Available online at www.sciencedirect.com
ScienceDirect
Vision Research 47 (2007) 1631–1644
Vision Research
www.elsevier.com/locate/visres

Explaining brightness illusions using spatial filtering and local response normalization

Alan E. Robinson [a,*], Paul S. Hammon [b], Virginia R. de Sa [a]

[a] Department of Cognitive Science, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0515, USA
[b] Department of Electrical and Computer Engineering, University of California, San Diego, USA

Received 13 October 2006; received in revised form 5 February 2007

Abstract

We introduce two new low-level computational models of brightness perception that account for a wide range of brightness illusions, including many variations on White's Effect [*Perception*, *8*, 1979, 413]. Our models extend Blakeslee and McCourt's ODOG model [*Vision Research*, *39*, 1999, 4361], which combines multiscale oriented difference-of-Gaussian filters and response normalization. We extend the response normalization to be more neurally plausible by constraining normalization to nearby receptive fields (models 1 and 2) and spatial frequencies (model 2), and show that both of these changes increase the effectiveness of the models at predicting brightness illusions.

Keywords: Brightness; White's effect; Contrast; Computational modeling

1. Introduction

One of the properties that the human visual system extracts is the brightness of surfaces in the visual scene. This is likely an important early stage of visual processing that impacts later stages, such as shape from shading, or even object recognition, where an object must be recognized independently of its illumination. It has been long known that the perceived brightness[1] of a surface depends on the brightness of neighboring surfaces. Fig. 1 shows several examples where identical gray patches appear lighter or darker depending on the immediate surround. These illusions show that the computation of brightness in the visual system requires more than just measuring the amount of light reflected from each surface. Rather, the context, or surrounding surfaces, influences the perceived brightness dramatically. The study of brightness perception, therefore, often concentrates on these kinds of illusions as a way to infer the underlying computational mechanism that drives brightness perception, even when there are no perceptual errors.

There are several different theories of brightness perception. These can be partitioned into high-level and low-level theories. High-level theories suggest that the visual scene is parsed into some kind of meaningful interpretation, and that brightness errors arise as a consequence of how the scene is interpreted. For instance, anchoring theory (Gilchrist et al., 1999), uses perceptual grouping to segment the scene into different visual frameworks, and then scales the perceived shade of gray of surfaces within each framework so that the brightest surface in each appears white, or, at the very least, brighter than without anchoring. Another example of the high-level

* Corresponding author.
E-mail address: robinson@cogsci.ucsd.edu (A.E. Robinson).

[1] Note that the terms *brightness* and *lightness* are sometimes used interchangeably in the literature, though many authors make the distinction that brightness is perceived luminance, and lightness is perceived reflectance (e.g., Gilchrist, 2006). Since the models we consider in this paper do not distinguish between perceived luminance and perceived reflectance, we have elected to just use the term brightness for reasons of simplicity.

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.02.017

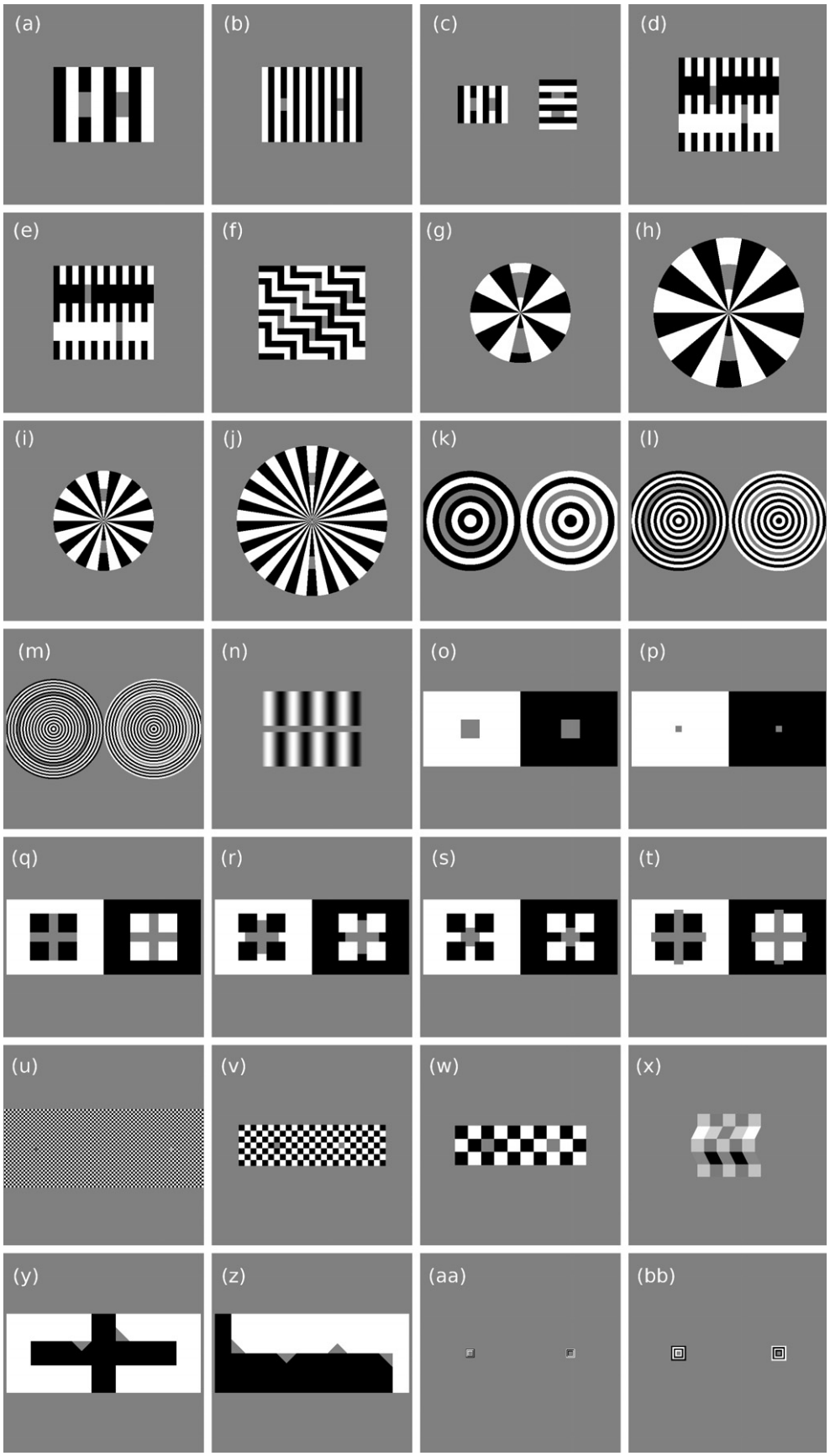


Fig. 1. Illusions tested, all shown to scale, except for (u). Each rectangle is $32 \times 32°$ of visual angle, except for (u) where the scale is twice as large ($16 \times 16°$ of visual angle) and only the central portion of the illusion is shown.

approach is Scission theory (Anderson & Winawer, 2005). According to this theory, surfaces in a scene are split into reflectance, transparency, and illumination layers. The visual system then infers the most probable decomposition into these three layers. Because the decomposition is not always correct, this theory also predicts brightness illusions.

In contrast, low-level theories suggest that brightness errors are caused by interactions of mechanisms in early visual areas which respond to simple features of the image, such as contrast edges, and that no interpretation of the global scene is necessary to cause these errors.

Blakeslee and McCourt (1999, 2001, 2004) and Blakeslee, Pasieka, and McCourt (2005) have introduced and extensively tested a low-level computational model based on filtering by oriented difference-of-Gaussian (ODOG) filters, and then applying global response normalization to equalize the amount of energy at each orientation across the entire visual field. The ODOG model is a compelling starting point because it accounts for many different illusions and uses low-level mechanisms that could be, at least in part, implemented by early visual areas, such as V1. The normalization step in the ODOG model, however, is not particularly neurally plausible, because it is computed globally.

In this work, we extend the ODOG model by exploring whether more neurally plausible normalization schemes will expand the range of illusions predicted by the model. We show that while the normalization step in the ODOG model is necessary to account for a family of illusions known as White's effect (Figs. 1a and b), which are characterized by a highly non-uniform distribution of energy at different orientations, normalization plays relatively little role in predicting non-White's type illusions. Furthermore, we demonstrate that the ODOG model fails on a variation of White's effect that has equal energy at most orientations when integrated across the entire image (Fig. 1c), and also on several previously published variations of White's effect that have a relatively uniform energy distribution across orientation (Figs. 1d–m). These illusions show that the global normalization step in the ODOG model cannot account for all variants of White's effect: instead, more localized normalization schemes may be necessary. A local model of contrast normalization also has the advantage of being more plausible for implementation in early visual areas such as V1.

We introduce two models that add local normalization to the original ODOG model. The first model is locally normalized ODOG (LODOG): instead of normalizing orientation energy across the entire scene, orientation energy is normalized within a local window spanning 4° of visual angle. The second model is frequency-specific locally normalized ODOG (FLODOG): instead of using a fixed window size, normalization is calculated separately for each

Table 1
Sources for illusions tested

Illusion	Fig. 1	Source/original	Test patch size ($w \times h$)	Strength (cd/m^2)
WE-thick	a	Blakeslee and McCourt (1999)/White (1979)	$2° \times 4°$	4.18
WE-thin-wide	b	Blakeslee and McCourt (1999)/White (1979)	$1° \times 2°$	4.6
WE-dual	c	New	$1° \times 2°$	
WE-Anderson	d	Blakeslee et al. (2005)/Anderson (2001)	$1° \times 3°$	6.43
WE-Howe	e	Blakeslee et al. (2005)/Howe (2001)	$1° \times 3°$	0
WE-zigzag	f	Based on Clifford and Spehar (2003)	$1° \times 3°$	
WE-radial-thick-small	g	Based on Anstis (2003)	$2° \times 4°$	
WE-radial-thick	h	Based on Anstis (2003)	$2° \times 4°$	
WE-radial-thin-small	i	Based on Anstis (2003)	$1° \times 2°$	
WE-radial-thin	j	Based on Anstis (2003)	$1° \times 2°$	
WE-circular1	k	Based on Howe (2005)	1° ring width	
WE-circular0.5	l	Based on Howe (2005)	0.5° ring width	
WE-circular0.25	m	Based on Howe (2005)	0.25° ring width	
Grating induction	n	Blakeslee and McCourt (1999)/McCourt (1982)	1° tall	6.23
SBC-large	o	Blakeslee and McCourt (1999)	$3° \times 3°$	11.35
SBC-small	p	Blakeslee and McCourt (1999)	$1° \times 1°$	19.78
Todorovic-equal	q	Blakeslee and McCourt (1999)/Pessoa et al. (1998)	Cross 8° long	2.2
Todorovic-in-large	r	Blakeslee and McCourt (1999)/Todorovic (1997)	Cross 5.3° long	2.4
Todorovic-in-small	s	Blakeslee and McCourt (1999)/Todorovic (1997)	Cross 3° long	4.4
Todorovic-out	t	Blakeslee and McCourt (1999)/Pessoa et al. (1998)	Cross 8.7° long	1.53
Checkerboard-0.16	u	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$0.156° \times 0.156°$	7.46
Checkerboard-0.94	v	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$0.938° \times 0.938°$	2.84
Checkerboard-2.1	w	Blakeslee and McCourt (2004)/DeValois and DeValois (1988)	$2.09° \times 2.09°$	5.67
Corrugated Mondrian	x	Blakeslee and McCourt (2001)/Adelson (1993)	$\sim 2° \times \sim 2°$	10.85
Benary cross	y	Blakeslee and McCourt (2001)/Benary (1924)	Hypotenuse 3°	9.2
Todorovic Benary 1–2	z	Blakeslee and McCourt (2001)/Todorovic (1997)	Hypotenuse 3°	11.95
Todorovic Benary 3–4	z	Blakeslee and McCourt (2001)/Todorovic (1997)	Hypotenuse 3°	9.55
Bullseye-thin	aa	Bindman and Chubb (2004)	Width 0.608°	
Bullseye-thick	bb	Bindman and Chubb (2004)	Width 0.608°	

Note each illusion is listed in the same order as shown in Fig. 1. Note that the illusion strength listed is the psychophysically measured difference between dark and light patches, averaged across subjects, as reported in the source paper.

frequency and orientation, and the window size depends on the spatial scale of the filter response that is being normalized. Furthermore, each frequency channel is normalized primarily by itself, with decreasing influence from nearby frequencies.

2. Illusions tested

We tested the new models on a wide range of illusions, including many that previous literature has tested with the ODOG model, with varying success. We included examples where the brightness of a test patch is shifted toward the brightness of the region that it shares the majority of its border with, and also examples where the brightness of the test patch is shifted away from the region that it shares the majority of its border with. We will refer to these effects as assimilation and contrast, respectively. Note that we use the terms contrast and assimilation to describe the direction of an illusion, not to indicate the underlying mechanistic cause of the illusion. The underlying causes of these illusions are still of debate.

Except where noted, we duplicated the exact dimensions of each illusion as published in the literature cited, and therefore we will only briefly summarize the relevant details for each illusion. Note that since our goal was to study the ODOG model, we elected to use the illusions as implemented in ODOG-related articles. For this reason, we cite the ODOG-related papers that describe the illusions, as well as the original empirical publication for that type of illusion (Table 1).

To facilitate comparisons between the models and people's perception of the illusions, we summarize here the psychophysical results published in papers by Blakeslee and McCourt. These papers all used a matching paradigm, where target patches on a gray background (Blakeslee & McCourt, 1999, 2001) or a checkerboard background (Blakeslee & McCourt, 2004; Blakeslee et al., 2005) are adjusted by the subjects to match the perceived brightness of test patches in the illusions. While the methods and subjects differ a bit between papers, on the whole the methods are much more similar between these papers than the other sources of illusions we used. This higher degree of methodological similarity allows at least tentative comparisons of the strength of illusions that were tested in different papers, although firmer conclusions can be drawn when comparing data points collected within a single study. In particular, the switch to checkerboard backgrounds around the target patch made some illusions appear as much as 50% stronger than when a gray background was used (Blakeslee & McCourt, 2001). Because such small differences in methodology can have large impacts on the psychophysical results, we elected to not include psychophysical results for illusions published by authors other than Blakeslee and McCourt.

Fig. 1 shows the illusions we tested. The first 13 illusions are all variations on White's effect, and each shall hereafter be referred to as WE-*type*. Except where noted all are seen as assimilation of varying strength. The first two illusions (Figs. 1a and b) are the canonical form of White's effect. Two versions are included because higher frequency versions have been shown to increase the strength of White's effect (Blakeslee & McCourt, 1999). The next 11 illusions are versions of White's effect where the amount of energy at each orientation is more evenly distributed than in the traditional White's illusion. WE-dual (Fig. 1c) is a new configuration of White's effect; the illusion on the right side is just a 90° rotation of the left.

Blakeslee et al. (2005) conducted psychophysical measurements of WE-Anderson (Fig. 1d) and WE-Howe (Fig. 1e). WE-Anderson was found to be weaker than a traditional White's illusion that was exactly matched in terms of test patch size and grating dimensions. With WE-Howe subjects saw either weak contrast or assimilation, with no consistent trend across eight subjects except that people who see White's effects as strong assimilation tend to see WE-Howe as weak assimilation. Note that methodological differences between Blakeslee et al. (2005) and Blakeslee and McCourt (1999) are likely the reason why WE-Anderson appears to be a stronger effect than WE-thin-wide when comparing results between the two papers.

WE-zigzag (Fig. 1f) is based on Clifford and Spehar (2003). This illusion is designed to have nearly equal horizontal and vertical orientation energy locally surrounding the test patches. This is in contrast to WE-dual were orientation energy is only equal when summed over the entire image.

WE-radial (Figs. 1g–j) is based on Anstis (2003). We created several new configurations of WE-radial; the 'thick' (Figs. 1g and h) and 'thin' (Figs. 1i and j) versions are designed to have test patches that are similar to WE-thick and WE-thin, respectively. The 'small' (Figs. 1g–i) and 'large' (Figs. 1h–j) versions denote the radius of the circular grating, which is 8° and 12°, respectively.

Howe (2005) studied a circular version of White's effect where the test patches are embedded in a circular grating shaped like a bull's-eye. The illusion remained when the test patches were extended in length so that they covered an entire ring (making the stimulus similar to that tested by Hong & Shevell, 2004), with almost no reduction in illusion strength. We elected to call this a variant of White's illusion, though the test 'patch' is no longer so analogous to those in a traditional White's effect. WE-circular (Figs. 1k–m) are parametric variations of the illusion, based on the version published in Howe (2005). The illusions are named for the width of the test 'patch' (ring). Subjectively, decreasing the width of the ring appears to increase the strength of the illusion.

We also tested a range of illusions that are not clearly related to White's illusion. We included the Todorovic variations on simultaneous brightness contrast (SBC) (Figs. 1q–t). Blakeslee and McCourt (1999) report that the test patch on the right side of the illusion appears lighter for all configurations except Todorovic-equal (Fig. 1q), where

the patch on the left appears lighter. Note that this does not agree with Pessoa, Baratoff, Neumann, and Todorovic (1998), who report that Todorovic-equal appears lighter on the right, so there is some ambiguity as to the proper prediction for this illusion. For consistency we follow Blakeslee and McCourt (1999).

We also tested several versions of the checkerboard illusion (Figs. 1u–w). The illusion flips between assimilation for Checkerboard-0.16 (Fig. 1u), and contrast at larger spatial scales, which is not captured by Fig. 1 because the figures have been reduced significantly in size relative to laboratory viewing. Note also that Fig. 1u is illegible when shown at the same scale as the other illusions, so in the figure we show just the central portion, enlarged by a factor of two.

For the Todorovic reconfiguration of the Benary cross (Fig. 1z) we list two illusions. This is because the image has four test patches, and our analysis depends on having two test patches per illusion. Thus we split the analysis of this illusion in two, summarizing the results for the two patches on the left and on the right separately. To make clear which test patches we are referring to, we number them 1–4, starting from the left.

The illusions selected here are a representative selection of the illusions that the ODOG model has been tested on previously. In general we elected to include one or two configurations of each illusion, rather than an exhaustive sweep of different scales and relative sizes. Our experience with the models, however, suggests that the results we present will generalize to reasonable variations in the configurations of the illusions.

3. The ODOG and UNODOG models

There are two major stages to the ODOG model. A flowchart of its mechanisms is shown in Fig. 2.

First, the input image is filtered by a set of 42 different filters (Figs. 2a and b). Each filter is a zero-sum difference of Gaussians; the center is circularly symmetric and positive, and the surround is negative and elongated in one direction by twice the extent of the center Gaussian. The filters span six orientations, spaced 30° apart, and seven scales (spatial frequencies), with octave spacing between scales. The largest filters have a central frequency of 6.5 cycles per degree. The filter responses are weighted by the spatial frequency of the filter (in cycles per degree) raised to the power 0.1. This function approximates the human contrast sensitivity function over the frequency range of the filters (Blakeslee & McCourt, 1999). Thus, higher frequency filters receive a higher weight.

In the second stage of the model, the 42 filter responses are summed across spatial scales, generating six different multiscale filter responses, one for each orientation (Fig. 2c). These summed filter responses are then normalized individually by dividing by an image-wide energy estimate calculated as the root mean square (RMS) of the pixels in that summed response (Fig. 2d). This makes the global energy for each orientation equal (Fig. 2e). Finally, the six normalized responses are added together, producing a point-by-point prediction of the relative perceived brightness of the input image (Fig. 2f).

The plausibility of the normalization step is somewhat questionable because, for each orientation, a single normalization factor is calculated over the entire input image. If this computation were to occur in V1, it would require lateral connections or feedback connections that are diffuse enough to allow any part of the visual field to influence responses in any other part of the visual field. It is much more likely that these influences are local, rather than global. Furthermore, in any moderately complex natural image the amount of energy at each orientation is relatively uniform, which would mean that the normalization step would only change the filter responses minimally.

For these reasons, we investigated to what extent the normalization step in the ODOG model is necessary by implementing the model without any normalization, which we call UNODOG (un-normalized ODOG). We then ran ODOG and UNODOG on the set of brightness illusions described in Section 2 to see where normalization played an important role.

3.1. Modeling details

Our implementation of the ODOG model contains two changes from the original Blakeslee and McCourt (1999) implementation. First, when filtering the input image we pad around the edges with gray. Whenever filtering an image, there is always the issue of how to treat the edges; we feel that extending the (gray) background to allow for valid filtering is the most plausible approach. When we tried to replicate Blakeslee and McCourt's exact results we found it necessary to use unpadded convolution, which in effect means the edges are extended by tiling the input. This does not seem particularly likely to occur in V1. In practice, our approach generally led to minor differences, with one exception discussed below.

The other difference is how we calculate the strength of the illusion. Blakeslee and McCourt use the average response along a line cutting through the center of the test patch. Instead, we take the average response for all pixels falling inside the test patch. We elected to use this measure because the values within a test patch are often quite non-uniform, and thus the orientation of the line cutting through the test patch can change the predicted illusion strength. Using all the pixels within the test patch is less arbitrary.

3.2. Results—ODOG and UNODOG

Table 2 shows the predicted illusion strength for the UNODOG and ODOG models (the results for the LODOG model will be discussed in Section 4). To derive a single value representing the predicted strength of each illusion we calculate the difference between the predicted value for the test patch that appears darker and the test

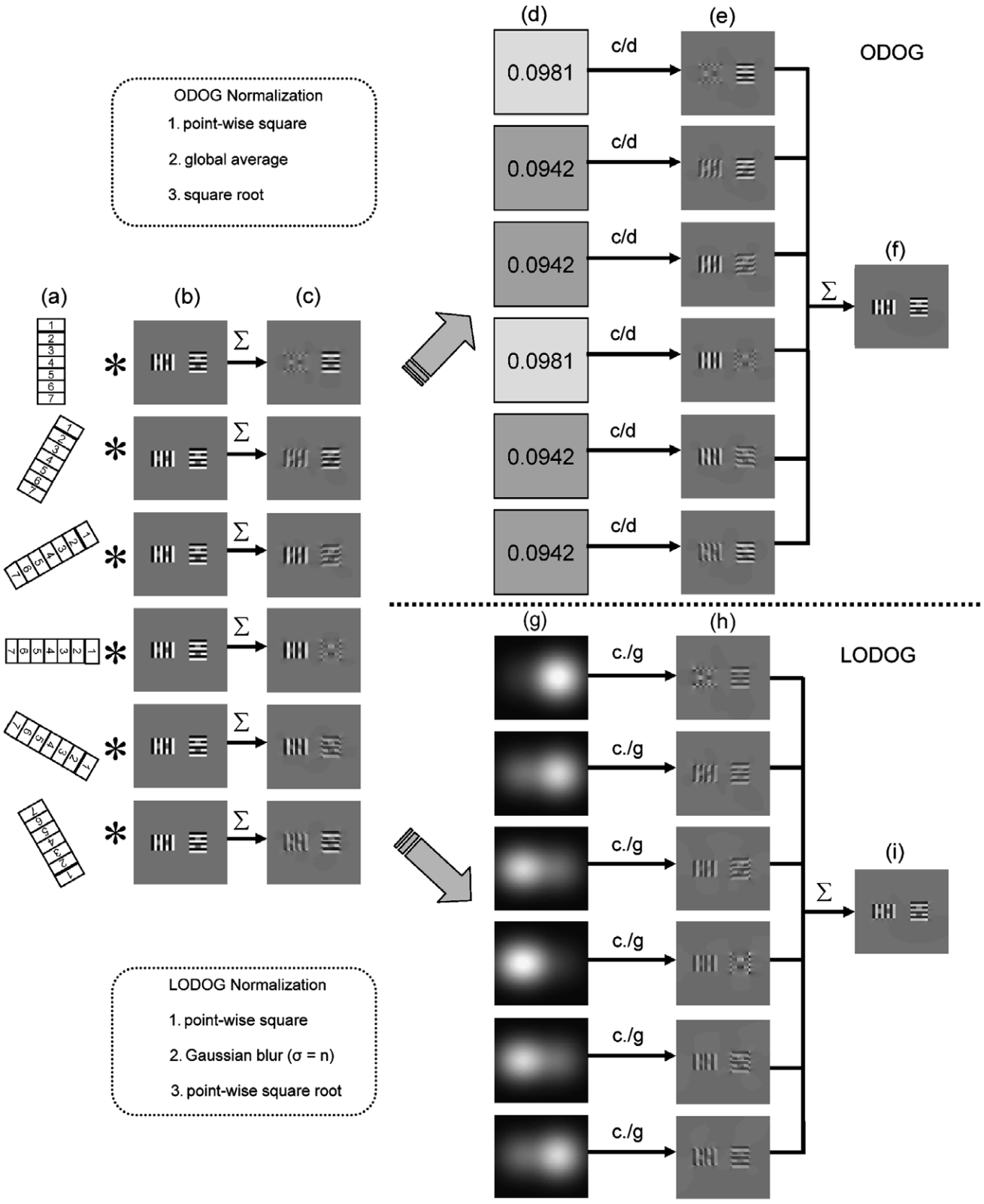


Fig. 2. ODOG and LODOG models. (a) Symbolic representation of the DoG filters at seven different scales and six orientations. (b) The input image. (c) The result of convolving ***a*** and ***b*** and summing the seven scales after weighting the result by a function of spatial frequency. (d) The normalization divisor ODOG calculates for each of the six orientations. (e) The result of applying normalization. (f) Final prediction of ODOG model, produced by summing up ***e***. (g) The point-by-point normalization mask formed by LODOG. (h) The result of dividing each point in ***c*** by each point in ***g*** (./ indicates point-wise division). (i) Final prediction of LODOG model, produced by summing up ***h***.

Table 2
Model results

	ODOG	UNODOG	LODOG $n=1$	LODOG $n=2$	LODOG $n=4$	Illusion Strength (human)
WE-thick	**1.00**	−0.36	**1.00**	**1.00**	**1.00**	1
WE-thin-wide	**2.08**	−0.62	**2.19**	**2.08**	**2.31**	1.1
WE-dual	−0.30	−0.26	**2.53**	**1.36**	**1.11**	
WE-Anderson	−0.15	−1.01	−0.64	−0.30	−0.25	*1.54*
WE-Howe	−0.42	−1.49	−1.99	−0.61	−0.47	*0*
WE-zigzag	−0.51	−0.49	−1.16	−0.76	−0.57	
WE-radial-thick-small	−0.67	−0.67	−0.67	−0.39	−0.55	
WE-radial-thick	−0.41	−0.70	−0.21	**0.01**	−0.29	
WE-radial-thin-small	−0.34	−0.42	**1.43**	**0.21**	−0.20	
WE-radial-thin	−0.22	−0.44	**2.13**	**0.83**	**0.05**	
WE-circular1	−0.82	−1.45	−2.63	−1.04	−1.00	
WE-circular0.5	−0.53	−1.00	−1.47	−0.67	−0.65	
WE-circular0.25	−0.38	−0.71	−1.05	−0.49	−0.48	
Grating induction	**2.03**	**0.17**	**2.32**	**1.69**	**1.77**	1.49
SBC-large	**4.75**	**4.93**	**14.80**	**7.56**	**6.33**	2.72
SBC-small	**6.22**	**6.05**	**26.56**	**14.94**	**9.19**	4.73
Todorovic-equal	−0.36	−0.56	−0.59	−0.26	−0.37	0.53
Todorovic-in-large	**0.49**	**0.77**	**1.63**	**0.55**	**0.52**	0.57
Todorovic-in-small	**0.80**	**1.28**	**2.68**	**0.95**	**0.86**	1.05
Todorovic-out	**0.35**	**0.54**	**1.05**	**0.38**	**0.40**	0.37
Checkerboard-0.16	**1.10**	**0.90**	**2.03**	**0.94**	**0.97**	*1.78*
Checkerboard-0.94	**0.40**	**0.48**	**0.80**	**0.35**	**0.35**	*0.68*
Checkerboard-2.1	**0.69**	**0.72**	**1.62**	**0.60**	**0.59**	*1.36*
Corrugated Mondrian	**0.95**	**0.44**	**2.58**	**0.91**	**0.73**	2.6
Benary cross	**0.09**	−0.12	**0.01**	**0.06**	**0.05**	2.2
Todorovic Benary 1–2	−0.12	−0.41	**0.20**	**0.55**	**0.54**	2.86
Todorovic Benary 3–4	−0.12	−0.41	**0.23**	**0.58**	**0.55**	2.28
Bullseye-thin	−0.74	−0.09	−0.44	−0.35	−0.56	
Bullseye-thick	−0.77	−0.24	−0.52	−0.38	−0.58	

Illusions are listed in the same order as in Fig. 1 and Table 1. Cells in bold indicate the model predicts that the illusion goes in the same direction people typically see it. Note that human values have been scaled so that 1.0 equals the average strength of WE-thick.

patch that appears lighter. We set the sign of the result to indicate whether the prediction matches the direction of the illusion that people see. Negative values indicate that the model predicts the opposite of what people see (i.e., contrast when people see assimilation, or assimilation when people see contrast).

As each model has different normalization steps, the raw numbers that they output are not comparable. To make the results easy to compare between models we scaled the output of each model so that the strength of the WE-thick illusion (Fig. 1a) equals 1, except for the UNODOG model. In this model, WE-thick is predicted in the reverse of what people see, and very weakly. We therefore, instead selected to scale the model's outputs to match the predictions of ODOG on the SBC illusions (Figs. 1o and p). Since the SBC illusion is nearly isotropic, the normalization step in the ODOG model has minimal influence on the strength of the illusion, making these values a good baseline for comparing ODOG to UNODOG.

Since the psychophysics values for these illusions were collected with different methods which impact the strength of the illusions there is no simple way to fairly scale the model output to match the scale of the human responses for all experiments. To enable rough comparison, however, we elected to scale the human data relative to the strength of WE-thick as measured in Blakeslee and McCourt (1999). Keep in mind this decreases how well the models can match to data from the papers after 2001 (shown in italics in Table 2).

The output of the models provides both the predicted direction of the effect (does the test patch get darker or lighter) and also a prediction of the magnitude of the effect. Thus, the models can be judged as to whether they predict the correct direction of the effect, and second, how well the strength of the illusion is predicted. Some care is necessary in using the second metric, as people are highly variable in how strongly they see these brightness illusions. For instance, in Blakeslee and McCourt (2004) data were collected on a White's illusion similar to the dimensions of WE-thin-wide (Fig. 1b). Out of eight subjects, the differences between the two test patches were perceived as small as 2.9 cd/m^2, and as large as 16 cd/m^2. Furthermore, as can be seen in the same paper, the relative strength of different illusions varies somewhat between subjects, even though most of the illusions tested in that paper are White's variants. On the other hand, subjects do tend to see illusions in the same direction, even if the strength varies. Thus, we count the models as being correct if they predict the correct direction of the illusion, although we will discuss the

cases below where the magnitude of the predictions appears to be beyond the range of variability found in human subjects.

By this standard, the ODOG model accounts for the two classic forms of White's effect (Figs. 1a and b). UNODOG, however, does not. It uniformly predicts that people will see contrast instead of assimilation. This shows concretely that the normalization step of ODOG is critical to its prediction of classic White's illusions.

When testing variants of White's illusion with roughly equal global energy at each orientation, we find that ODOG no longer predicts the correct direction of the illusion. These results show that equalizing the orientation energy in the input image makes ODOG fail. Examining the results from UNODOG, we find that it also fails to predict the illusions correctly. Note, however, that the two models do not make identical magnitude predictions, because the different illusion variants have slightly different amounts of energy at different orientations.

The prediction of UNODOG on the grating induction (Fig. 1n) illusion is much smaller than the ODOG prediction, showing that the prediction also depends on unequal energy. Indeed, this reveals that the ODOG model's prediction of grating induction is driven in part by similar mechanisms that make it predict White's effect, and that without normalization ODOG would significantly under-predict grating induction.

Both UNODOG and ODOG predict the correct direction of the SBC illusions (Figs. 1o and p). The models predict, however, that both SBC configurations are about five times stronger than WE-thick (Fig. 1a), when, in fact, the SBC configurations tested here are only slightly stronger than WE-thick (Blakeslee & McCourt, 1999). ODOG clearly over-predicts the strength of contrast, and this is an aspect of the model that needs additional research.

UNODOG and ODOG predict the checkerboard illusions (Figs. 1u and w), with very similar magnitudes. This makes sense since these stimuli are roughly isotropic. Both ODOG and UNODOG account for all of the Todorovic variations of SBC (Figs. 1q–t) except for Todorovic-equal (Fig. 1q), indicating that these predictions do not depend on normalization.

Interestingly, both models predict the corrugated Mondrian (Fig. 1x), though UNODOG predicts it to a smaller extent. This shows that the ODOG account of the Mondrian stimuli depends at least in part on the filters it uses, and not on the complexity of the normalization step. This is a decidedly simple account of an illusion that has been theorized to have high-level origins (Adelson, 1993). Note, however, that both models predict that the Mondrian is weaker than WE-thick, when in fact psychophysical measurements have suggested that it is stronger (comparing psychophysical measurements from Blakeslee & McCourt, 2001 to Blakeslee & McCourt, 1999).

Surprisingly, the Benary cross (Fig. 1y) was predicted correctly by ODOG, but not by UNODOG, revealing that normalization does play a role in predicting this illusion. Neither model predicts the illusion strongly, however, whereas Blakeslee and McCourt (2001) showed that this illusion is not much weaker than the corrugated Mondrian. Interestingly, in contrast to Blakeslee and McCourt (2001), we found that ODOG did not correctly predict the Todorovic version of the Benary cross (Fig. 1z). Upon investigation, we found that this is due to how we padded the input images; if we do not pad the image before filtering (effectively the same as padding the edges with a tiled copy of the illusion) the model makes the correct prediction. Since we find padding with gray to be more plausible, we argue that ODOG does not really account for this illusion. Unsurprisingly, neither does UNODOG.

Finally, ODOG and UNODOG cannot account for the Bullseye illusion (Figs. 1aa and bb), as noted in Bindman and Chubb (2004).

In summary, we found that normalization is key to explaining White's effect, but in general plays a small role in predicting most other illusions considered in this study. For variants of White's effect with more equal global orientation energy, ODOG fails.

4. Local normalization of ODOG

The failures of the UNODOG model show that the normalization step is important for ODOG to account for any of the variants of White's illusion. This normalization, however, is implausible in that normalization of each pixel depends on the energy across the entire scene. We implement a more neurally plausible, local normalization step for ODOG, which we call LODOG (locally normalized ODOG). The mechanisms of LODOG are explained in Fig. 2. The key change is that the summed filter responses are normalized by a local measure of RMS energy instead of a global measure (Fig. 2g). For each pixel, the normalizing RMS is calculated for a Gaussian weighted window centered on that pixel. The window size (n) is specified as the standard deviation of the Gaussian, measured in degrees of visual angle. We tested several different extents to see which sizes of local normalization windows would work as well as global normalization in ODOG.

Local normalization has other advantages as well. Consider the Dual White's illusion (Fig. 1c). While the illusion strength appears to be undiminished, the global energy is now nearly equal for each orientation, so the global normalization step of ODOG will have little effect. Since normalization is key to ODOG predicting White's illusion, this means that ODOG fails to make the correct prediction. Since LODOG uses a local window, each copy of White's illusion in Fig. 1c will be normalized relatively independently.

4.1. Results and discussion

We tested the LODOG model with Gaussian normalization windows of standard deviation $n = 1$, 2, or 4° of

visual angle. The predictions of the model are shown in Table 2. To facilitate comparison across models, the model outputs are scaled so that the illusion strength of WE-thick (Fig. 1a) equals 1.0.

For all the White's family of illusions that ODOG correctly predicts, LODOG also predicts that the illusions go in the same direction, independently of the window size tested. In contrast to ODOG, LODOG also can predict WE-dual in the correct direction, with the smaller window sizes predicting stronger illusions. LODOG does not, however, predict some of the equal energy variants like WE-Anderson (Fig. 1d), WE-Howe (Fig. 1e), or WE-zigzag (Fig. 1f). LODOG's performance is somewhat better on the WE-radial illusions (Figs. 1g–j), where some of the stimulus configurations are correctly predicted by some of the window sizes, but not all. LODOG also did not improve performance on WE-circular (Figs. 1k–m).

To summarize, LODOG does fix the simple case WE-dual, but for more complex equal-energy White's variants, its performance is only mixed, though it never does worse than ODOG.

LODOG also predicts SBC (Figs. 1o and p), but overestimates the illusion's strength, especially for small normalization windows. ODOG also over-predicts the strength of SBC, and the prediction for LODOG with a window of 4° is not much worse than ODOG predictions. Clearly, local normalization does not improve the ability to predict the strength of SBC. LODOG predicts the checkerboard illusion (Figs. 1u–w), with stronger predictions made when the window size is smaller. LODOG predicts the Todorovic variants of SBC (Figs. 1q–t) about as well as ODOG does.

LODOG predicts a smaller effect for the Benary cross (Fig. 1y) than does ODOG. For the Todorovic variation of the Benary cross (Fig. 1z), however, LODOG (with window sizes of 2° or 4°) predicts the illusion in the correct direction, something that ODOG, as we implemented it, does not. Thus, LODOG appears to be somewhat better than ODOG at predicting the Benary cross across variations in configuration, but is not a complete explanation of the effect, since it predicts a fairly small illusion.

Finally, LODOG's predictions on the Bullseye stimuli (Figs. 1aa and bb) are no better than ODOG's.

In addition to the mean illusion strength predicted by the LODOG models, we also examined the point-by-point predictions made by LODOG with differing window sizes. Fig. 3 shows the cross-section of the model's predictions for WE-thick. The cross-sections show that as the size of the normalization window decreases, the predicted uniformity of regions is decreased, relative to ODOG, and that smaller window sizes affect cross-sections by increasing the depth of valleys and the sharpness of peaks, with this effect becoming very extreme for ($n = 1°$). In Blakeslee and McCourt (1999), psychophysical data were collected for the test patches, and it was found that there was a non-uniform gradient of brightness across the test patch, which was similar to what ODOG predicted. In fact, the

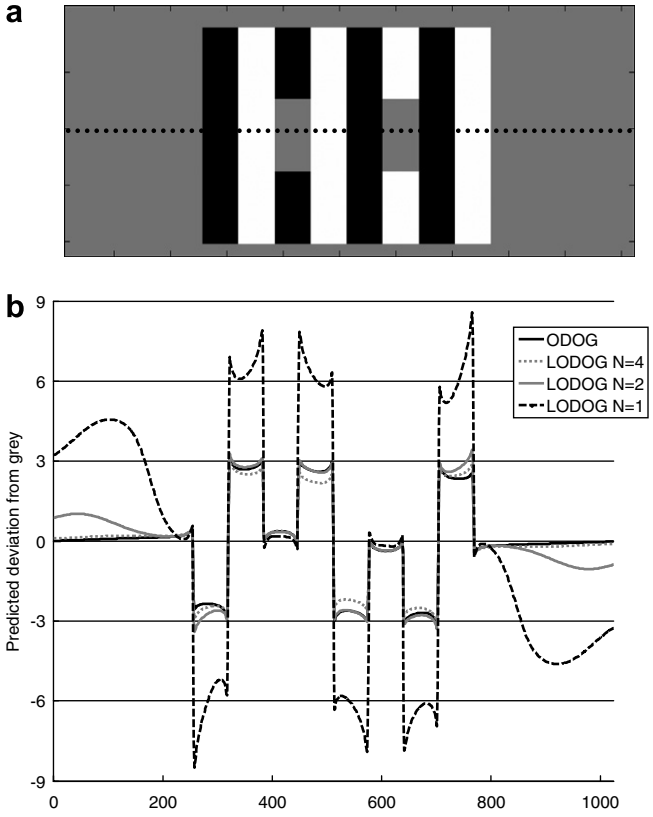


Fig. 3. Model predictions for WE-Thick. (a) Dotted line indicates location of cross-sections. (b) Perceived brightness predicted by ODOG and LODOG models along the cross-section.

psychophysical values suggested that the brightness profile had slightly more curved (i.e. deeper) valleys than ODOG predicted. Thus, the predictions of LODOG with a large window ($n = 4°$) may actually be closer to the psychophysical values than ODOG's.

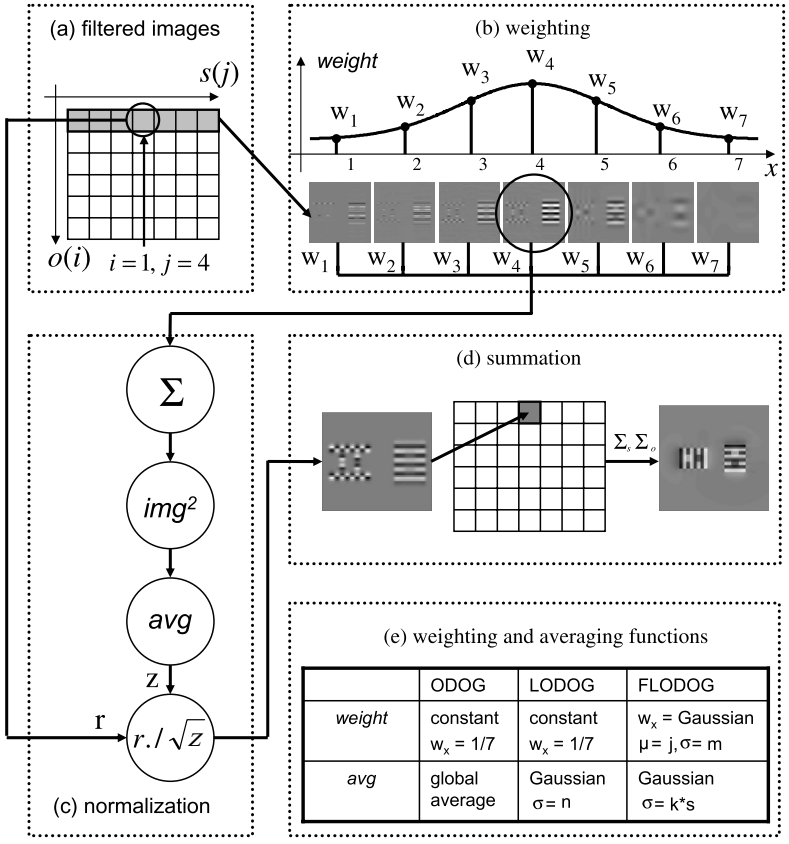
Taken together, these results show that LODOG, in general, works at least as well as ODOG, especially when the window size is larger, such as 4°. This shows that the ODOG normalization step can be made local without reducing the ability of the model to predict a range of brightness illusions. In fact, this more plausible localization scheme actually allows LODOG to predict some illusions that ODOG does not. There remain, however, many equal energy variants of White's illusion that LODOG does not account for. If a normalization-based model is to account for these illusions, a more complex extension to ODOG is necessary.

Examining the results we hypothesized that there were two extensions which would make the model more biologically plausible and which might also improve its ability to predict the illusions. One is that the size of the normalization window should not be constant; rather, very small-scale filters should be normalized by smaller local regions than large-scale filters. Second, it has been shown that neurons with similar spatial frequency preferences tend to cluster together in V1 (Issa, Trepel, & Stryker, 2000; Tootell,

Silverman, & De Valois, 1981; see also Sullivan & de Sa, 2003). Thus, lateral inhibition between neurons should be biased toward neurons of the same spatial frequency. This suggests a new way to normalize the response of a filter, which is local in spatial terms, and also localized to nearby frequencies. We implemented a new model called FLODOG (frequency-specific locally normalized ODOG), which implements these two changes. While all of these changes are reasonable, and likely to make the model more neurally plausible, they do increase the complexity of the model. In the next section, we will evaluate how well this new model compares with the simpler LODOG model.

5. FLODOG

The FLODOG model extends the LODOG model by adding frequency-dependent normalization windows and local weighting when summing across scales. Fig. 4 outlines how FLODOG works, and how it relates to the ODOG and LODOG models. The differences between the FLODOG and ODOG models start after the 42 filter responses have been generated and weighted by spatial frequency (Fig. 4a). For each filter response (r) a new normalization mask is created. Instead of summing across all scales for a single orientation, a weighted sum across frequencies is used. Each filter weight w is computed using a Gaussian function with standard deviation m, shifted so the highest point is centered on the filter being normalized (Fig. 4b). The sum of the weights is normalized to 1. The weighted sum is converted to a localized energy estimate by squaring the value at each point, blurring the whole image by a Gaussian of standard deviation n, and then taking the square root, point-by-point. n is calculated for each filter response by multiplying s (the standard deviation of the center of the DoG filter that generated r) by a scalar, k. This process generates a point-by-point local energy estimate for each r that also includes energy from nearby scales. r is normalized by dividing each point by square root of z, the local energy estimate for that point (Fig. 4c). Finally, each normalized filter response is summed up to produce a point-by-point estimate of the perceived brightness (Fig. 4d). Though it would be computationally inefficient, ODOG and LODOG can be thought of as operating in the same way as FLODOG does, except the weight w would be constant, and different averaging equations would be used (Fig. 4e). The averaging equation for ODOG is the global root mean square of the summed filters, whereas LODOG uses the same averaging equation as FLODOG except that n (the standard deviation of the blur) is a constant, independent of spatial scale.



	ODOG	LODOG	FLODOG
weight	constant w_x = 1/7	constant w_x = 1/7	w_x = Gaussian μ = j, σ = m
avg	global average	Gaussian σ = n	Gaussian σ = k*s

Fig. 4. The FLODOG model. (a) The 42 filter responses, which are generated identically to ODOG. (b) Example of weighted average calculated for a single filter response, at orientation $i = 1$ and scale $j = 4$. Note that the shape of the Gaussian changes for other values of j. (c) Normalization is applied to a single filter response, using the weighted average from (b) to calculate the local energy. (d) All 42 normalized filter responses are summed to produce the final prediction. (e) A comparison of the different weighting and averaging functions used by ODOG, LODOG, and FLODOG.

We tested a variety of FLODOG parameter combinations, crossing the size of the normalization window ($n = 2s$, $3s$, or $4s$) with the weighted sum across frequencies ($m = 0.25$, 0.5, 1, 1.5, 2, and 3). We tested these models on the same illusions we used for ODOG and LODOG.

5.1. Results

We found that FLODOG performed well across a wide range of parameters. The scaling of the normalization window between $n = 2s$ and $n = 4s$ had minimal effect on the predictions of the model for many of the illusions, with the notable exception of SBC and Bullseye. In contrast, the weighting of nearby frequencies (m) makes a big difference in which illusions are predicted correctly and also the magnitude of the predictions.

Due to space considerations we present a subset of models that we tested (Table 3). We include the model that accounted for the most illusions, FLODOG with a normalization window of $n = 4s$, and a weighting of nearby frequencies of $m = 0.5$. For comparison we also include FLODOG with $n = 2s$, $m = 0.5$, and $n = 4s$, $m = 3$. To allow comparison between different models we also include the response of ODOG, and of the most successful LODOG model (window size of 4°).

FLODOG with $m = 0.5$ accounts for all the White's illusions that LODOG does. In addition, it predicts the correct direction of illusion for WE-zigzag (Fig. 1f) and all of the WE-radial (Figs. 1g–j) and WE-circular (Figs. 1k–m) illusions. None of the models we tested, however, predicted assimilation for the Anderson (Fig. 1d) or Howe (Fig. 1e) versions of White's illusion. Blakeslee and McCourt (2004) found that WE-Howe does not lead to a consistent illusion direction across subjects, so the model's prediction of contrast actually matches what some people see. The model predicts contrast because the solid black and white horizontal bars that the test patches are on produce a strong contrast signal (note that part of the illusion is the same as SBC, an illusion the ODOG models see strongly). This contrast signal is bigger than the assimilation caused by the grating above and below the test patches, so the overall prediction is of contrast. While FLODOG predicts contrast for WE-Anderson, an illusion people consistently see as assimilation, it does predict that the illusion is closer to assimilation than is WE-Howe. This is because the test patches are offset from the contrast-inducing horizontal bars in the image. Since, however, the ODOG-based models are overly sensitive to contrast, the reduction in contrast from the offset of the test patches is still not

Table 3
Model results

	ODOG	LODOG $n = 4$	FLODOG $n = 2s$, $m = 0.5$	FLODOG $n = 4s$, $m = 0.5$	FLODOG $n = 4s$, $m = 3.0$	Illusion strength (human)
WE-thick	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	1
WE-thin-wide	**2.08**	**2.31**	**2.52**	**2.07**	**1.72**	1.1
WE-dual	−0.30	**1.11**	**1.93**	**1.67**	**1.58**	
WE-Anderson	−0.15	−0.25	−0.43	−0.03	−0.22	*1.54*
WE-Howe	−0.42	−0.47	−0.94	−0.27	−0.47	*0*
WE-zigzag	−0.51	−0.57	**1.26**	**0.91**	−0.28	
WE-radial-thick-small	−0.67	−0.55	**0.46**	**0.49**	−0.36	
WE-radial-thick	−0.41	−0.29	**0.18**	**0.18**	−0.38	
WE-radial-thin-small	−0.34	−0.20	**2.74**	**2.00**	**0.34**	
WE-radial-thin	−0.22	**0.05**	**3.24**	**2.31**	**0.43**	
WE-circular1	−0.82	−1.00	**0.28**	**0.49**	−1.36	
WE-circular0.5	−0.53	−0.65	**1.84**	**1.45**	−0.75	
WE-circular0.25	−0.38	−0.48	**3.64**	**2.58**	−0.07	
Grating induction	**2.03**	**1.77**	**0.66**	**0.41**	**1.32**	1.49
SBC-large	**4.75**	**6.33**	**3.96**	**2.37**	**6.35**	2.72
SBC-small	**6.22**	**9.19**	**5.96**	**4.01**	**10.27**	4.73
Todorovic-equal	−0.36	−0.37	**0.08**	−0.12	−0.18	0.53
Todorovic-in-large	**0.49**	**0.52**	**0.39**	**0.38**	**0.67**	0.57
Todorovic-in-small	**0.80**	**0.86**	**1.08**	**0.71**	**1.32**	1.05
Todorovic-out	**0.35**	**0.40**	**0.03**	**0.15**	**0.28**	0.37
Checkerboard-0.16	**1.10**	**0.97**	**8.03**	**6.13**	**1.36**	*1.78*
Checkerboard-0.94	**0.40**	**0.35**	−4.89	−4.05	**0.05**	*0.68*
Checkerboard-2.1	**0.69**	**0.59**	−1.48	−1.49	**0.19**	*1.36*
Corrugated Mondrian	**0.95**	**0.73**	**0.12**	−0.25	**0.09**	2.6
Benary cross	**0.09**	**0.05**	**0.05**	**0.03**	**0.06**	2.2
Todorovic Benary 1–2	−0.12	**0.54**	**0.11**	**0.10**	**0.34**	2.86
Todorovic Benary 3–4	−0.12	**0.55**	**0.14**	**0.11**	**0.36**	2.28
Bullseye-thin	−0.74	−0.56	**0.54**	**1.17**	**0.45**	
Bullseye-thick	−0.77	−0.58	**0.07**	**0.80**	**0.50**	

Illusions are listed in the same order as in Fig. 1 and Tables 1 and 2. Cells in bold indicate the model predicts that the illusion goes in the same direction people typically see it. Note that human values have been scaled so that 1.0 equals the average strength of WE-thick.

enough for the assimilation caused by the grating to dominate the overall prediction.

For SBC (Figs. 1o and p), the prediction FLODOG makes depends on the size of the normalization window, with larger windows predicting smaller illusions strengths. The original ODOG model predicts a much stronger SBC illusion than people tend to see, so the smaller prediction of the FLODOG model is more realistic, and reason to prefer the model with a larger normalization window. FLODOG accounts for the Todorovic SBC illusions (Figs. 1q–t) as well, except for Todorovic-equal (Fig. 1q). FLODOG with a 2*s* normalization window actually predicts it in the correct direction, but with a very small strength. As mentioned earlier, however, psychophysical measurements of this illusion have produced conflicting reports of which direction it goes, so it is unclear what the proper model output should be.

The FLODOG models with $m = 0.5$ do poorly on the checkerboard illusion (Figs. 1u–w). They predict that Checkerboard-0.16 (Fig. 1u) is much stronger than it really is, and predict the other two checkerboard illusions in the wrong direction. It is worth noting, however that the checkerboard illusion depends on spatial scale, and it switches from assimilation at small scales (i.e., Checkerboard-0.16) to contrast at larger scales, with the actual crossover point varying between subjects (Blakeslee & McCourt, 2004). The FLODOG model fails because it predicts assimilation at all these scales, and thus it could be failing because it has a different crossover point between assimilation and contrast. The trend (decreasing assimilation with increasing scale) is in the correct direction.

The same FLODOG models also have difficulty with the corrugated Mondrian (Fig. 1x). The 2*s* model predicts the illusion in the right direction, but predicts that it is much weaker than people see it, and the 4*s* model predicts the illusion in the opposite direction. While this does not support the model, the Mondrian stimuli may depend on more high-level factors that cannot be captured by a low-level model. Although ODOG does make a better prediction, given that the ODOG model is clearly incomplete, it is possible that its account of the Mondrian is erroneous.

The FLODOG models also predict both the Benary cross (Fig. 1y), and also the Todorovic version (Fig. 1z), which our implementation of the original ODOG model cannot account for. Note, however, that the strength of the illusion is predicted to be weaker than people see it. Finally, FLODOG predicts the Bullseye illusion (Figs. 1aa and bb), with the 4*s* version predicting a stronger illusion than the 2*s* version.

FLODOG can be made more similar to the LODOG model by setting the weighting of nearby frequencies (*m*) to a larger number, such as 3.0. With this parameter setting, energy at any scale within an orientation will influence the normalization of each filter. Table 3 shows how

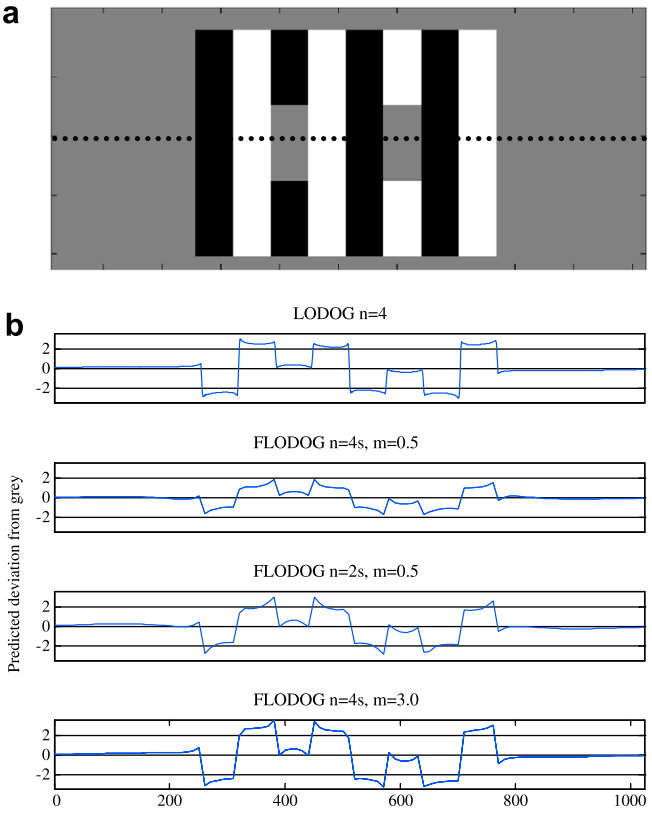


Fig. 5. Model predictions for WE-Thick. (a) Dotted line indicates location of cross-sections. (b) Perceived brightness predicted by LODOG and FLODOG models along the cross-section.

such a model performs. In contrast to the other FLODOG configurations, only the thin variants of WE-radial are predicted, no version of WE-circular is predicted, and WE-zigzag is not predicted. This configuration of FLODOG does, however, predict all versions of the checkerboard illusion, and weakly predicts the corrugated Mondrian illusion. Thus, we can see that a large part of FLODOG's success is due to normalizing each filter response primarily by itself, rather than by all the filters of the same orientation.

Cross-sections of the FLODOG model, shown in Fig. 5 for WE-thick, reveal that FLODOG predicts marked non-uniformity within regions of the input that are uniformly shaded. While no psychophysical experiments exist which directly contradict these variations, it is clear that the magnitude predicted by the model is larger than experienced when just looking at the input images. It is possible that much smaller-scale non-uniformities are measurable under psychophysical testing, and indeed Blakeslee and McCourt (1999) collected psychophysical measurements that showed the gray test patches in WE-thick do have non-uniform brightness, in the same direction as seen by the ODOG model. In addition, Blakeslee and McCourt (1997) found that the test patches in SBC and grating induction also have non-uniform brightness, of the same pattern ODOG predicts. At least for the grating induction illusion

(Fig. 1n), the non-uniform brightness predicted by ODOG is clearly visible to the untrained eye, without any psychophysical testing required. It is not known why this occurs with grating induction, and not other illusions. In any case, the non-uniform brightness perception predicted by FLODOG is not entirely wrong, though the magnitude is clearly too large.

The issue of non-uniform brightness responses cuts across different filter-based approaches, and deserves further study. While FLODOG and LODOG predict larger non-uniform responses than ODOG does for many illusions, it is worth noting that ODOG can also predict very non-uniform perceived brightness in some situations, such as SBC-large (Blakeslee & McCourt, 1999). One possibility is that these non-uniformities are only present in the early stages of brightness processing, and at some later stage the different brightness values are averaged within a region to produce a single perceived shade of gray, as suggested by Grossberg and Todorovic (1988).

In conclusion, the FLODOG model predicts many illusions that LODOG does not. The exact parameter values are not critical, but in general the quality of the predictions is better when a larger window is used for normalization (such as $4s$) and when each scale is normalized relatively independently (such as when $m = 0.5$). While different parameter settings do allow the model to account for other illusions, these settings predict the widest range of illusions and are the recommended values to use when testing the model in future work.

It seems that the biologically plausible changes to LODOG implemented in FLODOG do improve the predictive power of the model. FLODOG's failure to predict illusions that are correctly predicted by ODOG and LODOG raise the possibility that the success of these models depends on less plausible mechanisms (such as normalizing each filter response equally by all frequencies of the same orientation, and a fixed size normalization window). The true cause of these effects may be due to other mechanisms that operate before or after the steps modeled by FLODOG.

6. Conclusions

This paper explored the response normalization mechanism of the original ODOG model, and found that it could be extended to be both more neurally plausible and more effective at predicting brightness illusions.

Our simplest extension, LODOG, does not reduce the functionality of the ODOG model, but only minimally increases the number of illusions correctly predicted. At the least, this shows that the response normalization process can be successfully calculated locally.

Our more advanced model, FLODOG, is not only more plausible, but also increases the number of illusions correctly predicted. The success of this model suggests that many brightness illusions could be due to low-level mechanisms in early visual processing. For instance, an area like V1 could behave much like FLODOG if lateral interactions or feedback cause cells that respond to the same orientation and similar frequencies to inhibit each other. Indeed, Rossi and Paradiso (1999) have shown that there are a significant number of cells in V1 which respond to the perceived brightness of stimuli, instead of the actual luminance (though this has only been tested for SBC-like stimuli), which supports the idea that brightness perception could occur in early visual areas. Since FLODOG uses filters of much larger spatial extent than does V1, it is clear that FLODOG is not a model of V1 at the level of individual neurons, but might represent the combined activity of groups of neurons, either in V1, or distributed across multiple early visual areas.

Why might early visual areas perform the kind of calculation that FLODOG models? Perhaps it is part of the computational processing that produces lightness constancy. Another possibility is that it is a side effect of an entirely different calculation. Schwartz and Simoncelli (2001) have developed a model of the firing rate of neurons in V1, in which neurons are inhibited by neighboring neurons that have correlated variance of firing rate. When trained on natural scenes, where adjacent regions have similar orientations and spatial frequencies, the normalization resulting from this model weights nearby spatial frequencies and orientations most heavily, similar to FLODOG. The upside of such a model is that it makes the population activity more statistically independent, when exposed to a natural image. Increasing statistical independence is thought to produce a more optimal neural code (Barlow, 1961).

There may be additional modifications to the FLODOG model that could improve its performance, without radical changes to its mechanisms. One extension that we have considered is normalizing each orientation by nearby orientations, also using a Gaussian weighting analogous to how we currently weight nearby frequencies. Experiments with this extension, however, found that normalizing across nearby orientations does not improve the predictions of the model. Another alternative, which we have not explored, is that the orientation tuning of the model might be too broad, and that more than just six orientations should be used.

FLODOG is only a model of early stages of brightness processing. There is clearly a need for some form of anchoring to explain the fact that the lightest surface in a scene tends to look white (Gilchrist et al., 1999). In addition, the percept of transparency can change whether a surface looks white or black (Anderson & Winawer, 2005). A model like FLODOG cannot explain either of these types of effects. Our work does, however, suggest that low-level mechanisms could be a significant factor in many of the illusions studied here. By itself, however, the existence of a successful low-level model does not prove that higher-level mechanisms do not contribute as well. Further work will be necessary to develop variations of these illusions that pit aspects of the high or low-level theories against each other, to determine their relative contributions.

Finally, we would like to stress the utility of having a model that can be tested on an arbitrary input image with minimal assumptions.[2] Any grayscale image can be fed into the LODOG and FLODOG models, and a prediction of brightness produced. We look forward to applying the models to new illusions to see if our low-level approach can account for other brightness illusions not studied here.

Acknowledgments

We thank Micah Richert for his role in our first implementation of the ODOG model, and Sophie Soong for implementing many of the illusions we tested. This material is based upon work supported by the National Science Foundation under NSF Career Grant No. 0133996 to VR de Sa. AE Robinson and PS Hammon were supported by NSF IGERT Grant #DGE-0333451 to GW Cottrell.

References

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

Anderson, B. L. (2001). Contrasting theories of White's illusion. *Perception, 30*, 1499–1501.

Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature, 434*, 79–83.

Anstis, S. (2003). White's effect in brightness & color. Online Demonstration. http://psy.ucsd.edu/~sanstis/WhitesEffect.htm.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.

Benary, W. (1924). Beobachtungen zu einem experiment uber helligkeitskontrast. *Psychologische Forschung, 5*, 131–142.

Bindman, D., & Chubb, C. (2004). Brightness assimilation in bullseye displays. *Vision Research, 44*, 309–319.

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research, 37*, 2849–2869.

Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research, 39*, 4361–4377.

Blakeslee, B., & McCourt, M. E. (2001). A multiscale spatial filtering account of the Wertheimer-Benary effect and the corrugated Mondrian. *Vision Research, 41*, 2487–2502.

Blakeslee, B., & McCourt, M. E. (2004). A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization. *Vision Research, 44*, 2483–2503.

Blakeslee, B., Pasieka, W., & McCourt, M. E. (2005). Oriented multiscale spatial filtering and contrast normalization: a parsimonious model of brightness induction in a continuum of stimuli including White, Howe and simultaneous brightness contrast. *Vision Research, 45*, 607–615.

Clifford, C. W. G., & Spehar, B. (2003). Using colour to disambiguate contrast and assimilation in White's effect. *Journal of Vision, 3*, 294a.

DeValois, R. L., & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.

Gilchrist, A. (2006). *Seeing in Black and White*. New York: Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*, 795–834.

Grossberg, S., & Todorovic, D. (1988). Neural dynamics of 1-D and 2-D brightness perception: A unified model of classical and recent phenomena. *Perception & Psychophysics, 43*, 241–277.

Hong, S. W., & Shevell, S. K. (2004). Brightness contrast and assimilation from patterned inducing backgrounds. *Vision Research, 44*, 35–43.

Howe, P. D. L. (2001). A comment on the Anderson (1997), the Todorovic (1997), and the Ross and Pessoa (2000) explanations of White's effect. *Perception, 30*, 1023–1026.

Howe, P. D. L. (2005). White's effect: removing the junctions but preserving the strength of the illusion. *Perception, 34*, 557–564.

Issa, N. P., Trepel, C., & Stryker, M. P. (2000). Spatial frequency maps in cat visual cortex. *Journal of Neuroscience, 20*, 8504–8514.

McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research, 22*, 119–134.

Pessoa, L., Baratoff, G., Neumann, H., & Todorovic, D. (1998). Lightness and junctions: variations on White's display. *Investigative Ophthalmology and Visual Science (Supplement), 39*, S159.

Rossi, A. F., & Paradiso, M. A. (1999). Neural correlates of perceived brightness in the retina, lateral geniculate nucleus, and striate cortex. *Journal of Neuroscience, 19*, 6145–6156.

Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4*, 819–825.

Sullivan, T.J., & de Sa, V.R. (2003). Unsupervised learning of Complex Cell Behavior. Unpublished Tech Report. http://www.sullivan.to/Sullivan_CCell1.pdf.

Todorovic, D. (1997). Lightness and junctions. *Perception, 26*, 379–395.

Tootell, R. B., Silverman, M. S., & De Valois, R. L. (1981). Spatial frequency columns in primary visual cortex. *Science, 214*, 813–815.

White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*, 413–416.

[2] Unfortunately, implementing these models is not a trivial task. To aid further work on brightness perception we will make our code (MATLAB) available for research purposes to anybody that asks. In addition, we will make available compiled versions of these models so that access to MATLAB is not required.