Spatiotopic coding and remapping in humans

David C. Burr1,2,* and Maria Concetta Morrone3,4

1Department of Psychology, University of Florence, Via di San Salvi 12, Florence 50135, Italy
2CNR Institute of Neurosciences, Via Moruzzi 1, Pisa 56100, Italy
3Department of Physiological Sciences, University of Pisa, Via S. Zeno, 31, Pisa 56127, Italy
4Fondazione Stella Maris, Viale del Tirreno 331, Calambrone 56018, Pisa, Italy

How our perceptual experience of the world remains stable and continuous in the face of continuous rapid eye movements still remains a mystery. This review discusses some recent progress towards understanding the neural and psychophysical processes that accompany these eye movements. We firstly report recent evidence from imaging studies in humans showing that many brain regions are tuned in spatiotopic coordinates, but only for items that are actively attended. We then describe a series of experiments measuring the spatial and temporal phenomena that occur around the time of saccades, and discuss how these could be related to visual stability. Finally, we introduce the concept of the spatio-temporal receptive field to describe the local spatiotopicity exhibited by many neurons when the eyes move.

Keywords: vision; eye movements; spatiotopicity; remapping

1. INTRODUCTION

How a stable representation of the world is created from the output of mobile sensors is an old and venerable problem that has fascinated many scientists, including Descartes, Helmholtz, Mach and Sherrington, and indeed goes back to the eleventh century Persian scholar Abū ‘Alī al-Hasan ibn al-Hasan ibn al-Haytham (Latinized ‘Alhazen’): ‘For if the eye moves in front of visible objects while they are being contemplated, the form of every one of the objects facing the eye ... will move on the eyes as the latter moves. But sight has become accustomed to the motion of the objects’ forms on its surface when the objects are stationary, and therefore does not judge the objects to be in motion’ [1]. Although the problem of visual stability is far from solved, tantalizing progress has been made over the last few years.

2. SPATIOTOPICITY

Because of the spatial selectivity of individual neurons, the response of primary visual cortex forms a map [2], similar in principle (except for magnification distortions) to that imaged on the retinae. This retinotopic representation, which changes completely each time the eyes move, forms the input for all further representations in the brain. So a major question is how this retinotopic map becomes transformed into the spatiotopic representation of world that we perceive, anchored in stable real-world coordinates. Electrophysiological studies have shown that neurons in specific areas of associative visual cortex, including V6 [3] and ventral intraparietal area (VIP) [4], show the spatiotopic selectivity that we would expect to exist, with spatial tuning in real-world coordinates, invariant of gaze. Indeed, even primary cortex V1 is modulated to some extent by gaze [5], particularly the peripheral representation [6]. What about humans?

(a) Psychophysical evidence for spatiotopicity in humans

If there exist in the brain neural mechanisms tuned in spatiotopic coordinates, then these mechanisms should respond to stimuli in particular positions in space, irrespective of retinal projection. Melcher & Morrone [7] used a summation technique to investigate the spatiotopicity of neural mechanisms tuned to motion. They took advantage of the long integration time for motion stimuli [8] and showed that observers can integrate motion signals that are individually below threshold (and hence not perceived when presented alone) across saccades. Two periods of coherent horizontal motion, each lasting 150 ms, were presented successively, separated by sufficient time to allow for a saccadic eye movement between them. On some blocks of trials, subjects saccaded across the stimulus between the two motion intervals, while on others they maintained fixation above or below the stimulus. As figure 1 shows, sensitivity was similar in the two conditions, twice that for a single motion stimulus, showing that the two motion signals were integrated across the saccade—but only when the two motion signals were in the same position in space, indicating that the brain must use a mechanism anchored to external rather than retinal coordinates. Not only was trans-saccadic integration possible, it
occurred for longer durations than in fixation, 1.5 s as opposed to nearly 3 s. Importantly, the methodology excluded cognitive strategies or verbal coding as the motion signals presented before and after the saccade were each well below the conscious detection threshold: only by summating the two signals could motion be correctly discriminated. Interestingly, not all psychophysical studies show evidence for a spatio-temporal representation of motion. For example, the motion aftereffect seems to be anchored mainly in retinotopic coordinates [9–11].

Other experiments have shown that many attributes, such as form [12] and position [13], are integrated across saccades. This contrasts sharply with older claims that information is not integrated across saccades [14,15]. We (and many others) suggest that the reason for this is that the information integrated from one fixation to the next is information about features—such as motion and form—not light elements or pixels. As other reviews in this issue point out [16–21], the process is not like ‘sticking postage stamps on a tailor’s dummy’, integrating detailed ‘snapshots’ within a trans-saccadic buffer with an external metric. This would suggest that very early stages of analysis, such as V1, should not be spatiotopic, while higher centres responsible for motion and form (including middle temporal (MT) and medial superior temporal (MST) areas) might be spatiotopic. These areas would integrate motion signals or other elaborated information across saccades, not the detailed description of the individual dots that generated the motion signals.

Spatiotopicity has also been studied with vision after-effects: after inspecting for some time a specific stimulus, such as a slanted grating, a vertical grating presented to the same position will appear to be slanted in the other direction (see [22]). Melcher [23] adapted this technique to study spatiotopicity. His subjects adapted to various stimuli, then made a large saccade before presenting the test stimulus either in the same (spatiotopic) position as the adaptor, or a different, control position. He observed partial spatiotopic adaptation for orientation, form and face perception, pointing to the existence of spatiotopically encoded neural mechanisms for these attributes. On the other hand, contrast

Figure 1. Integration of motion signals across saccades. (a) Illustration of experimental setup. Subjects fixated above a field of dynamic random noise. When cued they saccaded to a target below this field. (b) Timeline of stimuli presentation. For two brief (150 ms) moments, a proportion of the dots moved coherently, both leftwards or rightwards. One motion signal was presented before the saccade, the other after. Subjects had to identify the direction of motion in two-alternative forced choice. (c) Coherence sensitivity (inverse of minimum coherence to support reliable direction discrimination) as a function of separation of the two pulses. The dashed line shows the sensitivity for just a single motion signal. Presenting two stimuli with brief separation, either in fixation or straddling saccades, doubles the sensitivity, implying total integration. The integration continues for longer when interspersed with a saccade than when presented during fixation (filled symbols indicate fixation and open symbols indicate saccade).
Aftereffects did not transfer across saccades at all, whether in the same spatial position or not. More recent experiments suggest that colour adaptation is spatiotopic [24]. Interestingly, the perception of event time is also subject to adaptation [25], and this adaptation has been shown to be almost entirely spatiotopic [9]. The clear implication is that low-level descriptive details of images such as local contrast are not integrated across saccades but high-level descriptions, such as orientation, form, time and colour are built up over saccades.

(b) Physiological evidence for spatiotopicity in humans

Functional magnetic resonance imaging has also indicated the existence of spatiotopic coding in human cortex, in lateral occipital area (LO) [26], an area involved in the analysis of objects, in VIP [27], a multi-sensory area and in MT+ [28]. We [29,30] have studied spatiotopicity of visual cortex by measuring blood oxygen level-dependent (BOLD) responses to random-dot motion stimuli presented to various positions while subjects maintained fixation at one of three different gaze directions (see inset to figure 2). In our first study [30], we reported that area MT, heavily involved in the perception of motion, showed a clear selectivity in external rather than retinal coordinates, whereas primary cortex V1 was retinotopically selective.

More recently, we repeated the experiment under two conditions: passive viewing (as before), where subjects simply maintained fixation (but were free to attend to the motion stimuli), and a dual-task attentive condition, where they performed a demanding detection task at the fovea [29]. Figure 2 shows preliminary data from one example subject. BOLD responses of motion area MT are plotted against the position of the stimulus (in screen coordinates) for three different fixation conditions. The responses at different fixations are strong for both the attention-to-fovea and passive-viewing conditions. However, there is an enormous difference in their spatial selectivity: in the passive-viewing condition the responses were similar for all fixations, tuned spatiotopically; but when attention was concentrated at the fovea, the responses become displaced in the direction of gaze, clearly retinotopic.

From responses like those of figure 2, we calculated a spatiotopicity index for each voxel (see figure 2 caption), which varies from −1 (blue in figure 2) for perfect spatiotopicity to +1 (yellow) for perfect retinotopicity. The insets show how the spatiotopicity of voxels in MT in the same subject are affected by attention. With passive fixation most of the region shows a clearly spatiotopic response. But performing the attention-demanding foveal task causes these same voxels to become retinotopic. Hence, we can draw two conclusions: first, area MT is involved in the perception of motion, showed a clear spatiotopic response. But performing the attention-demanding foveal task these voxels become strongly retinotopic (red/yellow code). The index was similar to that used by Gardner et al. [31]. It is the difference of the squared residuals differences in response amplitude for the three fixation conditions when they are in a spatiotopic alignment (residS) and retinotopic alignment (residR), normalized by the sum of the squared residuals SI = (residS − residR)/(residS + residR). The index is constrained between −1 (full spatiotopicity) and +1 (full retinotopicity).

et al. [31], who claimed that visual cortex is retinotopic rather than spatiotopic, and with our previous report [30], claiming the opposite: in our study subjects were free to attend to the stimuli (and for one experiment were required to attend to the stimuli), while the subjects of Gardner et al. maintained attention at the fovea. Attention is known to modulate BOLD responses in many areas, including V1 and associative cortex, particularly the dorsal pathway [32–36]. Directing attention to the fovea boosts the response to stimuli near the attended target, while suppressing that to irrelevant stimuli distant from the attended location. The effect of attention can even reshape and shift the receptive fields (RFs) of single cells in monkey’s MT [37] and in human V1 [38]. Is attention also instrumental in building spatiotopic selectivity? Evidence shows that attention can be allocated in retinal and spatiotopic coordinates [39,40] and could be an important mechanism mediating spatiotopic coding...
[41,42]. This raises the fascinating possibility that attention is an integral part of spatiotopicity. As it is well known that there exists a close link between attention and eye movements, it is not unreasonable that the two should work together in the creation of spatial maps. There is much evidence that only attended objects are remapped (e.g. [43]), and several other chapters in this issue highlight the importance of attention in visual stability and remapping, particularly Wurtz et al. [44] and Mathôt & Theeuws [45].

How may attention affect the spatial tuning of BOLD responses? One of the more successful models of attentional effects is the normalization model of Reynolds & Heeger [46], which predicts the reshaping and shift of RFs of single cells in monkey’s MT [37]. However, attention away from the fovea, or towards the stimulus, is not in itself sufficient to generate spatiotopicity. To obtain spatiotopic tuning from a normalization model, a gaze signal must contribute to the normalization factor, together with attention. In other words, attention should bind with gaze direction to generate a new functional entity that is spatiotopic. Interestingly gain changes contingent on gaze position (termed ‘gain fields’ [47,48]) have been described in much of associative visual cortex. T o date, no one has investigated whether gain field mechanisms are under attentional control, but it does not seem unreasonable that they should be.

The BOLD response of many areas of parietal, temporal (including MT) and frontal cortex [32–36,49], implicated in control of action, is clearly modulated by attention, even in the absence of visual stimuli [50,51]. Areas that are clearly implicated in eye movement control, like frontal eye field (FEF) and lateral intraparietal area (LIP) [52,53], are also very clearly involved with attention. And attention is anchored to motor programmes [54]: it is possible that the interaction between motor and attentional signals could yield a selectivity in external space.

Our results suggest that head-centred coding is more common in dorsal areas, which are implicated in action. Our studies on perceptual mislocalization have suggested that the action system seems to update spatial maps much later than the perceptual system [55,56]. Perhaps the updating of craniotopic maps takes time, but leads to more robust coding of information, explaining the resistance of the action system to saccadic mislocalization [57–59]. The perceptual system, on the other hand, may not always operate with a complete map anchored in external coordinates, and in some cases may be more efficient to operate on retinotopic coordinates.

3. REMAPPING OF TRANSIENT STIMULI AT THE TIME OF SACCADIES

As the reviews in this issue make clear, visual stability is almost certainly a complex process involving many neural mechanisms. Spatiotopic maps are certainly involved, but it is unlikely that these are the sole mechanism implicated in keeping the world stable. These maps are unlikely to constitute detailed representations of the whole world, but probably only items of importance, those that attract attention: a form of saliency map. Furthermore, a map of this sort must take time to compute, probably too long for rapid online interaction with the world. Most researchers now agree that besides the existence of spatiotopicity, transient processes that occur around the time of saccades are important for stability. Psychophysical studies have described a myriad of robust and at times bizarre perceptual effects to brief peri-saccadic stimuli: some (but not all) stimuli are suppressed; stimuli are poorly localized, shifted in the direction of the saccade, and also compressed in space; and stimuli are delayed and compressed in time.

Some of these phenomena have obvious functional correlates. For example, the very specific suppression of low-frequency, luminance-modulated stimuli could serve to dampen motion perception [60–62]. The eyes move very quickly during saccades, up to 600° s⁻¹, which smears-out much detail in the high- and mid-spatial frequency range, but the visual system can still resolve very low spatial frequencies associated with this large-field fast motion [63]. Under normal circumstances motion of this sort is most disturbing; but during saccades it is suppressed, reflected not only in raised thresholds, but also in that the motion itself is less salient, less disturbing [60,64]. There is good evidence from imaging studies for suppression at various levels of the visual system, during both saccades and blinks [65–67]. Suppression is also observed in many areas of the monkey’s visual system, particularly the motion areas MT and MST [68,69]. This suppression may reflect suppression in the superior colliculus, which responds to these low frequencies, feeds to area MT and is suppressed during saccades [70].

Suppression clearly plays an important role in reducing the sense of motion that the moving eyes would otherwise elicit. But suppression alone is not sufficient to account for stability and continuity of perception across saccades; even with the sense of motion damped, perception still has to link together images that are displaced on the retina. Another phenomenon, probably related to stability, is that stimuli briefly presented at the time of saccades are mislocalized. Matin et al. [71], Bischof & Kramer [72], Honda [73,74] and Mateeff [75] showed that stimuli presented within 50 ms of saccadic onset are not seen veridically, but mislocalized in the direction of the saccade. The explanation for these data was that they reveal the action of a corollary discharge signal that compensates for the change in eye position, a signal that does not follow the exact dynamics of the eye motion, but starts to act before the eyes actually move, producing the shift in perceived position of stimuli flashed briefly peri-saccadically.

However, as is often the case in neuroscience, the story is not so simple. Most early studies measured mislocalization of stimuli flashed to the same spatial position. However, more recent studies [76,77] showed that the magnitude—and indeed the sign—of mislocalization depend strongly on spatial position. Mislocalization is not always in the direction of the saccade—as ‘compensation’ theories would expect but, under most experimental conditions, towards the saccadic target. The net result is a strong peri-saccadic compression of the visual field, as shown in figure 3a: stimuli flashed to the left of the saccadic target (for a rightward saccade) are seen displaced.
rightwards, while stimuli flashed beyond it are displaced leftwards, in all cases towards the saccadic target. Indeed all stimuli flashed between $-10^\circ$ and $+20^\circ$ of visual space are all seen near the saccadic target at $+7^\circ$. This displacement clearly cannot result from the simple addition of a single 'efference copy' vector to the retinal eccentricity signal.

The compression is very real, causing multiple objects straddling the saccadic target to collapse down to a single bar. We displayed on the screen a variable number of vertical bars, ranging from zero to four, and asked subjects to report how many they saw. The square symbols of figure 4a show the number of bars reported as a function of time relative to saccadic onset, for the condition where there were actually four bars displayed. The data show clear evidence of compression, following a tight timecourse around saccadic onset. The bars were coloured and sharply defined, stimuli that were not suppressed. Indeed, no errors occurred when one bar was displayed (triangle symbols), nor were there false alarms when no bars were displayed (filled circles).

When the same experiments (localization or reporting bar number) were performed under conditions of 'simulated saccades' (by moving a mirror rapidly to mimic a saccade), the results were quite different [76]: very few errors were made in detecting bar number, and no compression was observed in localization. This, and the fact that the timecourses of peri-saccadic compression are so tight, suggests that they are driven by an extra-retinal signal, or corollary discharge, but that this is not simply an addition of a vector.

The phenomenon of saccadic compression has now been replicated in many laboratories, under various conditions.
conditions, and shown to be robust [78–81]. Interestingly, compression does not work at the level of details, or ‘pixels’, but at the level of features. For example, squares are displaced towards the saccadic target, but they do not become thinner, changing to rectangles [82–84]. Also, information about features such as colour remains, even when different coloured bars are compressed to the same point in space [85]. This is reminiscent of the nature of \textit{trans}-saccadic integration, which clearly operates at the level of features rather than fine detail.

4. FUNCTION OF COMPRESSION?

So what may be the function of peri-saccadic compression, and peri-saccadic displacement in general?—or what visual processes may it reflect? Most studies of saccadic mislocalization report the systematic errors, the systematic \textit{biases} of localization. Recently, however, we [86] have shown that peri-saccadically vision is not only biased, but also very imprecise. In a two alternative forced choice procedure, subjects reported whether a bar flashed peri-saccadically appeared to the left or right of one flashed some time before the saccade. These data yield psychometric functions (best fit cumulative Gaussians) like those of figure 5, whose median (50% point) estimates the bias of the judgement, and width (s.d.) reflects the precision. During fixation, visual localization is both accurate and precise: but peri-saccadically, localization becomes \textit{inaccurate}—with a systematic bias towards the saccadic target—and \textit{imprecise}, reflected by the broad curve. The bias is consistent with the many other studies showing localization errors, with non-forced choice techniques such as naming. But this is the first study to show that the precision for localization is also affected, about 10 times worse during saccades than in fixation (figure 5a).

As a further test of visual precision, we investigated how visual and auditory stimuli combined during saccades. Under normal conditions, when sight and sound are in conflict, vision wins: the so-called \textit{ventri-loquist effect}. The dominance of vision has been well explained by the popular Bayes-based model of optimal, reliability-based integration: as vision is the more reliable (precise) sense, it is given a much higher weight than audition when the two are combined [87]. However, when the reliability of vision is reduced by blurring the stimulus, audition can dominate. We measured localization of audio-visual stimuli presented at the time of the saccade, taking advantage of the fact that saccades have no effect on auditory localization, either accuracy or precision. The audio-visual results are shown by the triangular symbols of figure 5b; these stimuli are mislocalized in the direction of the saccades, but far less than the purely visual stimuli. The curves are also steeper, reflecting improved precision for the combined audio-visual signals, compared with either vision during saccades, or auditory localization.

These data clearly show that vision relaxes its precision for spatial localization at the time of saccades, to the point where visual localization is as poor as auditory localization. This would seem to be strongly linked to compression. If all stimuli presented over a wide range of positions are compressed to a single position, it follows that precision would be lost. But why should this occur, and how is it related to stability? As discussed above, and in other reviews of this issue, \textit{trans}-saccadic integration of features seems to be important for stability. How does the system decide what is appropriate to integrate? One criterion for integration could be spatial coincidence (in external space) before and after the saccade. But if the corollary discharge signal is only a coarse signal, as many suspect, or the saccades themselves are slightly off target, the pre- and post-saccadic images will not line up exactly. So it makes sense to relax the reliability of the positional information of features (probably only salient features) at the time of saccades, allowing for integration of stimuli that may not coincide precisely after the saccade. There is evidence in this direction.

*Phil. Trans. R. Soc. B* (2011)
Peri-saccadic mislocalization varies from individual to individual, and correlates strongly with eye velocity: individuals with faster saccadic velocities show more compression than do those with slower saccadic velocities [88]. The reason for the correlation is far from certain, but it known that it is not related to the faster retinal motion. It seems reasonable to assume that higher saccadic velocities are associated with lower reliability in localization, around the time of saccades, and hence the need of a stronger compression.

We are currently performing experiments to verify our hypothesis, measuring fusion and separation of stimuli presented before and after saccades. Note that this idea is similar to that of Deubel and colleagues (e.g. [89]), except they assume that the alignment is only for the saccadic landing point, while we make no such assumption and believe it affects all salient features in the visual field.

5. THE EFFECT OF SACCADIES ON TIME

One of the more fascinating psychophysical discoveries of recent times is the demonstration that saccades not only cause a shift and compression of space, but also affect time perception in a very similar way. Like space, time becomes both displaced and compressed around the time of the saccades. As figure 4b shows, a pair of peri-saccadic stimuli, actually separated in time by 100 ms, appear to be separated by only 50 ms when presented near a saccade [90]. The compression follows a similar timecourse to that for spatial compression (the curve is broader only because the stimulus spans 100 ms). Not only are pairs of stimuli temporally compressed, single stimuli are severely mislocalized in time [91]. Figure 3b shows how apparent time (measured with auditory markers) becomes distorted around the time of saccades. Stimuli near saccadic onset were delayed by up to 50 ms, resulting in a gross distortion of the perceptual timeline. The curve flattens out considerably during the saccade, so all stimuli presented during that period are perceived at the same time, after the eye has landed in its new fixation. The histogram on the left shows the frequency of stimuli appearing at a particular time (assuming a uniform distribution in real time). The effect of the saccade is to accumulate stimuli towards the beginning of the new fixation.

The peri-saccadic change in timing is robust, and at certain critical times it can cause pairs of stimuli to appear to be inverted in time [90]. The inversion in time can be predicted quantitatively from the saccade-induced distortions [91]. At very particular times, about 70 ms before saccadic onset, the first of two stimuli is accelerated with respect to other times, causing it to ‘overtake’ the second and arrive in consciousness first.

Saccades affect both space and time, and do so in the same way. The timecourses of mislocalization (when corrected for stimulus duration) are very similar. Binda et al. [91] asked subjects to localize flashed stimuli in space, and also in time (compared with an auditory marker) in the same experiment. Figure 3c plots the temporal mislocalization against the spatial mislocalization. The two are very strongly correlated, explaining 92 per cent of the variance. It is clear that saccades affect not just time, but space–time.

6. NEURONS WITH ‘SHIFTING RECEPTIVE FIELDS’

Neurons in many visual areas show clear saccade-related changes to their RF properties. In some areas, such as areas V6 [3] and VIP [4] there exist neurons with spatiotopic RFs, tuned to external not retinal space, providing a plausible neural substrate for a stable spatiotopic map. But this behaviour seems to be the exception, rather than the rule. In a landmark paper [92], it was reported that RFs of many visual neurons in area LIP change at the time of saccades, shifting in the direction of the saccade (figure 6a,b). This result has proven robust, and has been replicated in many other visual areas including superior colliculus, FEFs, area V3 and even, to a less extent, in V1 (see review of Wurtz et al. [44]).

Other reviews (especially [44]) describe how the RFs of many neurons in many cortical areas shift before saccades are executed, so they respond to the region of space that will be brought into view by the saccade, but do this before the eyes actually move. This behaviour is particularly evident in areas LIP and FEF, but has been reported in many areas including superior colliculus, V3 and even V1 and V2 (e.g. [93–95]).

It is tempting to jump immediately to the conclusion that the shifting RFs reflect neural mechanisms that lead to spatiotopicity, shifting the retinal images around so they can slot into a spatiotopic map. There is only one problem with this reasoning: the shift is in the wrong direction to compensate for the eye movement. The shift in the RFs of LIP and other areas is in the same direction as the saccade, while compensation for saccade-induced image shifts needs to go in the opposite direction. If the head moves rightwards, to maintain fixation at a given point the eye must move leftwards. Similarly, if the eyes move 10° rightward, the compensation must be leftwards. Suppose a neuron in LIP has a ‘classical RF’ centred at location 0 (straight ahead). Spatiotopicity requires that this neuron maintains its selectivity to location 0 in external space after the movement has occurred. But the physiology suggests that just before the movement, the RF shifts to a location in the same direction of the saccade, +10° to the right. The eyes then move 10° rightwards, so the RF (in external space) is now at +20°, twice as distant from the spatiotopic location as would have occurred by the eye movement alone! The tuning of the cell presumably returns then to its ‘classical’ location (in retinal space), but it is far from clear why the anticipatory shift, that seems to exacerbate the problem, should occur at all.

Many modellers glide over the problem, either ignoring (or pretending to ignore) the fact that the sign is wrong for their model, or recognizing the conflict and concluding that these RFs do not create spatiotopicity, but serve other stability-related functions, such as comparing pre- and post-saccadic responses. For example, Mathôt & Theeuwes [45]
believe that the ‘anticipatory RF shift’ allows neurons to take a ‘sneak peak’ at the location which will be brought into the RF. Similarly, Wurtz et al. [44] suggest ‘comparing the activity in the RF after the saccade to the activity in the future RF before the saccade is a potential mechanism that might underlie the perception of visual stability’. Similar ideas have been expressed by Nakamura & Colby [93], Kusunoki & Goldberg [96] and Melcher & Colby [42]. Heiser & Colby [97] have suggested a further role that updating activity in LIP could be brought into the RF*. Similarly, Wurtz et al. [44] suggest ‘comparing the activity in the RF after the saccade to the activity in the future RF before the saccade is a potential mechanism that might underlie the perception of visual stability’. Similar ideas have been expressed by Nakamura & Colby [93], Kusunoki & Goldberg [96] and Melcher & Colby [42]. Heiser & Colby [97] have suggested a further role that updating activity in LIP could be brought into the RF*. Similarly, Wurtz et al. [44] suggest ‘comparing the activity in the RF after the saccade to the activity in the future RF before the saccade is a potential mechanism that might underlie the perception of visual stability’. Similar ideas have been expressed by Nakamura & Colby [93], Kusunoki & Goldberg [96] and Melcher & Colby [42]. Heiser & Colby [97] have suggested a further role that updating activity in LIP could be brought into the RF*.

We [91] believe that the key to the mystery is that saccades have profound psychophysical consequences not only for the perception of space but also for the perception of time, and the two cannot be considered independently. The temporal dynamics of remapping is also interesting, as shown in figure 6. Figure 6a,b schematize the concept of the ‘future RF’. Suppose that during normal fixation a neuron has a RF selective to a region directly below the fovea. In the moments just prior to the monkey making a saccade to the right, while the eyes are still fixating the red spot, the RF shifts rightward to what is termed the future RF. Typical responses of an example FEF cell to stimulation in these two regions are shown in figure 4c,d [98]: during normal fixation the cell responded only to stimuli in the ‘RF’ (blue), while peri-saccadically it responded best to the future RF (pink). But what is interesting is that the latencies of the responses are quite different, those of the future RF being much longer. Similar results can be seen in recordings in LIP [96] and other visual areas [93].

Figure 6e is a cartoon of a space–time plot of the response of an LIP neuron, drawn from data of Wang et al. [99]. In an experiment similar to that reported above, they recorded from an LIP neuron to stimuli delivered to the future RF at various times before, during and after the saccade. Note that after the saccade, the eye movement will have brought the retinal projection of the future RF back to the original classical RF, under the fovea. In this cartoon the responses are all aligned to the saccade, and ordered according to stimulus presentation relative to it. With this representation, the first spikes of the responses to all stimuli occur at the same time, implying that pre- and post-saccadic stimulation to this part of space causes a spike train that is effectively identical. A higher order cell (or a neurophysiologist) monitoring the response has no way of distinguishing whether a particular spike results from early pre-saccadic stimulation to the future RF or later post-saccadic stimulation of the classic RF. By definition, the region in space–time that
elicits identical responses defines the RF of the cell, in space and in time. The RF at the time of the saccade of this hypothetical LIP cell is illustrated schematically in figure 6f. In retinal coordinates, it is oriented in space–time along the trajectory of the retinal projection of the saccadic target during the eye movement. All spikes from this spatio-temporal RF arrive at the same time, and are, therefore, indistinguishable. Objects falling within the spatio-temporal RF, either in the ‘future field’ (pink circle) or classical RF (light-blue circle) will be integrated, and coded as being in the same external location. The fact that the RF is aligned to the saccadic target trajectory effectively stabilizes the image, at least transiently. A similar argument has previously been developed for the mechanisms involved in the perception of spatial form of moving objects [100,101]. That is not to say that these neurons are spatiotopic: their RFs are not anchored in external space. At most they exhibit a form of local transient spatiotopicity, in space and in time achieved by delaying the future RF response. To achieve absolute spatiotopicity, the eye position signal must be combined in some way with the transient retinal position signalled by this cell. But perhaps local spatiotopicity is sufficient to offset the retinal motion and displacement caused by the saccade, providing an immediate compensation for the displacement, thereby allowing perception to proceed seamlessly. The oriented spatio-temporal RF will generate a signal of continuity between pre- and post-saccadic view of the same object, and integrate signals of the same object trans-saccadically, possibly over a longer time span compared with fixation. A true spatiotopic representation, invariant to eye and body position, may be constructed in other areas, by integrating successive corollary discharge vectors. Interestingly, the spatio-temporal behaviour of the LIP neurons described by Wang et al. [99] closely resemble the psychophysical results for transient stimuli around the time of saccades (figure 3b, histogram at left). Stimuli presented just before and during the saccade are delayed relative to those presented later, similar to the neural discharges of LIP cells. The result is that stimuli presented over a wide range of times are localized in time to appear just after the saccade (histogram at left), just as stimuli presented at different times to the ‘future’ RF of LIP cells all cause spike trains that arrive at a similar time, after the saccade has been completed. Interestingly, the thick curves that fit the data are derived from a model based on units that transiently change their impulse-response function to become oriented in space–time, similar to the hypothetical spatio-temporal RF cartooned in figure 6.

7. CONCLUDING REMARKS

While the problem of visual stability is far from solved, tantalizing progress has been made over the past few years. It is clear that there do exist spatiotopic representations in the human brain, and that the construction of these representations requires visual attention. It is also clear that many transient processes occur when the eyes move, in both space and in time, leading to a spatio-temporal tuning that could provide a quick transient local spatiotopicity for immediate interaction with the world. Exactly how all these different mechanisms interact to provide stability will be one of the main challenges for future research.

Supported by EEC framework 6 (MEMORY) and 7 (ERC: STANIB) and Italian Ministry of Universities and Research.

REFERENCES


