Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 1705–1720
Vision Research
www.elsevier.com/locate/visres

Blurred edges look faint, and faint edges look sharp: The effect of a gradient threshold in a multi-scale edge coding model

Keith A. May *,[1], Mark A. Georgeson

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

Received 19 April 2006; received in revised form 4 February 2007

Abstract

A multi-scale model of edge coding based on normalized Gaussian derivative filters successfully predicts perceived scale (blur) for a wide variety of edge profiles [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*]. Our 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. Peaks in the inverted normalized 3rd derivative across space and scale indicate the positions and scales of the edges. The edge contrast can be estimated from the height of the peak. The model provides a veridical estimate of the scale and contrast of edges that have a Gaussian integral profile. Therefore, since scale and contrast are independent stimulus parameters, the model predicts that the perceived value of either of these parameters should be unaffected by changes in the other. This prediction was found to be incorrect: reducing the contrast of an edge made it look sharper, and increasing its scale led to a decrease in the perceived contrast. Our model can account for these effects when the simple half-wave rectifier after the 1st derivative is replaced by a smoothed threshold function described by two parameters. For each subject, one pair of parameters provided a satisfactory fit to the data from all the experiments presented here and in the accompanying paper [May, K. A. & Georgeson, M. A. (2007). Added luminance ramp alters perceived edge blur and contrast: A critical test for derivative-based models of edge coding. *Vision Research, 47*, 1721–1731]. Thus, when we allow for the visual system's insensitivity to very shallow luminance gradients, our multi-scale model can be extended to edge coding over a wide range of contrasts and blurs.
© 2007 Elsevier Ltd. All rights reserved.

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

1. Introduction

Since the field of image processing first appeared, it has been considered that a useful function of early visual processing is to encode the image in terms of primitive features such as edges (Kirsch, Cahn, Ray, & Urban, 1957; Kovasznay & Joseph, 1953, 1955; Roberts, 1965). Edge maps are highly compact representations that make explicit many behaviourally relevant transitions in surface properties (Marr & Hildreth, 1980). If an edge map contains information about edge contrast, scale (i.e. blur), and mean luminance at the edge location, then a close approximation to the original image can be reconstructed, demonstrating that an edge map that includes these properties preserves most of the information in the image (Elder, 1999).

As well as helping to preserve the information content of the processed image, the representation of edge contrast and scale makes some of this information explicit in ways that may be useful in a number of higher-level computations. Edge contrast can indicate the magnitude of transitions in surface properties or illumination. Edge scale is also very informative, allowing the visual system make inferences about shadows, surface curvature, and optical

* 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, Bradford, West Yorkshire, BD7 1DP, UK.

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.02.012

blur/defocus (Elder & Zucker, 1998; Marshall, Burbeck, Ariely, Rolland, & Martin, 1996; Mather, 1996, 1997; Mather & Smith, 2002). This paper is concerned with how edge contrast and scale are encoded in human vision.

Many models of edge processing in biological vision have been based on derivative operations. Edges are defined as changes in image intensity, and are usually located by finding peaks in the 1st derivative or zero-crossings (ZCs) in the 2nd derivative, using derivative operators with a range of scales (i.e. sizes). The main alternative to derivative-based models is the energy model; 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 (Morrone & Burr, 1988; Morrone & Owens, 1987). The original energy model did not estimate the edge contrast or blur; Morrone, Burr, and Ross (1994) extended it to predict perceived edge contrast, but this new model did not estimate edge blur. Indeed, phase congruency itself provides "no information about blur" (Wang & Simoncelli, 2004). Wang and Simoncelli showed that phase information could, in principle, be used to measure blur, but their algorithm was designed to estimate the blur of whole images, and did not provide a means of estimating the blur of individual edges. We know of only one attempt to use an energy-based algorithm to estimate the blur of individual edges, and this was a machine vision algorithm that was not a plausible model of human vision (Kisworo, Venkatesh, & West, 1994).

Marr was the first to seriously consider how edges might be detected in biological visual systems (Marr, 1976; Marr & Hildreth, 1980), but his algorithms for estimating edge blur and contrast were not supported by any computational justification or empirical data.

Watt and Morgan (1983) argued that perceived edge blur was determined from the spatial separation between peak and trough in the 2nd derivative of the retinal image. Their main evidence for this was that blur discrimination thresholds for three different types of edge profile were closely matched when blur was defined in this way. Watt and Morgan's (1983) model of blur coding was adapted slightly and incorporated into MIRAGE, a model of feature coding in human vision (Watt & Morgan, 1985).

Georgeson (1994) proposed that perceived blur is a function of the ratio of 1st to 3rd derivative at the edge location. This model accurately predicted the results of an experiment in which subjects were asked to match the perceived blurs of Gaussian edges[2] and sine wave edges; Watt and Morgan's (1983) model overestimated the perceived blur of sine wave edges compared with Gaussian edges.

The derivative operators in MIRAGE were normalized so that they all had the same amplitude in the space domain, which would tend to boost the large-scale (low spatial frequency) channels. Several approaches to edge detection in machine vision have involved other types of normalization, which vary the amplitude of the operator as a function of scale. These approaches use the concept of a *scale-space representation* (Witkin, 1983), which has a dimension representing the operator scale, σ, in addition to the two spatial dimensions of the filtered image.

The changing response across scale can be used to measure edge blur, either by fitting a template to the response profile across scale (Dijk, van Ginkel, van Asselt, van Vliet, & Verbeek, 2003; Zhang & Bergholm, 1997) or by finding the scale of the operator with the strongest response (Korn, 1988; Lindeberg, 1998a; van Warmerdam & Algazi, 1989). The latter approach was pioneered by Korn (1988), using Gaussian 1st derivative operators, but his normalization factor ($\sigma\sqrt{2\pi}$, where σ is the operator scale) would only give rise to a peak across scale when the edge of interest was flanked by other nearby edges: the response profile across scale due to an isolated Gaussian edge increased monotonically with increasing operator scale (van Warmerdam & Algazi, 1989; Zhang & Bergholm, 1997). van Warmerdam and Algazi (1989) noted that a peak *does* occur if the normalization factor is "any power of σ *between* [but not including] zero and one." (p. 977). If the normalization factor is $\sigma^{1/2}$, then, for a Gaussian edge stimulus, the edge scale matches the scale of the most strongly responding operator (Lindeberg, 1998a).

Lindeberg (1998a) referred to the normalized 1st derivative as the *edge strength*. He presented an algorithm that marked edges at points in scale-space that were peaks in the edge strength across both space and scale. The position of the peak along the scale dimension gave the edge scale. This algorithm provides a good solution to the problem of how to deal with information from multiple scales. Small operators resolve fine details well, but produce a mass of broken line segments on blurred edges, such as shadows; large operators process blurred edges well, but distort the edge locations, and fail to capture fine details. Lindeberg proposed that each edge should be detected by an operator with the same scale as the edge.

1.1. A new model of edge detection and blur perception in human vision

Georgeson, May, Freeman, and Hesse (in press) proposed a model of edge processing in human vision along quite similar lines to the algorithms described by van Warmerdam and Algazi (1989) and Lindeberg (1998a). They presented the results of several psychophysical experiments in which subjects had to match the blur of a Gaussian edge to a non-Gaussian edge. The non-Gaussian stimuli included mixtures of two superimposed Gaussian edges of different scale, Gaussian edges sharpened with a non-linear transducer, Gaussian derivatives of different order, blurred square waves with different numbers of harmonics, half-period sine edges, a full period of a sine wave

[2] We define a *Gaussian edge* with scale σ as the integral of a Gaussian function with standard deviation σ.

grating, and multiple periods of a sine wave grating. Georgeson et al. examined the predictions of a family of Gaussian scale-space models, including Lindeberg's (1998a) 1st-derivative-based algorithm described above, and found that one model alone gave a strikingly good fit to the data. This model, which they named N_3^+, is described in detail in the next section. The subscript, 3, refers to the fact that the model looked for peaks in the 3rd derivative scale-space representation. The superscript, +, refers to the fact that the 1st derivative stage was followed by a half-wave rectifier, which only allowed positive values to pass through to the next stage.

Although edge blur measures based on the 3rd derivative have been proposed before (Lindeberg, 1998a; van Warmerdam & Algazi, 1989), no previous algorithm had *detected* edges by finding peaks in the 3rd derivative. Georgeson et al. argued in favour of this 3rd-derivative-based model because it provided a very good fit to their data, with no free parameters: models based on the 1st or 2nd derivative did not fit the data so well. The rectifier after the 1st derivative was needed because the 3rd derivative normally contains far more peaks than perceived edges: these unwanted peaks are removed by the rectifier. The rectifier also improved the model's fit to the blur-matching data in cases where the stimuli were periodic, as in a grating.

Most of the data that Georgeson et al. presented in favour of the N_3^+ model were from blur-matching experiments, but the 3rd-derivative approach is also supported by a new phenomenon: Mach Edges. Georgeson (2006) reported that in triangle-wave gratings, blurred to different extents, edges are seen at peaks in the 3rd derivative, without corresponding peaks in the 1st derivative, or ZCs in the 2nd derivative.

The N_3^+ model returns a veridical estimate of the scale (i.e. blur) of a Gaussian edge. Also, as shown in the next section, it can be extended to give a veridical contrast estimate. Therefore, since scale and contrast are independent parameters of the stimulus, the model predicts that the perception of scale should be independent of contrast and vice-versa. This prediction is at odds with Georgeson's (1994) finding that edges looked sharper when their contrast was reduced. This was a potentially serious problem for the N_3^+ model. Since there had been no other reports of the effect of contrast on perceived scale, we first set out to confirm and extend those earlier findings. We describe the effects of contrast on perceived edge scale (Experiment 1), the effects of edge scale on perceived contrast (Experiment 2), and we show how these effects can be explained by N_3^+ if the half-wave rectifier that follows the first derivative operation is replaced with a threshold-like function. Before describing the experiments, we describe the N_3^+ model in more detail, and explain how it can be extended to make estimates of edge contrast. We also give an alternative description of the model in terms of template matching because, although this description is formally identical to the description based on derivative operations, it gives a better insight into the behaviour of our model in the simulations of our experiments.

2. Details of the N_3^+ model

The N_3^+ model is illustrated in Fig. 1. The full version of the model uses filters that are directional derivatives of 2D circular Gaussians (the 2D shape of the filters was determined from an experiment on the effect of edge length on blur perception, reported by Georgeson et al., in press). However, for simplicity we describe a 1D version of the model below. For an infinitely long straight edge, the 2D model reduces exactly to the 1D model applied to the edge's cross-sectional profile. The experiments in this paper used 1D edges within windows that were large enough to ensure that the differences in the predictions of the 1D and 2D versions of the model were negligible.

The model consists of a set of channels, selective for edge scale. Within each channel, the processing occurs as follows. The image is filtered with a Gaussian 1st derivative operator with scale σ_1. Then the output is half-wave rectified and filtered with a Gaussian 2nd derivative operator with scale σ_2. Within a channel that has the same polarity as an isolated edge, the half-wave rectifier has no effect, so the two derivative operations combine to form a Gaussian 3rd derivative operation with scale $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$, which we call the *channel scale*. The channel scale varies across channels. In our current implementation, $\sigma_1 = \sigma/4$. The polarity of the Gaussian 2nd derivative operator is inverted, which inverts the 3rd derivative output, so positive-gradient edges give rise to peaks rather than troughs in the 3rd derivative. The output of each blur channel is normalized, by multiplying by $\sigma^{3/2}$, to give the edge strength, S, which is a function of space and scale. With this normalization factor, the peak in response to a Gaussian edge occurs in a channel with scale equal to the edge scale.

For a vertical Gaussian edge, with scale σ_e, Michelson contrast C, and horizontal position x_0, the edge strength at a point $(x;\sigma)$ in scale-space is given by

$$S(x;\sigma) = C\sqrt{\frac{2}{\pi}}\left(\frac{\sigma}{\sigma_e^2+\sigma^2}\right)^{3/2}\left(1-\frac{(x-x_0)^2}{\sigma_e^2+\sigma^2}\right)\exp\left(\frac{-(x-x_0)^2}{2(\sigma_e^2+\sigma^2)}\right).$$

This is derived assuming that the image that forms the input to the model is the function, f, in Eq. (1):

$$L(x) = L_0(1+f(x)), \tag{1}$$

where L is the stimulus luminance and L_0 is the background luminance. For a Gaussian edge, $f(x) = C(2\Phi(x;\sigma)-1)$, where $\Phi(\cdot;\sigma)$ is the integral of a unit-area Gaussian with scale (standard deviation) σ.

It can be shown that the edge strength, S, peaks at $(x_0;\sigma_e)$, correctly indicating the position and scale of the edge. The height of the peak, S_{peak}, is given by

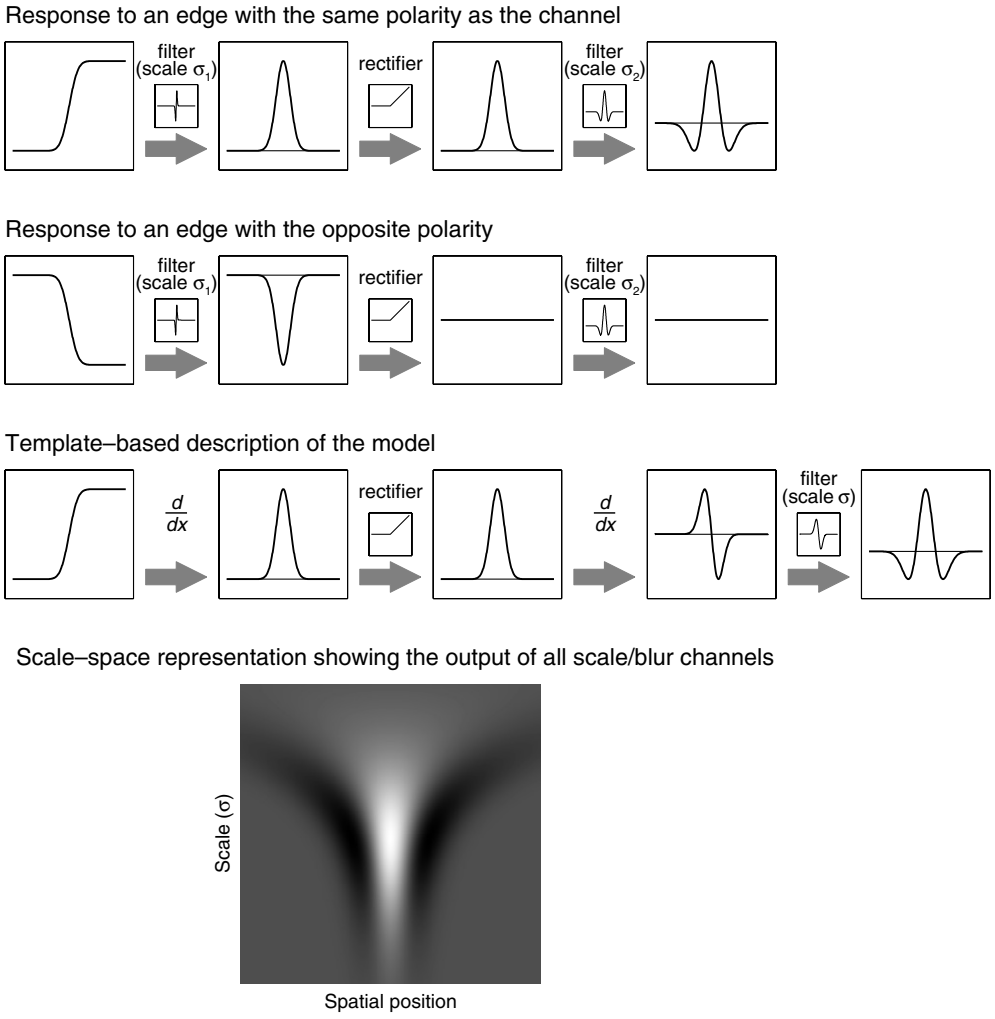

Fig. 1. The top three rows show examples of processing within one channel of the N_3^+ model. In the top row, the edge has the same polarity as the channel. For this stimulus, the half-wave rectifier has no effect, and the output gives a peak across space at the edge location. The outputs from all the channels form a scale-space representation (bottom), which has a peak across scale in the channel with scale equal to the edge scale. In the second row, the edge has the opposite polarity; its gradient in this channel is negative, so the signal is removed by the half-wave rectifier. This edge would be processed by a channel with opposite polarity, in which it would have a positive gradient. The third row shows the processing of a formally equivalent model, based on template matching. There are two pure 1st derivative operations, followed by filtering with a Gaussian 1st derivative with scale equal to the channel scale, $\sigma = \sqrt{\sigma_1^2 + \sigma_2^2}$. The Gaussian 1st derivative kernel acts as a template that is matched against the 2nd derivative of the edge: the normalized output at each position and scale in the scale-space representation represents the correlation between the 2nd derivative of the stimulus and a template with that position and scale. Peaks in correlation give the position and scale of the locally best-fitting template, so these peaks can be used both to detect the edge, and estimate its position and scale.

$$S_{\text{peak}} = \frac{\sigma_e^{-3/2} C}{2\sqrt{\pi}}.$$

Rearranging this equation gives a formula for estimating edge contrast ($\hat{C}$) from peak height and estimated edge scale ($\hat{\sigma}_e$):

$$\hat{C} = 2\sqrt{\pi}\hat{\sigma}_e^{3/2} S_{\text{peak}}. \quad (2)$$

It might appear from Eq. (2) that the estimated contrast would depend on the spatial units used to measure σ_e. This is not the case, because the Gaussian derivative operators in the model are derivatives of a Gaussian that has unit area in whichever spatial units are being used. If the spatial units were changed, then the amplitude of each operator would have to be scaled to maintain the unit area of the Gaussian from which it was derived. This would lead to a change in the value of S_{peak} which compensates exactly for the change in $\hat{\sigma}_e$.

2.1. Template matching

Lindeberg's (1998a) algorithm estimated the location and blur of each edge from the position of its peak in the Gaussian 1st derivative scale-space representation. He contrasted this algorithm with the template-matching approach of Zhang and Bergholm (1997), claming that "there is hence no need ... to fit a model to the data to estimate the degree of diffuseness." (p. 144). But in fact, as

noted by Lindeberg himself, we can view his algorithm as a "pattern matcher" (Lindeberg, 1998b, p. 83): the algorithm is formally equivalent to treating the 1st derivative of the edge as a spatial "signature", and then finding the best-fitting template. To see why this is the case, note that the Gaussian derivative operation is equivalent to a pure derivative operation, followed by filtering with a Gaussian kernel. The first operation finds the 1st derivative of the edge (the "signature"), and the second performs a correlation between the signature and the Gaussian kernel (the "template"), which is itself the 1st derivative of a Gaussian edge. It is necessary to normalize the correlation coefficients, allowing them to be compared across scale, so that the peak across scale correctly indicates the scale of the best-fitting template. Appendix A shows that the normalization factor required to normalize this correlation coefficient is $\sigma^{1/2}$, as used by Lindeberg (1998a). With this normalization factor, the peak of correlation will occur in the operator with a scale matched to the edge scale.

The N_3^+ model is also formally equivalent to the matching of a signature and template (see Fig. 1), a fact that will provide a crucial intuitive insight into the behaviour of the model in our simulations. In this case, the signature and template are the 2nd derivatives of edges. To understand this, note that the Gaussian 3rd derivative operation performed by N_3^+ is equivalent to a pure 2nd derivative operation followed by filtering with a Gaussian 1st derivative kernel. The first operation finds the 2nd derivative of the edge (the "signature"), and the second performs a correlation between the signature and the Gaussian 1st derivative kernel (the "template"), which is itself the 2nd derivative of a Gaussian edge. As with Lindeberg's algorithm, the correlation coefficient must be normalized, and Appendix B shows that the correct normalization factor in this case is $\sigma^{3/2}$, as used by Georgeson et al. (in press).

3. General psychophysical methods

3.1. Subjects

Three subjects participated in the experiments: the authors (KAM and MAG), and an experienced psychophysical observer who was unaware of the purposes of the experiments (PAA). All had corrected-to-normal vision.

3.2. Apparatus

The experiments were run on a Pentium III PC with a VSG 2/3 graphics card (Cambridge Research Systems). MATLAB (Mathworks) was used to generate images and control the experiments. The images were linearly scaled to fit the range 0–255 and stored in an 8-bit frame store on the VSG card. Stimuli were then scaled to the correct contrast, and gamma corrected, by mapping the 8-bit values onto 15-bit values. An analogue input to the monitor was generated from these 15-bit values by adding the outputs of two 8-bit digital-to-analogue converters in the VSG. Stimuli were displayed on an Eizo FlexScan 6600-M greyscale monitor at a frame rate of 110 Hz, with mean luminance 45 cd/m^2.

3.3. Stimuli

The stimuli in all the experiments were vertical edges that had the profile of the integral of a Gaussian with scale (i.e. standard deviation), σ. Examples are shown in Fig. 2. The luminance, L, at spatial position (x, y) is given by

$$L(x,y) = L_0[1 + Cw(x,y)(2\Phi(x;\sigma) - 1)], \quad (3)$$

where L_0 is the mean luminance (45 cd/m^2). C is the Michelson contrast: a positive value gives edges that are dark on the left, and a negative value gives edges that are dark on the right. w is the spatial envelope, which differed across the two experiments. 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 of half-period 25 ms, so the entire stimulus lasted for 300 ms. The value, C, in Eq. (3) refers to the contrast during the flat central period of the stimulus.

3.4. Methods

The purpose of each experiment was to find the perceived blur or contrast of a set of edges, referred to as *fixed edges*. The perception of each fixed edge was assessed by using a 1-up-1-down staircase (e.g., Levitt, 1971; Wetherill & Levitt, 1965) 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 to give a maximum-likelihood estimate of the adjustable edge scale (or contrast) that gave a perceptual match to the fixed edge.

On each trial there were two intervals separated in time by 500 ms: one randomly selected interval contained a fixed edge, and the other contained the corresponding adjustable edge, with the same polarity. Subjects clicked a mouse button to indicate which edge appeared higher in either blur or contrast.

Each staircase began at a random starting point between 10 and 14 dB above or below the fixed edge scale/contrast (an increase of 1 dB is multiplication by a factor of $10^{1/20}$). The step size started off at 8 dB, and halved with each reversal of direction, until it had reached 1 dB. The staircases terminated immediately after the 12th reversal of direction, and usually lasted for about 25 trials. Several staircases, corresponding to different conditions (i.e. different fixed edges), were interleaved within a session. On each trial, the probability that a particular staircase was selected was equal to the number of reversals remaining on that staircase divided by the total number of reversals remaining across all the staircases. For each condition of each experiment, there were eight staircases, distributed over two or more sessions, giving about 200 trials per condition.

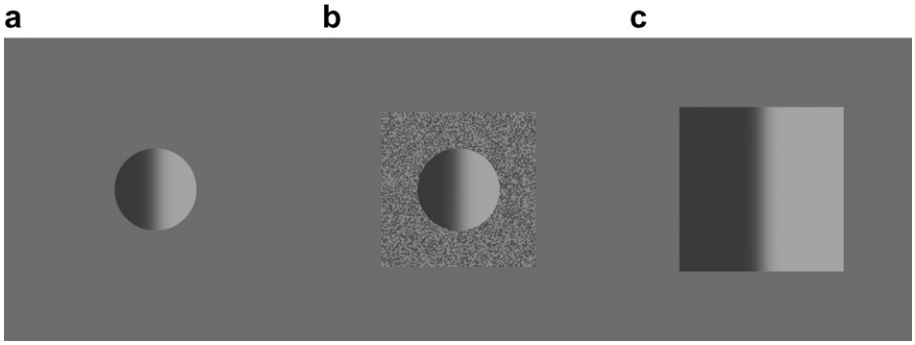

Fig. 2. Examples of stimuli used in Experiments 1 and 2. Each stimulus is a 16′ Gaussian edge. (a) A stimulus from the main part of Experiment 1. (b) A stimulus from the version of Experiment 1 with a noise surround. (c) A stimulus from Experiment 2.

All the data for each condition were pooled, to give a histogram of the proportion of times the adjustable edge was judged to be higher in scale or contrast than the fixed edge, for each adjustable edge scale/contrast level. A cumulative Gaussian psychometric function was fitted to the data using psychofit, version 2.5.1 (now renamed psignifit; see http://www.bootstrap-software.org/psignifit/), a software package which implements the maximum-likelihood method described by Wichmann and Hill (2001a). Log units were used for the domain of the psychometric function. The "lapse rate" parameters, γ and λ, which determine how far the minimum and maximum values deviate from 0 to 1, respectively, were allowed to vary between 0 and 0.05. After fitting the psychometric function, the 50% point of the function was found, which corresponded to the PSE. Confidence limits (5% and 95%) for the PSE were then found using the percentile bootstrap method implemented by psychofit (see Wichmann & Hill, 2001b); each confidence limit was based on 10,000 Monte-Carlo simulations. All error bars shown on the graphs in this paper indicate bootstrap confidence limits generated with this method.

4. Experiment 1: The effect of contrast on perceived blur

4.1. Method and stimuli

Experiment 1 had two independent variables: the contrast and scale of the fixed edge. Michelson contrast levels were 0.05, 0.1, 0.2, and 0.4; scale levels were 6, 8, 12, 16, 24, and 32 arcmin. The six levels of fixed edge scale were spread over three viewing distances, which are given in Table 1, along with the screen resolutions. 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. This ordering was intended to balance out any practice effects, and avoided two consecutive sessions at the same viewing distance. Within a session, subjects saw fixed edges with two different scales, at all four contrast levels. Each of these eight conditions was assigned to four staircases, which each controlled the scale of a different adjustable edge that had a contrast of 0.4. For one pair of these staircases, the stimuli had positive polarity (dark on the left); for the other pair, the stimuli had negative polarity. Within a pair, one staircase started above the fixed edge scale and the other started below. Data for each condition were collapsed across polarity, staircase start position, and session number, and analyzed as described earlier.

Edge stimuli were constructed according to Eq. (3). The windows (w) were circular with a flat profile and a sharp border. 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 are given in Table 1. An example stimulus is shown in Fig. 2a.

KAM performed an additional version of this experiment, in which the stimuli were surrounded by binary noise with contrast 0.2, which filled a 256-by-256 pixel square (width and height 10 deg at viewing distance 65 cm, 5 deg at viewing distance 130 cm, or 2.5 deg at 260 cm viewing distance). 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 $L_0 \times (1 \pm 0.2)$. An example is shown in Fig. 2b. The idea here was to keep the overall contrast of the stimulus

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

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.

moderately high, even when the edge contrast was very low, thereby limiting the possible action of contrast gain control mechanisms, which are not taken into account by our model.

4.2. Results

The results are shown in Fig. 3. The vertical axis of each panel represents the ratio of the adjustable edge scale to fixed edge scale when the edges perceptually matched in blur. Not surprisingly, when the adjustable and fixed edge contrasts were equal (0.4), the two edges appeared to match in blur when they were physically identical, giving a scale ratio of 1. When the contrast of the fixed edge was reduced, the scale of the perceptually matched adjustable edge decreased, as reported by Georgeson (1994), showing that edges look sharper when their contrast is reduced. As with Georgeson's (1994) data, this effect was stronger for larger-scale edges. The addition of a noise surround made little difference to the results.

4.3. Discussion—modelling a gradient threshold

Georgeson et al.'s (in press) N_3^+ model is linear with contrast level, so the positions of the peaks in the scale-space output (which determine the estimated blur) are unaffected by a change in contrast. The contrast-dependence of blur perception observed here suggests that we should incorporate some non-linearity in the response to contrast. One possibility was to replace the half-wave rectifier after the 1st derivative with a *biased rectifier*, or *threshold-linear function*, with the form $R(x) = \lfloor I(x) - T \rfloor$, where R is the output, I is the input (i.e. the output of the 1st derivative operator), T is the threshold, and $\lfloor a \rfloor = \max(a, 0)$. This approach was used in Kovasznay and Joseph's (1955) edge detector to remove noise. Fig. 4 illustrates how the threshold could make low-contrast edges appear sharper. As noted above, the N_3^+ model can be viewed as matching a template to the 2nd derivative of the edge. The threshold truncates the 1st-derivative signal, so that only the central (above-threshold) region survives. This causes the 2nd

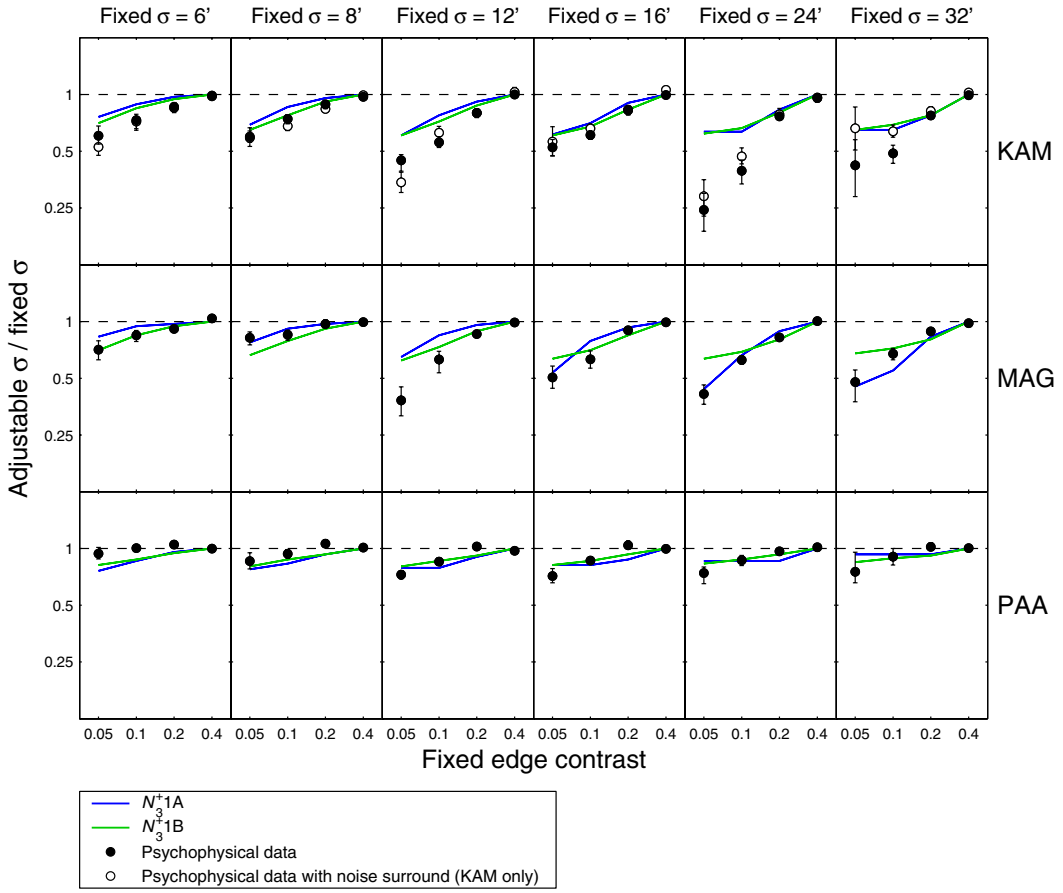

Fig. 3. Results of Experiment 1. Circles indicate psychophysical data; error bars indicate bootstrap confidence limits of 5 and 95%. The vertical axis gives the ratio of adjustable edge scale to fixed edge scale when the two edges perceptually matched in blur. Each row of panels gives data for one subject. Each column of panels gives the data for one fixed edge scale. The horizontal dashed lines indicate the results that would have occurred if perceived blur was independent of contrast, as predicted by the original version of the N_3^+ model (N_3^+0). The blue and green lines show the results predicted by N_3^+1A and N_3^+1B, in which the half-wave rectifier in N_3^+0 is replaced with Transducer A or B. The parameters of the transducers for each subject are given in Table 2. The numerical values plotted in this figure are given in the supplementary data file, "Fig3_data.txt". (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

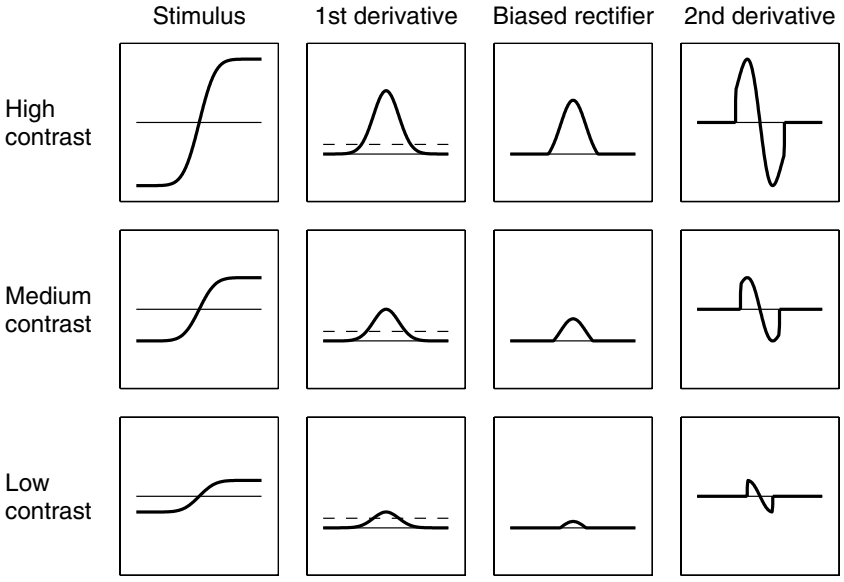

Fig. 4. An illustration of why half-wave rectification with a threshold makes low-contrast edges appear sharper. Each row shows the processing of an edge of different contrast. Each column shows a different processing stage. The threshold level is indicated by the dashed line in the 1st derivative stage. The threshold causes the 2nd derivative to be truncated. The truncation is more severe with the low-contrast edge, so the 2nd derivative is narrower, and fits a narrower template, and therefore looks sharper than the high-contrast edge.

derivative to be truncated, too, so that it fits best to a narrower template. With low-contrast edges, more of the 1st derivative falls below threshold, so the truncation is more severe, and the 2nd derivative matches a narrower template than the 2nd derivative from a higher-contrast edge. Thus, the low-contrast edge is perceived to be sharper than the high-contrast edge.

When we modelled the effect of the threshold, we found that the magnitude of the sharpening effect varied greatly with the scale of the edge—much more so than was found in the psychophysical data. We had much more success with smooth transducers which had the approximate shape of threshold functions, but which always gave above-zero outputs for above-zero inputs. These are described in the next section.

5. Two transducer models

There are two basic types of threshold function: a *soft threshold* and a *hard threshold*[3]. The soft threshold function is the standard threshold-linear function: $R = \lfloor I - T \rfloor$, whereas the output of the hard threshold function is equal to the input for inputs greater than or equal to the threshold, and zero otherwise. We investigated two transducers: a smoothed soft threshold function (Transducer A), and a smoothed hard threshold function (Transducer B).

For inputs between 0 and T, Transducer A is a power function, scaled so that its gradient is 1 at T. For inputs greater than or equal to T, the function continues as a straight line with gradient 1. Transducer A is formally defined as follows:

$$R = \begin{cases} \frac{I^p}{pT^{p-1}}, & 0 \leqslant I < T, \\ I - \frac{T(p-1)}{p}, & I \geqslant T. \end{cases}$$

Transducer B is described by the following function, for inputs $\geqslant 0$:

$$R = \frac{I^p}{I^{p-1} + T^{p-1}}.$$

Both transducers give zero outputs for negative inputs. They are illustrated in Fig. 5, for a range of exponents. In both cases, setting T to zero yields a standard half-wave rectifier, so the original N_3^+ model is a specific parameterisation of either of these models. As $p \to \infty$, the transducers tend towards pure soft (Transducer A) or hard (Transducer B) threshold functions, with threshold, T. Henceforth, we will refer to Georgeson et al.'s (in press) original N_3^+ model (containing a half-wave rectifier) as N_3^+0; the modified versions of N_3^+0, using Transducers A or B, will be referred to as N_3^+1A and N_3^+1B, respectively.

6. Simulation of Experiment 1 with N_3^+1A and N_3^+1B

6.1. Simulating Experiment 1

The model was implemented using MATLAB. Each channel differed in scale from the next by a factor of 1.02, or 2%. The spatial resolution of the input image

[3] Our usage of the terms "soft threshold" and "hard threshold" is standard in the field of wavelet theory (see, for example, Donoho & Johnstone, 1994), and has been used within the field of psychophysics (e.g., Langley, 2002; McIlhagga, 2004), but some authors use these terms differently, referring to a "soft threshold" as a smooth function.

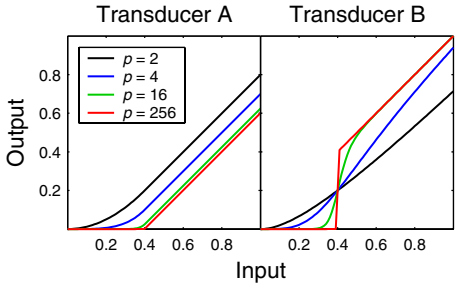

Fig. 5. Transducers A and B for a range of exponents, p. In each case, $T = 0.4$. For low inputs, Transducer B approximates a power function with exponent p; for high inputs, it asymptotes towards a pure half-wave rectifier. For this transducer, T is the input value for which the output is half the input, regardless of the value of p. As $p \to \infty$, Transducers A and B tend towards soft and hard threshold functions, respectively, with threshold, T.

was set so that the smallest operator in the smallest channel had a scale of 1 pixel. Within a channel, the processing proceeded as for N_3^+0, described in Section 2, except that the half-wave rectifier was replaced with Transducer A or B. For each condition, the estimated scale of the fixed edge was found from the position of its peak along the scale dimension. Then a binary search method was used to adjust the scale of the adjustable edge until it matched the fixed edge in estimated scale (i.e. until the peak occurred in the same scale channel for both edges).

To allow the simulations to be completed within an acceptable time period, a 1D version of the model was used. Given that the edges had a 1D profile, the results should be the same as for the 2D version. This was confirmed empirically, by running 1D and 2D versions of N_3^+1A and N_3^+1B on all the fixed edge stimuli. It has been suggested to us that, because 1D and 2D filters differ in their response to noise, we cannot simply generalize to 2D from 1D simulations. While this is true in general, it does not apply here because our simulations were noise-free, and in these circumstances the 1D model makes the same predictions as the 2D version.

6.2. Fitting the parameters

A similar simulation was conducted for Experiment 2 (described in more detail later). All three subjects also participated in further experiments reported in the accompanying paper (May & Georgeson, 2007), and each transducer was fitted to the data from the two papers (five experiments for KAM and MAG, and three for PAA), to give one pair of parameters for each transducer, for each subject.

The best-fitting pair of parameters was defined as the pair that minimized the root mean square (RMS) difference between the predicted and actual scale/contrast of the blur- or contrast-matched adjustable edge across all conditions of all experiments (120 conditions for KAM and MAG; 48 for PAA). For KAM in Experiment 1, the data were averaged across the noise and no-noise conditions, since the noise surround did not substantially affect the results. To make the units of error comparable across different experiments and conditions, all blur and contrast values were first converted to decibels using the formula, $x_{dB} = 20 \log_{10} x$, where x is the edge scale or Michelson contrast, and x_{dB} is the value converted to decibels. A given difference in decibels corresponds to a constant multiplicative factor, whether the units are edge scale or contrast.

The best-fitting parameters were found for each subject by first sampling the parameter space at regular intervals, and then using the best-fitting pair of parameters as the starting point of a simplex minimization procedure (Nelder & Mead, 1965).

6.3. Results and discussion

For each subject and transducer type, Table 2 gives the single pair of parameters that fitted best across all experiments from this paper and the accompanying paper (May & Georgeson, 2007). For the T parameter to be interpretable, it should be noted that, in our simulation, the spatial position was measured in deg visual angle. If the spatial position was to be measured in arcmin, then the gradient signal, and hence the T values, would be divided by 60. The exponent parameter, p, is independent of the spatial units. Table 2 also gives the RMS error of the fit of the same best-fitting model to Experiments 1 and 2 of this paper, and to the whole set of experiments from both papers. The overall RMS error of the fit to Experiment 1 varied between 0.95 and 2.18 dB for N_3^+1A, and between 0.73 and 2.07 dB for N_3^+1B, indicating an acceptable fit. The RMS error of the fit to all the experiments was less than 2 dB for each subject and each transducer type. The difference in goodness-of-fit between the two transducers was small.

The predicted results of Experiment 1, using the best-fitting parameters, are plotted in Fig. 3, along with each subject's data. The model fits most of the data points well, although there are a few isolated conditions for which the data deviate quite substantially from the predictions. These discrepancies occurred mainly for the lowest-contrast edges, for which the blur judgements would have been least reliable, so we attribute these discrepancies to experimental error (the subjects reported that the blur judgements felt much more difficult for the lowest-contrast edges). The size of the experimental error is underestimated by the bootstrap error bars, because bootstrapping does not take into account any non-stationarity in the subject across time (Wichmann & Hill, 2001b, p. 1316).

In general, incorporating a threshold-like transducer (A or B) into the N_3^+ model accurately accounts for the perceived sharpening of low-contrast edges, as well as the slight increase in the size of this effect with increasing edge scale. The increase in the effect size with increasing edge scale occurs because enlarging the edge scale reduces the gradient magnitude, so that the gradient signal is more severely truncated by the threshold-like behaviour of the transducers.

Table 2
Best-fitting parameters and RMS errors (in dB) for each subject

Subject	Transducer A					Transducer B				
	T	p	RMS error (dB)			T	p	RMS error (dB)		
			Ex 1	Ex 2	All			Ex1	Ex2	All
KAM	0.210	2.83	2.18	2.21	1.98	0.317	3.16	2.07	2.05	1.95
MAG	0.120	5.33	1.32	2.10	1.85	0.329	3.02	1.50	1.51	1.82
PAA	0.410	1.75	0.95	2.41	1.59	0.434	2.15	0.73	2.43	1.53

For each subject and transducer, the parameters were fitted to the complete set of data from all the experiments in this paper and the accompanying paper (May & Georgeson, 2007), to give a single pair of parameters across all experiments for each subject and transducer type. K.A.M. and M.A.G. participated in five experiments in total; PAA participated in three. The "All" RMS error column gives the RMS error across all experiments to which the model was fitted; the "Ex 1" and "Ex 2" columns give the RMS error for Experiments 1 and 2 of this paper, respectively.

Georgeson (1994) had previously reported the results of a similar experiment to Experiment 1, using different edge scales and contrasts. Georgeson's (1994) data were the means of two subjects, MAG and TCAF. Fig. 6 shows the data from MAG alone. We simulated this experiment in the same way as Experiment 1, using MAG's parameters, given in Table 2. Despite the fact that the model had not been fitted to these data, it predicted the results fairly accurately, with RMS error values of 1.37 dB for N_3^+1A, and 1.53 dB for N_3^+1B. The only previous attempt to model the effect of contrast on perceived blur was by Purushothaman, Lacassagne, Bedell, and Ogmen (2002), using a neural network model of retino-cortical dynamics. In contrast to our data, their model predicted that perceived blur would be an inverted U-shaped function of contrast, and that the size of the effect of contrast on perceived blur would *decrease* with increasing edge scale, the opposite of our empirical findings.

7. Experiment 2: The effect of edge scale on perceived contrast

7.1. Introduction

The threshold-like transducers, introduced to explain the results of Experiment 1, cause the N_3^+ model to predict that perceived contrast should decrease with increasing edge scale. This can be understood as follows. Eq. (2) shows that the more the scale is underestimated, the more the contrast is underestimated. Increasing the edge scale reduces its gradient, so the threshold has a greater effect, the scale is underestimated by a greater amount, and therefore the contrast is also underestimated by a greater amount. In addition, the size of this effect should be smaller for high-contrast edges, since these will be largely unaffected by any threshold-like effect on the gradient. Experiment 2 was designed to test these predictions.

7.2. Psychophysical methods and stimuli

In Experiment 2, fixed edge contrasts were 0.05, 0.1, 0.2, and 0.4, and scales were 2, 4, 8, and 16 arcmin. Each of the 16 conditions was assigned to a different staircase, which controlled the contrast of an adjustable edge that had a scale of 2 arcmin. All 16 conditions were tested within one session, and each subject completed eight similar sessions. The starting contrast for 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.

Edge stimuli were constructed according to Eq. (3). For all stimuli, the window function, w, was a sharp-edged square with width and height 320 arcmin (Fig. 2c). Viewing

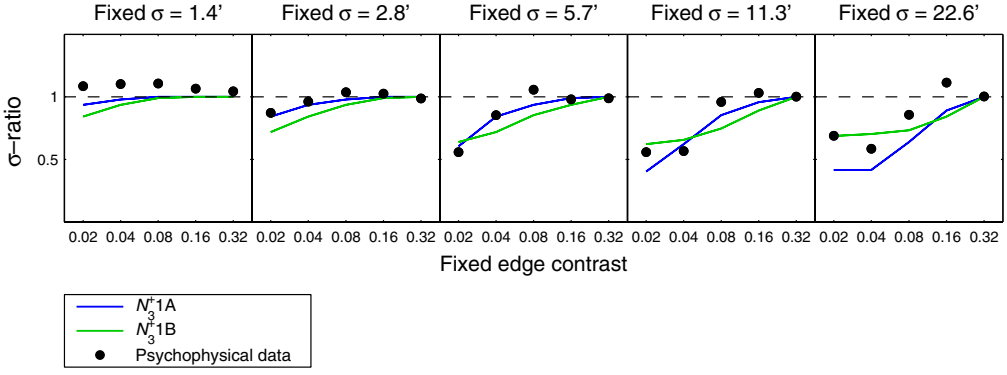

Fig. 6. Data from Georgeson's (1994) experiment that investigated the effect of contrast on perceived blur. The blue and green lines show the predictions of the N_3^+1A and N_3^+1B models, using the parameters (given in Table 2) that fitted best to MAG's data from Experiments 1 and 2 of this paper and Experiments 1, 2 and 3 of the accompanying paper (May & Georgeson, 2007). The model was not fitted to the data shown in this figure. The numerical values plotted in this figure are given in the supplementary data file, "Fig6_data.txt". (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distance was 152.3 cm, which gave a screen resolution of 1′ per pixel. For each fixed edge, we measured the contrast of an adjustable edge (with scale 2′) that matched the fixed edge in perceived contrast, using the method described in Section 3.4. Subjects were asked to judge the perceived contrast of the central edge in each image, and to ignore the rest of the stimulus. Unlike in Experiment 1, each stimulus was accompanied by a 440 Hz tone so that, even at very low contrast, the subject would always know when it was being presented.

7.3. Simulation method

Experiment 2 was simulated in a very similar way to Experiment 1, except that the binary search method adjusted the contrast, rather than the scale, of the adjustable edge. The binary search continued until the difference between the adjustable edge contrast and the true matching value was within 1% of the true value. The details of this method are given in Appendix C.

7.4. Results and discussion

Fig. 7 shows the results of the experiment, along with the predicted results using the best-fitting parameters from Table 2. As predicted, the perceived contrast decreased with increasing edge scale, and this effect was smaller for higher-contrast edges. For all subjects, N_3^+1A and N_3^+1B provided a satisfactory fit to the data, except for one very wayward point in PAA's data (fixed edge contrast 0.2 and scale 16′). This data point violates the pattern shown by the other subjects, and shown by PAA for the other fixed edge contrasts. The error bars indicate that this data point was very unreliable: despite the large deviation, the predicted value still fell within the confidence limits.

8. General discussion

In this paper, we have shown that blurred edges look faint (i.e. low in contrast), and faint edges look sharp. The effect of contrast on perceived blur increased with increasing edge scale; the effect of edge scale on perceived contrast increased with decreasing edge contrast. Neither of these effects are predicted by Georgeson et al.'s (in press) N_3^+0 model of edge coding, which has successfully predicted the results of many previous blur-matching experiments. We have shown that the model can account for these data if it is modified slightly by replacing the half-wave rectifier in each channel with a transducer that has the shape of a smoothed threshold function. Two different 2-parameter transducers were used, one that is a smoothed soft threshold function, and one that is a smoothed hard threshold function. The differences in the fits of these models to the data were small. For each subject, a single pair of parameters was sufficient to explain all of that subject's blur-matching and contrast-matching data presented in this paper and the accompanying paper (May & Georgeson, 2007), as well as the data from a blur-matching experiment reported by Georgeson (1994).

Threshold models have had a long history in machine vision, psychophysics and physiology. Many edge detection algorithms for machine vision use gradient thresholds to remove noise (Rakesh, Chaudhuri, & Murthy, 2004; Rosin, 1997). A threshold on the gradient is also a critical element of many computational models of lightness perception (Hurlbert, 1986). Psychophysical studies have suggested that the visual system applies a threshold when estimating stimulus contrast: perceived contrast is approximately a function of the physical stimulus contrast minus the detection threshold (Cannon, 1979; Georgeson, 1991; Kulikowski, 1976; McIlhagga, 2004). Neurophysiologists have often modelled simple cells in the visual cortex using threshold models, because each neuron has a threshold contrast level below which it does not fire (Albrecht & Hamilton, 1982; Dean, 1981; Ikeda & Wright, 1974; Tolhurst, Movshon, & Thompson, 1981). It has been suggested that the underlying membrane potential of a simple cell shows linear spatial summation, and that the overt response is an essentially linear function of membrane potential, except for half-wave rectification (Jagadeesh, Wheat, & Ferster, 1993; Movshon, Thompson, & Tolhurst, 1978; Schumer & Movshon, 1984). Many apparent non-linearities can be accounted for simply by assuming that the rectifier has a threshold (Andrews & Pollen, 1979; DeAngelis, Ohzawa, & Freeman, 1993; De Valois, Thorell, & Albrecht, 1985; Field & Tolhurst, 1986; Jagadeesh et al., 1993; Movshon et al., 1978; Robson, Tolhurst, Freeman, & Ohzawa, 1988; Schumer & Movshon, 1984; Tadmor & Tolhurst, 1989; Tolhurst & Dean, 1987; Tolhurst & Heeger, 1997).

A problem with the threshold-linear model of simple cells is that it incorrectly predicts that the measured tuning bandwidth of a cell should increase with increasing contrast, as the tails of the tuning curve rise above the threshold: in fact, the orientation and spatial frequency tuning curves of simple cells have been found to be largely contrast-invariant (Sclar & Freeman, 1982; Skottun, Bradley, Sclar, Ohzawa, & Freeman, 1987). This led Heeger (1992a, 1992b) to model a simple cell as a linear filter, followed by a power function transducer, since this arrangement predicts contrast invariance. More recently, Miller and Troyer (2002) proved that a power function is the *only* transducer that will predict contrast invariance. However, Carandini and Ferster (2000) measured instantaneous spike rate as a function of instantaneous membrane potential, and found that the threshold-linear model was well supported. This apparent inconsistency can be resolved by proposing that, although the *instantaneous* spike rate is a threshold-linear function of *instantaneous* membrane potential, the spike rate of threshold-linear units averaged over time can approximate a power function if noise is added to the membrane potential (Anderson, Lampl, Gillespie, & Ferster, 2000). Laming (1986) and Miller and Troyer (2002) showed that adding Gaussian noise to a half-wave rectifier or threshold-linear function resulted in

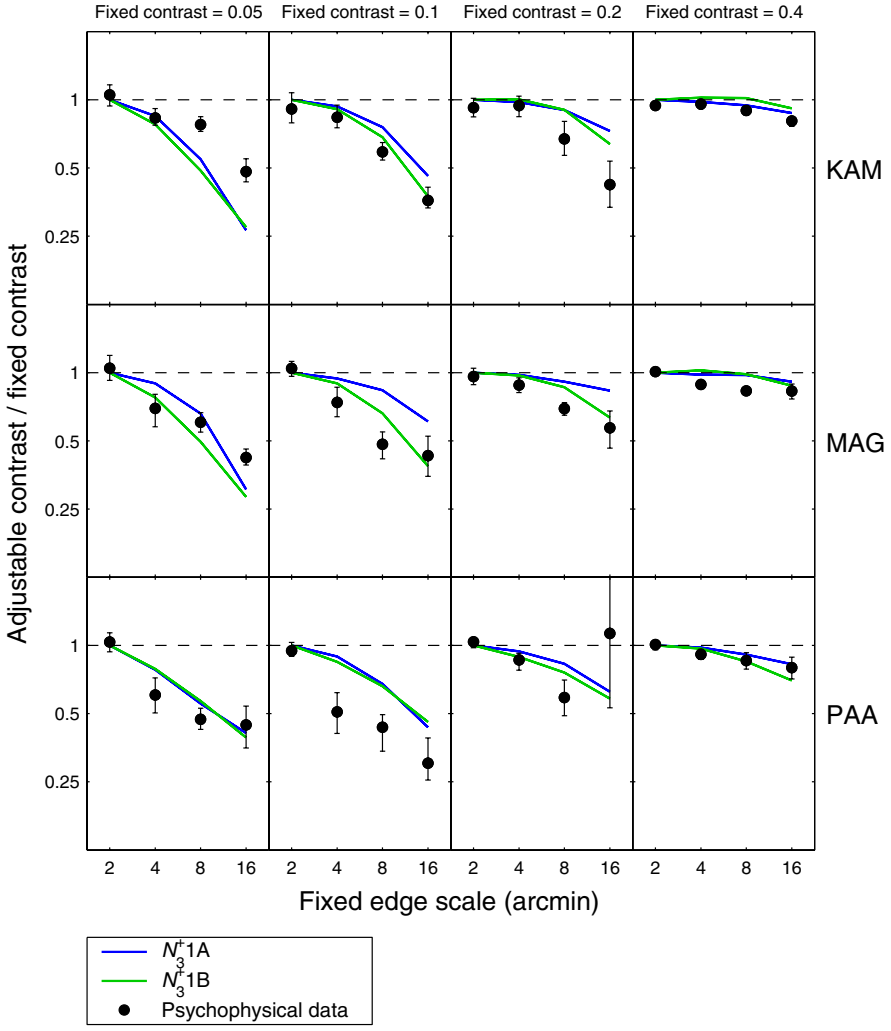

Fig. 7. Results of Experiment 2. Error bars indicate bootstrap confidence limits of 5 and 95%. The vertical axis gives the ratio of adjustable edge contrast to fixed edge contrast when the two edges perceptually matched in contrast. The horizontal dashed lines indicate the results that would have occurred if perceived contrast was independent of edge scale. The blue and green lines show the results predicted by N_3^+1A and N_3^+1B. The parameters of the transducers for each subject are given in Table 2. The numerical values plotted in this figure are given in the supplementary data file, "Fig7_data.txt". (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a transducer with a similar shape to our Transducer A. In fact, Suarez and Koch (1989) showed that, if the noise has a flat distribution between 0 and T, then the transducer is identical to Transducer A, with $p = 2$. Suarez and Koch's scheme can be extended to give Transducer A with any exponent, p, by making the probability distribution of the noise a power function with exponent $(p - 2)$ between 0 and T; alternatively, Transducer A could be constructed by summing across threshold-linear units with a distribution of *thresholds* that had this form.

Another problem with the threshold-linear model of simple cells is that the model does not show the saturation shown by real cells. Albrecht and Hamilton (1982) found that the contrast-response function of simple cells was well described by the Naka–Rushton function (Naka & Rushton, 1966). More complex normalization models of simple cells have been proposed, with a similar general form to the Naka–Rushton function (Albrecht & Geisler, 1991; Heeger, 1992a, 1992b). As well as showing saturation, Heeger's model also explains all the non-linearities previously explained using thresholds (Heeger, 1992a, 1992b; Tolhurst & Heeger, 1997). The Naka–Rushton function is the same as our Transducer B, except that each term has the same exponent, causing the function to saturate. Transducers with a similar form to Transducer B have also been used to explain the results of psychophysical experiments on contrast increment detection (Foley, 1994; Legge & Foley, 1980).

Given that a power function has approximately the same shape as a smoothed threshold-linear function (Carandini, Heeger, & Movshon, 1997), it might be supposed that a good fit to our data could be achieved with a simple power law transducer. In fact, a pure power law gives rise to contrast-invariant blur estimates. The proof of this result is essentially the same as the proof that a cell with a power law transducer will give contrast-invariant

tuning curves (Heeger, 1992b; Miller & Troyer, 2002): a change in contrast affects the *amplitude* of the transducer's output, but has no effect on the *shape* of the response profile across any domain (e.g. orientation, spatial frequency, or, in this case, scale and spatial position). Since the stages of processing that follow the transducer in our model are linear, the profile of the output of the model will also be unchanged by a change in contrast, except for a scaling in amplitude, so the positions of the peaks in edge strength will be unchanged, leading to contrast-invariant blur estimates.

This argument does not apply to the estimation of contrast, which is determined by the height of the peak in edge strength. A pure power law transducer has a similar effect on contrast estimation to a threshold function: perceived contrast reduces with increasing scale, although, unlike with the threshold function, the size of this effect is contrast-invariant.

The difference in the effects of a power law on blur and contrast estimation allows us to explain how the model accommodates one particular individual difference between the subjects. In Experiment 1, the effect of contrast on perceived blur was considerably smaller for PAA than for the other two subjects, whereas, in Experiment 2, all subjects showed a similar size of effect of blur on perceived contrast. The model accounted for this by fitting a larger T parameter to PAA's data. Thus, PAA's transducer had the shape of a power function over a larger range of inputs, and this reduced the effect of contrast on perceived blur, while maintaining a strong effect of blur on perceived contrast. In fact, with Transducer A, when the stimulus gradient falls below T, the transducer is effectively identical to a power law. At this point, the estimated blur will be completely unaffected by a reduction in contrast, and this explains why some of the graphs fitted to PAA's data in Fig. 3 are completely horizontal at low contrasts.

One might ask why the original N_3^+0 model had been so successful at predicting Georgeson et al.'s (in press) data, and whether the new transducers would prevent it from predicting these data so well. There are two reasons why it is likely that the experiments reported by Georgeson et al. were not very sensitive to the effects of the transducers. Firstly, contrast in those experiments was at a moderately high level (around 0.3), where the transducers have less impact. Secondly, most of Georgeson et al.'s experiments involved matching the blur of edges of similar contrast; thus, when the fixed and adjustable edges matched in blur, their gradients would have been quite similar, so the threshold-like transducer that followed the 1st derivative operation would have affected both edges in similar ways, cancelling out the effect. In the experiments reported here, blur matches were performed between edges of very different contrast, and contrast matches were performed between edges of very different scale. Thus, at the point of subjective equality, the fixed and adjustable edges would have had very different gradients, and the transducer would have affected the low-gradient edge much more than the high-gradient edge. Therefore, the experiments reported here were more sensitive to the effects of the transducer.

Our model follows the same strategy as Lindeberg's (1998a) algorithm: both detect edges using a filter whose scale is automatically matched to the scale of the edge. We make no claims about the computational optimality of this class of algorithm with respect to efficiency or reliability in noise. Indeed, there are many successful edge detection algorithms that select the filter scale in very different ways from our model. One approach, for example, is to choose, at each point in the image, the smallest filter that is responding at a statistically reliable level (Elder & Zucker, 1998; Fleck, 1992). Elder and Zucker's algorithm produces impressive results on digital images, coping well with both fine detail and blurred edges due to shadows and shading. For human perception, however, the strength of psychophysical evidence (Georgeson et al., in press; May & Georgeson, 2007) leads us to favour the N_3^+ algorithm as a model of edge coding.

Finally, we should consider alternative explanations of the effects reported in Experiments 1 and 2. It might be possible to explain the results by assuming that the visual system is performing a statistical inference about the world (James Elder, personal communication). Take the finding that low-contrast edges look sharp (Experiment 1). At low contrasts the signal-to-noise ratio is reduced, and the response, $\mathbf{R}$, of the visual system becomes a less reliable indicator of edge scale, σ, so $P(\sigma \mid \mathbf{R}) \approx P(\sigma)$. If the distribution of the prior, $P(\sigma)$, were skewed towards sharp edges (an unknown, but possible scenario, since many edges, such as those due to occlusion, are sharp), then the inferred scale would be small (i.e. sharp) at low contrasts, even for blurred edges. Now consider the finding that blurred edges look lower in contrast (Experiment 2). Many illusions of perceived brightness and contrast can be explained by proposing that the perceived brightness is affected by the inferred reflectance of the surface (e.g., Adelson, 1993). A blurred edge is likely to be interpreted as a shadow or shading, rather than a difference in reflectance between the two sides of the edge stimulus, and this reduced difference in inferred reflectance could reduce the perceived brightness difference between the two sides, making the edge appear lower in contrast.

The explanations of our results based on statistical inference are not necessarily in conflict with the model that we have presented. Instead, our model could be seen as an implementation of the statistical inference process. It might be fruitful to attempt to combine the two approaches in future research. This combined approach might ultimately prove to be the most satisfying, as it could provide a physiologically plausible, quantitative account of the data, with few free parameters, while also giving an insight into the ecological influences that have caused the visual system to develop in this way.

Acknowledgments

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

Appendix A

van Warmerdam and Algazi (1989) and Lindeberg (1998a) both presented identical algorithms that estimated the scale of an edge by locating the peak across scale in the output of a normalized Gaussian 1st derivative operation. Both algorithms allowed flexibility in the exponent of the normalization factor, but Lindeberg showed that it should be $\sigma^{1/2}$ for the peak to occur in operators with a scale matching the edge scale. This appendix proves that, if this algorithm is considered to be the matching of a Gaussian template to the 1st derivative signature of the edge, the normalization factor should indeed be $\sigma^{1/2}$. Let f refer to the filter kernel centred on a particular pixel, and let s refer to the region of the signature covered by the filter kernel. f and s can both be considered to be N-dimensional vectors, where N is the number of pixels in the filter kernel. The result, R, of the filtering process, at the pixel on which the kernel is centred, is given by

$$R = \sum_{i=1}^{N} f_i s_i.$$

R is called the *dot product* of f and s, written $f \cdot s$ (alternative terms are *scalar product* or *inner product*). A standard result is

$$f \cdot s = |f||s| \cos\theta,$$

where θ is the angle between vectors f and s. $|f|$ and $|s|$ are the magnitudes of the vectors, which, according to Pythagoras's theorem, are given by

$$|f| = \sqrt{\sum_{i=1}^{N} f_i^2},$$

$$|s| = \sqrt{\sum_{i=1}^{N} s_i^2}.$$

Thus, we can deduce that

$$\cos\theta = \frac{\sum_{i=1}^{N} f_i s_i}{\sqrt{\left(\sum_{i=1}^{N} f_i^2\right)\left(\sum_{i=1}^{N} s_i^2\right)}}. \tag{4}$$

The term $\cos\theta$ in (4) is called the *normalized correlation coefficient*. It is a measure of the similarity between the vectors f and s. The numerator in (4) is the result of filtering the signature with the Gaussian derivative kernel. The normalization factor is derived from the denominator as follows. We can ignore the sum of squares of s, since this is constant across scale, and will not affect where the peak occurs. The task is simply to find the integral of the squared filter kernel. Thus, for a 1D kernel, the normalization factor, m, is given by

$$\begin{aligned} m &= \frac{1}{\sqrt{\int_{-\infty}^{\infty} f^2(x)\,\mathrm{d}x}} \\ &= \frac{1}{\sqrt{\int_{-\infty}^{\infty} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-x^2}{2\sigma^2}\right)\right]^2 \mathrm{d}x}} \\ &= (8\pi)^{1/4}\sigma^{1/2}. \end{aligned}$$

The constant, $(8\pi)^{1/4}$ can be ignored, since it does not change the shape of the response profile across space or scale, so the normalization factor is $\sigma^{1/2}$. Derivation of the normalization factor using a 2D operator is somewhat more complicated, because what is required is a correlation of the signature with the 1D cross-section of the filter kernel, not with the filter kernel itself. However, the normalization factor is the same as in the 1D case.

Appendix B

This appendix proves that, if Georgeson et al.'s (in press) N_3^+ algorithm is considered to be the matching of a Gaussian derivative template to the 2nd derivative signature of the edge, the normalization factor should be $\sigma^{3/2}$. If f is the 1D Gaussian 1st derivative kernel, then the normalization factor, m can be derived as follows:

$$\begin{aligned} m &= \frac{1}{\sqrt{\int_{-\infty}^{\infty} f^2(x)\,\mathrm{d}x}} \\ &= \frac{1}{\sqrt{\int_{-\infty}^{\infty} \left[\frac{-x}{\sigma^3\sqrt{2\pi}} \exp\left(\frac{-x^2}{2\sigma^2}\right)\right]^2 \mathrm{d}x}} \\ &= 2\pi^{1/4}\sigma^{3/2}. \end{aligned}$$

The constant, $2\pi^{1/4}$, can be ignored, leaving a normalization factor of $\sigma^{3/2}$. As with the normalization factor derived in Appendix A, the derivation is more complex in the 2D case, but the result is the same.

Appendix C

With the binary search method, it is possible to estimate the matching physical contrast, C, to any specified degree of accuracy. After the first reversal in the binary search, there is a known interval (a,b) within which C must lie. The size of this interval halves after each step. After each comparison, the current value of the adjustable contrast will be at one end of the interval (either a or b). The size of the next step is $\Delta x = |b-a|/2$, which will take the adjustable contrast to a new value $x = (a+b)/2$. Thus,

$$\begin{aligned} a &= x - \Delta x \\ b &= x + \Delta x. \end{aligned}$$

Since C lies within this interval, we have

$$x - \Delta x < C < x + \Delta x, \tag{5}$$

and so,

$$|x - C| < \Delta x. \tag{6}$$

If x is the estimate of C at each step, then Δx is the largest possible error in estimation of C. If the search ends when $\Delta x \leqslant \Delta x_{\min}$, then we can be sure that the error in estimation of C is no higher than $\Delta x_{\min}$. Rather than setting an upper limit on the *absolute* size of the error, it is preferable to set an upper limit on the error as a *proportion*, p, of C. Thus we require

$$|x - C| < pC.$$

This can be guaranteed if the search is terminated when the following is true:

$$\Delta x \leqslant p(x - \Delta x). \tag{7}$$

This is proved as follows. From (5) and (7), it follows that

$$\Delta x < pC, \tag{8}$$

and from (6) and (8), we get $|x - C| < pC$, as required. The simulations of Experiments 2 and 3 used $p = 0.01$, and terminated when condition (7) was true. Thus, the estimated contrast of the contrast-matched adjustable stimulus was always within 1% of the correct value.

Appendix D. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.visres.2007.02.012.

References

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science, 262*, 2042–2044.

Albrecht, D. G., & Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience, 7*, 531–546.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48*, 217–237.

Anderson, J. S., Lampl, I., Gillespie, D. C., & Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science, 290*, 1968–1972.

Andrews, B. W., & Pollen, D. A. (1979). Relationship between spatial frequency selectivity and receptive field profile of simple cells. *Journal of Physiology, 287*, 163–176.

Cannon, M. W. (1979). Contrast sensation: A linear function of stimulus contrast. *Vision Research, 19*, 1045–1052.

Carandini, M., & Ferster, D. (2000). Membrane potential and firing rate in cat primary visual cortex. *Journal of Neuroscience, 20*, 470–484.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17*, 8621–8644.

Dean, A. F. (1981). The relationship between response amplitude and contrast for cat striate cortical neurones. *Journal of Physiology, 318*, 413–427.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology, 69*, 1118–1135.

De Valois, R. L., Thorell, L. G., & Albrecht, D. G. (1985). Periodicity of striate-cortex-cell receptive fields. *Journal of the Optical Society of America, A, 2*, 1115–1123.

Dijk, J., van Ginkel, M., van Asselt, R. J., van Vliet, L. J., & Verbeek, P. W. (2003). A new sharpness measure based on Gaussian lines and edges. *Lecture Notes in Computer Science, 2756*, 149–156.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika, 81*, 425–455.

Elder, J. H. (1999). Are edges incomplete? *International Journal of Computer Vision, 34*, 97–122.

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.

Field, D. J., & Tolhurst, D. J. (1986). The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex. *Proceedings of the Royal Society of London B, 228*, 379–400.

Fleck, M. M. (1992). Multiple widths yield reliable finite differences. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 14*, 412–429.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America, A, 11*, 1710–1719.

Georgeson, M. A. (1991). Over the limit: Encoding contrast above threshold in human vision. In J. J. Kulikowski, V. Walsh, & I. J. Murray (Eds.), *Vision and visual dysfunction*, Vol. 5: Limits of visual perception, (pp. 106–119).

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. (2006). Bars & edges: A multi-scale Gaussian derivative model for feature coding in human vision. *Journal of Vision, 6*, 191a.

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*.

Heeger, D. J. (1992a). Half-squaring in responses of cat striate cells. *Visual Neuroscience, 9*, 427–443.

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

Hurlbert, A. (1986). Formal connections between lightness algorithms. *Journal of the Optical Society of America, A, 3*, 1684–1693.

Ikeda, H., & Wright, M. J. (1974). Sensitivity of neurones in visual cortex (area 17) under different levels of anaesthesia. *Experimental Brain Research, 20*, 471–484.

Jagadeesh, B., Wheat, H. S., & Ferster, D. (1993). Linearity of summation of synaptic potentials underlying direction selectivity in simple cells of the cat visual cortex. *Science, 262*, 1901–1904.

Kirsch, R. A., Cahn, L., Ray, C., & Urban, G. H. (1957). Experiments in processing pictorial information with a digital computer. In Proceedings of the Eastern Joint Computer Conference (pp. 221–229). Washington, DC.

Kisworo, M., Venkatesh, S., & West, G. (1994). Modeling edges at subpixel accuracy using the local energy approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 16*, 405–410.

Korn, A. F. (1988). Toward a symbolic representation of intensity changes in images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 10*, 610–625.

Kovasznay, L. S. G., & Joseph, H. M. (1953). Processing of two-dimensional patterns by scanning techniques. *Science, 118*, 475–477.

Kovasznay, L. S. G., & Joseph, H. M. (1955). Image processing. *Proceedings of the IRE, 43*, 560–570.

Kulikowski, J. J. (1976). Effective contrast constancy and linearity of contrast sensation. *Vision Research, 16*, 1419–1431.

Laming, D. (1986). *Sensory analysis*. London: Academic Press.

Langley, K. (2002). A parametric account of contrast adaptation on contrast perception. *Spatial Vision, 16*, 77–93.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70*, 1458–1471.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 467–477.

Lindeberg, T. (1998a). Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision, 30*, 117–154.

Lindeberg, T. (1998b). Feature detection with automatic scale selection. *International Journal of Computer Vision, 30*, 79–116.

Marr, D. (1976). Early processing of visual information. *Philosophical Transactions of the Royal Society of London B, 275*, 483–519.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207*, 187–217.

Marshall, J. A., Burbeck, C. A., Ariely, D., Rolland, J. P., & Martin, K. E. (1996). Occlusion edge blur: A cue to relative visual depth. *Journal of the Optical Society of America, A, 13*, 681–688.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London Series B-Biological Sciences, 263*, 169–172.

Mather, G. (1997). The use of image blur as a depth cue. *Perception, 26*, 1147–1158.

Mather, G., & Smith, D. R. R. (2002). Blur discrimination and its relation to blur-mediated depth perception. *Perception, 31*, 1211–1219.

May, K. A., & Georgeson, M. A. (2007). Added luminance ramp alters perceived edge blur and contrast: A critical test for derivative-based models of edge coding. *Vision Research, 47*, 1721–1731.

McIlhagga, W. (2004). Denoising and contrast constancy. *Vision Research, 44*, 2659–2666.

Miller, K. D., & Troyer, T. W. (2002). Neural noise can explain expansive, power-law nonlinearities in neural response functions. *Journal of Neurophysiology, 87*, 653–659.

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.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. *Journal of Physiology, 283*, 53–77.

Naka, K. I., & Rushton, W. A. H. (1966). S-potentials from colour units in the retina of fish (cyprinidae). *Journal of Physiology, 185*, 536–555.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7*, 308–313.

Purushothaman, G., Lacassagne, D., Bedell, H. E., & Ogmen, H. (2002). Effect of exposure duration, contrast and base blur on coding and discrimination of edges. *Spatial Vision, 15*, 341–376.

Rakesh, R. R., Chaudhuri, P., & Murthy, C. A. (2004). Thresholding in edge detection: A statistical approach. *IEEE Transactions on Image Processing, 13*, 927–936.

Roberts, L. G. (1965). Machine perception of three-dimensional solids. In J. T. Tippett, D. A. Berkowitz, L. C. Clapp, C. J. Koester, & A. Vanderburgh (Eds.), *Optical and electro-optical information processing* (pp. 159–197). Cambridge, MA: MIT Press.

Robson, J. G., Tolhurst, D. J., Freeman, R. D., & Ohzawa, I. (1988). Simple cells in the visual cortex of the cat can be narrowly tuned for spatial frequency. *Visual Neuroscience, 1*, 415–419.

Rosin, P. L. (1997). Edges: Saliency measures and automatic thresholding. *Machine Vision and Applications, 9*, 139–159.

Schumer, R. A., & Movshon, J. A. (1984). Length summation in simple cells of cat striate cortex. *Vision Research, 24*, 565–571.

Sclar, G., & Freeman, R. D. (1982). Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Experimental Brain Research, 46*, 457–461.

Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behaviour. *Journal of Neurophysiology, 57*, 773–786.

Suarez, H., & Koch, C. (1989). Linking linear threshold units with quadratic models of motion perception. *Neural Computation, 1*, 318–320.

Tadmor, Y., & Tolhurst, D. J. (1989). The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning curve in simple cells of the cat's striate cortex. *Visual Neuroscience, 3*, 445–454.

Tolhurst, D. J., & Dean, A. F. (1987). Spatial summation by simple cells in the striate cortex of the cat. *Experimental Brain Research, 66*, 607–620.

Tolhurst, D. J., & Heeger, D. J. (1997). Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex. *Visual Neuroscience, 14*, 293–309.

Tolhurst, D. J., Movshon, J. A., & Thompson, I. D. (1981). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. *Experimental Brain Research, 41*, 414–419.

van Warmerdam, W. L. G., & Algazi, V. R. (1989). Describing 1-D intensity transitions with Gaussian derivatives at the resolutions matching the transition widths. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*, 973–977.

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.

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology, 18*, 1–10.

Wichmann, F. A., & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*, 1293–1313.

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

Witkin, A. P. (1983). Scale-space filtering. In Proceedings of the International Joint Conference on Artificial Intelligence (pp. 1019–1022). Karlsruhe, Germany.

Zhang, W., & Bergholm, F. (1997). Multi-scale blur estimation and edge type classification for scene analysis. *International Journal of Computer Vision, 24*, 219–250.