Relating spatial and temporal orientation pooling to population decoding solutions in human vision

Ben S. Webb *, Timothy Ledgeway, Paul V. McGraw

Visual Neuroscience Group, School of Psychology, University Park, University of Nottingham, Nottingham, NG7 2RD, UK

Abstract

Spatial pooling is often considered synonymous with averaging (or other statistical combinations) of local information contained within a complex visual image. We have recently shown, however, that spatial pooling of motion signals is better characterized in terms of optimal decoding of neuronal populations rather than image statistics (Webb et al., 2007). Here we ask which computations guide the spatial and temporal pooling of local orientation signals in human vision. The observers' task was to discriminate which of two texture patterns had a more clockwise global orientation. Standard textures had a common orientation; comparison textures were chosen independently from a skewed (asymmetrical) probability distribution with distinct spatial or temporal statistics. We simulated observers' performance using different estimators (vector average, winner-takes-all and maximum likelihood) to decode the orientation-tuned activity of a population of model neurons. Our results revealed that the perceived global orientation of texture patterns coincided with the mean (or vector average read-out) of orientation signals accumulated over both space and time. To reconcile these results with our previous work on direction pooling, we varied stimulus duration. Perceived global orientation was accurately predicted by a vector average read-out of orientation signals at relatively short stimulus durations and maximum likelihood read-out at longer durations. Moreover, decreasing the luminance contrast of texture patterns increased the duration of the transition from a vector average to maximum likelihood read-out. Our results suggest that direction and orientation pooling use similar probabilistic read-out strategies when sufficient time is available.

1. Introduction

The cortex accumulates sensory evidence from early visual areas in order to form purposeful decisions and initiate motor commands. To overcome the ambiguity inherent within early, noisy neural representations, cortical pathways combine (‘pool’) incoming signals. The visual system has to strike a delicate balance between combining signals from a common surface and segregating signals from the other surfaces and objects. Yet there still remains little consensus on the precise nature of the computations which govern how local visual signals are pooled across space and time.

One simple solution with substantial empirical support is that the visual system averages incoming signals in order to guide perception. (Ariely, 2001; Chong & Treisman, 2003; Cohen, Singh, & Maloney, 2008; Dakin, 1999; Dakin & Watt, 1997; Morgan, Chubb, & Solomon, 2008; Morgan, Ward, & Castet, 1998; Motoyoshi, Nishida, Sharan, & Adelson, 2007; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001; Pavlovskaya, Vol, & Blum, 1992; Sharan, Li, Motoyoshi, Nishida, & Adelson, 2008; Watamaniuk & Duchon, 1992; Watt, Morgan, & Ward, 1983; Whitaker, McGraw, Pacey, & Barrett, 1996; Williams & Sekuler, 1984; Wilson, Ferrera, & Yo, 1992; Zohary, Scase, & Braddick, 1996). For example, perception of a moving surface, such as a field of dots where local motion is pooled across space or a plaid pattern where the components are pooled across orientation, frequently coincides with the vector average direction or velocity of the local samples (Kim & Wilson, 1993; Mingolla, Todd, & Norman, 1992; Watamaniuk & Duchon, 1992; Williams & Sekuler, 1984; Wilson & Kim, 1994; Wilson et al., 1992; Zohary et al., 1996). Following earlier pioneering work on reaching direction (Georgopoulos, Kettner, & Schwartz, 1988; Georgopoulos, Schwartz, & Kettner, 1986), many physiological studies have demonstrated that neurons in the motion pathway compute a vector average of velocity signals in order to guide ocular following and smooth pursuit eye movements. (Ferrera, 2000; Groh, Born, & Newsome, 1997; Huang & Lisberger, 2009; Lisberger & Ferrera, 1997; Masson, 2004; Recanzone & Wurtz, 1999; Wallace, Stone, & Masson, 2005; Yang & Lisberger, 2009). In the spatial domain, texture perception also frequently coincides with spatial summary statistics.
statistics (Balas, 2006; Balas, Nakano, & Rosenholtz, 2009; Beck, 1983; Dakin, 1999; Dakin & Watt, 1997; Julesz, 1981; Keeble, Kingdom, Moulden, & Morgan, 1995; Kingdom, Hayes, & Field, 2001; Morgan et al., 1998, 2008; Parkes et al., 2001; Voorhees & Poggio, 1988). Analogous statistical processes have been invoked to explain other visual decisions, including the precision with which humans judge relative position (Watt et al., 1983; Whitaker et al., 1996), tilt (Morgan et al., 1998; Parkes et al., 2001), size (Ariely, 2001; Chong & Treisman, 2003) and surface reflectance (Motoyoshi et al., 2007; Sharan et al., 2008).

Although appealing because of its simplicity (Salinas & Abbott, 1994), linear operations like vector averaging can be biased estimators if the underlying detectors are irregularly spaced or narrowly tuned (Seung & Sompolinsky, 1993; Snippe, 1996) or the local samples are distributed asymmetrically (Webb, Ledgeway, & McGraw, 2007). For local orientation pooling with some types of moving plaid stimuli, the computations may be better characterized by the intersection of constraints (IOC) rule (Adelson & Movshon, 1982; Albright, 1984; Fennema & Thompson, 1979; Movshon, Adelson, Gizzi, & Newsome, 1986; Weiss, Simoncelli, & Adelson, 2002). The IOC is an accurate mathematical description of rigid motion, but is limited by its inability to explain non-rigid forms of motion.

An alternative to averaging and IOC is a “winner-takes-all” (WTA) or “max” rule which selects the preferred stimulus of a neuron or detector with the strongest response. This form of non-linear pooling has been successfully applied to many domains, including both spatial and motion processing (Anstis, 2009; Baldassi & Burr, 2004; Baldassi, Megna, & Burr, 2006; Baldassi & Verghe, 2002; Gheri & Baldassi, 2008; Palmer, 1994; Palmer, Ames, & Lindsey, 1993; Palmer, verghe, & Pavel, 2000; Salzman & Newsome, 1994; Shaw, 1980, 1982; Shiu & Pashler, 1995; Solomon, Lavie, & Morgan, 1997; Verghe & Stone, 1995; Webb et al., 2007). Unlike averaging, the fidelity of a WTA estimate is much less susceptible to changes in the spacing and number of detectors (Shamir, 2006). Moreover, it is simple to implement (Baldassi &

---

**Fig. 1.** Simulation of global orientation discrimination. We simulated trial-by-trial performance on a global orientation discrimination task. A bank of orientation tuned neurons responds to an asymmetrical distribution of orientations with a Gaussian sensitivity profile corrupted by Poisson noise. From the population response, we derive the maximum likelihood, winner-take-all and vector average read-out of orientation signals accumulated over space and time.
Vergheze, 2002) and often as least as accurate at predicting psychophysical behavior as so called “optimal decoders”, though the estimates themselves tend to be more variable (Webb et al., 2007).

A theoretical limitation with all of the above decoding solutions is that they collapse the distributed activity of a population of neurons down to a single value to represent the “best estimate” of a stimulus. Extracting a singular estimate may not be optimal, or even desirable, under all circumstances. For example, representing multi valued stimuli, such as certain forms of transparent motion, where at least two directions can be detected at any one time (Anderson, 1989; Edwards & Greenwood, 2005) may require a more principled decoding strategy (Treue, Hol, & Rauber, 2000).

A more parsimonious formulation might be to frame pooling as a statistical inference problem (Beck et al., 2008; Deneve, Latham, & Pouget, 1999; Foldiak, 1993; Gold & Shadlen, 2001; Jazayeri & Movshon, 2006; Knill & Pouget, 2004; Ma, Beck, Latham, & Pouget, 2006; Paradiso, 1988; Pouget, Dayan, & Zemel, 2000, 2003; Pouget, Zhang, Deneve, & Latham, 1998; Sanger, 1996; Seung & Sompolinsky, 1993; Weiss & Fleet, 2002; Zemel, Dayan, & Pouget, 1998), since this allows the cortex to compute and infer the probability that a wide range of stimuli are consistent with a neural response. Optimal decoding of the distributed activity across a population of neurons can then be computed as a likelihood function, which represents the probability that each of a range of stimuli gave rise to the neural response. With access to the full likelihood function, population decoders are efficient, unbiased estimators of performance on a wide range of perceptual tasks (Deneve et al., 1999; Foldiak, 1993; Paradiso, 1988; Sanger, 1996; Seung & Sompolinsky, 1993; Weiss & Fleet, 2002). The maximum likelihood decoder, for example, accurately predicts orientation discrimination (Regan & Beverley, 1985), perceived direction (Webb et al., 2007), perceived velocity (Weiss et al., 2002) and cue combination both within (Jacobs, 1999; Landy, Maloney, Johnston, & Young, 1995) and across modalities (Alais & Burr, 2004; Ernst & Banks, 2002).

We have developed a psychophysical paradigm that uses asymmetrical distributions of local visual signals to distinguish the contribution of different putative computations. Adopting this approach, we recently demonstrated (Webb et al., 2007) that spatial pooling of motion signals is poorly estimated by a vector average decoder, but accurately predicted by a maximum likelihood read-out of direction signals combined over space (see Fig. 1). Here we extend this paradigm to ask which class of algorithms guides the spatial and temporal pooling of local orientation signals in human vision. Our results suggest that orientation pooling uses different decoding strategies at different time scales.

2. Methods

2.1. Observers

Four observers with normal vision participated. Three were authors (BSW, PVM, TL) and one (HL) was naive to the purpose of the experiments.

2.2. Stimuli

Static and dynamic texture patterns were generated on a PC computer using software written in Python using components of Psychopy (Peirce, 2007). We displayed the texture patterns on a CRT monitor (LaCie Electron 22 Blue II or iiyama Pro Vision Master 514) at a viewing distance of 76.3 cm, resolution of 1280 x 1024 pixels and update rate of 75 Hz. Each texture pattern (see Fig. 2) was composed of 500 Gaussian lines (peak luminance ~80 cd/m², line envelope SD was 0.166 x 0.083°) presented on a uniform background (luminance ~40 cd/m²) within a circular window (diameter 12°). Static textures consisted of one image; dynamic textures consisted of 25 images displayed consecutively at 18.75 Hz (0.052 s image duration). On each image of a dynamic sequence, lines were randomly positioned inside the circular window at non-overlapping locations.

2.3. Procedure

We used static and dynamic texture patterns with distinct spatial and temporal statistics, respectively. Except where stated, the procedures with both forms of texture were the same. In a temporal two-alternative forced choice task, observers judged which of two textures had a more clockwise global orientation. On each trial, we presented a standard and comparison texture in a random temporal order. Static and dynamic textures were presented for 0.052 s (1 image) and 1.3 s (25 images), respectively and separated by 0.5 s interval containing a fixation cross on a uniform background. Lines in the standard texture had a common orientation, randomly assigned on each trial from a range spanning 180°. Line orientations in the comparison texture were chosen, with replacement, from either a symmetrical or asymmetrical (skewed) probability distribution with distinct measures of central tendency. In the static and dynamic patterns, orientations were sampled from the probability distributions over space and time, respectively. A schematic of the task is shown in Fig. 2.

To investigate spatial and temporal pooling of orientation, we conducted each of the following experiments with static and dynamic textures patterns, respectively. In the first experiment, line orientations of the comparison texture were discretely sampled (peak luminance ~80 cd/m², line envelope SD was 0.166 x 0.083°) presented on a uniform background (luminance ~40 cd/m²) within a circular window (diameter 12°). Static textures consisted of one image; dynamic textures consisted of 25 images displayed consecutively at 18.75 Hz (0.052 s image duration). On each image of a dynamic sequence, lines were randomly positioned inside the circular window at non-overlapping locations.
distributed line orientations. The standard deviation of the counter clockwise half of the Gaussian was 15°, 20°, 25° or 30°; the corresponding values for the clockwise half were 15°, 10°, 5°, or 0°.

In the second experiment, line orientations of the comparison texture were discretely sampled from a Gaussian with standard deviations of 15°, 25°, 35° or 45° for the clockwise and 6°, 10°, 14°, 18° for the counter clockwise halves. We sampled the counter clockwise and clockwise halves of the distribution at 2.5° and 0.5° intervals, respectively. This generated asymmetrical distributions of line orientations with the same mode and median but a different mean. For both experiments, the difference between the orientation of the standard texture and modal orientation of the comparison texture was varied according to the method of constant stimuli with nine levels.

In the third experiment, for the comparison texture we generated a uniform distribution of line orientations with a total range of 90°. We assigned each half of the distribution (i.e., orientations clockwise and counter clockwise to the median direction) a different range and sampling density. Line orientations for the counter clockwise half of the distribution were sampled at 2.5° intervals over a range of 45°, 55°, 65° or 75°; the corresponding values for the clockwise half were sampled over a range of 45°, 35°, 25° or 15°. This generated asymmetrical distributions of orientation with a different mean and median. The median orientation of the comparison was randomly chosen on each trial using the method of constant stimuli.

In the final two experiments, observers judged whether a static texture pattern composed of a distribution of orientations (shown in Fig. 6A; chosen from pilot work as diagnostic for distinguishing orientation space) was oriented clockwise or counter clockwise of implicit vertical. Each pattern was presented at seven durations, ranging between 0.05 and 3.33 s in logarithmic steps, and at three Michelson contrasts (0.25, 0.5, and 1). Global orientation was controlled via a method of constant stimuli.

2.4. Data analysis

For each condition, observers completed a minimum of 4 runs of 180 trials. Data were expressed as the percentage of trials on which observers judged the modal (exp. 1 and 2) or median (exp. 3–5) orientation of the comparison as more clockwise than the standard as a function of the angular difference between them and fitted with a logistic function:

\[
y = \frac{100}{1 + e^{x-y}}/\theta, \tag{1}
\]

where \(y\) is the percentage of clockwise judgements, \(\mu\) is the stimulus level at which observers perceived the orientation of the standard and comparison to be the same, and \(\theta\) is an estimate of discrimination threshold.

2.5. Simulations

We simulated observers’ performance on a trial-by-trial basis using the same stimulus parameters and methods described in the psychophysical procedure. The spacing and bandwidth of neurons in our model were chosen to give sufficient coverage of the orientation space. The model (shown in Fig. 1) consists of a bank of evenly spaced orientation tuned neurons spanning a 180° range. Each neuron responds to a limited range of orientations with a Gaussian sensitivity profile corrupted by Poisson noise. The separation between adjacent neurons was fixed at 1°. The sensitivity of the \(i\)th neuron, centered at \(\theta_i\), to orientation \(\theta\) is:

\[
S_i(\theta) = \exp \left( -[(\theta - \theta_i)^2] / h^2 \log 2 \right) \tag{2}
\]

where \(h\) is the bandwidth (half-height, half-width), fixed at 22.5° (David, Hayden, & Gallant, 2006). The response of the \(i\)th mechanism to stimulus \(Or\) with a distribution of orientations \(Or(\theta)\) is:

\[
R_i(Or) = k \sum_{\theta=1}^{180} S_i(\theta) pr(Or(\theta)) \quad \text{where } k = R_{\text{max}}t \tag{3}
\]

\(R_{\text{max}}\) is the maximum mean firing rate of the neuron in spikes/s (60), \(t\) is stimulus duration and \(\text{pr}(Or(\theta))\) is the proportion of orientations in the stimulus. The number of spikes \(n_i\) elicited in response to a stimulus on a given presentation is Poisson distributed with a mean of \(R_i(Or)\)

\[
p(n_i|Or) = \frac{R_i(Or)^{n_i}}{n_i!} \exp \left( -R_i(Or) \right) \tag{4}
\]

The log likelihood of any stimulus \(Or\) is computed as a weighted sum of the responses of the population of neurons, where the activity of each neuron is multiplied by the log of its tuning function (Jazayeri & Movshon, 2006; Seung & Sompolinsky, 1993):

\[
\log L(Or) = \sum_{i=1}^{180} n_i \log R_i(Or) \tag{5}
\]

The estimated orientation is the value of \(\theta_i\) for which \(\log L(Or)\) computed for all \(Or\) is maximal. To obtain the estimated \(Or\) of the comparison from a winner-take-all decoder, we read-off the value of \(\theta_i\) where \(n_i\) max. To obtain the corresponding estimate from a vector average decoder we calculated the average of the preferred orientation of all neurons weighted in proportion to their response magnitude:

\[
V_{\text{av}}(Or) = \tan^{-1} \left( \frac{\sum_{i=1}^{180} n_i \sin(\theta_i)}{\sum_{i=1}^{180} n_i \cos(\theta_i)} \right) \tag{6}
\]

3. Results

In the first set of experiments we investigated the pooling of orientation signals across space. We generated comparison texture patterns composed of asymmetrical distributions of orientations across space with distinct measures of central tendency. To quantify the relationship between perceived global orientation and different statistical measures of central tendency in the comparison stimulus, we estimated the point of subjective equality – the stimulus level at which observers perceived the global orientation of the comparison and standard texture to be the same (see Section 2). These data are plotted in Fig. 3 as a function of the clockwise standard deviation (A and B) or range (C) of the comparison. The perceived global orientation corresponded very closely to the mean orientation of the lines in the comparison texture. When the orientations in the comparison were drawn from a Gaussian with a clockwise standard deviation (SD) of 30° and a counter clockwise SD of 0° (comparison texture shown in Fig. 2), the modal direction of the comparison (represented by the dotted line in Fig. 3A) had to be rotated, on average, by approximately 20° to be indistinguishable from the standard (Fig. 3A). Similarly, when the comparison orientation distribution had a clockwise SD of 45° and a counter clockwise SD of 18°, the modal orientation of the comparison had to be rotated by approximately 10° to be indistinguishable from the standard (Fig. 3B). Similar effects were obtained with a skewed uniform distribution of orientations in the comparison. When the clockwise ranges were 75° and 15°, respectively, the median orientation (represented by the dashed line in Fig. 3C) had to be rotated by 13°, on average, to be indistinguishable from the standard. It is noteworthy that in Fig. 3B and C three subjects...
deviate slightly from the mean orientation of the comparison for the largest clockwise SD and range, respectively. The most likely reason is that some subjects reported small amounts of transparency (i.e. break down of the global orientation structure) at the largest SD and ranges.

We ran two control conditions to establish whether a vector average read-out holds when we added some orientation uncertainty to the standard stimulus. In two key conditions (comparison: counter clockwise SD 30° and clockwise SD 0°; comparison:

counter clockwise range 75° and CW range 15°) the standard orientations were sampled from either a symmetrical Gaussian with a standard deviation of 30° or a uniform distribution with a range of 90°. The results for four observers are plotted in Fig. 3A and C (squares on far right) and are quantitatively the same as we found when the standard was composed of a common orientation (Fig. 3A and C, circles on far right).

We used exactly the same methods and analysis as described above in the temporal pooling experiments, with the exception that the texture patterns were dynamic and the orientations in the comparison were asymmetrically distributed over time rather
than space. The perceived global orientation accumulated over time (point of subjective equality) is plotted in Fig. 4A–C as a function of the clockwise SD or range of the comparison texture. Although the data are slightly more variable, it is striking how similar the results are to the spatial case. The perceived global orientation corresponds very closely to the mean line orientation accumulated over time, diverging substantially from both the modal (Fig. 4A and B) and median temporal statistics (Fig. 4B and C).

We simulated observers’ performance on all of the experiments with a simple model, schematically illustrated in Fig. 1 and described in detail in Section 2. We read-off the perceived global orientation of texture patterns on each trial with a vector average, winner-takes-all and maximum likelihood decoder. The pattern of results we find in both the spatial and temporal experiments is accurately predicted by our simulations with a vector average, but not a maximum likelihood or winner-takes-all read-out of orientation signals. The results of the simulations for the spatial experiment are plotted in Fig. 5A–C. We only plot the estimates of the vector average (black squares) and maximum likelihood decoder, since winner-takes-all generated qualitatively similar predictions to maximum likelihood, but with higher variance. Open circles show the perceived global orientation in the spatial experiment for the four observers (mean ± SD) plotted and notated as in Figs. 3 and 4. The vector average response of the model neurons clearly provides a more accurate estimate of perceived global orientation than the other decoders. These results contrast with our previous work in which we found that both maximum likelihood and winner-takes-all provided a robust guide to the perceived direction of global motion (Webb et al., 2007).

In Fig. 5D, we plot the results of the experiment in our previous study which was most diagnostic for distinguishing the predictions of a vector average and maximum likelihood read-out of motion direction. The perceived direction of motion of four observers (open circles; mean ± SD) is plotted as a function of the ranges of the direction distributions used for the comparison stimulus. The filled circles show the global motion direction estimated by a maximum likelihood decoder, which clearly corresponds very closely to the perceived direction of the observers. It is notable how different these data are for the perceived direction of global dot motion (Fig. 5D) compared to that for global orientation (Fig. 5C), even though we used analogous stimulus distributions and psychophysical procedures in both studies.

One important difference between the current and previous study, however, was that we used very different stimulus durations for the global motion and the orientation experiments. To try and reconcile these apparently conflicting results, we ran a simple experiment in which we varied stimulus duration. The observers’ task was to judge whether a static texture pattern composed of a skewed uniform distribution of orientations (shown in Fig. 6A) was oriented clockwise or counter clockwise of implicit vertical. We chose this distribution because in the model simulations maximum likelihood and vector average decoders estimated the perceived global orientation to be clockwise and counter clockwise of vertical, respectively. In Fig. 6B, we plot for four observers the proportion of clockwise (maximum likelihood) and counter clockwise (vector average) judgements as a function of stimulus duration. We have not presented the corresponding prediction of a winner-takes-all decoder because it is very similar to that of maximum likelihood. Perceived global orientation was biased towards

---

**Fig. 5.** (A–C) Model simulations of the spatial pooling experiment. Plotted as in the same manner as Fig. 3. Vector average (black squares) and maximum likelihood (gray circles) estimate of the perceived global orientation of four observers (open circles; mean ± SD) in the spatial pooling experiment. (D) From Webb et al. (2007), perceived direction of motion of four observers (open circles; mean ± SD) plotted as a function of the ranges of the direction distributions used for the comparison stimulus. Gray circles show the global motion direction estimated by a maximum likelihood decoder.
Figure 6. Biases in perceived global orientation as a function of stimulus duration and luminance contrast. (A) Skewed uniform distribution of orientations for which maximum likelihood and vector average decoders estimate the perceived global orientation to be clockwise and counter clockwise of vertical, respectively. (B) Colored symbols show for four observers the proportion of clockwise (maximum likelihood) and counter clockwise (vector average) judgements plotted as a function of stimulus duration. (C) Gray symbols show for four observers (mean ± SEM) how luminance contrast modulates the duration of the transition from a vector average to maximum likelihood read-out.

4. Discussion

We have examined which computations accurately predict the perceived global orientation of signals accumulated over space and time. Using a global orientation discrimination task, we found that the perceived global orientation of texture patterns coincided with the mean (or vector average read-out) of orientation signals. This result is consistent with a large body of work demonstrating that different forms of texture perception are well characterized by image-based, summary statistics (Balas, 2006; Balas et al., 2009; Beck, 1983; Dakin, 1999; Dakin & Watt, 1997; Julesz, 1981; Keeble et al., 1995; Kingdom et al., 2001; Morgan et al., 1998, 2008; Parkes et al., 2001; Voorhees & Poggio, 1988). We have, however, previously shown that spatial pooling of local samples might be better characterized in terms of optimal decoding of neuronal populations rather than image-based statistics (Webb et al., 2007). To reconcile the current results with our previous work, we varied the stimulus duration of texture patterns. Perceived global orientation was accurately predicted by a vector average read-out of orientation signals at relatively short stimulus durations and maximum likelihood read-out at longer durations. Moreover, decreasing the luminance contrast of texture patterns increased the duration at which the read-out translated from vector average to maximum likelihood.

The dynamics of the read-out of global orientation are reminiscent of earlier psychophysical work with certain types of two-dimensional motion (Cropper et al., 1994; Lorenceau, Shiffrar, Wells, & Castet, 1993; Yo & Wilson, 1992). At short stimulus durations (e.g. less than 90 ms) type II plaids are perceived moving in a vector average direction, whereas at longer durations they are perceived to move in a direction predicted by an IOC rule (Cropper et al., 1994; Yo & Wilson, 1992). The transition from vector average to an IOC direction of plaid motion over time is also modulated by luminance contrast (Yo & Wilson, 1992). Subsequent work demonstrated that many of these dynamic two-dimensional motion effects (Bows, 1996; Burke & Wenderoth, 1993; Lorenceau et al., 1993; Stone & Thompson, 1992; Stone, Watson, & Mulligan, 1990; Yo & Wilson, 1992) can be explained within a Bayesian framework, provided one assumes that slow speeds are more likely (Weiss et al., 2002). We cannot rule out the possibility that an orientation equivalent of this framework with suitable, statistical a priori assumptions could predict our results.

One potential criticism of our interpretation is that the bias towards a maximum likelihood read-out at long stimulus durations can be explained by adaptation of neural responses in early visual cortex (Dragoi, Sharma, & Sur, 2000; Maffei, Fiorentini, & Bisti, 1973; Movshon & Lennie, 1979; Muller, Metha, Krauskopf, & Lennie, 1999). It is certainly true that the longer stimulus durations will have caused adaptation of the population response in early visual cortex. However, this would predict a relative reduction of the bias to the denser part of the orientation distribution (shown in Fig. 6A), causing a bias away from a maximum likelihood read-out at long durations. Similarly, surround suppression of neural responses in early visual cortex (Blakemore & Tobin, 1972; Cavanaugh, Bair, & Movshon, 2002; DeAngelis, Freeman, & Ohzawa, 1994; Webb, Dhruv, Solomon, Taliby, & Lennie, 2005) cannot account for the bias at long durations, since this would also reduce responses most to the denser part of the distribution.

Analogous changes to the nature of the pooling computation have also been found to occur over different temporal scales in single neurons in the middle temporal (MT) area (Pack & Born, 2001; Smith, Majaj, & Movshon, 2005). In these studies, the responses of
individual neurons were initially dominated by the component directions of a plaid pattern, and much later in the response by the pattern, or plaid direction of motion. These results closely mirror the psychophysics (Cropper et al., 1994; Yo & Wilson, 1992) and reinforce the notion that global stimulus selectivity takes time to evolve. At present, it is unknown whether or not the perceived direction of asymmetrical distributions of global motion direction is predicted by different forms of read-out at different stimulus durations. Based on the current results, we would predict that a vector average computation might be a better estimator of the perceived direction of global dot motion at very short stimulus durations. Ongoing work in our laboratory is testing this prediction (see below).

The concept of adaptive pooling – a flexible process in which the visual system recruits different estimators to address the prevailing computational demands – is gaining prominence in the literature (Amano, Edwards, Badcock, & Nishida, 2009; Bowns & Alais, 2006; Ferrera, 2000; Huang, Albright, & Stoner, 2007; Liu & Wang, 2008; Nichols & Newsome, 2002; Pack, Berezovskii, & Born, 2001; Pack & Born, 2001; Recanzone & Wurtz, 1999; Zohary et al., 1996). In a recent psychophysical demonstration of this phenomenon, Amano and colleagues found that the computations mediating spatial pooling of motion signals depend upon the available stimulus information. They showed that the visual system pools ambiguous local direction signals simultaneously across orientation and space in a manner consistent with IOC, whereas unambiguous motion is first pooled locally across orientation and then pooled globally across space according to a vector average computation. This reinforces the notion that pooling may not be a rigid process, but rather depends upon the nature of the stimulus and task demands.

Our current and previous results (Webb et al., 2007) suggest that spatial pooling of local direction and orientation use similar probabilistic read-out strategies when sufficient time is available. When time is limited, the visual system appears to adopt a parsimonious, but potentially biased read-out of visual signals. One possibility is that it generates its “best estimate” based upon the limited information available, but resorts to optimal pooling over longer time scales. For example, a vector average decoder might be deployed during the early phase of a neural response to rapidly initiate ocular motor systems (Ferrera, 2000; Groh et al., 1997; Huang & Lisberger, 2009; Lisberger & Ferrera, 1997; Masson, 2004; Osborne & Lisberger, 2009; Recanzone & Wurtz, 1999; Wallace et al., 2005; Yang & Lisberger, 2009) before complex stimulus selectivity has time to evolve (Cropper et al., 1994; Pack & Born, 2001; Smith et al., 2005; Yo & Wilson, 1992). Whereas, optimal, non-linear pooling can take time to evolve because of the dynamics of the underlying neural computations. Neural networks have to mitigate the effects of moment to moment fluctuations (i.e. neural noise) in the system, and it can take several iterations of activity before a recurrent network approaches an optimal decoding solution (Deneve et al., 1999).

An intriguing aspect of our results is the smooth transition from a vector average to maximum likelihood read-out over time. This gradual shift might reflect the implementation of two decoding operations that are weighted against each other. This sort of competitive inhibition implemented within a simple recurrent network can account for the gradual shift from a vector average to WTA read-out of ocular motor signals (Ferrera, 2000). An alternative proposal is that the transition might reflect a form of temporal summation. That is, each neuron in the population can only sum a fixed number of samples per unit time before its response reaches saturation – the orientation equivalent of a semi-saturation constant. Preliminary work in our laboratory suggests that implementing this form of temporal summation in our population decoding model can cause a maximum likelihood decoder to gradually change its read-out over time.

Acknowledgment

This research was funded by a Wellcome Trust Research Career Development Fellowship awarded to Ben Webb.

References


