Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control

Kevin J. MacKenzie
Wolfson Centre for Cognitive Neuroscience, School of Psychology, Bangor University, Wales, UK

David M. Hoffman
Vision Science Program, University of California, Berkeley, CA, USA

Simon J. Watt
Wolfson Centre for Cognitive Neuroscience, School of Psychology, Bangor University, Wales, UK

Most stereoscopic displays present images at a single focal plane, resulting in “conflicts” between the stimuli to vergence and accommodation. Minimizing these conflicts is beneficial because they can cause distorted depth percepts, visual fatigue, and reduced stereoscopic performance. One proposed solution is to present a sum of images at multiple focal planes and to vary focal depth continuously by distributing image intensity across planes—a technique referred to as depth filtering. We evaluated this digital approximation to real-world variations in focal depth by measuring accommodation responses to depth-filtered stimuli at various simulated distances. Specifically, we determined the maximum image-plane separation that supported accurate and reliable accommodation. We used an analysis of retinal image formation to predict when responses might be inaccurate. Accommodation to depth-filtered images was accurate and precise for image-plane separations up to ~1 diopter, suggesting that depth filtering can be used to precisely match accommodation and vergence demands in a practical display. At larger plane separations, responses broke down in a manner consistent with our analysis. We develop this approach to consider how different spatial frequencies contribute to accommodation control. The results suggest that higher spatial frequencies contribute less to the accommodation response than has previously been thought.

Keywords: accommodation, vergence, stereoscopic display, volumetric display, 3D graphics


Introduction

Conventional stereoscopic displays present images at a single focal plane. As a result, the focal distance to points in the image is usually inconsistent with the depicted scene, resulting in a mismatch or conflict between the stimulus to accommodation and the stimulus to vergence. This causes two significant problems. First, distance information from accommodation specifies the properties of the display surface, not the depicted scene, which can lead to distortions in perceived depth (Hoffman, Girshick, Akeley, & Banks, 2008; Watt, Akeley, Ernst, & Banks, 2005). Second, because the focal distance and vergence-specified distance typically differ, observers frequently have to accommodate at one distance while converging at another in order to see a single, sharp binocular image. Accommodation and vergence responses are synergistically coupled (Fincham & Walton, 1957; Martens & Ogle, 1959), and attempting to decouple them has been shown to lead to visual fatigue (Hoffman et al., 2008; see also Emoto, Niida, & Okano, 2005; Ukai, 2007; Ukai & Howarth, 2008; Wann & Mon-Williams, 1997; Yano, Emoto, & Mitsuhashi, 2004), difficulty fusing stereoscopic images (Akeley, Watt, Girshick, & Banks, 2004; Watt, Akeley, Girshick, & Banks, 2005), reduced stereoaquity (Hoffman et al., 2008), and unstable accommodation responses (Fukushima, Torii, Ukai, Wolffsohn, & Gilmartin, 2009).

In principle, these problems could be eliminated if one could construct a stereo display in which the focal distance to each point in the scene could be adjusted to match the simulated distance. In one proposed solution, each eye sees a sum of images at multiple image planes, placed at different focal distances; focal distance is varied continuously by distributing image intensity across planes, using a technique referred to as depth filtering (Akeley et al., 2004). Psychophysical studies have shown that depth-filtered stimuli can reduce some of the problems caused by incorrect focus cues (Akeley et al., 2004; Hoffman et al.,...
required in the world. The visual system’s sensitivity to focal depth is therefore a number of focal planes required to adequately sample the continuous range of focal distances in the real world. The visual system’s sensitivity to focal depth is much poorer than its spatial sensitivity (Campbell, 1957). Thus, the eye’s viewpoint is fixed relative to the display along the surface normal to the focal planes, a relatively small number of planes (compared to the number of pixels required in the x and y dimensions) could provide adequate resolution in focal depth (Akeley et al., 2004; Rolland, Krueger, & Goon, 1999). This is an important advantage of fixed-viewpoint volumetric displays over other approaches. However, all multiplane displays suffer from significant specific constraints on the maximum number of (and therefore minimum spacing between) focal planes, as well as general bandwidth limits imposed by the increased graphics computation required for multiple images (Akeley et al., 2004). The multiplane displays constructed to date have at most three or four focal planes with spacings greater than the effective depth of focus of the eye (which has been estimated to be ±0.25 to 0.3 diopters (D); Campbell, 1957; Charman & Whitefoot, 1977; Rolland et al., 1999). At first approximation, therefore, this approach appears likely to provide only a coarse approximation to continuous real-world variations in focal distance.

Depth filtering

In multiplane displays, a decision rule is required to determine how image intensities at each point in the image should be assigned to the focal planes. A simple rule that assigns 100% of image intensity at each point to its nearest focal plane results in visible discontinuities in the retinal image (Akeley et al., 2004). This problem is solved by a technique referred to as depth filtering, in which image intensity is distributed across more than one plane. This is schematized in Figure 1. Image intensity at each image plane is linearly weighted according to the ratio of the dioptic distances of each point in the scene from the two nearest planes, calculated along lines of sight (Figure 1, right panel). Because the weights sum to 1, the luminance of an object remains constant at all distances.

Akeley (2004) and Hoffman et al. (2008) have suggested that depth filtering might also provide a reasonable approximation to continuous variations in focal distance. That is, the eye could be stimulated to accommodate to distances between image planes, even though no light originates from that position. This is conceivable because similar retinal images can result from viewing depth-filtered stimuli and equivalent real targets (see The stimulus to accommodation from depth-filtered images section below). If correct, continuous variations in accommodation could be stimulated using significantly more widely spaced (and therefore fewer) image planes than a simple analysis of the eye’s depth of focus suggests.

Several psychophysical studies have examined stereoscopic performance when viewing depth-filtered images. Stereoacuity and the time required to fuse stereo pairs have been found to be similar when the stimuli to accommodation and vergence were matched precisely (by presenting images coincident with one focal plane), and when depth-filtering was used to simulate the dioptic midpoint between two focal planes spaced 0.67 D apart (Akeley et al., 2004; Hoffman et al., 2008; Watt, Akeley, Girshick et al., 2005). The apparatus used in these studies had fixed power optical element (Liu, Cheng, & Hua, 2008; Love et al., 2005). The apparatus used in these studies had fixed

Multiple-focal-plane displays

Multiple-focal-plane displays (hereafter referred to as multiplane displays) have been constructed using static optics such as mirrors, beam splitters, or prisms (Akeley et al., 2004; McDowall & Bolas, 1994), and using adaptive optics. In this latter approach, images at different focal distances are displayed in a time-multiplexed manner by synchronizing the image on the screen with a variable-power optical element (Liu, Cheng, & Hua, 2008; Love et al., 2009; Shevlin, 2005). Despite different implementations, these two approaches are conceptually similar in that each eye sees a sum of images presented on a number of discrete focal planes. They are volumetric displays because light comes from different (focal) distances.

A key design parameter is the spacing between, and therefore number of, focal planes required to adequately sample the continuous range of focal distances in the real world. The visual system’s sensitivity to focal depth is much poorer than its spatial sensitivity (Campbell, 1957). Thus, if the eye’s viewpoint is fixed relative to the display along the surface normal to the focal planes, a relatively small number of planes (compared to the number of pixels required in the x and y dimensions) could provide adequate
focal-plane positions, however, and so it was not possible to make direct comparisons between performance when viewing a depth-filtered stimulus and an equivalent “real” stimulus at the same distance. The fact that performance was not disrupted when viewing depth-filtered stimuli is consistent with observers accommodating at the simulated distance. However, if they instead accommodated to one or other image plane, the vergence–accommodation conflict may have been too small (0.33 D) to cause significant problems. Thus, while these studies indicate that depth filtering does not impair stereo performance, they do not provide direct evidence that depth-filtered images are effective accommodative stimuli.

In this study, we measured the accuracy and reliability of accommodation to depth-filtered images and assessed how plane spacing influences these measures. This approach is informative for several reasons. First, it provides a direct, objective measure of the extent to which the stimulus to accommodation can be matched with the stimulus to vergence in depth-filtered images and provides a straightforward way to explore the display parameters (in particular, the separation between image planes) required for this. Second, it allows us to evaluate our analysis of retinal-image formation, which makes predictions about when depth-filtered images should provide a good approximation to the real world and when they should fail. If this analysis is validated by our empirical data, this approach could be used to make predictions not only about accommodation responses to depth-filtered images, but also more general predictions about responses (and therefore the extent of accommodation–vergence conflicts) with other types of displays. Third, depth-filtered images provide a novel stimulus manipulation, which, in conjunction with our analysis of retinal-image formation, should allow us to learn more about the stimulus to accommodation.

### Accommodation responses to multiplane images

The accommodation literature has little to say about how the accommodation system might respond to depth-filtered images. The most similar situation that has been studied is the so-called “Mandelbaum effect”: when one tries to accommodate to a distant scene, viewed through a mesh screen, accommodation is often involuntarily driven to the distance of the screen (Mandelbaum, 1960). Owens (1979) measured accommodation to two plaids composed of 2.5-cpd rectangular-wave gratings. The plaids differed in orientation by 45 deg and were presented at focal distances differing by 2 D. Consistent with Mandelbaum, there was a strong tendency to accommodate to the plaid nearest the resting focus. Rosenfield and Ciuffreda (1991) conducted a similar experiment but reached different conclusions. They presented horizontal and vertical square-wave gratings (unspecified frequency) at distances differing by 2 or 4 D. Responses often fell in between the focal distances rather than at one or the other, suggesting that observers might accommodate in-between image planes when viewing depth-filtered images. There are two important differences between depth-filtered images and the Mandelbaum situation that make accommodation to intermediate distances more likely. First, the geometric patterns of the retinal images resulting from the near and far stimuli are highly dissimilar in the Mandelbaum situation while the patterns in depth-filtered images are identical. Second, the above studies used much larger separations between focal planes than existing multiplane displays (the planes were 2/3 D apart in Akeley et al.’s display, for example). In the following section, we explore why smaller plane separations make accommodation to intermediate distances more likely.

### The stimulus to accommodation from depth-filtered images

The accommodation system operates to optimize retinal-image contrast (Alpern, 1958; Fender, 1964; Heath, 1956; Kotulak & Schor, 1986a; Manny & Banks, 1984; Owens, 1980; Raymond, Lindblad, & Leibowitz, 1984; Switkes, Bradley, & Schor, 1990; Toates, 1972; see also Cheng, Thibos, & Bradley, 2003). We can therefore gain insight into how the system might respond to depth-filtered images by considering the retinal-image contrast that results from viewing a spatially broadband target at different simulated focal distances and with different plane separations.

Akeley (2004) developed a model of retinal-image formation that suggests depth-filtered images might pro-
provide a reasonable approximation to real-world variations in focal depth, under certain circumstances. This model used geometrical optics to extend an empirically measured well-focused point-spread function to situations in which the eye was incorrectly focused. Hoffman et al. (2008) used a different approach, using measurements of the optical aberrations of a real eye. This latter analysis better captures the effective depth of focus of the eye, and so we follow it here. The aberrations, pupil size, and accommodative state of the normal eye determine the quality of the retinal images for both depth-filtered images and real targets. The aberrations lower the peak contrast transmission of the eye but greatly expand the depth of focus (Marcos, Moreno, & Navarro, 1999). Therefore, it is important for us to consider real eye optics in our model of retinal-image formation. Our analysis is described in full in Appendix A. Because we did not have facilities to measure our particular observers’ aberrations and because we wished to draw general conclusions, we based the analysis on an eye with aberrations in the typical range, from a population of measurements made by Austin Roorda at UC Berkeley (see Appendix A). We also carried out our analysis using aberration data from several different eyes with varying optical quality to confirm that our optical model produced generalized predictions (see Appendix A for examples).

We computed the modulation transfer function (MTF) for a real eye for a range of accommodative responses. We then plotted the modulation transfer (ratio of retinal to incident contrast) for a subset of spatial frequencies. We plot these contrast ratios in Figure 2 for the situation of viewing a one-plane stimulus (equivalent to a real-world stimulus) and a two-plane, depth-filtered, simulation of the same focal distance, as a function of where the eye is focused. In Figure 2a, the stimulus is at the dioptric midpoint between two focal planes that are separated by 6/9 D. For a real target (dashed curves), the maximum contrast occurs, of course, when the eye is focused at or near the target distance (as a result of this eye’s high-order aberrations, maximum contrast is achieved by focusing at slightly different distances for different spatial frequencies; Green & Campbell, 1965). The degree to which the two-plane simulation (solid curves) approximates the real target depends strongly both on spatial frequency and on the separation between image planes. Figure 2a shows that at around 2 cpd, the 6/9 D plane separation produces a nearly perfect approximation wherever the eye is focused. The approximation gets poorer, however, with increasing spatial frequency. This effect is more pronounced with larger plane separations. Figure 2b shows a similar plot but with 10/9 D plane separation. At ~8 cpd, the maximum contrast no longer occurs at the midpoint, but rather at one or the other image plane. This effect is explored further in

Figure 2. Sample contrast-ratio functions for real targets and depth-filtered images with differing levels of defocus. The ordinate is the ratio of retinal-image contrast to stimulus (incident) contrast. The abscissa is the accommodation state, relative to the “simulated” focal distance. The different plots show different plane separations (vertical green lines) and two different far/near image-plane intensity ratios (see icon above each plot). The dashed curves show retinal-image contrast from viewing a real target, at the location shown by the dashed gray line. The solid curves show retinal contrast from viewing the depth-filtered approximation, presented at two image planes. Different colors denote different spatial frequencies (see (a)). The slight asymmetries in the contrast-ratio functions evident in (a) and (b) result from the aberrations in this particular eye.

Downloaded from jov.arvojournals.org on 12/26/2021
Figure 2c, which shows retinal contrast for a stimulus at three-quarter distance between the image planes (far/near image intensities of 75:25%). This shows that for simulated distances away from the dioptric midpoint of the image planes, there is a systematic bias in where the maximum contrast occurs. As spatial frequency increases, the peak contrast shifts progressively closer to the image plane nearest to the simulated distance.

The effect of plane separation on the accuracy of the depth-filtered approximation to a real target is summarized in Figure 3. The figure plots the accommodative distance at which contrast, at single spatial frequencies of the MTF, is maximized, as a function of simulated focal distance. The different panels show different image-plane separations. These plots were generated by first computing retinal-image contrast curves like those in Figure 2, for a near-continuous range of simulated distances. The accommodative distance at which peak retinal contrast occurred was then determined in each case. Figure 3 shows that with 4/9 D plane separation, the peak contrast occurs very close to the intended focal distance across the range of spatial frequencies plotted (2–16 cpd). With increasing plane separation, however, the curves become increasingly sigmoidal, indicating that peak contrast is shifted toward one of the image planes, particularly at higher spatial frequencies. With sufficient plane separation, at high spatial frequencies, peak contrast occurs only at the closest focal plane to the simulated distance (8 and 16 cpd, with plane separations greater than 10/9 D). (Note that Figure 3 does not show the magnitude of maximum contrast; although the signal to accommodation is increasingly biased with increasing spatial frequency, it also presumably reduces significantly, because contrast is highly attenuated [Figure 2].) Precise details of these effects differ with each eye’s particular optics, but the overall dependence of the accuracy of the stimulus to accommodation on (i) spatial frequency and (ii) image-plane separation is similar across eyes with a range of optical qualities (see Appendix A).

Based on this analysis, we expect that accommodation responses to multiplane stimuli and real targets will be very similar with smaller plane separations, but at larger separations they should be increasingly different. In particular, we expect responses to multiplane stimuli to resemble responses in the Mandelbaum situation because retinal-image contrast can only be maximized by accommodating at one or other image plane. Our model shows that the maximum image-plane separation that can be used depends critically on the contributions of different spatial frequencies to the accommodation response. Low spatial frequencies should drive the accommodation response to the simulated distance, whereas high spatial frequencies may increasingly drive the response to one or other image plane. Therefore, if low to mid-range spatial frequencies are the principal stimulus to accommodation (as suggested by Mathews & Kruger, 1994; Owens, 1980; Tucker, Charman, & Ward, 1986; Ward, 1987), large plane separations can be used. If higher frequencies are important (as suggested by Charman & Tucker, 1977, 1978; Kotulak & Schor, 1986b; Phillips, 1974), the eye may focus at one or the other image plane unless the plane separation is rather small. It has been hypothesized that accommodation is initially driven by low and mid-range spatial frequencies and that higher frequencies contribute by fine-tuning the response (Charman & Tucker, 1978). This makes sense because retinal contrast falls off quickly with defocus at higher spatial frequencies (Figure 2). Thus, high frequencies might provide a good signal for fine-tuning when the eye is already reasonably focused. Most empirical research has, however, found little effect on accommodation.
responses of spatial frequencies higher than 6–8 cpd. If this is correct, multiplane depth-filtered images should provide an appropriate stimulus to accommodation up to plane separations of 6/9 D or even 8/9 D.

Aims of the current study

Our first aim is to provide a proof-of-principle test of whether (as our analysis of the retinal image suggests) depth-filtered images can stimulate continuous, accurate variations in accommodation responses. If so, we aim to determine the maximum plane separation that results in a reasonable accommodation response. This should provide useful information regarding the feasibility of this approach for correcting accommodation–vergence conflicts. Our second aim is to evaluate our analysis of retinal data is complicated by the fact that peak retinal distance (e.g., Mathews & Kruger, 1994). Interpreting such accommodation in response to modulation in target focal distance (e.g., Charman & Tucker, 1977, 1978) or (ii) modulation in either (i) the steady-state accommodation response (e.g., Charman & Tucker, 1977, 1978) or (ii) the separation between the focal planes. We isolated the stimulus to accommodation) or to reduce the gain (amplitude) of the response to modulations in focal depth. Our analysis of the retinal image shows that spatially broadband depth-filtered stimuli have the interesting feature that there is a spatial-frequency-dependent bias in the focal distance at which peak retinal-image contrast occurs. We can therefore gain insight into the contribution of higher spatial frequencies by examining biases in the steady-state accommodation responses.

We examined steady-state accommodation to a spatially broadband target (a Maltese cross) as a function of (i) the ratio of image intensities at each of two focal planes and (ii) the separation between the focal planes. We isolated the influence of retinal image sharpness on accommodation by presenting stimuli monocularly (so vergence could not drive accommodation) and by keeping the retinal image always the same size (so responses could not be driven by looming). We characterized the extent to which depth-filtered images stimulated a “correct” accommodation response by determining whether changes in the ratio of near-plane to far-plane image intensity lead to linear changes in the accommodation response and by comparing responses to those to an equivalent real target (100% image intensity at one focal plane).

Methods

Observers

Five observers participated in the experiment, aged 18 to 25 years. None reported a history of ophthalmic abnormalities (besides myopia). All observers attended an optometrist within the last 18 months prior to data collection. Four had not been prescribed eyeglasses or other treatments, and one was prescribed a −3D correction but required no correction for astigmatism. This observer wore his normal soft, disposable contact lenses (Sauflon Bioclear, 55% water content) during the experiment. Each observer was screened for a normal accommodation response prior to testing. All observers showed nearly linear accommodation stimulus–response functions to real targets, and their responses were stable throughout a 5-s presentation (the duration of the experimental stimuli).

Multiple-focal-plane display

A schematic of the display is shown in Figure 4. Similar to the display constructed by Akeley et al. (2004), the eye sees the sum of aligned images drawn at three focal planes, created using a combination of beam splitters and a first-surface mirror. The images are presented on a 30” Samsung 30ST TFT monitor (2560 × 1600 pixel resolution, 0.25-mm dot pitch). Different regions of the screen are viewed via a combination of the beam splitters and/or mirror, giving rise to focal planes at different distances. The images are viewed through a Badal lens (spherical lens, focal length 60 cm, at 60-cm viewing distance; Badal, 1876). Using Badal optics is useful for two reasons. First, it ensures that a given image size on the monitor is always constant at the eye, independent of focal distance (the angular resolution of the focal planes is therefore also constant; Nyquist limit ~21 cpd). Second, between the lens and optical infinity, variations in dioptric distance are linear in physical distance. Thus, a given focal-plane separation is achieved by positioning the beam splitters/mirror a constant physical distance apart, independent of their overall optical distance. The Badal lenses are 15 cm wide × 12 cm high, giving a field of view of 14.3 × 11.4 deg. The mirror and beam splitters are mounted on an optical rail so the distance...
to the focal planes, as well as the inter-plane separation, is continuously adjustable from 1.33 to 0 D. The eye views the display through a carefully constructed aperture designed to eliminate visible sharp edges on the lens mounts and other apparatus, which could act as a stimulus to accommodation. (This is a particular issue with TFT monitors, which emit some light even when the luminance is set to zero.) The aperture consists of a series of baffles, close to the eye, made of gel diffusing material as used on stage lights.

**Measurement of accommodation**

The accommodative state of the observers’ right eye was sampled at 30 Hz, using a Grand Seiko WV-500 infrared (IR) autorefractor. This “open field” instrument has high validity and reliability (Wolffsohn, O’Donnell, Charman, & Gilmartin, 2004). The display was viewed through a visible-transmitting, IR-reflecting mirror in the head of the instrument (Figure 4). The Grand Seiko WV-500 displays a single objective measurement of accommodative state on a button push. To do this, the instrument analyzes a video image to calculate the separation of two sets of measurement bars, reflected off the retina (Figure 5a; for a detailed description, see Mallen, Wolffsohn, Gilmartin, & Tsujimura, 2001; Wolffsohn et al., 2004). Following the work of Wolffsohn et al., we used the instrument in a “continuous mode”, in which the measurement bars are continuously visible and updated at 30 Hz (Wolffsohn, Gilmartin, Mallen, & Tsujimura, 2001; Wolffsohn et al., 2004). The video is output as an NTSC signal, which we analyzed offline to compute the separation of the bars on each frame. Our analysis differed from that of Wolffsohn et al. (2004), so we detail it here and in Figure 5.

The NTSC signal was digitized using a Canopus ADVC 110 analog-to-digital video converter. Each frame was then analyzed using custom image-analysis software, written in MATLAB (Mathworks). A sample frame is shown in Figure 5a. First, the measurement bars were cropped from the image to remove extraneous image data, such as the text and focusing ring in the center of the image. To localize the center of each vertical measurement bar, we then computed the sum of the horizontal luminance profile across this cropped image in the region known to contain the measurement bar. The bars have a symmetrical, approximately Gaussian luminance profile, independent of their separation (Figure 5b). Finally, to compute the exact pixel location of the center of each bar, we computed the first derivative of the luminance profile and identified the zero-crossing (Figure 5c). A corresponding process was carried out for...
the horizontal bars. This method allowed us to measure the bars’ separation with subpixel accuracy, independent of the overall luminance in the images.

We determined the relationship between measurement bar separation (in pixels) and changes in refractive state by taking measurements of a model eye, of variable refractive power, in 1-D steps. When operating the autorefractor in standard mode (via a button push), the internal optics are refocused before each measurement. Because this is not possible during continuous recordings, the calibration measurements were made without refocusing the instrument (Wolffsohn et al., 2004). The results are shown in Figure 5d. The relationship between the power of the model eye and the pixel separation of the measurement bars was very linear. The variation in measurement bar separation with refractive power was 3.46 pixels/D for the horizontal bars and 3.16 pixels/D for the vertical bars. These data agree closely with those reported by Wolffsohn et al. (2004).

The slopes of the calibration functions in Figure 5d are very stable across different positions at which the instrument is focused and across positions of the autorefractor head relative to the eye. The intercepts, however, depend on these parameters. Consequently, we can quantify changes in accommodative state precisely and accurately, but absolute accommodative state cannot be precisely determined (the instrument displays objective measurements on a button push, but these are quantized to units of 0.12 D). We therefore normalized data within sessions, relative to the average measurement bar separation for each observer while accommodating to the farthest focal plane. Within sessions, the autorefractor position was held constant, and it was not refocused. To minimize error resulting from this normalization process, we screened all observers prior to testing and tested only those who showed a repeatable, approximately linear accommodation stimulus–response function throughout the stimulus range when the target was a real Maltese cross.

Changes in accommodative state were computed separately from the horizontal and vertical bars for each frame and averaged to give an overall estimate. We also monitored pupil size using an Eyelink II infrared eye tracker (SR Research), sampling at 250 Hz.

Lumincance calibration of the display

We carefully calibrated the luminance and color balance of our display to ensure that we could present white (i.e., spectrally broadband) depth-filtered stimuli correctly (note the mirrors and beam splitters are not wavelength neutral). For each focal plane, we measured gamma functions for the red, green, and blue color channels, using a Minolta CS-100 Chromameter at the eye’s position (i.e., viewing through all the optics, including the autorefractor mirror). These were fitted using second-order polynomials to provide a continuous luminance lookup table for each color channel at each focal plane. We then used the Chromameter to confirm empirically that we could produce a “white” (CIE chromaticity coordinates, x ≈ y ≈ 0.333), at any given luminance, on each focal plane. We also confirmed that our calibration was correct for different positions of the focal planes, because TFT monitors can have nonuniform luminance properties across their surface. This potential problem was minimized by using only a central strip of our monitor, avoiding the corners where luminance change is most prominent.

Alignment

For multiplane images to be effective, it is critical that images at the various focal planes are aligned with one another, with respect to the eye’s viewpoint, and that the viewpoint is fixed relative to the display. Accurate measurement of accommodation also requires that the eye’s location be stabilized. Observers were positioned using individually adjusted bite bars. We used a sighting device to locate their eyes at the correct distance (60 cm from the optical center of the Badal lens) along the surface normal from the center of the display (Hillis & Banks, 2001). At the beginning of each experimental session, observers also performed an alignment task, to account for any small errors in bite-bar alignment and to ensure near-perfect alignment of images at each focal plane. First, a wire loom was placed in front of the display, consisting of two wires that intersected at the surface normal from the center of the display. The observer centered a small dot (10 arc min), presented at one focal plane, on this intersection. They then performed a vernier alignment task to co-align images at the remaining focal planes. This consisted of aligning gratings comprised of fine vertical or horizontal lines (Akeley et al., 2004). The lines were drawn using anti-aliasing to achieve subpixel accuracy. This vernier task was completed five times, and the average of these settings was used to position the images in the experimental session. Observers repeated this task if the standard deviation of their settings was higher than 0.75 pixels. After this procedure, observers could not detect any misalignment of the two images when viewing the two-plane experimental stimuli.

Stimuli

All stimuli were presented monocularly, to the right eye only, using one or two image planes. The stimulus was a white-on-black Maltese cross (5.25 deg in diameter), chosen because it is a spatially broadband target. The luminance of the “white” regions was 60 cd/m² on a 0.27 cd/m² background. The room lights were turned off so only the stimulus was visible. We varied simulated focal distance and the separation between image planes. The image-plane separations were 4/9, 6/9, 8/9, 10/9, and 12/9 D, centered around 6/9 D. At each plane separation,
stimuli were presented at the near and far image planes ("real" targets, with 100% image intensity at one image plane), and at intermediate simulated focal distances, in increments of 1/9 D. All focal distances were presented by varyin the ratio of image intensities at two image planes, according to the depth-filtering rule described earlier.

**Procedure**

At the start of each trial, the display screen was blank. The observers initiated a trial with a button press, at which point the stimulus appeared for 5 s. Observers were instructed to fixate the center of the Maltese cross and to try to keep its image clear. They were also asked to refrain from blinking and were encouraged to take numerous breaks during a block of trials (pilot testing indicated that these measures, along with allowing observers to initiate trials, minimized blinking during stimulus presentations). Sessions were blocked by plane separation. Within each block, the simulated focal distance was randomized. The order of plane-separation blocks was randomized for each subject. Each focal distance, at each plane separation, was viewed a total of 10 times (a total of 450 trials per observer).

**Results**

For each trial, we computed the eye’s accommodative state at each sample throughout the 5-s period in which the stimulus was visible. Within a session, measurements were normalized according to the accommodative state measured at the farthest image plane. A sample trace from one trial is shown in Figure 6. We found that the accommodation response generally reached a steady state after 1 s. We computed a mean accommodation response for each trial by averaging across the final 4 s of each trace. We also removed blinks, which are detectable as blank periods in the accommodation trace; blinks occurred rarely. The resulting mean accommodation responses were averaged across the 10 repetitions of each trial type to give an average for each observer in each condition. We measured pupil size for the 5-s duration of a trial. The average pupil size across all trials and observers was 4.9 mm (SD = 0.49 mm); we used this value as a parameter in our analysis of retinal-image formation (Figures 2 and 3, Appendix A).

**Bias in accommodation responses**

The top row in Figure 7 shows steady-state accommodation responses as a function of simulated focal distance for the five plane separations. Each data point is the average across the five observers. In each plot, the vertical green lines denote the positions of the image planes. The data points lying on those lines are control conditions because they are accommodation responses to one-plane targets and are therefore equivalent to real-world stimuli. In pilot testing, our observers showed near-linear accommodation responses to a real target, through the range of focal distances tested here. Thus, the test of whether depth-filtered stimuli provide an accurate stimulus to accommodation is the extent to which accommodation responses to the two-plane stimuli varied linearly between these points.

Figure 7 (top row) shows that at the small plane separations there was a very linear relationship between the simulated focal distance and the accommodative response. Indeed, this remains the case for most of the plane separations tested (although the data became slightly noisier at larger plane separations), suggesting that depth-filtered stimuli can stimulate accurate continuous variations in the accommodation response. At the largest plane separation of 12/9 D, however, accommodation responses were clearly sigmoidal. This pattern is consistent with our predictions of how the stimulus to accommodation would break down with increasing separation between the image planes (Figure 3).

**Variability in accommodation responses**

We also measured variability in accommodation responses in each condition. It is important to do this because the pattern of mean accommodation responses (Figure 7, top row) could result from accommodation switching between the two image planes, in a manner
Figure 7. Bias and variability in accommodation responses to depth-filtered stimuli. (Top) Mean accommodation response (across all observers) as a function of simulated focal distance. Each plot shows a different image-plane separation. The vertical green lines denote the positions of the image planes. The green data points show the accommodation response to “real” targets, in which 100% image intensity was at one plane. The blue data points are the accommodation responses to two-plane stimuli. The dashed red line in each plot is a linear fit between the accommodation responses to the one-plane stimuli. Error bars indicate ±1 between-subjects SEM. (Bottom) Root-mean-square deviation (RMSD), averaged across observers, as a function of focal distance, at each plane separation. The color coding is the same as in the top row. Error bars again denote ±1 between-subjects SEM.
the mean squared error (MSE) around a linear fit to the average data at each plane separation. The linear fit is very good for the 4/9 D plane separation, less good for the intermediate plane separations (6/9–10/9 D), and becomes substantially poorer for the 12/9 D separation. The bottom panel in Figure 9 plots the average residual for each plane-separation condition and simulated distance for a linear fit constrained between the average accommodation response at the near and far image planes. The residuals in Figure 9b are plotted as a function of the percentage of luminance drawn on the near plane to allow for direct comparisons across conditions.

Figure 8. Sample average accommodation responses for one observer, averaged across 10 repetitions, for three different ratios of far/near image intensity (see key). The image planes were 8/9 D apart. Error bars denote ±1 SEM.

Figure 9. Changes in accommodation responses with increasing image-plane separation. (a) Mean squared error of an unconstrained linear fit through the mean accommodation response at each image-plane separation tested. (b) Average residuals for linear fits constrained between the average responses to one-plane stimuli. Different colors denote different image-plane separations. The abscissa is plotted in units of image-plane intensity (not dioptric distance) to allow direct comparisons between residuals at similar near-to-far image intensity ratios. Open symbols denote residuals that were significantly different from zero (one-sample t-test, p < 0.05). Error bars show ±1 SEM, plotted in one direction for clarity.

The effect of plane separation on response bias

The changes in the accommodation stimulus–response function with increasing image-plane separation are explored in more detail in Figure 9. The top panel plots proportional to the image intensity at each plane. If this occurred, the response variability would be higher for two-plane stimuli than for stimuli on a single image plane. We computed the root-mean-square deviation (RMSD) for each trial and averaged across repetitions to give a measure of variability for each observer in each condition. The bottom row in Figure 7 plots these RMSDs, averaged across observers. There was no evidence for increased variability in responses to the two-plane stimuli (blue bars) compared to the single-plane stimuli (green bars), indicating that accommodation responses to depth-filtered stimuli did not oscillate between planes. Separate one-way ANOVAs for each plane-separation condition indicated that there were no significant differences between the variabilities of responses to one-plane and two-plane targets (p > 0.05). We also checked that accommodation responses remained stable during the 4-s averaging period because systematic shifts in different directions would not be detected by our RMSD analysis. To do this, we examined each observer’s mean accommodation response in each condition across the last 4 s of each trial. Sample average responses are shown in Figure 8 for three different simulated distances and a plane separation of 8/9 D. Aside from the different overall means, the three traces are very similar, with no evidence of any systematic effects across conditions. This was similar for all observers in all conditions.
between plane-separation conditions. For plane separations up to and including 10/9 D, none of the two-plane simulations resulted in residuals that deviated significantly from zero, indicating no systematic deviation from linearity in accommodation responses in these conditions. At the 12/9 D plane separation, however, the residuals show a clear pattern that is consistent with a sigmoidal accommodation stimulus–response function. The difference of the residuals from zero was statistically significant around the one-quarter and three-quarter distances between the planes. This is consistent with our predictions, because the model of retinal-image contrast showed that the stimulus to accommodation is least accurate for stimuli at one-quarter and three-quarter distances between image planes (Figure 3).

**Discussion**

Depth-filtered simulations of variations in focal distance resulted in continuous near-linear accommodation distance responses for image-plane separations of up to 10/9 D. Accommodation responses to two-plane and one-plane stimuli were also similarly precise. At the larger plane separation (12/9 D), accommodation responses were biased toward one or other image plane in a manner predicted by our analysis of retinal-image formation.

**Implications for eliminating accommodation–vergence conflicts**

These results show that multiplane displays, in conjunction with depth filtering, can generate an accurate and continuous stimulus to accommodation, using relatively widely spaced image planes. Accommodation and vergence demands could therefore be precisely matched in stereo displays with relatively few image planes. This is significant for the development of such displays because there are fundamental limits on the number of focal planes that can be used. In our display, which uses mirrors and beam splitters to create the focal planes, the maximum luminance at a given image plane is given by the luminance of the monitor, divided by the number of image planes. So large numbers of planes result in low luminance. In adaptive optics approaches, each focal plane is displayed in a time-multiplexed manner, synchronized with a variable-power optical element that changes the optical distance to a fixed display (Liu et al., 2008; Love et al., 2009; Shevlin, 2005).

To avoid visible flicker at each image plane, the image source and optical element must both operate at 60–70 Hz times the number of planes. It is therefore extremely challenging to construct displays with large numbers of focal planes, without yielding noticeable flicker. Increasing the number of planes also results in a decrease in the “duty cycle” of each plane, reducing the effective luminance. In the future, these problems may be solved by developments in transparent display media such as organic light-emitting-diode (OLED) displays (Schneidenbach et al., 2009), which would allow multiple emissive displays to be stacked behind one another. However, there remains a general constraint in that any increase in the number of focal planes requires an increase in computational power to generate the images (Akeley et al., 2004). Therefore, minimizing the number of focal planes (or more generally, the focal-depth resolution) of a stereo display is highly desirable. In this respect, our results are encouraging. We found that the stimulus to accommodation from depth-filtered images remained accurate with image-plane separations of 8/9 D, and even 10/9 D. Just five image planes, separated by 8/9 D, could eliminate accommodation–vergence conflicts at all distances from 28 cm to infinity (3.55 to 0 D).

In practice, the visual system has some tolerance to accommodation–vergence conflicts, and so a somewhat inaccurate accommodative stimulus (i.e., wider image-plane separations) may sometimes be acceptable. Accommodation and vergence responses can differ by quite large amounts without excessive effort or discomfort (~0.4 D, as indicated by Percival’s zone of comfort; Hoffman et al., 2008; Howard & Rogers, 2002; Morgan, 1944; Percival, 1920). Experiments on stereoscopic depth constancy have shown, however, that an inaccurate stimulus to accommodation can cause biases in perceived depth from disparity (Hoffman et al., 2008; Watt, Akeley, Ernst et al., 2005). Thus, even if observers can decouple their accommodation and vergence responses without discomfort, it may nonetheless be preferable to present a correct stimulus to accommodation.

Our results indicate that depth filtering is effective for controlling accommodation. It should be noted, however, that images from real targets, and depth-filtered approximations to the same focal distance, may have a different appearance. In depth-filtered images, contrast is relatively attenuated at (higher) spatial frequencies that are well inside the range to which humans are sensitive (Campbell & Robson, 1968; Figure 2) because at least one image plane is defocused wherever the eye is accommodated. This effect is worse with larger plane separations, but it remains substantial with 8/9 and even 6/9 D plane separation. This places an upper limit on the “sharpness” of the overall image that can be presented (the effect is worst at dioptic midpoint between planes, while distances that coincide with focal planes are unaffected). It also means that the perceived blur in depth-filtered images will likely be different to that in equivalent real scenes. Evaluation of these effects is necessary to understand their implications for depth perception and visual acuity. It may, for instance, be necessary to place image planes closer together for applications in which high spatial frequency image content is critical (such as medical imaging). In general, however,
it seems likely that the disadvantages of these inaccuracies in the image will be outweighed by the advantages of eliminating accommodation–vergence conflicts.

**Analysis of retinal-image formation**

Our analysis of retinal-image contrast (Figures 2 and 3) shows that the accuracy of the stimulus to accommodation from depth-filtered images becomes increasingly poor with larger plane separation and with higher spatial frequencies. Specifically, it becomes nonlinear: the accommodative state at which peak contrast is achieved becomes biased toward the image plane closest to the simulated distance (Figure 3), and at large plane separations (≥10/9 D) and high spatial frequencies, these peak-contrast functions become discontinuous, with peak contrast achieved only by accommodating to either image plane. In our experiment, we found that with increasing plane separation, accommodation responses to a spatially broadband stimulus became sigmoidal (Figures 7 and 9) in a manner that was qualitatively similar to the peak-contrast functions for mid-range spatial frequencies in Figure 3. This is therefore consistent with observers accommodating to distances that maximize retinal-image contrast at low to mid-range spatial frequencies, with relatively little influence of high spatial frequencies on the response (see Implications for accommodation control section). The similarity between changes in the peak-contrast functions and our data, and the fact that observers can, as predicted, accommodate to distances between image planes, even though there is no light originating from this distance, validates our approach to analyzing the retinal image. In principle, this approach could be used not only to predict accommodation responses to depth-filtered images, but also to other types of displays. Indeed, we hope wavefront analysis may allow design-stage assessment of the likely implications of different display techniques for the accommodation–vergence conflict. To make quantitative predictions, however, it is necessary to understand the contribution of different spatial frequencies to the accommodation response. We consider this in the next section.

**Implications for accommodation control**

As discussed earlier, it seems likely on theoretical grounds that the accommodation system adopts a fine-tuning strategy, using increasingly high spatial frequencies to guide the response as defocus decreases (Charman & Tucker, 1977, 1978; Kotulak & Schor, 1986b; Phillips, 1974). As Figure 2 shows, retinal-image contrast falls off increasingly quickly with defocus at higher spatial frequencies. Thus when the eye is quite defocused, there is no signal from high spatial frequencies, so low to mid-range frequencies must drive the response. However, as the eye becomes better focused, increasingly high frequencies could provide a good signal for fine-tuning (Charman & Tucker, 1978). Previous studies have shown, however, that spatial frequencies above 6–8 cpd do not contribute significantly to the accommodation response (Mathews & Kruger, 1994; Owens, 1980; Tucker et al., 1986; Ward, 1987). Our results are also consistent with a quite low upper limit on the spatial frequencies that drive the accommodation response. Figure 3 shows that at 8/9 D plane separation, the peak retinal-image contrast at ~8 cpd is already quite biased toward one or other image plane. Thus, if the steady state of accommodation is in part determined by high spatial frequencies, we would expect to already see a sigmoidal accommodation stimulus–response function at this plane separation. Contrary to that expectation, the observed responses are linear, again suggesting that spatial frequencies above 6–8 cpd do not contribute significantly to the accommodation response.

We can gain insight into the apparent contradiction between the theoretical and empirical data by extending the analysis in Figures 2 and 3 to consider how retinal-image contrast changes as the system attempts to focus. Accommodation during target fixation fluctuates within an envelope of up to 0.5 D, in the frequency range ~0–3 Hz (Campbell, Robson, & Westheimer, 1959; Charman & Heron, 1988). It has been hypothesized that the changes in retinal contrast that result from these microfluctuations might provide a signal to accommodation (Alpern, 1958; Campbell et al., 1959; Charman & Heron, 1988; Hung, Semmlow, & Ciuffreda, 1982). We quantify this signal in Figure 10.

The analysis in Figure 10 is based on the same eye’s optics used in Figures 2 and 3. Figure 10a shows retinal contrast as a function of defocus for a range of spatial frequencies. The curves are based on the real-target MTF functions used to compute Figure 2, normalized by assuming an incident contrast of 1.0 at all spatial frequencies. Figure 10b plots the change in retinal contrast (∆C = Cmax − Cmin) that results from an accommodation microfluctuation of 0.5 D (±0.25 D) at each defocus state in Figure 10a. This analysis clarifies why different spatial frequencies might contribute differently to accommodation with different amounts of defocus. We assume that a greater magnitude of change in retinal contrast indicates a larger signal to accommodation. Thus, for an initially defocused (~1.5 D) eye, the retinal signal created by microfluctuations is highest at 2 cpd. As the eye focuses to ~0.75 D, the low frequency signal decreases and the mid-range (4–6 cpd) signal increases. With still better focus the signals at higher spatial frequencies (~8 cpd) increase significantly, and so would be expected to contribute more to the accommodation response. Critically, however, higher spatial frequencies do not provide an ever larger signal to accommodation as the eye nears correct focus. Instead, because high spatial frequencies are highly attenuated by the eye’s optics, the signal to accommodation decreases above ~5 cpd.

The overall pattern of changes in retinal-image contrast with accommodation microfluctuations is summarized in
Figure 10c, which plots change in retinal contrast as a function of defocus (Figure 10b), at all spatial frequencies. This plot suggests that the signal to accommodation is largest from spatial frequencies around 4–6 cpd, when the eye is defocused by \(~0.5\) D. We also conducted the same analysis on several other eyes, with differing optics, and found a very similar pattern of results (see Appendix B for examples).

This analysis seems to reconcile the fine-tuning model of accommodation control (Charman & Tucker, 1978) with the empirical data on the contribution of different spatial frequencies to the accommodation response. Consistent with the fine-tuning hypothesis, our analysis suggests the signal to accommodation must come from increasingly higher spatial frequencies as the eye becomes more focused. However, there is a roll-off in the useful range of spatial frequencies far below what has previously been estimated (Charman & Tucker, 1978). Our analysis suggests this roll-off occurs at around 4–6 cpd, which is highly consistent with the empirical finding that spatial frequencies above 6–8 cpd have little effect on accommodation responses (Mathews & Kruger, 1994; Owens, 1980; Tucker et al., 1986; Ward, 1987; our data).

It can be seen (most clearly in Figure 10b) that additional spatial frequencies above \(~8\) cpd add relatively little to the overall signal to accommodation as the eye nears perfect focus. Thus, additional high spatial frequencies would not be expected to significantly improve responses in the presence of mid-range frequencies. The fine-tuning hypothesis has been characterized as a process of gradually “handing over” the control of accommodation to increasingly higher spatial frequency channels, as the eye becomes more focused (Kotulak & Schor, 1986b). Our analysis suggests instead that the system could work effectively by summing across the whole range of relevant spatial frequencies, because the signal to accommodation from mid-range spatial frequencies continues to be relatively large as the eye approaches zero defocus.

Comparison to Charman and Tucker (1978)

Our analysis is similar to that of Charman and Tucker (1978) but produces quite different results due to two important differences. First, we characterized the process of retinal-image formation based on measured aberrations of a real eye—an approach that makes relatively few assumptions. Charman and Tucker (1978) modeled an aberration-free eye, leading them to underestimate significantly the extent to which high spatial frequencies, in particular, are attenuated when the eye is well focused (their Figure 4a). In turn, this lead to an overestimate of the change in retinal contrast at high spatial frequencies with changes in accommodative state. Second, we quantified the signal to accommodation by computing the absolute change in retinal contrast that would result from a microfluctuation of a given magnitude. Charman and Tucker...
Accommodation to natural images

Figure 10 describes the case in which there is equal contrast at all spatial frequencies. In natural images, the relative contrast at different spatial frequencies is well described by an approximately 1/f distribution (Field, 1987), meaning there is relatively low contrast at high spatial frequencies. Accommodation microfluctuations would therefore result in smaller retinal contrast changes at high spatial frequencies than shown in Figure 10. We would therefore expect that high spatial frequencies would have a smaller effect on the accommodation response when viewing natural images than when viewing gratings of equal contrast.

Similarly, retinal contrast in multiplane images is significantly attenuated at higher spatial frequencies, relative to equivalent real targets (Figure 2). The overall accommodation response to depth-filtered images presumably reflects a combination of the contributions of a reasonably accurate low spatial frequency signal (driving the response to the “correct” place), and an increasingly inaccurate mid-range to high spatial frequency signal (driving the response toward one or other image plane). Reduced contrast at high spatial frequencies, therefore, works in favor of large plane separations, because it reduces the inaccurate part of the signal to accommodation. It follows that higher spatial frequencies should have a greater influence on responses to real targets than to depth-filtered images.

Conclusions

Our results suggest that multiplane images, in conjunction with depth filtering, can produce continuous non-linear variations in the stimulus to accommodation. Critically, this is possible using image-plane separations large enough for a practical display. These findings support previous psychophysical studies suggesting that depth-filtered images can reduce many of the problems caused by accommodation–vergence conflicts in conventional stereo displays (Akeley et al., 2004; Hoffman et al., 2008; Watt, Akeley, Girshick et al., 2005). Indeed, our results suggest that this technique allows variations in accommodation and vergence demand to be precisely matched in stereo displays, just as they are in natural viewing.

We also described an analysis of retinal-image formation that allows us to make predictions about accommodation responses to depth-filtered images. We found that accommodation responses broke down with increasing image-plane separation in a manner consistent with this analysis, suggesting that this approach may be useful for understanding accommodation responses to current and future displays. Our analysis also suggests that the signal to accommodation from high spatial frequencies is much poorer than might be expected. This observation seems to reconcile apparently contradictory theoretical and empirical accounts of the contribution of different spatial frequencies to the control of accommodation.

Appendix A

Analysis of retinal-image formation

The analysis of retinal-image formation in Figures 2 and 3 is based on one eye, chosen because it had aberrations in the typical range (see Effect of different optics section). To characterize this eye’s optics, we measured its aberrations in a relaxed accommodation state using a Shack–Hartmann wavefront sensor and parameterized these aberrations with Zernike polynomials (Mahajan, 1994). We modeled variable accommodation and image-plane position by manipulating the Zernike defocus term (C_{2,0} coefficient) and the spherical aberration term (C_{4,0} coefficient). During accommodation, other higher order aberrations subtly change, but their changes tend to be mostly random (see Cheng et al., 2004; Plainis, Ginis, & Pallikaris, 2005). We assumed a change in the spherical aberration Zernike term (C_{4,0} coefficient) of 0.048 μm/D of accommodation (Plainis et al., 2005). Commensurate with the aberration changes, we modeled a constriction of the pupil diameter of 0.18 mm/D of accommodation (Plainis et al., 2005), with a pupil size of 4.9 mm when accommodation was 2/3 D (the average pupil size during our experiment). Note that in this case, the changes in the optics with accommodation and pupil size are relatively minor because of the small range of accommodative distances considered.

We computed the retinal-image contrast for the depth-filtered images by summing the point-spread functions (PSFs) corresponding to the accommodative state and the defocus of the two image planes, with weights proportional to the image intensity on each image plane (Hoffman et al., 2008). To include the eye’s chromatic aberration, we computed preliminary PSFs for a number of visible wavelengths with added chromatic defocus and summed them with the 1924 CIE Standard Observer V(λ) weighting (the crystalline lens is effectively 2.5 D stronger for 400-nm light.
than 700-nm light; Bedford & Wyszecki, 1957). We computed the optical transfer function (OTF) for each viewing situation by taking the 2D Fourier transform of the composite PSF. We collapse this function to a 1D modulation transfer function (MTF) by taking radial averages of the magnitude of the OTF. The MTF specifies the ratio of the contrasts of the retinal image to the incident stimulus image (retinal contrast/incident contrast) at every spatial frequency.

**Effect of different optics**

We confirmed that the conclusions we drew from Figures 2 and 3 were generalizable by performing the same calculations for eight additional eyes, with a range of optical qualities.

Figure A1 shows the third- to sixth-order Zernicke coefficients characterizing the higher order aberrations for the eye used in Figures 2 and 3 (BR) and two other examples (TP and EB). We omitted second-order aberrations (astigmatism and defocus) because they are correctable with simple spectacles. The optical quality of an eye is

![Figure A1](image-url)

Figure A1. Individual aberration data for different eyes with 5-mm pupil size. The symbols denote the values of the third- to sixth-order Zernike coefficients for BR (blue circles), TP (red diamonds), and EB (green squares). The analyses in Figures 2 and 3 (and Figure 10) are based on BR’s optics. The aberration measurements were made by Austin Roorda at UC Berkeley.

Figure A2. Retinal-image formation for eyes with different aberration patterns. The left and right panels show the analyses for TP and EB’s eyes, respectively. The left column in each panel shows contrast ratios for real targets and depth-filtered images as a function of where the eye is focused (as Figure 2). The dashed curves show retinal-image contrast from viewing a real target, at the location shown by the dashed gray line. The solid curves show retinal contrast from viewing the depth-filtered approximation (two-plane stimulus). The positions of the image planes are shown by the vertical green lines (see also icons above each subplot, which also show the near/far intensity ratio). Different colored curves denote different spatial frequencies The right panels show the accommodative state at which peak contrast is maximized, when viewing depth-filtered stimuli, as a function of simulated distance (as Figure 3). The top and bottom rows show image planes separated by 6/9 D and 10/9 D, respectively.
difficult to quantify with a single metric. An aggregate wavefront metric is the RMS, which specifies the root-mean-square perturbation of the wavefront from a planar surface. When we zeroed out the second-order aberrations, we computed residual RMS values of 0.16 for BR, 0.19 for TP, and 0.33 for EB. Based on these pupil-plane wavefront metrics, BR and TP had comparable optics, while EB was significantly worse. However, these pupil-plane metrics are poor predictors of image quality, which is best computed from the PSF. One image-quality metric is the Strehl ratio, which specifies the ratio of the peak of the aberrated PSF to the peak of the diffraction-limited PSF. With a 5-mm pupil, the Strehl ratio for BR was 0.109, 0.103 for TP, and 0.034 for EB.

BR and TP have comparable and slightly above average retinal-image quality, and EB has poorer retinal-image quality (Porter, Guirao, Cox, & Williams, 2001; Thibos, Bradley, & Hong, 2002). Choosing BR as our exemplar was conservative, because the stimulus to accommodation from depth-filtered images breaks down at smaller image-plane separations, and lower spatial frequencies, with better optics. Higher quality optics improve the peak contrast ratio for an eye at the expense of depth of field; defocus will decrease image quality more severely in a high-quality eye than a heavily aberrated eye. We included analyses of TP with typical aberrations and EB with higher than typical aberrations to demonstrate that our general conclusions are robust to the observers’ unique sets of aberrations.

Figure A2 shows the results of the analyses in Figures 2 and 3 for TP and EB’s eyes. The right subplots in each panel show the retinal-image contrast that results from viewing real and depth-filtered stimuli, as a function of where the eye is focused. The right subplots show the accommodative distance at which contrast is maximized, when viewing depth-filtered stimuli, as a function of simulated distance. Although there are substantial differences in the individual patterns of aberrations, the accuracy of the stimulus to accommodation is similarly dependent on image-plane separation and spatial frequency for all of the eyes. With image planes separated by 6/9 D, the position of maximum contrast changes nearly linearly with simulated position up to quite high spatial frequencies. At 10/9 D plane separation, the curves become nonlinear for spatial frequencies higher than 4 cpd. EB’s peak-contrast curves have vertical offsets at different spatial frequencies because of her unique aberrations, but these offsets are present for both depth-filtered and real targets (Green & Campbell, 1965; see left subplots for EB in Figure A2).

Appendix B

Stimulus to accommodation with different optics

Our analysis of the change in retinal-image contrast with accommodation microfluctuations (Figure 10) suggests that the peak signal to accommodation occurs at ~4–6 cpd (Figure 10c). Thus, there is a roll-off in the useful range of spatial frequencies at quite low frequencies. These conclusions are based on one eye’s optics (BR), and to examine their generality we carried out the analyses in Figure 10 for the eight other eyes for which we have aberration measurements, and all showed a similar roll-off. Figure B1 gives two examples, based on the same eyes described in Appendix A (TP and EB). Each panel in Figure B1 plots the change in retinal contrast with an accommodation microfluctuation (±0.25 D) as a function of defocus, at all spatial frequencies. These plots are in the same form as Figure 10c. Despite differences in the pattern of aberrations, and therefore

Figure B1. Change in retinal contrast as a function of defocus and spatial frequency for different eyes. These plots are in the same format as Figure 10c and show the change in retinal-image contrast for a ±0.25 D accommodation microfluctuation, as a function of defocus (abscissa) and spatial frequency (ordinate). The left and right panels show the analysis for TP and EB, respectively.
overall optical quality, the roll-off in the signal to accommodation is still present at a similar point (between 4 and 6 cpd). Across the nine eyes we analyzed, the peak contrast change occurred at a mean of 4.57 cpd, with a standard deviation of 0.53 cpd. Note that while EB's relatively poor optics result in lower overall retinal contrast changes than either BR or TP, the peak signal still occurs at around the same spatial frequency. Thus, our estimate of the roll-off in the contribution of higher spatial frequencies to the accommodation response appears to be relatively invariant across the range of normally aberrated eyes.

Acknowledgments

Thanks to Austin Roorda, for making the wavefront aberration measurements and for assisting with the accompanying analysis, and to Tecwyn Davies (Industrial Developments Bangor) for technical assistance building the display. Thanks to Kurt Akeley and Marty Banks, for many helpful discussions and for comments on the manuscript, and also to two anonymous reviewers. This work was supported by a grant to SJW from the Engineering and Physical Sciences Research Council.

Commercial relationships: none.
Corresponding author: Simon J. Watt.
Email: s.watt@bangor.ac.uk.
Address: School of Psychology, Bangor University, Adelaid Brigantia, Penrallt Rd., Bangor, Gwynedd, Wales LL57 2AS, UK.

References


