3D
Displays

Ernst Lueder

3D Displays

Ernst Lueder
University of Stuttgart, Germany & Independent Consultant, USA

Preface

Flat panel display technology and manufacture have now reached the level of maturity required to introduce 3D displays to the marketplace. The book covers five approaches to realize 3D perception, namely stereoscopic and autostereoscopic displays, integral imaging, holography and volumetric displays.

I owe thanks to Dr. Tony Lowe who with his thorough understanding of scientific trends very much supported the book on 3D technologies. I very much profited from Dan Schott's excellent knowledge about flat panel display technologies and I am very grateful for that. Based on his profound evaluation of new display technologies, Dr. Christof Zeile drew my attention to various new publications. I very much appreciate his support.

I would also like to express my appreciation of the excellent work performed by the typesetters.

The competent contribution to the index by Neil Manley is gratefully acknowledged.

As in earlier books, I am greatly indebted to Heidi Schuehle for diligently and observantly typing the manuscript and to Rene Troeger for the professional and accomplished drawing of the figures.

<div align="right">

Ernst Lueder

Scottsdale, USA, October 2011

</div>

Series Preface

Professor Lueder wrote his first book "Liquid Crystal Displays" for the Wiley-SID Series in Display Technology in the year 2000. That book went on to become the best seller in the entire series and is now in its second edition. I am therefore delighted to be writing a foreword to Ernst Lueder's newest work, this time on the topical subject of 3D Displays.

Most sighted human beings have a perception of what 3D means. We are familiar with what we see around us, that we perceive some objects to be nearer than others, that distant objects traversing our field of view appear to move more slowly than and are obscured by those nearer to us, and so on. A smaller but growing fraction of the population is familiar with 3D movies and television. However, a majority will have only a vague understanding of how our brains operate on visual stimuli to create our familiar three-dimensional view of the world. When it comes to creating 3D images on displays, further levels of complexity are required not only to avoid eye strain by displaying inconsistent or misleading visual cues, but to process prodigiously large quantities of data at sufficient speeds to enable real-time 3D visualisation.

This book sets out to present its subject in a manner which places it on a sound mathematical basis. After an overview of the physiology of 3D perception, there follow detailed descriptions of stereoscopic and autostereoscopic displays which are, after all, the most developed of 3D display technologies. Much attention is given to the synthesis of 3D from 2D content, a most important topic, given the quantity of 2D content already available. Quality issues are addressed next, with particular emphasis on methods to improve the visual quality of 3D imagery and to reduce the bandwidth required to transmit it, with special emphasis on a method known as depth image-based rendering. The book then describes three types of displays (integral imaging, holography and volumetric displays) which, although less developed than stereoscopic and autostereoscopic displays, are able to present real three-dimensional images in which the view changes - with nearer objects obscuring more distant ones - as the viewer changes position. This is in contrast to providing a mere illusion of three-dimensionality, as is the case with many stereoscopic images.

The book concludes with a chapter aptly named "A Shot at the Assessment of 3D Technologies" This is not so much a guess at what is coming next, but rather a logical *in futuro* extension of the technologies and methods already described and, to my reading, a credible one.

This is a complete book, full of the necessary equations, with many illustrations and replete with references. The subject matter, whilst complex, is very clearly presented and will provide readers with a sound technical basis from which to develop their skills further into the exciting field of three-dimensional display science.

Anthony Lowe
Braishfield, UK, 2011

Introduction

The design and manufacture of displays are now mature enough to introduce three-dimensional (3D) displays into the marketplace. This happened first with displays for mobile devices in the form of near-to-the-eye displays, but home TV will follow suit.

This book covers five approaches to realize 3D perception, namely, stereoscopic and autostereoscopic displays, integral imaging, holography, and volumetric displays.

The intention guiding the book is to promote a well-founded understanding of the electro-optic effects of 3D systems and of the addressing circuits. Equations are as a rule not simply stated but are derived, or, if not fully done so, at least hints for the derivation are given. An example of this concept is the explanation of the basics of holography by phasors, which will be outlined, but which are also known from electrical engineering or from the Jones vector. This renders complex facts associated with holograms easier to understand.

Emphasis is placed on stereoscopic and autostereoscopic displays as they are closest to being commercialized. The basic components of stereoscopic displays are patterned retarders and to a lesser degree wire grid polarizers. Autostereoscopic displays rely on beam splitters, lenticular lenses, parallax barriers, light guides and various types of 3D films. All of these elements are explained in detail.

The glasses required for stereoscopic displays distinguish between the left and the right eye views either by shutters or by circular polarization. Linearly polarized glasses have the disadvantage of being sensitive to tilting of the head.

Special attention is given to 3D systems working in a spatial or temporal multiplex, as well as in a combination of the two, and to novel fast addressing schemes. In order to suppress crosstalk and blur, a 240 Hz frame rate is preferred. The increased speed of addressing is handled by parallel processing and by the recently published interleaved addressing, which also parallels the images. Special care is taken to outline how the autostereoscopic approach is able to provide side views, the perspectives, of the object.

This paves the way for an understanding of integral images (IIs) with a pickup stage for information similar to the lenticular lenses of the autostereoscopic displays. Very naturally this leads to the ingenious design of an II projector working with real and virtual images where the viewer can walk around the displayed object, thus enjoying a first solution for a true 3D display.

The chapter on holography leads the reader on to digital computer-generated holography, which is not yet a real-time process.

Volumetric displays consist of a stack of LCDs, each of which is devoted to a particular depth, where also the limitations of the fusion of the images become noticeable.

Notably, Chapter 4 is devoted to familiarizing designers of flat panel displays with the work done by computer scientists on the assessment and improvement of 3D image quality. Algorithms are introduced for evaluating the properties of 3D displays based on objective and subjective criteria and on tracking the motion of selected special features. Special attention is drawn to establishing disparity maps and preparing a 3D image ready for transmission with a bandwidth-saving "depth image - based rendering" (DIBR). Head tracking for 3D reception by a group of single viewers is not included.

Contents

1

The Physiology of 3D Perception

1.1 Binocular Viewing or Human Stereopsis

As one eye is capable only of perceiving a planar image, 3D viewing is commonly achieved by the cooperation of both eyes in providing each eye with a view of the object. The images that the eyes receive from the same object are different according to the different locations of the eyes. This binocular viewing provides the perception of depth, the third dimension, as further explained by the horopter circle in Figure 1.1. This circle serves as a reference from which the depth is determined [1,2]. If the eyes are focusing, for which the synonyms fixating, accommodating, or converging are also used, on point M on the horopter circle, the ciliary muscles of the eyes rotate the eyeballs into such a position that the light from M passes the pupils parallel to the axes of the lenses in the eyes. The axes intersect at M. Then the light hits the retina in Figure 1.1 at the foveas m_l for the left eye and m_r for the right eye. The foveas are in the center of the retina and exhibit the highest density of light receptors. The rotation of the eyes is called the vergence. Obviously the axes of the eyes are no longer parallel, which will provide the depth information required by the brain [1,3]. In this situation light from point P hits the retinas at the points p_l for the left eye and p_r for the right eye. The angles α at the periphery of the circle are, as is known from geometry, the same for all points P on the circle above the distance b between the pupils. As a consequence, also all the angles γ for points on the horopter circle are equal [4]. The angle γ at the retina, measured as a rule in arcmin, is called the disparity or the parallax. As all the points M and P on the horopter circle have the same disparity γ in both eyes, the difference d in the disparities of all points on this circle is zero. The further P is away from M, but still on the horopter circle, the larger is the disparity [2,3]. Obviously the larger disparity is associated with a smaller depth. The disparity information is transferred to the brain, which translates it into a perceived depth. How the brain fuses the two disparities into a 3D image is not yet fully understood.

As all points on the horopter circle exhibit a zero difference in disparities, the circle serves as a reference for the depth. The fusion of the disparities and the depth perception as described works only in Panum's fusional area in Figure 1.1 [3]. In this area, reliable depth perception decreases monotonically with

---
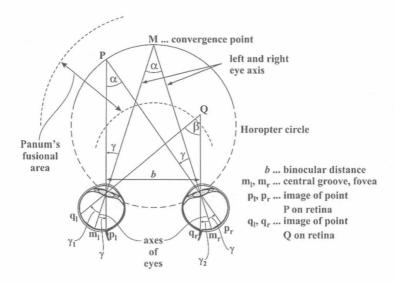
Figure 1.1 Horopter circle.

increasing magnitude of the disparity. This relationship is called the patent stereopsis. For a point Q in Figure 1.1 [3] not on the horopter circle but closer to the eyes and still in Panum's area, the disparities on the retina are given by the points q_l for the left eye and q_r for the right eye with the disparities γ_1 and γ_2. These points lie across the fovea on the other side of the retina and exhibit a so-called crossed disparity, while the points farther away than the horopter have an uncrossed disparity. Their image points corresponding to q_r and q_l for crossed disparities lie on the opposite side of the fovea.

For point Q the disparities γ_1 and γ_2 are no longer equal. The value $\gamma_1 - \gamma_2 \neq 0$ together with the disparities themselves provide information to the brain on how much the depth of Q is different from the depth on the horopter. However, how the brain copes with this difference of disparities is again not fully known.

When moving an object from the horopter closer to the eye, the patent stereopsis is finally lost at a distance of around 2 m or less from the eyes. Fusion of the images may no longer work and double images, called diplopia, appear [3]. Due to overlarge disparities, the eyes perceive the object they are trying to accommodate and its background separately. The brain unsuccessfully tries to suppress the background information. On the other hand, the further away from the horopter the object is, the smaller is the disparity, because the axes of the lenses become closer to being parallel. Finally, at distances beyond about 10 m the differences between the small disparities can no longer be resolved and the depth information is lost. This coincides with our inability to estimate the difference in depth of objects that are too far away.

The average distance b of the pupils in Figure 1.1 of adults in the USA is 6.5 cm, and for 90% of these adults it lies between 6 and 7 cm [5]. The total range of disparity is about 80 arcmin for the perception of spatial frequencies from 2 to 20 cycles per degree and about 8 arcdegrees for low spatial frequencies around 0.1 cycles per degree [3]. This means that for low spatial frequencies larger disparities are available than for larger spatial frequencies. As a consequence, the sensitivity of disparities for low spatial frequencies is larger than for larger spatial frequencies. The same facts apply also for lower and larger temporal frequencies of the luminance in an image.

The smallest still recognizable disparity, the stereoacuity D_{min}, is 20 arcsec in the spatial frequency range of about 2–20 cycles per degree, while the maximum perceivable disparity D_{max} is 40 arcmin for low spatial frequencies [3]. As the values for D_{min} and D_{max} apply to both the crossed and uncrossed disparities standing for different ranges of depths, the values can be added to a total of 80 arcmin for high and 8 arcdegrees for low spatial frequencies, as already given above [6,7]. Again this is also true for temporal

frequencies in dynamic images with a larger sensitivity of disparities for lower temporal frequencies and a lower sensitivity for larger temporal frequencies of luminance.

There are two visual pathways from the retina to the brain. The parvocellular-dominated dorsal–cortical path connects the central retina to the ventral–cortical areas in the visual cortex where spatial patterns and color are analyzed. The magno-cellular-dominated dorsal–cortical path leads from the central and peripheral retina to dorsal–cortical areas in the visual cortex, where optical flow information for heading control and biological motion are investigated. Further information on these paths can be found in [8–10].

The stereoanomalies are associated with defects in these paths of information where there are neurons sensitive to only crossed or uncrossed disparities. The perception of depth is thought to involve responses from both types of neurons. In stereoanomalous individuals, one type of these neurons fails to be sensitive to their information. Then the other type of neurons dominates the response to all disparity information. In the case where neurons are only sensitive to uncrossed disparities belonging to objects located further away than the horopter circle, the information from crossed disparities stemming from objects closer to the eye than the horopter is suppressed in favor of objects far away. The individual perceives the close-up information as far away information with a far away depth. When the neurons are only sensitive to crossed disparities, the individual perceives the far away information with a depth close to the eye [11,12].

Individuals who are stereoblind, as a rule resulting from a disease called strabismus, are assumed to be entirely lacking in disparity-sensitive neurons.

Under degraded stimulus conditions such as brief stimulus exposure, stereoanomalies are found in 30% of the population [13]. In addition, 6–8% of the population are stereoblind. The relatively large percentage of people incapable of perceiving a 3D image would merit more attention.

Another physiological disturbance is binocular rivalry. In this case an individual views a stereo display with a very large disparity or with interocular misalignment or distortion such that no fusion of the two eyes' image takes place [7,14]. One eye inhibits the visual activities of the other eye. One view may be visible, as the other eye's view is suppressed, which reverses over time. This is a problem which may be experienced with headworn displays, where two images from different sources may be misaligned or distorted [15].

Two physiological stimuli of depth can be detected by one eye alone. These are disparity and motion parallax. Under this parallax the shift of a moving object toward a still background is understood. The eye together with the brain extracts from this parallax a 3D perception with an associated depth.

Similar to motion parallax is Pulfrich's phenomenon [16]. One eye is covered with a filter which darkens the image. The processing of the dark image is delayed in relation to the processing of the bright image. This leads to disparity errors when the viewer moves relative to an object. However, it can also be used to provide a depth cue, as the delay renders the two eyes' images differently as usually caused by depth.

1.2 The Mismatch of Accommodation and Disparity and the Depths of Focus and of Field

Now we are ready to consider a phenomenon explicable with known stereoptic facts. As we shall see later, in stereoscopic and autostereoscopic displays the two required views of an object are presented next to each other on the screen of a display. The distance to the eyes of the viewer is constant for all scenes displayed. That is the cause of a problem, as the eyes accommodate to the two images with a vergence associated with the disparity. The disparity stimulates a depth perception in the brain. On the other hand, the accommodation of points on the screen also conveys depth information, which is the constant distance to the screen. The two depth details are contradictory, and are called the mismatch of accommodation and vergence or disparity. This may cause discomfort for viewers, manifested by eyestrain, blurred vision, or a slight headache [7]. Fortunately the problems stemming from this mismatch are experienced mainly for short viewing distances of around 0.5 m. A quick and obvious explanation is the already mentioned fact that for larger distances the disparities become smaller and are crowded together on the retina, so the
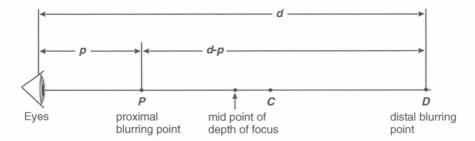
Figure 1.2 Depth of focus and depth of field.

resolution of depth associated with disparity is diminished. Therefore the depth information based on disparity no longer changes much with increasing distances and is more easily matched with the depth information based on accommodation. In practice it was found that a viewing distance of 2 m or more from a TV screen no longer leads to annoying discomfort [7].

A more thorough explanation is derived from the depth of focus and the depth of field, which is also important for the design of a 3D system for moving viewers [17]. We assume that the eyes have focused on an object at point C in Figure 1.2, providing a sharp image. The depth of focus describes the range of distance from a point P nearer to the eye than C to a point D further away than C in which an object can still be detected by applying a given criterion for detection. If the distance of point P is p and that of D is d then the depth of focus T in diopters is

$$T = \frac{1}{p} - \frac{1}{d} \tag{1.1}$$

where p and d are expressed in m. The depth of field is

$$F = d - p \tag{1.2}$$

also in m.

Diopters are defined by $1/f$, where f is the focal length of a lens in m; in our case the lens is the eye with that f where the eyes experience a sharp image.

Possible criteria for the detectability of features in a display are:

(a) the deterioration of visual acuity or of resolving power;
(b) the discrimination of least perceptible blurring of the image;
(c) the loss of visibility or detectability of target details through loss of contrast; and
(d) the perceptual tolerance to out-of-focus blur which results in a stimulus for a change in accommodation.

The first three criteria depend on the perception of out-of-tolerance blur, while the last one depends on physiological tolerance. Point P is called the proximal blurring point, while D is the distal blurring point. Below P and beyond D the image is no longer accepted.

The results reported now are based on criterion (a) and the out-of-focus blur in criterion (d) [17]. A checkerboard test pattern is used and test persons provide the percentage of correct answers in detecting the correct pattern. The test pattern had a size of 1.25 arcmin corresponding to a Snellen notation of 20/25. The diameter of the pupils was 4.6 mm. The test result is shown in Figure 1.3. The abscissa represents the displacement of the test pattern from the fixation point C measured in diopters. Hence the abscissa indicates in diopters the degree to which the test pattern is out of focus. The ordinate represents the percentage of the correct visual resolution perceived for the test pattern. This percentage exhibits a Gaussian probability density.
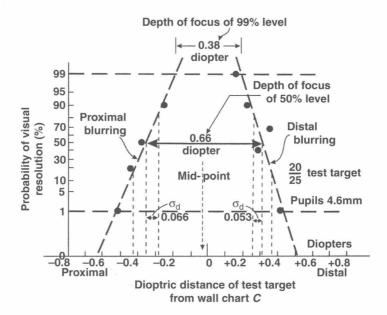
Figure 1.3 Percentage of correct resolution perceived versus displacement of the test pattern from the fixation point C in Figure 1.2.

The midpoint of the depth of focus is always slightly nearer to the eye than the focus point C.

For a 50% correct visual resolution, the depth of focus has a width of 0.66 diopters, whereas for 99% the width shrinks to 0.38 diopters. This shrinking is about 0.06 diopters for an increase in the visual resolution of 10% of the proximal blurring. The depth of focus at the 99% level is an important one for the out-of-focus blur at which the visual resolution begins to deteriorate.

The diagram in Figure 1.3 depends upon the location of the fixation point C. This is evident from Table 1.1 with measured distances for the fixation point C in m, the distances p of the proximal and d of the distal blur also in m, as well as the resulting depth of focus T in diopters. Only if T were constant for all points C would the diagram be independent of the location of C. The fixation point C for the diagram in Figure 1.3 is about 1 m from the eye. The depth of field, $d - p$, in m increases with increasing distance to the fixation point C; it can even become infinite.

Further results in [17] relate to the influence of luminance, pupil diameter, and size of object in arcmin on the depth of focus. The larger the luminance, the smaller the diameter of the pupil. At 0.03 cd/m^2 the diameter is 6 mm, at 30 cd/m^2 it is 3 mm, and at 300 cd/m^2 only 2 mm. A linear decrease in the diameter of the pupil is associated with a logarithmic increase in luminance. For a 1 mm decrease of this diameter the depth of focus increases by 0.12 diopters.

For an increase in the object by 0.25 arcmin the depth of focus increases by 0.35 diopters. At a size of 2 arcmin the depth of focus reaches 2 diopters.

Table 1.1 Dependence of proximal and distal blur as well as depth of focus T on location of C

Distance of fixation point C in m	Distance of proximal blur p in m	Distance of distal blur d in m	Depth of focus T in diopters
1	0.75	1.5	$1333 - 0.666 = 0.667$
2	1	5	$1 - 0.2 = 0.8$
3	1.5	∞	0.666

Table 1.2   Newer values for those in Table 1.1

Distance of fixation point C	Distance of low end of depth of field	Distance of high end of depth of field
0.5	0.4	0.67
1	0.67	z
2	1	∞

The results in Figure 1.3 are very important for those 3D displays where the viewer only has a sharp picture at a given distance from the screen. Figure 1.3 reveals how much the viewer has to move backward and forward while still perceiving an acceptable image.

Newer values for the depth of field depending on the distance of the fixation point C are given in Table 1.2 [18]. Obviously the depth of field increases strongly with increasing distance of the fixation point. So fixation or accommodation on a given point is no longer so important for larger distances. As a consequence for the mismatch of accommodation and disparity, accommodation plays a minor role which also alleviates discomfort. This is no longer true for a fixation point at 0.5 m or closer, meaning that discomfort certainly is a problem for near-to-the-eye displays. For regular 3D displays a viewing distance of at least 2 m should sufficiently minimize discomfort, as already stated above.

In view of this result, discomfort when viewing 3D movies from larger distances should not occur as a rule. This, however, is not the case, because there is a different effect identified as the cause of discomfort, as discussed in Section 1.6.

Stereoscopic and autostereoscopic displays provide only an illusion of 3D perception. This is among other effects due to the difficulty stemming from the mismatch of accommodation and disparity, resulting in a conflict of depth perception. Contrary to this, integral imaging, holography, and volumetric displays, which will be treated later, do not exhibit this mismatch. There, the viewer, when moving, has the impression of walking around the 3D object, thus experiencing true 3D. On the other hand the viewer would always see the same image in the case of stereoscopic solutions.

1.3   Distance Scaling of Disparity

In stereopsis there are two definitions of perceived distance or depth. The egocentric view refers to the conventional distance D between an observer and an object and is usually measured in m. On the other hand, relative depth is based on the depth interval between a viewer and the reference point on the horopter circle and is measured in radians of the disparity γ on the retina in Figure 1.1. The disparity information γ is connected to D by a strongly nonlinear relation stemming from the geometry shown in Figure 1.1. This relation has to be differently approximated or recalibrated or, in other words, scaled for different regions of distance D [19,20].

For obtaining a veridical or true value, egocentric distance information D together with the relative depth γ are needed by the brain. It is assumed that the brain combines binocular disparity γ with egocentric distance cues for the process of disparity scaling.

For a large distance D in real-world scenery, the magnitude of the disparity γ varies, as we have seen intuitively from the geometry in Figure 1.1, approximately with the inverse of D^2. It was found that γ is also proportional to the interpupillary distance b. This leads to the equation

$$\gamma = \frac{b\,d_0}{D^2} \tag{1.3}$$

in which d_0, with the dimensions of cm arcmin, is an experimentally determined proportionality factor, called the depth interval and sometimes also the predicted depth [21]; d_0 is different for each D and is approximated by a constant in an interval around D.

In stereoscopic displays the disparity γ depends approximately on the separation S between the two images required for 3D perception and inversely on the viewing distance D. Thus

$$\gamma = S/D. \tag{1.4}$$

For Equation 1.4, a symmetrical convergence and targets on the retina close to the midsaggital plane symmetrically dividing the body are required.

For a fixed value of S a stereoscopic display will provide the depth interval d_0 as

$$d_0 = \frac{DS}{b \pm S}. \tag{1.5}$$

In the denominator the $+$ sign applies for crossed disparities and the $-$ sign for uncrossed disparities. In [22] it was found that this determination of d_0 is very robust toward unequal luminances in the two views. If the luminance exceeds 0.63 cd/m^2 a 60% difference in the interocular luminance does not harm the perception of the depth interval d_0. However, for greater interocular luminance differences the perceived depth may be far away from the correct value. Another luminance-related effect is the discomfort created by interocular luminance differences of more than 25% [22]. Finally, interocular differences in contrast of up to 83% did not affect depth perception, while the threshold for discomfort in this case was between 25% and 50%.

1.4 Interocular Crosstalk

Information leaking from the view in one eye into that of the other eye is known as crosstalk, which as a rule severely damages the quality of the perceived image but can also affect the fusion of the two images. At no crosstalk the fusion is limited by 27 arcmin for crossed disparity and by 24 arcmin for uncrossed disparity. For a 200 ms stimulus, crosstalk has only a small effect on fusion, which is no longer true for a 2 s stimulus [23]. In this case, 2–7% crosstalk can already hamper fusion and can cause discomfort [24].

Autostereoscopic displays may apply spatial multiplexing of the two views, for which an array of lenticular lenses or parallax barriers is used. Lenticular lenses exhibit chromatic aberrations, while barriers produce diffraction by which image content can leak into the wrong eye. The remedy is to limit aberration and diffraction at least for a given position of the viewer.

For stereoscopic and autostereoscopic displays with temporal multiplexing, crosstalk occurs due to the persistence of a display, in which the image content of one eye's view is still visible in the next frame when that eye is exposed to a new view. This is shown in Figure 1.4. Temporal multiplexing can also induce flicker seen in the visual periphery. This disrupts vision in large field-of-view immersive displays. The cause is that these displays stimulate the magno-cellular-dominated dorsal–cortical area, which draws connections from the peripheral retina, and above all have a transient response and high temporal acuity, perceived as flicker. A remedy is a high frame rate enabling the visual system to integrate the intermittent information in the periphery [6].

A further, very strong source of crosstalk is blurring of the edges of a moving image. Blur occurs in all displays where the luminance of the image is held constant during the entire frame time. This occurs in liquid crystal displays (LCDs) and in organic light-emitting diode displays (OLED displays). A relatively brief description of this important phenomenon is given here, while a more detailed one can be found on pages 298–300 of [25].

Blur is explained in Figure 1.5a, where a black stripe at rest on an LCD screen can be seen, while Figure 1.5b shows the stripe moving to the right. The edges of the stripe in Figure 1.5a are perfectly sharp but are blurred by the movement in Figure 1.5b. The main cause is that an image on an LCD is held constant during the frame time T_f, which for a frame frequency of $f = 60$ Hz is given by $T = 1/f = 16.66$ ms.