Defining human contrast sensitivity and discrimination from complex imagery

S. Triantaphillidou, J. Jarvis, G. Gupta, H. Rana

S. Triantaphillidou, J. Jarvis, G. Gupta, H. Rana, "Defining human contrast sensitivity and discrimination from complex imagery," Proc. SPIE 8901, Optics and Photonics for Counterterrorism, Crime Fighting and Defence IX; and Optical Materials and Biomaterials in Security and Defence Systems Technology X, 89010C (16 October 2013); doi: 10.1117/12.2029194

SPIE.

Event: SPIE Security + Defence, 2013, Dresden, Germany

DEFINING HUMAN CONTRAST SENSITIVITY AND DISCRIMINATION FROM COMPLEX IMAGERY

S. Triantaphillidou*, J. Jarvis*, G. Gupta*, **H. Rana
*Imaging Technology Research Group, University of Westminster
Watford Road, HA1 3TP, Middlesex, UK
** Defence, Science and Technology Laboratory
Fort Halstead, Sevenoaks, TN14 7BP, UK

Shape, form and detail define image structure in our visual world. These attributes are dictated primarily by local variations in luminance contrast. Defining human contrast sensitivity (threshold of contrast perception) and contrast discrimination (ability to differentiate between variations in contrast) directly from real complex scenes is of outermost relevance to our understanding of spatial vision. The design and evaluation of imaging equipment, used in both field operations and security applications, require a full description of strengths and limitations of human spatial vision. This paper is concerned with the measurement of the following four human contrast sensitivity functions directly from images of complex scenes: i) *Isolated Contrast Sensitivity (detection) Function* (iCSF); ii) *Contextual Contrast Sensitivity (detection) Function* (cCSF); iii) *Isolated Visual Perception (discrimination) Function* (iVPF) and iv) C*ontextual Visual Perception (discrimination) Function* (cVPF). The paper also discusses the following areas: Barten's mathematical framework for modeling contrast sensitivity and discrimination; spatial decomposition of image stimuli to a number of spatial frequency bands (octaves); suitability of three different relevant image contrast metrics; experimental methodology for subjective tests; stimulus conditions. We finally present and discuss initial findings for all four measured sensitivities.

Keywords: contrast sensitivity, contrast discrimination, threshold contrast, suprathreshold contrast, contrast metrics, Barten detection model, Barten discrimination model, CSF, VPF

1. INTRODUCTION

This paper describes developments on research concerned with the determination of human contrast sensitivity and contrast discrimination functions from complex visual stimuli. We are interested in examining the relationship between the derived visual sensitivity functions and various attributes of the stimuli.

An initial survey of models of human sensitivity to spatial information, using results from similar investigations, suggests that the mathematical framework developed by Barten[1] for sinewave spatial stimuli can provide a sound position for our own modeling studies of complex image stimuli. Section 2 discusses the Barten contrast sensitivity and contrast discrimination models. In this section, we also theoretically extend the framework of Barten, to provide a contrast discrimination function covering all frequencies of visual interest, which we call the *Visual Perception Function* (VPF). Section 3 provides details on spatial frequency decomposition for obtaining band-limited images (i.e. images of different spatial frequency bands, or octaves), which are further manipulated to produce our test stimuli. It also considers the important issue of the quantification of image (and band) contrast and metrics for its determination. It compares three relevant contrast metrics and discusses their appropriateness for describing contrast sensitivity in our work. Section 4 presents the experimental outline, the setup for our psychophysical investigations and the range of stimulus conditions. Section 5 presents initial results, draws a summary of the research and discusses relevant further work.

Understanding how human vision behaves in low light conditions, when viewing complex scenes containing low contrast objects and noise is key to optimizing modern military viewing equipment. This is especially relevant for future developments in handheld lowlight color imagers, where the trade space is spatiotemporal resolution and color depth compared with size, weight and power consumption (SWAP). For example, understanding how much color depth is necessary to conduct a range of visual discrimination tasks may lead to low power displays and sensors.

*triants@westminster.ac.uk; phone +44 (0)20 7911 5000; fax +44 (0)20 7911 5943

Optics and Photonics for Counterterrorism, Crime Fighting and Defence IX; and Optical Materials and Biomaterials in Security and Defence Systems Technology X, edited by Douglas Burgess, Gari Owen, Roberto Zamboni, Francois Kajzar, Attila A. Szep, Proc. of SPIE Vol. 8901, 89010C · © 2013 SPIE · CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2029194

1.1 Contrast perception and masking

Much of our understanding of visual processing of spatial information is based on experiments employing simple stimuli, such as sinewave gratings and Gabor patches. In a typical experiment, the modulation (*Michelson contrast*)[2] is adjusted until the grating structure is at the *absolute* threshold of perception. If this visual threshold is measured at a range of spatial or temporal frequencies as appropriate, *contrast sensitivity functions* (CSFs) are obtained. For reviews of many of the previous studies of the human CSF see Barten[1] and for other species see Jarvis et al.[3,4].

One important reason why the CSF has been extensively studied, is that theoretically, it is capable of predicting spatial sensitivity to real complex images. However, limitations in the use of the CSF to predict image quality in real scenes have been observed as long ago as 1976[5]. More recently, Bex et al.[6] measured the CSF from a small target contained within a real image. They concluded that the standard threshold CSF is a poor indicator of sensitivity to structure in natural scenes. An important reason for this failure is that CSFs are based on the *threshold* of perceiving low contrast grating stimuli. Although such threshold measurements allow us to define the borders between visible and invisible, they are limited in describing sensitivity to small contrast differences that may exist between spatial signals which are already perceivable, i.e. *suprathreshold* signals.

Perceptible changes in suprathreshold contrast (referred to as contrast discrimination) and luminance are the basic building blocks leading to the perception of shape, form, and detail, whether such changes are observed in real scenes or pictorial representations of them i.e. photographic, or video. Early contrast matching experiments that employed suprathreshold sinewave gratings, found that the measured contrast discrimination functions were more low-pass in character compared with the CSF[7,8].

Any suppression, or interference of the contrast of a target by another spatial signal in close proximity in space or time is often referred to as *masking*. The most obvious mask is image noise and the impact of this on threshold grating contrast has been extensively studied. As expected, the presence of two dimensional static white noise will lead to a suppression of the CSF[5,9,10].

The mask can also be another grating. Legge & Foley[11] performed a classic study of contrast masking of a sinewave test signal, spatially superimposed upon a suprathreshold (*pedestal)* masking sinewave. This experiment draws attention to an important difference between contrast *detection* and contrast *discrimination.* In contrast detection (as in the case of the normal CSF), the masking contrast is zero. Contrast discrimination refers to the case when a just noticeable difference has to be observed between two nearly identical gratings that differ in contrast. It is sometimes referred to as *contrast increment detection*. In experiments of the type performed by Legge and Foley, *dipper* functions are obtained. These are plots of just noticeable contrast difference versus pedestal contrast. Typically, as pedestal contrast increases from low values, this JND initially decreases in value, and then finally increases with near Weber Law behavior.

1.2 Contrast perception in complex images

In the above sections, contrast perception has been described from a consideration of simple sinewave stimuli. These studies have provided important information regarding the perception of basic contrast signals subjected to a number of influences, which will occur in natural scenes, e.g. masking and general contextual effects. But they are still, esoteric stimuli. Here we examine studies where real, natural images have been used in the analysis of human spatial vision. Interestingly, there are relatively few, and most are recently published accounts. For static pictures, Bex, et al.[12], have employed a derivative of the Legge & Foley[11] technique to determine contrast discrimination in four natural scenes. Each image was presented on an LCD monitor and spatially filtered in a number of different ways. One modification involved firstly the removal of a 1-octave wide spatially band-filtered image from a given scene. Three peak frequencies were chosen in this filtration process (1, 2 and 4 c/deg). The band-filtered scene was then defined as the standard (pedestal) image. Each of these images had its RMS contrast fixed at one of 10 levels (pedestal contrast). A second test image was defined from the same band-pass image, but now its contrast could be varied. The images were viewed side by side. The increment in test contrast, which resulted in a just noticeable difference between standard and test, was then determined as pedestal contrast varied at each of the three spatial frequencies. *Dipper functions* were obtained. A second experimental modification (defined as a correlated condition) was made which was similar to the band-pass condition, but spatial frequencies outside of the band were not discarded. Again, a dipper function was obtained, and this revealed less discrimination sensitivity compared with the band-pass condition. This is not unexpected in itself, because the correlated condition contains more frequency-related contrast information compared with the band-pass condition, which is likely to offer masking effects on the target octave.

The band-pass condition for the Bex et al. study outlined above is interesting because although in frequency terms, represents only a small representation of a real image, it constitutes a sort of half-way position between an experiment performed on a single sinewave test and one performed on an individual band contained within the full, intact natural image. For the albeit, distorted complex band-pass image configuration of Bex et al., individual points on the dipper function define sensitivity to contrast information within a given frequency band. Only three bands are considered in their study, and at least 10-15 bands would be required for a full frequency response characterization of visual sensitivity, which we define as the *contextual visual perception function* (VPF – see section 4.1). Nevertheless, the filtration technique demonstrated in their band-pass condition will form the basis of our initial planned experiments. The concepts of band filtration and contrast discrimination threshold are still applicable to a dynamic (video) scene.

2. MODELING CONTRAST DETECTION AND DISCRIMINATION

2.1 Modeling Framework

Barten has presented two models relevant to our studies in his classical thesis on contrast sensitivity[1]: a model for *contrast detection* (in chapter 3) and a model for *contrast discrimination* (in chapter 7). We have extended the Barten model for contrast discrimination to provide a full description of the VPF for static images. The original model quantifies the relationship between basic contrast detection (the threshold for just perceiving a spatial contrast signal) and the ability to perceive any contrast changes once this signal is clearly detectable (suprathreshold discrimination). Our extension to this original model involves inclusion of spatial frequency as a variable and also incorporates Barten's model for contrast detection. This expresses *contrast detection,* c_t, as a function of spatial frequency, u, in cycles per degree (c/deg), by:

$$c_t(u) = \left[KO(u)H(u)A(u)\left(\Theta(u)\right)^{-0.5} \right]^{-1} \tag{1}$$

where, K is a constant, and the terms O,H,A and Θ are filter signal transfer characteristics associated with photon capture, lateral inhibition in the retina, and area summation respectively. The term Θ is a combination of both photon noise and internal neural noise. Equation 1 has been extensively studied and most equation parameter values are already known. The remainder, which are specific to particular stimulus conditions, can be readily calculated[13,14]. Our formulation for the VPF is given by:

$$\mathrm{VPF}(u) = \left[\sqrt{\frac{c_t^2(u) + 0.04k^2c_s^2(u)}{1 + 0.004k(c_s(u)/c_t(u))} + c_s^2(u)} - c_s(u) \right]^{-1} \tag{2}$$

Equation 2 shows that if contrast detection c_t is known, the VPF can be immediately calculated for a given pedestal contrast c_s. Potentially therefore the model provides a powerful tool in spatial vision. The constant k is the Crozier factor[15], and represents the minimum signal/noise level required for detecting a spatial signal. In human vision, this is approximately equal to 3.0[16,17]. The inverse of c_t gives the classic contrast sensitivity function. At a given frequency, the VPF represents the reciprocal of the difference between the contrast level of a suprathreshold sinewave signal (c_s) and a level with can just be perceived (discriminated) as different. i.e. VPF = $| c - c_s |$ or Δc. Although Equation 2 has been derived for spatial sinewave stimuli, our initial testing of the model (see section 2.2) indicates its viability for predicting contrast discrimination with narrow band-pass information in real scenes, which is of fundamental importance in our experimental work. Note that as pedestal contrast approaches zero, Δc approaches c_t. In other words, both the CSF and VPF collapse into a single function.

2.2 Testing the CSF and VPF models

In a recent publication by some authors of this paper[18], Barten's contrast sensitivity (detection) model (CSF model, see equation 1), and our extension of his discrimination model (VPF model, see equation 2) have been tested with published data from Bradley and Ohzawa[19], obtained from a sinewave-based experiment. Excellent agreement was seen between data and both models. Further testing of the VPF model using data from a different sinewave experiment, performed by

Yang and Makous[20], revealed that the reciprocal of the VPF can be used to successfully model dipper functions, at least for the four levels of stimulus luminance (25, 2.5, 0.25 and 0.025 cd/sqm) that the data ranged, and a frequency of 2 c/deg.

Finally, the VPF model was further tested with contrast discrimination data, this time obtained from complex images by Bex et all.[12]. The data were fitted to modelled dipper functions for two filtered spatial frequency bands (centers at 1 and 4 c/deg). Again, a good agreement between data and model was observed. This later finding was particularly encouraging, since the study of Bex and his colleagues is the closest published research to our own work (i.e. it employs complex pictorial images as visual stimuli in contrast discrimination experiments).

3. FREQUENCIES, CONTRAST METRICS AND CONTRAST SPECTRA

This section provides details on the spatial frequency decomposition we implement to obtain band-limited images. It also discusses the important issue of the quantification of image (and band) contrast and metrics for its determination.

3.1 Spatial frequency decomposition

We employ spatial frequency (u and v) domain band-pass filters to decompose the image's amplitude spectrum for isolating selected spatial frequency bands. Peli's cosine log cosine filters[21,22] of 1-octave bandwidth centered at frequency 2^i cycles/picture, satisfy our requirements of symmetrical shape on a log frequency axis, together with the property that the image is reconstructed back from all individual bands (plus low and high residuals) by simple addition.

The i^{th} order filter is expressed as:

$$C_i(u,v) = C_i(r) = \begin{cases} 0.5\left[1+\cos(\pi \log_2 r - \pi i\right], 2^{i-1} \le \pi \le 2^{i+1} \\ 0 \quad \text{elsewhere} \end{cases} \tag{3}$$

where r is the radial spatial frequency, $r = (u^2 + v^2)^{1/2}$.

The filters have a number of advantages with respect to our work: they are isotropic (without orientation sensitivity), meaning that they are suitable for our 1D modeling purposes; they are symmetric on a logarithmic scale, as visual channels are; they sum to unity, indicating that the decomposition is complete. They are applied in the spatial frequency domain.

3.2 Contrast metrics

Three relevant complex image contrast metrics are examined in our work. The root-mean-square calculation of image contrast, C_{rms}, has been used extensively in visual studies as well as in imaging applications[12,23,24]. For our purpose, it is calculated from the luminance image, for any selected spatial frequency band in the image, i, as follows:

$$C_{rms\_i} = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\frac{(L_j - \bar{L})^2}{\bar{L}^2}} \tag{4}$$

where N is the number of pixels in the band (equal to that of the original image), L_j is the sum of the three channel displayed luminances of the j^{th} pixel in the band image and $\bar{L}$ is the mean luminance of the image (which is similar, in most occasions, to that of the individual band). The resulting contrast value is in the range of 0 to 1.

Peli's local band-limited contrast[21] takes into account the intensity of a point in the band-pass filtered image and the local background luminance at that point, which varies from place to place in the image. Based on Peli's formulation, we calculate C_{lbl_i} for every pixel (x,y location) in the i^{th} band by:

$$C_{lbl\_i}(x,y) = \frac{|L_i(x,y)|}{L_{lp}(x,y)} \tag{5}$$

where $|L_i(x,y)|$ is the magnitude of the luminance fluctuation from the mean luminance at *x,y* pixel location in the *i*th band image, and $L_{lp}(x,y)$ is the corresponding local luminance in the low-pass filtered version of the image, containing all energy below the band (i.e. the sum of the lower frequency band energies and the low frequency residuals, including the DC which is removed from the pre-filtered image). Equation 5 returns a contrast value at each pixel location, i.e. a contrast image with size equal to that of the original image. Equation 5 can be considered as the Weberian equivalent contrast measurement of local contrast in images and provides for most images mean values between 0 and 1 (although local –point- contrast can exceed 1 in some occasions).

In equation 6, we present a modification to the local band limited contrast (shown in equation 5), where in the denominator the absolute luminance fluctuation is added to the mean luminance. This version returns suitable for the purpose values between 0 and 1 per pixel location; it is expressed by:

$$C_{Mlbl\_i}(x,y) = \frac{|L_i(x,y)|}{L_{pl}(x,y) + |L_i(x,y)|}$$

where $|L_i(x,y)|$ and $L_{lp}(x,y)$ are the same quantities as those in equation 5.

The way we formulate C_{rms}, C_{lbl}, and C_{Mlbl} is analogous to Michelson sinewave contrast, and provides a means of directly quantifying contrast sensitivities in real complex images.

3.3 Contrast spectra

When contrast values, obtained using a given contrast metric, are plotted for a range of frequency bands in an image, *contrast spectra* are created. In a recent publication[18], we used the metrics presented in equations 4 and 5, to derive contrast spectra for a database comprising of 64 digital images with varying scene content, luminance and contrast. It was shown that contrast spectra obtained with the two metrics are not dissimilar, but overall the C_{lbl} metric appeared to ne more sensitive to the individual scenes. In Figure 1 we present mean contrast spectra using equations 4, 5 and 6 for the same image database and a comparison between C_{rms} and C_{lbl}.

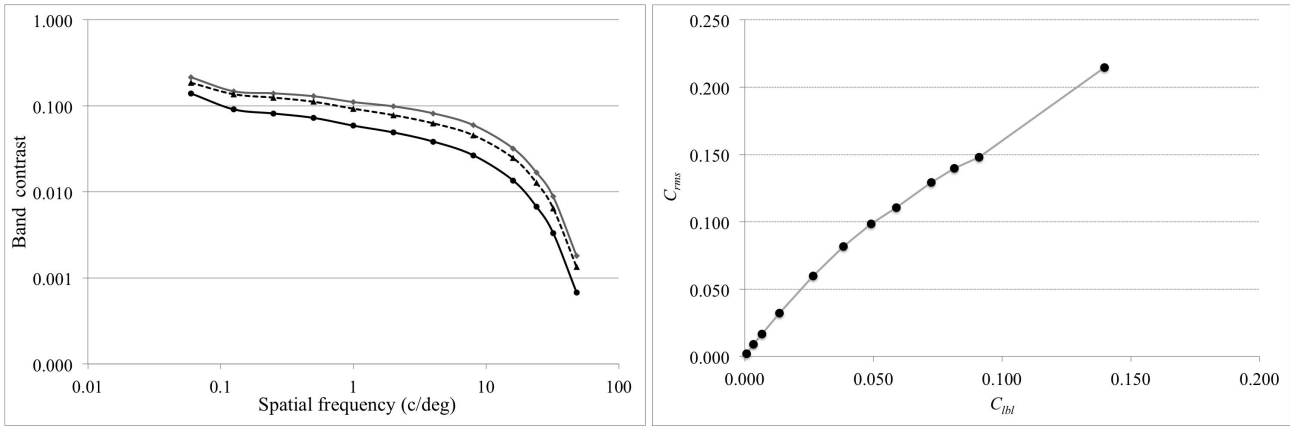

Figure 1. Left: Mean contrast spectra from 64 images[18], using C_{rms} (equation 4 – grey curve), C_{lbl} (equation 5 – black curve) and C_{Mlbl} (equation 6 – broken line curve). Right: Correlation between band C_{rms} and C_{lbl} mean values from digital 64 images[18], 12 spatial frequency bands ranging from 0.06 to 48 c/deg.

We notice that the three curves has very similar profiles and magnitudes, with the C_{lbl}, metric resulting to overall lower values than the other two metrics.

4. EXPERIMENTAL SET-UP FOR MEASURING DETECTION AND DISCRIMINATION

The sections below describe the experimental outline, the setup we have been examining for our psychophysical investigations and the range of stimulus conditions.

4.1 Experimental outline

We measure four visual response functions, as described below. All experiments involve the presentation of an unchanged *standard* stimulus and a modified *test* stimulus.

i) *Isolated Contrast Sensitivity* (iCSF), describes the ability of the visual system to detect any spatial signal in a given spatial frequency octave in isolation and is the closest equivalent to a conventional sinewave CSF. It is measured by presenting the standard, comprising a uniform field of mean luminance equal to the mean luminance of the image, against the test, comprising a variable increment of band contrast around the standard. The comparisons continue until the test image is perceived as just different from the standard. The value of the band contrast of the chosen test image is the *band contrast detection threshold.*

ii) *Contextual Contrast Sensitivity Function* (cCSF), describes the ability of the visual system to detect a spatial signal in a given octave contained within an image. The procedure for obtaining this measurement is the same as for the iCSF, but now both standard and test consist of an image containing all pictorial information outside the band of interest (i.e. the selected band is removed). As in the measurement of the iCSF, the test is shown with a variable increment of band contrast and the comparisons continue until the test is perceived as just different from the standard. Again, the value of the test band contrast is the *contrast detection threshold,* but this time defined for the contextual sensitivity conditions. A comparison between isolated and contextual band conditions indicates the extent that spatial information outside of the band of interest acts as a source of signal masking.

iii) *Isolated Visual Perception Function* (iVPF), describes visual sensitivity to changes in suprathreshold contrast of any spatial signal in a given spatial frequency octave in isolation. The procedure for obtaining this measurement is similar to that used for iCSF determination, but now both standard and test consist of a band in isolation, containing full pictorial information.

iv) C*ontextual Visual Perception Function* (cVPF), describes visual sensitivity to changes in suprathreshold contrast in an image. The procedure for obtaining this measurement is similar to that used for cCSF determination, but now both standard and test consist of an image containing full pictorial information in all bands. The contrast of the band of interest within the test image is progressively reduced until a difference between the test and the standard is just perceived. For a given band, the difference between the contrast of the standard and the contrast of the 'just perceived' test will produce the *contrast difference threshold*.

Figure 2 gives illustrative examples of a standard and a test image we used first for all experimental conditions i), ii), iii) and iv). Note that, at this first stage of our experimentation we have chosen a standard test image of a relatively low contrast. The contrast of the original image was lowered to avoid channel overflow during the filtering process[18] (see section 5 for the contrast spectrum of this image).

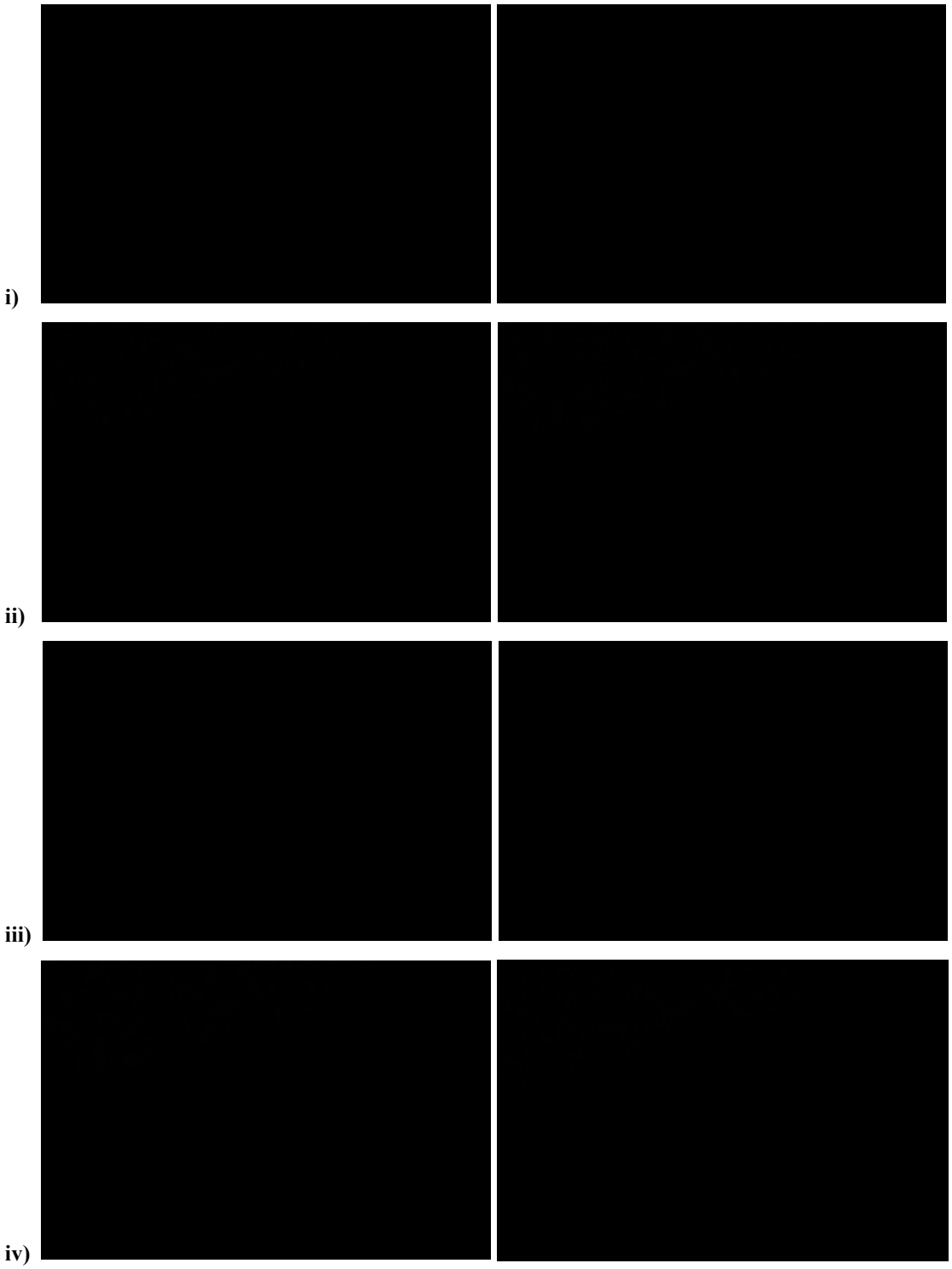

Figure 2: Standard images (left) and test images (right) presented for i) iCSF experiment, ii) cCSF experiment, iii) iVPF and iv) cVPF experiments, using the Low Contrast Gallery image.

4.2 Stimulus conditions

Careful thought has been put into the collection of suitable visual stimuli for our experiments. The acquisition and calibration of the visual stimuli, as well characterization and calibration of the imaging chain we used are detailed in reference 18.

We examine the following stimulus conditions for our experiments with static stimuli:

- Mean stimulus luminance approximately 5-50 cd/m^2 (mesopic and photopic viewing). The observer is adapted to the luminance of the stimulus viewed on the display, with dark surrounding conditions.
- Use of the same image stimuli, with and without colour information.
- Use of the same stimuli, with and without static and/or (shot) dynamic noise.
- Introduce variation in global contrast in the selected test stimuli to examine the role of pedestal contrast in detection and discrimination.
- Prioritise on five selected 'scene types': a) cityscape (including buildings/strong lines); b) leafy/bushy/busy landscape (including large amounts of high frequency information); c) landscape including mainly low frequency information; d) full-body person on a near uniform background e) large groups of people.

The stimulus size and viewing specifications for all conditions above are as follows: field size 15^o subtense; two modes of view: *global free-view scanning* and *local fixation spot*; viewing distance 1.8m; spatial frequencies centred up to 32 cycles/deg.

4.3 Stimulus presentation

Experiments are carried out at an observer viewing distance of 1.8m, which ensures blending of the display grid; the display visual field is 15 degrees. The angular pixel resolution is 0.0155 and 0.0086 degrees respectively. A two alternative-forced-choice (2AFC) interface has been designed to present standard and test images superimposed, with a temporal separation of 250 msc (with a mid-grey screen shown between stimuli presentation) in a random order – a design similar to that proposed by many workers. We have found that this arrangement, which uses one monitor and superimposed stimuli, provides a restricted eye-scanning movement and results in less observer fatigue than settings that present images side-by-side (which we had originally examined). It is important to note that, the presentation of the two stimuli, with a grey screen in between, is revealed to be the best arrangement for our purposes, but relies on short-term observer memory. The observers require initial training with each image stimulus so that i) they can first 'locate' where the contrast changes occur within the entire complex image and ii) fully understand the task and thus provide us with consistent answers.

We have developed a modified staircase method for the presentation of our stimuli, with an adaptive step size for adding/removing contrast in the band of interest during stimuli presentation. It eliminates human decision error while guaranteeing convergence within a predetermined error range. It is based on classical staircase presentations, but has been modified according to a classic perception training method used in neural networks[25]. The test image is separated from the standard by a contrast interval (or step size), *s,* belonging to the range 0 -100% of the full band energy. During the experiment, each observer response corresponds to an increment of time step t_n that marks convergence progress. The test image is initially positioned at the opposite end of this scale to the standard at time t_0, and the step size, *s,* is set to a fraction of the length of the entire scale. The test is then moved towards the standard by the step size at time t_1. After each repositioning of the test at time t_n, the step size is adjusted as a real-valued function of t_n and the observer's latest response patterns, allowing the observer to converge towards the threshold.

Images are presented on a calibrated wide-gamut 24" EIZO Colour Edge CG245W LCD with a mean luminance of 55 cd/sqm. The monitor displays 10-bit color from a 16-bit look-up table at a frequency of 60Hz. It employs built-in hardware calibration and ensures uniformity over 25 points across the screen. The auto-calibration is examined regularly with a Minolta CS-200 luminance and color meter. It is driven by a 1GB NVIDIA Quadro 2000 graphics card, capable to display 10-bit color. The choice of LCDs in visual work has been not advisable (especially the presentation of dynamic stimuli[26]), but much work has been carried out on LCDs in recent years[6,12,27,28], both because CRTs are not longer readily available, but also because both spatial and temporal properties of modern LCDs and associated graphics cards make them suitable for such use.

5. RESULTS AND CONCLUSIONS

5.1 Initial experimental results

Initial sensitivity measurements are presented here, obtained using a minimum of four trained observers. Figure 3 illustrates measurements for both iCSF (black data points) and cCSF (grey data points), obtained with the image stimulus illustrated in Figure 2. The contrast metric used to calculate sensitivities was C_{rms}. Measurements of iCSF closely follow the expected profile and magnitude of a sinewave CSF, appropriate for the stimulus size and mean luminance employed (broken line). This is not unexpected, since the stimuli used for the iCSF experiment contain pictorial information only from a narrow band at a time (pictorial content not dissimilar to images of gratings), on an otherwise uniform background. Measurements of cCSF are shown to be both lower in magnitude and produce a flatter function in profile, showing that basic band detection is masked by the presence of the picture. It is well established that spatial sinewave detection is generally masked when other spatial information is present, whether this be in the form of white, or band filtered 2D noise (Legge, et al, 1987[29]; Pelli, 1985[30]), or 1D compound grating structures (Stomeyer & Julesz, 1972)[10]. In most of these studies, the measured CSF reduces in magnitude, and its peak shifts to relatively high frequencies (10-20 c/deg). A likely candidate for the detection suppression of a given octave band of the cCSF is therefore noise generated from signal bands outside of this octave. Again, previous work on sinewave detection suggests that the majority of this noise will originate from a frequency range of + and – 1 octave from the band (Barten, 1999[1]; Stomeyer & Julesz, 1972))[10].

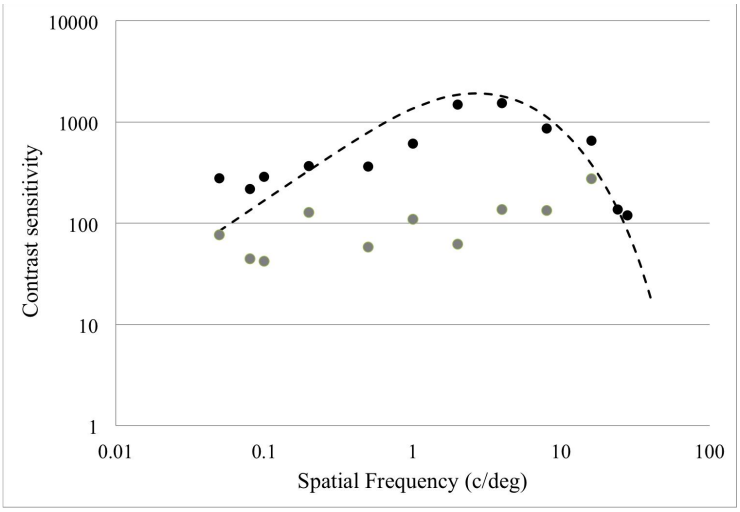

Figure 3: Measurements for iCSF (black data points) and cCSF (grey data points), obtained with the stimulus illustrated in Figure 2.

Contrast discrimination measurements of both iVPF (black data) and cVPF (grey data) from the same image are illustrated in Figure 4. For spatial frequencies above about 0.1c/deg, both functions reveal a gradual increase in discrimination sensitivity as frequency increases. This increase is consistent with previous measurements obtained from band-filtered stimuli. The "dipper" functions obtained by Bex et al (2012)[12] show that, as pedestal contrast (equivalent to our picture C_{rms} band contrast measurements shown in Figure 5) decreases to around 0.005, discrimination sensitivity increases. The low frequency up-turn in sensitivity shown by both data sets in Figure 4 can again be explained through the behavior of picture contrast. As spatial frequency increases to relatively high values, the VPF must peak and then reduce because of the optical limitations of the eye.

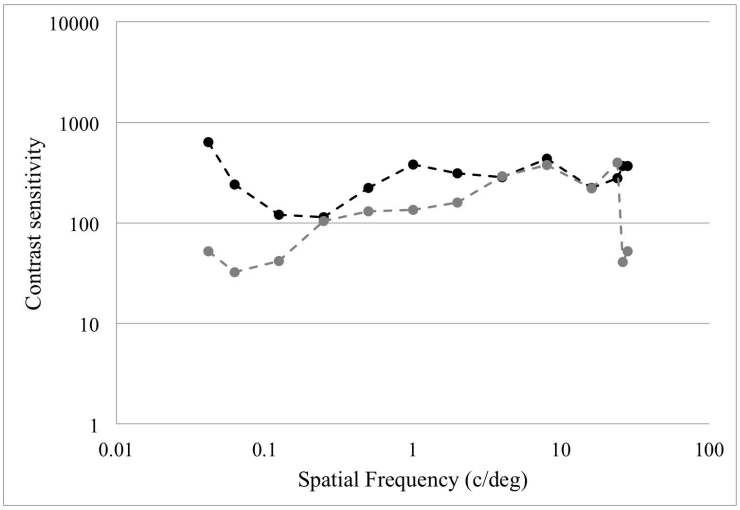

Figure 4: Measurements for iVPF (black data points) and cVPF (grey data points), obtained with the stimulus illustrated in Figure 2.

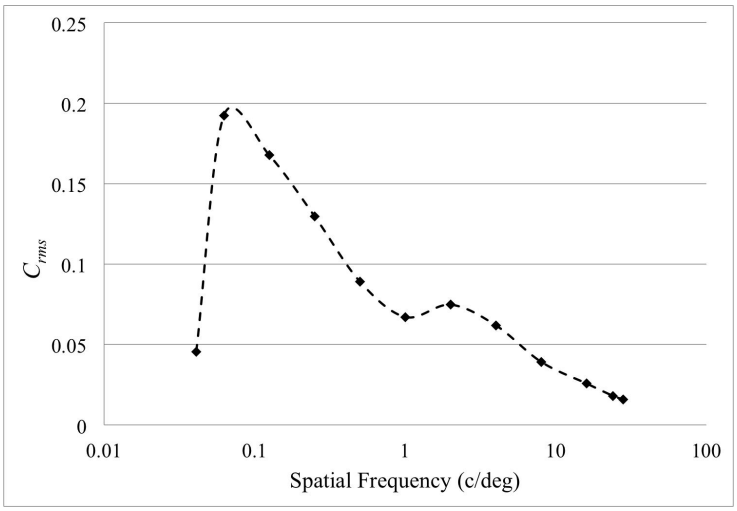

Figure 5: C_{rms} (black data points) measured for each spatial frequency of interest, providing the contrast spectrum for the pictorial image used in our initial experiments.

These findings reveal two important implications for both detection and discrimination of spatial information in pictorial scenes. Firstly, band discrimination sensitivity is likely to be improved if the overall contrast of a scene is reduced (as revealed from considerations of the "dipper" function). Secondly, detection sensitivity of a given band is also likely to improve, if its contextual reduction is due solely to noise masking from surrounding signal bands as argued above. The magnitude of this noise, is directly proportional to the square of the masking band contrast level (Barten, 1999)[1]; lower this contrast level and its ability to mask the absolute detection level of another band will reduce.

The direct consequences of this finding to the design of surveillance and/or field operation instruments would be: human sensitivity to shape and form detection and discrimination is likely to increase when the 'gain' in the instruments/displays is reduced. This may be especially important when the noise of the instrument is relatively high with respect to the signal (i.e. low signal-to-noise ratio). This has positive implications for developing low SWAP viewing instruments, especially hand-held and head-mounted types.

5.2 Summary and conclusions

The paper provides a summary background on human contrast sensitivity and contrast discrimination and discusses relevant measurements, using both simple and more complex visual stimuli. We have outlined our objectives and experimental paradigms for deriving contrast sensitivity and discrimination functions, using complex images and by following experimental procedures that are in-line with relevant work that uses simple visual stimuli. For modeling purposes, we have chosen the Barten models of contrast sensitivity and contrast discrimination. Furthermore, we have expressed Barten's discrimination model as a function of spatial frequency, to derive a model function for contrast discrimination (or visual perception - VPF) and tested it with published data obtained with sinewave gratings, and more recent data obtained with static images[18].

The suitability of contrast metrics used to describe perceived image contrast has been carefully considered, for expressing a relevant measure of sensitivity. The luminance rms (C_{rms}), Peli's local band limited contrast (C_{lbl}) and our modification of the later (C_{Mlbl}) have returned similar contrast values for a large number of test images, although C_{lbl} and our version of it were shown to be more sensitive to scene variability[18]. We provide details on chosen methods of experimentation, equipment and stimulus conditions for the derivation of isolated detection, contextual detection and contextual discrimination functions.

Initial measurements using one image stimulus indicate that the iCSF follows a typical CSF profile (bandpass) and magnitude, while the iVPF has a flatter profile and is overall lower in magnitude. Provided that this finding holds true for a number of image stimuli and different stimulus conditions, it implies that typical CSF derived from simple stimuli (such as sinewave gratings, Gabor functions and band-passed images) do not truly represent the human spatial contrast sensitivity to complex pictorial stimuli. Further, the measured cCSF and cVPF are shown to having profiles and magnitudes not dissimilar to the measured iVPF, especially in the mid-frequency range. Additional measurements using more images and involving different stimulus conditions (see section 4.2) are currently being undertaken in our laboratories, as well as work toward the implementation of Barten's framework to model our observational data.

Future work involves the implementation and testing of cCSF and VPF within image quality (IQ) models. The performance of a number of IQ models can be assessed with our derived cCSFs and VPFs as the visual weighting functions. This performance will be compared with that obtained from currently employed visual weighting functions[1,31,32]. The use of scenes that are relevant to defense and security applications will be made, including those that present the greatest challenge, that is, low contrast noisy images found in most lowlight scenarios.

The implications of this research for hand-held and head-mounted surveillance equipment are significant, with the potential of a reduction in size, weight and power consumption.

REFERENCES

[1] Barten, P. J. G., [Contrast Sensitivity of the Human Eye and its Effects on Image Quality], SPIE Press, Washington, chapters 2 and 7 (1999).

[2] Michelson, A. A., [Studies in Optics], Dover Publications Inc., New York (1995).

[3] Jarvis, J. R., Prescott, N.B. and Wathes, C. M., "A mechanistic inter-species comparison of flicker sensitivity", Vision Res., 43, 1723-1734 (2003).

[4] Jarvis, J. R & Wathes, C. M., "A mechanistic inter-species comparison of spatial contrast sensitivity", Vision Res., 48, 2284-2292 (2008)

[5] Kriss, M., O'Toole, J. and Kinard, J., "Information capacity as a measure of image quality", SPSE Toronto, Canada, 122, (1976).

[6] Bex, P. J., Solomon, S. G. and Dakin, S. C, "Contrast sensitivity in natural scenes depends on edge as well as spatial frequency structure", J. Vision, 9(10), 1-19 (2009).

[7] Watanabe, A., Mori, T., Nagata, S. and Hiwatashi, K., "Spatial sine-wave responses of the human visual system", Vision Res., 8, 1245-1263 (1968).

[8] Georgeson M. A. and Sullivan, G. D., "Contrast constancy: Deblurring in human vision by spatial frequency channels", J. Physiol., 252, 627-656 (1975).

[9] Rovamo, J., Kukkonen, H., Tiipana, K. and Nasanen, R., "Effects of luminance and exposure time on contrast sensitivity in spatial noise", Vision Res., 33, 1123-1129 (1993)

[10] Stromeyer, C.F. and Julesz, B., "Spatial frequency masking in vision: critical bands and spread of masking", J. Opt. Soc. Am. A, 62, 1221-1232 (1972).
[11] Legge, G.E. & Foley, J.M., "Contrast masking in human vision", J. Opt. Soc. Am. A, 70, 1458-1471 (1980).
[12] Bex, P. J., Mareschal, I. and Dakin, S. C., "Contrast gain control in natural scenes", J. Vision, 7(11),1-12 (2007).
[13] Jarvis, J.R. and Wathes, C.M., "On the calculation of optical performance factors from vertebrate spatial contrast sensitivity", Vision Res., 47, 2259-2271 (2007).
[14] Jarvis, J.R. & Wathes, C.M., "Mechanistic modeling of vertebrate spatial contrast sensitivity and acuity at low luminance", Visual Neurosci., 29, 169-181(2012).
[15] Crozier, W.J., "On the variability of critical illumination for flicker fusion and intensity discrimination", J. Gen.l Psychol., 19, 503-522 (1935).
[16] Schade, O., "Optical and photoelectric analog of the eye", J. Opt. Soc. Am, 46, 721-739 (1956)
[17] Roufs, J.A.J., "Dynamic properties of vision – V1. Stochastic threshold fluctuations and their effect on flash-to-flicker sensitivity ratio", Vision Res., 14, 871-888 (1974).
[18] Triantaphillidou S, Jarvis J. and Gupta G., "Contrast sensitivity and discrimination of complex scenes", Proc. SPIE 8653 (2013).
[19] Bradley, A. & Ohzawa, I. "A comparison of contrast detection and discrimination", Vision Res., 26, 991-997 (1986)
[20] Yang, J. & Makous, W. "Modelling pedestal experiments with amplitude instead of contrast", Vision Res., 35, 1979-1989 (1995)
[21] Peli, E., "Contrast in complex images", J. Opt. Soc. Am. A, 7, 2032-2040 (1990).
[22] Peli, E., "Feature detection algorithm based on a visual system model", Proc. IEEE, 90(1), 78-93 (2002).
[23] Frazor, R. A. and Geisler, W. S., "Local luminance and contrast in natural scenes", Vision Res., 46, 1585–1598 (2006).
[24] Triantaphillidou, S., Allen, E. and Jacobson, R., "Image quality of JPEG vs JPEG 2000 image compression schemes, Part 2: Scene dependency, scene analysis and classification", J. Imaging Sci. Techn., 51, 259-270 (2007).
[25] Rojas, R., [Neural networks: a systematic introduction], Springer-Verlag New York, Inc. New York, NY, USA (1996).
[26] Pelli, D. G., and Farell, B., "Psychophysical methods" In M. Bass, C. DeCusatis, J. Enoch, V. Lakshminarayanan, G. Li, C. MacDonald, V. Mahajan and E. V. Stryland (Eds.), [Handbook of Optics], chapter 3 (2010).
[27] Dorr, M. and Bex, P., "A gaze-contingent display to study contrast sensitivity under natural viewing conditions", Proc. SPIE 7865, 7867Y (2011).
[28] Wang, P. and Nicolic, D., "An LCD monitor with sufficiently precise timing for research in vision", Front. Hum. Neurosci., 5(85), 1-10 (2011).
[29] Legge, G.E., Kersten, D and Burgess, A.E, "Contrast discrimination in noise", J. Opt. Soc. Am., A, 4, 391-404 (1987).
[30] Pelli, D.G., "Uncertainty explains many aspects of visual contrast detection and discrimination", J. Opt. Soc. Am. A, 2, 1508-1532 (1985).
[31] Orfanidou, M., Triantaphillidou, S and Allen, E., "Predicting compressed image quality using a modular image difference model", Proc. SPIE 6808, F1-F12 (2008).
[32] Keelan, B. W., Jin, E.W. and Prokushkin, S., "Development of a perceptually calibrated objective metric of noise", Proc. SPIE 7867 (2011).