Spatiotopic temporal integration of visual motion across saccadic eye movements

David Melcher[1,3] & M Concetta Morrone[1,2]

Saccadic eye movements pose many challenges for stable and continuous vision, such as how information from successive fixations is amalgamated into a single precept. Here we show in humans that motion signals are temporally integrated across separate fixations, but only when the motion stimulus falls either on the same retinal region (retinotopic integration) or on different retinal positions that correspond to the same external spatial coordinates (spatiotopic integration). We used individual motion signals that were below detection threshold, implicating spatiotopic trans-saccadic integration in relatively early stages of visual processing such as the middle temporal area (MT) or V5 of visual cortex. The trans-saccadic buildup of important congruent visual information while irrelevant non-congruent information fades could provide a simple and robust strategy to stabilize perception during eye movements.

Although seldom aware of this, humans make an average of 3–5 saccadic eye movements per second. Although this is an efficient strategy to make maximum use of the high resolution of central vision[1], it raises the question of how information from separate glances is combined to give a stable perception of the world. One potential solution would be to combine visual information across saccades. However, this mechanism could be useful only if information from the same spatiotopic position is integrated. Previous studies of trans-saccadic integration show that visual information is not integrated across saccades, at least in the case of visual patterns[2–4] (but see ref. 5). Many simple visual tasks, such as contrast sensitivity and shape discrimination, have a limited integration time falling within the duration of a typical fixation[6]. Thus, the failure to find visual integration in these tasks does not necessarily indicate that spatiotopic integration does not occur; simple patterns might be detected in a single fixation.

Motion coherence thresholds, on the other hand, are mediated by mechanisms that summate information linearly for several seconds[7], a timeframe that will typically include several saccades. Temporal integration of motion allows a weak signal that would otherwise go undetected to be accurately perceived given a long enough viewing time. To improve motion discrimination from one fixation to another, a weak, sub-threshold motion signal viewed in one glance must be summated with a subsequent motion signal for the same object after the saccade, requiring spatiotopic integration. On the other hand, compulsory motion integration irrespective of spatial position could harm motion discrimination, as information from different objects or possibly from different motions would often be erroneously combined. Another possibility, suggested by previous studies of trans-saccadic integration, is that motion integration would be interrupted on each new fixation, implying an active signal that disrupts integration. Here we test all three hypotheses and show that

motion does integrate across saccades, but only if the two motion signals are presented either in the same retinal location or in the same spatial position. Whereas retinotopic integration of sub-threshold motion signals can be attained by the same mechanisms that are active during fixation, spatiotopic integration implies the existence of receptive fields built in the external rather than in the retinal coordinate system, like those seen in associative cortex of the monkey[8–11].

RESULTS

We devised an integration technique for motion that uses two brief (150-ms) horizontal motion signals embedded in 10 s of random noise (see **Fig. 1** for experimental details). The motion signals were diluted with random noise to determine coherence thresholds for direction discrimination[7], allowing us to probe directly trans-saccadic integration, its linearity and its spatial property.

Spatiotopic integration of two motion signals

We first investigated temporal integration with steady eccentric fixation (6° above or below the motion patch) by measuring coherence sensitivity (inverse of thresholds) as a function of the temporal separation of the two motion signals. For separations of less than about 1 s, sensitivity was twice that for a single patch of motion, implying perfect summation of the two signals (**Fig. 2a**, filled squares). At longer delays, performance dropped to the level of the single signal condition, implying that temporal integration was not possible over that duration (consistent with previous studies[7]).

We then measured motion discrimination sensitivities when observers made a 12° saccade from one fixation point to the other (**Fig. 2a**, open triangles). The saccade had very little effect on the results: summation remained perfect for up to about 1 s and decreased rapidly thereafter. This shows that integration of motion
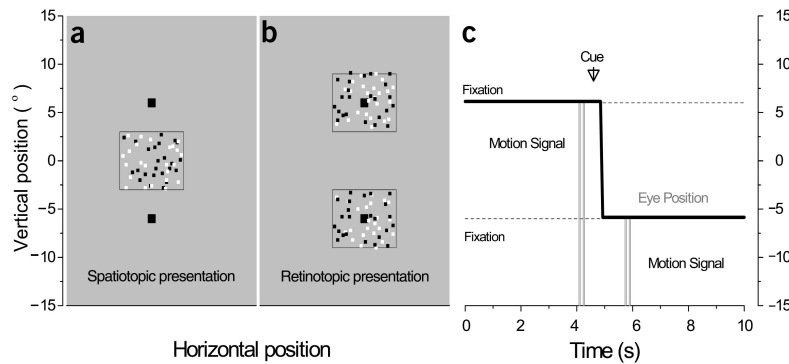
Figure 1 Schematic representation of the visual stimuli for experiment 1 (**a**) and experiment 2 (**b**), and relative time course of the events (**c**). In each trial, the first motion signal was presented, then a cue to saccade, then the second motion signal. The timing of the cueing varied randomly across trials, and the trials were later sorted according to saccadic onset time with respect to the onset of the two coherent motion signals.

signals is not interrupted by saccades, but continues without any attenuation. It also points to spatiotopic integration, as the saccade brought the stimulus to a new location on the retina.

Retinotopic integration of two motion signals

To test whether summation across saccades can also occur for stimuli that are displayed at the same retinal location (and hence different spatial location), we adapted the procedure so as to always stimulate the central visual field (**Fig. 1b**). For the non-saccade controls, observers fixated one or the other of the random dot patches and both motion stimuli were presented to that patch (as before, except that viewing was now central). This yields very similar results to before (**Fig. 2b**, squares), except the summation period is somewhat longer, presumably reflecting a difference in temporal integration in central and peripheral vision. In the saccade condition, the observer was cued to saccade from one dot pattern to the other, and the motion signals were presented to the region where he or she was fixating. Again, perfect summation occurred in these circumstances (**Fig. 2b**, triangles), showing motion integration across saccades—retinotopic summation in this case. Furthermore, integration across saccades was also found for retinotopically matched peripheral motion stimuli (arranged in an annulus of 4° and 6° diameters), inducing an improvement in sensitivity of 1.7 for two subjects: AB and DM.

**Figure 2b** shows data only for the condition where the saccade

occurred between stimulus presentations so as to position the coherent motion stimulus on the same retinal region. It has been reported that saccades are preceded by a shift of attention at the saccadic target[12]. We found, however, that summation occurred only when the second motion stimulus fell in the central visual field and not when it fell in the supposed peripheral attentional focus (**Fig. 3**). When the saccade occurred before the first motion signal or after the second, there was no advantage compared with the single motion signal. However, when the saccade occurred between the two motion signals, performance was consistently superior to that with only one signal, producing a halving of threshold (doubling of sensitivity). Thus, motion integration was not simply due to the intention to make a saccade or to very large receptive fields that encompass the whole area. Nor can the lack of summation for inappropriate eye movements be ascribed to a difference in peripheral and central sensitivity, as sensitivities were virtually identical for these different eccentricities for both subjects (6.0 vs. 5.5 for DM and 5.3 vs. 4.7 for MCM).

Control experiments

To exclude the possibility that the effects could have resulted from summation of a very large spatial pool, we performed two additional controls. The first condition was designed to mimic the first experiment,

Figure 2 Motion coherence sensitivity as a function of temporal delay between the two motion signals. (**a**) Performance for peripherally viewed targets (see **Fig. 1a**). Coherence sensitivity is defined as the ratio of the total dot number on each frame to the number of signal dots. Filled squares show performance when observers fixated above or below the motion patch. The open triangles show results when observers made a 12° saccade between motion signals. (**b**) Motion coherence sensitivity for centrally viewed stimuli as a function of the temporal delay between the two motion pulses (see **Fig. 1b**). The curves are best fit of the data during fixation with sigmoidal functions. Performance on fixation trials (filled squares in **a** and **b**) decreased as the temporal delay between the two motion signals was increased. There was total summation of sub-threshold signals for over 1 s: the maximum performance was about twice that for a single motion stimulus. In other words, at threshold the total number of coherently moving dots in both stimuli were the same as those in the single-interval stimulus. Similar performance was found on saccade trials (open triangles in **a** and **b**) for both conditions, both when the stimuli were retinotopically invariant (**b**) and more surprisingly when they were spatiotopically invariant (**a**). Data is reported only for trials in which the saccade was accurate and occurred during the interval between the two motion signals. The dotted line shows performance when only one motion signal was present during fixation trials. Data is shown for the two authors (DM and MCM) and a naïve subject (AB).
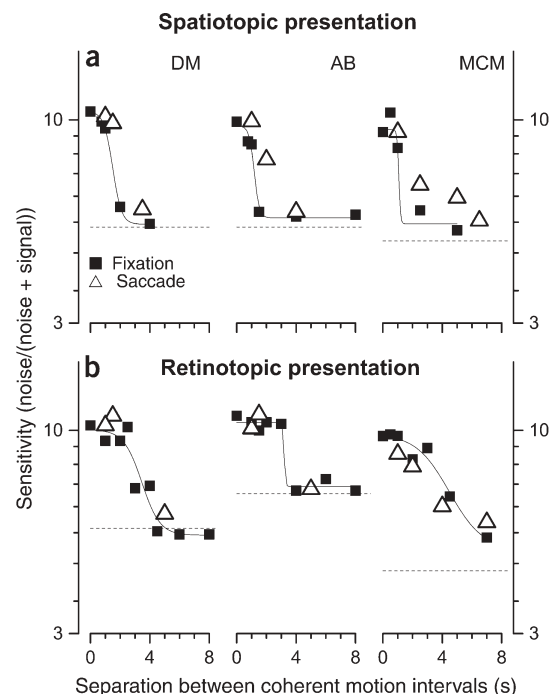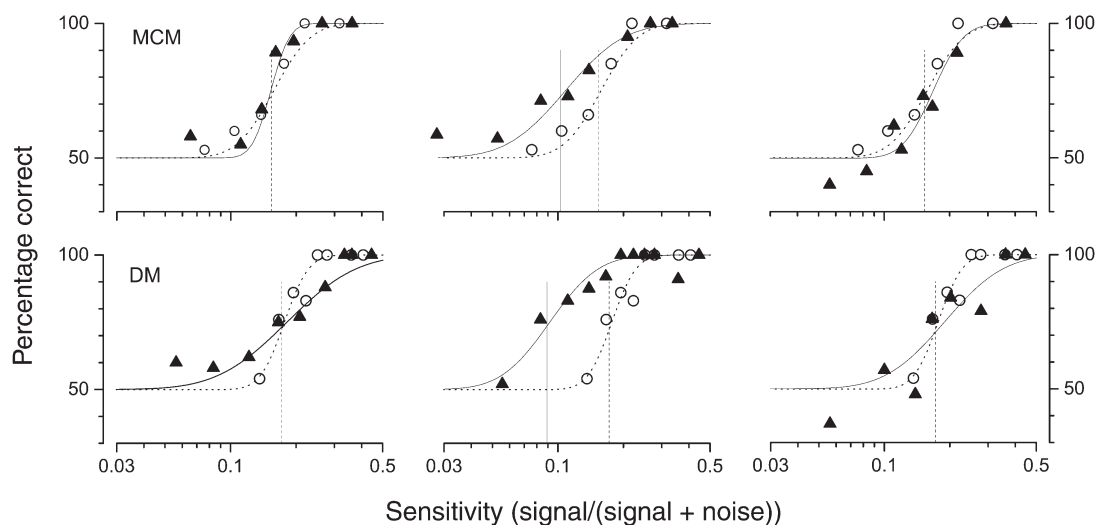
Figure 3 Percentage correct motion discrimination as a function of motion strength in two subjects (MCM, top row, and DM, bottom row) for retinotopic integration across saccades (see **Fig. 1b,c**). Performance was similar for subject AB. Open circles (and dotted line fit) show performance on control trials when only one signal was presented in the saccadic condition. Filled triangles indicate percentage correct for trials in which the first motion signal was presented in the top patch and the second motion signal in the bottom patch with a delay of 1.5 s. Left column, performance when a saccade from the top to bottom patches was made 100–4,000 ms before the first motion signal was presented. Middle column, performance when the saccade was made during the temporal interval between the two motion signals. Right column, motion discrimination when the saccade occurred 100–4,000 ms after the second motion signal was displayed. Only in the condition shown in the middle column were both signals presented to central vision; otherwise, one was central and the other 12° peripheral. The curves show the integral of a Gaussian function that best fitted the data; vertical lines show 75% correct threshold values. Note that summation in this condition occurred only when the saccade was performed between the two coherent motion signals (with minimum latency from coherent motion of 100 ms).

which showed spatiotopic integration (**Fig. 2a**). Motion signals were displayed consecutively to two motion patches 6° above and below fixation, separated by a 1-s interval. As with the first experiment, the second presentation was anticipated by an auditory cue, allowing observers to switch attention but not gaze. No summation occurred between the two motion signals, with sensitivity to the double presentation following exactly that to the single presentation (**Fig. 4a**). To ensure that the presence of a second patch or its spurious motion induced by the eye movement did not hinder performance, the same control was repeated but with only one patch visible at a time: first above fixation, and then (150 ms after the beep cue) below fixation (**Fig. 4b**). Again, the two motion signals were not temporally integrated.

As a final control, we investigated temporal integration of motion across saccades by testing motion signals that were neither spatiotopically nor retinotopically matched. Motion was presented in the patch 6º above fixation followed 1 s later by one 6° below fixation, with a 4° horizontal saccade between the two. No temporal integration of the two motion signals occurred in this condition (**Fig. 4c**). Taken together, the control conditions show that temporal integration of motion across saccadic eye movements cannot be explained simply by very large receptive fields during saccades, by saccadic suppression of noise or motion signal[13,14], by attentive selection of motion located in spatially different regions[15], nor by spurious motion introduced by the saccade. Summation of two temporally distinct motion signals across separate glances is specific to retinotopic or spatiotopic transsaccadic integration.

DISCUSSION

The results reported here provide evidence for trans-saccadic integration of an important visual attribute—global motion—and show that this integration can be both retinotopic and spatiotopic. Retinotopic

integration might be useful during smooth pursuit eye movements (possibly with saccadic catch-up), in which the spatial location of an object changes but its retinal location stays relatively constant; it would be less useful during normal saccades. On the other hand, spatiotopic integration of information at an early stage of visual processing could be instrumental in the perception of visual stability across eye movements, allowing incongruent retinotopic information to fade while carrying over only congruent information at an invariant spatial position.

Previous studies of trans-saccadic integration have produced inconsistent results. Some researchers report no trans-saccadic integration of visual patterns[2–4], although accurate spatial memory across saccades has been reported for simple geometric stimuli[5], and extraretinal eye movement signals have been shown to strongly influence the three-dimensional structure of moving images[16]. The transsaccadic integration found here may reflect a more general mechanism that combines information across glances. If so, then the brief integration times of the visual stimuli used previously might be the limiting factor for the failure to demonstrate trans-saccadic visual integration, while highlighting higher cognitive factors, like visual memory[17,18]. Another possibility is that there is a genuine difference in the effect of saccades on mechanisms mediating form perception versus motion perception. This is consistent with evidence showing that the magnocellular pathway is selectively suppressed during saccade, sparing the parvocellular pathway (for review, see ref. 19).

The motion integration observed here could not be explained by indiscriminate summation of all motion signals, irrespective of spatial location. Summation over large areas might have been expected given that the size of receptive field for translation are large during fixation[20], and there may be a further broadening induced by the dynamic spatial remapping observed in several primary and associa-
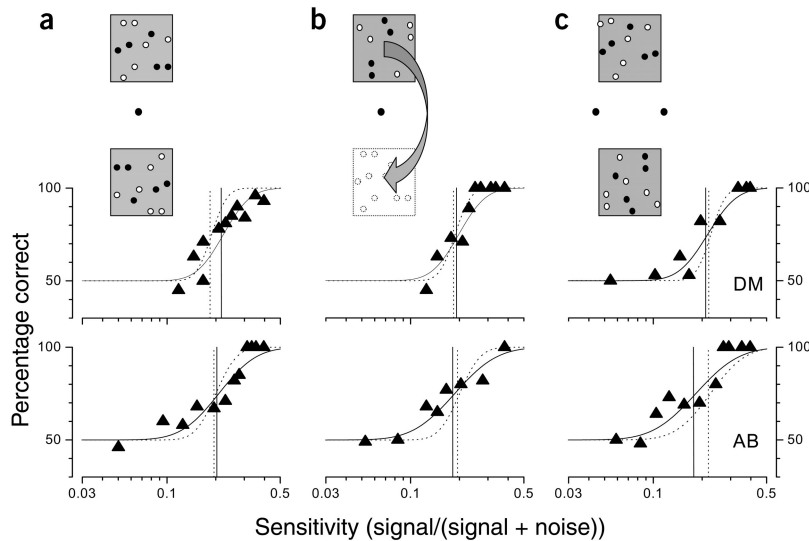
Figure 4 Accuracy of motion discrimination as a function of motion strength for control conditions for DM (top) and AB (bottom). Triangles show performance in the control task, with solid curves showing the integral of a Gaussian function that best fitted the data. Dotted curves show the best fits to the data (not shown) when only one motion signal was presented. Vertical lines show 75% correct threshold values. Note that performance in the control conditions with two motion signals was equivalent to that of a single motion signal viewed under the same conditions. (**a**) Maintained fixation in the center, with the first motion in the top patch and the second motion in the bottom patch. (**b**) Identical to **a**, except only one patch was visible at a time, with the patch changing to the bottom location at cue onset. (**c**) Identical stimulus as in **a**, with the addition of the instruction to make a 4° horizontal saccade after the cue.

tive areas around the time of saccades[21–23]. Some neurons shift their receptive fields before an eye movement, anticipating the change in receptive field position brought about by the saccade. Other neurons do not shift their receptive field, so at the moment of the saccade the average receptive field of the population includes both the shifted and the non-shifted regions, an effective elongation of the collective population receptive field. In humans, stimuli flashed briefly before or during a saccade tend to be misperceived as nearer to the saccadic target (saccadic compression[24]), consistent with an enlargement of the average or population receptive field. However, we tested directly, and refuted, the possibility that a large summation area would induce the spatiotopic integration.

The finding that integration is both retinotopic and spatiotopic, but always confined to a limited region of space, points to the existence of an active and precise selection of the spatial region in which trans-saccadic integration occurs. Dynamic remapping occurs only for stimuli very close to the saccade, preceding it by about 100 ms (refs. 21–23). We did not explicitly test motion perception around the time of the saccade in the current study. However, motion sensitivities of the two signals were unaffected by the saccade when measured for the same retinotopic condition, implying a smooth transition of motion processing from one fixation to another and suggesting a form of remapping.

Motion coherence sensitivity thresholds are correlated with the response of single MT/MST neurons in monkeys[25] and BOLD activity (detected by functional magnetic resonance imaging) in area MT+ in humans[26], implicating these areas in the current results. The increased sensitivity found here suggests that the same detector was able to summate the motion signal over different retinotopic regions. Although neurons with spatiotopic or craniotopic receptive fields have been reported in other visual areas[8–11], no studies to date have shown spatiotopic encoding in MT+. However, MT neurons have been shown to be strongly modulated by eye position signal[27], and this would be sufficient to generate a spatiotopic receptive field[28].

The existence of spatiotopic receptive fields in monkey has been demonstrated in association areas of the parietal cortex[9,10], implying their role in visuomotor coordination. It is notable that here we found spatiotopic summation of sub-threshold visual information with a basic visual task—direction discrimination of translational motion—

raising the possibility that trans-saccadic spatiotopic summation may be found also in other basic visual judgments. If so, spatiotopic receptive fields may have a more general function, providing an elegant means of perceiving objects as stable in the world, even when the eye moves.

METHODS

Subjects. Three subjects (the two authors and a naïve observer) with normal vision participated in the experiments. Informed written consent was obtained for all observers.

Stimuli. Stimuli were generated using VSG Framestore (Cambridge Research Systems) and displayed on a Barco Calibrator monitor, subtending $38° \times 28.5°$ when viewed from 60 cm (mean luminance of 28 cd/m²). Motion was confined within regions of $6° \times 6°$, either a single central patch (**Fig. 1a**) or two peripheral patches above and below the fixation point (**Fig. 1b**). The regions were filled with 58 dark and light dots, re-plotted in random positions at 63 Hz to give the impression of random motion. During periods of motion presentation, a subset of these dots moved coherently either to the right or left at 10° per second for 150 ms, and then resumed incoherent random motion. To balance the local space-time energy between coherent and random movement, each dot had a limited lifetime of two frames.

Procedure. In the first experiment, a single $6° \times 6°$ motion target was viewed in the periphery. On fixation trials, the subject was instructed (in separate blocks) to maintain fixation on one of the two fixations points. On saccade trials, the observer was cued by an auditory beep ('cue' in **Fig. 1c**) to make an eye movement from the upper to the lower fixation point. Note that in this condition, the second motion signal occurred after the saccade and hence occupied a different retinal position (from eccentricity +6° to –6° along the vertical meridian).

In the second experiment, two motion patches were presented 12° apart with a fixation point in the center of each patch. In the fixation trials, observers fixated one of these (varied between sessions), and both motion signals were presented to it. In the saccade condition, the observers were required to saccade between one and the other fixation point on auditory cue. In this condition, as well as in the three control conditions (**Fig. 4**, top), the first motion was presented to the top patch, and the second presented in the bottom patch. In the first control condition (**Fig. 4**, left), two patches were shown and fixation was maintained. In the second control (**Fig. 4**, middle), one patch was shown at a time, with the stimulus patch moving from top to bottom position when the cue occurred. In the final control condition (**Fig. 4**, right), a 4° saccade was cued from the left fixation point to the right fixation point in the center of the two patches.

The task of the observer was to indicate whether the perceived motion was rightward or leftward by pressing the appropriate button. The proportion of coherent dots varied from trial to trial to home in on the observer threshold, using the adaptive QUEST procedure[29]. A minimum of four QUEST sessions of 40 trials each were run for each data point, and the final threshold estimate obtained by fitting the data with a cumulative Gaussian function weighted by the number of trials for each $S/(S + N)$ value, using a simplex algorithm, and evaluating threshold at 75% correct response. For all conditions, we also measured threshold when only one coherent motion signal was embedded within 10 s of noise. In all sessions, eye-position was monitored at 240 Hz by an ASL Model 514 Remote Eye Tracker (Applied Science Laboratories) and stored for later analysis.

1. Land, M.F. & Nilsson, D.E. *Animal Eyes* (Oxford Univ. Press, Oxford, 2001).
2. Jonides, J., Irwin, D. & Yantis, S. Failure to integration visual information from successive fixations. *Science* **222**, 188 (1983).
3. Bridgeman, B. & Mayer, M. Failure to integrate visual information from successive fixations. *Bull. Psychonomic Soc.* **21**, 285–286 (1983).
4. Irwin, D.E. Information integration across saccadic eye movements. *Cognit. Psychol.* **23**, 420–456 (1991).
5. Hayhoe, M., Lachter, J. & Feldman, J. Integration of form across saccadic eye movements. *Perception* **20**, 393–402 (1991).
6. Barlow, H.B. Temporal and spatial summation in human vision at different background intensities. *J. Physiol. (Lond.)* **141**, 337–350 (1958).
7. Burr, D.C. & Santoro, L. Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Res.* **41**, 1891–1899 (2001).
8. Graziano, M.S., Yap, G.S. & Gross, C.G. Coding of visual space by premotor neurons. *Science* **266**, 1054–1057 (1994).
9. Snyder, L.H., Grieve, K.L., Brotchie, P. & Andersen, R.A. Separate body- and world-referenced representations of visual space in parietal cortex. *Nature* **394**, 887–891 (1998).
10. Duhamel, J., Bremmer, F., BenHamed, S. & Graf, W. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* **389**, 845–848 (1997).
11. Galletti, C., Battaglini, P.P. & Fattori, P. Parietal neurons encoding spatial locations in craniotopic coordinates. *Exp. Brain Res.* **96**, 221–229 (1993).
12. Deubel, H. & Schneider, W.X. Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision Res.* **36**, 1827–1837 (1996).
13. Burr, D.C., Holt, J., Johnstone, J.R. & Ross, J. Selective depression of motion sensitivity during saccades. *J. Physiol. (Lond.)* **333**, 1–15 (1982).
14. Shiori, S. & Cavanagh, P. Saccadic suppression of low-level motion. *Vision Res.* **29**, 915–928 (1989).
15. Saenz, M., Buracas, G.T. & Boynton, G.M. Global effects of feature-based attention in human visual cortex. *Nat. Neurosci.* **5**, 631–632 (2002).
16. Wexler, M., Panerai, F., Lamouret, I. & Droulez, J. Self-motion and the perception of stationary objects. *Nature* **409**, 85–88 (2001).
17. Melcher, D. Persistence of visual memory for scenes. *Nature* **412**, 401 (2001).
18. Brockmole, J.R., Wang, R.F. & Irwin, D.E. Temporal integration between visual images and visual percepts. *J. Exp. Psychol. Hum. Percept. Perform.* **28**, 315–334 (2002).
19. Ross, J., Morrone, M.C., Goldberg, M.E. & Burr, D.C. Changes in visual perception at the time of saccades. *Trends Neurosci.* **24**, 113–121 (2001).
20. Burr, D.C., Morrone, M.C. & Vaina, L.M. Large receptive fields for optic flow in humans. *Vision Res.* **38**, 1731–1743 (1998).
21. Duhamel, J.R., Colby, C.L. & Goldberg, M.E. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**, 90–92 (1992).
22. Colby, C.L., Duhamel, J.R. & Goldberg, M.E. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J. Neurophysiol.* **76**, 2841–2852 (1996).
23. Nakamura, K. & Colby, C.L. Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc. Natl. Acad. Sci. USA* **99**, 4026–4031 (2002).
24. Ross, J., Morrone, M.C. & Burr, D.C. Compression of visual space before saccades. *Nature* **386**, 598–601 (1997).
25. Salzman, C.D., Britten, K.H. & Newsome, W.T. Cortical microstimulation influences perceptual judgments of motion direction. *Nature* **346**, 106 (1990).
26. Rees, G., Friston, K. & Koch, C. A direct quantitative relationship between the functional properties of human and macaque V5. *Nat. Neurosci.* **3**, 716–723 (2000).
27. Bremmer, F., Ilg, U.J., Thiele, A., Distler, C. & Hoffman, K.P. Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. *J. Neurophysiol.* **77**, 944–961 (1997).
28. Zipser, D. & Andersen, R.A. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* **331**, 679–684 (1988).
29. Watson, A.B. & Pelli, D.G. QUEST: a Bayesian adaptive psychometric method. *Percept. Psychophys.* **33**, 113–120 (1983).