Multiresolution wavelet framework models brightness induction effects.

Xavier Otazu[1], Maria Vanrell[1,2], C. Alejandro Párraga[1]

[1] *Computer Vision Center, Campus UAB, Cerdanyola del Vallès, 08193 Barcelona, Spain*

[2] *Computer Science Dept., Universitat Autònoma de Barcelona, Cerdanyola del Vallès, 08193 Barcelona, Spain*

Abstract

Two types of brightness induction effects (named contrast and assimilation) have been described depending on the interactions between a test stimulus and its surroundings. Brightness contrast describes a shift of the test stimulus brightness away from its surroundings and brightness assimilation describes the opposite (the brightness of the test stimulus shifts towards that of its surroundings). These two effects have been accounted by a single framework (Blakeslee and McCourt, 1999) called ODOG, using multiresolution decompositions and receptors modelled as oriented difference-of-Gaussians. In this work we present a new multiresolution wavelet framework which describes assimilation and contrast effects in a unified formulation, including known psychophysical and physiological attributes of the primate visual system (such as the contrast sensitivity function), its receptive fields (spatial frequency and orientation bandwidths) and the visual cortex (contrast non-linearities). This formulation reproduces known visual effects which were reproduced by previous models (like the ODOG model) such as simultaneous contrast, the White effect, grating induction and Todorovic effect. Furthermore, our framework reproduces these effects and also some additional brightness induction effects such as Mach bands, Chevreul effect, Adelson-Logvinenko snake and some previously unaccounted effects such as the Dungeon illusion in a very intuitive way.

Key words: human visual system, brightness induction, White effect, Mach bands, Chevreul effect, Adelson-Logvinenko snake, wavelet transform

1 Introduction

In visual perception, the term brightness refers to a non-quantitative perception of light elicited by the luminance of a visual target. This (perceived) 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 these changes of appearance due to the surrounding light and its effects are classified according to the perceptual direction of the changes. When the change in brightness of the visual target goes away from the surrounding brightness, it is called *brightness contrast* and when the change goes towards that of the surrounding brightness, it is called *brightness assimilation.*

1.1 Brightness induction effects

One of the oldest known examples of this phenomenon (already mentioned by Alhazen in his book Optics -circa 1025AD-) is called simultaneous brightness contrast (SBC). Fig. 7 shows an example where a grey patch looks darker when placed over a white background than when placed over a black background (Wallack, 1948; Heinmann, 1955). 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 nucleous (LGN) neurons in monkey (DeValois & DeValois, 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; Spillman & Werner, 1996).

A second known example of brightness induction is the so called grating induction (GI) where a spatial variation of luminance (such as a grating) produces a counter-phase brightness variation on an adjacent extended text field (McCourt, 1982) (see Fig. 10). 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 & McCourt, 1985). The induced

Email address: xotazu@cvc.uab.es, maria@cvc.uab.es, aparraga@cvc.uab.es (Xavier Otazu[1], Maria Vanrell[1,2], C. Alejandro Párraga[1]).

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 its physiological basis are related to the discovery of cortical neurons in cat (Rossi et al., 1996) and monkey (Gilbert et al., 1996) that integrate over relatively large distances.

Another well known brightness effect is the White effect (White, 1979) (see Fig. 8), 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 of 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 Todorovic (1997) where the brightness shift seems to be independent of the amount of black or white background in contact with the test patch (see Fig. 11). Several explanations both at the receptor-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 receptoral level needs both elongated cortical filters (White, 1981; Foley & McCourt, 1985) 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.

The Mach bands (named after the physicist Ernst Mach) are brightness maxima and minima perceived at the beginning and end of luminance gradients respectively (Mach, 1865) (see Fig. 12). They have been interpreted in terms of lateral inhibition of retinal ganglion cells (Goldstein, 1996) and more recently as a consequence of the physical properties of real world luminance gradients (Lotto et al., 1999). 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 (see Fig. 13). 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 (Peromaa and Laurinen, 2004; Morrone et al., 1994). However, there are also alternative explanations of this effect in terms of filling-in process triggered by the edges at the different spatial scales (Pessoa et al., 1995).

The Adelson tile illusion (Adelson, 1993), appears when a “wall made from

homogeneous blocks" (see Fig. 14) is spatially modulated by a horizontal dark stripe in such 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. A way of showing the need to incorporate long-ranging receptor field interactions to explain the Adelson tile illusion consists on rearranging the pattern to make the effect disappear while keeping the local contrast around diamonds the same (Adelson, 1993). 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 is a wide variety of explanations for the Adelson illusions, ranging from "low-level" explanations based on local contrast and multiscale spatial filtering (Cornsweet, 1970; Blakeslee and McCourt, 1999) to those based on the role of borders or luminance junctions between or across strips (Adelson, 1993, 2000; Anderson, 1997), "high-level" explanations where the explanation s is based on how the visual system deals with illumination (Gilchrist et al., 1999; Logvinenko & Ross, 2005) and "multi-level" explanations (Kindom, 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 & 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 (DeValois & DeValois, 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).

Current interpretations of these brightness induction effects are either based on low-level sensory processes (sensitive to contrast) or higher level processes involving assumptions about the illumination and the configuration of the visual display as a whole. These two lines are identified with the work of 19th century scientists Hermann (1870) and Helmholtz (1867). The "Hering approach" was concerned with adaptation and local interactions and favoured a connection between physiology and psychophysics (e.g. a Laplacian derivative operator modelled retinal cells with centre-surround opponency). Centre-surround opponency (or lateral inhibition) can go a long way explaining some illusions especially those involving sharp transitions in light intensity (such as the edge in the Craik-O'Brien-Cornsweet illusion (Cornsweet, 1970)). The Helmholtz

approach is based on higher level knowledge of the world. He proposed that our perceptions are "best guesses" of what it is in the world, based on previous experience and low level data. Under this view, brightness induction effects are the results of the whole visual system trying to interpret natural scenes.

In the following section we review several attempts to model brightness induction in computational frameworks.

1.2 Modelling attempts

One of the earliest computational models to explain some of these phenomena was developed by Land & McCann (1971) and called Retinex. It applies a derivative operator to obtain edge information of a given image. This model takes into account the spatial differences between changes in reflectance (which tend to change abruptly at the edges of objects) and illuminance (which tend to change gradually in space) and attempts to remove illuminant variations, thus incorporating higher level knowledge of the statistics of the natural world. Subsequent models of brightness perception have been developed using multiscale approaches to low level vision. Some 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 (Tolhurst, 1972; Morrone & Burr, 1988; Fiorentini et al., 1990; duBuff, 1994). These models may differ in the way these operators interact with each other. For example, both the models of Tolhurst (1972) and Morrone & Burr (1988) employ pairs of orthogonal operators but the first applies mutual inhibition between them while the second pools their responses. The model of Fiorentini et al. (1990) employs a single filter type at different spatial scales while the model proposed by duBuff (1994) uses operators that resemble pairs of simple cells centred at the same location but in quadrature.

A second type of models 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 & Moulden (1991) 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 models 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). The latest 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)).

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 & 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 (Moulden & Kingdom, 1991; Kingdom & Moulden, 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 anisotropic filters (oriented difference of Gaussians or ODOG), non-linearly pooled, was developed to account for a variety of brightness effects that require oriented filters such as the White effect (Blakeslee and McCourt, 1999; Blakeslee & McCurt, 2001; Blakeslee and McCourt, 2004; Blakeslee et al., 2005). Another multi-resolution perceptual model is the one developed by D'Zmura & Singer (1998, 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 Perceptual Wavelet Model (PWM) we present here shares some similarities with both the ODOG and the D'Zmura & Singer (1998, 1999) models. Although a multi-resolution 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. Also in contrast to the previous models, a precise dependency on the contrast energy of the surround compared to the central stimulus is introduced here. In the ODOG model the interaction of the central stimulus and its surround is performed through a normalisation of the total visual space (and in the case of the D'Zmura model this explicit comparison is not performed). We decided to use the wavelet transform as the main framework for this work. 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; DeValois et al., 1982) and no single basis set can capture this variability, wavelets provide basis functions that are well localized (both in space and frequency) among other mathematical properties (such as self-similarity) which makes them popular among modelers (Zetzsche & Nuding, 2005; Van Rullen & Thorpe, 2001; Field, 1999; Olshausen & Field, 1997).

2 The Perceptual Wavelet Model

In this work we propose a new low-level brightness induction model (the PWM) that combines three important contributions, namely spatial frequency, spatial orientation and surround contrast to brightness assimilation/contrast phenomena.

At the basis of the PWM there is a multiresolution wavelet decomposition which separates an achromatic image in different spatial frequency (reminiscent of parvocellular s.f. channels) and orientation components. The recovery of the perceptual brightness image is done by weighting the wavelet coefficients using a modified version of the human 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.

We aim to produce the simplest mathematical formulation (i.e. including the least possible number of free parameters) which models these three properties in a manner compatible with current physiological and psychophysical research.

With this simple architecture, we approximately reproduce several brightness induction effects such as the Simultaneous Brightness Contrast, White effect, Grating Induction, Todorovic effect, Mach Bands effect, Chevreul illusion, Adelson-Logvinenko snake illusion, Dungeon illusion and Checkerboard effect.

2.1 Contrast Sensitivity Function

The detection threshold for sinusoidal gratings depends on the grating s.f. (Mullen (1985) and Simpson & McFadden (2005)) and this relationship is described by the contrast sensitivity function ($CSF(\nu)$) which is band pass for achromatic stimuli. Fig. 1 shows an approximate representation of the CSF.

Experiments with square-wave periodic patterns have shown that brightness

assimilation effects increase when the s.f. of the target feature is higher than certain induction threshold. For brightness induction, this transition point ν_{thr} was estimated to be near 4 cpd (Walker, 1978) and for chromatic contrast induction between 4 - 6.7 cpd (Fach et al, 1986) and 4 cpd (Smith et al., 2001). **ALEJ:This point needs to be made clear: why chose 6 cpd when the literature says 4 cpd???**

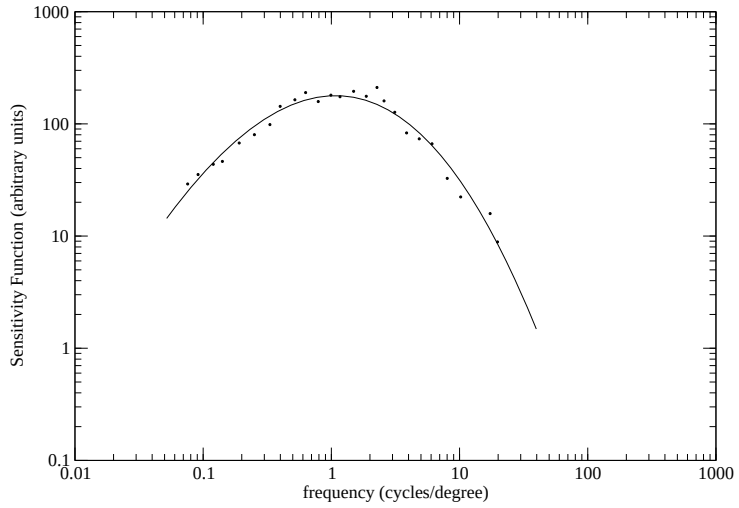


Fig. 1. Generic profile of the $CSF(\nu)$. Dots are the original data obtained by Mullen (1985). Spatial frequency is in cycles per degree (cpd)

ALEJ: This CSF seems a bit strange... why is it peaking at 6 cpd when it is generally lower (see Mullen (1985)???

Physiological and psychophysical studies (Blakemore & Campbell, 1969; Graham & Nachmias, 1971; DeValois et al., 1982; Wilson et al., 1983; Werner, 2003) show that spatial frequencies differing by a factor of around 2 (i.e. one octave) are processed by narrowly-tuned, nearly independent s.f. channels. These channels have a bandwidth of about one octave and the CSF acts as an overall envelope for them (Wilson et al., 1983).

The ordinate axis of the CSF is usually defined in terms of cycles per degree of visual angle. Since visual angle is related to both the object's physical size and its distance d from the observer, we can associate a spatial scale s to a certain visual angle ν (see Appendix A for details). Therefore, given an observation distance d the psychophysical $CSF(\nu)$ function shown in Fig. 1 can be defined in the scale space s as a function $CSF_d(s)$ and approximated by a piecewise function defined by two Gaussians (see eq. (B.1) in Appendix B). We can also define a particular scale s_{thr} associated to the particular $\nu_{\text{thr}} = 3$ cpd spatial frequency (see Appendix A) where the CSF peaks **XOP: HEmos puesto 3**

cpd en el maximo, pero si miramos la grafica de Mullen diria que puede estar incluso en el 1cpd !!!. This allows us to define the function $CSF_d(\dot{s})$, 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

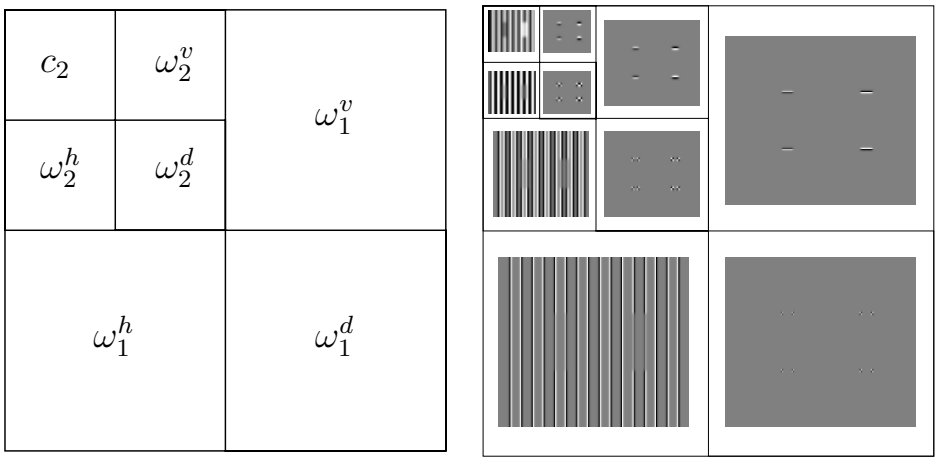
A multiresolution analysis allows us to describe an image in terms of different spatial resolution (i.e. spatial frequency) components. The particular multiresolution framework used in this work is similar to the one described by the Mallat decomposition algorithm (Mallat, 1989, 1998) which is mainly used to perform orthogonal wavelet transforms, and has been modified here to produce non-orthogonal basis with a smooth and symmetric profile. Despite these mathematical differences, the main concept and philosophy of these two algorithms is the same. The Mallat algorithm decomposes the 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° orientation bandwidths, i.e. in vertical, horizontal and diagonal orientations. In Fig. 2 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. 2 b). After being decomposed, the original image I is represented as a summa of wavelet planes of different s.f. and orientation as follows

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

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

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

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



a) b)

Fig. 2. Mallat multiresolution decomposition: a) The White effect image from Fig. 3 a) is decomposed into several wavelet planes ω which contain features of a certain s.f. and orientation. In this case only 2 multiresolution levels are shown.

In our model, these s.f. channels (wavelet planes ω_s) have a bandwidth and layout similar to those of the human visual system (h.v.s.) channels that determine the shape of $CSF(\nu)$.

2.3 Hypothesis

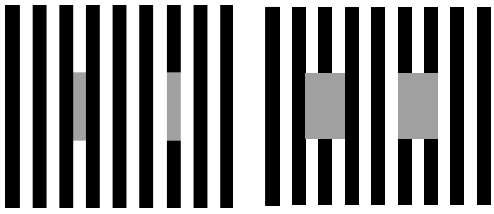
As mentioned before, there is ample evidence that our perception of a central stimulus can be modified by the spatial content of the surroundings (Heeger, 1992; DeValois et al., 1982; Chubb et al., 1989; Werner, 2003; D'Zmura & Singer, 1998, 1999; Yu et al., 2001, 2002). In this section describe how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

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 (Werner, 2003; D'Zmura & Singer, 1998, 1999; 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; Graham & Nachmias, 1971; DeValois et al.,

1982; Wilson et al., 1983; Werner, 2003; D'Zmura & Singer, 1998, 1999; Yu et al., 2001, 2002). Panel (a) in Fig. 3 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 effect, we define the CSF function by defining the first of the three assumptions in our model:

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



a) b)

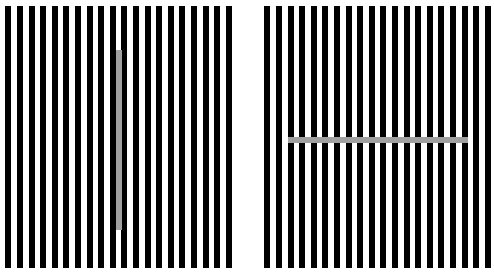
Fig. 3. a) Example of the influence of surround spatial frequency. Grey patches have different brightness because of their different local spatial information content. Both grey patches and vertical stripes have the same width (i.e. horizontal size or spatial scale) and this produces a strong induction 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.

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 make the $CSF_d(\dot{s})$.

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 et al., 1993; Yu et al., 2001, 2002, 2003) show that brightness assimilation in 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. 4 (Panel a) where the grey stripe has the same spatial

orientation as its surrounding black and white stripes. Conversely, the grey stripe in Fig. 3 b) has a different spatial orientation than its surrounding stripes and is perceived brighter than the grey stripe in Fig. 3 a). Following this, we define our second assumption:



a) b)

Fig. 4. 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 degrees) to the black and white stripes, brightness assimilation does not occur.

Assumption 2: Brightness assimilation is strongest when the central stimulus 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 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 important 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 (Nachmias & Sansbury, 1974; Cannon & Fullenkamp, 1991; Chubb et al., 1989; Ejima & Takahashi, 1985; Ellemberg et al., 1998; Klein et al., 1974; Mackay, 1973; 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. 5 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. 5) 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:

Assumption 3: When the brightness contrast of the surround features increases, brightness assimilation increases, i.e. brightness contrast decreases, and viceversa.

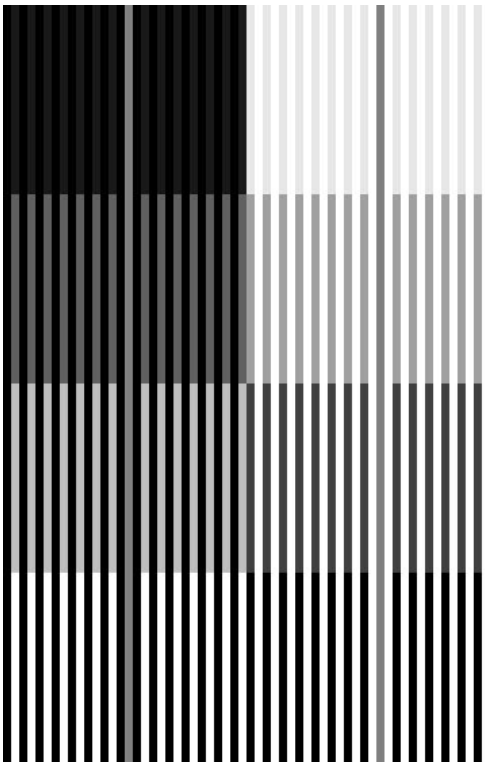


Fig. 5. In this series of images, the surround contrast of the two vertical grey lines is increased downwards in 4 steps. The left grey line is in contact with black stripes, and the right grey line is in contact with white stripes. As the line extends downwards (and surround contrast increases) brightness induction increases, i.e. the left grey line becomes darker and the right grey line becomes brighter.

To translate this assumption to the model we will need to insert the surround contrast as a new parameter of the model. We will see in the next section how to perform this.

2.4 Multiresolution brightness induction model

All the assumptions mentioned above are implemented in our model as a reconstruction process from the decomposed original image towards the perceptual or induced image, i.e. modifying the wavelet coefficients in eq. (2). To accomplish this perceptual recovery, we can modify the wavelet coefficients applying a weighting factor function to. This function should reflect some perceptual

properties of the HVS. As a first approximation, we assume that this function has a shape similar to the human CSF, defining

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

being I_{percep} the recovered perceptual image.

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). This way, *Assumption 1* and *Assumption 2* are naturally achieved 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

$$r = \frac{\sigma_{\text{cen}}^2}{\sigma_{\text{sur}}^2} \ , \tag{4}$$

being σ_{cen}^2 and σ_{sur}^2 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 3 times larger than the inner region, an approximate ratio suggested by Spitzer & Semo (2002) and Shapley & Enroth-Cugell (1984) and psychophysically measured by Yu et al. (2001). Althrough 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 pioneering study by Nachmias & 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 close to the Naka & Rushton (1966) function (which has also been successful in reproducing the responses of cortical V1 neurons (Albrecht &

Hamilton, 1982; Sclar et al., 1990; Tolhurst & Heeger, 1997)). The same relationship has been used to model non-linear dynamic feedback between a single channel and the all the others (across s.f. and orientations), implying that visual channels are subject to gain control (Heeger (1992)). 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:

$$z_{\text{ctr}} = \frac{r^2}{1 + r^2} \ , \tag{5}$$

where z_{ctr} is non-linear and fulfils $z_{\text{ctr}}(x, y; s, o) \in [0, 1]$. The previous expression is a non-linearisation of the r variable in eq. (4).

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. 5). Since r in eq. (4) is an estimation of the central feature contrast relative to its surround contrast, z_{ctr} in eq. (5) 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 $CSF_d(\dot{s})$ we use this variable $z_{\text{ctr}}(x, y; s, o)$ where s represents the s.f. and o is orientation involved. A new CSF' can be written as follows:

$$CSF'(\dot{s}, z_{ctr}) = z_{\text{ctr}} \cdot CSF_d(\dot{s}) + CSF_{min}(\dot{s}) \tag{6}$$

In this expression, $CSF'(\dot{s}, z_{\text{ctr}})$ is minimum when $z_{ctr} = 0$ (i.e. minimum brightness contrast or maximum brightness assimilation). To avoid $CSF'(\dot{s}, z_{\text{ctr}})$ becoming null for some spatial frequencies s (mainly for low spatial frequencies -see Fig. 6) we have introduced the term $CSF_{min}(\dot{s})$ in Eq. (6). The exact mathematical expression we used for $CSF_{min}(\dot{s})$ is shown in Appendix B. In this way $CSF'(\dot{s}, z_{\text{ctr}}) \rightarrow CSF_{min}(\dot{s})$ when $z_{\text{ctr}} \rightarrow 0$. This avoids a high degree of assimilation being performed at low s.f. (i.e. large scale features), which would make them become unseen. We show this function $CSF_d(\dot{s})$ in Fig. 6. In the opposite situation, $CSF'(\dot{s}, z_{ctr})$ is maximum when $z_{ctr} = 1$. Spitzer & Semo (2002) estimates that the maximum enhancement factor, i.e. maximum perceived brightness contrast, in the HVS is around 1.5. It is the peak value of $CSF'(\dot{s}, z_{ctr})$ (see Fig. 6).

Another important property of the $CSF'(\dot{s}, z_{ctr})$ 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 (Mallat, 1998). The function that defines the

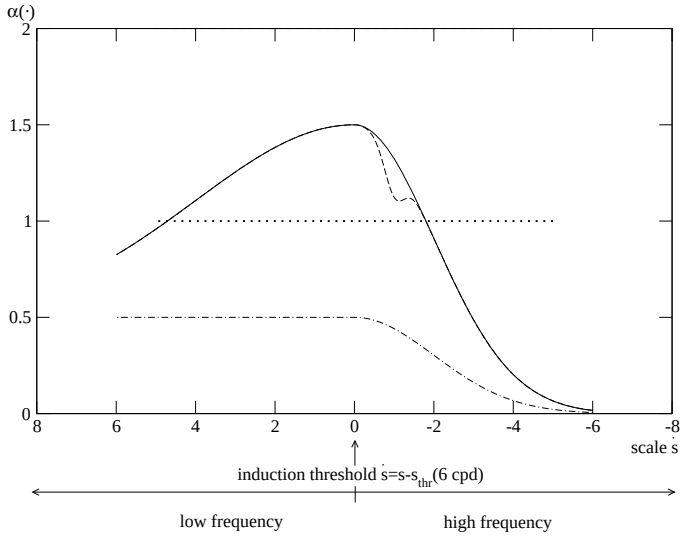


Fig. 6. (*continuous function*): Same function as Fig. 1. (*dotted function*): Profile of $CSF'(\dot{s}; z_{ctr}(\cdot))$ with $z_{ctr}(x, y; s, o) = 0.75$. That is, when applying $z_{ctr}(\cdot) = 0.75$ just on a particular s wavelet plane which fulfils $\dot{s} \equiv s - s_{\text{thr}} = -1$, see text for details.

spread's influence is defined by averaging the Wigner-Wille distribution (Mallat, 1998, chap. 4), which can itself be approximated by a Gaussian function. In order to show what is the effect on the continuous $CSF'(\cdot)$ 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. 6). 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. 6 shows how the $CSF'(\cdot)$ 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 & Nachmias (1971); Nachmias & Sansbury (1974) for grating adaptation and masking effects, using a mathematical expression qualitatively equivalent the ours (eq. (5)).

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'

$$\alpha(s, \cdot) \equiv CSF'(\dot{s}, z_{ctr}(\cdot)) \tag{7}$$

We obtain

$$I_{\text{percep}}(x,y) = \sum_{s=1}^{n} \sum_{o=v,h,d} CSF'(\dot{s}, z_{ctr}(x,y;s,o)) \cdot \omega_s^o(x,y) + c_n(x,y) \qquad (8)$$

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

3 Model predictions

Once we have introduced the PWM model, let us to show some predictions on a set of images we have already commented in section 1 that show some known brightness induction effects. We are going to show quantitatively and qualitatively how the model predicts some visual effects as the Simultaneous Brightness Contrast (SBC), the White effect (W), the Grating Induction (GI), the Todorovic effect (T), the the Mach bands (MB), the Chevreul effect (C), the Adelson-Logvinenko snake, the Dungeon illusion and the checkerboard illusion. Many models have been proposed to account for these individual effects, but few are general enough to explain all of them. Our modelling framework is similar in philosophy to the ODOG model (Blakeslee and McCourt, 1999) and can reproduce all these visual effects including some effects which were previously unexplained at low level such as the Dungeon illusion and the Checkerboard effect.

To be able to compare PWM predictions to the psychophysical data available from the literature, we adjusted the input image and the model's parameters to be consistent with the physical dimensions (size, visual angle, observation distance, etc.) of the actual stimuli. For the SBC, GI, W and T effects, we compared our model to published psychophysical results (Blakeslee and McCourt, 1999) kindly supplied by M.E. McCourt. For the MB and the C effect we obtained data from Lu & Sperling (1995) (Table 2).

3.1 SBC and White effect

Panels (a) and (c) in Fig. 7 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.

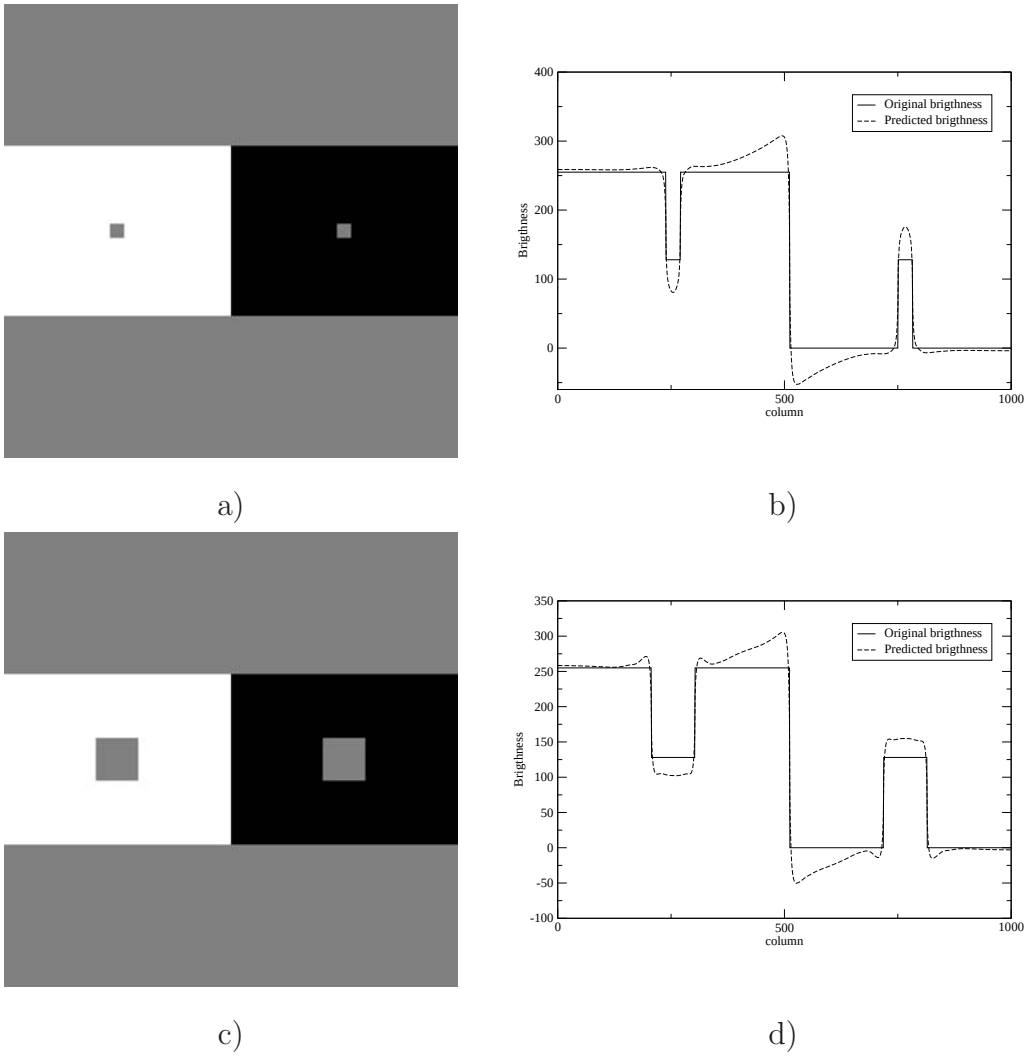


Fig. 7. Panels (a) and (c) show an example of the simultaneous contrast effect and panels (b) and (d) show our model's results. The solid lines in the plots on the right show the profile of the central row taken from the corresponding left panel. The dashed lines show our model's predicted brightness profile (darker left grey patch and a brighter right grey patch.)

The continuous function in Fig. 7, (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 PWM predicts the increase/decrease of the perceived brightness of this grey patch over its original value. Furthermore, the model predicts the perceived maxima and minima that appears to run parallel to the vertical light/dark edge at the centre of figures (a) and (c). The operation of PWM can be summarised as follows. Consider the left grey patch in 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 on it at this particular spatial scale. Given that the grey square is darker than its local surround, it becomes even more darker (perceptually). The same reasoning can be applied to the other grey patch on top of the dark background, which becomes perceptually lighter (and the bottom panels in Fig. 7).

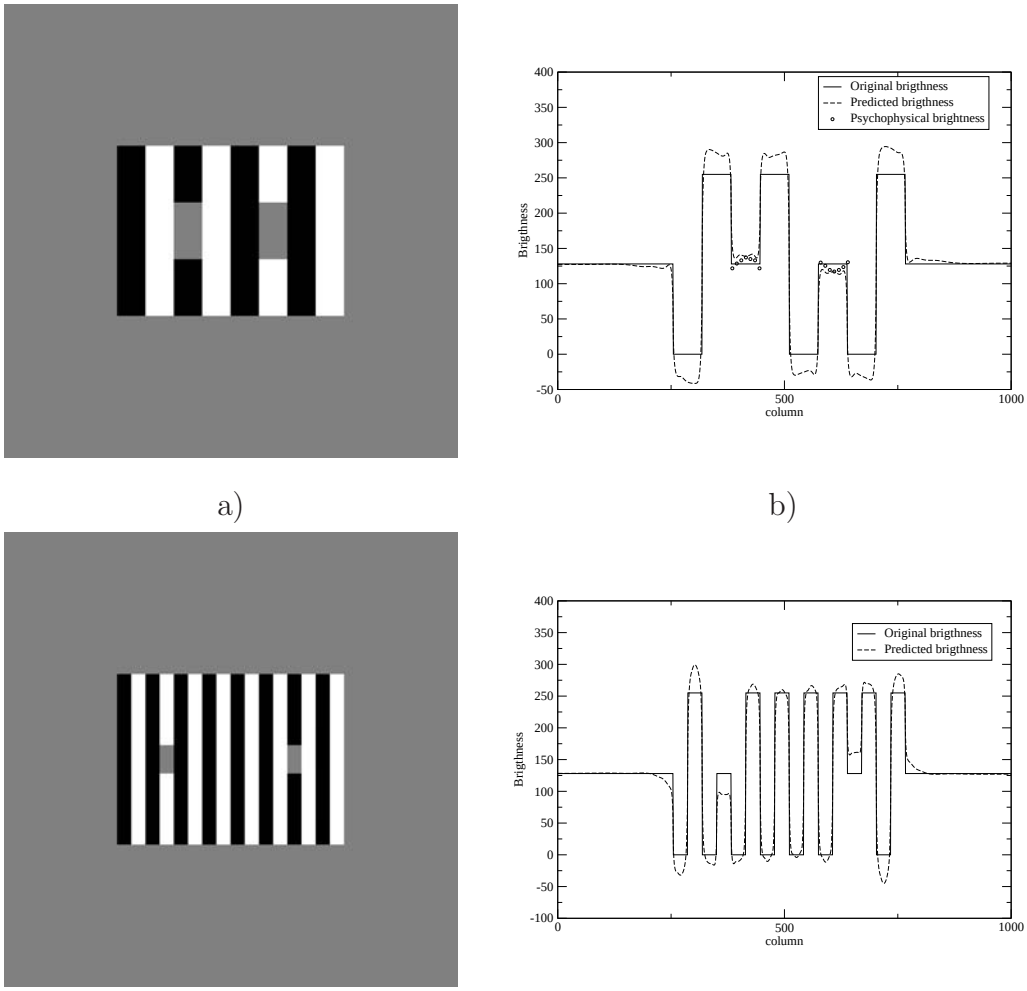
Another extensively studied effect (shown in Fig. 8 panels (a) and (c), where the left grey rectangle is seen brighter than the right one in panel (a) and darker in panel (c)) is the White effect. This effect is generally considered a particular case of SBC (Zaidi, 1989; Moulden & Kingdom, 1991) and can be explained using spatially oriented filters (Blakeslee and 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 functions in Fig. 8 (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 PWM. We see that our method correctly predicts that left grey patches are perceived darker and right patches brighter. The model also predicts that the vertical white stripes are predicted to be lighter, and the black stripes darker.

As a comparison, we have added in Fig. 8 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.

Our model reproduces the White effect in a similar way as before. 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 an optimal wavelet plane where this patch is well represented because of 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



a) b)

c) d)

Fig. 8. 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).

brightness contrast. The total perceived brightness is a combination of these orientation-dependent interactions.

Fig. 8 (panels (b) and (d)) also shows that the model's prediction for perceived brightness of bright vertical stripes is higher that the original value and similarly, the predicted brightness 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. 9 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.

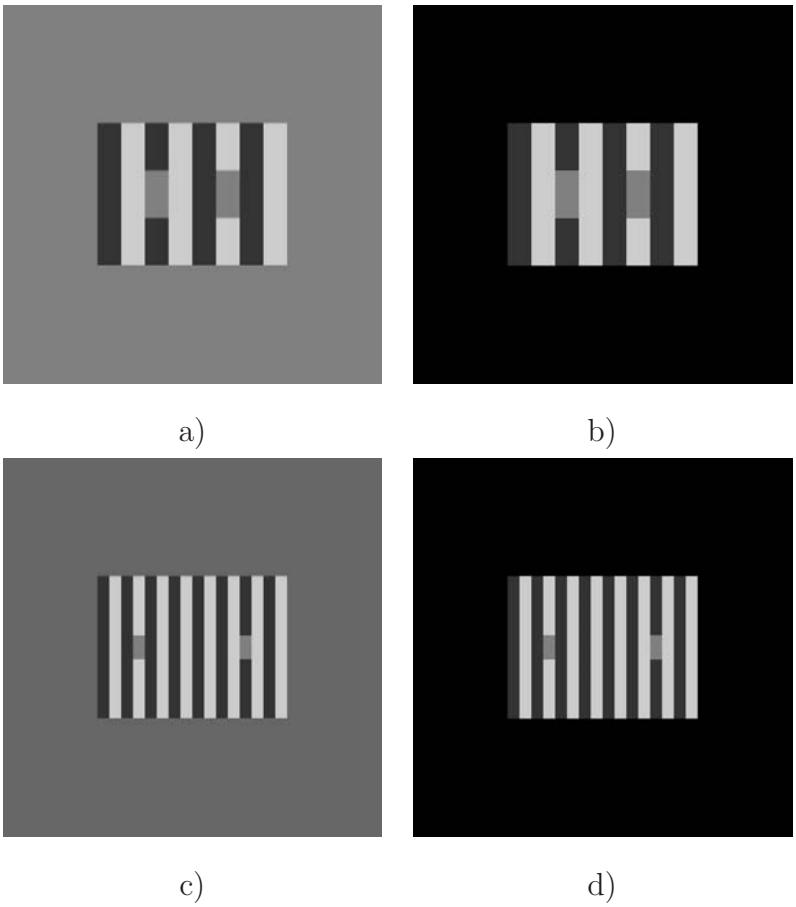


Fig. 9. Influence of a dark background on the White effect. Images (b) and (d) are the same as (a) and (c) but surrounded by a dark background. The perceived brightness of vertical dark stripes in dark background surroundings is different than in the case of the grey background.

3.2 Grating induction

Grating Induction (GI) (McCourt, 1982) is a brightness effect that produces a perceived brightness variation (a grating) on an spatially extended test field, see Fig. 10 (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 & McCourt (1997) (who modelled it with their ODOG model), this effect may be interpreted as a particular case of brightness contrast.

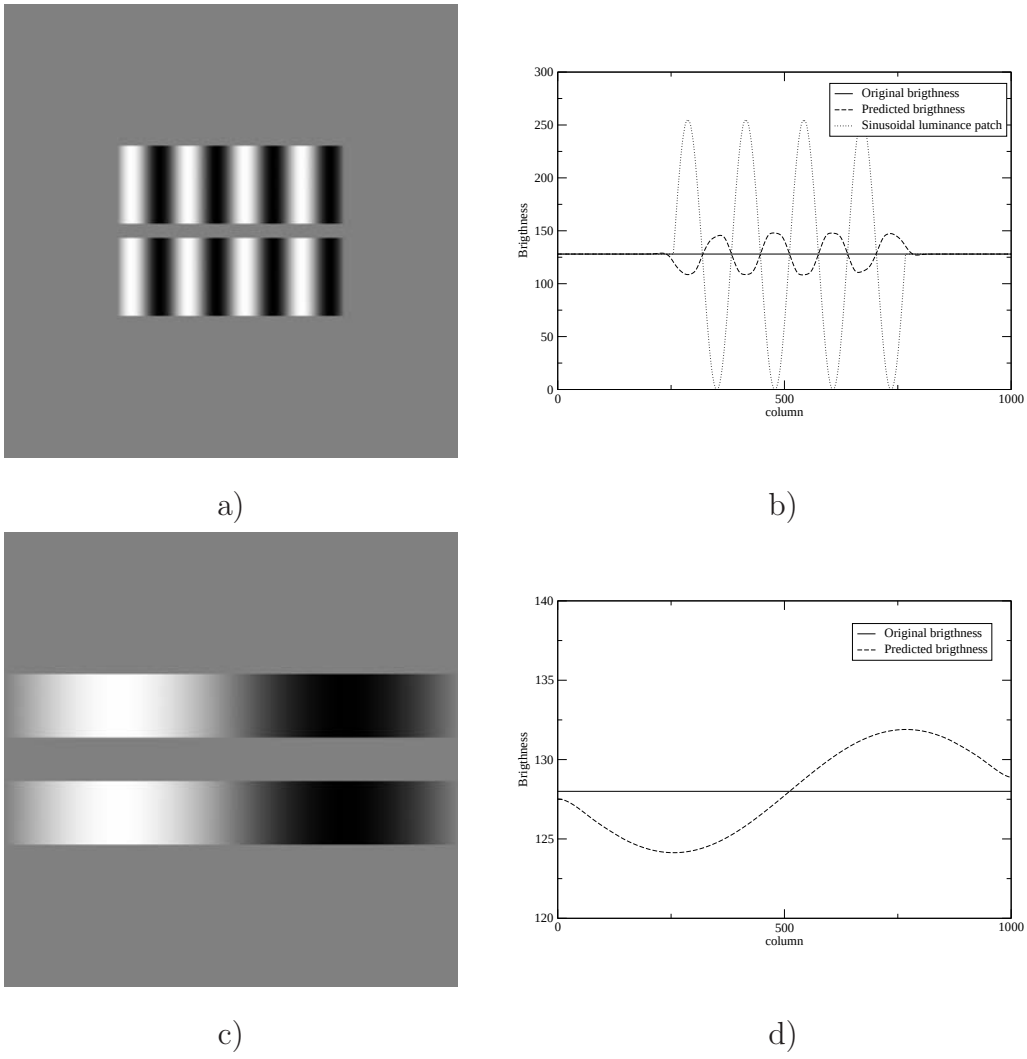


Fig. 10. Panel (a) In the GI effect, the thin horizontal stripe with constant luminance between the horizontal sinusoidal gratings is perceived as a sinusoidal brightness stripe in counterphase with the grating. Panel (b) Profile of a row of panel (a) showing the constant luminance of the horizontal 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).

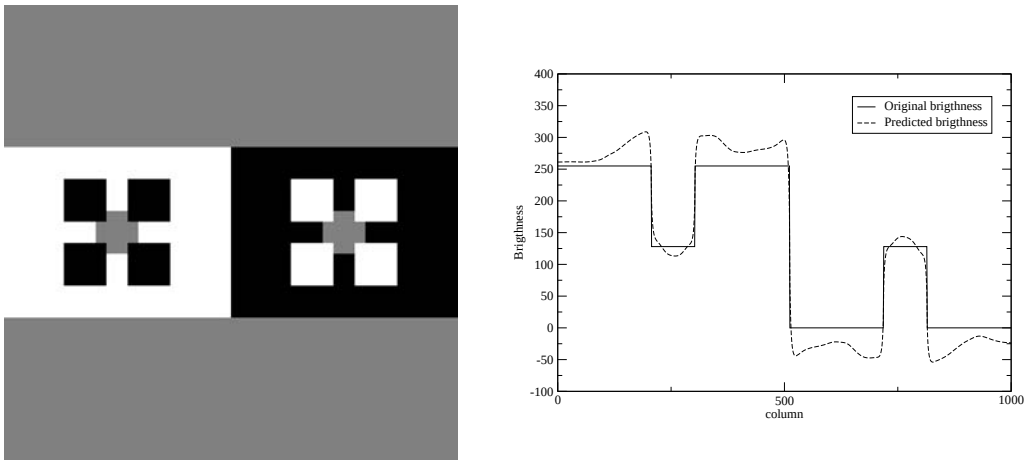
Fig. 10 Panel (b) shows the profile of the central row of both the uniform brightness of the central thin horizontal patch, the brightness profile predicted by our method and a row of the extended sinusoidal luminance grating. As we can see, the perceived brightness of the central grey stripe is predicted 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, 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 Todorovic effect

In Fig. 11 (a) we show a version of the Todorovic effect (Todorovic, 1997). This image is the same as in Fig. 7 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.



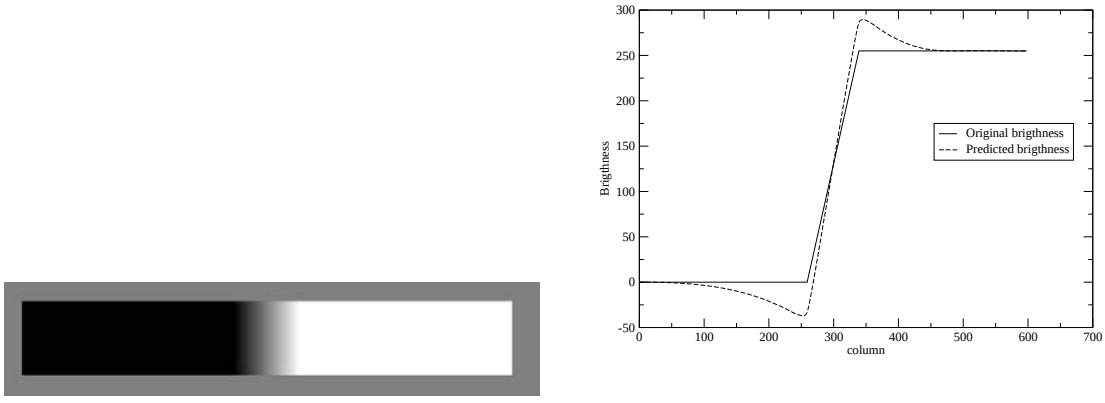
a) b)

Fig. 11. Panel (a): in the Todorovic effect (Todorovic, 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.

In Fig. 11, 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. 12 (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. Models that reproduce this effect can be classified as feature-based models (Tolhurst, 1972; Morrone & Burr, 1988; Fiorentini et al., 1990), rule-based models (Watt & Morgan, 1985; Kingdom & Moulden, 1991), filling-in models (Pessoa et al., 1995; Pessoa, 1996) and gradient-based models Keil et al. (2006).



a) b)

Fig. 12. Panel (a): example of the Mach bands effect. Panel (b): original data luminance (solid lines) and perceptual brightness predicted by our model (dashed line).

Panel (a) in Fig. 12 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 effect is 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 particular 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 perceived cusps.

3.5 Chevreul effect

In the Chevreul illusion (see Fig. 13), 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).

In Fig. 13, Panel (b) we show the original staircase profile (continuous function) and the brightness predicted by our method (dashed function), which correctly 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.

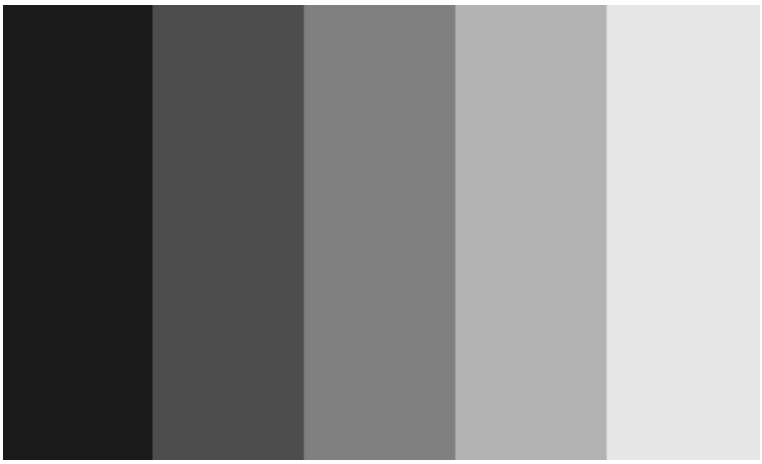
3.6 Adelson-Logvinenko snake

In Fig. 14 (Panel (a) we show a version of the Adelson snake pattern created by Logvinenko (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 of these top surfaces is 134 digital number (DN). Our method predicts 134 DN for the apparently light surfaces and 99 DN for the apparently dark ones.

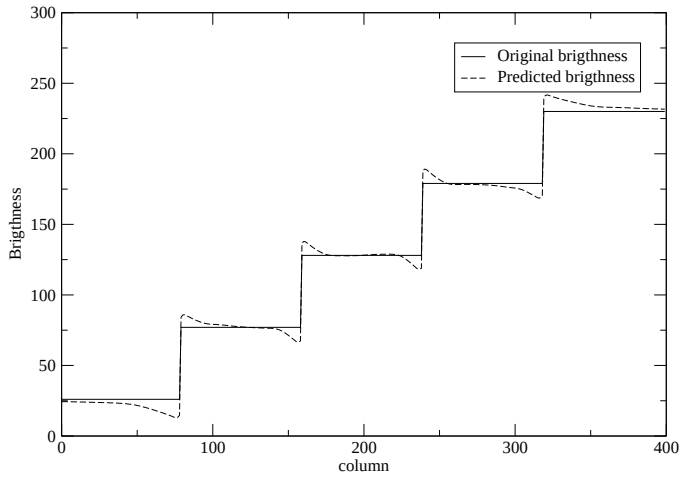
The presence of the vertical (low s.f.) modulating sinusoidal grating means that there will be a particular vertically-oriented wavelet plane where this feature will be represented best with very little influence of the rest of the image. This will induce a strong brightness contrast effect in the rows 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



a)



b)

Fig. 13. 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.

is the dungeon illusion (Fig. 15 Bressan (2001) where the perceived difference between the grey squares on the left and the right of the picture is the opposite as one would predict from analysis of individual squares and their immediate surroundings (Gilchrist, 2006) . Panel (a) in Fig. 15 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

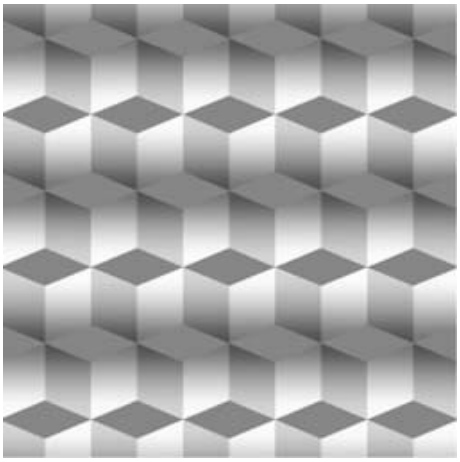


Fig. 14. Panel (a): example of the Adelson snake. The parts of the cubes at the crests and valleys of the modulating sinusoidal have all the same luminance but are perceived differently. The model we present qualitatively predicts these differences(see text).

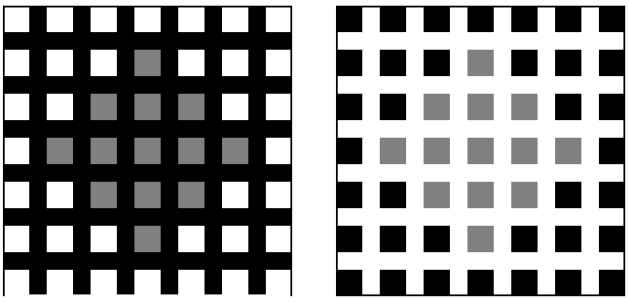
to its size, both vertically, horizontally and diagonally. Therefore, they will be strongly represented in the same wavelet planes, leading to a brightness assimilation effect. This in turn will produce darker rectangles in the left side of the image (where the rectangles' brightness will tend towards that of the bars) and slightly brighter rectangles on the right side.

3.8 Checkerboard

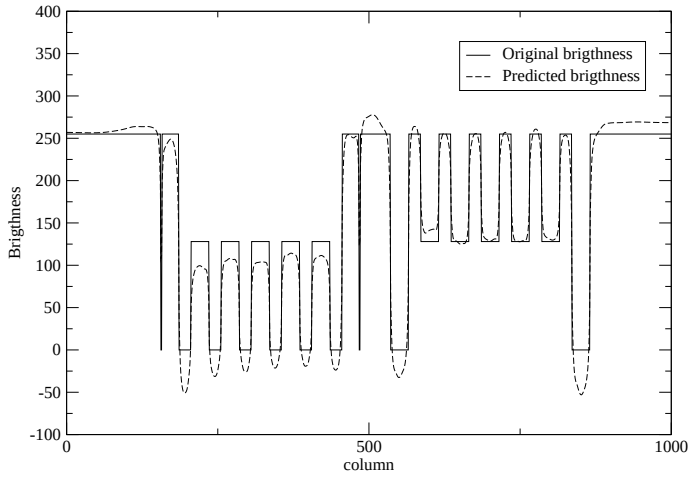
Another example of a complete reversal of contrast is the checkerboard illusion (shown in Panel (a) of Fig. 16). Here, the grey square in contact with white squares is perceived brighter that the grey square in contact with the black squares (an effect similar to the dungeon illusion). A simplified explanation 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 rela-



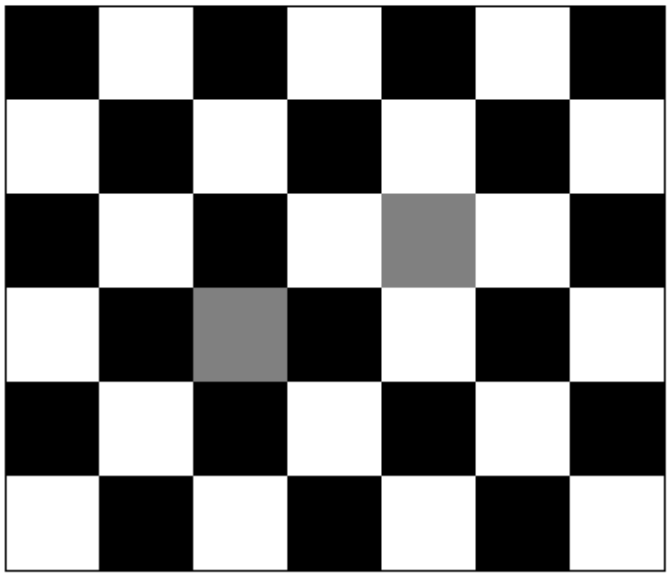
a)



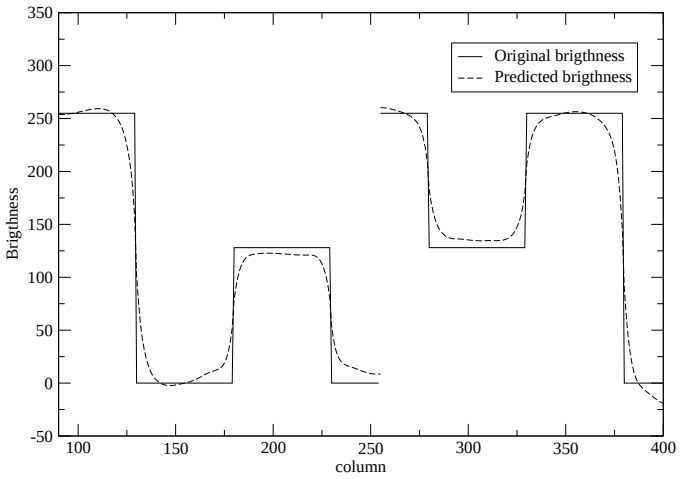
b)

Fig. 15. 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. 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.

tive 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. Figures 17 and 18 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 Todorovic effect). In Fig. 17, all experimental values (and their associated 95% confidence limit error bars) are represented by bars (dotted bars for observer MM and solid bars for observer BB in Blakeslee



a)



b)

Fig. 16. 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.

and McCourt (1999)). 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, consistently with Blakeslee and 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 modifying CSF' individually for these two sets of results, 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.

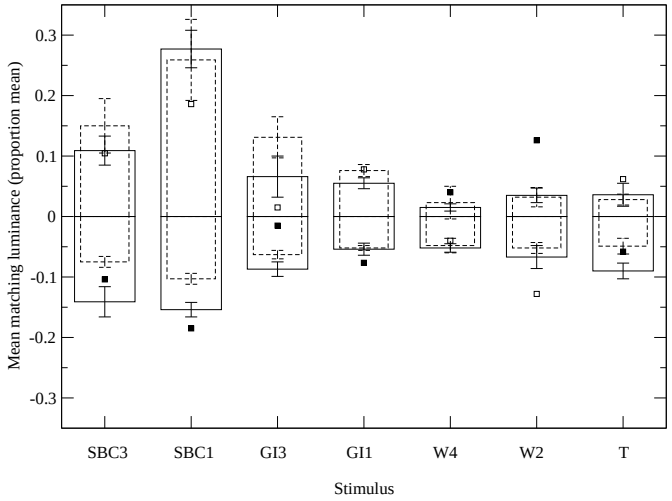


Fig. 17. Representation of the values predicted by our model (empty squares for patches on dark background and filled squares for patches on bright background) and the psychophysical values obtained by Blakeslee and McCourt (1999) (dotted bars for MM and continuous bars for BB) for Simultaneous Brightness Contrast (SBC), Grating Induction (GI), White effect (W) and Todorovic effect (T).

In Fig. 18 we show a plot of the values predicted by our model (abcissae) 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 aproximately linear behaviour. The solid line represents the best fitting line (linear regression) with a slope of around 0.9 and a correlation coefficient r = 0.87.

We also used psychophysical data by Lu & Sperling (1995) in order to ana-

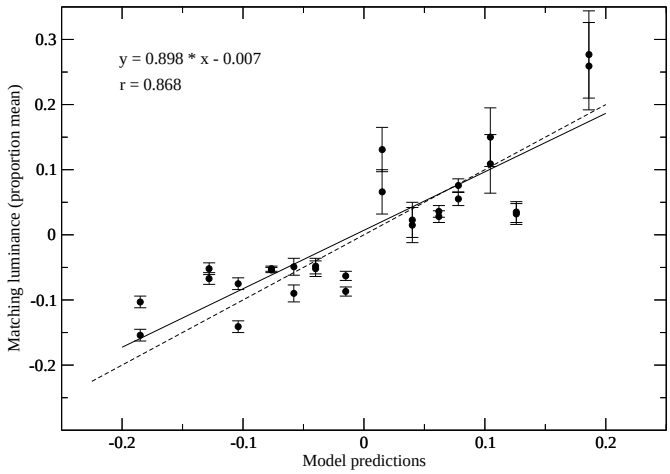


Fig. 18. Representation of the values predicted by our model (filled circles) and the psychophysical values obtained by Blakeslee and McCourt (1999) for all the visual effects considered in the figure above.

lyze the predictions of our model for the Mach bands and the Chevreul effect. Similarly to the previous examples, we simulated in our model the physical conditions of Lu & Sperling (1995). 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 is different from the observed 54% value.

Plots in Fig. 17 and 18 and the 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 and orientation provides guidelines as to how to interpret these effects.

5 Conclusions

Our simplistic model of brightness induction in the HVS is based on three main features of visual scenes: spatial scale (i.e. spatial frequency), 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 to unify 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. SBC, White effect, GI, Todorovic effect, Mach bands, Chevreul effect, Adelson-Logvinenko snake, dungeon illusion and checkerboard illusion, which were not explained by a single (unified) framework before.

But we believe our most important contribution is to explicitly incorporate observation distance, which dramatically changes the perceived effect, to include physiologically-plausible features in the model and to balance the weight of surround features in favour of those that are spatially matched and similarly oriented to the central elements. At this stage we have avoided more complex issues such as global contrast normalisation, which may have introduced more degrees of freedom at the cost of extra complexity.

Acknowledgements

Authors want to thank M.E. McCourt for kindly supplying psychophysical data. Xavier Otazu and Alejandro Párraga were funded by the Ramon y Cajal and the Juan de la Cierva programmes respectively. This work has been partially supported by project TIN2004-02970 of Spanish MEC (Ministery of Science). Thanks to Matthias S. Keil for his helpful comments.

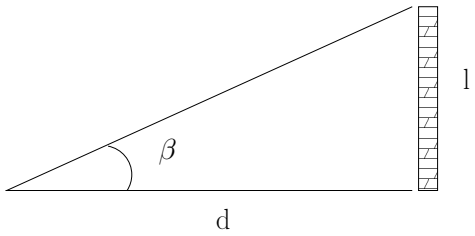


Fig. A.1. An object of size l observed from a distance d has a visual angle β.

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 (see Fig. A.1):

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

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 $\nu(s)$, i.e. to a period $T = \frac{1}{\nu(s)}$. This relation is defined by $2^s = T = \frac{l}{l_p}$, where $\frac{l}{l_p}$ 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 6 cycles of a certain spatial scale in the same longitude, we can define

$$6\,T \equiv 6 \cdot 2^{s_{thr}} = \frac{l}{l_p} \; . \tag{A.2}$$

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

$$s_{thr} = \log_2 \left(\frac{d \cdot \tan 1^\circ}{6 \cdot l_p} \right) \; . \tag{A.3}$$

The s_{thr} factor is the image scale associated to the $\nu_{\text{thr}} = 6$ cpd induction threshold value when observing an image with a pixel size l_p from a distance d.

B Piecewise CSF

We have approximated the human CSF by a piecewise function defined by two Gaussians of the form:

$$CSF(s) = \begin{cases} \exp\left\{-\frac{s^2}{2\sigma_1^2}\right\} & s \leq s_{\text{thr}} \\ \exp\left\{-\frac{s^2}{2\sigma_2^2}\right\} & s > s_{\text{thr}} \end{cases} . \tag{B.1}$$

Parameters σ_1 and σ_2 are the standard deviation of the piecewise Gaussian function for $s \leq s_{\text{thr}}$ and $s_{\text{thr}} < s$, respectively. When $\sigma_1 = \sigma_2$, eq. (B.1) is a symmetric Gaussian. To reproduce the approximate profile of the psychophysical $CSF(s)$ functions obtained from the literature citepMullen85, we made $\sigma_2 = 2\,\sigma_1$ and $\sigma_1 = 2$.

$CSF_{min}(s)$ was defined as:

$$CSF_{min}(s) = \begin{cases} \frac{1}{2}\ \exp\left\{-\frac{s^2}{2\sigma_1^2}\right\} & s \leq s_{\text{thr}} \\ \frac{1}{2} & s > s_{\text{thr}} \end{cases} . \tag{B.2}$$

References

E.H Adelson, Perceptual Organization and the Judgment of Brightness, *Science*, **262(5142)**, 2042-2044 (1993).

E.H Adelson, Lightness perception and lightness illusions, in The New Cognitive Neurosciences, 2nd edn, M. Gazzaniga (Ed.), pp. 339-351. MIT Press, Cambridge, MA, USA, 2000.

T. Agostini and A. Galmonte, Perceptual organisation overcomes the effect of local surround in determining simultaneous lightness contrast *Psychol. Sci.*, **13**, 88-92 (2002).

D.G. Albrecht and D.B. Hamilton, Striate cortex of monkey an cat: contrast response function, *J. Neurophysiol.*, **48**, 217-237 (1982).

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

B. Anderson, Perceptual organization and White's illusion, *Perception*, **32**, 269-284 (2003).

C. Blakemore and F.W. Campbell, On the existence of neurones on the human visual system selectively sensitive to the orientation and size of retinal images, *J. Physiol.*, **203**, 237-260 (1969).

B. Blakeslee and M. E. McCourt, Similar mechanisms underlie simultaneous

brightness contrast and grating induction, *Vision Research*, **37**, 2849-2869 (1997).

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

B. Blakeslee and M. E. McCourt, A multiscale spatial filtering account of the Wertheimer-Benary effect and the corrugated Mondrian, *Vis. Res.*, **41**, 2487 (2001).

B. Blakeslee and M.E. McCourt, A unified theory of brightness contrast and assimilation incorporating oriented multiscale spatial filtering and contrast normalization, *Vis. Res.*, **44(21)**, 2483-2503 (1992).

B. Blakeslee, W. Pasieka and M. E. McCourt, 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 (2005).

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

W.M. Cannon and S.C. Fullenkamp, Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations, *Vis. Res.*, **31**, 1985-1998 (1991).

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

C. Chubb, G. Sperling and J.A. Solomon, Texture interactions determine perceived contrast, *Proc. Nat. Acad. Sci. USA*, **86**, 9631-9635 (1989).

T.N. Cornsweet, Visual Perception, Academic Press, San Diego, 1970.

R.L. DeValois, R.L. Albrecht and L.G. Thorell, Spatial frequency selectivity of cells in macaque visual cortex, *Vis. Res.*, **22**, 545-559 (1982).

R.L. DeValois and K.K. DeValois, Spatial vision, Oxford University Press, New York, 1988.

J. du Buff, Ramp edges, Mach bands, and the functional significance of the simple cell assembly, *Biological Cybernetics*, **69**, 449-461 (1994).

M. D'Zmura and B. Singer, Color contrast gain control, in In *Color Vision, Perspectives from different disciplines* (W.G.K. Backhaus, R. Kiegl and J.S. Werner, eds.), New York: Walter de Gruyter and Co., 9369-385 (1998).

M. D'Zmura and B. Singer, Contrast gain control, in *Color Vision: From Genes to Perception* (K.R. Gegenfurtner and L.T. Sharpe, eds.), Cambridge: Cambridge University Press,9369-385 (1999).

Y. Ejima and S. Takahashi, Apparent contrast of a sinusoidal grating in the simultaneous presence of peripheral gratings, *Vis. Res.*, **25**, 1223-1232 (1985).

D. Ellemberg F.E. Wilkinson, H.R. Wilson and A.S. Arsenault, Apparent contrast and spatial frequency of local texture elements, *J. Opt. Soc. Am. A*, **15**, 1733-1739 (1998).

C. Fach and L. T. Sharpe Assimilative hue shifts in color gratings depend on bar width *Perception and Psychophysics*, **40**, 412418 (1986).

D.J. Field, Wavelets, vision and the statistics of natural scenes, *Philosophi-*

cal Transactions of the Royal Society of London A, **357(1760)**, 2527-2542 (1999).

A. Fiorentini, G. Baumgartner, S. Magnussen, P. Schiller and J. Thomas, The perception of brightness and darkness: Relations to neuronal receptive fields, 129-161, in: L. Spillmann and J. Werner, eds., Visual perception: the neurophysiological foundations, Academic Press, San Diego, CA, 1990.

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

C. D. Gilbert, A. Das, M. Ito, M. Kapadia and G. Westheimer, Spatial integration and cortical dynamics, *Proceedings of the National Academy of Sciences USA*, **93**, 615-622 (1996).

A. Gilchrist, C. Kossyfidis, F. Bonato, T. Agostini, J.Cataliotti, X. Li, B. Spehar, V. Annan and E. Economou, An anchoring theory of lightness perception, *Psychol. Rev.*, **106**, 795-834 (1999).

A. Gilchrist, Seeing Black and White, Oxford University Press, New York, 2006.

L.B. Goldstein, Sensation and Perception, Brooks-Cole, Pacific Grove, CA, 4th ed, 1996.

N. Graham and J. Nachmias, Detection of grating patterns containing two spatial frequencies: a comparison of single channel and multiple-channel models, *Vis. Res.*, **11**, 251-259 (1971).

D.J. Heeger, Normalization of cell responses in cat striate cortex, *Vis. Neurosci.*, **9**, 181-197 (1992).

D.J. Heeger and R.R. Robinson, IS simultaneous contrast divisive?, *Investigative Ophtalmology & Visual Science (Suppl.)*, **35**, 2006 (1994).

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

H. von Helmholtz, Handbuch der Physiologischen Optik. Voss, , Leipzig, 1867.

L. Hermann, Eine Erscheinung simultanen Contrastes, *Pflgers Archiv fr die gesamte Physiologie*, **3**, 13-15 (1870).

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

E.N. Johnson, M.J. Hawken and R. Shapley, The spatial transformation of color in the primary visual cortex of the macaque monkey, *Nature Neuoscience*, **4**, 409-416 (2001).

J. Jones and L. Palmer, An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex, *J. Neurophysiol*, **58**, 1233-1258 (1987).

M. S. Keil, G. Cristobal and H. Neumann, Gradient representation and perception in the early visual system- A novel account of Mach band formation, *Vision Research*, **46**, 1659-2674 (2006).

M. S. Keil, Smooth gradient representations as a unifying account of Chevreul's illusion, Mach bands and a variant of the Ehrenstein disk, *Neural*

computation, **18**, 871-903 (2006).

F. Kingdom and B. Moulden, A multi-channel approach to brightness coding, *Vis. Res.*, **32**, 1565-1582 (1992).

F.A.A. Kingdom, Levels of Brightness Perception, Levels of Perception, L. Harris and M. Jenkin.(ed), Springer-Verlag, New York, 2003.

S.A. Klein C.F. Stromeyer and L. Ganz, The simultaneous spatial frequency shift: a dissociation between the detection and perception of gratings, *Vis. Res.*, **14**, 1421-1432 (1974).

E.H. Land and J.J. McCann, Lightness and the retinex theory, *JOSA*, **61**, 1-1 (1971).

Z.-L. Lu and G. Sperling, Second-order illusions: Mach Bands, Chevreul, and Craig-O'Brien Cornsweet, *Vis. Res.*, **36**, 559-572 (1995).

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

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

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

E. Mach, , *Classe Kaiserlichen Akad. Wiss.*, **52**, 303-322 (1865).

D.M. Mackay, Lateral interaction between neural channels sensitive to texture density, *Nature (Lond.)*, **245**, 159-161 (1973).

S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. on Pattern Anal. and Machine Intell.*, **11**, 674-693 (1989).

S. Mallat, A wavelet tour of signal processing, 2nd edition, Academic Press, San Diego, 1998.

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

J.A. McArthur and B. Moulden, A two-dimensional of brightness perception based on spatial filtering consistent with retinal processing, *Vis. Res.*, **39**, 1199-1219 (1999).

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

M.C. Morrone and D.C. Burr, Feature detection in human vision: a phase-dependent energy model, *Proc. R. Soc Lond. B*, **235**, 221-245 (1988).

M.C. Morrone, D.C. Burr and J. Ross, Illusory brigtness step in the Chevreul illusion, *Vis. Res.*, **34**, 1567-1574 (1994).

B. Moulden and F. Kingdom, The local border mechanism in grating induction, *Vis. Res.*, **31**, 1999-2008 (1991).

K.T. Mullen, The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings, *J. Physiol*, **359**, 381-400 (1985).

J. Nachmias and R.V. Sansbury, Grating contrast: discrimination may be better than detection, *Vis. Res.*, **14**, 1039-1042 (1974).

K.I. Naka and W.A. Rushton, S-potentials from colour units in the retina of fish, *J.Physiol*, **185**, 584-599 (1966).

B.A. Olshausen and D.J. Field, Sparse coding with an overcomplete basis set: a strategy employed by V1?, *Vision Res.*, **37**, 3311-3325 (1997).

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

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

L. Pessoa, E. Mingolla and H. Neumann, A contrast and luminance-driven multiscale network model of brightness perception, *Vis. Res.*, **35**, 2201-2223 (1995).

L. Pessoa, Mach-bands: How many models are possible? Recent experimental findings and modelling attempts, *Vis. Res.*, **36**(19), 3205-3277 (1996).

A.F. Rossi, C. D. Rittenhouse and M.A. Paradiso, The representation of brightness in primary visual cortex., *Science* , **273(5278)**, 1055-1056 (1996).

S.J. Schein and R. Desimone, Spectral properties of V4 neurones in the macaque, *Journal of Neuroscience*, **10**, 3369-3389 (1990).

D. Schluppeck and S.A. Engel, Color Opponent Neurons in V1: A Review and Model Reconciling Results from Imaging and Single-Unit Recording, *J. of Vision*, **2**, 480-492 (2002).

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

R. Shapley and C. Enroth-Cugell, Visual adaptation and retinal gain controls, *Prog. Retinal Res.*, **3**, 263 (1984).

W.A. Simpson and S.M McFadden, Spatial frequency channels derived from individual differences, *Vis. Res.*, **45**, 2723-2727 (2005).

V.C. Smith, P.Q. Jin and Joel Pokorny, The role of spatial frequency in color induction, *Vision Research*, **41**, 1007 (2001).

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

L. Spillmann and J.S. Werner, Long-range interactions in visual perception, *Trends in Neuroscience*, **19**, 429-434 (1996).

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

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

D. Tolhurst, On the possible existence of edge detector neurons in the human visual system, *Vis. Res.*, **12**, 797-804 (1972).

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

D.J. Tolhurst and D.J. Heeger, Comparison of contrast-normalization and threshold model of the responses of simple cells in cat striate cortex, *Vis. Neurosci.*, **3**, 445-454 (1989).

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

J. T. Walker, Brightness enhancement and the Talbot level in stationary gratings *Perception and Psychophysics*, **23**, 1356359 (1978).

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

R. Watt and M. Morgan, A theory of the primitive spatial code in human vision, *Vis. Res.*, **25**, 1661-1674 (1985).

A. Werner, The spatial tuning of chromatic adaptation., *Vision Resarch*, **43**, 1611 (2003).

S. Westland, H. Owens, V. Cheung and I. Paterson-Stephens, Model of Luminance-Contrast Sensitivity Function for Application to Image Assessment, *Color Research and Application*, **31**, 315-319 (2006).

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

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

H.R. Wilson, D.K. McFarlane, G.C. Phillips, Spatial frequency tuning of orientation selective units estimated by oblique masking, *Vis. Res.*, **23**, 873-882 (1983).

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

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

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

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

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

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