Do cortical neurons process luminance or contrast to encode surface properties?

Abbreviated title: Surface coding in visual cortex

Tony Vladusich\textsuperscript{1}, Marcel P. Lucassen\textsuperscript{2} and Frans W. Cornelissen\textsuperscript{1}

\textsuperscript{1}Laboratory of Experimental Ophthalmology & NeuroImaging Centre, School of Behavioural and Cognitive Neurosciences, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands

\textsuperscript{2}Department of Human Interfaces, TNO Human Factors, Soesterberg, The Netherlands

Corresponding author: T. Vladusich, Laboratory of Experimental Ophthalmology & NeuroImaging Centre, School of Behavioural and Cognitive Neurosciences, University Medical Centre Groningen, University of Groningen, PO Box 30.001, Groningen 9700 RB, The Netherlands, t.vladusich@med.umcg.nl

Number of figures and tables: 8 Figs

Number of pages: 45

Keywords: surface; luminance; contrast; integration; model

Acknowledgements: We thank Professors Kinoshita and Komatsu for kindly providing the data on which this study was based. The criticisms of four reviewers led to substantial improvements in several aspects of the manuscript.

Grants: This work was supported by grant 051.02.080 of the Cognition program of the Netherlands Organization for Scientific Research (NWO).
Abstract

Studies of human surface perception emphasise the complementary nature of contrast and luminance processing. On the one hand, contrast signals provide information about surface properties, such as reflectance, and patchy illumination conditions, such as shadows. On the other hand, processing of luminance signals may provide information about global light levels, such as the difference between sunny and cloudy days. We devised models of contrast and luminance processing, using principles of logarithmic signal coding and half-wave rectification. We fit each model to individual response profiles obtained from 67 surface-responsive macaque V1 neurons in a centre-surround paradigm similar to those used in human psychophysical studies (Kinoshita & Komatsu, 2001, Journal of Neurophysiology, 86, 2559-2570). The most general forms of the luminance and contrast models explained, on average, 73% and 87% of the response variance over the sample population, respectively. We used a statistical technique, known as Akaike’s information criterion, to quantify goodness-of-fit relative to number of model parameters, giving the relative probability of each model being correct. Luminance models, having fewer parameters than contrast models, performed substantially better in the vast majority of neurons, while contrast models performed similarly well in only a small minority of neurons. These results suggest that the processing of local and mean scene luminance predominates over contrast integration in surface-responsive neurons of the primary visual cortex. The sluggish dynamics of luminance-related cortical activity may provide a neural basis for the recent psychophysical demonstration that luminance

**Introduction**

How does the brain transform light signals registered on the retina into visual surface representations? Classical psychophysical studies of brightness and color constancy imply that surface representations are formed through the long-range spatial integration of visual information (Land, 1959, 1977, 1983, 1986). One version of the well-known retinex model (Land & McCann, 1971), for example, posits that the brain spatially integrates the contrasts, or log luminance ratios, formed at reflectance borders, to discount global illumination conditions. Related computational approaches advocate the filling-in of contrast information within the regions defined by object boundaries (Cohen & Grossberg, 1984; Gerrits & Vendrik, 1970; Grossberg & Todorovic, 1988), a process that is generally assumed to occur in the visual cortex (Grossberg & Mingolla, 1985; Komatsu et al. 2000, 2002; Pessoa et al. 1998). Recent psychophysical studies, however, have shown that the tendency toward brightness constancy can be relatively weak (Masin, 2003), even in complex displays (Arend & Spehar, 1993a,b; Robilotto & Zaidi, 2004). Furthermore, classical experiments with Ganzfeld stimuli indicate that human observers can act like photometers in the absence of contrast information (Barlow & Verrillo, 1976). These psychophysical findings are consistent with the informal observation that humans are easily able to perceive global illumination changes, such as when a cloud passes overhead. It is important to note, however, that the scaling of local luminance to mean scene luminance has been proposed as a mechanism to underlie the
tendency toward brightness constancy (Helson & Himelstein, 1955; Robilotto & Zaidi, 2004). In summary, understanding the roles of contrast and luminance processing may shed light on the nature of the cortical computations underlying surface perception.

How do the concepts of contrast and luminance processing compare with neurophysiological data? The centre-surround properties of receptive fields (RFs) in the mammalian retina and Lateral Geniculate Nucleus (LGN) are generally taken as evidence that early visual processes transmit contrast information, rather than luminance information, to visual cortex. Indeed, the preponderance of contrast-responsive neurons in retina and LGN forms the basis for some general computational approaches to vision (Grossberg & Mingolla, 1985). Evidence from the cat (e.g. Barlow & Levick, 1969; Mante et al. 2005; Rossi & Paradiso, 1999), however, shows that many retinal and LGN neurons encode luminance information, in addition to contrast information. In monkey primary visual cortex (V1), there is strong evidence that some neurons encode the luminance of Ganzfeld stimuli (Kayama et al. 1979; Maguire & Baizer, 1982). Thus, contrary to the textbook view, the neural building blocks for the processing of luminance information are clearly available in visual cortex of monkey and cat.

In visual cortex, the responses of a small proportion of neurons qualitatively mirror aspects of human brightness constancy (MacEvoy & Paradiso, 2001), simultaneous brightness contrast (Kinoshita & Komatsu, 2001; Rossi et al. 1996; Rossi & Paradiso, 1999) and brightness filling-in (Hung et al. 2001; Komatsu et al. 2000; Roe et al. 2005). Some authors have found that few neurons in cat striate cortex (Hung et al. 2001;
MacEvoy & Paradiso, 2001) and monkey V1 (Friedman et al. 2003) respond vigorously to changes in surface luminance that are unaccompanied by concurrent stimulation of the classical RF with contrast stimuli. The apparent sparsity of neurons sensitive to the properties of uniform surfaces has even led some authors to conclude that surface brightness and color are exclusively encoded by *border-responsive* neurons (Friedman et al. 2003; Zhou et al. 2000; see also Blakeslee & McCourt, 1999). Kinoshita and Komatsu (2001) have recently described *surface-responsive* neurons that integrate information over large regions of the visual field and respond vigorously to uniform surfaces in the absence of local-contrast changes, and in some cases even local-luminance changes, in the classical RF (see also Roe et al. 2005; Peng & Van Essen, 2005).

Here we develop a detailed computational-statistical framework to *quantitatively* assess whether cortical neurons process luminance or contrast information. We demonstrate the utility of the framework by analysing the data described in the Kinoshita and Komatsu (2001) study. Our approach combines elements familiar in the areas of computational vision modelling and model selection. First, we assume logarithmic processing of luminance and contrast signals (Barlow & Verrillo, 1976; Land & McCann, 1971). To a first approximation, logarithmic processing captures the highly non-linear signal processing that occurs in the early visual pathway. Second, we implement half-wave rectified (HWR) processing on the inputs and outputs of modelled neurons, which captures the notion that neurons cannot generate negative spike rates (Grossberg & Mingolla, 1985; Heeger, 1993). Third, we apply a statistical technique (Burnham & Anderson, 2002), known as Akaike’s information criterion (AIC), to examine multiple
versions of the contrast and luminance models. The AIC method trades-off the number of free model parameters against fit quality, thereby introducing the concept of parsimony into model analysis. The AIC approach also overcomes some known limitations to conventional statistical methods for analysing model performance, allowing batch analysis of both nested and non-nested models (e.g. contrast versus luminance) without the need to specify null hypotheses or make corrections for multiple comparisons.

**Materials and Methods**

**Kinoshita-Komatsu paradigm and classification scheme**

Simulations were carried out on cells classified in the original publication of Kinoshita and Komatsu (2001). We first briefly describe the methods of this study. For details we refer to the original publication. Neuronal responses were recorded in two separate conditions. In one condition, the luminance of a central square, embedded in a background of constant luminance, was varied in 7 equal log steps ranging from 0.1 cd/m² (candelas per square metre) to 100 cd/m². The size of the centre stimulus was always much larger than the hand-mapped classical RF (which was itself often tuned to oriented stimuli). In a second condition, the luminance of an annulus surrounding the central stimulus was varied through the same range as was used for the centre-change tests. In this case, both the central square and the background surface had constant luminance. Importantly, the luminances of the square and the background were tailored to optimally excite a given neuron. As such, the square and background luminances were generally different, meaning that for most neurons there was always a border present in the image. We note here the similarity between the stimuli used in the Kinoshita-
Komatsu study and those typically used in studies of human brightness perception (Arend & Spehar, 1993a,b; Bindman & Chubb, 2004a,b; Hong & Shevell, 2004a,b; Reid & Shapley, 1988; Rudd & Arrington, 2001; Rudd & Zemach, 2004; Shapley & Reid, 1985; see also Boucard et al. 2005; Huang et al. 2002).

The raw data for each neuron were provided to us by the authors. The sample consisted of 67 single- and multi-unit recordings (hereafter referred to as ‘neurons’) classified into six groups, according to their response profiles. The original dataset included 76 neurons, but nine were excluded because they were not subjected to the same experimental conditions as the other neurons. One class of neurons (Bright Type 1) responded to increasing surface luminance with monotonic increases in firing rates, but were unaffected by changes in the luminance of the annulus surrounding the central surface. A second class of neurons (Bright Type 2) responded to changes in the luminance of the central square and of the annulus in a manner consistent with human brightness contrast. That is, these neurons increased their firing rates in response to increasing central-surface luminance but decreased their firing rates in response to increasing annulus luminance (in which case humans would perceive a darkening of the central surface). The third class (Bright Type 3) responded to increases in both central-surface and annulus luminances with increased firing rates. Kinoshita and Komatsu also described Dark-type neurons, with complementary response profiles to those of the Bright-type neurons.

The number of cells for each classification were: Bright Type 1 cells \( (n = 13) \), Bright Type 2 cells \( (n = 14) \), Bright Type 3 cells \( (n = 22) \), Dark Type 1 cells \( (n = 5) \), Dark Type
2 cells \((n = 11)\), and Dark Type 3 cells \((n = 2)\). These RF classifications were based on
the slopes of the response functions, observed during the later part of the response phase.
For most neurons, the recording phase was between 520 and 1020ms after stimulus onset
(see Discussion).

**Half-wave rectified contrast model**

We first specify the most general form of the contrast model and then derive two simpler
forms that we also test. The contrast model is an adapted version of the retinex model
(Land & McCann, 1971). The model computes edge signals as weighted log luminance
ratios. It has previously been used to fit psychophysical brightness data (Rudd &
Arrington, 2001; Rudd & Zemach, 2004) on the effects of variable annulus luminance
and width in displays similar to those used in the annulus-change condition of the
Kinoshita-Komatsu study. To adapt the Rudd et al. version of the contrast model to the
current context, we implemented half-wave rectification of the input signals, such that
there were four inputs (Fig. 1). Two inputs represented *polarity-specific* signals from the
inner border, while the other two represented polarity-specific signals from the outer
border (Fig. 1a). Each input was weighted by coefficients \(w_j\) that were fit by the method
described below. The neural spike rate \(x\) was derived by adding the contributions of the
four input kernels with a fifth free parameter \(C\) that set the baseline activation around
which the inputs modulated:

\[
x = \left[ w_1 \left[ \log \left( \frac{L_c}{L_{r1}} \right) \right]^+ + w_2 \left[ \log \left( \frac{L_{r1}}{L_c} \right) \right]^+ + w_3 \left[ \log \left( \frac{L_{r2}}{L_c} \right) \right]^+ + w_4 \left[ \log \left( \frac{L_{r2}}{L_{r1}} \right) \right]^+ + C \right]^+ ,
\]

(1)
where $L_c$ is the centre luminance, $L_{r1}$ is the luminance of the spatial region immediately surrounding the centre (the region corresponding to the annulus in the annulus-change condition) and $L_{r2}$ is the luminance of the region immediately surrounding region $r1$ (the region corresponding to the background in the annulus-change condition). In the centre-change condition, $L_{r1} = L_{r2}$, and only the first two terms in Eqn. (1) are relevant since,

$$\log\left(\frac{L_{r1}}{L_{r2}}\right) = \log\left(\frac{L_{r2}}{L_{r1}}\right) = 0.$$  

In the annulus-change condition, $L_{r1}$ does not equal $L_{r2}$ and so only the edge pathways weighted by $w_3$ and $w_4$ are active. The positively-signed brackets $[\ ]^+$ denote half-wave rectification, meaning that spike rate could be either positive or zero, but not negative. Depending on the values of the fitted weights ($w_j$), the response function, $f(x)$, associated with Eqn. (1), can be either monotonic or non-monotonic (where $f(x)$ represents the vector of neural responses to all stimuli in the set). For example, all else being equal, positive values for $w_1$ and $w_2$ give rise to a V-shaped function with respect to variable $L_c$, as Kinoshita and Komatsu (2001) observed in some surface-responsive neurons especially during the early phases of the response (<120ms).

We built two simplified versions of the contrast model. In one simplification, we constrained the parameters multiplying each edge polarity to equal one another. We did this separately for inner and outer edges ($w_1 = -w_2$, $w_3 = -w_4$), thereby deriving the contrast model without half-wave rectification. This constraint prevented the model from generating different slopes for opposite-edge polarities but it also reduced the number of free parameters to three—the same number as the mean-luminance model described
below. In a second simplification, we constrained the contrast model such that only the
inner edge provided the inputs ($w_3 = w_4 = 0$), also giving a three-parameter model.

**Half-wave rectified luminance models**

We built three versions of the luminance model. The most general form was the mean-
luminance model (Fig. 1b) with input half-wave rectification. The other two models were
derived through simplifications of this model. The mean-luminance model is based on the
idea that local luminance signals are scaled by an estimate of mean luminance (e.g.
Robilotto & Zaidi, 2004). We modelled this theoretical idea in terms of the logarithmic
ratio of the local luminance and the mean luminance, weighting each of the inputs
separately by means of a power exponent. We added the input kernels, together with a
third free parameter ($C$), to give

$$x = \left[ \log\left( \frac{L_c}{L_{\text{mean}}} \right)^{w_1} + C \right]^{+}, \tag{2}$$

where $L_c$ and $L_{\text{mean}}$ correspond to the luminance of the centre stimulus and the mean
luminance,

$$L_{\text{mean}} = \frac{L_c \cdot \text{area}(c) + L_{r1} \cdot \text{area}(r1) + L_{r2} \cdot \text{area}(r2)}{\text{area(total)}}. \tag{3}$$

The mathematical operations embodied in Eqn. (2) are equivalent to additive interactions
between the logarithms of the centre- and mean-luminance values, multiplied by the
values of their respective weights, $w_1 \log(L_c) - w_2 \log(L_{\text{mean}})$. Eqn. (2) is quite flexible in
the interactions it can entertain. The sign of the relationship between centre- and mean-
luminance values, for example, can be inverted $\log\left( \frac{L_{\text{mean}}}{L_c} \right)$, with respect to Eqn. (2), if
both weights are fitted as negative values. Indeed, the relationship can also be multiplicative, as in 
\[ \log\left(\frac{L_{c}^{w_1}L_{\text{mean}}^{w_2}}{L_{c}^{w_1}L_{\text{mean}}^{w_2}}\right) \] when \( w_1 \) and \( w_2 \) are positive and negative, respectively, or as in 
\[ \log\left(\frac{1}{L_{c}^{w_1}L_{\text{mean}}^{w_2}}\right) \] when \( w_1 \) is negative and \( w_2 \) is positive.

We constructed two local-luminance models, in which only the luminance of the centre patch was processed, by constraining the weight for the mean-luminance term to equal zero (\( w_2 = 0 \)). The local-luminance models therefore had only two free parameters, and the only difference between the two versions was that we omitted the inner half-wave rectification brackets in one model and kept the brackets in the other. We expect that the local-luminance models would outperform the mean-luminance model in the case of Type 1 neurons, but not Type 2 and 3 neurons, since the classification as Type 1 implies that these neurons encode only local luminance.

**Fitting**

We fit the models to data from each neuron individually by means of a non-linear least squares optimisation procedure. Fitting was done on the median values, calculated over all trials, in each of the 14 stimulus conditions. We did not fit to individual trials (or means) for the following reasons: (a) the data for some stimulus conditions were highly skewed, thereby rendering the assumption of independent, gaussian-distributed residuals (that underlies regression) implausible, and (b) the variance associated with each condition varied considerably within a neuron, violating the assumption of uniformity of variance across conditions. Although it may be possible to deal with non-uniformity of variance (by weighting each condition, assuming the variance-stimulus relationship is
predictable), our choice of regression to the medians was seen as the more conservative option. This choice is perhaps justified by the zero-mean gaussian-distributed residuals we obtained for the vast majority of neurons.

All models usually converged well, often generating high values for the variance explained, $R^2$, defined as

$$R^2 = 100(1 - \frac{SS_{fit}}{SS_{total}}),$$

where $SS_{fit}$ is the sum-of-squares derived from the fit, and $SS_{total}$ is the sum-of-squares derived from a flat line given by the mean of the fitted residuals. In instances where $R^2 < 40\%$, we attempted to obtain better fits by randomly toggling the starting parameter values. We found that, in a few cases where $R^2 = 0\%$, changing the starting parameter led to an excellent fit. However, these cases were rare. We always took pains to obtain convergence solutions that represented global, rather than local, minima.

**Performance analysis using Akaike’s information criterion**

We analysed the performance of each model, that is, the goodness-of-fit relative to the number of parameters, using Akaike’s information criterion (AIC) with sample-size correction ($AIC_c$). While the AIC approach is certainly not new (Burnham & Anderson, 2002), only in relatively recent times has the method begun to be applied in certain scientific contexts, such as neuroscience (Averbeck & Lee, 2003; Elder & Sachs, 2004; Schall et al. 2004) and phylogenetics (Posada & Buckley, 2004). The core idea of the approach is to estimate the ‘loss of information’ that occurs when one attempts to construct a model of reality. The measure of information loss consists of a mathematical
term estimating the goodness-of-fit to a data set (e.g. sum-of-squares) and a term estimating the effect of the number of estimated parameters (i.e. complexity). In this sense, AIC embodies a statistical principle of parsimony. Formally, we have

$$AIC_c = N \ln\left(\frac{SS_{fit}}{N}\right) + 2K + \frac{2K(K + 1)}{N - K - 1},$$

where $N$ is the number of data points, $SS_{fit}$ is the fitted sum-of-squares, and $K$ is the number of fitted model parameters. Generally speaking, the smaller the value of $AIC_c$ the better the model has performed. By comparing $AIC_c$ values for $i$th model to a comparison model ($C$), as in

$$\Delta AIC^i_c = AIC^i_c - AIC^C_c,$$

and computing the ratio of these differences relative to the sum of all the models ($r$) in the set ($R = \text{number of models}$), one obtains the relative probabilities of each model being correct, also known as Akaike’s weights

$$p_i = \frac{e^{-0.5(\Delta AIC^i_c)}}{\sum_{i=1}^{R} e^{-0.5(\Delta AIC^i_c)}}.$$

For comparison, we also compute a second criterion, known as the Bayesian Information Criterion (BIC),

$$BIC = N \ln\left(\frac{SS_{fit}}{N}\right) + K\delta\ln(N).$$

where $\delta = 2.5$ represents a factor correcting for small sample size (Ball, 2001). An analogous computation underlies the calculation of relative probabilities associated with the BIC method. A detailed discussion and comparison of the AIC and BIC approaches is provided in Burnham and Anderson (2002).
The general AIC approach has several advantages over tests conventionally used to compare models with different numbers of parameters (Burnham & Anderson, 2002). Of particular interest here, the AIC method is valid even when models are not ‘nested’—that is, when they cannot be derived directly from one another, as is the case with the contrast and luminance models examined here. Additionally, the AIC approach does not depend on arbitrary critical ($\alpha$) values for accepting or rejecting hypotheses and does not require adjustments for multiple comparisons of the sort familiar in conventional statistical inference (e.g. Bonferroni correction). (General discussions of the problems associated with conventional methods of statistical inference can be found elsewhere (Goodman, 1999a,b; Sterne & Smith, 2001)). The AIC approach also allows one to compute evidence ratios with selected models (ratios of relative probabilities of each model being correct) or to add together the relative probabilities associated with specific models to examine the importance of parameters common to the selected models. We make particular use of the additive property in our analysis to compare the joint relative probabilities associated with the local and mean-luminance models.

In common with conventional methods, the AIC approach depends on the assumption that model residuals are gaussian-distributed with zero mean. We tested this assumption for all model fits using the D’Agostino-Pearson test for skewness and kurtosis, and the Students t-test for differences of the mean residuals from zero, respectively. In only a few isolated cases did the p-values associated with either test fall below 0.05. These instances are discussed in the text, without being dismissed outright (since p-values of less than 0.05 hold no privileged status).
Simulation methods

All simulations were performed on an Apple Mac G5 dual 2.0Gh machine using software implemented in Matlab (Version 7.0.4, The Mathworks Inc.). For the mean-luminance model, we assumed that the stimulus was a 129x129 lattice of luminances (in candelas per square metre, or cd/m²), with a central square of 41x41 pixels and an annulus of 101x101 pixels on which the square was superimposed (the contrast model, being agnostic to stimulus area, did not require this assumption). The dimensions of these stimuli conformed to the average stimulus dimensions used in the Kinoshita-Komatsu study.

Results

Examples of model fits

We now analyse the data of Kinoshita and Komatsu (2001) using the modelling framework detailed above. In this section, we limit ourselves to showing example fits obtained with the most general forms of the contrast and luminance models, since these provided the best fits. Fig. 2a shows the responses of a neuron classified as Bright Type 3, along with the best-fitting function, \( f(x) \), with 95% confidence intervals (CIs), of the contrast model. The left panel corresponds to the conditions in which the central patch changed luminance but the background luminance (dashed vertical line) remained constant (centre-change condition). The error bars for each data point represent the first and third quartiles associated with the median for that point. The fitted spike rates are a bit below the data points when the square is darker than the background (dashed
horizontal line represents isoluminance), meaning that the model predicts slightly less ongoing activity than observed. For both model and data, spike rate increases linearly (in log space) as the square’s luminance increases above the background luminance. The difference in slopes of the two components of the model response function arises due to different weightings of the opposite contrast polarities. That is, the weight associated with the decremental inner edge (square darker than the background) is small ($w_2 = 1.32 \pm 5.7$, the later value being the 95% confidence interval on the parameter estimate), while the weight associated with the inner incremental edge is large ($w_1 = 24.26 \pm 4.01$; square brighter than the background).

The story is more complicated (Fig. 2b) when the luminance of the square is kept constant (dotted vertical line) and the luminance of the surrounding annulus changes (annulus-change condition). In this condition, the annulus is itself surrounded by a background surface of constant luminance (dashed vertical line), meaning that there is an inner and an outer border in the image. At the lowest annulus luminance values, the inner incremental weight is active because the annulus is much darker than the centre. At the same time, the weight associated with the outer decremental edge is large and negative ($w_4 = -38.23 \pm 4.23$), which, combined with the large contrast ratio between annulus and the background, generates strong suppression. As this contrast decreases, the suppression declines and there is a steep increase in spike rate up until the point where the annulus and background are isoluminant. When the annulus becomes brighter than the background, there is a sudden dip in the model response function. This occurs because the contrasts associated with the inner and outer borders are both very small. Thus, even
though the inner and outer incremental weights ($w_3 = 16.05 \pm 3.64$) are positive, the small contrast ratios lead to a dip. As the annulus luminance increases above the centre luminance, the inner edge becomes a weakly-weighted decrement, which, when combined with the strongly-weighted outer incremental edge, results in a secondary peak in the response function at the highest annulus luminance.

As expected of a model with more degrees of freedom, the fit for the contrast model ($R^2 = 93.72$) was better than for the mean-luminance model ($R^2 = 65.13$). As we demonstrate in the following section, this fit is sufficiently better to justify the two extra parameters in the contrast model. In other words, the data strongly support the contrast model over the mean-luminance model. The fit associated with the mean-luminance model is shown in Fig. 2c,d. The shape of the response function in the centre-change condition arises as follows. Both the weighted centre luminance ($w_1 = 7.98 \pm 6.51$) and weighted mean luminance ($w_2 = 12.08 \pm 9.80$) are small at the lowest centre-luminance values. Since both weights are positive, the interaction between the two input signals is multiplicative, resulting in an increase in spike rate with centre luminance. In the annulus-change condition, the centre luminance remains constant, meaning that only the weighted mean-luminance signal varies. Since only one input source increases, the response function increases more slowly than in the centre-change condition.

A second example, a Dark Type 2 neuron, is shown in Fig. 3. The fit associated with the contrast model is only slightly better ($R^2 = 93.91$) than that associated with the mean-luminance model ($R^2 = 92.47$). In this instance, our performance analysis (see below)
indicates that the mean-luminance is more likely to be correct, since it achieves a comparably good fit with fewer parameters. In the present example, the fitted response functions are decreasing in the centre-change condition and increasing in the annulus-change condition. The explanations of the functions are similar to those given above. With respect to the contrast model, the main difference is that the outer edge in the annulus-change condition has only a relatively weak effect. In the case of the mean-luminance model, the weighted centre luminance is negative while the weighted mean luminance is positive.

**Comparison of model fits**

The variance explained ($R^2$) values for all fitted neurons are shown in separate histograms for each model in Fig. 4. Each panel also shows the breakdown of fits across the three classified types of neurons in Kinoshita and Komatsu (2001). The contrast model (Fig. 4a) generated far better fits than the mean-luminance model (Fig. 4b). The median fits over the entire population were 73% and 87% for the mean-luminance and contrast models, respectively. The contrast model generated the better fit in 60/67 neurons (90%), with a mean improvement of 15.8% associated with these 60 neurons. In 64% of neurons, fits associated with the contrast model exceeded 80%, while the mean-luminance model performed similarly in only 37% of neurons. Interestingly, the quality of the fits do not appear to differ substantially with the Kinoshita-Komatsu classification of RF type for either model. We now turn to the question of whether the extra variance explained by the contrast model justifies the need for two extra free parameters.
Comparison of model performance

We analysed the performance of the general contrast and mean-luminance models using Akaike’s corrected information criterion (see Materials and Methods), a technique which trades-off fit quality against number of parameters to give the probability of a given model being correct. This analysis was carried out for the entire population of 67 surface-responsive neurons. Fig. 5a shows a histogram of the proportions of neurons versus the relative probabilities of the mean-luminance model being correct (the corresponding graph for the contrast model is simply Fig. 5a with left-right reversal). Relative probabilities are represented in percentages (partly to underline that our relative probabilities should not be interpreted in the classical statistical sense). The bins are 10% units wide. As a guide, we shall interpret frequencies in the 90-100% bin as providing relatively strong evidence in favour of the mean-luminance model, with values in the 0-10% bin providing relatively strong evidence in favour of the contrast model. We should like to emphasize, however, that such percentiles do not represent arbitrary statistical criteria for accepting or rejecting either model. The figure clearly shows that the mean-luminance model is favoured in around 50% of neurons (34/67 neurons in the 90-100% bin, of which 18/67 fall in a 99-100% bin), with the contrast model performing well in only about 10% of neurons (6/67). In the remaining cases, there is no strong evidence in favour of either model. Similar results obtain using the BIC method (Fig. 5b). The better performance of the mean-luminance model can be traced to the fact that it has two fewer free parameters than the contrast model and so is penalised less in the calculation of relative probabilities in both the AIC and BIC analyses. For simplicity, we restrict the remainder of our analysis to the AIC method.
Since the validity of our performance analysis depends on the assumption of gaussian-distributed residuals with zero mean, we checked that all neurons met these criteria (see Materials and Methods). We found that in no case did the means of the residuals differ from zero. In six instances, the residuals associated with the mean-luminance model were found to be significantly non-gaussian at $\alpha < 0.05$. Only one of these cases corresponded to a neuron deemed highly likely to be correct (90-100% bin). Likewise, there were seven cases of non-gaussian residuals associated with the contrast model and only one such instance related to a neuron deemed highly likely to be correct (0-10% bin). Setting the statistical threshold at $\alpha < 0.01$, we found that, for both models simultaneously, only four neurons did not fulfil the gaussian requirement. Thus, the results of our analysis are not greatly affected by mild violations of the assumptions underlying the analysis.

**Simplifying the contrast model**

Due to the flexibility of AIC approach, we were able to simplify both the contrast and mean-luminance models and to make multiple comparisons between various models without the constraints associated with classical statistical analyses. (Note that classical F-test comparisons between the luminance and contrast models are not, in any case, valid, since these models are not nested.) We implemented two simplified versions of the general contrast model (see Materials and Methods). We then fit these constrained models to the entire data set and calculated the combined relative probabilities of these models being correct relative to the mean-luminance model. We found that the performance of the mean-luminance model far exceeded that of the two constrained contrast models (Fig. 6). The performance of the constrained model without half-wave
rectification (Fig. 6a) was roughly the same as that of the general contrast model (Fig. 5). The performance of the other constrained model, in which only the inner edge contributed to the response, was worse than that of the general contrast model for Type 1 and 3 neurons (Fig. 6b). The constrained contrast model improved slightly in performance for Type 2 neurons, however.

**Simplifying the mean-luminance model**

We also implemented two simplified versions of the mean-luminance model and compared the performance of these models against that of the mean-luminance model. In these simplified models, the weight associated with the mean-luminance input was set to zero. We therefore refer to these simplified models as *local-luminance models*. The only difference between the two local-luminance models was the absence of half-wave rectification of the input signal in one model. The consequence of omitting the half-wave rectification bracket is to allow the input to become negative. We added together the relative probabilities associated with the two local-luminance models (addition of relative probabilities is possible within the AIC framework) and compared these values against the relative probabilities (multiplied by two) associated with the mean-luminance model.

We found that the local-luminance models convincingly outperformed the mean-luminance model in around 30% of Type 1 and Type 2 neurons (Fig. 7). These results are somewhat surprising since one may have expected that all Type 1 neurons (and no Type 2 neurons) would be better explained by the local-luminance models. The mean-luminance model, by comparison, performed very well in about 50% of Type 2 neurons and 20% of Type 3 neurons. We are led to conclude that, for many neurons originally
classified as Type 3, and to a lesser extent, those classified as Type 2, the processing of local luminance alone provides an adequate quantitative description of neuronal response profiles. Taken together with the poor overall performance of the contrast models, the present analysis suggests that the majority of surface-responsive neurons process local luminance (together with mean luminance in many cases).

**Are the fitted parameters functionally interpretable?**

We checked whether the estimated parameter values from our analysis of the mean-luminance model could be interpreted physiologically. Fig. 8 shows the best-fitting parameters and associated 95% confidence intervals. Importantly, across the entire population of neurons \( j \), we found strong correlations between the parameters \( w_1^j \) and \( C^j \) common to the mean-luminance model and the two local-luminance models \( r^2 > 0.89, p = 1, \text{ in all cases} \). Thus, studying the parameter values derived from the mean-luminance model is meaningful for the majority of neurons. In Fig. 8, the different panels represent \( w_1^j, w_2^j \) and \( C^j \), respectively. The color of the background encodes the neuron type. We suggest that the weighting parameters \( w_1^j \) and \( w_2^j \) find a natural physical interpretation in terms of excitatory and inhibitory weightings of the input sources to neurons. The interpretation of \( C^j \), however, seems more difficult. For example, \( C^j \) could represent spontaneous firing rates, in which case we would expect \( C^j \) to be always positive. Fig. 8 indicates that this was not always the case. Another problem with this interpretation is that many neurons are likely to process local-luminance information, meaning that accurate estimates of spontaneous firing rates must come from recordings in complete darkness. The default (baseline) conditions in Kinoshita and
Komatsu (2001), however, generally involved showing animals uniform grey stimuli on the experimental display. Thus, it may be that $C^j$ represents the combination of spontaneous firing rates and the influence of the background luminance stimuli. Another possibility is that $C^j$ represents a factor analogous to the relationship (Carandini et al. 2000) between resting membrane potential (RMP) and spiking threshold (ST). If RMP is above ST, on the one hand, a neuron exhibits positive spontaneous firing rates. If RMP is below ST, on the other hand, a neuron requires additional input excitation to overcome the ST. According to this interpretation, $C^j$ represents the value of ST relative to RMP for each neuron.

Does the luminance approach provide any new insights into the functional properties of surface coding in visual cortex? To answer this question, we plotted the weights $w_1^j$ and $w_2^j$ against each other (Fig. 8d). This enabled us to examine the relative contribution of local- and mean-luminance to the firing rate of each neuron. We found a significant positive linear relationship between $w_1^j$ and $w_2^j$ for bright neurons ($r = 0.67, p < 0.0001$), and a non-significant linear relationship for dark neurons ($r = 0.36, p = 0.15$). The population as a whole exhibited a V-shaped distribution, with bright neurons to the right side of the zero value for the local-luminance weight and dark neurons to the left. Here we focus only on bright neurons. When the local-luminance weight is zero ($w_1^j = 0$), surface-responsive neurons encode the weighted mean log luminance, as in $\log(L_{mean}^{w_2})$. As the value of the local-luminance weight becomes more positive ($w_1^j \rightarrow \infty$), we first pass through a region of parameter space where local- and mean-luminance terms are positive and negative, respectively, as in $\log(L_c^{w_1} L_{mean}^{w_2})$. These neurons would correspond
to Type 3 neurons in the Kinoshita-Komatsu classification scheme. As noted previously, however, strong AIC evidence in favour of the Type 3 classification emerged in only a few neurons. One possible functional interpretation of neurons exhibiting strong evidence for positive and negative local- and mean-luminance weights is that, in reality, these neurons sum luminance signals over the entire visual field, with greater weight assigned to local luminance signals. As $w_1^j$ increases further, the mean-luminance weights first approach zero ($w_2^j \rightarrow 0$), corresponding to the local-luminance coding regime, as in $\log(Lc^w)$, before taking on positive values ($w_2^j \rightarrow \infty$), thereby encoding weighted local luminance relative to weighted mean luminance, as in $\log(\frac{Lc^w}{L_{\text{mean}}^{w_2^j}})$. The latter neurons may partially discount illumination through ratio processing (Type 2 neurons), although the efficiency of the discounting will depend on the precise values of $w_1^j$ and $w_2^j$: discounting is 100% efficient when $w_1^j = w_2^j \neq 0$. To summarize, our analysis suggests that most neurons encode either local or mean luminance, and, in some instances, local luminance relative mean luminance. These functional types would therefore appear to provide useful information about complementary aspects of the stimulus.

**Discussion**

In this study, we have developed models to examine whether the known properties of surface-responsive V1 neurons (Kinoshita & Komatsu, 2001) could be better explained by assuming processing of luminance or contrast signals. Our results indicate that, even though the full contrast model provides the better fits, the mean-luminance model performs better in the majority of neurons. The full contrast model clearly performs better
in only a small minority of neurons. Simplification of the mean-luminance model, but not
the full contrast model, leads to improvements in performance in a sizable proportion of
neurons. Not all neurons are well fit by any of the tested models, however, and in some
instances the data do not strongly support one or another model. In so far as we
acknowledge that all models are necessarily wrong—that is, they all fall short of
describing reality—we view our methods and results as a useful guide for interpreting
experimental data on surface-responsive neurons (Hung et al. 2001, 2002; Peng & Van
Essen, 2005; Roe et al. 2005). Since our results depend on a statistical technique that
appears to be relatively unfamiliar in visual neuroscience, we first briefly discuss the
approach.

**Akaike’s information criterion and model selection**

Determining how well a model describes reality, or selecting between competing models
of reality, is a fundamental problem in science. The approach utilized herein—Akaike’s
information criterion, or AIC—is based on considerations from information theory
(Burnham & Anderson, 2002). Briefly, AIC measures the amount of information lost
when a model is used to approximate reality. Of a candidate set of models, the model that
minimizes the loss of information is deemed most likely to be correct. The loss of
information is quantified in the trade-off between goodness of fit and number of free
parameters (see Materials and Methods). Importantly, AIC is *not* derived under the
assumption that the true model is included in the candidate set of models.

The AIC approach is becoming increasingly popular in the biological sciences, due
largely to its simplicity and flexibility (Elder & Sachs, 2004; Posada & Buckley, 2004).
Multimodel analyses of the type performed herein, for example, are not possible with conventional statistical approaches, such as the F-ratio test (Burnham & Anderson, 2002). The present application represents one of the first uses of AIC in the context of modelling neurophysiological data (Averbeck & Lee, 2003; Schall et al. 2004). Averbeck and Lee (2003), for example, used a simpler version of the AIC approach (one not corrected for sample size) in their study of neural coding in the supplementary motor area of the rhesus monkey.

**A reappraisal of surface coding in V1**

Kinoshita and Komatsu (2001) classified surface-responsive neurons according to the slopes of their response functions in centre- and annulus-change conditions. We adopted a complementary approach in our analysis, asking whether neurons’ responses were better explained by luminance or contrast information, regardless of the respective slopes of the response functions. Our analysis indicates that neurons classified as different based on analysis of response slopes, may, in fact, share hidden similarities. Neurons for which we found strong evidence of contrast processing, for example, were of both the Type 2 and Type 3 varieties. Conversely, we found that only 30% of neurons classified by Kinoshita and Komatsu (2001) as Type 1 were better explained by our local-luminance models, relative to the mean-luminance model (although no Type 1 neurons were better explained by the mean-luminance model than the local-luminance models). We conclude that, in order in order to draw meaningful functional conclusions concerning the information processed by surface-responsive neurons in the Kinoshita-Komatsu experiments, one needs to take into account more than the slopes of the response functions.
**Distinguishing between contrast and luminance responses**

Roe et al. (2005) examined surface-responsive neurons using stimuli well-suited to delineating between contrast and luminance responses (see also Hung et al. 2001). Their stimuli consisted of a bipartite field in which left and right halves modulated in counter phase over time, keeping mean luminance constant. Either the left or right half of the stimulus was placed over the mapped local RF of a neuron. In one condition, the luminance values of the entire left and right hemifields modulated in time (real luminance change). In a second condition, the authors used Cornsweet stimuli to modulate only luminance values near the border—a stimulus that elicits illusory brightness filling-in in humans—while keeping local luminance within the RF, and also mean scene luminance, constant.

The findings of Roe et al. (2005) are in agreement with our conclusion that the vast majority of surface-responsive neurons process luminance information (note that the stimuli of Roe et al. cannot be used to distinguish local- and mean-luminance models, since mean luminance was always constant). These authors found that around 50% of sampled V1 and V2 neurons were significantly modulated by local-luminance changes (i.e. surface responsive). In comparison, no surface-responsive V1 neurons, and only around 10% of surface-responsive V2 neurons, were significantly modulated by the Cornsweet stimulus. Optical imaging revealed significant activation in V2 thin stripes for both real and illusory brightness changes, but no significant signal was obtained in V1 for either stimulus type. Roe et al. concluded that the computations underlying surface brightness are likely to occur in V2 but not V1.
The results of Roe et al. (2005) support the generality of our conclusion that luminance processing predominates over contrast integration in V1 neurons. This point is of particular importance since the stimuli used by Roe et al. differed in key ways from those of Kinoshita and Komatsu (2001). The luminance stimuli of Roe et al. varied sinusoidally in time, and the authors’ analyses were based on the spike rates derived over the entire presentation period. In contrast, the stimuli of Kinoshita and Komatsu remained static over the entire presentation period, and our analyses were based only on the latter part of the response. It remains to be seen whether application of our analysis to V2 neurons would support contrast integration.

**Tuning for luminance stimuli**

Peng and Van Essen (2005) reported that around 10-30% of surface-responsive neurons in macaque V1 and V2 are tuned for luminance stimuli in a manner analogous to the way edge-responsive neurons are tuned to spatial frequency. In the context of the experiments of Kinoshita and Komatsu (2001), luminance tuning corresponds to peak firing rates at luminance values away from the maximum or minimum values used in the experiments. Although we did not test any models that incorporate luminance tuning, we observed informally that instances where both models fit the data poorly were largely due to non-monotonic relationships between firing rate and centre (or annulus) luminance, as would be expected with luminance tuning. The methods developed herein could be naturally extended to quantitatively assess evidence for various forms of luminance tuning in future modelling studies.
Roles of contrast and luminance

Our article began with an overview of the two main stimulus cues in achromatic surface perception: contrast and luminance. We had expected to find strong evidence in favour of contrast integration in many cortical neurons, since this notion is at the core of many models of surface perception (Land & McCann, 1971). Psychophysical evidence points to a critical role for contrast in determining surface brightness (Davey et al. 1999; Hong & Shevell, 2004a, b; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996; Rudd & Arrington, 2001; Rudd & Zemach, 2004, Shapiro et al. 2004). Yet, as indicated in our introductory remarks, contrast processing does not obviate the need for luminance processing. Indeed, theoretical studies support roles for both local luminance (Gilchrist, 1999; Pessoa et al. 1995) and mean luminance (Land & McCann, 1971; Robilotto & Zaidi, 2004) in determining aspects of brightness and lightness perception. More generally, if no luminance information were to reach visual cortex, animals would be unable to estimate overall light level (Barlow & Verrillo, 1976; Masin, 2003). Previous reports of cortical neurons that respond to Ganzfield luminance stimuli (Kayama et al. 1979; Maguire & Baizer, 1982) are also consistent with the notion that both luminance and contrast play important roles in determining surface brightness. The precise nature of the putative contribution of surface-responsive neurons to brightness perception, however, remains unclear, particularly in light of the availability of alternative approaches to brightness that do not involve assumptions of spatial isomorphism (Blakeslee & McCourt, 1999; Friedman et al. 2003; Zhou et al. 2000). Consistent with the present analysis, and with approaches eschewing spatial isomorphism, a recent fMRI study (Cornelissen et al. 2005) has shown that temporal changes in local luminance, but
not brightness changes induced through temporal modulation of a surround field, are correlated with retinotopic activity in early human visual cortex.

**Limitations and additional considerations**

Our analysis is based on responses obtained (for most neurons) in the epoch 520-1020ms following stimulus onset. How then might we reconcile the temporal discrepancy between the latency of neural responses and psychophysical data indicating that human brightness percepts emerge after about 120ms (Davey et al. 1998; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996) or even earlier (McCourt & Foxe, 2004)? One answer is that the temporal discrepancy is not as great as it first appears. It is clear, for example, from Figs. 3 and 8 of Kinoshita and Komatsu (2001), that responses characteristic of surface coding, such as polarity selectivity and the modulatory effects of annuli, actually emerged at around 120-220ms in most neurons. Thus, our analysis of the 520-1020ms epoch may actually generalize to earlier epochs which appear more consistent with classical psychophysical estimates of the time course of brightness perception (Davey et al. 1998; Paradiso & Nakayama, 1991; Rossi & Paradiso, 1996). A second answer is that recent psychophysical evidence indicates that temporal aspects of brightness perception may consist of separate contrast and luminance components with different time courses. Shapiro et al. (2004), for example, provide evidence to indicate that luminance signals primarily determine brightness at temporal frequencies around 1Hz, while contrast signals primarily determine brightness at higher temporal frequencies. The relatively sluggish luminance-based responses studied here would therefore appear to fit well with the slow luminance component of brightness perception. One potential explanation for the sluggish nature of these responses is that local-luminance signals must be extracted
through cortical computations (Neumann, 1996) from the combined luminance/contrast signals encoded by LGN input neurons (e.g. Barlow & Levick, 1969; Mante et al. 2005; Rossi & Paradiso, 1999). As indicated previously, luminance signals may also contribute to the anchoring of lightness percepts (Gilchrist et al. 1999). Since the temporal properties of lightness anchoring are not known, it remains possible that the time course of luminance responses in visual cortex may be consistent with anchoring. None of these speculations, however, shed light on the manner in which luminance and contrast signals might combine to determine brightness and lightness within a framework that does not depend upon spatial isomorphism.

Another potential limitation of the present study concerns the use of multi-unit recordings in our analysis. Forty percent (30/76) of the recordings in the Kinoshita and Komatsu (2001) study were obtained under conditions where it was not possible to isolate single neurons. This may have led to an averaging-out of the response properties of the contributing neurons, leading to uncertainties concerning the classification of recordings into functional types. Recordings classified as Type 1, for instance, may actually have arisen through averaging of individual Type 2 and 3 neuronal responses. This is because pooling spikes from Type 2 and Type 3 neurons, which respond with opposite sign to changes in annulus luminance, would tend to flatten out the response function in the annulus-change condition, thereby making multi-unit recording traces appear more like Type 1 ‘neurons’. The use of multi-unit recordings in our study probably does not affect our main conclusion that the mean-luminance model outperforms the contrast model, since identical results were obtained for all functional classes. It is nonetheless possible
that the comparison between local- and mean-luminance models may have been partially
distorted by the use of multi-unit traces. Table 2 of Kinoshita and Komatsu (2001),
however, shows that only 12 of 30 multi-unit recordings were classified as Type 1,
implying that multi-unit recordings largely agree with traces obtained from single
neurons.

Conclusions

We conclude that no single model of surface coding captures the heterogenous nature of
cortical surface computations. Our theoretical approach, in which log luminance ratio
processing is combined with half-wave rectification, provides a simple and novel
mathematical framework for examining specific variants of surface-coding models. The
challenge for future research will be to further refine our understanding of the myriad
cortical mechanisms underlying surface coding in order to link them to brightness and
lightness perception.
A

General contrast model

B

Mean-luminance model

local luminance

mean scene luminance
**Figure 1.** Schematic illustration of the models. (a) Contrast model: the log luminance ratios computed at image borders are rectified and summed by the surface-responsive neuron. Edges of opposite polarity and the same spatial position are weighted separately. (b) Mean-luminance model: the logarithm of the local luminance is computed relative to the logarithm of the mean scene luminance, by subtraction of the weighted log mean signal from the weighted log local signal. Since the luminance values are transformed in log space, this is equivalent to computing ratios of the local and mean luminance values to the power of their respective weights (see Materials and Methods). The general contrast model has five free parameters while the mean-luminance model has only three parameters. Examples of the stimuli used in the Kinoshita-Komatsu study are shown below the contrast model.
Figure 2. One example of fits generated by contrast and mean-luminance models for a Type 3 Bright neuron. (a) Fits (thick grey line) and 95% confidence intervals (thin grey lines) of contrast model in the centre-change condition. Each data point represents the neuron’s median firing rate over trials, with 1st and 3rd quartiles indicated by error bars. The vertical dashed line represents the (constant) luminance of the background, while the horizontal dashed line represents the actual neuron’s median firing rate when presented with only the background. (b) Contrast fit in the annulus-change condition. The dashed vertical line represents background luminance; the dotted vertical line is the centre luminance. (c) Similar plot for the mean-luminance model in the centre-change condition. (d) Fit of the mean-luminance model in the annulus-change condition. The figure illustrates that, for this neuron, the contrast model fits the data better ($R^2 = 93.72$) than the
mean-luminance model ($R^2 = 65.13$). The icons below illustrate the general form of the stimuli and do not correspond directly to any of the stimulus conditions.
Figure 3. Example of fits generated for a Dark Type 2 neuron. Panels (a) and (b) represent the contrast-model fit, while (c) and (d) represent the fit for the mean-luminance model. The fit is again better for the contrast model but the mean-luminance model actually performs better because it has fewer parameters. See Results for details.
Figure 4. Histograms of the fits associated with the general contrast and mean-luminance models. (a) The contrast model generated excellent fits in most cases. (b) The mean-luminance model generated poorer fits on average.
Figure 5. Comparison of general contrast and mean-luminance models using Akaike’s information criterion (AIC) and Bayesian information criterion (BIC). (a) With the AIC method, the mean-luminance model convincingly outperforms the contrast model in ~50% of neurons (90-100% bin) while the contrast model performs similarly in only ~10% of neurons (bin 0-10). The better performance of the mean-luminance model can be traced back to the fact that it has two fewer parameters. (b) The BIC method produces similar results. We present the remaining results using the AIC method.
Figure 6. Comparisons between two constrained contrast models and the mean-luminance model. (a) Unrectified version of contrast model performs about the same as the general contrast model. (b) The mean-luminance model completely outperforms the inner-edge contrast model (note the different scale).
Figure 7. Comparison between local- and mean-luminance models indicates that the mean-luminance model performs poorly for most Type 1 neurons. While around 50% of Type 2 neurons and 20% of Type 3 neurons are highly consistent with the mean-luminance model, a surprisingly large proportion of these neurons are reasonably well-described by the local-luminance models. These quantitative results suggest that many Type 2 and 3 neurons were incorrectly classified in the Kinoshita-Komatsu study.
Figure 8. All fitted parameters for the mean-luminance model for each neuron. Weights associated with (a) the local-luminance signal, and (b) the mean-luminance signal. (c) Additive constants. Shaded red, green and blue regions correspond to Type 1, 2 and 3 neurons, respectively. The lighter shading represents Bright neurons, while the darker regions indicate Dark neurons. The weighting parameters are readily interpretable in terms of excitatory and inhibitory inputs to neurons, whereas the additive constant is more difficult to interpret. (d) Scatter plot relating local- and mean-luminance weights. Different functional classes are interpreted in terms of different combinations of local- and mean-luminance weights. See text for further discussion.
References


