Adaptation to blurred and sharpened video

Andrew M. Haun
Schepens Eye Research Institute, Massachusetts Eye & Ear, Harvard Medical School, Boston, MA, USA

Eli Peli
Schepens Eye Research Institute, Massachusetts Eye & Ear, Harvard Medical School, Boston, MA, USA

The visual system can distinguish different levels of blur and different levels of excess sharpness. Adaptation alters this capacity so that the adapted blur (or sharp) level looks more like a normal, properly focused image. Here, we describe the more general pattern of aftereffects of blur and sharp adaptation by measuring matching functions, using video clips from a DVD movie as stimuli. Results show that blur and sharp adaptation are selective: The sharpening aftereffects of blur adaptation are strongest for blurry videos while the blurring aftereffects of sharp adaptation are strongest for sharp videos. Despite the spatiotemporal variability of our adaptor and test stimuli, we found adaptation effects similar in magnitude to previous studies using invariant static images. A recent model of blur adaptation can be simplified to explain the form of our data, leading us to conclude that what we see as blur/sharp adaptation is a consequence of narrowband contrast adaptation.

Introduction

When image details are attenuated physically by blurring and lost to perception, the visual system adapts so that some of the attenuated details can be recovered. This process of blur adaptation is known to occur over long and brief time scales. The principal example of the former is the emmetropization of the developing eye (Wallman & Winawer, 2004), which consists of a change in the physical focus of the retinal image. The latter process is entirely neural and can be demonstrated psychophysically in a matter of seconds (Webster, Georgeson, & Webster, 2002). The speed of the adaptation described by Webster et al. is one of many reasons to suppose that the process is a direct outcome of contrast adaptation (Elliott, Georgeson, & Webster, 2011; Ohlendorf & Schaeffel, 2009), which has a time course on the order of seconds (Greenlee, Georgeson, Magnussen, & Harris, 1991). Adaptation over such brief time scales is thought to permit effective neural representation of the current stimulus distribu-

tion given a limited response range, thus allowing for the fact that consequent contrast distributions may be very different—this benefit is often referred to as *normalization* (Frazor & Geisler, 2006; Heeger, 1992; Wainwright, 1999). A further link between contrast sensation and blur adaptation has been demonstrated through the correlation between individual contrast sensitivity functions and blur/sharp adaptation magnitudes (Vera-Diaz, Kark, Woods, & Peli, 2011).

Studies of contrast adaptation using spatial two-alternative, forced-choice matching or discrimination methods with a blank-adapted standard (Abbonizio, Langley, & Clifford, 2002; Blakemore, Muncey, & Ridley, 1973; Elliott et al., 2011; Georgeson, 1985; Ross & Speed, 1996; Snowden & Hammett, 1996) have established that psychophysical contrast adaptation has a retinotopically localized effect (otherwise such experimental methods would be useless) although the limits of this localization has never been established. Webster et al.'s (2002) study of blur adaptation using different images as tests and adaptors suggested incomplete transfer of adaptation when the spatial phase structure of the adaptor and test stimuli were different. Vera-Diaz, Goldstein, and Peli (2008) also found incomplete transfer of adaptation between dissimilar adaptors and tests. Most other studies applying the same paradigm have used similar adaptor and test stimuli (Elliott et al., 2011; Sawides et al., 2010; Vera-Diaz, Woods, & Peli, 2010; Webster et al., 2002). Some studies of blur adaptation have induced adaptation with optical defocus of videos or other realistic scenery and tested its effects on blur sensitivity (Wang, Ciuffreda, & Vasudevan, 2006), acuity (Rosenfield, Hong, & George, 2004), or contrast sensitivity (Rajeev & Metha, 2010) using simple stimuli. Yet it is still unclear just how important it is for blur adaptation—or for contrast adaptation—that test and adaptor contrasts have the same retinal location or spatial phase.

For blur adaptation to be useful in normal visual experience, it should generalize across successively

viewed images; it does not seem reasonable that the system should have to readapt to the same level of blur from fixation to fixation. The reason why transfer of adaptation seems incomplete in the cases in which it has been sought experimentally may be due to the relatively low probability that two randomly selected scenes will have many overlapping areas of contrast, so test contrasts will often fall in unadapted regions of the visual field. In normal visual experience, no point in a scene is fixated for more than a few hundred milliseconds, so we should expect that if the adaptors are spatiotemporally variable enough we should see transfer of adaptation to any test (as in the defocus blur studies cited above, in which transfer was seen from videos or natural scenes to acuity or contrast sensitivity targets).

In the current study, we used a matching task (similar to Elliott et al., 2011) to measure the effects of blur and sharp adaptation, enabling us to measure not only normalization (a shift in the perceived normal toward the quality of the adaptor), but also the aftereffects of adaptation over a range of blur and sharp levels. In our experiments, both test and adaptor stimuli were video clips from DVD movies that varied within and between trials. Despite the fact that our adaptors and tests were never the same, we measured adaptation effects of the same magnitude as those measured in single-pattern experiments in similar paradigms (Elliott et al., 2011; Vera-Diaz et al., 2010; Webster et al., 2002). We found that relative to adaptation to normal video, adaptation to blurred or sharpened video resulted in both a shift and a compression of the matching functions, and adaptation to a blank screen had a similar effect to adapting to blurred or sharpened video. Using a quick method of measuring blur-sharp matching functions (Lesmes, Jeon, Lu, & Dosher, 2006), we then measured blur/sharp adaptation in a larger number of subjects to test the generality of the finding.

To explain the observed pattern of results, we modified the model of blur adaptation proposed by Elliott et al. (2011) so that it could be used to predict the effects of blur, normal, oversharp, and blank adaptation within our experimental design. Our analysis suggests that there is no blank-field "long-standing normal adaptation" of the type invoked by Elliott et al., and a simpler model can capture all the basic features of blur adaptation with more complex stimuli.

General methods

Subjects

Three subjects participated in the main experiment. S1 was the first author while S2 and S3 were naive to the purposes of the experiment and relatively inexperienced as psychophysical observers. Nine more subjects were added later; these were a mix of experienced and inexperienced subjects, and all were naive to the purposes of the experiment. The age range for all subjects was 22 to 49 with a median of 27. All subjects had normal or corrected-to-normal visual acuity.

Stimuli

Video stimuli were video segments from a DVD movie. Fifty 3-s (72-frame) segments not containing scene cuts were extracted, color information was discarded, and the central 360×360 pixel region of each segment was cropped from the frame and processed for use in the experiments. While the paper was in review, we found that we had used the wrong aspect ratio—horizontally narrower by a factor of 1.2—to convert the H.264 video to .avi, resulting in the videos having a slightly squashed appearance, which had gone unnoticed by us (but which was apparent to an observant reviewer in the frames shown in the original version of Figure 1). Each component in the Fourier transform of each frame of each movie segment was multiplied by $f^{\Delta s}$: f is the spatial frequency of the fast Fourier transform component, and Δs is the specified change in log-log slope of the amplitude spectrum, ranging from -1.0 (blurred) to $+1.0$ (sharpened) in steps of 0.1. The spatial transform of the altered spectrum was then set to have the same mean and standard deviation (RMS contrast) as the original frame. The unaltered s values ($\Delta s = 0$) of the video frames averaged $-1.46 \pm .22$, in the normal range for real-world images (Bex, Solomon, & Dakin, 2009; Field, 1987; Hansen & Essock, 2005). Stimuli were displayed on a linearized CRT display (Trinitron) set at 2.1 pixels/mm and a 144 Hz refresh rate, viewed at 1 m, so each video subtended H/V 9.8° of visual angle. Matlab and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997) were used to design and run the experiments.

Procedure

The adaptation period was 30 s (10 3-s video segments) prior to the first trial. Presentation duration in the trial interval was 500 ms, and the subjects responded as quickly as possible. During the response period, the screen was blank; the next adaptor appeared directly after the subject response. Subsequent trials were preceded by a 3-s readaptation (top-up) period (the same scheme as in Vera-Diaz et al. [2010] but shorter than the 180:6 s adapt:top-up of Webster et al. [2002] and the 120:6 s scheme of Elliott et al. [2011]). Stimuli in each trial consisted of a pair of 12-frame sequences presented in the blank- and video-
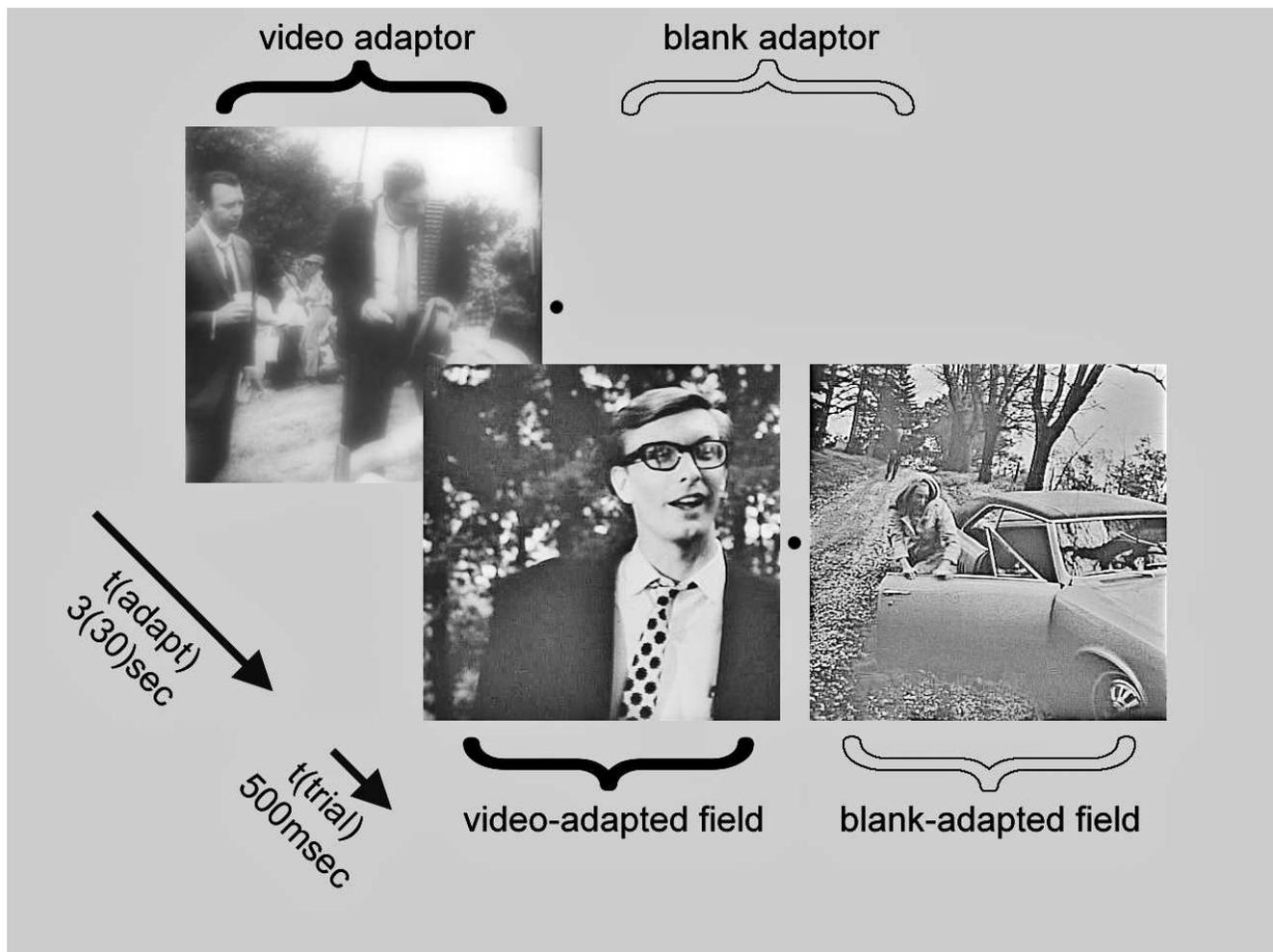
Figure 1. Experimental trial design. On each trial, the subject adapted for 3 seconds (30 seconds before the first trial) by fixating a spot between two adaptors. The video adaptor could be blurred (as shown here), sharpened, or unaltered, depending on the condition/block. The blank adaptor was an empty mean-luminance area. After the adaptation period, two different stimulus videos appeared in the two adapted fields for 500 ms. The subject chose one of the two stimuli as "sharper," and the next adapt period began. Frames are from *Night of the Living Dead* (1968), which is in the public domain; different movies were used in the experiments.

adapted regions of the display (Figure 1), selected with a random starting point from within a randomly selected 72-frame source segment. Within a trial, stimuli were never drawn from the same segment (i.e., each trial used three different video segments). Subjects maintained fixation on a green spot at the center of the display and responded by indicating which of the two trial videos seemed *sharper*.

Trials were blocked by an adaptor condition with each block repeated three times by each subject. Posture and head position were maintained with a chin rest. In separate blocks of trials, the subject adapted to blurred, normal, or sharpened video (Δs values of −0.5, 0.0, or +0.5, respectively) with subject S1 adapting to two additional levels of blur and sharp (−1.0 and +1.0). We devised the following scheme for interleaving multiple standard Δs levels on both sides of the display (video and blank adapted) so that matches were made

in both locations during the same block of trials (illustrated in Figure 2). For each adaptor condition, six Δs matches between the blank- and video-adapted fields were obtained through six randomly interleaved one-up, one-down staircases, running for 30 trials each with a step size of $\Delta s = 0.1$. Three of the staircases varied the Δs of stimuli in the video-adapted field to estimate matches to one of three "standard" Δs values (−0.2, 0.0, +0.2) in the blank-adapted field. The other three staircases followed the opposite arrangement, varying the Δs of stimuli in the blank-adapted field paired with standards in the video-adapted field. Subjects were instructed to choose the video that appeared sharper. Trials thus collected were fit with logistic functions that represented the probability that the test stimulus was chosen as the sharper of the two (test and standard):
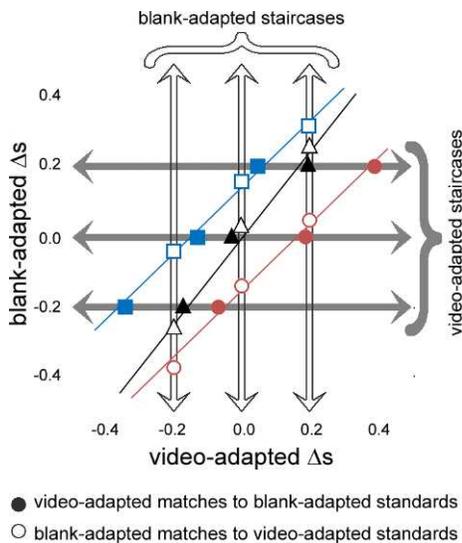
Figure 2. Staircase design. In one block of trials, six interleaved staircases (horizontal and vertical arrows) controlled the Δs values in one of two regions of the visual field. Blank-adapted staircases (open vertical arrows) controlled the Δs of videos in the blank-adapted field with respect to three fixed Δs standards in the video-adapted field, and video-adapted staircases (solid horizontal arrows) controlled the Δs of stimuli in the video-adapted field with respect to standards in the blank-adapted field. Solid symbols indicate video-adapted matches to blank-adapted standards, and open symbols indicate blank-adapted matches to video-adapted standards for three hypothetical underlying matching functions obtained under three video adapting conditions (thin colored lines).

$$p(\Delta s_{test}) = \gamma_{blur} + \frac{1 - \gamma_{sharp} - \gamma_{blur}}{1 + \exp\left(-(\Delta s_{test} - \Delta s_{match})/\lambda\right)} \quad (1)$$

The value of Δs_{test} yielding 50% in Equation 1 was taken as the match to the corresponding standard. The value of λ set the spread of the fitted psychometric function. The lapse rate parameters γ_{sharp} and γ_{blur} set the upper and lower bounds for the psychometric functions; trial blocks in which either lapse parameter exceeded 5% were repeated and excluded from data analysis (these were rare). Final estimates were obtained from the full set of trial data (90 trials per data point). This procedure resulted, for each adapting condition, in six matches made between video-adapted and blank-adapted regions of the visual field. These matches fall along the *line of subjective equality* between stimuli in the video-adapted and blank-adapted fields (blue, black, and red lines as illustrated in Figure 2).

Results

Matching functions

Results for three subjects are shown in Figure 3. Video-adapted Δs is plotted along the horizontal axis and blank-adapted Δs along the vertical axis. Straight lines were fitted to the matching data, with the form
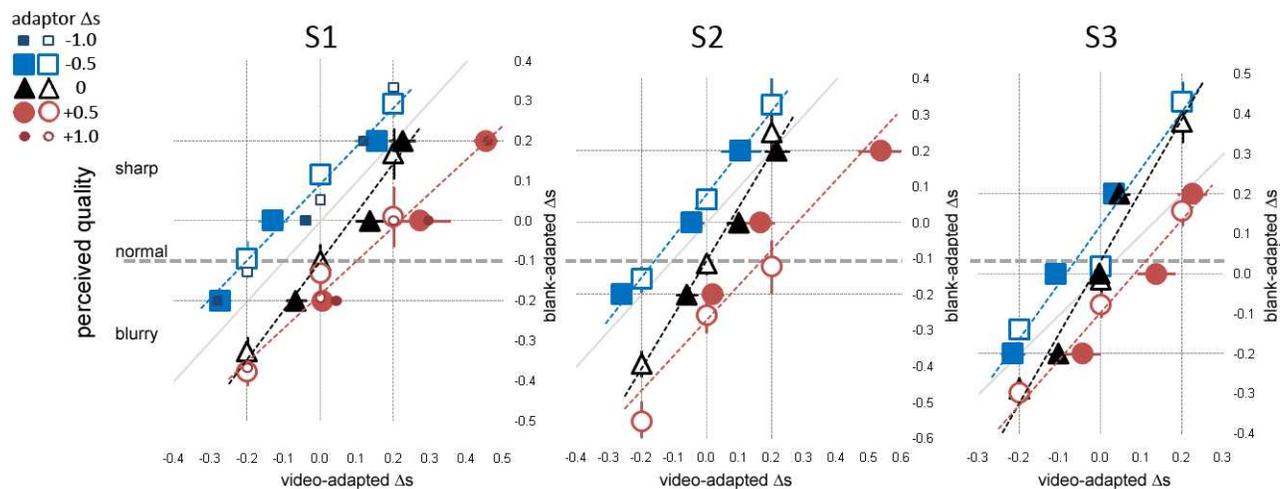

Figure 3. Matching functions measured for three subjects. Different symbols represent different video adaptor conditions as identified in the legend. Video-adapted Δs is represented on the abscissa, blank-adapted Δs on the ordinate. Increasing values on both axes indicate increasing physical sharpness (or decreasing physical blur) with the data points indicating where stimuli have the same apparent quality. Thin gridlines represent the standard Δs values. Solid symbols are video-adapted matches made to blank-adapted standards and thus are located along the horizontal gridlines. Open symbols are blank-adapted matches to video-adapted standards and are located along vertical gridlines. Error bars, when visible, are bootstrapped 68% confidence intervals (Wichmann & Hill, 2001). Individual plots are aligned vertically to each subject's perceived normal, estimated using the fitted lines and represented by the dashed gray line. If it is valid to suppose that the blank-adapted field is the same in all conditions (see text), then the ordinate axis can be understood as indicating a fixed perceptual continuum from blurry through normal (dashed gray line) to oversharp.

$$\Delta s_{blank} = m\Delta s_{video} + b \qquad (2)$$

In Equation 2, the subscripts identify the adapting field. When the video adaptor $\Delta s = 0$, we assumed that a stimulus in the video-adapted field with $\Delta s = 0$ should appear *normal*, and Equation 2 shows that this will be matched to a blank-adapted stimulus with $\Delta s = b$. The intention of the two-field matching design is that the blank-adapted field provides a fixed "external standard" to which the effects of different video-adapted states can be compared. If we assume that a blank has consistent adaptive properties across all video adaptor conditions, then we can consider the vertical axis of Figure 3 to represent a perceptual blur-normal-over-sharp continuum running from blurry through normal to oversharp. We can then mark out, on the vertical axis, the blank-adapted match to the normal-adapted normal (i.e., b when the adaptor $\Delta s = 0$) in the matching plots: This is represented by the dashed gray horizontal line in each panel of Figure 3. This line thus illustrates the location of the perceived normal on the vertical axis. The individual subject data axes in Figure 3 have been shifted vertically so that the perceived normal is aligned for all three subjects.

Horizontal differences between differently adapted matching functions in Figure 3 can be understood as changes, from one state of adaptation to another, in the stimulus on the video-adapted axis that are necessary to maintain a match to the perceptual quality of a stimulus on the blank-adapted axis. These *normalization effects* are measured directly in the "internal standard" paradigm of Webster et al. (2002); we discuss the relationship between the different paradigms in the Appendix. Vertical differences between different matching functions can be understood as changes in the perceptual quality of a stimulus on the video-adapted axis that result from changes in the state of adaptation; these are often referred to as the *aftereffects* of adaptation. For example, a $\Delta s = 0$ stimulus falls on the "perceived normal" line during normal adaptation (adaptation to $\Delta s = 0$, black triangles) but is sharpened (shifted upward) during blur ($\Delta s = -0.5$) adaptation (blue squares). Blur-adapted matching functions were shifted completely above the normal-adapted functions, meaning that adaptation to blur made stimuli appear sharper. This sharpening aftereffect was greater for blurred stimuli than for sharpened stimuli, indicating selective adaptation (Elliott et al., 2011). Sharp adaptation (adaptation to $\Delta s = +0.5$, red circles) had the opposite effect, shifting the matching function downward, meaning that sharp adaptation made stimuli appear less sharp with the blurring effect stronger for sharpened than for blurred stimuli.

The identity line in Figure 3 (diagonal gray line) can be taken to represent "blank matching," which is where the matches would fall if video adaptation yielded the same perceptual effect as blank adaptation. For two of
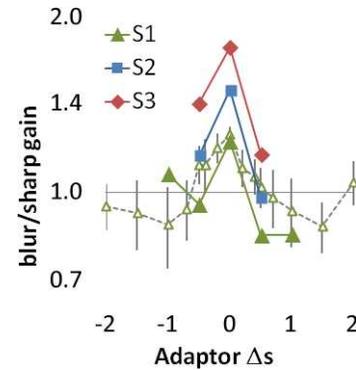

Figure 4. Vertical axis shows the perceived change in blank-adapted Δs with perceived change in video-adapted Δs, parameter, the *blur-sharp gain*. Horizontal axis shows the adaptor Δs value. Solid symbols show slopes for regression lines plotted against individual data in Figure 3; open symbols show slopes measured for subject S1 using the adaptive method described below.

the three subjects (S1, S2), the normal-adapted matching functions were below the identity line (b was negative). This suggests that for these subjects blank adaptation was akin to blur adaptation with the perceived quality of blank-adapted stimuli sharpened relative to normal-adapted stimuli. For the third subject, the blank-adapted perceived normal was not significantly different from the video-adapted perceived normal (the dashed gray line passed within measurement error of the origin).

Individual variation in video adaptation

The slope of the matching functions (m in Equation 2) can be taken to describe a subject's "blur-sharp gain" during a given adaptation state, i.e., how much of a *perceptual* change results from a given *physical* change. The (hypothetical) blank-adapted matching function, by definition, has a gain of 1.0. The gains of the matching functions in Figure 3 are plotted in Figure 4—a peak in gain for normal adaptors (the steeper matching functions of Figure 3) is apparent for all three subjects. The three individual gain functions are very similar except for a vertical shift; a simple explanation for this is that all three subjects adapted similarly to the video adaptors but differently to the blank standard that sets the position of the y-axis. At any rate, this pattern makes it appear that the blur-sharp gain is highest when the subject is adapted to an in-focus image. We wanted to determine more conclusively whether this normal-adapted gain peak is a general feature of human performance in this task, and so we obtained data from more subjects.

The method used to collect the data in Figure 3 is time consuming: Each matching function required at

least 2 hr of testing time for completion. Conveniently, the linearity of the matching functions gives us a simple model of matching performance—i.e., a straight line with offset and slope—whose parameters can be sought strategically. We employed an adaptive algorithm to directly measure the function parameters of interest (Lesmes et al., 2006)—we will refer to this quick method as "QM". Use of this method requires that the model calculate the probability of a response given a particular pair of stimuli (video- and blank-adapted). We used the linear model of Equation 2 to set the mean of the same logistic function used to estimate individual matches from the original staircase data (Equation 1), computing response probabilities with respect to the blank-adapted field:

$$
p(\Delta s_{blank})
$$
$$
= \frac{1 - 2\gamma}{1 + \exp\left(-\left(\Delta s_{blank} - (m\Delta s_{video} + b)\right)/\lambda\right)} + \gamma \quad (3)
$$

Here, m is the blur-sharp gain as described above, b is the adaptor aftereffect on a normal video-adapted stimulus (the y-intercept of the lines in Figure 3), and the parameter λ is the spread of the psychometric function. The normalized Δs value (see Appendix) is equal to $-b/m$ and denoted as β below. The expected lapse rate γ was fixed at 2%, close to the average value of 2.4% yielded by the three subjects in the first experiment (excluding rates for excluded data, cf. *procedure*). This model, with two stimulus values and three free model parameters, thus, has the same dimensionality as the QuickTvC model of Lesmes et al. (2006). Simulations applying model-matching functions similar to those obtained in the main experiment were used to adjust the grain of the QM to make it efficient and accurate in estimating the parameters of interest (Lesmes et al., 2006). Blank-adapted test values (Δs_{blank}) were not selected from the QM distribution, but from a uniform distribution of Δs values ranging from −.6 to +.6 in steps of 0.2. This was done because the algorithm otherwise would try to find the straight-line model parameters by placing trials only at the most extreme available stimulus pairs, which would limit the interpretability of the data as "matching functions." Subject S1 performed the QM task for a large number of adaptor values (open symbols in Figure 4)—the coherence of the results speaks to the repeatability of the QM measurements.

Matching function slopes (m) and normalization offsets (β) were obtained for nine additional subjects for three adaptor conditions ($\Delta s = -0.5$, 0.0, and +0.5) and are illustrated by the box and whisker plots in the top row of Figure 5 (the other symbols are for model data as explained in the next section). Parameter values from the original three subjects are also included in the data used for the box plots. The peaked-at-normal
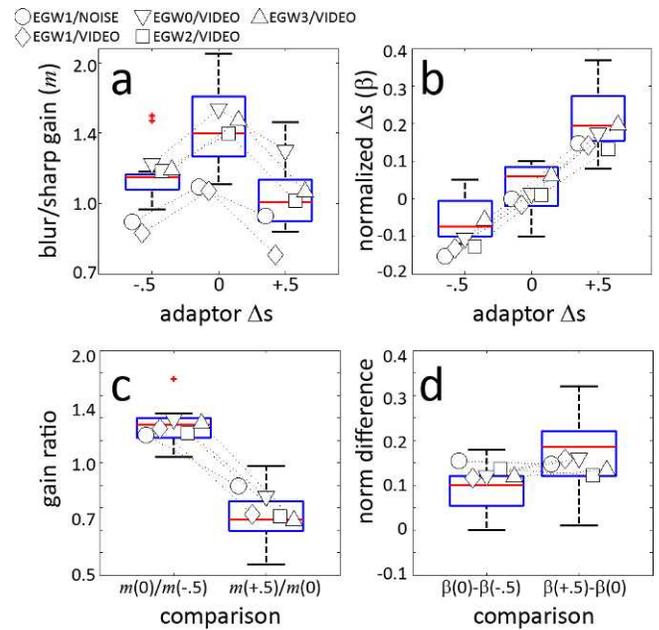

Figure 5. (a and b) Box-and-whisker plots illustrating the distribution of linear matching function parameters for 12 subjects (three from the original experiment shown in Figure 3, nine with the QM). Red lines are median values; upper and lower box bounds are upper and lower quartile values; whiskers are values closest to 1.5 times the upper and lower quartiles; red plus signs are outliers. Symbols are parameters from model-matching functions (identified in the legend at top). Symbols are spread out horizontally to improve visibility. Models EGW2 and EGW3 are discussed in the Appendix. (c and d) Illustration of the distribution of subject-by-subject differences in each parameter. (c) Ratios of normal-adapted to blur-adapted function slopes ($m_0/m_{-.5}$) and of sharp-adapted to normal-adapted function slopes ($m_{+.5}/m_0$). The steepest matching function was always the normal-adapted function despite considerable variance. (d) Normalization offset differences between adaptor conditions. More positive Δs adaptors always had more positive offsets.

blur-sharp gain ($m_{adaptor}$) pattern (Figure 5a) was replicated for every subject: As illustrated in Figure 5c, the ratio $m_{\Delta s=0}/m_{\Delta s=-.5}$ was always greater than 1.0, and the ratio of $m_{\Delta s=+.5}/m_{\Delta s=0}$ was always less than 1.0. The normalization offset pattern (Figure 5b) was also consistent: The differences $\beta_{\Delta s=0} - \beta_{\Delta s=-0.5}$ and $\beta_{\Delta s=+0.5} - \beta_{\Delta s=0}$ were always greater than zero (Figure 5d). Psychometric function slopes λ did not vary systematically with adaptor condition and averaged around 0.15 (not shown).

Simulating blur adaptation

Elliott et al. (2011) carried out an experiment similar to ours using static images (spatial noise or checker-
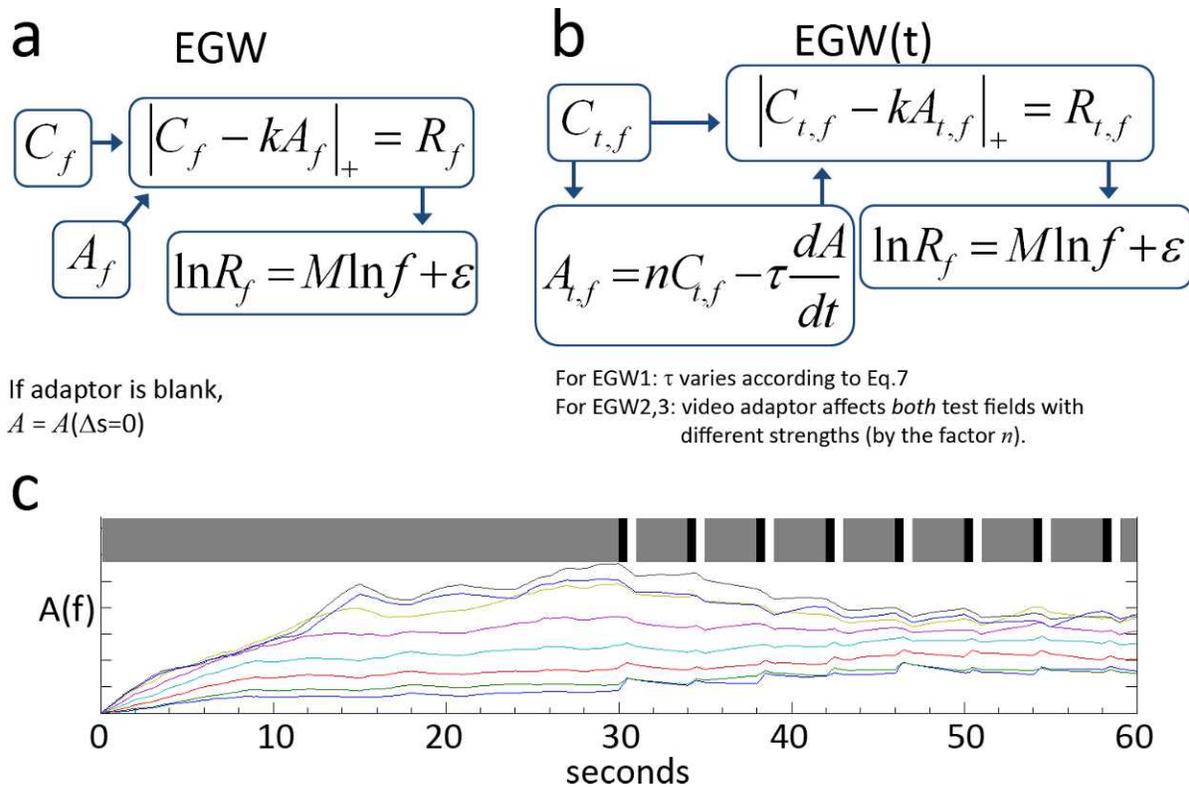
Figure 6. Schematics for two versions of the blur-adaptation model. (a) Elliott et al.'s (2011) adaptation model (EGW) takes, as two inputs, test (C_f) and adaptor (A_f) contrasts, subtracts a fixed proportion k of the adaptor from the test, and takes the average change in log difference with log frequency as a measure of image blur or sharpness. When the adaptor is blank, it is replaced with $\Delta s = 0$ contrasts ("long-standing normal adaptation"). (b) An elaborated version of the model that handles time-varying tests and adaptors. There is one input (C_f) that varies over time. This input builds up in an adaptation term (A_f), which is subtracted from the input as in the original model. (c) Time course of the first 60 s of the experiment (bars) and an example of the buildup of adaptation (colored lines) for the EGW3 model (starting from a zeroed adaptation state). A video-adapted field with $\Delta s = +1.0$ is represented, so the uppermost colored lines correspond to higher spatial frequency A values. Gray bars represent adaptor periods; black bars represent test periods; blank interstices represent ISIs. Note how these correspond to the plots of $A(f)$ in the same axes; after each adaptation period, $A(f)$ tends to converge because the random tests will be closer to $\Delta s = 0$ than the adaptor, and A drops at all frequencies during the ISI.

boards in different conditions) and proposed a simple model to explain their results. In their model (call it the EGW), the test image is decomposed into multiple scales (f) by band-pass filters, and the standard deviation of each band is measured to give an estimate of the mean linear contrast response. This gives us a vector of contrast responses for a particular Δs, e.g., $C_f(\Delta s)$. Adaptation is the subtraction of a constant proportion k of the adaptor responses A from the test responses C. The relative quality M (blurriness or sharpness) of the test image is decoded as the log change in adapted response with log spatial frequency f (obtained through linear regression):

$$\ln|C_f - kA_f|_+ = M\ln f + \varepsilon \qquad (4)$$

In Equation 4, the brackets $|x|_+$ indicate the operation max(x, 0) or half-wave rectification with zero adapted responses then left out of the regression. Elliott et al.

(2011) used a small value in the max operation rather than zero and included this in their calculations of M; we found that this resulted in unstable and unpredictable model performance. The reason for this is that for positive responses there is a relatively constant log-log relationship between response and frequency, which makes the linear descriptor of Equation 4 appropriate; however, introducing a response floor causes an artificial discontinuity in the response/frequency relationship, making the linear descriptor less appropriate and surprisingly sensitive to arbitrary changes in the chosen small value. The constant ε is not used in the model but could be understood as reflecting the overall perceived contrast of the stimulus. This system is essentially a model of multichannel contrast adaptation with logarithmic response compression (after subtractive contrast adaptation) and a hard threshold. With this model, the matching function between two adapting fields $A1$ and $A2$ (e.g., blurry and blank) is the

set of stimulus pairs C1 and C2 that satisfy the equation $M(C1, A1) = M(C2, A2)$. Elliott et al. found that this model (depicted in Figure 6a), with the parameter k set to a value around 0.4, was able to reproduce their matching results *if* it was assumed that exposure of the system to zero contrast caused the most recent state to be held in adaptive "memory."

Our data are very similar to those of Elliott et al. (2011) except that our normal-adapted matching functions do not fall along the identity line. However, our experiment differs in important ways that mean that we cannot fit the EGW model analytically to our data to test its performance, at least not while staying true to our experimental stimuli. In Elliott et al.'s experiments, within a condition, the test and adaptor images were the same throughout each condition (although in some conditions they used different random-phase noise patterns on each trial, we would still qualify these as having homogeneous spatial structure within and across trials). In our experiment, the adaptors and tests were never the same on any trial in either their phase or amplitude spectra (even disregarding the experiment-imposed Δs change), and furthermore, they varied across trials. So, although test and adaptor stimuli would have specified Δs values, their actual s (spectral slope) values—and thus their C_f values—were highly variable. Values of s and C_f even varied *within* the video stimuli from frame to frame. To test the model in a way that the random variation of stimulus properties could be properly incorporated, we used simulations of the experiment, running the model through 15 blocks (instead of three as the human subjects had done) of 180 trials for each of three adaptor Δs (−0.5, 0, + 0.5). The simulated experiments used the same staircase design as described in the General methods.

The simulation allows us to use random stimuli with the model, but it does not tell us how the adaptors and the tests should interact. To connect the two stimuli, we used a dynamic version of the EGW model (Elliott et al. 2011, supplemental materials) in which the adaptation term is determined by a first-order differential equation with a time constant τ of 10 s (Equation 5).

$$A_t = C_t - \tau \frac{dA}{dt} \tag{5}$$

The time constant value is broadly consistent with the sparse literature on temporal dynamics of contrast adaptation (Blakemore & Campbell, 1969; Greenlee et al., 1991; Hammett, Snowden, & Smith, 1994), but we chose it as a post-hoc justification of the 30:3 s adaptation:top-up period scheme and in recognition of the fact that most studies of contrast adaptation have employed top-up periods longer than 3 s (the median adapt:top-up scheme over all the studies we cite that used this design is 120:6 s). It is worth noting that a single time constant is certainly inadequate to describe the time course of contrast adaptation (Greenlee et al., 1991). Although there is an added parameter, Equation 5 can also be considered a simplification of the original model in that there is now a single input contributing both to adaptation and to a testable response (Figure 6b).

We incorporated a full representation of the temporal structure of the experiment in the simulation: an initial 30 s period of adaptation, followed by a 500 ms test interval, followed by a 500 ms blank "subject response" interstimulus interval (ISI), followed by 3 s of top-up adaptation, etc. Adaptor, test, and ISI all contributed to the adaptation term. The response to the test was the summed adapted test response:

$$R_f = \sum_{t=1}^{T} |C(t) - kA(t)|_+ \tag{6}$$

Note that Figure 6 and Equations 4 through 6 represent a single adapted field. It is likely that the adaptation process takes place over a much smaller area, but as we used a single stimulus size, we have no information about what the scale of the process might be.

The long-standing adaptation was implemented as follows: first, we set $A(t = 0)$ to the mean (over all 50 video clips) $\Delta s = 0$ video contrast values so that the model observer began each experiment in a state of "normal adaptation" (presuming that "normal," in this context, is the appearance of an unaltered DVD video presented on a CRT display). Next, the value of the time constant τ was made dependent on the contrast response C in a given channel, according to the following rule:

$$\tau_f = 10 + 600\left(1 - \frac{C_f}{C_f + .0001}\right) \tag{7}$$

So, for most contrasts (well above the small value of .0001), τ would be near the set value of 10 s, but as C approached zero, τ would increase rapidly to a very high value closer to 10 min. Exposing this system to a blank image—a string of zero contrasts—would cause the adaptation state to freeze. Elliott et al. (2011) proposed a similar implementation in the form of a step function that sets τ to infinity when the contrast response is zero.

Elliott et al. (2011) used four oriented filters in their model and summed across orientation after adaptation. In our tests, we found no useful difference in model performance between a version with oriented filters versus a version with one isotropic filter (at each of eight spatial frequencies), so all our subsequent modeling used an isotropic filter (raised cosine of log frequency with one-octave bandwidth) (Peli, 1990, 2002) for the sake of simplifying the computations.
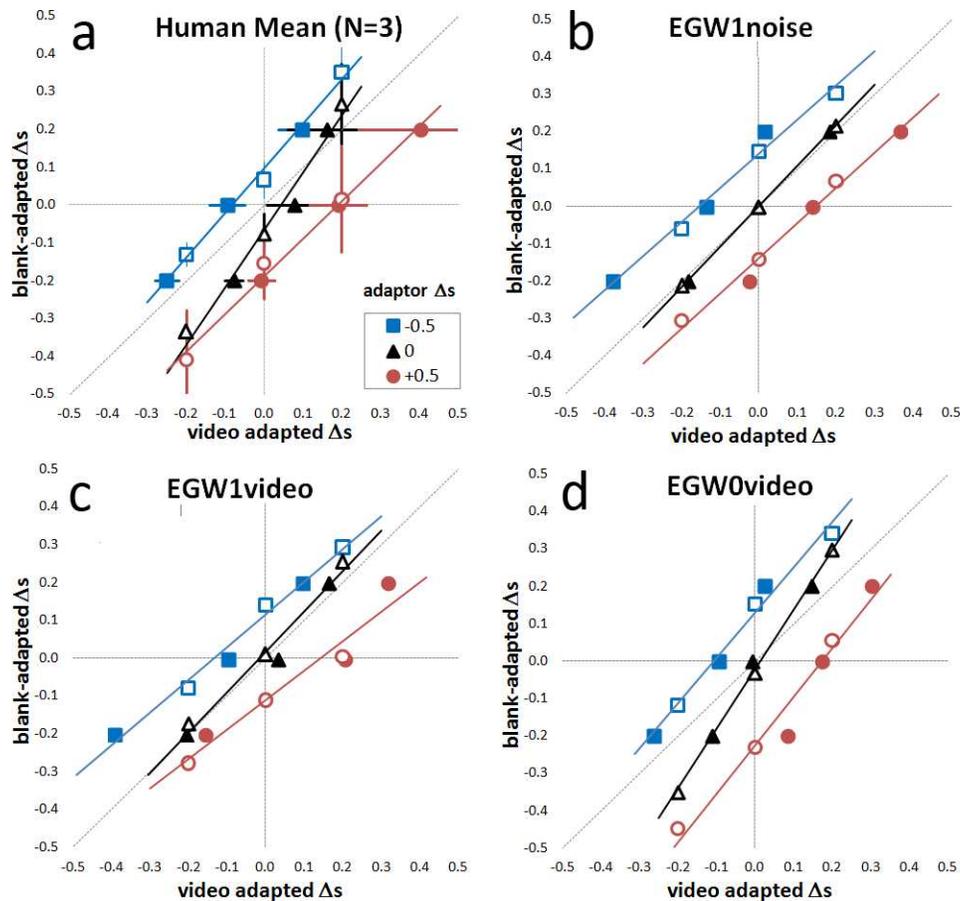
Figure 7. Matching functions generated by: (a) the human subjects shown in Figure 3; (b) the dynamic version of the original EGW1 with noise stimuli and $k = 0.4$; (c) same as b but with video stimuli; (d) the EGW model ($k = 0.4$) with no long-standing adaptation (blank-adapt to zero) and video stimuli.

Finally, we added Gaussian noise with 0.1 variance to each calculated M value. This was necessary for the noise-stimulus simulations described below because the amplitude spectra of these stimuli had virtually no variance across samples; without system noise, the simulated experiment would only be able to measure coarsely quantized matches. The noise had no effect on model simulations when the video contrasts were used; the internal variance was overwhelmed by the external variance. All the parameter values cited below (and in the caption to Figure 7) were obtained through simple iterative searches through the parameter spaces and manual adjustments.

When we test this model (call it the EGW1) with spatial noise samples for which the $\Delta s = 0$ stimulus has an amplitude spectrum slope of exactly $s = -1$, similar to one of the conditions originally tested by Elliott et al. (2011), the results are similar to theirs (with $k = .4$, Figure 7b). The normal-adapted matching function falls on the identity line with selectively adapted (lower gain) functions shifted above or below by blur and sharp adaptation. With video inputs, the model

performs similarly (Figure 7c) but differently from the human subjects in our experiment (the mean of the three original subjects is shown in Figure 7a). Refer back to Figure 5 in which the matching function parameter distributions are illustrated; model parameters have been plotted in the same diagrams (circle and diamond symbols corresponding to the first two model results shown in Figure 7b and c). While these and most other versions of this model do well at predicting the pattern of normalization (the values plotted in Figure 5b), the basic version of the model pins the normal-adapted matching function to the identity line (Figure 5a, a matching function slope of 1.0 is the blank-adapted blur-sharp gain), thus uniformly underestimating the blur-sharp gain. This is a result of the long-standing adaptation in the EGW1, which makes blank-adaptation nearly equivalent with normal adaptation.

Removing the long-standing adaptation by setting $\tau = 10$, independent of contrast response, results in the performance shown in Figure 7d. This model—labeled as EGW0—does well at capturing the basic features of our human subject data although it is rather inflexible:

A good fit cannot be obtained to both the slopes and the offsets by varying k, so here we show results for the model with $k = 0.4$, which captures the magnitude of the normalization offsets (the down-pointing triangles in Figure 5b) while leaving the blur-sharp gains all a little steep (Figure 5a). With this model, it is impossible to shift the normal-adapted $\Delta s = 0$ away from the origin, so this feature of the human data (i.e., in Figure 5b, most of the normal-adapted offsets were greater than zero) cannot be reproduced. The Appendix describes some more speculative modifications to the model that can produce simulated data that closely match human performance.

Discussion

Our purpose in doing these experiments was to test whether the blur (and sharp) adaptation effects measured in other studies can be generalized to more complex, realistic stimuli. We find that the effects of adaptation to random video clips are similar in magnitude and form to the effects of adaptation to static, nonvarying images as measured by others (Elliott et al., 2011; Vera-Diaz et al., 2010; Webster et al., 2002). To a large extent, we have replicated Elliott et al.'s recent study, in which they found that the effects of blur and sharp adaptation were selective, with blurred adaptors having stronger effects on blurred tests and vice versa. They were able to explain their results with a simple subtractive model of adaptation (the EGW). Their model included an interesting and crucial feature: When the model system is exposed to a blank field, the dynamic adaptive mechanisms under-lying blur adaptation stick in place, retaining the state of adaptation reached before the system was exposed to a field of zero contrast. Meanwhile, if the system is exposed to a sequence of nonzero contrasts, it follows these relatively quickly on a time scale of seconds. It is difficult to reconcile this with what is known of contrast adaptation: Detection thresholds for a target grating will be higher after adaptation to a high-contrast adaptor with similar spatial features as the target, and the lowest thresholds will be obtained after adaptation to a blank screen (Georgeson & Harris, 1984; Snowden, 1994). The decay of psychophysical adaptation to contrast, usually measured by placing a near-threshold target in a blank field with some variable delay after the adaptor offset (Blakemore & Campbell, 1969; Greenlee et al., 1991; Lorenceau, 1987), seems rapid enough to largely eliminate any stored adaptive state within a few tens of seconds.

So why the difference between our data and those of Elliott et al. (2011)? Why did they require the long-standing adaptor? It is probably not enough to suggest that the subjects were different because none of our 12 subjects showed the same pattern of blur-sharp gain as Elliott et al.'s three subjects. It is more likely that the difference is in the adaptor and test stimuli. In our simulations, the blank-adapted field, in the absence of a long-standing adaptation, still built up a low level of adaptation to the blank-adapted *test* stimuli. This test residual tended to resemble normal adaptation simply because the staircase design distributed a broad range of Δs values (tests and standards) in the blank-adapted field. In Elliott et al.'s similar experiment, this buildup may have been much more significant if only because the spatial structure of all the tests (random noise or phase-reversible checkerboards) was so similar from trial to trial. It is also possible that the decay of adaptation is different for static or constant-phase stimuli than for moving, random stimuli—most contrast adaptation studies with gratings use drifting stimuli, and we used moving adaptors and tests as well. We know already that the strength of some types of contrast gain control varies strongly with stimulus speed (Meese & Holmes, 2007); perhaps the persistence of contrast adaptation varies along the same dimension, which would allow for an automatic transition between the behavior of models EGW0 and EGW1.

Keywords: blur adaptation, contrast adaptation, contrast sensitivity, contrast enhancement, 1/f amplitude spectra, individual differences

Commercial relationships: none.
Corresponding author: Andrew M. Haun.
Email: andrew_haun@meei.harvard.edu.
Address: Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA.

References

Abbonizio, G., Langley, K., & Clifford, C. W. G. (2002). Contrast adaptation may enhance contrast discrimination. *Spatial Vision*, *16*(1), 45–58.

Bex, P. J., Solomon, S. G., & Dakin, S. C. (2009). Contrast sensitivity in natural scenes depends on

edge as well as spatial frequency structure. *Journal of Vision*, 9(10):1, 1–19, http://www.journalofvision.org/content/9/10/1, doi:10.1167/9.10.1. [PubMed] [Article]

Blakemore, C., & Campbell, F. W. (1969). On existence of neurones in human visual system selectively sensitive to orientation and size of retinal images. *Journal of Physiology-London*, 203(1), 237–260.

Blakemore, C., Muncey, J. P. J., & Ridley, R. M. (1973). Stimulus specificity in human visual system. *Vision Research*, 13(10), 1915–1931.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436.

Elliott, S. L., Georgeson, M. A., & Webster, M. A. (2011). Response normalization and blur adaptation: Data and multi-scale model. *Journal of Vision*, 11(2):7, 1–18, http://www.journalofvision.org/content/11/2/7, doi:10.1167/11.2.7.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical-cells. *Journal of the Optical Society of America a-Optics Image Science and Vision, 4*(12), 2379–2394.

Frazor, R. A., & Geisler, W. S. (2006). Local luminance and contrast in natural images. *Vision Research, 46*(10), 1585–1598.

Georgeson, M. A. (1985). The effect of spatial adaptation on perceived contrast. *Spatial Vision, 1*(2), 103–112.

Georgeson, M. A., & Harris, M. G. (1984). Spatial selectivity of contrast adaptation: Models and data. *Vision Research, 24*(7), 729–741.

Greenlee, M. W., Georgeson, M. A., Magnussen, S., & Harris, J. P. (1991). The time course of adaptation to spatial contrast. *Vision Research, 31*(2), 223–236.

Hammett, S. T., Snowden, R. J., & Smith, A. T. (1994). Perceived contrast as a function of adaptation duration. *Vision Research, 34*(1), 31–40.

Hansen, B. C., & Essock, E. A. (2005). Influence of scale and orientation on the visual perception of natural scenes. *Visual Cognition, 12*(6), 1199–1234.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*(2), 181–197.

Lesmes, L. A., Jeon, S. T., Lu, Z. L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research, 46*(19), 3160–3176.

Lorenceau, J. (1987). Recovery from contrast adapta-

tion: Effects of spatial and temporal frequency. *Vision Research, 27*(12), 2185–2191.

Meese, T. S., & Holmes, D. J. (2007). Spatial and temporal dependencies of cross-orientation suppression in human vision. *Proceedings of the Royal Society B-Biological Sciences, 274*(1606), 127–136.

Ohlendorf, A., & Schaeffel, F. (2009). Contrast adaptation induced by defocus: A possible error signal for emmetropization? *Vision Research, 49*(2), 249–256.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 7(10), 2032–2040.

Peli, E. (2002). Feature detection algorithm based on a visual system model. *Proceedings of the IEEE*, 90(1), 78–93.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10(4), 437–442.

Rajeev, N., & Metha, A. (2010). Enhanced contrast sensitivity confirms active compensation in blur adaptation. *Investigative Ophthalmology & Visual Science, 51*(2), 1242–1246, http://www.iovs.org/content/51/2/1242. [PubMed] [Article]

Rosenfield, M., Hong, S. E., & George, S. (2004). Blur adaptation in myopes. *Optometry and Vision Science, 81*(9), 657–662.

Ross, J., & Speed, H. D. (1996). Perceived contrast following adaptation to gratings of different orientations. *Vision Research, 36*(12), 1811–1818.

Sawides, L., Marcos, S., Ravikumar, S., Thibos, L., Bradley, A., & Webster, M. (2010). Adaptation to astigmatic blur. *Journal of Vision*, 10(12):22, 1–15, http://www.journalofvision.org/content/10/12/22, doi:10.1167/10.12.22. [PubMed] [Article]

Snowden, R. J. (1994). Adaptability of the visual system is inversely related to its sensitivity. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 11(1), 25–32.

Snowden, R. J., & Hammett, S. T. (1996). Spatial frequency adaptation: Threshold elevation and perceived contrast. *Vision Research*, 36(12), 1797–1809.

Vera-Diaz, F. A., Goldstein, R. B., & Peli, E. (2008). *Asymmetrical adaptation to highpass versus lowpass filtered images*. Paper presented at the Vision Sciences Society Annual Meeting, Naples, FL.

Vera-Diaz, F. A., Kark, S., Woods, R. L., & Peli, E. (2011). *Blur adaptation and contrast sensitivity*. Paper presented at the American Academy of Optometry Annual Meeting, Boston, MA.

Vera-Diaz, F. A., Woods, R. L., & Peli, E. (2010).

Shape and individual variability of the blur adaptation curve. *Vision Research*, *50*(15), 1452–1461.

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research*, *39*(23), 3960–3974.

Wallman, J., & Winawer, J. (2004). Homeostasis of eye growth and the question of myopia. *Neuron*, *43*(4), 447–468.

Wang, B., Ciuffreda, K. J., & Vasudevan, B. (2006). Effect of blur adaptation on blur sensitivity in myopes. *Vision Research*, *46*(21), 3634–3641.

Webster, M. A., Georgeson, M. A., & Webster, S. M. (2002). Neural adjustments to image blur. *Nature Neuroscience*, *5*(9), 839–840.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Attention, Perception, & Psychophysics*, *63*(8), 1314–1329.

Appendix

Comparing adaptation between internal and external standard paradigms

The video-adapted Δs values that appear normal as a function of adaptor Δs—*normalization functions*—are plotted as open symbols in Figure A1a. These can be estimated from the matching functions of Figure 3 as $\beta = -b/m$ if we assume that the perceived normal is the dashed horizontal line in each panel of Figure 3 (i.e., for each adaptation condition, the perceived-normal Δs is where the matching function intersects the perceived-normal line). For comparison, we also measured normalization functions with the internal-standard method of Webster et al. (2002). The stimulus conditions were similar to the matching experiment, except that there was a single adapting field, centrally fixated by subjects rather than two adapting fields
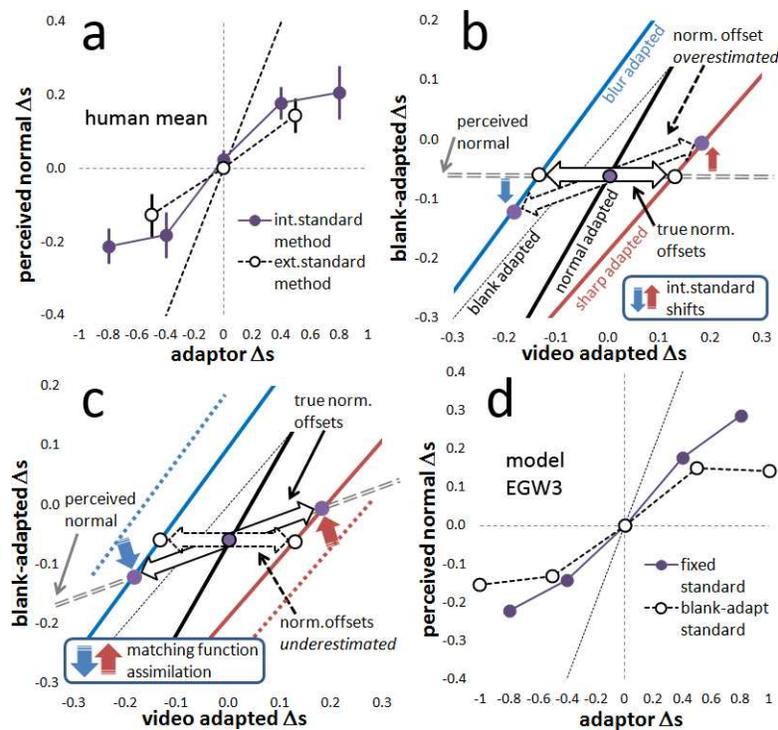

Figure A1. (a) Perceived normal represented as a function of adaptor Δs. Open symbols are the video-adapted offsets of the averaged matching functions shown in Figure 3, indicating shifts in the perceived normal toward the adaptor Δs. Solid symbols are data from the internal-standard (Webster et al., 2002) paradigm using video clips averaged for the same three subjects. Note that the adaptation/normalization effect is slightly larger when measured with the internal standard method. Although small, the possible reasons for the difference are an interesting contrast; (b) shows how the discrepancy could arise from assimilation of the internal standard with the adaptor (illustrated by the colored arrows), so that, e.g., matching a blur-adapted test to a blurry internal standard will result in a blurrier match (blue arrow). (c) If the blank-adapted field is influenced by the video adaptor, then it is not appropriate to treat the ordinate as a perceptual quality axis, and the perceived-normal blank-adapted Δs will vary with adaptor condition (double dashed line), and the measured matching functions will be closer together than they would be with a truly fixed standard (colored arrows). (d) Simulated data from the EGW3 model described in the text; "fixed-standard" data emulate the internal standard data of Figure A1a while "blank-adapt standard" data are estimated from the same model producing the data in Figure A2b.
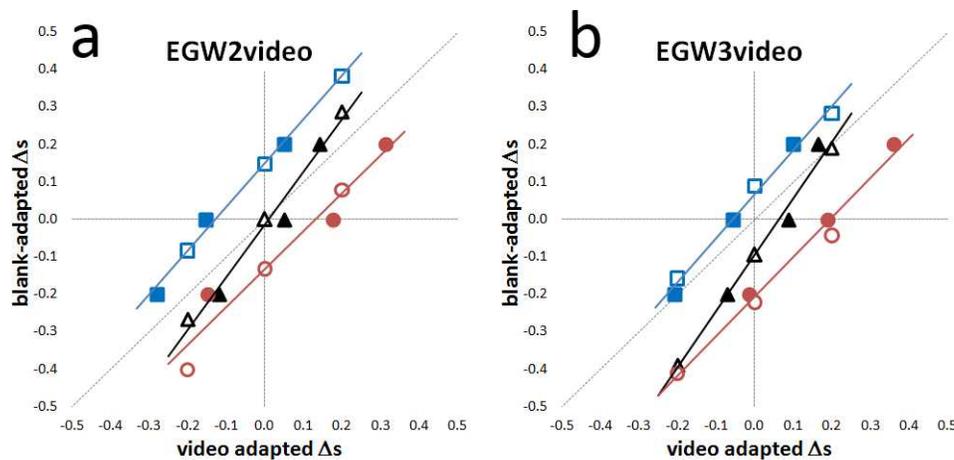
Figure A2. More speculative extensions of the EGW model. (a) Allowing the video adaptor to have a small (proportional) effect on the blank-adapted field results in compression of the matching functions toward the identity line. This compression of effect size results in the different measures of normalized blur as shown in Figure A1d, which may explain the difference in performance of our three subjects shown in Figure A1a. (b) Adjusting the relative strength of the spread of adaptation according to spatial frequency allows the matching functions to be shifted up or down on the y-axis so that the off-origin effects of adaptation to normal video (shown in Figure 5d) can be reproduced.

immediately flanking fixation. Subjects judged whether test stimuli presented in the video-adapted field were blurrier or sharper than what they knew to be "normal" (Vera-Diaz et al., 2008; Webster et al., 2002), and their responses drove a one-up, one-down staircase, which adjusted the Δs value of the test on the next trial. The null point, at which tests were equally likely to be judged as sharper or blurrier, was obtained by the same method as the matching point as described in the General methods. These points are plotted as the solid symbols in Figure A1a. The normalization offsets measured with the external standard (the matching method) were slightly smaller than those measured with the internal standard.

The discrepancy in effect size between the external- and internal-standard methods is interesting. As illustrated in Figure A1b, it could be that the effects (distances along the video-adapted horizontal axis) are inflated with the internal-standard method because of a drift in the internal standard toward the adaptor. The internal-standard design requires that the subject identify stimuli as blurry or sharp relative to *what they know as normal*, but it is possible that subjects may allow their judgments to be influenced by the repeatedly seen adaptor (e.g., by identifying a stimulus as sharp or blurry *relative to the adaptor*). Figure A1b shows that if a subject's internal reference drifts in the direction of the quality of the adaptor (as shown by the colored arrows), the internal reference method could yield inflated adaptation results (the dashed-edge arrows extend further horizontally than the solid-edge arrows). We had thought that this was an advantage of the external-standard design.

However, it may be that in the external-standard method, the effect of the video adaptor spreads spatially into the blank-adapted field. If this is so, the measured matching functions will be pushed closer to the identity line (Figure A1c as illustrated by the colored arrows) than they would have been if the blank-adapted field were truly an independent standard. This is not a fatal problem to the external-standard design, but it does greatly complicate our interpretation of the results: If true, then the ordinate axis of the plots in Figure 3 cannot be assumed to be identical with a constant perceptual continuum as we (and Elliott et al. [2011] in their similar study) had intended. In particular, there will be no unique blank-adapted Δs value that corresponds to the perceived normal; instead, the blank-adapted perceived normal corresponds to a different Δs depending on the adaptor condition as illustrated by the dashed double line in Figure A1c.

We also considered whether adaptation to the trial stimuli in the blank-adapted field might somehow explain this result. Although the standards were set symmetrically around $\Delta s = 0$ (at -0.2, 0, and $+0.2$), the tests were determined by adaptive staircases, so they tended to be distributed around the points of subjective equality (to the video-adapted standards). Because video-adapted standards would be matched by blank-adapted tests with the quality *opposite to* the quality of the adaptor (e.g., blur-adapted standards would be matched by sharpened tests), the blank-adapted field would tend to accumulate adaptation opposite to the video adaptor, which would tend to *increase* effect sizes. Because we found slightly reduced effect sizes in the matching experiment relative to the internal

standard experiment, trial-phase adaptation in the blank-adapted field cannot explain the observed pattern of results.

Looking only at the matching experiment, the EGW0 does well at simulating human performance although it cannot reproduce the small shift in the normal-adapted matching functions (Figure 5b). We found that a low-frequency bias in the extent of the spatial spread of adaptation suggested above would allow the EGW model to reproduce both the off-origin matching functions and the difference in normalization strength between the internal- and external-standard paradigms. The spread of adaptation from the video-adapted to the blank-adapted field consists of setting $C_{blank} = nC_{video}$ during the adaptation phase (refer back to Figure 6b) with $n < 1$. In effect, this means that the video adaptor acts on both adapting fields but that k in the blank-adapted field is equal to $n \cdot k$. This would imply that adaptation is pooled over a very large area; however, because the EGW already pools over the entire stimulus area before affecting adaptation, this can be considered a relatively minor adjustment. The matching performance of this model (the EGW2 with $k = 0.55$ and $n = 0.25$) is shown in Figures A2a and 6 (upward triangles). Two improvements are gained: Values of n and k can be found that obtain matching function slopes and offsets in the normal range of the human subjects, and an explanation for the discrepancy in offset magnitude shown in Figure A1a emerges. The

spread of adaptation into the blank-adapted field effectively squeezes all the matching functions closer together, so the normalization offsets are underestimated. We demonstrate this after the next step.

We implemented a low-frequency bias in the spatial spread of adaptation as $n = 0.8/(1 + \log2[f])$ (where f is in cycles per picture). This bias shifts all the matching functions downward (Figure A2b, compare with Figure A2a). We also get a realistic relative difference in normalization strength between the blank-adapted standard method (the matching simulation as described above) and a "fixed-standard" method in which the model was run through a simulation of the single-adaptor (Webster et al., 2002) method (Figure A1d). So the most successful model of blur adaptation that we have found seems to support the notion that the internal standard used by subjects in the single-adaptor paradigm is indeed stable (refer back to the discussion around Figure A1b and c) and that there may be some form of contamination between the two stimulus fields in the external-standard paradigm. The idea that there is a low-frequency spread in adaptation between the two closely placed test fields is consistent with blur adaptation being contrast adaptation at multiple spatial scales: Neurons that prefer lower spatial frequencies have larger receptive field extent and must sometimes overlap between the video- and blank-adapted fields.