CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Nottingham ePrints

Size-induced distortions in perceptual maps of visual space

Paul V. McGraw Visual Neuroscience Group, School of Psychology, The University of Nottingham, Nottingham, UK

Neil W. Roach Visual Neuroscience Group, School of Psychology, The University of Nottingham, Nottingham, UK

David R. Badcock School of Psychology, The University of Western Australia, Crawley, WA, Australia

David Whitaker Bradford School of Optometry and Vision Sciences, University of Bradford, Bradford, UK

In order to interact with our environment, the human brain constructs maps of visual space. The orderly mapping of external space across the retinal surface, termed retinotopy, is maintained at subsequent levels of visual cortical processing and underpins our capacity to make precise and reliable judgments about the relative location of objects around us. While these maps, at least in the visual system, support high precision judgments about the relative location of objects, they are prone to significant perceptual distortion. Here, we ask observers to estimate the separation of two visual stimuli—a spatial interval discrimination task. We show that large stimulus sizes require much greater separation in order to be perceived as having the same separation as small stimulus sizes. The relationship is linear, task independent, and unrelated to the perceived position of object edges. We also show that this type of spatial distortion is not restricted to the object itself but can also be revealed by changing the spatial scale of the background, while object size remains constant. These results indicate that fundamental spatial properties, such as retinal image size or the scale at which an object is analyzed, exert a marked influence on spatial coding.

Keywords: distortions, perceptual maps, visual space

Citation: McGraw, P. V., Roach, N. W., Badcock, D. R., & Whitaker, D. (2012). Size-induced distortions in perceptual maps of visual space. *Journal of Vision, 12*(4):8, 1–14, http://www.journalofvision.org/content/12/4/8, doi:10.1167/12.4.8.

Introduction

The orderly representation of visual space is one of the fundamental organizing principles of visual analysis. The presence of this organization, referred to as a "retinotopic map" is evident at the earliest levels of visual sensory transduction. Reflected light from objects in our visual field is imaged on the retina in such a way that the spatial arrangement of the scene is preserved on the retinal surface: Adjacent points in the scene are analyzed or sampled by adjacent detectors in the photoreceptor mosaic. This orderly mapping of space is maintained at subsequent stages of visual processing. Unique retinotopic maps are found at the level of retinal ganglion cells (RGC), lateral geniculate nucleus (LGN), and the primary visual cortex, although at the level of the cortex, increasing space is devoted to the central region of the visual field (Daniel & Whitteridge, 1961; Schwartz, 1980). However, despite this regional change in the scale of the map in the cortex, the same general principle applies: The receptive fields of neighboring neurons sample adjacent regions of space. This form of cerebral cartography where adjacent sensory receptors stimulate adjacent cortical locations is not unique to the visual system and is replicated in other areas of sensory cortex. For example, the representation of the fingertips in somatosensory cortex also mirrors their spatial layout on the hand (Sanchez-Panchuelo, Francis, Bowtell, & Schluppeck, 2010).

Given the existence of these orderly and detailed maps of visual space, the task of determining the location of an object, or indeed its position relative to another object, appears somewhat trivial. Provided the visual system can encode spatially localized activity and has a means to establish where the activity originates in space (a labeled line), the task of determining the relationships of objects should be computationally straightforward. This reasoning was embraced in the early models of visual position coding, most notably in that described by Rudolph Hermann Lotze (1852; see also related discussion in section 6 of "Molyneux's Question" by Morgan, 2009). In the "local signs" model, Lotze outlined a scheme where stimulation

doi: 10.1167/12.4.8 Received July 28, 2011; published April 17, 2012 ISSN 1534-7362 © ARVO

arising at the retina is transferred to the brain, via the optic nerve, with a unique identifier (or local sign) that tags where it has arisen from in space—in the same way a postcode might be used to identify a particular street in a large town. Such a scheme supports the reassembly of spatial maps at subsequent stages of visual processing. Even though the retina is rarely static due to eye movements (e.g., fixation instability), the relative separation between objects are preserved, and as a result, the accuracy of spatial judgments are largely unaffected (Badcock & Wong, 1990). This simple scheme, although appropriate for a 2D representation, is problematic when it comes to 3D stimuli, as adjacent retinal elements may have to analyze information that arises from different relative depths.

Under normal conditions, when spatial maps are fully developed, they remain accurate and stable entities throughout life. In light of this, subsequent models of position coding underwent a change in focus, and rather than concentrating on the absolute fidelity of spatial maps, they instead attempted to address a number of new questions. How accurate are judgments of relative position? Do the mechanisms that underpin positional sensitivity exhibit lawful behavior? We now have a good understanding of these issues. The human visual system is capable of localizing objects in visual space with a precision finer than that offered by the sampling mosaic of retinal receptors (Westheimer, 1975). Contemporary models of positional coding rely on differential responses of spatial frequency- and orientation-tuned cortical mechanisms to explain the observed thresholds and have been highly successful in this regard for abutting or closely spaced stimuli matched in contrast (Klein & Levi, 1985; Regan & Beverly, 1983; Wilson, 1986) but much less so for separated elements where individual position labels determine geometric relationships (Burbeck, 1987; Morgan & Regan, 1987). Thresholds for positional tasks, such as foveal spatial interval discrimination, have also been shown to display systematic behavior where thresholds are directly proportional to the interval itself (Levi & Klein, 1992; Levi, Klein, & Yap, 1988). Furthermore, positional information is encoded independently of many other stimulus attributes for widely separated objects. Both Burbeck (1987) and Kooi, De Valois, and Switkes (1991) have shown that positional thresholds are essentially immune to changes in spatial frequency, orientation, and color of their internal spatial structure.

The fact that positional thresholds reach truly impressive levels and are robust to many stimulus manipulations should not lead one to conclude that under these circumstances the visual system displays a high degree of absolute accuracy. Thresholds describe the precision with which an observer can distinguish small changes in the physical position of visual stimuli, either in isolation or relative to other stimuli. This ability can remain impressively good yet coexist with large perceptual errors (Badcock & Westheimer, 1985; Morgan, Hole, & Glennester, 1990; Morgan, Ward, & Hole, 1990). The decoupling between sensitivity and bias strikingly revealed several well-documented geometrical illusions that involve judgments of length (e.g., Muller–Lyer), angle (e.g., Zollner), and size (e.g., Baldwin; see Robinson, 1972 for a detailed list). For example, in the Muller–Lyer illusion, the addition of fins to a line of fixed length can make the line appear longer or shorter depending on fin configuration, yet this manipulation has no measurable impact on the subject's threshold for length judgments (Morgan, Hole et al., 1990; Morgan, Ward et al., 1990). Similarly, when observers are instructed to bisect a line flanked by boxes of different sizes, they tend to place the bisection marker closer to the smaller, or further from the larger, of the two boxes (Baldwin, 1895). While a wide variety of theories based around perspective, spatial assimilation, or feature contrast have been advanced over the years, none unfortunately offer a comprehensive account of geometrical illusions.

More recently, a number of influential studies have begun to reveal new situations in which positional coding is distinctly non-veridical. The most widely investigated errors are those that involve motion. Several studies have shown that motion within a stationary object causes the object itself to appear displaced in the direction of motion (Anstis & Ramachandran, 1990; De Valois & De Valois, 1991; Edwards & Badcock, 2003; Nishida & Johnston, 1999; Whitaker, McGraw, & Pearson, 1999). Similarly, motion that occurs in the vicinity of briefly presented static objects causes a flashed object to appear offset in the direction of nearby motion (Roach & McGraw, 2009; Whitney & Cavanagh, 2000). This effect occurs across surprisingly large regions of space, where the inducing motion and stationary objects occupy geographically distinct regions of the visual field (Whitney & Cavanagh, 2000). Distortions of positional maps are also found when retinal motion is introduced via ballistic eye movements (saccades). Just before, or early in the saccadic sequence, briefly presented targets are perceived in new illusory positions (see Ross, Morrone, Goldberg, & Burr, 2001). Targets are commonly mislocalized in the direction of the saccade (Honda, 1989) and the space between sequential targets is often perceptually compressed (Ross, Morrone, & Burr, 1997).

The large body of classic geometrical illusions described over the last century, combined with more recent motion-related distortions of visual space, provide compelling evidence that spatial maps are not fixed analyzers but are in fact mutable: dynamically altered by the context in which we view a visual stimulus. We present evidence to show that fundamental spatial properties, such as retinal image size or the scale at which an object is analyzed, exert a marked influence on spatial coding.

Methods and results

A total of seven observers participated across four experiments: the four authors and three inexperienced

psychophysical observers who were naive to the purposes of the experiment. Observations were carried out under dim room illumination, using binocular viewing conditions and with one practice session prior to data collection. All subjects had normal or corrected-to-normal visual acuity. Visual stimuli were generated using MATLAB and displayed via a Cambridge Research Systems ViSaGe on a gamma-corrected Mitsubishi Diamond Pro 2045U monitor (frame rate 100 Hz, 1 pixel subtends 2 arcmin at a viewing distance of 65 cm, mean luminance of 47 cd m^{-2}).

Experiment 1: Spatial interval discrimination

Observers were presented with two horizontally separated luminance-defined Gaussian patches. The mathematical description of the stimuli is given by

$$L(x,y) = L_{\text{mean}} + A\ \exp(-(x^2+y^2)/2\sigma^2), \quad (1)$$

where L_{mean} is the mean luminance of the background, A is the luminance amplitude and σ is the standard deviation of the Gaussian envelope. The vertical and horizontal distances from the peak of the Gaussian envelope are denoted by x and y. In addition, in separate experiments, observers were presented with Gabor patches that varied in their spatial frequency content (carrier frequency range of 0.89–7.63 c deg^{-1}), and Gabor patches of a fixed spatial frequency (7.63 c deg^{-1}). The Gabor patches are described by

$$L(x,y) = L_{\text{mean}} + \text{A}/2\ \sin\left((2\pi Nx/\sigma) + \phi\right) \cdot \exp(-(x^2+y^2)/2\sigma^2), \quad (2)$$

where N is the number of cycles of carrier grating per standard deviation and the random variable ϕ represents the phase of the carrier grating. All other parameters are as before. Each pair of patches was separated by a baseline separation of 480 arcmin. Patch size (defined as the standard deviation of the Gaussian envelope) varied from 3.2 to 51.2 arcmin. A peak Weber contrast (Gaussian) and a Michelson contrast (Gabors) of 84% were used throughout. Examples of the Gaussian and Gabor stimuli—both variable and fixed spatial frequency-are presented in Figures 1a–1c, respectively.

On a single 200-ms presentation, any one of seven predetermined separations and seven patch sizes was displayed. Following each presentation, the subject was required to indicate whether the separation between patch centers was perceived to be greater or smaller than the mean of all the previous presentations. The results of the first 49 trials were discarded to allow observers to construct their own internal metric with which to compare each trial. Tasks of this type, in which a single standard is maintained across a dimension, can be performed with relative ease and high accuracy (Morgan, 1992; Morgan, Watamaniuk, & McKee, 2000). The resulting psychometric functions for each patch size were fitted with a logistic function of the following form:

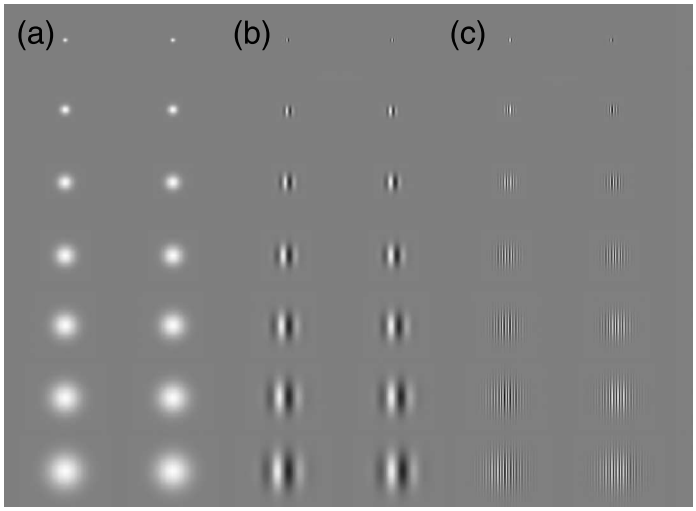

Figure 1. Examples of the three types of stimuli used in Experiment 1: (a) Gaussian patches, (b) Gabor patches that vary in carrier spatial frequency, and (c) Gabor patches with a fixed carrier spatial frequency.

$$Y = \left(100/\left(1+\left(\exp^{-((x-\phi)/\theta)}\right)\right)\right), \quad (3)$$

where ϕ is the offset corresponding to the 50% level on the psychometric function and θ provides an estimate of spatial interval discrimination threshold (half the offset between the 27% and 73% levels on the psychometric function approximately).

Figure 2 summarizes the basic observation. Large patch sizes require much greater separation in order to be perceived as having the same center-to-center separation as small patch sizes. The relationship is linear, with perceived separation changing by ~2 arcmin for every 1 arcmin variation in patch size.

One possible explanation is that changes in stimulus size affect perceived distance from the observer (large objects appear closer). If so, this could potentially result in a given retinal separation being interpreted as a greater physical separation as the size of the object is reduced. Perceived depth, or constancy scaling, has long been implicated as a major factor underlying classic geometric illusions (Gregory, 1963, 1968). In order to investigate the role of perceived distance, we replaced the Gaussian

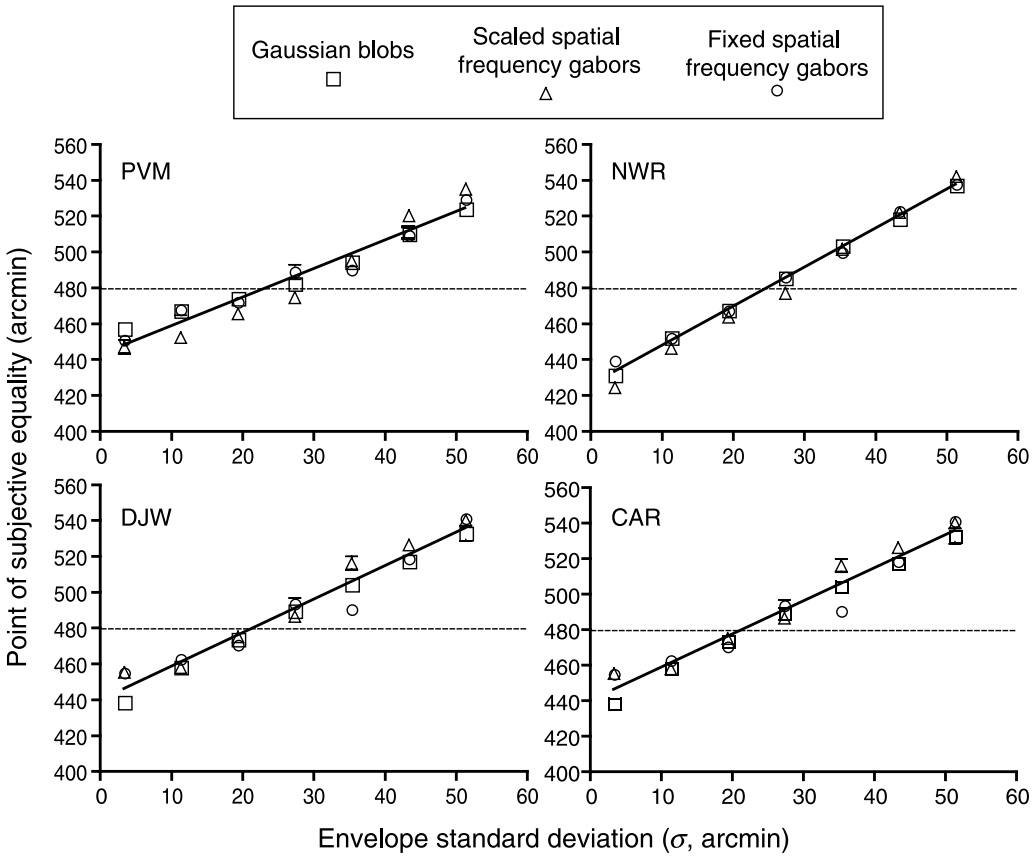

Figure 2. The physical separation of the two patches required for the stimuli to appear equal to the mean separation as a function of patch size for four observers. Data from three stimulus conditions are presented—Gaussian patches (squares), fixed-frequency Gabor patches (circles), and variable frequency Gabor patches (triangles). The dashed line represents the veridical mean separation of 480 arcmin. The linear fit is to the entire data sets for all three conditions. Error bars in this and subsequent figures represent the *SEM* for each estimate.

patches with Gabor patches but with two very different configurations. In one case, the spatial frequency of the patches was held constant (fixed condition—Figure 1c), while in the other the spatial frequency varied inversely with patch size (variable condition—Figure 1b). The latter condition simulates a situation where relative size and linear perspective cues are created and produces a much stronger sense of depth as a function of patch size and might, therefore, be expected to produce a larger separation misperception. Figure 2 shows that this is not the case—both types of Gabor patch and the Gaussian patches each result in an almost identical misperception of separation. This result demonstrates that it is the envelope representation that forms the basis for the misperception rather than the spatial frequency content of the stimuli. It has previously been shown that object frequency, rather than retinal frequency, is important in certain types of spatial judgment. For example, Burbeck (1987) found that the precision of spatial frequency discrimination, for identical objects placed at different distances from the observer, were very similar suggesting that changes in retinal frequency had a negligible effect on overall performance. Estimates of spatial interval thresholds here were approximately equivalent for each stimulus type and remained relatively stable across the range of sizes tested.

Experiment 2: Bisection acuity

In order to confirm that the previous results were not specific to the chosen task, psychophysical method, or baseline separation, we conducted another experiment using a three-element bisection task in which the horizontal position of a central line probe had to be judged relative to two outer Gaussian patches that differed in size. Stimuli were presented for 500 ms. Three baseline separations (7.5 deg, 15 deg and 30 deg) and four levels of patch size asymmetry (54.4/6.4, 46.4/14.4, 38.4/22.4 and 30.4/30.4 arcmin) were investigated. Within any experimental run, perceived offset was established for reference stimuli of equal but opposite size asymmetry (i.e., one stimulus and its mirror image), and either could occur with

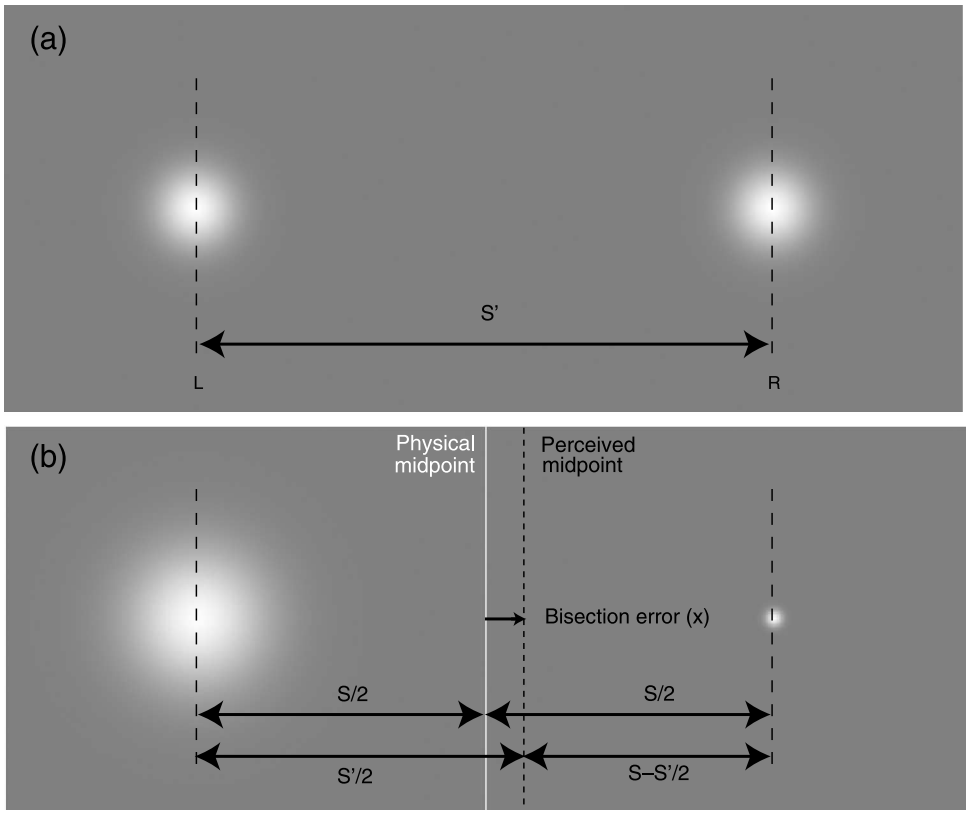

Figure 3. (a) Spatial interval discrimination task: In order for the perceived separation of two Gaussian patches to be maintained at S, their physical separation needs to be set to S', where $S' = S + k(\sigma - \sigma_{mean})$. k is a constant of proportionality, σ is the size of each patch (defined in terms of the standard deviation of the Gaussian envelope), and σ_{mean} is a constant representing the mean size of the stimulus ensemble. A plot of separation against σ should, therefore, reveal a gradient of k. (b) Bisection discrimination: The perceived midpoint of two blobs that differ in size will be displaced from the true physical midpoint such that $x = (S'/2) - (S/2) = ((S/2) + k(\sigma_L - \sigma_{mean})) - (S/2) = k(\sigma_L - \sigma_{mean})$, where σ_L is the size of the left-hand patch and σ_{mean} is the mean size of the two patches. However, the patch size, $\sigma_{mean} = (\sigma_L + \sigma_R)/2$, so that $x = k(\sigma_L/2 - \sigma_R/2) = k/2(\sigma_L - \sigma_R)$. A plot of bisection error against σ_L should, therefore, reveal a gradient of $k/2$—half the gradient found for the spatial interval discrimination task.

equal probability on any trial. The sum of the two patch sizes in the asymmetric and symmetric conditions always remained constant. The central probe was presented at any one of the seven offsets, and the proportion of "rightward" responses was calculated for each offset. The resulting psychometric functions for each patch size asymmetry were again fitted with a logistic function to reveal the point of subjective bisection. A schematic example of the bisection task, with asymmetrically sized stimuli, is presented in Figure 3b. Geometric comparison of this configuration, where the observer judges whether a line bisecting two objects appears closer to the larger or smaller object, suggests that any bisection misperception should vary as a function of Gaussian patch size with a gradient half that of the spatial interval task (Figure 3a).

When plotted against the size of the left-hand patch, the bisection error varied linearly (see Figure 4). As the left patch increased in size, the perceived midpoint shifted away from that element. The gradient for all observers was as predicted, approximately half that of the spatial interval task. This relationship was present for all baseline separations, suggesting that the biases reported in the previous experiment are genuine perceptual phenomena and not peculiar to the comparison of spatial intervals stored in memory. Despite marked changes in perceptual bias, bisection thresholds remained relatively constant as a function of the envelope size difference. Thresholds were elevated as separation was increased, as would be expected (Burbeck, 1987; Whitaker, Bradley, Barret, & McGraw, 2002).

Experiment 3: Controlling for perceived edges

It is conceivable that observers might judge the size of the interval based on the proximity of the nearest edges rather than the distance between the centroids or peak

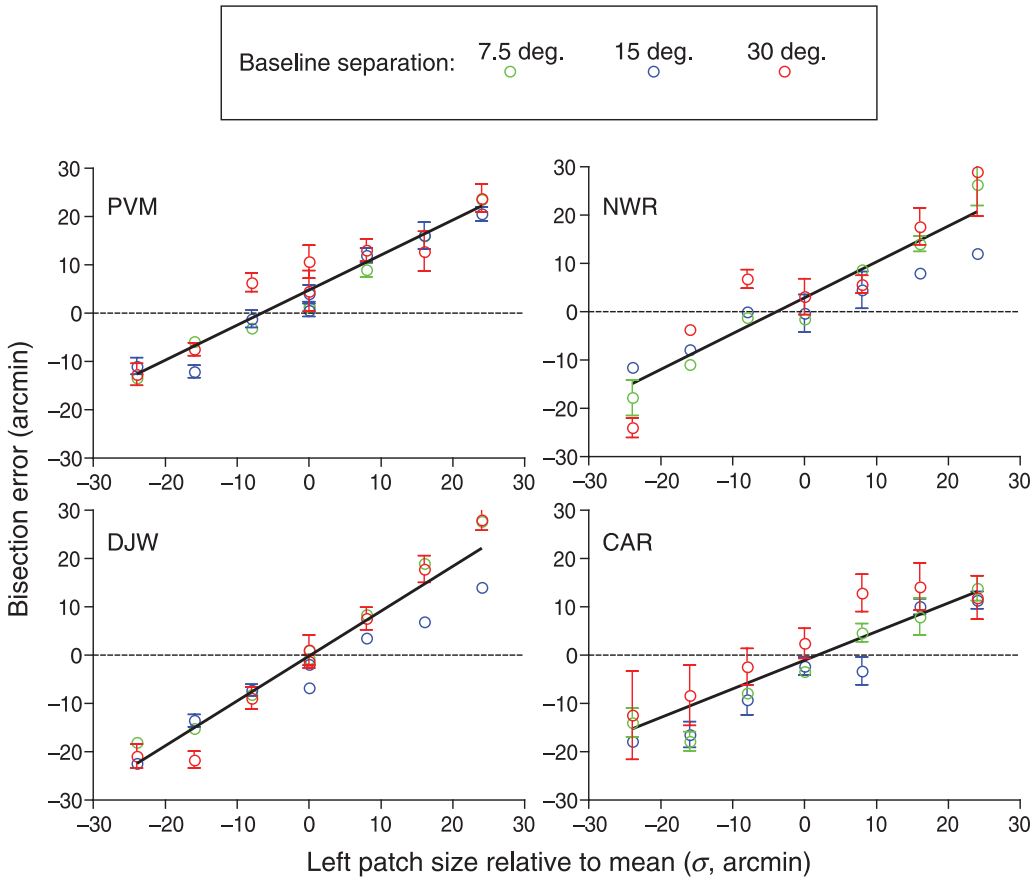

Figure 4. Bisection error is plotted as a function of left-hand blob size for four observers. As the left-hand patch increases in size, the perceived midpoint shifts away from the larger patch—positive bisection errors represent rightward offsets. Data for three baseline separations are presented: 7.5 deg (green circles), 15 deg (blue circles), and 30 deg (red circles). The linear fit is to the combined data sets for all three separations.

luminances. Such a strategy would predict that larger patches would appear closer together and smaller patches further apart—consistent with the previous data. To rule this out, we performed two control experiments.

First, we asked the observers to perform a spatial interval discrimination task between Gaussian patches of equal but opposite asymmetry (see Figure 5a). When the standard deviation either side of the peak is unequal, the centroid of the distribution and its perceived location are shifted in the direction of the larger standard deviation (Whitaker, McGraw, Pacey, & Barrett, 1996). We took the middle patch size from Experiment 1 (27.2 arcmin) and then increased the standard deviation by 26% on one side of the peak and reduced it by the same proportion on the other. Two patch configurations were used, one in which the edges are displaced inward and a horizontally flipped mirror image pair in which the edges were displaced outward. All stimuli were then adjusted in horizontal position so that the distance between the centroids of both configurations was a constant separation (480 arcmin). As before, any one of seven predetermined separations was displayed, and the subject indicated whether the separation between the patches was perceived to be greater or smaller than the mean of all the previous presentations.

In the second experiment, the horizontal standard deviation ($\sigma_{horizontal}$) of the Gaussian patches remained constant (32 arcmin), but we varied the vertical standard deviation ($\sigma_{vertical}$) from 56 arcmin to 8 arcmin (see Figure 5b). Stimuli were presented using a two-interval forced choice method of constant stimuli. On any trial, subjects judged the relative separation between a pair of symmetrical patches (aspect ratio of 1) and a pair in which the vertical standard deviation had been altered (aspect ratios of 1.75, 1, and 0.25). The center-to-center separation of the symmetrical comparison patches was again fixed at 480 arcmin. For each aspect ratio, any one of seven predetermined separations was displayed, and the subject indicated which interval contained the largest separation.

The results of both experiments shown in Figure 5 effectively eliminate any explanations based on the proximity of perceived edges. In the first condition, physically realigning the centroid of each patch resulted in approximately equal (and veridical) estimates of perceived separation, despite the physical manipulation

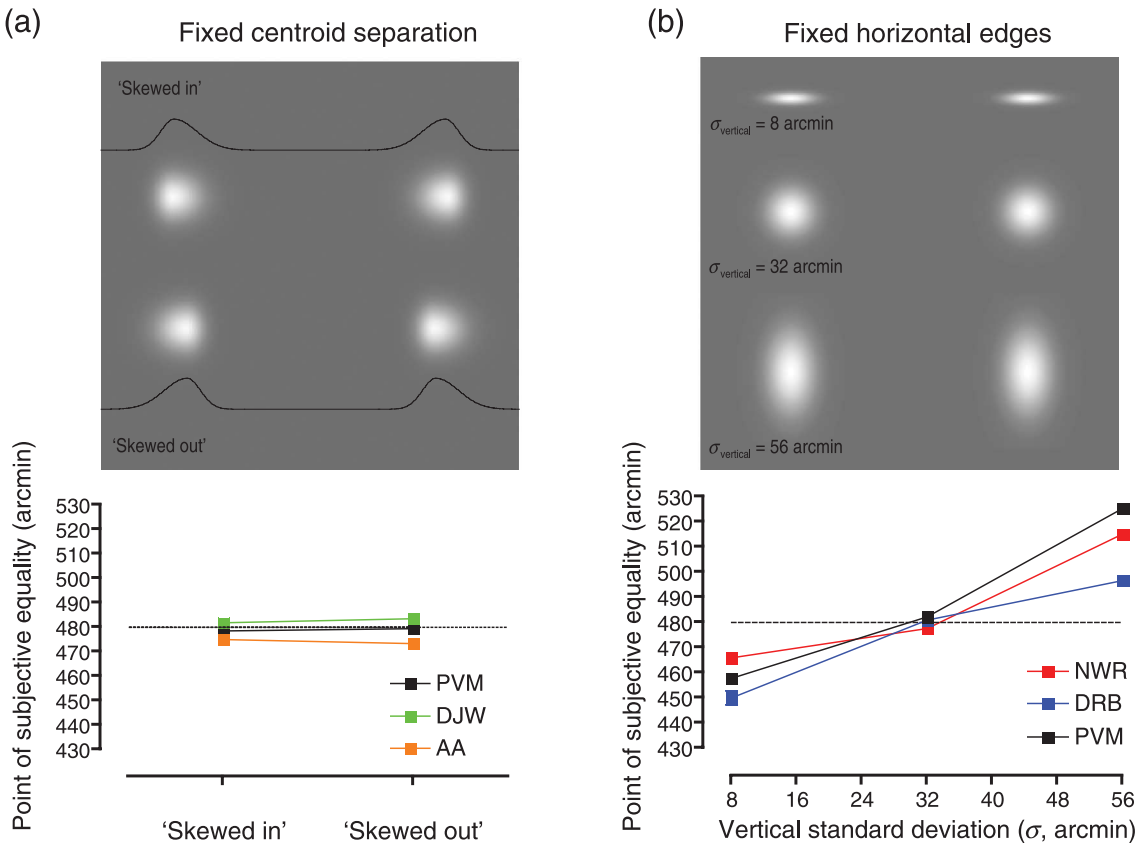

Figure 5. Controlling for perceived edge position. (a) Skewing the luminance profile of stimulus patches in opposite directions while maintaining their centroid locations results in no change in perceived separation. (b) Manipulating the vertical extent of the pairs of patches results in a systematic change in perceived separation despite no change in the horizontal edge position.

of horizontal edge position. Second, when the horizontal extent of the patch is fixed, but its vertical dimension varied, perceived separation is reliably modulated in all three subjects despite the fact that perceived horizontal edge position remains unchanged in each configuration. It again remains possible that spatial judgments are contaminated by inappropriate constancy scaling (Gregory, 1963), operating on the average size of the target rather than target size along the dimension relevant to the judgment (horizontal).

Experiment 4: Spatial interval discrimination judgments with contrast-defined stimuli

The visual system is also adept at localizing objects defined by variations in contrast or texture (Badcock & Derrington, 1985; Derrington, Badcock, & Henning, 1993). Contrast-defined stimuli are commonly composed of a low-frequency envelope (object) superimposed onto a "carrier" pattern of higher spatial frequency and are thought to be encoded by specialized filters (Mareschal & Baker, 1998; Zhou & Baker, 1993). In this experiment, we investigate the influence of stimulus size (envelope) and background spatial frequency (carrier) on the perceived separation of contrast-defined stimuli.

Observers were presented with two horizontally separated Gaussian-windowed contrast modulations. The mathematical description of the stimuli is given by

$$L_{\text{mean}} + [(L_{\text{mean}}\sin((2\pi Nx/\sigma) + \phi))/2) * (\text{contrast} - \text{contrast} * \exp(-(x^2 + y^2)/2\sigma^2)], \quad (4)$$

where L_{mean} is the mean luminance of the background, σ is the standard deviation of the Gaussian envelope, N is the number of cycles of carrier grating per standard deviation (spatial frequency), and ϕ represents the phase of the carrier grating. The vertical and horizontal distances from the peak of the Gaussian envelope are denoted by x and y (see Figure 6).

In the first part of this experiment, the carrier spatial frequency was fixed at 2 c deg^{-1} with a Michelson contrast of 0.5. To avoid the formation of retinal afterimages, the background carrier was contrast reversed (5 Hz). Gaussian-windowed contrast modulations (decrements with

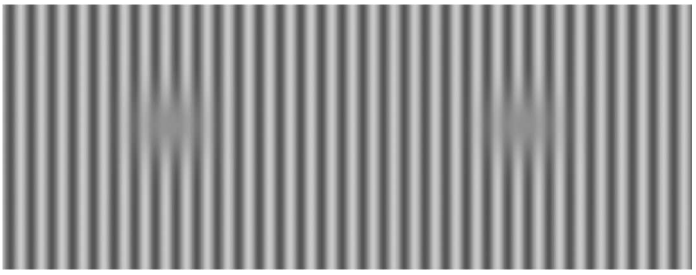

Figure 6. Example of the contrast-defined patches used in Experiment 4. Gaussian-windowed contrast modulations (decrements with zero contrast at peak) were presented at a range of sizes (σ = 3.2–51.2 arcmin), and to avoid the formation of retinal afterimages, the background carrier was contrast reversed (5 Hz). Both horizontal and vertical carrier orientations were investigated.

zero contrast at peak) were presented at a range of sizes (σ = 3.2–51.2 arcmin). The experiment then followed the protocol of Experiment 1, where on any single presentation, one of seven predetermined separations and seven patch sizes was displayed. Following each presentation, the subject was required to indicate whether the separation between the patches was perceived to be greater or smaller than the mean of all the previous presentations. The baseline separation was again 480 arcmin and both horizontal and vertical carrier orientations were investigated.

In a separate experiment, we examined the influence of changing carrier spatial frequency (background) on the perceived separation of targets with a fixed envelope size (27.2 arcmin). For each observer, the detection threshold of individual carrier frequencies (0.5, 1, 2, 4, and 8 c deg^{-1}) was established using a 2AFC method of constant stimuli. All carrier frequencies were then adjusted to a contrast level that was 60 times their respective detection threshold level, ensuring that each was equated in terms of distance above threshold. Then following the same procedures as the previous experiment, we measured changes in perceived separation that result from systematically manipulating the scale of the background.

When background spatial frequency is fixed and the size of the contrast envelope systematically varied, the results are remarkably different to those found for luminance-defined stimuli. As shown in Figure 7, large changes in envelope size have little or no influence on perceived separation and this is observed for both vertical and

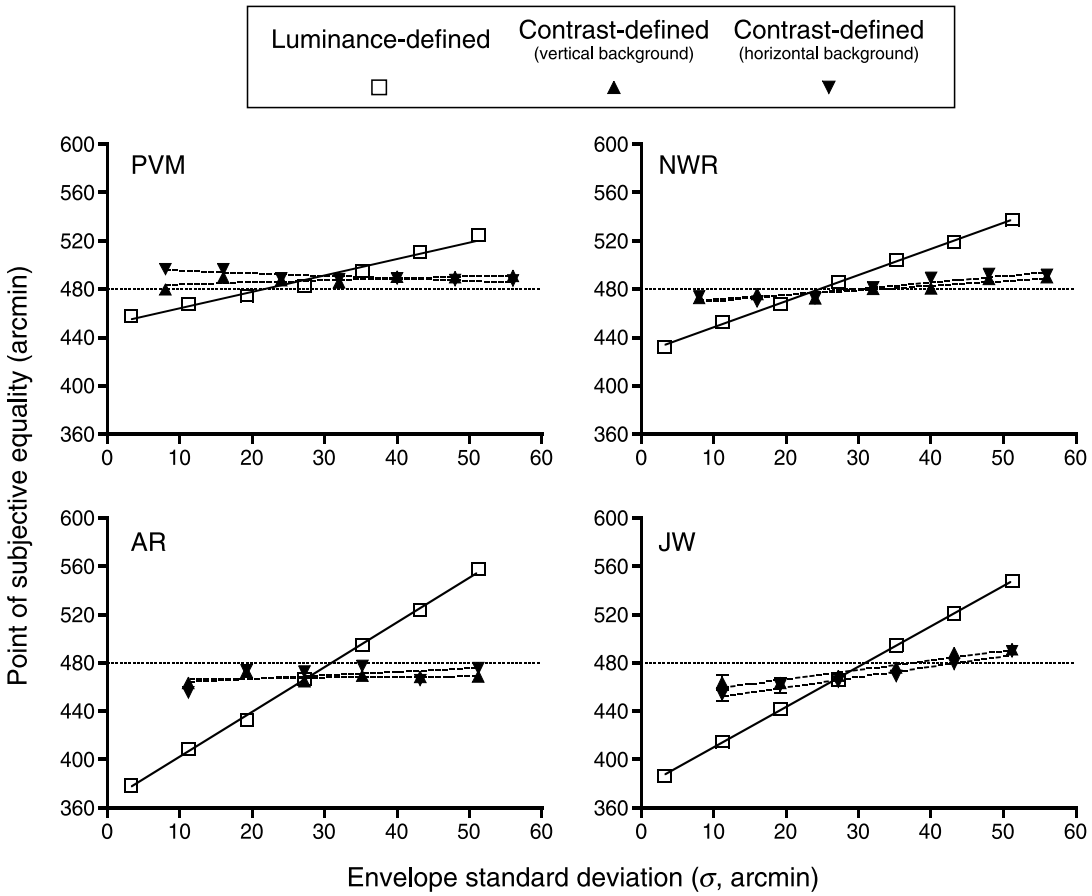

Figure 7. The physical separation of the two patches required for the stimuli to appear equal to the mean separation as a function of patch size for four observers. Data from three stimulus conditions are presented: luminance-defined Gaussian patches (open squares) and contrast-defined patches with either a vertical (upright triangle) or horizontal (inverted triangle) background carrier.

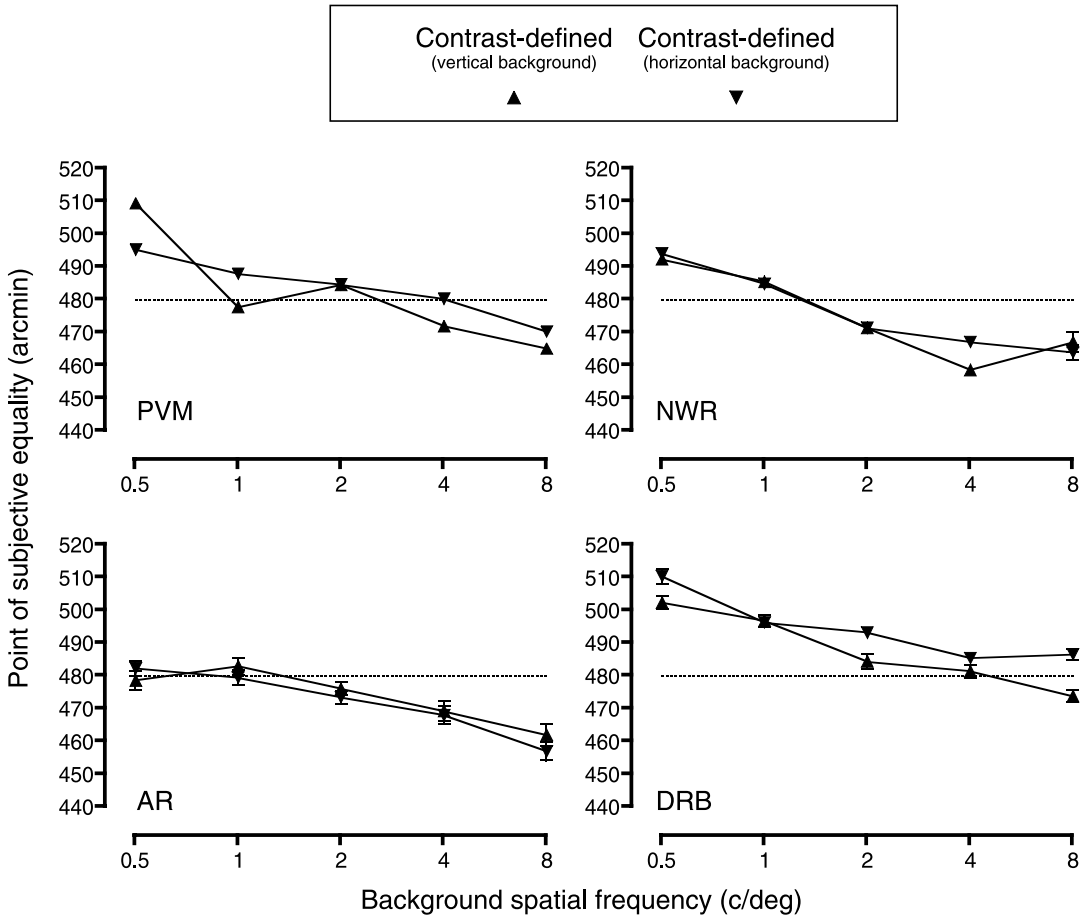

Figure 8. The physical separation of the two patches required for the stimuli to appear equal to the mean separation as a function of background spatial frequency for four observers. Data from two different stimulus conditions are presented: contrast-defined patches with either a vertical (upright triangle) or horizontal (inverted triangle) background carrier.

horizontal background carriers. Perception of the spatial interval itself was largely veridical, with estimates falling close to the mean separation (indicated by the dashed line). Alignment thresholds, like their luminance-based counterparts, were unaffected by changes in stimulus size for either carrier orientation.

Next, we varied the spatial frequency and orientation of the background carrier while keeping the size of the contrast envelope fixed. This prevented subjects from simply counting the cycles or using the variable number of cycles between elements as a cue. In marked contrast to the results presented above, we found systematic changes in perceived separation as the background spatial frequency was manipulated (Figure 8). The perceived spatial interval changed by about 20–30 arcmin across different background frequencies. This change, while smaller than that found for luminance-defined stimuli, is still large relative to the error associated with each measure of the interval and is likely to be constrained by the range of spatial frequencies investigated. For two of the observers (AR and DRB), a small bias is evident but opposite in direction in each case, which is likely to be attributable to individual differences (imbalances) in the formation of a mean internal representation. Spatial interval thresholds, on the other hand, showed a shallow but systematic change with manipulations of background spatial frequency, despite the fact that they were equated in terms of multiples of contrast threshold. Specifically, thresholds were best for the lowest frequency carrier (0.5 c deg^{-1}) and the change in sensitivity was well described by a power function with an exponent of around 0.15 (horizontal carrier = 0.162, R = 0.839; vertical carrier = 0.143, R = 0.944).

Discussion

The results of the present study show that for luminance-defined visual objects, the physical size of the stimuli exerts a marked influence on their perceived separation. This is a rather surprising outcome given our current understanding of spatial coding and retinotopic mapping.

The receptive fields of neurons in early parts of the visual pathway vary dramatically in size and typically

show a band-pass response to different spatial frequencies (De Valois, Albrecht, & Thorell, 1982; Movshon, Thompson, & Tollhurst, 1978). This has led to the initial stage of visual coding being characterized as a multi-scaled bank of spatial frequency filters that are generally well matched to the image statistics of the visual environment (Field, 1987). In order to localize objects in visual space, or to make other spatial judgments for that matter, the visual system must extract the most informative features from the initial pattern of stimulation. There are a number of potential attributes that can be extracted and a range of schemes that describe the mechanisms of feature extraction (see Burr & Morrone, 1994; Georgeson, 1992; Marr & Hildreth, 1980; Watt & Morgan, 1985). For example, the edges of an object, its luminance peak, or the center of mass (centroid) of its luminance distribution are considered to be key features (Watt & Morgan, 1983; Whitaker et al., 1996). To explain feature extraction, models that are capable of accurately characterizing edge location for a broad range of luminance profiles have been developed based on Gaussian scale–space theory (Georgeson & Freeman, 1997; Georgeson, May, Freeman, & Hesse, 2007). The same general principles can be used to find the peak or centroid of the image, although a different combination of Gaussian derivative filters would be required. Importantly, in all cases, the output of filtering operations of this type is unbiased. That is, the perceived edges and peak will closely match the physical properties of the image. Our results show, however, that the representation of space can be modified without changing either centroid (or luminance peak) position or edge position. This suggests that the marked changes in the representation of visual space that we report must happen after basic feature extraction has occurred. Consistent with this idea, a number of position coding schemes have been developed that employ a set of coincidence detectors, which provide a secondary representation that encodes the activity pattern of early filter-based processes in a way that retains information on the distance between features (Burbeck, 1987; Klein & Levi, 1985; Morgan, Hole et al., 1990; Morgan & Regan, 1987; Morgan, Ward et al., 1990).

Assuming that the early representation of features is unbiased, another possibility is that contextual modulation, from higher visual areas, regulates the representation of space. For example, a commonly advanced explanation of the Baldwin illusion, where the perceived length of a line is distorted by the addition of flanking boxes, is that the "framing ratio" (the size of the entire figure relative to the size of the focal stimuli) is predictive of the spatial distortion of line length (Brigell, Uhlarik, & Goldhorn, 1977; Kunnapas, 1955). Distortions from the point of objective equality (i.e., perceived line length versus physical length) as a function of box size—an analogue of our changes in Gaussian envelope size—has been shown to increase initially and then gradually reduce to baseline (no box condition) with further increases in size. Importantly, when line length is changed, the contextual width must be altered accordingly to produce a constant framing ratio and thus maintain the pattern of distortion. Other judgments of linear extent, such as the divided line and Muller–Lyer illusions, show a very similar pattern of distortion and are also thought to represent situations where the proximal extent of focal stimuli and the overall width of the contextual figure interact (Brigell et al., 1977). A simple channel model has been proposed to account for the pattern of data in the Baldwin illusion, where the focal stimulus (line) and global stimulus (total figure) activate overlapping populations of length-selective units. The misperceptions are caused by the interaction between these distributions: A maximal effect is found for a particular ratio of focal to contextual length (~3:2) that declines as the offset between the population responses to each element increases or decreases (Brigell et al., 1977). While the misperceptions we report are in qualitative agreement with some general aspects of the Baldwin illusion, e.g., the perceptual expansion of space closer to smaller visual stimuli, our results are clearly inconsistent with a scheme in which the framing ratio determines the magnitude of distortion. The data we report for a fixed baseline separation did not show a non-monotonic relationship between stimulus size and the magnitude of perceptual distortion. Further, when we changed baseline separation (Experiment 2), a manipulation that markedly alters the framing ratio of the stimulus, the spatial error remained constant. Finally, changes in vertical size (Experiment 3) of the stimulus altered perceived separation, yet this manipulation is known to have no effect on the Baldwin illusion (Brigell et al., 1977).

More recent evidence has shown that retinotopic mapping can be dramatically altered by changes in perceived depth (Murray, Boyaci, & Kersten, 2006). Using functional magnetic resonance imaging (fMRI), Murray et al. (2006) showed that two objects, each subtending the same visual angle, activate quite different volumes in early visual cortex (V1) when their perceived distance is changed. Specifically, objects that are made to appear more distant activate larger areas of V1 than identical objects that appear closer. This suggests that depth and size information are combined at the level of V1 and that the influence of viewing context can act on relatively early stages of visual processing. Similarly, a more recent study has reported that objects of different size can induce a bias in perceived disparity-defined distance (Lugtiheid & Welchman, 2010). The authors suggest that this effect is mediated by probabilistic mapping between retinal image size and object distance (i.e., objects that subtend a smaller visual angle are perceived as being further away and vice versa). Critically however, the influence of retinal size on estimates of distance is not found when presenting two objects of equal size. Therefore, this effect cannot explain the results of our spatial interval task, where the elements defining the interval were always identical in size. Moreover, when the stimuli were configured in a way that simulates changes in

perceived distance (spatial frequency changing inversely with patch size), the degree of spatial distortion was identical to that found for stimuli that were not simply magnified versions of each other. Therefore, perceived distance or the influence of retinal image size on apparent proximity is probably unrelated to the effects we report. That said, at this point, we cannot completely rule out the possibility that the visual system is inappropriately using size constancy mechanisms (Gregory, 1963) when making spatial interval judgments.

Burbeck and Hadden (1993) reported that estimates of spatial interval were biased by placing distracter elements either side of the stimuli delimiting the interval. The magnitude of the induced distortion was dependent on the distance between element and distracter and present for distances at least as large as the target interval. They interpreted this finding as evidence for an increasing integration zone for positional information as object separation is increased and suggested that this process underpinned Weber's law for separation, where separation discrimination thresholds increase in proportion to baseline separation. This issue was revisited by Hess and Badcock (1995), who demonstrated that both over- and underestimations of the interval was possible, the direction of the effect dictated by the relative proximity of the distracter. If the distracter-element distance was shorter than the target spatial interval, then the interval itself was perceptually overestimated and vice versa. The changes in perceived separation induced by distracter elements were dependent on the similarity between target and distracter—not present in situations where they differed in size or carrier spatial frequency. To explain this effect, the authors suggested that the spatial interval to be judged is implicitly compared to the distracter-element interval, where the scale of the former influences the latter.

Could the same reasoning be used to explain the spatial distortions we report? Where comparison intervals are not available, it is conceivable that the perceived distance between stimuli is judged relative to the size of the objects themselves or the underlying filter required for recovering their spatial structure. In a spatial interval task, the first goal of the visual system is to recover the contrast envelopes of the Gaussian-windowed patches. Following envelope recovery and localization, the system can then estimate the spatial distance between the patches. If the scale of visual analysis is determined by the size of filter (or underlying receptive field) required for veridical envelope recovery, this may dominate in the absence of other structure in the visual field against which space can be calibrated.

In the case of second-order stimuli, filter scale may be selected relative to the background carrier spatial frequency (or texture). The textural elements that define a second-order image are initially analyzed by linear filters tuned to a relatively high spatial frequency. The output of these filters is then subjected to a non-linear transform, such as rectification, and then conveyed to second-stage linear filters tuned to a coarser spatial frequency. This filter–rectify–filter cascade allows the visual system to detect boundaries defined by contrast but not by luminance (Baker & Mareschal, 2001). The results of the present experiments suggests that for spatial interval judgments, where the interval is defined by changes in contrast rather than luminance, the scale of analysis is set by the early filter in this processing mechanism, since changes in the size of the contrast-defined patch have little or no influence on perceived separation.

Early attempts to model the extremely high levels of positional sensitivity shown by human observers explored the possibility of using a distributed set of orientation- and frequency-selective filters (Carlson & Klopfenstein, 1985; Klein & Levi, 1985; Sullivan, Oatley, & Sutherland, 1972; Wilson & Gelb, 1984). Although this general class of filter distribution model varies in their implementation, a common assumption is that the spatial relationship between objects is encoded by a filter large enough to encompass both simultaneously. The advantage of such schemes is that the relative position is encoded without the need for information about the receptive field position. However, the requirement for filters large enough to span both objects has meant that this approach essentially fails for separations larger than about a degree. Moreover, this class of model cannot account for several other aspects of behavioral data such as the influence of flanking targets or contrast polarity effects (Burbeck, 1987; Morgan & Ward, 1985). These problems led to the development of hybrid filter-local sign models, where a filter-based approach provides an initial representation and a second stage of analysis computes the distance between features (e.g., Burbeck, 1987). Here, local sign information provides retinotopic information (Burbeck, 1987) or guides the intelligent selection of filters (Morgan, Hole et al., 1990; Morgan, Ward et al., 1990). More recent population coding models of visual space have adopted an implementation where stimulus locations are represented intrinsically rather than via an external coordinate frame of reference, such as a retinotopic map (Lehky & Sereno, 2011). An intrinsic coding scheme, where only relative position is important, has the advantage that visual representations are largely invariant of changes in scale or viewing position, transformations that the visual system must commonly accommodate. However, purely intrinsic schemes present a considerable challenge for the visual guidance of motor commands, where eye or hand movements must be directed toward specific physical locations. Intrinsic spatial coding models predict that populations with only relatively small receptive fields available, such as those typically found in striate cortex (V1), result in markedly distorted spatial representations (Lehky & Sereno, 2011). While most of the data generated from intrinsic coding models have been derived from single stimulus presentations, some model predictions have been

produced for situations in which two stimuli are presented either one after the other or simultaneously in visual space. Interestingly, the types of distortion that result from simultaneous presentation are qualitatively in agreement with previous psychophysical observations (e.g., Badcock & Westheimer, 1985). Whether intrinsic coding schemes can be adapted to predict the size-related distortions of visual space we report here remains a future computational challenge.

The distortions of visual space induced by objects of different size may also provide a potential explanation for other paradoxical visual effects. For example, Brown (1931) showed the perceived velocity of an object is dependent on its size, with smaller objects seeming to speed up while larger objects appeared to slow down. Such an effect is qualitatively predictable from the perceptual distortions we report. If visual space surrounding a large object is perceptually compressed, this object will appear to travel over a shorter distance in a given time interval, i.e., it will appear to slow down. Previous reports have also shown that perceived velocity can be altered by changing the spatial frequency content of the background over which stimuli move (Blakemore & Snowden, 2000). However, it remains to be seen whether the distortions of visual space that arise from simply changing target size or background spatial frequency are directly predictive of the perceived speed of a moving object: This is currently the focus of ongoing work in our laboratory.

Acknowledgments

This work was supported by grants from the Wellcome Trust and the Australian Research Council (DP1097003 & DP110104553).

Commercial relationships: none.
Corresponding author: Paul V. McGraw.
Email: pvm@psychology.nottingham.ac.uk.
Address: Visual Neuroscience Group, School of Psychology, The University of Nottingham, University Park, Nottingham NG5 2RD, UK.

References

Anstis, S. M., & Ramachandran, V. S. (1990). Illusory displacement of kinetic edges. *Perception, 19,* 611–616. [PubMed]

Badcock, D. R., & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research, 25,* 1253–1258. [PubMed]

Badcock, D. R., & Westheimer, G. (1985). Spatial location and hyperacuity: The centre/surround localization contribution function has two substrates. *Vision Research, 25,* 1259–1267. [PubMed]

Badcock, D. R., & Wong, T. (1990). The sensitivity of separation discrimination to spatiotemporal jitter. *Vision Research, 30,* 1555–1560. [PubMed]

Baker, C. L., & Mareschal, I. (2001). Processing of second-order stimuli in the visual cortex. *Progress in Brain Research, 134,* 171–191. [PubMed]

Baldwin, J. M. (1895). The effect of size-contrast upon judgments of position in the retinal field. *Psychological Review, 2,* 244–259.

Blakemore, M. R., & Snowden, R. J. (2000). Textured backgrounds alter perceived speed. *Vision Research, 40,* 629–638. [PubMed]

Brigell, M., Uhlarik, J., & Goldhorn, P. (1977). Contextual influences on judgements of linear extent. *Journal of Experimental Psychology, 3,* 105–118. [PubMed]

Brown, J. F. (1931). The visual perception of velocity. *Psychological Research, 14,* 199–232.

Burbeck, C. A. (1987). Position and spatial frequency in large-scale localization judgements. *Vision Research, 27,* 417–427. [PubMed]

Burbeck, C. A., & Hadden, S. (1993). Scaled position integration areas: Accounting for Weber's law for separation. *Journal of the Optical Society of America A, 10,* 5–14.

Burr, D. C., & Morrone, M. C. (1994). The role of features in structuring visual images. In G. R. Bock & J. A. Goode (Eds.), *Higher order processing in the visual system, Ciba Foundation Symposium 184* (pp. 129–146). Chichester, UK: Wiley. [PubMed]

Carlson, C. R., & Klopfenstein, R. W. (1985). Spatial frequency model for hyperacuity. *Journal of the Optical Society of America, 2,* 1747–1751.

Daniel, P. M., & Whitteridge, D. (1961). The representation of the visual field in the cerebral cortex in monkeys. *The Journal of Physiology, 159,* 203–221. [PubMed]

Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research, 33,* 1785–1794. [PubMed]

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22,* 545–559. [PubMed]

De Valois, R. L., & De Valois, K. K. (1991). Vernier acuity with stationary moving Gabors. *Vision Research, 31,* 1619–1626. [PubMed]

Edwards, M., & Badcock, D. R. (2003). Motion distorts perceived depth. *Vision Research, 43,* 1799–1804. [PubMed]

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, 4,* 2379–2394. [PubMed]

Georgeson, M. A. (1992). Human vision combines oriented filters to compute edges. *Proceedings of the Royal Society B, 249,* 235–245. [PubMed]

Georgeson, M. A., & Freeman, T. C. (1997). Perceived location of bars and edges in one-dimensional images: Computational models and human vision. *Vision Research, 37,* 127–142. [PubMed]

Georgeson, M. A., May, K. A., Freeman, T. C., & Hesse, G. S. (2007). From filters to features: Scale–space analysis of edge and blur coding in human vision. *Journal of Vision, 7*(13):7, 1–21, http://www.journalofvision.org/content/7/13/7, doi:10.1167/7.13.7. [PubMed] [Article]

Gregory, R. L. (1963). Distortion of visual space as inappropriate constancy scaling. *Nature, 204,* 678–680. [PubMed]

Gregory, R. L. (1968). Visual illusions. *Scientific American, 219,* 66–67. [PubMed]

Hess, R. F., & Badcock, D. R. (1995). Metric for separation discrimination by the human visual system. *Journal of the Optical Society of America A, 12,* 3–16. [PubMed]

Honda, H. (1989), Perceptual localization of visual stimuli flashed during saccades. *Perception & Psychophysics, 46,* 162–174. [PubMed]

Klein, S. A., & Levi, D. M. (1985). Hyperacuity thresholds of 1 sec: Theoretical predictions and empirical validation. *Journal of the Optical Society of America A, 7,* 1170–1190. [PubMed]

Kooi, F. L., De Valois, R. L., & Switkes, E. (1991). Spatial localization across channels. *Vision Research, 31,* 1627–1631. [PubMed]

Kunnapas, T. M. (1955). Influence of frame size on apparent length of a line. *Journal of Experimental Psychology, 50,* 168–170. [PubMed]

Lehky, S. R., & Sereno, A. B. (2011). Population coding of visual space: Modeling. *Frontiers in Computational Neuroscience, 4,* 1–25. [PubMed]

Levi, D. M., & Klein, S. A. (1992). "Webers law" for position: The role of spatial frequency and contrast. *Vision Research, 32,* 2235–2250. [PubMed]

Levi, D. M., Klein, S. A., & Yap, Y. L. (1988). "Webers law" for position: Unconfounding the role of separation and eccentricity. *Vision Research, 28,* 597–603. [PubMed]

Lotze, R. H. (1852). *Medicinische Psychologie Oder Physiologie Der Seele*. Leipzig, Germany: Weidmann'sche Buchhandlung.

Lugtiheid, A., & Welchman, A. (2010). A surprising influence of retinal size on disparity-defined judgments [Abstract]. *Journal of Vision, 10*(7):63, 63a, http://www.journalofvision.org/content/10/7/63, doi:10.1167/10.7.63.

Mareschal, I., & Baker, C. L. (1998). Temporal and spatial response to second-order stimuli in cat area 18. *Journal of Neurophysiology, 80,* 2811–2823. [PubMed]

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B, 207,* 187–217.

Morgan, M. J. (1992). On the scaling of size judgements by orientational cues. *Vision Research, 32,* 1433–1445. [PubMed]

Morgan, M. J. (2009). *Moleyneux's question—Vision, touch and the philosophy of perception*. Cambridge, UK: Cambridge University Press.

Morgan, M. J., Hole, G. J., & Glennester, A. (1990). Biases and sensitivities in geometrical illusions. *Vision Research, 30,* 1793–1810. [PubMed]

Morgan, M. J., & Regan, D. M. (1987). Opponent model for line interval discrimination: Interval and vernier performance compared. *Vision Research, 27,* 107–118. [PubMed]

Morgan, M. J., & Ward, R. M. (1985). Spatial and spatial-frequency primitives in spatial-interval discrimination. *Journal of the Optical Society of America A, 2,* 1205–1210. [PubMed]

Morgan, M. J., Ward, R. M., & Hole, G. J. (1990). Evidence for position coding in hyperacuity. *Journal of the Optical Society of America A, 7,* 297–304. [PubMed]

Morgan, M. J., Watamaniuk, S. N. J., & McKee, S. P. (2000). The use of an implicit standard for measuring discrimination thresholds. *Vision Research, 40,* 2341–2349. [PubMed]

Movshon, J. A., Thompson, I. A., & Tollhurst, D. J. (1978). Spatial summation in the receptive fields of simple cells in the cat's striate cortex. The *Journal of Physiology, 283,* 53–77. [PubMed]

Murray, S. O., Boyaci, H., & Kersten, D. (2006). The representation of perceived angular size in human primary visual cortex. *Nature Neuroscience, 9,* 429–434. [PubMed]

Nishida, S., & Johnston, A. (1999). Influence of motion signals on the perceived position of spatial pattern. *Nature, 397,* 610–612. [PubMed]

Regan, D., & Beverly, K. I. (1983). Spatial-frequency discrimination and detection: Comparison of post adaptation thresholds. *Journal of the Optical Society of America A, 73,* 1684–1690. [PubMed]

Roach, N. W., & McGraw, P. V. (2009). Dynamics of spatial distortions reveal multiple time scales of motion adaptation. *Journal of Neurophysiology, 102,* 3619–3626. [PubMed]

Robinson, J. O. (1972). *The psychology of visual illusions*. London: Hutchinson.

Ross, J., Morrone, C. M., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature, 386,* 598–601. [PubMed]

Ross, J., Morrone, C. M., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences, 24,* 113–121. [PubMed]

Sanchez-Panchuelo, R. M., Francis, S., Bowtell, R., & Schluppeck, D. (2010). Mapping human somatosensory cortex in individual subjects with 7T functional MRI. *Journal of Neurophysiology, 105,* 3042–3053. [PubMed]

Schwartz, E. L. (1980). Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. *Vision Research, 20,* 645–669. [PubMed]

Sullivan, G. D., Oatley, K., & Sutherland, N. S. (1972). Vernier acuity as affected by target length and separation. *Perception & Psychophysics, 12,* 438–444.

Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: Evidence for spatial primitives in human vision. *Vision Research, 23,* 97–109. [PubMed]

Watt, R. J., & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research, 25,* 1661–1674. [PubMed]

Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology and Visual Science, 14,* 570–572.

Whitaker, D., Bradley, A., Barret, B. T., & McGraw, P. V. (2002). Isolation of stimulus characteristics contributing to Weber's law for position. *Vision Research, 42,* 1137–1148. [PubMed]

Whitaker, D., McGraw, P. V., Pacey, I., & Barrett, B. T. (1996). Centroid analysis predicts the visual localization of first- and second-order stimuli. *Vision Research, 36,* 2957–2970. [PubMed]

Whitaker, D., McGraw, P. V., & Pearson, S. (1999). Non-veridical size perception of expanding and contracting objects. *Vision Research, 39,* 2999–3009. [PubMed]

Whitney, D., & Cavanagh, P. (2000). Motion distorts visual space: Shifting the perceived position of remote stationary objects. *Nature Neuroscience, 3,* 954–959. [PubMed]

Wilson, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research, 26,* 453–469. [PubMed]

Wilson, H. R., & Gelb, D. J. (1984). Modified line-element theory for spatial frequency and width discrimination. *Journal of the Optical Society of America A, 1,* 124–131. [PubMed]

Zhou, Y.-X., & Baker, C. L. (1993). A processing stream in mammalian visual cortex neurons for non-Fourier responses. *Science, 261,* 98–101. [PubMed]