Multiple imaging technique for extending depth of focus in retinal displays

Marc von Waldkirch, Paul Lukowicz, Gerhard Tröster

Wearable Computing Lab, ETH Zurich, Gloriastrasse 35, CH-8092 Zurich, Switzerland

waldkirch@ife.ee.ethz.ch

Abstract: Limited depth of focus is among the main problems of todays see-through head-mounted displays. In this paper we propose and evaluate a new solution to this problem: the use of the coherent multiple imaging technique in a retinal projection display by incorporating an appropriate phase-only mask. The evaluation is based on a schematic eye model and on the partial coherence simulation tool SPLAT which allows us to calculate the projected retinal images of a text target. Objective image quality criteria demonstrate that this approach is promising provided that partially coherent illumination light is used. In this case, psychometric measurements reveal that the depth of focus for reading text can be extended by a factor of up to 3.2. For fully coherent and incoherent illumination, however, the retinal images suffer from structural and contrast degradation effects, respectively.

OCIS codes: (120.2820) Heads-up displays; (110.4190) Multiple imaging; (110.4980) Partial coherence in imaging; (330.4460) Ophthalmic optics.

References and links

1. 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.
2. G. Edgar, J. Pope, and I. Craig, "Visual accommodation problems with head-up and helmet-mounted displays?," Displays **15**(2), 68–75 (1994).
3. J. Kollin and M. Tidwell, "Optical Engineering Challenges of the Virtual Retinal Display," in *Novel Optical Systems Design and Optimization*, J.M. Sasian, ed., Proc. SPIE **2537**, 48–60 (1995).
4. R. Johnston and S. Willey, "Development of a Commercial Retinal Scanning Display," in *Proc. of Helmet- and Head-Mounted Displays and Symbology Design*, 2–13 (W. Stephens and L.A. Haworth (Eds.), 1995).
5. G. de Wit, "A Retinal Scanning Display for Virtual Reality," Ph.D. thesis, TU Delft (1997).
6. T. Tomono, "Spectacle-type Wearable Display," Opt. Commun. **180**, 205–210 (2000).
7. 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).
8. 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**, 106–123 (Springer-Verlag, 2004).
9. G. Westheimer, "The maxwellian view," Vision Res. **6**, 669–682 (1966).
10. 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), http://www.opticsexpress.org/abstract.cfm?URI=OPEX-11-24-3220
11. J. Rolland, M. Krueger, and A. Goon, "Multifocal planes head-mounted displays," Appl. Opt. **39**(19), 3209–3215 (2000).
12. L. Marran and C. Schor, "Multiaccomodative Stimuli in VR systems: Problems & Solutions," Human Factors **39**(3), 382–388 (1997).

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

13. T. Sugihara and T. Miyasato, "A Lightweight 3-D HMD with Accommodative Compensation," in *Proc. 29th Soc. Information Display (SID98)*, vol. **XXIX**, 927–930 (San Jose, CA, 1998).
14. 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).
15. S. Inoué and K.R. Spring, "Microscope image formation - Principles of Köhler illumination" in *Video Microscopy: the Fundamentals* (Plenum Press, New York, 1997).
16. 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).
17. 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).
18. A. Gullstrand, Appendix II in *Handbuch der Physiologischen Optik* (Voss, Hamburg, 1909).
19. A. Popiolek-Masajada and H. Kasprzak, "Model of the optical system of the human eye during accommodation," Ophthal. Physiol. Opt. **22**, 201–208 (2002).
20. 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).
21. 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).
22. OSLO is a registered trademark of Lambda Research Corp.
23. K. Toh and A. Neureuther, "Identifying and Monitoring Effects of Lens Aberrations in Projection Printing," in *Optical Microlithography VI*, H.L. Stover, ed., Proc. SPIE **772**, 202–209 (1987).
24. R.J. Becherer and G.B. Parrent, "Nonlinearity in optical imaging systems," J. Opt Soc. **57**(12), 1479–1486 (1967).
25. 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).
26. H. Lieberman and A. Pentland, "Microcomputer-based estimation of psychophysical thresholds: the best PEST," Behaviour Research Methods & Instrumentation **14**(1), 21–25 (1982).
27. C. Kaernbach, "Adaptive threshold estimation with unforced-choice tasks," Perception & Psychophysics **63**(8), 1377–1388 (2001).
28. M. Bass, ed., *Handbook of Optics*, vol. **1** (McGraw-Hill, Inc., New York, 1995).
29. J.W. Goodman, *Introduction to Fourier Optics*, (McGraw-Hill, Inc., New York, 1996).

1. Introduction

Extending the depth of focus (DOF) of two-dimensional head-mounted displays (2D-HMDs) has been the goal of many researchers over the past few years. In particular, head-mounted displays with a see-through mode should provide a high DOF. This mode allows to overlay virtual images over the user's real view being of great importance in many applications in the field of augmented reality [1]: For example, see-through 2D-HMDs may provide additional information to the user by attaching virtual data to objects in the real scene. However, a really unobtrusive and comfortable use of this feature is guaranteed only, when the image quality of the virtual scene is almost insensitive to natural accommodation changes of the user's human eye. Otherwise, the overlaid virtual data become blurred as soon as the user's eye is changing its accommodation in order to focus on real objects at various distances. This leads to discomfort or causes even severe accommodation problems as shown in experimental studies by Edgar et al. [2]. Thus, the display's DOF should be large enough so that the image quality of the overlaid computer output becomes independent of the distance of the object in the real world the user is accommodating at. A good target value for the DOF is a value larger than $5\,\mathrm{diopters}\,(\mathrm{D})$. Such a value implies that the image quality would be unaffected by eye accommodation for any viewing distances between infinity and at least $20\,\mathrm{cm}$.

1.1. Related work

Conventional 2D-HMD systems are usually based on a miniature LCD placed close to the eye. The miniature display is then imaged through magnifying optics to create a virtual image at some fixed distance in space. To extend the DOF in such systems, apodization methods can be used by e.g. introducing an additional aperture stop into the optical system. However, these apodization-based techniques suffer from a substantial decrease of display illuminance

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

making the readability of the display in bright ambient lighting condition difficult. To avoid reduction in illuminance, the display's image may be projected directly onto the user's retina, either by scanning the image as discussed in [3, 4, 5] or by a Maxwellian-view optical system as proposed in [6, 7, 8]. In both setups, the retinal projection increases the light efficiency so that the display's illuminance is less sensitive to apodizing [9]. However, apodization in these setups results in a decrease of lateral image resolution. Without more apodizing, we have recently shown by means of text readability measurements that the DOF of a Maxwellian-view retinal projection display can be extended by illuminating the LCD by partially coherent light instead of incoherent light [7].

Another approach for extending the DOF of head-mounted displays has been discussed in [11] in the context of stereoscopic HMDs: The system proposed there 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. [12] have outlined several focus solutions, such as adding a different monocular lens in front of each eye of a biocular 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 has been discussed by Sugihara et al. [13] by proposing a real-time adaption of the display's focus to the current eye accommodation. This, however, requires a fast and accurate tracking of the eye accommodation leading to an enhanced technical complexity.

1.2. Paper contribution

In this paper, we discuss a novel approach which is based on the partially coherent multiple imaging technique known from optical microlithography [14]. In this method a pupil phase mask is introduced at the aperture stop plane of the partially coherent retinal projection display as discussed in [8]. 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 determine 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 final 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.

2. Theoretical background

2.1. System overview

Figure 1 shows the principal setup of the considered retinal projection display representing a "Maxwellian-view optical system". Except for the additional phase mask, it corresponds to the system discussed in [7]: 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 'Kohler illumination' [15]. Through the lens L1, the light is focused onto the focal plane P1 where the system's aperture stop is located. Here the additional phase pupil mask (PM) is integrated. 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. Note 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. 1).

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

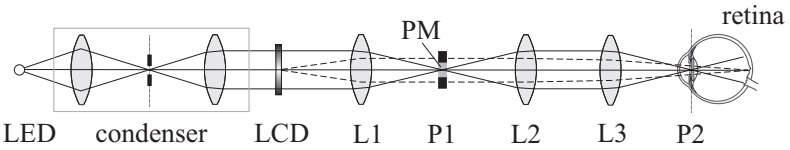


Fig. 1. Principal setup of the retinal projection display. P1 signifies the L_1-focal plane with the aperture stop and the pupil phase mask (PM) for the multiple imaging. L_1 to L_3 represent lenses and P_2 the eye pupil plane. The straight lines show the axial illumination rays while the dotted lines indicate the LCD-imaging.

As common in lithography, the coherence level of the illumination light is described by σ as the ratio between the numerical apertures of the condenser and the objective lenses. When $\sigma \to 0$ or $\sigma \to \infty$, the illumination is coherent or incoherent, respectively [16]. Finally, in this paper the setup is assumed to be best adjusted for an unaccommodated emmetropic eye.

2.2. Theoretical approach

For the mathematical description of the system we apply the Hopkins theory on the image formation in partially coherent systems [16, 17] (see Appendix). The most relevant parameter in this theory for defocusing considerations is the complex coherent transfer function $f(x,y)$.

For a perfect, aberration-free system at best focus, $f(x,y)$ becomes the system's exit pupil function $P(x,y)$ that equals 1 within the clear aperture and 0 outside. Thus, for aberrated systems the transfer function can generally be split into the exit pupil function $P(x,y)$ and into aberration factors introduced first by the display's optical system and secondly by the ocular system:

$$f(x,y) = P(x,y) \cdot e^{i\psi(x,y)} = P(x,y) \cdot e^{i(\psi_{disp}(x,y)+\psi_{eye}(x,y))} \tag{1}$$

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 $\psi_{disp}(x,y)$ keeps 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 paper 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:

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

where $\rho^2 = x^2 + y^2$ is the radial coordinate 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:

$$\psi_{disp}(\rho) = angle\left[\sum_{l=0}^{L} w_l\, e^{i\phi_l} e^{i\left(\frac{\pi}{\lambda}\delta_l\rho^2\right)}\right] \tag{3}$$

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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. (1)) limits the mask PM to phase-only masks without apodization.

3. Simulation background

3.1. Accommodation-dependent eye model

Several schematic eye models have been proposed in the literature from the simple and famous Gullstrand eye [18] up to a sophisticated one which models the eye's anatomy very accurately [19]. 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 paper are based on the accommodation-dependent wide-angle schematic eye model as proposed by Navarro et al. and Escudero et al. [20, 21].

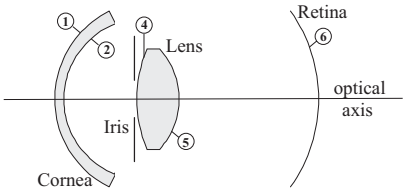


Fig. 2. The accommodation-dependent schematic eye model as used for all simulations.

This model - depicted in Fig. 2 - consists of three conical optical surfaces (surfaces 1,4,5) and two spherical surfaces (2,6), the latter one representing the retina. The parameters defining the eye's geometry in the unaccommodated 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 wavelength $\lambda = 635\,\mathrm{nm}$. The focal length f of the schematic eye is $22.18\,\mathrm{mm}$ in the image space; the refractive power in the unaccommodated state is 60.2 diopters (D).

When the natural eye accommodates [1], the increment in refractive power is mainly caused by two contributions: first, a geometrical variation of the eye lens shape and secondly, a change of the graded-index structure within the eye lens. In the eye model used, the geometrical shape variations are incorporated accurately as proposed by [20]. 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. The explicit parameter values as well as the modeling of the accommodation dependence are discussed in detail in [10].

This schematic eye model has been implemented in the ray-tracing software package OSLO [22] in order to calculate the ocular wavefront error map $\psi_{eye}(x,y)$ as a function of eye accommodation, iris diameter and eye orientation. The wavefront error map has been expressed by the Zernike polynomial expansion (up to the 36th Zernike polynomial, thus describing all aberrations up to the ninth order). The resulting Zernike coefficients act as input for the subsequent simulation of the spatially partial coherence.

3.2. Simulation of partial coherence

The calculation of the retinal images is mainly based on the Eqs. (1-3) and (6-7) and are carried out on a slightly adapted version of the SPLAT 5.0 software package. This program has

[1] for illustration: a change of the eye's refractive power by e.g. $\Delta D = 4\,\mathrm{D}$ corresponds to a change of the viewing distance from ∞ to $25\,\mathrm{cm}$.

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

been developed by the University of California at Berkeley for simulating two-dimensional projection-printing with partial coherence [23]. The system is described by several 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 $(d_l, w_l, \phi_l)_{l=0..L}$ defines the phase mask PM according to Eq. (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.

3.3. Simulation parameters

For the design of the phase mask the system parameters according to Fig. 1 were defined as follows:

Illumination wavelength	$\lambda = 635\,\text{nm}$
Diameter of eye's iris	$\varnothing_{\text{iris}} = 3\,\text{mm}$ (bright-adapted iris)
Diameter of aperture stop P1	$\varnothing_1 = 2.1\,\text{mm}$
Magnification of relay lens system	$M = 0.65$

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. This box defines the area where the eye can freely move without loosing the LCD image due to iris-blocking 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 $\varnothing_{\text{exit}}$, but is a function of the eye's iris diameter $\varnothing_{\text{iris}}$. The exit pupil controls, in this case, the maximum retinal resolution. This can be illustrated by the following considerations: In a coherent Maxwellian-view system, a Fourier transform pattern of the LCD image is formed on the plane of the iris pupil. Thereby, the angular spatial frequency F_x (in cycles/rad) is situated at a distance $x = \pm\lambda F_x$ from the pupil center, independent of any system parameters (in particular independent of the exit pupil) (see left drawing in Fig. 3) [9]. 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.

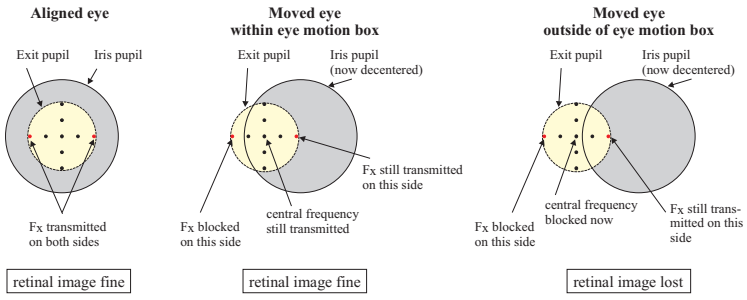


Fig. 3. Schematic illustration of eye motion in case of fully coherent illumination. The dots within the exit pupil depict the Fourier transform pattern of the LCD image.

When the eye moves, the iris pupil will block some spatial frequencies, but only at one side of the symmetric spectrum as illustrated in the middle drawing of Fig. 3. This does not affect

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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 iris. This happens, when the eye moves more than the current radius of its iris pupil (see right drawing in Fig. 3). Thus, in fully coherent imaging, the eye motion box depends on $\varnothing_{\text{iris}}$ 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 iris 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 iris 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 iris, preventing thus the iris-blocking effects [5].

4. Design of the phase mask

4.1. Evaluation method

Due to the nonlinearities in partially coherent imaging [24] 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 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 [5]. Thus, the test image defined for the subsequent evaluations consists of a simple 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 [25]:

$$C = \frac{2 s_x s_y}{s_x^2 + s_y^2} \qquad \text{and} \qquad S = \frac{s_{xy}}{s_x s_y} \tag{4}$$

[2] To illustrate this notion: a viewing angle of e.g. $\alpha_v = 0.4$ 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)

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

with

$$s_x^2 = \frac{1}{N-1}\sum_{m,n=1}^{N}(x_{m,n}-\overline{x})^2 \qquad \text{and} \qquad s_{xy} = \frac{1}{N-1}\sum_{m,n=1}^{N}(x_{m,n}-\overline{x})(y_{m,n}-\overline{y})$$

C and S are first calculated locally in a sliding window of size $N \times N$ (in our case, N=30), 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. The contrast comparison function C compares the contrast of the test image X with regard to the reference image Y while the structural function 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. 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 for all pixels (m,n) where b is constant.

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

Basically, the phase mask, which optimizes the display's DOF, should now be found by varying the coefficients $(\delta_l, w_l, \phi_l)_{l=0..L}$. However, 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 to be equally distributed (i.e. $\phi_l = l\,\pi/L$). Thus, for our specific case, Eq. (3) turns out to be:

$$\psi_{disp}(\rho) = angle\left[\sum_{l=0}^{L} e^{i\pi l\left(\frac{1}{L}+\frac{\Delta\delta}{\lambda}\rho^2\right)}\right] + \frac{\pi}{\lambda}\delta_0\rho^2 \tag{5}$$

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 $(\varepsilon, \Delta\delta, \overline{\delta})$. Here $\varepsilon = L\Delta\delta = \delta_L - \delta_0$ signifies the dioptric extent over which the $L+1$ planes are distributed and $\overline{\delta} = \delta_0 + \varepsilon/2 = (\delta_L + \delta_0)/2$ indicate the mean value of all focal planes. Consequently, the focal planes δ_l are equally distributed in the range $[\overline{\delta} - \varepsilon/2, \overline{\delta} + \varepsilon/2]$.

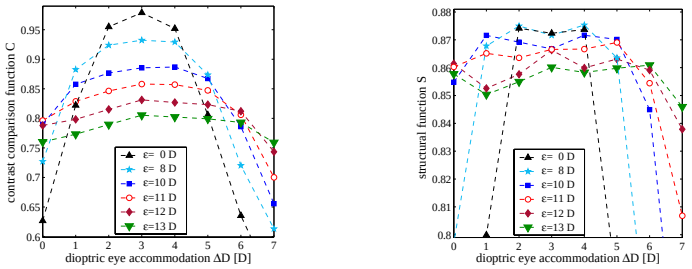
The potential of this approach has been considered separately for three different coherence levels, $\sigma = 0$ (coherent), $\sigma = 0.5$ (partially coherent) and $\sigma = \infty$ (incoherent).

4.3. 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 D while ε was varied. The mean value $\overline{\delta}$ was set to 3.5 D.

Figure 4 shows the results for various ε-values. The curves for $\varepsilon = 0$ represent the corresponding unifocal system for comparison reasons. All data were calculated with partially coherent illumination of $\sigma = 0.5$ and a font size subtending a viewing angle $\alpha_v = 0.4$ 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 $\Delta D \approx 0$ D and $\Delta D \approx 5.5$ D. This improvement occurs at the cost of the image quality for accommodation values in the intermediate range (2 D $\lesssim \Delta D \lesssim$ 4 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 D, an ε-value is required at which the image quality in the range between 5 D and 6 D is similar to the one at

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

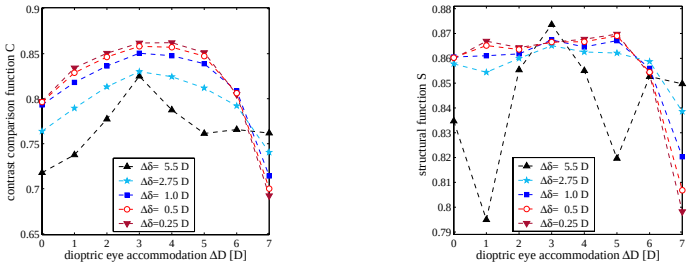


(a) Contrast function C (b) Structural function S

Fig. 4. 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 $\alpha_v = 0.4\,\mathrm{deg}$.

$\Delta D \approx 0\,\mathrm{D}$. 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, $\varepsilon = 11\,\mathrm{D}$ 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 intermediate part. With $\varepsilon = 11\,\mathrm{D}$, 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 $\varepsilon = 11\,\mathrm{D}$ constant. The results are shown in Fig. 5. Both, the contrast and the structural similarity are improved



(a) Contrast function C (b) Structural function S

Fig. 5. Image quality functions C and S in terms of ΔD for various values of $\Delta\delta$ with $(\varepsilon, \overline{\delta}) = (11\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\sigma = 0.5$ and $\alpha_v = 0.4\,\mathrm{deg}$.

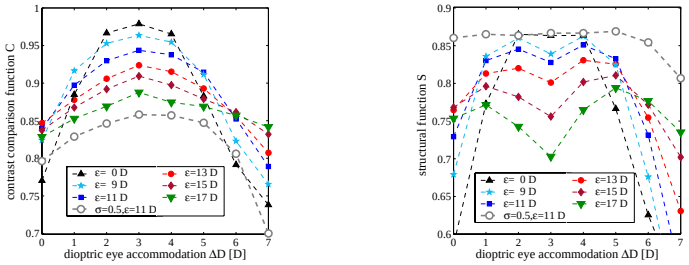
when reducing $\Delta\delta$ down to $\Delta\delta \approx 0.5\,\mathrm{D}$. For $\Delta\delta$-values smaller than $0.5\,\mathrm{D}$ the image quality stays rather constant (i.e. $\Delta C < |0.014|$ and $\Delta S < |0.005|$) within the accommodation range of interest. This is not surprising as any 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.

4.4. Fully coherent illumination $\sigma = 0$

The same analysis can be made for fully coherent light. Figure 6 shows the corresponding results when the dioptric extent ε is varied. Again, the other coefficients were kept constant:

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

$\Delta\delta = 0.5\,\mathrm{D}$ and $\overline{\delta} = 3.5\,\mathrm{D}$. Obviously, the contrast results are better than those for the partially



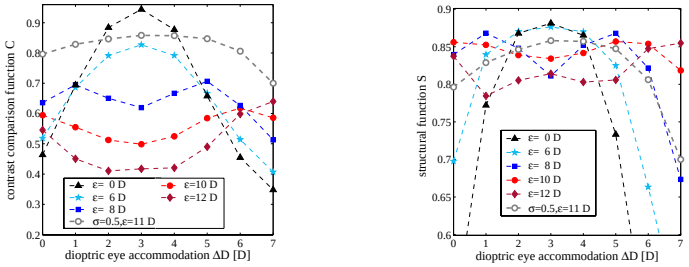
(a) Contrast function C (b) Structural function S

Fig. 6. Image quality functions C and S for $\sigma = 0$ in terms of ΔD for various values of ε and with $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_v = 0.4\,\mathrm{deg}$. The grey unfilled symbols show the former results for $\varepsilon = 11\,\mathrm{D}$ with partially coherent illumination for comparison.

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 ($\varepsilon \lesssim 9\,\mathrm{D}$) the structural similarity is comparable to those with $\sigma = 0.5$ at least for accommodation values between about 2 D and 4 D. But at the edges of the considered range the structural quality decreases substantially. For higher ε-values ($\varepsilon \gtrsim 15\,\mathrm{D}$) 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 D. These effects can mainly be attributed to disturbing coherent interferences between the simultaneously projected images.

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. 7: The structural quality for $\varepsilon = 8\,\mathrm{D}$ turns out to be similar to the



(a) Contrast function C (b) Structural function S

Fig. 7. Image quality functions C and S for $\sigma = \infty$ in terms of ΔD for various values of ε and with $(\Delta\delta, \overline{\delta}) = (0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_v = 0.4\,\mathrm{deg}$. The grey unfilled symbols show the former results for $\varepsilon = 11\,\mathrm{D}$ with partially coherent illumination for comparison.

corresponding one in the partially coherent case for $\varepsilon = 11\,\mathrm{D}$. However, now the image quality suffers from a severe loss in contrast.

4.6. *Final phase mask profile*

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

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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 $(\varepsilon, \Delta\delta, \overline{\delta}) = (11\,\mathrm{D}, 0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. The corresponding phase profile $\psi_{disp}(\rho)$ of the phase mask is shown in Fig. 8. Note again, that the transmission rate is 100% all over the mask.

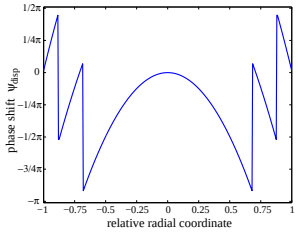


Fig. 8. Phase profile ψ_{disp} of the rotationally symmetric phase-only mask PM.

To give an idea of the image quality, Fig. 9 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 $\varepsilon = 0$ and $\sigma = 0.5$. Figure 9 confirms the results as

Fig. 9. First row: retinal images for $\varepsilon = 0\,\mathrm{D}$ and $\sigma = 0.5$ for comparison. The other three rows show the retinal images for the three coherence levels when the designed phase-mask *PM* (see Fig. 8) is applied. The labels below the images indicate $(\sigma/\Delta D)$. Again, $\alpha_v = 0.4\,\mathrm{deg}$.

obtained by the contrast and structural evaluation: While the fully coherent case suffers from structural degradation, the fully incoherent case is characterized by a reduced contrast due to the phase mask. Again, the improvement for the partial coherent case (second row in Fig. 9) is evident.

5. Evaluation

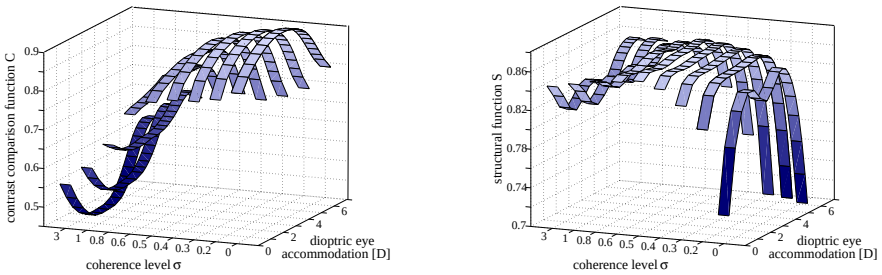
So far, we have designed the phase mask on the basis of three coherence levels $\sigma = 0/0.5/\infty$ and for one font size viewing angle $\alpha_v = 0.4\,\mathrm{deg}$. Now, we will explore the potential of this designed phase mask (as defined in Fig. 8) in more detail. To this end, the retinal images of the text samples are calculated for three font size viewing angles α_v and for more coherence levels

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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.

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). Figure 10 shows the corresponding results for various coherence levels σ and eye accommodation values ΔD. The



(a) Contrast function C (b) Structural function S

Fig. 10. Image quality functions C and S in terms of coherence level σ and eye accommodation ΔD. The phase-mask PM used is described by $(\varepsilon, \Delta\delta, \overline{\delta}) = (11\,\mathrm{D}, 0.5\,\mathrm{D}, 3.5\,\mathrm{D})$. Again, $\alpha_v = 0.4\,\mathrm{deg}$.

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

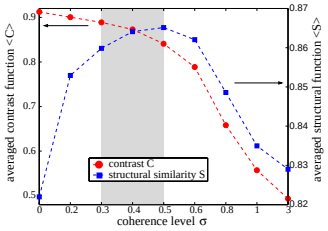


Fig. 11. Contrast and structural image quality averaged over the considered range of accommodation $[0\,\mathrm{D}, 6\,\mathrm{D}]$ in terms of σ.

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. 11. These averaged values represent well the grade of insensitivity of the corresponding quality 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 $\sigma \approx 0.5$.

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

5.2. *Psychometric experiments*

Objective image quality criteria are valuable tools to characterize 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 assess the quality of the simulated retinal images relative to a given reference image.

5.2.1. 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 [26, 27]. 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 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 [28]. In our setup, the Rayleigh criterion gives a defocus of 0.85 D. The reference image for $\alpha_v = 0.4\,\mathrm{deg}$ is shown in the inset of Fig. 12(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 characterized by one of the 12 combinations (α_v, σ) with the font size viewing angles[3] $\alpha_v = 0.34/0.4/0.46\,\mathrm{deg}$ 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.

5.2.2. 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

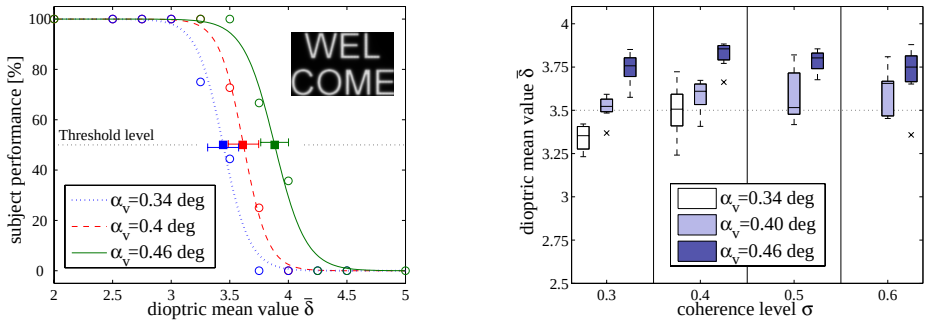
[3] Note that these viewing angles correspond to font sizes $F_S = 12/14/16\,\mathrm{pt}$ on a screen when viewed from a distance of 50 cm

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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 12(a) illustrates exemplarily the estimated psychometric functions and the derived thresholds - indicated by squares (■) - for one subject and all three font size viewing angles $\alpha_v = 0.34/0.4/0.46\,\mathrm{deg}$ at $\sigma = 0.4$. Other psychometric curves are similar.

Figure 12(b) shows the summarized 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 ($\alpha_v = 0.34\,\mathrm{deg}$) 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 $\sigma-$values.

The threshold results for $\alpha_v = 0.4\,\mathrm{deg}$ and $\alpha_v = 0.46\,\mathrm{deg}$ lie all well above $\overline{\delta} = 3.5\,\mathrm{D}$. Consequently, setting $\overline{\delta} = 3.5\,\mathrm{D}$ guarantees a satisfactory image quality at $\Delta D = 0\,\mathrm{D}$. However, for $\alpha_v = 0.34\,\mathrm{deg}$, the thresholds lie rather close or even below $\overline{\delta} = 3.5\,\mathrm{D}$ so that for this size the image quality at $\Delta D = 0\,\mathrm{D}$ is critical, if $\overline{\delta} = 3.5\,\mathrm{D}$. If a better quality for $\alpha_v = 0.34\,\mathrm{deg}$ 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 $\alpha_v = 0.34\,\mathrm{deg}$ the image quality at $\Delta D = 0\,\mathrm{D}$ is critical.



(a) (b)

Fig. 12. (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 (○) while the squares (■) signify the derived thresholds. The threshold error bars indicate the 95%-confidence interval. The inset shows the reference image for $\alpha_v = 0.4\,\mathrm{deg}$. (b) shows the results of the first experiment based on 7 subjects (see text).

5.2.3. 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 vari-

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

ous 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 assessed to be better or at least equal as the reference image (except for $(\alpha_v, \sigma) = (0.34\,\mathrm{D}, 0.3)$). Figure 13 depicts the results of this second psychometric experiment

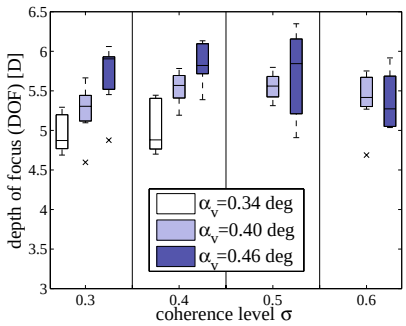


Fig. 13. Results of the second experiment (estimation of the DOF) based on 7 subjects.

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 the depth of focus to about $5.5\,\mathrm{D}$ for $\alpha_v \geq 0.4\,\mathrm{deg}$. The impact of the coherence level ranging from 0.3 to 0.6 is small as already expected from Fig. 11. For a smaller font size viewing angle, $\alpha_v = 0.34\,\mathrm{deg}$, 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.

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 partial coherent simulation tool. 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, however, 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 $\alpha_v \geq 0.4\,\mathrm{deg}$ 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 ($\alpha_v \leq 0.34\,\mathrm{deg}$), 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

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009

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.

Appendix

Here, we provide the formula for calculating the retinal images according to the Hopkins theory [16, 17]. According to this theory, the retinal intensity distribution $I(u',v')$ can be described as:

$$I(u',v') \propto \sum_{m,n,p,q} C_{m,n,p,q} a_{m,n} a^*_{p,q} e^{2\pi i \frac{NA'}{NA}[(m-p)u'+(n-q)v']} \tag{6}$$

with the transmission cross coefficients $C_{m,n,p,q}$ defined as:

$$C_{m,n,p,q} \propto \int_{-\infty}^{\infty} \gamma(x,y) f(x+\beta m, y+\beta n) f^*(x+\beta p, y+\beta q)\, dxdy \tag{7}$$

In these equations, $a_{m,n}$ are the Fourier coefficients of the LCD image and $\gamma(x,y)$ the effective illumination source. NA and NA' are the system's numerical aperture in the object and image space, respectively. Note that according to Fourier optics theory the coherent transfer function $f(x,y)$ at best focus is proportional to the system's exit pupil function [29]. The factor β stands for $\beta = \lambda / NA$ where λ is the wavelength of the illumination light. (u',v') are coordinates in the retinal image plane; (x,y) coordinates in the exit plane.

#5728 - $15.00 USD Received 12 Nov 2004; revised 2 Dec 2004; accepted 6 Dec 2004; published 20 Nov 2009