

Added luminance ramp alters perceived edge blur and contrast: A critical test for derivative-based models of edge coding

Keith A. May ^{*,1}, Mark A. Georgeson

School of Life & Health Sciences, Aston University, Birmingham B4 7ET, UK

Received 20 April 2006; received in revised form 4 February 2007

Abstract

In many models of edge analysis in biological vision, the initial stage is a linear 2nd derivative operation. Such models predict that adding a linear luminance ramp to an edge will have no effect on the edge's appearance, since the ramp has no effect on the 2nd derivative. Our experiments did not support this prediction: adding a negative-going ramp to a positive-going edge (or vice-versa) greatly reduced the perceived blur and contrast of the edge. The effects on a fairly sharp edge were accurately predicted by a nonlinear multi-scale model of edge processing [Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (in press). From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision*], in which a half-wave rectifier comes after the 1st derivative filter. But we also found that the ramp affected perceived blur more profoundly when the edge blur was large, and this greater effect was not predicted by the existing model. The model's fit to these data was much improved when the simple half-wave rectifier was replaced by a threshold-like transducer [May, K. A. & Georgeson, M. A. (2007). Blurred edges look faint, and faint edges look sharp: The effect of a gradient threshold in a multi-scale edge coding model. *Vision Research*, 47, 1705–1720.]. This modified model correctly predicted that the interaction between ramp gradient and edge scale would be much larger for blur perception than for contrast perception. In our model, the ramp narrows an internal representation of the gradient profile, leading to a reduction in perceived blur. This in turn reduces perceived contrast because estimated blur plays a role in the model's estimation of contrast. Interestingly, the model predicts that analogous effects should occur when the width of the window containing the edge is made narrower. This has already been confirmed for blur perception; here, we further support the model by showing a similar effect for contrast perception.

© 2007 Elsevier Ltd. All rights reserved.

Keywords: Blur; Contrast; Edge; Gaussian derivatives; Scale; Template; Psychophysics; Human vision

1. Introduction

The problem of how to detect edges in images has been studied by vision scientists for over half a century. A very large proportion of edge detection algorithms involve some kind of derivative operation, and many different derivative operators have been proposed. In this paper, we briefly review the arguments for different derivative operators in edge detection, and we report some new experiments on

perceived edge blur and contrast that provide evidence favouring one particular form of derivative as a model of edge detection in human vision.

1.1. Detecting edges using the 1st or 2nd derivative

Many edge detection algorithms define an edge as a peak of luminance gradient magnitude. According to this definition, edges can be found by locating peaks in the 1st derivative, or zero-crossings (ZCs) in the 2nd derivative. The first serious attempt at a biologically plausible edge detection algorithm used a 2nd derivative operator, the Laplacian of Gaussian (Marr & Hildreth, 1980). This work was extremely influential, and several subsequent

* Corresponding author. Fax: +44 (0) 1274 235570.

E-mail address: keith@keithmay.org (K.A. May).

¹ Present address: Department of Optometry, University of Bradford, Richmond Road, West Yorkshire, Bradford BD7 1DP, UK.

models of feature detection in human vision used a 2nd-derivative operation as the initial stage (e.g., Georgeson, 1992; Kingdom & Moulden, 1992; Watt & Morgan, 1985).

Marr and Hildreth gave no computational justification for using the 2nd derivative rather than the 1st derivative. Boie, Cox, and Rehak (1986) argued that a Gaussian 2nd derivative, such as the Laplacian, was optimal for edge localization, but a Gaussian 1st derivative was optimal for edge detection. Several other researchers have presented different arguments that a Gaussian 1st derivative is optimal (or close to optimal) for edge detection (e.g., Canny, 1986; Sarkar & Boyer, 1991; Tagare & deFigueiredo, 1990). These approaches have assumed that an edge detection filter should give rise to a peak or a high signal-to-noise ratio (SNR) at the edge location. This assumption favours the 1st derivative, which produces a high signal at the edge location, and rules out the 2nd derivative, for which the expected signal at the edge location is zero. While it is clear that, for good detection, the SNR should be high at some location in the vicinity of the edge, none of these authors justified their assumption that the SNR should be high at the edge location itself, rather than in the regions on each side of the edge, as given by the 2nd derivative. Elder and Zucker (1998) argued that having a high SNR at the edge location does not guarantee that peaks in the 1st derivative are reliable. They described an algorithm for machine vision that selected, at each image location, the smallest operators with response levels that could be statistically distinguished from zero. If ZCs are found in the 2nd derivative at the minimum reliable scale, then these will identify points at which the 2nd derivative reliably changes sign, indicating positions of reliable peaks in the 1st derivative. Their argument in favour of the 2nd derivative could also be applied to any other method that rejects signals that are too weak to be reliable (e.g., a threshold).

1.2. A new approach based on the 3rd derivative

Georgeson, May, Freeman, and Hesse (in press) presented a new model of edge processing in human vision, called N_3^+ . The model spatially differentiates the luminance profile, half-wave rectifies the 1st derivative, and then differentiates twice more, to give the 3rd derivative of all regions with a positive gradient. This process is implemented by a set of Gaussian derivative filters with a range of scales. Each peak in the inverted normalized 3rd derivative across space and scale indicates the position and scale (i.e., blur) of an edge. The edge contrast can be estimated from the height of the peak (see May & Georgeson, 2007). The model explains a great deal of psychophysical data with no free parameters (Georgeson et al., in press), but cannot account for the fact that blurred edges look faint (i.e., low in contrast), and faint edges look sharp (May & Georgeson, 2007). May and Georgeson showed that both of these effects can be

explained by replacing the half-wave rectifier with a smooth threshold-like transducer function that is described by two parameters. They described two alternative models, N_3^+1A and N_3^+1B , which had slightly different transducer functions, but gave similar results. We refer to the original model as N_3^+0 . The family of models that includes N_3^+0 , N_3^+1A , and N_3^+1B will be referred to collectively as the N_3^+ models.

1.3. A test between different derivative-based mechanisms

We noted above that many derivative-based models of edge detection in biological vision have a 2nd derivative as the initial stage. This hypothesis can be tested by examining the effects of adding a linear luminance ramp to an edge. Adding a linear ramp has no effect on the 2nd derivative, because the 2nd derivative of the ramp is zero everywhere except at the ends of the ramp itself: if the appearance of the edge (e.g., contrast and blur) is determined by local cues in the 2nd derivative, then its appearance should be unchanged by the addition of the ramp. This is illustrated in Fig. 1.

Fig. 2 shows how adding a positive or negative ramp would affect the processing of a Gaussian edge in a simplified version of May and Georgeson's (2007) N_3^+1A model, using a biased rectifier after the 1st derivative, rather than a smooth transducer. May and Georgeson showed that the N_3^+ models are equivalent to fitting a 2nd-derivative edge template to the 2nd derivative profile of the edge. The 2nd derivative of a Gaussian-blurred edge is truncated by the threshold of the biased rectifier (indicated by the dashed lines in Fig. 2). When a positive ramp is added, the gradient profile is lifted above the threshold, and is no longer truncated. Thus, the positive ramp causes the 2nd derivative to match a wider template, so the edge should look more blurred. The negative ramp causes the 2nd derivative to be highly truncated by the rectifier, so it would match a narrower template, and look sharper. The rectifier in the N_3^+0 model has no threshold, so this model would predict no effect of the positive ramp, but would still predict a strong effect of the negative ramp. In the N_3^+ models, perceived contrast is an increasing function of perceived blur (May & Georgeson, 2007, Eq. (2)), so the effects of the ramp on perceived contrast should mirror the effects on perceived blur. Experiment 1 examined the effect of an added ramp on perceived blur, and Experiment 2 looked at the effect on perceived contrast. The stimuli for these two experiments were the same, and the methods were very similar, so they are described together.

2. Experiments 1 and 2

2.1. Subjects

The subjects were the two authors, KAM and MAG. Both had corrected-to-normal vision.

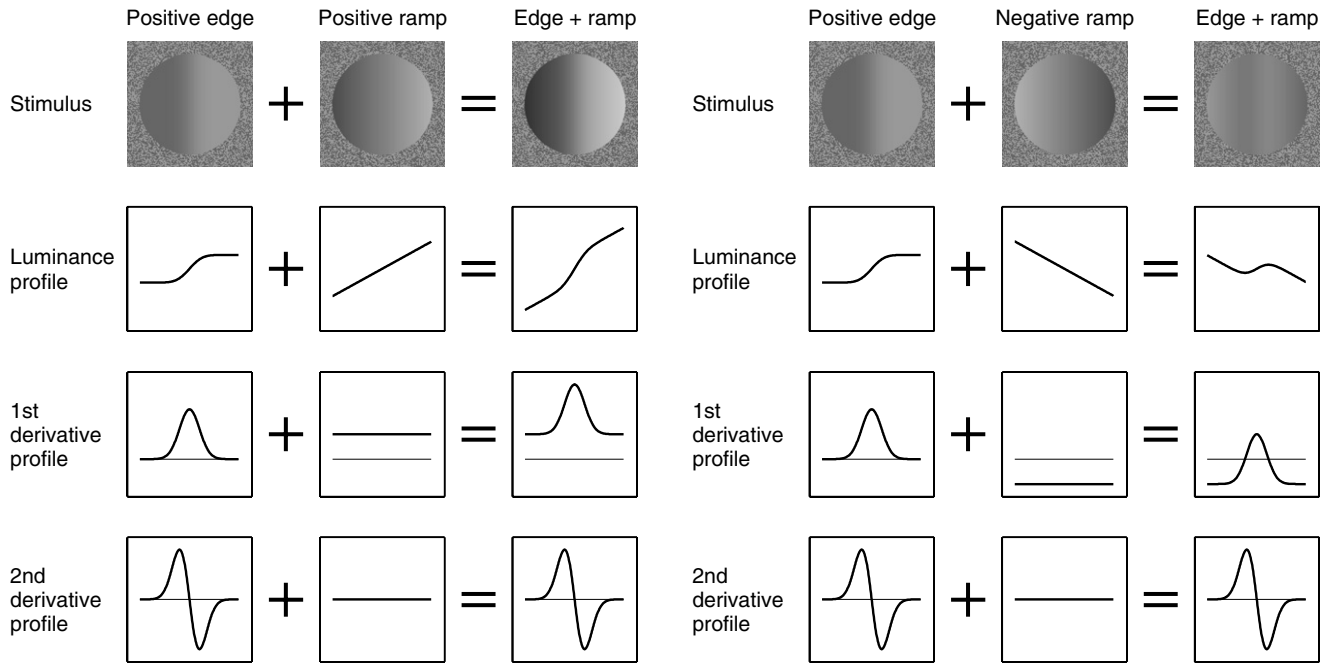


Fig. 1. Adding a linear ramp to a Gaussian edge. The left half of this figure shows the effect of adding a ramp with the same polarity as the edge; the right half shows the effect of a ramp with opposite polarity. Although the luminance profiles of the resulting stimuli are quite different, their gradient profiles are identical except for a vertical shift. Therefore, the 2nd derivative of each stimulus is the same except at the ends of the ramp. If the perceived edge properties are determined by local cues in the 2nd derivative that lie within the spatial extent of the ramp, then adding the ramp will not affect these properties. In these examples, the ramp gradient is half the peak edge gradient, so the gradient profile is shifted up or down by half its height. In the experiments, the steepest negative ramp gradient was three-quarters of the peak edge gradient, and the steepest positive ramp gradient was half the peak edge gradient.

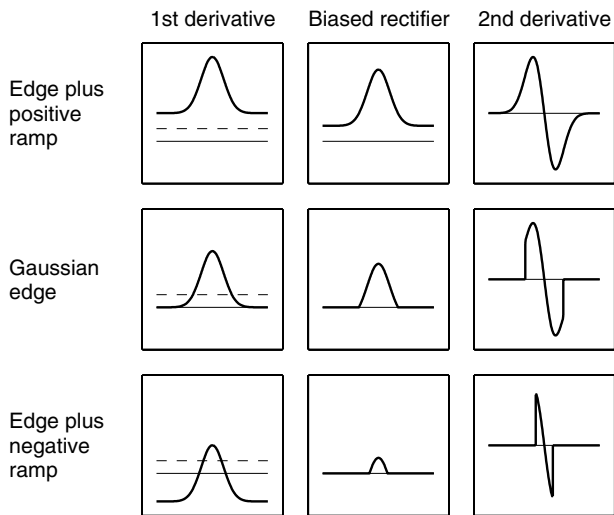


Fig. 2. The effect of adding a positive or negative ramp on the processing of a Gaussian edge in the N_3^+ model. See text for details.

2.2. Apparatus

The apparatus was identical to that described in the accompanying paper (May & Georgeson, 2007).

2.3. Stimuli

The stimuli for both experiments were vertical edges that had the profile of a Gaussian integral with scale (i.e.,

standard deviation), σ , with an added ramp of gradient r . Examples of the stimuli are given in Fig. 1. The luminance, L , at spatial position (x, y) is given by

$$L(x, y) = L_0[1 + w(x, y)(C(2\Phi(x; \sigma) - 1) + rx)]. \quad (1)$$

L_0 is the mean luminance of 45 cd/m². C is the Michelson contrast of the edge, which was always 0.2. $\Phi(\cdot; \sigma)$ is the integral of a unit-area Gaussian with scale (standard deviation) σ . w is the spatial window, which was circular with a flat profile and a sharp border. The value outside the circular border of the window was 0; the value inside was ± 1 : a value of +1 gives edges that are dark on the left, and -1 gives edges that are dark on the right. The window border was antialiased by giving it a raised sine profile with half-period equal to the diagonal distance across a pixel. Window diameters depended on the edge scale, σ , and are given in Table 1. All stimuli were invariant over time in every aspect except their contrast. The contrast had a temporal profile that was flat for the central 250 ms, and flanked by raised sine edges with half-period 25 ms, so the entire stimulus lasted for 300 ms. The value, C , in Eq. (1) refers to the contrast during the flat central period of the stimulus.

From Eq. (1), it can be seen that positive values of r correspond to ramps with gradients with the same sign as the Gaussian edge, and negative values correspond to ramps with gradients of opposite sign. Throughout this paper, the terms “positive ramp” and “negative ramp” will refer to the sign of r .

Table 1
Window diameter, viewing distance, and horizontal and vertical screen resolution for each fixed edge scale in Experiments 1 and 2

Fixed edge scale (arcmin)	Fixed and adjustable window diameter (arcmin)	Viewing distance (cm)	Screen resolution (arcmin per pixel)
6	80	260	0.586
8	80	260	0.586
12	160	130	1.17
16	160	130	1.17
24	320	65	2.34
32	320	65	2.34

The 6', 12' and 24' fixed stimuli were physically identical on the screen, and differed in scale only because of the different viewing distances; the same applied to the 8', 16' and 32' fixed stimuli.

Subject KAM performed two versions of Experiment 1. In one version, the area of the screen outside of the spatial window had a uniform luminance of L_0 . In the other version, the stimuli were surrounded by binary noise with Michelson contrast 0.2. The noise filled a 256×256 pixel square (size 10 deg at viewing distance 65 cm, 5 deg at viewing distance 130 cm, and 2.5 deg at viewing distance 260 cm). The noise was generated by dividing the 256×256 pixel square into 2×2 pixel squares and assigning each 2×2 square a randomly selected luminance of $L_0 \times (1 \pm 0.2)$. The example stimuli in Fig. 1 have a noise border generated in this way, although the border in these examples is narrower than was used in the experiment. One purpose of the noise surround was to de-emphasize the edges of the window: the negative ramp stimuli were quite complex, with several features, and we wanted the subjects to focus their attention on the central edge in each stimulus. Subject MAG only performed the noise-surround version of Experiment 1. Experiment 2 did not use noise surrounds.

2.4. Psychophysical method

The purpose of the experiments was to find the perceived blur (Experiment 1) or perceived contrast (Experiment 2) of the central edge in several different edge + ramp stimuli. These experiments used the same general procedure and data analysis as the accompanying paper (May & Georgeson, 2007, Section 3.4). The perceived blur or contrast of a *fixed edge* (with particular scale and ramp gradient) was assessed by using a 1-up-1-down staircase to bring the scale (or contrast) of an *adjustable edge* close to the point of subjective equality, and then fitting a psychometric function to the data from several staircases to give a maximum-likelihood estimate of the adjustable edge scale (or contrast) that gave a perceptual match to the fixed edge. Bootstrap confidence limits of 5% and 95% were calculated (these are indicated on the graphs in Figs. 3, 4, and 6).

The adjustable edge was a standard Gaussian edge (created by setting $r = 0$ in Eq. (1)), with the same polarity and

window as the corresponding fixed edge. In Experiment 1, the adjustable edge had the same contrast as the fixed edge before addition of the ramp (always 0.2); in Experiment 2, the adjustable edge had the same scale as the fixed edge before the addition of the ramp (the edge scale varied across conditions). In Experiment 2 (but not Experiment 1), each stimulus was accompanied by a tone in case the contrast of the adjustable edge fell below detection threshold.

In both experiments, there were two independent variables: fixed edge scale and ramp gradient. Scales were 6, 8, 12, 16, 24, 32 arcmin. These six scales were spread over three viewing distances, which are given in Table 1, along with the screen resolutions. There were six different ramp gradients. Each ramp gradient, r , was expressed as a proportion of the peak gradient of the Gaussian edge to which the ramp had been added. This scaled ramp gradient, r' , was calculated as follows:

$$r' = (r\sigma/C)\sqrt{\pi/2}. \quad (2)$$

Expressing the ramp gradient as a proportion of the edge gradient meant that, for a given ramp gradient, r' , a change of fixed edge scale did not change the shape of the luminance profile, except for spatial scaling, so the ramp levels were comparable across different fixed edge scales; in fact, many models, including N_3^+0 , make the same predictions for each fixed edge scale when the ramp gradient is expressed in this way. The levels of ramp gradient, as a proportion of peak edge gradient, were -0.75 , -0.50 , -0.25 , 0 , 0.25 and 0.50 .

In Experiment 1, two sessions were conducted at each viewing distance. The sessions were run in the order ABCBCA, where a different viewing distance was assigned (randomly for each subject) to each of the letters A, B, and C. The six ramp gradients were each tested in different blocks, which were randomly ordered within a session. Within each block, subjects saw fixed edges with two different scales (determined by the viewing distance for that block—see Table 1). Within a block, each condition was assigned to four staircases that differed in edge polarity and staircase start position, as in May and Georgeson (2007, Experiment 1).

In Experiment 2, four sessions were completed at each viewing distance. They were run in the order ABCACB-CBABCA, where the three viewing distances were assigned to the letters A, B and C, randomly for each subject. This order ensured that A, B, and C had the same mean ordinal position, and were distributed evenly throughout the sequence, with no viewing distance appearing twice in succession. Within each session, all the ramp gradients were randomly interleaved. Two fixed edge scales were used in each session, as in Experiment 1. Within a session, each condition was assigned to two staircases, one starting above the fixed edge contrast and the other starting below. The polarity of the stimuli was selected randomly on each trial.

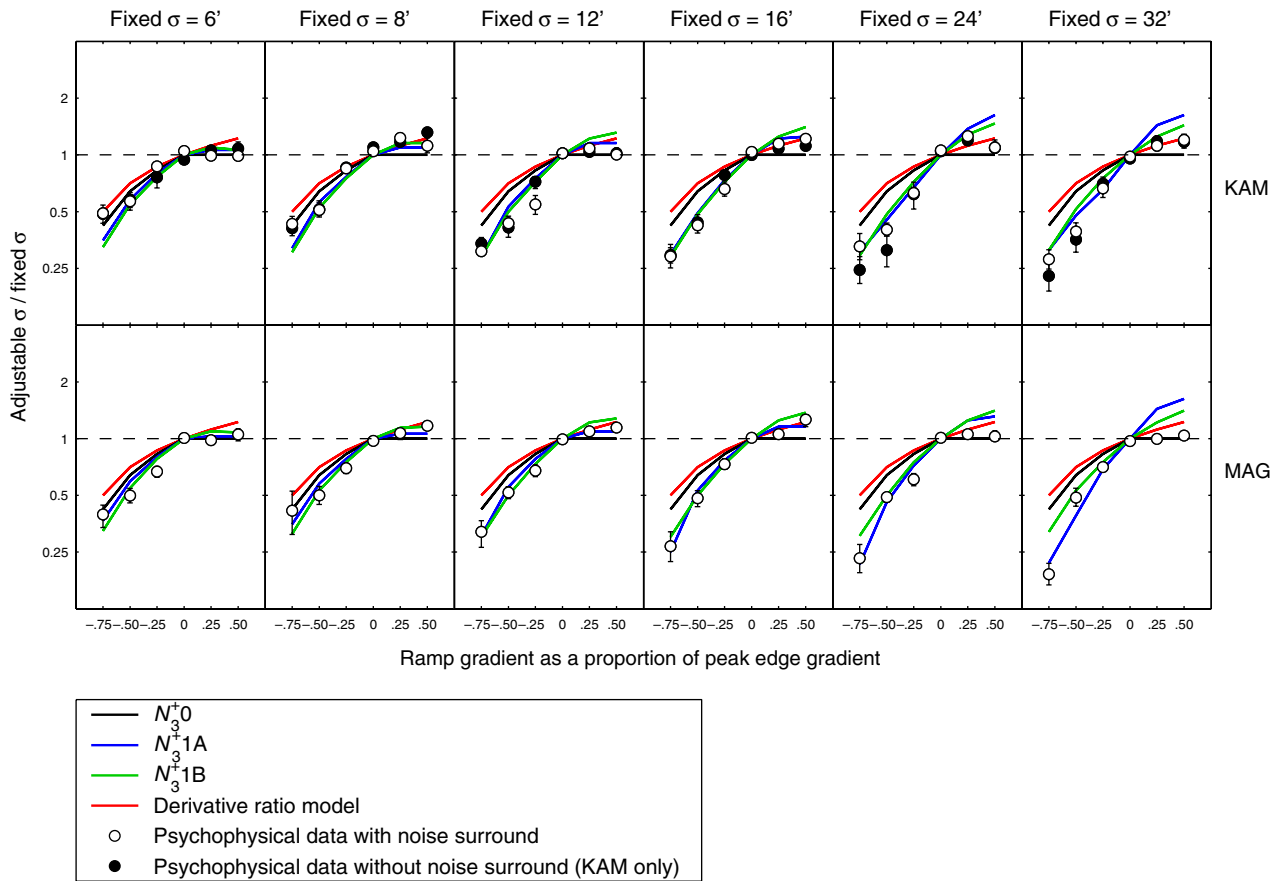


Fig. 3. Results of Experiment 1. Circles indicate psychophysical data; error bars indicate bootstrap confidence limits of 5% and 95%. The vertical axis gives the ratio of adjustable edge scale to fixed edge scale when the two edges perceptually matched in blur. Each column of panels gives the data for one fixed edge scale; each row gives data for one subject. The horizontal dashed lines indicate the results that would have occurred if perceived blur was unaffected by the added ramp. The red lines show the predictions of Georgeson's (1994) derivative ratio model. The black lines show the predictions of N_3^+0 . The blue and green lines show the predictions of N_3^+1A and N_3^+1B with the best-fitting parameters. The numerical values plotted in this figure are given in the [supplementary data](#).

2.5. Psychophysical results

The results of Experiments 1 and 2 are shown in Figs. 3 and 4, respectively. KAM's data in Fig. 3 show that the effect of the noise surround was minimal. Results were very similar for the two observers, except that the positive ramp caused a slight increase in perceived blur and contrast for KAM, but with little or no effect for MAG. On the other hand, the negative ramp caused a considerable reduction in perception of both blur and contrast: the steepest negative ramp reduced the matched blur by a factor of between 2 and 4, and reduced the matched contrast by a factor of between 4 and 8. A model in which the initial stage is a 2nd derivative operation would predict that the ramp would have no effect, so this is strong evidence against this general class of models.

2.6. Simulation results

Experiments 1 and 2 were simulated using Georgeson et al.'s (in press) parameter-free N_3^+0 model and May and

Georgeson's (2007) modified versions of this model (N_3^+1A and N_3^+1B).

For Experiment 1 (blur perception), we also examined two other types of model: Georgeson's (1994) derivative ratio model, and models in which the perceived edge blur is a function of the spatial separation between peak and trough in the 2nd derivative (e.g., Elder & Zucker, 1998; Watt & Morgan, 1983, 1985). Most models of the latter type were ruled out by the argument, given in the introduction, that a 2nd derivative operator shows no response to a linear ramp. However, because the ramp necessarily comes to an end fairly close to the edge, a very large 2nd derivative operator might be affected by the edges at the ends of the ramp. We therefore investigated the effect of the ramp on peak-to-trough spatial separation in the outputs of larger 2nd derivative operators.

For Experiment 2 (contrast perception), we looked at the predictions of the energy model (Morrone & Burr, 1988; Morrone, Burr, & Ross, 1994), and a simple model in which the perceived contrast was function of the local peak-to-trough difference in luminance.

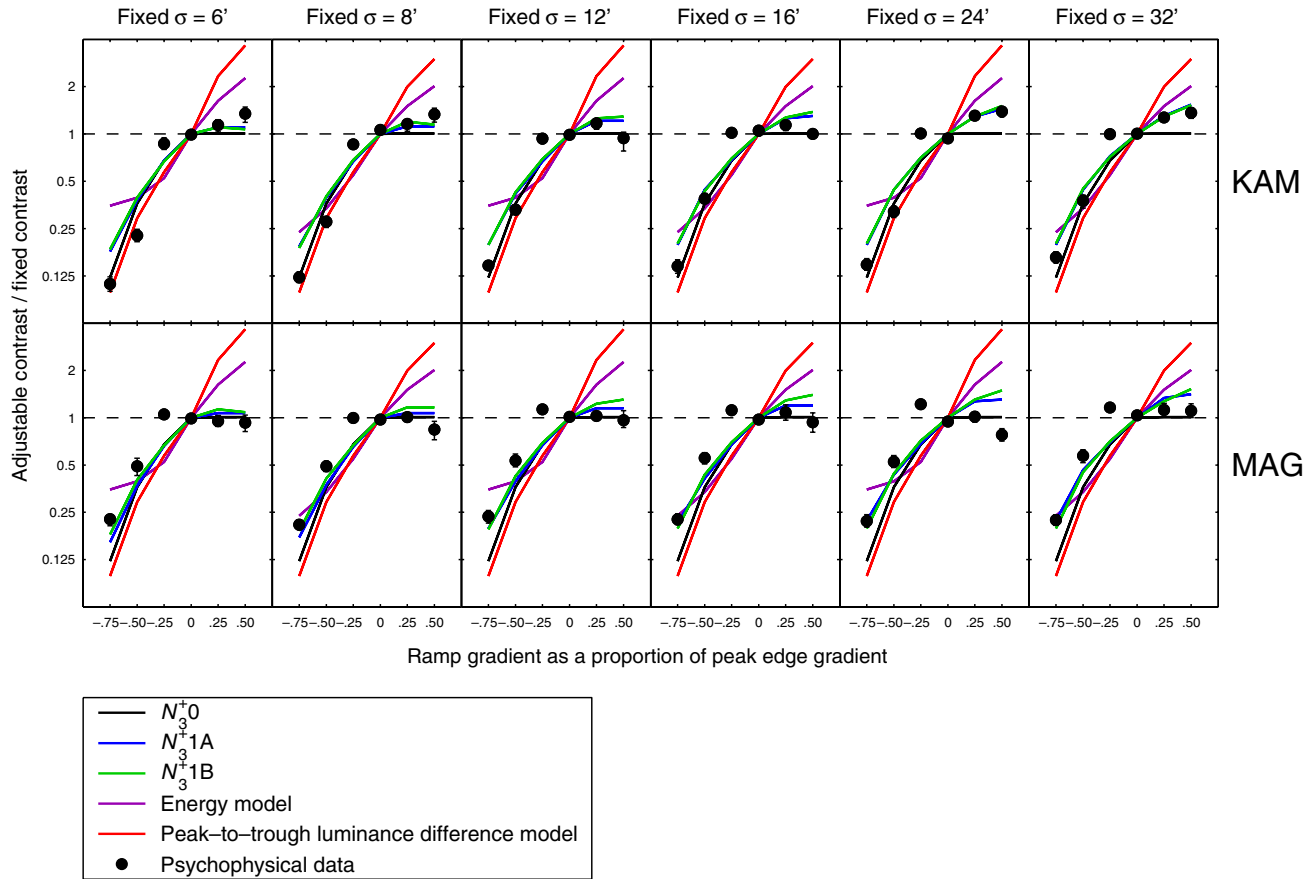


Fig. 4. Results of Experiment 2. Circles indicate psychophysical data. The vertical axis gives the ratio of adjustable edge contrast to fixed edge contrast when the two edges perceptually matched in contrast. The horizontal dashed lines indicate the results that would have occurred if perceived contrast was unaffected by the added ramp. The red lines show the predicted results if perceived contrast of the central edge was a function of the local peak-to-trough luminance difference. The purple lines show the predictions of the best-fitting energy model. The black lines show the predictions of N_3^+0 . The blue and green lines show the predictions of N_3^+1A and N_3^+1B . The numerical values plotted in this figure are given in the [supplementary data](#).

2.6.1. The N_3^+ models

N_3^+1A and N_3^+1B each have a 2-parameter transducer in place of the half-wave rectifier found in the N_3^+0 model. For each subject, we found the single pair of parameters that fitted best across all three experiments reported in this paper and the accompanying paper (May & Georgeson, 2007). Details of the simulation methods, parameter-fitting, and parameter values are given in May and Georgeson (2007, Section 6). Table 2 gives the RMS error of the fit of each model to the data.

N_3^+0 gave a reasonable fit to most of the data. The RMS error for Experiment 1 was 2.11 dB for KAM and 2.07 dB for MAG, and for Experiment 2 it was 1.93 dB for KAM and 3.14 dB for MAG. However, the model failed to account for two aspects of the data. First, the predictions are scale-invariant, i.e., the predicted graph is the same for each fixed edge scale. Although the effect of the negative ramp on perceived contrast (Fig. 4) was largely scale-invariant, its effect on perceived blur (Fig. 3) increased substantially with increasing perceived edge scale, for both subjects, so that N_3^+0 only fitted the data well for the sharpest edges. Second, N_3^+0 predicted that the positive ramp would have

no effect but, for KAM, the positive ramp caused a slight increase in perceived blur and contrast.

Both of the above aspects of the data are explained by N_3^+1A and N_3^+1B . First, consider the effect of the negative ramp on perceived blur (Experiment 1). Fig. 3 shows that the best-fitting models accurately predicted the size of the effect at all fixed edge scales, correctly explaining the increase in effect size with increasing scale. The RMS error for this experiment was no more than 1.40 dB for any subject or transducer (see Table 2). Interestingly, Fig. 4 shows that, as well as explaining the increase in effect size with increasing scale in Experiment 1, the best-fitting models were also able to account for the fact that the effect of the negative ramp on perceived contrast was largely scale-invariant (Experiment 2). These models also account for the slight increase in perceived blur and contrast caused by the positive ramp. Fig. 2 shows that this occurs because the positive ramp lifts the 1st derivative above the threshold-like nonlinearity, so the signal is not truncated, and fits a wider template, so the edge looks more blurred. The effect on perceived blur is then transferred to perceived contrast because, in the models, estimated contrast is an increasing

Table 2
RMS errors (in dB) of the fit of six models to the data

	Experiment 1		Experiment 2		Experiment 3			All experiments		
	KAM	MAG	KAM	MAG	KAM	MAG	PAA	KAM	MAG	PAA
N_3^+0	2.11	2.07	1.93	3.14	3.55	2.96	2.13	2.99	3.02	2.93
N_3^+1A	1.36	1.24	2.15	2.40	2.35	2.09	0.90	1.98	1.85	1.59
N_3^+1B	1.40	1.39	2.23	2.46	2.16	1.61	0.76	1.95	1.82	1.53
Derivative ratio	2.62	2.69	—	—	—	—	—	—	—	—
Energy model	—	—	4.25	4.62	0.796	2.26	1.82	—	—	—
Peak-to-trough luminance difference	—	—	2.33	4.31	2.68	2.08	1.22	—	—	—

N_3^+1A and N_3^+1B had the parameters reported by May and Georgeson (2007). N_3^+0 is parameter-free. The conditions with a positive ramp were excluded from the RMS error measure for the peak-to-trough luminance difference model in Experiment 2. The “All Experiments” column gives the RMS errors across all experiments in this paper and May and Georgeson (2007).

function of estimated edge scale (May & Georgeson, 2007, Eq. (2)). The predicted increase in perceived blur and contrast matches the data well for KAM, but overestimates the effect for MAG, who showed virtually no effect of positive ramp. However, the predicted effect of the positive ramp is very small, except in the case of the highest edge scale in Experiment 1, so the discrepancy is mainly based on a couple of data points, and could easily be due to noise in the data.

To demonstrate the key role of the rectifier/transducers in allowing the N_3^+ models to explain the data, we ran N_3^+0 without the half-wave rectifier. This linear filter model showed a negligible effect of the ramp.

One aspect of the data not explained by any of the models is that the shallowest negative ramp ($r' = -0.25$) had virtually no effect on perceived contrast. All versions of the N_3^+ model predicted that this ramp gradient would cause a reduction in perceived blur and contrast. This prediction was verified for perceived blur, but not contrast, and this largely accounts for the higher RMS error for the models fits in Experiment 2. We do not have an explanation for this discrepancy.

2.6.2. Estimating blur from the spatial separation between peak and trough in the 2nd derivative

As argued earlier, the response of a pure 2nd derivative operator at the edge location would be completely unaffected by the added ramp, but a larger 2nd derivative operator might be affected by the ends of the ramp. However, we found that, even for a Gaussian 2nd derivative operator with scale equal to the edge scale, the ramp had only a small effect on the peak-to-trough separation. With operator scale matched to the edge scale, the effect of the ramp was negligible for the 6', 12', and 24' edges; for the 8', 16', and 32' edges (which had a smaller ratio of window width to edge scale), the steepest negative ramp reduced the peak-to-trough separation by about 5%, which was much smaller than the observed effect on perceived blur in Experiment 1. We simulated Experiment 1 with 65 different operator scales between 1 and 2 times the edge scale. No scale gave a good fit to the data across all the ramp gradients when estimating blur from peak-to-trough

separation in the 2nd derivative. Operators with scale greater than 1.61 times the edge scale were unable to detect some of the target edges. It remains possible that a model using peak-to-trough separation might be able to explain the data by selecting a different operator scale for each stimulus, but we know of no model that would do so: Elder and Zucker's (1998) algorithm selects the *smallest* reliable operator scale at each point, and we have shown that even quite large operators are barely affected by the ends of the ramp. The 2nd derivative operators in MIRAGE (Watt & Morgan, 1985) are also too small to be affected by the ends of the ramp: the largest operator in MIRAGE has a scale of 2.8 arcmin. This is less than half the scale of the sharpest edge in Experiment 1, so the effect of the ramp on the peak-to-trough separation in MIRAGE would be negligible.

2.6.3. Estimating blur with the derivative ratio model

In Georgeson's (1994) derivative ratio model, perceived blur is a function of the ratio of the 1st to 3rd derivative at the edge location. This model predicts that, when the adjustable and fixed edges in Experiment 1 match in perceived blur, the ratio of their scales (plotted on the vertical axis in Fig. 3) will be given by

$$\frac{\sigma_{\text{adjustable}}}{\sigma_{\text{fixed}}} = \sqrt{1 + r'},$$

where r' is the ratio of ramp gradient to peak edge gradient, defined in Eq. (2) (this formula is derived in Appendix A). As shown in Fig. 3, the derivative ratio model gave scale-invariant predictions that were quite similar to the predictions of N_3^+0 : it fitted the data quite well for the 6' edge, but the overall fit to the data was worse than any version of N_3^+ (see Table 2).

2.6.4. Estimating contrast with the energy model

An alternative to the derivative-based edge coding models considered so far is the energy model (Morrone & Burr, 1988; Morrone & Owens, 1987). This model defines features as peaks of phase congruency, and detects them by detecting peaks in the energy function, found by adding the squared outputs of odd and even filters. Different types of feature (e.g., edge, bar, etc.) can be classified by

comparing the outputs of the odd and even filters at the energy peaks (Venkatesh & Owens, 1990).

The original energy model only *detected* edges, without estimating contrast or blur, but Morrone et al. (1994) extended the energy model to estimate edge contrast. We implemented a very similar model to Morrone et al. Like their model, ours was implemented in 1D. The amplitude spectra of the filters were Gaussian functions of log frequency. The amplitude, A , at spatial frequency, f , is given by

$$A(f) = \exp\left(\frac{-(2\ln(f/p))^2}{\ln(2)b^2}\right),$$

where p is the frequency of the peak in the amplitude spectrum, and b is the bandwidth in octaves (full width at half height). The filter kernels were constructed by adding sinusoids together, to give the following impulse response function, F :

$$F(n) = \frac{1}{N} \sum_{f=1}^{N/2} A(f) \cos(2\pi fn/N + \phi).$$

n is the pixel location, N is the number of pixels in the impulse response function (4096 in our implementation), and f is the spatial frequency of each component, in cycles per image. The phase, ϕ , was 0 for the even filters, and $-\pi/2$ for the odd filters. Stimuli were 4096 pixels wide: the central 512 pixels contained the normalized stimulus, found by dividing the expression in Eq. (1) by L_0 , and subtracting 1; the remaining pixels were set to zero. Each stimulus was filtered with a wide range of filters with different peak frequency; the step between channels was an eighth of an octave. As in Morrone et al.'s (1994) implementation, all the filters gave the same response amplitude to a sinusoid with the peak frequency. The model found all the spatial frequency channels with an energy peak at the edge location (and an odd-symmetric filter response of the correct sign); then perceived contrast was determined by the strongest odd-symmetric response from the selected channels. Filter bandwidth was a free parameter of the model, and was constant across all channels. Both subjects' data from Experiment 2 were fit best by the same bandwidth: 4.75 octaves. With this bandwidth, the energy model fitted KAM's data quite well (RMS error = 2.07 dB), but deviated somewhat from MAG's data (RMS error = 4.05 dB). However, this bandwidth is implausibly high (De Valois, Albrecht, & Thorell, 1982; Wilson, McFarlane, & Phillips, 1983), and a bandwidth of this size prevents the energy model from detecting the central edge in some of the stimuli in Experiment 3 (described later). We found the bandwidth that fitted best across all the conditions of Experiments 2 and 3. The best-fitting bandwidth, for both subjects, was a more plausible 2.25 octaves. With this bandwidth, the energy model's predictions were qualitatively in the right direction (see Fig. 4), but the error in the model's fit to the data from Experiment 2 was worse than any of the other models

tested, and was substantially worse than any of the N_3^+ models (see Table 2).

2.6.5. Estimating contrast from peak-to-trough luminance difference

We also modelled Experiment 2 using a model in which the perceived contrast was a function of the local peak-to-trough difference in luminance. This model gives a fairly accurate prediction of the perceived contrast for the negative ramp stimuli, showing that perception is nearly veridical in these conditions (see Fig. 4). However, the peak-to-trough model clearly could not account for the fact that the effect of the positive ramp is small because, with a positive ramp, the nearest peak and trough in luminance are the edges of the window, which differ greatly in luminance when the positive ramp gradient is high. Table 2 shows the RMS error of the peak-to-trough model's predictions for the negative and zero ramp stimuli only. For KAM's data, the RMS error for the peak-to-trough model was only slightly worse than the various versions of the N_3^+ model; for MAG's data, the peak-to-trough model was clearly worse than the N_3^+ models. It is interesting that the N_3^+ model can give a contrast estimation that is a reasonably close match to the local peak-to-trough contrast: the estimated contrast in N_3^+ is not based on the local peaks and troughs at all, and there is no *a priori* reason to suppose that the N_3^+ models should provide veridical estimates of edge profiles other than Gaussian integrals.

3. Experiment 3

In the N_3^+ models, the reduction in perceived blur due to the negative ramp occurs because the 1st derivative signal is truncated by the threshold-like transducer. The reduction in perceived contrast occurs because, in the models, estimated contrast is an increasing function of estimated blur. Our explanation of the effect of the negative ramp predicts that a similar effect should be obtained by physically truncating the stimulus. Georgeson (2001) has already confirmed this for edge blur. The purpose of Experiment 3 was to test this prediction for perceived contrast.

3.1. Methods and stimuli

The subjects were KAM and MAG, and an experienced observer, PAA, who was unaware of the purposes of the experiment. All subjects had corrected-to-normal vision. The apparatus was the same as in Experiments 1 and 2.

Experiment 3 had two independent variables: fixed edge scale (2, 4, 8, and 16 arcmin) and window width (20 or 40 arcmin). All eight fixed edge stimuli are shown in Fig. 5. The experiment used the same contrast-matching method as Experiment 2: for each fixed edge, we found the contrast of an adjustable edge that matched the fixed edge in perceived contrast. Each of the eight conditions was assigned to a different staircase, which controlled the contrast of the adjustable edge. The starting contrast for

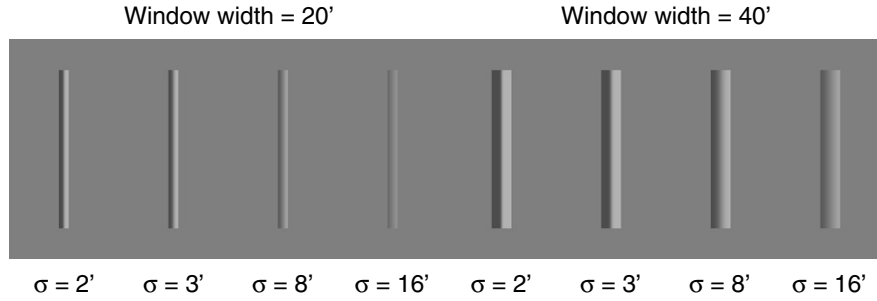


Fig. 5. The fixed edge stimuli used in Experiment 3. The adjustable edge in each condition had the same scale and window width as the leftmost stimulus in this figure (scale 2', window width 20'), but its contrast was varied by the staircase.

each staircase was 10–14 dB above or below the fixed edge contrast, and the sign of this offset was balanced within and between sessions. All eight conditions were tested within one session. Each subject completed eight similar sessions.

Edge stimuli were standard Gaussian edges constructed according to Eq. (1), with r set to 0. The window function, w , was a sharp-edged rectangle with height 320 arcmin. Window width and scale of the fixed edge were independent variables manipulated by the experiment. The Michel-

son contrast of the fixed edge was always 0.2 before windowing, but the window caused the peak-to-trough luminance difference to decrease with decreasing window width and increasing edge scale. The adjustable edge had a scale of 2' and a window width of 20'. Viewing distance was 152.3 cm, which gave a screen resolution of 1' per pixel.

3.2. Results and discussion

Fig. 6 shows the results of the experiment, along with the model predictions. As the fixed edge scale increased, the peak-to-trough luminance difference was substantially reduced for the 20' fixed edge window, and less so for the 40' window.

N_3^+1A and N_3^+1B gave a reasonably good fit to the data. As with Experiment 2, the predictions of these models were surprisingly close to those of the peak-to-trough luminance difference model. The energy model also accurately predicted the results of this experiment. As reported earlier, the best-fitting bandwidth across Experiments 2 and 3 was 2.25 octaves for both KAM and MAG. Subject PAA only participated in Experiment 3, and his best-fitting bandwidth was 0.5 octaves. This was the narrowest bandwidth that we used, since anything narrower would have been quite implausible (De Valois et al., 1982; Wilson et al., 1983). The mean RMS error for the energy model across the three subjects (1.63 dB) was slightly better than N_3^+1A (1.78 dB) but slightly worse than N_3^+1B (1.51 dB). N_3^+0 gave a substantially worse fit to the data than the other models we considered. In particular, the estimated contrast actually *increased* when the window width was 5 times the edges scale, despite a slight decrease in peak-to-trough luminance difference.

4. General discussion

We began by considering the arguments for different types of derivative operator. Elder and Zucker (1998) presented a persuasive argument that unreliable gradient peaks are easier to remove from the 2nd derivative than the 1st derivative. Experiments 1 and 2 were devised as a test of the hypothesis that the first stage of edge processing

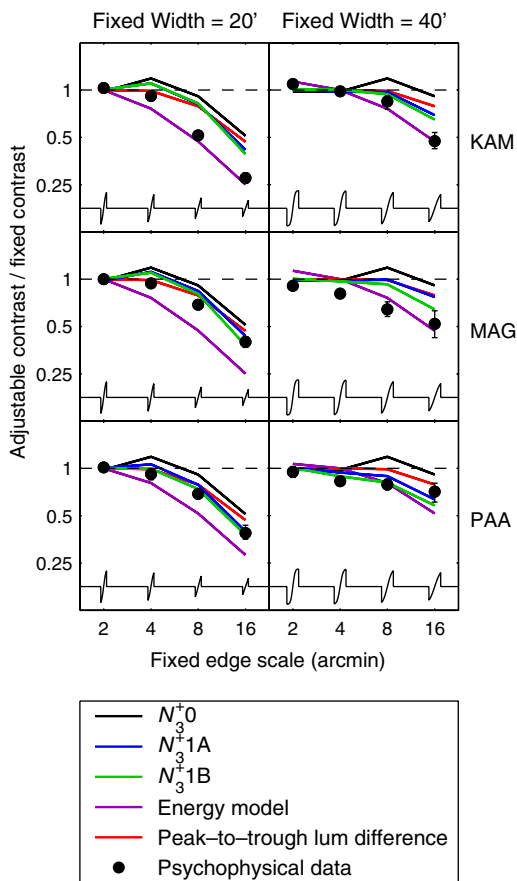


Fig. 6. Results of Experiment 3. The graphs show the psychophysical data, and the predictions of five models: N_3^+0 , N_3^+1A , N_3^+1B , the best-fitting energy model, and the peak-to-trough luminance difference model. The traces along the bottom of each panel show the profile of the fixed edge in each condition. The numerical values plotted in this figure are given in the supplementary data.

in human vision is a 2nd derivative operation. This hypothesis predicts that, if edge blur and contrast are estimated from local cues, then these perceptual variables will be unaffected by the addition of a linear ramp. This prediction was strongly violated. The perceptually matched blur and contrast were markedly reduced by the negative ramp and, in some cases, were slightly increased by the positive ramp. The data are readily explained by placing a half-wave rectifier, or similar transducer, after the 1st derivative, as in the N_3^+ models (Georgeson et al., in press; May & Georgeson, 2007). The rectifier has its effect by truncating the 1st derivative signal when the negative ramp is added. This explanation predicts that a similar effect on perceived blur and contrast could be achieved by physically truncating the stimulus. This had already been confirmed for perception of edge blur, and Experiment 3 confirmed this prediction for perceived contrast. N_3^+1A and N_3^+1B gave a reasonably good fit to these data.

Although our data rule out the 2nd-derivative-based algorithm of Elder and Zucker (1998) as a model of human vision, their argument that unreliable gradient peaks are easily removed from the 2nd derivative could still be applied to our model. In the N_3^+ models, the transducer after the 1st derivative is followed by a 2nd derivative operation. This operation could be separated into two separate 1st derivative operations, so the full sequence of operations would be derivative \rightarrow transducer \rightarrow derivative \rightarrow derivative. A threshold applied to the output of the second derivative operation could remove unreliable gradient peaks, which would otherwise have caused spurious peaks in the 3rd derivative. In our noise-free model, we obtained a satisfactory fit to our data without a further transducer at this stage: the RMS error of the model's fit to all the data from this paper and the accompanying paper (May & Georgeson, 2007) was less than 2 dB for each subject, which is acceptable, given the diversity of the stimuli, the small number of parameters (2 per subject), and the large number of data points (120 for KAM and MAG, and 48 for PAA). We could probably improve the fit of the model to the data by adding further transducers at other stages of the model, or by increasing the number of parameters in each transducer. However, it is doubtful that this would give a substantially better insight into the key mechanisms underlying the effects that we have described.

As an alternative to derivative-based models, we also considered the energy model. Using a similar implementation to Morrone et al. (1994), we found that the energy model fitted the data from Experiment 3 about as well as N_3^+1A and N_3^+1B , but it deviated somewhat from the data from Experiment 2. There has never been a biologically plausible attempt to use the energy model to estimate edge blur. As pointed out by Wang and Simoncelli (2004), phase congruency measures are unaffected by the blur of an edge. Some steps have been taken towards producing models of blur estimation using phase information. For example, Wang and Simoncelli (2004) found that the phases of fine-scale wavelet coefficients are well predicted from

larger-scale coefficients in the vicinity of a sharp edge, but not a blurred edge, showing that blur could, in principle, be estimated from the phase prediction error. However, their technique was designed to estimate the blur of whole images; they did not produce a blur metric that could be used to predict the perceived blur of a given image feature. Therefore, the energy model is currently unable to make predictions for experiments on the perception of edge blur. Given that N_3^+1A and N_3^+1B can account for perceived blur as well as contrast, and that they gave an overall better fit to our data than the energy model, despite being fitted to five experiments, rather than just two, we conclude that N_3^+1A and N_3^+1B provide a more wide-ranging and accurate model of perceived edge appearance than any current version of the energy model.

5. Conclusion

The effect of an added luminance ramp on perceived blur and contrast of an edge is difficult to explain using edge processing models in which the initial stage is a 2nd derivative operation (e.g., Elder & Zucker, 1998; Georgeson, 1992; Kingdom & Moulden, 1992; Marr & Hildreth, 1980; Watt & Morgan, 1985). The energy model (Morrone & Burr, 1988; Morrone et al., 1994) makes no predictions about perceived edge blur, and it predicted the effect of the ramp on perceived contrast for only one subject and, for this, required a filter bandwidth that was implausibly wide. The N_3^+1A and N_3^+1B models are, to our knowledge, the only ones proposed so far that accurately account for the striking variations in perceived blur and contrast that result from the addition of a linear ramp to an edge.

Acknowledgments

This work was supported by grants from the Wellcome Trust (056093/B/98) and the EPSRC (GR/S07261/01) to MAG, and by an Aston University studentship to KAM.

Appendix A

Here, we derive predictions of Georgeson's (1994) derivative ratio model of blur coding for Experiment 1. In the model, perceived blur is a function of the ratio of 1st to 3rd derivatives at the edge location. Without loss of generality, consider the edge located at $x = 0$, described in Eq. (1). We need the 1st and 3rd derivatives, in the x direction, of the stimulus, L , at $x = 0$. We can ignore the window function, $w(x, y)$, since it does not affect the derivatives at $x = 0$. The 1st derivative, $L_x(\cdot; \sigma)$, is given by

$$\begin{aligned} L_x(x; \sigma) &= \frac{2CL_0}{\sigma\sqrt{2\pi}} \exp\left(\frac{-x^2}{2\sigma^2}\right) + L_0r \\ &= \frac{2CL_0}{\sigma\sqrt{2\pi}} \left[\exp\left(\frac{-x^2}{2\sigma^2}\right) + r' \right], \end{aligned}$$

where r' is defined in Eq. (2). The 3rd derivative, $L_{xxx}(\cdot; \sigma)$, is given by

$$L_{xxx}(x; \sigma) = \frac{2CL_0}{\sigma^3\sqrt{2\pi}} \left(\frac{x^2}{\sigma^2} - 1 \right) \exp\left(\frac{-x^2}{2\sigma^2}\right).$$

So the derivative ratio, L_x/L_{xxx} , at the edge location, $x = 0$, is given by $-\sigma^2(1 + r')$. For the fixed edge, $L_x/L_{xxx} = -\sigma_{\text{fixed}}^2(1 + r')$. For the adjustable edge, the ramp gradient, r' , is zero, so $L_x/L_{xxx} = -\sigma_{\text{adjustable}}^2$. The derivative ratio model predicts that the fixed and adjustable edges will match in perceived blur when their L_x/L_{xxx} ratios are equal, i.e., when $-\sigma_{\text{adjustable}}^2 = -\sigma_{\text{fixed}}^2(1 + r')$. This gives

$$\frac{\sigma_{\text{adjustable}}}{\sigma_{\text{fixed}}} = \sqrt{1 + r'}.$$

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.visres.2007.02.018](https://doi.org/10.1016/j.visres.2007.02.018).

References

- Boie, R. A., Cox, I. J., & Rehak, P. (1986). On optimum edge recognition using matched filters. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 100–108).
- Canny, J. F. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 679–698.
- De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22, 545–559.
- Elder, J. H., & Zucker, S. W. (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20, 699–716.
- Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society of London B*, 249, 235–245.
- Georgeson, M. A. (1994). From filters to features: Location, orientation, contrast and blur. In *Higher-Order Processing in the Visual System (Ciba Foundation Symposium 184)*, pp. 147–165. Chichester: Wiley.
- Georgeson, M. A. (2001). Seeing edge blur: Receptive fields as multi-scale neural templates. *Journal of Vision*, 1, 438a.
- Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision* (in press).
- Kingdom, F., & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, 32, 1565–1582.
- Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, 207, 187–217.
- May, K. A., & Georgeson, M. A. (2007). Blurred edges look faint, and faint edges look sharp: The effect of a gradient threshold in a multi-scale edge coding model. *Vision Research*, 47, 1705–1720.
- Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B*, 235, 221–245.
- Morrone, C., Burr, D., & Ross, J. (1994). Illusory brightness step in the Chevreul illusion. *Vision Research*, 34, 1567–1574.
- Morrone, M. C., & Owens, R. A. (1987). Feature detection from local energy. *Pattern Recognition Letters*, 6, 303–313.
- Sarkar, S., & Boyer, K. L. (1991). On optimal infinite impulse response edge detection filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13, 1154–1171.
- Tagare, H. D., & deFigueiredo, R. J. P. (1990). On the localization performance measure and optimal edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12, 1186–1190.
- Venkatesh, S., & Owens, R. (1990). On the classification of image features. *Pattern Recognition Letters*, 11, 339–349.
- Wang, Z., & Simoncelli, E. P. (2004). Local phase coherence and the perception of blur. *Advances in Neural Information Processing Systems*, 16.
- Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: Evidence for spatial primitives in human vision. *Vision Research*, 23, 1465–1477.
- Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, 25, 1661–1674.
- Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23, 873–882.