letters to nature

8. Rasnow, B. The effects of simple objects on the electric field of *Apteronotus. J. Comp. Physiol. A.* **178**, 397–411 (1996).
9. Howard, I. P. & Rogers, B. J. *Binocular Vision and Stereopsis* (Oxford Univ. Press, New York, 1995).
10. Collett, T. Stereopsis in toads. *Nature* **267**, 349–351 (1977).
11. Kral, K. & Poteser, M. Motion parallax as a source of distance information in locusts and mantids. *J. Insect Behav.* **10**, 145–163 (1997).
12. Lehrer, M., Wehner, R. & Srinivasan, M. V. Motion cues provide the bee's visual world with a third dimension. *Nature* **332**, 356–357 (1988).
13. Schuif, A. & Hawkins, A. D. Acoustic distance discrimination by the cod. *Nature* **302**, 143–144 (1983).
14. Bleckmann, H. *Reception of Hydrodynamic Stimuli in Aquatic and Semiaquatic Animals* (Gustav Fischer, Stuttgart, 1994).
15. Atema, J. Eddy chemotaxis and odor landscapes: exploration of nature with animal sensors. *Biol. Bull.* **191**, 129–138 (1996).
16. Moller, P., Serrier, J., Belbenoit, P. & Push, S. Notes on the ethology and ecology of the Swashi river mormyrids (Lake Kainji, Nigeria). *Behav. Ecol. Sociobiol.* **4**, 357–368 (1979).
17. von der Emde, G. The sensing of electric capacitances by weakly electric mormyrid fish: effects of water conductivity. *J. Exp. Biol.* **181**, 157–173 (1993).
18. Tippler, P. A. *Physik* (Spektrum Akademischer, Heidelberg, 1994).
19. von der Emde, G. & Bleckmann, H. Finding food: senses involved in foraging for insect larvae in the electric fish, *Gnathonemus petersii. J. Exp. Biol.* **201**, 969–980 (1998).
20. Wagner, H. Flow-field variables trigger landing in flies. *Nature* **297**, 147–148 (1982).
21. Hassan, E. S. in *The Mechanosensory Lateral Line. Neurobiology and Evolution* (eds Coombs, S., Görner, P. & Münz, H.) 217–228 (Springer, Berlin, 1989).
22. Bleckmann, H., Tittel, G. & Blübaum-Gronau, E. in *The Mechanosensory Lateral Line. Neurobiology and Evolution* (eds Coombs, S., Görner, P. & Münz, H.) 501–526 (Springer, Berlin, 1989).
23. Harkness, L. Chameleons use accommodation cues to judge distance. *Nature* **267**, 346–349 (1977).
24. Hopkins, C. D. Temporal structures of non-propagated electric communication signals. *Brain Behav. Evol.* **28**, 43–59 (1986).
25. Suga, N., Butman, J. A., Teng, H., Yan, J. & Olsen, J. F. in *Active Hearing* (eds Flock, A., Ottoson, D. & Ulfendahl, M.) 13–30 (Pergamon, New York, 1995).
26. Schnitzler, H.-U., Menne, D. & Hackbarth, H. in *Time Resolution in Auditory Systems* (ed. Michelsen, A.) 180–204 (Springer, Berlin, 1985).
27. Dear, S. P., Simmons, J. A. & Fritz, J. A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. *Nature* **364**, 620–623 (1993).
28. Cialo, S., Gordon, J. & Moller, P. Spectral sensitivity of the weakly discharging electric fish *Gnathonemus petersii* using its electric organ discharges as the response measure. *J. Fish Biol.* **50**, 1074–1087 (1997).
29. Zar, J. H. *Biostatistical Analysis* (Prentice Hall, Englewood Cliffs, 1984).
30. Hoerl, A. E. Jr in *Chemical Business Handbook* (ed. Perry, J. H.) 20–50 (McGraw-Hill, London, 1954).

Seeing biological motion

Peter Neri*, **M. Concetta Morrone**† & **David C. Burr**†‡

* *University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK*
† *Istituto di Neurofisiologia del CNR, Via S. Zeno 51, Pisa 56127, Italy*
‡ *Department of Psychology, University of Florence, Florence 50125, Italy*

One of the more stunning examples of the resourcefulness of human vision is the ability to see 'biological motion', which was first shown[1] with an adaptation of earlier cinematic work[2]: illumination of only the joints of a walking person is enough to convey a vivid, compelling impression of human animation, although the percept collapses to a jumble of meaningless lights when the walker stands still. The information is sufficient to discriminate the sex and other details of the walker[3,4], and can be interpreted by young infants[5]. Here we measure the ability of the visual system to integrate this type of motion information over space and time, and compare this capacity with that for viewing simple translational motion. Sensitivity to biological motion increases rapidly with the number of illuminated joints, far more rapidly than for simple motion. Furthermore, this information is summed over extended temporal intervals of up to 3 seconds (eight times longer than for simple motion). The steepness of the summation curves indicates that the mechanisms that analyse biological motion do not integrate linearly over space and time with constant efficiency, as may occur for other forms of complex motion[6], but instead adapt to the nature of the stimulus.

Biological motion was produced on a video display by an adaptation of a standard cyclic algorithm[7] that emulates the motion of the 11 main joints of a person walking on a treadmill. Subjects were required either to detect the presence of the walker or to discriminate the direction of walking, in the presence of variable amounts of dynamic random noise. Similar tasks were performed for simple translational motion, for which a random pattern of dots was continuously displaced horizontally within a window of dimensions similar to that of the walker. We adapted the standard biological-motion technique so that each point had a 'limited lifetime' of two frames, after which it disappeared and was redrawn in another randomly chosen position (see Fig. 1 and Methods). This allowed for dynamic undersampling of the entire dot pattern, ensuring that, even at the lowest sampling rates, all 11 joints were likely to be represented during the 1,200-ms exposure time.

Figure 2 shows how sensitivity (expressed as maximum tolerable noise) for detection and direction discrimination varied with the number of displayed points for the two types of motion. For translational motion, sensitivity for both detection and discrimination increased with dot number at a very similar rate, almost linearly for both subjects tested (log–log slope is roughly at unity). For biological motion, the summation curves for detection were very similar to those for detection of translation, in both slope and absolute sensitivity. But the direction-discrimination curves were far steeper, with log–log slopes of about 3 or more (indicating a cubic or higher relationship).

Figure 3 shows how sensitivity to the walker varied with exposure time (temporal summation). As before, we used a limited-lifetime paradigm, with six dots displayed. Sensitivity to both simple and biological motion increased with time, first rapidly then more gradually. In comparable studies, it was thought that the initial rapid increase reflects physiological summation, whereas the gradual phase reflects 'probability summation' (a statistical improvement based on recruitment of independent detectors rather than summation within a single unit[8]). The intersection of these two curves gives an estimate of the time constant of the
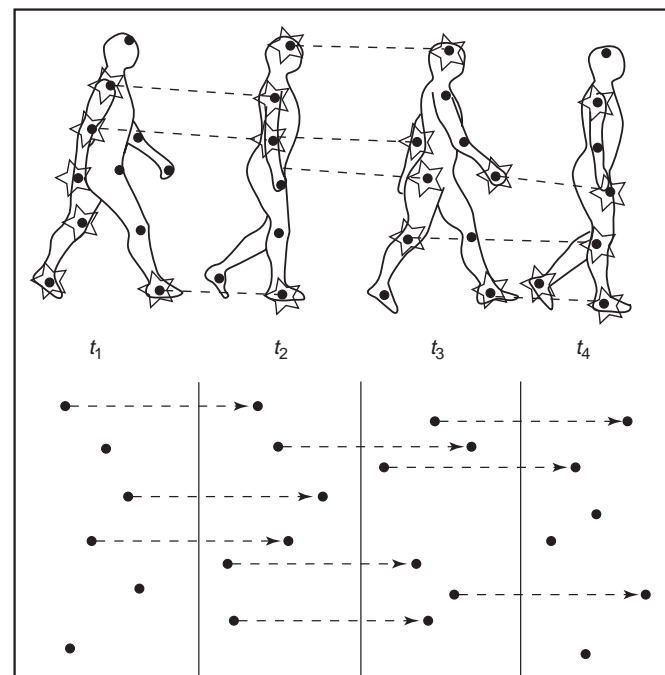

Figure 1 The limited-lifetime technique for studying the ability to see biological motion (upper figures) and translational motion (lower figures), for six-dot sampling (over a two-frame running average). The starred points of the walker indicate those that are actually in motion; the others, the possible positions to be sampled. Half of the dots move from frames 1 to 2, and the other half from frames 2 to 3. The starred points sometimes undergo occlusion (when required by the algorithm[7]), and are not visible at these times.

Figure 3 Sensitivity for direction discrimination, for both translation and biological motion. The following range of image speeds was used: 1.3 deg s⁻¹ (filled squares) and 7.3 deg s⁻¹ (open squares) for translation; and 1.3 deg s⁻¹ (filled circles) and 3.8 deg s⁻¹ (open circles) for biological motion, using the limited-lifetime procedure (with six joints) (see Methods). The dashed vertical lines indicate the transition from physiological to probability summation, calculated by best-fit of a two-segment template of variable slope. The summation constant for translation was relatively short (600 ms) and invariant with speed, whereas summation for biological motion was longer and varied inversely with image speed. The initial steep slope of ~2 for all curves is predicted by a linear integrator given the stimulus was embedded in a 7-s noise sequence. We obtained similar results with another subject (S.M.).
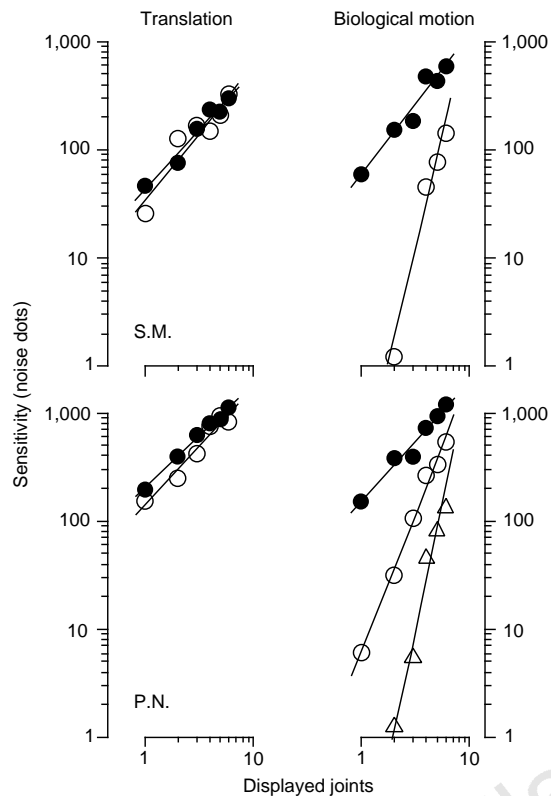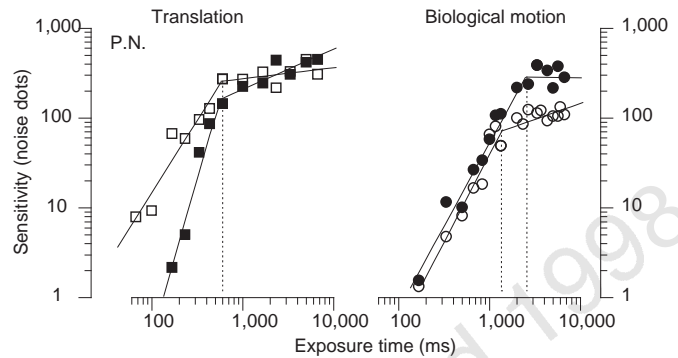

Figure 2 Sensitivity, expressed as the number of noise dots at threshold, for detection (filled circles) and direction discrimination (open circles), for both simple translation and biological motion. The open triangles indicate results of a study in which the observer was required to discriminate 'coherent' from 'incoherent' motion; in this case, the upper and lower body moved either in the same or in opposite directions[20]. The steepness of the log–log curves (calculated by linear regression) for subjects S.M. and P.N. (respectively) were: 1.10 and 0.96 for translation detection; 1.24 and 1.08 for translation discrimination; 1.29 and 1.12 for detection of biological motion; 4.23 and 2.55 for discrimination of biological motion; and 4.48 for discriminating coherent from incoherent motion (P.N.). S.M. was naïve of the goals of the study.

physiological summation along local trajectories[9]. For simple translation, the time constant of summation was ~600 ms, and was nearly invariant with velocity. For biological motion, however, the time constants were much larger, up to 2,800 ms. Furthermore, the estimate of summation depended on image velocity, in roughly inverse proportion, indicating that the limiting factor was the number of cycles presented rather than time passed *per se*.

The results indicate several new facts about biological motion. They demonstrate objectively the robustness of this curious stimulus; the visual system is very sensitive to fragments of natural motion scenes: observers reliably detected the direction of motion of point-lit figures in greatly degraded conditions, when half of the points were of opposite polarity, local-motion trajectories were given by two-frame sequences, and up to 1,000 similar dots were sprinkled over the limited area. Indeed, with adequate spatial sampling, sensitivity for discriminating the direction of ambulation was as high as that for its detection, or for detection of discrimination of direction of simple translation. Temporal summation for biological motion occurs over much longer periods than for simple translation, and seems to be limited by the number of cycles of ambulation, rather than by total duration of the stimulus.

Theoretically, sensitivity might be expected to increase linearly with the spatial sampling density[10], as observed in discrimination of Glass patterns[11,12], simple translational motion[13,14] and complex optic-flow patterns[6,15]. The same relationship occurred here for detection and discrimination of translation, and for detection of

biological motion. This latter result may reflect independent integration of signals along the trajectory of each joint by local-motion mechanisms, given that the sampling rate of each trajectory will increase directly with the sampling rate of the joints, and that detection of a single joint is sufficient to reveal the location of the walker.

Direction and coherence discrimination of biological motion varied more with undersampling than did detection, so, at low sampling rates, direction could not be discerned at noise levels that were nearly two logarithmic units below the detection threshold. If there were a linear integrator (or matched filter) that summates the motion signal along all joint trajectories, then sensitivity for direction and coherence discrimination should increase linearly with sampling rate, as it does for the examples mentioned above. The far steeper slope observed for biological motion may indicate that these types of task are mediated by detectors with efficiency that varies dynamically with the signal strength. It has been suggested that biological motion depends on virtual links between appropriate joints, requiring the simultaneous analysis of both joints[1,16–18]. This idea is intuitively appealing, but the requirement of coincidence does not in itself predict the steep summation, because it affects both signal and noise equally, so the theoretical prediction is still a curve of unit slope (or an even shallower slope, depending on the assumptions about the coincidences produced by random noise).

Our results indicate that there is probably no specialized hardware for biological motion that acts like an ideal detector, as seems to exist for other forms of simple and complex motion or spatial vision[6,11–15]. Instead, the results show that biological motion may be analysed by very sensitive, but flexible, mechanisms with variable efficiency. One method of achieving variable efficiency would be by a system that adjusts its internal noise (or threshold) to respond optimally to the stimulus under all conditions. The advantage of such a system over the seemingly simpler matched-filter systems may be that it can optimize a limited neural resource for the analysis of the much wider range of stimuli that can yield information about spatial structure from motion. □

....

Methods

Both biological and translational motion were created with a 'limited-lifetime' technique, at a sampling rate of 30 Hz (Fig. 1). Each signal dot moved to the next position in the motion sequence, then disappeared; it was 'reborn' at a

joint chosen randomly from those not sampled in the previous sequence. All signal and noise dots were 7 arcmin in diameter, and were half-white and half-black against a grey background of 20 cd m^{-2} (thus causing no change in mean luminance). The walker was generated by Cutting's algorithm[7] (with no net translation) from a randomly chosen starting position, usually at 0.75 gait-cycles per second, with the individual dots moving at an average speed of 1.8 deg s^{-1}. For translation, dots were placed in random positions over the area of the walker, and all moved at 1.8 deg s^{-1} (to obtain results shown in Fig. 2). The walker was symmetrical about the vertical midline (particularly the upper body), so spatial cues alone could not aid discrimination.

Observers fixated the centre of a Barco Calibrator monitor (frame rate 180 Hz) from a distance of 80 cm. After a warning signal, the target was presented to either the left or the right of fixation, with an appropriate density control on the other side (both stimuli subtended $3.4 \times 5.7°$, centred $2.4°$ from fixation). For the walker, the density control was derived from the walking algorithm by randomizing the order of the frames presented. For translation, the control dots were displayed in new random positions within the region on each frame. Thus both target and control were dynamic, but only in the target was the motion coherent and smooth. Detection of either class of stimuli could be based on a judgement of smoothness of motion of individual dots. Discrimination was a two-stage process, in which observers first selected which side contained the target, and then identified the direction of the moving dots (for translation), the direction of ambulation (for discrimination of biological motion), or whether the upper and lower body of the walker moved coherently. As the discrimination thresholds for translation were similar to those for detection, it is unlikely that the two-stage task was an impediment to performance.

Dynamic random noise, comprising dots of similar size and colour, was scattered over the entire screen. The density of the noise increased or decreased in each trial, depending on the correctness of the observer's response (following the adaptive procedure QUEST[19], without feedback). There were 200–400 trials for each condition, with sensitivity defined as the noise level at which 75% correct responses are made; sensitivity was calculated by fitting a raised cumulative gaussian curve (with asymptotes at 0.5 and 1) to the psychometric functions. For spatial summation (Fig. 2), stimuli were presented for 1,200 ms (almost one complete gait-cycle, comprising 40 frames) centred within a noise window of 1,400 ms. For temporal summation (Fig. 3), the stimulus interval varied within a noise window of 7 s.

1. Johansson, G. Visual perception of biological motion and a model for its analysis. *Percept. Psychoph.* **14**, 201–211 (1973).
2. Marey, E.-J. *Le Mouvement* (Masson, Paris, 1894).
3. Mather, G. & Murdoch, L. Gender discrimination in biological motion displays based on dynamic cues. *Proc. R. Soc. Lond. B* **259**, 273–279 (1994).
4. Dittrich, W. H., Troscianko, T., Lea, S. & Morgan, D. Perception of emotion from dynamic point-light displays represented in dance. *Perception* **25**, 727–738 (1996).
5. Fox, R. & McDaniel, C. The perception of biological motion by human infants. *Science* **218**, 486–487 (1982).
6. Morrone, M. C., Burr, D. C. & Vaina, L. Two stages of visual processing for radial and circular motion. *Nature* **376**, 507–509 (1995).
7. Cutting, J. A program to generate synthetic walkers as dynamic point-light displays. *Behav. Res. Methods Instrument* **10**, 91–94 (1978).
8. Watson, A. B. Probability summation over time. *Vision Res.* **19**, 515–522 (1979).
9. Burr, D. C. Temporal summation of moving images by the human visual system. *Proc. R. Soc. Lond. B* **211**, 321–339 (1981).
10. Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966).
11. Maloney, R. K., Mitchison, G. J. & Barlow, H. B. Limit to the detection of Glass patterns in the presence of noise. *J. Opt. Soc. Am. A* **4**, 2336–2341 (1987).
12. Wilson, H. R., Wilkinson, F. & Asaad, W. Concentric orientation summation in human form vision. *Vision Res.* **37**, 2325–2330 (1997).
13. Lappin, J. S. & Bell, H. H. The detection of coherence in moving random-dot patterns. *Vision Res.* **16**, 161–168 (1976).
14. Barlow, H. B. & Tripathy, S. P. Correspondence noise and signal pooling in the detection of coherent visual motion. *J. Neurosci.* **17**, 7954–7966 (1997).
15. Burr, D. C., Morrone, M. C. & Vaina, L. Large receptive fields for optic flow direction in humans. *Vision Res.* **38**, 1731–1743 (1998).
16. Cutting, J. E. Coding theory adapted to gait perception. *J. Exp. Psychol.* **7**, 71–87 (1981).
17. Marr, D. & Vaina, L. Representation and recognition of the movements of shapes. *Proc. R. Soc. Lond. B* **214**, 501–524 (1982).
18. Ullman, S. in *Human and Machine Vision* (eds Beck, J., Hope, B. & Rosenfeld, A.) 459–480 (Academic, New York, 1983).
19. Watson, A. B. & Pelli, D. G. QUEST: A Bayesian adaptive psychometric method. *Percept. Psychophys* **33**, 113–120 (1983).
20. Mather, G., Radford, K. & West, S. Low-level visual processing of biological motion. *Proc. R. Soc. Lond. B* **249**, 149–155 (1992).

Correspondence and request for materials should be addressed to D.C.B. (e-mail: daye@in.pi.cnr.it).

Temporal dynamics of chromatic tuning in macaque primary visual cortex

Nicolas P. Cottaris* & Russell L. De Valois*†

* Program in Vision Science and † Department of Psychology, University of California Berkeley, 3210 Tolman Hall, California 94720, USA

The ability to distinguish colour from intensity variations is a difficult computational problem for the visual system because each of the three cone photoreceptor types absorb all wavelengths of light, although their peak sensitivities are at relatively short (S cones), medium (M cones), or long (L cones) wavelengths. The first stage in colour processing is the comparison of the outputs of different cone types by spectrally opponent neurons in the retina and upstream in the lateral geniculate nucleus[1–3]. Some neurons receive opponent inputs from L and M cones, whereas others receive input from S cones opposed by combined signals from L and M cones. Here we report how the outputs of the L/M- and S-opponent geniculate cell types are combined in time at the next stage of colour processing, in the macaque primary visual cortex (V1). Some V1 neurons respond to a single chromatic region, with either a short (68–95 ms) or a longer (96–135 ms) latency, whereas others respond to two chromatic regions with a difference in latency of 20–30 ms. Across all types, short latency responses are mostly evoked by L/M-opponent inputs whereas longer latency responses are evoked mostly by S-opponent inputs. Furthermore, neurons with late S-cone inputs exhibit dynamic changes in the sharpness of their chromatic tuning over time. We propose that the sparse, S-opponent signal in the lateral geniculate nucleus is amplified in area V1, possibly through recurrent excitatory networks. This results in a delayed, sluggish cortical S-cone signal which is then integrated with L/M-opponent signals to rotate the lateral geniculate nucleus chromatic axes[4–5].

The term 'receptive field' is used traditionally to characterize how a neuron responds to stimuli in different spatial locations, thus referring to a spatial receptive field[6]. Here we are concerned with determining how a cortical neuron responds to stimuli consisting of chromatic shifts from a white point to different locations in colour space, and how this responsivity develops over time. We are therefore studying a chromatic–temporal receptive field. To examine the cortical transformation of lateral geniculate nucleus (LGN) chromatic signals directly, we specified stimulus chromaticity in the MacLeod–Boynton–Derrington–Krauskopf–Lennie (MBDKL) isoluminant plane[7,8]. This plane is defined by two axes whose chromaticities isolate responses from the two types of LGN opponent neuron[8]: the $0°$ to $180°$ axis isolates responses from L/M-opponent LGN neurons ($0°$: L − M, 'pinkish-red'; $180°$: M − L, 'cyan'); and the $90°$ to $270°$ axis isolates responses from S-opponent LGN neurons ($90°$: S − (L + M), 'violet'; $270°$: − S + (L + M), 'greenish-yellow'). A cortical cell that integrates signals from both types of colour-coding geniculate cell, as most cortical cells do, would have a preferred chromaticity at an intermediate angle, depending on the relative weights and timings of the cell's inputs.

To study the structure of the chromatic–temporal receptive field, we probed neurons with spatially uniform chromatic stimuli presented in a fast (30-ms flashes), pseudorandom sequence and analysed the neuronal response using the reverse-correlation procedure[9,10]. Stimulus dimensions were one to two times the classical receptive-field dimensions, as measured with *m*-sequence receptive-field mapping[11]. For each action potential (spike) fired, we determined which chromatic stimulus had been presented at various preceding times. Spikes were accumulated in a two-dimensional