Spatial invariance of visual receptive fields in parietal cortex neurons

Jean-René Duhamel*, Frank Bremmer**, Suliann BenHamed* & Werner Graf†

* Laboratoire de Physiologie de la Perception et de l’Action, CNRS- Collège de France, 11 Place Marcelin Berthelot, 75005 Paris, France
† Department of Zoology and Neurobiology, Ruhr-University Bochum, D-44780 Bochum, Germany

Spatial information is conveyed to the primary visual cortex in retinal coordinates. Movement trajectory programming, however, requires a transformation from this sensory frame of reference into a frame appropriate for the selected part of the body, such as the eye, head or arms. To achieve this transformation, visual information must be combined with information from other sources: for instance, the location of an object of interest can be defined with respect to the observer’s head if the position of the eyes in the orbit is known and is added to the object’s retinal coordinates. Here we show that in a subdivision of the monkey parietal lobe, the ventral intraparietal area (VIP), the activity of visual neurons is modulated by eye-position signals, as in many other areas of the cortical visual system. We find that individual receptive fields of a population of VIP neurons are organized along a continuum, from eye to head coordinates. In the latter case, neurons encode the azimuth and/or elevation of a visual stimulus, independently of the direction in which the eyes are looking, thus representing spatial locations explicitly in at least a head-centred frame of reference.

The occipito–parietal cortical visual pathways contain several anatomically and functionally distinct areas which have been shown in monkeys to participate in specific aspects of spatial analysis and in movement preparation. One of these areas is the ventral intraparietal area. Area VIP receives a major visual projection from the middle temporal area (MT), the principal cortical relay for motion information. Consistent with this input, most of its neurons are sensitive to the direction and speed of moving visual stimuli. Although the prevalent sensory modality is visual, VIP neurons are often multimodal, responding to visual, tactile and vestibular stimulation, with parallel preferred directions for all three modalities. The somatic representation emphasizes the face region, and neurons with visual and tactile responses have spatially congruent receptive fields (RFs): that is, cells with central visual RFs have tactile RFs located in the middle of the face, and cells with peripheral visual RFs have somatic RFs located on the side of the head. VIP is reciprocally connected to portions of the premotor cortex responsible for head movements, oral prehension and coordinated hand/mouth actions. Thus, this area offers an opportunity to investigate multimodal encoding of reference frames and associated coordinate transformations by parietal neurons.

Previous studies that have addressed these issues have found that neurons in most visual areas of the cortex have an RF that remains anchored to the retina. There is some evidence, however, that direct encoding of non-retinal coordinate frames exists at the single-cell level. In the lateral intraparietal area (LIP), an area that participates in saccadic eye movements, neurons signal an eye-movement trajectory towards a location in space, even when it is not equivalent to the retinal coordinates of this location. Another study using hand mapping of RF contours described neurons in area PO/V6 responding to a position in space, irrespective of gaze direction. Finally, in bimodal visual–tactile premotor neurons, visual responses were found to remain anchored to the body part containing the tactile RF, and not to a particular part of the visual field, thus encoding spatial locations in ‘body-parts’ coordinates.

We have now investigated the RF of area VIP neurons by using a protocol designed to dissociate a retinal, or eye-centred coordinate system from a head-centred coordinate system. We applied a...
receptive field mapping procedure that allows rapid stimulation of a large number of locations in the visual field of head-fixed monkeys who were trained to look at one of several possible fixation targets. Different results were expected in this task depending on whether a neuron's RF encoded eye-centred or head-centred coordinates. An eye-centred RF should move rigidly from one fixation location to the next, in the same direction and with the same amplitude as the eyes (Fig. 1a, left). By contrast, a head-centred RF should remain fixed in space despite changes in eye position (Fig. 1a, right). RFs of single neurons were calculated for each fixation location by converting the raw spike trains evoked by each stimulus into a mean firing rate and, after data interpolation, by plotting the results as a colour-coded contour map (Fig. 1b).

We found a wide range of RF types. Some neurons had an RF that moved rigidly with the eyes, whereas other neurons encoded the same location in space irrespective of eye position. The plots in Fig. 2a show, for a single VIP neuron, the distribution of neural activity over the stimulated screen area for nine different eye fixation positions. In each map, the most active region is located in the upper left quadrant of the screen. The fact that the cell's RF remains fixed relative to the stimulation screen indicates that it does not encode a fixed retinal location, but that the retinal locations that drive this neuron vary with eye position. RFs measured at different eye positions spaced 20° apart can be directly compared in Fig. 2b, where contour plots traced at half the maximum firing rate are drawn in both head (screen) and eye coordinates. The best alignment is obtained for RF contours represented in head coordinates. Neurons like the one shown in Fig. 2 exhibited complete compensation for eye displacements, but other neurons compensated partially, sometimes to different degrees for the horizontal and vertical components of eye position, a pattern also observed in a small number of area V6/PO neurons. Figure 3 shows a neuron whose RF remained fixed relative to the screen when the eyes moved vertically, but shifted partially when the eyes moved horizontally. The upper horizontal borders of the RFs overlap precisely when plotted in screen coordinates (Fig. 3b, left) but appear scattered when plotted in eye coordinates (Fig. 3b, right). The right vertical borders of the RF move in conjunction with the horizontal eye displacements. However, because they do so only by about half of the amplitude of the eye displacement, the vertical borders do not overlap better in eye-centred than in screen coordinates. This neuron therefore encodes the elevation of a stimulus in head-
centred coordinates but it encodes azimuth in an intermediate manner, which is neither fully head-centred nor fully eye-centred.

To quantify the extent to which an RF moves with the eyes, we applied two-dimensional cross-correlation between the nine different maps obtained for the same cell. This procedure allows us to define how much an RF map obtained at a given fixation must be shifted relative to a map obtained at a different position in order to maximize the correlation coefficient. For head-centred RFs, the highest correlation is obtained when the two maps are perfectly superimposed (null shift), whereas for eye-centred RFs, it is obtained when the maps are shifted by an amount equal to the difference in eye position between the two mapping conditions. Thus, the expected ratio of the computed shift to the difference in eye position is $1$ for head-centred RFs, and $0$ for eye-centred RFs. For the neuron shown in Fig. 2, the values of the mean shift ratios were $0.04$ for vertical and $0.01$ for horizontal eye displacements, reflecting a stable position in space of this cell’s RF. For the neuron shown in Fig. 3, the shift ratios were $0.09$ and $0.55$ for vertical and horizontal eye displacements, respectively. The results for 66 VIP neurons are summarized in Fig. 4. There is a continuum between cells whose RF is fixed to the eyes and cells whose RF is fixed to the screen. Between the two extremes were cells whose RF shifted partially or asymmetrically with respect to horizontal and vertical eye positions. We tested whether the distribution of RF shift index was bimodal by using cluster analysis, with a two-means clustering algorithm and horizontal and vertical shift ratios as input variables. This procedure grouped the cells in two normally distributed clusters that were statistically highly different from each other. It seems that although RF shifts as a function of eye position are measured on a continuous scale, the population of VIP neurons studied might actually contain two distinct subpopulations of RF types, linked by an area of gradual transition between retinal and head-centred coordinates.

As already mentioned, for a cortical area to process a head-centred coordinate frame, retinal and eye position input must be combined. How these two sources of information interact is still unclear. One possibility is that higher-order forms of space representation may be computed from the distributed activity of populations of retinotopic visual neurons whose sensitivity is gated by eye position. Over half of the neurons (35/66) recorded here showed a significant eye-position effect (at the 0.05 level or better). In most cases (26/35), the effect takes the form of a planar ‘gain field’ which can be approximated by two-dimensional linear regression. Such eye-position effects occurred across the whole continuum of eye- and head-centred neurons. The cell illustrated in Fig. 3 is a typical example, in which the size of the visually evoked discharge in the centre of the RF increases monotonically as the fixation point moves from the left to the right part of the screen. Several cortical areas, including V3A, MT, MST, PO, 7A, LIP and the premotor cortex, contain neurons whose overall activity is gated by a coarse eye position in a manner similar to our VIP neurons. According to the population coding hypothesis, all of these areas carry the necessary signals for generating a distributed head-centred representation of visual space. However, area VIP and probably also area PO/V6 (ref. 25) stand out because there head-centred locations are encoded at the single neuron level. This suggests that space may be represented in the cortex both at the population level and at the single cell level. It is also possible that the eye-position signals that are widely distributed throughout the visual system are used in certain areas to perform spatial transformations while they are merely being passed on by the others.

The functional heterogeneity of the posterior parietal cortex suggests that space is not represented as a single map but instead as a collection of interconnected, specialized areas that play a role in different behaviours. VIP neurons are sensitive to tactile stimuli on the head, to near or approaching visual stimuli, and to vestibular head-motion signals. Although no physiological evidence has as yet linked area VIP to a specific motor behaviour, head-centred coordinates could provide a common reference frame to encode the multisensory signals that converge onto it, and provide targeting and feedback information to premotor areas in the frontal lobes, where neurons describe goals for motor acts using the same coordinate system. The frame of reference encoded by VIP neurons is at least head-centred. However, it is not possible at this point to assert that this represents an end-stage computation. Head-centred RFs may themselves correspond to an intermediate step towards other reference frames. Oculomotor coordinates could be computed before eye movements from a target’s head-centred location by subtracting out eye position. Neck proprioception and vestibular signals could also make as-yet unrecognized contributions. As in our experiments the head was fixed relative to the body and the body fixed relative to the laboratory environment, future studies will have to examine whether individual parietal neurons can provide even higher-order, body-centred or environment-centred reference frames.

Methods

General. We recorded from three hemispheres of two macaque monkeys (Macaca fascicularis and Macaca mulatta). Monkey care, training and surgical procedures were carried out in accordance with local and European Community directives. The monkeys were trained to maintain stable fixation on a small spot of light with their heads fixed in return for a liquid reward. Visual stimuli were computer-generated and back-projected by a liquid crystal display system on a translucent screen. Eye position was continuously monitored and the monkeys were required to maintain the eyes within a $2^\circ$ wide tolerance window for uninterrupted periods lasting up to 3.5 s. Single unit activity was recorded extracellularly with tungsten microelectrodes. Histological verification of electrode penetrations is not yet available because the animals are still being used for experiments. The location of area VIP within the recording area was identified on the basis of a set of physiological criteria.

Figure 4 Ratio of horizontal and vertical displacement of the RF to horizontal and vertical displacement of the eyes for all VIP neurons studied. Individual points in the scatter plot represent the mean normalized shift averaged over all possible pairwise comparisons (n = 36) of the nine RFs obtained for a given neuron (total: n = 66). Open and closed circles correspond to two subpopulations identified through cluster analysis. The mean horizontal and vertical shift indices were 0.20 and 0.00 for the first cluster, and 0.88 and 0.79 for the second cluster (horizontal $F(1,64) = 121.3$; vertical $F(1,64) = 94.7$; both $P < 0.0001$). On top and to the right of the scatterplot are shown the distribution of horizontal and vertical shift ratios, and the normal curve fits to each subpopulation satisfying the Kolmogorov–Smirnov normality test.
described previously\textsuperscript{17,18,20}. Briefly, area VIP was found in the fundus of the intraparietal sulcus, at ~7 mm from the cortical surface during electrode penetrations made parallel to the sulcus; the area extends for ~2–3 mm. Access from the lateral bank shows an abrupt transition from non-direction-selective visual and sacadic-related activity which characterizes the lateral intraparietal area (LIP), to direction-selective visual responses often accompanied by direction-selective somatosensory response in the face and the head region. Access to VIP from the medial bank of the sulcus through area 5 and the medial intraparietal area is characterized by an abrupt transition from purely hand or arm somatosensory activity, to brisk direction-selective visual and face somatosensory responses.

**Receptive field mapping.** Each recorded neuron’s visual RF was initially located manually and optimal stimulus parameters were identified. The RF was subsequently measured quantitatively while the monkey fixated at one of nine possible locations spaced 10° apart, from 10° up to and the left of the screen centre, to 10° down and to the right. The visual stimulation area covered a 70°-by-70° surface, divided into a virtual square grid of 49 non-overlapping subregions, each one subtending 10° by 10°. A single stimulus consisted of a 10°-by-1° white bar appearing at one edge of a given subregion, moving in the optimal direction perpendicular to the orientation of the bar at a constant velocity of 100 deg·s\(^{-1}\) for 100 ms and disappearing, thus covering exactly 10° during a single sweep. The stimulus reappeared at a different location 300 ms later, moved for 100 ms, disappeared, and so on. A stationary 10°-by-10° white square was used when a moving bar did not elicit a reliable response but a flashed stationary stimulus did. Six to eight stimuli were presented in this rapid sequence at randomly selected locations in the course of a single successful fixation trial. The location of the fixation spot also varied randomly from one trial to the next. Nine separate RF maps were thus obtained concurrently during a single experiment. For each fixation location, we recorded a complete stimulation grid with 6–10 stimulus presentations per grid position.

RF maps were constructed off-line by counting the total number of spikes evoked by stimulating a given grid subregion using a shifted temporal window adjusted to the cell’s response latency. Spike count was averaged over the number of presentations and converted into mean firing rates. The resulting 7 × 7 raw matrix was subsequently converted into a 25 × 25 matrix by linear interpolation of three new points between each original data point, and plotted as colour-coded contours connecting isofrequency regions. The number of contours in the plot correspond to the discretization of the firing rate, scaled to the maximum firing rate over the nine RF maps obtained for a given neuron. The contour outlines plotted in panels b of Figs 2 and 3 correspond to a line traced around the area for which firing is at least 50% of the peak rate for the corresponding RF map.

**Cross-correlation analysis.** An estimation of RF displacement in conjunction with changes in eye position was obtained using a cross-correlation analysis. This approach allows us to make use of all the information available in the neural activity maps. It provides a more robust estimate of RF displacement than measures based on peak location or centre of mass, especially when an eccentric portion of the RF exceeds the stimulation area. It is also insensitive to variation in the absolute level of activity associated with the presence of a gain field. In essence, the procedure involves calculating correlation coefficients between two maps (for example, the interpolated activity matrices) that are systematically displaced relative to one another step by step, row-wise and column-wise. The shift associated with the highest correlation coefficient (\(t_{\text{max}}\)) was normalized with respect to the difference in eye position (\(D_{\text{eye}}\)), yielding the ratio \(t_{\text{max}}/D_{\text{eye}}\). For a neuron strictly encoding eye-centred coordinates, the shift that maximizes the correlation coefficient corresponds to the difference in eye position associated with the maps. For example, when cross-correlating a pair of maps obtained during fixation at 10° to the left and at 10° to the right, \(t_{\text{max}}\) is obtained with a pure horizontal 20° shift. As \(D_{\text{eye}}\) is also equal to 20°, \(t_{\text{max}}/D_{\text{eye}}\) will be equal to 1.00. By contrast, in the case of a neuron whose RF is spatially invariant, the non-shifted alignment yields the highest correlation, resulting in \(t_{\text{max}}\) and \(t_{\text{max}}/D_{\text{eye}}\) of zero. A mean shift ratio was computed from the 36 different pairs formed by nine different RF maps.

**Eye-position effects.** A global influence of eye position can be observed on baseline activity in darkness or on visually evoked activity. To measure the effects on baseline activity, we counted the number of spikes present during the 400 ms fixation period preceding the onset of the first stimulus in each trial. Analyses conducted on visually evoked activity were based on the number of spikes in the post-stimulus window for all locations evoking discharges above 50% of the maximum firing rate. For both sets of measurements, differences in firing rate between the nine fixation locations were screened with a distribution-free ANOVA. Neurons with a significant effect were further examined using a two-dimensional linear regression in order to determine if a simple oriented plane could accurately describe the observed variation. Analyses of visually evoked activity were restricted to cells whose RF remained within the boundaries of the mapped area at all eye positions, in order to avoid identifying artefactual eye-position effects. For this subpopulation, a comparison between results obtained with baseline and visually evoked activity showed similar proportions of cells with and without significant effects, and of cells with reliable planar gain field, in agreement with previous observations\textsuperscript{2}. Baseline activity can therefore be considered a reliable indicator of eye-positioning, and the results reported for the full population were based on this parameter.