Probing intermediate stages of shape processing

Gunter Loffler

Department of Life Sciences,
Glasgow Caledonian University, Glasgow, UK

The visual system provides a representation of what and where objects are. This entails parsing the visual scene into distinct objects. Initially, the visual system encodes information locally. While interactions between adjacent cells can explain how local fragments of an object’s contour are extracted from a scene, such computations are ill suited to capture extended objects. This article reviews some of the evidence in favor of intermediate-level computations, tuned to the shape of an object, in the transformation from discrete local sampling to representation of complex objects. Two main paradigms, employed to study how information about the position and orientation of local signals are combined at intermediate levels, are considered here: a shape detection task (measuring the number of signal elements required to detect a shape in noise) and a shape discrimination task (requiring observers to discriminate between shapes). Results support the notion of global mechanisms that integrate information beyond neighboring cells and are optimally tuned to a range of different shapes. These intermediate processing stages appear vulnerable to damage. Diverse clinical conditions (amblyopia, macular disease, migraine, premature birth) show specific deficits for these tasks. Taken together, evidence is converging in favor of intermediate levels of processing, at which sensitivity to the global shape of objects emerges.

Introduction

In order to safely navigate through and purposefully interact with our environment, humans require access to a reasonably accurate representation of the outside world. An important aspect of this process concerns the location and identification of individual objects. To support this, the visual system has to be able to parse visual input into distinct objects. At the early stages of visual processing, however, the coding of the sensory input is discrete and highly localized: The map of activity corresponds to light emerging from small parts of the visual scene, with neighboring neurons responding to adjacent parts of the visual field. Such discrete and localized sampling presents a fundamental computational problem (Wallach, 1935). Neurons receiving input from only a tiny fraction of the visual scene are unable to distinguish between extended objects. The activity of an individual cell in primary visual cortex (V1) reflects contour orientation and spatial scale at a particular retinotopic location (Hubel & Wiesel, 1968). This activity is consistent with any of an infinite number of object contours. Such localized information must be disambiguated using information from other parts of the visual scene.

Based on image statistics, the likelihood that two adjacent points in the visual field belong to the same object is high (Elder & Goldberg, 2002; Geisler, Perry, Super, & Gallogly, 2001). Therefore, an obvious starting point for signal integration is long-range lateral interactions (Figure 1A, ‘+’) between neighboring V1 neurons with nonoverlapping receptive fields. Such horizontal connections between V1 cells in close proximity are well documented (Gilbert & Wiesel, 1981). Behaviorally, geometric rules (e.g., proximity, co-alignment; Wertheimer, 1923) have been inferred from studies on collinear facilitation that describe the circumstances where these interactions are effective (e.g., Polat & Sagi, 1993).

Object contours may then be extracted from the visual input by chains of mutually excitatory connections between cells (Figure 1B). Such connections could form the building blocks of contour integration by allowing subsequent stages to bind cell responses and thereby group parts of a scene into extended, coherent contours.

These computations are insufficient, however, to determine whether two signals originate from the same object in the case of overlap or abrupt changes in edge orientation (e.g., corners). Where objects overlap, and due to the resulting contour intersections, the visual system has to distinguish between parts of a scene that should be combined (Figure 1B, ‘+’) from those that should be kept separate (Figure 1B, ‘-‘). This decision is impossible to make on a local basis, and computational rules more complex than simple proximity have to be considered. Experimental evi-
Evidence has confirmed the existence of global mechanisms that integrate information beyond neighboring cells (Figure 1C, “Σ”).

Following the detection of a shape embedded within a scene, the visual system must be able to discriminate it from other shapes (Figure 1D) in order to enable object identification, recognition, and categorization. These processes are likely to depend on the way that shapes are represented. One popular proposal is a reference-based coding strategy, whereby objects are represented within a multidimensional space according to how much they differ from a reference (a prototype or mean). Evidence for norm-based representation has been reported for a number of shapes, including squares, rectangles (Regan & Hamstra, 1992), circles (Habak, Wilkinson, & Wilson, 2006; Wilson, Loffler, & Wilkinson, 2002), as well as more complex objects such as faces (Figure 1E; Kayaert, Biederman, & Vogels, 2003; Loffler, Gordon, Wilkinson, Goren, & Wilson, 2005; Loffler, Yourganov, Wilkinson, & Wilson, 2005).

This is in accordance with physiological studies that have shown an increase in computational complexity along the hierarchy of processing stages. While results on spatial facilitation have been explained by known anatomy and physiology in primary visual cortex, more global computations are likely processed in extrastriate areas. Recordings from area V2 in the macaque are consistent with processing of angles (Hegde & Van Essen, 2000; Ito & Komatsu, 2004), combining outputs from multiple orientations-selective neurons in striate cortex. Neurons in V4 have been shown to exhibit selectivity to curved shapes including concentric circles (Dumoulin & Hess, 2007; Gallant, Braun, & Vanessen, 1993; Wilkinson et al., 2000). The response characteristics of these cells require pooling of information from detectors tuned to a wide range of orientations centered at different positions of the visual field. Moreover, cell recordings from V4 are consistent with a population code for complex curved shapes that is sensitive to the location of convex curvature extrema (Pasupathy &Connor, 2001, 2002; see also Wilson & Wilkinson, 2015). Based on these insights, V2 and V4 are considered to be at an intermediate level, in the sense that they encode more complex object features than

Figure 1. Overview of the putative processes involved in shape processing. (A) Long-range lateral interactions (“+”) between neighboring V1 neurons with nonoverlapping receptive fields (shown by ellipses) can be used to respond to contour fragments. Geometric rules (e.g., proximity, co-alignment) have been inferred from studies on collinear facilitation that describe the circumstances when these interactions are effective. (B) Chains of such interactions might be building blocks for contour integration. A computation problem in this process is to determine those parts of a scene that should be combined (“+”) and those that should be kept separate (“−”). (C) This problem cannot entirely be solved on a local basis, and experimental evidence points towards global mechanisms that integrate information beyond neighboring cells (“Σ”). (D) Following the detection of a global shape embedded in a scene, the visual system must be able to discriminate it from other shapes. (E) These processes are likely to depend upon the way the brain represents objects. One popular proposal is a reference-based coding strategy, whereby objects are represented within a multidimensional space depending on how much they differ from a reference (a prototype or mean). Evidence for such norm-based representations has been reported for a number of shapes, including circles and triangles as well as more complex objects such as faces. In the latter case, individual faces might be encoded within a multidimensional face space, where the distance from a mean face determines the facial distinctiveness and the direction its identity.

Reproduced with permission from Loffler, 2008.
edge orientation but more elementary features than meaningful objects such as faces (Loffler, 2008).

This article is concerned with the midlevel stage (Figure 1C) that sits between the early (e.g., local contour extraction; Figure 1A) and late (complex object representation; Figure 1E) stages of visual processing, i.e., the point where global processes integrate local information to derive a global shape representation. The existence of intermediate-level representations tuned to object shape is gaining support from physiological, perceptual, and brain-imaging studies (see other articles in this issue).

Rather than providing a comprehensive review of the literature on shape perception (for this, the reader is referred to Loffler, 2008), the following outlines a selection of behavioral studies that aimed to elucidate the characteristics of these intermediate stages. It is broadly split into two prominent behavioral tasks: detecting a shape embedded in noise (shape detection) and discriminating one shape from another (shape discrimination).

If these intermediate stages play a fundamental role in the processing of visual objects, abnormalities in the underlying mechanisms may be manifest in developmental and acquired pathologies. In agreement with this, studies have reported shape-processing deficits in a number of clinical conditions, including those affecting the very early stages of visual processing (age-related macular degeneration; Kennedy, Baird, McGarva, Abady, & Loffler, 2014; Wang, Wilson, Locke, & Edwards, 2002), visual development (amblyopia: Hess, Wang, Demanis et al., 1999; Kennedy et al., 2014), cortical visual dysfunction in preterm children (Atkinson & Braddick, 2007; Macintyre-Beon et al., 2013), and other neurological conditions (migraine: Wagner, Manahilov, Gordon, & Loffler, 2013).

**Shape detection**

Measuring the ability of observers to detect a stimulus embedded in noise has a long history. Describing the effect of external noise on the visibility of a target has been used to determine the sensitivity and characteristics of the mechanisms responsible for processing the stimulus. Motion coherence thresholds, for example, are a standard tool for investigating motion perception (e.g., Newsome & Pare, 1988).

**Shape coherence thresholds**

This approach has been utilized to test detectability of a contour. A contour, sampled by a variable number of oriented elements, is embedded in an array of otherwise randomly oriented elements (Figure 2A). Detection requires the visual system to link the aligned contour elements while avoiding linkage of noise elements. Studies have provided insight into the computations underlying this process (e.g., Field, Hayes, & Hess, 1993). The main determinant for successful linkage follows from certain geometric relationships between adjacent elements, notably alignment and collinearity. Behavioral data can be explained by an “association field” model (Field et al., 1993). According to this model, association between
filters is strong along an axis given by the filter’s preferred orientation, i.e., between elements with orientations tangential to a smooth contour. It has been suggested that many of these contextual effects may be mediated by long-range interactions between cells in V1 (e.g., Li & Gilbert, 2002).

Long-range interactions and the association model consider adjacent detectors, but the detectability of contours embedded in noise is unlikely to be mediated solely by interactions between neighboring elements. For example, the number of elements sampling a contour modulates detectability. Contours composed of too few elements are undetectable (Braun, 1999; Kovacs & Julesz, 1993). It has therefore been proposed that the presence of extended contours may be signaled by chains of aligned elements (Figure 1B). The saliency of a contour depends not only on elements’ being aligned with immediate neighbors but also on those neighbors’ likewise having aligned neighbors (Braun, 1999; Field et al., 1993; Kovacs & Julesz, 1993).

A number of studies have looked at the influence of information from outside the immediate neighborhood of individual elements. One important feature in this regard is closure. Closed contours are more easily detected than open ones (Elder & Zucker, 1993; Kovacs & Julesz, 1993; Pettet, 1999; Pettet, McKee, & Grzywacz, 1998). For example, an S-shaped contour is harder to detect than a circular one (Pettet, 1999; see Figure 2B). This points toward a need for global computations, as information about contour closure is not available locally.

One way to selectively target global aspects of signal integration is to render local information useless, so that observers are forced to base their decisions on information that is not available in a restricted local region. This strategy has been used successfully for texture processing (Figure 2C; Achtman, Hess, & Wang, 2003; Dakin, 1997; Glass, 1969; Wilson, Loffler, Wilkinson, & Thistlethwaite, 2001; Wilson & Wilkinson, 1998; Wilson & Wilkinson, 2015). A similar approach has been used to study shape detection (Figure 2B; Achtman et al., 2003; Loffler, 2008; Schmidtmann, Gordon, Bennett, & Loffler, 2013). For example, one study determined the amount of signal pooling and its dependence on the actual shape of the contour (Schmidtmann et al., 2013). The shapes were radial frequency patterns (Wilkinson, Wilson, & Habak, 1998), a class of smooth shapes that have been used widely to study shape processing (see discussions later).

Best performance of about 10% coherence was found for circular shapes (Figure 3). Substituting the circle for different radial frequency shapes had a systematic effect on coherence thresholds: Thresholds increased approximately proportionally to the square of the shape frequency (number of lobes). As the maximum curvature of the shapes is proportional to the square of the shape frequency, it was suggested that detectability of these shapes is inversely proportional to their maximum curvature (Schmidtmann et al., 2013).

There are two obvious ways in which these patterns might be processed. The first is a texture detector, which integrates information across space from elements so long as their orientation is consistent with the specific type of texture for which it is tuned. This computation is sensitive to the orientation but not position of local elements. For example, a concentric texture detector sums information from any element within its receptive field as long as the orientation of that element is perpendicular to radial lines emerging from the texture’s center (Figure 4C). An alternative mechanism is a shape detector. As with the texture detector, it is tuned to the orientation of local elements. Unlike the texture detector, however, it is also sensitive to their position (Figure 4C). One can distinguish between these two hypothetical mechanisms experimentally by constraining the location of signal elements. If the underlying mechanism sums information from anywhere in its receptive field (texture detector), the number of signal elements required to reach threshold should be independent of their distribution. On the other hand, a shape detector would require fewer signal elements if those were sampled from within specific annular regions.

Figure 3. Detecting different shapes embedded in noise. The stimuli were sampled concentric shapes (Figure 2C). The icons above the data illustrate the general shape of the concentric contours that were sampled by Gabor filters. Data are detection thresholds (the percentage of signal elements aligned to the contours relative to all elements in the array) as a function of the number of lobes (RF) of the shape. Detection thresholds for low RFs were ~10%, and thresholds rose for higher RFs. Thresholds increased approximately with the square of the shape frequency (solid gray line). Adapted with permission from Schmidtmann et al. (2013).
Experimental evidence (Schmidtmann et al., 2013) favors the shape detector. Observers require only about half the number of signal elements when they fall within an annulus compared to when they are spread randomly across the stimulus area (Figure 4). The resulting sensitivity (3%–4% coherence thresholds) for annuli is remarkably low. This has implications for the underlying mechanisms that are responsible for detecting these patterns. Depending on the diameter of the shape, adjacent signal elements are separated, on average, by five noise elements (Figure 4C). To be able to detect the signal in these displays, an analyzer has to sum information efficiently and globally along the entire contour. This process has to operate in the absence of reliable local image statistics such as alignment or collinearity of adjacent elements.

The comparatively poor performance seen when elements are spread across the display is, to a first approximation, captured by the probabilistic combination of the outputs from multiple, independent shape detectors. This can explain the greater number of signal elements required for textured patterns. The same pattern of behavior is also seen for other shapes (Schmidtmann et al., 2013), suggesting the existence of highly sensitive analyzers, which sum information globally but only within shaped annuli.

**Position versus orientation: A shape illusion**

Subsequent studies have added external noise to either the orientation or the position of the elements to determine the tuning profile of these putative shape
detectors (Schmidtmann et al., 2013). Different shapes exhibit similar tuning with regard to orientation and position, suggesting the existence of a range of detectors that are tuned to different shapes (Schmidtmann et al., 2013). In all cases, the tuning for position is broader than for orientation. A relatively broad tuning for position makes an interesting prediction: If elements were taken from, e.g., a pentagon and placed on a circular ring, the resulting percept might be driven by element orientation, largely ignoring element position. That is, these patterns might be perceived as a pentagon rather than a jagged circle. This has been confirmed experimentally and provides a compelling shape illusion (Figure 5; Day & Loffler, 2009). In the study, observers were presented with ambiguous patterns where element position was consistent with one shape (circle) and element orientation with another (e.g., pentagon). In these cases, observers could perceive either the illusory pentagon with elements incorrectly seen at different distances from the shape’s center or a jagged circle. Which of the two percepts prevailed depended on a number of parameters (including total number of elements) but was relatively independent of characteristics such as element scale, phase, and polarity, indicating that the effect is largely unaffected by interactions at the early stages of visual processing.

**Shape discrimination**

Physiological evidence has been presented in favor of extrastriate processing for angles and smoothly curved contour shapes (see Introduction). A number of behavioral studies have used pattern discrimination to probe these intermediate processing stages. Extrastriate involvement in these tasks has been implied by performance that cannot fully be captured by information that is available locally and instead requires global signal integration.

**Angles**

Points of maximum curvature and angles are considered particularly important for object perception (Attneave, 1954). The classical demonstration by Attneave showed that an object could be recognized when its contour was reduced to straight lines connecting points of maximum curvature. Angles are appealing features for general object representation because they are scale invariant (Kennedy, Orbach, Gordon, & Loffler, 2008; Milner, 1974) and can, in theory, be easily computed by combining outputs from filters tuned to edge orientations like those found in V1 (Boynton & Hegde, 2004). Physiological evidence supports this notion (Hegde & Van Essen, 2000).

Although angles are completely defined locally at the point of intersection of two lines, behavioral data point to more complex operations than the mere determination of the difference in orientation of two lines. For example, human sensitivity to angles can be better than sensitivity to orientation (Chen & Levi, 1996; Heeley & Buchanan-Smith, 1996; Kennedy, Orbach, & Loffler, 2008; Regan, Gray, & Hamstra, 1996; Snippe & Koenderink, 1994). Moreover, sensitivity to angle discrimination shows a dependence on the global, triangular shape that contains the angle (Figure 6A). Sensitivity is substantially better for symmetrical isosceles triangles compared to scalene shapes (Kennedy, Orbach et al. 2006; Figure 6B). The overall stimulus shape also affects the appearance of angles (Kennedy, Orbach, & Loffler, 2008). An angle presented in an isosceles triangle is judged to be
substantially larger than the same angle embedded in a scalene triangle (Figure 6C). That sensitivity and appearance of angles are dependent on the overall stimulus geometry implies that angular processing is not wholly determined by local computations but instead is influenced by more global aspects of the stimulus.

Smooth contours

Radial frequency (RF) patterns (Wilkinson et al., 1998) are widely used stimuli in the study of shape perception. RF patterns are generated by applying a sinusoidal modulation to the radius of a circle in polar coordinates:

\[ r(\theta) = r_{\text{mean}} \left( 1 + A \sin(\omega \theta + \varphi) \right), \]

where \( r_{\text{mean}} \) represents the mean radius (size), \( \omega \) the frequency (number of cycles or corners), and \( A \) the modulation amplitude (pointedness of each corner). Icons at the top of Figure 3 illustrate a range of RF frequencies (e.g., \( \omega = 0 = \text{circle}; \quad \omega = 5 = \text{pentagon}; \quad \omega = 13 = 13\text{-lobed star} \)). Icons on the right in Figure 7 show two RF amplitudes for the case of an RF5 pattern (\( A = 5\% = \text{pentagon shape without concavities}; \quad A = 20\% = \text{five-sided star} \)).

To quantitatively analyze shape sensitivity, studies have often measured the minimum amplitude of an RF pattern for it to be discriminable from a circle. Typical thresholds for this task fall in the hyperacuity range, with differences between circle and RF shape at threshold of less than 10°–15° (Schmidtmann, Kennedy, Orbach, & Loffler, 2012; Wilkinson et al., 1998). Such exquisite performance requires highly sensitive mechanisms. Studies have investigated the premise that this high sensitivity is the result of efficient signal integration (Bell & Badcock, 2009; Bell, Badcock, Wilson, & Wilkinson, 2007; Bell, Dickinson, & Badcock, 2008; Habak et al., 2006; Hess, Achtman, & Wang, 2001; Hess, Wang, & Dakin, 1999; Jeffrey, Wang, & Birch, 2002; Loffler, Wilson, & Wilkinson, 2003; Schmidtmann et al., 2012; Wilkinson et al., 1998). One strategy to describe the magnitude of global pooling is to compare sensitivity when different amounts of a stimulus contain signal. In discriminating contours from circular, decreasing the number of modulated...
cycles (with the remainder of the pattern circular) decreases the amount of signal (Figure 7).

The improvement in sensitivity with increasing amounts of signal can be compared to predictions of processes that either have access only to local information or can utilize information globally (see Figure 7). Although studies differ in detail, the consensus is that sensitivity to RF discrimination cannot be explained by local processes or inefficient combination of local information. Rather, evidence is accumulating in favor of global processing strategies whereby detectors have access to, and efficiently integrate, information from the entire extent of a shape (Bell et al., 2007; Dickinson, McGinty, Webster, & Badcock, 2012; Hess, Wang, & Dakin, 1999; Loffler et al., 2003; Schmidtmann et al., 2012; Wilkinson et al., 1998). In the case of a recent study that tested a wide range of shapes (Schmidtmann et al., 2012), the typical data exhibit two regimes (Figure 7). First, there is an initial shallow improvement when the amount of modulated contour increases. Once the modulation is applied to the entire contour, the improvement is substantial. The shallow region is well captured by local processes. The steep region is consistent with global pooling.

While evidence for global pooling is seen for a range of shapes, it is not a universal feature. It is only found for shapes with a few lobes, up to about RF8–RF10 (Jeffrey et al., 2002; Loffler et al., 2003). If shapes have more lobes (higher radial frequencies), no evidence for global summation is seen, and sensitivity appears to be limited by local computations. In summary, these studies have provided evidence in favor of mechanisms that achieve their high sensitivities by efficiently combining information from across disjoint parts of a contour shape.

**Shape channels**

A particularly appealing feature of RF patterns is that, in combination, they can be used to represent complex natural shapes such as fruits and vegetables (Wilson & Wilkinson, 2002), human head contours (Loffler et al., 2003; Wilson et al., 2002; Wilson, Wilkinson, Lin, & Castillo, 2000), and animal shapes and torsos (Alter & Schwartz, 1988; Wilson & Wilkinson, 2002). A number of studies have investigated the hypothesis that the visual system may decompose and represent complex shapes by simpler components. Differences in the nature or strength of global pooling for different RF shapes have been linked to a range of shape channels (Hess, Wang, & Dakin, 1999; Jeffrey et al., 2002; Loffler et al., 2003; Schmidtmann et al., 2012; Wilkinson et al., 1998). Adaptation (Anderson, Habak, Wilkinson, & Wilson, 2012).
Population code for shape processing

Based on these behavioral results, models have been devised for shape processing that incorporate a population code strategy (Kempgens, Loffler, & Orbach, 2013; Poirier & Wilson, 2006; Wilson & Wilkinson, 2015). This follows cell recordings from V4, which have been shown to be consistent with a population code for complex curved shapes (Pasupathy & Connor, 2001, 2002). These cells show tuning for curvature, distance from the shape center, and polar angle corresponding to the location of curvature extrema.

One model (Kempgens et al., 2013) arose from a psychophysical investigation into the ability to detect local changes in sampled RF contours. Observers were required to detect a deviation in orientation of one element from tangential to the sampled shape. Such imperfections or heterogeneities in shapes often correlate with points of interest in a visual scene. The resulting data could be predicted by a model in which individual shapes are represented by a population code (Figure 8). According to this model, shapes (Figure 8A) are processed initially by a bank of curvature units (Figure 8B) that extract local convex and concave curvature extrema from the contour. The outputs are combined with information about the center of the contour by global arc units. These arc units are therefore tuned to the location of a shape’s points of curvature extrema relative to its center. For example, in Figure 8C, individual arc units respond to a convexity at 9 o’clock or a concavity at 7 o’clock. Individual shapes are represented by a population code of these arc units (Figure 8D). Activation of units within this population code depends on the shape of the stimulus, and shapes can be differentiated on the basis of the pattern of active units. The model offers an explicit framework for how a population code for shape processing that is consistent with physiology can be applied to psychophysical data.

Shape perception in clinical populations

Intermediate-stages of visual processing in both dorsal and ventral streams are vulnerable to damage (e.g., Atkinson & Braddick, 2007; Ellemberg, Hess, & Arsenault, 2002; Taylor, Jakobson, Maurer, & Lewis, 2009). Tests targeting shape processing reveal deficits in a range of clinical conditions, including amblyopia, macular disease (e.g., age-related macular degeneration), migraine, and premature birth. In some cases, these deficits can be explained by defects at early levels, where an impoverished signal has a detrimental effect on subsequent processing. In others, the deficits appear to originate at intermediate levels in the absence of low-level damage. In all cases, subjects with these conditions performed less well than subjects without them in shape detection or discrimination tasks.

Amblyopia

Amblyopia (lazy eye) is a heterogeneous developmental disorder that is characterized by reduced visual acuity, typically in one eye (Ciuffreda, Levi, & Selenow, 1991). The typical presentation is associated with a misalignment of the two eyes (squint, strabismic amblyopia) or a substantial difference in refractive error between the two eyes (anisometropic amblyopia). The prevalence of amblyopia is 2%–4%. It is the most common pediatric ophthalmic disorder in the industrialized world (for a review, see Levi & Li, 2009). The condition is a result of physiological changes in visual cortex that are caused by abnormal binocular visual experience during a critical period of development.

Although reduced visual acuity in an otherwise healthy eye is the presenting clinical feature, it has been
Figure 8. Overview of a model for shape processing. (A) A high-amplitude RF4 shape with convex and concave points of curvature serves as a sample stimulus. (B) Local curvature processing. The icon shows a fraction of the sample shape with superimposed triplets of V1 orientation filters used to extract local curvature. Curvature processing is supposed to be locally antagonistic (convex vs. concave); “active” triplets are shown in high contrast and their responses are combined multiplicatively. For the RF4 shape, a “convex” triplet responds to the orange segment and a “concave” triplet to the green segment. (C) Global arc units. The responses of local curvature units (B) are integrated (Σ) along the contour up to the points where the contour’s curvature changes sign. The arc units combine the responses from local curvature units with information about the center of the contour and the distance from the contour to the center. These arc units are therefore tuned to the location of a shape’s points of local curvature extrema relative to its center (the relation to the center is symbolized by showing arc units as closed shapes with an obvious center), consistent with V4 physiology. For the example shown, one arc unit is sensitive to a convexity at 9 o’clock (orange) and another to a concavity at 7 o’clock.
shown that deficits in amblyopia are not restricted to acuity (for a review, see Levi, 2006). For example, people with amblyopia exhibit defects in low-level tasks, such as contrast sensitivity, vernier acuity (McKee, Levi, & Movshon, 2003), and symmetry detection (Levi & Saarinen, 2004), as well as intermediate tasks, including efficient signal integration for moving (Simmers, Ledgeway, Hess, & McGraw, 2003) and static (Levi, Klein, & Sharma, 1999) targets. In many cases, these cannot be explained solely by low-level processing defects, such as those occurring in V1. Instead, evidence supports additional or amplified processing abnormalities at extrastriate cortical regions (Levi, 2006).

The ability to discriminate shapes is also affected in amblyopia (Dallala, Wang, & Hess, 2010; Hess, Wang, Demanins et al., 1996). People with amblyopia perform significantly more poorly with their amblyopic eye in an RF discrimination task than with their nonamblyopic eye. Recently, we investigated the extent to which this deficit can be explained by a lack of global pooling (Kennedy et al., 2014). Comparing the data for the amblyopic and nonamblyopic eyes of an individual observer indicates that local processing is largely intact but global pooling is ineffective in the amblyopic eye (Figure 9A). Sensitivity for the amblyopic eye improves with increasing number of modulated cycles, but the improvement remains modest. The shallow increase in sensitivity is consistent with inefficient signal integration (see Figure 7). In contrast, the nonamblyopic eye exhibits the typical sharp improvement when the entire pattern is modulated, as a result of efficient signal integration. The amblyopic eye’s sensitivity appears to be limited by local processing, lacking the advantage of global summation.

**Macular disease**

Shape discrimination deficits have also been reported in macular disease (Wang et al., 2002). The macula is the highly sensitive region at the center of the retina and extends across about 15°–20° of visual angle. Even small disruptions to the architecture of this area can have a profound effect on central vision (Curcio, Medeiros, & Millican, 1996). A range of conditions can affect the functioning of the macular region. The most common presentation of macular disease is due to age-related macular degeneration, which causes a disruption of the central photoreceptor mosaic. Other presentations include swelling at the macula (cystoid macular edema), diabetic changes (diabetic maculopathy), and choroidal pathologies affecting the overlying retina (e.g., chorioretinitis; Kanski & Bowling, 2011). Maculopathies ultimately affect visual acuity, and reduced acuity can have a negative impact on high-level visual functioning and perceived quality of life (e.g., McCulloch et al., 2011), but patients often report subtle distortions as the earliest symptoms of pathological changes.

Studies have applied shape discrimination as a tool to detect macular disease psychophysically (Wang et al., 2002). This is based on the distorting feature present in the early stages of macular disease and on the nature of a task that requires observers to detect subtle deformation to a circle as measured in RF discrimination. An appealing feature of this test is the high sensitivity of observers without macular disease, who perform in the hyperacuity range; hence even very subtle or early pathology might result in a measurable deviation from typical performance. In agreement with this hypothesis, Wang et al. (2002) observed poorer performance in subjects with age-related macular degeneration compared to healthy age-matched controls. Differing clinical severities of age-related macular degeneration correlated with different sensitivities on the shape discrimination test even across cases where visual acuity was similar. This shows that shape discrimination sensitivity can be directly related to macular function.

The defective shape processing in macular disease can be linked to a lack of global signal pooling (Kennedy et al., 2014). The data in Figure 9B show the results for one observer with a chorioretinal scar. A chorioretinal scar is an area of atrophy of the choroid and overlying retina. The fundus picture shows this as a lighter region, due to the increased visibility of the sclera. This area corresponds to a scotoma (blind spot). In this patient, the scar extends from the optic nerve head superiorly and temporally, including the macula but sparing the central fovea. Correspondingly, central visual acuity is largely unaffected. Shape discrimination in the unaffected eye follows the typical pattern.
is a shallow improvement with increasing contour deformation, followed by a substantial increase in sensitivity when the entire contour is deformed. The eye with the chorioretinal scar follows the shallow improvement throughout, suggesting a lack of global integration similar to that for an amblyopic eye (Figure 9B). In the case here, global pooling might be ineffective because local information about shape modulation is missing. It is nevertheless interesting that two very different pathologies, macular disease and amblyopia, result in such similar patterns of behavior.

Migraine

Migraine is a debilitating neurological condition that affects approximately 11% of the adult population (Steiner et al., 2003). Visual symptoms are common in migraine (e.g., Wilkinson, 2004). Visual aura, the most commonly reported symptom, is a visual hallucination that can take a number of forms. These include simple points of light (phosphenes) in the visual field and white or gold zigzag lines surrounding a scotoma which spreads across the visual field (the “classic” fortification spectra). The nature of these patterns makes it likely that they originate in the occipital cortex, possibly V1 (Wilkinson, 2004).

It has been proposed that migraine auras may be the result of cortical hyperexcitability (Welch, D’Andrea, Tepley, Barkley, & Ramadan, 1990) due to a lack of inhibitory control (Wilkins et al., 1984). This reduced-inhibition hypothesis has been tested in a number of studies, but results have been inconclusive. While some have found performance of people with migraine consistent with defective inhibition (e.g., Palmer, Chronicle, Rolan, & Mulleners, 2000), others have found them to operate at a near-typical level (McColl & Wilkinson, 2000; Shepherd, 2000, 2001; Shepherd, Palmer, & Davis, 2002; Shepherd, Wyatt, & Tibber, 2011; Wilkinson & Crotogino, 2000). An alternative explanation for perceptual differences between people

Figure 9. Shape discrimination sensitivity for an observer with amblyopia (A) and an observer with macular pathology (B). The task was to discriminate an RF5 test pattern from a circular reference. Thresholds are given as the minimum amplitude of the test required for reliable discrimination, plotted as a function of the number of modulated contour cycles (one, three, four, and all five cycles). The icons below the data show sample patterns with modulation applied to various fractions of the contours. (A) Data for a sample observer (mixed amblyope with 2 diopters of anisometropia and strabismus). The data for her nonamblyopic eye show the typical pattern (see Figure 7), with an initial shallow improvement in sensitivity with increasing number of modulated cycles followed by stark improvement when the entire pattern is modulated. Data for the amblyopic eye follow the shallow improvement throughout. It appears that the stark improvement, presumably due to efficient global signal integration, is absent for the amblyopic eye, suggesting that local processes limit sensitivity. (B) Data for a patient with a chorioretinal scar. The fundus picture shows the location and the extent of the scar, extending superior-temporally from the blind spot. The affected eye shows lower sensitivity across the entire range of modulated cycles. Under the assumption that few modulation cycles are essentially encoded by local processes, a deficit for, e.g., a single cycle would be expected in a case with early retinal pathology. This deficit is, however, amplified when the entire shape is modulated. As in the case of amblyopia, performance for the affected eye shows a lack of global summation that is evident for the fellow eye.
with and without migraine can be based on increased internal noise, a possible consequence of cortical hyperexcitability (McKendrick & Badcock, 2004b; Wagner et al., 2013). People with migraine perform worse than people without in a range of visual tasks, consistent with increased levels of internal noise. These include low-level tasks such as discrimination of spatial frequency (Shepherd, 2000), orientation (Wilkinson & Crotogino, 2000), flicker (McKendrick & Badcock, 2004a), color (McKendrick, Ciolfi, & Johnson, 2002; Shepherd, 2006), luminance (Wagner, Manahilov, Loffler, Gordon, & Dutton, 2010; Webster, Dickinson, Battista, McKendrick, & Badcock, 2011), and long-range inhibition (Wagner, Manahilov, Gordon, & Storch, 2012).

Deficits in migraine are not limited to low-level processing. Studies targeting extrastriate processing have also shown impaired sensitivity in migraine (e.g., McKendrick, Badcock, Badock, & Gurgone, 2006), including shape processing (Ditchfield, McKendrick, & Badcock, 2006; McKendrick, Badock, & Gurgone, 2006; Wagner et al., 2013). A recent study measured shape discrimination on a population of people with migraine (Wagner et al., 2013). The task required the discrimination of RF patterns in the absence and presence of backward masking. Backward masking describes the phenomenon in which the sensitivity to a test stimulus is reduced when a target is followed by a second, task-irrelevant masking stimulus. Testing performance with and without a mask allows experimenters to distinguish between defective inhibition and increased internal noise as possible causes for defects in migraine.

The magnitude of backward masking typically depends on the timing of the mask (Figure 10): Little masking is seen when the target and mask are presented simultaneously (signal onset asynchrony [SOA] = 0) or when they are sufficiently separated in time. Strong masking occurs for SOAs of about 100 ms. The resulting masking curve exhibits a typical inverted U-shape as a function of SOA. Under the assumption that masking is a consequence of inhibition, the reduced-inhibition hypothesis would predict that the performance of people with migraine is similar to that of people without migraine without a mask but better with a mask. If, on the other hand, people with migraine have raised internal noise levels, they should perform more poorly than people without in both conditions. The results supported neither prediction. People with migraine (especially those who experience visual auras) performed slightly worse than people without in the absence of a mask (Figure 10A). Even between attacks, people with migraine are less sensitive to shapes than are people without. This subtle difference is amplified when patterns are backward masked (Figure 10B). This is consistent with an extrastriate deficit in migraine that cannot be completely explained by defective inhibition or raised internal noise.

Preterm children

Prematurely born babies are at risk of a wide range of health-related problems, including visual impairment. Visual impairment in preterm and very-low-

Figure 10. Shape discrimination thresholds for a group of people with migraine who have visual auras (MA), people with migraine who do not have auras (MO), and age-matched people without migraine. (A) Thresholds for discriminating an RF5 test from a circular reference (see insets) are slightly elevated in people with migraine, significantly for those who experience visual auras (MA). (B) The same task, but in the presence of a mask (RF pattern of the same frequency) presented at different times relative to the target contour (SOA). Typical backward masking behavior is evident from the control group: The target is substantially harder to discriminate when the mask follows the target, with maximum masking occurring at an SOA of 66–100 ms. People with migraine show the same pattern, but the magnitude of the masking deficit is significantly amplified, especially for the MA group and at SOAs at which the mask is most disruptive. Adapted with permission from Wagner et al. (2013).
birth-weight babies can be caused by defects in the ocular structures but may also occur in the absence of ocular pathology. The latter includes damage to the optic radiations (periventricular white matter lesions; Fazzi et al., 2004) or other retrogeniculate brain structures, resulting in cerebral visual impairment (Good, Jan, Burden, Skoczenski, & Candy, 2001).

Although up to a third of children with cerebral visual impairment are born prematurely (Dutton & Jacobson, 2001; Milligan, 2010), the prevalence of visual dysfunction in premature children is less clear, in part due to the heterogeneous nature of this group. In very early preterm children, visual problems have been documented (Atkinson & Braddick, 2007). These children typically also present with a range of nonvisual problems and represent the most affected group.

Visual function in the less affected preterm children has received little attention. Understanding any visual problems offers the potential to advise on coping strategies and thereby minimize the impact of those problems on many aspects of life, including education. To establish the extent of visual problems in that population, a recent study targeted a group of high-functioning preterm children, all attending mainstream school (Macintyre-Beon et al., 2013).

As a group, the premature children performed marginally below the level of the age-matched controls (Figure 11). However, many of the preterm children performed within the typical range. Hence, considering preterm children as a group camouflages the fact that the poorer overall performance is essentially driven by a fraction who perform considerably outside the typical range. To further investigate this, a cluster analysis was conducted, based on a question inventory that was completed by the preterm children or their parents (Macintyre-Beon et al., 2013). The inventory contained a range of questions probing specific behavioral features of cerebral visual dysfunction. For example, one of the questions asked, “Does your child trip over toys and obstacles on the floor?” This question aimed to address the presence of a visual field impairment or impairment of visual attention. Another question (“Does your child get lost in places where there is a lot to see?”) targeted difficulties with handling complexity in a visual scene (Macintyre-Beon et al., 2013).

Answers to these questions were “never” or “rarely” for the majority of the preterm children. However, the answers for about a third of the children revealed visual problems. Relating the performance in the visual tests to the questionnaire showed a strong correlation. While the former cohort performed as well as the control group on the visual tests, the latter showed substantially poorer sensitivity. Hence, a substantial fraction of preterm children experience considerable visual and visual-attention problems (including motion and face as well as shape perception) in the absence of impaired low-level vision (visual acuity, contrast sensitivity).
Detailed analysis suggests a prevalence of cortical visual problems in 21%–47% (95% confidence interval) of high-functioning prematurely born children attending mainstream school (MacIntyre-Beon et al., 2013).

Summary

A central question in this article concerns the transformation of basic local image characteristics—such as contour orientation—to representations of complex image properties such as face identity. The evidence favors an intermediate representation that is tuned to object shape. In typical observers, tasks requiring the detection and discrimination of shapes reveal global processing that achieves remarkably high sensitivities by efficiently integrating information widely across space. This behavior is inconsistent with the limits imposed by early detectors and their immediate neighbors. Instead, it is suggestive of operations at intermediate levels of visual processing within extrastriate regions. Results are consistent with the existence of a range of shape channels, tuned to simple contour configurations. Such channels would permit the visual system to decompose more complex contours and represent them by means of a population code of simpler components.

If object shape were encoded by intermediate stage mechanisms, one might expect these to be vulnerable to damage. In support of this premise, deficits for shape detection and discrimination have been shown in a diverse range of developmental and acquired conditions. These include peripheral pathology (macular disorders), abnormalities in visual development (amblyopia, premature birth), and neurological conditions (migraine). Measurable deficits in a wide range of conditions indicate that shape detection and discrimination can be utilized as powerful tests to detect abnormality in midlevel visual function. In some cases, the defect may lie within midlevel processing, whereas in others it may be a consequence of lower level defects that become apparent or amplified in testing shape sensitivity. Wherever the origin of the defect, the exquisitely high sensitivity of these processes to tests promises advantages in their applications as tools to detect and monitor diseases and the effects of intervention.

Keywords: shape perception, midlevel vision, detection/discrimination, shape and contour

Acknowledgments

Much of the work described here has been the result of collaborations, and thanks are due to David Badcock, Stewart Baird, Jason Bell, David Bennett, Mhairi Day, Gordon Dutton, Gael Gordon, Hussein Ibrahim, Graeme Kennedy, Andrew Logan, Catriona MacIntyre-Beon, Emily McGarva, Harry Orbach, Gunnar Schmidtmann, Doreen Wagner, Fran Wilkinson, and Hugh Wilson. I am grateful to Gael Gordon and Harry Orbach for valuable comments on an early draft of the manuscript and to Jon Peirce for organizing the symposium that ultimately led to this special issue.

Commercial relationships: none.
Corresponding author: Gunter Loffler.
Email: G.Loffler@gcu.ac.uk.
Address: Department of Life Sciences, Glasgow Caledonian University, Glasgow, UK.

References


The detecting of structure in Glass


Kovacs, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure ground segmentation. Proceeding of the National Academy of Sciences, USA, 90(16), 7495–7497.


