Retinal projection displays for accommodation-insensitive viewing

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Doctor of Sciences

presented by

Marc von Waldkirch
Dipl. Phys. ETH
born 16th November 1974
citizen of Schaffhausen SH and Zürich ZH

accepted on the recommendation of

Prof. Dr. Gerhard Tröster, examiner
PD Dr. Marino Menozzi, co-examiner

2005

Vorwort des Herausgebers

Die Kombination artifizieller mit realer Information ist eine zukunftsträchtige Technik, die z. B. als Erweiterung sinnesphysiologischer oder kognitiver Fähigkeiten eines Anwenders eingesetzt wird. Schwerpunktmässig liegen hierzu die Forschungs- und Anwendungsarbeiten im Bereich der Erweiterung der visuellen Funktion. Die Kombinationstechnik hat durch das Schlagwort 'augmented reality' einen gewissen Bekanntheitsgrad erreicht.

Sowohl die reale als auch die artifizielle Information werden in augmented reality durch dasselbe dioptrische System des Auges betrachtet. Die scharfe Abbildung der realen und der artifiziellen Objekte ins Auge bedingt entweder ein Angleichen der Sehdistanzen für beide Informationstypen oder eine Anpassung der Brechkraft des Auges an die jeweilige Sehdistanz. Beide Forderungen schränken die Anwendung von augmented reality empfindlich ein. Alterssichtige Personen werden von diesem Handicap besonders betroffen sein.

In der vorliegenden Dissertationsarbeit untersucht Herr von Waldkirch Möglichkeiten, oben beschriebene Einschränkungen bei der Kombination realer und artifizieller Information zu umgehen. Zur Verbesserung der Schärfentiefe des Auges schlägt Herr von Waldkirch mehrere Prinzipien vor. Eine Methode beruht auf der Verwendung von kohärentem Licht. Ähnliche Resultate betreffend der Schärfentiefe erzielt Herr von Waldkirch durch die Einführung einer oszillierenden Linse in die Apparatur. Eine weitere Methode lehnt an die Erzeugung von Mehrfachabbildungen an verschiedenen Orten längs der optischen Augenachse an, ähnlich wie Vorstellungen über den Sehvorgang, die Hermann von Helmholtz in seinem Handbuch der physiologischen Optik Ende 1800 zitiert hat.

Dank rasant forschreitender Miniaturisierung wird es demnächst möglich sein, die in dieser Arbeit entwickelten Ideen im Dienste eines erweiterten Sehens im Alltag umzusetzen. Vielleicht steht die Brillenentwicklung vor einem Quantensprung.

Zürich, im Januar 2005
PD Dr. Marino Menozzi

To Franziska

Contents

ii

iv

Abstract

In this thesis we propose and evaluate several optical approaches which aim to substantially extend the depth of focus (DOF) of see-through head-mounted displays (HMD). The main objective is to find a display setup, for which the perceived quality of the virtual images is almost insensitive to changes of the ocular accommodation.

Head-mounted displays promise to become a valuable and unobtrusive alternative to conventional displays, as widely used for laptop computers, PDAs and mobile phones. Additionally, the see-through mode allows the user to overlay the computer data over the real view and thus to combine real and virtual scene. This is of crucial importance for many applications in the area of wearable computing and augmented reality. However, a really unobtrusive and comfortable use of the see-through feature is guaranteed only, if the image quality of the virtual scene is almost insensitive to natural accommodation changes of the user's human eye. Otherwise, the overlaid virtual image gets often out of focus and thus becomes blurred. To avoid this blurring, the DOF of future see-through HMDs should be significantly extended compared to today's systems.

In our work, we propose a novel LCD-based partially coherent retinal projection display which is characterised by a direct projection of a miniature LCD onto the user's retina. The LCD is illuminated by partially coherent light (instead of common incoherent light) leading to a first improvement of the defocusing properties and thus, to an extended depth of focus. Additionally, we evaluate the use of a variable-focus fluid lens in the retinal projection display. This oscillating lens projects the LCD screen sequentially at different foci onto the retina, so that the user perceives a fusion of in-focus and defocused images. We have determined optimum oscillating parameters and discuss the advantages and drawbacks of this method. The most promising approach proposed in this thesis is mainly based on a similar idea: In the so-called multiple image technique an incorporated phase-only mask projects the LCD screen simultaneously at different foci onto the retina so that the multiple images superpose partially coherently to each other. We have designed an appropriate phase-only mask and show the promising potential of this technique to extend the DOF.

In order to characterise the defocusing properties we have carried out contrast experiments as well as measurements of the image quality. In the latter case, the retinal images of short text targets have been assessed by a combination of qualitative image quality criteria and psychometric experiments with subjects. Finally, in a study we propose a miniaturised and compact design of this novel retinal projection display in order to integrate it into common eyeglasses.

An alternative to the LCD-based retinal projection is the retinal scanning technology where the image is scanned pixel-by-pixel onto the retina. For this technology we have carried out defocusing simulations on the basis of an accommodation-dependent eye model and discuss the trade-off between image resolution, DOF and maximum field of view.

In conclusion, we have been able to lay the basis for solutions of the problem of accommodation-insensitive HMDs within the range of accommodation of adult human beings. This should allow further applications of the see-through display mode.

Zusammenfassung

Die vorliegende Arbeit diskutiert und evaluiert verschiedene optische Methoden, um die Schärfentiefe von sogenannten See-through Head-mounted Displays (HMDs) wesentlich zu vergrössern. Das Hauptziel dabei ist es, einen optischen Aufbau zu finden, bei welchem die wahrgenommene Bildqualität möglichst unabhängig ist von der Akkommodation des Auges.

HMDs stellen eine vielversprechende und bequeme Alternative dar zu heute üblichen Bildschirmen, wie sie oft in Laptops, Taschencomputern (PDAs) und Mobiltelephonen verwendet werden. Die zusätzliche Überblend-Möglichkeit ('see-through mode') erlaubt es dem Benutzer, die dargestellten Computer-Informationen über die jeweilige reale Sicht zu legen und damit virtuelle und reale Welt zu verknüpfen. Dies ist entscheidend für zahlreiche Anwendungen im Wearable Computing oder auch in Augmented Reality. Diese Überblendung ist jedoch für den Benutzer nur dann von wirklichem Nutzen, wenn die Bildqualität des virtuellen Bildes möglichst unabhängig ist von der Augen-Akkommodation. Andernfalls wird die überblendete Computer-Information oft verschwommen und unscharf wahrgenommen. Um dies zu verhindern, sollte die Schärfentiefe zukünftiger See-through-Displays entscheidend vergrössert werden gegenüber heutigen Systemen.

In unserer Arbeit diskutieren wir einen neuartigen LCD-basierten, partiell kohärenten Projektions-Display, der sich dadurch auszeichnet, dass das LCD-Bild direkt auf die Netzhaut des Betrachters projiziert wird. Dabei wird für die Beleuchtung des LCDs statt inkohärentem Licht partiell kohärentes Licht verwendet, was zu einem verbesserten Defokussier-Verhalten führt. Zusätzlich analysieren wir die Verwendung einer neuartigen oszillierenden Flüssig-Linse. Diese oszillierende Linse projiziert das LCD-Bild rasch aufeinanderfolgend auf verschiedene Fokalebenen, so dass der Benutzer nur die inkohärente Überlagerung von scharfen und unscharfen Bildern wahrnimmt, nicht aber die Einzelbilder. Für diesen Aufbau analysieren wir die optimalen Oszillations-Parameter und diskutieren dabei Vorteile und Nachteile dieser Methode. Der vielversprechendste Ansatz - die Mehrfach-Abbildungs-Technik - basiert grösstenteils auf einer ähnlichen Idee: In diesem Fall wird das LCD-Bild mittels einer Phasenmaske gleichzeitig auf verschiedene Fo-

kalebenen projiziert. Dadurch überlagern sich die einzelnen Bilder partiell kohärent miteinander. In dieser Arbeit entwerfen wir eine passende Phasenmaske und zeigen das grosse Potential dieser Technik.

Um das Defokussier-Verhalten des Systems zu charakterisieren, haben wir Kontrast- wie auch Bildqualitätsmessungen durchgeführt. Die Beurteilung der Textproben erfolgte dabei durch eine Kombination von objektiven Qualitätsmassstäben und psychometrischen Experimenten mit Testpersonen. Schliesslich analysieren wir in einer Studie die Machbarkeit einer Miniaturisierung dieses neuartigen Displays mit dem Ziel, dieses in herkömmliche Brillengestelle zu integrieren.

Das pixelweise Scannen der Bilder auf die Netzhaut stellt eine mögliche Alternative zum LCD-basierten Projektions-Display dar. Für diese Technologie analysieren wir auf der Basis eines Augenmodells ebenfalls das Defokussier-Verhalten. Basierend auf Simulationen diskutieren wir den auftretenden Zielkonflikt zwischen maximaler Bildauflösung, Schärfentiefe und maximalem Gesichtsfeld.

Zusammenfassend ist es gelungen, die Basis zu legen für technische Lösungen akkommodations-unabhängiger HMDs, und zwar im Umfang, wie es durch das menschliche Akkommodationsvermögen vorgegeben ist. Dadurch sollte der breiteren Anwendung der Überblend-Möglichkeit in HMDs nichts mehr im Wege stehen.

1

Introduction to head-mounted displays

The first part of this introduction will give a short overview on today's common types of head-mounted displays (HMDs), categorising them according to their technology and to their functionality, respectively. Some design challenges are briefly discussed. The second part of the chapter focuses on the accommodation conflict occurring in see-through 2D-displays. The chapter concludes with the discussion of former work done to solve or at least to reduce this conflict.

1.1. Requirements to HMDs

The vision of wearable computing as well as of augmented reality demands that the visual computer output is provided in a completely modified form as compared to today's common, more or less bulky screens [1, 2]. From the ergonomic point of view, future displays, as appropriate for wearable computing, should not only be wearable and lightweight, but they should also be unobtrusive and guarantee the largest possible freedom in head and eye movement. To meet the specification of a large freedom in head movement, most of the proposed displays can - in any kind - be mounted on the user's head. Thus, they are commonly referred to as *'head-mounted displays' (HMDs)*.

In addition to the ergonomic requirements, comfortable HMDs should further satisfy the following visual features:

- Contrast and resolution

- Retinal illuminance

- Field of view (FOV)

- Colour

- Focus adjustment / Accommodation conflict (see section 1.4)

Depending on the specific functionality and intended purpose of the corresponding HMD, some features are more important than others. Consequently, it is not possible to define general minimum figures for all the features. For applications in the area of augmented reality an additional feature is of great importance: the so-called see-through mode, which enables to overlay the display's output over the user's real view [3, 4, 5].

In recent years, extensive research in display technology has been carried out to satisfy the visual and ergonomic requirements. Several different approaches have been proposed in the literature. In 1968, Sutherland [6] proposed one of the first head-mounted displays for three-dimensional viewing on the basis of two small cathode ray tubes (CRT) mounted on a head band and placed closely in front of the user's eyes. Today, there is a wide variety of display devices to choose from, each having its specific spatial resolution, contrast, colour range, brightness, weight and power consumption. Thus, the following sections

attempt a categorisation of these various technologies and functionalities to give a short overview on the current state of the art in head-mounted display research. A good general reference on a wide range of HMD issues is given by Melzer et al. [7].

1.2. Technological categorisation

From the technological point of view, head-mounted displays can roughly be split into three main categories, according to the way the image is provided to the user (see Fig. 1.1): The first category (referred to as *'screen-based HMDs'*) comprises all HMDs whose picture elements are created in a spatially adjacent way such as LCDs and CRTs. An
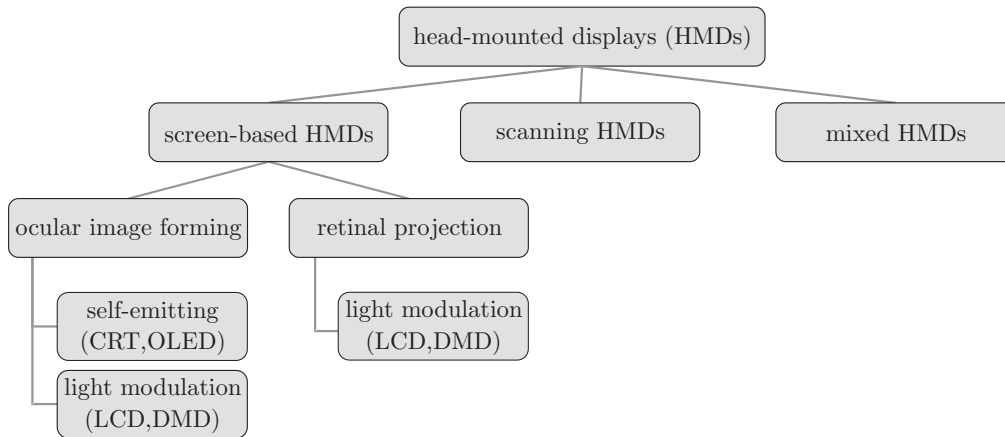
Figure 1.1:   *Head-mounted displays categorised according to the technology used*

alternative to screen-based displays (and thus the second category) are scanning displays, where the image is scanned pixel-by-pixel e.g. directly onto the retina; a method which is widely known from conventional TV systems. Finally, a combination of scanning technology and screen-based systems is possible, defining thus the third category.

1.2.1. Screen-based HMDs

The first category comprises most of today's commercial HMDs. This category can further be subdivided first into those systems where the user is just looking at the display while forming the retinal image by the imaging capabilities of the human eye, and second to those systems where the display image is projected onto the retina. The differences

(a) Direct viewing

(b) Direct viewing with use of eye-piece
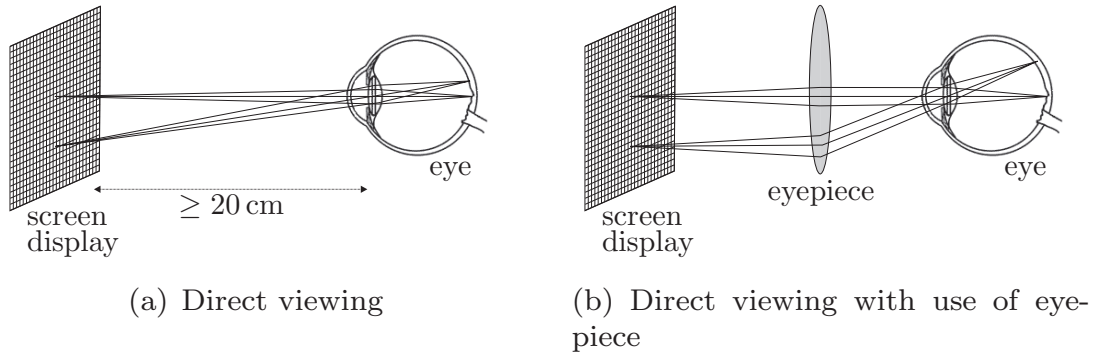
Figure 1.2: *Ocular image forming*

between these two modes should be analysed in more details:

Ocular image forming

In the first case, the eye, in principal, is able to form a retinal image of the display by using its own optics, only. However, an eye can only accommodate correctly if the screen is positioned in at least about 20 cm distance from the eye (see Fig. 1.2(a)). But this large distance leads to three main disadvantages: First, to a relatively small angular image field, second to a continuously accommodated eye (which is straining) and finally to implementation problems if the display should be small and unobtrusive. To overcome these disadvantages, an optical lens system (called 'eyepiece') can be inserted between the screen and the eye as depicted in Fig. 1.2(b). This eyepiece generates a virtual image of the screen at a specific viewing distance which the eye can easily accommodate to. This reduces eye strain and increases the field of view.

Within this sub-category, several screen technologies have been proposed in the literature, starting from the CRT screens [8, 9] up to novel technologies like digital mirror devices (DMD) [10, 11]. Most of today's HMDs on the market, however, are based on transmissive or reflective liquid crystal minidisplays (LCD) [12, 13, 14, 15]. Miniature LCDs are available for relatively low prices and provide an appropriate resolution (of up to SXGA resolution of 1280×1024 pixels) while being lightweight [16]. However, LCD screens suffer from low light output since only a portion of the illumination light is transmitted through the LCD (or reflected by the LCD, respectively) [16]. Thus, LCD-based HMDs suffer usually from luminance problems, particularly in see-through ap-

plications since see-through displays inherently require higher intensity image sources.

A promising new display technology may arise on the basis of organic light emitting diodes (OLEDs) [17, 18]. Such OLED displays [19, 20] are specially characterised by a high luminance and low energy consumption, thus probably having the potential to overcome the problem of limited luminance.

Retinal projection

In the projection method, the second sub-category within the screen-based HMDs, the display image is projected directly onto the retina in the same manner as a slide is projected onto the screen (see Fig. 1.3).
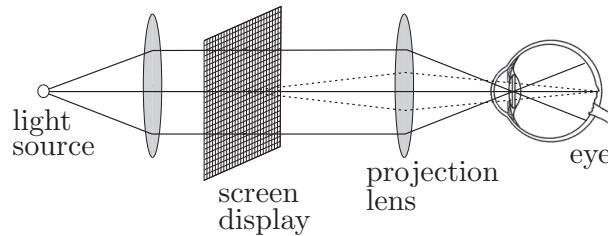
Figure 1.3: *Retinal projection. The straight lines show the axial illumination rays, while the dashed lines indicate the display imaging.*

To this end, the optical imaging capabilities of the human eye are not used and thus, no virtual image must be formed the user is looking at. This results mainly in improved light efficiency. Projection displays are normally designed in the form of a 'Maxwellian-view optical system' [21], where the screen plane is optically conjugated to the retina and the illumination source is conjugated to the eye's pupil plane. Consequently, such displays can only be implemented on the basis of illuminated screens (like LCD and DMD), but not with self-emitting technologies like OLED and CRT. Chapter 2 will discuss this kind of displays in detail.

1.2.2. Retinal scanning display

In the category of scanning displays, the retinal scanning display (RSD) is the most important one. The RSD technology (sometimes also referred to as 'Virtual Retinal Display (VRD)') bases on the same idea as the scanning laser ophthalmoscope (SLO) [22, 23] and was first proposed in 1992 by Sony in a patent application entitled 'Direct viewing

picture image display apparatus' [24]. Since 1992, researchers of the Human Interface Technology Lab, Washington DC, have been developing this technology towards a commercial product. In 2003, they presented together with the US-based company Microvision the first commercial RSD-display called 'Nomad'. [25, 26].

The RSD technology is based on a light beam which is scanned in x- and y-direction directly onto the retina in a raster pattern while being intensity modulated to form an image. The image refresh rate must lie above the critical flicker fusion frequency of the human vision (of about 60 Hz). The principle of the RSD technology is depicted schematically in Fig. 1.4. Defocusing properties of this display type are thoroughly discussed in chapter 7.
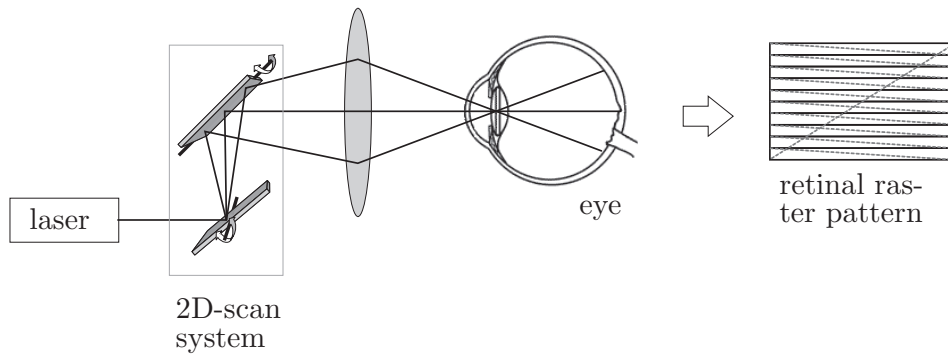
Figure 1.4: *Retinal scanning technology. The figure on the right hand side shows schematically the retinal raster pattern.*

1.2.3. Mixed displays

A display combining scanning and screen-based technologies was proposed in 1991 by Pausch et al. [27]. Their 'private eye system' contains a one-dimensional array of light-emitting diodes (LEDs) which is scanned in one direction to get a raster pattern. Similar systems based on a LCD, which is scanned in one direction, have been proposed in [28, 29].

1.3. Functional categorisation

Beside the technological categorisation, the various HMD solutions can also be classified according to their functionality or intended purpose. Figure 1.5 shows a rough functional categorisation. Here, HMDs can be

```
                    ┌─────────────────────────────┐
                    │ head-mounted displays (HMDs) │
                    └─────────────────────────────┘
         ┌───────────────────┬───────────────────┐
   ┌───────────┐      ┌──────────────┐     ┌──────────────┐
   │ monocular │      │ biocular (2D)│     │ binocular (3D)│
   └───────────┘      └──────────────┘     └──────────────┘
   ┌────────────┐     ┌────────────┐       ┌────────────┐
   │ closed-view│     │ closed-view│       │ closed-view│
   └────────────┘     └────────────┘       └────────────┘
   ┌────────────┐     ┌────────────┐       ┌────────────┐
   │ see-through│     │ see-through│       │ see-through│
   └────────────┘     └────────────┘       └────────────┘
```
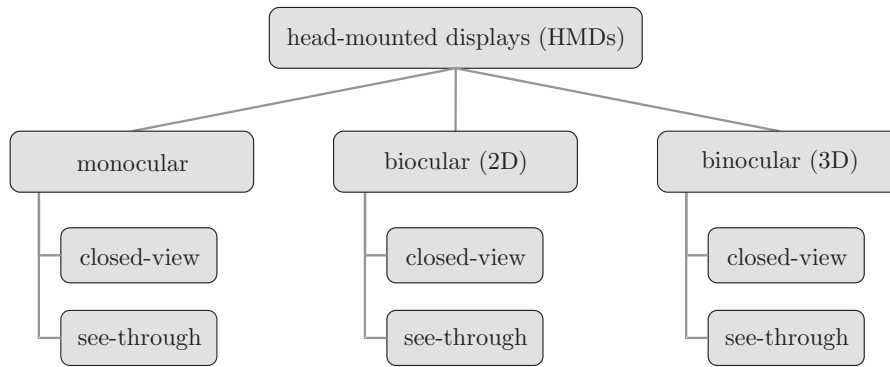
Figure 1.5: *Head-mounted displays categorised according to the functionality*

divided into monocular, biocular and binocular displays. Additionally, all these types can provide the image in a closed-view and a see-through mode (see section 1.3.2).

Monocular displays have only one display source and thus, provide the image to one eye, only [30]. This makes the display lighter weight and less costly than either the biocular or the binocular HMDs. But in a monocular HMD, the view seen by the other eye may produce binocular rivalry with the image seen through the HMD [31]. Biocular HMD have two displays with separate displays and optics paths, but both eyes see exactly the same image. So, the biocular HMD is heavier than the monocular HMD and both eyes must be properly aligned to the system. Since both eyes see the same image, the user perceives a 2D-image, only, similar to a computer screen.

1.3.1. Stereoscopic viewing

Only the binocular display allows stereoscopic viewing that provides 3D-depth perception. To this end, two displays and two sets of electronics will be required presenting slightly different images to the right and the left eye. To produce the stereoscopic view, the two images must mutually overlap. Binocular HMDs suffer normally from the so-called convergence/accommodation conflict [32]: The eyes naturally turn inward (converge) to view close objects and outward (diverge) to view far objects. This natural vergence eye movement is required to gather stereoscopic depth information of objects at varying viewing distances. Under normal viewing conditions, accommodation and vergence eye movements vary synchronously and are dependent on object distance.

In contrast, when a stereoscopic display is used, the eyes must maintain accommodation on the fixed LCD screens, despite the presence of depth cues that necessitate vergence eye movements in the virtual scene. That means, that in stereoscopic displays the close link between accommodation and vergence is partially broken up causing discomfort and unrealistic perception of the 3D-image.

1.3.2. See-through mode

In addition, HMDs can also be subdivided according to whether or not they have 'see-through' capability. See-through HMDs let the user see the real world with virtual objects superimposed by optical or video technologies. In contrast, closed-view HMDs do not allow a direct view of the real world. The see-through mode has two significant advantages: On the one hand, it allows the user to access computer information while (at least) partially paying attention to his environment and performing other tasks. On the other hand, this mode enables to create a mixture between the real and the virtual world by attaching computer-generated data to real life objects (see Fig. 1.6).

Optical see-through HMDs are based on optical combiners placed in front of the user's eyes. Because these combiners are partially transmissive and partially reflective, the user can look at both the real world and the computed data. In optical terms, the combiners can either be achieved by the use of half mirrors or by holographic optical elements (HOE). The latter can accomplish to reflect only the specific range of the display's wavelengths, while being transparent for all other visible wavelengths coming from the real scene [33]. Optical see-through HMDs must have sufficient luminance to be compatible with the real scene luminance. In particular in case of outdoor applications, the see-through HMD should be visible even for bright ambient lighting conditions. If the see-through HMD is binocular, then vergence and accommodation for seeing the virtual image must be compatible with the real world objects that will be combined with the HMD image. For example, if the HMD image is to be combined with features of a relatively close real-world object, then the HMD images must cause the eyes to converge to the same distance as the real object; otherwise double vision will occur - of either the real-world object or the overlaid virtual scene.

As alternative to the optical see-through HMDs, video see-through HMDs combine a closed-view HMD with one or two head-mounted video cameras that capture the user's view of the real world [34]. The

output from these cameras is combined with the virtual images from the computer and finally displayed together on the HMD screen.

1.3.3. Low vision aid

People with low vision have reduced visual acuity and a significant loss of contrast sensitivity, often in combination with visual field loss. These impairments cause a number of disabilities including difficulty with reading, writing, recognising faces, watching television, and completing activities of daily living. However, many visually impaired patients can interpret much detail if it is sufficiently magnified since then the details can be read by undamaged photo-receptors of their retina. The HMD technology can provide assistance to those people by capturing an image of the real scene (via a camera) and relaying it to an HMD. The HMD provides then a magnified and higher-contrast image of the captured real scene. Several low vision HMD systems have been proposed, either on the basis of screen-based displays [35, 36, 37] or on the retinal scanning technology [38].

1.4. Accommodation conflict

This work focuses mainly on monocular and biocular see-through HMDs which provide a 2D-virtual image overlaid over the real scene, only. Thus, vergence conflicts as discussed above are not a main issue. Nevertheless, for such 2D-displays severe accommodation problems may occur in case of see-through applications. These frictions with the optical properties of the human eye shall be discussed in this section in more detail:

Ocular accommodation means the variation of the total refracting power of the human eye allowing an observer to clearly see objects at different distances. The limits of accommodation are the distances of the nearest and farthest points that can be focused clearly by the eyes of an observer. These limits change with age. A 40-year old person, for example, can usually accommodate to all viewing distances between infinity and about 20 cm - corresponding to a range of accommodation of about 5 diopters (D) [1].

This ocular property has direct implications on the see-through feature of the display: The overlaid information is commonly provided by

---

[1]Diopter (D) is, in our context, the reciprocal of the viewing distance (in meter) (e.g. accommodation of 2 D ≅ 50 cm viewing distance) (see Appendix C).

the display at a fixed viewing distance. Thus, as soon as the eye adapts to a real object at an other viewing distance, the overlaid features get out of focus and thus blurred. The accommodation range, within which the overlaid features are perceived as sharply imaged and thus be comfortably viewed, is referred to as 'depth of focus (DOF)' of the display.

A comfortable use of the see-through feature requires, that the overlaid information is perceived sharp and in focus, independently of the object in the real world, to which the user is accommodating. Otherwise, severe disturbing effects are caused as demonstrated in experiments by Elgar et al. [39]. This crucial specification, however, is mostly not met by today's see-through HMDs since these provide only a quite restricted depth of focus of less than $0.3\,\mathrm{D}$, normally[2].

Figure 1.6 illustrates the problem considering a simple scenario: It shows a hiker in the mountain holding a map in his hands and looking at the landscape. His see-through HMD provides additional geographical and general information (such as names, current time and position). As long as the hiker is looking at the mountains far-away (corresponding to eye accommodation to infinity) the display information are in-focus and well readable. However, when the hiker is gazing at the map in his hands and thus accommodates to this closer viewing distance, the overlaid information gets out of focus causing inconvenience and discomfort.

1.5. Related work

The example above illustrates that either the focus of the see-through display should be adapted in real-time to the eye's current accommodation or the display should provide a depth of focus that equals or exceeds the eye's range of accommodation. For the latter, a good target value for the DOF would be about 5 diopters (D). That value would imply, that the image quality stays unaffected by variations of the eye accommodation for any viewing distances between infinity and at least $20\,\mathrm{cm}$.

As discussed in section 1.2.1, today's most common HMDs consist of a microdisplay unit placed close to the eye and illuminated by spatially incoherent backlight. The miniature display is then imaged through magnifying optics to create a virtual image at some fixed distance in space [40, 41]. To extend the DOF in such systems, normally apodisation methods are used by e.g. introducing an additional aperture stop

---

[2]according to Rayleigh's quarter-wavelength criterion [21] and assuming a light-adapted eye pupil diameter of $3\,\mathrm{mm}$.
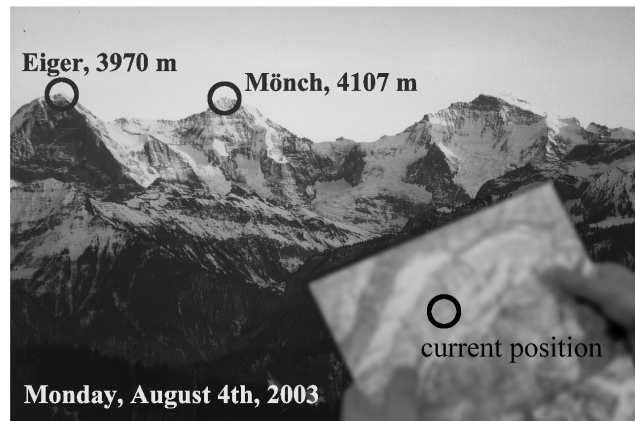
(a) Focusing on the mountains

(b) Focusing on the map

Figure 1.6: *Simple scenario illustrating the accommodation conflict in a see-through display with restricted depth of focus. In this example, the display's focus is assumed to be adjusted to infinity.*

into the optical system [42]. This aperture has the property of blocking light rays that cause blurring. This is widely used in photo cameras. However, these apodisation-based techniques result in a substantial decrease of retinal illuminance making the readability of the display in bright ambient lighting conditions difficult.

Another approach for extending the DOF of head-mounted displays has been discussed in Ref. [43] in the context of stereoscopic HMDs: This system is based on a multiplane display consisting of a stack of planar arrays so that all virtual objects would be displayed ideally at the appropriate distances. Finally, Marran et al. [44] have outlined several focus solutions, as adding a different monocular lens in front of each eye of a binocular HMD to focus each eye at a different depth, or using bifocal lenses to provide the virtual image at two different focal planes in the lower and upper part of the visual field.

A completely different approach was discussed by Sugihara and Miyasato [45] by proposing a real-time adaption of the display's focus to the current eye accommodation. Their display provides a movable LCD-screen so that the LCD-position can be varied to match the current eye's accommodation within $0.3\,$s. This, however, requires a fast and accurate tracking of the eye accommodation leading to an enhanced technical complexity.

2

LCD-based coherent retinal projection display[*]

In this chapter, the concept of a novel LCD-based wearable display is discussed. It is characterised by the direct projection of a LCD image onto the eye's retina and by the use of partially or even fully coherent light. The first part of the chapter gives a system overview, the second part discusses theoretical aspects on coherent imaging in the context of defocusing.

[*]This chapter is mainly based on Ref. [46, 47, 48]

2.1. System overview

Figure 2.1 shows the principal setup of the retinal projection display representing a 'Maxwellian-view optical system' [21, 49]. Partially coherent light from a light-emitting diode (LED) illuminates the LCD through a condenser which provides an homogeneous illumination. This mode of illumination is widely used in microscopy and lithography and normally referred to as Köhler illumination [50]. Through the lens L1 the illumination light is focused onto the focus plane P1 where a diaphragm is located. This diaphragm acts as the system's aperture stop and as a spatial filter, filtering out any undesired higher diffraction patterns of the LCD-image. The subsequent lenses (L2, L3) form a relay lens system with magnification M, imaging the intensity pattern of plane P1 to the eye's pupil plane P2. A geometric image is finally formed at the retina. Consider that the LCD-plane is optically conjugate to the retina as long as the eye is well accommodated to the display (see the dotted line in Fig. 2.1).
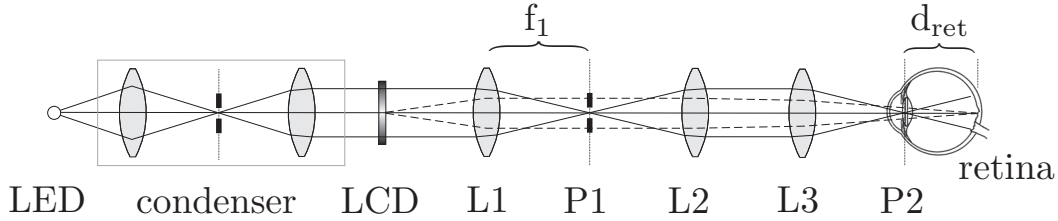
Figure 2.1: *Principal setup of the retinal projection display. L1-L3 represent the lenses, P1 the L1-focal plane with the diaphragm and P2 the eye's pupil plane. f_1 denotes the focal length of lens L1 and d_{ret} the distance from the plane P2 to the retina. The straight lines show the axial illumination rays. The dashed lines indicate the LCD-imaging.*

Note for the subsequent considerations that the system's exit pupil is located at the pupil plane P2 and corresponds to the aperture stop at P1 scaled by the magnification M. Apodisation techniques (i.e. reducing the diameter of the exit pupil) would be a common and well-known method to increase the DOF of this system. However, a smaller exit pupil reduces the retinal image resolution so that this is not a promising approach. In this work, consequently, the DOF should be extended by adjusting the degree of spatial coherence of the illumination light in place of any kinds of apodising.

The degree of spatial coherence is described by the parameter σ which has an illustrative link to the experimental setup: σ is defined

as the numerical aperture of the condenser system divided by the numerical aperture of the objective lens of the system [51, 52, 53]. Consequently, $\sigma \to 0$ indicates the limit of fully spatial coherence while $\sigma \to \infty$ signifies the fully incoherent limit. This σ-parameter is often used in the areas of lithography, microscopy and coherent imagery. Details as well as a motivation for this definition of σ are given in Appendix B.

2.2. Defocusing analysis

The theoretical analysis of the defocusing properties of the retinal projection display is limited to the discussion of the two limits of fully coherent light ($\sigma \to 0$) and fully incoherent light ($\sigma \to \infty$). These two limits give a good understanding of the influence of the spatial coherence on the display's defocusing properties.

2.2.1. Amplitude spread function

Referring to the optical setup in Fig. 2.1 we are interested in formulating a mathematical relationship between the LCD transmission function (given as $\tau_{LCD}(u,v)$) and the retinal field amplitude distribution $\iota_{ret}(u',v')$. Here, (u,v) are the geometrical coordinates in the LCD-plane, while (u',v') denote the corresponding coordinates on the retina.

To find this general mathematical formula, we first restrict $\tau_{LCD}(u,v)$ to a δ-function (point source) at the LCD-coordinates (u_0,v_0) and search the retinal field amplitude at (u',v') produced by this point source. This field amplitude, denoted by $h(u',v';u_0,v_0)$, is commonly referred to as the system's amplitude spread function (ASF). It turns out - as derived in [54] in details - that $h(u',v';u_0,v_0)$ corresponds to the scaled Fraunhofer diffraction pattern of the system's exit pupil function $P(x,y)$[1]:

$$h(u',v';u_0,v_0) \propto \iint\limits_{-\infty}^{\infty} P(x,y)e^{-i\frac{2\pi}{\lambda}\left[(\frac{u'}{d_{ret}}+\frac{u_0}{Mf_1})x+(\frac{v'}{d_{ret}}+\frac{v_0}{Mf_1})y\right]}dxdy$$

(2.1)

where f_1 is the focal length of lens L1 and (x,y) are coordinates in the exit pupil plane P2. λ represents the wavelength of the used illu-

---

[1]Note that $P(x,y)$ describes the exit pupil at the plane P2 being unity inside the pupil aperture and zero otherwise.

mination light while d_{ret} denotes the distance from the exit plane P2 to the retina. Consider that due to the relay optics (L2, L3) the LCD coordinates (u_0, v_0) in eq. 2.1 are scaled by the relay magnification M; the corresponding retinal coordinates (u', v') are scaled by d_{ret}.

However, the eq. 2.1 is only valid as long as the eye is well accommodated to the LCD-screen. When the eye accommodates[2] at a viewing distance $\zeta < \infty$, the wavefront that leaves the eye lens is deformed. Thus, it is common to replace the aperture function $P(x, y)$ in eq. 2.1 by a complex generalised pupil function $\mathcal{P}(x, y)$ which considers the additional phase factor $\psi(x, y)$ being picked up at the pupil plane P2 due to the dioptric eye accommodation $\Delta D = 1/\zeta$ of the eye lens [21]:

$$\mathcal{P}(x, y) = P(x, y)e^{i\psi(x,y)} = P(x, y)e^{-i\frac{\pi}{\lambda}\Delta D(x^2+y^2)} \qquad (2.2)$$

Here, $\psi(x, y)$ is the wavefront aberration error at the exit pupil plane point (x, y) due to the focus error ΔD.

The Fourier spectrum of the ASF $h(u', v'; u, v)$ in terms of the object coordinates (u, v) is commonly referred to as the coherent transfer function (CTF) $H(\nu_u, \nu_v)$ [55]. According to eqs. 2.1 and 2.2, $H(\nu_u, \nu_v)$ is found to be proportional to the scaled generalised pupil function $\mathcal{P}(x, y)$:

$$H(\nu_u, \nu_v) = \mathcal{F}\{h(u', v'; u, v)\} \propto \mathcal{P}(\lambda M f_1 \nu_u, \lambda M f_1 \nu_v) \qquad (2.3)$$

Note that (ν_u, ν_v) signify the spatial frequencies of the object LCD-pattern in both directions (u, v).

So far, the two cases - the incoherent and the coherent limits - can be treated together. To get now the retinal field amplitude distribution $\imath_{ret}(u', v')$ (or the retinal irradiance distribution $I_{ret}(u', v') \equiv |\imath_{ret}(u', v')|^2$, respectively) produced by a general LCD transmission function $\tau_{LCD}(u, v)$, the two limits must be discussed separately.

2.2.2. Incoherent limit ($\sigma \to \infty$)

An incoherent imaging system is linear in intensity. Consequently, the general retinal irradiance $I_{ret}(u', v')$ produced by $\tau_{LCD}(u, v)$ can be expressed as the incoherent superposition of many point source irradiance distributions $|h|^2$. Hence, $I_{ret}(u', v')$ is given by the following

---

[2]The system is assumed to be adjusted for an emmetropic, relaxed eye (i.e., for an eye focusing at the viewing distance $\zeta = \infty$).

convolution:

$$
\begin{aligned}
I_{ret}(u',v') &= \iint\limits_{-\infty}^{\infty} |h(u',v';u,v)|^2 |\tau_{LCD}(u,v)|^2 \, du \, dv \\
&= \underbrace{|h|^2}_{PSF} \otimes |\tau_{LCD}|^2
\end{aligned}
\qquad (2.4)
$$

where $\otimes$ stands for the convolution. The squared magnitude of the amplitude spread function h is commonly called the point spread function (PSF). The PSF denotes the irradiance distribution produced by a point source [55].

By means of the convolution theorem [56], the frequency spectrum of the retinal irradiance $I_{ret}(u',v')$ is now:

$$
\mathcal{F}\{I_{ret}\} = [H \star H]\,\mathcal{F}\{|\tau_{LCD}|^2\} \propto \underbrace{[\mathcal{P} \star \mathcal{P}]}_{OTF}\mathcal{F}\{|\tau_{LCD}|^2\}
\qquad (2.5)
$$

where the symbol $\star$ represents the autocorrelation integral and H is the coherent transfer function (CTF) as defined in eq. 2.3. Thus, in the incoherent case the spectrum of $I_{ret}(u',v')$ is connected to the spectrum of $\tau_{LCD}(u,v)$ by the autocorrelation integral of the generalised pupil function $\mathcal{P}(x,y)$. This autocorrelation function is called the optical transfer function OTF.

2.2.3. Coherent limit ($\sigma \to 0$)

A coherent optical system is linear in complex amplitude rather than in intensity. This implies, that the retinal irradiance distribution $I_{ret}(u',v')$ is the squared magnitude of the convolution of the amplitude spread function $h(u',v';u,v)$ with the LCD-transmission function $\tau_{LCD}(u,v)$:

$$
I_{ret}(u',v') = |\imath_{ret}(u',v')|^2 = |h \otimes \tau_{LCD}|^2
\qquad (2.6)
$$

By means of the convolution theorem and eq. 2.3 the following result for the spectrum of the retinal amplitude distribution $\imath_{ret}(u',v')$ is achieved:

$$
\mathcal{F}\{\imath_{ret}\} = H\mathcal{F}\{\tau_{LCD}\} \propto \mathcal{P}(\lambda M f_1 \nu_u, \lambda M f_1 \nu_v)\mathcal{F}\{\tau_{LCD}\}(\nu_u, \nu_v)
\qquad (2.7)
$$

where $\mathcal{F}\{\tau_{LCD}\}$ is the spectrum of τ_{LCD} and $H(\nu_u, \nu_v)$ the coherent transfer function as defined in eq. 2.3.

Equation 2.7 allows to understand the special defocusing properties of the coherent limit: The frequency spectrum of the retinal amplitude distribution $\imath_{ret}(u', v')$ corresponds to the one of the LCD transmission function $\tau_{LCD}(u, v)$ but it is band-limited due to the finite extent of the exit pupil $\mathcal{P}(x, y)$. In other words, the optical system acts as low-pass filter characterised by a cut-off frequency:

$$(\nu_u^{cut}, \nu_v^{cut}) = \frac{1}{\lambda M f_1}(r_x, r_y) \tag{2.8}$$

where (r_x, r_y) stands for the radii of the exit pupil in x- and y-direction. Defocusing causes additional phase shifts of the individual frequency components, as can be seen with eq. 2.2.

Regarding the current setup this can be interpreted as follows: Lens L1 in Fig. 2.1 performs the Fourier transformation $\mathcal{F}\{\tau_{LCD}\}$ of the LCD-image τ_{LCD} at the focus plane P1. Consequently, the diaphragm in plane P1 acts as a spatial filter filtering out higher image frequency components as well as higher diffraction patterns. Then, the relay optics image the Fourier transformation pattern to the pupil plane P2 while scaling it with magnification M. Now when the eye accommodates, the frequency components pick up the additional phase factor $\psi(x, y)$ (cf. the phase term of the generalised pupil function $\mathcal{P}(x, y)$ in eq. 2.2). Finally, a second Fourier transformation within the eye forms the retinal image $I_{ret}(u', v')$.

2.2.4. Difference between $\sigma \to 0$ and $\sigma \to \infty$

The only effect of defocusing in the case of coherent limit is the introduction of phase distortions in the frequency spectrum. Thus, the defocusing properties of a coherent imaging system differ much from the incoherent case where the defocusing phase shifts affect the retinal irradiance distribution indirectly via the autocorrelation function. To make a rough estimate of the different impact on image quality, consider a classical sine pattern transmission mask where the transmission is modulated sinusoidally with a specific spatial frequency ν_0. The phase shift of this frequency ν_0 due to defocus is signified by ψ_ν.

When defocusing in the incoherent case, the phase shift ψ_ν causes a decline in the sine pattern maxima while the minima rise, as illustrated in Fig. 2.2(a). At a specific defocus value the contrast is finally completely lost ('blurring effect'). In the coherent case, however, increasing defocus ΔD leads to an increasing phase shift ψ_ν added to the frequency components $\pm\nu_0$. This results in increasing intermediate peaks
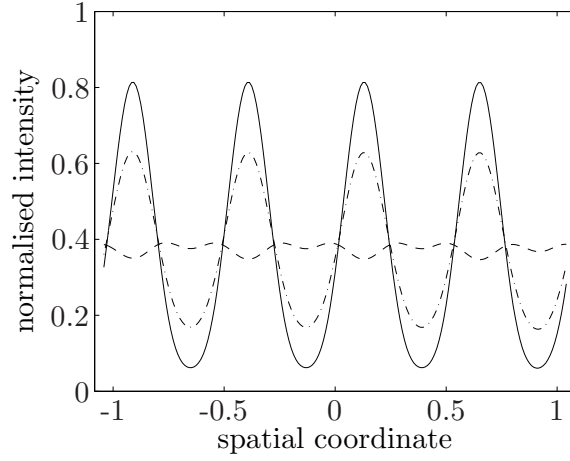
(a) Incoherent limit
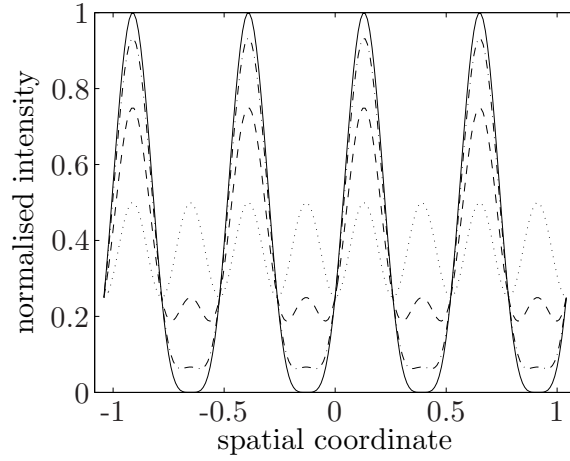
(b) Coherent limit

Figure 2.2: *Sinusoidal example pattern while defocusing for a) the incoherent and b) the coherent case. The solid line (—) represents the image pattern at best focus ($\psi_\nu = 0$), while the dash-dotted line ($-\cdot-$) and the dashed line ($--$) indicate the image pattern at $\psi_\nu = 0.33 * \pi/2$ and $\psi_\nu = 0.66 * \pi/2$. The dotted line ($\cdots$) at $\psi_\nu = \pi/2$ shows the frequency doubling effect for the coherent case, only.*

at the minima of the original sine pattern combined with a decrease of the maxima (see Fig. 2.2(b)). At $\psi_\nu = \pi/2$, the image sine pattern even suffers a complete frequency doubling.
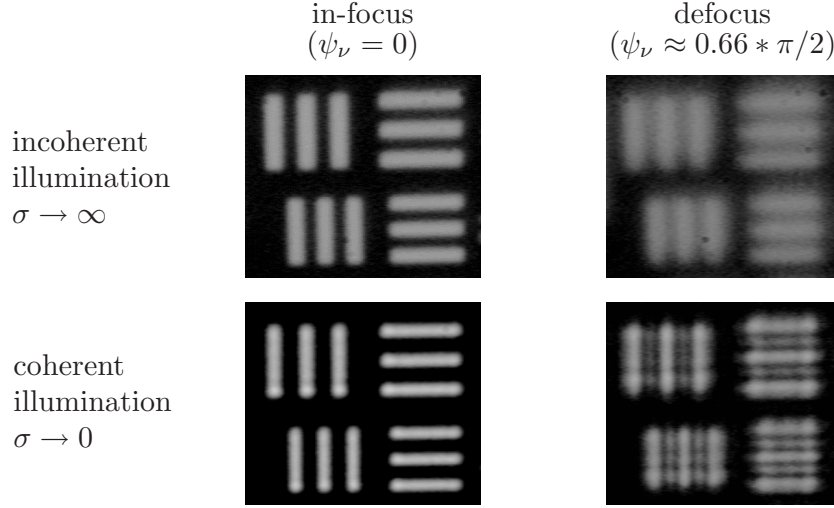
Figure 2.3:   *Incoherent and coherent imaging of horizontal and vertical bars*

For illustration, Fig. 2.3 shows some images at best focus as well as at a specific defocus (with $\psi_\nu \approx 0.66 * \pi/2$) for both limits of coherence level. The four images illustrate evidently the gain in contrast when increasing the coherence level from incoherent to coherent illumination. In addition, the occurring intermediate defocusing peaks for coherent illumination are clearly observable. Note further that for in-focus imaging ($\psi_\nu = 0$) the contrast is equal to 1 in the coherent case while in the incoherent case the contrast is already lowered.

To conclude, it can be stated that the contrast in the coherent limit is better than that in the incoherent limit at same defocusing values. This indicates that the defocusing properties of a display could be improved when increasing the spatial coherence of the illumination light. However, this improvement comes along with the occurrence of coherent intermediate peaks at the irradiance minima when defocusing. These intermediate peaks might degrade the perceived quality of the defocused image.

2.3. Coherence effects

Coherent illumination, however, does not only alter the defocusing properties, but also causes various noise effects which might affect the

image quality. Three of them are of special importance - the speckle noise, the coherent noise, and the ringing effect - and thus, are discussed briefly in the following.

2.3.1. Speckle noise

Speckle noise on the image appears whenever fairly coherent light illuminates a diffuse object whose surface is rough on an optical wavelength scale [57, 58]. In this case the image is found to have a granular appearance with a multitude of bright and dark spots (see Fig. 2.4(a)). Speckle noise is caused by the fact that the various amplitude spread functions, whose superposition forms the image at a given point, are randomly dephased to each other due to the surface roughness of the object. This results in a random interference pattern. The characteristic granular size of the speckle has the same order of magnitude as the imaging resolution limit so that the image quality is mainly affected at high spatial frequencies [59, 60]. One common method to reduce disturbing speckle patterns is simply to reduce the spatial coherence of the light [61]. In our setup, however, the speckle noise is negligible - even with highly coherent illumination - as the LCD can be considered as a non-diffusing transmission object and no stationary diffuser is included in the optical pathway. Thus, the coherent light can not be dephased randomly and no speckles occur.

2.3.2. Coherent noise

Coherent noise is essentially different from speckle noise. Although highly coherent light gives rise to both noise types, the speckle patterns are formed by diffused light, while the coherent noise appears with the use of straight light (see Fig. 2.4(b)). Thus, coherent noise may degrade the perceived image quality in our setup.

The origin of coherent noise is mainly the diffraction of the coherent light by inhomogeneities and imperfections in the optical setup (such as dust particles and scratches on the lenses) and of the ocular media. Contrast sensitivity measurements by Felipe et al. [59] revealed that in contrast to speckle noise, the coherent noise affects the image quality in the low-frequency range. On the other hand, the degradation effect turned out to be quantitatively much less important than the one introduced by speckle noise. The most common method to avoid coherence noise is to incorporate a stationary diffuser before illuminating the LCD. However, to improve the defocusing properties this

(a) Speckle noise   (b) Coherent noise
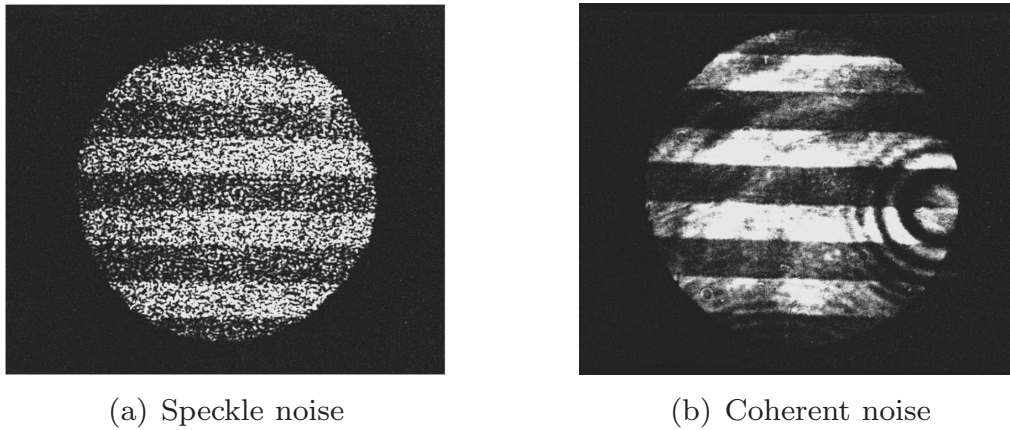
Figure 2.4: *Appearance of a grating test pattern with speckle and coherent noise when illuminated by spatially coherent light (Images from [59]).*

method is not beneficial for two reasons. First, the transfer function of a diffuse coherent illumination is not the coherent, but the incoherent transfer function, as demonstrated by Lowenthal and Arsenault [62]. Thus, when incorporating a diffuser into the optical path, the promising coherent defocusing properties would be lost. Second, the diffuser gives rise to speckle noise, which affects the image quality even more than coherence noise. A different method to reduce coherent noise is the reduction of the spatial coherence.

2.3.3. Edge ringing and shifting effect

With coherent illumination, sharp edges in the object are imaged with pronounced intensity oscillations (referred to as 'edge ringing') as illustrated exemplarily in Figs. 2.5 and 2.6 [63].

This effect is caused by the abrupt fall of the in-focus coherent transfer function at the cut-off frequency. This is absent in incoherent imaging because the incoherent transfer function does not have a sharp cut-off. It degrades the image quality especially for objects with straight lines (such as text targets).

In addition, the coherent image of sharp edges crosses the location of the actual edge with only 1/4 of its asymptotic value in intensity, while the incoherent image crosses with a value of 1/2 of its asymptotic value (see Fig. 2.6). This results in an edge shifting and thus in an undesired broadening of bright bars. The edge shifting phenomenon is a result of the basic nonlinearity in intensity of the coherent imaging
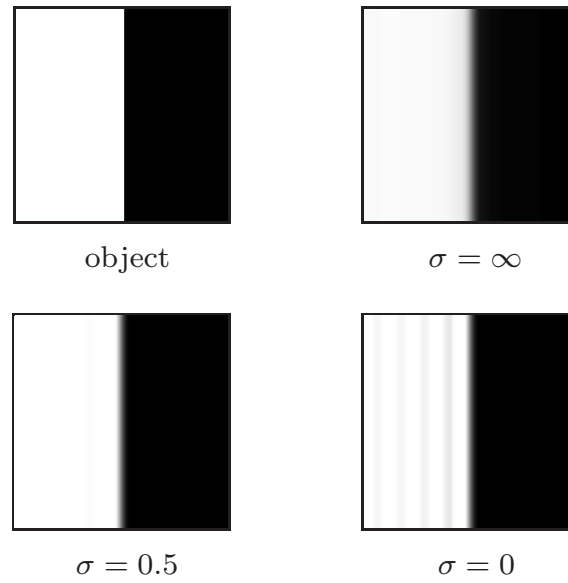
Figure 2.5: *Image of an edge when illuminated by light of different spatial coherence levels*
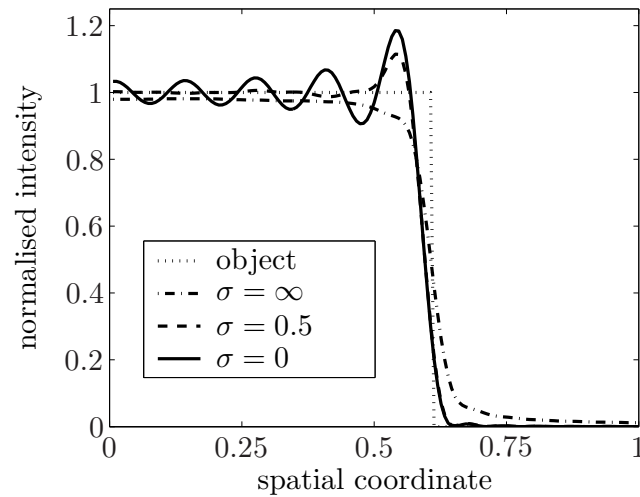
Figure 2.6: *Intensity profile for the image of an edge when illuminated by light of different spatial coherence levels*

process. Thus, both effects - edge ringing and shifting - may affect the image quality in our display. The ringing effect can be reduced by two methods. The first method consists of adjusting the coherent transfer function by apodising with an amplitude mask [64]. This, however, also reduces the in-focus image contrast of higher-frequency patterns. The second method is to reduce the coherence of the illumination light which also enables us to control the edge shifting.

2.4. Conclusion

To conclude, the theoretical considerations indicate that the contrast of a defocused image can be substantially enhanced by increasing the spatial coherence of the illumination light. However, intermediate peaks that occur while defocusing as well as coherence effects (especially coherent noise and edge ringing) may possibly affect the perceived image quality and thus, may partially compensate the gain in contrast.

3.1. Experimental setup

3.1.1. Bench model

For the experimental measurements, a bench model of the projection display - as discussed in the previous chapter - was assembled on an optical breadboard. The detailed optical layout is depicted in Fig. 3.1. For the experiments, the retina was replaced by a CCD camera. A further diaphragm at the pupil plane acts as an artificial eye pupil (diameter $\varnothing_{pup} = 2\,\text{mm}$).
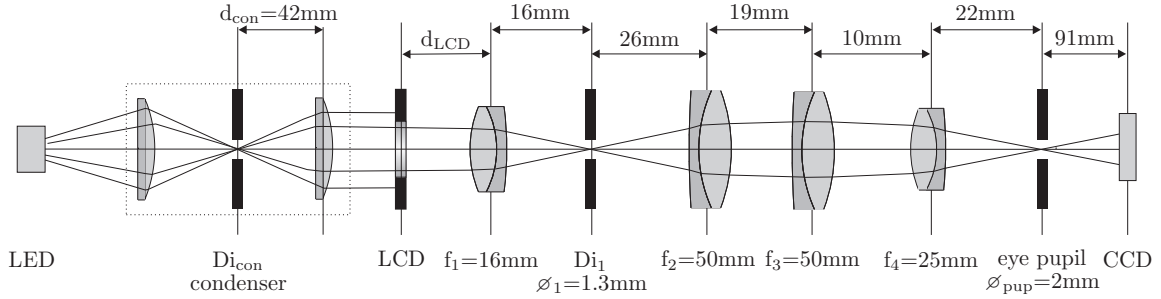
Figure 3.1: *Optical setup of the assembled bench model. Di represents the diaphragms, CCD the CCD camera which replaces the retina. The orientation of the achromat lenses were chosen so that spherical aberration effects are minimised.*

As the light source a commercial superbright red LED by Agilent Technology was chosen (wavelength $\lambda = 626\,\text{nm}$, maximum luminous intensity $I = 18.4\,\text{cd}$). To vary the spatial coherence level of the illumination light, the condenser setup contains a diaphragm Di_{con}, whose diameter $\varnothing_{\text{con}}$ can be varied. This mode of illumination is widely used in microscopy and lithography and normally referred to as 'Köhler illumination' [50]. With this setup, the illumination light can be altered between nearly fully incoherent up to nearly coherent. As the LCD a transmissive active matrix liquid crystal microdisplay (CyberDisplay 320 Mono by Kopin [65]) was used, providing 320 x 240 pixels with a pixel pitch of $15\,\mu\text{m}$. This corresponds to a maximum spatial display frequency ν_{cut} of $33.3\,\text{cyc/mm}$. The three achromats after the diaphragm Di_1 form the relay optics with magnification $M = 0.63$. The diameter of the diaphragm Di_1 in the focal plane of lens L1 is about $1.3\,\text{mm}$, so that Di_1 acts as the aperture stop of the system. Thus, the cut-off frequency of the coherent transfer function (CTF) is about $64\,\text{cyc/mm}$

(corresponding to an angular frequency of 11.3 cyc/deg on the retina) while the cut-off of the incoherent optical transfer function (OTF) is $\approx$ 128 cyc/mm (or 22.5 cyc/deg, respectively). These limits are well beyond the resolution limits of the LCD and thus have no significant impact on the image quality. Note, finally, that the diameter of the eye pupil neither affects the displayed image quality nor the defocusing properties, as the exit pupil of the optical system ($\varnothing$=820 μm) is much smaller than any typical eye pupil diameter.

The distance from the eye pupil plane to the CCD was set to 91 mm so that the CCD pixel size (of about 11 μm) has no limiting effect on any experiments. With the human eye instead of the CCD camera, this distance would be given at about 20.6 mm [21]. Except for the condenser setup, the bench model works with commercial achromat lenses to reduce aberration effects. Nevertheless, all results were evaluated in the centre part of the image. To obtain high resolution in the entire LCD-area, hologram lenses that correct aberration may be used.

Since no apodisation techniques are applied, it is important to note again that the aperture stop Di_1 is kept constant for all experiments (and thus also the system's exit pupil). To control the spatial coherence of the illumination light the aperture Di_{con} of the condenser is varied.

3.1.2. Evaluation method

To better control the defocusing, the LCD-position d_{LCD} was shifted by δd_{LCD} in place of a variable 'artificial eye lens'. If the LCD is displaced by δd_{LCD} then - using the Newtonian form of the lens formula - the corresponding change ΔD in the dioptric eye power turns out to be [21]:

$$\Delta D = -\frac{\delta d_{LCD}}{f_1^2 M^2} \tag{3.1}$$

This is valid for any coherence levels.

To control the degree of spatial coherence σ of the illumination light, the diameter $\varnothing_{con}$ of the diaphragm Di_{con} can be altered. As discussed in appendix B, σ is defined as the ratio of the numerical apertures of the condenser system NA_0 and of the objective lens NA, respectively. Consequently, in our setup, σ is given as follows:

$$\sigma = \frac{NA_0}{NA} = \frac{\varnothing_{con}}{\varnothing_1} \frac{f_1}{d_{con}} \tag{3.2}$$

where $\varnothing_1$ indicates the diameter of the aperture stop Di_1 and $d_{con} = 42$ mm stands for the distance from the diaphragm Di_{con} to the second

condenser lens (see Fig. 3.1). Table 3.1 shows the $\varnothing_{\mathrm{con}}$-values selected for the experiments and their corresponding coherence levels.

$\varnothing_{\mathrm{con}}$ [mm]	0.5	1.2	1.7	3.5	-
$\sigma(\pm 20\%)$	0.15	0.35	0.5	1.0	>3.0

Table 3.1: *The first row shows the diameter values $\varnothing_{\mathrm{con}}$ of the diaphragm Di_{con} as selected for the subsequent measurements. The second row shows the corresponding coherence levels σ. Note that $\sigma > 3.0$ can be considered as fully incoherent.*

To determine the coherence level experimentally, the in-focus contrast function in terms of spatial frequency was measured and fitted to the theoretical predictions. Estimated errors in σ of $\pm 20\%$ are based on this analysis.

3.1.3. Experiments

To evaluate the impact of the spatial coherence on the defocusing properties of the retinal projection display, two different measurements were performed: First, the measurement of the contrast function (CF) in terms of defocusing and coherence level, and second, the assessment of the image quality of short text samples of different sizes when defocusing and altering the coherence. The CF enables an objective, but technical evaluation of the image contrast when defocusing. However, the CF measurement is practically insensitive to any coherence effects that might affect the perceived image quality. The assessment of the image quality of text targets, on the other hand, is closer to the real use of the display and consider any disturbing coherence effects. However, it is subjective. Consequently, the assessment of the image quality is carried out by two independent methods: by an image quality metrics, on the one hand, and by direct visual assessment, on the other.

3.2. Contrast measurements

For the experimental measurement of the contrast, the LCD-display in the bench model of Fig. 3.1 was replaced by an USAF 1951 resolution test target. This transmission mask contains sets of horizontal and vertical bars at different spatial frequencies. It enables an accurate

measurement of the CF-values for specific, well defined spatial frequencies. The pixel pitch of the LCD corresponds to the 32 cyc/mm USAF target (group 5, element 1), so that this target is of special interest.

The MTF-like contrast function (CF) is defined as the ratio of the contrast of the object (c) and the image (c') with respect to spatial frequency ν and defocus ΔD:

$$CF(\nu, \Delta D) = \frac{c'(\nu, \Delta D)}{c} \tag{3.3}$$

where c and c' stand for the respective Michelson contrast:

$$c = (I_{\max} - I_{\min})/(I_{\max} + I_{\min}) \tag{3.4}$$
$$c' = (I'_{\max} - I'_{\min})/(I'_{\max} + I'_{\min})$$

Here, $I_{\max}$ and $I_{\min}$ indicate the maximum and minimum of the irradiance level in the object. $I'_{\max}(\nu, \Delta D)$ and $I'_{\min}(\nu, \Delta D)$ indicate the respective levels in the projected image. The CF-values were measured experimentally by analysing the different tone levels in the CCD images. The CCD camera without any further objective lenses was positioned as indicated in Fig. 3.1. The projection image at the CCD was stored as a 8-bit tiff-image file. To control the defocus, the USAF test target was shifted according to eq. 3.1.

Figure 3.2(a) shows the through-frequency CF at a defocus $\Delta D = 3.2\,\mathrm{D}$ for the different coherence levels σ. In addition, Fig. 3.2(b) displays the through-focus CF for an angular frequency $\nu = 5.6\,\mathrm{cyc/deg}$. The curves representing the fully coherent case ($\sigma \rightarrow 0$) in Fig. 3.2 are shown for comparison. They are based on measurements when illuminating the USAF test target by a laser diode rather than by a LED through a condenser system [46]. Both figures show data from the vertical bars. However, the data of the horizontal direction are quite similar, as the optical system is rotationally symmetric.

Particularly, there are two results to note: First, at in-focus ($\Delta D = 0$) the CF-values increase and tend to 1 when the spatial coherence is increased, as can be seen exemplarily in Fig. 3.2(b) for $\nu = 5.6\,\mathrm{cyc/deg}$. This holds for any spatial frequencies ν below the coherent cut-off frequency which is at $\approx 11.3\,\mathrm{cyc/deg}$. Second, the decrease in the CF-values due to defocusing is significantly reduced when increasing the coherence level. This implies, for instance, that the defocusing value where the CF falls below 0.5, raises from $\approx 2\,\mathrm{D}$ to $\approx 5\,\mathrm{D}$ for $\nu = 5.6\,\mathrm{cyc/deg}$ (see Fig. 3.2(b)).

(a) Through-frequency CF
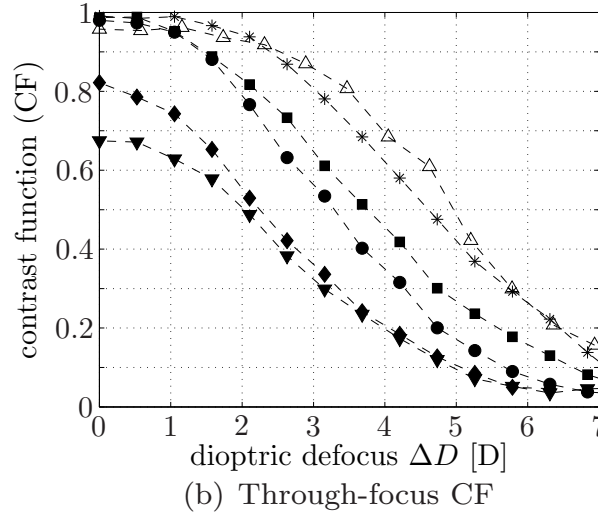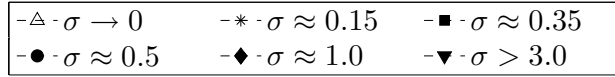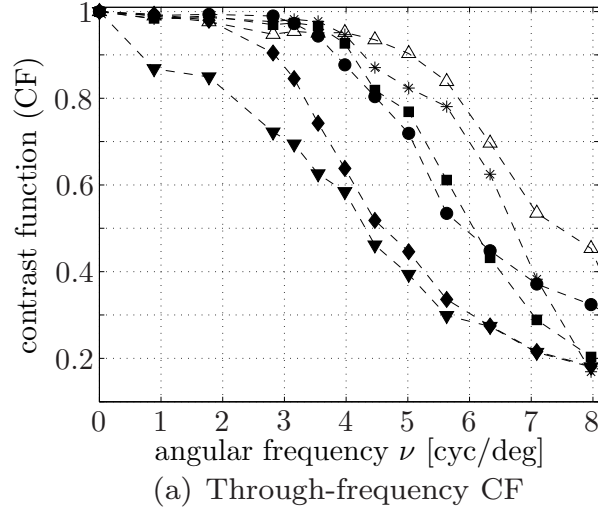
(b) Through-focus CF

Figure 3.2: *a) Through-frequency CF at a defocus $\Delta D = 3.2\,D$ for different coherence levels σ. b) Through-focus CF at an angular frequency $\nu = 5.6\,cyc/deg$ for different σ-values. The coherence level was varied as explained in the text. The open-shape symbols represent the fully coherent case and are based on illumination by a laser diode [46]. All curves represent data from pattern in vertical direction. The dashed lines should act as eye-guide, only.*

3.3. Image quality of text targets

The CF-measurements indicate a significant increase in the depth of focus in terms of image contrast when increasing the coherence of the illumination light. However, they consider only the image contrast when defocusing, but hardly the quality of the defocused image. The perceived quality might be affected by the occurrence of coherent intermediate peaks while defocusing (as seen in Fig. 2.2) and by coherence effects. Thus, as second measurements, the image quality of text targets was evaluated as a function of defocus ΔD and coherence level σ. This second evaluation is well suited as reading text is considered as one of the most stringent visual tasks in virtual reality [66].

3.3.1. Experimental setup

For these measurements, the LCD in Fig. 3.1 was replaced by a binary transmission slide which consists of short text samples written in arial font (normal face, capital bright letters on black background) of different font sizes (see Fig. 3.3). In retinal projection setups, the deci-

WELCOME TO ETH

WELCOME TO ETH

WELCOME TO ETH

WELCOME TO ETH

WELCOME TO ETH

Figure 3.3: *Binary transmission slide consisting of short text samples*

sive measure for the font size is not the original size of the letters on the object slide, but the viewing angle in degree, the letters subtend. Therefore, in this work, we specify the fonts by their respective viewing angles[1] α_{v}.

---

[1]To illustrate this notion: a viewing angle of e.g. $\alpha_{\mathrm{v}} = 0.46 \deg$ corresponds to common 16-pt capital letters on a screen when viewed from a distance of 50 cm. (The height of standard 16-pt capital letters on a screen is 4 mm [67])

In a manner analogous to that already discussed, the text slide was shifted by δd_{LCD} using a micrometer screw for controlling the defocus according to eq. 3.1. For each coherence level σ and font size viewing angle α_{v}, the projected image at the CCD was stored as a tiff-file at different defocusing values. Based on these image series $(\alpha_{\mathrm{v}}, \sigma)$ the image quality was assessed. The judgement of image quality is a subjective issue. Thus, to get reliable results, two different assessment methods were used: First, an objective image quality metrics which assesses the image quality in relation to a reference image, and second, the direct visual evaluation of the image quality by different subjects.

3.3.2. Metrics-based assessment

As image quality metrics the algorithm proposed by Wang and Bovik was used [68, 69]. This algorithm evaluates the quality of the test image Y in comparison to an original reference image X and models any image degradations as a combination of three different factors: loss of correlation, irradiance distortion, and contrast distortion. As extended experiments by Wang and Bovik indicated, this index performs substantially better than other common image quality metrics and exhibits very consistent correlation with subjective measures for various types of image degradations.

This quality index Q is defined as follows:

$$Q = \frac{4 s_{XY} \overline{X}\,\overline{Y}}{(s_X^2 + s_Y^2)(\overline{X}^2 + \overline{Y}^2)} \tag{3.5}$$

where $\overline{X}$ signifies the mean irradiance, s_X the standard deviation of the irradiance[2], and s_{XY} the cross correlation between the images X and Y. To consider local distortions the index Q is first calculated locally in a sliding window of size B x B pixels, which moves pixel by pixel horizontally and vertically through all rows and columns of the image. Finally, the overall image quality index Q is given as the arithmetic mean of all these local indices.

Figure 3.4 shows the image quality index for $\alpha_{\mathrm{v}} = 0.46\,\mathrm{deg}$ as a function of defocus ΔD for various spatial coherence levels σ. The plot indicates that partially coherent illumination (at $\sigma \approx 0.5$) results in a better image quality as almost fully coherent or fully incoherent illumination. This is caused by the discussed trade-off between higher

---

[2]this can be viewed as estimate of the image contrast.

contrast, on the one hand, and more pronounced intermediate peaks, on the other hand, when going to more coherent illumination.
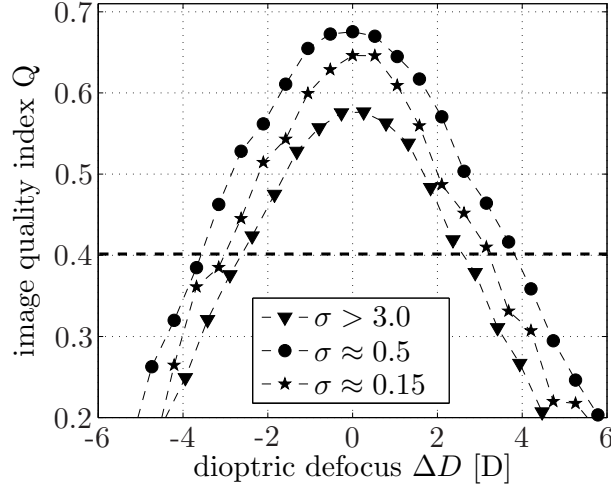
Figure 3.4: *Image quality index Q (as defined in eq. 3.5) as a function of defocus ΔD for text targets with $\alpha_\mathrm{v} = 0.46$ deg. For clarity, the plots were limited to three selected coherence levels. The horizontal dashed line at 0.4 shows the quality limit as explained in the text.*

From these results, the DOF is obtained as the difference between the two defocusing values (on both sides of the optimum focus) where the corresponding quality index equals 0.4. This quality limit of 0.4 has been found by direct visual assessments of the image quality. It is kept constant for all index-based evaluations. The result of this evaluation for all coherence levels and three selected viewing angles α_v is depicted in Fig. 3.5.

3.3.3. Direct visual assessment

For the direct visual evaluation of the image quality, a modified version of the psychophysical 'method of limits' was used [70]: In this evaluation, five subjects were asked to compare the quality of the image series $(\alpha_\mathrm{v}, \sigma)$ within and between the σ-classes. Based on that, the subjects determined for each $(\alpha_\mathrm{v}, \sigma)$-combination two boundaries on both sides of the focus where the corresponding text sample still seems to be in-focus (see Fig. 3.6): The difference between these two boundaries corresponds to the searched DOF. This procedure was repeated several times for each $(\alpha_\mathrm{v}, \sigma)$-combination to reduce any subjective errors. Fi-
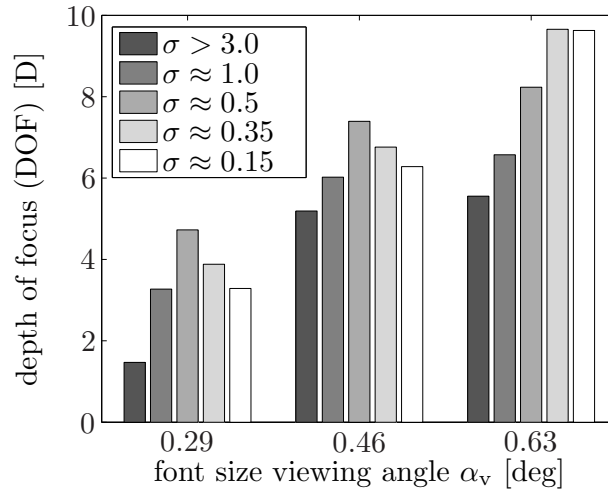
Figure 3.5: *DOF for various font size viewing angles α_v when illuminated by light of different coherence levels σ. The data are based on the quality index assessment.*

nally, from all these DOF-results the arithmetic mean and the standard deviation were calculated for each (α_v, σ).

Figure 3.7 shows the results of the direct visual assessment for all five evaluated coherence levels. The error bars are defined by the standard deviation of all the evaluation runs. The data in Fig. 3.7 are in good agreement with the results obtained by the objective quality index method (see Fig. 3.5). However, for larger viewing angles α_v, the index-based method tends to exhibit higher absolute DOF-results than the visual assessment.

3.3.4. Discussion

Several results are revealed by Figs. 3.5 and 3.7: First, both figures show that the DOF can be extended by increasing the spatial coherence as expected by the contrast analysis. However, it also shows that for most viewing angles the optimum in DOF is not reached with fully coherent light but with partially coherent light. For instance, for $\alpha_v = 0.46\,\mathrm{deg}$, the maximum DOF of about $6.7\,\mathrm{D}$ is reached with a spatial coherence level of $\sigma \approx 0.5$. For higher coherent illumination, the DOF drops to approximately $5.4\,\mathrm{D}$. This effect might be attributed to the disturbing impact of defocusing intermediate peaks when the illumination light is highly coherent. As a third result of Fig. 3.7, it can be seen that the optimum coherence level σ depends on the font size viewing angle α_v: For small α_v-values, the optimum shifts to higher σ-levels while for
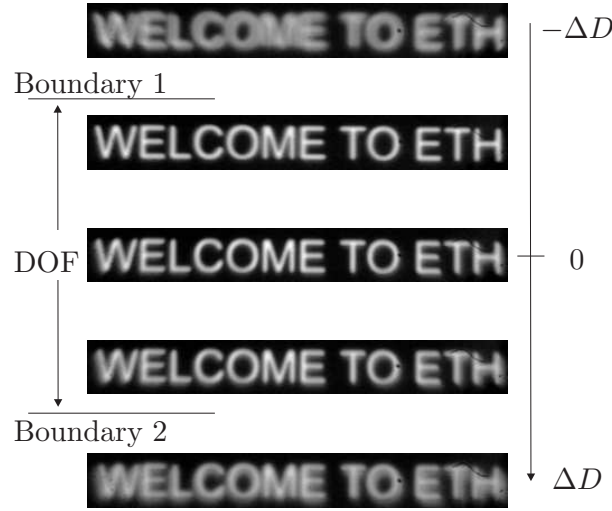
Figure 3.6: *Part of a $(\alpha_{\mathrm{v}}, \sigma)$-series as a function of defocus ΔD. The subjects were asked to determine the boundaries 1 & 2 by comparing the image quality as explained in the text.*
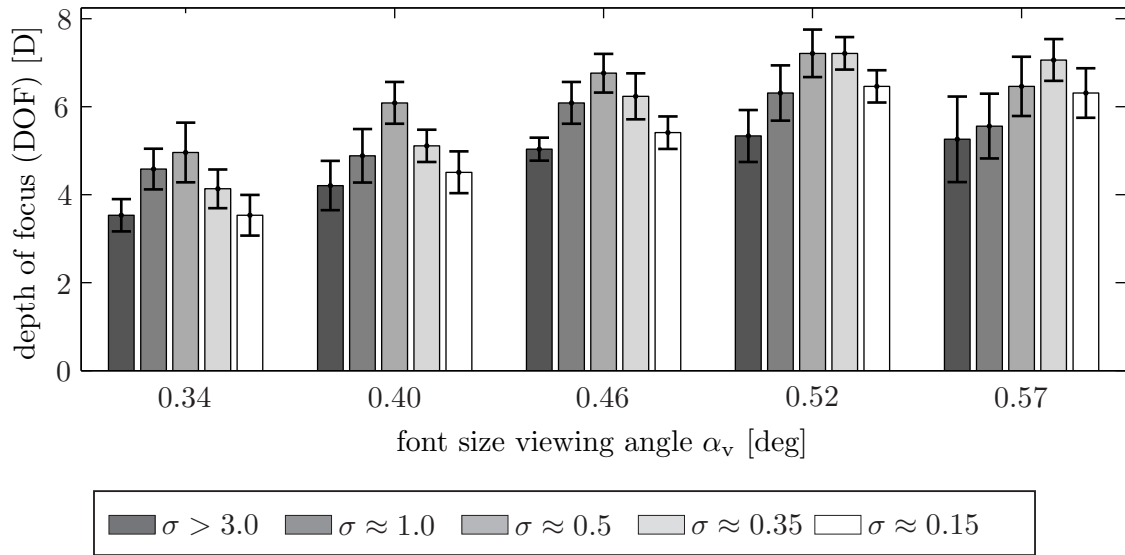
Figure 3.7: *DOF for various font size viewing angles α_{v} when illuminated by light of different coherence levels. The data are based on the direct visual assessment. Note that the α_{v}-values on the x-axis correspond to font sizes $F_s = 12/14/16/18/20\,pt$ on a screen when viewed from a distance of $50\,cm$.*

larger α_v-values the coherent illumination gets better and better. This is presumably caused by two different facts: First, the cut-off frequency for the coherent case is only half of the corresponding incoherent value. Thus, small structures are better imaged in the incoherent limit than in the coherent one. Second, the coherent intermediate peaks depend quadratically on the spatial frequency. This implies that the disturbing impact of these intermediate peaks seems to be higher for small structures than for larger ones.

The coherent effects, especially the coherent noise and the edge ringing (see section 2.3), turned out to have little impact on the image quality. The reason is probably that the characteristic wavelengths of the edge ringing oscillations are normally larger than the typical width of a letter line so that the irradiance oscillations on the letters are negligible.

3.4. Conclusion

We have discussed the impact of light coherence on the depth of focus of a 'Maxwellian-view' retinal projection display. Thereby, contrast as well as image quality measurements were carried out on a bench model, revealing the following results:

- Increasing the spatial coherence level of the illumination light enhances the contrast of a defocused image and consequently, results in an increase of the DOF. However, coherence that is too high leads to the occurrence of intermediate peaks at the irradiance minima when defocusing. These peaks degrade the image quality and reduce the improvements of the defocusing properties reached by using higher coherent light.

- Image quality measurements of text targets showed that the best DOF is reached with partially coherent light of $0.35 \lesssim \sigma \lesssim 0.5$. The optimum σ-level depends slightly on the targeted font size viewing angle α_v: for smaller structures the optimum shifts to a lower coherence level while for larger structures the best DOF is reached with higher spatially coherent light.

- Coherence effects such as edge ringing and coherent noise have little effects only on the perceived image quality. Speckle noise does not occur as no diffuser is used in the optical path.

4

Multiple imaging technique[*]

To extend the DOF further, this chapter discusses the use of the multiple imaging technique. In this technique, an appropriate phase-only mask produces a series of images of the LCD at various focal planes on the retina. Based on a schematic eye model and on a partial coherence simulation tool, the projected retinal images of a text target are calculated. The evaluation of the retinal images shows that this approach is promising provided that partially coherent light is used. In this case, psychometric measurements with subjects reveal that the depth of focus for reading text can be extended by a factor of up to 3.2.

---

[*]This chapter is mainly based on Ref. [71]

4.1. Approach

In this chapter, we discuss a novel approach which is based on the partially coherent multiple imaging technique known from optical microlithography [72, 73, 74]. In this method a pupil phase mask is introduced at the aperture stop plane of the partially coherent retinal projection display. This additional mask produces a series of images of the LCD at various focal planes, shifted to each other by an individual phase. Due to the use of partially coherent light, the multiple images are added to one another partially coherently, so that their individual phase and amplitude distribution determines the final retinal image. Provided that the introduced phase mask is a phase-only mask and has, thus, full transmission within the aperture, this approach allows to significantly extend the depth of focus without reducing the lateral resolution.

To examine the potential of adding a pupil phase mask into a retinal projection display, the images on the retina are simulated on the basis of an accommodation-dependent eye model for various values of coherence level and eye accommodation. The resulting retinal images are then evaluated using a combination of objective image quality criteria and psychometric measurements.

4.2. Theoretical background

4.2.1. System overview

Figure 4.1 shows the principal setup of the considered retinal projection display representing a "Maxwellian-view optical system". Except for the additional phase mask, the setup is identical to the one discussed in chapter 2.
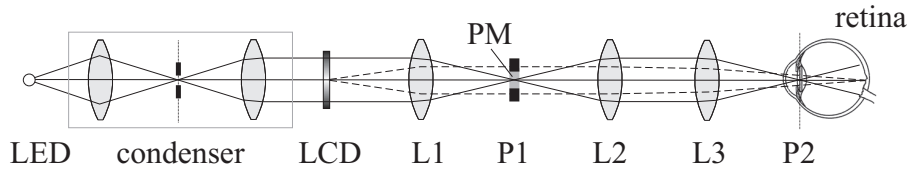
Figure 4.1: *Principal setup of the modified retinal projection display. It differs from the setup shown in Fig. 2.1 in the additional phase mask PM at the aperture stop plane P1.*

4.2.2. Theoretical approach

For the mathematical description of the system we apply the Hopkins theory on image formation in partially coherent systems [53, 75]. A detailed discussion of this theory is given in appendix B.

The most relevant parameter in this theory for defocusing considerations is the complex coherent transfer function $H(x, y)$. As discussed in section 2.2.1, the coherent transfer function is proportional to the generalised exit pupil function $\mathcal{P}(x, y)$. For our setup, the phase factor $\psi(x, y)$ of $\mathcal{P}(x, y)$ (as introduced in eq. 2.2) can be split into those aberration factors introduced by the display's optical system and those by the ocular system. Thus, the coherent transfer function can be expressed as:

$$H(x, y) \propto P(x, y) \cdot e^{i\psi(x,y)} = P(x, y) \cdot e^{i(\psi_{disp}(x,y)+\psi_{eye}(x,y))} \qquad (4.1)$$

Here $\psi(x, y)$ represents the total wavefront aberration error at the exit pupil point (x, y). The ocular wavefront error $\psi_{eye}(x, y)$ is given by the human eye and considers the accommodation of the eye as well as ocular aberrations. The display part $\psi_{disp}(x, y)$ is mainly defined by the phase pupil mask PM and thus, can be determined to some extent in the design process.

Since for ideal imaging the total wavefront error $\psi(x, y)$ should be 0 for all (x, y)-values, the challenge is now to design the phase mask in a way, that the sum of $\psi_{eye}(x, y)$ and $\psi_{disp}(x, y)$ remains small for eye accommodation values within the intended DOF range. To achieve that, the idea is to simultaneously project a series of images of the LCD at various eye accommodation values. Evidently, this DOF-improvement can only be achieved at the cost of a degraded image quality of the in-focus image. It is one of the challenges of this chapter to explore this trade-off.

In mathematical terms, the design process can be started as follows: To adjust the display to one specific dioptric accommodation δ, the phase map $\psi_{disp}(x, y)$ may be formed as (cf. eq. 2.2):

$$\psi_{disp}(\rho) = \frac{\pi}{\lambda}\delta\rho^2 \qquad (4.2)$$

where $\rho^2 = x^2 + y^2$ is the radial coordinate in the exit pupil and λ the wavelength of the illumination light. Thus, for multiple imaging, where the mask PM should produce a series of images of the LCD at various focal planes (described by the corresponding dioptric focus values δ_l),

shifted to each other by an individual phase ϕ_l and weighted by the real values w_l, the phase mask PM can generally be described as:

$$\psi_{disp}(\rho) = angle \left[\sum_{l=0}^{L} w_l \, e^{i\phi_l} e^{i\left(\frac{\pi}{\lambda} \delta_l \rho^2 \right)} \right] \tag{4.3}$$

Here, the total number of different image planes is $L + 1$. The function *'angle'* takes the phase angle of the subsequent sum. Note that this definition (in combination with eq. 4.1) limits the mask PM to phase-only masks without apodisation.

4.3. Simulation background

4.3.1. Accommodation-dependent eye model

For our purpose, ideally, the eye model should reproduce both anatomy and optical properties (such as first order aberrations) with a minimum of fitting parameters. Moreover, the model should incorporate the increment of refractive power of the eye during accommodation. Therefore, the simulations presented in this chapter are based on the accommodation-dependent wide-angle schematic eye model as proposed by Navarro et al. and Escudero et al. [76, 77]. A detailed discussion of this eye model is given in Appendix A.

This schematic eye model has been implemented in the ray-tracing software package OSLO [78] in order to calculate the ocular wavefront error map $\psi_{eye}(x, y)$ as a function of eye accommodation, eye pupil diameter and eye orientation. The wavefront error map has been expressed by the Zernike polynomial expansion[1]. The resulting Zernike coefficients act as input for the subsequent simulation of the spatially partial coherence.

4.3.2. Simulation of partial coherence

The calculation of the retinal images is mainly based on the eqs. 4.3 and B.12 and is carried out on a slightly adapted version of the SPLAT 5.0 software package. This program has been developed by the University of California at Berkeley for simulating two-dimensional projection-printing with partial coherence [79]. The system is described by several

---

[1] up to the 36th Zernike polynomial, thus describing all aberrations up to the ninth order.

input parameters, such as wavelength λ, numerical aperture NA and coherence level σ. The LCD image is defined by its Fourier coefficients $a_{m,n}$ while the parameter set $(\delta_l, w_l, \phi_l)_{l=0..L}$ defines the phase mask PM according to eq. 4.3. Finally, the ocular wavefront map $\psi_{eye}(x,y)$ is read in by their Zernike polynomial coefficients calculated on the basis of the eye model as discussed above.

4.3.3. Simulation parameters

For the design of the phase mask the system parameters according to Fig. 4.1 were defined as follows:

Illumination wavelength	$\lambda = 635\,\mathrm{nm}$
Diameter of eye pupil	$\varnothing_{\mathrm{pup}} = 3\,\mathrm{mm}$ (bright-adapted)
Diameter of aperture stop Di_1	$\varnothing_1 = 2.1\,\mathrm{mm}$
Magnification of relay lens system	$M = 0.65$

4.3.4. Exit pupil and eye motion

The diameter $\varnothing_1$ of the aperture stop (and thus the exit pupil's diameter) may seem to be rather small. This appears to be contrary to common HMD designs where normally a large exit pupil is desired in order to increase the so-called eye motion box [80]. This box defines the area where the eye can freely move without loosing the LCD image due to vignetting effects. However, it is important to note that this direct link between exit pupil diameter and eye motion box does not hold for partially coherent or even fully coherent projection displays: In case of fully coherent light, the eye motion box does not depend on the exit pupil's diameter $\varnothing_{\mathrm{exit}}$, but is a function of the eye pupil's diameter $\varnothing_{\mathrm{pup}}$. The exit pupil controls, in this case, the maximum retinal resolution. This can be illustrated by the following considerations: In a fully coherent Maxwellian-view system, a Fourier transform pattern of the LCD image is formed on the plane of the eye pupil. Thereby, the angular spatial frequency ν_x (in cycles/rad) is situated at a distance $x = \pm\lambda\nu_x$ from the pupil centre, independent of any system parameters (in particular independent of the exit pupil) (see left drawing in Fig. 4.2) [49]. The exit pupil diameter determines the cut-off frequency and thus, the retinal resolution: The larger the exit pupil, the more angular frequencies are transmitted.

Aligned eye

exit pupil   eye pupil

ν_x transmitted
on both side

retinal image fine

**Moved eye
within eye motion box**

exit pupil   eye pupil
                    (now decentered)

ν_x still transmitted
on this side

ν_x blocked   central frequency
on this side   still transmitted

retinal image fine

**Moved eye
outside of eye motion box**

exit pupil   eye pupil
                    (now decentered)

central frequency ν_x still trans-
blocked now   mitted on
                    this side
ν_x blocked
on this side

retinal image lost

Figure 4.2: *Schematic illustration of eye motion in case of fully coherent illumination. The dots within the exit pupil depict the Fourier transform pattern of the LCD image.*

When the eye moves, the eye pupil will block some spatial frequencies, but only at one side of the symmetric spectrum as illustrated in the middle drawing of Fig. 4.2. This does not affect the image quality significantly, since the corresponding frequency is still transmitted on the other side. However, when the eye moves further, the retinal image gets lost, as soon as the central frequency at x=0 is blocked by the eye pupil. This happens, when the eye moves more than the current radius of its pupil (see right drawing in Fig. 4.2). Thus, in fully coherent imaging, the eye motion box depends on $\varnothing_{\text{pup}}$ and not on $\varnothing_{\text{exit}}$.

For partially coherent light, the loss of the retinal image is less abrupt. Its overall intensity rather fades out, when the eye moves. In this case, we consider the eye motion box to end when the overall image intensity is reduced by 50%. Additionally, in the partially coherent case the eye motion box depends slightly on $\varnothing_{\text{exit}}$: In our setup with e.g. $\sigma = 0.4$, simulations (as described above) revealed that the eye motion box is about 3 mm in extent so that the eye can move within about ± 10 deg without loosing the image. If the exit pupil was expanded to 8 mm in diameter, the eye motion box would only increase to ≈ 3.2 mm or ± 10.7 deg, respectively. In contrast, in case of a darker-adapted eye pupil of 6 mm in diameter, the motion box is increased to 6 mm or $\approx \pm 19.5$ deg. These simulation results show that in case of partial coherence the eye motion box is still mainly determined by the eye pupil and can not be extended significantly by expanding the exit pupil. Consequently, expanding the exit pupil in our setup to e.g. 8 mm would lead

to an higher retinal image resolution, but at the cost of degraded defocusing properties. The eye motion box, however, would stay practically unchanged.

To achieve a substantially larger eye motion box, a more promising method is to apply a pupil tracking system which adjusts the display's exit pupil dynamically to the current position of the eye pupil, preventing thus vignetting effects [66].

4.4. Design of the phase mask

4.4.1. Evaluation method

Due to the nonlinearities in partially coherent imaging [81], the quality of the retinal images depends on the specific image and cannot be described generally by a transfer function as widely used for incoherent and fully coherent systems. In particular, the common modulation transfer function (MTF) is unsuited as the MTF is inherently linked to systems being linear in intensities. This, however, is not fulfilled in partially coherent systems so that the MTF is not appropriate to describe the quality of the imaging process. Given a specific object and a fixed optical imaging configuration, one would need to propagate the mutual coherence function to evaluate the image quality. For a set of objects, other metrics may more conveniently be defined as further described in this paper. Consequently, a test-image is defined which is representative for a likely application scenario of the display. For see-through HMDs, reading text is considered as one of the most likely and stringent tasks in any scenarios [66]. Thus, the test image defined for the subsequent evaluations consists of a short text pattern written in arial font and capital letters at various font sizes. In retinal projection setups, the decisive measure for the fonts is not the original size of the letters on the LCD, but the viewing angle α_v (in degree), the letters subtend. Therefore, the font sizes in the subsequent simulations are indicated by their viewing angles [2].

In the design process of the phase mask, the quality of the projected retinal image is evaluated by calculating the contrast comparison function C as well as the structural function S between the retinal image X and the original reference image Y as suggested by Wang et al. [82].

---

[2]To illustrate this notion: a viewing angle of e.g. $\alpha_v = 0.4\,\mathrm{deg}$ corresponds to common 14-pt capital letters on a screen when viewed from a distance of 50 cm. (The height of standard 14-pt capital letters on a screen is 3.5mm)

The contrast comparison function C compares the contrast of the test image X with regard to the reference image Y while the structural function S measures the degree of structural similarity between X and Y. Thus, these two indices act well as a metric for the quality of the test image X. In mathematical terms, the two functions are defined as follows:

$$C = \frac{2s_X s_Y}{s_X^2 + s_Y^2} \qquad \text{and} \qquad S = \frac{s_{XY}}{s_X s_Y} \tag{4.4}$$

with

$$s_X^2 \;\; = \;\; \frac{1}{B-1} \sum_{m,n=1}^{B} (X_{m,n} - \overline{X})^2$$

$$s_{XY} \;\; = \;\; \frac{1}{B-1} \sum_{m,n=1}^{B} (X_{m,n} - \overline{X})(Y_{m,n} - \overline{Y})$$

C and S are first calculated locally in a sliding window of size $B \times B$ [3], which moves pixel by pixel horizontally and vertically through all rows and columns of the image. Finally, the overall function value is given as the arithmetic mean of all these local results. $X_{m,n}$ and $Y_{m,n}$ are the values of the (m,n)-pixel in the sliding window of X and Y, respectively. $\overline{X}$ and $\overline{Y}$ denote the arithmetic mean values and s_X and s_Y are the corresponding variance values. Their dynamic ranges are $[-1, 1]$; the best value 1 for both functions is only obtained when $Y_{m,n} = X_{m,n} + b$ for all pixels (m, n) where b is constant.

4.4.2. Finding PM coefficients $(\delta_l, w_l, \phi_l)_{l=0..L}$

Basically, the phase mask, which optimises the display's DOF, should now be found by varying the coefficients $(\delta_l, w_l, \phi_l)_{l=0..L}$. In order to reduce the dimensionality of the parameter space, the search will be limited to focal planes which are equally weighted (i.e. $w_l = 1$, $\forall l$) and equally distributed (with a constant spacing $\Delta\delta$) (i.e. $\delta_l = l\,\Delta\delta + \delta_0$, $l = 0..L$). Thus, the planes are equally distributed within a range $[\delta_0, \delta_0 + L\Delta\delta]$. Furthermore, the phase shifts are limited to the range $[0, \pi]$ so that the images from the two extreme focal planes at δ_0 and δ_L superpose destructively. Within this range, the phases are assumed

---

[3] in our case, B=30

to be equally distributed (i.e. $\phi_l = l\,\pi/L$). Thus, for our specific case, eq. 4.3 turns out to be:

$$\psi_{disp}(\rho) = angle\left[\sum_{l=0}^{L} e^{i\pi l\left(\frac{1}{L} + \frac{\Delta\delta}{\lambda}\rho^2\right)}\right] + \frac{\pi}{\lambda}\delta_0\rho^2 \qquad (4.5)$$

The set of free coefficients is reduced to $(\Delta\delta, L, \delta_0)$. With regard to the subsequent simulations it is more illustrative to replace this set by the equivalent set $(\epsilon, \Delta\delta, \overline{\delta})$. Here, $\epsilon = L\Delta\delta = \delta_L - \delta_0$ signifies the dioptric extent over which the $L + 1$ planes are distributed and $\overline{\delta} = \delta_0 + \epsilon/2 = (\delta_L + \delta_0)/2$ indicate the mean value of all focal planes. Consequently, the focal planes δ_l are equally distributed in the range $[\overline{\delta} - \epsilon/2, \overline{\delta} + \epsilon/2]$.

The potential of this approach has been considered separately for three different coherence levels, $\sigma = 0$ (coherent), $\sigma = 0.5$ (partially coherent) and $\sigma = \infty$ (incoherent).

4.4.3. Partially coherent illumination $\sigma = 0.5$

We will start with the partially coherent case $\sigma = 0.5$. First, the focal plane spacing $\Delta\delta$ was set to $0.5\,\mathrm{D}$ while ϵ was varied. The mean value $\overline{\delta}$ was set to $3.5\,\mathrm{D}$.

Figure 4.3 shows the results for various ϵ-values. The curves for $\epsilon = 0$ represent the corresponding unifocal system for comparison. All data were calculated with partially coherent illumination of $\sigma = 0.5$ and a font size subtending a viewing angle $\alpha_v = 0.4\,\mathrm{deg}$. When increasing the extent ϵ over which the focal planes are equally distributed, both the contrast (described by C) and the structural quality (described by S) are strongly improved for $\Delta D \approx 0\,\mathrm{D}$ and $\Delta D \approx 5.5\,\mathrm{D}$. This improvement occurs at the cost of the image quality for accommodation values in the intermediate range $(2\,\mathrm{D} \lesssim \Delta D \lesssim 4\,\mathrm{D})$. However, for the structural quality S the degradation is small $(\Delta S \approx -0.02)$ while the reduction in contrast is more substantial $(\Delta C \approx -0.17)$. As the display's DOF should ideally be larger than $5\,\mathrm{D}$, an ϵ-value is required at which the image quality in the range between $5\,\mathrm{D}$ and $6\,\mathrm{D}$ is similar to the one at $\Delta D \approx 0\,\mathrm{D}$. Furthermore, the image quality for an unaccommodated eye $(\Delta D = 0\,\mathrm{D})$ is of special importance since viewing at infinity is more likely to occur in typical see-through application scenarios than viewing at very close objects. Thus, $\epsilon = 11\,\mathrm{D}$ turns out to be a good compromise between good image quality at the edges of the considered accommodation range $[0\,\mathrm{D}, 6\,\mathrm{D}]$ and reduced quality in the

(a) contrast function C   (b) structural function S
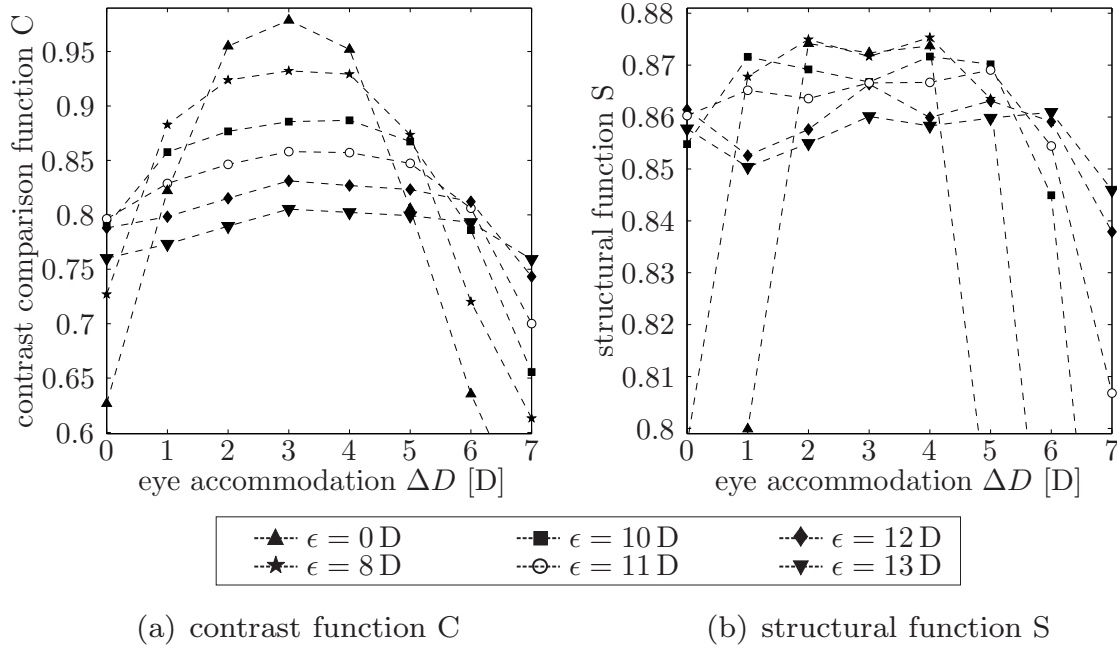
Figure 4.3: *Image quality functions C and S in terms of eye accommodation ΔD for various values of ϵ and with $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. The data were calculated with $\sigma = 0.5$ and a font size viewing angle $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$.*

intermediate part. With $\epsilon = 11\,\mathrm{D}$, this phase mask consists of 23 image planes equally distributed with a spacing of $0.5\,\mathrm{D}$ between $-2\,\mathrm{D}$ and $9\,\mathrm{D}$.

Based on this phase mask, we now vary the focal spacing $\Delta\delta$ while keeping $\epsilon = 11\,\mathrm{D}$ constant. The results are shown in Fig. 4.4. Both, the contrast and the structural similarity are improved when reducing the focal spacing down to $\Delta\delta \approx 0.5\,\mathrm{D}$. For $\Delta\delta$-values smaller than $0.5\,\mathrm{D}$ the image quality stays rather constant (i.e. $|\Delta C| < 0.014$ and $|\Delta S| < 0.005$) within the accommodation range of interest. This is not surprising as any values $0 < \Delta\delta \leq 0.5\,\mathrm{D}$ leads to practically identical phase mask profiles. Consequently, a spacing value $\Delta\delta = 0.5\,\mathrm{D}$ turns out to be a good choice. Smaller values $\Delta\delta \leq 0.5$ are also possible and produce nearly identical results. Finally, the third free parameter, the mean value of all focal planes $\overline{\delta}$, may be used to shift the S- and C-curves relative to the eye accommodation. This will be discussed in the context of the psychometric experiments.

(a) contrast function C

(b) structural function S
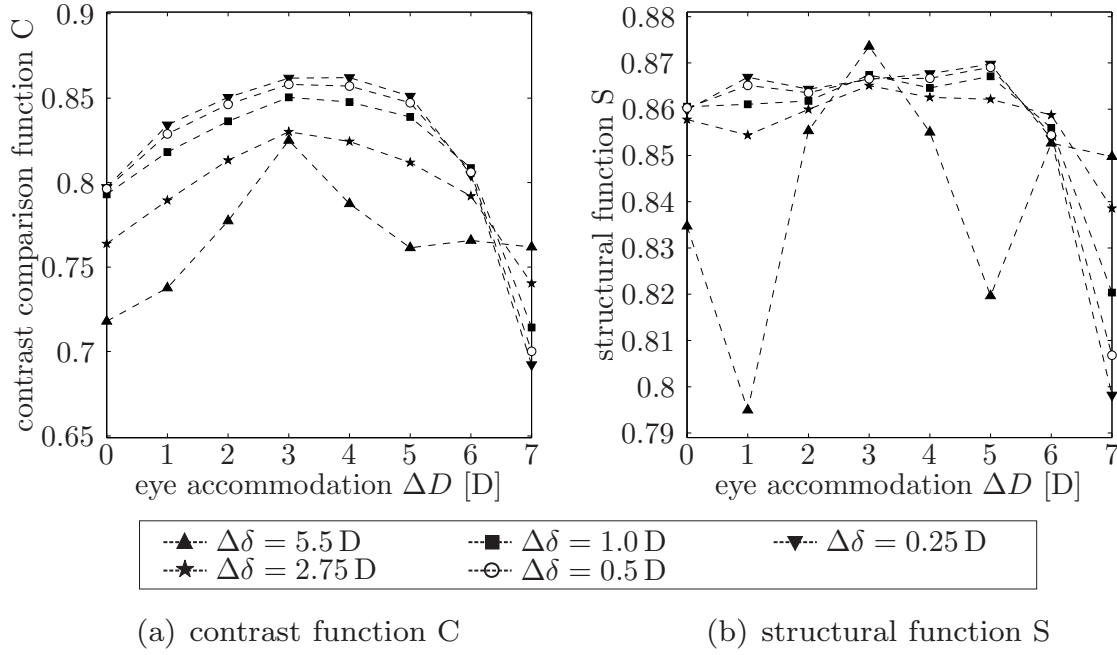
Figure 4.4: *Image quality functions C and S in terms of eye accommodation ΔD for various values of $\Delta \delta$ with $(\epsilon, \overline{\delta}) = (11\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\sigma = 0.5$ and $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$.*

4.4.4. Fully coherent illumination $\sigma = 0$

The same analysis can be made for fully coherent light. Figure 4.5 shows the corresponding results when the dioptric extent ϵ is varied. Again, the other coefficients were kept constant: $\Delta \delta = 0.5\,\mathrm{D}$ and $\overline{\delta} = 3.5\,\mathrm{D}$. Obviously, the contrast results are better than those for the partially coherent case. This is not surprising when considering that coherent imaging even provides contrast 1 for in-focus imaging. However, the structural quality is substantially more degraded than for $\sigma = 0.5$: For small ϵ-values ($\epsilon \lesssim 9\,\mathrm{D}$) the structural similarity is comparable to those with $\sigma = 0.5$ at least for accommodation values between about $2\,\mathrm{D}$ and $4\,\mathrm{D}$. But at the edges of the considered range the structural quality decreases substantially. For higher ϵ-values ($\epsilon \gtrsim 15\,\mathrm{D}$) the degradation at the edges is reduced, but at a much lower overall level ($\Delta S \approx -0.1$) and at the cost of a decrease for accommodation values around $3\,\mathrm{D}$. These effects can mainly be attributed to disturbing coherent interferences between the simultaneously projected images.

0.7

(a) contrast function C   (b) structural function S
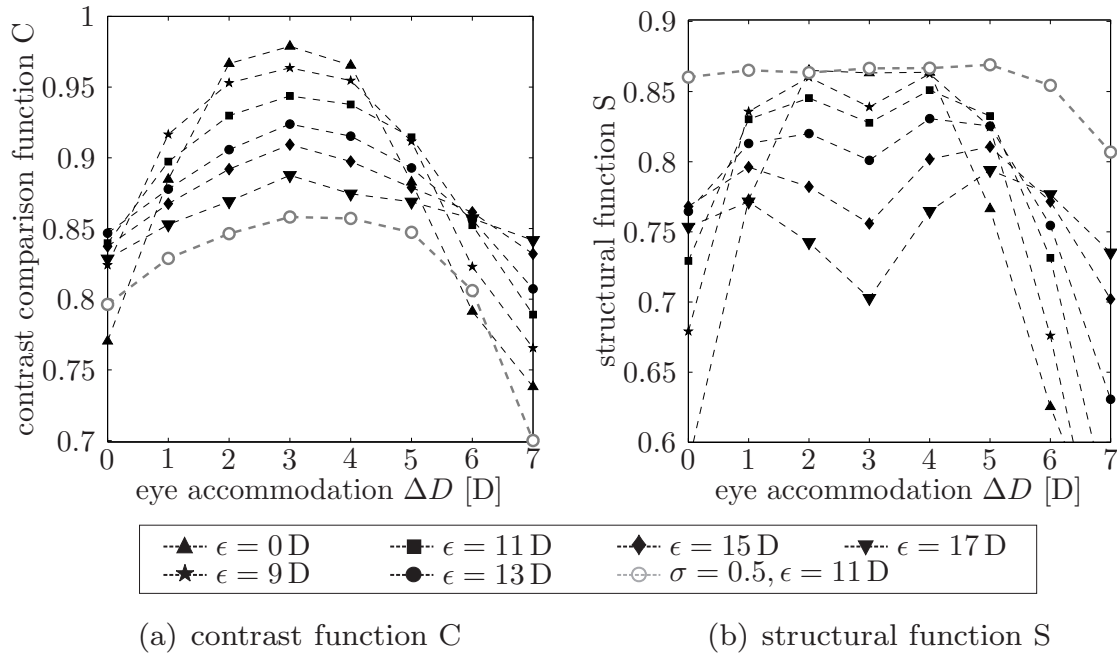
Figure 4.5: *Image quality functions C and S for $\sigma = 0$ in terms of eye accommodation ΔD for various values of ϵ and with $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$. The grey unfilled symbols show the former results for $\epsilon = 11\,\mathrm{D}$ with partially coherent illumination for comparison.*

4.4.5. Fully incoherent illumination $\sigma = \infty$

For fully incoherent light, the results for the quality functions C and S seem to be swapped to each other, as shown in Fig. 4.6: The structural

(a) contrast function C   (b) structural function S
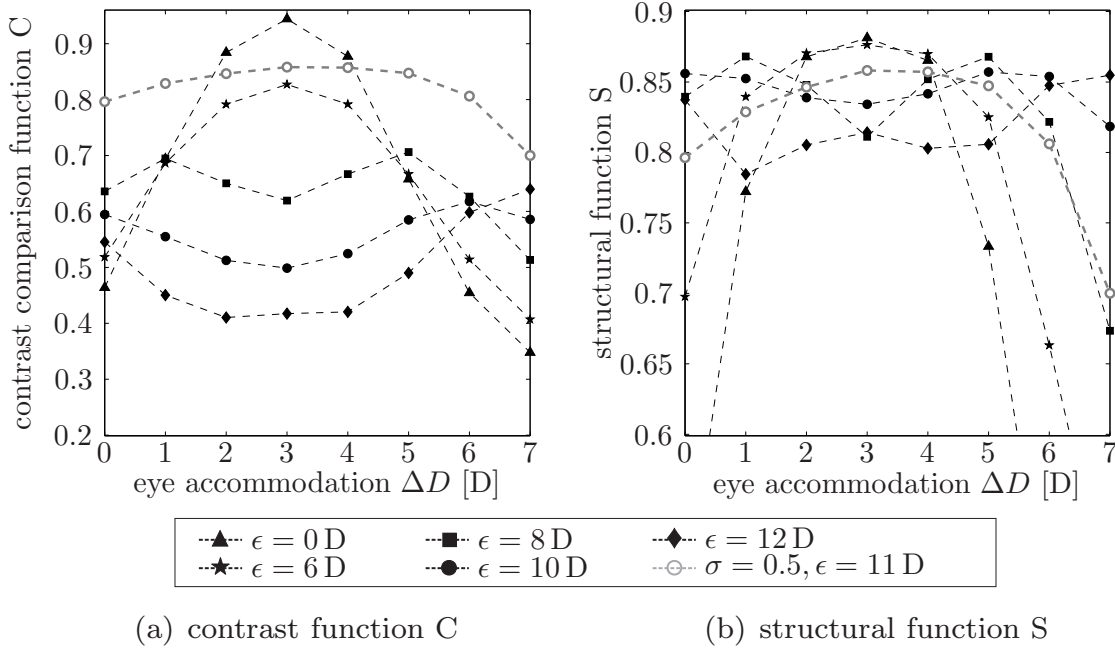
Figure 4.6: *Image quality functions C and S for $\sigma = \infty$ in terms of eye accommodation ΔD for various values of ϵ and with $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$. The grey unfilled symbols show the former results for $\epsilon = 11\,\mathrm{D}$ with partially coherent illumination for comparison.*

quality for $\epsilon = 8\,\mathrm{D}$ turns out to be similar to the corresponding one in the partially coherent case for $\epsilon = 11\,\mathrm{D}$. However, now the image quality suffers from a severe loss in contrast.

4.4.6. Final phase mask profile

Given some restrictions, the first analysis has revealed that multiple imaging is most promising in combination with partially coherent illumination where the degradations in contrast and structural quality due to the extended depth of focus are limited. Consequently, the following explorations focus on partially coherent light and the use of the phase mask as described by the three coefficients $(\epsilon, \Delta\delta, \overline{\delta}) = (11\,\mathrm{D}, 0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. The corresponding phase profile $\psi_{disp}(\rho)$ of the

phase mask is shown in Fig. 4.7. Note again, that the transmission rate is 100% all over the mask.
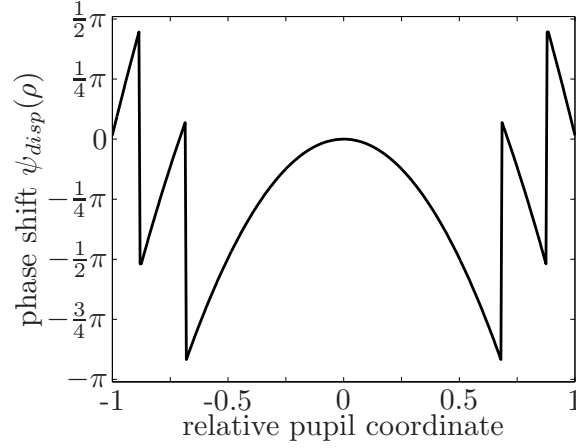
Figure 4.7: *Phase profile $\psi_{disp}(\rho)$ of the rotationally symmetric phase-only mask PM.*

To give an idea of the image quality, Fig. 4.8 shows the retinal images for the three considered coherence levels at four different eye accommodation values. For comparison, the first row contains the corresponding images for the unifocal case ($\epsilon = 0$) and incoherent light ($\sigma = \infty$). Figure 4.8 confirms the results as obtained by the contrast and structural evaluation: While the fully coherent case suffers from structural degradation, the fully incoherent case is characterised by a reduced contrast due to the phase mask. Again, the improvement for the partial coherent case (second row in Fig. 4.8) is evident.

4.5. Evaluation

So far, we have designed the phase mask on the basis of three coherence levels $\sigma = 0/0.5/\infty$ and for one font size viewing angle $\alpha_v = 0.4\,\mathrm{deg}$. Now, we will explore the potential of this designed phase mask (as defined in Fig. 4.7) in more detail. To this end, the retinal images of the text samples are calculated for three font size viewing angles α_v and for more coherence levels σ and accommodation states ΔD. Subsequently, the quality of the retinal images are evaluated using a combination of objective image quality criteria and psychometric experiments.

	$\Delta D = 0\,\mathrm{D}$	$\Delta D = 3\,\mathrm{D}$	$\Delta D = 5\,\mathrm{D}$	$\Delta D = 6\,\mathrm{D}$
reference $(\epsilon = 0\,\mathrm{D}, \sigma = \infty)$				
partial coherence $(\epsilon = 11\,\mathrm{D}, \sigma = 0.5)$				
full coherence $(\epsilon = 11\,\mathrm{D}, \sigma = 0)$				
full incoherence $(\epsilon = 11\,\mathrm{D}, \sigma = \infty)$				
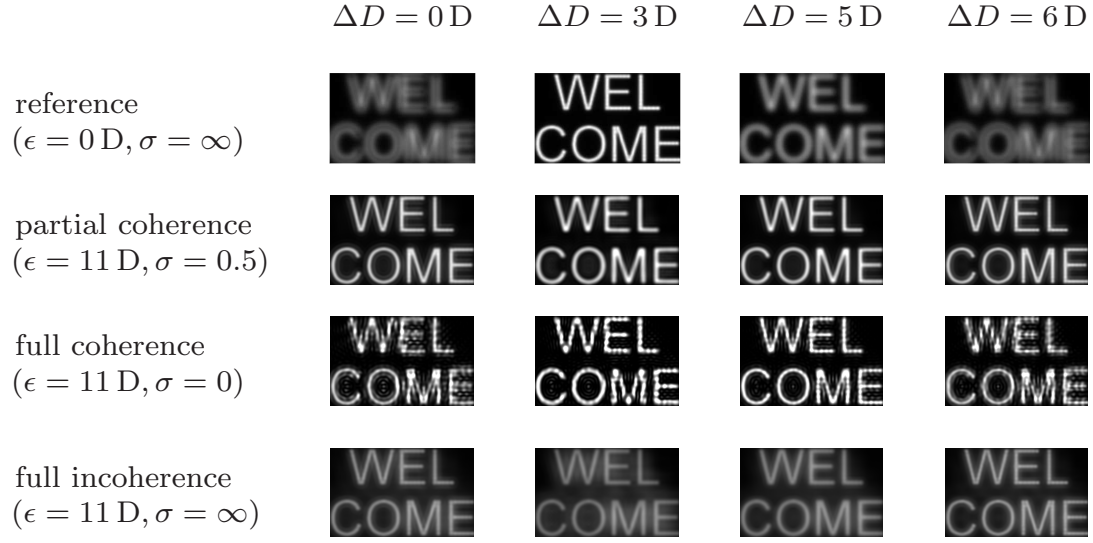
Figure 4.8: *First row: retinal images for $\epsilon = 0\,\mathrm{D}$ and incoherent light $(\sigma = \infty)$ for comparison. The other three rows show the retinal images for the three coherence levels when the designed phase-mask PM (see Fig. 4.7) is applied. The columns show the images for four different eye accommodation values ΔD. Again, $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$ and $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$.*

4.5.1. Metrics-based Evaluation

For the objective image quality evaluation, the same criteria are applied as above: the contrast comparison function C and the structural function S as defined in eq. 4.4. Figure 4.9 shows the corresponding results for various coherence levels σ and eye accommodation values ΔD.

The contrast values decrease continuously when the coherence of the light is reduced. However, the decrease is small as long as $\sigma \lesssim 0.4$. The severe contrast decrease in the middle of the considered accommodation range appears at about $\sigma \approx 0.8$. Concerning the structural quality, the best values are achieved for intermediate σ-values between about 0.3 and 0.6. For high spatial coherence (i.e. $\sigma \lesssim 0.2$) the decrease for intermediate accommodation values occurs, as already discussed above in case of fully coherent light. For low coherence (i.e. $\sigma \gtrsim 1.0$), the structural quality is reduced for $\Delta D \approx 1.5\,\mathrm{D}$ and $\Delta D \approx 4\,\mathrm{D}$. To evaluate the trade-off between contrast and structural quality, the function values C and S can be averaged over the considered accommodation range $[0\,\mathrm{D}, 6\,\mathrm{D}]$. The results are given in Fig. 4.10. These averaged values represent well the grade of insensitivity of the corresponding quality

(a) contrast function C

(b) structural function S
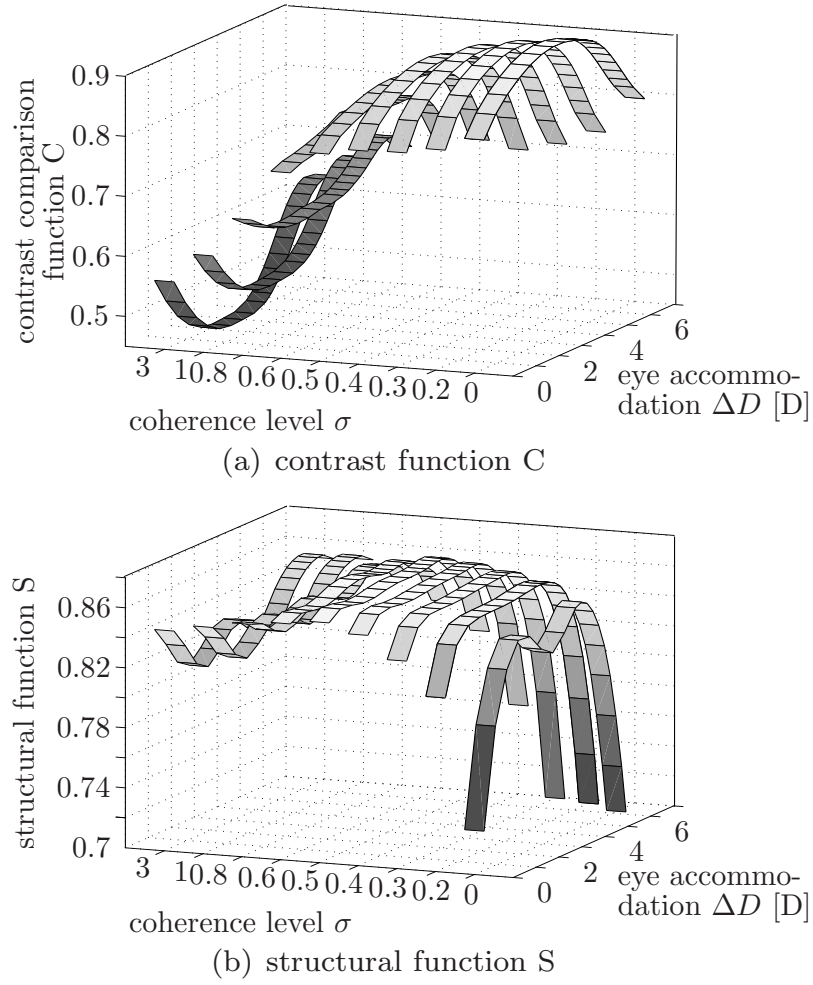
Figure 4.9: *Image quality functions C and S in terms of coherence level σ and eye accommodation ΔD. The phase-mask PM used is described by $(\epsilon, \Delta\delta, \overline{\delta}) = (11\,\mathrm{D}, 0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$.*
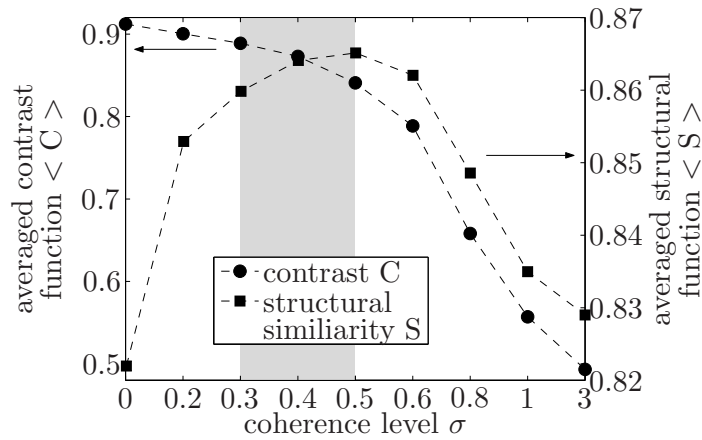
Figure 4.10: *Contrast and structural image quality averaged over the considered range of accommodation $[0\,\mathrm{D}, 6\,\mathrm{D}]$ in terms of σ.*

to eye accommodation. The highest image quality and insensitivity to eye accommodation is obviously achieved for a coherence level between $\sigma \approx 0.3$ and $\sigma \approx 0.5$.

4.5.2. Psychometric experiments

Objective image quality criteria are valuable tools to characterise the quality. However, they cannot replace the direct (but also subjective) assessment of the image quality by people. Consequently, in addition to the objective quality criteria, psychometric measurements were carried out to rate the quality of the simulated retinal images relative to a given reference image.

Method

To this end, eight subjects (mean age: 28.5 years, variance: 2.8 years) with normal or corrected-to-normal vision participated in the tests. All were naive to the purpose of the experiments. For the tests, we used the so-called 'two-alternative unforced-choice task' in combination with the adaptive procedure called PEST [83, 84]. The subjects were sitting in a distance of 50 cm in front of a common computer screen. On the screen, two of the simulated retinal images were displayed simultaneously as stimuli to the subjects using Matlab with the Image Processing Toolbox extension. On each trial, the subjects were asked to judge which of the two images, concurrently presented side-by-side, appeared to be of better quality. As we used an unforced-choice task an additional response alternative *'equal quality'* was offered. One of the presented images was the reference image which was kept constant from trial to trial, but changed the side randomly, while the other was one image out of a specific query image series which was defined individually according to the experiment's purpose. After each subject's response, the PEST algorithm adjusted the query image for the subsequent trial. The PEST algorithm is based on a statistical estimation of the subject's threshold by fitting a logistic psychometric function to all results obtained from task start-up.

The reference image was chosen due to the following considerations: To evaluate the impact of the phase mask onto the depth of focus, the query images should be compared to an image which represents a common display setup (without phase mask) at the edge of the respective depth of focus. Thus, as reference image we took that retinal image which results in the simulation for incoherent light, without a phase

mask and at a defocus corresponding to Rayleigh's quarter-wavelength criterion [21]. In our setup, the Rayleigh criterion gives a defocus of $0.85 \, \mathrm{D}$. The reference image for $\alpha_\mathrm{v} = 0.4 \, \mathrm{deg}$ is shown in the inset of Fig. 4.11(a).

All eight subjects performed two different experiments: The first experiment had the purpose to evaluate the image quality at $\Delta D = 0 \, \mathrm{D}$ by varying the dioptric mean value $\overline{\delta}$ of the phase mask. The second experiment was to measure the ΔD-range for which the query images are assessed to be of at least equal quality as the reference image. Both experiments included 12 separate test runs, each of them characterised by one of the 12 combinations $(\alpha_\mathrm{v}, \sigma)$ with the font size viewing angles[4] $\alpha_\mathrm{v} = 0.34/0.4/0.46 \, \mathrm{deg}$ and the coherence levels $\sigma = 0.3/0.4/0.5/0.6$. Note finally, that the eye accommodation values ΔD refer - as before - to the accommodation of the schematic eye model in the simulations and not to the current eye accommodation of the subjects during the experiments. The subjects' eye accommodation is fixed due to the constant distance to the computer screen.

First experiment: Image quality at $\Delta D = 0 \, \mathrm{D}$

The image quality for an unaccommodated eye (corresponding to $\Delta D = 0 \, \mathrm{D}$) is of special importance since viewing at infinity is more likely to occur in typical see-through application scenarios than viewing at very close objects. The quality at $\Delta D = 0 \, \mathrm{D}$ can be controlled by adjusting the dioptric mean value $\overline{\delta}$ (as discussed above) since variations in $\overline{\delta}$ do not alter the shape of the C- and S-curves, but shift the curves relative to the eye accommodation ΔD, only. Thereby, lower $\overline{\delta}$-values improve the image quality at $\Delta D = 0 \, \mathrm{D}$. Thus, the first psychometric experiment was carried out to find an appropriate $\overline{\delta}$-value. To this end, the query image series used for this experiment consisted of the retinal images at $\Delta D = 0 \, \mathrm{D}$ for various $\overline{\delta}$-values. The threshold was defined as that $\overline{\delta}$-value which yields 50% of answers *'better quality'* or *'equal quality'* compared to the reference image. Figure 4.11(a) illustrates exemplarily the estimated psychometric functions and the derived thresholds - indicated by squares (■) - for one subject and all three font size viewing angles $\alpha_\mathrm{v} = 0.34/0.4/0.46 \, \mathrm{deg}$ at $\sigma = 0.4$. Other psychometric curves are similar.

---

[4]Note that these viewing angles correspond to font sizes $F_s = 12/14/16 \, \mathrm{pt}$ on a screen when viewed from a distance of $50 \, \mathrm{cm}$.

Figure 4.11(b) shows the summarised results for all values of α_v and σ tested. The boxes show the lower quartile, median, and upper quartile values of the individual threshold results. The dashed lines extending from each end of the boxes indicate the extent of the rest of the data within 1.5 units of interquartile range. The crosses represent outliers beyond this limit. For the plot, the results of seven out of eight subjects were considered. For one subject the logistic psychometric function could not be fitted to the subject's response, as this subject preferred smoother image patterns so that he assessed slightly defocused images better than in-focus images. For the smallest font size ($\alpha_v = 0.34\,\text{deg}$) and $\sigma = 0.5/0.6$, no threshold could be determined since none of the query images was assessed to be better or equal than the reference image due to the reduced contrast for these σ−values.

The threshold results for $\alpha_v = 0.4\,\text{deg}$ and $\alpha_v = 0.46\,\text{deg}$ lie all well above $\overline{\delta} = 3.5\,\text{D}$. Consequently, setting $\overline{\delta} = 3.5\,\text{D}$ guarantees a satisfactory image quality at $\Delta D = 0\,\text{D}$. However, for $\alpha_v = 0.34\,\text{deg}$, the thresholds lie rather close or even below $\overline{\delta} = 3.5\,\text{D}$ so that for this size the image quality at $\Delta D = 0\,\text{D}$ is critical, if $\overline{\delta} = 3.5\,\text{D}$. If a better quality for $\alpha_v = 0.34\,\text{deg}$ is desired, $\overline{\delta}$ might be lowered. Such a focal shift, however, would reduce the DOF-values (as obtained in the second experiment) for all font sizes accordingly. For the following, we will set $\overline{\delta}$ to 3.5 D, but keeping in mind, that for $\alpha_v = 0.34\,\text{deg}$ the image quality at $\Delta D = 0\,\text{D}$ is critical.

Second experiment: Estimation of depth of focus

The second psychometric experiment had the purpose to measure the dioptric range for which the simulated retinal images are considered as being at least of equal quality as the reference image. In this experiment, the query image series included the retinal images for various ΔD-values. The threshold was defined to be that ΔD-value which yields 50% of answers *'better quality'* or *'equal quality'* compared to the reference image. This ΔD-value represents the display's depth of focus since the first experiment has shown that the image quality at $\Delta D = 0\,\text{D}$ is also assessed to be better or at least equal as the reference image (except for $(\alpha_v, \sigma) = (0.34\,\text{D}, 0.3)$). Figure 4.12 depicts the results of this second psychometric experiment for all three viewing angles α_v and the four coherence levels σ tested. Again, only seven out of eight subjects were considered.

The results indicate that the multiple imaging phase mask extends

(a)

(b)
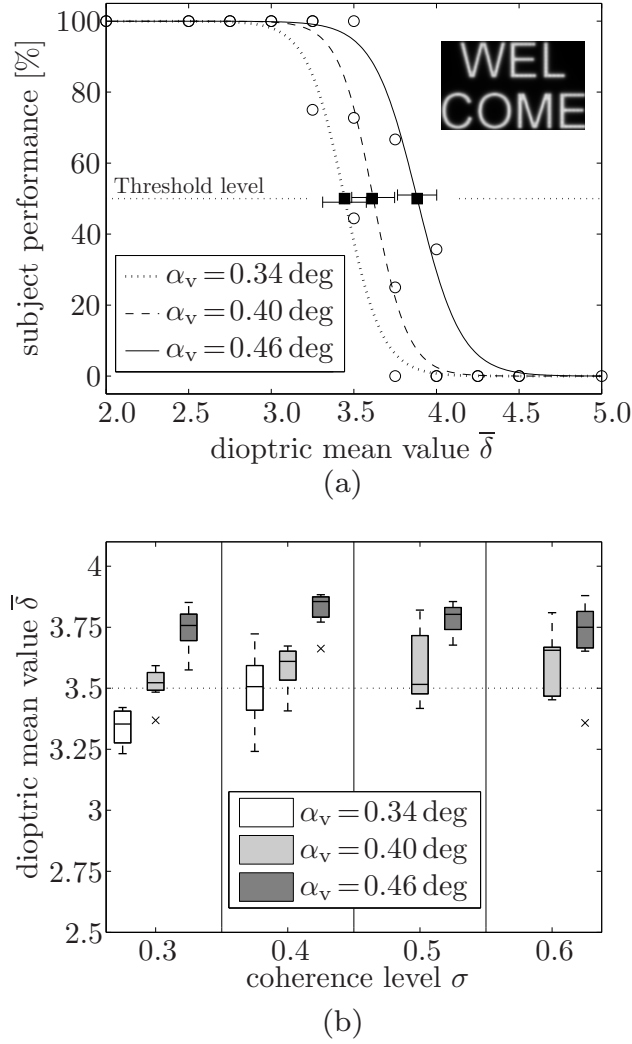
Figure 4.11: *(a) shows the psychometric functions of one subject for the first experiment for all three α_v-values and $\sigma = 0.4$. The ordinate shows the subject's performance as proportion of answers 'better or equal quality'. The measured performances are indicated by circles ($\circ$) while the squares ($\blacksquare$) signify the derived thresholds. The threshold error bars indicate the 95%-confidence interval. The inset shows the reference image for $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$. (b) shows the results of the first experiment based on 7 subjects (see text).*
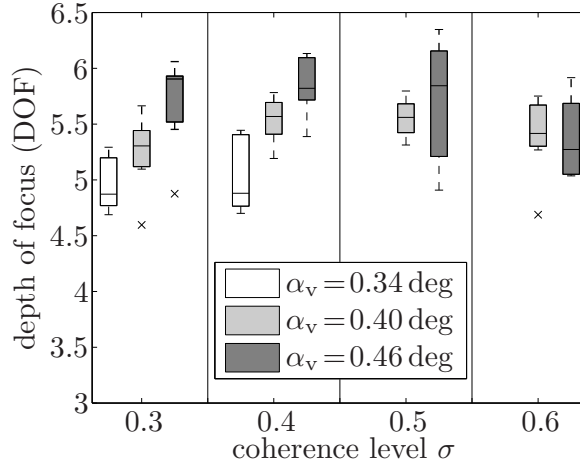
Figure 4.12: *Results of the second experiment (estimation of the DOF) based on 7 subjects.*

the depth of focus to about $5.5\,\mathrm{D}$ for $\alpha_\mathrm{v} \geq 0.4\,\mathrm{deg}$. The impact of the coherence level ranging from 0.3 to 0.6 is small as already expected from Fig. 4.10. For a smaller font size viewing angle, $\alpha_\mathrm{v} = 0.34\,\mathrm{deg}$, the DOF is limited to about $4.8\,\mathrm{D}$ for coherence levels of 0.3 and 0.4. As in the first experiment, no results could be obtained with this small α_v-value for $\sigma = 0.5/0.6$ since none of the query images was assessed to be better or equal than the reference image.

4.6. Conclusion

We have discussed the application of the multiple imaging technique for extending the depth of focus of retinal projection displays. Thereby, the retinal images were calculated by means of a schematic eye model and a simulation tool for partially coherent imaging. An appropriate phase-only mask has been designed and evaluated on the basis of objective quality criteria and psychometric measurements revealing the following results:

- The depth of focus can be extended substantially provided that partially coherent illumination light is used $(0.3 \lesssim \sigma \lesssim 0.6)$. For other coherence values the retinal images suffer from structural and contrast degradation effects, respectively.

- Psychometric measurements have revealed that a depth of focus of about 5.5D for font size viewing angles $\alpha_\mathrm{v} \geq 0.4\,\mathrm{deg}$ can be

achieved. Thus, the DOF can be extended by a factor of up to 3.2 compared to the corresponding Rayleigh criterion. This viewing angles correspond to common font sizes $F_s \geq 14\,\mathrm{pt}$ on a screen when viewed from a distance of $50\,\mathrm{cm}$.

- For smaller font size viewing angles ($\alpha_\mathrm{v} \leq 0.34\,\mathrm{deg}$), the used phase mask should be redesigned to guarantee a satisfactory image quality for the unaccommodated eye.

Although the analysis is based on an accurate and sophisticated eye model, it is not able to replace real measurements on a implemented display with the corresponding phase mask and with the real human eye. An implementation would also offer the opportunity to gain more experience about the psychological reaction of users when they are provided with an accommodation-insensitive virtual scene overlaid over the real view.

5

Oscillating fluid lens[*]

In this chapter, we discuss the use of a variable-focus lens in the retinal projection display. In contrast to the work by Sugihara et al. [45], the display's focus in our approach is not adapted in real-time to the current accommodation, but the focus is oscillating periodically. Thus, the user is viewing an incoherent superposition of in-focus and defocused images on his retina. As in the previous chapter, the retinal images are simulated and then evaluated using a combination of objective image quality criteria and psychometric measurements.

5.1. System consideration

5.1.1. Optical setup

The principal setup of the considered retinal projection display is identical to the one discussed in chapter 2, except that a variable-focus lens is integrated at the aperture stop plane P1 (see Fig. 5.1).
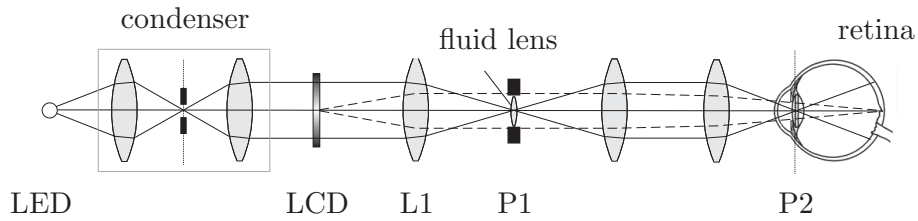
Figure 5.1: *Principal setup of the retinal projection display. It is almost identical to the one discussed in chapter 2. Additionally, a variable-focus lens at P1 is integrated.*

The variable-focus lens as used here has recently been proposed by Kuiper and Hendriks [86]. This miniature lens is based on two immiscible liquids with different refractive indices; one of them is electrically conductive, whereas the other is insulating. By electro-wetting, i.e. by applying a voltage, the curvature of the meniscus between the two liquids can be varied leading to a change in the dioptric power of the lens. The prototype device - as presented in [86] - can change its focal length from 5 cm to infinity in less than 10 ms. This corresponds to a dioptric power range of 20 D. The prototype lens has a diameter of 3 mm and a length of 2.2 mm, thus being well-suited for an integration into a retinal projection display. Durability tests showed that the lens has performed over 1 million focusing operations with no loss of optical performance [86].

5.1.2. Approach

The approach in our setup is now to oscillate the dioptric power of the fluid lens periodically in time (described by a periodic dioptric function $\mathcal{D}(t)$). Thus, the LCD image is projected sequentially at different foci onto the retina. In case that the characteristic frequency of $\mathcal{D}(t)$

lies above the critical fusion frequency (CFF)[1] (of about 60 Hz), the user does not perceive the individual images but a temporal fusion of several defocused and in-focus images. In optical terms, the temporally sequential projection of the images gives rise to an incoherent superposition of the individual images - regardless of the coherence level of the illumination light.

Apparently, the temporal fusion of defocused and in-focus images is expected to extend the display's depth of focus. But this will come at the cost of lowered contrast and sharpness compared to best-focused imaging. This weakness, however, can be reduced when using partially coherent illumination light rather than incoherent one since partially coherent imaging achieves higher contrast values than the corresponding incoherent imaging (cf. chapter 2). Thus, the following discussion will focus on the trade-off between extended DOF and reduced overall image quality when applying an oscillating fluid lens and partially coherent illumination.

5.1.3. Theoretical aspects

Since the individual images superpose incoherently to each other, the fused retinal irradiance at a point (u', v') can be expressed as follows:

$$
\begin{aligned}
I_{\text{fused}}(u', v') \quad &\propto \quad \frac{1}{\Theta} \int_0^\Theta I(u', v', \mathcal{D}(t)) dt \\
&\propto \quad \frac{1}{\Theta} \int_{\mathcal{D}(0)}^{\mathcal{D}(\Theta)} I(u', v', \mathcal{D}) \left| \frac{\partial \mathcal{D}(t)}{\partial t} \right|^{-1} d\mathcal{D} \quad (5.1)
\end{aligned}
$$

Here, $I(u', v', \mathcal{D}(t))$ defines the retinal irradiance at (u', v') of that individual image projected at time t when the fluid lens has a dioptric power $\mathcal{D}(t)$. (u', v') are geometrical coordinates on the retina. The integration limit Θ is the periodic time of $\mathcal{D}(t)$. Thus, the fused retinal distribution is an incoherent superposition of the individual images at all $\mathcal{D}$-values, weighted by the inverse derivation of the dioptric function $\mathcal{D}(t)$.

The retinal irradiance of the individual images $I(u', v', \mathcal{D}(t))$ can be calculated by means of the Hopkins theory of partially coherent imaging [53] (see appendix B). In this theory, the optical properties of the display as well as of the eye are described by their wavefront error

---

[1]The CFF is defined as the frequency at which all flicker of an intermittent light stimulus disappears.

maps $\psi_{disp}(x,y)$ and $\psi_{eye}(x,y)$ giving the respective phase error at the exit pupil point (x,y) due to defocus and aberrations. The wavefront error map $\psi_{disp}(x,y)$ is mainly determined by the fluid lens. Hence, $\psi_{disp}(x,y)$ depends on the lens' current dioptric power $\mathcal{D}(t)$ and can be expressed as follows [54] (see eq. 2.2):

$$\psi_{disp}(x,y,t) = \frac{\pi}{\lambda}\mathcal{D}(t)(x^2 + y^2) \tag{5.2}$$

Here, λ denotes the wavelength of the illumination light and (x,y) are geometrical coordinates in the exit pupil plane. The ocular wavefront map $\psi_{eye}(x,y)$ depends mainly on the accommodation of the eye and considers further ocular aberration effects.

5.2. Simulation background

For calculating the retinal images, the same simulation setup has been used as in the previous chapter for the multiple imaging technique. Again, the ocular wavefront error map $\psi_{eye}(x,y)$ is based on the accommodation-dependent eye model as discussed in appendix A. The partially coherent imaging simulations were carried out on the slightly adapted version of the SPLAT 5.0 software package as discussed in chapter 4.

The dioptric power function $\mathcal{D}(t)$ was assumed to be sinusoidal for any subsequent simulations:

$$\mathcal{D}(t) \propto \mathcal{D}_0 + \frac{\xi}{2} \cdot \sin\left(\frac{2\pi}{\Theta}t\right) \tag{5.3}$$

Here, $\mathcal{D}_0$ indicates the dioptric offset of the oscillating fluid lens while ξ is the dioptric amplitude of the lens' oscillation.

In the simulations, the retinal images $I(u',v',\mathcal{D}(t))$ are calculated at N discrete, equally spaced values $\mathcal{D}_j = \mathcal{D}(t_j)$. According to eq. 5.1 the fused retinal image $I_{\text{fused}}(u',v')$ can then be approximated as follows:

$$I_{\text{fused}}(u',v') \propto \sum_{j=1}^{N} I(u',v',\mathcal{D}_j)\left|\frac{\partial\mathcal{D}(t_j)}{\partial t}\right|^{-1}\Delta\mathcal{D} \tag{5.4}$$

where $\Delta\mathcal{D}$ is the spacing between the discrete $\mathcal{D}_j-$values. Note that the periodic time Θ has no influence on $I_{\text{fused}}(u,v)$ provided that $\Theta^{-1} > CFF \approx 60\,\text{Hz}$.

For the simulations, the system parameters according to Fig. 5.1 were defined in analogous manner as in chapter 4:

Illumination wavelength	$\lambda = 635\,\text{nm}$
Diameter of eye pupil	$\varnothing_{\text{pup}} = 3\,\text{mm}$ (bright-adapted)
Diameter of aperture stop Di_1	$\varnothing_1 = 2.1\,\text{mm}$
Magnification of relay lens system	$M = 0.65$

Note that the relations between the exit pupil diameter and the eye motion box are discussed in detail in section 4.3.4.

5.3. Evaluation

5.3.1. Fluid lens amplitude ξ

For characterising the image quality objectively, the same quality functions C and S were applied as discussed in eqs. 4.4 and 4.5. The most important parameter to be found is the amplitude ξ of the oscillating fluid lens (defined as in the sinusoidal dioptric function $\mathcal{D}(t)$ in eq. 5.3). The best ξ-value should optimise the trade-off between the depth of focus and the reduction of the overall image quality. For the first simulations, the oscillation offset $\mathcal{D}_0$ in eq. 5.3 was set to 3.5 D. This parameter will be used later to shift the results relative to the eye accommodation ΔD.

Figure 5.2 shows the contrast and structural quality functions C and S in terms of the dioptric amplitude ξ for various eye accommodation values ΔD. The data were calculated with partially coherent illumination of $\sigma = 0.3$ and for a viewing angle $\alpha_{\text{v}} = 0.4\,\text{deg}$ [2]. Results for other σ- or α_{v}-values are qualitatively similar. The results were obtained with $\Delta\mathcal{D} = 0.5\,\text{D}$ as the spacing between the discrete $\mathcal{D}_j-$values (eq. 5.4). The simulation will not become more accurate with smaller $\Delta\mathcal{D}$-values as corresponding simulations with $\Delta\mathcal{D} = 0.25\,\text{D}$ have shown.

Principally, the plots in Fig. 5.2 confirm the trade-off: On the one hand, the C- and S-values for eye accommodation values ΔD between 2D and 4D get lower and lower when the amplitude ξ of the oscillating fluid lens increases. On the other hand, the corresponding C- and S-results for ΔD between 0 D and 1 D and 5 D and 6 D, respectively, get improved. In more detail, however, the figures reveal some differences between the contrast and the structural quality: The structural qual-

---

[2] see section 4.4.1 for a detailed discussion, how the font sizes are defined in this work.

(a) contrast function C   (b) structural function S
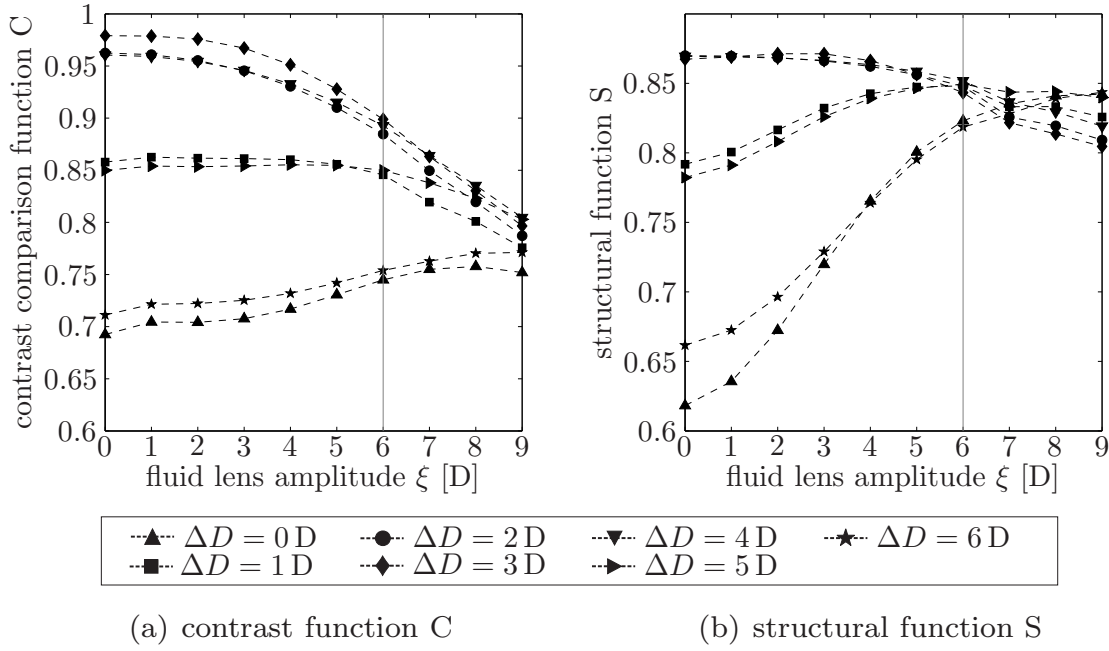
Figure 5.2: *Image quality functions C and S in terms of the dioptric amplitude ξ for various values of ΔD and with $\mathcal{D}_0 = 3.5\,\text{D}$. The data were calculated with $\sigma = 0.3$ and $\alpha_v = 0.4\,\text{deg}$. The vertical grey lines indicate the values at $\xi = 6\,\text{D}$ (see text)*

ity is substantially improved ($\Delta S \approx 0.22$) for $\Delta D = 0\,\text{D}$, whereas the structural degradation for $\Delta D = 3\,\text{D}$ is limited ($\Delta S \approx -0.03$). The situation for the contrast quality is less favourable: Here, the degradation for $\Delta D = 3\,\text{D}$ is much higher ($\Delta C \approx -0.18$) than the small contrast gain for $\Delta D = 0\,\text{D}$ ($\Delta C \approx 0.06$).

Ideally, ξ should be raised to about $9\,\text{D}$ where both functions become practically unaffected by any changes of ΔD ranging from $0\,\text{D}$ to $5\,\text{D}$. This, however, will not be possible, since this would result in a low-contrast image. In fact, the optimum ξ-value should be as small as possible to limit the contrast loss and large enough to profit from the gains in structural quality. To explore this trade-off, Fig. 5.3 shows the mean value of the structural function S over the considered accommodation range $[0\,\text{D}, 5\,\text{D}]$ in terms of the amplitude ξ. Additionally, the standard deviation of S over the same accommodation range is plotted. Both curves represent the coherence level $\sigma = 0.3$ and the viewing angle $\alpha_v = 0.4\,\text{deg}$. For other σ- and α_v-values qualitatively similar results are obtained.

The curves show that the mean value of the S-function is maximised for amplitude values ξ of about $5 - 6\,\text{D}$ while the standard deviation
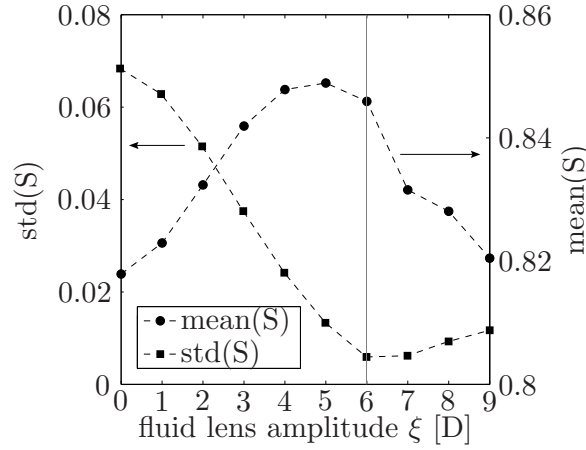
Figure 5.3: *The mean value as well as the standard deviation of the structural image quality S calculated over the accommodation range $[0\,\mathrm{D}, 5\,\mathrm{D}]$ in terms of ξ. Again, $\sigma = 0.3$ and $\alpha_{\mathrm{v}} = 0.4\,\mathrm{deg}$.*

is minimised for $\xi \approx 6 - 7\,\mathrm{D}$. Consequently, $\xi = 6\,\mathrm{D}$ is an optimum for this trade-off. The psychometric evaluation with subjects will show whether the reduction in contrast at $\xi = 6\,\mathrm{D}$ is acceptable.

5.3.2. Coherence level σ

While keeping $\xi = 6\,\mathrm{D}$ constant, the coherence level σ of the illumination light is now varied. Again, the quality of the temporally fused retinal images is evaluated by the contrast comparison function C and the structural function S. Figure 5.4 shows the corresponding results. The results indicate, that the contrast values increase continuously when the spatial coherence of the illumination light is raised (corresponding to lower σ-values). This is not surprising when considering again, that coherent imaging is known to provide higher contrast than the respective incoherent imaging. In case of the structural quality S, the best results are achieved for partially coherent light of about $\sigma \approx 0.5$. For lower σ-values increasingly dominant coherence effects degrade the quality (as discussed in more detail in chapter 3) . As a consequence, the psychometric evaluation in the subsequent section can be limited to coherence levels between 0 and 0.5 since at higher σ-values both functions C and S are less favourable.

trast function

(a) contrast function C

(b) structural function S

Figure 5.4: *Image quality functions C and S in terms of coherence level σ and of eye accommodation ΔD. The amplitude of the oscillation lens was set to $\xi = 6\,\mathrm{D}$ while $\mathcal{D}_0 = 3.5\,\mathrm{D}$. Again, $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$.*

5.4. Psychometric evaluation

Objective image quality criteria are valuable tools to characterise the quality. However, they cannot replace the direct (but also subjective) ass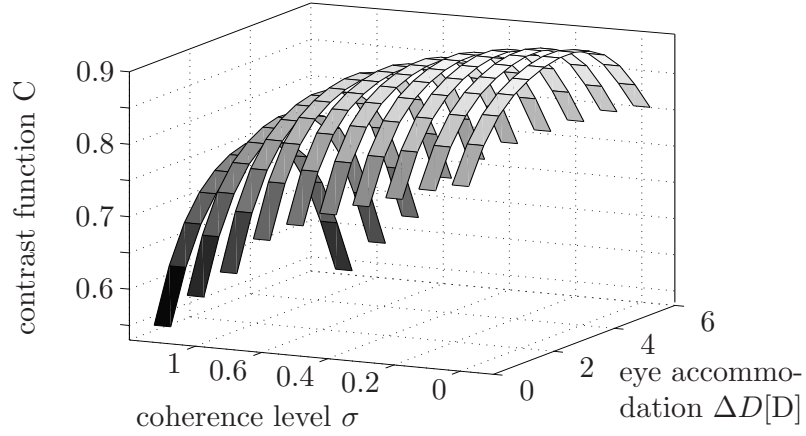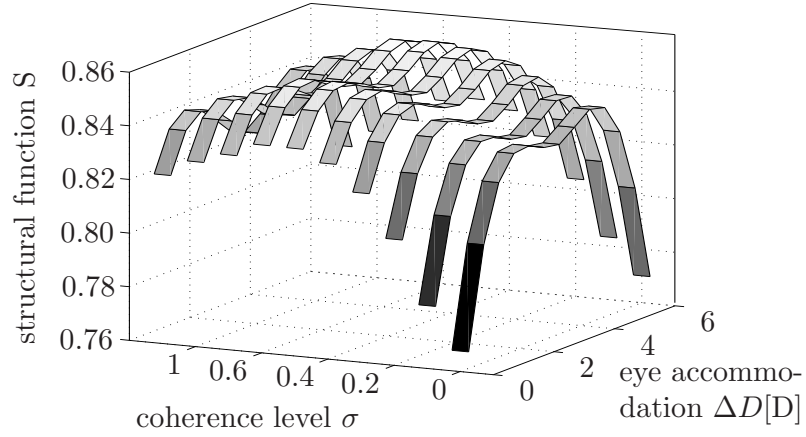essment of the image quality by subjects. Consequently, in addition to the objective quality criteria, psychometric measurements were carried out to rate the quality of the simulated retinal images relative to a given reference image.

5.4.1. Methods

To this end, eight subjects (mean age: 31 years, variance: 8 years) with normal or corrected-to-normal vision participated in the tests. We applied the same test procedure and reference image as in the previous chapter. The procedure is based on the 'two-alternative unforced-choice task' in combination with the adaptive procedure called PEST. Again, all eight subjects performed two different experiments: The first experiment had the purpose to evaluate the image quality at $\Delta D = 0\,\mathrm{D}$ by varying the oscillation offset $\mathcal{D}_0$. The second experiment was to measure the ΔD-range for which the query images are assessed to be of at least equal quality as the reference image. Both experiments included 12 separate test runs, each of them characterised by one of the 12 combinations $(\alpha_\mathrm{v}, \sigma)$ with font size viewing angles[3] $\alpha_\mathrm{v} = 0.4$ and $0.46\,\mathrm{deg}$ and coherence levels $\sigma = 0/0.1/0.2/0.3/0.4/0.5$.

5.4.2. First experiment: Image quality at $\Delta D = 0\,\mathrm{D}$

The quality at $\Delta D = 0\,\mathrm{D}$ can be controlled by adjusting the oscillation offset $\mathcal{D}_0$ (see eq. 5.3). Note that variations in $\mathcal{D}_0$ do not alter the shape of the C- and S-curves, but shift the curves relative to the eye accommodation ΔD, only. Thereby, lower $\mathcal{D}_0$-values improve the image quality at $\Delta D = 0\,\mathrm{D}$. Thus, first psychometric experiments were carried out to find an appropriate $\mathcal{D}_0$-value. To this end, the query image series used for this experiment consisted of the retinal images at $\Delta D = 0\,\mathrm{D}$ for various $\mathcal{D}_0$-values. The threshold was defined as that $\mathcal{D}_0$-value which yields 50% of answers *'better quality'* or *'equal quality'* compared to the reference image.

Figure 5.5 depicts the results of this first psychometric experiment for both font size viewing angles $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$ and $\alpha_\mathrm{v} = 0.46\,\mathrm{deg}$ and all

---

[3]Note that these viewing angles correspond to font sizes $F_s = 14$ and $16\,\mathrm{pt}$ on a screen when viewed from a distance of $50\,\mathrm{cm}$.
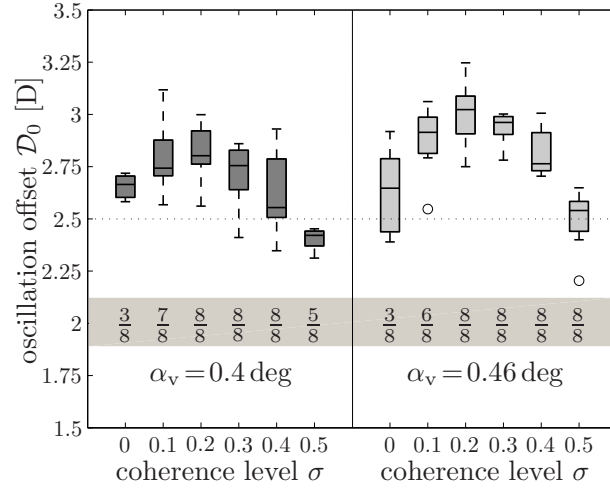
Figure 5.5: *Results of the first psychometric experiment based on eight subjects. The meaning of the fractions below the boxes are explained in the text.*

six coherence levels σ. The boxes show the lower quartile, median and upper quartile values of the threshold-results of the eight subjects. The dashed lines extending from each end of the boxes indicate the extent of the rest of the data within 1.5 units of interquartile range. The open circles represent outliers beyond this limit. The fractions below the boxes indicate the portion of the eight subjects whose results could be considered for the corresponding box. It shows that for extreme σ-values ($\sigma = 0/0.1/0.5$) some subjects rated the quality of all query images as worse than the reference image due to structural and contrast deficits.

The experimental results in Fig. 5.5 show that $\mathcal{D}_0$ should be set to 2.5 D. With this $\mathcal{D}_0$-value the subjects assessed the quality of the fused image at $\Delta D = 0$ D to be better (or at least as equal) than the one of the reference image. However, for $\sigma = 0.5$, the thresholds lie rather close or even below $\mathcal{D}_0 = 2.5$ D so that for this size the image quality at $\Delta D = 0$ D is critical, if $\mathcal{D}_0 = 2.5$ D. If a better quality for $\sigma = 0.5$ is desired, $\mathcal{D}_0$ might be lowered. However, such a focal shift reduces the DOF-values (as obtained in the second experiment) for all font sizes and coherence levels accordingly. For the second experiment, we set $\mathcal{D}_0 = 2.5$ D so that the dioptric power of the fluid lens oscillates sinusoidally between -0.5 D and 5.5 D (according to eq. 5.3).

5.4.3. Second experiment: Estimation of depth of focus

Analogous to the tests in the previous chapter, the second psychometric experiment had the purpose to measure the dioptric range for which the simulated retinal images are considered as being at least of equal quality as the reference image. In this experiment, the query image series included the retinal images for various ΔD-values. The threshold was defined to be that ΔD-value which yields 50% of answers *'better quality'* or *'equal quality'* compared to the reference image. This ΔD-value represents the display's depth of focus since the first experiment has shown that the image quality at $\Delta D = 0\,\mathrm{D}$ is also rated to be better or at least equal as the reference image (except for $\sigma = 0.5$). Figure 5.6 illustrates the results of this second psychometric experiment for both font size viewing angles and all six coherence levels. In case of full coherence ($\sigma = 0$) and $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$, none of the subjects judged any of the query images to be better so that no results are obtained.
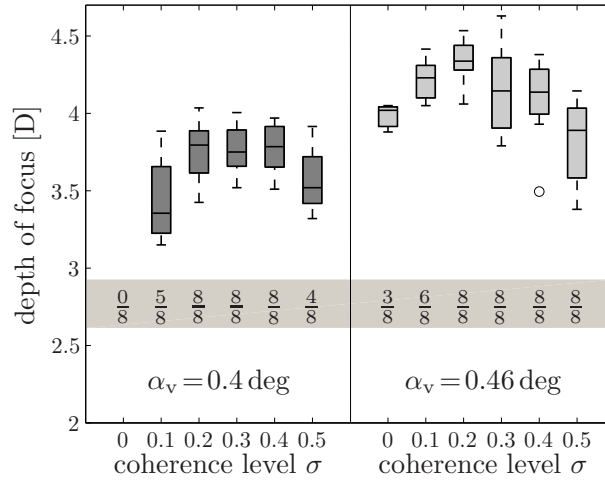
Figure 5.6: *Results of the second psychometric experiment based on eight subjects.*

   In principal, the results in Fig. 5.6 illustrate that an oscillating fluid lens has the potential to extend the depth of focus of a retinal projection display. However, due to contrast and structural deficits, the coherence level of the illumination light must be properly chosen within a range of $0.2 \lesssim \sigma \lesssim 0.4$. In this case, the DOF can be extended up to $3.8\,\mathrm{D}$ for $\alpha_\mathrm{v} = 0.4\,\mathrm{deg}$ and even up to $4.2\,\mathrm{D}$ for $\alpha_\mathrm{v} = 0.46\,\mathrm{deg}$. This is an extension by a factor of 2.2 compared to the corresponding Rayleigh

criterion.

Similar measurements with smaller font viewing angles (e.g. $\alpha_{\mathrm{v}} = 0.34\,\mathrm{deg}$) revealed that then the oscillation offset $\mathcal{D}_0$ must be lowered further to guarantee a satisfactory image quality at $\Delta D = 0\,\mathrm{D}$. However, this would also reduce the DOF for larger font sizes. In general, the experiments for $\alpha_{\mathrm{v}} = 0.34\,\mathrm{deg}$ showed that the use of an oscillating lens is not ideal for font viewing angles smaller than $0.4\,\mathrm{deg}$ since contrast and structural deficits are more dominant in those cases.

5.5. Retinal images

Finally, for illustrating the experiments, Fig. 5.7 shows the retinal images for three coherence levels σ and for three accommodation values ΔD when applying the oscillating lens with $\xi = 6\,\mathrm{D}$ and $\mathcal{D}_0 = 2.5\,\mathrm{D}$. For comparison, the first row contains the corresponding images with a fix lens (i.e. with $\xi = 0\,\mathrm{D}$ and $\mathcal{D}_0 = 2.5\,\mathrm{D}$).
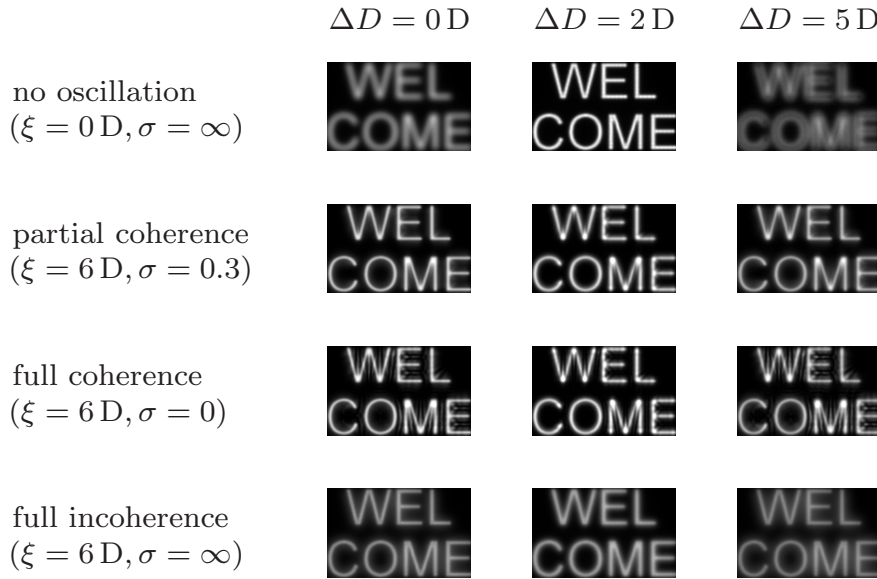
Figure 5.7: *First row: retinal images without lens oscillation ($\xi = 0\,D$) and with incoherent illumination for comparison. The other three rows show the retinal images for three coherence levels $\sigma = 0/0.3/\infty$ when the fluid lens is oscillating (with $\xi = 6\,D$). Again, $\alpha_{\mathrm{v}} = 0.4\,deg$ and $\mathcal{D}_0 = 2.5\,D$.*

Figure 5.7 confirms the results obtained by both the index-based and the psychometric evaluation: The oscillating fluid lens can extend

the depth of focus, but only for partially coherent light. For fully coherent light, the contrast is improved, but degradations due to coherence effects are dominant. For fully incoherent light, the contrast is substantially reduced due to the fusion of several defocused images. Additionally, Fig. 5.7 illustrates that the extended DOF for $\sigma = 0.3$ is achieved at the cost of a reduction in contrast for ΔD-values around 2 D. However, the reduced contrast is still higher than the corresponding incoherent counterpart. The reason is that partially coherent imaging achieves originally higher contrast values for in-focus imaging than the respective incoherent imaging.

5.6. Conclusion

We have discussed the benefits of an oscillating fluid lens in a retinal projection display for extending the DOF. Since the fluid lens' dioptric power oscillates periodically in time above the critical fusion frequency of the human eye, the image - perceived by the user - is an incoherent superposition of in-focus and defocused images, and thus less sensitive to changes in the eye accommodation. On the basis of an accommodation-dependent eye model, partially coherent simulations of a simple text sample have been performed. The evaluation of the retinal images by objective quality criteria as well as psychometric tests revealed the following results:

- The depth of focus can be extended provided that partially coherent illumination light of $0.2 \lesssim \sigma \lesssim 0.4$ is used. For other coherence values, however, structural and contrast degradation effects are dominant.

- Optimised results are obtained when the fluid lens oscillates sinusoidally between $-0.5\,\mathrm{D}$ and $5.5\,\mathrm{D}$. Higher oscillation amplitude values would reduce the overall image contrast below a satisfactory level.

- Psychometric measurements have shown that a depth of focus of about $3.8\,\mathrm{D}$ for font size viewing angles α_v of $0.4\,\mathrm{deg}$ or larger can be achieved. Thus, the oscillating fluid lens extends the DOF by a factor of about 2.2 compared to the corresponding Rayleigh criterion.

- For smaller font sizes, the parameters for the lens' oscillation (either the oscillation amplitude ξ or the offset $\mathcal{D}_0$) must be lowered

to guarantee a satisfactory image quality for the unaccommo-
dated eye. This, however, reduces the DOF for all font sizes.

6

Spectacle-Based Design[*]

For a really wearable use, compactness and weight are essential factors. Thus, based on the previous studies further effort has been made to design an lightweight and compact system, but without loosing the improved defocusing properties. In this section we focus on design specification and conception and propose a compact and lightweight design for an unobtrusive integration into normal spectacles. Furthermore, eye safety in case of direct retinal projection is discussed.

6.1. System Specifications

For the system design, we focus first on the system setup as illustrated in Fig. 2.1, i.e. on a system without a phase mask or an oscillating fluid lens as incorporated in the chapters 4 and 5. In a second step, the integration of the phase mask (as discussed in chapter 4) or of the oscillating fluid lens (chapter 5) and its consequences on the design are briefly discussed.

To prove the concept in the bench experiments (see chapter 3) a rather limited resolution of 320 x 240 pixels (with a 15 μm-pixel pitch) was adequate. For a convenient use in see-through applications it will be necessary to have at least a VGA-resolution of 640 x 480 pixels. Thus, the miniature display selected is a 9.6-mm backlighting monochrome AMLCD (active-matrix LCD) with VGA-resolution and a 12 μm pixel pitch (provided by 'Kopin's CyberDisplay 640') [16].

Furthermore, the display should have a depth of focus of 5 D or more to provide accommodation-insensitive imaging. The experimental part has shown that this can be achieved for font size viewing angles $\alpha_v \geq$ 0.4 deg. Thus, the design should be aimed for displaying text of this size properly. This requires a minimal retinal resolution of ≈ 12.5 cyc/deg.

Table 6.1: *Optical design specifications*

	Parameter	Specification
1.	retinal resolution @ LCD 12 μm-pixel pitch	12.5 cyc/deg
2.	cut-off frequency ν^{cut}	12.5 cyc/deg
3.	eye relief d_{eye}	25 mm
4.	adaption to eye accommodation	2.5 D
5.	coherence level of the light σ	≈ 0.5

It is the goal to make the wearable display as unobtrusive as possible. This can be best achieved by an integration of the optical system into spectacles. Thus, the eye relief - that is the distance from the last surface lens to the eye's pupil - should be set to $d_{\text{eye}} = 25$ mm to allow for common types of eyeglasses [88].

Finally, to take best benefit from the large DOF of more than $5\,\mathrm{D}$, the display should be adapted to an eye accommodation which lies in the middle of this DOF-range. That means, that the display should provide best focusing at an eye accommodation of $\approx 2.5\,\mathrm{D}$ what corresponds to a viewing distance of about $40\,\mathrm{cm}$. All these system specifications are summarised in Tab. 6.1.

6.2. Design Analysis

Within the scope of these specifications, the display system can now be designed. Figure 6.1 illustrates again the setup schematically. To reduce complexity the relay system, which was formed by several lenses in Fig. 2.1, is considered here to consist of one single lens with focal length f_2, only. A split of the lenses L1 and L2, respectively, into multiple lens system can be made in a later step without lost of generality, since it does not affect the remaining system. Furthermore note, that the condenser, which provides an homogeneous illumination of desired spatial coherence level $\sigma = 0.5$, can be designed completely independently. Therefore, the condenser part is depicted in Fig. 6.1 as a black box.
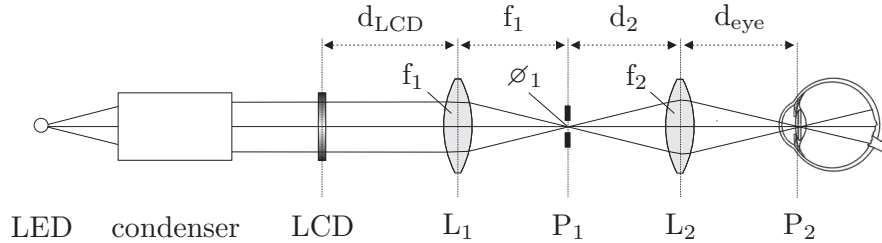
Figure 6.1: *Five parameters (f_1,f_2,d_{LCD},d_2 and $\varnothing_1$) are free to choose in the design process*

As shown in Fig. 6.1, there are principally five parameters which can be varied: the diameter of the aperture P1 (signified by $\varnothing_1$), the focal lengths of the lenses L1 and L2 (indicated as f_1 and f_2, respectively) as well as the separations d_{LCD} and d_2. The distance from lens L1 to the aperture P1 corresponds to the focal length f_1 as the aperture must lie in the focal plane of lens L1.

However, the specifications defined in Tab. 6.1 reduce the degrees of freedom: First, the focal lengths f_1 and f_2 are connected to each other

in order to meet the first requirement of a specific retinal resolution for the given pixel pitch of $12\,\mu$m. Secondly, the diameter $\varnothing_1$ is connected to the focal length f_1 by the requirement that the cut-off frequency is given at $\nu^{cut} = 12.5\,$cyc/deg (specification 2). Finally, the separations d_{LCD} and d_2 are linked with the focal lengths f_1, f_2 by the general and well-known lens formula (see eq. 6.1) in order to fulfil the obligations 3 and 4.

$$\frac{1}{g} + \frac{1}{b} = \frac{1}{f} \qquad \text{with} \qquad \begin{cases} \text{g=object distance} \\ \text{b=image distance} \\ \text{f=focal length} \end{cases} \qquad (6.1)$$

Again, the specification 5 in Tab. 6.1 can be treated separately when defining the condenser.

That means that considering all specifications the system keeps one degree of freedom, represented - for instance - by the focal length f_2. As soon as f_2 is fixed, all other parameters are given by meeting the specifications above. The goal of this design step is now to find an optimum value for f_2 which minimises the system length L, defined by the sum of d_{LCD}, f_1 and d_2, but considering practicability limits - especially with regard to the focal lengths f_1 and f_2.

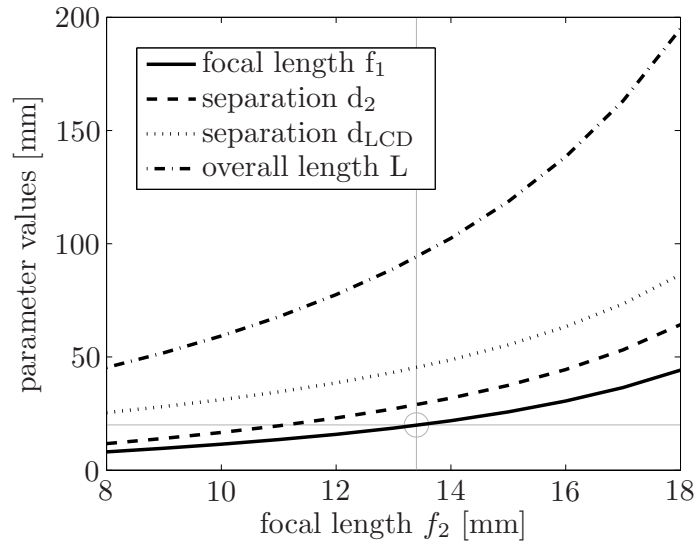**Figure 6.2:** *Parameter values in terms of the focal length f_2 considering the system specifications in Tab. 6.1.*

The results of this study are illustrated in Fig. 6.2 showing the values for the parameters f_1, d_{LCD}, d_2 and finally for the overall length L in terms of the focal length f_2. It turns out that the system becomes more compact when reducing the focal lengths f_1 and f_2, accordingly. However, smaller focus lengths cause normally larger aberration effects, if keeping the lens aperture constant. The aperture of lens L1, on its part, must be kept constant, as it has to be at least of the same size as the diagonal dimension of the display device[1]. Consequently, smaller focal lengths require more sophisticated lens systems L1 and L2 to control the growing aberration effects. More complex, multiple lens systems, however, reduce the gain in compactness and increase weight and cost.

Consequently, compactness and low aberration effects are two contradictory, but desirable requirements. It is now the challenge to find an optimum balance between these two goals. Concerning the aberrations a good guideline is the so called f-number which is defined as the ratio of the lens' focal length to the diameter of the lens' aperture [88]. As a rule of thumb, the f-number should not fall below ≈ 2.0 since otherwise the design complexity jumps up for getting high quality lenses [89]. This implies a value of $20\,\mathrm{mm}$ for the focal length f_1 since the aperture of lens L1 must be larger than the LCD's diagonal dimension of $9.6\,\mathrm{mm}$. Thus, all other parameters are fixed according to Fig. 6.2: The focal length f_2 is $13.5\,\mathrm{mm}$, the LCD separation $d_{LCD} = 45\,\mathrm{mm}$ and the overall system length L (from the LCD up to lens L2) turns out to be about $94\,\mathrm{mm}$.

6.3. Additional aperture lens

The LCD separation d_{LCD} (and thus also the system length L) can further be reduced by incorporating an additional lens at the aperture plane P1. This aperture lens is less sensitive to aberration effects than the lenses L1 and L2, since its clear aperture $\varnothing_1$ is much smaller. In addition, this lens can easily be combined with the phase-only mask as discussed in chapter 4 or the oscillating fluid lens as in chapter 5. Figure 6.3 illustrates the reduced values for the overall system length L for various focal values f_{ap} of the additional aperture lens. All other parameters (i.e. f_1,d_2 and d_{eye}) stay unchanged. Note that the combination of small f_{ap}-values (e.g. $f_{ap} = 10\,\mathrm{mm}$) and large f_2-values (e.g. $f_{ap} = 16\,\mathrm{mm}$) is not possible as then d_{LCD} would become negative.

---

[1]for the used Kopin display: $9.6\,\mathrm{mm}$
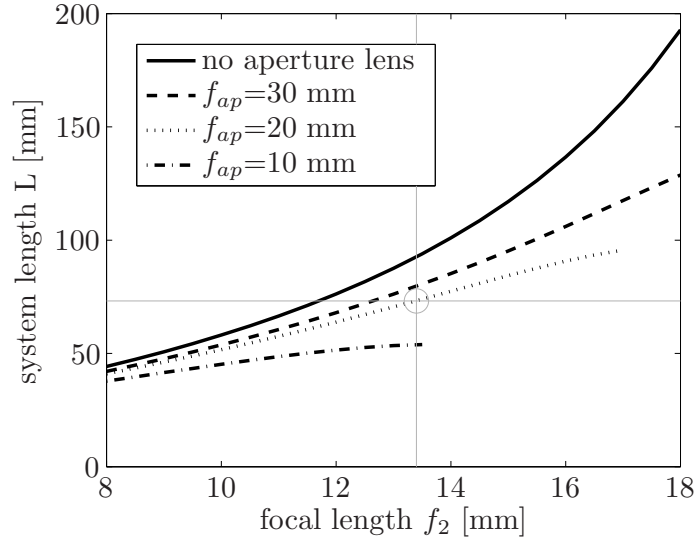
Figure 6.3: *Reduction of the overall system length L by integrating an additional aperture lens at plane P1 with focal length f_{ap}.*

A realistic value for the focal length f_{ap} is about 20 mm. Including such an aperture lens, the LCD separation d_{LCD} can be reduced down to about 25 mm, leading to a shorter system length $L \approx 74$ mm.

6.4. Proposed Design

Based on these considerations and the results in the Figs. 6.2 and 6.3, the design as illustrated in Fig. 6.4 can be proposed. It is one possible way for an unobtrusive integration of the display into normal eyeglasses. Other configurations are also feasible.

To control the aberration effects the lens L1 is proposed to be split into a so-called Petzval lens system consisting of two achromatic doublet lenses. This lens type is known to show significantly less aberrations than any singlet lenses for the same f-number. A more sophisticated, so-called 'double Gauss lens system' would perform even better, leading, however, to more complexity and weight as this system consists of more single lenses than the Petzval design [89].

The aperture lens at plane P1 can - as already mentioned above - easily be combined with the phase-only mask or the fluid lens. The integration of those elements does not affect the optical setup further.

The lens L2 is formed as a diffractive optical element (DOE or holographic lens). The use of DOEs at this position has some significant advantages over conventional refractive optics: First, they form

① condenser $\sigma=0.5$
② LCD
③ Petzval lenses (L1)
④ mirror
⑤ aperture lens
(optionally with
phase mask or fluid lens)
⑥ holographic lens
and combiner (L2)
⑦ eye
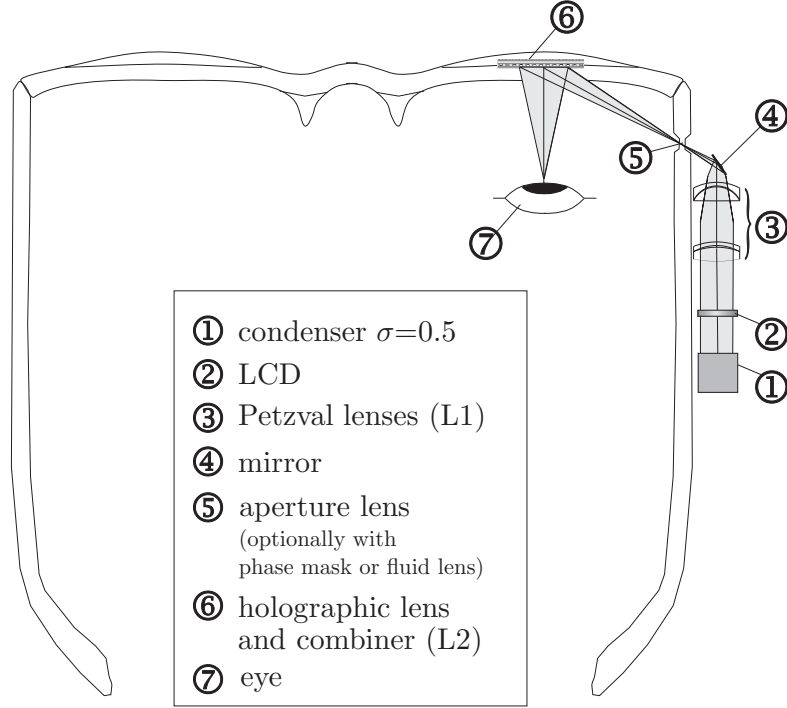
Figure 6.4: *Possible integration of the novel display system into normal eyeglasses.*

large-aperture and lightweight optical elements which can be integrated easily into an eyeglass. Second, by adopting the first diffraction order the DOE can simultaneously bend and converge the display rays at the eye (as illustrated in Fig. 6.4). In addition, DOEs provide high flexibility on aberration correction, even at high optical power, and so reducing significantly system complexity and cost. Finally, DOE can also act as optical combiner for the see-through display as DOE can accomplish to diffract only the specific wavelength rays while be completely transparent for all other visible rays [33, 90, 91]. All other parameters of

Table 6.2: *Parameter values in the proposed configuration*

$f_1 = 20\,\mathrm{mm}$	$f_2 = 13.5\,\mathrm{mm}$	$f_{ap} = 20\,\mathrm{mm}$
$d_{\mathrm{LCD}} = 25\,\mathrm{mm}$	$d_1 = 20\,\mathrm{mm}$	$d_2 = 29\,\mathrm{mm}$
$\varnothing_1 = 1.6\,\mathrm{mm}$		

the design are defined by meeting the requirements $1 - 4$, as discussed above. The values are listed in Tab. 6.2. Finally, Tab. 6.3 summarises the resulting properties of this new accommodation-free see-through display.

Table 6.3: *System properties*

Parameter	Specification
Display type	Maxwellian-view retinal projection display
Display mode	see-through
Field of view (FOV)	31.2 deg diagonally[a] 25.2 deg x 19.0 deg
Resolution	640 x 480 pixels (VGA-resolution)
Colour	monochrome red ($\lambda = 626\,\mathrm{nm}$)
Depth of focus (DOF)	> 5 diopters [b] [c] corresponds to viewing distances from ∞ to 20 cm
Number of refractive lenses	2 doublets
Number of diffractive optical elements (DOE)	1 holographic lens

[a]corresponds to a 11"-screen viewed from 50 cm
[b]for font size viewing angles $\alpha_\mathrm{v} \geq 0.4\,\mathrm{deg}$
[c]provided that a multiple imaging phase-mask is applied

6.5. Laser Safety Considerations

The eye - and especially the retina - is the organ most susceptible to damage by exposure to high radiation. Therefore, it is the goal of this section to discuss briefly the retinal exposure limits for a partially coherent LCD-based retinal projection display. The biological effects induced by optical radiation are essentially the same for both spatially coherent and incoherent sources. However, if fully coherent laser radiation is imaged onto the retina, the diameter of the irradiated retinal spot can be as small as $10 - 20\,\mu$m producing high retinal irradiances. Consequently, laser radiation must be treated as the most severe case because few conventional optical sources can produce the irradiances achieved by e.g. lasers. The retinal projection display with partially coherent illumination is a mixture between the coherent and incoherent case. Hence, it is a safe way to primarily consider the stringent standards for coherent radiation. Internationally accepted standards for the safe use of lasers are e.g. the "guidelines on limits of exposure to laser radiation" presented by the International Commission on Non-Ionising Radiation Protection (ICNIRP) in 1996 and revised in 2000 [92, 93]. Additionally, the limits as provided by the same commission for exposure to broad-band incoherent optical radiation will be discussed [94]. The exposure limit results are presented in a normalised form, i.e. in units of W/deg^2, a quantity which is independent of image size or field of view, respectively.

6.5.1. Mechanism of interaction with biological tissue

Generally, radiation affects that kind of tissue, which absorbs the radiation. The biological effects are the result of one or more competing biophysical interaction mechanisms - thermal, acoustic, optical and photochemical - which vary depending upon spectral region and exposure duration (see Fig. 6.5) [93, 95]. In the 400 - 1400 nm band (visible (VIS) and near-infrared (IRA) light), thermal injury to the retina is the principal effect for exposure durations less than 10 s. When the temperature of the pigment epithelium is increased above a critical temperature, proteins are denaturated. Photochemical injury predominates in the UV spectral region (at wavelengths of less than 400 nm) and is also the principal type of injury resulting from lengthy exposure (10 s and more) to shorter wavelength visible radiation. Photochemical damage can occur, when photon energies are sufficiently high to cause direct damage to macromolecules of retinal cells. It is typical for photochem-

ical damage that the degree of damage depends on the time-integrated radiant exposure (measured in J/m^2, i.e. the accumulated energy over time), so that a short duration exposure (seconds) with high irradiance has the same effect as a long term exposure (hours) with a correspondingly smaller irradiance. Finally, thermo-acoustic injury occurs at short pulse durations less than about 0.1 ms while optical breakdown and plasma formation become important for sub-nanosecond exposures.
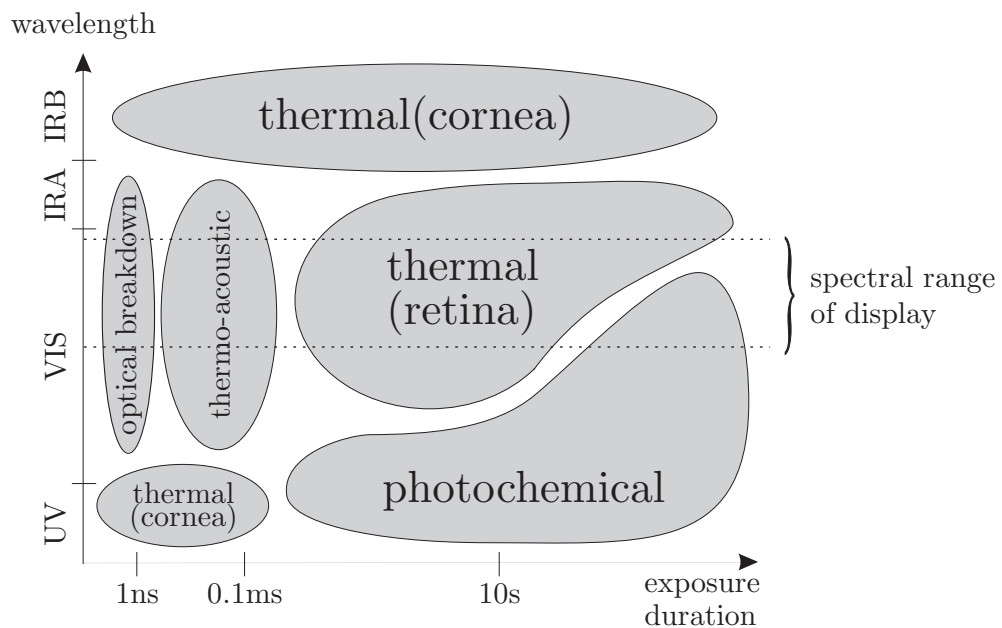
Figure 6.5: *The predominant biological effect for eye exposure to laser light of different spectral regions and exposure durations [92]*

Due to the transparency of the ocular media and - in particular - the inherent focusing properties of the eye, the retina is much more susceptible to damage by radiation in the visible spectral region than any other part of the eye. Therefore, the primary effect on the eye of visible radiation is damage to the retina. For a retinal projection display working in the visible region and with longer exposure durations (several hours), retinal thermal injuries and photochemical hazards are the predominant mechanism.

In the previous version of the ICNIRP guidelines [92] only one 'combined' exposure limit existed considering both hazard mechanism. To reflect the two damaging mechanism more accurately, the revised guidelines for coherent radiation [93] as well as the corresponding guidelines

for broadband incoherent radiation [94] provide now dual limits for exposure duration of longer than 10 s for photochemical and thermal injuries. Both limits must not be exceeded.

6.5.2. Exposure limits for coherent radiation

For wavelengths between 400 and 1400 nm the exposure limit (EL) for coherent radiation depends upon the viewing condition. The ICNIRP standards describe three different types of ocular exposures:

- **Intrabeam viewing:** the eye is positioned inside the collimated beam giving a nearly diffraction-limited image on the retina.

- **Extended source ocular exposure:** Extended source conditions apply to sources that subtend a visual angle at the eye greater than a minimum visual angle $\hat{\alpha}_{min} = 1.5$ mrad.

- **Pulsed exposures:** the light source is repetitively turned on and off.

The specific case of a LCD-based retinal projection display falls into the second category of extended sources with a continuous exposure duration of up to $T = 8\,\text{h} = 3 \cdot 10^4\,\text{s}$. The exposure limits given in the guidelines are always specified at a plane tangential to the cornea, i.e. the focusing properties of the eye are accounted for in the derivation of the limits. Moreover, all EL-values are averaged over an eye pupil size of 7 mm in diameter. The size of this averaging aperture corresponds to the maximum diameter of a dark-adapted pupil. Surely, this value is unrealistically large for lengthy viewing durations, but considers the worst case, leading so to a still greater safety factor. Table 6.4 shows the realistic assumptions that were made for the calculation of the exposure limit of the retinal projection display.

Thermal limit
The thermal limits for extended sources depend on the exposure duration T as well as on the apparent angular subtense $\hat{\alpha}$ of the source. For larger images ($\hat{\alpha} > 100$ mrad) the thermal EL decreases with increasing exposure duration up to the time $T_2 = 100$ s, and keeps constant for exposure durations larger than T_2. The time T_2 reflects the influence of eye movements which distribute the laser energy over a larger area on the retina and thus compensate for a possible temperature increase due to prolonged exposure. The apparent angular subtense of the source is

Table 6.4: *Assumptions for exposure limit calculation*

Parameter	Assumption
Wavelength range	$\lambda = 400 - 650\,\text{nm}$ (visible range)
Exposure duration	$T = 100s - 8\,\text{h} = 100\,\text{s} - 3 \cdot 10^4\,\text{s}$
Aperture size	$A_{pup} = 0.385\,\text{cm}^2$ corresponds to a pupil diameter of 7 mm according to the guidelines
Angular subtense of source	$\hat{\alpha} > 110\,\text{mrad}$ corresponds to a field of view (FOV) larger than $6.3 \times 6.3\,\text{deg}$

directly related to the diameter of the irradiated area on the retina. For a given power, an increase in image size reduces the retinal exposure per area and consequently decreases the injury potential. Therefore if the EL-values are specified as irradiance at the cornea, the EL for retinal thermal injury would be expected to vary with the angular subtense $\hat{\alpha}$ of the source. However, for extended sources smaller than 100 mrad this relation is reduced to a linear dependence on $\hat{\alpha}$, as cooling to surrounding non-irradiated tissue is much more effective for small images than for large images. For $\hat{\alpha} > 100\,\text{mrad}$, the dependence of the heat flow on the image size does not longer apply and the limit can be expressed as a constant radiance limit:

$$EL_{th} = \begin{cases} 15.3 \cdot T^{-0.25}\,\text{W/cm}^2\text{sr} & : \quad T < T_2 = 100\,\text{s} \\ 15.3 \cdot T_2^{-0.25}\,\text{W/cm}^2\text{sr} & : \quad T > T_2 = 100\,\text{s} \end{cases} \tag{6.2}$$

Photochemical limit

In contrast to the thermal limits, the photochemical limits do strongly depend on wavelength which is described by a correction factor C_B in the visible range:

$$C_B = \begin{cases} 1 & : \quad 400\,\text{nm} < \lambda < 450\,\text{nm} \\ 10^{0.02(\lambda-450)} & : \quad 450\,\text{nm} < \lambda < 700\,\text{nm} \end{cases} \tag{6.3}$$

Due to the additive nature of photochemical interaction, the photochemical limit can be expressed as a time-integrated radiance of $100 \cdot C_B \, \mathrm{J/cm^2 sr}$ for exposure durations up to $10^4 \, \mathrm{s}$. For longer exposure durations the limit is expressed as $10 \cdot C_B \, \mathrm{mW/cm^2 sr}$ as the law of additivity holds only for times of up to $10^4 \, \mathrm{s}$. Generally, this radiance limits are defined for a limiting cone angle $\hat{\gamma}$ which increases with exposure duration T up to a maximum value of $110 \, \mathrm{mrad}$ in order to consider the larger eye movements. However, for source sizes larger than $110 \, \mathrm{mrad}$ the limiting cone angle has no direct impact [96, 97].

Figure 6.6 shows the normalised exposure limits (in $\mathrm{mW/deg^2}$) for both hazard mechanisms for a retinal projection display (according to the ICNIRP guidelines 2000 [93]). Photochemical hazards limit the exposure in the shorter visible wavelength range as well as for longer exposure durations while the thermal limit is more stringent for the red-end visible spectrum. For a monochrome display at $\lambda = 635 \, \mathrm{nm}$ with a field of view (FOV) of $20 \times 15 \, \mathrm{deg^2}$ and a long exposure duration $T = 8 \, \mathrm{h}$ the exposure power should not exceed $170 \, \mathrm{mW}$ (thermal limit). For a similar display with green monochrome illumination light at $\lambda = 532 \, \mathrm{nm}$ the exposure limit is already substantially reduced to $14 \, \mathrm{mW}$ due to photochemical hazards.

6.5.3. Exposure limits for broad-band incoherent radiation

The limits for coherent radiation in the previous section are assumed to be more stringent than those for incoherent radiation. Nevertheless, the corresponding exposure limits for broad-band incoherent radiation (according to Ref. [94]) should be discussed in this section for comparison. As above the guidelines provide two limits, one for thermal hazards, the other for photochemical hazards. In contrast to monochrome laser radiation, the exposure limit here is not given for a specific wavelength, but the weighted sum of the spectral radiance L_λ of the source must not exceed the provided limit EL:

$$\sum_\lambda L_\lambda \cdot R(\lambda) \cdot \Delta\lambda < EL \tag{6.4}$$

where $R(\lambda)$ is the corresponding retinal hazard function which is listed for thermal hazards as well as for photochemical hazards in the guidelines. The EL for $T > 10 \, \mathrm{s}$ and $\hat{\alpha} > 100 \, \mathrm{mrad}$ for both damaging
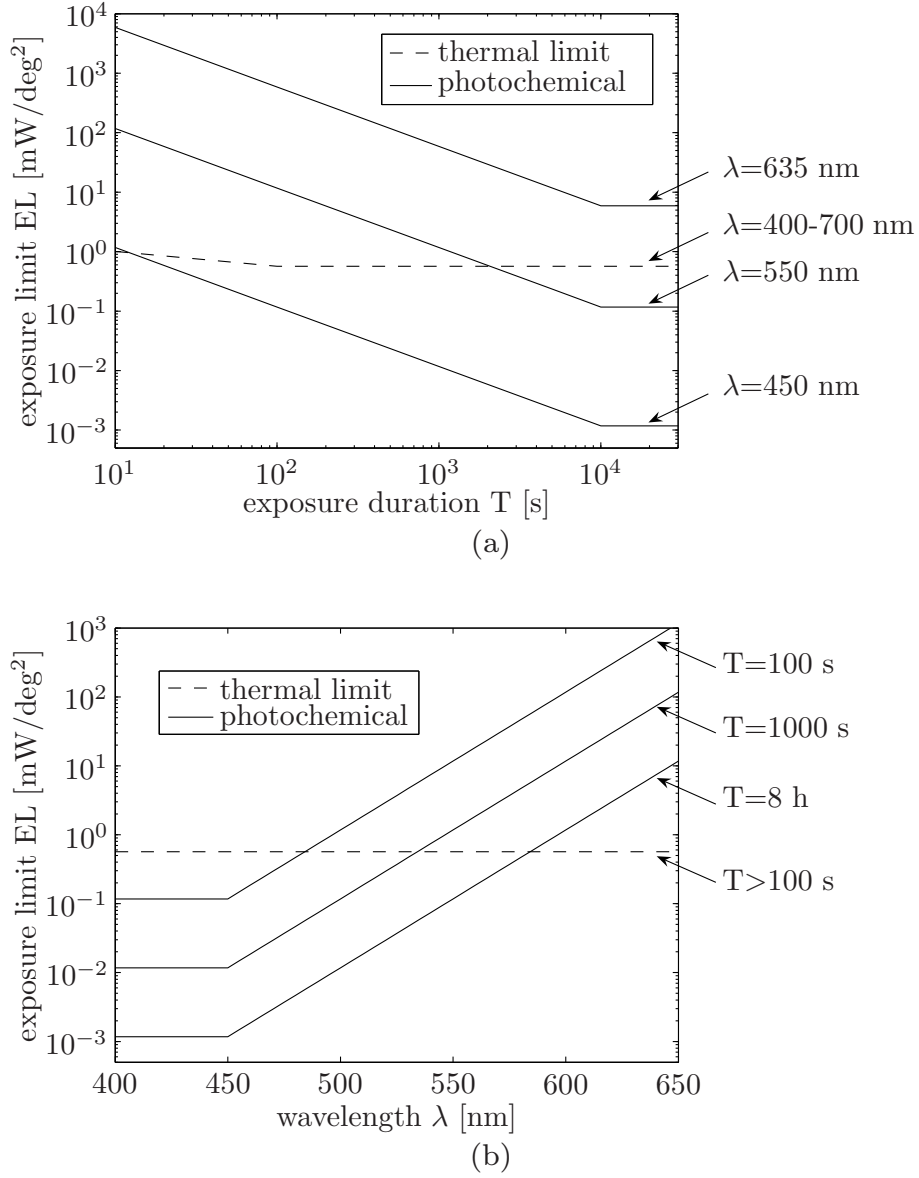
=550 nm

(a)

(b)

Figure 6.6: *The normalised exposure limits (in units of mW/deg^2) for coherent radiation as a function of a) the exposure duration T and b) the wavelength λ for a retinal projection display with a FOV larger than 6.3×6.3 degrees. Both limits must not be exceeded.*

mechanism are expressed as follows:

$$EL_{th} \quad = \quad 28.1\,\mathrm{W/cm^2 sr} \tag{6.5}$$

$$EL_{ph} \quad = \quad \begin{cases} 100 \cdot T^{-1}\,\mathrm{W/cm^2 sr} & : \quad 10\,\mathrm{s} < T < 10^4\,\mathrm{s} \\ 10\,\mathrm{mW/cm^2 sr} & : \quad T > 10^4\,\mathrm{s} \end{cases} \tag{6.6}$$

Similar to the laser guidelines, the averaging aperture diameter is 7 mm considering the worst case of a dark-adapted eye pupil. The limiting cone angle $\hat{\gamma}$ for the photochemical limit is defined also in the same manner as above.

In the specific case of the retinal projection display the used LED is still rather narrow-band around $\lambda = 635\,\mathrm{nm}$ so that the hazard functions $R(\lambda)$ can be replaced by their respective values at $\lambda = 635\,\mathrm{nm}$. Based on that, the normalised thermal limit for prolonged exposure to broad-band radiation turns out to be $3.3\,\mathrm{mW/deg^2}$ (leading to 985 mW for a display with a FOV of $15 \times 20\,\mathrm{deg^2}$). The corresponding photo-chemical limit is $1.2\,\mathrm{mW/deg^2}$ (and 350.4 mW for a FOV of $15 \times 20\,\mathrm{deg^2}$, respectively).

As assumed, these exposure limits are less stringent than the corresponding values for coherent radiation although the photochemical value is smaller than for coherent radiation. This is caused by the fact, that the photochemical hazard function $R(\lambda)$ is kept constant for all wavelengths larger than 600 nm while the wavelength dependence for laser radiation is described by an exponential function.

6.5.4. Required power in bright ambient lighting condition

For comparison, this subsection will shortly estimate the exposure power needed to provide a retinal projection perceived as bright enough even in bright lighting conditions.

For outdoor applications at a sunny day with see-through optics that superimpose the display image on a bright natural environment, the luminance required may be as high as $3000\,\mathrm{cd/m^2}$ [98]. For a monochrome projection display at $\lambda = 635\,\mathrm{nm}$ and a FOV of $15 \times 20\,\mathrm{deg^2}$ this minimum brightness corresponds to an exposure power of about $6\,\mu\mathrm{W}$. This theoretical value matches very well to experimental results about the perceived brightness of a see-through scanning system by Menozzi et al. [99]. For a green display at $\lambda = 532\,\mathrm{nm}$ the needed exposure power is even lower (about $1.5\,\mu\mathrm{W}$) as green light is perceived brighter at same irradiance than the red spectrum range [21].

The pupil diameter was assumed to be light-adapted in this situation (2 mm). This rough estimate shows that the necessary exposure powers are orders of magnitude lower than the corresponding exposure limits. Hence, safety considerations do not set limits to the application of HMDs.

7

Defocusing simulations on a scanning RSD[*]

In this chapter we explore the potential of the retinal scanning technology for providing a see-through display with high depth of focus. The defocusing properties of a retinal scanning display (RSD) are evaluated by using an accommodation-dependent schematic eye model. For characterising the image quality, the contrast function in terms of eye accommodation, spatial frequency and eye rotation is calculated and discussed. Based on that design trade-offs are derived.

---

[*]This chapter is mainly based on Ref. [100]

7.1. Simulation background

7.1.1. Retinal scanning display

The retinal scanning technology has already been discussed briefly in chapter 1. It is based on a light beam which is scanned directly onto the retina in a raster pattern while being intensity modulated to form an image. The principle is depicted schematically in Fig. 7.1. As modulated light source a collimated, low-power laser diode is normally used. The tiny laser beam is subsequently deflected in u- and v-direction by two uniaxial scanners. The horizontal scanner (u-direction) operates at several kHz. The vertical scanner frequency (v-direction) defines the image refresh rate which must exceed the critical fusion frequency (CFF) of about 60 Hz to prevent any flickering effects. Finally, a viewing optics projects the laser beam through the centre of the eye pupil onto the retina.
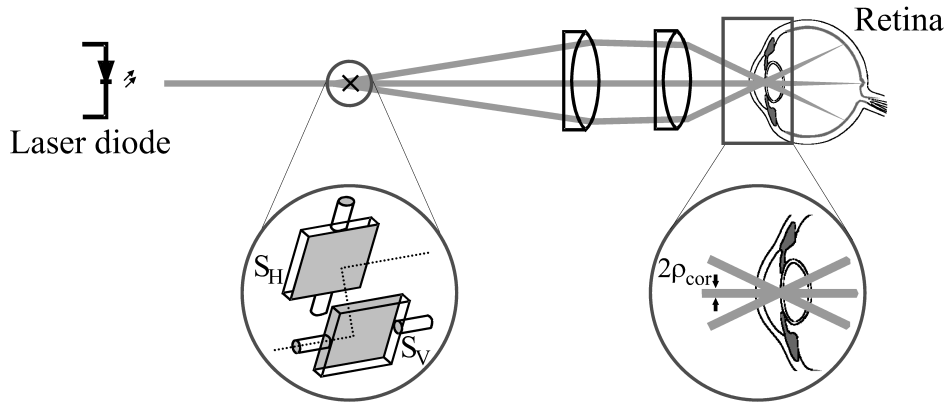
Figure 7.1: *Schematical illustration of the RSD principle (after [99]). The left zoom shows the scanning mirror setup (with S_H and S_V indicating the horizontal and vertical mirror, respectively). The right zoom focuses on the region, where the laser beam enters the eye. ρ_{cor} signifies the radius of the beam at the eye's cornea. This parameter will play an important role in the subsequent simulations.*

In the RSD system the pixels are projected serially in time on the retina. As the image refresh rate lies above the temporal resolution limit of the human eye, the user does not perceive any flickering effects. In optical terms, however, the temporally sequential pixel projection has

the important effect that the pixels are mutually incoherent despite the use of coherent laser light. This has significant influence on the analysis.

7.1.2. Accommodation-dependent schematic eye model

For our purpose, ideally, the eye model should reproduce both anatomy and optical properties (such as first order aberrations) for on- and off-axis beams with a minimum of fitting parameters. Moreover, the model should incorporate the increment of refractive power of the eye during accommodation. Therefore, the simulations are based on the accommodation-dependent wide-angle schematic eye model as proposed by Navarro et al. and Escudero et al. [76, 77]. This model consists of three conical optical surfaces (anterior cornea surface and both surfaces of the eye lens) as well as two spherical surfaces (posterior cornea surface and the retina). Details about the model as well as the explicit parameter values are discussed in appendix A.

7.1.3. Retinal properties

The resolution of the human retina differs from point to point [101, 102]. The highest retinal resolution is found in the retinal centre (known as the fovea) where the receptors are most densely concentrated. The fovea's diameter is approximately 5 deg of visual angle. Just a few degrees from the fovea the resolution falls by an order of magnitude. Thus, when looking at a picture, the eye moves quickly from the current point of gaze to a new location to bring selected image parts to the fovea (so-called 'saccades') [103]. When looking at a scanned image from a RSD

(a) centred image point   (b) peripheral image point
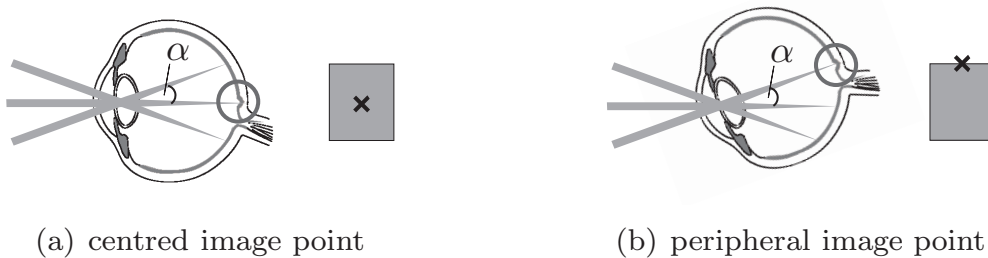
Figure 7.2: *Due to the high foveal resolution the eye moves to bring selected image parts to the fovea. The cross in the grey rectangle indicates the corresponding image part which the eye is currently gazing at.*

system, these eye rotations are also needed to gaze at all parts of the scanned image. In other words: to gaze at peripheral image points with a field angle α, the eye has to rotate by α to make sure that these peripheral beams hit the retina in the fovea (illustrated in Fig. 7.2). Note, however, that these fast saccades do not interfere with the scanning process as the image refresh time ($\approx 20\,\mathrm{ms}$) is about an order of magnitude faster than the characteristic saccadic latency ($\approx 200\,\mathrm{ms}$) [104].

7.2. Contrast function

To characterise the quality of the scanned image and especially the defocusing properties of the retinal scanning display, a MTF-similar[1] contrast function has been calculated in terms of eye accommodation ΔD, spatial frequency ν and corneal beam radius ρ_{cor}. This contrast function (CF) is defined as the ratio of the contrast of the retinal scanned image (c') and the original image (c):

$$CF(\nu, \Delta D, \rho_{cor}) = \frac{c'(\nu, \Delta D, \rho_{cor})}{c} \qquad (7.1)$$

where c and c' stand for the respective Michelson contrast:

$$c = (I_{\max} - I_{\min})/(I_{\max} + I_{\min}) \qquad (7.2)$$
$$c' = (I'_{\max} - I'_{\min})/(I'_{\max} + I'_{\min})$$

Here $I_{\max}$ and $I_{\min}$ indicate the maximum and minimum of the irradiance level in the original image. $I'_{\max}(\nu, \Delta D, \rho_{cor})$ and $I'_{\min}(\nu, \Delta D, \rho_{cor})$ indicate the respective levels in the scanned retinal image.

The calculation of the CF is processed in two steps: In a first step for each image pixel the accurate beam shape and position on the retina is evaluated. These beam properties determine the local resolution and contrast of the scanned image. In a second step the CF is calculated based on these retinal beam data. The viewing optics (as illustrated in Fig. 7.1) is not considered in the simulations since it turned out that possible aberrations do not affect the display's defocusing properties significantly due to the tiny laser beam.

---

[1]MTF=Modulation transfer function

7.2.1. Beam raytracing

The goal of the first step is to calculate the shape and position of the scanning beam on the retina. Thus, the incoming scanned beam is ray-traced from the cornea to the retina in the eye model. To perform this process the ray-tracing software package OSLO [78] is used. Note that due to the scanning mode these raytracing results correspond to the local point spread function $PSF(u', v')$ at a certain retinal pixel point (u', v'). For all simulations it is assumed that the incoming scanned laser beam is Gaussian beam shaped with a variable beam radius ρ_{cor} at the corneal plane (see Fig. 7.1). This ray-tracing calculation is processed for all retinal pixel positions of the scanned image. Further, the refractive power of the modelled eye as well as the eye orientation in terms of rotational movements are varied.

7.2.2. Contrast calculation

In a second step, the CF has been calculated based on these local PSF-values. Considering that the human eye has a time-integrating behaviour while scanning the beam onto the retina, the fused retinal irradiance $I(u', v')$ can be written as an integral over the time of one frame ϑ [105]:

$$I(u', v') \propto \frac{1}{\vartheta} \int_0^\vartheta g(t) PSF(u' - \tilde{u}(t), v' - \tilde{v}(t)) dt \qquad (7.3)$$

with $g(t)$ being the temporal intensity function of the scanning beam. $(\tilde{u}(t), \tilde{v}(t))$ indicate the retinal position of the scanning beam at the time t. Every time a u-line is finished, a new line is started at $\tilde{u} = 0$ but with a slight shift $\delta\tilde{v}$ in v-direction. Therefore, the eq. 7.3 can be rewritten as

$$I(u', v') \propto \frac{1}{\vartheta} \sum_{\ell=1}^n \int_0^{u_{max}} g(\tilde{u}, \ell\delta\tilde{v}) PSF(u' - \tilde{u}, v - \ell\delta\tilde{v}) \frac{d\tilde{u}}{V(\tilde{u}, \ell\delta\tilde{v})} \quad (7.4)$$

where n signifies the total number of lines in one frame and $V(\tilde{u}, \tilde{v})$ the horizontal scanning velocity at the retinal point $(\tilde{u}, \tilde{v})$ [2].

---

[2] For the subsequent analysis it is assumed that $V(\tilde{u}, \tilde{v})$ is constant over one scan line.

CF in u-direction

For the CF-calculation in u-direction, we assume that vertically oriented, dark and bright bars of contrast 1 with a certain spatial frequency ν_u are scanned onto the retina. The corresponding intensity function $g(\tilde{u}, \tilde{v})$ is defined as follows:

$$g(\tilde{u}, \tilde{v}) = \begin{cases} 1 & \text{for} \quad (4m-2)\Delta u \leq \tilde{u} < 4m\Delta u, m \in \mathbb{Z}, \forall \tilde{v} \\ 0 & \text{for} \quad 4m\Delta u \leq \tilde{u} < (4m+2)\Delta u, m \in \mathbb{Z}, \forall \tilde{v} \end{cases} \quad (7.5)$$

Note that the spatial frequency of this bar pattern is $1/(4\Delta u)$. According to eq. 7.1 the contrast can now be received by calculating the irradiance levels $I(u', v')$ of the dark and bright bars in the scanned retinal image. Symmetry considerations show that the maximum level is reached at the middle of the bright bar ($u' = -\Delta u$) while the minimum irradiance value lies at the middle of the dark bar ($u' = \Delta u$), respectively.

Based on eq. 7.4 and eq. 7.5 and leaving out constant factors the maximum irradiance at $u = -\Delta u$ can be expressed as:

$$
\begin{aligned}
I_{max}^u(v') &= I(u' = -\Delta u, v') && (7.6) \\
&\propto \sum_\ell \int g(\tilde{u}, \ell\delta\tilde{v}) PSF(-\Delta u - \tilde{u}, v' - \ell\delta\tilde{v}) \, d\tilde{u}
\end{aligned}
$$

Due to the discrete scanning steps $\delta\tilde{v}$ in v-direction, the irradiance distribution in v-direction varies slightly within the range of $\delta\tilde{v}$ and, thus, it is not totally homogeneous. Consequently, the irradiance level $I_{max}^u(v')$ is averaged in v-direction over a range of $\delta\tilde{v}$. Further, it is assumed that $PSF(u', v') = PSF_u(u') \cdot PSF_v(v')$ which is true for every Gaussian-beam shaped PSF-functions.

$$< I_{max}^u > \propto \int g(\tilde{u}) PSF_u(-\Delta u - \tilde{u}) \, d\tilde{u} \cdot \frac{1}{\delta\tilde{v}} \int_0^{\delta\tilde{v}} \sum_\ell PSF_v(v' - \ell\delta\tilde{v}) \, dv'$$
$$(7.7)$$

After inserting $g(\tilde{u}, \tilde{v})$ as defined in eq. 7.5 and after some substitutions, $< I_{max}^u >$ is given as follows:

$$< I_{max}^u > \propto \sum_{m \in \mathbb{Z}} \int_{-1}^1 PSF_u(\Delta u(\eta + 4m)) \, d\eta \cdot \frac{1}{\delta\tilde{v}} \int_0^{\delta\tilde{v}} \sum_\ell PSF_v(v' - \ell\delta\tilde{v}) \, dv'$$
$$(7.8)$$

Here, $\tilde{u}$ was substituted by $\eta = -\frac{\tilde{u}}{\Delta u} + 4m - 1$.

In analogous manner $< I^u_{min} >$ turns out to be:

$$< I^u_{min} > \propto \sum_{m \in \mathbb{Z}} \int_1^3 PSF_u(\Delta u(\eta + 4m)) \, d\eta \cdot \frac{1}{\delta\tilde{v}} \int_0^{\delta\tilde{v}} \sum_\ell PSF_v(v' - \ell\delta\tilde{v}) \, dv'$$

(7.9)

where $\eta = -\frac{\tilde{u}}{\Delta u} + 4m + 1$. The contrast function can now be calculated by using eqs. 7.1 and 7.2. Thereby, the integrals over v' in eqs. 7.8 and 7.9 are cancelled.

CF in v-direction

In similar manner the CF in v-direction can be calculated by defining an intensity function $g(\tilde{u}, \tilde{v})$ with bright and dark horizontally oriented bars. However, due to the discrete scanning step $\delta\tilde{v}$ in v-direction, the thickness of the bars must be an integer multiple of $\delta\tilde{v}$. Thus, $g(\tilde{u}, \tilde{v})$ can be set as follows and extended periodically:

$$g(\tilde{u}, \tilde{v}) = \begin{cases} 1 & \text{for} \quad -q\delta\tilde{v} \leq \tilde{v} \leq -\delta\tilde{v}, \forall \tilde{u} \\ 0 & \text{for} \quad 0 \leq \tilde{v} \leq (q-1)\delta\tilde{v}, \forall \tilde{u} \end{cases}$$

(7.10)

with $q \in \mathbb{N}_+$. This pattern corresponds to a spatial frequency of $\nu_v = 1/(2q\delta\tilde{v})$.

Based on eq. 7.4 the maximum and minimum irradiance $I^v(u', v')$ at $v' = -\frac{q+1}{2}\delta\tilde{v}$ and $v' = \frac{q-1}{2}\delta\tilde{v}$ can now be figured out with similar considerations as above. It follows:

$$I^v_{max} \quad \propto \quad \sum_{n \in \mathbb{Z}} \sum_{\ell=1}^q PSF_v((-\frac{q+1}{2} - 2nq + \ell)\delta\tilde{v})$$

(7.11)

$$I^v_{min} \quad \propto \quad \sum_{n \in \mathbb{Z}} \sum_{\ell=1}^q PSF_v((\frac{q-1}{2} - 2nq + \ell)\delta\tilde{v})$$

(7.12)

7.3. Contrast analysis

7.3.1. General considerations

The goal of the contrast analysis is to explore the potential of the retinal scanning technology to provide a wearable display with high depth of focus. The exploration space is defined by the following four parameters: spatial frequency $\nu = (\nu_u, \nu_v)$, corneal beam radius ρ_{cor}, change in the eye accommodation ΔD and eye rotation angle $\alpha =$

(α_u, α_v) to gaze at peripheral points. By varying these four parameters the potential, the limits and the design trade-offs will be investigated. Sometimes the spatial frequency will be kept constant. In these cases the contrast values are evaluated and shown for the spatial frequency $\nu = 6.2\,\mathrm{cyc/deg}$ referred as 'standard frequency'[3].

To quantify the depth of focus (DOF) several different criteria are used in the literature. Frequently, the DOF is defined as the dioptric range for which the contrast value does not fall below a certain amount (normally between 0.5 and 0.8) relative to its in-focus optimal value [106, 107]. In our case this definition is not very helpful as the in-focus contrast value depends substantially on the used corneal beam radius. Thus, the definition adopted in this work is slightly different: the DOF here is defined as the dioptric range for which the absolute contrast value does not fall below 0.7 [108]. The limit of 0.7 seems to be rather arbitrary, but a different limit has an effect on the quantitative results only, but not on any qualitative ones.

Note further that the display is assumed to provide best focusing for an eye accommodated at infinity. Consequently, the DOF in this chapter is measured as the dioptric power difference from optical focusing to the maximum defocusing, only [4]. By adjusting the best focusing distance to closer values than infinity the depth of focus could be increased since in this case depth of focus is defined as the dioptric power difference between the maximum defocusing values on both sides around optimal focus.

For the first CF-simulation it is assumed that the modelled eye is unrotated and perfectly aligned to the scanning optics (i.e. $\alpha = 0$). That means that the exit pupil of the scanning optics is concentrical to the entrance pupil of the modelled eye and the eye is gazing at a central image point. The eye pupil diameter is set to $3\,\mathrm{mm}$. Note that this diameter has no effect on the contrast results as long as the eye is aligned and the corneal beam radius ρ_{cor} is much smaller than the pupil radius. Variable parameters are, thus, the corneal beam radius ρ_{cor}, the change in the eye's refractive power ΔD due to accommodation and the spatial frequency ν. $\Delta D = 0$ is set to the unaccommodated, emmetropic eye with $D = 60.2\,\mathrm{D}$, where the eye is focused to infinity.

---

[3]For illustration, note that this standard frequency corresponds approximately to the frequency occurring at the capital letter 'E' with a font size viewing angle $\alpha_v = 0.4\,\mathrm{deg}$ and can, thus, be considered as an acceptable benchmark for text targets.

[4]Note that this definition of DOF is different to the one for the retinal projection displays in the previous chapters.

7.3.2. Resolution analysis

As the CF in u- and v-direction are not identical (see eqs. 7.7-7.11), both directions have to be calculated. For the analysis the respective lower value is to be considered. Figure 7.3(a) shows the corresponding through-focus CF in u- and v-direction in terms of changes in eye accommodation ΔD for three different spatial frequencies ν at a fixed corneal beam radius $\rho_{cor} = 350\,\mu$m.

As it can be seen from Fig. 7.3(a) the contrast-values in u-direction are lower than the ones in v-direction for any values of accommodation and spatial frequency. It turns out that this relation is valid for any values of the corneal beam radius and for any scanning steps $\delta\tilde{v} > 0$, as long as the eye is perfectly aligned to the optics. This asymmetry between u- and v-direction is caused by the fact that the scan pattern in v-direction is discrete while in u-direction the image is scanned continuously. Thus, the subsequent simulation results for central image points deal with the CF in u-direction, only. For illustration, Fig. 7.3(b) shows the CF in u-direction in terms of eye accommodation ΔD and spatial frequency ν_u.

Note that the resolution in v-direction is limited due to the discrete scanning step $\delta\tilde{v}$. The cut-off frequency ν_{cut}^v in v-direction is $\nu_{cut}^v = 11.9\,$cyc/deg for $\delta\tilde{v} = 15\,\mu$m. Certainly, the cut-off frequency can be extended when reducing the scanning step $\delta\tilde{v}$. However, this results in either a decrease of the vertical image field or in a higher u-scanning frequency. A reduction of the image refresh rate - as the third possibility - is not feasible as this leads to a flickering image.

7.3.3. Corneal beam variation

It is well known that the depth of focus of a scanning system can be increased when decreasing the corneal beam radius. However, this comes along with a decrease of the in-focus image contrast since the retinal radius of a well focused beam increases due to diffraction effects.

Figure 7.4 illustrates this effect for the standard frequency $\nu_u = 6.2\,$cyc/deg: The lower the corneal beam radius is chosen, the higher is the depth of focus which results in a lower decrease of the contrast values when the eye changes its refractive power. But similarly, the contrast values for $\Delta D \lesssim 1.5\,$D decrease.

To evaluate this trade-off, Fig. 7.5(a) shows the isolines for different contrast values as a function of ΔD and ρ_{cor}. The bold line in Fig. 7.5(a) symbolises the depth of focus (DOF) at CF=0.7, as defined
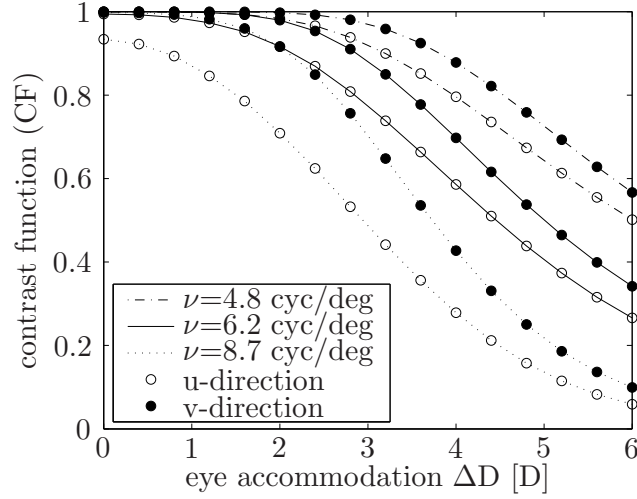
(a) Through-focus CF
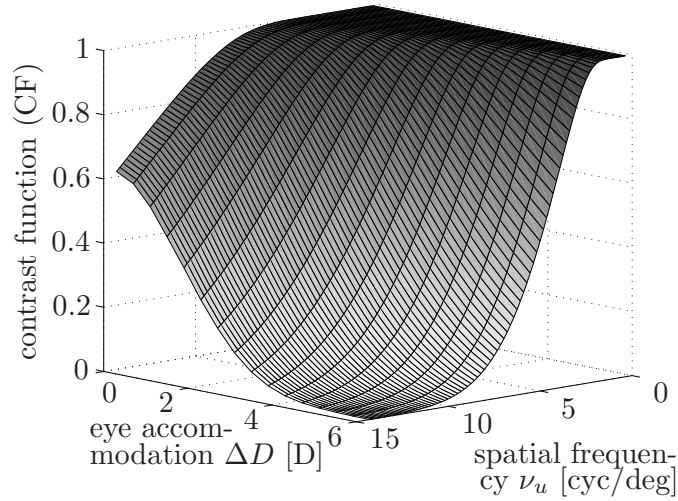
(b) CF in u-direction

Figure 7.3: *(a) shows the through-focus CF in u and v-direction for three different spatial frequencies. The bright circles (○) indicate the u-direction whereas the filled ones (●) signify the v-direction. Here the scanning step $\delta\tilde{v}$ was $15\,\mu m$. (b) shows the respective CF in u-direction in terms of eye accommodation and spatial frequency. In both cases the corneal beam radius ρ_{cor} has been set to $350\,\mu m$.*
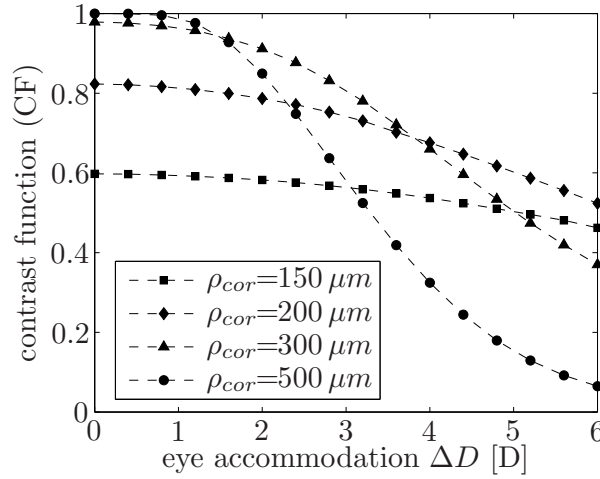
Figure 7.4: *The through-focus CF for different values of the corneal beam radius ρ_{cor}. The spatial frequency ν_u was set to the standard frequency $\nu_u = 6.2\,cyc/deg$.*

above. As expected the DOF increases when reducing ρ_{cor} as long as $\rho_{cor} \geq 250\,\mu$m. Reducing the corneal beam radius further, however, does not result in a further increase in DOF, but in an abrupt decrease as diffraction effects become more and more important. Thus, for each spatial frequency an optimum beam radius ρ_{cor} exists where the DOF is maximised. As presented exemplarily in Fig. 7.5(a) the optimum for a resolution target of $\nu_u = 6.2\,$cyc/deg lies at $\rho_{cor} \approx 230\,\mu$m where a maximum DOF of about $4\,$D is achieved.

This analysis can be repeated for any spatial frequencies. The results are depicted in Fig. 7.5(b): The circles (lower plot, right axis) show the corneal beam radius at which the system is optimised for the corresponding spatial frequency. The triangles (upper plot, left axis) indicate the resulting DOF at that resolution and beam radius. It can clearly be seen that if the system should be optimised to higher spatial frequencies the corneal beam must be widened to reduce diffraction effects. This, however, results in a lower depth of focus. Consequently, Fig. 7.5(b) shows the fundamental limits of the retinal scanning technology with regard to the trade-off between large depth of focus and high resolution and thus, gives first important criteria for the display design.
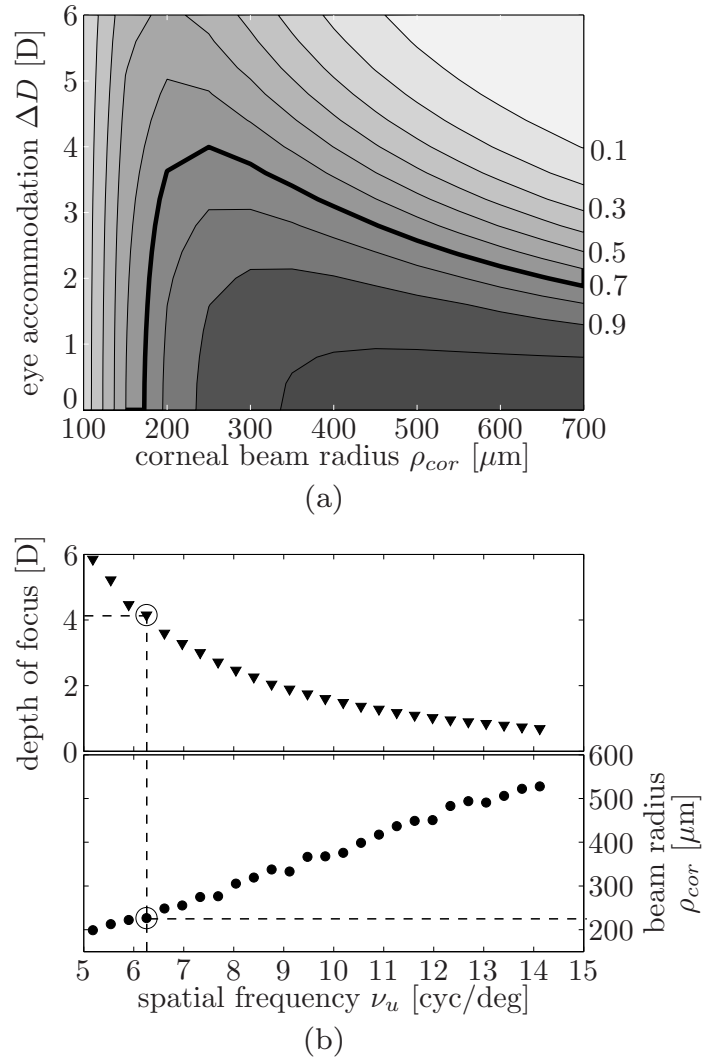
Figure 7.5: *a) CF-isolines for $\nu_u = 6.2\,cyc/deg$ as a function of corneal beam radius and changes in eye accommodation. The CF-level values are indicated at the right edge. The bold line signifies the DOF at CF=0.7. b) The circles ($\bullet$) show the optimum corneal beam radius in terms of spatial frequency as explained in the text. The triangles ($\blacktriangledown$) indicate the corresponding maximum depth of focus. The dashed lines symbolise the discussed example at $\nu_u = 6.2\,cyc/deg$.*

7.4. Image field analysis

So far, only the contrast for the unrotated eye - and consequently for the central part of the scanned image - has been analysed. Thus, the following section will focus on effects arising when gazing at peripheral points. Of special importance will be the maximum angle α which the eye can be rotated horizontally (u-direction) and vertically (v-direction) without a significant loss in image quality. These angles define approximately the maximum image field within which the eye can still completely recognise the image without any pupil tracking systems.

The simulation procedure is identical to the section before. However, in contrast to the previous simulations the beams of interest are no longer the beams from the middle of the image, but from peripheral points where the eye has to be rotated in order to gaze at. Furthermore, the eye pupil diameter now affects the image quality as it might block the scanned beam. Consequently, the smaller the pupil, the more it affects the image quality when the eye rotates. Therefore, for any subsequent simulations the eye pupil diameter was set to $3\,\mathrm{mm}$ which corresponds to the minimum value occurring in the human eye for normal ambient lighting conditions [104].

7.4.1. In-focus analysis

First, the effect of eye rotations on the in-focus CF is analysed. When the eye is rotating the entering laser beams diffracts more and more at the eye pupil edge causing a widening of the beam at the retina. For horizontal eye rotations the retinal beam is widened in u-direction, while for vertical rotations the same happens in v-direction, correspondingly.

Figure 7.6 shows the CF-isolines in terms of spatial frequency ν and corresponding eye rotation α. The corneal beam radius was set to $350\,\mu\mathrm{m}$ and the eye accommodation to $0\,\mathrm{D}$. Qualitatively similar patterns can be seen for any other values of beam radius. For low spatial frequencies ($\nu \lesssim 6\,\mathrm{cyc/mm}$) the eye rotating has no significant effect on the contrast up to the limit of $\alpha_{max} \approx 12\,\mathrm{deg}$ where the laser beam begins to be blocked by the eye pupil. This results in an abrupt fall of the CF to zero. For higher spatial frequencies this abrupt decrease of the CF is less pronounced. In return, however, the decrease in the contrast starts already at smaller rotating angles. Analysing this for all corneal beam radii leads to Fig. 7.7 where the vertical and horizontal maximum image field is plotted as a function of ρ_{cor} for the standard frequency
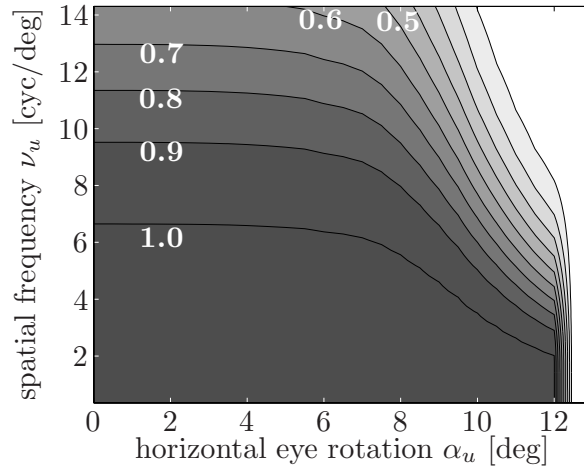
(a) CF in u-direction
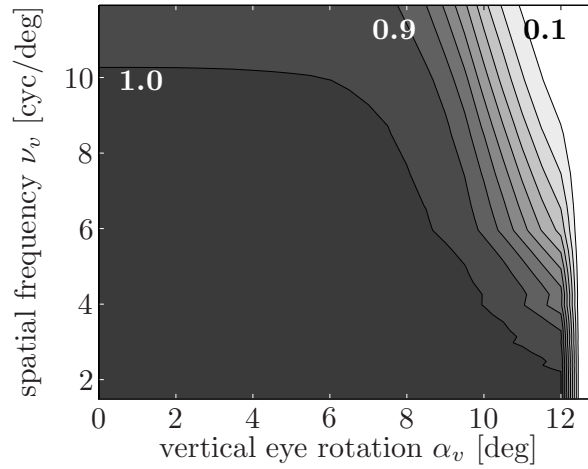
(b) CF in v-direction

Figure 7.6: *Isolines of the CF for $\Delta D = 0$ and with $\rho_{cor} = 350\,\mu m$ as a function of spatial frequency and eye rotation. The CF level values are indicated in the plot. The eye pupil diameter was set to $3\,mm$.*
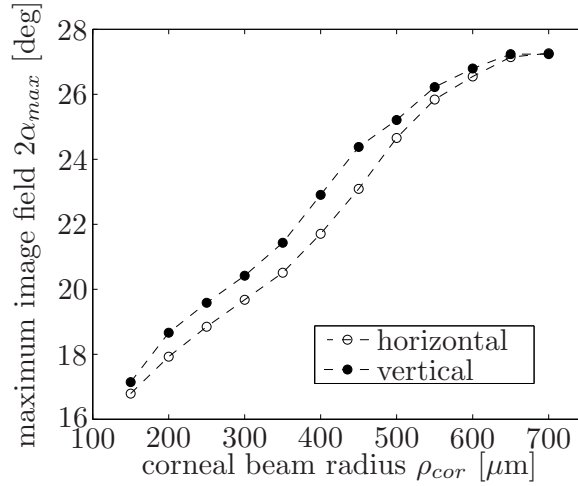
Figure 7.7: *The maximum image field as a function of corneal beam radius for a retinal scanning system without pupil tracking. The data are evaluated for $\nu = 6.2\,cyc/deg$ with an eye pupil diameter of $3\,mm$. The maximum image field is defined as twice the rotating angle where the CF has decreased to 0.7.*

$\nu = 6.2\,\text{cyc/deg}$. Note that the maximum image field is defined as twice the rotating angle α_{max} where the contrast value has decreased to 0.7.

7.4.2. Defocusing analysis

In a second step, the defocusing properties for peripheral points should be investigated. Thus, Fig. 7.8 shows the contrast values in u-direction for $\nu_u = 6.2\,\text{cyc/deg}$ in terms of eye accommodation and horizontal eye rotation α_u for two selected corneal beam radii. For $\rho_{cor} = 200\,\mu\text{m}$ the CF is rather constant for any eye rotations up to the maximum rotating angle $\alpha_{max} \approx 9\,\text{deg}$. As known from the in-focus analysis, at α_{max} the beams are blocked by the eye pupil resulting in a abrupt fall of the contrast and the loss of the scanned image. Note especially that the maximum rotating angle (and consequently the image field) are nearly independent of the eye's refractive power. For $\rho_{cor} = 350\,\mu\text{m}$ the CF looks differently: For small eye rotations the CF is also rather constant. However, for higher eye accommodation and for eye rotations close to the limit α_{max} the contrast values increase before falling abruptly. This comes from the fact that first only a part of the laser beam is blocked resulting effectively in a smaller beam radius and thus a higher depth of focus. For higher corneal beam radii this effect is even more pronounced. Again, the maximum rotation angle is not dependent on the eye accommodation.
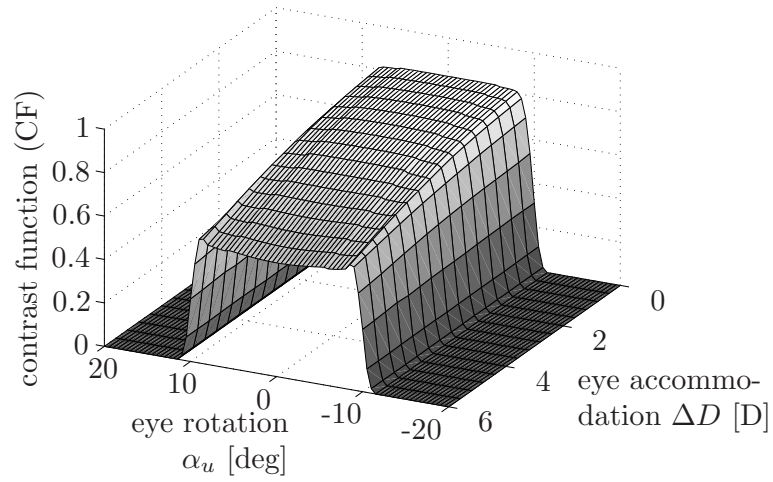
(a) $\rho_{cor} = 200\,\mu$m
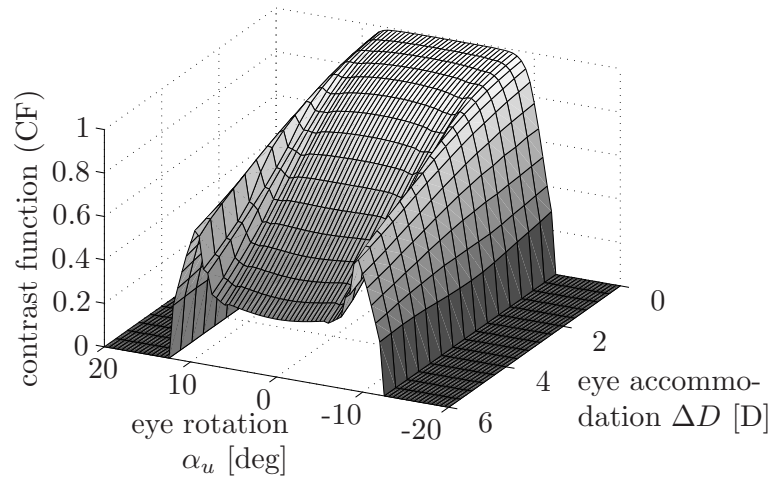
(b) $\rho_{cor} = 350\,\mu$m

Figure 7.8: *CF in horizontal direction in terms of eye rotation α_u and changes in eye accommodation ΔD for the standard frequency $\nu = 6.2\,cyc/deg$.*

7.5. Final trade-off between resolution, DOF and image field

The results on the depth of focus from the previous section can finally be combined with the latest image field results. Figure 7.9 shows this fundamental trade-off between resolution, depth of focus and maximum image field for different spatial frequencies. For instance, for $\nu_u = 6.2\,\text{cyc/deg}$ the maximum DOF in u-direction is $\approx 4\,\text{D}$ and is reached - as seen in the Fig. 7.5(b)- for $\rho_{cor} \approx 230\mu\text{m}$, but with a limited image field of $\approx 18\,\text{deg}$ (see Fig. 7.7). This image field can be extended to e.g. $\approx 23\,\text{deg}$ at $\rho_{cor} \approx 450\,\mu\text{m}$, but at the expense of the depth of focus which decreases to $2.8\,\text{D}$. Similar results have been evaluated for other frequencies.
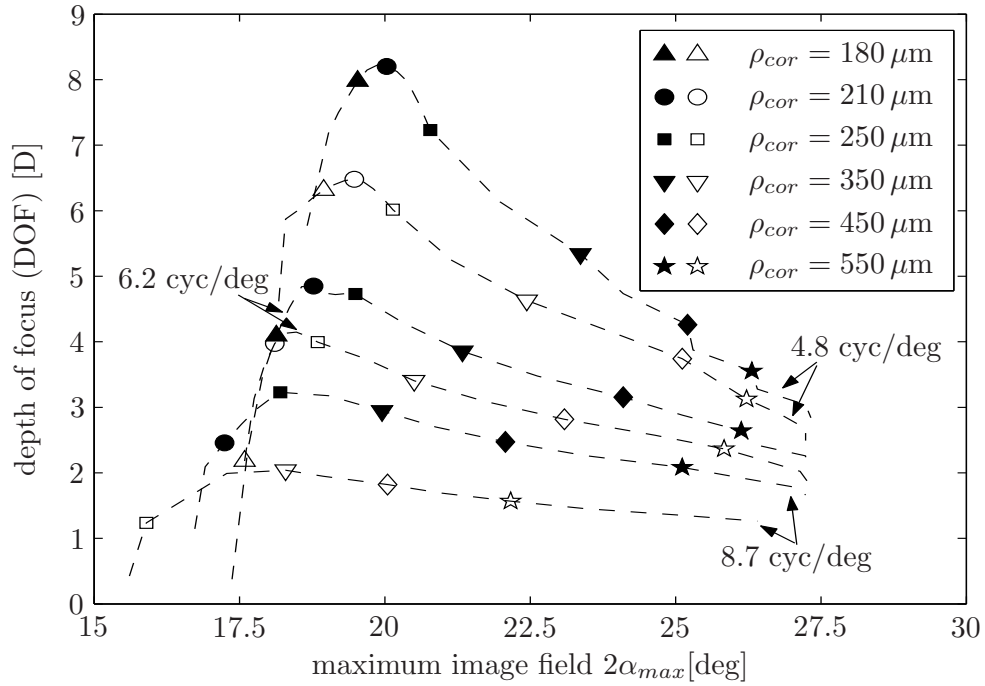
Figure 7.9: *Trade-off between depth of focus (DOF) and maximum image field for three different spatial frequencies. The lines connect points of the same frequency, but with different ρ_{cor}-values. In contrast, the symbols indicate points with the same ρ_{cor}-value at different frequencies. Unfilled symbols indicate data in u-direction, whereas filled symbols signify the v-direction. The selected frequencies correspond to the capital letter 'E' with font size viewing angles $\alpha_v = 0.28/0.40/0.52\,\text{deg}$.*

Further, it can be seen, that the DOF in v-direction is normally larger by approximately 1 D due to the discrete scanning pattern in v-direction. This has beneficial effects on the design process: e.g. a display optimised for $\nu = 6.2\,\text{cyc/mm}$ can be designed with an elliptical corneal beam shape ($\rho_{cor}^u \approx 230\,\mu\text{m}, \rho_{cor}^v \approx 310\,\mu\text{m}$) rather than with a circular one. Such a display provides a DOF of about 4 D with an image field of 18.4 x 20.7deg. For higher spatial frequencies (e.g. $\nu = 8.7\,\text{cyc/mm}$) the use of elliptical beam shapes has even more effect.

7.6. Conclusion and outlook

In conclusion, the following results of the analysis can be highlighted:

- There is a design trade-off between in-focus resolution, depth of focus and maximum image field caused by diffraction effects and retinal properties.

- Due to the asymmetry in the scan pattern between u- and v-direction, this trade-off can be optimised by using an elliptically-shaped scanning beam. In this case a display with a DOF of 4 D is achieved for spatial frequencies up to 6.2 cyc/deg. This frequencies allow to properly display text targets with font size viewing angles $\alpha_v \geq 0.4\,\text{deg}$. The image field is limited to 18.4 x 20.7 deg. Thus, for font sizes of $\alpha_v \geq 0.4\,\text{deg}$ the display can provide practically entirely accommodation-insensitive images for any viewing distances between infinity and the typical reading distance of about 25 cm.

- When adjusting the best focusing distance to 50 cm [5] rather than infinity, a total DOF of 4 D can be achieved even for 8.7 cyc/deg. In this case, the image field is limited to 18.2 x 25.2 deg[6] when using an elliptical beam shape of $350\,\mu\text{m}$ x $550\,\mu\text{m}$.

When the human eye ages, accommodation decreases gradually, so that spectacles are needed to focus on closer objects. Further simulations have shown that such spectacles do not affect the defocusing properties of the retinal scanning display significantly.

Although the analysis is based on an accurate and sophisticated eye model it is not able to replace real measurement with a corresponding

---

[5] this corresponds to $-2\,\text{D}$.

[6] this corresponds approximately to a 11"-screen viewed from a distance of 50 cm.

display implementation and with the real human eye. In particular, some human eye factors such as the appearance of eye floaters - as reported in [80] for small exit pupils - are not considered in this analysis. Thus, to take all these additional factors into account, extended tests with a bench model and a statistical set of subjects will be needed. Note further that by integrating a sophisticated pupil tracking system which adjusts the display's exit pupil to the current eye's position in real-time the image field constraint can be extended. However, this causes additional technical expenses and solves only the image field constraint, but not the trade-off between depth of focus and in-focus resolution.

8

Conclusion

8.1. Summary and Achievements

In the last years, see-through head-mounted displays have gained increasing interest in a wide area of applications, particularly in the fields of wearable computing as well as of augmented reality. For wearable computers, see-through head-mounted displays (HMDs) promise to become a valuable output device that is unobtrusive and allows to combine the real scene with virtual data. However, the see-through mode in today's common HMDs provokes often a severe accommodation conflict as soon as the user changes his accommodation to focus on a real object at a different distance. In that case, namely, the overlaid virtual data - as provided from the HMD - gets out of focus and thus becomes blurred. The reason for that lies in the limited depth of focus (DOF) of today's HMDs. This work focuses on this well-known accommodation conflict. The objective of this work is the evaluation of several optical techniques in order to extend the depth of focus of a see-through HMD substantially. Ideally, the perceived quality of the virtual information should become insensitive to ocular accommodation changes. This, however, should be achieved without substantially degrading other important visual specifications such as display luminance, field of view or

resolution. The major contributions of the thesis are the following:

- We have proposed a partially coherent retinal projection display system which is mainly characterised by a direct retinal projection of the LCD image onto the retina ('Maxwellian-view optical system') as well as by the use of partially coherent illumination light instead of incoherent one as usual. The retinal projection improves the light efficiency and thus, increases the retinal irradiance and display luminance, respectively. This is a decisive fact when applying the see-through display in bright ambient lighting conditions. The spatially partial coherence of the illumination light improves the defocusing properties of the displays in comparison to a common incoherent system.

- When increasing the spatial coherence of the illumination light, there is a trade-off between contrast gains due to coherent imaging and increasingly disturbing coherence effects. We have discussed and explored this trade-off in order to find an optimum spatial coherence level where the depth of focus is maximised. To this end, measurements of both the contrast function and the image quality of text targets have been carried out on the basis of a bench model. The quality of the projected text samples were evaluated by a combination of objective image quality criteria and direct visual assessments by subjects. The experiments revealed that the depth of focus can be extended by a factor of about 1.5 compared to a corresponding incoherent system when applying partially coherent light with a coherence level of $0.35 \lesssim \sigma \lesssim 0.5$.

- To extend the DOF of such a retinal projection display further, we have evaluated two additional optical techniques on the basis of simulations: First, the use of the multiple imaging technique by means of a phase-only mask and second, the incorporation of an oscillating fluid lens at the display's aperture stop.

  * For applying the multiple imaging technique, a phase-only mask is introduced at the system's aperture stop. This additional phase mask produces a series of images of the LCD at various focal planes, shifted to each other by an individual phase. The images are projected simultaneously, so that they superpose coherently to each other. Based on an accommodation-dependent eye model and a partial coherence simulator we have designed an appropriate phase mask

and have evaluated its potential in terms of extending DOF for reading text. The evaluation of simulated retinal text samples by objective image quality criteria as well as psychometric experiments has shown that the use of the multiple imaging technique is very promising provided partially coherent light is used. However, the quality of the retinal images is degraded by structural and contrast deficit if fully coherent or incoherent light is used. The psychometric evaluation has further shown that the multiple imaging technique can extend the depth of focus by a substantial factor of up to 3.2 compared to the corresponding Rayleigh criterion. Consequently, the DOF achieved lies at about 5.5 D for font sizes subtending viewing angles of 0.4 deg or larger. This DOF-value exceeds the typical range of accommodation of adult human beings (of about 5 D) providing, thus, really accommodation-insensitive viewing.

∗ The sinusoidally oscillating fluid lens projects the LCD image sequentially at different foci onto the retina. As a consequence, the retinal image as perceived by the user corresponds to an incoherent fusion of several defocused and in-focus images. Simulations of text samples based on the accommodation-dependent eye model have shown, that this approach is promising provided that partially coherent illumination light ($0.2 \lesssim \sigma \lesssim 0.4$) is used. In that case the display's depth of focus can be extended by a factor of up to 2.2 compared to the corresponding Rayleigh criterion if the oscillation parameters of the fluid lens are chosen appropriately. The DOF achieved lies in this system at about 3.8 D for font sizes subtending viewing angles of 0.4 deg or larger. For fully coherent and fully incoherent light, the fused retinal images suffer from structural and contrast degradations, respectively.

Table 8.1 sums up all optical methods evaluated in this work and the corresponding results.

• For a wearable, unobtrusive use, compactness and weight are essential factors. Thus, we have proposed a lightweight and compact spectacle-based design of this retinal projection display. The design consists of a commercially available miniature LCD-device and an holographic optical element acting as optical combiner

for fusing the real scene with the virtual data. To guarantee un-obtrusiveness the system can be fully integrated into common eyeglasses.

- Finally, we have explored the potential of the retinal scanning technology for providing accommodation-insensitive viewing. To this end, we have carried out defocusing simulations in terms of eye accommodation, eye rotation and spatial frequency. The design parameter in this system is the radius of the Gaussian-shaped laser scanning beam at the eye's cornea. The simulations have shown that there is a design trade-off between in-focus resolution, depth of focus and maximum field of view caused by diffraction effects and retinal properties. Due to the asymmetry in the scan pattern between the x- and y-direction, this trade-off is optimised by choosing an elliptically-shaped scanning beam rather than a circular beam shape.

Table 8.1: *Summary of all methods evaluated to extend the DOF of the LCD-based retinal projection display*

optical method	DOF extension factor[a]	optimum coherence level
partially coherent light	1.5	$0.35 \lesssim \sigma \lesssim 0.5$
multiple imaging technique	3.2	$0.3 \lesssim \sigma \lesssim 0.6$
oscillating fluid lens	2.2	$0.2 \lesssim \sigma \lesssim 0.4$

[a]compared to the Rayleigh criterion with incoherent light

8.2. Outlook

There are several starting points for further research, which we briefly discuss in the following paragraph:

- The evaluations presented in this work for the oscillating fluid lens as well as the multiple imaging technique are based on simulations, only. This method has the advantage that all system

parameters can be controlled properly and that objective image quality criteria can be applied. However, for further research an implementation of the proposed fluid lens and phase-only mask will be required. On the basis of an implemented display prototype, further psychometric experiments with a set of subjects may help to explore the promising potential of these techniques in more detail.

- Additionally, a prototype will offer the opportunity to gain more experience about the psychological reaction of users when they are provided with an accommodation-insensitive virtual scene overlaid over the real view.

- In technical terms, a further approach could be the combination of the multiple imaging technique and the retinal scanning technology. In this case, the final scanning laser beam, entering the eye at the cornea, is a coherent superposition of beams of different foci. Similar as discussed in chapter 4 for the projection display, the multiple imaging technique is expected to result in a degraded image quality for in-focus scanned images, but in a significant improvement for defocused images. The resulting potential extension of the depth of focus can also be used to increase the radius of the scanning beam. This would - as discussed in chapter 7 - lead to a larger field of view and higher image resolution due to reduced vignetting and diffraction effects, respectively.

- The most restrictive parameter for all discussed display types is still the limited field of view due to vignetting effects in case of retinal projection. To evaluate this restriction, further research will be needed. The integration of a pupil tracking system which adjusts the display's exit pupil to the current eye pupil position will surely help to overcome this shortcoming [109, 110, 111].

A

Accommodation-dependent schematic eye model

The simulations in the chapters [4],[5] and [7] are based on an accommodation-dependent schematic eye model. This appendix will discuss the eye model and the modelling of the ocular accommodation.

A.1. The emmetropic eye model

Several different schematic eye models have been proposed in the litera-ture[1] - from the simple and famous Gullstrand eye [112] up to a sophis-ticated one which models the eye's anatomy very accurately [113]. For our purpose, ideally, the eye model should reproduce both anatomy and optical properties (such as first order aberrations) with a minimum of fitting parameters. Moreover, the model should incorporate the incre-ment of refractive power of the eye during accommodation. Therefore, all simulations presented in this thesis are based on the accommodation-dependent wide-angle schematic eye model as proposed by Navarro et al. and Escudero et al. [76, 77].
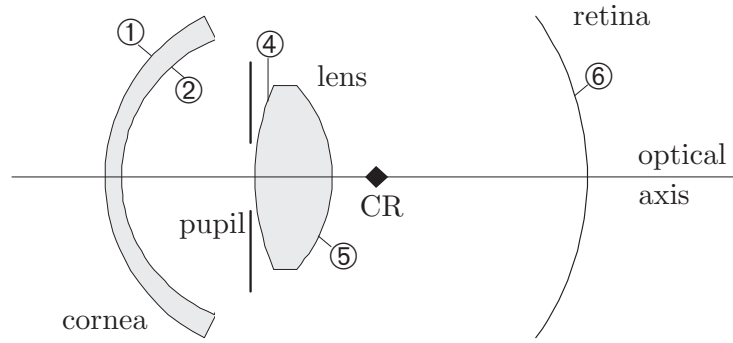
Figure A.1: *The accommodation-dependent schematic eye model. The numbered labels indicate the surfaces as parameterised in Tab. A.1. CR indicates the centre of rotation as explained in the text.*

This model - depicted in Fig. A.1 - consists of three conical optical surfaces (surfaces 1,4,5) and two spherical surfaces (surfaces 2,6) whose second one represents the retina. Table A.1 shows the geometrical pa-rameters and the used refractive indices in the unaccommodated state for a reference wavelength $\lambda = 632.8$ nm. Note in particular, that in this model all parameters defining the eye's geometry in the unaccom-modated state are based on anatomical data so that there is no need to fit original values to match experimental results.

The image surface is intersected by the optical axis at the paraxial focus for the reference wavelength $\lambda = 632.8$ nm. The focal length of

---

[1] a good overview is given in [106].

Table A.1: *Geometry of the schematic wide-angle eye model in the unaccommodated state*

Surface	Anatomy	Conic Constant	Radius [mm]	Thick-ness [mm]	Optical Medium	Refractive Index (@ 632.8 nm)
1	Anterior Cornea	-0.26	7.72	0.55	Cornea	1.3747
2	Posterior Cornea	0	6.50	3.05	Aqueous	1.336
Stop	Pupil	0	Infinity	0	Aqueous	1.336
4	Anterior Lens	-3.1316	10.20	4.00	Lens	1.4183
5	Posterior Lens	-1.0	-6.00	16.4562	Vitreous	1.3347
Image	Retina	0	-12.00			

this schematic eye is 22.18 mm in the image space; the refractive power in the unaccommodated state is 60.2 diopters (D).

A.2. Accommodation modelling

When the natural eye accommodates, the increment in refractive power is mainly caused by two contributions: first, a geometrical variation of the eye lens shape and second, a change of the graded-index structure within the eye lens. In this eye model, the geometrical shape variations are incorporated accurately as proposed by [76]. They are listed in Tab. A.2. The second contribution, the variation in the graded-index structure, however, is replaced by an effective, accommodation-dependent refractive index $n_3(\Delta D)$, for simplification.

To get the dependence of the effective index $n_3(\Delta D)$, a cubic adjustment has been made to fit the refractive powers for 2 D, 4 D and 6 D of accommodation at the reference wavelength $\lambda = 632.8$ nm, similar to the procedure in Ref. [76] (cf. Tab. A.2). Consequently, it is the only parameter which is not based on anatomical data. For details of the modelling process as well as comparisons of simulated aberration results with experimental data see Ref. [76, 77].

A.3. Eye rotation

The natural eye rotates in its socket under the action of six muscles. Because of the way these muscles are positioned and operate, there is

Table A.2:   *Dependence of the lens parameters on changes of the eye accommodation ΔD (in diopters)*

Lens Parameter	Accommodation Dependence
Anterior lens radius	$R_3(\Delta D) = 10.2 - 1.75 \cdot ln(\Delta D + 1)$
Posterior lens radius	$R_4(\Delta D) = -6 + 0.2294 \cdot ln(\Delta D + 1)$
Aqueous thickness	$D_2(\Delta D) = 3.05 - 0.05 \cdot ln(\Delta D + 1)$
Lens thickness	$D_3(\Delta D) = 4 + 0.1 \cdot ln(\Delta D + 1)$
Anterior lens conic	$Q_3(\Delta D) = -3.1316 - 0.34 \cdot ln(\Delta D + 1)$
Posterior lens conic	$Q_4(\Delta D) = -1 - 0.125 \cdot ln(\Delta D + 1)$
Lens refractive index	
(for $\lambda = 632.8$ nm)	$n_3(\Delta D) = 1.4183 - (1.66\Delta D - 0.839 * \Delta D^2 + 0.062\Delta D^3) \cdot 10^{-3}$

no unique centre of rotation (CR). However, a mean position for this point can be assumed to lie 8.5 mm behind the eye pupil and along the optical axis (as indicated in Fig. A.1) [66].

B

Theory on partial coherence

The goal of this appendix is to give a short review on the theory of partial coherence. In particular, the Hopkins formula are derived which the simulations of the retinal images in the chapters 4 and 5 are based on. Further, the definition of the coherence level σ as the ratio of the numerical apertures of the condenser and the optical system is motivated. In this review we follow mainly the work by Hopkins [53].

B.1. Definition of spatial coherence

The spatial coherence is usually defined by means of a phase-coherence factor Γ which specifies the correlation between wave disturbances at any two points. For two points P_1, P_2 of an illuminated surface, Γ is defined as:

$$\Gamma_{12} = \frac{1}{\sqrt{I_1 I_2}} \int_{\Sigma} \imath_1 \imath_2^* d\sigma \tag{B.1}$$

where I_1, I_2 are the intensities at P_1, P_2 produced by the source Σ. $\imath_1, \imath_2$ are the complex amplitudes at these points associated with an element $d\sigma$ of the source. If the source has an intensity $\gamma(\hat{x}, \hat{y})$ at the relative source point $(\hat{x}, \hat{y})$, the coherence factor Γ between points P_1 and P_2 of the object plane illuminated by the source can be expressed as:

$$\Gamma(\hat{u}_1 - \hat{u}_2, \hat{v}_1 - \hat{v}_2) = \frac{1}{\sqrt{I_1 I_2}} \iint\limits_{-\infty}^{\infty} \gamma(\hat{x}, \hat{y}) e^{i(\hat{x}(\hat{u}_1 - \hat{u}_2) + \hat{y}(\hat{v}_1 - \hat{v}_2))} d\hat{x} d\hat{y} \tag{B.2}$$

Here, $\gamma(\hat{x}, \hat{y})$ is taken to be zero at points exterior to the source. Further, $(\hat{u}_1, \hat{v}_1)$ and $(\hat{u}_2, \hat{v}_2)$ are rectangular reduced coordinates of P_1 and P_2. If we scale the unit of intensity to make $\sqrt{I_1 I_2} = 2\pi$ and assume that the plane is uniformly illuminated, eq. B.2 defines the coherence factor Γ to be the Fourier transform of the source map $\gamma(\hat{x}, \hat{y})$.

B.2. Notation and Scaling

In the following we consider a simple optical system as given schematically in Fig. B.1: The source is considered to be placed at plane E_0 at some distance from the object plane O. Subsequent to the transmissive object at O an image-forming optical system (described by their entrance and exit pupils E and E') projects the object to the image plane at O'. $E_0 O E E'' O'$ is the axis of the optical system.

The reduced variables $(\hat{u}, \hat{v})$ and $(\hat{u}', \hat{v}')$ will now be defined with reference to Fig. B.1: Let the marginal ray (i.e. the ray from the object axial point O which just passes through the aperture stop) make angle of α with the axis so that $NA = N \sin \alpha$ defines the numerical aperture [1] in the object space. With (u, v) the geometrical distances in the object, the reduced variable $(\hat{u}, \hat{v})$ is defined by:

$$\hat{u} = kNAu \qquad\qquad \hat{v} = kNAv \tag{B.3}$$

---

[1] N is the refractive index of the object space.
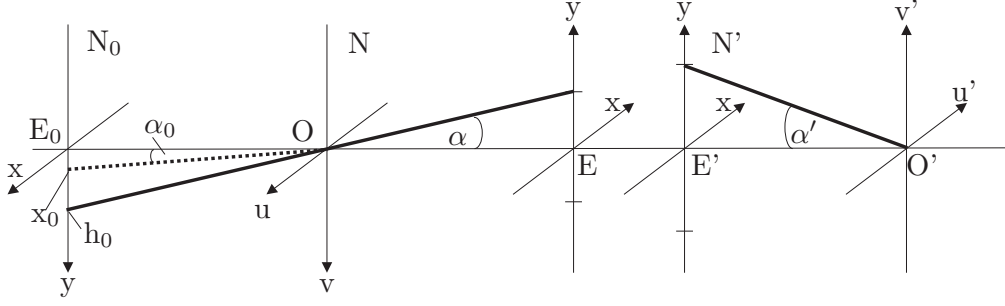
Figure B.1: *Arrangement and notation of the optical system. The bold line shows the marginal ray which passes through the rim of the entrance and exit pupil. The dashed line shows the case of a point source placed at $(x_0, 0)$ as referred to in section B.4*

Similarly, for the image plane,

$$\hat{u}' = kNA'u' \qquad\qquad \hat{v}' = kNA'v' \qquad\qquad \text{(B.4)}$$

Usually, the optical system has a circular aperture stop. Thus, the co-ordinates $(\hat{x}, \hat{y})$ in the entrance and exit pupils E, E' are defined as relative coordinates so that the two pupils have a radius of 1. Similarly, the coordinates in the source plane E_0 are normalised by the distance h_0 from the axis, at which the marginal ray, continued backwards to the source, meets the source plane (see Fig. B.1).

B.3. Image forming

Let $H(\hat{x}, \hat{y})$ denote the coherent transfer function of the optical system, that is between the entrance and exit pupil plane E and E'. $H(\hat{x}, \hat{y})$ includes both the aberrations of the system and a term corresponding to defects of focus. It is convenient to define $H(\hat{x}, \hat{y})$ to be zero at points exterior to the pupil. Consider further that the amplitude spread function $h(\hat{u}, \hat{v})$ may be obtained as the Fourier transform of $H(\hat{x}, \hat{y})$.

Let a complex function $\tau(\hat{u}, \hat{v})$ specify the complex transmission of the object at the object plane O. If an element $d\sigma$ of the source produces a complex amplitude $\imath_1$ at $(\hat{u}_1, \hat{v}_1)$, the amplitude at $(\hat{u}', \hat{v}')$ in the image plane due to light from $d\sigma$ is given by:

$$\imath(\hat{u}', \hat{v}') = \iint\limits_{-\infty}^{\infty} \imath_1 \tau(\hat{u}_1, \hat{v}_1) h(\hat{u}' - \hat{u}_1, \hat{v}' - \hat{v}_1) d\hat{u}_1 d\hat{v}_1 \qquad \text{(B.5)}$$

where $h(\hat{u}, \hat{v})$ denotes the amplitude spread function. Thus, the partial intensity in the image which is associated with light from $d\sigma$ is then given by

$$
\begin{aligned}
dI(\hat{u}', \hat{v}') \quad &= d\sigma \iiiint_{-\infty}^{\infty} \imath_1 \tau(\hat{u}_1, \hat{v}_1) h(\hat{u}' - \hat{u}_1, \hat{v}' - \hat{v}_1) \imath_2^* \\
&\times \tau^*(\hat{u}_2, \hat{v}_2) h^*(\hat{u}' - \hat{u}_2, \hat{v}' - \hat{v}_2) d\hat{u}_1 d\hat{v}_1 d\hat{u}_2 d\hat{v}_2 \quad \text{(B.6)}
\end{aligned}
$$

Integrating over the whole source, and noting the definition B.1, the total intensity in the image plane O' turns out to be:

$$
\begin{aligned}
I(\hat{u}', \hat{v}') \quad &= \iiiint_{-\infty}^{\infty} \Gamma(\hat{u}_1 - \hat{u}_2, \hat{v}_1 - \hat{v}_2) \tau(\hat{u}_1, \hat{v}_1) h(\hat{u}' - \hat{u}_1, \hat{v}' - \hat{v}_1) \\
&\times \tau^*(\hat{u}_2, \hat{v}_2) h^*(\hat{u}' - \hat{u}_2, \hat{v}' - \hat{v}_2) d\hat{u}_1 d\hat{v}_1 d\hat{u}_2 d\hat{v}_2 \quad \text{(B.7)}
\end{aligned}
$$

The transmission function $\tau(\hat{u}, \hat{v})$ of the object can be replaced by its Fourier series:

$$
\tau(\hat{u}, \hat{v}) = \frac{1}{2\pi} \sum_{m,n} a_{m,n} e^{2\pi i(m\hat{u} + n\hat{v})} \quad \text{(B.8)}
$$

Combining now eq. B.7 and eq. B.8 and changing then the variables to $\tilde{u}_1 = \hat{u}' - \hat{u}_1$, $\tilde{u}_2 = \hat{u}_2 - \hat{u}'$ results in:

$$
I(\hat{u}', \hat{v}') = \left(\frac{1}{2\pi}\right)^2 \sum_{m,n,p,q} C_{m,n,p,q} a_{m,n} a_{p,q}^* e^{2\pi i\{(m+p)\hat{u}' + (n+q)\hat{v}'\}} \quad \text{(B.9)}
$$

where

$$
\begin{aligned}
C_{m,n,p,q} \quad &= \iiiint_{-\infty}^{\infty} \Gamma(-\tilde{u}_1 - \tilde{u}_2, -\tilde{v}_1 - \tilde{v}_2) e^{-2\pi i(m\tilde{u}_1 + n\tilde{v}_1)} h(\tilde{u}_1, \tilde{v}_1) \\
&\times e^{2\pi i(p\tilde{u}_2 + q\tilde{v}_2)} h^*(-\tilde{u}_2, -\tilde{v}_2) d\tilde{u}_1 d\tilde{u}_2 d\tilde{v}_1 d\tilde{v}_2 \quad \text{(B.10)}
\end{aligned}
$$

These equations can be further simplified by several mathematical manipulations, considering eq. B.2 as well as the fact that $h(\hat{u}, \hat{v})$ and $H(\hat{x}, \hat{y})$ are Fourier transforms to each other:

$$
I(\hat{u}', \hat{v}') = \left(\frac{1}{2\pi}\right)^2 \sum_{m,n,p,q} C_{m,n,q,p} a_{m,n} a_{p,q}^* e^{2\pi i\{(m-p)\hat{u}' + (n-q)\hat{v}'\}}
$$

$$
\text{(B.11)}
$$

with

$$
C_{m,n,p,q} = 2\pi \iint_{-\infty}^{\infty} \gamma(\hat{x}, \hat{y}) H(\hat{x} + 2\pi m, \hat{y} + 2\pi n) H^*(\hat{x} + 2\pi p, \hat{y} + 2\pi q) d\hat{x} d\hat{y}
$$

Re-transforming to the original geometrical coordinates (u, v) and (u', v') in the object and image space leads to the final result:

$$I(u', v') = \left(\frac{1}{2\pi}\right)^2 \sum_{m,n,p,q} C_{m,n,q,p} a_{m,n} a^*_{p,q} e^{2\pi i \frac{NA'}{NA}\{(m-p)u' + (n-q)v'\}}$$

(B.12)

with

$$C_{m,n,p,q} = 2\pi \iint\limits_{-\infty}^{\infty} \gamma(\hat{x}, \hat{y}) H(\hat{x} + \beta m, \hat{y} + \beta n) H^*(\hat{x} + \beta p, \hat{y} + \beta q) d\hat{x} d\hat{y}$$

where $\beta = \lambda/NA$. Thus, a general solution of the optical problem is given by the eq. B.12.

B.4. Definition of coherence level σ

For deriving the notation σ, let consider a one-dimensional line structure in u-direction so that $a(m, n) = 0$ ($\forall n \neq 0$), and $a^*(p, q) = 0$ ($\forall q \neq 0$). Let further the object be coherently illuminated by a point source placed in the source plane E_0 at $(\hat{x}_0, 0) = (\frac{x_0}{h_0}, 0)$ (see Fig. B.1). Note that $\hat{x}_0$ is normalised so that $\hat{x}_0 \leq 1$. Then, eq. B.11 turns to be

$$C_{m,p} = H(\hat{x}_0 + 2\pi m, 0) H^*(\hat{x}_0 + 2\pi p, 0)$$

Thus, $C_{m,p} = 0$ when either of the conditions $|\hat{x}_0 + 2\pi m| < 1$, $|\hat{x}_0 + 2\pi p| < 1$ is not satisfied since $H(\hat{x}, \hat{y})$ equals 0 outside of the unit circle. The maximum values of $|m|, |p|$ which satisfy these conditions are equal to $\frac{1}{2\pi}(1 + |\hat{x}_0|)$. Since, according to eq. B.8, the length of period of the frequency m is given by $m\hat{u} = 1$, $\hat{u}$ being defined by eq. B.3, the limit of resolution occurs when

$$u_{min} = \frac{\lambda}{(1 + \sigma)NA} \qquad \text{for } \sigma < 1 \qquad (B.13)$$

where $\sigma = |\hat{x}_0|$ is the position of the point source normalised by h_0 as defined in section B.2. Thus, σ is equal to the ratio of the numerical apertures of the condenser and the optical system:

$$\sigma = |\hat{x}_0| = \left|\frac{x_0}{h_0}\right| = \frac{N_0 \sin \alpha_0}{N \sin \alpha} = \frac{NA_0}{NA} \qquad (B.14)$$

where α_0 is the angular distance of the effective point source from the axis and N_0 is the refractive index of the space between the source and the object plane (see Fig.B.1).

The relation B.14 holds also for $\sigma > 1$. Then, however, the resolution is limited by

$$u_{min} = \frac{\lambda}{2NA} \qquad \text{for } \sigma > 1 \tag{B.15}$$

as well known from Abbé's formula.

C

Definitions

Amplitude spread function

The amplitude spread function (ASF) is the field amplitude distribution about the image of a point source. The Fourier transform of the ASF corresponds to the coherent transfer function (CTF), while the squared magnitude of the ASF is the point spread function (PSF) [114].

Apodisation

The use of a variable transmission filter at the aperture stop of a optical system to modify its diffraction pattern. Reduced transmission at the edge to block higher spatial frequencies is a common method to extend the depth of focus of an optical system [115].

Augmented reality

Augmented reality (AR) allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. Thus, AR supplements reality, rather than completely replacing it. Ideally, it would appear to the user that the virtual and real objects coexisted in the same space [1].

Binocular rivalry

The alternation (or even suppression) of a percept over time between one eye's view and the other eye's view when the two eyes view very different imagery [7].

Coherence

Coherence is a property of waves that measures the ability of the waves to interfere with each other. Two waves that are coherent can be combined to produce an unmoving distribution of constructive and destructive interference (a visible interference pattern) depending on the relative phase of the waves at their meeting point [116].
Spatial coherence describes the correlation between waves at different points in space. It is related to the size of the light source.
Temporal coherence describes the correlation or predictable relationship between waves observed at different moments in time. The temporal coherence of a wave is related to the spectral bandwidth of the source.

Coherent transfer function

The coherent transfer function (CTF) is the transfer function of a coherent imaging system. The CTF measures the loss in amplitude in the image of a sinusoidal target as well as any phase shifts. The CTF is the Fourier transform of the amplitude spread function ASF and has the same shape as the system's exit pupil function $P(x, y)$ [114].

Critical fusion frequency

The critical fusion frequency (CFF) is the frequency that is needed just to produce complete image fusion so that all flicker of an intermittent light stimulus disappears. The CFF is usually measured in Hertz (Hz) [115].

Depth of focus

The depth of focus (DOF) signifies the distance in terms of refractive power value over which an image is perceived as 'in focus' or sharply imaged. Mostly, the depth of focus is measured in diopters.

Diopters

A unit of optical measurement that expresses the refractive power of a lens system or the human eye. In a lens system, the dioptric value signifies the reciprocal of the focal length in meters [115]. In this thesis, we consider only dioptric changes ΔD in the ocular refractive power, relative to the unaccommodated eye viewing at infinity. Thus, in this context, the ΔD-values correspond to the reciprocal of the eye's viewing distance. (e.g. $\Delta D = 2\,\mathrm{D} \cong$ viewing distance of $50\,\mathrm{cm}$)

Emmetropic eye

The normal condition of vision where an object at infinity is in sharp focus with the eye lens in a unaccommodated or relaxed state [115]. Any accommodation values ΔD are measured relative to this state. Thus, the emmetropic eye has an accommodation value $\Delta D = 0\,\mathrm{D}$.

Exit pupil

In a lens or other optical system, the exit pupil is the image of the aperture stop as seen from image space. The **exit plane** is the plane where the exit pupil is located [115].

Eye motion box

The eye motion box defines the area where the eye can move laterally while full FOV is visible and no vignetting is taking place.

Eye relief

The eye relief is the distance between the vertex of the last optical surface of a visual system and the system's exit pupil [115]. In retinal projection display, the exit pupil plane coincides normally with the eye pupil plane.

Field of view

The field of view (FOV) is the maximum angular subtense of the retinal image as provided from an HMD. The FOV is expressed in degrees or radiant.

Illuminance

Illuminance is the luminous flux per unit area incident on a surface and is measured in lux (=lumen/m^2)[117]. The illuminance characterises an illuminated surface (in contrast to the luminance, which characterises the emitting surface). It is the photometric counterpart to the radiometric irradiance.

Irradiance

Irradiance is a measure of radiometric flux per unit area, or flux density. Irradiance is typically expressed in W/m^2 and characterises an illuminated surface [118].

Köhler illumination

A method of illuminating the object in a projection imaging system whereby a condenser lens forms an image of the illumination source at the entrance pupil of the objective lens, and the mask is at the exit pupil of the condenser lens. The Köhler illumination is widely used in projection lithography due to its superior uniformity [50].

Luminance

Luminance is the luminous flux emitted from a surface per unit solid angle and per unit of area (projected on a plane normal to the direction of propagation). It is most often used to characterise the perceived brightness of flat emitting or reflecting surfaces and is measured in candela per square meter (cd/m^2 = lm/(m^{2}sr)) [117].

Luminous flux

The luminous flux is the part of the power radiated from a source over visible wavelengths (from about 330 nm to 780 nm). Since the human eye does not respond equally to all visible wavelengths, the luminous flux is a weighted average of the radiant power according to the sensitivity of the human eye. Luminous flux is expressed in lumen (lm) [117].

Maxwellian-view optical system

In a Maxwellian-view optical system, the illumination source is made optically conjugate to the pupil of the eye. As a result, the object (e.g.

the LCD) is not placed at the source, but rather at a separate plane optically conjugate to the retina [21].

Numerical aperture *NA*

The measure of the light-gathering power of an optical system. It is equal to the sine of half the angle subtended by the aperture at an object point times the index of refraction of the medium between the object and the objective lens [119].

Ocular accommodation

The variation of the total refracting power of the human eye that allows an observer to clearly see objects at different distances. This process is chiefly effected by changes in the convexity of the crystalline lens. The limits of accommodation are the distances of the nearest and farthest points, usually about 20 cm to infinity, that can be focused clearly by the eyes of an observer. There is a reduction in focusing ability with age (presbyopia) [115, 120].

Optical transfer function

The optical transfer function (OTF) is the transfer function of an incoherent imaging system. It measures the loss in contrast in the image of a sinusoidal target, as well as any phase shifts. The OTF is the Fourier transform of the point-spread function (PSF). The absolute value of the optical transfer function corresponds to the modulation transfer function (MTF) (i.e. MTF = |OTF|) [114].

Optically conjugate planes

Two planes in an optical system so positioned that light emitted from any point of the either plane will be focused onto the other [115].

Point spread function

The point spread function (PSF) is the intensity distribution about the image of a point light source [115]. The PSF corresponds to the squared magnitude of the amplitude spread function (ASF) and is the Fourier transform of the OTF.

Radiance

Radiance is a measure of the flux emitted from a surface per unit solid angle and per unit of area. It is expressed in $W/(m^2sr)$ and is the radiometric counterpart to the luminance [118].

Rayleigh criterion

A widely used criterion for characterising in-focus, unaberrated imaging. It states that any wavefront errors should be restricted to within a quarter of a wavelength of a true spherical surface [21].

Relay lens

A lens or lens system used to transfer a real image from one point within an optical system to another, with or without magnifying it [115].

See-through mode

Additional mode in head-mounted displays which allows to combine the real world with virtual objects superimposed by optical or video technologies. The optical see-through mode is normally achieved by an optical combiner such as half mirrors or diffractive optical elements [7].

Vergence

Coordinated turning inward (convergence) and turning outward (divergence) of the eyes to view near or far objects. If the eyes do not converge and diverge correctly, double vision or the suppression of one eye or the other will occur [7].

Glossary

Symbols

A_{pup}	aperture size [m^2]
$a_{m,n}$	Fourier coefficient of object transmission function $\tau(u,v)$
B	size of sliding window in image quality index Q
C	contrast comparison function
C_B	correction factor for photochemical limit
$C_{m,n,p,q}$	transmission cross coefficient
c	Michelson contrast of object
c'	Michelson contrast of image
$\mathcal{D}(t)$	periodic dioptric function of lens' oscillation [D]
$\mathcal{D}_0$	dioptric offset of oscillating lens [D]
$\mathcal{D}_j$	dioptric power of oscillating lens at time t_j [D]
d_2	distance from aperture stop plane to lens L2 [m]
d_{con}	distance from the condenser's diaphragm to the second condenser lens [m]
d_{eye}	eye relief [m]
d_{LCD}	distance from LCD to objective lens [m]
d_{ret}	distance from eye pupil plane to retina [m]
EL_{ph}	photochemical exposure limit [W/m^{2}sr]
EL_{th}	thermal exposure limit [W/m^{2}sr and W/deg^2]
F_s	font size [points]
f_i	focal length of lens i [m]
$g(x,y)/g(t)$	intensity function of scanning beam
$H(\nu_u, \nu_v)$	coherent transfer function (CTF)
$h(u',v';u,v)$	amplitude spread function (ASF)
I	irradiance or power density [W/m^2]

I_{fused}	fused irradiance [W/m^2]
I_{ret}	retinal irradiance [W/m^2]
$\imath_{\text{ret}}$	retinal field amplitude [W$^{0.5}$/m]
k	wave vector [m^{-1}]
L	number of focal planes
M	magnification of the relay lens system
N	refractive index of object space
N'	refractive index of image space
N_0	refractive index of space between source and object plane
NA	numerical aperture of object plane
NA'	numerical aperture of image plane
NA_0	numerical aperture of condenser
$P1$	aperture stop plane
$P2$	exit pupil plane or eye pupil plane
$P(x,y)$	exit pupil function
$\mathcal{P}(x,y)$	complex generalised pupil function
$PSF(x,y)$	point spread function
Q	image quality index
$R(\lambda)$	retinal hazard function
(r_x, r_y)	radius of exit pupil [m]
S	structural function
s_X, s_Y	variance of images X and Y, respectively
s_{XY}	cross correlation function of images X and Y
T	exposure time [s]
(u, v)	geometrical coordinates in the LCD or object plane [m]
(u', v')	geometrical coordinates in the retinal or image plane [m]
$(\hat{u}, \hat{v})$	reduced coordinates in the LCD or object plane
$(\hat{u}', \hat{v}')$	reduced coordinates in the retinal or image plane
$(\tilde{u}, \tilde{v})$	retinal position of scanning beam [m]

u_{min}	limit of resolution in u-direction [m]
$V(u', v')$	scanning velocity of beam on the retina [m/s]
w_l	real weight of focal plane l
$X_{m,n}, Y_{m,n}$	value of (m,n)-pixel in image X and Y, respectively
$\overline{X}, \overline{Y}$	mean value of image X and Y, respectively
(x, y)	geometrical coordinates in the exit pupil plane [m]
$(\hat{x}, \hat{y})$	relative coordinates in the exit pupil plane
α	angle of eye rotation [rad or deg]
$\hat{\alpha}$	angular subtense of source [rad or deg]
Γ	phase-coherence factor
$\gamma(x, y)$	source map
$\hat{\gamma}(x, y)$	cone angle [sr]
ΔC	difference in contrast comparison function C
ΔD	dioptric eye accommodation [D]
$\Delta \mathcal{D}$	spacing between discrete $\mathcal{D}_j$-values [D]
δ_l	focal value of focal plane l [D]
$\overline{\delta}$	mean value of focal planes [D]
$\Delta \delta$	spacing between focal planes [D]
δd_{LCD}	shift in LCD position [m]
ΔS	difference in structural function S
Δu	width of horizontal bar pattern [m]
$\delta \tilde{v}$	discrete scanning step in v-direction [m]
ϵ	dioptric extent of focal planes [D]
ζ	viewing distance [m]
Θ	periodic time of oscillation function $\mathcal{D}(t)$ [s]
ϑ	scanning time for one frame [s]
λ	wavelength [m]
ν	spatial frequency [cyc/m or cyc/deg]
(ν_u, ν_v)	spatial frequency in u- and v-direction [cyc/m or cyc/deg]

ν^{cut}	cut-off frequency [cyc/m or cyc/deg]
ξ	dioptric amplitude of lens' oscillation [D]
ρ	radial exit pupil coordinate ($\rho^2 = x^2 + y^2$) [m]
ρ_{cor}	corneal beam radius [m]
σ	spatial coherence level
$\tau(u, v)$	complex transmission function of the object or LCD
ϕ_l	phase shift of focal plane l
ψ	wavefront phase error
ψ_{disp}	wavefront phase error caused by the display
ψ_{eye}	wavefront phase error caused by ocular aberrations
$\varnothing_i$	diameter of diaphragm i [m]
$\varnothing_{\text{pup}}$	diameter of eye pupil [m]

Abbreviations

AMLCD	Active matrix LCD
AR	Augmented reality
ASF	Amplitude spread function
CCD	Charge-coupled device
CF	Contrast function
CFF	Critical fusion frequency
CMI	Coherent multiple imaging
CR	Centre of rotation
CRT	Cathode ray tube

CTF	Coherent transfer function
cyc	cycles
deg	degree
D	diopter
Di	Diaphragm
DMD	Digital micromirror device
DOE	Diffractive optical element
DOF	Depth of focus
EL	Exposure limit
$\mathcal{F}(.)$	Fourier transformation
FOV	Field of view
HMD	Head mounted display
HOE	Holographic optical element
ICNIRP	International Commission on non-ionising radiation protection
IRA	Near-infrared light range
LCD	Liquid crystal display
LED	Light emitting diode
MTF	Modulation transfer function
NA	Numerical aperture
OLED	Organic light emitting device
OTF	Optical transfer function
P	Plane
PDA	Personal digital assistant
PEST	Parameter estimation by sequential testing
PM	Phase-only mask
PSF	Point spread function
pt	points
RSD	Retinal scanning display
SXGA	Super Extended Graphics Array (resolution of 1280×1024 pixels)
TV	Television

UV	Ultraviolet light range
VGA	video graphics array (resolution of 640×480 pixels)
VIS	Visible light range
VRD	Virtual retinal display

Bibliography

[1] R. Azuma, "A survey of augmented reality," Presence: 'Teleoperators and Virtual Environments' **6**(4), 355–385 (1997).

[2] S. Feiner, "Augmented reality: A new way of seeing," Scientific American **286**(4), 48–55 (2002).

[3] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, "Recent advances in augmented reality," IEEE Computer Graphics and Applications **21**(6), 34–47 (2001) and references therein.

[4] H. Hua, C. Gao, L. Brown, N. Ahuja, and J. Rolland, "Using a head-mounted projective display in interactive augmented environments," in *Proc. IEEE and ACM Int. Symp. on Augmented Reality*, pp. 217–223 (2001).

[5] D. Drascic and P. Milgram, "Perceptual issues in augmented reality," in *Proc. SPIE, 'Stereoscopic Displays and Virtual Reality Systems III'*, vol. 2653, pp. 123–134 (1996).

[6] I. Sutherland, "A head-mounted three-dimensional display," in *Proc. Fall Joint Computer Conference, AFIPS Conf. Proc.*, vol. 33, pp. 757–764 (1968).

[7] J. Melzer and K. Moffitt, *Head-mounted displays: Designing for the user* (McGraw Hill, Inc., New York, 1997).

[8] S. Agyare, S. Darlington, and C. Saunders, "Helmet-mounted display CRT developments," in *SID International Symposium (SID'98), Digest of Technical Papers*, vol. XXIX, pp. 65–69 (San Jose, CA, 1998).

[9] J. Leyland, F. Walters, and D. Etherington, "Developments in CRT's for HMD applications," in *Proc. SPIE, 'Helmet- and Head-Mounted Displays II'*, vol. 1290, pp. 30–40 (1990).

[10] L. Hornbeck, "128 x 128 deformable mirror device," IEEE Trans. on Electron Devices **ED-30**(5), 539–545 (1983).

[11] J. Younse, "Mirrors on a chip," IEEE Spectrum **30**(11), 27–31 (1993).

[12] R. Leinenwever, L. Best, and B. Ericksen, "Low-cost color LCD helmet display," in *Proc. SPIE, 'Helmet-Mounted Displays III'*, vol. 1695, pp. 68–71 (1992).

[13] J. Rabin and R. Wiley, "Dynamic visual performance comparison between helmet-mounted CRTs and LCDs," Journal of the Society for Information Display **3**(3), 97–100 (1995).

[14] S. Lee, J. Song, J. Lee, T. Park, and M. Park, "Head mounted display with a single LCD panel," in *Proc. of 18th Internat. Display Research Conference (Asia Display'98)*, pp. 471–474 (1998).

[15] T. Holzel, "Are head-mounted displays going anywhere?" Information Display **15**(10), 16–18 (1999).

[16] Homepage of Kopin Corp., CyberDisplay$^{\text{TM}}$(2005); http://www.kopin.com/products/index_cyberdisplay.html.

[17] J. Sheats, H. Antoniadis, M. Hueschen, W. Leonard, J. Miller, R. Moon, D. Roitman, and A. Stocking, "Organic electroluminescent devices," Science **273**, 884–888 (1996).

[18] Z. Shen, P. Burrows, V. Bulovic, S. Forrest, and M. Thompson, "Three-color, tunable, organic light-emitting devices," Science **276**, 2009–2011 (1997).

[19] O. Prache, "Active matrix molecular OLED microdisplays," Displays **22**(2), 49–56 (2001).

[20] A. Ghosh, "Full color OLED on silicon microdisplay," in *Proc. SPIE, 'Organic Light-Emitting Materials and Devices V'*, vol. 4464, pp. 1–10 (2002).

[21] M. Bass, ed., *Handbook of optics*, vol. 1 (McGraw-Hill, Inc., New York, 1995).

[22] R. Webb, G. Hughes, and O. Pomerantzeff, "Flying spot TV ophthalmoscope ," Appl. Opt. **19**(17), 2991–2997 (1980).

[23] R. Webb and G. W. Hughes, "Scanning laser ophthalmoscope," IEEE Trans. on Biomedical Engineering **28**, 488–492 (1981).

[24] K. Ashizaki, M. Yamamoto, S. Miyaoka, and S. Tamada, *Direct viewing picture image display apparatus*, European Patent 0,473,343 (1992).

[25] J. Kollin, "A retinal display for virtual-environment applications," in *SID International Symposium (SID'93), Digest of Technical Papers*, vol. XXIV, pp. 827–830 (Playa del Rey, CA, 1993).

[26] J. Kollin and M. Tidwell, "Optical engineering challenges of the virtual retinal display," in *Proc. SPIE. 'Novel Optical Systems Design and Optimization'*, vol. 2537, pp. 48–60 (1995).

[27] R. Pausch, P. Dwivedi, and J. A.C. Long, "A practical, low-cost stereo head-mounted display," in *Proc. SPIE, 'Stereoscopic Displays and Applications II'*, vol. 1457, pp. 198–208 (1991).

[28] N. Fruehauf, T. Aye, K. Yua, Y. Zou, and G. Savant, "Liquid crystal digital scanner-based HMD," in *Proc. SPIE, 'Helmet- and Head-Mounted Displays V'*, vol. 4021, pp. 2–10 (2000).

[29] T. Aye, K. Yu, A. Kostrzewski, G. Savant, and J. Jannson, "Line-image-scanning head mounted display," in *SID International Symposium (SID'98), Digest of Technical Papers*, vol. XXIX, pp. 564–567 (San Jose, CA, 1998).

[30] M. Spitzer, P. Aquilino, R. McClelland, M. Olson, N. Rensing, C. DiMarzio, P. Zavracky, A. Lemoncelli, and J. Hilliar, "Optical approaches to incorporation of displays within eyeglasses," in *SID International Symposium (SID'98), Digest of Technical Papers*, vol. XXIX, pp. 1007–1009 (San Jose, CA, 1998).

[31] R. Laramee and C. Ware, "Rivalry and interference with a head-mounted display," ACM Trans. on Computer-Human Interaction **9**(3), 238–251 (2002).

[32] J. Wann, S. Rushton, and M. Mon-Williams, "Natural problems for stereoscopic depth perception in virtual environments," Vision Res. **35**(19), 2731–2736 (1995).

[33] T. Ando, K. Yamasaki, M. Okamoto, and E. Shimizu, "Head mounted display using holographic optical element," in *Proc. SPIE, 'Practical Holography XII'*, vol. 3293, pp. 183–189 (1998).

[34] A. Takagi, S. Yamazaki, Y. Saito, and N. Taniguchi, "Development of a stereo video see-through HMD for AR systems," in *Proc. of IEEE/ACM International Symposium on Augmented Reality (ISAR 2000)*, pp. 68–77 (2000).

[35] E. Peli, "Head-mounted display as a low vision aid," in *Proc. of the 2nd Annual International Conference on Virtual Reality and Persons with Disabilities*, pp. 115–122 (Northridge, CA., 1994).

[36] M. Everingham, B. Thomas, and T. Troscianko, "Head-mounted mobility aid for low vision using scene classification techniques," International Journal of Virtual Reality **3**(4), 3–12 (1999).

[37] F. Vargas-Martin and E. Peli, "Augmented view for tunnel vision: Device testing by patients in real environments," in *SID International Symposium (SID'01), Digest of Technical Papers*, pp. 602–605 (2001).

[38] S.V. Lin and E.J. Seibel and T.A. Furness III, "Testing visual search performance using retinal light scanning as a future wearable low vision aid," International Journal of Human-Computer Interaction **15**(3), 245–263 (2003).

[39] G. Edgar, J. Pope, and I. Craig, "Visual accommodation problems with head-up and helmet-mounted displays ?" Displays **15**(2), 68–75 (1994).

[40] M. Spitzer, N. Rensing, R. McClelland, and P. Aquilino, "Eyeglass-based systems for wearable computing," in *Proc. of IEEE International Symposium on Wearable Computers (ISWC'97)*, pp. 48–51 (Cambridge, MA, 1997).

[41] I. Kasai, Y. Tanijiri, T. Endo, and H. Ueda, "A forgettable near eye display," in *Proc. of IEEE International Symposium on Wearable Computers (ISWC'00)*, pp. 115–118 (Atlanta, GA, 2000).

[42] J. Ojeda-Castaneda and L. B. Valdós, "Arbitrarily high focal depth with finite apertures," Opt. Lett. **13**, 183–185 (1988).

[43] J. Rolland, M. Krueger, and A. Goon, "Multifocal planes head-mounted displays," Appl. Opt. **39**(19), 3209–3215 (2000).

[44] L. Marran and C. Schor, "Multiaccomodative stimuli in VR systems: Problems & solutions," Human Factors **39**(3), 382–388 (1997).

[45] T. Sugihara and T. Miyasato, "A lightweight 3-D HMD with accommodative compensation," in *SID International Symposium (SID'98), Digest of Technical Papers*, vol. XXIX, pp. 927–930 (San Jose, CA, 1998).

[46] M. von Waldkirch, P. Lukowicz, and G. Tröster, "LCD-based coherent wearable projection display for quasi accommodation-free imaging," Opt. Commun. **217**, 133–140 (2003).

[47] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Impact of light coherence on depth of focus of wearable retinal displays," in *Proc. SPIE, 'Design of Efficient Illumination Systems'*, vol. 5186, pp. 5–14 (2003).

[48] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Effect of light coherence on depth of focus in head-mounted retinal projection displays," Opt. Eng. **43**(7), 1552–1560 (2004).

[49] G. Westheimer, "The maxwellian view," Vision Res. **6**, 669–682 (1966).

[50] S. Inoué and K. Spring, *Video microscopy: the fundamentals*, chapter 'Microscope image formation - Principles of Köhler illumination', pp. 22–26, 2nd ed. (Plenum Press, New York, 1997).

[51] N. Beaudry and T. Milster, "Effects of object roughness on partially coherent image formation," Opt. Lett. **25**(7), 454–456 (2000).

[52] M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1980).

[53] H. Hopkins, "On the diffraction theory of optical images," Proc. of the Royal Society of London. Series A, Mathematical and Physical Sciences **217**(1130), 408–432 (1953).

[54] J. Goodman, *Introduction to Fourier optics* (McGraw-Hill, New York, 1996).

[55] J. Gaskill, *Linear systems, Fourier transforms, and optics* (John Wiley & Sons, New York, 1978).

[56] I. Bronstein, K. Semendjajew, G. Musiol, and H. Mühlig, *Taschenbuch der Mathematik* (Verlag Harri Deutsch, Thun, 2001).

[57] J. Goodman, "Some fundamental properties of speckle," J. Opt. Soc. Am. **66**(11), 1145–1149 (1976).

[58] J. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Topics*, vol. 9 of *Topics in Applied Physics*, pp. 9–75 (Springer-Verlag, Berlin Heidelberg, 1984).

[59] A. Felipe, J. Artigas, A. Pons, and M. Buades, "Human contrast sensitivity in coherent Maxwellian view: effect of coherent noise and comparison with speckle," J. Opt. Soc. Am. A **14**(5), 972–983 (1997).

[60] J. Artigas, A. Felipe, and M. Buades, "Contrast sensitivity of the visual system in speckle imagery," J. Opt. Soc. Am. A **11**(9), 2345–2349 (1994).

[61] J. Dainty, ed., *Laser Speckle and Related Phenomena* (Springer-Verlag, Berlin, 1975).

[62] S. Lowenthal and H. Arsenault, "Image formation for coherent diffuse objects: Statistical properties," J. Opt. Soc. Am. A **60**(11), 1478–1483 (1970).

[63] P. Considine, "Effects of coherence on imaging systems," J. Opt. Soc. Am. A **56**(8), 1001–1009 (1966).

[64] L. Gruber and B. Thompson, "On the apodization of coherent imaging systems," Opt. Eng. **13**(5), 451–454 (1974).

[65] H. Ong and R. Gale, "Small displays have a big future," Information-Display **14**(12), 18–22 (1998).

[66] G. de Wit, "A retinal scanning display for virtual reality," Ph.D. thesis, TU Delft (1997).

[67] "Drucktechnik - Schriftgrösse - Teil 2: Digitaler Satz und verwandte Techniken," DIN 16507-2 (1999).

[68] Z. Wang and A. Bovik, "A universal image quality index," IEEE Signal Processing Letters **9**(3), 81–84 (2002).

[69] Z. Wang and A. Bovik, "Why is image quality assessment so difficult ?" in *Proc. of IEEE International Conference on Acoustics, Speech, & Signal Processing (ICASSP '02)*, vol. 4, pp. 3313–3316 (2002).

[70] G. Gescheider, *Psychophysics, the fundamentals* (Erlbaum Inc., NJ, 1997).

[71] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Multiple imaging technique for extending depth of focus in retinal displays," Opt. Express **12**(25), 6350–6365 (2004).

[72] M. Erdélyi, Z. Bor, W. Wilson, M. Smayling, and F. Tittel, "Simulation of coherent multiple imaging by means of pupil-plane filtering in optical microlithography," J. Opt. Soc. Am. A **16**(8), 1909–1914 (1999).

[73] Z. Horváth, M. Erdélyi, G. Szabó, Z. Bor, F. Tittel, and J. Cavallaro, "Generation of nearly nondiffracting Bessel beams with a Fabry Perot interferometer," J. Opt. Soc. Am. A **14**(11), 3009–3013 (1997).

[74] M. Erdélyi, Z. Horváth, G. Szabó, Z. Bor, F. Tittel, J. Cavallaro, and M. Smayling, "Generation of diffraction-free beams for applications in optical microlithography," J. Vac. Sci. Technol. B **15**, 287–292 (1997).

[75] H. Hopkins, "The concept of partial coherence in optics," Proc. of the Royal Society of London. Series A, Mathematical and Physical Sciences **208**(1093), 263–277 (1951).

[76] R. Navarro, J. Santamaría, and J. Bescós, "Accommodation-dependent model of the human eye with aspherics," J. Opt. Soc. Am. A **2**(8), 1273–1281 (1985).

[77] I. Escudero-Sanz and R. Navarro, "Off-axis aberrations of a wide-angle schematic eye model," J. Opt. Soc. Am. A **16**(8), 1881–1891 (1999).

[78] OSLO is a registered trademark of Lambda Research Corp.

[79] K. Toh and A. Neureuther, "Identifying and monitoring effects of lens aberrations in projection printing," in *Proc. SPIE, 'Optical Microlithography VI'*, vol. 772, pp. 202–209 (1987).

[80] G. C. de Wit and R. Beek, "Effects of a small exit pupil in a virtual reality display system," Opt. Eng. **36**(8), 2158–2162 (1997).

[81] R. Becherer and G. Parrent, "Nonlinearity in optical imaging systems," J. Opt Soc. **57**(12), 1479–1486 (1967).

[82] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. on Image Processing **13**(4), 600–612 (2004).

[83] H. Lieberman and A. Pentland, "Microcomputer-based estimation of psychophysical thresholds: the best PEST," Behaviour Research Methods & Instrumentation **14**(1), 21–25 (1982).

[84] C. Kaernbach, "Adaptive threshold estimation with unforced-choice tasks," Perception & Psychophysics **63**(8), 1377–1388 (2001).

[85] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Oscillating fluid lens in coherent retinal projection display for extending depth of focus," submitted to Opt. Commun. (2005).

[86] S. Kuiper and B. Hendriks, "Variable-focus liquid lens for miniature cameras," Appl. Phys. Lett. **85**(7), 1128–1130 (2004).

[87] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Spectacle-based display design for accommodation-free viewing," in *Proc. of 2nd International Conference on Pervasive Computing (Pervasive 2004)*, LNCS Vol. 3001, pp. 106–123 (Springer-Verlag, 2004).

[88] R. Fischer, *Optical System Design*, SPIE press monograph; PM87 (McGraw-Hill, Inc., New York, 2000).

[89] W. Smith, *Modern Lens Design: A resource manual* (McGraw-Hill, Inc., New York, 1992).

[90] T. Ando, T. Matsumoto, H.Takahasihi, and E. Shimizu, "Head mounted display for mixed reality using holographic optical elements," in *Memoirs of the Faculty of Engineering Osaka City University*, vol. 40, pp. 1–7 (1999).

[91] A. Ashmead, M. Popovich, J. Clark, S. Sagan, and R. Smith, "Application specific integrated lenses (ASILs) for the next generation of wearable displays," in *Proc. SPIE, 'Helmet- and Head-Mounted Displays V'*, vol. 4021, pp. 180–186 (2000).

[92] International Commission on Non-Ionizing Radiation Protection (ICNIRP), "Guidelines on limits of exposure to laser radiation of wavelengths between 180 nm and 1 mm," Health Physics **71**(5), 804–819 (1996).

[93] International Commission on Non-Ionizing Radiation Protection (ICNIRP), "Revision of guidelines on limits of exposure to laser radiation of wavelengths between 400 nm and 1.4 μm," Health Physics **79**(4), 431–440 (2000).

[94] International Commission on Non-Ionizing Radiation Protection (ICNIRP), "Guidelines on limits of exposure to broad-band incoherent optical radiation (0.38 to 3 μm)," Health Physics **73**(3), 539–554 (1997).

[95] D. Sliney and M. Wolbarsht, *Safety with lasers and other optical sources* (Plenum Press, New York, 1980).

[96] K. Schulmeister, "Concepts in dosimetry related to laser safety and optical radiation hazard evaluation," in *Proc. SPIE, 'Laser and Noncoherent Light Ocular Effects'*, vol. 4246, pp. 104–116 (2001).

[97] K. Schulmeister, *Handbook on industrial laser safety*, chapter 1: 'Radiation' (2000).

[98] W. Howard and O. Prache, "Microdisplays based upon organic light-emitting diodes," IBM J. Res. & Dev. **45**, 115–127 (2001).

[99] M. Menozzi, H. Krüger, P. Lukowicz, and G. Tröster, "Perception of brightness with a virtual retinal display using badal projection," Biomed. Technik **46**, 55–62 (2001).

[100] M. von Waldkirch, P. Lukowicz, and G. Tröster, "Defocusing simulations on a retinal scanning display for quasi accommodation-free viewing," Opt. Express **11**(24), 3220–3233 (2003).

[101] B. Wandell, *Foundations of vision* (Sinauer Associates Inc., Sunderland, MA, 1995).

[102] L. Thibos, D. Still, and A. Bradley, "Characterization of spatial aliasing and contrast sensitivity in peripheral vision," Vision Res. **36**(2), 249–258 (1996) and references therein.

[103] P. Hallett, "Eye movements," in *Handbook of perception and human performance*, vol. 1 (John Wiley & Sons, New York, 1986).

[104] M. Bass, ed., *Handbook of optics*, vol. 1, chapter 24 (McGraw-Hill, Inc., New York, 1995).

[105] G. C. de Wit, "Resolution matching in a retinal scanning display," Appl. Opt. **36**(22), 5587–5593 (1997).

[106] D. Atchison and G. Smith, *Optics of the human eye* (Butterworth-Heinemann, Oxford, 2000).

[107] S. Marcos, E. Moreno, and R. Navarro, "The depth-of-field of the human eye from objective and subjective measurements," Vision Res. **39**, 2039–2049 (1999).

[108] J. Tucker and W. Charman, "Depth of focus and accommodation for sinusoidal gratings as a function of luminance," Am. J. Optom. Physiol. Opt. **63**, 58–70 (1986).

[109] C. Morimoto, D. Koons, A. Amir, and M. Flickner, *Pupil detection and tracking using multiple light sources* (Technical Report RJ-10117, IBM Almaden Research Center, IBM Almaden Research Center, San Jose, CA, 1998).

[110] A. Haro, M. Flicknerz, and I. Essay, "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance," in *Proc. of Conference on Computer Vision and Pattern Recognition (CVPR'00)*, pp. 1163–1169 (Hilton Head, 2000).

[111] A. Clarke, C. Steineke, and H. Emanuel, "High image rate eye movement measurement: A novel approach using CMOS sensors and dedicated FPGA devices," in *T. Lehmann (Eds.), Bildverarbeitung für die Medizin*, pp. 398–403 (Springer-Verlag, Berlin, 2000).

[112] A. Gullstrand, *Appendix II in 'Handbuch der Physiologischen Optik'* (Voss, Hamburg, 1909).

[113] A. Popiolek-Masajada and H. Kasprzak, "Model of the optical system of the human eye during accommodation ," Ophthal. Physiol. Opt. **22**, 201–208 (2002).

[114] C. Williams and O. Becklund, *Introduction to the optical transfer function* (John Wiley & Sons, New York, 1989).

[115] *Photonics Dictionary, A dictionary of photonics terminology* (Laurin Publishing Co. Inc., 1996-2004). Online available at http://www.photonics.com/dictionary/XQ/ASP/QX/index.htm.

[116] *Encyclopedia of 'TheFreeDictionary.com'* (Farlex Inc., 2004). Online available at http://encyclopedia.thefreedictionary.com.

[117] W. Smith, *Modern optical engineering - The design of optical systems* (McGraw-Hill, Inc., New York, 2000).

[118] A. Ryer, *Light measurement handbook* (International Light Inc., 1997). Online available at http://www.intl-light.com/handbook/registered.html.

[119] *Encyclopdia Britannica Online* (Encyclopdia Britannica, Inc., 2004). Online available at http://www.britannica.com.

[120] *Merriam-Webster's Collegiate Dictionary, 11th Edition* (Merriam-Webster, Inc., Springfield, Mass., 2003).

Acknowledgements

This work would not have been possible without the great support and encouragement of many people and friends:

First of all, I would like to express my sincere gratitude to Professor Dr. Gerhard Tröster for giving me the opportunity to write my Ph.D. thesis at the Wearable Computing Lab. I have enjoyed the open and inspiring research atmosphere at his institute and his great enthusiasm for the vision of wearable computing.

Special thanks go to my associate advisor PD Dr. Marino Menozzi for co-examining my work and for interesting discussions and many valuable inputs he has given me during my Ph.D. dissertation.

My sincere appreciation goes to my supervisor Professor Dr. Paul Lukowicz for uncountable ideas and inspiring support. I am deeply impressed by his inexhaustible enthusiasm and his commitment to the research at our lab.

I would like to thank all members of the 'Tröster group' for their support and friendship and for many interesting (and sometimes controversial discussions) at the coffee corner. In particular, let me thank to those colleagues who spent much time for performing my image quality test runs. Special thanks go to my 'pioneering colleagues' - Urs Anliker, Thomas von Büren, Holger Junker and Mathias Stäger. It was a great pleasure for me to launch the adventure 'Wearable Computing' together with them. I am also grateful to Ruth Zähringer and Evelyne Isenschmid for their administrative support.

Finally, great thanks go to my parents for their continuous encouragement, interest and support during my studies. Last, but not least I am deeply grateful to Franziska Diederichs for her understanding, listening and patience and for just being the way she is.

Zurich, December 2004   Marc von Waldkirch

Curriculum Vitae

Personal Information

Marc von Waldkirch
Born November 16, 1974
Citizen of Schaffhausen SH and Zürich ZH,
Switzerland

Education

1981–1987: Primary school in Küsnacht ZH, Switzerland

1987–1993: Literargymnasium Rämibühl, Zurich, Switzerland
Final degree: Matura Type A

1994–1999: M.Sc. (Dipl. Phys. ETH) in experimental physics
at the Swiss Federal Institute of Technology
(ETH) Zurich, with distinction

2000–2004: Ph.D. student at the Swiss Federal Institute of
Technology (ETH) Zurich

Work

2000: Internship at Centre Suisse d'Electronique et de
Microtechnique (CSEM), Neuchâtel, Switzerland:
Nanotechnology and Nano-Optics Group

2000–2004: Teaching and research assistant, Wearable Com-
puting Lab, ETH Zurich