

# The lowest spatial frequency channel determines brightness perception

A. Perna<sup>a,b,\*</sup>, M.C. Morrone<sup>c</sup>

<sup>a</sup> *Scuola Normale Superiore, Piazza dei Cavalieri 7, 56100 Pisa, Italy*

<sup>b</sup> *Istituto di Neuroscienze CNR, via Moruzzi 1, 56127 Pisa, Italy*

<sup>c</sup> *Università 'Vita-Salute' San Raffaele, via Olgettina 58, 20132 Milan, Italy*

Received 27 February 2006; received in revised form 5 January 2007

## Abstract

This study investigates the role played by individual spatial scales in determining the apparent brightness of greyscale patterns. We measured the perceived difference in brightness across an edge in the presence of notch filtering and high-pass filtering for two stimulus configurations, one that elicits the perception of transparency and one that appears opaque. For both stimulus configurations, the apparent brightness of the surfaces delimited by the border decreased monotonically with progressive (ideal) high-pass filtering, with a critical cut-off at 1 c/deg. Using two octave ideal notch filtering, the maximum detrimental effect on apparent brightness was observed at about 1 c/deg. Critical frequencies for apparent brightness did not vary with contrast, viewing distance, or surface size, suggesting that apparent brightness is determined by the channel tuned at 1 c/deg. Modelling the data with the local energy model [Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase dependent energy model. *Proceedings of the Royal Society (London)*, B235, 221–245] at 1 c/deg confirmed the suggestion that this channel mediates apparent brightness for both opaque and transparent borders, with no need for pooling or integration across spatial channels.

© 2007 Elsevier Ltd. All rights reserved.

**Keywords:** Brightness; Edges and lines detection; Spatial frequency channels; Filling-in; Feature detection

## 1. Introduction

The visual system is known to elaborate in parallel many of the attributes of visual images. While parallel analysis guarantees faster processing and stability, it poses the problem of how these independent, sometimes incongruent estimates of scene characteristics can support a unitary explicit perception.

An example of parallel analysis in the visual system is the processing of different spatial frequency bands by independent channels (Campbell & Robson, 1968), providing independent estimates of image attributes at the same retinal location. For near-threshold contrast-levels, scene detection can be simulated by evaluating the probability summation of information across spatial scales (Graham,

1977), implying independent analysis. However, independent analysis may not necessarily hold at supra-threshold contrasts. To predict scene appearance many models assume that the outputs of processing at each scale are summed, with constant weighting independently of the task to be performed (e.g., MIRAGE, Watt & Morgan, 1985; MIDAAS, Kingdom & Moulden, 1992).

At supra-threshold contrast values, contrast appearance does not scale with the size of the stimulus pattern, suggesting that all spatial scales play a similar role in determining the perceptual appearance of contrast and brightness (Georgeson & Sullivan, 1975). However, there are visual scenes whose perception seems to be mediated by a single channel, irrespective to it being the more suitable for the current task or not. Such tasks involve letter identification (Majaj, Pelli, Kurshan, & Palomares, 2002; Solomon & Pelli, 1994) and face recognition (Pelli, 1999). Recently, Peromaa and Laurinen (2004) showed that the appearance of a brightness illusion such as Chevreul staircase is medi-

\* Corresponding author. Address: UMR 5169, Bat IVR3, 118 route de Narbonne F-31062, Toulouse cedex 09, France. Fax: +39058446249.

E-mail address: [perna@cict.fr](mailto:perna@cict.fr) (A. Perna).

ated by the low spatial-frequency components of the pattern, while the visibility of the edges is mainly determined by higher spatial frequencies.

In general, the integration of responses across channels can be described in terms of a Minkowski sum, where the total activity  $A_{\text{tot}}$  is linked to the activities at the single scales  $A_i$  through

$$A_{\text{tot}} = \sqrt[m]{\sum_{i=1}^s w_i \cdot |A_i|^m} \quad (1)$$

Depending on the value of  $m$ , the equation describes a linear sum of responses across scales, Pythagorean sum, different kinds of probability summation, or winner take all (for very large  $m$ ). The weights  $w_i$  can be equal for all scales, or favour some scales over the others.

An alternative strategy to integration across scales would be to select an optimal channel depending on the task to be performed, for example, on the basis of the strongest response or the highest signal to noise ratio. This type of strategy have been shown to be used by our visual system, at least under some circumstances (Solomon, 2000).

The apparent brightness difference across an edge is known to depend on many factors, such as the luminance profile of the edge (Cornsweet, 1970; Craik, 1966; O'Brien, 1958) and the characteristics of other regions in the visual image, that interfere with brightness in a variety of different ways, including simultaneous contrast (Heinemann, 1955) and assimilation (Shapley & Reid, 1985). In addition, brightness is also influenced by the three-dimensional arrangement of surfaces and image segmentation, both in the presence and in the absence of transparency illusion (Adelson, 1993; Anderson & Winawer, 2005; Singh & Anderson, 2002). In the present experiment, we investigate how the contributions from different spatial frequency channels are integrated to determine the difference in apparent brightness across a luminance-defined edge that give rise to a simultaneous brightness illusion. In order to reach a general finding, we measured both opaque edges and edges that give rise to a transparency illusion.

## 2. General methods

### 2.1. Apparatus and stimuli

Stimuli were prepared with MATLAB and displayed with a Cambridge Research System VSG 2/3 graphics card on a high-resolution 21 in. BARCO monitor (spatial resolution  $656 \times 507$  pixels, frame refresh rate 126 Hz, mean display luminance  $16.6 \text{ cd/m}^2$ ). The whole stimulus measured  $25 \times 25 \text{ cm}$ , subtending  $23.5 \times 23.5 \text{ deg}$  of visual angle at the viewing distance of 60 cm (with the exception of one experiment where viewing distances of 30 and 120 cm were tested).

The basic stimulus configuration comprising the test and the match patch is illustrated in Fig. 1A. The test stimulus was constructed of two regions of different uniform intensity above and below the mean intensity, marked in the figure by the letters  $P$  and  $Q$ , and subtended  $0.93 \text{ w} \times 3.72 \text{ h deg}$  of visual angle.  $P$  and  $Q$  regions were aligned vertically

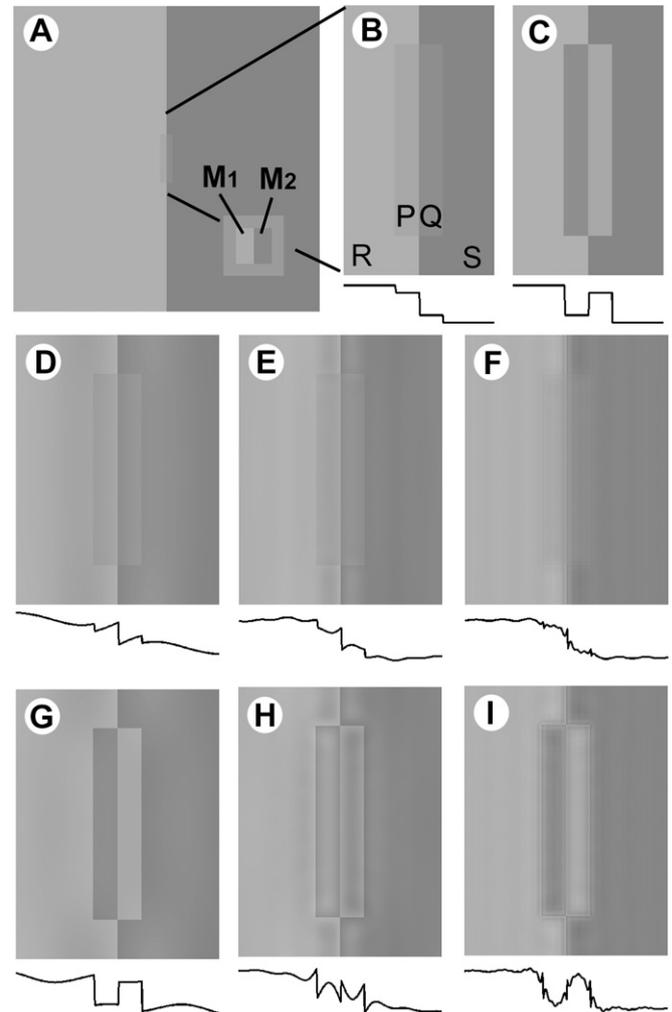


Fig. 1. Configuration of the stimuli used in the experiment. (A) Basic stimulus configuration comprising the test and the match patch. The central region is represented with increased size in (B). Subjects had to compare the contrast of the edge between region  $P$  and region  $Q$  with the contrast of the edge in the lower right part of the screen (the edge separating  $M1$  from  $M2$ ). Luminance of regions  $R$  and  $S$  were the same for all stimuli, while luminance of regions  $P$  and  $Q$  was variable (in all the stimuli represented in this figure,  $P$ - $Q$  contrast was 17.6%). Opaque configuration stimuli, such as that shown in (C) were obtained inverting contrast polarity of the edge between  $P$  and  $Q$ , but not of the edge between  $R$  and  $S$ . Both transparent and opaque stimuli could be spatial filtered in different bands. Examples of the stimuli obtained are shown in (D-I) corresponding to a notch-filter of two octaves positioned at 0.3 (D), 0.9 (E), and 4 (F) c/deg (centre frequency) applied to a “transparent” stimulus and to the same filters applied to an “opaque” stimulus (0.3 c/deg (G), 0.9 c/deg (H), 4 c/deg (I)).

against a larger pair of background regions of different uniform intensity again above and below the mean intensity ( $R$  and  $S$  (luminance values:  $S = 11.7 \text{ cd/m}^2$   $R = 21.5 \text{ cd/m}^2$ , contrast 0.29).

The match stimulus was made up of two regions of different uniform intensity above and below the mean, marked in the figure by the letters  $M1$  and  $M2$  and subtending  $1.38 \text{ w} \times 2.75 \text{ h deg}$  of visual angle each. It was located in the right corner of the screen and centred at a distance of  $9.5 \text{ deg}$  of visual angle from the test patch along the diagonal and was isolated from the other regions by a  $0.93 \text{ deg}$  wide white paper mask put on the screen to perceptually segregate it from the display ( $8.5 \text{ cd/m}^2$ ).

The contrast and contrast polarity across the  $P$ – $Q$  edge was changed for the different conditions. In particular, two qualitatively different configurations could be identified, one where the  $P$ – $Q$  and the  $R$ – $S$  edge had the same polarity and one where the two edges had opposite polarity. We termed the stimulus configurations where the  $P$ – $Q$  and the  $R$ – $S$  edge had the same contrast polarity “transparent” (in agreement with e.g., Anderson, 1997; Beck, Prazdny, & Ivry, 1984; Metelli, 1974) and the stimulus configurations where  $P$ – $Q$  and  $R$ – $S$  edge had opposite polarity “opaque”. The subjects, however, did not perform evaluations on the strength of transparency illusion.

“Transparent” and “opaque” stimuli were presented as such, or after notch or high-pass filtering at different positions in the spectrum, as will be described in more detail in the procedure section. Examples of the filtered stimuli are shown in Fig. 1D–I. The match stimulus was not filtered in any condition.

## 2.2. Procedure

The general aim of the experiment was to estimate the difference of brightness of the  $P$  and  $Q$  regions (as indicated in Fig. 1) at different values of edge contrast, contrast polarity, filtered range of spatial scales and stimulus size, and to simulate the psychophysical results with a filter model.

The subjects participated in three or more sessions of 40 trials for each configuration (contrast, contrast polarity, filter frequency) of the test stimulus. Throughout the whole session the test stimulus remained unchanged, while the contrast of the comparison patch was adjusted by the subjects via an adaptive QUEST procedure (Watson & Pelli, 1983) to match the contrast of the test stimulus. For each presentation of test and match stimulus, the subjects selected with a button the stimulus that appeared to have higher brightness contrast.

The fractions of “yes” psychophysical responses as a function of comparison contrast were fitted with a cumulative Gaussian function, and the perceived contrast of the stimulus estimated by the contrast of the comparator at 50% performance.

We chose to display the test and the match in the same presentation, in different regions of the screen. In a preliminary set of data we verified that contrast matches in this condition were similar to those measured when stimulus and match were presented sequentially in the centre of the screen with a delay of 1 s. The subjects were allowed to move their eyes from the stimulus to the comparison patch, and the stimulus remained on the screen until a response button was pressed.

Filtered stimuli often show a sharp rebound closer to the central edges, and ripples in the region  $P$  and  $Q$ , that can be of higher amplitude than that the luminance difference between the two regions. The subjects were explicitly instructed not to rely on these local cues, but to compare the difference of brightness of the two  $P$  and  $Q$  panels. In performing the judgment the subject were required to fixate the central edge and to wait for a stable percept of the two regions to generate. They were asked to reproduce the difference in brightness by varying the single step edge contrast (region  $M1$  and  $M2$ ), again fixating the central edge. The subjects found the task quite easy: the ripple inside the  $P$  and  $Q$  regions posed no particular difficulties, without resorting to “cognitive” strategies, like evaluating the average luminance profile over regions. Informal observations suggest that matching the local edge contrast would lead to different results, in particular for filtered stimuli.

In a preliminary set of data, the brightness of regions  $P$  and  $Q$  in a given stimulus was independently estimated by individually matching  $P$  and  $Q$ —in independent sessions—with a homogenous region. Then, the normalized difference of the matched  $P$ – $Q$  luminance was computed. Given that the two methods yield similar results, we collected the remaining of the data using the simultaneous match with the  $M1$  and  $M2$  region of the step-edge.

Starting from the basic stimulus pattern, four different experiments were performed:

### (1) Non-filtered transparent and opaque stimuli

In a first experiment, the contrast across the  $P$ – $Q$  edge of transparent and opaque stimuli was increased from zero to 29.5% (the same as the  $R$ – $S$  contrast). Stimuli used were either in the “transparent” configuration

(when the  $P$ – $Q$  edge has the same polarity of the  $R$ – $S$  edge), or in the “opaque” configuration (when the  $P$ – $Q$  edge and the  $R$ – $S$  edge have opposite polarity). For the transparent configuration, changes in contrast can also be described in terms of changes in the Metelli transmittance of the transparent layer from 0% to 100%.

### (2) Notch-filtered transparent and opaque stimuli

Three opaque and three transparent stimuli, with  $P$ – $Q$  Michelson contrasts of 5.9%, 17.6%, and 23.5% were filtered with ideal notch-filters of two-octaves bandwidth (examples of the stimuli are reported in Fig. 1D–F for the transparent configuration and Fig. 1G–I for the opaque configuration, see caption for details). For the transparent stimuli, the three contrast values corresponded to Metelli’s transmittance values of 0.2, 0.6, and 0.8, respectively.

### (3) High-pass filtered transparent stimuli

Two “transparent” stimuli, with  $P$ – $Q$  Michelson contrasts of 11.7% and 17.6% (Metelli’s transmittances 0.4 and 0.6) were filtered with ideal high-pass filters.

### (4) Effect of viewing distance and stimulus size

In an additional series of experiments the transparent stimulus of 17.6%  $P$ – $Q$  contrast (transmittance 60%) was presented at different viewing distances and on-screen sizes, to obtain retinal images of the  $P$  plus  $Q$  patch of  $0.47 \times 1.86$  deg (half the size used in the previous experiments—viewing distance 60 cm; patch width 0.48 cm),  $0.93 \times 3.72$  deg (viewing distance 60 cm; patch width 0.98 cm), and  $3.72 \times 14.88$  deg (four times the size of the previous experiments—viewing distance 30 cm; patch width 1.95 cm). The retinal size of the comparison patch regions  $M1$  and  $M2$  was held constant at  $2.75 \times 5.5$  deg each (twice the size used in all other conditions). In all conditions, the stimulus was high-pass filtered and the perceived brightness contrast was measured.

The complete set of data was collected on two subjects, the authors; however preliminary and important data were confirmed in five observers. All subjects had a normal or corrected-to-normal vision.

## 2.3. Determination of the lowest spatial frequency channel in subjects’ contrast sensitivity

We used a standard masking method (Wilson, McFarlane, & Phillips, 1983) to identify the lowest channel in subjects’ contrast-sensitivity, for the luminance level and other setup conditions used in the experiment.

We measured detection-thresholds for a luminance-modulated sine-wave grating oriented vertically and of spatial frequency of either 0.5 or 0.7 c/deg. The grating was presented alone, or superimposed to a sine-wave grating mask of spatial frequencies of 0.5, 0.7, 1, 1.4, 2, 4 c/deg. In order to average out spatial-phase effects, the masking grating differed in orientation by  $\pm 14$  deg. The Michelson contrast of the mask was kept constant at 30%, while the contrast of the test patch was adjusted by the subjects until they reached a visibility threshold. Threshold elevation curves were computed as the ratio of detection thresholds for the masked test, over thresholds for the non-masked test. In order to estimate the actual channel bandwidth from the shape of the curve, it was assumed that thresholds are related to masking contrast with a power function of exponent 0.6 (as computed by Legge & Foley, 1980). Following this assumption, in order to determine the selectivity of the spatial frequency filter, we raised the measured threshold-elevation of the test to the reciprocal of 0.6.

## 2.4. Modelling with local energy

The data were simulated using the output of local energy model implemented at a single scale. The horizontal luminance profile of the stimuli was filtered with a quadrature pair filters, whose amplitude spectrum  $a(u)$  follows a Gaussian function of log spatial frequency, described by the equation

$$a(u) = e^{-\frac{\log^2(|u|/p)}{2\sigma^2}} \quad (2)$$

where the peak frequencies  $p$  was 1.3 c/deg, and the parameter  $\sigma$  was set to 0.55 log-units. for AP, and 0.4 log-units. for MCM. The phase spectra of the filters were constant and equal to 0 for the odd filter and 90 deg for the even filter.

At each scale, a local energy profile was computed as the Pythagorean sum of the outputs of the even- and odd-symmetric detector. Local maxima in the energy function were detected and the energy maximum corresponding to the central edge was directly used to evaluate the local brightness contrast. In very few cases no local energy maximum was present in correspondence to the central edge at the scale of 1 c/deg. In these rare occasions the scale of 2.5 c/deg was used.

The phase value at the location of the energy maximum was computed by analysing the relative outputs of the even- and odd-symmetric detectors.

### 3. Results

#### 3.1. Psychophysics

We measured the perceived brightness difference across the  $P$ – $Q$  edge in the stimuli for different values of contrast, size and contrast–polarity of the central edge and for different values of spatial frequency filtering.

##### (1) Non-filtered stimuli

In a first experiment, we varied the contrast across the  $P$ – $Q$  edge, for stimuli in the “transparent” configuration (when the  $P$ – $Q$  edge had the same polarity of the  $R$ – $S$  edge), and in the “opaque” configuration (when the  $P$ – $Q$  edge and the  $R$ – $S$  edge had opposite polarity).

Fig. 2 plots the perceived brightness contrast of the  $P$ – $Q$  edge as a function of the physical contrast of the edge. Data for transparent stimuli are reported in the left column and data for the opaque stimuli in the right column, for two subjects. The dashed line in this and in the following figures represents the detection threshold for the  $M1$ – $M2$  step-edge, while the continuous curve is the simulation of experimental data with the local energy model (see Section 3.3).

Neither for the transparent nor for the opaque stimulus did the perceived brightness contrast follow the real contrast of the edge (shown by the diagonal dotted line). In particular the perceived brightness contrast was generally lower than physical contrast for the transparent configuration and generally higher for the opaque configuration. This pattern of results was confirmed on other two naïve subjects, whose brightness contrast was maximally affected (maximal ratio of real contrast over measured brightness contrast) at physical  $P$ – $Q$  edge contrasts of 17.6% (subject HA, perceived brightness contrast 10.1%) and 20.6% (subject AV, perceived brightness contrast 9.8%) for transparent configuration stimuli.

The small filled circles for the subject AP (Fig. 2, bottom) show the results obtained when the apparent brightness of each region ( $P$  and  $Q$ ) was matched individually (see Section 2). When the individual matches were transformed in Michelson contrast, the data reproduced closely the same pattern of data obtained by matching simultaneously the two regions. This result validates the technique and shows that the subjects were basing their judgements

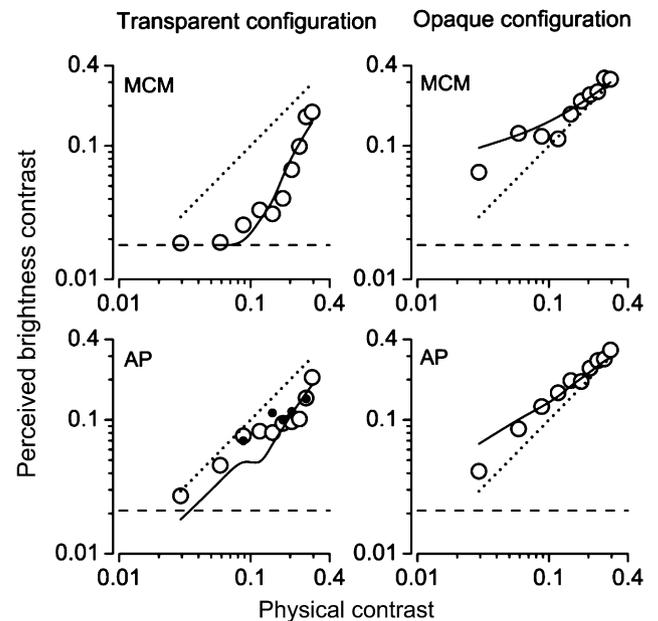


Fig. 2. Open circles report the measured apparent brightness contrast of the stimulus as a function of physical edge contrast, for transparent (left column) and opaque stimuli (right column). The filled small circles for subject AP show the data obtained by matching separately the brightness of the two regions (see Section 2). The thin dotted line marks the real Michelson contrast of the edge. The continuous line is the value predicted by local energy model, as described in the main text. The detection threshold for perceiving the comparison patch is signalled by the horizontal dashed line.

on the apparent luminance rather than on the apparent contrast of the step edge between the  $P$  and  $Q$  regions.

##### (2) Notch-filtered transparent and opaque stimuli

In order to investigate if one spatial-frequency channel is more important than others in determining the contrast of brightness, we notch-filtered at different frequencies three transparent and three opaque stimuli.

Examples of the stimuli used are shown in Fig. 1D–I. Ideal notch-filters introduce ripples and overshoot in the luminance profiles. For example the opaque stimulus filtered at 0.9 c/deg shows clear luminance maxima and minima at the centre of the  $P$  and  $Q$  region (Fig. 1H) that could locally appear lighter or darker of the overall brightness of the single region. However the presence of these ripples did not make the task of the subject particularly difficult or impossible and the subject could easily match the overall brightness contrast of the two regions. Similar effects are present in several other brightness illusion, like for example the missing-fundamental square wave (Campbell, Howell, & Robson, 1971; Maffei, Morrone, Pirchio, & Sandini, 1979). Fig. 3 reports measurements of brightness contrast as a function of the filtered spatial frequency for edges in the “transparent” (columns on the left) or “opaque” (columns on the right) configuration for two subjects.

The stimuli had a Michelson contrast for the  $P$ – $Q$  edge of 5.9% (top row), 17.6% (middle row), and 23.5% (bottom row) and were notch-filtered with an ideal band of two octaves around the central spatial-frequency plotted in

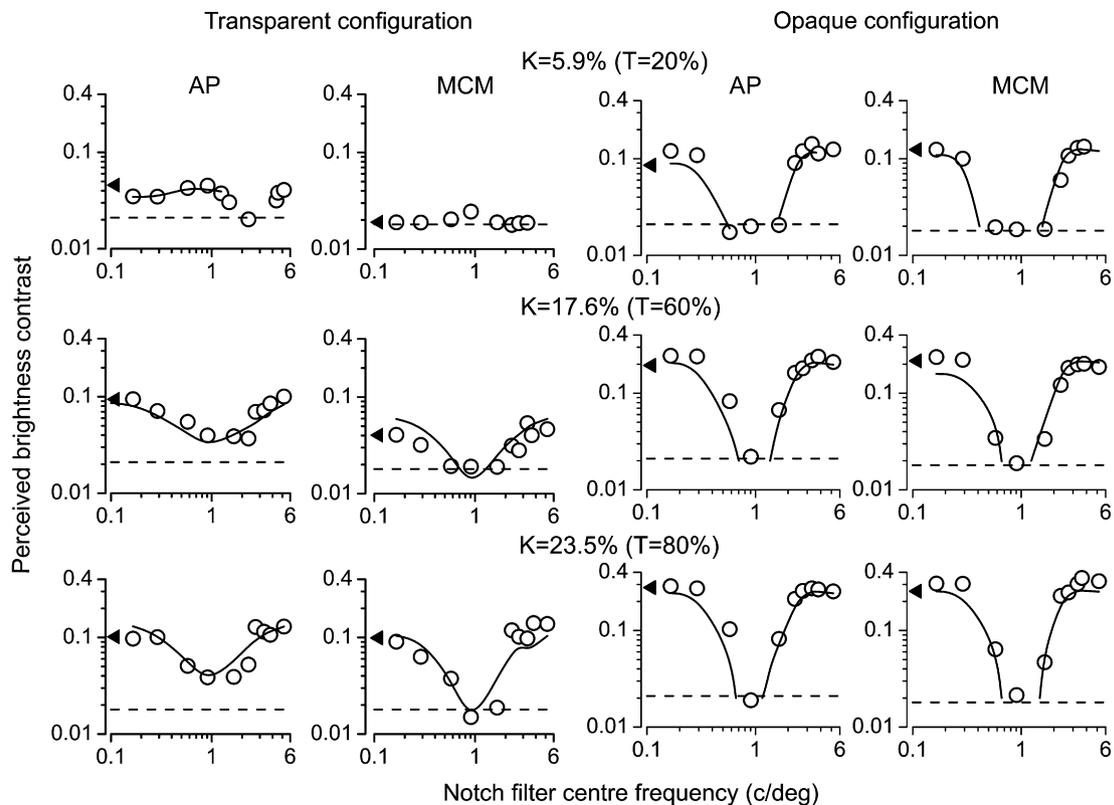


Fig. 3. “Transparent” (left columns) and “opaque” (right columns) stimuli notch filtered in a two-octave band centred on the value in abscissa for two subjects.  $K$  is the Michelson contrast of the non-filtered  $P$ – $Q$  edge;  $T$  is the contrast for each filtered pattern. The function describing the perceived brightness contrast forms a U-shaped curve, with a minimum around 1 c/deg. The triangles mark the threshold for the unfiltered image, the dashed line the detection threshold for the comparison patch, the continuous line the fit with the local energy model.

the abscissa. The measured brightness contrast for the corresponding non-filtered stimuli is indicated in the graphs by filled triangles.

In all stimulus conditions except low contrast “transparent” stimuli, the perceived brightness contrast formed a U-shaped curve presenting a minimum around 1 c/deg. Removal of higher or lower spatial scales did not affect the perceived brightness contrast. Some stimuli at or close to the minima of the U-shaped curves elicited an unstable percept, with rivalry between seeing two regions of different brightness (step waveform) and a roof-like brightness variation (triangular waveform). In these cases the subject was instructed to match the brightness of the step waveform.

With the transparent stimulus of 5.9% Michelson contrast, one subject reached the detection threshold level of the comparison patch at all filtering scales, while data from the subject AP showed a minimum around 2.5 c/deg.

With the exception of only this set of data, the remaining results suggest that the filter tuned at 1 c/deg may be critical for coding the apparent brightness contrast between surfaces. For the contrast and viewing distances used in this experiment, subjects never reported an inversion of the polarity of the opaque stimuli, which sometime occurred at higher contrast.

Data for five naïve subjects (including subjects HA and AV from Experiment 1) with stimuli at 17.6% contrast con-

firmed the major results of Fig. 3 showing a similar pattern (Fig. 4). The perceived brightness contrast of “transparent” stimuli had a maximum reduction for filtering at 0.91 c/deg, and the effect was similar, but stronger for “opaque” stimuli.

### 3.1.1. High-pass filtered transparent stimuli

To explore further the possibility that a narrow range of spatial frequencies around 1 c/deg, rather than high spatial frequencies *per se*, cause the brightness appearance of the stimuli we high-pass filtered stimuli of 11.7% and 17.6% Michelson contrast in the “transparent” configuration (transmittance values of 0.4 and 0.6, respectively).

In this paradigm, we expected the perceived contrast of brightness to be similar to that measured with non-filtered stimuli, as long as information important for brightness computation was not filtered out. The apparent brightness contrast should start to decrease when the energy from the first spatial channel (or whichever mechanism) involved in brightness computation (minus its half-bandwidth) is attenuated and should reach zero when the energy from the last channel involved in brightness computation (plus its half-bandwidth) is totally eliminated.

The results of the contrast matches for these stimuli are reported in Fig. 5 for the two values of edge contrast tested and for the two subjects. The perceived brightness–contrast

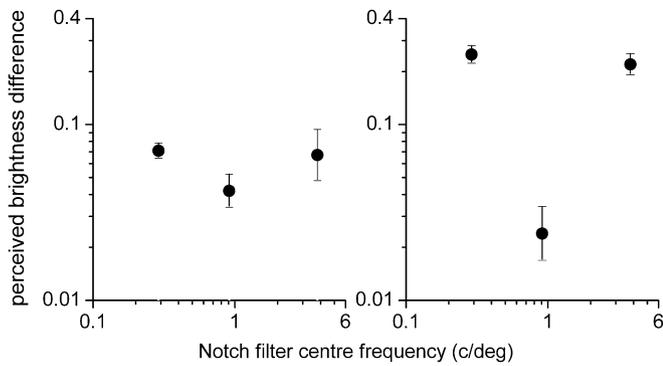


Fig. 4. As Fig. 3 for naïve subjects and the  $P$ – $Q$  edges of  $K = 17.6\%$ . Each symbol shows the average with SEM of five subjects.

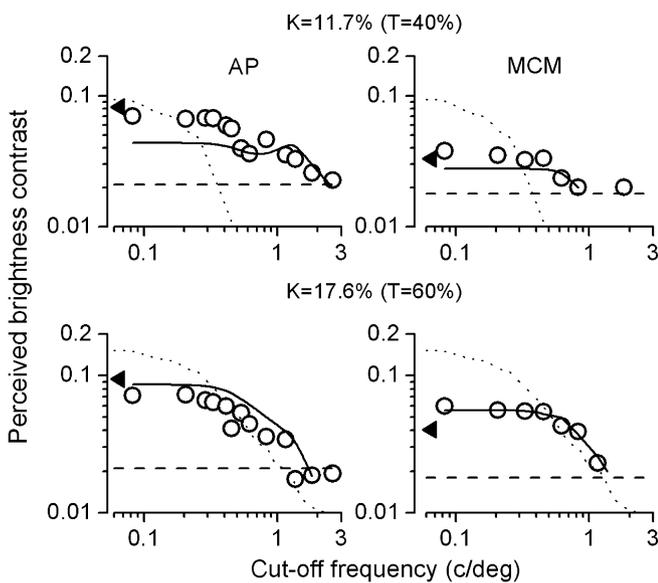


Fig. 5. Perceived-brightness contrast across the stimulus  $P$ – $Q$  edge as a function of high-pass filtering (the value in abscissa represents the starting filtering frequency). The stimuli are in the “transparent” configuration.  $K$  is the physical Michelson contrast of the stimulus;  $T$  is the corresponding Metelli transmittance of the stimulus before filtering. The continuous line reports the prediction of the energy model at 1.3 c/deg, the dotted line the contrast computed from the mean luminance within the  $P$  and  $Q$  regions.

starts to decrease at about 0.5 c/deg, reaching rapidly the detection threshold for perceiving the comparison patch.

(3) *Effect of viewing distance and stimulus size*

The dependency of our data on the critical frequency of 1 c/deg could either be related to the size of the  $P$ – $Q$  stimulus or of the comparison patch, or reflect a basic property of the visual system. To test these possibilities, the stimulus in the transparent configuration, with a  $P$ – $Q$  contrast of 17.6% (transmittance 60%), was presented at different viewing distances and screen sizes, to obtain retinal images of the patch varying from  $0.47 \times 1.86$  deg (half the size used in the previous experiments: Fig. 6, bottom row), through  $0.93 \times 3.72$  deg (Fig. 6, middle row), to  $3.72 \times 14.88$  deg (four times the size of the previous experiments: Fig. 6, top row). The retinal size of the comparison patch regions

$M1$  and  $M2$  was kept constant at  $2.75 \text{ deg} \times 5.5 \text{ deg}$  each (twice the size used in all other conditions). In all conditions, the stimulus was high-pass filtered and the perceived contrast was measured.

Interestingly, despite the wide range in the stimulus sizes used, the perceived brightness contrast shows a similar dependency on the high-pass filtering, starting to decrease around 0.5 c/deg in each of the three conditions.

Doubling the size of the comparison patch did not alter the data, as can be evaluated by comparing the bottom row of Fig. 5 with the middle row of Fig. 6. This result indicates that the non-linear dependence of the apparent brightness contrast on the real contrast of the  $P$ – $Q$  patch is not affected by a non-linear dependence of the appearance of the  $M1$ – $M2$  comparison patch on its physical contrast. This control is important, given that the match patch was presented on a dark surround that in principle could interfere with its apparent contrast (see Section 3.3).

3.2. *Determination of the lowest spatial frequency channel by masking of contrast sensitivity*

The range of frequencies that elicit the strongest effect on perceived brightness is roughly consistent with the peak frequency and bandwidth of the lowest channel in human contrast sensitivity reported in literature.

With a standard masking technique (Wilson et al., 1983), we measured the peak frequency and bandwidth of

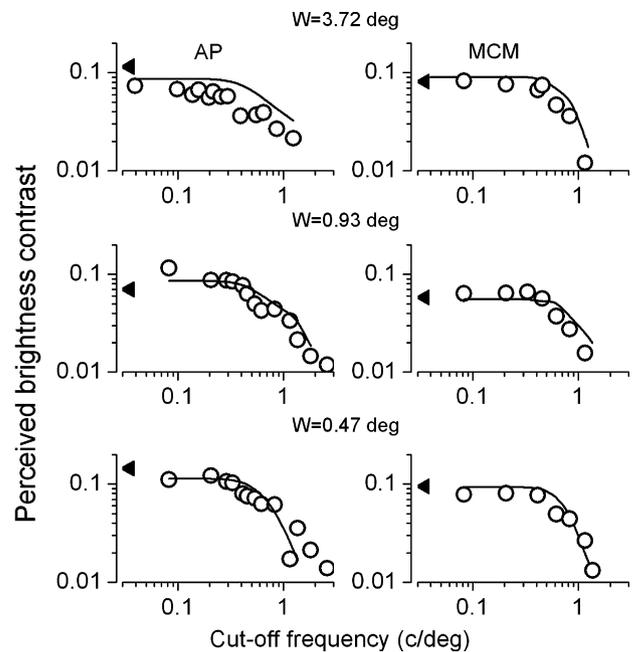


Fig. 6. Apparent brightness contrast of high-pass filtered stimuli in the transparent configuration. The stimuli have a  $P$ – $Q$  Michelson contrast of 17.6% (Metelli transmittance 60%). Each row represents a different stimulus retinal size ( $W$  = width of  $P$  plus  $Q$  patch), obtained by varying both viewing distance and stimulus size. The continuous line is the value predicted by local energy model. Note that the size of the comparison regions  $M1$  and  $M2$  was  $2.75 \times 5.5$  deg each, instead of  $1.38 \times 2.75$  deg.

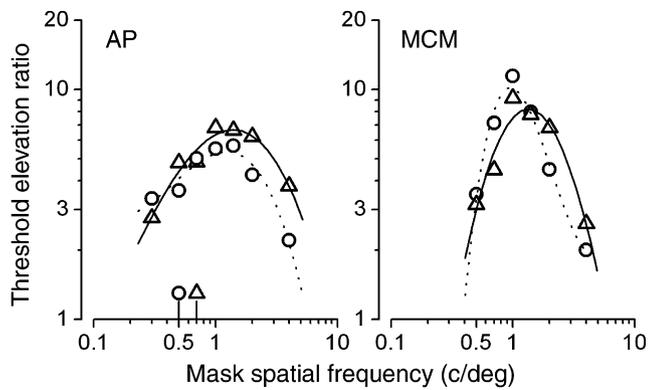


Fig. 7. Masking functions obtained with test frequencies of 0.5 (circles) and 0.7 (triangles) c/deg. For both subjects all curves peak around 1 c/deg, suggesting that the lowest spatial-frequency mechanism is at this spatial scale, at the luminance conditions used in the experiment.

the lowest channel of the contrast sensitivity for the same two subjects (Fig. 7), at the same luminance conditions and long exposure time used in the brightness match experiment. The triangles and circles of Fig. 7 show the elevation curves for two different test frequencies of 0.5 and 0.7 c/deg. For both test frequencies the maximum detrimental effect of the mask occurred around 1 c/deg, indicating that also in the present condition the lowest operating channel peaks at that frequency. The channel bandwidth was estimated from the shape of the curve, as described in Section 2 section. In subject AP the channel had an estimated half-height bandwidth of about 0.7 log-units, while in subject MCM the channel bandwidth was about 0.4 log-units.

### 3.3. Model simulation

The horizontal luminance profile of each stimulus was convolved with even and odd symmetric detectors (see Section 2), with central frequency at 1.3 c/deg and bandwidth corresponding to 0.55 log-units at half-height for AP, and 0.4 log-units for MCM. Both peak and bandwidth were left as free parameters to be determined by minimizing the residual to the experimental data. Interestingly both parameters agree very well with the psychophysical estimate of the lowest channel, particularly for subject MCM.

Only the local energy value at the position of the  $P$ - $Q$  edge and at the scale of 1.3 c/deg was used in the simulation. If the central edge did not determine a maximum in the local energy at this spatial scale at the appropriate position (for instance because the energy peaks determined by the lateral edges masked it), the local energy at the next higher spatial scale at 2.5 c/deg was taken. In practice, for subject MCM this was never necessary; for subject AP the local energy at 2.5 c/deg was used to fit the data of transparent stimuli, with transmittance values ranging from 0.1 to 0.3, both non-filtered and notch filtered (transmittance 20%,  $K$  5.9% graph in Fig. 3).

All data points, both transparent and opaque, are well fitted (continuous line superimposed to all graphs) by the

energy maximum  $E_C$  after rescaling by a multiplicative constant independently determined for the two subjects:

$$K_{\text{app}} = cE_C \quad (3)$$

The constant  $c$  was equal to  $6.25 \times 10^{-3}$  for MCM and  $5.56 \times 10^{-3}$  for AP.

This simple equation captures the essential aspect of all the data obtained for all different filtering frequencies and sizes. For the majority of data points the fit is very good. Somewhat worse predictions of the experimental data were obtained at low contrasts of the  $P$ - $Q$  edge. For these data points, the detectability of the  $P$ - $Q$  edge by the spatial mechanisms tuned around at 1.3 c/deg is close to threshold values of the  $M1$ - $M2$  edge, and for this reason the output values are strongly dependent on small differences in the frequency tuning and channel bandwidth used in the simulation. Given that the subjects' task was to match the  $P$ - $Q$  region with that of the  $M1$ - $M2$  region, it is important to assess that the same model would predict a linear dependence of the  $M1$ - $M2$  matching stimulus on its physical contrast. We used the same filters centred at 1.3 c/deg and applied them to a waveform that reproduced the luminance profile of the comparison patch, the dark surround given by the cartoon windowing and the mean luminance surround given by the filtered  $S$  region. The local energy computed from the 1.3 c/deg channel showed a linear increase with contrast in the whole range of 0.01–1 that was above threshold for both subjects.

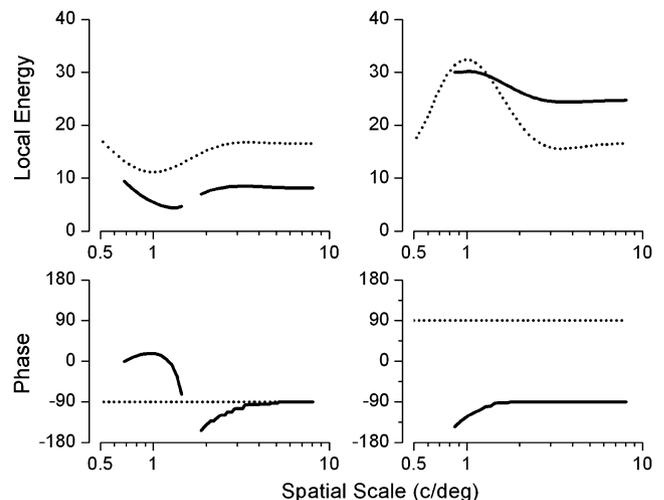


Fig. 8. The values of energy and phase at the position of the  $P$ - $Q$  edge (dotted line), and of the  $Q$ - $S$  edge (continuous line) are reported as a function of the detector spatial scale for stimulus in Fig. 1B (left) and Fig. 1C (right) ( $P$ - $Q$  contrast equal to 17.6%). The curve signalling the edges of the opaque patch are continuous across scales, indicating that both detectors tuned to high- and to low-spatial frequencies coherently signal the feature. In the transparent configuration two distinct curves contemporarily signal the edges of the patch, one at the low spatial scales and one at the high scales. The phase diagram (bottom row) shows that detectors tuned to 1 c/deg signal the lateral edge with a phase close to 0, signalling the presence of a line.

In all the transparent stimuli in the range of transmittance from 0.3 to 0.7 (those for which the transparency illusion is strongest) we observed an intriguing characteristic: the channel tuned to 1 c/deg signals the presence of a feature with a different phase in correspondence with the “lateral edges” of the transparent patch; the higher spatial scales still signal the edge with the correct polarity.

Fig. 8 plots the energy and phase for the central edge and one lateral edge at all scales for one transparent and one opaque non-filtered stimulus.

While simulated receptive fields at all scales coherently signal the local phase of the opaque central and lateral edges, the energy and phase corresponding to the edge of the transparent patch is incoherently signalled from low- to high-spatial scales: low spatial detectors signal features with zero phase corresponding to a line, while higher spatial scales still signal the edge with the real polarity. Interestingly this phase incongruence was present only for  $P-Q$  patches that elicit a transparency illusion. In a  $P-Q$  patch corresponding to a contrast of 5.9 ( $T=0.2$ ) or less, the phases of the lateral edges were constant at  $-90$  deg. If carefully observed with steady fixation, these stimuli sometimes appear to have a luminance overshoot near the edge.

The information from a single channel was sufficient to fit all data, with no need for a pooling of information among channels. However, we also tested other mechanisms of integration, such as probability-summation or averaging across channels. None of the integration mechanisms tested could explain the sharp fall of apparent-contrast when spatial-frequencies around 1 c/deg are removed from the image in both notch-filtered and high-pass filtered stimuli. Indeed, if a pooling mechanism existed, integration of information across two or more scales would compensate for the loss of contrast at the filtered scales, leaving the simulated brightness almost unaffected by notch-filtering, and predicting much longer slopes for the high-pass filtered stimuli. We also tried to replicate the brightness contrast data using a simple linear average of the luminance inside the single  $P$  and  $Q$  region. The dotted curves of Fig. 5 report an example of the results obtained with these simulations and the fit is quite poor for both patch contrasts.

#### 4. Discussion

We measured the perceived contrast of brightness across an edge in a variety of conditions. For notch-filtering, the brightness contrast between regions  $P$  and  $Q$  was strongly altered when spatial frequencies around 1 c/deg were removed from the image, but stayed almost unaffected when lower or higher spatial frequencies were filtered out. When the stimuli were high-pass filtered, the perceived difference of brightness was not affected by the removal of frequencies below 0.5 c/deg, but thereafter started to drop until it reached detection threshold around 1 c/deg. High-pass filtering affected brightness judgements in a similar fashion, independently of edge contrast and patch size, suggesting that the spatial frequency dependency reflects prop-

erties of the visual system and not attributes of the visual image (like the size).

Data of notch filtering suggest that a band-pass mechanism spatially tuned at 1 c/deg mediates the perceived difference of brightness in the stimuli. Removal of lower or higher spatial frequencies did not affect the brightness.

The sharp fall of perceived brightness contrast for high-pass filtering around 0.5 c/deg fully supports this hypothesis, excluding a role of spatial mechanisms higher than about 1 c/deg. If the contribution from high spatial frequencies is present, it must to be limited to small effects, given that the presence in the image of spatial frequencies above 2 c/deg is never sufficient to raise the perceived difference of brightness above threshold. The high-spatial frequency filtering results could leave open the possibility that mechanism tuned to frequencies lower than 1 c/deg could mediate the effect. However, in the same condition of luminance and temporal presentation of the apparent contrast experiment, we could not demonstrate the existence of very low tuned spatial frequency mechanisms, in agreement with previous studies (Wilson et al., 1983). In addition, when spatial frequencies in the range from 0 to 0.5 c/deg are removed from the image, the data of perceived brightness difference are similar to those collected with non-filtered stimuli (marked in the figures by the black triangle).

The measurement of characteristics of the lowest spatial frequency channel of subjects' contrast sensitivity and the modelling of psychophysical results with the local energy model at 1.3 c/deg confirm that the lowest channel to detect the edge can account for the brightness appearance of the pattern. The lowest channel is usually the one around 1.3 c/deg. However, in a few cases (Fig. 3, first row) the channel was at a higher frequency, around 2.5 c/deg. Interestingly, for those data points no energy local maxima at the position of the central edge was found at the spatial scale tuned to 1.3 c/deg, and the corresponding brightness contrast could effectively be simulated using the energy at higher scale.

Letter recognition and face perception (Hayes, Morrone, Burr, & Ross, 1986; Pelli, 1999; Solomon & Pelli, 1994) are optimally performed within a narrow range of spatial frequency content of the stimulus and probably mediated by a single channel. However, the range of important spatial frequencies scales with viewing distance (Hayes et al., 1986) indicating that it is the letter dimension that determines the optimal channel mediating recognition. Also for letter recognition the optimal frequency depends on distance and size, but the scaling is not complete. For the brightness perception of the present study we observed total invariance between stimulus size and critical band: over a range of more than a factor of eight the lowest channel is always mediating the task. This conclusion is in agreement with the recent report by Peromaa and Laurinen (2004) that the illusory scalloping associated with Chevreul staircases is heavily reduced when the low spatial harmonics are masked by band-pass noise, although the edges remain visible. However the study did not address whether the illusion scaled with stimulus size.

Our findings are consistent with models for feature detection that suggest there is no integration of information across different spatial scales, like the local energy model (Morrone & Burr, 1988) and the model by Pessoa, Mingolla, and Neumann (1995). They are in contrast with models that perform feature extraction after integration of various scales, like MIRAGE (Watt & Morgan, 1985) and MIDAAS (Kingdom & Moulden, 1992) or models that summate information across scales (Blakeslee & McCourt, 1999). When the stimuli are notch or band-pass filtered, the missing information at one scale would be compensated in these models by the information still available at lower or higher spatial scales, hence predicting nearly flat curves for notch filtering and slight slopes for high-pass filtering.

Our data suggest that the activity at a single spatial scale can describe the appearance of both transparent and opaque stimuli. Modeling with local energy at the lowest spatial scale provides a good fit to most experimental points, suggesting that the information available at this stage is sufficient to account for the perceived brightness of our stimuli.

For transparent stimuli, but not for opaque stimuli, some discrepancies between the experimental data and simulation with the energy model are observed. It is possible that these discrepancies result from influences of transparency illusion on apparent brightness, not accounted for by the present implementation of the energy model. However, since the bigger discrepancies were found for values near detection threshold of the 1 c/deg channel, it seems likely that a better implementation of the model, that includes contrast gain and a detection threshold of the low-frequency spatial channels would provide a better fit to the data, with no need for top down effects due to the transparency or a lateral interference between edges.

The brightness difference between the  $P$  and  $Q$  regions varies strongly between the transparent and opaque configuration. Simultaneous brightness contrast (SBC) qualitatively predicts a decrease of brightness difference across the  $P$ – $Q$  edge for transparent objects and an increase for opaque objects, based on the luminance relationships present in the stimuli, reflecting what is observed in our stimuli. However, our data differ somehow from what one would expect on the basis of previous quantifications of simultaneous contrast on different stimuli. Namely, the effect of induced brightness is known to be stronger when the difference in luminance between inducing and test field is bigger (Heinemann, 1955). In our data we would expect a maximal discrepancy between experimental data and real contrast of transparent objects at low values of  $P$ – $Q$  contrast, when the inducing contrast of the lateral edges is correspondingly higher. The present data do not confirm this prediction, showing the greatest discrepancy at intermediate contrast values of 0.24 where each subject matched the physical contrast with an attenuation greater than 2.3 (see Fig. 2). This finding is in agreement with the widespread opinion that the brightness appearance of more complex patterns, such as the transparent stimuli used in this experiment, cannot be completely accounted by the

simple summation of individual SBC contributions (e.g., D'Zmura & Singer, 1999).

One possibility is that of invoking an effect of high-level inferences of transparency on brightness perception. Such effects are particularly evident in complex stimuli that facilitate the segregation of perceptual layers. Adelson (1993) described the “argyle illusion”, where two equiluminant regions are perceived as having different lightness as long as a transparency illusion is produced. Modifications to the basic pattern that destroyed the sense of transparency also destroy the lightness effect and vice versa. Recently Anderson and Winawer (2005) introduced a new lightness illusion demonstrating that the computations underlying surface segmentation can cause identical texture patches to appear either dark or bright to a greater extent than what could be explained by contrast effects only, postulating a transparency induction effect. Singh and Anderson (2002) suggested that also phenomena traditionally explained in terms of lateral inhibition (Chubb, Sperling, & Solomon, 1989; Solomon, Sperling, & Chubb, 1993) could be more adequately described as transparency illusion phenomena. In our stimuli, however, the low-level simulation implemented by the local-energy model catches most—though perhaps not all—of the brightness difference between transparent and opaque stimuli, suggesting that with the simple Metelli-like patterns used in our experiment, there is no need to appeal to high-level effects of transparency to explain the appearance of brightness.

Phase incongruence between low- and high-spatial scales was shown to induce a strong illusion of transparency. Morrone and Burr (1997) had shown that a phase shift of  $\pi$  between high and low spatial harmonics of coarse-quantized images induced a perception of transparency, annulling the blocking illusion. These data suggested that a transparency signal might be already detected at an early level of image processing, by evaluating the phase congruency across different spatial scales at the edge positions. In the class of transparent stimuli we studied, different phases of the feature at the external edges of transparent patch are revealed at high and at low spatial scales, providing the visual system with incoherent information about the characteristics of these features. The phase congruency across spatial scales computed in the proximity of important spatial features may play a role in triggering a transparency illusion both in complex scene (Morrone & Burr, 1997) and in simple stimuli such as those used in the present study.

### Acknowledgments

We are grateful to David Burr for fruitful discussion and comments on the manuscript. This research was supported by MIUR PRIN grant.

### References

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

- 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.
- Anderson, B. L., & Winawer, J. (2005). Image segmentation and lightness perception. *Nature*, 434, 79–82.
- Beck, J., Prazdny, K., & Ivry, R. (1984). The perception of transparency with achromatic colors. *Perception & Psychophysics*, 35(5), 407–422.
- Blakeslee, B., & McCourt, M. E. (1999). A multiscale spatial filtering account of the White effect, simultaneous brightness contrast and grating induction. *Vision Research*, 39(26), 4361–4377.
- Campbell, F. W., Howell, E. R., & Robson, J. G. (1971). The appearance of gratings with and without the fundamental Fourier component. *Journal of Physiology*, 217(1), 17–18.
- Campbell, F. W., & Robson, J. G. (1968). On the application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, 197, 551–556.
- 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*. New York: Academic Press.
- Craik, K. (1966). *The nature of psychology*. Cambridge: Cambridge University Press.
- D’Zmura, M., & Singer, B. (1999). Contrast gain control. In L. T. Sharpe & K. R. Gegenfurtner (Eds.), *Color vision: From genes to perception* (pp. 369–385). Cambridge: Cambridge University Press.
- Georgeson, M. A., & Sullivan, G. D. (1975). Contrast constancy: deblurring in human vision by spatial frequency channels. *Journal of Physiology (London)*, 252, 627–656.
- Graham, N. (1977). Visual detection of aperiodic stimuli by probability summation among narrow band channels. *Vision Research*, 17, 637–652.
- Hayes, A., Morrone, M. C., Burr, D. C., & Ross, J. (1986). Recognition of positive and negative bandpass-filtered images. *Perception*, 15, 595–602.
- Heinemann, E. G. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology*, 50, 89–96.
- Kingdom, F., & Moulden, B. (1992). A multi-channel approach to brightness coding. *Vision Research*, 32, 1565–1582.
- Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, 70, 1458–1471.
- Maffei, L., Morrone, M. C., Pirchio, M., & Sandini, G. (1979). A perceptual phenomenon and its neurophysiological correlates. *Perception*, 8, 43–46.
- Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research*, 42(9), 1165–1184.
- Metelli, F. (1974). The perception of transparency. *Scientific American*, 230, 91–98.
- Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase dependent energy model. *Proceedings of the Royal Society (London)*, B235, 221–245.
- Morrone, M. C., & Burr, D. C. (1997). Capture and transparency in coarse quantized images. *Vision Research*, 37, 2609–2629.
- O’Brien, V. (1958). Contour perception, illusion and reality. *Journal of the Optical Society of America*, 48, 112–119.
- Pelli, D. G. (1999). Close encounters—an artist shows that size affects shape. *Science*, 285(5429), 844–846.
- 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*, 15, 2201–2224.
- Shapley, R., & Reid, R. C. (1985). Contrast and assimilation in the perception of brightness. *Proceedings of the National Academy of Sciences of the United States of America*, 82, 5598–5983.
- Singh, M., & Anderson, B. L. (2002). Toward a perceptual theory of transparency. *Psychological Review*, 109(3), 492–519.
- Solomon, J. A. (2000). Channel selection with non-white-noise masks. *Journal of the Optical Society of America A-Optics Image Science and Vision*, 17(6), 986–993.
- Solomon, J. A., & Pelli, D. G. (1994). The visual filter mediating letter identification. *Nature*, 369(6479), 395–397.
- 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.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2), 113–120.
- Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, 25, 1167–1661.
- Wilson, H. R., McFarlane, D. R., & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23, 873–882.