Vision Research xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

A perceptual space of local image statistics

Jonathan D. Victor *, Daniel J. Thengone, Syed M. Rizvi, Mary M. Conte

Brain and Mind Research Institute, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10065, United States

ARTICLE INFO

Article history:
Received 29 October 2014
Received in revised form 28 May 2015
Accepted 30 May 2015
Available online xxxx

Keywords:
Local features
Visual textures
Metamers
Multipoint correlations
Intermediate vision

ABSTRACT

Local image statistics are important for visual analysis of textures, surfaces, and form. There are many kinds of local statistics, including those that capture luminance distributions, spatial contrast, oriented segments, and corners. While sensitivity to each of these kinds of statistics have been well-studied, much less is known about visual processing when multiple kinds of statistics are relevant, in large part because the dimensionality of the problem is high and different kinds of statistics interact. To approach this problem, we focused on binary images on a square lattice – a reduced set of stimuli which nevertheless taps many kinds of local statistics. In this 10-parameter space, we determined psychophysical thresholds to each kind of statistic (16 observers) and all of their pairwise combinations (4 observers). Sensitivities and isodiscrimination contours were consistent across observers. Isodiscrimination contours were elliptical, implying a quadratic interaction rule, which in turn determined ellipsoidal isodiscrimination surfaces in the full 10-dimensional space, and made predictions for sensitivities to complex combinations of statistics. These predictions, including the prediction of a combination of statistics that was metameric to random, were verified experimentally. Finally, check size had only a mild effect on sensitivities over the range from 2.8 to 14 min, but sensitivities to second- and higher-order statistics was substantially lower at 1.4 min. In sum, local image statistics form a perceptual space that is highly stereotyped across observers, in which different kinds of statistics interact according to simple rules.

1. Introduction

The analysis of image statistics underlies many key components of intermediate visual processing, including not only visual texture, but also visual characterization of surfaces and segmentation of images into objects. Although each of these tasks might at first seem deterministic, each is fundamentally statistical. For example, identification of surface materials (such as wood, grass, or hair) is not carried out by matching the image to a stored sample, but rather, by their image statistics, such as the range of contrasts and colors and the distribution of oriented contours at different scales (Karklin & Lewicki, 2009). Segmentation of an image is a statistical task as well, because it is fundamentally ambiguous: multiple scene interpretations are consistent with a single image, and image statistics play a key role in assessing which one is chosen as the most plausible. For example, contours due to a shadow or change in illumination are not typically coincident with a change in material properties, while real object boundaries typically have such changes, and hence, changes in image statistics.

Thus, understanding the processing of image statistics has broad importance as part of a foundation for understanding many aspects of intermediate visual processing. Visual textures, the focus here, present image statistics in their purest form.

While natural textures are characterized by many kinds of statistical features, systematic approaches to studying visual texture (with few exceptions (Motoyoshi & Kingdom, 2007; Saarela & Landy, 2012; Victor, Chubb, & Conte, 2005)) usually explore just one kind of feature, such as luminance distributions (Chubb, Econopouly, & Landy, 1994; Chubb, Landy, & Econopouly, 2004), color (Li & Lennie, 1997), orientation (Landy & Oruc, 2002; Wolfson & Landy, 1995, 1998), or curvature (Ben-Shahar & Zucker, 2004). There are two main reasons for this. One is the high dimensionality of the problem: if all kinds of statistical features were explored, the number of parameters (i.e., the number of different image statistics) would be impractically large. The other is that image statistics exhibit a high degree of interdependency. Edges cannot exist without local changes in luminance, and corners cannot exist without edges at multiple orientations, so these statistics cannot be considered to be independent attributes. Here, we attempt to address both issues, by constructing a texture space of large but manageable dimension (10), whose coordinates take into account the interactions implied by geometry. The data show that once these steps are taken, the perceptual interactions of image statistics obey simple rules that (a) are highly consistent across subjects, (b) accurately predict sensitivity to complex

* Corresponding author. Fax: +212 746 8050.
E-mail address: jdvicto@med.cornell.edu (J.D. Victor).

http://dx.doi.org/10.1016/j.visres.2015.05.018
0042-6989/

combinations of image statistics, and (c) are approximately preserved across a range of spatial scales.

To overcome the problem of high dimensionality (specifically, that an image statistic can be defined from the joint probabilities of any set of gray levels at any configuration of nearby points), we restricted consideration to black-and-white images on a checkerboard. By restricting the analysis to a single scale and only two luminance levels, we can then consider *all* possible local images statistics – i.e., the probabilities of all configurations of black and white checks within a 2×2 neighborhood. This set of image statistics has 10 free parameters (summarized here in Methods; detailed in (Victor & Conte, 2012)). It encompasses not only the intuitively-important features of luminance, contrast, edge, and corner, but also, its four-point correlations are independently informative for natural images (Tkačik et al., 2010). Thus, although it is a reduced space, it has image statistics of many different types and levels of complexity.

To overcome the second hurdle, the interdependency of different kinds of stimulus features, we used a "maximum-entropy" approach. That is, we specify stimuli by the prevalence of one or more elementary features, and then synthesize an ensemble of images that meet these specifications but are otherwise as random as possible. This limits the interdependence of features to what is implied by geometry, so that observed interactions at the level of neural or perceptual responses can be more readily interpreted.

1.1. Texture space and color space: their geometry and its implications

The above considerations lead to the construction of a "texture space", in which each point corresponds to a specific combination of image statistics that together specify luminance distributions and the prevalence of edges and corners at different orientations (Victor & Conte, 2012). The experiments presented here determine the perceptual distances in this space, focusing on the region near its origin.

The analogy with trichromatic color space provides a helpful geometrical framework. In both color space and texture space, points represent stimuli and the origin represents the neutral point (in color space, a white light; here, the random texture). The present experiments, which consist of measuring thresholds for perceiving that a texture is not random, correspond to measuring thresholds to changes in color and intensity near the white point. In both spaces, a line segment space represents mixtures. In color space, the points on a line segment are the colors that can be created by mixing the lights that correspond to the endpoints. In the space of local image statistics, the points on a line segment are the textures that can be created by mixing the textures that correspond to the endpoint. In color space, mixtures are created by physical mixing of lights; here, mixtures are created at the level of statistics: at the level of the frequency of each way that a 2×2 block can be colored with black and white checks (as described in (Victor & Conte, 2012). In color space and in texture space, a ray emanating from the origin corresponds to a set of stimuli that are progressively more saturated. Thus, determining the point along this ray that is first discriminable from the origin is a way of quantifying sensitivity to the combination of features represented by the direction of the ray. By determining the thresholds for rays that emanate from the origin in many directions, one can map out the "isodiscrimination surface," which summarizes the perceptual sensitivities in the neighborhood of the origin. In the case of color space, the isodiscrimination surfaces are approximately ellipsoids (the "Macadam ellipses" (Macadam, 1942)), and below we find that this holds in texture space as well.

The notion of navigating the space by moving along a straight line trajectory brings up an important mathematical distinction between the geometries of the two spaces. In color space, moving along a line is straightforward: it corresponds to increasing or decreasing the intensity of a light. For textures, this is not the case. For example, increasing the number of edges may also increase the number of intersections, and the proportionality between corners and intersections is typically nonlinear. These nonlinear dependencies underlie the maximum entropy approach (Victor & Conte, 2012) for navigating the space: a direction in the space corresponds to a specified coordinate, and movement along this direction may take a curved path to minimize the introduction of further structure. That is, the maximum-entropy approach yields a locally flattened coordinate system. Here, since we are studying discrimination thresholds, we work in these local coordinates, and ignore the impact of global curvature.

Color space and texture space have other important differences, and these allow us to interpret the sensitivity measurements in a way that has no immediate analogy in color space. The differences go beyond the difference in dimensionality or global curvature, and trace back to a fundamental difference in the way that the coordinate systems are defined. For color space, the origin of the coordinate system – the white point – is defined subjectively. For image statistics, the origin of the texture space has an *a priori* mathematical definition: it is the texture in which each check is randomly and independently assigned to black or white. A similar distinction applies to the axes: for color space, axes are defined empirically based on cone excitations (MacLeod & Boynton, 1979) or combinations motivated by physiological and psychophysical measurements (Derrington, Krauskopf, & Lennie, 1984); for image statistics, axes are defined *a priori* mathematically, in terms of correlations.

The kind of geometry that applies to the two spaces is also different (Zaidi et al., 2013). In color space, any of several coordinate systems (Derrington, Krauskopf, & Lennie, 1984; MacLeod & Boynton, 1979; Wyszecki & Stiles, 1967), each based on its own set of empirical observations, are equally valid descriptions of the space. Changing from one set of axes to another is a general linear transformation, which means that distances and angles computed from the coordinates in one system (via the Pythagorean rule and dot-products) need not match values computed with another. In the space of local image statistics, the coordinates are defined by mathematical considerations. This means that there is a standard definition of distance, and a "sphere" is a well-defined term: it is the locus of points that are at an equal distance from its center.

Because of the mathematics underlying the texture-space coordinates, spheres centered at the origin have another interpretation. Specifically, spheres are the isodiscrimination surfaces for an ideal observer ((Victor & Conte, 2012), Appendix B), i.e., an observer who is able to make full use of all image statistics. Of course we do not anticipate that human performance will resemble this. Rather, we expect that human observers will be selective, and make use of some image statistics more efficiently than others. This will distort the human isodiscrimination surface away from a spherical shape. For example, if sensitivity is reduced along one axis, then the isodiscrimination surface will become elongated in that direction, into an ellipsoid. If sensitivity is different for positive vs. negative changes in a coordinate, the surface will be asymmetrically distorted (i.e., it will become egg-shaped). If cues along different coordinates are not combined, the shape of the isodiscrimination surface will become squared-off. But as the results show, only the first kind of distortion is prominent, and this enables a concise, predictively accurate model for sensitivity to complex combinations of image statistics.

2. Methods

2.1. The stimulus space

The goal of these experiments is to determine visual sensitivity to local image statistics, individually and in combination. To do

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

this, we use stimuli in which one or more local statistics are specified, and all other aspects of the image are as random as possible, subject to these specifications. We consider only binary (black-and-white) images; for such images, there are 16 ways in which a 2×2 block can be colored ($16 = 2^{2\times 2}$). These 16 potential degrees of freedom are reduced to 10, because the 2×2 blocks must match where they overlap. Here we describe 10 natural coordinates that capture these degrees of freedom, and how they are combined to generate the stimuli used in the experiments. For a complete description of the space and details of the algorithms to construct the stimuli within it, see (Victor & Conte, 2012).

The 10 coordinates each range from -1 to 1, and a completely random binary image corresponds to a value of 0 for each coordinate (Fig. 1A). The coordinates fall into four groups, based on the order of the correlation that they describe. (The "order" of a correlation means the number of image points that must be simultaneously considered to determine the correlation's value.)

There is one first-order coordinate, γ. This captures the overall luminance bias of the image: it is the difference between the probability of a white check and the probability of a black check. Thus $\gamma = 1$ means that all checks are white, $\gamma = -1$ means that all checks are black, and $\gamma = 0$ means that both colors are equally likely.

There are four second-order coordinates (denoted $\beta_{\_}$, $\beta_{|}$, $\beta_{\backslash}$, and $\beta_{/}$), each of which describe a two-point correlation at the orientation indicated by the subscript. Each is the difference between the probability that two neighboring checks match (i.e., both are white or both are black), and the probability that they do not match (i.e., one is white and one is black). We use the convention that white checks are denoted by 1, and black checks are denoted by 0. So for example, $\beta_{\_} = 1$ means that all 1×2 blocks are either $(0\ \ 0)$ or $(1\ \ 1)$; in such images, horizontal rows will be either all black or all white, but there will be no relationship between the colors of adjacent rows. $\beta_{\_} = -1$ means that all 1×2 blocks are either $(0\ \ 1)$ or $(1\ \ 0)$ and none are $(0\ \ 0)$ or $(1\ \ 1)$; in such images, horizontal rows will have alternating black and white checks. Intermediate values of $\beta_{\_}$ indicate a bias toward matching neighbors (positive values) or mismatching neighbors (negative values), and $\beta_{\_} = 0$ means that matching and mismatching neighbors are equally likely. Similarly, $\beta_{|}$, $\beta_{\backslash}$, and $\beta_{/}$ describe the two-point correlations in the vertical and two diagonal directions. Note that for $\beta_{\_}$ and $\beta_{|}$ (which we call the "cardinal" β's), the two checks involved in the correlation are abutting – so $\beta_{\_}$ and $\beta_{|}$ control the frequency of edges. In contrast, for $\beta_{\backslash}$, and $\beta_{/}$, (which we call the "diagonal" β's) the involved checks only share a common vertex. Thus, we may anticipate that there are differences between the visual impact of the cardinal β's and the diagonal β's. These differences need not be related to the classical oblique effect (Doi, Balcan, & Lewicki, 2007) – since they describe different *kinds* of correlations, not merely correlations that differ by a rotation.

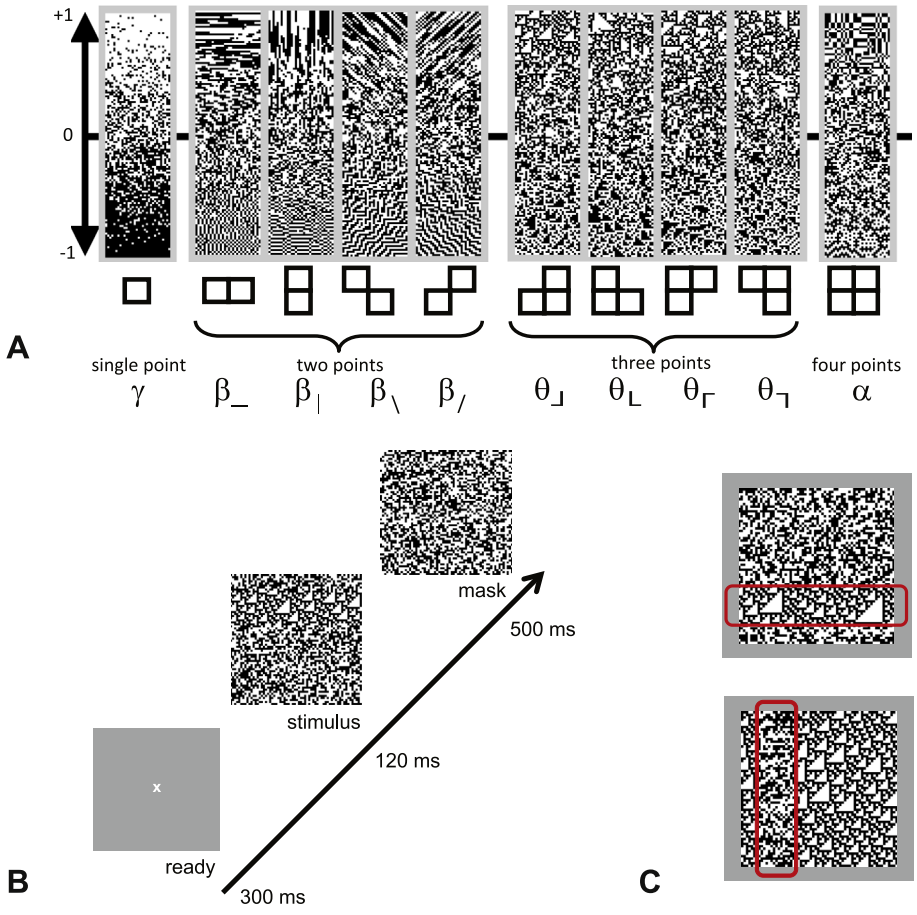

Fig. 1. (A) The gamuts of the 10 image statistics. In each case, the coordinate value (vertical scale) indicates the correlation strength, which can range from -1 to 1. (B) Timecourse of a typical trial. Stimuli were typically presented for 120 ms, followed by a random mask. (C) The stimulus was a 64×64 array of checks that contained a 16×64 target positioned 8 checks away from one of the four edges. Stimuli were of two types: *top*, random background with structured target (here, $\theta_{\lrcorner} = 1$), or *bottom*, structured background with random target. Red rectangles indicate the target, and were not visible during the trials. Panels B and C reproduced, with permission of the copyright holder, The Association for Research in Vision and Ophthalmology, from (Victor, Thengone, & Conte, 2013). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

Next, there are four third-order coordinates, $\theta_{\lrcorner}$, $\theta_{\llcorner}$, $\theta_{\ulcorner}$, and $\theta_{\urcorner}$. Each quantifies a three-point correlation within an L-shaped region, by comparing the probability that the region contains an even number of white checks, vs. an odd number of white checks. A value of +1 means that only an odd number of white checks (one or three) are present. For example, $\theta_{\llcorner} = 1$ means that only the configurations $\begin{pmatrix}1 & \\ 1 & 1\end{pmatrix}$, $\begin{pmatrix}1 & \\ 0 & 0\end{pmatrix}$, $\begin{pmatrix}0 & \\ 1 & 0\end{pmatrix}$ or $\begin{pmatrix}0 & \\ 0 & 1\end{pmatrix}$ are present; this will lead to images with prominent white triangular-shaped regions pointing downward and to the left. $\theta_{\llcorner} = -1$ means that only the complementary configurations $\begin{pmatrix}0 & \\ 0 & 0\end{pmatrix}$, $\begin{pmatrix}0 & \\ 1 & 1\end{pmatrix}$, $\begin{pmatrix}0 & \\ 0 & 1\end{pmatrix}$ or $\begin{pmatrix}1 & \\ 1 & 0\end{pmatrix}$ are present; these images have prominent black triangular-shaped regions.

The final coordinate, α, quantifies the four-point correlation among the checks in a 2×2 block: $\alpha = 1$ means that an even number of them are white, and $\alpha = -1$ means that an odd number are white. This is the same image statistic that has been the subject of much previous work (Julesz, Gilbert, & Victor, 1978; Victor, Chubb, & Conte, 2005; Victor & Conte, 1989, 1991, 1996, 2004).

Together, the ten coordinates $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$ fully specify the distribution of colorings in 2×2 blocks.

2.2. Stimuli

All experiments used the texture segmentation paradigm (Fig. 1B and C) first developed by Chubb and coworkers for the study of textures in which each check's luminance is independently chosen from the same distribution (Chubb, Landy, & Econopouly, 2004), and later used for correlated textures (Victor, Chubb, & Conte, 2005; Victor & Conte, 2012; Victor, Thengone, & Conte, 2013). The psychophysical paradigm and stimulus layout is taken from the latter studies, and is summarized here.

The basic stimulus (Fig. 1B and C) consisted of a 64×64 array of checks, which contained a 16×64 rectangular target, positioned 8 checks away from one of the four edges of the array. The target was distinguished from the remainder of the array by its statistics, i.e., by one or more values of $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$. The subject's task was to identify the location of the target via key-press on a response box.

To ensure that the subject identified the target by segmenting it from the background rather than, for example, by identifying a texture gradient (Wolfson & Landy, 1998), two types of stimuli were constructed (Fig. 1C): (1) a random background array, with a target that had a nonzero value of one or more image statistics, and (2) a background array in which the image statistics had a nonzero value, with a target that was random. Our rationale (both here and in previous studies mentioned above using this design) for requiring segmentation rather than just gradient detection was to reduce intra- and inter-subject variability due to strategy choice. These two types of stimuli were randomly interleaved, and our analyses are based on the pooled responses.

There were two kinds of experiments. In the first, we measured sensitivity to individual image statistics and their pairwise combinations. These measurements were used to construct a phenomenological model, and we used the model to predict sensitivities to combinations of multiple image statistics. In the second kind of experiment, we measured sensitivities to these combinations.

For the first kind of experiment, each session examined a single plane of the stimulus space (i.e., a set of images in which two of the coordinates $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$ were given specific nonzero values.) Since there are 10 coordinates, there are $45 = 10 \times 9/2$ different planes, one corresponding to each coordinate pair. However, many of these planes differ only by a 90° rotation or mirror reflection. Pilot studies and previous work (Victor, Thengone, & Conte, 2013) showed that visual sensitivities were not affected by these transformations, so we focused on a 15-plane subset (the ones shown in Fig. 4) that included all combinations, once these symmetries are taken into account.

Each of these 15 planes was explored in a radial fashion, in most cases by choosing points at several distances from the origin along 8 rays. For the four rays along the coordinate axes (the positive and negative directions of the two axes in the plane), we used five equally-spaced values, with the maximal values chosen based on pilot experiments to ensure that performance would span the range from floor to ceiling: ±0.25 for γ, ± 0.45 for the cardinal β's, ±0.75 for the diagonal β's, ±1.0 for the θ's, and ±0.85 for α. The other four rays pointed into each of the four quadrants, with maximal values chosen in approximate proportion to the above maximum values. These rays were sampled at two points: the maximal point, and a point in the same direction at a relative distance of 0.7 from the origin. In two of the planes $(\gamma, \theta_{\lrcorner})$ and $(\theta_{\lrcorner}, \alpha)$, we used a minor modification for consistency with an earlier pilot dataset, in which there were two rays pointing into each quadrant, and all 12 rays (4 along the axes, and 8 oblique) were sampled at three equally-spaced points. After assigning the in-plane coordinates as just described, the 8 unspecified coordinates were determined (see (Victor & Conte, 2012) Table 2) by first, setting all values of lower-order coordinates to zero, and then, setting the remaining coordinates to values that maximized the entropy of the resulting images. Details on construction of these stimuli are provided in (Victor & Conte, 2012).

The second kind of experiment measured sensitivity to combinations of multiple image statistics. We considered 12 such combinations, each corresponding to a vector $\vec{c}$ (specified below) in the 10-coordinate space. There were four kinds of sessions, each of which focused on 3 of these "out-of-sample" directions, in two opposite rays (i.e., $+\vec{c}$ and $-\vec{c}$). The four subjects who participated in these experiments included two subjects who also participated in the first set of experiments 6 months to 2 years previously (MC and DT), and two who had not (SR and KP). Each session therefore also included stimuli along two of the coordinate axes (in positive and negative directions, as described above); for MC and DT, this served as a check for a change in sensitivity over time, for the others, it served as a measure of overall sensitivity to the on-axis stimuli to scale the model predictions.

The 12 out-of-sample directions were chosen to explore (i) the range of predicted sensitivities, (ii) the combinations of statistics typically present in natural images, and (iii) the combinations that determine the topological properties of a texture. For (i), we made use of the finding that the in-plane measurements suggested a quadratic perceptual metric, which can be captured in a symmetric matrix (see below). The eigenvectors corresponding to the largest and smallest eigenvalues of this matrix are thus the predicted directions of maximal and minimal sensitivities. Since there are 10 dimensions, there are 10 eigenvectors in all, covering the range of predicted sensitivities (which, as shown below, included sensitivities predicted to be both above and below the range of the in-plane measurements). Three of these 10 eigenvectors were not tested, since they were included in the stimuli used to build the model (one is in the $(\beta_{\_}, \beta_{|})$-plane and two are in planes spanned by the θ's; see "Eigenvector classes" below), The remaining 7 out-of-sample eigenvectors (listed in Table 3), which included the largest of the ten eigenvalues (sym1) and the smallest (hvi2), constituted the first 7 out-of-sample test directions $\vec{c}$. Note that since these are the eigenvectors of a symmetric matrix, they are necessarily orthogonal – and thus, probe different directions in the perceptual space.

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

For the two subjects that participated in both kinds of experiments, we used the eigenvectors calculated from their sensitivities in the 15 coordinate planes. For the two naïve subjects, we used the eigenvectors calculated from the average of the four subjects who participated in the 15-plane experiments (which were in good agreement, see Fig. 5 and Table 3). These choices also satisfied criterion (ii), as the predicted directions of maximal and minimal sensitivity correspond closely to axes of greatest and least variation in natural images (Hermundstad et al., 2014). Note that the specific directions were determined from experimental data, but the directions themselves did not correspond to any of the data used to fit the model.

To satisfy criterion (iii), i.e., to sample the combinations that determine the topological properties of a texture, we made use of the Minkowski image functionals (Michielsen & De Raedt, 2001), a series of measures that can be applied to images to extract their topological features. For binary images, there are three cardinal functionals, typically denoted A, U, and χ. Considering the images to represent material (with the black checks representing substance, and the white checks representing empty space), the three functionals have simple meanings: A is density (amount of substance per unit area), U is perimeter length (amount of boundary per unit area), and χ is porosity (number of holes per unit area). The first two of these functionals correspond to axes or planes in our coordinate system: $A = (1-\gamma)/2$ and $U = 1-(\beta_{\_}+\beta_{|})/2$, and therefore would not provide out-of-sample tests. So we focused on χ, which, other than an additive offset, is a linear combination of multiple coordinates. This choice is also motivated by recent findings that demonstrate human sensitivity to this image attribute(Barbosa, Bubna-Litic, & Maddess, 2013). The precise relationship of χ to the coordinates depends on how one defines connectivity on the checkerboard lattice: 4-connected (i.e., the material is contiguous across two checks that share a common side) vs. 8-connected (i.e., the material is contiguous across two checks that share a common side or a common vertex). The relationship can be obtained by transforming standard formulae (Eq. (5) of (Michielsen & De Raedt, 2001)) into our coordinates (see (Barbosa, Bubna-Litic, & Maddess, 2013) Appendices A and B for a related calculation):

$$\chi^{[4]} - \frac{1}{16} = \frac{1}{16}(-4\gamma - 2\beta_{\_} - 2\beta_{|} + \beta_{/} + \beta_{\backslash} + \theta_{\lrcorner} + \theta_{\llcorner} + \theta_{\ulcorner} + \theta_{\urcorner} + \alpha) \tag{1}$$

and

$$\chi^{[8]} + \frac{1}{16} = \frac{1}{16}(-4\gamma + 2\beta_{\_} + 2\beta_{|} - \beta_{/} - \beta_{\backslash} + \theta_{\lrcorner} + \theta_{\llcorner} + \theta_{\ulcorner} + \theta_{\urcorner} - \alpha). \tag{2}$$

We therefore used these directions (the right hand sides of Eqs. (1) and (2) as out-of-sample test directions $\vec{c}$, as well as their sums and differences:

$$\chi^{[4]} + \chi^{[8]} = \frac{1}{8}(-4\gamma + \theta_{\lrcorner} + \theta_{\llcorner} + \theta_{\ulcorner} + \theta_{\urcorner}), \tag{3}$$

$$\chi^{[4]} - \chi^{[8]} = \frac{1}{8}(1 - 2\beta_{\_} - 2\beta_{|} + \beta_{/} + \beta_{\backslash} + \alpha). \tag{4}$$

For completeness, we also considered a 6-connected lattice, in which there is connectivity across sides and two of the vertices (6L: vertices running from upper left to lower right, and 6R, vertices running from upper right to lower left). These yield:

$$\chi^{[6L]} = \frac{1}{8}(-2\gamma + \theta_{\llcorner} + \theta_{\urcorner}) \text{ and } \chi^{[6R]} = \frac{1}{8}(-2\gamma + \theta_{\ulcorner} + \theta_{\lrcorner}). \tag{5}$$

To obtain mixtures of multiple coordinates, our final out-of-sample test direction $\vec{c}$ was the difference between these quantities,

$$\chi^{[6L]} - \chi^{[6R]} = \frac{1}{8}(-\theta_{\lrcorner} + \theta_{\llcorner} - \theta_{\ulcorner} + \theta_{\urcorner}). \tag{6}$$

(The sum of these quantities was not used, as $\chi^{[6L]} + \chi^{[6R]} = \chi^{[4]} + \chi^{[8]}$, duplicating Eq. (3)). Thus, the Minkowski image functionals, along with several choices for lattice connectivity, yielded 5 out-of sample test directions, which are listed in Table 1. The corresponding stimuli were generated by iterative applications of the "donut algorithm" of (Victor & Conte, 2012) to mix in-plane textures.

Stimuli were presented on a mean-gray background, followed by a random mask (Fig. 1B). The display size was 15 × 15°, contrast was 1.0, and viewing distance was 1 m. The standard check size was 14 min (10 × 10 hardware pixels). In some experiments, check size was varied in steps down to 1.4 min (1 × 1 hardware pixel); in these cases, the stimulus still consisted of a 64 × 64 array with a 16 × 64 target; it was centered at fixation contained and surrounded by mid-level gray. Studies were carried out on an LCD monitor with a mean luminance of 23 cd/m^2, a refresh rate of 100 Hz, and a presentation duration of 120 ms unless otherwise specified, driven by a Cambridge Research ViSaGe system.

2.3. Subjects

Studies were conducted in 16 normal subjects (8 male, 8 female), ages 21–54. Four subjects (MC, DT, JD, DF) completed the experiments to assess sensitivities to all 10 image statistics and 15 of their pairwise combinations. Four subjects, including two of the above (MC and DT) and two additional subjects (SR and KP) completed the experiments to test model predictions with out-of-sample stimuli. Six subjects (MC, DT, SR, KP, RS, and SP) completed the experiments that examined the effect of check size. Three subjects (MC, DT, DF) completed the experiments that examined the effect of stimulus duration. The remaining eight subjects did not complete any of the above experiments, and their data are only used to assess sensitivities along two or more coordinate axes at the standard check size of 14 min.

Of the 16 subjects, MC is an experienced psychophysical observer, and the other subjects had no prior viewing experience. All subjects other than MC, DT, SR, and AA were naïve to the purposes

Table 1
Directions in stimulus space corresponding to the Minkowski image functional χ of (Michielsen & De Raedt, 2001), which describes porosity, and several derived quantities. M4 and M8 correspond to the functionals $\chi^{[4]}$ and $\chi^{[8]}$ on a 4-connected or 8-connected lattice; M4 + 8 and M4–8 are their sums and differences, and M6L–R is derived from a 6-connected lattice (see "Stimuli" in Methods for details). All eigenvectors are normalized to Euclidean length 1.

Designation	γ	$\beta_{\_}$	$\beta_{\|}$	$\beta_{\backslash}$	$\beta_{/}$	$\theta_{\lrcorner}$	$\theta_{\llcorner}$	$\theta_{\ulcorner}$	$\theta_{\urcorner}$	α
M4	−0.718	−0.359	−0.359	0.180	0.180	0.180	0.180	0.180	0.180	0.180
M8	−0.718	0.359	0.359	−0.180	−0.180	0.180	0.180	0.180	0.180	−0.180
M4 + 8	−0.894	0.000	0.000	0.000	0.000	0.224	0.224	0.224	0.224	0.000
M4–8	0.000	−0.603	−0.603	0.302	0.302	0.000	0.000	0.000	0.000	0.302
M6L–R	0.000	0.000	0.000	0.000	0.000	−0.500	0.500	−0.500	0.500	0.000

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

of the experiment. All subjects had visual acuities (corrected if necessary) of 20/20 or better.

This work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki), and with the approval of the Institutional Review Board of Weill Cornell, and with the consents of the individual subjects.

2.4. Procedure

Subjects were asked to identify the position of the target, in a four-alternative forced choice (4-AFC) texture segregation task (Fig. 1B and C). They were informed that the target was equally likely to appear in any of four locations (top, right, bottom, left), and were shown examples of stimuli of both types: target structured/background random and target random/background structured. They were asked to maintain central fixation, rather than to attempt to scan the stimulus. Auditory feedback for incorrect responses was given during training trials. After performance stabilized (approx. 2 h for a new subject), blocks of trials (with trials presented in randomized order) were presented. Block order was counterbalanced across subjects. Feedback was not given during data collection to minimize the likelihood of learning over the course of the experiment, as collection of complete datasets occupied approximately a year (see below). Thresholds for conditions that were tested at the beginning and end of the testing period were in good agreement.

Experiments were organized into several kinds of blocks. In the blocks used to probe sensitivity in coordinate planes, stimuli were placed along the positive and negative directions on each of two axes (4 directions, 5 strengths), and in four oblique directions in the plane that these axes determined, at 2 strengths, with the latter repeated twice, for a total of $36 = 4 \times 5 + 4 \times 2 \times 2$ stimulus specifications. This design was used for all coordinate planes except for (γ, α) and $(\theta_{\lrcorner}, \alpha)$; in this case, we tested the four directions along the two axes and 8 oblique directions, each at 3 strengths, also for a total of $36 = 4 \times 3 + 8 \times 3$ stimulus specifications. In the blocks used to test sensitivity to complex mixtures, we tested 10 directions, each at 4 strengths ($40 = 10 \times 4$ stimulus specifications). Each stimulus specification was used eight times per block: once in each of four target locations, and in two configurations: target structured/background random and target random/background structured (Fig. 1C). This resulted in $288 = 8 \times 36$ to $320 = 8 \times 40$ trials per block. We collected 8 blocks per subject in the check size experiments at each of 4 check sizes ($9216 = 288 \times 4 \times 8$ trials). For the other experiments, we collected 15 blocks per subject per condition ($4320 = 288 \times 15$ to $4800 = 320 \times 15$ trials). This yielded 64–240 responses per coordinate in the stimulus space. A complete set of measurements of sensitivities along all axes and in each of the 15 coordinate planes required $64800 = 15 \times 4320$ trials per subject, and was carried out over approximately 1 year.

2.5. Analysis

2.5.1. Determination of thresholds from psychophysical data

The first step in data analysis consisted of determining thresholds in each tested plane (i.e., along each ray emanating from the origin). We adapted the procedure of Victor, Chubb, & Conte (2005) and Victor, Thengone, & Conte (2013), as summarized here. Along each ray r, we fit measured values of the fraction correct (FC) fit to Weibull functions via maximum likelihood:

$$\mathrm{FC}(x) = \frac{1}{4} + \frac{3}{4}\left(1 - 2^{-(x/a_r)^{b_r}}\right), \tag{7}$$

where x is the distance from the origin, a_r is the fitted threshold (i.e., the value of x at which FC = 0.625, halfway between chance, 0.25, and perfect, 1.0), and b_r is the Weibull shape parameter. The distance x is determined in the ordinary Euclidean fashion: on the coordinate axes, it is the absolute value of the coordinate; on the oblique rays, it is calculated as $x = \sqrt{c_y^2 + c_z^2}$, where c_y and c_z are the values of the two coordinates, each drawn from $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$ that specify the stimulus. In nearly all cases, the exponent b_r had 95% confidence limits that included the range 2.2 to 2.7. Since our focus was on thresholds, we therefore refit the data from all rays within a coordinate plane by a set of Weibull functions constrained to share a common exponent b, but with the threshold parameter a_r allowed to vary across rays. 95% confidence limits for a_r were determined via 1000-sample bootstraps. Sensitivity is defined as 1/threshold, with corresponding confidence limits. On some rays, performance was close to chance, and the upper confidence limit of these bootstraps was large (e.g., $>10^5$); in these cases, sensitivity was taken to be zero.

Averages across subjects of sensitivities or thresholds are computed as the geometric means, and statistics (standard deviations, t-tests) are computed on the logarithms of the raw values.

2.5.2. Determination of parameters of the quadratic sensitivity model

The second step in the analysis was the construction of a model that incorporates all of the measured thresholds, and predicts thresholds in out-of-sample directions (i.e., in directions that were not contained in the coordinate planes). Model parameters were determined separately for each of the four subjects in which we obtained threshold measurements in all 15 coordinate planes.

As in (Victor, Thengone, & Conte, 2013), we used a quadratic cue combination rule (Macadam, 1942; Poirson et al., 1990; Saarela & Landy, 2012), since the isodiscrimination contours within each plane were approximately elliptical (see Fig. 4, and see (Victor, Thengone, & Conte, 2013) for further discussion for the rationale of the quadratic model). Specifically, we postulated that texture segregation was based on a decision variable:

$$V(\vec{c}) = \sum_{i,j} Q_{i,j} c_i c_j, \tag{8}$$

where the c_i's are the values of the 10 coordinates $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$, and that threshold is reached when $V(\vec{c}) = 1$. The parameters $Q_{i,j}$, a symmetric matrix, specify the model, and indicate how the image statistics combine and interact.

We determined the parameters $Q_{i,j}$ by adjusting them so that $V(\vec{c})$ was as close as possible to 1 at the thresholds a_r measured along each ray r. To apply Eq. (8) for this purpose, we write $\vec{c}(a_r, r) = (c_1(a_r, r), \ldots, c_{10}(a_r, r))$, a vector consisting of the values of the texture coordinates when threshold is reached along the ray r. This vector typically contains only two nonzero values (the specifying coordinates within the plane), but in some planes, other coordinates also have small nonzero values (see (Victor & Conte, 2012) and its Table 2), reflecting the geometry of the image statistics. The criterion that the decision variables $V(\vec{c}(a_r, r))$ are as close as possible to 1 was formalized by minimizing:

$$\begin{aligned} F &= \sum_r \left(V(\vec{c}(a_r, r)) - 1\right)^2 \\ &= \sum_r \left(\left(\sum_{i,j} Q_{i,j} c_i(r, a_r) c_j(r, a_r) \right) - 1 \right)^2, \end{aligned} \tag{9}$$

where the second equality follows from Eq. (8), and the sum is over all rays r. Note that $F = 0$ only if the threshold a_r along each ray r is exactly predicted by the quadratic model. Since F is a quadratic function of the $Q_{i,j}$ and is bounded below by 0, a unique minimum is guaranteed. We found this minimum by solving the linear system of equations $\frac{\partial F}{\partial Q_{i,j}} = 0$.

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

Table 2
Summary statistics for the fit of the phenomenological model for individual subjects (MC, DT, DF, JD), and the average across subjects (AVG). Mean deviation: average Euclidean distance between the measured threshold and the model (positive values: measured threshold higher than model); units are image statistic coordinates (range, −1 to 1). RMS deviation: root-mean-squared deviation between measured and model thresholds, in image-statistic units. RMS *z*-score: root-mean-squared deviation between measured and model thresholds as a *z*-score, i.e., the ratio of the threshold difference to the standard error of measurement of the threshold. For each pair of columns, statistics are computed for all thresholds considered individually ("All"), and for the average of thresholds in opposite direction ("Symmetrized"). "On-axis": measurements on the coordinate axes; "Off-axis": values in oblique directions in the coordinate planes; "All": on-axis and off-axis combined. For the few directions in which model thresholds were >1 (beyond the range that could be determined experimentally), measured values were set to 1 for computation of statistics.

Subject	Direction	Mean deviation		RMS deviation		RMS *z*-score	
	Type	All	Symmetrized	All	Symmetrized	All	Symmetrized
MC	On-axis	−0.0160	−0.0122	0.0332	0.0142	2.36	1.12
	Off-axis	−0.0040	0.0037	0.0608	0.0330	3.42	2.34
	All	−0.0096	−0.0037	0.0485	0.0261	2.84	1.93
DT	On-axis	−0.0246	−0.0131	0.0613	0.0224	2.49	1.03
	Off-axis	−0.0109	0.0039	0.0948	0.0334	3.48	1.73
	All	−0.0162	−0.0030	0.0799	0.0277	3.00	1.49
DF	On-axis	−0.0236	−0.0131	0.0532	0.0321	2.06	1.47
	Off-axis	−0.0199	−0.0066	0.0951	0.0535	4.50	2.70
	All	−0.0206	−0.0073	0.0789	0.0430	3.58	2.16
JD	On-axis	−0.0274	−0.0120	0.0765	0.0238	1.97	0.62
	Off-axis	−0.0236	−0.0041	0.1149	0.0567	2.78	1.80
	All	−0.0226	−0.0059	0.0955	0.0434	2.39	1.41
AVG	On-axis	−0.0229	−0.0126	0.0561	0.0231	2.22	1.06
	Off-axis	−0.0146	−0.0008	0.0914	0.0442	3.55	2.14
	All	−0.0173	−0.0050	0.0757	0.0351	2.96	1.75

For the rays r that were along a coordinate axis, replicate measurements of thresholds a_r were available from each plane that contained this axis; these values were pooled (by averaging) across planes prior to use in Eq. (9). Similarly, thresholds were pooled across coordinates that were related by rotational symmetry (the two cardinal β's, the two diagonal β's, and the four θ's), since there were no significant differences between these thresholds (Victor, Thengone, & Conte, 2013). In total, there were 20 unique free parameters $Q_{i,j}$ to be determined (five unique values of $Q_{i,j}$ corresponding to γ, the cardinal β's, the diagonal β's, θ, and α, and 15 unique values for the 15 coordinate planes). There were 78 unique thresholds available to constrain this fit (10 values in the positive and negative directions on the 5 kinds of coordinate axes, and 68 unique oblique directions in the coordinate planes).

The parameters $Q_{i,j}$ determined by minimizing Eq. (9) provide a prediction of the threshold along any ray specified by an arbitrary unit vector $\vec{u} = (u_1, \ldots, u_{10})$: the predicted threshold is the value a for which $V(a\vec{u}) = 1$, i.e., the value of a for which:

$$a^2 \sum_{i,j} Q_{i,j} u_i u_j = 1. \tag{10}$$

Confidence limits on the $Q_{i,j}$ and on quantities derived from them (e.g., eigenvalues of the matrix $Q_{i,j}$, thresholds in mixture directions predicted by Eq. (10), and dot-products between eigenvectors of the Q-matrices for different subjects) were determined by a parametric bootstrap with 100 resamplings. Each bootstrap consisted of repeating the above determination of $Q_{i,j}$ using threshold values a drawn according to the distribution found along each ray in the 1000-sample bootstrap procedure described following Eq. (7). Specifically, for each ray, we determined the Gaussian distribution that matched the mean and standard deviation of the sensitivity (=1/threshold) values, drew randomly from this Gaussian, and set the threshold to 1/sensitivity. We worked in terms of sensitivities rather than directly in terms of thresholds to avoid outlier effects due to large upper confidence limits for some thresholds. Confidence limits for the $Q_{i,j}$ and quantities derived from them were then set at the 2.5% and 97.5% quantiles (interpolated by Matlab's quantile.m) of these 100 resamplings.

2.5.3. Eigenvector classes

The matrix Q is symmetric (i.e., $Q_{i,j} = Q_{j,i}$), and therefore, is fully characterized by its eigenvalues and eigenvectors. Because we assume that coordinate axes related by rotational symmetry are equivalent (see empirical evidence for this in (Victor, Thengone, & Conte, 2013) and below), there is a further induced symmetry on Q. We use this symmetry to classify its eigenvectors into subsets, and distinguish the quantities that are determined from the data, from those that are constrained by this assumed symmetry. (This classification is a natural one that emerges from standard group-theoretic procedure: decomposing the 10-dimensional space of image statistics according to the irreducible representations of the symmetry group of 90° rotations and reflections in the plane. For background on the theory of irreducible group representations, see for example (Serre, 1977)).

This machinery leads to the following decomposition of the 10 eigenvectors, and several guarantees about their symmetry and coordinates. First, there is a subset of 5 eigenvectors, each of which is symmetric with respect to 90° rotation (that is, these eigenvectors are linear combinations of the image statistics whose values are unchanged if the image is rotated by 90°). We designate these as sym1 through sym5, in descending order of their eigenvalues. In this subset and the subspace that they span, γ and α are unconstrained, but symmetry requires that, $\beta_{\_} = \beta_{|}$, $\beta_{\backslash} = \beta_{/}$, and $\theta_{⌟} = \theta_{⌞} = \theta_{⌜} = \theta_{⌝}$. Next, there is a pair of eigenvectors that span a subspace with a somewhat surprising property: for textures in this subspace, image statistic values are replaced by their negative if the texture is replaced by its horizontal or vertical mirror-image. We designate these eigenvectors as hvi1 and hvi2, also in descending order of their eigenvalues. In this subspace, symmetry forces, $\gamma = \beta_{\_} = \beta_{|} = \alpha = 0$, $\beta_{\backslash} = -\beta_{/}$, and $\theta_{⌟} = -\theta_{⌞} = \theta_{⌜} = -\theta_{⌝}$. The final three eigenvectors are constrained by symmetry to lie in specific coordinate directions or coordinate planes. They are $\beta_{\_} = -\beta_{|}$ (other coordinates zero), $\theta_{⌟} = -\theta_{⌜}$ (other coordinates zero), and $\theta_{⌞} = -\theta_{⌝}$ (other coordinates zero). The first we designate dii because in this subspace, image statistic values are replaced by their negative if the texture is replaced by its diagonal mirror image. The latter pair we designate rot_A and rot_B because they are rotations of each other. Symmetry forces rot_A and rot_B to

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

have the same eigenvalue. These final three eigenvectors are all within coordinate planes, and were therefore not used as out-of-sample tests of the model.

2.6. Previously reported work

A summary of findings for 4 subjects in 11 of the 15 planes at standard check sizes is included in a paper that compares these sensitivities to the statistics of natural images (Hermundstad et al., 2014). A partially overlapping portion of 6 subjects' data in 8 of those planes was also previously reported (Victor, Thengone, & Conte, 2013).

3. Results

3.1. Overview

Our immediate goal is to measure visual sensitivity to a set of image statistics chosen to capture the basic features of contrast, edge, and corner, and their interactions. As described in Methods and (Victor & Conte, 2012), these statistics parameterize a 10-dimensional space of black-and-white textures constructed on a square lattice. We begin with sensitivities to individual image statistics and then consider their pairwise interactions. Next, we use these data to constrain a model for sensitivities to combinations of multiple image statistics, and we then test the model with out-of-sample stimuli that contain such combinations. Finally, we show how the above sensitivities, determined for a check size of 14 min and a presentation time of 120 ms, change as a function of these parameters.

3.2. Individual image statistics

Fig. 2 shows representative psychometric functions for each kind of image statistic in the 4-AFC segmentation task. For all subjects, the thresholds follow the same ranking: threshold is lowest for the fraction of white vs. black checks (γ), next-lowest for two-point correlations in the cardinal direction ($\beta_{\_}$) and the diagonal direction ($\beta_{\backslash}$), next-lowest for four-point correlations, (α), and then finally highest for three-point correlations ($\theta_{\lrcorner}$). As mentioned in the Introduction, the stimuli are calibrated so that for an ideal observer, sensitivities along each axis are identical. Thus, the dependence of visual thresholds on the type of image statistic is a direct measure of the selectivity of visual processing. This ratio (threshold for $\theta_{\lrcorner}$ vs. threshold for γ) is approximately 4.5:1. Note that thresholds do not simply covary with the order of the statistic: the threshold for the third-order statistic, $\theta_{\lrcorner}$, is higher than for the fourth-order statistic, α. In all cases, the thresholds for positive values of the image statistic were similar to the thresholds for negative values. All of the psychophysical functions had approximately the same shape (corresponding to a Weibull exponent typically in the range 2.2–2.7), so the comparisons of relative thresholds are independent of the choice of threshold criterion. Subjects included an individual with extensive viewing experience (MC: thousands of hours, most with similar stimuli and tasks) to individuals for whom this was their first psychophysical experiment. While there is some variability in overall thresholds across subjects, this variability was on the order of 20%, and did not appear to depend on viewing experience.

The examples of Fig. 2 typify the results across all 16 subjects (Fig. 3). The lowest thresholds are for γ (mean 0.166, SD ± 11%); highest thresholds are for $\theta_{\lrcorner}$ (mean 0.793, SD ± 16%). Consistency across subjects is high: all had the same rank-order of thresholds ($\gamma < \beta_{\_} < \beta_{\backslash} < \theta_{\lrcorner} < \alpha$), and the scatter of thresholds for each kind of image statistic ranges from SD ±9.5% for $\beta_{\backslash}$ to ±18% to α). This intersubject variability is largely a matter of overall sensitivity differences. With intersubject variability of overall sensitivity removed, by adjusting each subject's set of thresholds up or down by factors ranging from 0.81 to 1.24, the scatter of thresholds for each kind of image statistic is 10% or less. The variability between subjects was not a matter of familiarity with the purpose of the experiment (threshold ratios for the 4 non-naïve observers compared to the 12 naïve observers ranged from 1.13 to 0.93 across the five types of statistics; no ratios were significantly different from 1, two-tailed un-paired t-tests $p = 0.06$ to $p = 0.97$).

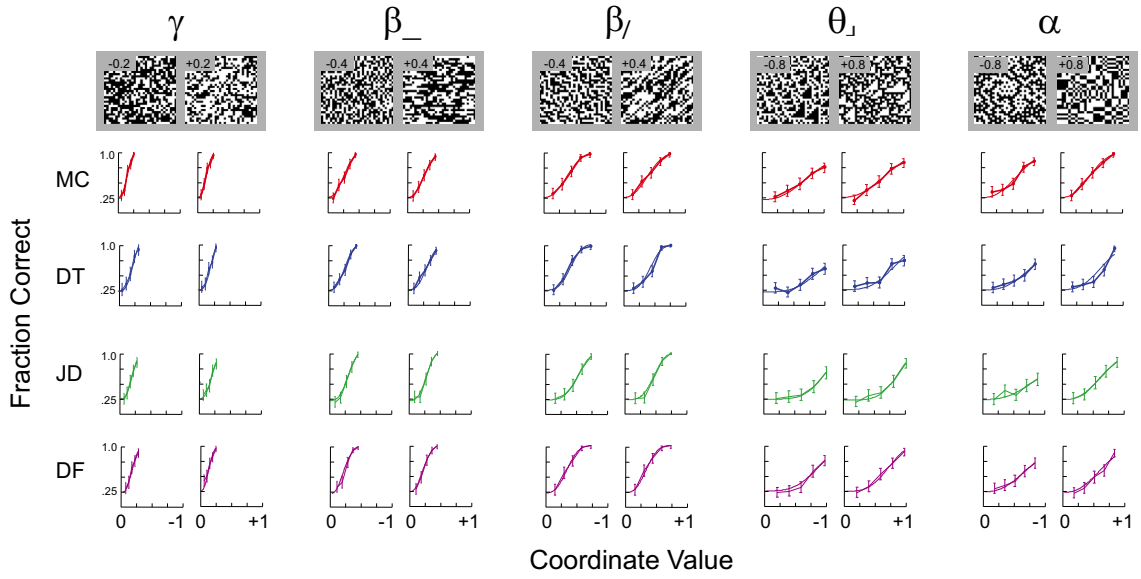

Fig. 2. Psychometric functions along five kinds of coordinates: first-order (γ), cardinal second-order ($\beta_{\_}$), diagonal second-order ($\beta_{/}$), third-order ($\theta_{\lrcorner}$), and fourth-order (α). For each coordinate, psychometric functions are shown for negative excursions (left element of each pair) and positive excursions (right element of each pair). Chance performance is 0.25; error bars indicate 95% confidence limits based on binomial statistics. The patches above the psychometric functions show typical 32 × 32 samples of images along each coordinate axis, constructed with $\gamma = \pm 0.2$, $\beta_{\_} = \pm 0.4$, $\beta_{/} = \pm 0.4$, $\theta_{\lrcorner} = \pm 0.8$, and $\alpha = \pm 0.8$. Subjects MC, DT, JD, DF. Data for subjects MC and DT for $\beta_{\_}$, $\beta_{/}$, and $\theta_{\lrcorner}$ reproduced, with permission of the copyright holder, The Association for Research in Vision and Ophthalmology, from (Victor, Thengone, & Conte, 2013).

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

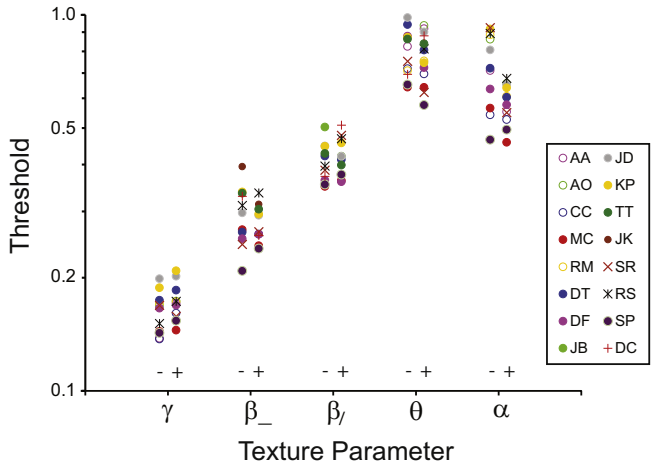

Fig. 3. Thresholds for positive and negative excursions for the five kinds of coordinates shown in Fig. 2, for $N = 16$ subjects. All subjects had the same rank-order of thresholds. $\gamma < \beta_{\_} < \beta_{\backslash} < \alpha < \theta$.

The consistency across subjects allows for identification of some subtle asymmetries. Thresholds for detecting deviations of first-order statistics were larger in the positive (mean 0.172, SD ± 11%) direction than in the negative (mean 0.161, SD ± 13%) direction (two-tailed paired t-test: $p = 0.017$, $N = 12$), meaning that an increase in the fraction of black checks was easier to detect than an increase in the fraction of white checks – but this difference in sensitivities was a small effect (<7%). Thresholds for detecting deviations of the four-point correlation α was smaller in the positive (mean 0.579, SD ± 13%) direction than in the negative (mean 0.726, SD ± 13%) direction ($p = 0.001$, $N = 12$); these sensitivities differed by approximately 25%. There were no asymmetries ($p = 0.08$ to $p = 0.36$, $N = 11$ to $N = 14$) between detection of positive vs. negative two-point or three-point correlations ($\beta_{\_}$, $\beta_{\backslash}$, or $\theta_{\lrcorner}$. Across subjects, we also did not detect any asymmetries in thresholds for image statistics that differed by a rotation ($\beta_{\_}$ vs. $\beta_{|}$, $\beta_{\backslash}$ vs. $\beta_{/}$, $\theta_{\lrcorner}$ vs. $\theta_{\llcorner}$ vs. $\theta_{\urcorner}$: threshold ratios were within 10% of each other, and not significantly different ($p = 0.06$ to $p = 0.98$, $N = 6$ to $N = 14$).

3.3. Pairwise interactions of image statistics

We next consider pairwise interactions of the image statistics. Since there are 10 independent image statistics, $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$, there are $45 = 10 \times 9/2$ pairwise combinations. To simplify the task of exploring these interactions, we exploit the above observation that thresholds for individual

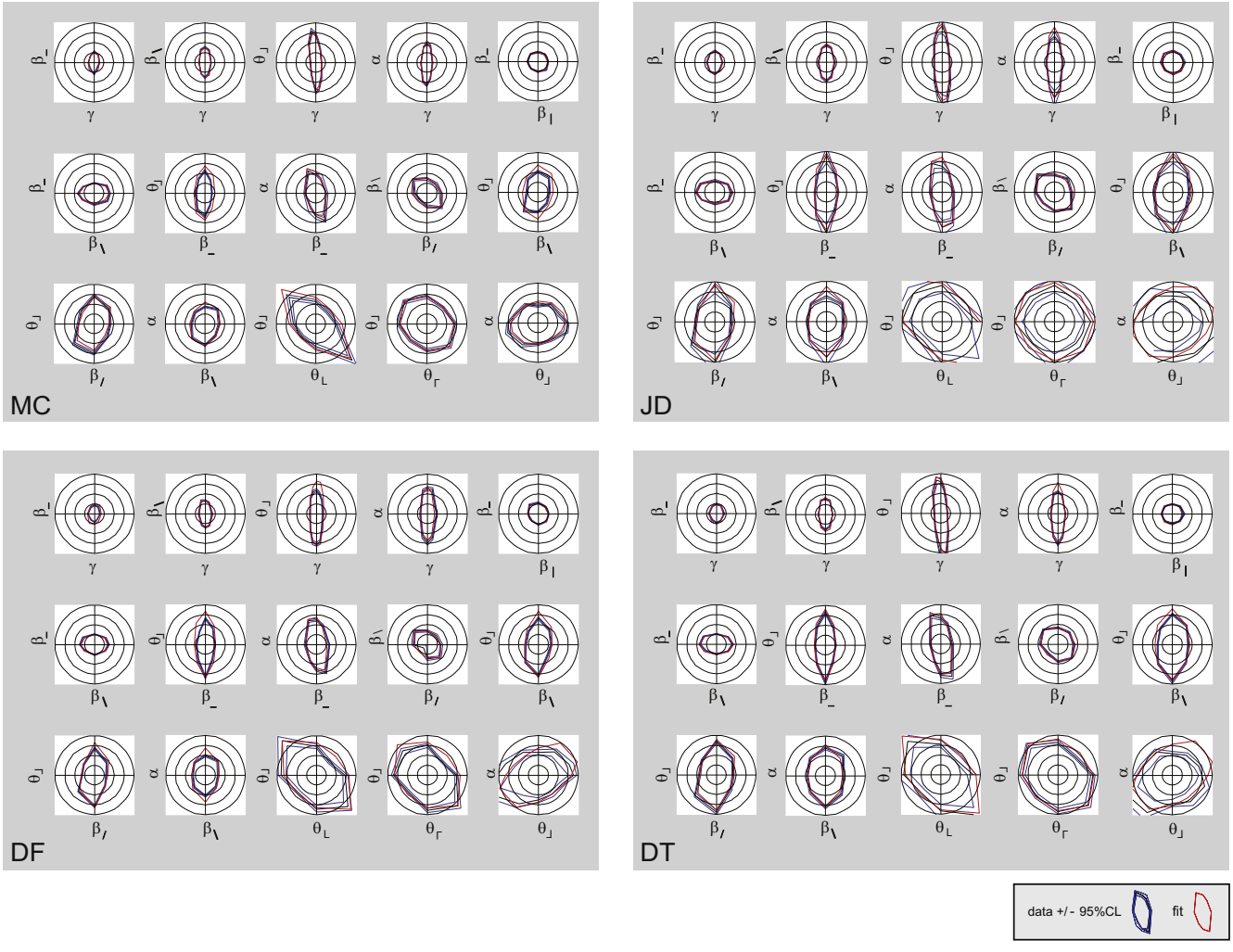

Fig. 4. Isodiscrimination contours in the 15 coordinate planes. Black: locus corresponding to a fraction correct of 0.625, halfway between chance and perfect. Blue: 95% confidence limits via bootstrap. Red: isodiscrimination contours for the ellipsoidal model fitted to each subject. The 15 coordinate planes include, up to symmetry, all pairwise combinations of the 10 coordinates $\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\}$. Subjects MC, JD, DF, DT. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

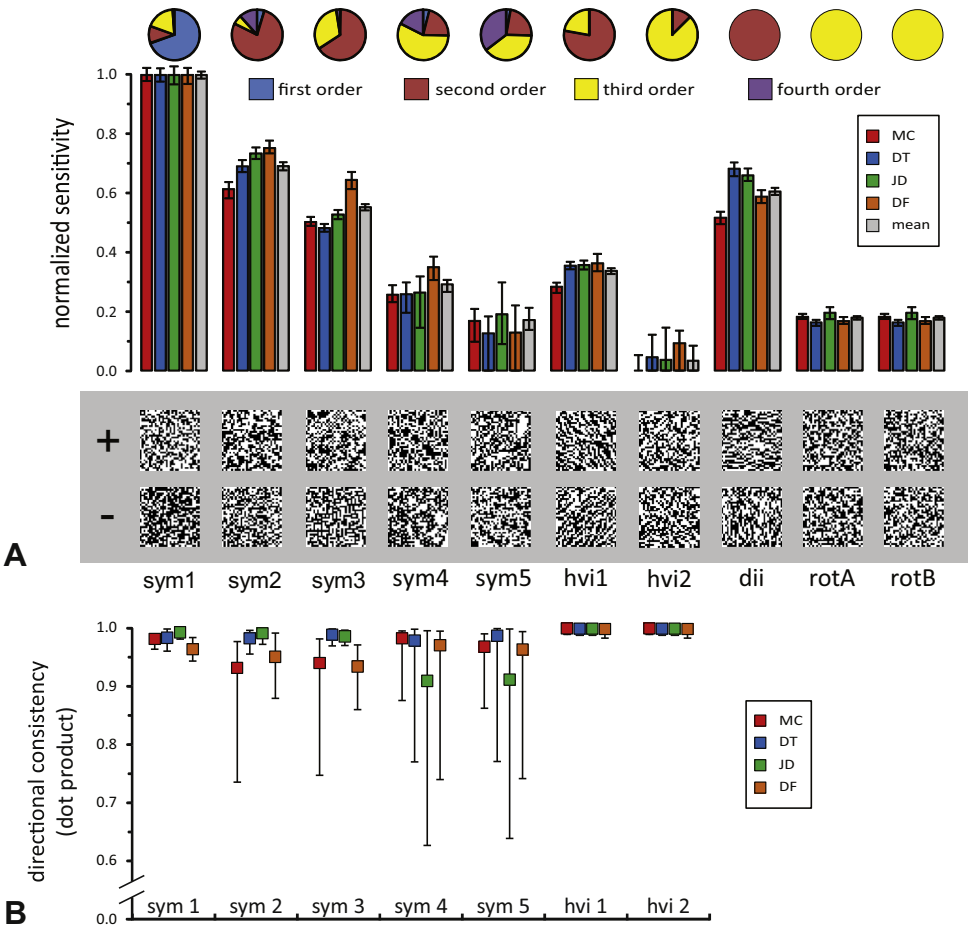

Fig. 5. (A) Eigenvalues of the sensitivity matrix Q for each subject of Fig. 4 (colored bars) and the average across these subjects (gray bar); each eigenvalue is normalized by dividing by the subject's largest eigenvalue (sym1). Error bars indicate 95% confidence limits, via parametric bootstrap of the sensitivity measurements. The pie graphs show the fractional contribution of each order of statistic to the eigenvectors. Eigenvectors are named according to symmetry class (see Eigenvector Classes in Methods), and, within each symmetry class (sym, hvi, dii, and rot), are labeled in order of decreasing eigenvalue. Texture samples are the eigenvectors of the averaged Q-matrix at distances of 0.18 (sym1 and sym2) and 0.36 (other eigenvectors) from the origin. (B) Dot-products of the eigenvectors of the Q-matrices for each of the four subjects of Fig. 4, with the eigenvectors of the average Q-matrix of the other three subjects. Dot-products are not shown for the last three eigenvectors, since symmetry considerations force these values to 1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

image statistics are invariant under rotation. Based on this, we assume that their pairwise interactions are invariant under rotation as well. As examples, we assume that the interaction involving $(\gamma, \beta_{\_})$ is identical to the interaction involving $(\gamma, \beta_{|})$ since these combinations differ by a 90° rotation; that the interaction involving $(\theta_{\lrcorner}, \theta_{\llcorner})$ is identical to the interaction involving $(\theta_{\ulcorner}, \theta_{\urcorner})$ since these combinations differ by a 180° rotation; and that the interaction involving $(\beta_{\diagdown}, \alpha)$ is identical to the interaction involving $(\beta_{\diagup}, \alpha)$, since these combinations differ by a 90° rotation. These assumptions reduce the number of distinct pairwise interactions from 45 to 15 (see Methods and Table 2 of (Victor & Conte, 2012) for details).

In four subjects, we measured these pairwise interactions, by determining the isodiscrimination contours in the corresponding 15 coordinate planes (Fig. 4). Isodiscrimination contours were typically symmetric about the origin: the threshold along a ray in one direction was approximately equal to the threshold along a ray in the opposite direction. This extends the above observation for thresholds along the coordinate axes to thresholds in the oblique directions within coordinate planes. We tested a total of 34 oblique orientations (2 oblique orientations in 13 of the planes, 4 oblique orientations in the $(\gamma, \theta_{\lrcorner})$ and $(\theta_{\lrcorner}, \alpha)$ planes); 18 of these 34 included either γ or α as one of their components, which, as noted above, manifested modest asymmetries. Considering the thresholds along opposite directions in each of these 34 oblique orientations, we only found statistically significant differences in threshold in three of them ($p < 0.05$ by paired t-test, false-discovery-rate corrected (Benjamini & Hochberg, 2001)) across subjects: the same-sign direction in the $(\beta_{\diagup}, \theta_{\lrcorner})$ and $(\beta_{\diagdown}, \beta_{\diagup})$-planes, and the opposite-sign direction in the $(\beta_{\_}, \alpha)$-plane. As noted previously (Victor, Thengone, & Conte, 2013), the $(\beta_{\diagdown}, \beta_{\diagup})$-asymmetry (the fourth plane in the second row of each of the datasets of Fig. 4) was large: across subjects, thresholds when both parameters were positive was 1.47 times (geometric mean) higher than when both parameters were negative. All other asymmetries were <25%, and most (18/34) were <10%.

The isodiscrimination contours had an approximately elliptical shape, suggesting that the image statistics combine via a quadratic combination rule. However, the axes of the isodiscrimination contours were not necessarily aligned with the coordinate axes. A tilted isodiscrimination contour means that the direction of lowest threshold (the direction from the origin to the closest approach of the contour) corresponds to a combination of the statistics, rather than either one in isolation. It also means that the interaction between two image statistics depends on their relative sign – facilitating detection for one sign combination, and hindering it for the other. These tilts can be subtle (for example, $(\gamma, \theta_{\lrcorner})$ – the third plane in the first row of each dataset of Fig. 4) or large (for example,

$(\theta_\lrcorner, \theta_\llcorner)$ – the third plane in the third row of each dataset of Fig. 4). To formalize this observation without postulating a specific shape for the isodiscrimination contours, we compared thresholds for same-sign vs.opposite-sign combinations of image statistics. This showed that of the 15 planes, eight had contours with a significant ($p < 0.05$) tilt with respect to the axes (two-tailed paired t-test across subjects, with false discovery rate correction.) In six planes, thresholds were lower for combinations of statistics that are of the same sign, than for combinations that are of opposite sign, corresponding to a counterclockwise tilt: $(\gamma, \theta_\lrcorner)$, $(\beta_\_, \beta_|)$, $(\beta_\_, \alpha)$, $(\beta_\diagdown, \beta_\diagup)$, $(\theta_\lrcorner, \theta_\llcorner)$, and $(\theta_\lrcorner, \theta_\ulcorner)$. In two planes, thresholds were higher for same-sign combinations than for opposite-sign combinations, corresponding to a clockwise tilt: $(\beta_\diagdown, \theta_\lrcorner)$ and $(\beta_\diagup, \theta_\lrcorner)$.

3.4. A phenomenological model for interaction of image statistics

The results above indicate that isodiscrimination contours in the coordinate planes had approximately elliptical shapes that were nearly symmetric with respect to the origin (i.e., that thresholds were similar along opposite rays), but were often tilted with respect to the coordinate axes. Based on these observations, we framed a phenomenological model for how the image statistics interact to determine the perceptual threshold.

The basic idea is that the isodiscrimination contours form an ellipse in each coordinate plane, and that these ellipses, taken together, determine an ellipsoidal isodiscrimination contour in the entire 10-dimensional space. In any single plane (e.g., a plane corresponding to coordinates c_u and c_v, where c_u and c_v are chosen from the 10 coordinates $\{\gamma, \beta_\_, \beta_|, \beta_\diagdown, \beta_\diagup, \theta_\lrcorner, \theta_\llcorner, \theta_\ulcorner, \theta_\urcorner, \alpha\}$), an elliptical isodiscrimination contour can be described by the locus of points (c_u, c_v) where:

$$Q_{u,u}c_u^2 + 2Q_{u,v}c_u c_v + Q_{v,v}c_v^2 = 1. \tag{11}$$

The parameters Q describe the size and shape of the ellipse: $1/\sqrt{Q_{u,u}}$ and $1/\sqrt{Q_{v,v}}$ are the thresholds for c_u and c_v, since $(\pm 1/\sqrt{Q_{u,u}}, 0)$ and $(0, \pm 1/\sqrt{Q_{v,v}})$ are the points where the isodiscrimination contour (11) intersects the axes. The third parameter, $Q_{u,v}$, describes their interaction. We can then use a single equation, generalizing Eq. (11), to represent these contours in all coordinate planes:

$$\sum_{i,j} Q_{i,j}c_i c_j = 1. \tag{12}$$

Note that Eq. (12) simplifies to Eq. (11) in any plane (i.e., if only two of the c_k are nonzero). In Eq. (12), $Q_{i,j}$ is a symmetric matrix, with $1/\sqrt{Q_{i,i}}$ corresponding to the threshold along axis c_i and $Q_{i,j} = Q_{j,i}$ describing the interaction of c_i and c_j. Geometrically, Eq. (12) describes the unique 10-dimensional ellipsoid whose intersection with each of the coordinate planes yields the ellipses of Eq. (11). Q, a 10×10 symmetric matrix, has 20 independent parameters, since the symmetry considerations reduce the number of unique diagonal elements to 5 (sensitivities for γ, the cardinal β's, the diagonal β's, the θ's, and α), and the number of unique off-diagonal elements to 15 (one for each coordinate plane of Fig. 4). There are 78 unique measured thresholds: the 10 on-axis thresholds, and 68 off-axis thresholds.

Fig. 4 shows the fit of this model in each of the planes, and Table 2 summarizes the statistics of the fit across subjects. Overall, the ellipsoidal shape provides a reasonable fit: the root-mean-squared (RMS) error in thresholds is 0.076 across subjects (Table 2 bottom row, RMS deviation "all" column). The main source of model error is that there are modest differences in thresholds for positive and negative values of the image statistics (as noted above, a 25% lower threshold for $\alpha > 0$ than for $\alpha < 0$, smaller asymmetries for the other parameters). No quadratic model can account for such asymmetries. When the thresholds in opposite directions are replaced by their averaged values – so the effect of this asymmetry is removed – the RMS error in the model fit is 0.035 (Table 2 bottom row, RMS deviation "symmetrized" column).

Though this error is small, we note that it is more than can be accounted for from errors in the experimental measure of threshold. This is quantified via a z-score: the ratio of model error in each direction to the measurement error (1 s.d. of the bootstrapped distribution of fitted thresholds). Root-mean-squared z-scores were 2.95 for all thresholds, and 1.75 for the symmetrized thresholds, indicating that the model error is between 2 and 3-fold higher than could be accounted for by the uncertainty of the psychophysical measurements.

Most of this excess error can be attributed to the off-diagonal thresholds, i.e., to the model's prediction of pairwise interactions. As seen in Table 2, the RMS z-score for the symmetrized thresholds was 1.06 on-axis, indicating that the model error was only marginally larger than the uncertainty of psychophysical thresholds, but the z-score was 2.14 off-axis, i.e., for the pairwise interactions. Overall, the phenomenological model slightly overestimated on-axis thresholds, compared to off-axis thresholds (-0.023 compared to -0.015 considering all thresholds, -0.013 compared to -0.001 for the symmetrized thresholds).

In sum, while there are detectable systematic deviations of the phenomenological model from an ideal fit (there are asymmetries of thresholds in opposite directions that account for model errors of approximately 0.04 and there is a deviation from the closest-fitting ellipsoidal shape that accounts for errors of 0.01–0.02), the model provides a reasonable summary of the shape of the isodiscrimination surfaces. We therefore use its parameters, the sensitivity matrix Q, to compare the shape of the isodiscrimination surfaces across subjects, and, in the next section, to predict sensitivities for combinations of several image statistics.

Since Q is a symmetric matrix, we can use its eigenvalues and eigenvectors to compare it across subjects. Fig. 5A shows the eigenvalues of the matrix Q determined from each of the four subjects, and the mean across subjects. Interestingly, the eigenvalue corresponding to the eigenvector hvi2 is experimentally indistinguishable from 0 (its confidence limits include 0 for all subjects and the cross-subject average). This means that the phenomenological model predicts that perceptual sensitivity is zero for this combination of image statistics. Thus, the model predicts a combination of image statistics that is metameric to random – even though each of the image statistics, individually, is perceptually salient. This prediction will be tested below.

Fig. 5B compares the directions of the eigenvectors across subjects. All dot-products are close to 1, indicating that across the subjects, the principal axes of the ellipsoid (i.e., the directions of greatest and least sensitivities) are consistent.

3.5. Prediction of sensitivities for combinations of multiple image statistics

We next asked whether the phenomenological model, which was constructed to capture pairwise interactions of image statistics, could also account for more complex combinations. We tested this with two kinds of out-of-sample measurements: first, with combinations of statistics that were predicted to yield highest and lowest sensitivities, and second, combinations of statistics that might have functional importance, as they are associated with material properties. The first set of combinations consists of the eigenvectors of the matrix Q determined above, and are listed in Table 3. One of these directions, hvi2, is of particular interest, since the predicted sensitivity in this direction is zero (Fig. 5). For the

Table 3
Empirical eigenvectors of the ellipsoidal model of sensitivity data for individual subjects (MC, DT, DF, JD), and the average across subjects (AVG). Symmetries of the model place constraints on the 10 eigenvectors, leaving only 7 that are free to vary (see "Eigenvector classes" in Methods for further details). These consist of (i) five eigenvectors that are symmetric with respect to 90° rotation, and therefore have $\beta_{-} = \beta_{|}$, $\beta_{\backslash} = \beta_{/}$, $\theta_{\lrcorner} = \theta_{\llcorner} = \theta_{\ulcorner} = \theta_{\urcorner}$, designated sym1 through sym5, in order of decreasing eigenvalue; and (ii); two eigenvectors that invert when the image is mirrored in the horizontal or vertical plane ($\beta_{\backslash} = -\beta_{/}$, and $\theta_{\lrcorner} = -\theta_{\llcorner} = \theta_{\ulcorner} = -\theta_{\urcorner}$), with remaining coordinates zero, designated hvi1 and hvi2, in order of decreasing eigenvalues. The final three eigenvectors are determined by symmetry, and are not shown. All eigenvectors are normalized to Euclidean length 1.

Eigenvector	Subject	γ	β_{-}	$\beta_{\|}$	$\beta_{\backslash}$	$\beta_{/}$	$\theta_{\lrcorner}$	$\theta_{\llcorner}$	$\theta_{\ulcorner}$	$\theta_{\urcorner}$	α
sym1	MC	0.971	−0.090	−0.090	0.001	0.001	0.100	0.100	0.100	0.100	−0.043
	DT	0.983	−0.001	−0.001	0.002	0.002	0.092	0.092	0.092	0.092	0.001
	DF	0.979	−0.066	−0.066	0.126	0.126	−0.011	−0.011	−0.011	−0.011	−0.010
	JD	0.996	−0.021	−0.021	−0.013	−0.013	0.040	0.040	0.040	0.040	0.011
	AVG	0.988	−0.057	−0.057	0.022	0.022	0.063	0.063	0.063	0.063	−0.017
sym2	MC	0.122	0.661	0.661	−0.091	−0.091	0.032	0.032	0.032	0.032	0.303
	DT	0.012	0.683	0.683	0.058	0.058	−0.029	−0.029	−0.029	−0.029	0.239
	DF	0.055	0.681	0.681	0.135	0.135	−0.052	−0.052	−0.052	−0.052	0.150
	JD	0.036	0.692	0.692	0.062	0.062	−0.042	−0.042	−0.042	−0.042	0.157
	AVG	0.088	0.686	0.686	0.044	0.044	−0.026	−0.026	−0.026	−0.026	0.214
sym3	MC	0.101	0.105	0.105	0.633	0.633	−0.203	−0.203	−0.203	−0.203	−0.038
	DT	0.049	−0.062	−0.062	0.675	0.675	−0.138	−0.138	−0.138	−0.138	−0.039
	DF	−0.194	−0.136	−0.136	0.650	0.650	−0.138	−0.138	−0.138	−0.138	−0.048
	JD	0.036	−0.071	−0.071	0.680	0.680	−0.122	−0.122	−0.122	−0.122	−0.054
	AVG	0.008	−0.047	−0.047	0.663	0.663	−0.168	−0.168	−0.168	−0.168	−0.055
sym4	MC	−0.158	0.124	0.124	0.223	0.223	0.384	0.384	0.384	0.384	−0.506
	DT	−0.155	0.102	0.102	0.173	0.173	0.413	0.413	0.413	0.413	−0.463
	DF	−0.017	0.090	0.090	0.165	0.165	0.397	0.397	0.397	0.397	−0.548
	JD	−0.056	0.095	0.095	0.140	0.140	0.416	0.416	0.416	0.416	−0.496
	AVG	−0.110	0.100	0.100	0.195	0.195	0.407	0.407	0.407	0.407	−0.478
sym5	MC	−0.090	−0.173	−0.173	0.205	0.205	0.225	0.225	0.225	0.225	0.804
	DT	−0.088	−0.137	−0.137	0.108	0.108	0.225	0.225	0.225	0.225	0.854
	DF	−0.022	−0.074	−0.074	0.125	0.125	0.266	0.266	0.266	0.266	0.822
	JD	−0.049	−0.077	−0.077	0.116	0.116	0.241	0.241	0.241	0.241	0.852
	AVG	−0.064	−0.121	−0.121	0.142	0.142	0.226	0.226	0.226	0.226	0.850
hvi1	MC				0.692	−0.692	0.104	−0.104	0.104	−0.104	
	DT				0.692	−0.692	0.103	−0.103	0.103	−0.103	
	DF				0.695	−0.695	0.092	−0.092	0.092	−0.092	
	JD				0.696	−0.696	0.087	−0.087	0.087	−0.087	
	AVG				0.693	−0.693	0.098	−0.098	0.098	−0.098	
hvi2	MC				−0.147	0.147	0.489	−0.489	0.489	−0.489	
	DT				−0.145	0.145	0.489	−0.489	0.489	−0.489	
	DF				−0.131	0.131	0.491	−0.491	0.491	−0.491	
	JD				−0.124	0.124	0.492	−0.492	0.492	−0.492	
	AVG				−0.138	0.138	0.490	−0.490	0.490	−0.490	

second set, we used the combinations of image statistics that extremize the surface porosity, i.e., the number of holes per unit area. The specific combinations of image statistics that determine porosity depend on whether each check is considered to be connected to its four nearest neighbors (the M4 combination) or its eight nearest neighbors (the M8 combination). We tested both of these and related combinations (Table 1), as detailed in Methods. For each combination, we determined thresholds in opposite directions from the origin of the space. Measurements were made in four subjects. For the two subjects, MC and DT, who participated in all experiments, we used their personalized eigenvectors (Table 3). For the other two subjects (SR and KP), for which we did not measure sensitivities in the coordinate planes, we used the eigenvectors determined from the average Q-matrix (AVG in Table 3).

Results are shown in Fig. 6. The overall sensitivities are well-captured by the phenomenological model. Of note, in the predicted metameric direction, hvi2, all of the subjects have a 95% confidence limit for threshold that extends beyond the boundary of the space – corresponding to a predicted sensitivity of zero. We also note that the predicted sensitivities are accurate in the Minkowski directions, known to tap topological properties of visual textures (Barbosa, Bubna-Litic, & Maddess, 2013). We calculated the statistics of the model predictions with no free parameters (upper half of Table 4), and with a single scale factor to match the on-axis thresholds measured in trials interleaved with the out-of-sample measurements (lower half of Table 4). Overall RMS prediction error was comparable, approximately 0.070. As noted above, the phenomenological model cannot account for asymmetries in thresholds for positive and negative variations along each of these test directions. This asymmetry accounted for approximately half of the model error (middle columns, bottom row of Table 4); the symmetrized thresholds were predicted with an RMS error of approximately 0.037. This is comparable to the in-sample RMS prediction error of the model, 0.035 (Table 2). The ratio by which model predictions exceeded the uncertainty of the psychophysical thresholds was also comparable for the out-of-sample directions (RMS z-score of 1.68, Table 4) and for the in-sample directions (RMS z-score of 1.75, Table 2). This ratio was comparable for the eigenvector directions and the Minkowski directions (RMS z-scores of 1.45 and 1.78).

Finally, we note that the eigenvector directions and the Minkowski directions represent very different tests of the model, as they probe diverse directions in the space. Considering the seven eigenvector directions and the five Minkowski directions, the 35 (= 7 × 5) dot products are exactly zero in 13 cases (guaranteed by symmetry), and the absolutes values of the other 22 dot products are approximately uniformly distributed between 0 and 1, with a median 0.38–0.44, depending on subject.

In sum, the ellipsoidal model constructed to fit the pairwise thresholds was able to predict sensitivities to out-of-sample combinations of multiple image statistics. The size and nature of the errors (Table 4) was comparable to the in-sample fits of the model, shown in Table 2. As was the case for the in-sample fits, about half

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

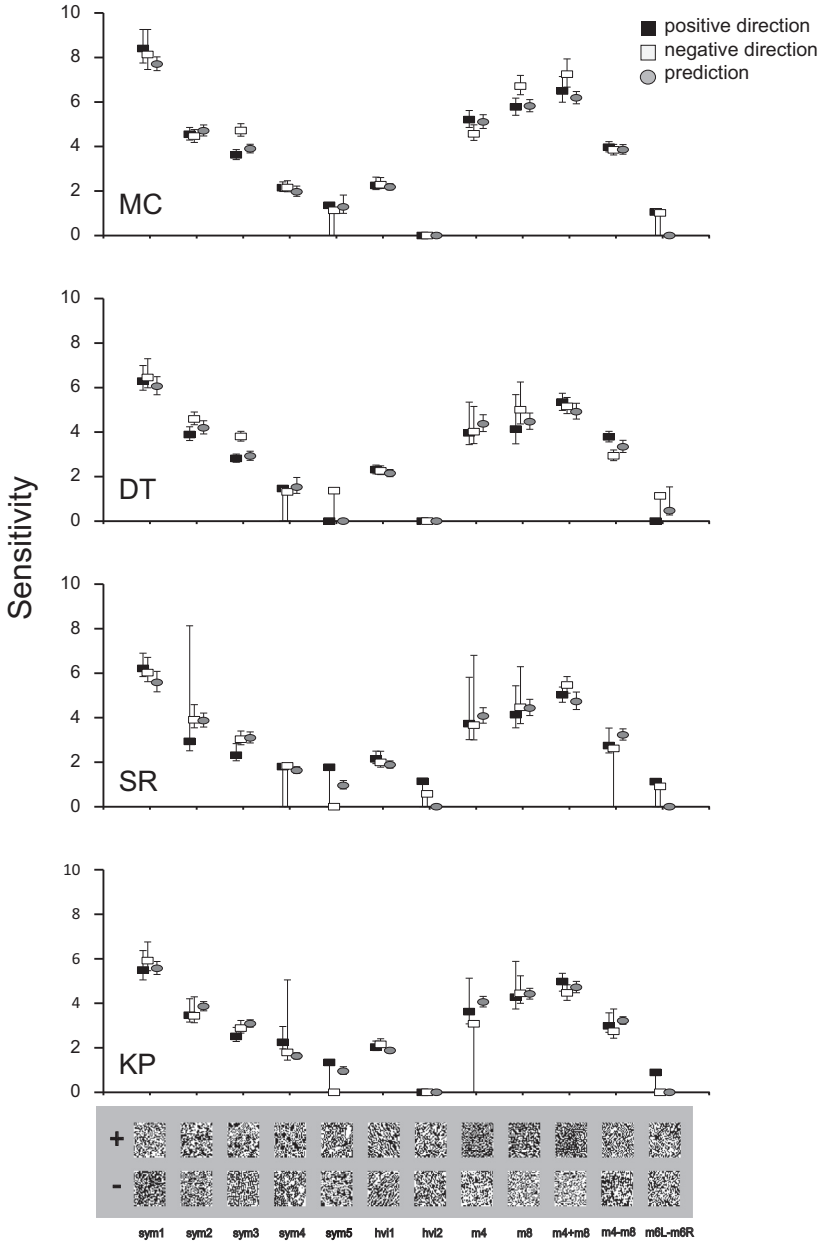

Fig. 6. Measured and predicted sensitivities along the eigenvectors (Table 3) and in the Minkowski directions (Table 1) for four subjects. Thresholds are measured for positive (filled symbols) and negative (open symbols) directions, where the sign is defined relative to the eigenvectors tabulated in Table 3 and the Minkowski directions in Table 1. The prediction of the ellipsoidal model is necessarily the same for both directions, and is shown in gray. Error bars show 95% confidence limits via bootstrap. Texture samples are at distances of 0.18 for sym1 and sym2, 0.36 for the other eigenvectors, and 0.3 for the Minkowski directions.

of the model error was due to its assumed symmetry, which precluded the model from predicting differences in thresholds along opposite directions. The remaining error was within a factor of two of the experimental error in determining the symmetrized thresholds (the RMS z-scores of Tables 2 and 4).

3.6. Scaling

So far we have considered stimuli constructed with a single check size, a 14-min square. We next examine how the perceptual sensitivities to individual statistics and their interactions depend on this scale. To keep the statistical demands of the task identical, we change the stimulus size in proportion to the check size, so that the number of checks in the target and background are unchanged.

Fig. 7 shows the results. For the statistics of order two or more (β_{-}, $\beta_{\backslash}$, $\theta_{\lrcorner}$ and α), there is a monotonic decrease in sensitivity as check size decreases over the range from 14 min to 1.4 min. The decrease is gradual over the range from 14 min to 2.8 min and more marked from 2.8 min to 1.4 min. For some subjects, sensitivities to $\theta_{\lrcorner}$ or α drop to zero (i.e., thresholds are greater than 1) for the smallest check size. In contrast, for the first-order statistic γ, there is little change in sensitivity over the entire range, and three subjects may even show a slight increase in sensitivity at the smallest check size. This difference in scaling behavior between the first-order statistic γ and the statistics of order two or more is not surprising: γ, which corresponds to the fraction of white vs. black checks, can be judged even if the individual checks cannot be resolved, while the other statistics require visibility of the individual checks.

Fig. 8 shows how isodiscrimination contours depend on check size. This corroborates the findings of Fig. 7 and extends them to pairwise mixtures of the image statistics. For combinations of

Table 4
Summary statistics for the quadratic model's predictions for thresholds for the combinations of image statistics shown in Fig. 6. In the top half of the table, predictions have no free parameters. In the bottom half of the table, the predictions use an overall scale factor to match the on-axis thresholds for the individual subjects (for MC and DT) or the group mean (for KP and SR) with the on-axis thresholds determined during the sessions used for measuring thresholds for combinations of image statistics (the out-of-sample stimuli). The scale factors used were 0.895 (MC), 0.927 (DT), 1.047 (SR), 1.045 (KP). Other details as in Table 2.

Subject	Direction Type	Mean deviation All	Mean deviation Symmetrized	RMS deviation All	RMS deviation Symmetrized	RMS *z*-score All	RMS *z*-score Symmetrized
Without global sensitivity adjustment							
MC	Eigenvectors	−0.0320	−0.0320	0.0370	0.0245	5.48	4.97
	Minkowski	−0.0356	−0.0365	0.0566	0.0274	5.03	4.20
	All	−0.0341	−0.0341	0.0487	0.0262	5.23	4.56
DT	Eigenvectors	−0.0285	−0.0170	0.0340	0.0237	3.86	2.76
	Minkowski	−0.0394	−0.0202	0.0276	0.0194	5.76	3.96
	All	−0.0349	−0.0189	0.0304	0.0213	4.90	3.41
SR	Eigenvectors	0.0084	0.0181	0.0600	0.0562	1.66	1.39
	Minkowski	−0.0230	0.0169	0.1540	0.1687	1.81	1.82
	All	−0.0099	0.0174	0.1238	0.1339	1.74	1.65
KP	Eigenvectors	0.0248	0.0248	0.0221	0.0148	1.28	1.09
	Minkowski	−0.0101	0.0082	0.0942	0.1104	2.16	1.87
	All	0.0045	0.0151	0.0734	0.0849	1.85	1.63
AVG	Eigenvectors	−0.0068	−0.0015	0.0383	0.0298	3.07	2.55
	Minkowski	−0.0270	−0.0077	0.0838	0.0815	3.69	2.97
	All	−0.0186	−0.0051	0.0691	0.0666	3.43	2.81
With global sensitivity adjustment							
MC	Eigenvectors	−0.0131	−0.0131	0.0212	0.0174	2.50	1.73
	Minkowski	0.0029	0.0029	0.0557	0.0470	2.62	1.45
	All	−0.0038	−0.0038	0.0447	0.0376	2.57	1.58
DT	Eigenvectors	−0.0131	−0.0016	0.0428	0.0111	2.46	1.24
	Minkowski	−0.0184	0.0008	0.0823	0.0317	3.90	1.75
	All	−0.0162	−0.0002	0.0687	0.0253	3.26	1.52
SR	Eigenvectors	−0.0008	0.0089	0.0462	0.0283	2.24	2.17
	Minkowski	−0.0418	−0.0019	0.1221	0.0461	2.12	2.03
	All	−0.0247	0.0026	0.0979	0.0397	2.18	2.09
KP	Eigenvectors	0.0161	0.0161	0.0295	0.0259	0.99	0.66
	Minkowski	−0.0278	−0.0095	0.0837	0.0574	2.27	1.87
	All	−0.0095	0.0012	0.0667	0.0469	1.85	1.54
AVG	Eigenvectors	−0.0027	0.0026	0.0349	0.0207	2.05	1.45
	Minkowski	−0.0213	−0.0019	0.0859	0.0456	2.73	1.78
	All	−0.0136	−0.0001	0.0695	0.0374	2.46	1.68

second- and higher-order image statistics (first two columns: $(\theta_{\lrcorner}, \theta_{\llcorner})$-plane and $(\beta_{\_}, \alpha)$-plane), isodiscrimination contours gradually enlarge as check size decreases from 14 to 2.8 min. They maintain the same shape and tilt, indicating that relative sensitivities to the image statistics and their combinations are approximately independent of check size. Similar results were found in the $(\beta_{\backslash}, \theta_{\lrcorner})$-plane and the $(\theta_{\lrcorner}, \alpha)$-plane (not shown). In planes corresponding to a combination of the first-order statistic and higher-order statistics ($(\gamma, \beta_{\_})$-plane in Fig. 8, last column; and $(\gamma, \theta_{\lrcorner})$- and (γ, α)-planes, not shown), isodiscrimination contours become progressively elongated as check size decreases, especially between 2.8 and 1.4 min. This is the expected behavior based on Fig. 7, since the sensitivity to γ is preserved at 1.4 min, but the sensitivities to the other statistics decrease markedly at the smallest check size.

These observations are summarized in Table 5. An ANOVA was conducted for each image statistic separately, with factors consisting of subject, check size, and polarity (image statistic value >0 vs. <0). For $\beta_{\_}$, $\beta_{\backslash}$, $\theta_{\lrcorner}$, and α, check size was the main source of variance; for γ, it was significant but a smaller source of variance than subject. Polarity (image statistic >0 vs. < 0) was a statistically significant source of variance for $\theta_{\lrcorner}$ and α, but only substantial for α, consistent with the findings of Fig. 3. There was no interaction of check size and polarity. Subject-to-subject differences in the effect of check size (subject × size interactions) were present, but these only accounted for a small fraction of the variance. There was no significant interaction of subject with polarity and the three-way interaction (subject × size × polarity) was also not significant.

The nature of the dependence on check size was quantified by the scaling exponent, i.e., the slope of the dependency of log(sensitivity) on log(check size). This dependence was smaller for γ (0.094) than for the other statistics (range, 0.119–0.441). Importantly, this difference was largely due to the behavior at the smallest check size: when only the three largest check sizes are considered, the scaling exponents for all image statistics were small (0.028–0.252).

3.7. Time course of sensitivity

The results presented above showed that the isodiscrimination surface is approximately elliptical, and the axes of the ellipsoid are consistent across subjects (Fig. 5) and check size (Fig. 8). While ellipsoidal contours need not be indicative of underlying mechanisms (Nielsen & Wandell, 1988; Poirson et al., 1990), we considered the possibility that such a relationship was present. Specifically, we hypothesized that there are a discrete set of neural mechanisms that are tuned along the axes of the ellipsoid, and that stimuli that are aligned to these hypothetical physiological axes might be processed more readily than stimuli that are oblique to them. We therefore measured sensitivities in three image-statistic planes at progressively shorter stimulus durations,

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

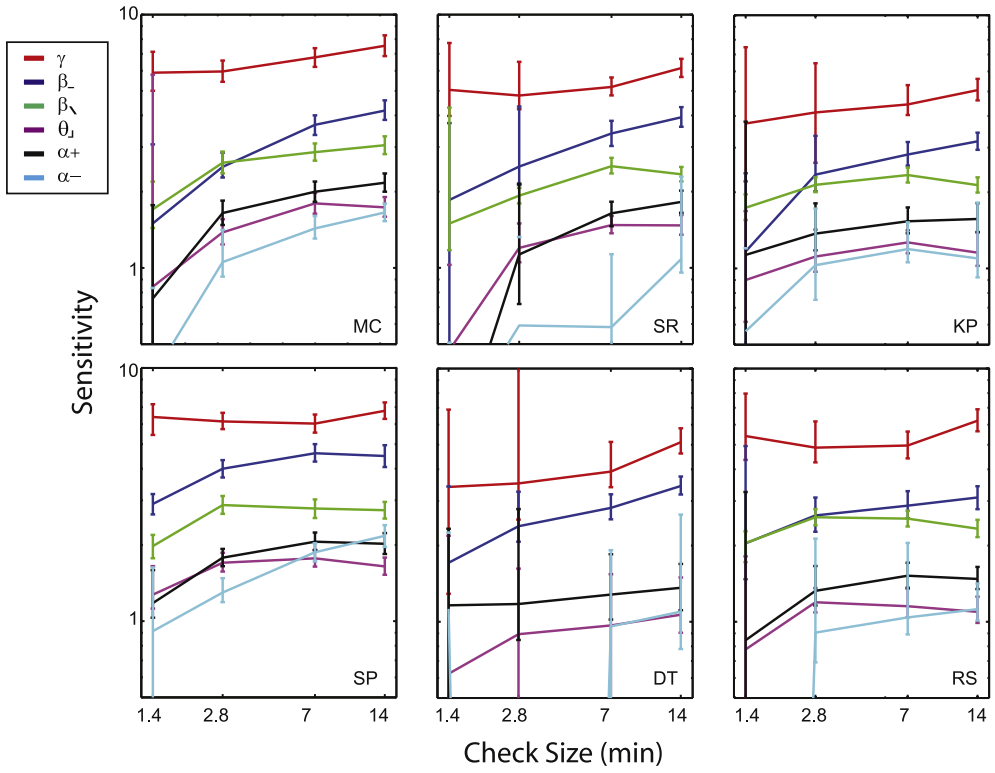

Fig. 7. Sensitivity as a function of check size for the five kinds of coordinates: first-order (γ), cardinal second-order ($\beta_{\_}$), diagonal second-order ($\beta_{\backslash}$), third-order ($\theta_{\lrcorner}$), and fourth-order (α). $N = 6$ subjects for γ, $\beta_{\_}$, $\theta_{\lrcorner}$, and α; $N = 5$ for $\beta_{\backslash}$ (all but DT). For α, sensitivities for positive and negative excursions are plotted separately as α_+ and α_-, since these were systematically different across subjects. Subjects MC, SR, KP, SP, DT, RS.

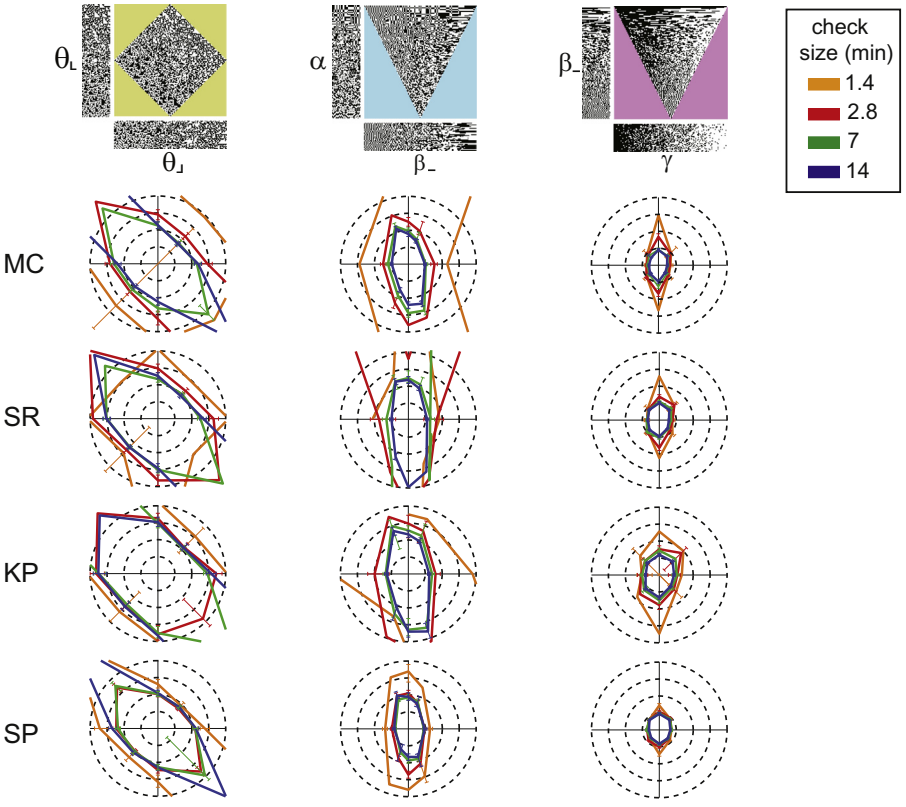

Fig. 8. Isodiscrimination contours as a function of check size for three pairwise combinations of the image statistics, in $N = 4$ subjects. Heavy lines indicate the loci corresponding to a fraction correct of 0.625, halfway between chance and perfect, for each check size. Thin lines indicate 95% confidence limits via bootstrap, and are omitted where some bootstrap resamplings had thresholds greater than 1. Subjects MC, SR, KP, and SP.

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

Table 5
Analysis of scaling. Upper portion of the table: ANOVA of sensitivities. * indicates $p < 0.05$, ** indicates $p < 0.01$, *** indicates $p < 0.001$, Bonferroni-corrected. Polarity refers to positive vs. negative values of coordinates, and Size refers to check size. Lower portion of table: scaling exponents, determined from regression slopes on log–log coordinates. Values quoted are means and range across all subjects with nonzero thresholds over the specified range of check sizes ($N = 6$ for γ, β_-, $\theta_\lrcorner$, and α_+; $N = 5$ for $\beta_\backslash$, $N = 4$ for α_-).

Analysis of Variance

Source	γ		β_-		$\beta_\backslash$		$\theta_\lrcorner$		α	
	frac SS	F (dof)	Frac SS	F (dof)	frac SS	F (dof)	frac SS	F (dof)	frac SS	F (dof)
Main effects										
Subject	0.420	42.349 (5)***	0.216	19.474 (5)**	0.148	3.399 (4)	0.235	45.427 (5)***	0.178	30.140 (5)***
Size	0.126	21.277 (3)***	0.367	55.091 (3)***	0.305	9.305 (3)**	0.256	82.382 (3)***	0.243	68.459 (3)***
Polarity	0.016	7.905 (1)	0.008	3.635 (1)	0.001	0.056 (1)	0.042	40.633 (1)***	0.111	94.238 (1)***
Interactions										
Size × Polarity	0.005	0.844 (3)	0.003	0.407 (3)	0.017	0.533 (3)	0.003	1.001 (3)	0.007	1.982 (3)
Subject × Size	0.030	1.018 (15)	0.047	1.407 (15)	0.070	0.532 (12)	0.089	5.741 (15)***	0.073	4.111 (15)***
Subject × Polarity	0.004	0.433 (5)	0.010	17.143 (5)	0.004	0.102 (4)	0.017	3.302 (5)	0.009	1.599 (5)
Subject × Size × Polarity	0.005	0.178 (15)	0.019	31.488 (15)	0.018	0.141 (12)	0.013	0.839 (15)	0.030	1.666 (15)

Scaling Exponents

	γ	β_-	$\beta_\backslash$	$\theta_\lrcorner$	α_+	α_-
All check sizes (1.4–14 min)						
Mean	0.094	0.302	0.139	0.219	0.343	0.466
Min	0.018	0.170	0.047	0.102	0.073	0.270
Max	0.169	0.403	0.208	0.456	0.969	0.610
Three largest check sizes (2.8–14 min)						
Mean	0.141	0.199	0.028	0.058	0.135	0.252
Min	0.056	0.076	−0.062	−0.053	0.072	0.045
Max	0.227	0.323	0.130	0.149	0.303	0.355

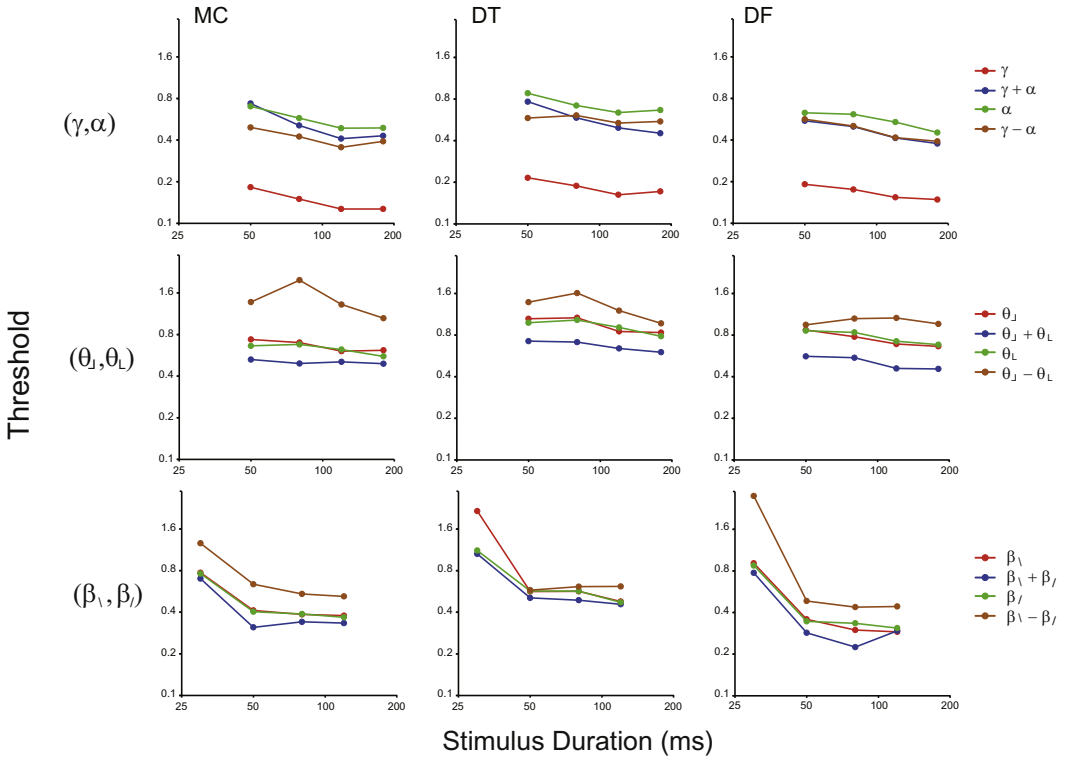

Fig. 9. Thresholds as a function of stimulus duration in planes sampling pairwise combinations of the image statistics (γ, α) and $(\theta_\lrcorner, \theta_\llcorner)$-plane, 50–180 ms; $(\beta_\backslash, \beta_/)$-plane, 30–120 ms). Each plot shows threshold vs. time in one direction of the indicated coordinate plane. Red trace: along the first coordinate; green trace: along the second coordinate; blue trace: along the diagonal in which both coordinates have the same sign; brown: along the diagonal in which coordinates have opposite signs. Subjects MC, DT, DF. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

to determine whether image statistics that were aligned to the axes of the isodiscrimination contours were processed more rapidly than those that are oblique.

Fig. 9 shows that no such difference could be identified. As expected, thresholds increased as stimulus duration decreased, but the threshold-vs.-time curves were parallel for each of the directions tested, and showed no systematic difference between directions that were parallel to, vs. oblique to, the axes of the isodiscrimination ellipsoid. This includes the (γ, α)-plane (in which the isodiscrimination ellipse is aligned with the coordinate axes), the $(\theta_{\lrcorner}, \theta_{\llcorner})$-plane (in which the isodiscrimination ellipse is tilted by 45 deg with respect to the coordinate axes), and the $(\beta_{\backslash}, \beta_{/})$-plane, in which the isodiscrimination ellipse is approximately circular. Similar results were obtained in four other planes $((\beta_{\_}, \alpha)$, $(\theta_{\lrcorner}, \theta_{\ulcorner})$, $(\beta_{\_}, \beta_{|})$, $(\beta_{\_}, \beta_{\backslash}))$ in one or two of these subjects. Thus, threshold-vs.-time characteristics are similar in all tested directions, and do not suggest the presence of a discrete set of underlying physiological mechanisms.

4. Discussion

4.1. Results summary

Above we presented a series of measurements of sensitivity to local image statistics – including luminance distribution, frequency of edges at different orientations, and frequency of corners at different orientations. Our main findings are, (i), a specific set of sensitivities to individual image statistics, in rank order, $\gamma > \{\beta_{\_}, \beta_{|}\} > \{\beta_{/}, \beta_{\backslash}\} > \alpha > \theta$, with an approximate fivefold difference between the largest (γ) and smallest (θ) sensitivities, (ii) similar sensitivities to positive and negative correlations, with the only exception that sensitivities to positive excursions of the four-point correlation α was approximately 25% greater than sensitivities to negative excursions; (iii) interactions between pairs of individual image statistics that were well-described by a quadratic combination rule and an ellipsoidal isodiscrimination surface, which provided accurate predictions of sensitivities to complex combinations of image statistics, (iv) approximate independence of sensitivities to individual image statistics and their combinations over the check size range from 2.8 min to 14 min, and (v) similar dependence of sensitivities to individual image statistics and their combinations on stimulus duration. All of these findings were consistent across a subject pool that included naïve and non-naïve observers, as well as an expert and novices.

4.2. A strategy for characterizing sensitivity to image statistics and their interactions

To enable the analysis of psychophysical sensitivity to multiple different kinds of image statistics and their interactions, we took two steps. First, we reduced the dimensionality of the problem by restricting consideration to black-and-white images on a lattice and to image statistics defined within a 2×2 block. Second, we used a coordinate system that was based on a set of parameters $(\{\gamma, \beta_{\_}, \beta_{|}, \beta_{\backslash}, \beta_{/}, \theta_{\lrcorner}, \theta_{\llcorner}, \theta_{\ulcorner}, \theta_{\urcorner}, \alpha\})$ that eliminated many of the dependencies between intuitively-defined image statistics. As described in Methods and (Victor & Conte, 2012), these statistics parameterize a 10-dimensional space of textures. To examine sensitivities to one or more image statistics, we used stimuli that were constrained by the specific image statistics of interest but were otherwise as random as possible.

The construction of the stimulus space has several implications for interpreting our results. As with the familiar three-dimensional color space, the points in the space correspond to specific stimuli, line segments between two points correspond to mixtures between the corresponding stimuli, and the dimensions of the space represent parameters that can be varied independently. But the underlying geometries of the spaces are different. Color space is generally thought of as an affine space (Zaidi et al., 2013), meaning that there is no intrinsic notion of length or angle. The reason for this is that measurements of lengths and angles depend on the coordinate system used (since they are computed from the dot product). In color space, there are many reasonable choices for coordinate systems (Derrington, Krauskopf, & Lennie, 1984; MacLeod & Boynton, 1979; Wyszecki & Stiles, 1967); each choice leads to its own set of distances and angles. In contrast, there is an intrinsic notion of length and angle in the space of image statistics: length corresponds to performance of the ideal observer, and a right angle corresponds to statistical independence (Victor & Conte, 2012). As a corollary, the ideal observer will have a spherical isodiscrimination surface.

The existence of this intrinsic geometry gives functional significance to the shape of the human isodiscrimination surface. Just as a spherical surface corresponds to the performance of an ideal observer who makes full use of all image statistics, distortions from the spherical shape are statements about the selective sensitivity of the human visual system. But only some of the many possible kinds of distortions are actually prominent. For example, sensitivities to positive and negative excursions along an axis (i.e., excursions in opposite directions from the origin) might differ, but with only one exception (positive vs. negative values of α, a 25% difference), substantial deviations of this kind do not occur. The main distortion from sphericity is due to selective sensitivity to different kinds of image statistics: as mentioned above, sensitivities to the image statistics represented by the coordinate axes differ by approximately a factor of 5, producing a corresponding distortion of the surface.

The way that different kinds of image statistics combine determines the overall shape of the isodiscrimination surface. In individual coordinate planes, we found that cues combine in a quadratic fashion – in analogy with color space (Macadam, 1942; Poirson et al., 1990), previous work in color and orientation (Saarela & Landy, 2012), and extending our recent study in a portion of this space (Victor, Thengone, & Conte, 2013). That is, the intersection of the isodiscrimination surface with each coordinate plane is approximately an ellipse.

To move from measurements of pairwise interactions of image statistics within each plane to a testable model for thresholds for complex combinations, we assumed that the quadratic combination rule held in general, and not only within the coordinate planes. This in turn implies that the isodiscrimination surface is an ellipsoid. As with any ellipsoid, an ellipsoidal isodiscrimination contour is completely characterized by its principal axes, which lie in a set of orthogonal directions that includes the directions of greatest and least sensitivity. While these axes may be oblique to all of the coordinate planes, they nevertheless are uniquely determined by the intersection of the ellipsoid with each coordinate plane. We used this approach (Eqs. (11) and (12)) to determine the principal axes, and, thereby, to estimate the full isodiscrimination surface. This led to a predicted direction of maximal sensitivity (direction sym1, Table 3) that was not aligned to any of the coordinate axes, but instead consisted of a texture with an excess of white checks ($\gamma > 0$), along with an excess of horizontal and vertical edges (negative nearest-neighbor pairwise interactions $\beta_{\_} < 0$), and small amounts of other image statistics. The analysis also predicted that there was a unique direction in which sensitivity was essentially 0 (direction hvi2, Table 3) – effectively, a texture that was metameric to random. The predicted sensitivities along these and all other principal axes were close to the values determined experimentally (Fig. 6 and Table 4). The quadratic model had similar predictive accuracy in other selected directions of the texture space. These

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018

other directions were unaligned with the coordinate axes or the principal axes, and were based on topological properties of the corresponding textures.

The present approach to analysis of image statistics differs from that of Portilla and Simoncelli (Portilla & Simoncelli, 2000) in several respects. The Portilla and Simoncelli approach uses hundreds of parameters across many octaves of scale and multiple orientations. Here, the number of parameters is strictly limited, and we consider only a single spatial scale, four orientations, and two luminance values. These differences reflect a difference in goals. The Portilla and Simoncelli approach supports representation and synthesis of naturalistic textures with high fidelity; here, our focus is to understand how image statistics interact and the structure of the resulting perceptual space. For the latter purpose, it is crucial to be able to manipulate individual image statistics independently. In the Portilla and Simoncelli approach, this is not possible – the representation is overcomplete, and there are implicit nonlinear relationships between the parameters. In contrast, the 10 image statistics used here correspond to 10 degrees of freedom that can be independently accessed via texture synthesis algorithms (Victor & Conte, 2012). The resulting textures are highly artificial, but because of this artificiality, they provide independent probes of the main image statistics fundamental to form vision (contrast, edge, and corner).

While each set of measurements here considered only a single spatial scale of statistics (i.e., a single check size), this is not likely to place a significant limitation on our conclusions. Over the range of check sizes from 2.8 to 14 min, the pattern of sensitivities showed an approximate scale-invariance (Figs. 7 and 8, and Table 5). At the smallest check size studied (1.4 min), sensitivity to pairwise and higher-order image statistics was substantially reduced – an expected finding, since individual checks of this size are no longer clearly visible. We did not study combinations of image statistics at two different spatial scales (e.g., β_- at one scale and α at another), but we hypothesize that a quadratic combination rule will hold in this case as well.

4.3. Symmetry and asymmetry with respect to contrast polarity and sign of correlation

Our results show that for first-order statistics (γ), there is a modest difference between sensitivity to positive and negative values: on average, a 7% superiority for $\gamma < 0$ vs. $\gamma > 0$, significant ($p < 0.02$) across subjects. Negative vs. positive values of γ correspond to an excess of black checks vs. an excess of white checks. This modest superiority for darks over lights is consistent with the well-recognized superiority for darks over lights in terms of spatial resolution ((Kremkow et al., 2014) and review therein; (Zemon, Gordon, & Welch, 1988)), and temporal resolution (Komban, Alonso, & Zaidi, 2011; Komban et al., 2014), and may be related to the observation of a perceptual mechanism sensitive to "blackshot" (Chubb, Landy, & Econopouly, 2004). Physiological differences between ON and OFF cells that are likely contributors to these asymmetries are known at the retinal (Chichilnisky & Kalmar, 2002; Pandarinath, Victor, & Nirenberg, 2010) and cortical (Xing, Yeh, & Shapley, 2010; Yeh, Xing, & Shapley, 2009) levels. However, one should be cautious about making a direct connection: the limiting factor in processing image statistics is not directly related to resolution; indeed, larger receptive fields may be more effective than smaller ones in gathering the necessary information.

The more substantial asymmetry in the present data – the 25% advantage of positive over. negative values for α – is likely unrelated to black/white asymmetries. This is because α is computed from the product of the contrast value at four points, so inverting contrast of the image (i.e., swapping black for white) does not change its value. Conversely, for both $\alpha > 0$ and $\alpha < 0$, on-axis images have a 50:50 mix of blacks and whites. We hypothesize that the asymmetry for α has another origin: that visual analysis takes place across many spatial scales, and that multiscale analysis has different implications for positive vs. negative correlations. To see what these implications are, consider analyzing a checkerboard texture at a scale that is equal to twice the check size. That is, instead of computing α from four contiguous checks (e.g., from checks at positions (i,j), $(i+1,j)$, $(i,j+1)$, and $(i+1,j+1)$), we compute it from the four checks at positions (i,j), $(i+2,j)$, $(i,j+2)$, and $(i+2,j+2)$. It can be shown that the larger-scale value of α is related to the original value by $\alpha_{\text{largescale}} = \alpha^4$ (Appendix 1 of (Victor & Conte, 1989)). Thus, a texture generated with $\alpha > 0$ will have positive values of α when analyzed at this larger scale as well. But, a texture generated with $\alpha < 0$ will also have positive values at this larger scale, as $\alpha_{\text{largescale}} = \alpha^4$ is positive even when α is negative. In sum, positive values of α at one scale will be reinforced by analysis at larger scales, but negative values of α will tend to be canceled by *positive* values when analyzed at a larger scale. We hypothesize that this cancellation results in a somewhat lower sensitivity when statistics are pooled across spatial scales. Of note, larger signal sizes for $\alpha = +1$ ("even texture") than for $\alpha = -1$ were also found in the visual evoked potential (Victor & Conte, 1989). No comparable asymmetry between positive and negative pairwise correlations (the β's) was seen. We hypothesize that this is because pooling across spatial scales is less prominent for analysis of two-point interactions.

5. Conclusion

Local image statistics form a domain that is complex both because of its high dimensionality and geometric considerations that make them interdependent. Restricting attention to two luminance values and a single spatial scale leads to a 10-parameter space of image statistics that includes variations in luminance, contrast, orientation content, and corner. Within this space, visual sensitivities are stereotyped across human observers, and a simple quadratic combination rule suffices to account for thresholds to combinations of multiple statistics.

Acknowledgments

Portions of this work were presented at the Vision Sciences Society (2012 and 2013: Naples, FL; 2014: St. Petersburg, FL), and the Society for Neuroscience (2012: New Orleans, LA; 2013: San Diego, CA). This work was supported by NIH NEI EY7977. We are grateful to Ted Maddess especially for the suggestion to use the Minkowski functionals, and to both Ted and Charles F. Chubb for their many very helpful discussions and insights. We are also very grateful to our subjects.

References

Barbosa, M. S., Bubna-Litic, A., & Maddess, T. (2013). Locally countable properties and the perceptual salience of textures. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 30*(8), 1687–1697. http://dx.doi.org/10.1364/JOSAA.30.001687.

Benjamini, Y., & Hochberg, Y. (2001). The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics, 29*(4), 1165–1188.

Ben-Shahar, O., & Zucker, S. W. (2004). Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research, 44*, 257–277.

Chichilnisky, E. J., & Kalmar, R. S. (2002). Functional asymmetries in ON and OFF ganglion cells of primate retina. *Journal of Neuroscience, 22*(7), 2737–2747. doi: 20026215.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 11*(9), 2350–2374.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research, 44*(27), 3223–3232.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology, 357*, 241–265.

Doi, E., Balcan, D. C., & Lewicki, M. S. (2007). Robust coding over noisy overcomplete channels. *IEEE Transactions on Image Processing, 16*(2), 442–452.

Hermundstad, A. M., Briguglio, J. J., Conte, M. M., Victor, J. D., Balasubramanian, V., & Tkacik, G. (2014). Variance predicts salience in central sensory processing. *eLife, 3*. http://dx.doi.org/10.7554/eLife. 03722.

Julesz, B., Gilbert, E. N., & Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics, 31*(3), 137–140.

Karklin, Y., & Lewicki, M. S. (2009). Emergence of complex cell properties by learning to generalize in natural scenes. *Nature, 457*(7225), 83–86. http://dx.doi.org/10.1038/nature07481.

Komban, S. J., Alonso, J. M., & Zaidi, Q. (2011). Darks are processed faster than lights. *Journal of Neuroscience, 31*(23), 8654–8658. http://dx.doi.org/10.1523/JNEUROSCI.0504-11.2011.

Komban, S. J., Kremkow, J., Jin, J., Wang, Y., Lashgari, R., Li, X., et al. (2014). Neuronal and perceptual differences in the temporal processing of darks and lights. *Neuron, 82*(1), 224–234. http://dx.doi.org/10.1016/j.neuron.2014.02.020.

Kremkow, J., Jin, J., Komban, S. J., Wang, Y., Lashgari, R., Li, X., et al. (2014). Neuronal nonlinearity explains greater visual spatial resolution for darks than lights. *Proceedings of the National academy of Sciences of the United States of America, 111*(8), 3170–3175. http://dx.doi.org/10.1073/pnas.1310442111.

Landy, M. S., & Oruc, I. (2002). Properties of second-order spatial frequency channels. *Vision Research, 42*(19), 2311–2329.

Li, A., & Lennie, P. (1997). Mechanisms underlying segmentation of colored textures. *Vision Research, 37*(1), 83–97.

Macadam, D. L. (1942). Visual sensitivities to color differences in daylight. *Journal of the Optical Society of America, 32*, 247–273.

MacLeod, D. I., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *The Journal of the Optical Society of America, 69*(8), 1183–1186.

Michielsen, K., & De Raedt, H. (2001). Integral-geometry morphological image analysis. *Physics Reports-Review Section of Physics Letters, 347*(6), 462–538.

Motoyoshi, I., & Kingdom, F. A. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Research, 2007*, 2047–2054.

Nielsen, K. R., & Wandell, B. A. (1988). Discrete analysis of spatial-sensitivity models. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 5*(5), 743–755.

Pandarinath, C., Victor, J. D., & Nirenberg, S. (2010). Symmetry breakdown in the ON and OFF pathways of the retina at night: Functional implications. *Journal of Neuroscience, 30*(30), 10006–10014. http://dx.doi.org/10.1523/JNEUROSCI.5616-09.2010.

Poirson, A., Wandell, B., Varner, D., & Brainard, D. (1990). Surface characterizations of color thresholds. *The Journal of the Optical Society of America, A7*(783–789).

Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision, 40*(1), 49–71.

Saarela, T., & Landy, M. S. (2012). Combination of texture and color cues in visual segmentation. *Vision Research, 58*, 59–67.

Serre, J.-P. (1977). *Linear representations of finite groups*. New York: Springer.

Tkačik, G., Prentice, J. S., Victor, J. D., & Balasubramanian, V. (2010). Local statistics in natural scenes predict the saliency of synthetic textures. *Proceedings of the National academy of Sciences of the United States of America, 107*(42), 18149–18154.

Victor, J. D., Chubb, C., & Conte, M. M. (2005). Interaction of luminance and higher-order statistics in texture discrimination. *Vision Research, 45*(3), 311–328.

Victor, J. D., & Conte, M. M. (1989). Cortical interactions in texture processing: Scale and dynamics. *Visual Neuroscience, 2*(3), 297–313.

Victor, J. D., & Conte, M. M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research, 31*(9), 1457–1488.

Victor, J. D., & Conte, M. M. (1996). The role of high-order phase correlations in texture processing. *Vision Research, 36*(11), 1615–1631.

Victor, J. D., & Conte, M. M. (2004). Visual working memory for image statistics. *Vision Research, 44*(6), 541–556.

Victor, J. D., & Conte, M. M. (2012). Local image statistics: Maximum-entropy constructions and perceptual salience. *Journal of the Optical Society of America A: Optics, Image Science, and Vision, 29*(7), 1313–1345. http://dx.doi.org/10.1364/JOSAA.29.001313.

Victor, J. D., Thengone, D. J., & Conte, M. M. (2013). Perception of second- and third-order orientation signals and their interactions. *Journal of Vision, 13*(4), 21. http://dx.doi.org/10.1167/13.4.21.

Wolfson, S. S., & Landy, M. S. (1995). Discrimination of orientation-defined texture edges. *Vision Research, 35*(20), 2863–2877.

Wolfson, S. S., & Landy, M. S. (1998). Examining edge- and region-based texture analysis mechanisms. *Vision Research, 38*, 439–446.

Wyszecki, G., & Stiles, W. (1967). *Color science: Concepts and methods, quantitative data and formulas*. New York: Wiley.

Xing, D., Yeh, C. I., & Shapley, R. M. (2010). Generation of black-dominant responses in V1 cortex. *Journal of Neuroscience, 30*(40), 13504–13512. http://dx.doi.org/10.1523/JNEUROSCI.2473-10.2010.

Yeh, C. I., Xing, D., & Shapley, R. M. (2009). "Black" responses dominate macaque primary visual cortex v1. *Journal of Neuroscience, 29*(38), 11753–11760. http://dx.doi.org/10.1523/JNEUROSCI.1991-09.2009.

Zaidi, Q., Victor, J., McDermott, J., Geffen, M., Bensmaia, S., & Cleland, T. A. (2013). Perceptual spaces: Mathematical structures to neural mechanisms. *Journal of Neuroscience, 33*(45), 17597–17602. http://dx.doi.org/10.1523/jneurosci.3343-13.2013.

Zemon, V., Gordon, J., & Welch, J. (1988). Asymmetries in ON and OFF visual pathways of humans revealed using contrast-evoked cortical potentials. *Visual Neuroscience, 1*(1), 145–150.

Please cite this article in press as: Victor, J. D., et al. A perceptual space of local image statistics. *Vision Research* (2015), http://dx.doi.org/10.1016/j.visres.2015.05.018