Contrast discrimination in images of natural scenes

Triantaphillidou, S., Jarvis, J. and Gupta, G.

This is an author's accepted manuscript of an article published in the Journal of the Optical Society of America A 39 (6), pp. B50-B64 2022.

The final definitive version is available online at:

https://doi.org/10.1364/JOSAA.447390

Contrast discrimination in images of natural scenes

JOHN JARVIS[1], SOPHIE TRIANTAPHILLIDOU[1]*, GAURAV GUPTA[2]

[1] *University of Westminster, W1W 6UW, London, UK*
[2] *Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

**triants@westminster.ac.uk*

Contrast discrimination determines the threshold contrast required to distinguish between two suprathreshold visual stimuli. It is typically measured using sine-wave gratings. We first present a modification to Barten's semi-mechanistic contrast discrimination model to account for spatial frequency effects and demonstrate how the model can successfully predict visual thresholds obtained from published classical contrast discrimination studies. Contrast discrimination functions are then measured from images of natural scenes, using a psychophysical paradigm based on that employed in our previous study of contrast detection sensitivity. The proposed discrimination model modification is shown to successfully predict discrimination thresholds for structurally very different types of natural image stimuli. A comparison of results shows that for normal contrast levels in natural scene viewing, contextual contrast detection and discrimination are approximately the same and almost independent of spatial frequency within the range of 1-20 c/deg. At higher frequencies both sensitivities decrease in magnitude due to optical limitations of the eye. The results are discussed in relation to current Image Quality models.

1. INTRODUCTION

The spatial contrast sensitivity function (CSF) is a measure for quantifying spatial vision. Typically, the CSF is determined using luminance modulated sine-wave gratings, where visual sensitivity at a given grating frequency is defined as the inverse of Michelson contrast at the threshold of grating detection. Over the years, CSF studies have proved extremely valuable in elucidating basic mechanisms underlying spatial vision in both humans and animals from photon capture through to signal detection [1-9]. The CSF concept has also been adapted to provide a rapid spatial vision assessment tool in clinical studies [10, 11] and applied as a visual weighting factor in models designed to measure the visual image quality of reproduced natural scenes [12-15]. However, an important factor to be considered in our understanding of spatial vision, is that images in the natural environment are considerably more complex in structure compared with traditional stimuli used to determine the CSF and contain clearly visible (suprathreshold) contrast information. As a consequence, the ultimate relevance of the CSF, which is based on threshold vision, as derived from simple sine-wave patterns has been questioned in its ability to provide an understanding of the visual processing of natural images [16]. Further, the use of the CSF as a weighting tool in image quality (IQ) modelling [17-20], which aims to predict the visual quality of reproduced images, has been a subject of much debate.

In an attempt to address some of these reservations relating to the CSF, we previously conducted a study designed to directly measure this sensitivity function from a number of natural images [15]. In this study, measurements of the CSF were obtained through spatial

frequency decomposition of a selection of pictorial scenes into ten 1-octave wide frequency bands. The psychophysical experimentation involved an observer free viewing both a standard image and test image spatially overlapping, but separated in time by 300ms. The standard was defined by the original scene, but with all contrast energy removed from the frequency band from which contrast sensitivity was to be determined. The test image was similarly defined, except that now, the band contrast could be increased to the point where a difference in appearance between standard and test could just be perceived. This threshold test band contrast was defined in terms of the RMS contrast metric, this being considered an appropriate band contrast measure for complex imagery, as discussed in our previous communication [15]. The inverse of the threshold RMS contrast value defines band contrast detection sensitivity. In addition, measurements were taken with each standard and test image defined from a single frequency band with no other band information present; a condition close to that associated with the use of a simple sine-wave grating stimulus.

The results showed that the spatial contrast sensitivity profile obtained from individual frequency bands presented within the complete image (the *contextual* CSF, cCSF) displayed a slow increase in sensitivity from low frequencies up to a turning point of around 16 c/deg, whereas the isolated band contrast sensitivity function (iCSF) showed the classic band-pass profile so characteristic of measurements obtained from simple grating stimuli. The mechanistic Barten model for sine-wave contrast detection [6] offered a good fit to the iCSF data, and we derived a further model that could satisfactorily predict cCSF using both the iCSF and the image contrast spectrum as input. We argued that spatial noise generated from frequency bands at +/- 1-octave from a specific band was the major factor in determining its contextual detection level [15].

Our work has now been extended to include the measurement and modelling of contrast discrimination in natural scenes. Contrast discrimination measures the threshold contrast (just noticeable difference) required to distinguish between two visual suprathreshold stimuli. Changes in suprathreshold luminance contrasts are basic building blocks leading to the discrimination of shape, form and detail. This paper addresses the important issue of sensitivity to changes in suprathreshold contrast in natural scene viewing. Modelled contextual contrast discrimination functions, describing *suprathreshold contrast sensitivity* at a given spatial frequency in the presence of suprathreshold signals at other frequencies, are also compared with contextual contrast sensitivity (detection) functions (cCSFs) derived in our previous work [15]. These latter functions are fundamentally different from discrimination functions (i.e. they describe the *absolute threshold contrast detection* at a given spatial frequency (from no signal to the minimum signal detection). From this comparison conclusions are then drawn regarding the importance of contrast discrimination sensitivity compared with contrast detection sensitivity in the understanding of how the human visual system responds to structural changes in natural scenes.

The following sections first discuss the general concept of contrast discrimination and its modelling. We then present an extension of Barten's contrast discrimination model that accounts for spatial frequency effects and test this using visual data from classical discrimination studies. Further on we present measurements of contrast discrimination functions obtained from natural image stimuli at two different contrast levels; results are compared against our model predictions. Modelled contextual discrimination functions are also compared with contextual detection sensitivity functions derived in our earlier work. Finally, we discuss our findings and their implications in IQ modelling.

1.1 The dipper function

As in the case of the CSF, most previous studies of basic contrast discrimination have involved the use of sine-wave grating stimuli [21-29]. For contrast discrimination, a just noticeable difference in contrast has to be perceived between two nearly identical suprathreshold gratings (test and standard stimuli) that differ only in Michelson contrast value. When this occurs, the threshold level for discrimination is then defined as Δc, the algebraic difference between test and standard grating contrasts. The fixed contrast level of the standard is usually termed the "pedestal" contrast. Note that a value of Δc for zero pedestal contrast represents the contrast detection level.

The conventional way of presenting contrast discrimination data is by plotting Δc versus pedestal contrast at a given spatial frequency u c/deg. It is found that for a grating with a mean luminance above about 2.0 cd/m^2 and spatial frequency above 1.0 c/deg, as pedestal contrast increases from zero, values of Δc remain fairly constant and then start to decrease to a minimum value at a pedestal contrast of around 0.01 [29]. The function then has a turning point and as pedestal contrast level increases, Δc increases with a relatively constant slope. From about a pedestal contrast of 0.5 and above, the function then shows a slow decrease in slope. This general profile is observed in most contrast discrimination studies and has led to the term "dipper" function. In other words, the "dip" in the function shows that as pedestal contrast increases from low levels, discrimination sensitivity (as defined as the inverse of Δc) becomes higher than that at the absolute threshold detection level and then progressively decreases. Examples of this dipper function (which were produced from the model explained and discussed in Section 2) are shown by the continuous curves in Fig. 1(a,b,c). If the dipper function is determined in the presence of 2D spatial noise, low pedestal contrast values of Δc increase significantly and the "dip" point shifts to a higher pedestal value [6].

1.2. Spatial frequency effects

Early contrast matching experiments using suprathreshold sine-wave gratings showed that perceived contrast was relatively independent of spatial frequency, providing Michelson contrast was above 0.1 [2, 30]. This form of visual adaptation to contrast was later confirmed from other experimentation [31, 32]. Bradley and Ohzawa [28] determined sine-wave contrast discrimination as a function of spatial frequency at a fixed pedestal contrast of 0.25 and obtained a curve of slightly flatter profile than their companion detection threshold (CSF) curve. This form of visual adaptation to contrast, is referred to as contrast constancy. Contrast constancy can be considered as a form of contrast gain control, which has also been advanced to account for the phenomenon of *contextual modulation* in contrast perception [33]. This phenomenon measures how both threshold contrast detection [34-36] and suprathreshold apparent contrast [37, 38] can be suppressed, or enhanced when the sinusoidal grating is presented within a surrounding image. Contrast constancy is an important feature when studying the perception of real complex images and plays a significant role in visual IQ assessment.

Previous studies suggest that as spatial frequency increases from about 1.0 c/deg, the overall slope of the dipper function remains approximately constant. Both Legge [24] and Bradley and Ohzawa [28] found that in the spatial frequency range 1.0 to 16.0 c/deg, the initial slope of the function for pedestal contrast values above the dip, was consistently around 0.9; a result close to Weber law behaviour. Yang and Makous [29] obtained slightly lower slopes of 0.77 and 0.65 at 6 and 2 c/deg respectively. As spatial frequency reduces below 1.0 c/deg, and pedestal contrast increases above about 0.2-0.3, a substantial reduction in the dipper slope occurs.

Values of slope of around 0.4 at 0.8 c/deg [27], and 0.2-0.3 at 0.25 c/deg [24, 39] have been obtained. Kulikowski [39] argued that the shallower slopes at very low frequency are because detection is mediated by transient mechanisms, which do not adapt and have shallower increment threshold functions compared to high frequency sustained mechanisms. Legge [24] challenged the detail of this explanation, but agreed that the differences between low and high frequency dipper slopes probably do relate to operating differences between the two types of mechanism [40, 41]. Certainly, it is now well established that the magnocellular (transient) pathway dominates spatial information transfer at frequencies below 1.0 c/deg [42].

1.3 Contrast discrimination in natural scenes

Of all the determinations of the dipper function conducted over the last 50 years, only a few have used natural images. These include the study by Bex, Mareschal and Dakin [43], who employed a derivative of the Legge and Foley technique [25] to determine contrast discrimination in images of four natural scenes. In their work, spatially filtered images were presented in a number of different ways. One modification involved the presentation of a 1-octave wide, spatially band-filtered image of a given scene, where three peak frequencies were chosen in the filtration process. A second experimental modification was made, similar to the band-pass condition, but in that spatial frequencies outside of the band were not discarded, i.e. the band-image was presented in the context of the remaining frequency bands. Contrast discrimination data were measured at three spatial frequencies (1.0, 2.0 and 4.0 c/deg). Dipper functions were obtained from this experiment, with overall profiles similar to those obtained from sine-wave grating experiments.

Tolhurst et al [44] measured sensitivity to increments of contrast in a range of natural scenes and again dipper functions were used to express the data. This particular study related contrast discrimination to firing rates in striate cortex neurones.

1.4 Contrast discrimination models

Of all the models generated over the years to account for the profile of the dipper function, the one which has received most attention is that proposed by Legge and Foley [25]. This incorporates 5 stages of signal processing: an early linear filter, followed by a non-linear transduction process, with final stages which include the addition of neural noise and eventual decision making. A similar cascade model appeared at the same time by Wilson [45]. Whittle [46,47] suggested that the earliest stage in the cascade must in fact be a non-linear process, a concept later supported by the work of McIlhagga [48] in his study of sine-wave contrast discrimination. Kane and Bertalmio [49] re-examined the data set generated by Whittle and developed a model construct which incorporated a divisive contrast gain control mechanism.

In complete contrast to these models, Pelli [27], described the nature of the dipper function in terms of uncertainty. Here, it is supposed that the brain monitors many channels of information, of which only a small group are relevant to the detection of a contrast change. One plausible mechanism is where the pedestal allows the observer to ignore some of the irrelevant information channels, resulting in a decrease in the uncertainty level associated with the signal detection process. Blackwell [50] employed aspects of Pelli's uncertainty model to account for improvements in contrast detection when measured in the presence of low contrast noise. Sanborn and Dayan [51] formulated yet another alternative dipper function model based on Bayesian inference.

The different attempts at modelling contrast discrimination using either the non-linear transduction or the uncertainty concept have produced radically different mathematical

constructs. The models vary considerably in the number and meaning of equation parameters and a detailed side by side comparison is beyond the scope of this communication. A detailed comparison of these model genres is, however, available elsewhere [52]. It is important to note however that, none of these model investigations have resulted in any firm conclusions being drawn regarding the actual visual mechanisms underlying contrast discrimination. This has resulted in much debate on the issue [44, 53-58].

1.5 Barten contrast discrimination model

Although most of the models mentioned in section 1.4 have been successful in describing the general nature and profile of the dipper function, they are all examples of a particular genre which is limited in its ability to link through to underlying neurophysiological mechanisms. Barten [6], in his theoretical treatment of spatial vision, has produced models for both contrast detection and contrast discrimination, the former being used in our detection study [15].

Barten's detection model is characterised by the incorporation of the key physiological loci in the visual system, which impact contrast detection sensitivity. Specifically, stages quantifying optical attenuation, quantum efficiency, receptor sampling, lateral inhibition, area signal summation effects, neural noise and final cortical detection are incorporated into the model framework. The model is both mechanistic and predictive, minimally relying on the use of measured contrast sensitivity data in its formulation. Values of equation parameters in the model are in most instances, derived directly from both physical and electrophysiological measurements. Barten demonstrated the predictive power of this model using measured contrast sensitivity data obtained from a substantial number of independent studies [1, 2, 59-70]. Further, Barten [71,72] and others [17,73-75] showed how his detection model can be incorporated in predictive models of subjective (visual) image quality.

As shown in section 2.1, Barten's contrast detection model defines the foundation for his contrast discrimination model and as a consequence, this latter model also directly links through to underlying visual neurophysiology. Two basic features of the discrimination model are the embodiment of a non-linear transduction process in the visual system and the concept that discrimination is simply a threshold *detection* process in the presence of a noise source, namely the pedestal contrast itself. It is this model variant that we extend and use in relation to our contrast discrimination data obtained from natural scenes.

The derivation of the Barten contrast discrimination model is outlined and discussed in detail in Appendix A. For the case when test and standard stimuli are both sinusoidal and have the same spatial frequency, u c/deg, the generalized Barten discrimination model for Δc may be written as:

$$\Delta c(u) = \left[\left(\frac{c_o^2(u)+0.04k^2c_s^2(u)}{\Psi^2(u)}\right) + c_s^2(u)\right]^{0.5} - c_s(u) \quad (1)$$

where

$$\Psi(u) = \left[1 + 0.004k\,\frac{c_s(u)}{c_o(u)}\right]^{0.5} \quad (2)$$

In equations 1 and 2, c_s denotes the Michelson contrast value of the standard (defining the pedestal value). The term c_o is the threshold level for contrast detection and as a function of spatial frequency, the value of $(c_o)^{-1}$ defines the CSF. The function $\Psi(u)$ represents the operation of a non-linear transduction process at high pedestal contrast values. The constant k represents

the Crozier factor which is considered a basic signal/noise ratio for detection [76-80] and has a nominal value of around 3.0 [6]. At any fixed value of u c/deg, equation 1 defines a dipper function.

Barten showed the dipper represented by equation 1 to be a good fit to measured discrimination data obtained from a number of published sine-wave grating studies at fixed pedestal frequencies above 1.0 c/deg. [6]. Moreover, for discrimination data measured with a 4 c/deg grating presented in the presence of spatial noise [81], the model again offered good predictions. In these particular model calculations (see Appendix A), c_o was replaced by $(c_o)_n$ where:

$$\left(c_o(u)\right)_n = [c_o(u)^2 + k^2 c_n(u)^2]^{0.5} \tag{3}$$

The term $(c_o)_n$ denotes the contrast detection threshold level in the presence of a noise source, with an average sine-wave component contrast c_n.

2. METHODS

2.1 Barten model extension

None of the discrimination models outlined in Section 1.4, nor the Barten model described by equations 1 and 2 can account for the reduced dipper slope at low frequencies and high pedestal contrasts. Whatever the exact nature of the visual mechanism responsible for this dipper behaviour, it is mathematically characterised by the combination of both a contrast and a frequency response function; the first operating to substantially lower the dipper slope at pedestal contrasts above around 0.2-0.3 and the second to predominantly restrict the slope reduction to below 1.0 c/deg. As now shown, both of these response characteristics can be successfully modelled by the inclusion of two Butterworth squared functions [82] (designated B and B^* in equations 4-7), into the function Ψ, which controls the dipper behaviour at high contrasts. Accordingly, Ψ is re-defined as:

$$\Psi(u) = \left[1 + \{f(c_s(u))B^*(u) + 0.004\}k\frac{c_s(u)}{c_o(u)}\right]^{0.5} \tag{4}$$

where

$$f\left(c_s(u)\right) = \beta\left[1 - B\left(c_s(u)\right)\right] \tag{5}$$

and

$$B\left(c_s(u)\right) = \left[1 + \left(\frac{c_s(u)}{c^*}\right)^{2n}\right]^{-1} \tag{6}$$

The term β represents a dimensionless non-zero scaling parameter and c^* is a contrast constant. The function B^* is given by:

$$B^*(u) = \left[1 + \left(\frac{u}{u^*}\right)^{2m}\right]^{-1} \tag{7}$$

where u^* is a spatial frequency constant.

2.1.1 Initial model testing

This modification to the Barten model was initially tested using sine-wave contrast discrimination data from the Legge [24], Yang & Makous [29] and Bradley and Ohzawa [28] studies. The results of this comparison are shown in Fig. 1(a,b,c,d). The symbols shown in Fig. 1(a) are measured contrast discrimination values from the Legge study at a pedestal frequency of 0.25 c/deg. Symbols shown in Fig. 1(b) are values from the Yang and Makous study for a pedestal of 0.8 c/deg. The continuous curves in Fig. 1(a,b) represent dipper functions, calculated using the modified Barten model (equations 1 and 4), with k fixed at its nominal value 3.0. Values chosen for the B and B^* parameters (n, m, u^*, c^*) are (2.0, 1.0, 3.0 c/deg, 0.2), with β set at 0.02 (Legge data) and 0.015 (Yang and Makous data). The similarity of the value of β (which through our work has ranged between 0.01 and 0.02 for isolated signal viewing and, although larger, has remained stable also for contextual viewing) is important for the general application of the model (see Table 1, section 3.2).

The low frequency results given in Fig. 1(a,b) show that the modified model yields good predictions of the measured data from both studies, using the same set of equation parameter values. With these values, the appropriate criteria described above for both contrast and frequency response functions are fulfilled, i.e. f starts to significantly decrease in magnitude as pedestal contrast reduces from below about 0.2 and accordingly, the product $f.B^*$ tends toward zero with equation 4 becoming equivalent to equation 2 (the original Barten formulation for Ψ, equation 2). The function B^* diminishes from its maximum scaling value of β as spatial frequency increases above 1.0 c/deg, with again the product $f.B^*$ tending toward zero. The hatch curves in Fig. 1(a,b) represent model calculations using the original Barten formulation (equations 1 and 2) which show a poorer fit to the data at higher pedestal contrasts. RMSE values calculated to quantify the goodness of the model fits, for the original Barten and the modified models, confirm this (RMSE 0.0144 (original) vs 0.0066 (modified) in Fig. 1(a); RMSE 0.0158 (original) vs 0.0021 (modified) in Fig. 1(b)). This expected reduction in slope in the dipper function at low frequencies and high pedestal contrasts can only be achieved with the modified model.

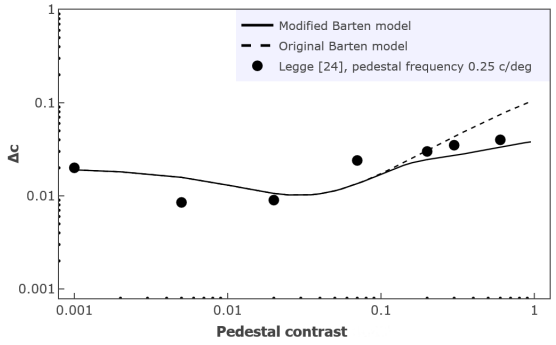
Fig. 1(c) shows the modified model prediction for contrast discrimination data obtained at 2.0 c/deg (Yang & Makous [29]) and 4.0 c/deg (Legge [24]). In the model calculations, c_o was estimated by extrapolation from the measured contrast discrimination value at the lowest pedestal. The single curve represents model calculations using the modified Barten model with β equal to 0.02 (for Legge data [24]) and 0.015 (for Yang & Makous data [29]). Note that the two model calculations yield the same curve. This is because for spatial frequencies above about 1.0 c/deg the term $f.B^*$ is near to zero and so the modified model becomes equivalent to the original version and independent of β.

The data points in Fig. 1(d) represent the measured CSF and the companion contrast discrimination function obtained by Bradley and Ohzawa [28]. The upper continuous curve represents the Barten CSF model [6] fitted to the data. The lower curve is obtained from our modified Barten discrimination model, with parameters (n, m, u^*, c^*) set at the *same* values used in the dipper modelling shown in Fig. 1(a,b,c) and with values of β and k of 0.01 and 4.0 respectively (i.e. a little higher than the nominal k value, but well within the range suggested by Barten, i.e. 2.7 -4.5) [6, p. 59] . In the calculation of discrimination, values of c_o were defined from the Barten CSF and c_s fixed at 0.25.

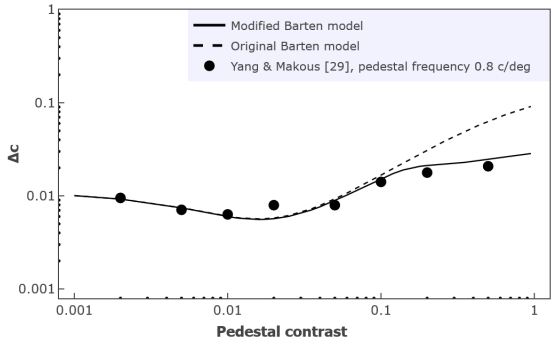
The constancy found for the equation parameter values in this examination of three experimental studies is extremely important in a theoretical model of this type. It is a necessary feature which enables performance trends to be predicted without extensive curve fitting

procedures, which is the case for many computational, or black box models that rely on large amounts of data for fitting purposes. This examination of previously measured sine-wave contrast discrimination data strongly indicates that the modified Barten model can offer good performance predictions with the use of consistent parameter values in equations 4-7 at all pedestal frequencies.

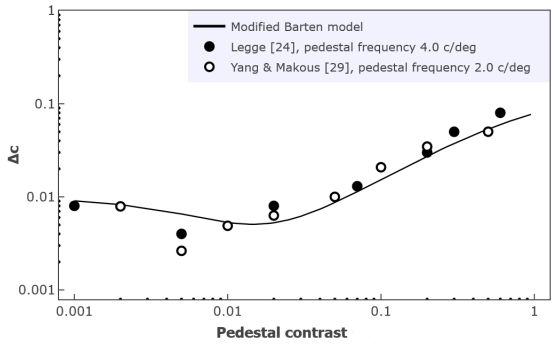
(a)



(b)



(c)



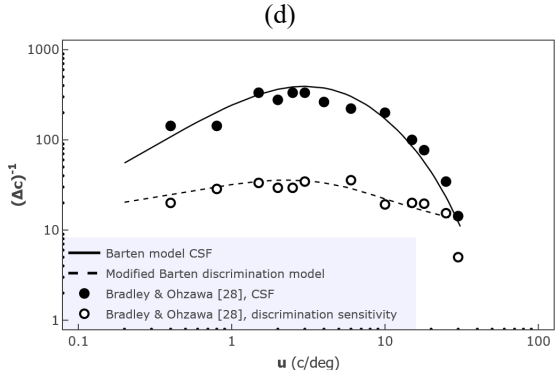


Fig. 1. Comparison of contrast discrimination data Δ*c*, from classical literature studies presented in [24] (a), [29] (b) and [24, 29] (c), together with our modified Barten contrast discrimination model predictions. (d) Measured contrast detection and discrimination data obtained from [28] versus the Barten CSF model and the modified Barten discrimination model.

Note that the scaling factor β is particularly important for obtaining an absolute fit between the model and measured contrast discrimination data. Since the numerical value of β was empirically adjusted to give a good fit to the data obtained from three distinctly separate studies, the expectation would be that some variation in its value would occur. In this initial testing of the model, the fact that this variation was found to be relatively small (β= 0.020 (for Legge data [24]), β=0.015 (for Yang & Makous data [29]), β=0.010 (for Bradley & Ohzawa data [28])), is particularly important in demonstrating the model's predictive stability. A summary of the β and k values used for these sine-wave studies, using Michelson contrasts, is given in Table 1(a).

The following sections will explore the validity of the model for predicting contrast discrimination using natural images. In both our studies with natural images (the current discrimination study and detection study in [15]), the RMS contrast metric was used. The choice of the RMS metric for such work is fully discussed in [15]. For sinusoidal gratings, comparing Michelson with RMS is not specifically a problem, the former measure is simply $\sqrt{2}$ RMS [83] resulting to only shifting the functions by a factor of $\sqrt{2}$ without any change in their shape. Nonetheless, it is necessary to be cautious when comparing studies with different contrast measures.

2.2 Stimuli and their features

The image stimuli used in our measurements of contrast discrimination from natural scenes were the same three employed in our previous study of contextual contrast sensitivity functions derived from natural scenes (cCSF) [15]. Full specification of the images, their capture and image processing are detailed in this communication.

Fig. 2(a,b,c) shows the images, designated *gallery* (containing angular and linear structural features), *park* (defined by open natural spaces, with foreground and background detail) and *people* (structurally complex group scene) respectively; all three offering substantial differences in structural composition. Two overall pictorial contrast levels designated "normal" and "low" were chosen for the experimentation as per [15]. The low contrast version was defined from band contrast values of just half those for the normal image; their mean image luminance was kept constant.

Fig. 2. Normal contrast version of the test images: (a) *gallery*, (b) *park*, (c) *people* [15].

Briefly, filtering techniques used for spatial frequency decomposition employed Peli's cosine log filters [84] of 1-octave bandwidth, centered at frequency 2_i cycles per picture (*cpp*). The filters satisfied our requirements of symmetrical shape on a log frequency axis and the condition of additive reconstruction [15]. Ten filters, centered at the corresponding retinal frequencies of 0.125, 0.25, 0.50, 1, 2, 4, 8, 16, 24 and 32 c/deg were produced for the purpose. They have a number of convenient properties and have been employed in a variety of digital imaging applications [84-88].

The RMS band contrast spectra of the images, calculated as per [15], reduced with increasing spatial frequency according to the inverse power law function (*park*) or a logarithmic function (*gallery* and *people*). Fig. 3 shows band contrast spectra for the normal image versions.

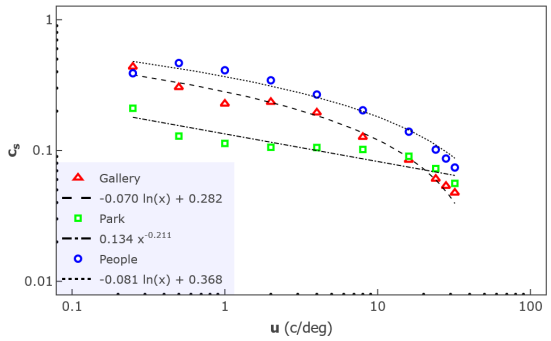


Fig. 3. *RMS* band contrasts, c_s, of normal versions of the three stimuli used in the contrast discrimination experiments. The dashed lines indicate regression curves, used in the modeling of contrast discrimination.

2.3 Stimulus presentation and psychophysical paradigm

The modified adaptive staircase method used to determine contrast thresholds in this study, is described in full detail in our previous communication [15, Appendix C]. The psychophysical display and observer selection are also discussed in detail in [15].

In summary, visual experiments were carried out using two identical high quality, fully characterized, wide gamut 24″ EIZO LCDs. The displays incorporated digital uniformity equalizers, ensuring luminance uniformity across the screen. The display settings and stimulus parameters are summarized below:

- Display pixel dimensions: 1794 (H) x 1196 (V) pixels
- Display mean luminance: 55 cd/m^2.
- Display output: 10 bits per channel, from 16-bit per channel input images, linearized via LUTs so that each color channel was displayed with 1024 equally spaced luminance steps.
- Display temporal frequency: 60 Hz.
- Display pixel pitch: 0.270 mm square.
- Maximum display spatial frequency: 1.85 mm^{-1}.
- Observer viewing distance: 1800 mm.
- Maximum retinal frequency of 58 cpd.
- Stimulus pixel size: 1794 (H) x 1196 (V) pixels (full screen).
- Stimulus field size: 16.5 (H) x 11.0 (V) deg.
- Stimulus mean luminance: image dependent, varied between 14 cd/m^2 and 20 cd/m^2.

The same observer pool from our previous study [15] was used also in this study. Observers were given initial training with each image stimulus and all frequency bands, so that they (i) fully understood the task, (ii) got used to the free-viewing mode that was employed in the experiments. Examination of the intra-class correlation coefficient in software R led to the selection of 10 observers, a mixture of males and females, with ages ranging from 18 to 35 years old.

The filtering and preparation of the image stimuli is the same as in our work on the detection (CSF) studies; it is described in detail in [15]. However, the presentation of the stimuli in the current work on the discrimination sensitivity was different from that study (i.e. in discrimination studies, both reference and test stimuli are suprathreshold, in detection the reference is subthreshold). In the current study, for defining the *isolated band contrast discrimination sensitivity*, the standard consisted of the band of interest in its full contrast, presented in isolation (i.e. outside the remaining image contents, within a uniform field of mean luminance equal to the mean luminance of the image). The test was originally defined as identical to the standard, then the RMS contrast of the test band selected for investigation was adjusted (via the adaptive staircase method [15]) until the observer just perceived a difference between test and standard. For defining the *contextual band contrast discrimination sensitivity*, the standard consisted of the complete original (unfiltered) image. The test was again initially defined as identical to the standard, then the RMS contrast of the selected test band was adjusted until the observer just perceived a difference between test and standard images.

This way, for any given frequency band, the *RMS contrast discrimination threshold,* Δc, was identified in the absence of the contrasts in the remaining frequency bands of the image (isolated discrimination), or in their presence (contextual discrimination). Fig. 4 (a,b) illustrates

an example of a standard and a test stimulus respectively for the contextual discrimination experiment.

(a) (b)

Fig. 4: (a) Standard and (b) test stimulus example for the contextual discrimination experiment, test image *park*.

During the experiments, the observers had their head positioned on a chin rest. They free-viewed the displayed images by only moving their eyes. Standard and test stimuli were displayed at full display, in random order in alteration for 3 seconds each, with a temporal separation of 300 msc and a mid-gray screen displayed between stimuli. They continued being displayed in alternation until the observer chose to answer *"yes"* or *"no"* to the question "*are the two images different?*". Observers operated blindly two keys on a nearby keyboard for answering.

Measurements were obtained from a number of 1-octave wide bands with central frequencies ranging from 0.1 to 32 c/deg. Band contrast discrimination sensitivity was defined as the inverse of the RMS contrast discrimination threshold (i.e. Δc^{-1}). The standard band contrast in this study with complex images is equivalent to the pedestal contrast (c_s) in a sine-wave discrimination experiment.

Δc^{-1} values for each spatial frequency in the resulting functions were averages from minimum 3 and maximum 7 observations. The standard error, which accounts the number of observations, was calculated for each point; it is presented as error bars in all relevant figures.

3. RESULTS

3.1 Isolated contrast discrimination

To supplement the testing of our modification to Barten's model, we obtained a limited amount of contrast discrimination data from a scene containing isolated pictorial bands, which as described in the introduction, produces a stimulus close to an individual sine-wave grating. Data were obtained from just the *gallery* image, simply to determine if the dipper-predicted low frequency rise in discrimination sensitivity at high contrast band levels was still evident. The results are shown in Fig. 5, where data points are measured values of *isolated band contrast discrimination sensitivity* $(\Delta c_i)^{-1}$ as a function of spatial frequency, obtained from both normal and low contrast versions of the image.

For a given spatial frequency above 1.0 c/deg, contrast discrimination sensitivity is lower for the normal contrast image, as predicted from sine-wave dipper behaviour. As spatial frequency reduces below 1.0 c/deg, discrimination sensitivity for the normal contrast image progressively

moves toward that for the low contrast version. This result is consistent with the dipper results shown in Fig. 1(a,b) for the two low frequency sine-wave cases. The solid and hatch curves in Fig. 5 are modelled values of contrast discrimination sensitivity obtained from the modified Barten model (equations 1 and 4), for the normal and low contrast versions respectively.

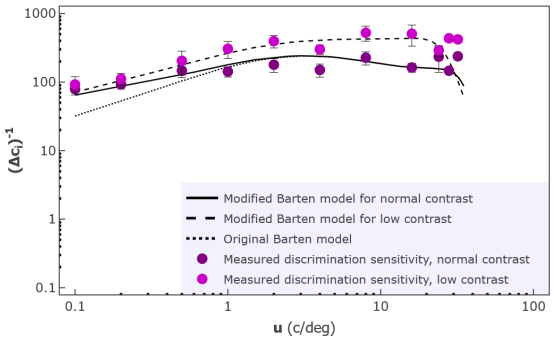


Fig. 5. Measured and modelled isolated band contrast discrimination sensitivity, $(\Delta c_i)^{-1}$, for the normal and low contrast versions of scene *gallery*.

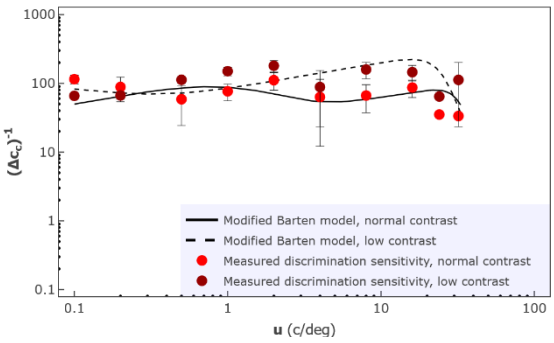
Parameter values (*n, m, u*, c**) for the evaluation of Ψ are the same as used for the modelling shown in Fig. 1. The value for β is 0.015, the same as for the modelled dipper data from the Yang and Makous study (Fig. 1b) and k is 2.8 (see Table 1(a, b) for comparisons). Required values of c_s used in the modelling were obtained from the regression line fit to the *gallery* contrast spectrum shown in Fig. 3. We chose the regression line curve to represent c_s for mathematical consistency within equation 1, where c_o was defined from the Barten CSF appropriate for natural images [15] and both B and B^* are all continuous Butterworth squared functions (see section 2.1).

The modified version of the Barten discrimination model is seen to offer a good fit to the data of Fig. 5 for both normal and low contrast images at all spatial frequencies. The dotted curve shows discrimination sensitivity calculated from the original Barten formulation (equations 1 and 2), which undershoots the normal contrast level data below 1.0 c/deg.

3.2 Contextual contrast discrimination

The data points in Fig. 6(a,b,c) show measured *contextual band contrast discrimination sensitivity* $(\Delta c_c)^{-1}$ as a function of band spatial frequency for all three images, for both normal and low contrast versions.

(a)



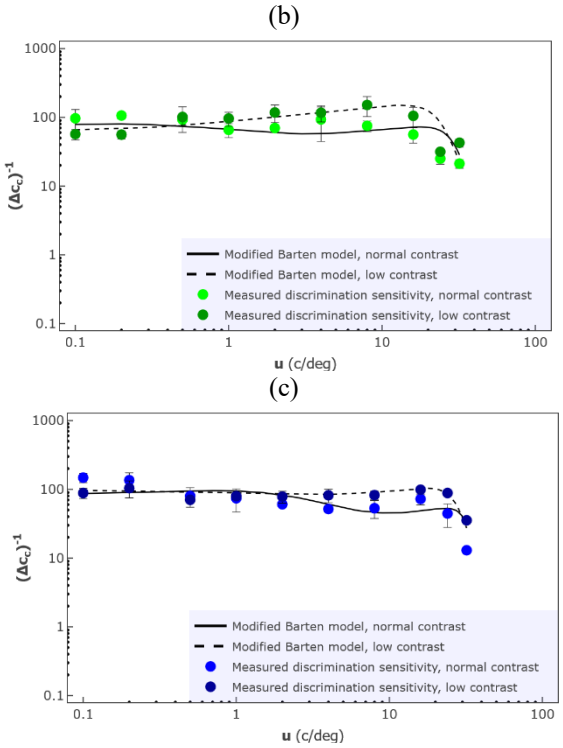


Fig. 6. Measured and modelled contextual band contrast discrimination sensitivity, $(\Delta c_c)^{-1}$, for all three images and both normal and low contrast versions. (a) *gallery* (b) *park* (c) *people*.

For all three images, above about 2.0 c/deg, the normal contrast image yields a lower discrimination sensitivity compared with the low contrast version. As frequency reduces below this value, contrast discrimination gradually improves with the normal image, until it matches, or even slightly surpasses that for the low contrast version.

The continuous and hatch curves in Fig. 6 are again calculated from the modified Barten model for normal and low contrast conditions respectively. Equation 4 parameter values are as given above with the exception of β, which are for *gallery* [0.2 (low contrast image), 0.25 (normal contrast image)], *park* [0.2 both low and normal image contrast] and *people* [0.2 both low and normal image contrast]. Values of the Crozier factor k for *gallery*, *park* and *people* are 3.0, 3.0 and 2.5 respectively (see also Table 1(a,b,c) for comparisons between all experiments).

In the use of equation 1, c_o is replaced by previously modelled values of contextual contrast detection as a function of u for all three images [15]. Contextual contrast detection reflects the impact of spatial noise generated from the image as described in the introduction and can be calculated directly from c_o and c_s as shown in our previous communication [15]. Incorporating image noise into the contrast detection term in this fashion, is directly equivalent to Barten's use of equations 1 and 3 in his successful modelling of sine-wave contrast discrimination in the presence of spatial noise [6]. Values of c_s used in the modelling of the data in Fig. 5 were again obtained from regression line fits to the image contrast spectra shown in Fig. 3.

Our adaptation of the Barten model is seen to offer a good fit to the measured discrimination sensitivity values for all three images and throughout the measured frequency range.

The numerical values for both k and β model parameters for all our discrimination studies are summarised in Table 1 for comparative purposes. Table 1(a) lists values for the sine-wave derived data (Figure 1, section 2.1.1); Table 1(b) list values for the isolated image band derived data (Figure 5, section 3.1); and Table 1(c) lists values for the contextual image band derived data (Figure 6, current se,ction). Note the similarity in β values between the *gallery* isolated data and the sine-wave data, suggesting that isolated image band discrimination is analogous to sine-wave discrimination. On the other hand, the β values for the contextual studies are around ten times higher than those for the sine-wave studies. This discrepancy is due to the two different modes of viewing (contextual vs isolated) and is discussed further in section 4. The values of k (Crozier factor) fall within the expected range [59, 76-80].

Table 1: Numerical values for k and β model parameters for all studies.

(a)

Sine-wave studies	**Legge [24] 0.25 cpd**	**Legge [24] 4.0 cpd**	**Yang & Makous [29] 0.8 cpd**	**Yang & Makous [29] 2.0 cpd**	**Bradley & Ohzawa [28] c_s = 0.25**
β	0.020	0.020	0.015	0.015	0.010
k	3.0	3.0	3.0	3.0	4.0

(b)

Natural-scene isolated band studies	***gallery***
β	0.015 (0.015)
k	2.8

(c)

Natural-scene contextual band studies	***gallery***	***park***	***people***
β	0.25 (0.20)	0.20 (0.20)	0.20 (0.20)
k	3.0	3.0	2.5

(a) sine-wave studies, Michelson contrast; (b) natural-scene studies with isolated image bands, RMS contrast; (c) natural-scene studies image bands in context, RMS contrast. The β values in parenthesis in (b) and (c) correspond to the low image contrast versions and those outside the parenthesis to the normal image contrast versions.

Appendix B presents measures of deviation (RMSE and MAE) between data and the original Barten and the modified discrimination models for all our studies.

3.3 Comparison of contrast detection and discrimination

Fig. 7(a,b,c) compares modelled contrast detection sensitivities obtained from our previous study [15] and modelled discrimination sensitivities, as depicted in Fig. 6, for all three image stimuli and both contrast versions. We used the modelled values of contextual contrast detection sensitivity for visual clarity, since they show a close fit to the measured detection data [15]. When contrast detection is directly compared with contrast discrimination, some interesting features emerge. First, for *gallery* and *park* scenes, at spatial frequency above about 2.0- 3.0 c/deg, and for each overall contrast level, detection and discrimination sensitivity levels are similar. This is approximately the case for *people* above about 6.0 c/deg, although discrimination sensitivity is slightly higher than for detection. Second, as spatial frequency reduces below 2.0 c/deg, contrast discrimination sensitivity progressively increases to a level above that for detection sensitivity in both normal and low image contrast conditions.

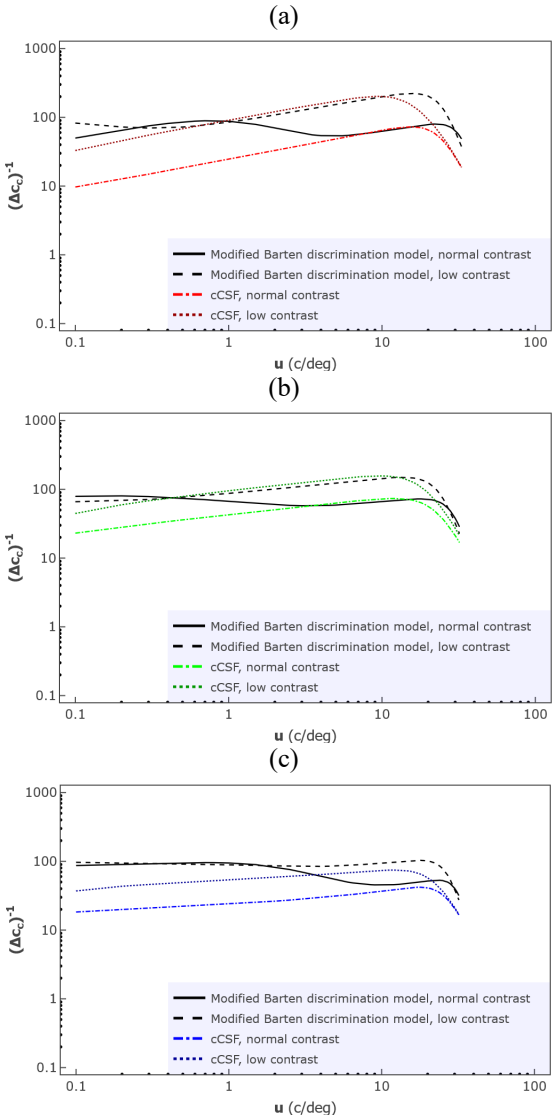


Fig 7: Modeled contextual detection sensitivity and discrimination sensitivity for (a) gallery, (b) park and (c) people image stimuli at normal and low contrast levels.

4. DISCUSSION

The modified Barten discrimination model introduced in section 2.1 has been found to offer a realistic framework for describing contrast discrimination data obtained from traditional sine-wave grating patterns with variations in both spatial frequency and pedestal contrast. More important to our study is that, the model is capable of describing sensitivity to changes in suprathreshold contrast in 1-octave wide spatial frequency bands within three radically different natural scenes. The model structure addresses non-linear operations in the visual system and embodies the concept *that discrimination is simply detection in the presence of spatial noise, this being defined by the pedestal contrast*. In natural scenes, a further consideration is contrast noise generated from signals immediately outside of the band whose discrimination threshold is being measured [15].

For both sine-wave grating and isolated band stimuli within a natural image, numerical parameter values within the model remain at constant levels, which is important when using the model as a predictive tool. However, with natural scenes, for the model to give an absolute fit to contextual contrast discrimination data, the value of the scaler term β (equation 5), is around ten times higher than the value required in the case of both the isolated band sensitivity measurements and the sine-wave data (see Table 1). This, itself, is not a problem providing the numerical value of β remains stable as scene type changes, as was found to be case in our study. Since the required values of contextual detection in the model can be readily calculated from combining image contrast spectrum data with a standard Barten contrast detection curve (CSF for a given image luminance and field size), the model framework offers a straightforward method for the prediction of band contrast discrimination sensitivity in natural scenes.

There are, of course, other pictorial factors likely to affect final contextual band contrast sensitivity, such as the spatial distribution of contrasts, edge density and colour [36, 85]. Nevertheless, the model is shown to satisfactorily predict absolute measurements for at least the three distinctly different natural scenes studied. As indicated in [15], the image contrast spectra are significant influencers of the shape and magnitude of the detection (and in consequence the discrimination) sensitivity curves in free viewing images. The natural scene images we used in both our studies follow the general *1/frequency* profile of natural scene spectra, but also span over a relatively large range in terms of magnitude and gradient, when compared with measured scene contrast spectra of a large number of scenes with important variations of scene contents, as shown in [15]. Although the model predictions require further validation from more images, including synthetic images where spatial frequency content can be controlled, the results obtained in our study provide a very solid foundation for future experimentation. Of primary interest would be an investigation of the numerical stability of the model parameter β for radical changes in image type.

At a given spatial frequency above about 2.0 c/deg, the difference in contrast discrimination sensitivity between normal and low image contrast versions shown in Fig. 6 can be conceptualised by the near Weber law behaviour normally demonstrated in the sine-wave dipper function at pedestal contrasts higher than the "dip". Certainly, the image contrast spectra shown in Fig. 3 indicate that pedestal (band) contrasts at moderate to high frequencies are large enough for this to be the case for both scene contrast levels. The behaviour shown by the measured contrast discrimination at a value below 2.0 c/deg can be related to the dipper profile found in previous sine-wave studies at low frequency [24, 29, 39], where the function's slope reduces considerably beyond the dip as discussed in section 2.3. This feature of the dipper response has been incorporated into the Barten model with the visual mechanism responsible for the phenomenon characterised by the product of both a contrast and a frequency response function $f.B^*$. As discussed in section 1.4, it has been suggested that the visual mechanism is, in fact, simply a switch from signal transfer through sustained visual processing channels to transient channels as spatial frequency reduces [24, 39-41]. Our model variant with the incorporation of $f.B^*$ into the function Ψ is a simple single channel construct and therefore does not directly address the detail of this possible crossover mechanism.

Fig. 6 and Fig. 7 show that low frequency contrast discrimination sensitivity is higher than detection sensitivity and for the normal contrast version of our images, this increase is substantial, being around 5.0-10.0 times at 0.1 c/deg. This may be because at low frequencies with our normal contrast level images, noise defined by adjacent band signal information at +1/-1 octaves, (which is likely to be the most significant pictorial noise source [15, 50, 89]) will be near its maximum level as indicated from the contrast spectra of Fig. 3. Previous dipper function determinations have revealed higher sine-wave contrast discrimination sensitivity

compared with detection sensitivity for pedestal contrasts between 0.05 and 1.0 when measured in the presence of substantial levels of noise [27].

Both our contextual contrast discrimination and detection sensitivity [15] findings have important implications for general everyday vision. If, for example, spatial frequencies from 1.0-15 c/deg are the most critical for simply perceiving objects in the natural world, as suggested by Norton et al [90], then the results indicate that for this frequency range, and at a given overall level of scene contrast, discrimination and detection sensitivity are approximately the *same* and almost *independent* of spatial frequency. The most important factor in controlling band spatial contrast sensitivity then becomes simply the overall level of scene contrast itself. Field and Chandler [91] introduced an approach to understanding spatial contrast sensitivity which focuses on a neurons vector magnitude. This study, together with results from a contrast matching experiment performed with log-Gabor functions, led them to conclude that contrast sensitivity to natural scenes will be flat up to around 30 c/deg. Our findings are in broad agreement with this, although a turning point nearer to 20 c/deg is suggested. The decrease in sensitivity at higher frequencies then being mainly due to optical limitations of the eye and receptor sampling effects.

4.1 Implications in IQ modelling

The results also provide implications for IQ modelling which routinely uses the luminance CSF (a contrast detection function derived from simple stimuli presented *in isolation* from other signals) as a visual weighting tool. Our findings for both *contextual* contrast detection [15] and discrimination over the frequency range (0.1-32 c/deg) suggest that, any relevant visual weighting function should be relatively flat up to around 20 c/deg and then reduce in magnitude. This contradicts the typical bandpass CSF implementation, but is in line with suggestions from Peli [92] who states that low pass functions that are flat at low frequencies are better suited for quality modeling, as well as the more recent conclusions by Field and Chandler [91].

Debates on the relevance of the CSF, as well as contrast suprathreshold models in IQ and fidelity modelling have been extensive and go back as far as 1976 [93]. Amongst others, Haun and Peli [12], Triantaphillidou et al. [13] and Chandler [94,95] have provided reviews on the subject. Both the linear systems approach and the relevance of threshold contrast models in image quality evaluation have been questioned for almost as long as they have been in use. Almost three decades ago Ahamuda and colleagues [96,97] stated that in image viewing, the CSF is largely outweighed by contrast masking. Around the same period, Silverstein and Farrell [98] declared that suprathreshold judgments of IQ are unrelated to contrast threshold judgments. In contrast, Haun and Peli [12] argued that in estimating the visual quality of images, contrast thresholds are of principle importance, whilst perceived suprathreshold magnitudes are relatively less important. In any case, contrast is routinely mathematically treated in IQ modelling, it is of basic perceptual importance, yet overall perceived pictorial contrast is difficult to quantify or predict [12, 99, 100].

Many more have queried the implementation of threshold detection models in metrics evaluating visual IQ [17, 101-103]. Nonetheless, the CSF is widely used and currently included in several standard visual metrics, for example metrics proposed in IEEE Standard for Camera Phone Image Quality (CPIQ) [104] and the current ISO12233 [105]. Our finding that, at a given level of overall contrast in a natural scene, contrast discrimination and detection [15] are similar over the visually significant spatial frequency range has implications in IQ modelling. It suggests that both contextual detection or discrimination visual weighting functions may be equally effective. Even so, given that the human contrast sensitivity weighting function in IQ

modelling is in a normalised form [74, 75, 94], it cannot account for the important changes in detection and discrimination sensitivity incurred by a variation in overall image contrast level.

The proposed predictive detection and discrimination models presented in [15] and this communication are both scene and process dependent. They may therefore offer a more viable alternative to CSF models implemented in current IQ modelling and relevant standards. Initial research by Fry et al. [75] compared the performance of IQ models with the standard CSF and with the contextual detection and discrimination sensitivity models outlined in [15] and in this study. They showed that, although current computational as well as signal-transfer IQ metrics that were calibrated with observer data performed better with the standard bandpass CSF, other, well established non-calibrated signal transfer metrics ('purer' in form metrics that do not rely on observers' data fittings) performed better with our proposed model which accounts for the natural scene spectra.

In other words, contextual contrast detection and discrimination functions have the potential to represent more meaningfully human visual responses within IQ metrics than classical CSFs, since they represent responses to the frequency contents of target image frequencies within the context of all remaining frequencies. Also, given their inherent scene dependent nature, they can lead to metric design and calibration that do not rely on observer responses to image datasets. Considering their implementation, despite the scene dependency in our determinations of contrast discrimination, in broad terms the contextual detection and discrimination functions we derived basically resemble low pass filters. This is in support of the abovementioned claims [91, 92] and the argument that, the visual weighting function used in IQ models from signal transfer theory could simply be represented by a low pass filter which decays at high frequency according to the optical transfer function of the eye.

The suitability of the proposed visual models for IQ purposes should be further tested with a larger variety of scenes and implemented in different genres of IQ metrics. Nevertheless, since these models are essentially mechanistic, i.e. they derive predictions from visual functions directly from natural scene contrasts, they should provide a robust basis for IQ modelling of natural scenes.

5. CONCLUSIONS

The following conclusions can be drawn from the research presented in this paper.

For all three image stimuli, above about 2.0 c/deg, the normal contrast image yields a lower discrimination sensitivity compared with the low contrast version. As frequency reduces below this value, contrast discrimination gradually improves with the normal image, until it matches, or even slightly surpasses that for the low contrast version.

Our adaptation of the Barten model is seen to offer a good fit to the measured discrimination sensitivity values for all three images used in this study and throughout the measured frequency range. Values of the parameter β and the Crozier factor k for *gallery*, *park* and *people* are very similar, demonstrating the robustness of the model, which accounts for different scene spectra content.

Contextual contrast detection reflects the impact of spatial noise generated from the image as described in the introduction and can be calculated directly from c_o and c_s as shown in our previous communication [15].

Comparison of contrast detection and discrimination showed some interesting features. For two out of three scenes, at spatial frequency above about 2.0-3.0 c/deg and a given overall contrast level, modeled detection and discrimination sensitivity levels are similar. This is also nearly the case for the third scene, but not at all spatial frequencies.

In summary, it is clear that important predictions can be seen from the theoretical model curves:

- The collapse of low frequency/high contrast sine-wave dipper functions (as shown in Fig 1 and reported in the literature).
- At normal contrast levels, lower discrimination sensitivity at moderate to high frequencies are shown compared to low contrasts (both contextual and isolated bands).
- Equalization of low frequency discrimination sensitivity at both image contrast levels in isolated and contextual bands.
- General flattening of discrimination sensitivity before optical attenuation takes hold (contrast gain control in operation)
- Overall magnitude of discrimination sensitivity is predicted with stable β values.

Perceptual image contrast is an intrinsic factor in IQ and affects all IQ attributes. The proposed predictive detection and discrimination models presented in [15] and here, being scene and process dependent, may offer a more viable alternative to CSF models implemented in current IQ modelling standards.

It is important to note that these models are largely mechanistic and are not expected to fit perfectly observer response data. Their purpose is to make predictions from visual functions which are based on natural scene data. This should make them suitable for incorporating in to IQ modelling of natural scenes.

AKNOWLEGMENTS

We thank the UK Defence Science and Technology Laboratory (DSTL) for funding this work and all our observers for their participation to the visual experiments. We confirm that the funding source had no involvement in the study design, in the collection, analysis and interpretation of data, in the writing of the report; and in the decision to submit the article for publication. We also confirm that the funding source has not imposed any restrictions with respect to the publication of the funded research.

DISCLOSURE

The authors declare no conflicts of interest.

DATA AVAILABILITY

Data underlying the results presented in this paper are not available in the public domain, they may be obtained from the authors upon request.

REFERENCES

1. F. W. Campbell and J. G. Robson, “Application of Fourier analysis to the visibility of gratings”, J.Physiol, **197,** 551-566 (1968)

2. A. Watanabe, T. Mori, S. Nagata and K. Hiwatashi, "Spatial sine-wave responses of the human visual system", Vision Res, **8**, 1245-1263 (1968)
3. D. H. Kelly, "Visual contrast sensitivity", Opt. Acta. **24**, 107-129 (1977)
4. R. L. De Valois and K.K. De Valois, *Spatial Vision. Oxford Psychology Series 14*, (New York: Oxford University Press, 1990)
5. J. Rovamo, J. Mustonen, and R. Nasanen, "Modelling contrast sensitivity as a function of retinal illuminance and grating area" Vision Res, **34**, 1301-1314 (1994)
6. P. J. G. Barten, *Contrast sensitivity of the human eye and its effects on image quality* (SPIE Press, 1999)
7. A.B. Watson and A.J. Ahumada, "A standard model for foveal detection of spatial contrast", J. Vision. **5,** 717-740 (2005)
8. J. R. Jarvis and C. M. Wathes, "On the calculation of optical performance factors from vertebrate spatial contrast sensitivity" Vision Res, **47,** 2259-2271 (2007)
9. J. R. Jarvis and C. M. Wathes, "Mechanistic modeling of vertebrate spatial contrast sensitivity and acuity at low luminance" Visual Neuroscience, **29**, 169-181 (2012)
10. D. Regan. The Charles F. Prentice Award Lecture 1990: specific tests and specific blindnesses: keys, locks, and parallel processing. Optom. Vis. Sci. **68**, 489–512 (1991)
11. J.Tardif, M.R.Watson, D.Giaschi and G.Gosselin. "The Curve Visible on the Campbell-Robson Chart is not the contrast sensitivity function". Frontiers in Neuroscience. **15,** 1-10 (2021)
12. A. Haun and E. Peli, "Is image quality a function of contrast perception?", Proc SPIE **8651,** 86510C (2013)
13. S. Triantaphillidou, J. Jarvis and G. Gupta, "Spatial contrast sensitivity and discrimination in pictorial scenes", Proc SPIE **9016,** 901604 (2014)
14. E. Fry, S. Triantaphillidou, J. Jarvis and G. Gupta "Image quality optimization via application of contextual contrast sensitivity and discrimination functions", Proc SPIE **9396,** Image Quality and System Performance X11 (2015)
15. S. Triantaphillidou, J. Jarvis, A. Psarrou and G. Gupta, "Contrast detection in images of natural scenes" Signal Processing: Image Communication, **75**, 64-75 (2019)
16. B. A. Olshausen and D. J. Field, "How close are we to understanding v1?" Neural Computation, **17,** 1665-1699 (2005)
17. R. E. Jacobson, "An evaluation of image quality metrics", J. Photogr. Sci. **43,** 7-16 (1995)
18. B. E. Rogowitz, T. N. Pappas and J. P. Allebach, "Human vision and electronic imaging", J. Electronic Im, **10,** 10-19 (2001)
19. S. Bouzit and L. W. MacDonald, "Sharpness enhancement through spatial frequency decomposition", in *Proceedings of PICS 2001: Image Processing, Image Quality, Image Capture, Systems Conference* [Society for Imaging Science and Technology, 2001], pp377-381 (2001)
20. Z. Wang, A. Bovik, H. Sheikh and E. Simoncelli "Image quality assessment: From error visibility to structural similarity" IEEE Trans.Image Process, **13,** 600-612 (2004)
21. F. W. Campbell and J. J Kulikowski, "Orientation selectivity of the human visual system", J. Physiol, Lond, **187,** 437-445 (1966)
22. Y. Kohayakawa, "Contrast-difference thresholds with sinusoidal gratings", J. Opt. Soc. Am. **62,** 584-587 (1972)
23. J. Nachmias and R. V. Sansbury, "Grating contrast: discrimination may be better than detection", Vision Res, **14,** 1039-1042 (1974)
24. G. E. Legge, "Spatial frequency masking in human vision: binocular interactions", J. Opt. Soc. Am. **69,** 838-847 (1979)
25. G. E. Legge and J. M. Foley, "Contrast masking in human vision", J. Opt. Soc. Am. **70,** 1458-1471 (1980)
26. G. E. Legge, "Binocular contrast summation-II. Quadratic summation", Vision Res, **24,** 385-394 (1984)

27. D.G. Pelli, “Uncertainty explains many aspects of contrast detection and discrimination”, J. Opt. Soc. Am. A, **2,** 1508-1532 (1985)
28. A. Bradley and I. Ohzawa, “A comparison of contrast detection and discrimination”, Vision Res. **26,** 991-997 (1986)
29. J. Yang and W. Makous, “Modeling pedestal experiments with amplitude instead of contrast”, Vision Res. **35,** 1979-1989 (1995)
30. M. A. Georgeson and G. D. Sullivan, “Contrast constancy: Deblurring in human vision by spatial frequency channels”, J. Physiol, **252,** 627-656 (1975)
31. M. W. Cannon, “Contrast sensation: a linear function of stimulus contrast”, Vision Res, **19,** 1045-1052 (1979)
32. N. Brady and D. J. Field, “What’s constant in contrast constancy? The effects of scaling on the perceived contrast of bandpass patterns”, Vision Res, **35,** 739-756 (1995)
33. R. A. Frazor and W. S. Geisler, “Local luminance and contrast in natural scenes”, Vision Res. **46**, 1585–1598 (2006).
34. U. Polat and D. Sagi, “Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments”, Vision Res. **33**, 993-999 (1993).
35. J. A. Solomon and M. J. Morgan, “Facilitation from collinear flanks is cancelled by non-collinear flanks”, Vision Res. **40**, 279-286 (2000).
36. P. J. Bex, S.G. Solomon and S.C. Dakin, “Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure”, J. Vision, **9,** 1-19 (2009)
37. C. Chubb, G. Sperling, and J. A. Solomon, “Texture interactions determine perceived contrast”, in *Proceedings of the National Academy of Science***,** (1989), pp. 9631-9635.
38. J. Xing and D. J. Heeger, “Measurement and modeling of center- surround suppression and enhancement, Vision Res. **41,** 571-583 (2001).
39. J. J. Kulikowski, “Effective contrast constancy and linearity of contrast sensation”, Vision Res, **16,** 1419-1432 (1976)
40. A. B. Watson and J. Nachmias, “Patterns of temporal integration in the detection of gratings”, Vision Res, **17,** 893-902 (1977)
41. G. E. Legge, “Sustained and transient mechanisms in human vision: temporal and spatial properties”, Vision Res, **18,** 69-81 (1978)
42. A. Leonova, J. Pokorny and V.C. Smith, “Spatial frequency processing in inferred PC- and MC-pathways”, Vision Res, **43,** 2133-2139 (2003)
43. P. J. Bex, I. Mareschal and S. C. Dakin, “Contrast gain control in natural scenes”, J.Vision, **7,** 1-12 (2007)
44. D. J. Tolhurst, M. P. S. To, M. Chirimuuta, T. Troscianko, P-Y., Chua and P.G. Lovell, “Magnitude of perceived change in natural images may be linearly proportional to differences in neuronal firing rates”, Seeing and Perceiving, **23,** 349-372 (2010)
45. H. R. Wilson, “A transducer function for threshold and suprathreshold human vision”, Biological Cybernetics, **38,** 171-178 (1980)
46. P. Whittle.”Increments and decrements: Luminance discrimination”. Vision Res. **26,** 1677-1691 (1986)
47. P. Whittle. “Brightness, discriminability and the crispening effect”. Vision Res. **32,** 1493-1507 (1992)
48. W. McIlhagga “Classification images for contrast discrimination”. Vision Res. **182,** 36-45 (2021)
49. D. Kane and M. Bertalmio. “A re-evaluation of Whittle (1986,1992) reveals the link between detection thresholds, discrimination thresholds and brightness perception”. Journal of Vision. **19(1):16**, 1-13 (2019)
50. K. T. Blackwell, “The effect of white and filtered noise on contrast detection thresholds”, Vision Res, **38,** 267-280 (1998)
51. A. N. Sanborn and P. Dayan, “Optimal decisions for contrast discrimination”, J. Vision, **11,** 1-13 (2011)

52. J. A. Solomon, “The history of dipper functions”, Atten. Percept. Psychophys., **71**, 435-443 (2009).
53. A. Gorea and D. Sagi, “Disentangling signal from noise in visual contrast discrimination”, Nature Neuroscience, **4,** 1146-1150 (2001)
54. L. L. Kontsevich, C-C. Chen, C.W. Tyler, “Separating the effects of response nonlinearity and internal noise psychophysically”, Vision Res, **42,** 1771-1784 (2002)
55. M. Katkov., M. Tsodyks and D. Sagi, “Analysis of a two-alternative forced-choice signal detection theory model”, J. Mathematical Psychology, **50,** 411-420 (2006)
56. S. A. Klein, “Separating transducer non-linearities and multiplicative noise in contrast discrimination”, Vision Res, **46,** 4279-4293 (2006)
57. M. A. Georgeson and T. S. Meese, “Fixed or variable noise in contrast discrimination?”, Vision Res, **46,** 4294-4303 (2006)
58. J. A. Solomon, “Contrast discrimination: Second responses reveal the relationship between the mean and variance of visual signals”, Vision Res, **47,** 3247-3258 (2007)
59. J. J. Depalma and E. M. Lowry, “Sine-Wave Response of the Visual System. II. Sine-Wave and Square-Wave Contrast Sensitivity", J. Opt. Soc. Am. **52**, 328-335 (1962)
60. A. S. Patel, “Spatial resolution by the human visual system. The effect of mean retinal luminance”. J.Opt. Soc. Am, **56,** 689-694 (1966)
61. J. G. Robson. “Spatial and temporal contrast sensitivity functions of the visual system”. J.Opt.Soc.Am. **56**, 1141-1142 (1966)
62. F. L. van Nes and M.A. Bouman. “Spatial modulation transfer in the human eye”. J.Opt.Soc.Am. **57,** 401-406 (1967)
63. M. B. Sachs, J. Nachmias and J. G. Robson. “Spatial-frequency channels in human vision”. J.Opt.Soc.Am. **61,** 1176-1186 (1971)
64. van Meeteren and J.J. Vos. “Resolution and contrast sensitivity at low luminance levels”. Vision Res. **12,** 825-833 (1972)
65. E.R. Howell and R.F. Hess. “The functional area for summation to threshold for sinusoidal gratings”. Vision Res. **18,** 369-374 (1978)
66. V. Virsu and J. Rovamo. “Visual resolution, contrast sensitivity and the cortical magnification factor”. Experimental Brain Research. **37,** 475-494 (1979)
67. C.R. Carlson. “Sine-wave threshold contrast-sensitivity function: dependence on display size. RCA Review. **43,** 675-683 (1982)
68. J.Rovamo, R. Franssila and R. Nasanen. “Contrast sensitivity as a function of spatial frequency, viewing distance and eccentricity with and without spatial noise”. Vision Res. **32,** 631-637 (1992)
69. J. Rovamo, H. Kukkonen, K. Tippana and R. Nasanen. “Effects of luminance and exposure time on contrast sensitivity in spatial noise”. Vision Res. **33,** 1123-1129 (1993a)
70. J. Rovamo, O. Luntinen and R. Nasanen. “Modelling the dependence of contrast sensitivity on grating area and spatial frequency”. Vision Res. **33**, 2773-2788 (1993b)
71. P. G. J. Barten, "Evaluation of subjective image quality with the square-root integral method," J. Opt. Soc. Am. A **7**, 2024-2031 (1990)
72. P. G. J. Barten, “Effects of quantization and pixel structure on the image quality of color matrix displays”, J. of. SID, **1**, 147-153 (1993)
73. R. B. Jenkin, S. Triantaphillidou, and M. A. Richardson "Effective pictorial information capacity as an image quality metric", Proc. SPIE 6494, Image Quality and System Performance IV, 64940O (2007)
74. G.M. Johnson, M.D. Fairchild, On contrast sensitivity in an image difference model, in: Proc. PICS 2002: Image Processing, Image Quality, Image Capture, Systems Conference, Society for Imaging Science and Technology, 18–23 (2002).
75. E. Fry, S. Triantaphillidou, R. Jenkin, R. E. Jacobson, and J. Jarvis, “Scene-and-Process-Dependent Spatial Image Quality Metrics”. Journal of Imaging Science and Technology **63**, 60407-1-60407-13 (2019).

76. W. J. Crozier, “On the variability of critical illumination for flicker fusion and intensity discrimination”, J.General Physiology, **19,** 503-522 (1935)
77. O. Schade, “Otical and photoelectric analog of the eye”, J. Opt. Soc. Am, **46,** 721-739 (1956)
78. J. A. J. Roufs, “Dynamic properties of vision-V1. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ration”, Vision Res, **14,** 871-888 (1974)
79. L. T. Maloney, “The slope of the psychometric function at different wavelengths”, Vision Res, **30,** 129-136 (1990)
80. R. Shani and D. Sagi, “Psychometric curves of lateral facilitation”, Spatial Vision, **19,** 413-426 (2006)
81. A. van Meeteren and J.M. Valeton, “Effects of pictorial noise interfering with visual detection”, J. Opt. Soc. Am., A, **5,** 438-444 (1988)
82. G. C. Holst, *CCD Arrays, cameras and displays,* SPIE Optical Engineering Press (1996)
83. H. Kukkonen, J. Rovamo, K. Tiippana, R. Näsänen, “Michelson contrast, RMS contrast and energy of various spatial stimuli at threshold”, Vision Res., 33, 1431-1436 (1993).
84. E. Peli, “Contrast in complex images”, J. Opt. Soc. Am. A **7***,* 2032-2040 (1990).
85. A. M. Haun and E. Peli, “Perceived contrast in complex images”, J. Vision, **13,** 1-21 (2013)
86. E. Peli, “Feature detection algorithm based on a visual system model”, Proc. IEEE 90 78–93 (2002).
87. M. Nezamabadi, E.D. Montag, R.S. Berns, “An investigation of the effect of image size on the color appearance of softcopy reproductions using a contrast matching technique”, Proc. SPIE 6493, 649309 (2007).
88. M. Pedersen, G. Simone, M. Gong, I. Farup, “A total variation based color image quality metric with perceptual contrast filtering”, in: International conference on Pervasive Computing, Signal Processing and Applications (2011).
89. C. F. Stromeyer and B. Julesz, “Spatial frequency masking in vision: critical bands and spread of masking”, J. Opt. Soc. Am., A, **62,** 1221-1232 (1972)
90. T. T. Norton, D.A. Corliss and J.E. Bailey, *The psychophysical measurement of visual function.* Butterworth Heinemann, Elsevier Science (2002)
91. D. J. Field and D. M. Chandler. “A new model and a new demonstration of contrast sensitivity” Journal of Vision, **20,** 811 (2020)
92. E. Peli, Contrast sensitivity function and image discrimination, J. Opt. Soc. Amer. A **18** (2001) 283–293.
93. M. Kriss, J. O’Toole, J. Kinard, “Information capacity as a measure of image quality”, Proc. SPSE conference on Image Analysis and Evaluation, 122–133 (1976).
94. D.M. Chandler, “Seven challenges in image quality assessment: past, present, and future research”, ISRN Signal Process. **53** (2013).
95. D.M. Chandler, M.M. Alam, T.D. Phan, “Seven challenges for image quality research”, Proc. SPIE 9014 (2014) 901402.
96. A.J. Ahumada Jr, Computational image quality metrics a review, Proc. SID **24** 305–308 (1993).
97. A.J. Ahumada Jr, A.B. Watson, A.M. Rohaly, Models of human image discrimi- nation predict object detection in natural backgrounds, Proc. SPIE 2411, 355-352 (1995).
98. D.A. Silverstein, E.J. Farrell, “The relationship between image fidelity and image quality”, Proc. IEEE International Conference on Image Processing I, IEEE, 881–884 (1996).
99. G. Simone, M. Pedersen, J.Y. Hardeberg, “Measuring perceptual contrast in digital images”, J. Vis. Commun. Image R. **23,** 491–506 (2012).
100. A. Beghdadi, M. C. Larabi, A. Bouzerdoum & K. M. Iftekharuddin, "A survey of perceptual image processing methods," Signal Processing: Image Communication, **28**, 811-831, (2013).

101. B.E. Rogowitz, T.N. Pappas, J.P. Allebach, Human vision and electronic imaging, J. Electronic Image. **10** 10–19 (2001).
102. S. Bouzit, L.W. MacDonald, Sharpness enhancement through spatial frequency decomposition, in: Proceedings of PICS 2001: Image Processing, Image Quality, Image Capture, Systems Conference, Society for Imaging Science and Technology, 377–381 (2001).
103. Z. Wang, A. Bovik, H. Sheikh, E, Simoncelli image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Process. **13** 600–612, (2004).
104. IEEE P1858, IEEE Standard for Camera Phone Image Quality (CPIQ) (2017).
105. ISO 12233 - 2017, Photography - Electronic still picture imaging - Resolution and spatial frequency response (2017).
106. J. M. Foley and G. E. Legge, "Contrast detection and near-threshold discrimination in human vision", Vision Res, **21,** 1041-1053 (1981)
107. T. S. Meese, "Area summation and masking", J. Vision, **4,** 930-943 (2004).
108. C. D. Schunn and D. Wallach, "Evaluating goodness-of-fit in comparison of models to data", InW. Tack (Ed.), Psychologie der Kognition: Reden and Vorträge anlässlich der Emeritierung vonWerner Tack. Saarbrueken, Germany: University of Saarland Press (2005).Available from https://www.researchgate.net/publication/282237248_Evaluating_Goodness-of-Fit_in_Comparison_of_Models_to_Data [accessed Mar 13 2022].

APPENDIX A: Barten contrast discrimination model

In the development of both his contrast detection (i.e. contrast sensitivity function - CSF) and discrimination models, Barten first considered the mathematical nature of the psychometric function [6]. This expresses the detection probability P of a signal of strength s, and is represented by a cumulative Gaussian probability function with an associated standard deviation σ. In luminance perception, Crozier [76] experimentally found that the ratio between the signal strength at the threshold of detection s_o and σ was effectively constant over a wide range of signal conditions. This relationship is often referred to as *Crozier's Law*. In the following, this ratio will be designated k. Regarded as a basic signal/noise ratio for detection, values of between about 2.0 and 4.0 have been experimentally determined from a number of studies after Crozier [77-80]. In Barten's studies the range was found to be between 2.7 and 4.5 [6, p. 59]. The constant k relates to the slope of the psychometric function and Barten showed that at the value of the signal for threshold detection (s_o):

$$\left(\frac{dP}{ds}\right) = \frac{k}{s_o\sqrt{2\pi}} \tag{A.1}$$

In the case of sine-wave contrast detection, s is represented by Michelson contrast (c) and s_o by the Michelson contrast threshold level (c_o).

When the psychometric function obtained from a sine-wave contrast detection experiment is compared with that obtained from a contrast discrimination experiment (with the signal defined as the test contrast at the threshold of discrimination), higher slopes are observed with the latter. For the case of contrast detection, the slope indicates k values near to the expected value of 3.0. However, when Barten examined the discrimination psychometric functions measured by Foley and Legge [106] and Legge [26], k values of 5.6 and 8.3 respectively, were found; far too high for *Crozier's law*. Re-plotting this data with the signal variable defined by the contrast difference metric Δc gave slopes with a k of around 1.0, which is lower than predicted by *Crozier's law*. However, when the signal was defined as a contrast difference given by $(c_t^2 - c_s^2)^{0.5}$, where c_s is the standard (pedestal) contrast and c_t the test contrast at the threshold of discrimination, k values near to the nominal Crozier value were found. This led Barten to conclude that the functional parameter in the visual contrast discrimination process is $(c_t^2 - c_s^2)^{0.5}$ and not Δc.

Conceptually, Barten also considered contrast discrimination as equivalent to contrast detection in the presence of a noise source defined by the pedestal itself. Now, in the presence of noise, Barten found that the sine-wave contrast detection threshold $(c_o)_n$ is given by:

$$(c_o)_n = (c_o^2 + k^2 c_n^2)^{0.5} \tag{A.2}$$

where c_n denotes the average sine-wave component contrast of the noise. Equation A2. was empirically deduced by Barten from sine-wave contrast detection data with added 2D noise as measured by van Meeteren and Valeton [81][*]. Barten therefore presented as his basic model construct for contrast discrimination the following expression:

$$(c_t^2 - c_s^2)^{0.5} = (c_o^2 + k^2 c_n^2)^{0.5} \tag{A.3}$$

[*] Equation 2 is, in fact, functionally equivalent to the *linear amplification model* [15, 50] when the noise is expressed in terms of contrast rather than spectral density.

With the assumption that the pedestal defines the noise source, then for a single sine-wave component, Barten showed that c_n can be approximated by 0.2 c_s .The normally employed contrast discrimination metric Δc was also introduced into equation A.3 using the fact that $(c_t^{\,2} - c_s^2)^{\,0.5}$ is mathematically equivalent to $[(c_s + \Delta c)^2 - c_s^2\,]^{0.5}$.

A final addition to equation A.3 was made by Barten to account for the slow decrease in gradient displayed in the dipper function at high contrasts [25, 107] which was assumed due to the profile of Wilson's transducer function [45]. This particular non-linearity component was mathematically approximated by dividing the right-hand side of equation A.3 by a factor Ψ given by:

$$\Psi = \left(1 + 0.004k\frac{c_s}{c_o}\right)^{0.5} \tag{A.4}$$

If we now allow for a variation of pedestal spatial frequency u c/deg, we may describe the generalised Barten model (expressed in terms of Δc) as:

$$\Delta c(u) = \left[\left(\frac{c_o^2(u)+0.04k^2c_s^2(u)}{\Psi^2(u)}\right) + c_s^2(u)\right]^{0.5} - c_s(u) \tag{A.5}$$

which is equation 1 in the main text.

APPENDIX B: Goodness of fit

Table B.1 presents measures of deviation (RMSE and MAE) between data and i) the modified discrimination model and the original Barten model for the sine-wave discrimination studies (Fig. 1), ii) the modified discrimination model and the original Barten model for the isolated band discrimination study (Fig. 5), iii) the modified discrimination model for the contextual band discrimination study (Fig. 6).

Table B1: Measures of deviation between the data and the model(s) for all studies

Figure 1a	**Legge [24] 0.25 c/deg (data in Δc)**	
	modified Barten discrimination	*Barten discrimination*
RMSE	0.0066	0.0144
MAE	0.0080	0.0092
Figure 1b	**Young & Macus 0.8 c/deg [29] (data in Δc)**	
	modified Barten discrimination	*Barten discrimination*
RMSE	0.0021	0.0158
MAE	0.0059	0.0078
Figure 1c	**Legge [24] 4 c/deg (data in Δc)**	
	modified Barten discrimination	*Barten discrimination*
RMSE	0.0092	n/a
MAE	0.0062	n/a
Figure 1d	**Bradley & Osawa [28] (data in Δc^{-1})**	
	modified Barten discrimination	*Barten discrimination*
RMSE	4.43	n/a
MAE	3.74	n/a
Figure 5	**Low contrast gallery (data in Δc^{-1})**	
	modified Barten discrimination	*Barten discrimination*
RMSE	64.85	n/a
MAE	54.18	n/a
Figure 5	**Normal contrast gallery (data in Δc^{-1})**	
	modified Barten discrimination	*Barten discrimination*
RMSE	54.79	57.77
MAE	39.88	46.41
Figure 6a	**Low contrast gallery (data in Δc^{-1})**	**Normal contrast gallery (data in Δc^{-1})**
	modified Barten discrimination	*modified Barten discrimination*
RMSE	55.77	31.56
MAE	48.95	26.15
Figure 6b	**Low contrast park (in Δc^{-1})**	**Normal contrast park (data in Δc^{-1})**
	modified Barten discrimination	*modified Barten discrimination*
RMSE	24.86	21.56
MAE	19.90	18.26
Figure 6c	**Low contrast people (in Δc^{-1})**	**Normal contrast people (data in Δc^{-1})**
	modified Barten discrimination	*modified Barten discrimination*
RMSE	9.44	28.28
MAE	7.71	23.03

The measures in Table B.1 are provided to help reading the figures, but it is important to bear in mind that, both Barten's original and the proposed discrimination models are largely mechanistic and follow theoretical expectations (as discussed throughout the paper and

summarized in the conclusions), so interpretations cannot be necessarily derived from such statistical measures [108].