Available online at www.sciencedirect.com
ScienceDirect
Vision Research 48 (2008) 733–751
Vision Research
www.elsevier.com/locate/visres

Multiresolution wavelet framework models brightness induction effects

Xavier Otazu *, Maria Vanrell, C. Alejandro Párraga

Computer Vision Center/Computer Science Department, Universitat Autònoma de Barcelona, Campus UAB, Cerdanyola del Vallès, 08193 Barcelona, Spain

Received 10 May 2007; received in revised form 4 December 2007

Abstract

A new multiresolution wavelet model is presented here, which accounts for brightness assimilation and contrast effects in a unified framework, and includes known psychophysical and physiological attributes of the primate visual system (such as spatial frequency channels, oriented receptive fields, contrast sensitivity function, contrast non-linearities, and a unified set of parameters). Like other low-level models, such as the ODOG model [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], this formulation reproduces visual effects such as simultaneous contrast, the White effect, grating induction, the Todorović effect, Mach bands, the Chevreul effect and the Adelson–Logvinenko tile effects, but it also reproduces other previously unexplained effects such as the dungeon illusion, all using a single set of parameters.

Keywords: Visual system; Brightness induction; Wavelet transform

1. Introduction

In visual perception, the term brightness often refers to a non-quantitative perception of light elicited by the luminance of a visual target (see Gilchrist, 2006, p. 6). This brightness depends not only on the light reaching the retina from the visual target but also on the spatial distribution of light on its surroundings. *Brightness induction* refers to this change of appearance due to the surrounding light and its effects are classified according to the perceptual direction of the change. When the change in brightness of the visual target goes away from the surrounding brightness, it is called *brightness contrast* (Heinemann, 1955) and when the change goes towards that of the surrounding brightness, it is called *brightness assimilation* (Helson, 1963).

In the following section, we review several brightness induction effects which are widely studied in the literature.

1.1. Brightness induction effects

One of the oldest known examples of brightness induction is the simultaneous brightness contrast (SBC) effect (Heinemann, 1955; Wallach, 1948). This effect decreases for increasing test field size, but is still strong for test fields as large as 10 deg (Yund & Armington, 1975). Since this is far larger than receptive fields of retinal and lateral geniculate nucleus (LGN) neurons in monkey (De Valois & De Valois, 1988), it suggests that other types of neurons may be involved in the process. Such neurons with small excitatory centres and large inhibitory surrounds (which may be suited for the task of shifting brightness towards or away from a large test field) were found in area V4 of the primate visual cortex (Schein & Desimone, 1990; Spillmann & Werner, 1996).

A second known example of brightness induction is the so called grating induction (GI) effect (McCourt, 1982). The perceived contrast of the induced grating again decreases with increasing test field, but also decreases with the spatial frequency (s.f.) of the inducing grating (Foley &

* Corresponding author.
E-mail address: xotazu@cvc.uab.es (X. Otazu).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.12.008

McCourt, 1985). The induced grating is still perceived in test patches as large as 6 deg (Blakeslee & McCourt, 1997). It has been argued that these two phenomena (SBC and GI) are just manifestations of the same underlying mechanisms (Blakeslee & McCourt, 1997) and their physiological basis are related to the discovery of cortical neurons in cat (Rossi, Rittenhouse, & Paradiso, 1996) and monkey (Gilbert, Das, Ito, Kapadia, & Westheimer, 1996) that integrate over relatively large distances.

Another well known brightness effect is the White effect (White, 1979), where grey test patches of the same luminance appear to have different brightness when placed on top of the black and white bars of a square grating. Here, the brightness shift is independent of the aspect ratio of the test patch (i.e. it does not depend on the amount of white or black border near or in contact with the test patch). What makes this effect even more interesting is that the contrast between the grey patch and its borders (or surrounding area) seems to be less important than the contrast with the bar upon which it is situated.

A similar effect was described by Todorović (1997) where the brightness shift seems to be independent of the amount of black or white background in contact with the test patch. Several explanations both at the receptive-cortical level and at higher perceptual levels have been attempted to explain the White effect (see below). However, it is clear that the most plausible explanation for this effect at receptive level needs both elongated cortical filters (Foley & McCourt, 1985; White, 1981) and the operation of spatially extensive neuronal mechanisms, as opposed to isotropic receptive fields and shorter range spatial interactions such as those found in the retina.

Mach bands are brightness maxima and minima perceived at the beginning and end of luminance gradients, respectively (Mach, 1865). They have been interpreted in terms of lateral inhibition of retinal ganglion cells (Goldstein, 2002) and more recently as a consequence of the physical properties of real world luminance gradients (Lotto, Williams, & Purves, 1999).

The Chevreul illusion (Chevreul, 1890) is the name given to the brightness minima and maxima, respectively, perceived at the foot and tip of each step in a luminance staircase. There have been attempts to explain this illusion in terms of single channel and the contrast sensitivity function (Cornsweet, 1970) but this explanation has been abandoned in favour of multi-channel models and local features within the steps (Morrone, Burr, & Ross, 1994; Peromaa & Laurinen, 2004). However, there are also alternative explanations of this effect in terms of a filling-in process triggered by edges at the different spatial scales (Pessoa, Mingolla, & Neumann, 1995).

The Adelson tile illusion (Adelson, 1993), appears when a "wall made from homogeneous blocks" is spatially modulated by a horizontal dark stripe in such a way that some of the diamonds that form the top of the blocks fall within the brighter part of the wave and some fall within the darker part. The top of the blocks (horizontal diamond shapes) are constructed to be physically the same (i.e. they reflect the same amounts of light), but the diamonds that fall in the light strip look darker than the diamonds in the dark strip. By rearranging the pattern to make the effect disappear while keeping the local contrast around diamonds the same, Adelson (1993) demonstrated that explanations need to incorporate long-ranging receptive field interactions.

A modification of the Adelson tile illusion was introduced by Logvinenko (1999), who blurred the contrast edges of the horizontal strips, thus removing any apparent transparency (and verifying that the illusion still holds). There are a wide variety of explanations for these Adelson illusions, ranging from "low-level" explanations based on local contrast and multi-scale spatial filtering (Blakeslee & McCourt, 1999; Cornsweet, 1970) to those based on the role of borders or luminance junctions between or across strips (Adelson, 1993; Adelson, 2000; Anderson, 1997), "high-level" explanations where the explanation is based on how the visual system deals with illumination (Gilchrist et al., 1999; Logvinenko & Ross, 2005) and "multi-level" explanations (Kingdom, 2003).

There are particularly striking cases where simple predictions from the SBC effect (black surroundings induce lighter targets, etc.) seem to be completely reversed. One example of these is the Necker cube presented by Agostini and Galmonte (2002) whose dashed sides are perceived lighter even when they are completely surrounded by white background and vice versa. Other examples are the dungeon illusion (Bressan, 2001) and the Checkerboard contrast illusion (De Valois & De Valois, 1988), where grey features surrounded by white look lighter and grey features surrounded by white look darker. These were interpreted in terms of higher level "grouping factors" where, for example each set of dashes in the Necker cube is *anchored* by the cube (Gilchrist, 2006).

In the following section, we review several attempts to model brightness induction within a computational framework.

1.2. Modelling attempts

Some of the most successful computational models of brightness perception were developed using multi-scale approaches to low-level vision. They postulate that edges and lines are the driving features of early vision and a set of operators (receptive fields) are in charge of detecting these (du Buf, 1994; Fiorentini, Baumgartner, Magnussen, Schiller, & Thomas, 1990; Morrone & Burr, 1988; Tolhurst, 1972). These models may differ in the way these operators interact with each other. For example, both the models of Tolhurst (1972) and Morrone and Burr (1988) employ pairs of orthogonal operators but the former applies mutual inhibition between them while the latter pools their responses. The model of Fiorentini et al. (1990) employs a single filter type at different spatial scales while the model proposed by du Buf (1994) uses operators

that resemble pairs of simple cells centred at the same location but in quadrature.

A second type of model is based on the framework originally proposed by Marr (1982). An example of these is MIRAGE (Watt & Morgan, 1985) which filters the stimuli at various spatial scales and generates a list of "primitives" and uses a set of rules to detect lines and edges. A more sophisticated version was proposed by Kingdom and Moulden (1992) and called MIDAAS, which includes a gain control mechanism (light adaptation), spatial scale filtering, thresholding and symbolic descriptions at each spatial scale before applying a set of rules and combining the outputs across scales.

A third type of model propose that the main task of the visual system is not to extract salient features of scenes (as do the other two types of models) but to build perceptual representations that keep the geometric structure of scenes (Pessoa et al., 1995). This model uses a contrast-driven and a luminance-driven representation. The first representation is then filtered to produce boundaries. The filtering overshoots and undershoots trapped by these boundaries are filled-in before the contrast and luminance signals are recombined to provide the model's output, which is meant to resemble the spatial distribution of the percept. These models can account for several brightness induction effects such as the Mach bands and the Chevreul illusion with varying degrees of accuracy (for a review see Pessoa et al., 1995 & Gilchrist, 2006).

A unified brightness model based on low-level isotropic filters (difference of Gaussians or DOG) sensitive to contrast at multiple spatial scales was proposed by Blakeslee and McCourt (1997) to explain the GI effect which, they argue, cannot be explained by a fill-in type of model. The main difference between this model and the previous ones (Kingdom & Moulden, 1992; Moulden & Kingdom, 1991) was the presence of more s.f. filters (sensitive to lower spatial frequencies) and a weighting scheme adjusted to match the psychophysical data. This simple model is capable of accounting for other brightness effects such as SBC and the Hermann grid illusion. A more sophisticated version, which includes non-linearly pooled anisotropic filters (oriented difference of Gaussians or ODOG) and a normalisation to equalise the global response at each orientation, was developed to account for a variety of brightness effects that require oriented filters such as the White effect (Blakeslee & McCourt, 1999; Blakeslee & McCourt, 2001; Blakeslee & McCourt, 2004; Blakeslee, Pasieka, & McCourt, 2005). The latest extension of ODOG was made by Robinson, Hammon, and de Sa (2007), who constrained normalisation to make it more neurally plausible and expanded the range of illusions predicted by the model. Another multiresolution perceptual model is the one developed by D'Zmura and Singer (1998) and D'Zmura and Singer (1999). Here the visual space is decomposed into s.f. and orientation axis, and subdivided into several regions according to their spatial properties. The authors use four (octave-wide) s.f. channels and six orientations of 30 deg width. The contrast of the surround is introduced in this model as a Gaussian blurring of the full-wave rectified frequency channel (called spatial pooling of contrast).

The *Brightness Induction Wavelet Model* (BIWaM) we present here shares some similarities with both the ODOG and the D'Zmura and Singer (1998, 1999) models. Although a multiresolution decomposition of the stimulus is performed, the output of the s.f. channels is processed differently (see below) and the contrast sensitivity function and stimulus distance are introduced explicitly. In the ODOG model the interaction of the central stimulus and its surround is performed through a normalisation of the total visual space (in the case of the D'Zmura model this explicit comparison is not performed). In our model, we introduced a precise dependency on the contrast energy of the surround compared to the central stimulus. Another crucial feature of our model is that for each of the illusion simulations described below all parameters were held constant.

2. The Brightness Induction Wavelet Model (BIWaM)

In this work, we propose a new low-level brightness induction model (the BIWaM) that combines three important stimulus features, namely spatial frequency, spatial orientation and surround contrast to explain brightness assimilation/contrast phenomena. This is done through a multiresolution wavelet decomposition which separates the achromatic input image into different spatial frequency and orientation components (reminiscent of parvocellular s.f. channels and cortical orientation-selective receptive fields). The recovery of the perceptual brightness image is done by weighting the wavelet coefficients using a modified version of the contrast sensitivity function (CSF). This modified CSF takes into account the (spatial) surround information, so that the value of the contrast sensitivity increases when surround contrast decreases and vice versa. Observation distance is also taken into account to generalise the model.

The choice of the wavelet transform as the main framework for this work was motivated by the fact that wavelets share several mathematical properties that fit nicely with those of the early visual system (e.g. two-dimensional receptive field profiles are well described by two-dimensional Gabor functions (Jones & Palmer, 1987)). Although there is considerable variability in the receptive field shapes across neurons (Tolhurst & Thompson, 1982; De Valois, Albrecht, & Thorell, 1982) and no single basis set can capture this variability, wavelets provide basis functions that are well localised (both in space and frequency) among other mathematical properties (such as self-similarity, ease of mathematical representation, etc.) which make them popular among modellers (Field, 1999; Harvey & Doan, 1990; Olshausen & Field, 1997; Van Rullen & Thorpe, 2001; Zetzsche & Nuding, 2005).

We aim to produce the simplest mathematical formulation (and to include the least possible number of

free parameters) which models these three properties in a manner compatible with current physiological and psychophysical research with a unified set of parameters.

Despite sharing many of its characteristics and philosophy with other multiresolution models of the same type (mainly with the ODOG and D'Zmura's models), BIWaM has some important differences from those, such as:

- It is based on a wavelet decomposition which allows a full reconstruction of the original image (i.e. an invertible transform) with basis that have a profile similar to a Gabor function.
- It relates brightness induction to the CSF by weighting each spatial scale (see next section).
- It accounts for the observation distance, which is not explicitly defined in the other models (it predicts different brightness induction effects depending on the observer's physical position, which is the true behaviour of the human visual system (HVS)). This issue was recognised as problematic by other authors (e.g. FLODOG model, Robinson et al., 2007).
- It explicitly introduces the stimulus-surround contrast energy. ODOG family models use a normalisation, either global or local, and the D'Zmura's model performs a local gain control in order to modify the different scales.
- Its parameters are fixed and account (simultaneously) for all the results described below.

2.1. Contrast sensitivity function

The detection threshold for sinusoidal gratings depends on the grating s.f. and this relationship is described by the contrast sensitivity function $C(v)$, where v represents the spatial frequency, which is band pass for achromatic stimuli (Mullen, 1985; Simpson & McFadden, 2005).

Experiments with square-wave periodic patterns have shown that brightness assimilation effects increase when the s.f. of the target feature is higher than a certain induction threshold. For brightness induction, this transition point v_{thr} was estimated to be near 4 cpd (Smith, Jin, & Pokorny, 2001; Walker, 1978) and for chromatic contrast induction between 4 and 6.7 cpd (Fach & Sharpe, 1986; Mullen, 1985).

Given an observation distance d the psychophysical CSF $C(v)$ function can be defined in the scale space as a function $C_d(s)$, where s is the spatial scale, and approximated by a piecewise function defined by two Gaussians. We can also define a particular scale s_{thr} associated to v_{thr}, the spatial frequency (see Appendix A) where the CSF peaks. A choice of $v_{\text{thr}} = 4$ cpd was made taking into account the changes that occur to the retinal CSF with mean field luminance (its peak varies between 1 cpd at low luminance levels and 8 cpd for intense photopic backgrounds) (Vannes & Bouman, 1967).

This allows us to define the function $C_d(\dot{s})$, being $\dot{s} = s - s_{\text{thr}}$.

The spatial decomposition of the visual stimuli into the one octave-bandwidth independent channels that form the basis of the CSF is modelled by a multiresolution wavelet transform, as described in the next section.

2.2. Multiresolution wavelet analysis

A common approach to model the responses of visual cortical areas involves the use of Gabor functions (Daugmann, 1980), which are good descriptions of the early visual system's receptive fields profiles but have the disadvantage of being non-invertible (i.e. the original image cannot be fully recovered). Our wavelet basis functions (i.e. the mother wavelet) are not strictly Gabor functions, but have a similar profile (smooth, symmetric, and highly concentrated in both space and frequency). They are based on an algorithm halfway between the Mallat decomposition (Mallat, 1989; Mallat, 1998) and the à *trous* algorithm (Holschneider, Kronland-Martinet, Morlet, & Tchamitchian, 1989; Otazu & Vanrell, 2006). We decided against using a full Mallat decomposition (such as the Daubechies wavelets, which are orthogonal and compact) mainly because its basis functions are not smooth or symmetric like the Gabor function. On the other side, the à *trous* algorithm is a very flexible multiresolution wavelet (equivalent to a bi-orthogonal wavelet decomposition) that allows us to define smooth analysis filters that may lead to smooth and symmetric wavelets. Furthermore, it does not require us to define the synthesis filter, because the synthesis of the original image is performed by simple addition. It has also the advantage of being an undecimated algorithm which, in contrast to decimated algorithms, allows us to obtain a translation invariant decomposition. For the present work, we used a decimated version of the à *trous* algorithm in order to reduce the computational complexity of the model (a common approach in computer vision applications, e.g. JPEG2000 image compression) and to simplify future implementations. Despite these mathematical differences, the main concept and philosophy of the two algorithms is the same.

BIWaM decomposes its input in a series of new images ω_s^o (or wavelet planes) that contain features of the original image at different spatial frequencies (indexed by s), and spatial orientations (indexed by o). Our algorithm also decomposes the image in 3 orientations with 45 deg orientation bandwidths (i.e. vertical, horizontal and diagonal orientations). In Fig. 1 we show a graphical scheme of this decomposition. The terms ω_1^v, ω_1^h and ω_1^d represent the highest frequency components of the input image for the vertical, horizontal and diagonal orientations, respectively. The terms ω_2^v, ω_2^h and ω_2^d represent similar orientation wavelet planes of s.f. one octave lower. The c_2 image is a residual plane, which is a smoothed version of the original image and can be similarly decomposed, as shown in Fig. 1(b). After being decomposed, the original image I is

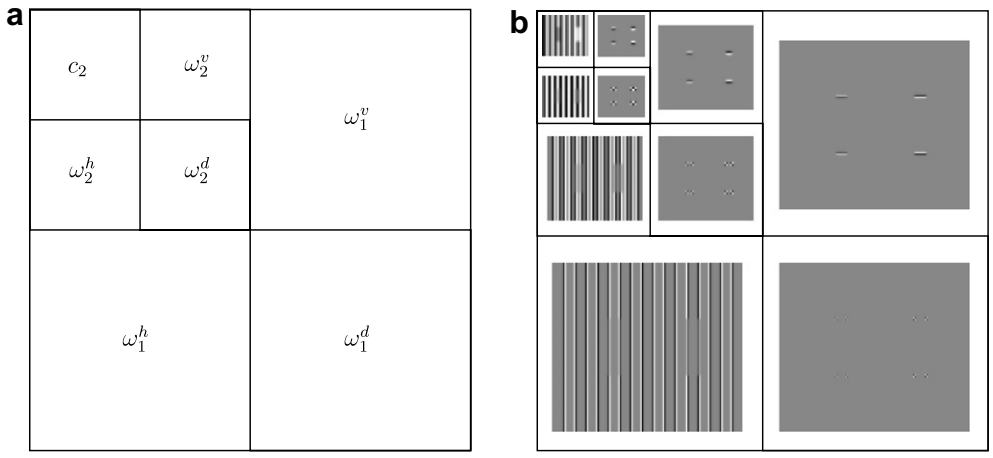

Fig. 1. Multiresolution decomposition. Panel (a) the White effect image from Fig. 2(a) is decomposed into several wavelet planes ω which contain features of a certain s.f. and orientation. Panel (b) shows the representations of these planes (only 3 multiresolution levels are shown).

represented as a summa of wavelet planes of different s.f. and orientation as follows:

$$I = \sum_{s=1}^{n} \left(\omega_s^v + \omega_s^h + \omega_s^d\right) + c_n, \tag{1}$$

where n is the number of wavelet planes. The term c_n is the residual plane, which is a low resolution version of the original image. This expression can be written more compactly as

$$I = \sum_{s=1}^{n} \sum_{o=v,h,d} \omega_s^o + c_n, \tag{2}$$

being the index o the several orientations vertical, horizontal and diagonal, i.e. $o = v, h, d$.

The s.f. channels (or wavelet planes) of our model have a bandwidth and layout similar to the visual system channels that determine the shape of $C(v)$.

2.3. Assumptions

As mentioned before, there is ample evidence that the perception of a central stimulus can be modified by the spatial content of the surroundings (Chubb, Sperling, & Solomon, 1989; D'Zmura & Singer, 1998, 1999; Heeger, 1992; De Valois et al., 1982; Werner, 2003; Yu, Klein, & Levi, 2001; Yu, Klein, & Levi, 2002). In this section we describe how the centre–surround interaction of the three main stimulus properties: spatial frequency, spatial orientation and contrast was modelled.

2.3.1. Stimulus-surround relative spatial frequency

The spatial frequency content of the surroundings is one of the main contributors to the perceived brightness changes in a central stimulus. As shown by grating perception studies (Chubb et al., 1989; D'Zmura & Singer, 1998, 1999; Werner, 2003; Yu et al., 2001, 2002), when the spatial frequencies of both central and surround stimulus are similar, brightness contrast of the central stimulus is reduced (brightness assimilation) and when these frequencies are different the central stimulus contrast is enhanced (brightness contrast). Therefore, brightness assimilation is only performed when both central and surround stimuli have similar spatial frequencies within a frequency range of about an octave (Blakemore & Campbell, 1969; D'Zmura & Singer, 1998, 1999; Graham & Nachmias, 1971; De Valois et al., 1982; Werner, 2003; Wilson, McFarlane, & Phillips, 1983; Yu et al., 2001, 2002). Panel (a) in Fig. 2 shows this property. In this figure, the grey patches have the same horizontal s.f. as the surrounding black and white stripes. The left patch is perceived darker because of the induced brightness assimilation with the contiguous dark vertical stripes, and similarly for the right grey patch among contiguous white stripes which is perceived brighter. Panel (b) shows how doubling the s.f. of the grey patches (i.e. one octave difference with the background), weakens the effect. Considering this, we modify the CSF according to the first of the three assumptions of our model:

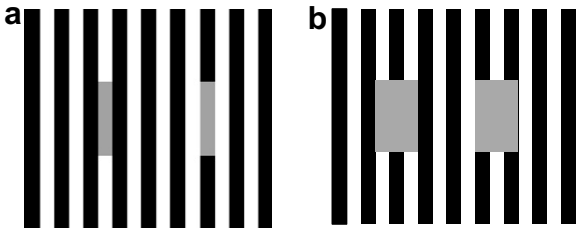

Fig. 2. (a) Example of the influence of surround spatial frequency. Grey patches have different brightness because of their different local spatial information content. Both grey patches and vertical stripes have the same width (i.e. horizontal size or spatial scale) and this produces a strong induction effect. (b) When the widths of the grey patches are different to that of the black and white stripes, the brightness induction effect is largely reduced.

Assumption 1. Brightness assimilation is only performed when both central and surround stimuli have similar spatial frequencies within a frequency range of about one octave.

The multiresolution wavelet framework allows us to decompose the visual stimulus into one octave-bandwidth spatial frequency components and estimate the influence of every spatial feature on features of the same spatial scale. The wavelet scales s here are related to the s.f. channels that constitute the $C_d(\dot{s})$ (observer-target distance is included in this CSF, see definition above).

2.3.2. Stimulus-surround relative spatial orientation

Another important contribution to brightness assimilation in gratings comes from the relative orientation of central and surround stimulus. Several studies (Cannon & Fullenkamp, 1991; Solomon, Sperling, & Chubb, 1993; Yu et al., 2001, 2002; Yu, Klein, & Levi, 2003) show that brightness assimilation of a central stimulus is strongest when this and the surround stimulus have identical orientations. On the contrary, when the relative spatial orientations are orthogonal, brightness assimilation of the central stimulus is weakest (brightness contrast is strongest). This effect can be observed in Fig. 3. Following this, we define our second assumption:

Assumption 2. Brightness assimilation is strongest when the central stimulus and the surround stimulus have identical orientations. Furthermore, when the relative spatial orientations are orthogonal, brightness assimilation of the central stimulus is weakest (brightness contrast is strongest).

This assumption was implemented by weighting and pooling the different orientation components of the wavelet transform, e.g. $\omega_s^h, \omega_s^v, \omega_s^d$.

2.3.3. Stimulus-surround relative contrast energy

Surround contrast is the third contribution to brightness induction considered by our model. It has been shown that the contrast of the surround stimuli plays an important role in brightness assimilation effects (Cannon & Fullenkamp, 1991; Chubb et al., 1989; Ejima & Takahashi, 1985; Ellemberg, Wilkinson, Wilson, & Arsenault, 1998; Klein, Stromeyer, & Ganz, 1974; MacKay, 1973; Nachmias & Sansbury, 1974; Yu et al., 2001, 2002, 2003). Brightness assimilation in a central test stimulus increases as its surround contrast increases (and vice versa), before reaching a saturation state. This effect can be observed in Fig. 4 where the two vertical grey lines are placed in backgrounds with different contrast. The left grey line is always in contact with dark stripes, and right grey line is always in contact with white stripes. When the surrounding luminance is uniform (i.e. surround contrast is null), brightness contrast is induced in the lines with the left line being perceived brighter and the right line darker (simultaneous contrast). When the contrast of the surrounding vertical stripes increases (downwards direction in Fig. 4) our perception of the grey stripes changes, clearly reversing their previously perceived difference when the contrast of the surrounding bars is maximum, i.e. when they are black and white. This way, we define the third assumption of our method as:

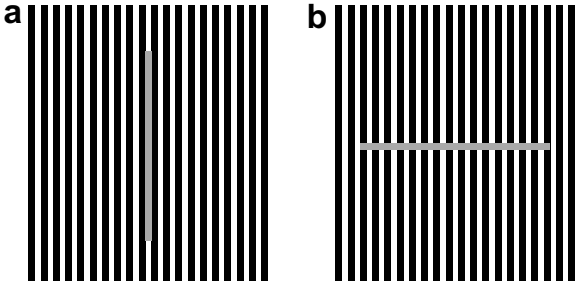

Fig. 3. The vertical grey stripe (a) is perceived darker than the horizontal grey stripe (b) because of brightness assimilation with its surrounding vertical black and white stripes (i.e. classical White effect). If the grey patch is orthogonal (i.e. at 90 deg) to the black and white stripes, brightness assimilation does not occur.

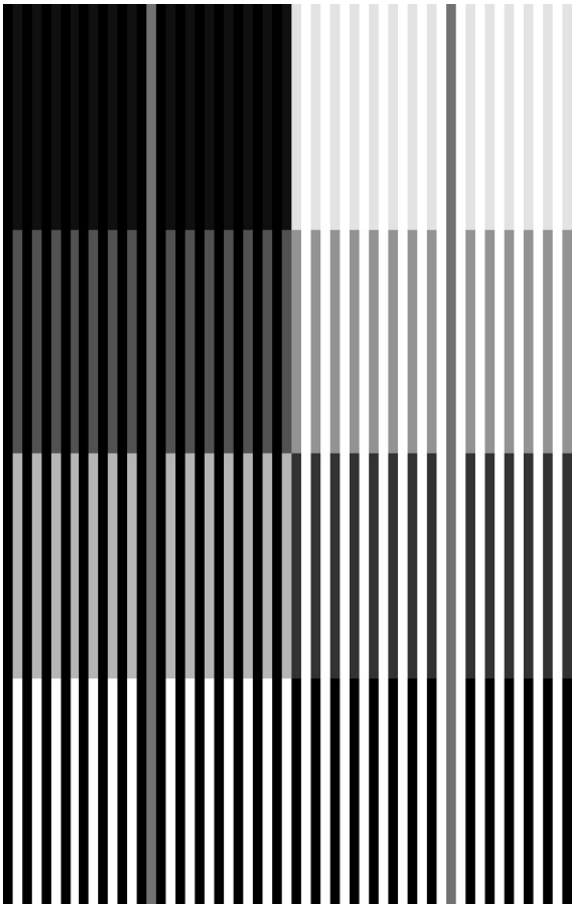

Fig. 4. In this series of images, the surround contrast of the two vertical grey lines is increased downwards in four steps. The left grey line is in contact with black stripes, and the right grey line is in contact with white stripes. As the line extends downwards (and surround contrast increases) brightness induction increases, i.e. the left grey line becomes darker and the right grey line becomes brighter.

Assumption 3. When the brightness contrast of the surround features increases, brightness assimilation increases (i.e. brightness contrast decreases) and vice versa.

To translate this assumption to the model we will need to add the contribution of the surround contrast. The next section shows how we do this.

2.4. Recovery of the "perceived" image

All the assumptions mentioned above are implemented in our model by modifying the wavelet coefficients in Eq. (2), where the goal was reconstructing the perceived or induced image from the decomposed original image. This weighting function attempts to emulate some perceptual properties. As a first approximation, we assume it has a shape similar to the CSF, such as:

$$I_{\text{percep}} = \sum_{s=1}^{n} \sum_{o=v,h,d} \left(C_d^o(\dot{s}) \cdot \omega_s^o \right) + c_n. \tag{3}$$

being I_{percep} the recovered perceptual image, and indexes s and o represent the multiple spatial frequency and orientation planes.

As seen before, the relative s.f. and orientation between central and surround stimulus are important contributors to brightness induction. In the multiresolution framework, image components are grouped by both their s.f. and orientation, giving us a representation in which similar features are grouped into the same data set (i.e. wavelet plane). In this way, Assumption 1 and Assumption 2 are naturally implemented within this framework.

On the other hand, Assumption 3 introduces the concept of surround contrast. Since the coefficients at spatial scale s and orientation o of the wavelet decomposition represent the variation of these image features at a certain scale and orientation around a mean value, the measured energy of these coefficients is related to the contrast energy of the corresponding feature. Therefore, we can easily estimate the relative contrast of a central feature compared to the contrast of its surround features by defining $r = \sigma^2_{\text{cen}}/\sigma^2_{\text{sur}}$, being σ^2_{cen} and σ^2_{sur} the standard deviation of the wavelet coefficients on two concentric annuli that represent a centre–surround interaction around each point (x,y). The regions determined by these annuli (referred as Φ and Ψ, respectively) were modelled as squares of 5×5 and 13×13 points wide, enclosing N_Φ and N_Ψ points inside, respectively. Region Φ was chosen to be 5×5 points wide in order to include 2 complete Nyquist periods (T_s), when measuring the variation of the central region $d_s(x,y)$, therefore, $N_\Phi = 25$. Its surrounding region Ψ is 13×13 points wide, that is about three times larger than the inner region, an approximate ratio suggested by Spitzer and Semo (2002) and Shapley and Enroth-Cugell (1984) and psychophysically measured by Yu et al. (2001). Although the surround region Ψ is centred in the same point, it does not overlap with the inner region Φ (it includes only $N_\Psi = 144$ points).

A study by Nachmias and Sansbury (1974) on how contrast masking varied with mask contrast suggested the presence of contrast non-linearities in visual s.f. channels. These contrast non-linearities were modelled with a function similar to the Naka and Rushton (1966) function (which was also successful in reproducing the responses of cortical V1 neurons (Albrecht & Hamilton, 1982; Sclar, Maunsell, & Lennie, 1990; Tolhurst & Heeger, 1997)). Since we are attempting to model the influence of surrounding image features on the perception of a central stimulus (in ways that are related to grating contrast masking) we defined a similar non-linearity which provides an acceleration at sub-threshold contrast levels and a compression at supra-threshold:

$$z_{\text{ctr}} = \frac{r^2}{1 + r^2}, \tag{4}$$

where z_{ctr} is non-linear and fulfils $0 \leqslant z_{\text{ctr}}(x,y;s,o) \leqslant 1$. The previous expression can be seen as a non-linearisation of the r variable.

As seen in the previous section, local contrast of a central test feature decreases as the contrast of its surround features increases and vice versa (see Fig. 4). Since r is an estimation of the central feature contrast relative to its surround contrast, z_{ctr} in Eq. (4) it can be interpreted as a non-linear estimation of the degree of brightness contrast induced by the surround contrast into a central feature.

To introduce the effect of surround contrast features into the $C_d(\dot{s})$ we use the variable $z_{\text{ctr}}(x,y;s,o)$ where s represents the s.f. and o is the orientation involved. A new CSF C' can be written as follows:

$$C'(\dot{s}, z_{\text{ctr}}) = z_{\text{ctr}} \cdot C_d(\dot{s}) + C_{\min}(\dot{s}). \tag{5}$$

In this expression, $C'(\dot{s}, z_{\text{ctr}})$ reaches its minimum when $z_{\text{ctr}} = 0$ (i.e. minimum brightness contrast or maximum brightness assimilation). To avoid $C'(\dot{s}, z_{\text{ctr}})$ becoming null for some spatial frequencies s (mainly for low spatial frequencies—see Fig. 5) we have introduced the term $C_{\min}(\dot{s})$ in Eq. (5), defining

$$C_d(\dot{s}) = \begin{cases} \exp\left\{-\frac{\dot{s}^2}{2\sigma_1^2}\right\}, & \dot{s} \equiv \dot{s} - s_{\text{thr}} \leqslant 0, \\ \exp\left\{-\frac{\dot{s}^2}{2\sigma_2^2}\right\}, & \dot{s} \equiv \dot{s} - s_{\text{thr}} > 0. \end{cases} \tag{6}$$

where parameters σ_1 and σ_2 are the standard deviation of the piecewise Gaussian function for $s \leqslant s_{\text{thr}}$ and $s_{\text{thr}} < s$, respectively. To reproduce the approximate profile of the psychophysical $C(s)$ functions obtained from the literature (Mullen, 1985), we made $\sigma_2 = 2\sigma_1$ and $\sigma_1 = 2$. We defined $C_{\min}(\dot{s})$ as

$$C_{\min}(\dot{s}) = \begin{cases} \frac{1}{2}\exp\left\{-\frac{\dot{s}^2}{2\sigma_1^2}\right\}, & \dot{s} \equiv \dot{s} - s_{\text{thr}} \leqslant 0. \\ \frac{1}{2}, & \dot{s} \equiv \dot{s} - s_{\text{thr}} > 0 \end{cases} \tag{7}$$

In this way, $C'(\dot{s}, z_{\text{ctr}})$ tends to $C_{\min}(\dot{s})$ when z_{ctr} tends to 0. This avoids a high degree of assimilation being performed at low s.f. (i.e. large scale features), which would make

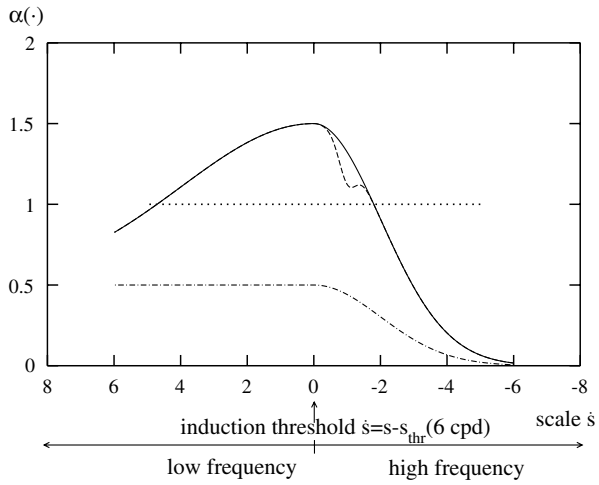

Fig. 5. *Continuous function*: contrast sensitivity function. *Dashed function*: profile of CSF $C'(\dot{s}; z_{\text{ctr}}(\cdot))$ with $z_{\text{ctr}}(x,y;s,o) = 0.75$. That is, when applying $z_{\text{ctr}}(\cdot) = 0.75$ just on a particular s wavelet plane which fulfils $\dot{s} \equiv s - s_{\text{thr}} = -1$, see text for details. *Dashed-dotted function*: profile of $C_{\min}(\dot{s})$. *Dotted line*: values above this value implies brightness contrast, and values below it implies brightness assimilation.

them invisible. We show this function $C_d(\dot{s})$ in Fig. 5. In the opposite situation, $C'(\dot{s}, z_{\text{ctr}})$ is maximum when $z_{\text{ctr}} = 1$. Spitzer and Semo (2002) estimate that the maximum enhancement factor, i.e. maximum perceived brightness contrast, is around 1.5. It is the peak value of $C'(\dot{s}, z_{\text{ctr}})$ (see Fig. 5).

Another important property of the $C'(\dot{s}, z_{\text{ctr}})$ is that it reproduces the dip function for grating adaptation and masking effects. A wavelet basis is usually represented in the s.f. plane by a Heinsenberg box with certain bandwidth in both space and frequency. The function that defines the spread's influence is defined by averaging the Wigner–Wille distribution (Mallat, 1998, Chapter 4), which can itself be approximated by a Gaussian function. In order to show what is the effect on the continuous $C'(\cdot)$ function when weighting down an individual wavelet coefficient from a particular discrete scale, we multiplied a Gaussian function with one octave frequency bandwidth by this weight (see dashed line in Fig. 5). To construct this function we defined a certain observation distance d and obtained the corresponding s_{thr} value (see Appendix A). Following this, we found the particular wavelet plane s that fulfils $(\dot{s} = s_{\text{thr}} - s = -1)$. The dashed line in Fig. 5 shows how the $C'(\cdot)$ is modified when we force $z_{\text{ctr}} < 1$ for a certain feature belonging to this particular wavelet plane. The resulting plot is very similar to the dip function obtained by Graham and Nachmias (1971) and Nachmias and Sansbury (1974) for grating adaptation and masking effects, using a mathematical expression qualitatively equivalent the ours (Eq. (4)).

Eq. (3) shows the general expression to recover a perceptual image I_{percep} represented by a set of wavelet planes ω_s^o. Replacing the set of weights α by our own weighted CSF C'

$$\alpha(s,\cdot) \equiv C'(\dot{s}, z_{\text{ctr}}(\cdot)). \tag{8}$$

We obtain

$$I_{\text{percep}}(x,y) = \sum_{s=1}^{n} \sum_{o=v,h,d} C'(\dot{s}, z_{\text{ctr}}(x,y;s,o)) \cdot \omega_s^o(x,y) + c_n(x,y) \tag{9}$$

which defines the perceptual image recovered from the wavelet components of the original image.

3. Model predictions

This section shows the model's predictions (both quantitatively and qualitatively) for the brightness induction effects mentioned in Section 1 (e.g. the simultaneous brightness contrast (SBC), the White effect (W), the grating induction (GI), the Todorović effect (T), the Mach bands (MB), the Chevreul effect (C), the Adelson–Logvinenko tile, the dungeon illusion and the checkerboard illusion).

To be able to compare BIWaM predictions to psychophysical data, we adjusted the input image and the observers distance parameter to be consistent with the physical dimensions (size, visual angle, observation distance, etc.) reported by the experimenters for their actual stimuli. For the SBC, GI, W and T effects, we compared our model to published psychophysical results (Blakeslee & McCourt, 1999) supplied by McCourt. For the MB and the C effect we obtained data from Lu and Sperling (1996) (Table 2).

3.1. SBC and White effect

Panels (a) and (c) in Fig. 6 show two versions of the SBC effect. The grey rectangle is seen darker when it is in front of a bright background, and brighter when it is in front of a dark background. For this example, we used the same stimulus geometry and observation distance as Blakeslee and McCourt (1999), and obtained a value of $s_{\text{thr}} = 2.65$ for our model.

The continuous line in Fig. 6 (panels (b) and (d)) shows the luminance profile of the central row from panels (a) and (c) which contains the grey patch surrounded by the light/dark uniform background. The dashed lines in the plots show the perceptual profile predicted by our model. We see that the BIWaM predicts the increase/decrease of the perceived brightness of this grey patch over its original value. The operation of the BIWaM can be summarised as follows. Consider the grey patch at the left of panel (a), which is relatively well represented in the wavelet plane that best corresponds to its s.f. and orientation. Since the grey patch is not surrounded by similar features (or it may be said that is surrounded by similar features with zero contrast), brightness contrast is induced at this particular spatial scale. Given that the grey square is darker than its local surround, it becomes even darker (perceptually). The same reasoning can be applied to the other grey patch upon the darker background, which becomes perceptually lighter (and the bottom panels in Fig. 6). The model also predicts a local maxima and minima running parallel to

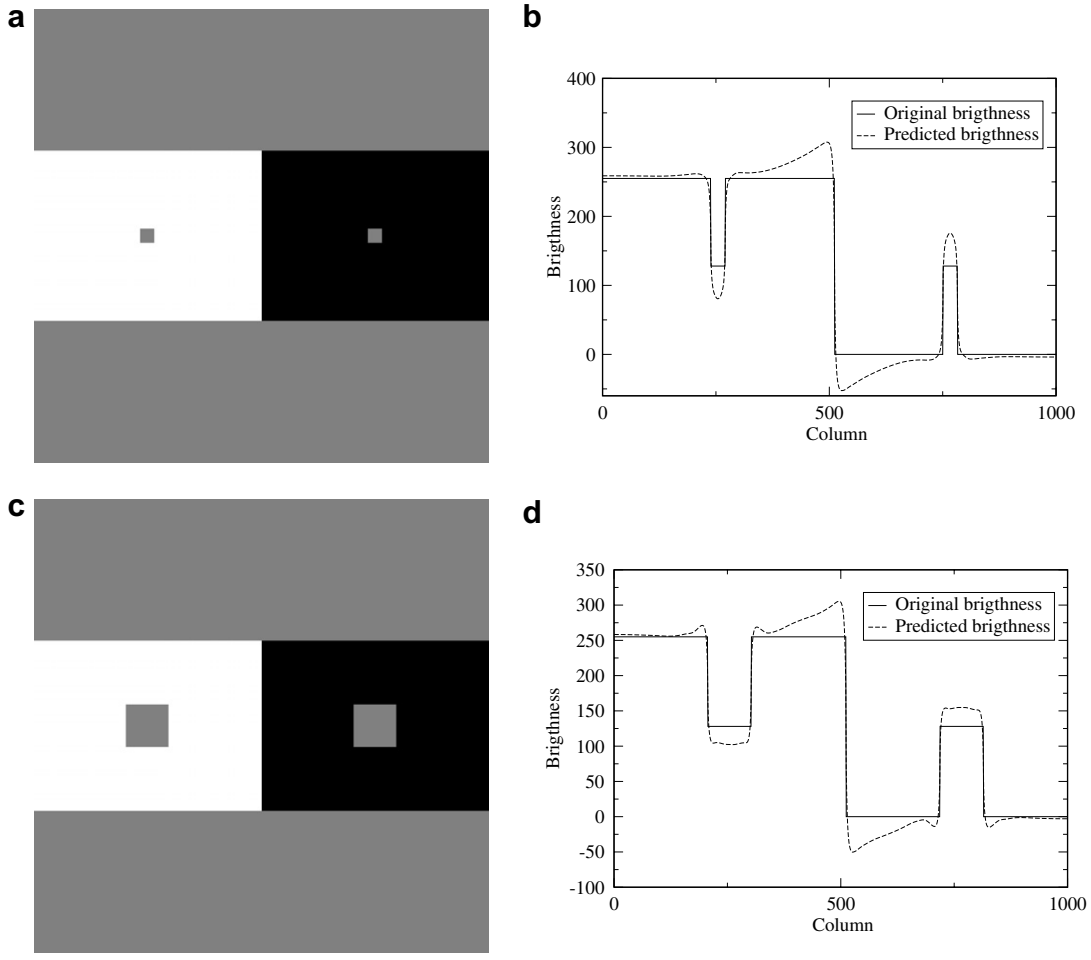

Fig. 6. Panels (a) and (c) show an example of the simultaneous contrast effect and panels (b) and (d) show our model's results. The solid lines in the plots on the right show the profile of the central row taken from the corresponding left panel. The dashed lines show our model's predicted brightness profile (darker left grey patch and a brighter right grey patch).

the vertical light/dark edge at the centre of figures (a) and (c), i.e. around column 500. In this image, this *crisping* effect is not perceived by the observer. This edge and the two black and white plateaus are defined by several spatial frequency components. The BIWaM considers some of these frequency components (mainly the highest s.f.), as isolated features of a given spatial frequency, i.e. brightness contrast is induced. It implies that different weighting factors have been applied to the different spatial frequency components that define both the edge and the plateaus, thus the final brightness profile predicted by the BIWaM in this edge is not psychophysically perceived by the observer. In some cases, this *crisping* effect can be perceived by the observer. For example, on the edge between the wide white rectangle and the grey background, we perceive a darker zone where the grey background is closer to the white patch, and similarly we perceive a brighter zone where the grey background is closer to the dark rectangle.

Another extensively studied effect (shown in Fig. 7 panels (a) and (c), where the left grey rectangle is perceptually brighter than the right one in panel (a) and darker in panel (c)) is the White effect. This effect is generally considered a particular case of SBC (Moulden & Kingdom, 1991; Zaidi, 1989) and can be explained using spatially-oriented filters (Blakeslee & McCourt, 1999; Blakeslee et al., 2005).

It has been suggested (Blakeslee et al., 2005) that the White effect is not an assimilation effect because the direction of brightness change is kept, even when the height of the test patch is reduced so that it has more border contact with the bar on which it is situated. In our formulation, the White effect can be modelled as a spatially-oriented brightness assimilation.

The continuous lines in Fig. 7 (panels (b) and (d)) show the luminance profile of a row from panels (a) and (c), respectively, containing the grey patches. The dashed line in the same figure shows the perceptual profile predicted by the BIWaM. We see that our method correctly predicts that left grey patches are perceived darker and right patches brighter.

As a comparison, we have added in Fig. 7 the psychophysical brightness values obtained by Blakeslee and McCourt (1999) for the same image. We can see that the brightness predicted by the present method fits the psychophysical data well.

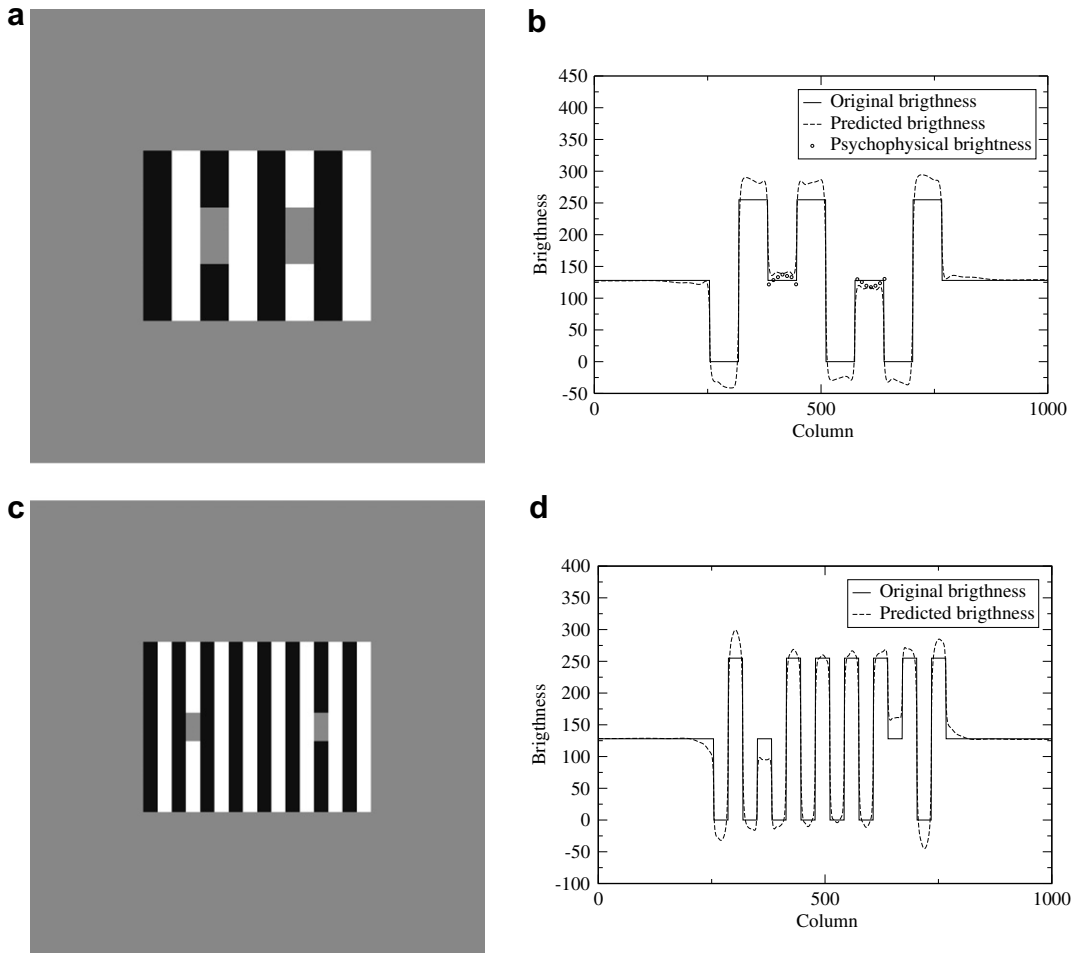

Fig. 7. Panels (a) and (c) show an example of the White effect, where the grey bars are equal but perceived with different brightness because of their different surrounds. Panels (b) and (d) show the profile of a row from (a) and (c), respectively, containing the grey patches. The dashed lines are the model's predictions, showing a perceptual brightness increase of the left grey patch and a decrease on the right patch. Similarly, perception of the vertical stripes is modified (i.e. brightness increase of white vertical stripes and brightness decrease of black vertical stripes). The same effects can be observed in panels (c) and (d).

Our model reproduces the White effect in a similar way. Consider for example, the patch on the right side of panel (a) (i.e. the one in front of white vertical stripes and in lateral contact with dark vertical stripes). There is a wavelet plane where this patch is well represented given its particular s.f. and orientation. This is the same wavelet plane where the background grating is also best represented, since it shares the same horizontal s.f. Since here the surround contrast is greater than the local contrast, brightness assimilation is induced on the grey patch (it becomes perceptually darker). On wavelet planes corresponding to different orientations (e.g. vertical) the opposite interaction may occur, since at these orientations the background is not best represented in the same wavelet plane as the grey patch. As a result, horizontal features induce brightness assimilation and vertical features induce brightness contrast. The total perceived brightness is a combination of these orientation-dependent interactions.

Fig. 7 (panels (b) and (d)) also shows that the model's prediction for bright vertical stripes is brighter that the original value and similarly, the prediction for dark vertical stripes is darker. The reason for this may be the brightness contrast induced by the surrounding (low s.f.) plain grey background on the vertical stripes. Fig. 8 illustrates this effect, where the original images (grey background) are represented on the left while a version of the same images on black background is on the right. The presence of a black background induces a brightness contrast effect on the vertical stripes, making the perceived light bars lighter and the perceived dark bars darker.

3.2. Grating induction

Grating induction (GI) (McCourt, 1982) produces a perceived brightness variation (a grating) on an spatially extended test field, see Fig. 9(a). The central thin horizontal test patch has constant luminance, but its brightness is perceived as an horizontal sinusoidal in counterphase with the upper and lower sinusoidal extended patches. As shown by Blakeslee and McCourt (1997) (who modelled it with their

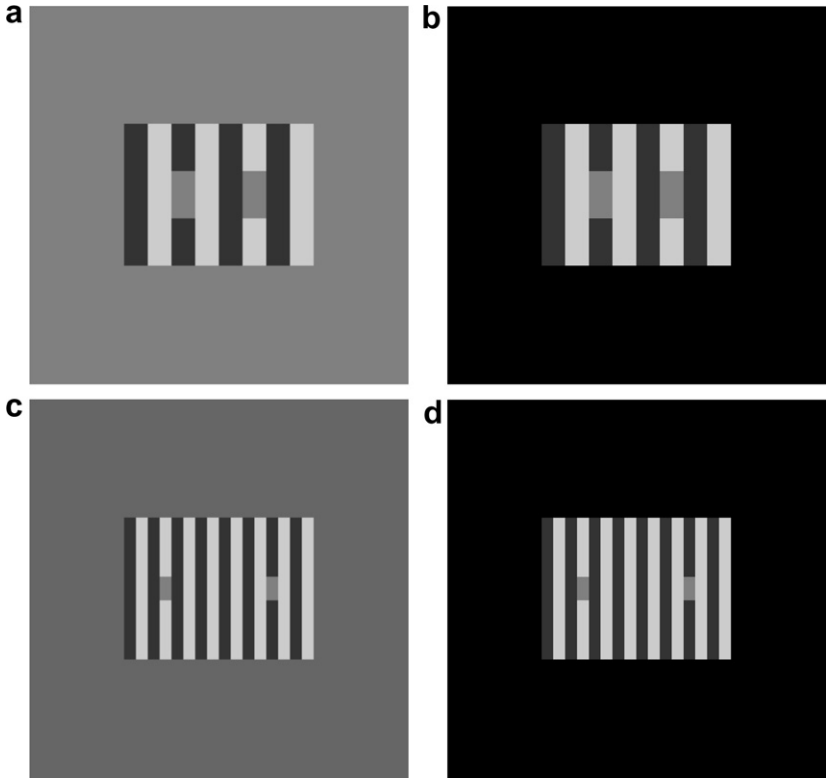

Fig. 8. Influence of a dark background on the White effect. Images (b) and (d) are the same as (a) and (c) but surrounded by a dark background. The perceived brightness of vertical dark stripes in dark background surroundings is different than in the case of the grey background.

ODOG model), this effect may be interpreted as a particular case of brightness contrast.

Fig. 9(b) shows the profile of the central row of both the uniform brightness of the central thin horizontal patch, the brightness profile predicted by our method and a row of the extended sinusoidal luminance grating. As we can see, the perceived brightness of the central grey stripe is predicted as a sinusoidal brightness profile in counterphase with the extended patches. Panels (c) and (d) show a similar example for a lower s.f. sinusoidal patch.

In the GI example, the situation is similar to the simultaneous contrast: since the horizontal grey patch is orthogonal to the grating, it is not well represented in the same vertically-oriented wavelets planes as the grating and therefore a contrast effect is induced. The overall result is a sinusoidal brightness grating in counterphase with the sinusoidal luminance grating.

3.3. Todorović effect

In Fig. 10, panel (a) we show a version of the Todorović effect (Todorović, 1997). This image is the same as in Fig. 6 panel (c), except for the superimposed black and white squares which make the grey patches take the form of a cross (bordered by equal amounts of black and white). The test patch on the black background appears brighter than the test patch on the white background despite the fact that both patches have the same amount of black and white border contact.

In Fig. 10, panel (b) we show the brightness predicted by our model. It correctly predicts that the grey patch on the white background is perceived darker than the grey patch on the black background. In this example, the grey patches do not share features with the rest of the figure within the same orientation and spatial-scale wavelet planes, and therefore brightness contrast is induced. All other squares do share some similarities and therefore have a tendency to be assimilated.

3.4. Mach bands

In the Mach bands effect (Mach, 1865), see Fig. 11(a), bright and dark bands are perceived near the brighter and darker border, respectively, of a ramp edge between two uniform regions of different luminance.

Panel (a) in Fig. 11 shows two plateaus of different luminance with a wedge between them. A brighter vertical cusp is perceived where the wedge meets the brighter plateau and a darker vertical cusp is perceived where the wedge meets the dark plateau. Our model correctly predicts this behaviour, as shown in panel (b). The same panel also shows the luminance profile for the central row of both the original image and the perceptual brightness predicted by our method. The Mach bands are reproduced at the region between the central wedge and the lateral plateaus.

In this case the s.f. features defined by the edges of the wedge and the two plateaus will be prominent in a partic-

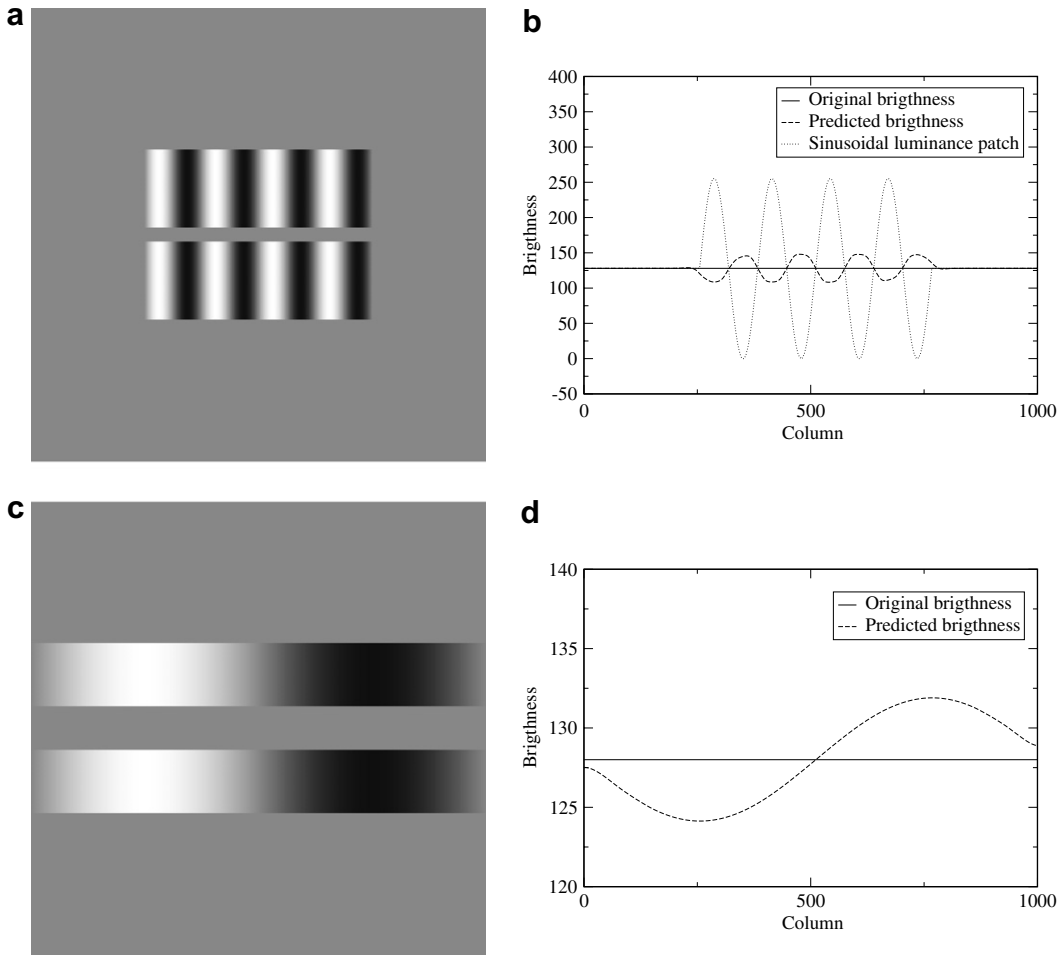

Fig. 9. Panel (a) in the GI effect, the thin horizontal stripe with constant luminance between the horizontal sinusoidal gratings is perceived as a sinusoidal brightness stripe in counterphase with the grating. Panel (b) profile of a row of panel (a) showing the constant luminance of the horizontal stripe (continuous line) and the brightness predicted by our method (dashed function) in counterphase with a row of the horizontal sinusoidal luminance grating (dotted function). A similar effect is shown in panels (c) and (d).

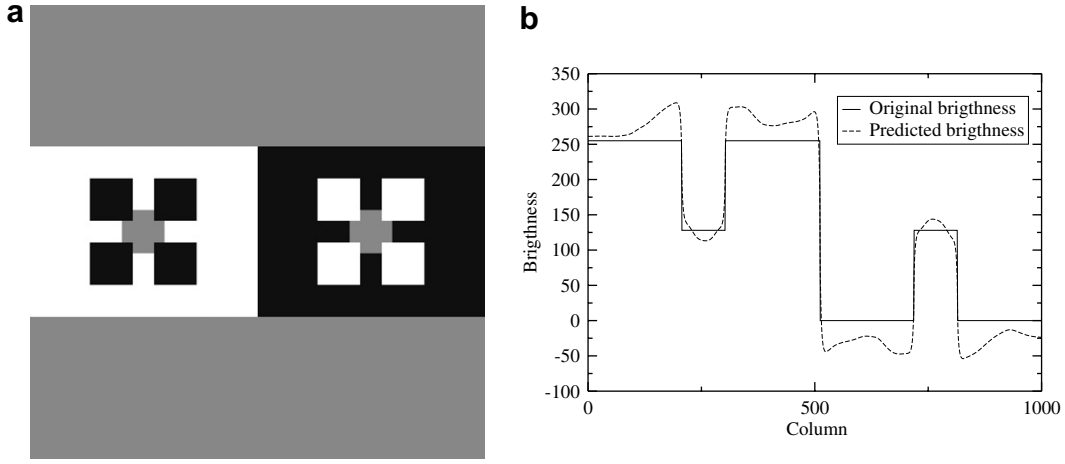

Fig. 10. Panel (a) in the Todorović effect (Todorović, 1997) the left grey patch is perceived darker that right one, even when they have the same amount of border contact with black and white surfaces. Panel (b) predicted brightness from our model.

ular group of wavelet planes of the optimal spatial scale and orientation. There will be no other such prominent feature in the same wavelet planes and this will determine a brightness contrast induction, producing the cusps.

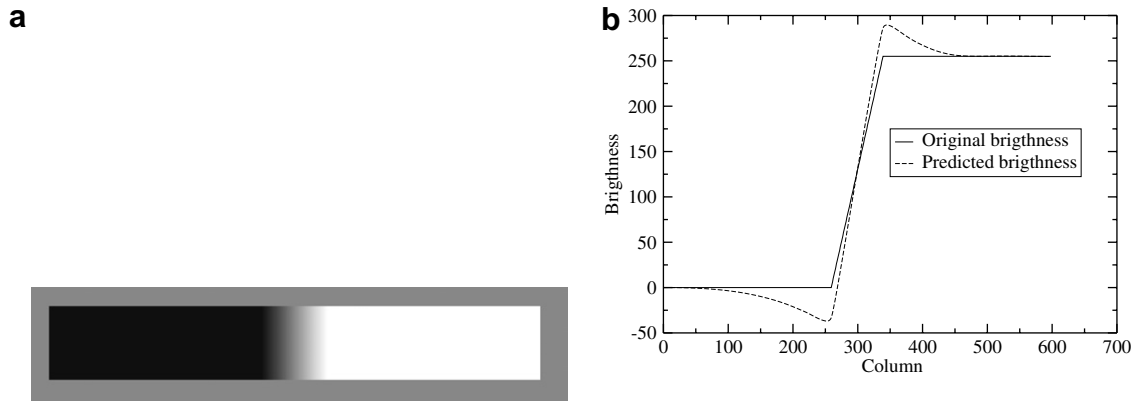

Fig. 11. Panel (a) example of the Mach bands effect. Panel (b) original data luminance (solid lines) and perceptual brightness predicted by our model (dashed line).

3.5. Chevreul effect

In the Chevreul effect (see Fig. 12), a series of stripes with staircase profile is perceived as a sawtooth, that is, each stripe is perceived with a brightness increasing regularly from one stripe to the next. This effect has been modelled with various accuracy levels (Keil, 2006; Morrone & Burr, 1988; Morrone et al., 1994; McArthur & Moulden, 1999).

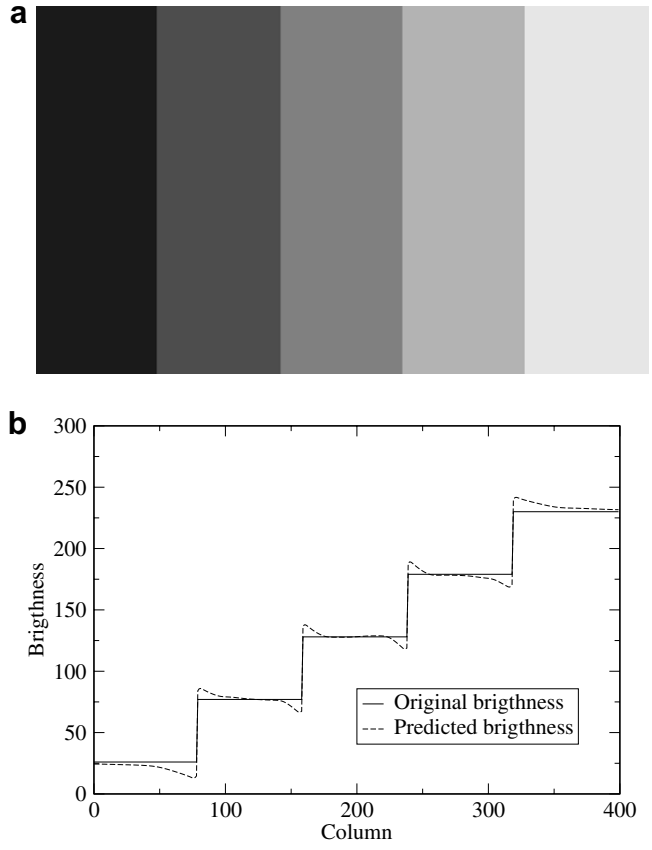

Fig. 12. Panel (a) example of the Chevreul effect. The original image has a luminance staircase profile, but it is perceived with a sawtooth profile. Panel (b) shows the original data values and the perceptual brightness predicted by our model.

In Fig. 12, panel (b) we show the original staircase profile (continuous function) and the brightness predicted by our method (dashed function), which follows an approximately sawtooth profile.

The luminance step between two patches is outlined by several s.f. components (mainly high s.f. components) that will feature highly in a group of wavelet planes of the optimal spatial scale and horizontal orientation. Since at these particularly high spatial frequencies they are not surrounded by similar components (the size of the steps make interactions between edges weak), brightness contrast is induced in the edges, leading to the final sawtooth profile.

3.6. Adelson–Logvinenko tile

In Fig. 13, panel (a) we show a version of the Adelson–Logvinenko tile pattern (Logvinenko, 1999). This image consists of a 2D representation of several 3D cubes modulated by a horizontal sinusoidal grating, where the upper (or lower, depending on how the observer solves the cube's ambiguity) sides of the cubes have equal luminance but are perceived with different brightness. The grey level value of these top surfaces is 134. Our method predicts a grey level value of 134 for the apparently light surfaces and 99 for the apparently dark ones.

The presence of the vertical (low s.f.) modulating sinusoidal grating means that there will be a particular vertically-oriented wavelet plane where this feature will be represented best with very little influence of the rest of the image. This will induce a strong brightness contrast effect in the rows that are coincident with the peaks and valleys of this sinusoidal (precisely where the tops of the cubes are located), therefore producing the final perceived effect.

3.7. Dungeon illusion

There is a subset of illusions where the direction of contrast does not fit the one predicted by traditional contrast theories (Gilchrist, 2006). One of these is the dungeon illu-

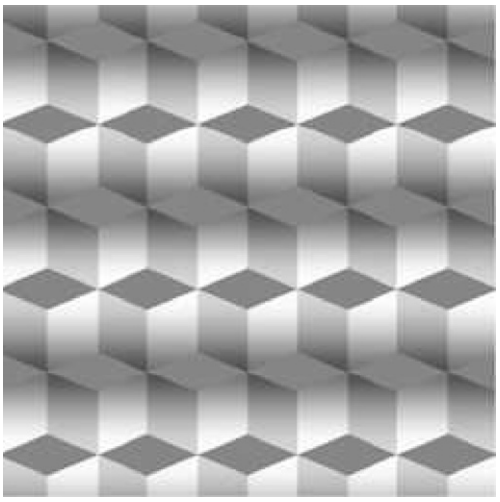

Fig. 13. Panel (a) example of the Adelson–Logvinenko tile. The parts of the cubes at the crests and valleys of the modulating sinusoidal have all the same luminance but are perceived differently. The model we present qualitatively predicts these differences (see text).

sion (Fig. 14, Bressan, 2001) where the perceived difference between the grey squares on the left and the right of the picture is the opposite of what one would predict from analysis of individual squares and their immediate surroundings (Gilchrist, 2006). Panel (a) in Fig. 14 shows an example of the dungeon illusion and panel (b) shows the brightness predicted by our model for the central row, where all grey rectangles on the left side are represented with a darker shade of grey than those on the right side of the image. This effect can be qualitatively explained by the fact that all grey rectangles are surrounded by other rectangles of the same size and at a distance similar to its size, both vertically, horizontally and diagonally. Therefore, they will be strongly represented in the same wavelet planes, leading to a brightness assimilation effect. This in turn will produce darker rectangles in the left side of the image (where the rectangles' brightness will tend towards that of the bars) and slightly brighter rectangles on the right side.

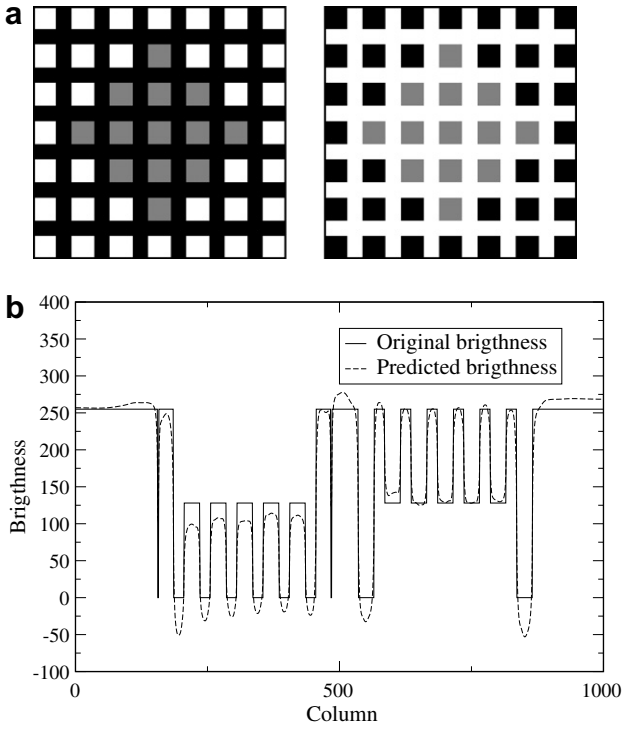

Fig. 14. Panel (a) shows an example of the dungeon effect. Individual grey squares on the left side are completely surrounded by black pixels and should be seen lighter than individual grey squares on the right side, which are in turn surrounded by white pixels, in practice the opposite occurs. Panel (b) shows our model's predictions for the central row of the figure in panel (a), demonstrating the power of a multiresolution wavelet approach to provide a qualitative explanation of the effect.

3.8. Checkerboard

Another example of a complete reversal of contrast is the checkerboard illusion (shown in panel (a) of Fig. 15). Here, the grey square in contact with white squares is perceived brighter than the grey square in contact with the black squares (an effect similar to the dungeon illusion). A simplified explanation can be given in terms of the features that surround each of the squares, since the grey squares are horizontally and vertically surrounded by elements of equal size and high contrast, they will again be represented in the same spatial scales and orientations wavelet planes which will induce brightness assimilation on them. If the square is surrounded by black squares (left), its brightness will tend to go in the direction of the local surroundings (darker). The other grey square will be assimilated towards the other end (it will look brighter).

4. Discussion: Comparison with psychophysics

To make a more quantitative assessment of our model's predictions, we tested our model against psychophysical measures from the literature of the relative brightness increase/decrease produced by brightness induction. We simulated the physical conditions (image size and observer's distance) in our model and produced a set of predictions that were compared to the measurements. All other parameters were kept the same for all conditions. Figs. 16 and 17 show the psychophysically-measured values published by Blakeslee and McCourt (1999) and our predicted values for some of the visual effects described above (e.g. simultaneous brightness contrast, grating induction, White effect and Todorović effect).

In Fig. 16, experimental values (and their associated 95% confidence limit error bars) are represented by bars. Our predicted values are represented by squares (empty squares for test patches on dark background, and solid squares for test patches on bright background). The ordinate axis shows the difference between the matching luminance and the mean luminance expressed as a proportion of the mean luminance, consistent with Blakeslee and

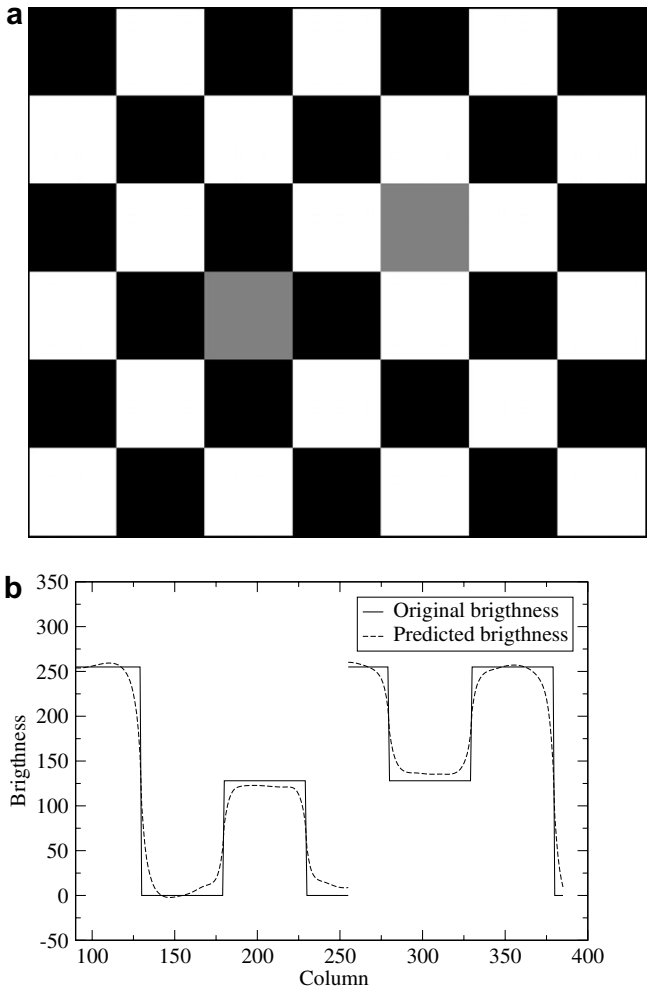

Fig. 15. Panel (a) shows an example of the checkerboard contrast effect. The image shows two grey squares (of the same luminance) embedded in a checkerboard, that are perceived differently by the observer. The right square is perceived brighter than the left one. Panel (b) shows the real and perceptual brightness profiles of these two squares as predicted by our model.

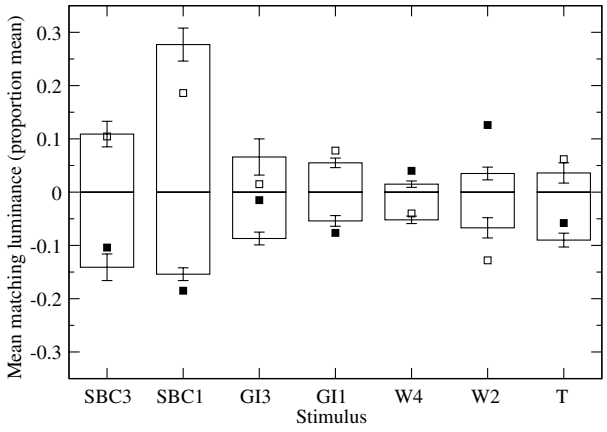

Fig. 16. Representation of the values predicted by our model (empty squares for patches on dark background and filled squares for patches on bright background) and the psychophysical values obtained by the BB observer in Blakeslee and McCourt (1999) (bars) for simultaneous brightness contrast (SBC), grating induction (GI), White effect (W) and Todorović effect (T).

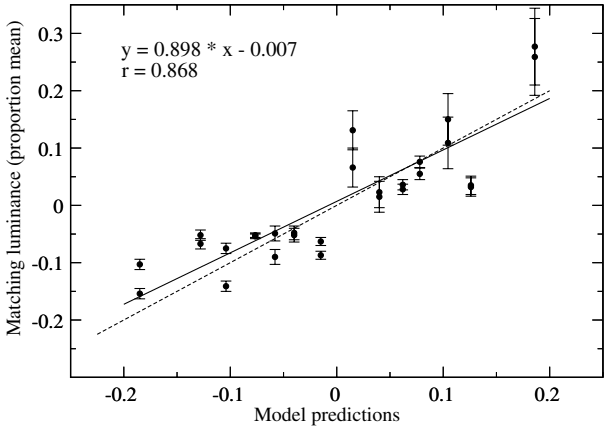

Fig. 17. Representation of the values predicted by our model (abcissa axis) and the psychophysical values obtained by Blakeslee and McCourt (1999) (ordinate axis) for all the visual effects considered in Fig. 16.

McCourt (1999). The 0.0 value represents the luminance of the test patches. We can see that our method approximately predicts the direction and magnitude of the brightness induction. The greatest deviation from the psychophysical values is for the GI3 (our model underestimates the effect) and W2 (our model overestimates the effect) results. It is possible to obtain better fits by adding more degrees of freedom (e.g. modifying the CSF C' to adjust for these two sets of results) to our model, but for the sake of consistency and simplicity, we prefer to keep the lowest number of degrees of freedom (and the simplest mathematical expression) in all cases.

In Fig. 17 we show a plot of the values predicted by our model (abscissa) versus the psychophysical values (ordinate) for all considered visual effects. Each point in the plot represents an observer (either MM or BB) from Blakeslee and McCourt (1999). We also show the diagonal line (dotted line) where all points would lie should our model's predictions be 100% accurate. The points show an approximately linear behaviour. The solid line represents the best fitting line (linear regression) with a slope of around 0.9, a correlation coefficient $r = 0.87$, and the sum of squares of the residuals being 0.077.

We also used psychophysical data by Lu and Sperling (1996) in order to analyse the predictions of our model for the Mach bands and the Chevreul effect. As with the previous examples, we simulated in our model the physical conditions of Lu and Sperling (1996). In the case of the Mach bands effect, these authors report a brightness increment/decrement (in the bright and dark plateau, respectively) which is around 8% of the luminance difference between the bright and the dark plateau. Our model also predicts a brightness increment around 13%. For the Chevreul effect, these authors report an increment/decrement of the brightness around 54% of the luminance difference between consecutive steps. Our model predicts an

increment/decrement around 17%, which differs from the observed value.

Plots in Figs. 16 and 17 (comparison with psychophysical data) show that despite its simplicity, our wavelet model is capable of predicting both the direction and magnitude (except for the Chevreul effect) of the psychophysical data. The analysis presented in the previous section also shows a qualitative agreement between the model's prediction and the spatial distribution of the brightness changes of the observed phenomena. It is also worth pointing out that a simple explanation based on common features in terms of spatial scale, spatial orientation and surround contrast provides guidelines as to how to interpret these effects.

However, BIWaM has been unable to reproduce some other brightness induction effects, such as the Wertheimer–Benary cross (Benary, 1924) (our method cannot predict different brightness values for the two grey triangles) and the shadow-incompatible luminance gradient (Logvinenko, 2003) (where the BIWaM predicts different brightness values for the equally perceived test patches). Another brightness induction effect that our model cannot reproduce is the glare effect (Zavagno, 1999; Zavagno & Caputo, 2001; Zavagno & Caputo, 2005) (a white patch is perceived lighter when surrounded by a smooth luminance profile).

Since the performance of our model when considering these effects (and many others) has not been included in Fig. 17, we acknowledge our model's "goodness" criteria is biased and BIWaM can provide only a partial explanation to whole range of brightness induction effects studied in the literature. Even more, our model is far better at explaining Hering's type of simultaneous brightness contrast effects than it is at explaining Helmoltz's (see (Logvinenko & Kane, 2004) for a review and classification of these effects).

5. Conclusions

Our simplistic model of visual brightness induction is based on three main features of visual scenes: spatial scale, spatial orientation and surround contrast. We selected these not only because there is evidence that they are highly relevant to brightness perception phenomena, but also because there is evidence (both psychophysical and physiological) that these attributes are processed in parallel by pre-cortical and cortical semi-independent channels. In our framework, we assume that brightness induction is performed mostly between features of similar s.f. and spatial orientation (i.e. within the same wavelet plane) and the effect is also dependent on the contrast of surround features compared to the central test feature for each spatial scale and orientation. The model also makes use of a psychophysically determined contrast sensitivity function and explicitly includes the observation distance to be able to relate the different spatial scales (wavelet planes) to the actual world. This simple set of assumptions allow the unification of brightness assimilation and brightness contrast in a single mathematical framework, and to reproduce (qualitatively in all cases, quantitatively in some) several known brightness induction effects, e.g. simultaneous brightness contrast, White effect, grating induction, Todorović effect, Mach bands, Chevreul effect, Adelson–Logvinenko tile, dungeon illusion and checkerboard illusion, which were not previously explained by a single (unified) framework, using a unique set of parameters.

Our most important contribution is to show that many brightness induction effects can be modelled and reproduced using only three assumptions (described in Section 2.3). A secondary contribution is the incorporation of Assumption 3, which states that when the brightness contrast of the surround features increases, brightness assimilation increases, i.e. brightness contrast decreases, and vice versa. This assumption is the key point which allows, using a single unified mathematical formulation, the model to perform either brightness assimilation or brightness contrast depending on the centre/surround spatial brightness distribution and the observers distance to the stimuli. A final contribution is to produce a model which can simultaneously reproduce brightness induction effects without the need to adjust its parameters for each particular case, in a manner consistent with the behaviour of the human visual system.

The extension of this model into colour opponent space is straightforward and so far it has proven capable of reproducing some chromatic induction effects.

Acknowledgments

Authors wish to thank M.E. McCourt for kindly supplying psychophysical data. Xavier Otazu and Alejandro Párraga were funded by the Ramon y Cajal program. This work has been partially supported by projects TIN2004-02970 and TIN2007-64577 of Spanish MEC (Ministery of Science). Thanks to Matthias S. Keil, George Lovell and two anonymous reviewers for their very helpful comments.

Appendix A. Induction threshold s_{thr}

Let d be the distance from the observer to the stimulus (printout, computer screen, etc.) If a given feature subtends a visual angle β when observed from distance d, the feature's size l is

$$l = d \cdot \tan \beta. \tag{A.1}$$

This projection is measured on the image space as a spatial measure, which in turn can be related to a period, i.e. a cycle, at a given spatial frequency. By definition, a scale s is related to a certain frequency $v(s)$, i.e. to a period $T = 1/v(s)$. This relation is defined by $2^s = T = l/l_p$, where l/l_p is the number of pixels into one frequency period T, and l_p is the image pixel size.

If instead of 1 cycle, i.e. 1 period, we want to include 4 cycles of a certain spatial scale in the same longitude, we can define

$$4T \equiv 4 \cdot 2^{s_{\text{thr}}} = \frac{l}{l_p}. \tag{A.2}$$

being s_{thr} this particular scale. If we take the particular case of a feature that show 1 visual degree when observed at distance d, using Eq. (A.1) into Eq. (A.2) we obtain

$$s_{\text{thr}} = \log_2\left(\frac{d \cdot \tan 1^\circ}{4 \cdot l_p}\right). \tag{A.3}$$

The s_{thr} factor is the image scale associated to the $v_{\text{thr}} = 4$ cpd induction threshold value when observing an image with a pixel size l_p from a distance d.

References

Adelson, E. (1993). Perceptual organization and the judgment of brightness. *Science, 262*(5142), 2042–2044.

Adelson, E. (2000). The new cognitive neurosciences. In *Lightness perception and lightness illusions* (2nd ed., pp. 339–351). Cambridge, MA, USA: MIT Press.

Agostini, T., & Galmonte, A. (2002). Perceptual organization overcomes the effects of local surround in determining simultaneous lightness contrast. *Psychological Science, 13*(1), 89–93.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology, 48*(1), 217–237.

Anderson, B. L. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception, 26*(4), 419–453.

Benary, W. (1924). Beobachtungen zu einem experiment uber helligkeitskontrast. *Psychologische Forschung, 5*, 131142.

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

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

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

Blakeslee, B., & McCourt, M. E. (1997). Similar mechanisms underlie simultaneous brightness contrast and grating induction. *Vision Research, 37*(20), 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., Pasieka, W., & McCourt, M. (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*(5), 607–615.

Bressan, P. (2001). Explaining lightness illusions. *Perception, 30*, 1031–1046.

Cannon, M. W., & Fullenkamp, S. C. (1991). Spatial interactions in apparent contrast: Inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research, 31*(11), 1985–1998.

Chevreul, M. E. (1890). The principles of harmony and contrast of colors. George Bell and Sons, Bohn's Library.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America, 86*(23), 9631–9635.

Cornsweet, T. N. (1970). Visual perception. San Diego, USA: Academic Press.

Daugmann, J. G. (1980). Two-dimensional spectral analysis of cortical receptive field profile. *Vision Research, 20*, 847–856.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22*(5), 545–559.

De Valois, R. L., & De Valois, K. K. (1988). Spat. vis. New York, USA: Oxford University Press.

du Buff, J. M. H. (1994). Ramp edges, mach bands, and the functional significance of the simple cell assembly. *Biological Cybernetics, 70*(5), 449–461.

D'Zmura, M., & Singer, B. (1998). Color vision, perspectives from different disciplines. In *Color contrast gain control* (pp. 369–385). New York, USA: Walter de Gruyter and Co..

D'Zmura, M., & Singer, B. (1999). Color vision: From genes to perception. In *Contrast gain control* (pp. 369–385). Cambridge University Press.

Ejima, Y., & Takahashi, S. (1985). Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Research, 25*(9), 1223–1232.

Ellemberg, D., Wilkinson, F., Wilson, H. R., & Arsenault, A. S. (1998). Apparent contrast and spatial frequency of local texture elements. *Journal of the Optical Society of America, 15*(7), 1733–1739.

Fach, C., & Sharpe, L. T. (1986). Assimilative hue shifts in color gratings depend on bar width. *Perception & Psychophysics, 40*(7), 412–418.

Field, D. J. (1999). Wavelets, vision and the statistics of natural scenes. *Philosophical Transactions of the Royal Society of London. Series A, 357*(1760), 2527–2542.

Fiorentini, A., Baumgartner, G., Magnussen, S., Schiller, P., & Thomas, J. (1990). Visual perception: The neurophysiological foundations. In *The perception of brightness and darkness: Relations to neuronal receptive fields* (pp. 129–161). San Diego, CA: Academic Press.

Foley, J., & McCourt, M. (1985). Visual grating induction. *Journal of the Optical Society of America A, 2*, 1220–1230.

Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 93*(2), 615–622.

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 106*(4), 795–834.

Gilchrist, A. L. (2006). *Seeing black and white [Oxford psychology series]*. Oxford: Oxford University Press.

Goldstein, E. B. (2002). *Sensation and perception* (6th ed.). Pacific Grove, CA: Wadsworth-Thomson Learning.

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single-channel and multiple-channels models. *Vision Research, 11*(3), 251–259.

Harvey, L. O., Jr., & Doan, V. V. (1990). Visual masking at different polar angles in the two-dimensional fourier plane. *Journal of the Optical Society of America A, 7*(1), 116–127.

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

Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology, 50*(2), 89–96.

Helson, H. (1963). Studies of anomalous contrast and assimilation. *Journal of the Optical Society of America, 53*, 179–184.

Holschneider, M., Kronland-Martinet, R., Morlet, J., & Tchamitchian, P. (1989). A real-time algorithm for signal analysis with the help of the wavelet transform. In J. Combes, A. Grossmann, & P. Tchamitchian (Eds.), *Wavelets, time-frequency methods and phase space* (pp. 286–297). Berlin, New York: Springer-Verlag.

Jones, J. P., & Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology, 58*(6), 1233–1258.

Keil, M. S. (2006). Smooth gradient representations as a unifying account of Chevreul's illusion, mach bands, and a variant of the ehrenstein disk. *Neural Computation, 18*(4), 871–903.

Kingdom, F., & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research, 32*(8), 1565–1582.

Kingdom, F. A. A. (2003). Levels of brightness perception. In *Levels of perception*. New York, London: Springer-Verlag.

Klein, S., Stromeyer, C. F. R., & Ganz, L. (1974). The simultaneous spatial frequency shift: A dissociation between the detection and perception of gratings. *Vision Research, 14*(12), 1421–1432.

Logvinenko, A. D. (1999). Lightness induction revisited. *Perception, 28*(7), 803–816.

Logvinenko, A. D. (2003). A fair test of the effect of a shadow-incompatible luminance gradient on the simultaneous lightness contrast. *Perception, 32*(6), 717–720 (Discussion 721–730).

Logvinenko, A. D., & Kane, J. (2004). Hering's and helmholtz's types of simultaneous lightness contrast. *Journal of Vision, 4*(12), 1102–1110.

Logvinenko, A. D., & Ross, D. A. (2005). Adelson's tile and snake illusions: A helmholtzian type of simultaneous lightness contrast. *Spatial Vision, 18*(1), 25–72.

Lotto, R. B., Williams, S. M., & Purves, D. (1999). An empirical basis for mach bands. *Proceedings of the National Academy of Sciences of the United States of America, 96*(9), 5239–5244.

Lu, Z. L., & Sperling, G. (1996). Second-order illusions: Mach bands, Chevreul, and craik-o'brien-cornsweet. *Vision Research, 36*(4), 559–572.

Mach, E. (1865). Classe kaiserlichen akad. *Wissenschaftsgesch, 52*, 303–322.

MacKay, D. M. (1973). Lateral interaction between neural channels sensitive to texture density? *Nature, 245*(5421), 159–161.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(7), 674–693.

Mallat, S. G. (1998). *A wavelet tour of signal processing*. San Diego: Academic Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

McArthur, J. A., & Moulden, B. (1999). A two-dimensional model of brightness perception based on spatial filtering consistent with retinal processing. *Vision Research, 39*(6), 1199–1219.

McCourt, M. E. (1982). A spatial frequency dependent grating-induction effect. *Vision Research, 22*(1), 119–134.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society Series B-Biological Sciences, 235*(1280), 221–245.

Morrone, M. C., Burr, D. C., & Ross, J. (1994). Illusory brightness step in the Chevreul illusion. *Vision Research, 34*(12), 1567–1574.

Moulden, B., & Kingdom, F. (1991). The local border mechanism in grating induction. *Vision Research, 31*(11), 1999–2008.

Mullen, K. T. (1985). The contrast sensitivity of human color vision to red-green and blue-yellow chromatic gratings. *Journal of Physiology, 359*, 381–400.

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast discrimination may be better than detection. *Vision Research, 14*, 1039–1042.

Naka, K. I., & Rushton, W. A. (1966). S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology, 185*(3), 587–599.

Olshausen, B. A., & Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by v1? *Vision Research, 37*(23), 3311–3325.

Otazu, X., Vanrell, M. (2006). Decimated multiresolution framework with surround induction function to unify assimilation and contrast mechanisms. Tech. Rep., Computer Vision Center. Available from http://www.cvc.uab.es.

Peromaa, T. L., & Laurinen, P. I. (2004). Separation of edge detection and brightness perception. *Vision Research, 44*(16), 1919–1925.

Pessoa, L., Mingolla, E., & Neumann, H. (1995). A contrast- and luminance-driven multiscale network model of brightness perception. *Vision Research, 35*(15), 2201–2223.

Robinson, A. E., Hammon, P. S., & de Sa, V. R. (2007). Explaining brightness illusions using spatial filtering and local response normalisation. *Vision Research, 47*, 1631–1644.

Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science, 273*(5278), 1104–1107.

Schein, S. J., & Desimone, R. (1990). Spectral properties of v4 neurons in the macaque. *The Journal of Neuroscience, 10*(10), 3369–3389.

Sclar, G., Maunsell, J. H., & Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Research, 30*(1), 1–10.

Shapley, R., & Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Progress in retinal research* (Vol. 3). Oxford: Pergamon Press.

Simpson, W. A., & McFadden, S. M. (2005). Spatial frequency channels derived from individual differences. *Vision Research, 45*(21), 2723–2727.

Smith, V. C., Jin, P. Q., & Pokorny, J. (2001). The role of spatial frequency in color induction. *Vision Research, 41*(8), 1007–1021.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research, 33*(18), 2671–2683.

Spillmann, L., & Werner, J. S. (1996). Long-range interactions in visual perception. *Trends in Neurosciences, 19*(10), 428–434.

Spitzer, H., & Semo, S. (2002). Color constancy: A biological model and its application for still and video images. *Pattern Recognition, 35*(8), 1645–1659.

Todorović, D. (1997). Lightness and junctions. *Perception, 26*(4), 379–394.

Tolhurst, D. J. (1972). On the possible existance of edge detector neurones in the human visual system. *Vision Research, 12*(5), 797–804.

Tolhurst, D. J., & Heeger, D. J. (1997). Contrast normalization and a linear model for the directional selectivity of simple cells in cat striate cortex. *Visual Neuroscience, 14*(1), 19–25.

Tolhurst, D. J., & Thompson, I. D. (1982). Organization of neurons preferring similar spatial-frequencies in cat striate cortex. *Experimental Brain Research, 48*(2), 217–227.

Van Rullen, R., & Thorpe, S. J. (2001). Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex. *Neural Computation, 13*(6), 1255–1283.

Vannes, F. L., & Bouman, M. A. (1967). Spatial modulation transfer in human eye. *Journal of the Optical Society of America, 57*(3), 401–406.

Walker, J. T. (1978). Brightness enhancement and the talbot level in stationary gratings. *Perception & Psychophysics, 23*(4), 356–359.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology, 38*(3), 310–324.

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25*(11), 1661–1674.

Werner, A. (2003). The spatial tuning of chromatic adaptation. *Vision Research, 43*(15), 1611–1623.

White, M. (1979). A new effect of pattern on perceived lightness. *Perception, 8*(4), 413–416.

White, M. (1981). The effect of the nature of the surround on the perceived lightness of grey bars within square-wave test gratings. *Perception, 10*(2), 215–230.

Wilson, H. R., McFarlane, D. K., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research, 23*(9), 873–882.

Yu, C., Klein, S. A., & Levi, D. M. (2001). Surround modulation of perceived contrast and the role of brightness induction. *Journal of Vision, 1*(1), 18–31.

Yu, C., Klein, S. A., & Levi, D. M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *Journal of Vision, 2*(3), 243–255.

Yu, C., Klein, S. A., & Levi, D. M. (2003). Cross- and iso-oriented surrounds modulate the contrast response function: The effect of surround contrast. *Journal of Vision, 3*(8), 527–540.

Yund, E. W., & Armington, J. C. (1975). Color and brightness contrast effects as a function of spatial variables. *Vision Research, 15*, 917–929.

Zaidi, Q. (1989). Local and distal factors in visual grating induction. *Vision Research, 29*(6), 691–697.

Zavagno, D. (1999). Some new luminance-gradient effects. *Perception, 28*, 835–838.

Zavagno, D., & Caputo, G. (2001). The glare effect and the perception of luminosity. *Perception, 30*, 209–222.

Zavagno, D., & Caputo, G. (2005). Glowing greys and surface-white: The photo-geometric factors of luminosity perception. *Perception, 34*, 261–274.

Zetzsche, C., & Nuding, U. (2005). Nonlinear and higher-order approaches to the encoding of natural scenes. *Network-Computation in Neural Systems, 16*(2–3), 191–221.