Neither here nor there: localizing conflicting visual attributes

Paul V. McGraw   Department of Optometry, University of Bradford, Bradford BD7 1DP, West Yorks, UK

David Whitaker   Department of Optometry, University of Bradford, Bradford BD7 1DP, West Yorks, UK

David R. Badcock   School of Psychology, University of Western Australia, Nedlands, WA 6009, Australia

Jennifer Skillen   Department of Optometry, University of Bradford, Bradford BD7 1DP, West Yorks, UK

Natural visual scenes are a rich source of information. Objects often carry luminance, colour, motion, depth and textural cues, each of which can serve to aid detection and localization of the object within a scene. Contemporary neuroscience presumes a modular approach to visual analysis in which each of these attributes are processed within ostensibly independent visual streams and are transmitted to geographically distinct and functionally dedicated centres in visual cortex (van Essen & Maunsell, 1983; Zihl, von Cramon & Mai, 1983; Maunsell & Newsome, 1987; Tootell, Hadjikhani, Mendola, Marrett & Dale, 1998). In the present study we ask how the visual system localizes objects within this framework. Specifically, we investigate how the visual system assigns a unitary location to objects defined by multiple stimulus attributes, where such attributes provide conflicting positional cues. The results show that conflicting sources of visual information can be effortlessly combined to form a global estimate of spatial position, yet, this conflation of visual attributes is achieved at a cost to localization accuracy. Furthermore, our results suggest that the visual system assigns more perceptual weight (Landy, 1993; Landy & Kojima, 2001) to visual attributes which are reliably related to object contours.

Introduction

Natural visual scenes contain an abundance of cues that can be used to perform visual tasks such as object detection, object discrimination or localization of one object relative to another. For example, cues such as luminance, colour, disparity, texture and motion information may all be used by the visual system to effectuate visual tasks. There is now compelling evidence for the existence of separate visual processing streams and functionally specialised cortical areas (van Essen & Maunsell, 1983; Zihl, von Cramon & Mai, 1983; Maunsell & Newsome, 1987). Each visual stream is thought to be involved in the analysis of a particular sensory cue (e.g. luminance, texture, motion or colour) which in turn contributes to a particular aspect of our everyday perceptual experience. Cumulative evidence for this modular framework to visual analysis is derived from studies which examine lesion-induced deficits, response properties of neuronal populations, and the architecture of anatomical connections within the cortex (Pearlman, Birch & Meadows, 1979; Damasio, A., et al. 1980; Zihl et al., 1983; Zeki & Shipp, 1988; Desimone & Ungerleider,

1989; De Yoe et al., 1990). Striking examples of this cortical "division of labour" result when localized areas of the human cerebral cortex suffer bilateral damage. For instance, individuals can suffer a complete loss in sensitivity for a particular visual attribute such as colour (Damasio et al., 1980) or motion (Zihl et al., 1983), yet show little or no deficit in processing other types of visual information.

In addition to the major sub-divisions outlined above, other examples of functional subdivisions within and between early cortical areas (V1 & V2) also exist. Recently, both physiological (Zhou & Baker, 1993; Mareschal & Baker, 1998), and psychophysical (Badcock & Derrington, 1985; Chubb & Sperling, 1988; Derrington, Badcock & Henning, 1993; Ledgeway & Smith, 1994; Li-Ming & Wilson, 1996; Whitaker, McGraw & Levi, 1997; McGraw, Levi & Whitaker, 1999) investigations of early visual coding have focused on how the visual system analyses both luminance-defined and contrast-defined image components using linear and non-linear processes respectively. Neurophysiological investigations have shown that the visual system contains many neurons which signal differences in average

luminance between the excitatory and inhibitory sub-regions of their receptive field in a linear manner. However, such linear neurons are not ideally suited to the analysis of texture- or contrast-defined stimuli where the spatial extent of luminance variations can be small relative to the receptive field size. In this situation, linear summation of luminance increments and decrements across the extent of the receptive field may produce no net variation in luminance relative to the surround. For this reason large linear cortical neurons are unable to signal the presence of such texture-defined visual stimuli. In order for neurons to be able to detect image features such as contrast variations, the outputs of relatively small scale initial filters must be subjected to a form of non-linearity (such as rectification) before they become amenable to conventional linear processing. Contemporary models of non-linear visual processing are therefore based upon an initial linear filtering stage, followed by a non-linear step (rectification stage), and subsequent linear filtering at a relatively coarse spatial scale. This processing cascade has previously been identified in the striate cortex of cats (Zhou & Baker, 1993) and psychophysically in humans (Wilson, Ferrera & Yo, 1992; Graham, Beck & Sutter, 1992).

The present study examines how the visual system localizes stimuli that are composed of both luminance and textural information. A combination of luminance and texture was chosen for two reasons. Firstly, this combination is one of the most commonly encountered in our visual environment. Secondly, there now exists convincing physiological (Olavarria et al., 1992; Zhou & Baker, 1993: Mareschal & Baker, 1998; see also Shapley, 1994) and psychophysical evidence (Badcock & Derrington, 1985; Chubb & Sperling, 1988; Derrington, Badcock & Henning, 1993; Ledgeway & Smith, 1994; Li-Ming & Wilson, 1996; Whitaker, McGraw & Levi, 1997; McGraw, Levi & Whitaker, 1999; Badcock & Khuu, 2001) to suggest that each attribute is processed by a dedicated cortical stream. Objects defined by variations in luminance are detected by linear neurons located in the primary visual cortex (V1) (Hubel & Wiesel, 1962; Movshon, Thompson & Tolhurst, 1978). Texture, or contrast-defined objects, constructed from balanced increments and decrements in luminance are invisible to linear V1 neurons, and are recovered via a non-linear operation carried out by a dedicated neural population located in V1 and V2 (Zhou & Baker, 1993; Mareschal & Baker, 1998; Mareschal & Baker, 1999). These neurons, which differ in their response output to image features, provide the physiological framework for the luminance and texture processing streams in the human visual system. We ask how the visual system localizes objects within a framework of seemingly autonomous visual processing streams. Specifically, we investigate how the visual system assigns a unitary location to objects defined by multiple stimulus attributes, where such attributes provide conflicting positional cues. Issues related to this question have been examined previously. For example, Rivest and Cavanagh (1996) determined how the precision of visual localization changes when multiple attributes (such as luminance, colour and texture) are combined at a single location. However, in their experimental arrangement stimulus attributes always provided harmonious positional information. Similar findings were reported by Gray and Regan (1997). Landy and co-workers (Landy, 1993; Landy & Kojima, 2001) investigated vernier alignment of texture-defined edges, where the edge location could be signalled by differences in orientation, contrast or spatial frequency. The results showed that perceived edge location is determined from a weighted average of individual component estimates. In some of their conditions the edge location signalled by one cue was displaced relative to that of another in order to obtain estimates of the individual cue weights. In the current experiment, rather than varying the contrast of one cue whilst the other remains fixed, we reduce the contrast of one cue in proportion to an increase in the other in an attempt to keep the global contrast of the combined stimulus constant.

Methods

Subjects

Three of the authors acted as observers, and wore their optimal refractive correction where necessary.

Apparatus and Stimuli

The stimulus elements were composed of an additive combination of luminance contrast and texture contrast components. The luminance component is

$$\text{if } x \le 0 : S_{lum} = \exp\left(-\frac{x^2 + y^2}{2\sigma_1^2}\right)$$
$$\text{if } x > 0 : S_{lum} = \exp\left(-\frac{x^2 + y^2}{2\sigma_2^2}\right) \qquad (1)$$

The texture component is

$$\text{if } x \le 0 : S_{text} = \text{rand}(x, y)\exp\left(-\frac{x^2 + y^2}{2\sigma_2^2}\right)$$
$$\text{if } x > 0 : S_{text} = \text{rand}(x, y)\exp\left(-\frac{x^2 + y^2}{2\sigma_1^2}\right) \qquad (2)$$

where rand(x,y) is uniformly distributed on the interval [-1,1] and uncorrelated across the array of texture elements, which consisted of 2-by-2 pixel squares of diameter 3.22 arcmin. The parameters σ_1 and σ_2 represent the standard deviations of the stimulus envelope either side of the midline ($\sigma_1 + \sigma_2$ = 48.32 arcmin), whilst x and y are the

respective horizontal and vertical distances from the centre of the stimulus. For the symmetric condition σ_1 = σ_2 = 24.16 arcmin.

The stimulus was an additive mixture of these two patterns, having a contrast C_{lum} for the luminance pattern and C_{text} = 1 - C_{lum} for the texture pattern. Thus,

$$L(x,y) = L_{mean}\left(1 + C_{lum}S_{lum}(x,y) + C_{text}S_{text}(x,y)\right) \quad (3)$$

such that variations in the parameter C_{lum} determined the relative contrast of luminance and texture components within the stimulus. An alternative approach would have been to fix the contrast of one component whilst varying that of the other. However, this latter approach results in marked changes to the visibility of the object as a whole.

Generation and control of stimuli was performed using the macro capabilities of the public domain software NIH Image 1.61 (developed at the U.S. National Institutes of Health and available from http://rsb.info.nih.gov/nih-image/ or on floppy disk from the National Technical Information Service, Springfield, Virginia, part number PB95-500195GEI). Stimuli were presented on a Mitsubishi 21 inch d2 Colour Display Monitor with a mean luminance, L, of 38.3 cd m^{-2} and a frame rate of 75 Hz. The non-linear luminance response of the display was linearized by using the inverse function of the luminance response as measured with a Minolta CS-100 photometer. The host computer was a Motorola Starmax 4000/200 PowerPC. Sufficient contrast resolution for the measurement of

contrast detection thresholds was achieved by the use of a video summation device (Pelli & Zhang, 1991).

Procedures

Observers were asked to perform a three patch Vernier alignment task in which the horizontal position of the central element had to be judged with reference to two vertically separated reference patches (Figure 1). The vertical separation between each of the elements was 3.44 deg. Two conditions were investigated. In one the Gaussian envelopes of all three elements were symmetric, whilst in the second, the central element of the stimulus was composed of asymmetric Gaussian profiles; i.e. they contained a different Gaussian standard deviation (σ) either side of the midline. However, the overall size of each component remained constant – a reduction in the standard deviation on one side was balanced by a proportionate increase on the other. Importantly, the polarity of asymmetry was opposite for luminance and texture information, resulting in the centroid for each type of information being offset in opposite directions (see Figure 1B). The technique of offsetting stimulus centroids has previously been applied to motion displacement thresholds (Morgan, Ward & Cleary, 1994), and measures of stereoscopic disparity (Harris & Morgan, 1993). The perceived offset of the central blob, for both symmetric and asymmetric stimuli, was established using a method of constant stimuli with two alternatives, right or left. Within any experimental run, perceived offset was established for two stimuli of equal but opposite
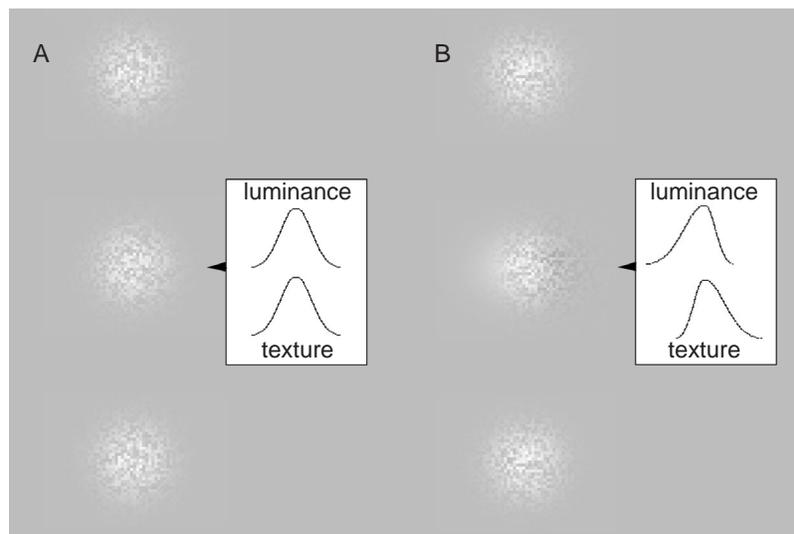

Figure 1. Examples of the stimuli used in the present experiments. Stimulus elements were composed of an additive combination of luminance and texture components. The Gaussian distribution of each component could be manipulated independently. In A the Gaussian envelopes are symmetric, but in B the middle envelope has been made asymmetric, with the polarity of asymmetry opposite for the luminance and texture components. The reader should be able to confirm how the relative visibility of each component changes the perceived position of the central element. If the figure is viewed from close up (~0.5m) the textural component tends to dominate and the central element should appear offset rightwards relative to the outer references. However, if the figure is viewed from a distance (~2m), or if it is blurred, the luminance component should now dominate and the central element appears offset in the opposite direction.

asymmetry (i.e. one stimulus and its mirror image), and either of these could occur with equal probability on any one trial. Each of the stimuli could be presented at any one of seven offsets, equally spaced around an alignment position determined by an initial method of adjustment. A step size of 3.22 arcmin between each of the seven offsets produced an appropriate ambit of responses ranging from approximately 100% rightwards to 100% leftwards. Stimuli were presented within a rectangular temporal window of 500msec duration. The results of the first 20 trials were discarded to allow subjects to familiarise themselves with the task. Following these, 80 trials were presented at each of the seven offsets and the proportion of "rightward" responses was calculated for each offset. The resulting data were fitted with a logistic function of the form

$$y = \frac{100}{1 + \exp\left(-\dfrac{x - \mu}{\theta}\right)} \qquad (4)$$

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

The relative amplitude of modulation of the luminance and texture components was obtained by varying C_{lum} between 0 and 1, and the perceived location of the central element was established as a function of this parameter. In addition, luminance and noise detection thresholds were measured to examine the role of stimulus visibility. Detection thresholds were established using a two alternative forced choice method of constant stimuli. Nine levels of contrast were used each separated by 0.05 log units. The task of the subject was to decide which of the two 500 msec intervals contained the stimulus. Twenty trials were randomly presented at each of the nine contrast levels. The data were then fitted with a logistic function in order to reveal the contrast level resulting in a 75% correct response level. All procedures followed the tenets of the Declaration of Helsinki.

Results

When both luminance and texture components are symmetrical and superimposed (Figure 2A-C, open circles), no perceived offset of the patch as a whole is observed, and, the stimulus is veridically perceived as at or near it's centroid (Figure 2A-C, solid lines). The perceived location of asymmetric patches composed of either luminance or texture information alone were also consistent with the calculated centroid positions of their stimulus envelopes (Figure 2A-C, dashed lines). The results of the present study support the view that observers locate luminance- and texture-only patches at or very close to the centroid of their respective distributions.

When the luminance and texture components of the central patch are made asymmetric, but skewed in opposite directions, a modulation in amplitude of one type of information relative to that of the other produces a smooth change in the perceived location of the object as a whole (Figure 2A-C, filled symbols). This effect could result from two very different modes of positional analysis. The representation of the object as a whole may be subjected to a nonlinear transformation (such as rectification) early in the visual pathway, following which a single positional cue is extracted. Alternatively, the visual system may extract a positional signal from both sources of visual information which are subsequently combined in order to extract a global representation of spatial position - the smooth change in position is a consequence of variation in the relative salience of each attribute as the 1st- and 2nd-order contrasts are traded off prior to the combination stage.

If the visual system is asked to locate either luminance or texture in isolation (luminance contrast of either 1 or 0 in Figure 2D-F), thresholds are similar in both the symmetric and asymmetric conditions. In the symmetric condition a reduction in the contrast of one component is offset by the increase in contrast of the other, resulting in an approximately linear function. However, when observers have to make localization judgements on patches composed of conflicting luminance and textural information (the asymmetric condition), localization accuracy is compromised as shown by the peak in thresholds near a luminance and texture contrast value of 0.5. The data in Figure 2D-F have been fitted by a function that allows quantification of this threshold elevation, as described in the legend to this figure.

For the inter-element separation used in this study, reducing the contrast of each cue in isolation from 100% to 50% has little effect on alignment thresholds. Nevertheless, whilst the accuracy of individual alignment measures shows a certain degree of contrast independence, the data presented in Figure 2A-C show that co-varying the relative supra-threshold contrasts of each component produces a marked, systematic change in perceived position, indicating that the *relative* level of supra-threshold contrast is of critical importance. In order to examine the role of visibility we measured threshold detection for asymmetric luminance and texture patches alone. Single-cue detection thresholds for both attributes were then used to express the point of perceived alignment as a multiple of its respective single-cue detection threshold (i.e. the contrast of a single cue in the combined stimulus is expressed as a function of its single-cue detection threshold at the point where no offset is perceived). The results are presented in Table 1. A much larger multiple of luminance detection threshold is required to balance textural information in situations where the two types of visual information provide conflicting positional cues.
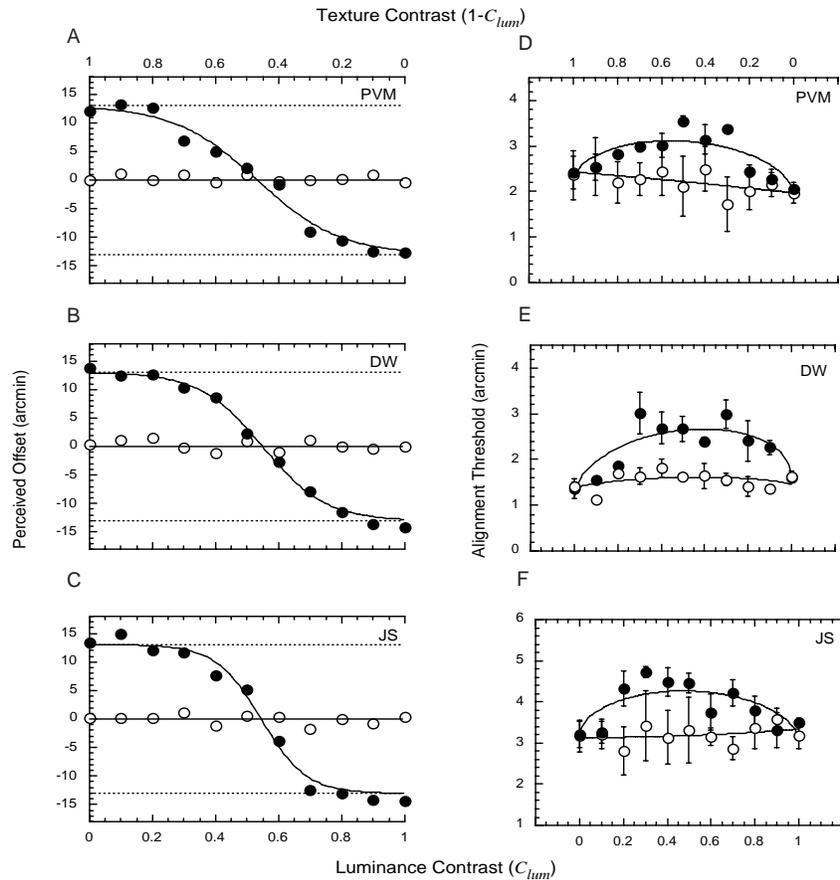
Figure 2. (A-C). Changes in perceived position for two different envelope asymmetry ratios: the symmetric condition (open circles) and an envelope asymmetry ratio of 0.5 – where one standard deviation is twice the size of the other (filled circles). The contrast of the texture component was coupled to that of the luminance component such that luminance contrast + texture contrast = 1. The dashed lines represent the calculated centroid positions of the luminance and texture envelopes for the asymmetric envelope condition. The asymmetric envelope (filled circles) shows a smooth change in perceived location as the modulation amplitude of one component is varied relative to that of the other. The curve fits are logistic functions constrained to the calculated centroid positions for the asymmetric luminance and texture envelopes. (D-F). Alignment thresholds as a function of luminance and texture contrast. Thresholds for the symmetric condition (open circles) and for an envelope asymmetry ratio of 0.5 (filled circles) are presented. Curves fitted to the data represent a least squares fit of the function $((T_L*x)^k+(T_T*(1-x))^k)^{(1/k)}$, where T_L and T_T represent the thresholds for luminance and texture components alone, and k is a parameter describing the degree of linearity. If k equals 1 a straight line joins T_L and T_T, values of k less than 1 reflect increasing amounts of threshold elevation in the mid-region of the data. The values of k for the symmetric and asymmetric conditions for each subject were as follows: PVM (k_{sym}= 0.99, k_{asym}= 0.64); DW (k_{sym}= 0.84, k_{asym}= 0.52); JS (k_{sym}= 1.03, k_{asym}= 0.7). Error bars were calculated from the parameter covariance matrix and represent one S.D. either side of the parameter value.

Discussion

   The question of what aspect of an object actually defines its apparent position has been of considerable interest for some time. There are a number of cues or 'location tags' which an observer can use to locate the relative position of objects within a visual scene. These include the peak of the object's luminance or contrast distribution, points of inflexion or zero crossings in the luminance distribution, the position at which edges of the object reach threshold, and the weighted mean or centroid of the distribution. Previous studies have

suggested that the most likely candidate is that of the centroid or 'centre of gravity' of the stimulus envelope for both luminance-defined and contrast-defined objects (Westheimer & McKee, 1977; Watt & Morgan, 1983; Morgan & Aiba 1985a; Morgan & Glennerster, 1991; Morgan, Ward & Cleary, 1994; Whitaker et al, 1996). The results of the present study support this assertion — the perceived location of both asymmetric luminance- and texture-defined patches was found to agree very closely with the calculated centroid position for each distribution.

Subject	Luminance	Texture
DW	19.93	11.34
PVM	16.93	12.34
JS	17.65	11.27

Table 1. The point of subjective alignment, i.e. the point where luminance and texture information exactly offset each other to produce perceptual alignment, are presented in terms of multiples of their respective detection thresholds. It can be seen that the luminance component must be a higher multiple of threshold to balance the texture component, suggesting that the visual system assigns more weight to texture information in the determination of perceived position.

When the luminance and texture components of the central patch provide conflicting positional cues, a modulation in amplitude of one type of information relative to that of the other produces a smooth change in the perceived location of the object as a whole. This indicates that a global estimate of object location is extracted after the visual system combines positional signals from both sources of visual information. In an elegant series of experiments, Rivest and Cavanagh (1996) showed that a contour defined by one visual attribute (e.g. luminance, colour, texture or motion) could influence the perceived location of that defined by another attribute. Furthermore, they showed that combining different attributes at a common location improves the accuracy of localization. The results of Rivest and Cavanagh's study strongly suggest that information from different visual attributes are combined at a common neural site prior to the level at which a localization decision is reached. Inspection of Figure 2D-F confirms that this is likely to be the case. The accuracy of relative localization for luminance or texture in isolation is very similar for both symmetric and asymmetric distributions. However, when observers are asked to locate patches composed of conflicting luminance and texture cues, localization thresholds are elevated, reaching a maximum in threshold elevation near a luminance and texture contrast value of 0.5. What might be the reason for this threshold elevation? Internal noise is likely to affect the relative salience of the two components from one trial to the next. For the symmetric condition this will have no effect since the positional cue provided by each component is in exact spatial registration. However, in the asymmetric condition, where each individual component provides a unique positional signal, trial-to-trial fluctuations in the relative strength of each signal constitute a significant additional source of variance. The localization noise for each individual component is no worse in the asymmetric condition compared to the symmetric condition, it is simply that the noise results in an increased response variance only when components signal conflicting positional estimates. A model of positional analysis which employed an early nonlinear transformation, followed by

the extraction of a single positional estimate, would not contain this additional source of variance. Furthermore, the results of both Rivest and Cavanagh (1996) and Gray and Regan (1997), suggest that the rule for combining positional signals derived from different visual attributes is consistent with probability summation between independent channels.

Alternative potential explanations exist for the elevation in localization thresholds for stimuli consisting of conflicting luminance and texture information. One possibility involves changes in the overall stimulus profile produced by combining two individual asymmetric profiles. Morgan and Aiba (1985b) have demonstrated that the precision with which the mean of a distribution can be extracted is dependent upon both the width of the distribution and its area. Our methodology ensured that asymmetric patches consisting of either luminance or texture alone differed from their symmetric counterparts in neither width nor area, since increases in the standard deviation of the patches on one side were counterbalanced by decreases on the other. It is reassuring, therefore, that asymmetric patches of either luminance or texture can be located with the same precision as their symmetric counterparts (Figure 2D-F). For stimuli consisting of an asymmetric combination of luminance and texture, however, it is important to eliminate potential changes in overall width and area as contributors to the threshold elevation in this specific region (Figure 2D-F). We therefore performed a control experiment in which alignment thresholds were measured for a combination of two asymmetric luminance profiles (each of contrast = 0.5) skewed in opposite directions, and also two asymmetric texture profiles. This allows us to directly compare performance against that for the combination of asymmetric luminance and texture profiles (shown in Figure2D-F, C_{lum}=0.5). Results are shown in the table below.

Subject	Asym Lum + Text (arcmin ± SD)	Asym Lum + Lum (arcmin ± SD)	Asym Text + Text (arcmin ± SD)
DW	2.67 ± 0.28	1.76 ± 0.08	1.38 ± 0.006
PVM	3.55 ± 0.65	1.56 ± 0.13	1.72 ± 0.17
JS	4.45 ± 0.8	2.51 ± 0.40	2.81 ± 0.30

Table 2. Alignment thresholds for three different asymmetric distribution combinations: luminance + texture; luminance + luminance; texture + texture.

For each observer, localization performance for combinations of the same type of information (i.e. luminance + luminance or texture + texture) are similar, and comparable with thresholds for the symmetric conditions (Figure2D-F). Thresholds for the combination of disparate sources of information (luminance + texture) are consistently higher, indicating that this threshold elevation reflects a true cost of disparate cue combination.

It might be argued that the threshold elevation is a result of a reduction in contrast of the individual luminance and texture components, i.e. at the extremes of Figure 2D-F (luminance contrast of C_{lum}= 0 and 1) either the luminance or texture component is at maximum contrast, whilst in the region of greatest threshold elevation both components are present at half their maximum contrast levels. However, if threshold elevation were a result of reduced individual component contrast, one would expect the same threshold elevation for the symmetric condition. This proves not to be the case, and indicates that the elevation in localization thresholds is likely to be a direct result of combining disparate sources of visual information.

Perceived alignment for patches composed of competing luminance and textural cues was obtained when the physical contrast of the luminance component was approximately equivalent to that of the texture component. This might seem to suggest that both components play an equivalent role in dictating the perceived position of the overall patch. However, this would only be the case if the visual system were equally adept at detecting the presence of luminance and textural information. In order to examine the role of visibility we measured threshold detection for asymmetric luminance and texture patches alone. Detection thresholds for both attributes were then used to express the point of perceived alignment (i.e. where no offset is perceived) as a multiple of its respective detection threshold. The results are presented in Table 1. A much larger multiple of luminance detection threshold is required to balance textural information when the two sources provide conflicting positional cues. It follows that, if both luminance and texture components were presented at an equal multiple of their detection thresholds, then the perceived location of the entire patch should appear offset in the direction of the textural component, which is indeed the case. An asymmetry in perceptual weights might be taken to indicate that the visual system does not treat all attributes equally but rather primacy is given to textural information over luminance information. This is in contrast to previous reports suggesting that luminance information was the dominant attribute in contour localization tasks (Livingstone & Hubel, 1988; Grossberg & Mingolla, 1985). Rivest and Cavanagh (1996) reported that when different attributes are combined at a single location, each providing concordant information, localization thresholds improve by an equivalent and statistically predictable amount as each attribute is added. This implies an equal role for each visual attribute. On the other hand, evidence for the unequal weighting of visual attributes has been suggested previously (Landy, Maloney, Johnston & Young, 1995; Mather & Smith, 2000; Landy & Kojima, 2001). For example, in the localization of texture-defined edges, Landy (1993) presents a model in which separate location estimates are made for each visual attribute, the attributes themselves

are then weighted, and the overall location is derived from the average of the weighted attributes. Within this framework, there is scope to assign larger weights to estimates derived from the particular visual cues that are the most robust and thus provide the most reliable estimate of the edge location. For example, in regions of a visual scene that contain little or no textural information preference might be given to more abundant visual attributes. Therefore, the final weighting of visual attributes is likely to be a product of the reliability of a particular cue and its availability. The quality of information provided by a visual attribute can vary not only from location to location but also over time, and the visual system needs to be able to accommodate such dynamic changes. However, the question of *how* the visual system weights different attributes remains. Landy et al. (1995) suggest that the weighting factors are derived from subsidiary cues which in isolation do not aid edge localization but do comment directly on the reliability of information provided by a particular attribute.

The results of the present study show that the visual cortex is able to effortlessly integrate disparate sources of visual information to form a global estimate of object position, although this conflation of visual attributes results in a modest loss of localization accuracy. Analogous effects have been reported in the motion domain, where the integration of luminance and chromatic information results in either enhancement or disruption of the motion percept depending on whether each attribute conflicts or concurs (Cavanagh, Arguin, von Grünau, 1989; Morgan & Ingle, 1994; Edwards & Badcock, 1996). Mis-matches between luminance and texture information are commonplace in the real world, where textured objects often vary in luminance across their surface, as a result of shadows or changes in illuminant position. The results of the present study suggest that texture information may be a more potent indicator of object position, implying that the human visual system gives more weight to visual attributes that are reliably related to the contours of objects. It is likely that visual experience plays an important role in shaping the weighting map of visual attributes, and conceivable that this weighting might be modified in different visual environments. Consider for example the mottled illumination of the forest floor. Textural difference between foliage can be small, and local luminance can change dramatically due to shadows, introducing luminance 'noise' to the scene. In such an environment, chromatic differences or colour cues, which are not subject to the same variability, may be particularly important. Therefore, the weighting map of visual attributes might reflect the evolutionary pressures imposed by the visual environment.

Acknowledgements

References

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

Badcock, D. R., Khuu, S. K. (2001) Independent first- and second-order motion energy analyses of optic flow. *Psychological Research*. *65*, 50-56. [PubMed]

Cavanagh, P., Arguin, M., von Grünau, M. (1989). Interattribute apparent motion. *Vision Research*, *29*, 1197-1204. [PubMed]

Chubb, C., Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986-2007. [PubMed]

Damasio, A., Yamada, T, Damasio, H., Corbett, J., McKee, J. (1980). Central achromatopsia: Behavioral, anatomical and physiologic aspects. *Neurology*, *30*, 1064-1071. [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]

Desimone, R., Ungerleider, L. G. (1989). Neural mechanisms of visual processing in monkeys. In Boller, F., Grafman, J. (Eds.) *Handbook of Neuropsychology* (Vol. 2, pp 267-299). New York: Elsevier.

De Yoe, E. G., Hockfield, S., Garren, H., van Essen, D. C. (1990). Antibody labeling of functional subdivisions in visual cortex: Cat-301 immunoreactivity in striate and extrastriate cortex of the macaque monkey. *Visual Neuroscience*, *5*, 67-81. [PubMed]

Edwards, M., Badcock, D. R. (1996). Global-motion perception: interaction of chromatic and luminance signals. *Vision Research*, *36*, 2423-2431. [PubMed]

van Essen, D. C., Maunsell, J. H. R.(1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neuroscience*, *4*, 370-375.

Graham. N., Beck, J., Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Research*, *32*, 719-743. [PubMed]

Gray, R., Regan, D. (1997). Vernier step acuity and bisection acuity for texture-defined form. *Vision Research*, *37*, 1713-1723. [PubMed]

Grossberg, S., Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review*, *92*, 137-211. [PubMed]

Harris, J. M., Morgan, M. J. (1993). Stereo and motion disparities interfere with positional averaging. *Vision Research*, *33*, 309-312. [PubMed]

Hubel, D. H., Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture of the visual cortex. *Journal of Physiology (London)*, *160*, 106-154.

Landy, M. S. (1993). Combining multiple cues in texture edge localization. *Proceedings of the SPIE*, *1913*, 506-517.

Landy, M. S. & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A*, *18*, 2307-2320. [PubMed]

Landy, M. S., Maloney, M. T., Johnston, E. B., Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research*, *35*, 389-412. [PubMed]

Ledgeway, T., Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, *34*, 2727-2740. [PubMed]

Li-Ming, L., Wilson, H. R. (1996). Fourier and non-Fourier pattern discrimination. *Vision Research*, *36*, 1907-1918. [PubMed]

Livingstone, M., Hubel, D. H. (1988). Segregation of form, color, movement and depth: Anatomy, physiology and perception. *Science*, *240*, 740-749. [PubMed]

Mareschal, I., Baker, C. L. (1998). A cortical locus for the processing of contrast-defined contours. *Nature Neuroscience, 1*, 150-154. [PubMed]

Mareschal, I., Baker, C. L. (1999). Cortical processing of second-order motion. *Visual Neuroscience, 16*, 527-540. [PubMed]

Mather, G., Smith, D. R. R. (2000). Depth cue integration: Stereopsis and image blur. *Vision Research*, *40*, 3501-3506. [PubMed]

Maunsell, J. H. R., Newsome, W. T. (1987). Visual processing in the monkey extrastriate cortex. *Annual Review of Neuroscience, 10*, 363-401. [PubMed]

McGraw, P. V., Levi, D. M., Whitaker, D. (1999). Spatial characteristics of the second-order visual pathway revealed by positional adaptation. *Nature Neuroscience, 2*, 479-484. [PubMed]

Morgan, M. J., Aiba, T. S. (1985a). Vernier acuity predicted from changes in the retinal light distribution of the retinal image. *Spatial Vision, 1*, 151-171. [PubMed]

Morgan, M. J., Aiba, T. S. (1985b). Positional acuity with chromatic stimuli. *Vision Research, 25*, 689-695. [PubMed]

Morgan, M. J., Glennerster, A. (1991). Efficiency of locating centers of dot-clusters by human observers. *Vision Research, 31*, 2075-2083. [PubMed]

Morgan, M. J., Ingle, G. (1994). What direction of motion do we see if luminance but not colour contrast is reversed during displacement? Psychophysical evidence for a signed-colour input to motion detection. *Vision Research, 34*, 2527-2535. [PubMed]

Morgan, M. J., Ward, R. M., Cleary, R. F. (1994). Motion displacement thresholds for compound stimuli predicted by the displacement of centroids. Vision Research, *34*, 747-749. [PubMed]

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

Olavarria, J. F., DeYoe,, E. A., Knierim, J. J., Fox, J. M., Van Essen, D. C. (1992). Neural responses to visual texture patterns in the middle temporal area of the macaque monkey. *Journal of Neurophysiology, 68*, 164-181. [PubMed]

Pearlman, A. L., Birch, J., Meadows, J. C. (1979). Cerebral color blindness: An acquired defect in hue discrimination. *Annals of Neurology, 5*, 253-261. [PubMed]

Pelli, D. G., Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research, 31*, 1337-1350. [PubMed]

Rivest, J., Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Research, 36*, 53-66. [PubMed]

Shapley, R. M. (1994). Linearity and non-linearity in cortical receptive fields. *In Higher-order visual processing in the visual system. Ciba Foundation Symposia*, (pp 71-87). John Wiley & Sons Ltd.: New York.

Tootell, R. B. H., Hadjikhani, N. K., Mendola, J. D., Marrett, S., & Dale, A. M. (1998). From retinotopy to recognition: fMRI in human visual cortex. *Trends in Cognitive Sciences, 2*, 174-183.

Watt, R. J., Morgan, M. J. (1983). Mechanisms responsible for the assessment of visual location: Theory and evidence. Vision Research, *23*, 97-109. [PubMed]

Westheimer, G., McKee, S. P. (1977). Integration regions for visual hyperacuity. *Vision Research, 17*, 89-93. [PubMed]

Whitaker, D., McGraw, P. V., Levi, D. M. (1997). The influence of adaptation on perceived visual location. *Vision Research, 37*, 2207-2216. [PubMed]

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

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience, 9*, 79-97. [PubMed]

Zeki, S., Shipp, S. (1988). The functional logic of cortical connections. *Nature, 335*, 311-317. [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]

Zihl, J., von Cramon, D., Mai, N. (1983). Selective disturbance of movement vision after bilateral brain damage. *Brain, 106*, 313-340. [PubMed]