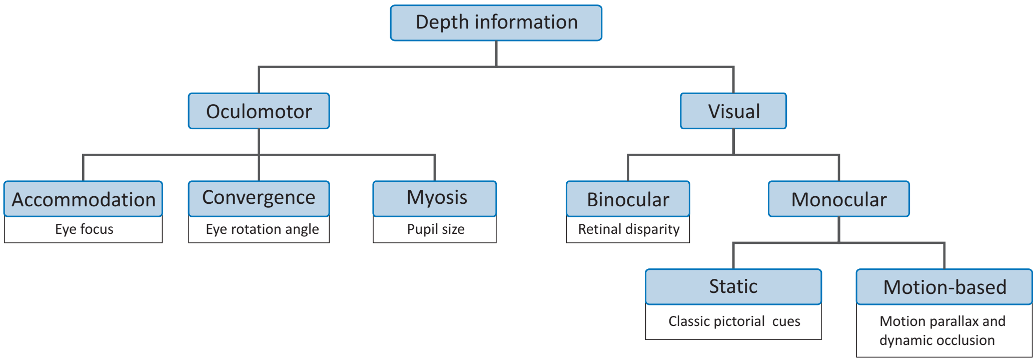


Figure 1. Overview and classification of depth cues.[1]

2.1 Visual depth cues

Visual depth cues can be classified into monocular and binocular cues. Monocular depth cues are subdivided into pictorial depth cues and motion cues. Even flat images can provide static depth cues such as interposition, linear perspective, relative and known size, texture gradient, heights in picture plane, light and shadow distribution, and aerial perspective – these so-called *pictorial depth cues* have been applied in visual arts for centuries. Motion-based cues involve shifts on the retinal image and are induced by relative movements between observer and objects. Among them are motion parallax, kinetic depth effect, and dynamic occlusion. Motion-based cues play an important role for depth perception; particularly in static scenery motion-parallax provides a fast and reliable depth estimate.

2.2 Oculomotor depth cues

A fixation of near targets evokes three oculomotor responses: accommodation, convergence, and pupillary constriction, which is in combination referred to as the *ocular near triad*. In a generally accepted view, the mutual interplay between accommodation and convergence is modeled as two dual and parallel feedback control systems that are connected via cross links. Both feedback control systems are attributed to the same physical input, that is, fixation to a point or region that differs in distance to a previously fixated object. But the primary effects are different in terms of monocular or binocular description of the situation. While accommodation is mainly driven by retinal blur (the monocular cue, working for each single eye), convergence is principally triggered by retinal disparity (the binocular cue, only for two eyes working properly together). However, due to the cross-link interacting of the control systems, the situation becomes more complex. Primary purpose of the interaction is to provide both sharp and comfortable, binocular single vision. For a more descriptive explanation of the accommodation-convergence model, see for example Lambooij et al.[2] and references therein.

Measures of the cross-linked behavior between accommodation and convergence are the so-called AC/A or CA/C ratios. The AC/A ratio describes the accommodation-induced alteration in convergence related to the accommodation change (without retinal disparity being present). On the other hand, the CA/C ratio describes the convergence-induced alteration in accommodation related to the convergence change (without retinal blur being present).

2.2.1 Accommodation and monocular acuity

Accommodation is the mechanism by which the human eye alters its optical power to hold objects at different distances into sharp focus on the retina. The power change is induced by the ciliary muscles, which steepen the crystalline lens' curvature for objects at closer distances. When an object of interest is fixated by the eye, the accommodation is adjusted such that a sharp image is perceived onto the retina. Full accommodation response requires a minimum fixation time of one second or longer. But the human eye can tolerate a certain amount of retinal defocus without readjusting accommodation, although the criteria for goodness of focus depend on the observed object and vary from individual to individual. In optometry, the corresponding optical power difference is called the ocular *depth of focus*. It is related to image space and usually given in diopter. That means the object of interest may be slightly shifted along the line of sight between a distal point P_d and a proximal point P_p without perceiving image blur. The term *depth of field* refers then to the corresponding distance between P_d and P_p in object space.

The total geometric *depth of focus* at a fixed accommodation state is defined as[3]

$$T = \frac{1}{p} - \frac{1}{d}, \tag{1}$$

with p and the d being the *proximal* (near) and *distal* (far) positions in meters at which a test target can be seen sharply. Note that the ocular depth of focus is often given in terms of $\pm\frac{T}{2}$, and thus it is related to the fixation distance z. For an estimated depth-of-focus tolerance T, proximal and distal distances can be calculated according to

$$p = \frac{2z}{2 + zT} \qquad \text{or} \qquad d = \frac{2z}{2 - zT}, \tag{2}$$

respectively. Another useful concept for depth-of-field estimation is that of the *hyperfocal distance* z_h, which is defined as the fixation distance for which the depth-of-field's distal position becomes infinite ($d \to \infty$). When the eye is fixated at z_h, all objects beyond will be still in focus. From Eq. (2) this yields for the hyperfocal distance

$$z_\mathrm{h} = \frac{2}{T}. \tag{3}$$

To compensate for the natural inaccuracies of accommodation, the optical power of the eye fluctuates with less than 0.25 diopters at a frequency of some Hertz. Combined with pupil diameter variation, that way the retinal image blur is minimized by depth-of-focus fine-adjustment. Within the physiological optics literature, numerous estimates for the depth of focus of the human eye have been given – albeit the variation in its quantity is quite large, ranging for a 2 mm pupil diameter between $0.126\,\mathrm{D} \leq T \leq 0.86\,\mathrm{D}$.[4] Any depth-of-field calculation based on those estimates might be therefore debatable.

But before we continue that discussion, we would briefly address the factors, which are influencing depth of focus and thus depth of field or hyperfocal distance:

Pupil size. As the pupil serves as a variable aperture stop of the eye, it controls the luminous flux and quality of the retinal image by balancing out the effects of diffraction, aberration, and depth of focus. The depth of focus is generally influenced by the size of the pupil, which is, in turn, mainly affected by the luminance level of the target. Under normal photopic conditions ($> 3\,\mathrm{cd/m^2}$), the pupil diameter is in the range 2...6 mm. The depth of focus increases as the size of the pupil decreases. Largest depth of field is thus obtained with a pupil diameter of approx. 2 mm, which is also the diameter where the eye reaches diffraction-limited performance.[5] Pupil size also depends on whether an object is viewed monocular or binocular with slightly smaller pupils for binocular situation. Moreover, the pupil constricts when focused and converged at a near object.

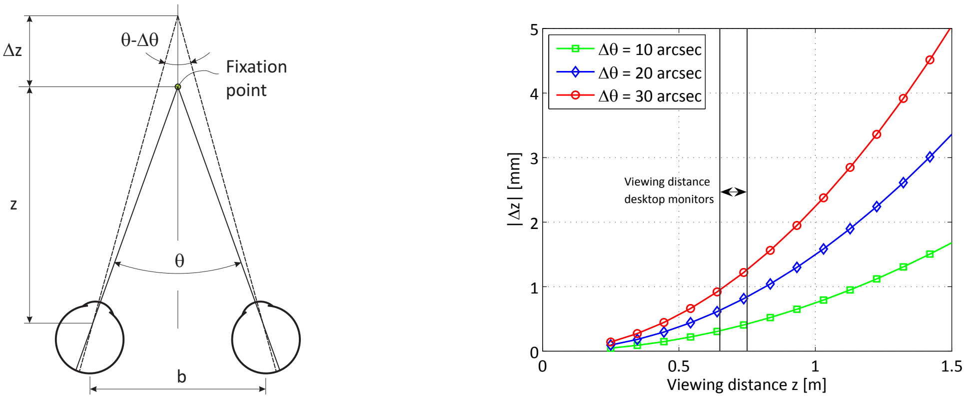
Contrast and spatial frequency of the target. Accommodation response is triggered by retinal image blur, but apart from a minimum fixation time, the amount of accommodation depends on contrast and spatial frequency of the target, too. Objects having low contrast or low spatial frequencies represent a weaker stimulus for accommodation because the retinal image may tolerate a bit more defocus than for an object having fine, high-contrast details. The normal visual acuity is considered to be 1 arcmin, which is also the value for the angular resolution of the eye. Under ideal conditions, the acuity threshold is about 30 arcsec for grating acuity and about 5 arcsec for vernier acuity, respectively.

Aberrations. The eye is not a perfect optical system, however; there are monochromatic as well as chromatic errors, which vary individually. The merit of accommodation can be impaired in the presence of aberrations such as defocus or astigmatism. But even with an ideally corrected eye, the inherent spherical aberration will in part diminish the eye's performance, especially when the pupil becomes larger at lower luminance levels. For pupils larger than 3 mm, there is also a chromatic aberration of approx. 0.75 D.[6]

Both accommodation ability and maximum pupil diameter decline with age. While children aged ten have a large accommodation ability of approx. 14 D, the accommodation is reduced to approx. 4 D within the forth decade and finally levels below 2 D in older age. For similar age groups, the maximum pupil diameter reduces from 8 mm over 6 mm to approx. 2...3 mm.

2.2.2 Convergence and stereoscopic acuity

Since the human eyes are horizontally separated, each eye sees a slightly different perspective of a natural scene. The retinal images are thus slightly different with so-called *crossed* or *uncrossed disparity* for objects in front or behind the fixation point, respectively. *Stereopsis* is based on the different perspective of the two retinal images, which provides a major cue for relative depth perception.



(a) (b)

Figure 2. Depth discrimination by stereoscopic acuity. (a) Geometry of stereopsis illustrating stereoscopic depth. (b) Theoretical minimum detectable difference $|\Delta z|$ for near range object distances z. Assumed are a median interpupillary distance of $b = 65\,\mathrm{mm}$ and three estimates for $\Delta\theta$.

The geometry of stereopsis is shown in Fig. 2(a). It corresponds to the experimental setup of the two-needle test.[7] According to the figure, the theoretical depth discrimination Δz can be derived from the definition of convergence angle θ by applying $d\theta/dz$, which gives

$$\Delta z = -\frac{z^2}{b}\left(1 + \frac{b^2}{4z^2}\right) \cdot \Delta\theta, \tag{4}$$

with b being the interpupillary distance and z being the mean object distance. With $4z^2 \gg b^2$, which is valid even for near point distance $z_{\text{near}} = 250\,\mathrm{mm}$, Eq. (4) simplifies to the common form of $\Delta z \approx -z^2\Delta\theta/b$.

Under photopic lighting conditions, a stereo acuity $\Delta\theta = 10\,\mathrm{arcsec}$ is assumed to be a reasonable value. The minimum detectable depth difference is for a fixation distance of 650 mm thus in the order of 0.3 mm. Among other factors, stereoscopic acuity is mainly affected by the following parameters:

Target luminance. Stereoscopic acuity strongly depends on the luminance contrast of the target. At a low luminance level, the stereoscopic acuity declines to approx. 40 arcsec, whereas under optimal conditions even 2...5 arcsec may be achieved.

Spatial frequency of the target. With higher spatial frequencies of the target, stereo acuity improves until the individual optimal level is reached.

Angular distance from fixation. With increasing angular distance from fixation, i.e. with growing retinal eccentricity the stereo acuity quickly narrows down. For example, at an eccentricity of 2°, $\Delta\theta$ becomes larger than 30 arcsec.

Observation time. The sensitivity to small disparity differences is improved with extended fixation time. Thus, fast temporally-changing scenes are perceived with a reduced stereo acuity.

3. QUALITY AND VISUAL COMFORT OF 3D DISPLAYS

Before we continue with a comparison of 3D-related characteristics of different display types, we briefly recapitulate common technologies for 3D visualization. We will attempt to explain basic principles of the underlying visualization technologies rather than specific display designs or embodiments. There are comprehensive reviews of current 3D display technologies.[8–13]

3.1 Types of 3D displays and related 3D characteristics

Various technologies for visualizing three-dimensional scenes on displays have been technologically demonstrated and refined. Their developmental history dates back to the Wheatstone stereoscope, invented in 1838.[14] Among them are the most common family of stereoscopic and multi-view displays with their further enhancements of autostereoscopic and super-multiview displays. Integral imaging is another technology, which becomes more and more popular. Volumetric displays are different from the other 3D display types as they have not a single screen but a truly enclosed 3D volume. Finally, holographic displays are commonly considered the golden standard for 3D visualization since they are capable of reproducing a scene in a most natural looking way. Each of the technologies has its particular advantages and disadvantages. When discussing the 3D-related performance, the following generic classification of currently dominating 3D display technologies might be useful:

Binocular displays. These displays utilize the conventional stereo principle, that is delivering two views of a scene to the viewer's left and right eye. Per frame, only one set of images is presented. Binocular separation of the views is created by multiplexing methods, utilizing e.g. wavelength-division (anaglyph type), space/direction-division (parallax-barrier type, lenticular type), time-division, polarization-division or combinations of it. When eyewear (passive: anaglyph or polarization; active: LCD-shutter glasses) is needed, the displays are called stereoscopic, whereas autostereoscopic displays require no bothersome glasses. The latter type is realized by creating a fixed viewing zone for each eye (parallax-barrier or lenticular) or, more advanced, is combined with tracking for eye detection and viewing zone movement (shifting barriers or lenticulars, steerable backlight). It is common to all that only binocular parallax is provided as depth cue.

Multi-view displays. Despite still being stereoscopic, multi-view displays create a discrete set of perspective views per frame and distribute them across the viewing field. This gives in the first place viewing freedom for one or more observer. Additionally, it provides motion parallax to a certain degree as the observer moves to the adjacent viewing zone. Spatially multiplexed designs with slanted lenticulars alleviate the effect of image flipping and black mask visibility, but even so, display resolution and thus image resolution has to be divided according the number of views.

Integral imaging displays. Integral imaging displays make use of a set of 2D elemental images taken with different perspective to create a 3D image. Advantages of integral imaging are that they inherently provide stereo parallax as well as full and continuous motion parallax and thus, allow multiple viewing positions for several viewers. Furthermore, they operate in incoherent light. Disadvantages are their low lateral and depth resolution together with a limited depth range.

Volumetric displays. In true volumetric displays, each point of a scene is created at its actual position in space, which can be realized by either directing laser beams or layered images on moving screens or by employing focused or intersecting laser beams that create voxel-emitting dots via fluorescence or scattering. In doing so, they provide a wide viewing range and correct accommodation cues. Swept-volume-type volumetric displays are possibly best-known, where very fast rotating or translating screens are illuminated by laser beams to create a translucent or contour image. Advantage of such scanning-type volumetric displays is that they generate imagery in true, albeit enclosed 3D space and thus provide full and continuous motion parallax. In multiplanar displays, multiple display planes are stacked to display depth-separated layers of the scene on different screens. However, depth resolution is limited by the number of display layers. With only few planes, they pose the risk of the cardboard effect; but contrary, brightness might suffer as the number of layers is increased.

Table 1. 3D display types and provided depth cues. Legend: ✓ = present; (✓) = with limitations

Type of 3D display	Implementation	Provided depth cues			Remarks on possible limitations
		Disparity	Motion parallax	Focus	
Binocular	Stereoscopic	✓			Limited comfort depth range (depth squeezing) Shear distortion
	Autostereoscopic	✓			Limited comfort depth range (depth squeezing) Shear distortion
Multi-view		✓	(✓)		Limited comfort depth range (depth squeezing) Only horizontal motion parallax Picket-fence effect (parallax barrier type) Image flipping, Limited resolution
Integral imaging	1D lenticular lens	✓	(✓)		Only horizontal motion parallax Limited depth range (depth squeezing) Limited resolution
	2D spherical lens array	✓	✓		Limited depth range (depth squeezing) Limited resolution
Volumetric	Stacked multi plane	✓	(✓)	(✓)	Few depth planes (cardboard effect)
	Swept volume	✓	✓	✓	Limited 3D volume Image transparency (translucent 3D objects)
Holographic	Horizontal parallax only (classic)	✓	(✓)	✓	Only horizontal motion parallax Limited 3D volume
	Tracked single parallax	✓	✓	✓	Fast tracking means required
	Full parallax (classic)	✓	✓	✓	Limited 3D volume
	Tracked full parallax	✓	✓	✓	Fast tracking means required

Holographic displays. Holography is a diffraction-based coherent imaging technique in which a 3D scene can be reproduced from a flat, two-dimensional screen having a complex amplitude transparency (amplitude and phase values). Holographic displays reconstruct the wave field of a 3D scene in space by modulating coherent light, e.g. with a spatial light modulator. Because of its superior capabilities, real-time holography is commonly considered the ideal 3D technique.

3.2 Crucial depth cues

Because of the ongoing boom in 3D displays and the awareness of potential limitations involved with the different technologies, it is not surprising that the investigation of human factors and visual comfort is an active area of research. There is a vast number of specific studies and general reviews on this topic, while most of them deal with stereoscopic displays.[15] Though some of the following factors may affect viewing comfort considerably, the discussion of typical stereoscopic artifacts,[16] such as *binocular rivalry* (excessive disparity), *frame cancellation* (near-edge cut-off for objects with front depth), *shear distortion* (perspective distortion with viewpoint changing), *keystone distortion* (unnatural vertical disparity), *binocular crosstalk* (ghost images), *cardboard effect* (few discrete depth planes) is beyond the scope of this review – especially as most of these imperfections can be handled with careful content preparation and suited display implementations.

A general survey of visual discomfort and visual fatigue involved with stereoscopic displays has been given recently by Lambooij et al.[2] To draw a distinction in terminology, the authors apply visual fatigue for objectively measurable symptoms, whereas visual discomfort refers to subjective estimates. They argue for a stereo comfort zone and derive its boundaries for a disparity limit of 1°, which is assumed to hold true for both crossed and uncrossed disparity. Nevertheless, they recognized that visual discomfort may still occur within this comfort zone in case of fast object movements across the depth range and unnatural blur.

Based on our experience, natural, full-parallax motion cues and consistent oculomotor cues of vergence and accommodation are likely to be the most decisive factors for a comfortable 3D viewing experience. This is particularly relevant to displays with short observer distance such as desktop monitors, notebooks or hand-held devices. Even TV systems with their increased viewing distance compared to desktop monitors suffer from the effect if you consider that at a distance observer to display of 2 m the tolerable total depth in a Super Bowl game is only 1 m.

3.2.1 Natural motion parallax

The term *motion parallax* refers to the effect that the retinal images of objects, located in front or behind the fixation point, moves in different direction and speed across the retina as a relative movement between observer and environment occurs. Objects closer to the observer than the fixation point appear to move faster and in opposite direction to the movement of the observer, whereas objects farther away move slower and in the same direction. While this enables the viewer to look around objects, at the same time it provides a strong cue to relative depth perception.[17,18] The underlying similarity between motion parallax and binocular disparity (stereo parallax) has been early recognized by Helmholtz in his fundamental work.[19] Obviously, if one eye moves horizontally about the interpupillary distance while maintaining fixation, the two monocular views would be the same as with binocular viewing, that is with disparity in perspective.[20] As with stereo acuity, motion parallax is therefore more effective and accurate in near range. It should be noted, however, that motion parallax works equally in all directions, whereas binocular disparity most of the time happens in horizontal direction only – naturally, we are in upright position.

Binocular stereo displays suffer from shear distortion, which refers to the effect that the image appears to follow as the observer moves. Depending on the concrete embodiment, conclusive motion-based depth cues can be implemented by super-multi view, integral imaging, volumetric, and holographic 3D displays (cf. Table 1). Smooth motion parallax is an important supplement to stereopsis and accommodation, which is useful for static 3D imagery in particular.

3.2.2 Consistent oculomotor cues – Matched vergence and accommodation

The difference between normal viewing and stereoscopic viewing with conventional 3D displays is illustrated in Fig. 3. Natural viewing provides real stimuli; the viewer is both fixated and focused on the object, i.e. accommodation distance and vergence distance are exactly matched. But the situation changes for stereoscopic 3D displays. Though the viewer is still fixated to the object with the same vergence as in natural viewing, his eye focus is now at the display and not where the object seems to be. That is because the viewer's eyes always focus on the brightest point or highest contrast. Hence with stereoscopic displays, the correlation between vergence and accommodation is disrupted[21,22] as depth remains an optical illusion. This mismatch requires the viewer to override the physiologically coupled oculomotor processes of vergence and accommodation that may cause visual discomfort and fatigue.[15] Although stereo 3D can work well for some applications, for example cinema with observing distance farther away or cell phones with short viewing time, it causes significant human factor risks for mainstream products as desktop monitors or TV displays.

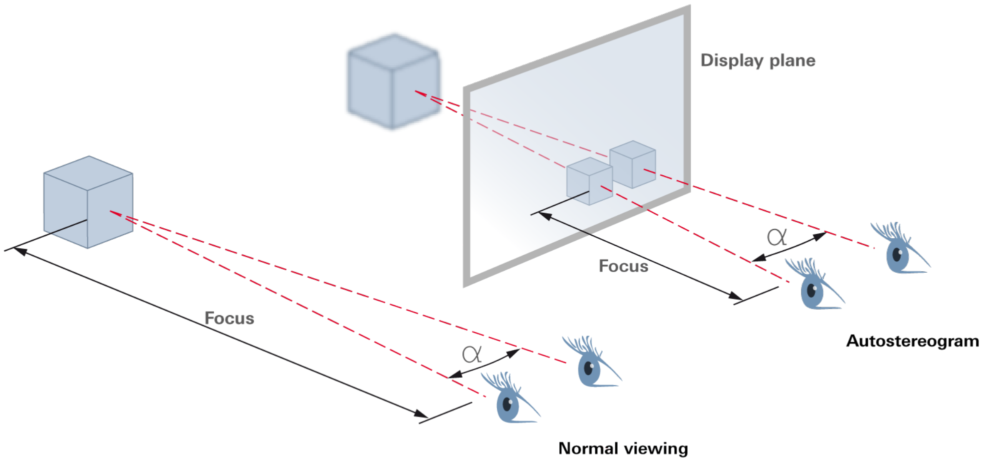


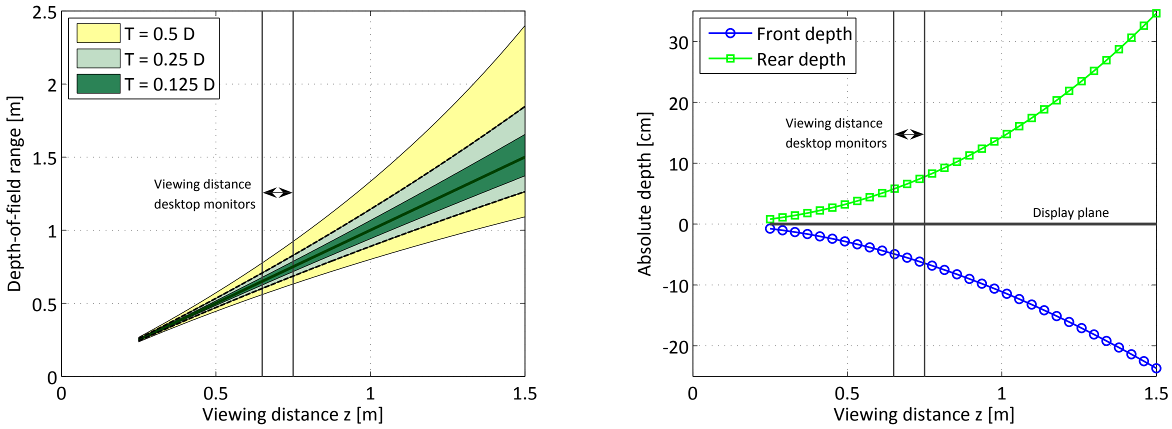
Figure 3. Comparison between natural viewing (left) and stereoscopic viewing with a 3D stereo display (right). Normal viewing and also holographic displays provide correct convergence and accommodation stimuli when fixing on an object. A stereogram only provides correct convergence whereas the eye focus remains unchanged to the display plane.

Because of its significant implications on 3D perception with displays,[23] numerous studies deal with the accommodation/convergence rivalry. Subjective[24] as well as objective assessments[25] have been carried out to understand and characterize visual fatigue and discomfort in order to minimize adverse effects and develop more qualified displays.[26] The oculomotor synergy is best characterized by AC/A and CA/C ratios.[25,27]

The inherent conflict common to all stereo displays between vergence and accommodation is a fundamental one that cannot be solved by technological means. The usual trade-off for stereoscopic displays is to artificially reduce the depth of a scene while centering it close to the display plane. This area near the display plane is called the "comfort depth range" or "comfort zone" of stereoscopic displays and its non-rigorous definition is based on the assumption that the human visual system can tolerate a certain amount of mismatch between convergence and accommodation. It is argued that even with binocular displays a comfortable viewing may provided – if only depth is chosen carefully. This might be correct for 3D cinema theaters when 3D objects appear to be not closer than approx. 8 m and hence fixation is possible with un-accommodated, *emmetropic* eyes. But for stereo displays utilizing 3D depth within accommodative-effective distances of the eye, the issue of comfort depth range needs careful consideration.

In numerous previous publications, the comfort depth range of stereo 3D displays is assumed to be equal to the depth of field of the human eye (cf. section 2.2.1). Depth-of-field estimates again were derived from depth-of-focus values given in the physiological optics literature. However, in addition to the reported large individual variations, the underlying values were often determined for monocular vision. As demonstrated by Campbell, accommodation between both eyes is correlated and fluctuations are much more stable when a target is binocularly fixed at steady convergence.[28] Because of the accommodative dynamics, we therefore question the value of the widespread model of a stereo comfort zone deduced from a static, comparatively large depth of focus. Depending on how optimistic design recommendations for stereo displays have been made, even depth-of-focus tolerances up to one diopter have been suggested.[29] But past experience with stereo 3D displays has shown that the comfort zone is much narrower than the zone of clear single vision. Whoever had experienced stereo 3D displays for more than a few minutes may probably know the arisen discomfort. We would now like to consider how vision is corrected by prescription glasses or contact lenses for individuals having focus defects of *myopia* or *hyperopia*. The common increment of corrective glasses is $^1/_4$ diopter, in exceptional cases even $^1/_8$ diopter. For example, an individual developing either near- or far-sightedness will become aware of it when refractive error is approaching a tolerance level of either plus or minus 0.25 diopter. In all probability, clear vision is still possible for that person but visual acuity will be perceptibly improved when the refractive error is fully corrected. Keep in mind, moreover, that the visual system is quite adaptable. Alterations of the eye's optical system can be corrected to some extent by powerful image processing in the brain. Such perceptual adaption have been studied for example with up-down/left-right inverting glasses, but it equally happens to a person who gets a new eyeglass prescription (e.g. progressive addition lenses). But perceptual adaption is a gradually changing process. Depending on how large the alteration is, it takes up to several (straining) days or weeks to become accustomed to the change. What all this suggests is that most likely the human vision system is much more sensitive to uncommon visual stimuli. To avoid any overstraining and to address the dynamic nature of accommodation fluctuations, it would be therefore prudent to design 3D stereo displays such that the common optometry tolerance of 0.25 diopter is not exceeded. What this means for the utilizable 3D range of stereo displays is shown in the figures below. Fig. 4(a) shows the proximal and distal depth-of-field limits p and d for near range viewing distances $250\,\text{mm} \leq z \leq 1.5\,\text{m}$ for three different depth-of-focus tolerances T. Narrow green areas represent the here proposed "safety depth range", where the mismatch between convergence and accommodation will probably cause no visual impairments. The yellow area indicates the smooth transition to the depth range where clear single binocular vision is still possible even though with increasing effort. The absolute front and rear depth according to the 0.25 diopter criterion is illustrated in Fig. 4(b). For desktop displays with common viewing distances of 650...750 mm, the depth range with usable front depths of 5...6 cm and rear depths of 6...8 cm becomes quite small. Looking at current stereoscopic 3D displays and prototypes and even movie theaters it can also be observed that even the $^1/_4$ diopter is scarcely utilized, limiting usable depth range even further.

Practical consequence of the limited depth range of stereo displays is that the 3D scene has to be squeezed in depth if the scene depth is larger than the available depth range. Leaving aside the fact that non-proportional



(a)

(b)

Figure 4. Proposed “safety zone” for 3D stereo displays. Depicted are (a) 3D depth range and (b) absolute front or rear depth for $T = 0.25\,\mathrm{D}$ as a function of viewing distance z (eye-to-display distance; corresponds to accommodation distance). Assumed are a tolerable depth-of-focus ranges of $T = 0.125\,\mathrm{D}$ (dark green area) and $T = 0.25\,\mathrm{D}$ (green area). Visual discomfort can arise when yellow regions are reached.

depth representation isn’t acceptable in medical or scientific applications such depth squeezing might contribute to an uncommon visual sensation too.

Other trade-offs for stereo 3D involve content and viewer-related aspects such as insertion of artificial blur and region-of-interest mapping to depth planes near the display plane. Both approaches proceed from the assumption that a viewer is most attracted by special features of a displayed scene such as faces, main characters or moving objects. A simple way is to artificially blur those areas of a 3D scene, which are away from the depth plane where the object-of-interest is located. By applying a Gaussian blur filter to the fore- or background, an accommodation stimuli is supposed to be simulated by giving priority to the main object. In a more advanced way, the depth range is divided into regions of main interest and additional near and far range regions. The depth region of interest may then proportionally be mapped within the narrow comfort zone.[30,31] However, both artificial blurring and tailored region-of-interest depth mapping might cause conflicts as the viewer is imposed upon to look at the predetermined regions. As soon as the viewer tries to fixate other than the predetermined areas, those simulated depth cues will be perceived as unnatural.

For all efforts, the conflict between accommodation and convergence, which is inherent in all 3D stereo displays can only be weakened but not resolved. As summarized in Table 1, true focus cues are provided by volumetric and holographic displays only.

4. SEEREAL’S HOLOGRAPHIC DISPLAY TECHNOLOGY

We have developed and demonstrated a novel approach to real-time display holography based on a sub-hologram encoding technique and a tracked viewing-window technology.[32] The solution is capable to provide both consistent accommodation/convergence and motion-based depth cues and thus fulfills the observer’s expectations on real depth perception.[33] Moreover, it can be built by using existing technology.[34,35]

The fundamental idea of our concept is rather simple when considering holography from an information point-of-view. As pointed out above, all human visual acuity and perception is limited by the capabilities of the eye and its subsequent image processing in the brain. When considering the human vision system regarding to where the image of a natural environment is received by a viewer, it becomes clear that only a limited angular spectrum of any object contributes to the retinal image. In fact, it is limited by the pupil’s aperture of some millimeters. If the positions of both eyes are known, it therefore would be wasteful to reconstruct a holographic scene or object that has an extended angular spectrum as it is common practice in classical holography. The key idea of our solution to electro-holography is to reconstruct a limited angular spectrum of the wave field of the 3D object,

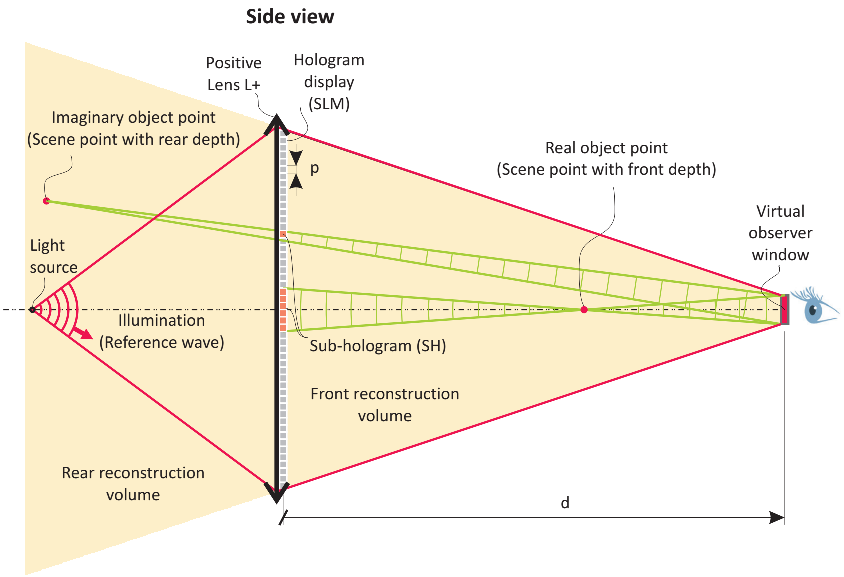
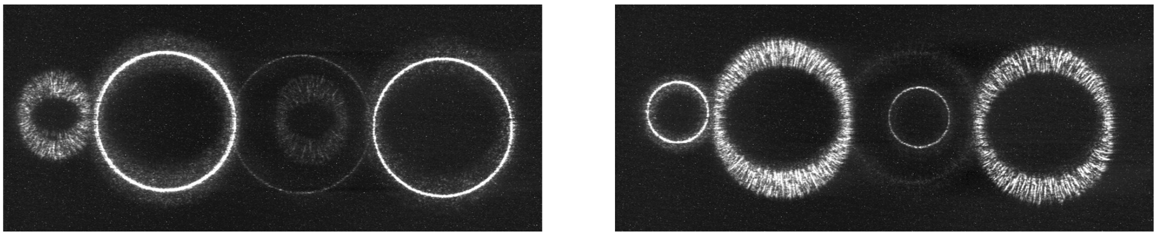


Figure 5. Schematic principle of the sub-hologram concept (side view). L+, positive lens; SLM, spatial light modulator; SH, sub-hologram. Tracking means are not shown.

which is adapted in size to about the human's eye entrance pupil. According to our solution, the highest priority is to reconstruct the wave field at the observer's eyes and not the three-dimensional object itself. The designated area in the viewing plane, i.e. the virtual *Viewing Window* from which an observer can see the proper holographic reconstruction, is located at the Fourier plane of the holographic display. The holographic code (i.e. the complex amplitude transmittance) of each scene point is encoded on a designated area on the hologram that is limited in size. This area in the hologram plane is called a *Sub-Hologram*. There is one sub-hologram per scene point, but owing to the diffractive nature of holography, sub-holograms of different object points are super-positioned without loss of information. Binocular parallax is provided by delivering different holographic reconstructions with the proper difference in perspective to left and right eye, respectively. For this, the techniques of spatial or temporal multiplexing can be utilized. Fortunately, dynamic or real-time video holography offers an additional degree of freedom with regard to quick hologram update. By incorporating a tracking system, which detects the eye positions of one or more viewers very fast and precisely and repositions the viewing window accordingly, a dynamic 3D holographic display providing full motion parallax can be realized. This way, all problems involved with the classic approach to holography can be circumvented. The steering of the viewing window can be done in different ways, utilizing for example light source shifting or beam-steering principles. The fundamentals of our holographic 3D display technology are schematically summarized in Fig. 5.


(a)

(b)

Figure 6. Photographs of holographic reconstructions taken with camera focus at different depth planes. (a) Focus 50 cm in front of display plane, (b) 50 cm behind display plane. (Note that the circles are actually in different colors, however the camera is monochrome only.)

To demonstrate the focus cue capabilities of our holographic display technology, we show in Fig. 6 the reconstruction of a 3D object that is extended in depth. The photographs were taken at a 8-inch holographic projection display.[36] The images were captured with the camera lens positioned in the virtual observer window dedicated for one eye. The 3D object consists of circles represented at two depth planes, which are separated by 1 m. Three large circles of different brightness are imaged sharply as the camera focus was set to 50 cm in front of the display plane, see Fig. 6(a). As the camera focuses through the scene to a distance 50 cm behind the display, large circles become blurred, whereas two small circles come in focus, Fig. 6(b). It is evident that the circles are reconstructed at different depths. Only the circles on which the focus of the camera lens is set are in focus, whereas the circles at other depths are not in focus. The same also holds for an observer eye: an observer perceives parts of the 3D object that are in focus, whereas other parts of the 3D object are not in focus. Thus, depth cue of focus information can be optimally provided by a holographic display.

5. CONCLUDING REMARKS

While lack of technical progress has hindered mass-introduction of stereoscopic 3D displays to the consumer markets until recently, visual conflicts will have strong impact on future acceptance and success. Inherent visual conflicts of stereo 3D to natural viewing are always going to limit capabilities and features if not market saturation. Non-consistent convergence-accommodation cues and unnatural motion parallax cause less viewing comfort and will hinder the general use of stereoscopic displays, especially for use at short viewing distances. The 3D scene depth of stereoscopic displays has to be limited in order to keep this mismatch at an acceptable level. We have proposed a “safety depth zone” for stereo displays, which has to compromise considerably on utilizable depth. Holographic displays, on the other hand, can show natural 3D scenes with unlimited depth within a non-enclosed 3D volume enabling a grasp experience. Therefore, we consider holography as the preferred option for 3D displays. The progress in full-color real-time holographic displays demonstrates that holography is technically feasible with a broad range of applications in all kinds of computer displays and TV.

Which 3D technology becomes finally successful, will be crucially depend upon its performance. Above all, they will be judged how realistically natural viewing can be mimicked. One important step along these lines is to establish objective assessment and standardized evaluation criteria, which are universally applicable for all types of 3D displays.

REFERENCES

1. M. J. Tovée, *An introduction to the visual system*, Cambridge Univ. Press, 1996.
2. M. Lambooij, W. IJsselsteijn, M. Fortuin, and I. Heynderickx, “Visual discomfort and visual fatigue of stereoscopic displays: a review,” *Journal of Imaging Science and Technology* **53**, pp. 030201–1–030201–14, 2009.
3. K. N. Ogle and J. T. Schwartz, “Depth of focus of the human eye,” *J. Opt. Soc. Am.* **49**(3), pp. 273–280, 1959.
4. F. W. Campbell, “The depth of field of the human eye,” *Opt. Acta* **4**(4), pp. 157–164, 1957.
5. F. W. Campbell and D. G. Green, “Optical and retinal factors affecting visual resolution.,” *The Journal of Physiology* **181**(3), pp. 576–593, 1965.
6. D. G. Green, M. K. Powers, and M. S. Banks, “Depth of focus, eye size and visual acuity,” *Vision Research* **20**(10), pp. 827–835, 1980.
7. K. N. Ogle, “Some aspects of stereoscopic depth perception,” *J. Opt. Soc. Am.* **57**(9), pp. 1073–1081, 1967.
8. S. Pastoor and M. Wöpking, “3-D displays: A review of current technologies,” *Displays* **17**(2), pp. 100–110, 1997.
9. I. Sexton and P. Surman, “Stereoscopic and autostereoscopic display systems,” *IEEE Signal Processing Magazine* **16**(3), pp. 85–99, 1999.
10. B. Javidi and F. Okano, *Three-dimensional television, video and display technology*, Springer Verlag, 2002.
11. N. Holliman, *Handbook of Optoelectronics*, vol. I, ch. Three-Dimensional Display Systems, pp. 1067–1100. Taylor and Francis, 2006.
12. P. Benzie, J. Watson, P. Surman, I. Rakkolainen, K. Hopf, H. Urey, V. Sainov, and C. von Kopylow, “A survey of 3DTV displays: techniques and technologies,” *IEEE Transactions on Circuits and Systems for Video Technology* **17**(11), pp. 1647–1658, 2007.

13. B. Lee and J. Park, "Overview of 3D/2D switchable liquid crystal display technologies," in *Proceedings of SPIE*, **7618**, p. 761806, 2010.
14. C. Wheatstone, "On some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philosophical Transactions of the Royal Society of London* **128**, pp. 371–394, 1838.
15. J. P. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," *Vision Research* **35**(19), pp. 2731–2736, 1995.
16. L. M. J. Meesters, W. A. IJsselsteijn, and P. J. H. Seuntiens, "A survey of perceptual evaluations and requirements of three-dimensional TV," *IEEE Transactions on Circuits and Systems for Video Technology* **14**(3), pp. 381–391, 2004.
17. E. J. Gibson, J. J. Gibson, O. W. Smith, and H. Flock, "Motion parallax as a determinant of perceived depth," *Journal of Experimental Psychology* **58**(1), pp. 40–51, 1959.
18. M. Ono, J. Rivest, and H. Ono, "Depth perception as a function of motion parallax and absolute-distance information," *Journal of Experimental Psychology: Human Perception and Performance* **12**(3), pp. 331–337, 1986.
19. H. von Helmholtz, *Handbuch der physiologischen Optik*, Leopold Voss, Leipzig, 1867.
20. B. J. Rogers and M. Graham, "Similarities between motion parallax and stereopsis in human depth perception," *Vision Research* **22**(2), pp. 261–270, 1982.
21. S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks, "Focus cues affect perceived depth," *Journal of Vision* **5**(10), pp. 834–862, 2005.
22. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue," *Journal of Vision* **8**(3), pp. 1–30, 2008.
23. T. Inoue and H. Ohzu, "Accommodative responses to stereoscopic three-dimensional display," *Appl. Opt.* **36**(19), pp. 4509–4515, 1997.
24. M. Emoto, T. Niida, and F. Okano, "Repeated vergence adaptation causes the decline of visual functions in watching stereoscopic television," *Journal of Display Technology* **1**(2), p. 328, 2005.
25. T. Fukushima, M. Torii, K. Ukai, J. Wolffsohn, and B. Gilmartin, "The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images," *Journal of Vision* **9**(13), pp. 1–13, 2009.
26. T. Shibata, "Stereoscopic 3-D display with optical correction for the reduction of the discrepancy between accommodation and convergence," *Journal of the SID* , p. 665, 2005.
27. H. Ripps, N. B. Chin, I. M. Siegel, and G. M. Breinin, "The effect of pupil size on accommodation, convergence, and the AC/A ratio," *Investigative Ophthalmology & Visual Science* **1**(1), p. 127, 1962.
28. F. W. Campbell, "Correlation of accommodation between the two eyes," *J. Opt. Soc. Am.* **50**(7), p. 738, 1960.
29. R. Patterson, "Human factors of stereo display: An update," *Journal of the SID* **17**(12), pp. 987–996, 2009.
30. N. S. Holliman, "Mapping perceived depth to regions of interest in stereoscopic images," in *Stereoscopic Displays and Virtual Reality Systems XI*, A. J. Woods, J. O. Merritt, S. A. Benton, and M. T. Bolas, eds., **5291**(1), pp. 117–128, SPIE, 2004.
31. N. Holliman, "Smoothing region boundaries in variable depth mapping for real-time stereoscopic images," **5664**, pp. 281–292, 2005.
32. A. Schwerdtner, N. Leister, and R. Häussler, "A new approach to electro-holography for TV and projection displays," in *SID-Proc.*, p. 32.3, 2007.
33. R. Häussler, A. Schwerdtner, and N. Leister, "Large holographic displays as an alternative to stereoscopic displays," *Stereoscopic Displays and Applications XIX* **6803**(1), p. 68030M, SPIE, 2008.
34. S. Reichelt, R. Häussler, N. Leister, G. Fütterer, and A. Schwerdtner, "Large Holographic 3D Displays for Tomorrow's TV and Monitors – Solutions, Challenges, and Prospects," in *Proc. 21st Annual Meeting of the IEEE Lasers and Electro-Optics Society LEOS 2008*, pp. 194 – 195, 2008. (invited).
35. S. Reichelt, R. Häussler, N. Leister, G. Fütterer, H. Stolle, and A. Schwerdtner, *Advances in Lasers and Electro Optics*, ch. Holographic 3-D Displays – Electro-holography within the grasp of commercialization, pp. 1–28. IN-TECH, 2010. (in press).
36. A. Schwerdtner, J.-C. Olaya, R. Häussler, and N. Leister, "Digital holography - Principles and challenges of holographic projection systems," in *IMID-Proc.*, 2007.