One major difficulty with this approach is that it is based on a verbal description of image features rather than on the raw intensity values in the image itself. This makes it difficult to test under more general conditions. In order to apply this analysis to a more general class of images, it would first be necessary to construct operators that extract the feature descriptions being invoked—a task that has yet to be accomplished. Before embarking on such a difficult approach it is worth asking whether simpler extracted properties, such as those derivable from linear filters, will suffice (see refs 4–7).

When inspecting this texture one may observe that the Xs look smaller than the Ls, and that they break up the background differently. This suggests that very simple size-tuned mechanisms, such as cells with centre-surround receptive fields, could play an important role in the discrimination. We changed the relative sizes of the Xs and the Ls to see whether we could increase and decrease the discriminability of the patterns.

Figure 1*b* shows the result of lengthening the bars of the Ls by 25%. The bar intensities have been compensated so that the overall density of the micropatterns (that is, the equivalent amount of ink in each) is unchanged. The discrimination becomes easier. Thus, although the micropatterns still have the same number of terminators, corners, and so on, the manipulation of size has a significant impact on the discriminability of the texture.

Figure 1*c* shows the result of making the bars of the Ls 25% shorter than in the original textures, again with compensation in the intensity of the bars. Now the discrimination is more difficult.

Figure 1*d*–*f* shows the response of perhaps the simplest size-tuned mechanism we can construct: a linear centre-surround receptive field followed by full-wave rectification. Figure 1*d* shows the response to the stimulus of Fig. 1*a*; the mechanism responds more strongly to the patch in the centre. Figure 1*e* shows the response to Fig. 1*b*; now the differences are even more apparent. Figure 1*f* shows the response to Fig. 1*c*; in this case the size-tuned mechanism gives responses of similar strength to the two textures.

For this particular set of textures, then, the discriminability can be predicted fairly well from the activities of size-tuned units, without reference to more feature-like properties of the micropatterns. We suggest that the visual system uses a two-stage cascade of local energy measures (similar to the cascade of orientation measures discussed by Knuttson and Granlund[5].) In the first stage, linear filters are followed by a rectifying nonlinearity (as in fig. 1*d*–*f*); spatial averaging provides primary energy measures. These responses are then treated as image arrays for input to a further layer of linear filters, which compute secondary energy measures that indicate the locations of texture boundaries.

Models for texture perception that are based on concepts such as 'terminators' and 'corners' have been important in motivating research in early vision, but the models have proven difficult to formalize in such a way that they can be applied to wide classes of textures. Although we do not present a full model of texture perception here, the above demonstration indicates that simple filtering processes operating directly on the image intensities can sometimes have surprisingly good explanatory power. The accompanying paper by Voorhees and Poggio[8], based on a computational investigation into texture analysis, offers an example of a more fully elaborated theory and further demonstrates the potential power of simple processes in early vision.

Received 20 January; accepted 26 February 1988.

1. Beck, J. *Percept. Psychophys.* **1**, 300–302 (1966).
2. Beck, J. *Percept. Psychophys.* **2**, 491–495 (1967).
3. Bergen, J. R. & Julesz, B. *Nature* **303**, 696–698 (1983).
4. Caelli, T. *Spatial Vision* **1**, 19–30 (1985).
5. Knuttson, H. & Granlund, G. H. (1983) *IEEE Workshop of Computer Architecture for Pattern Analysis and Image Data Base Management* 12–14, Pasadena (IEEE Computer Society Press, Washington DC, 1983).
6. Voorhees, H. & Poggio, T. *Proceedings, First International Conference on Computer Vision*, 250–258, London (IEEE Computer Society Press, Washington DC, 1987).
7. Bergen, J. R. & Adelson, E. H. *J. opt. Soc. Am.* **A3**, 99 (1986).
8. Voorhees, H. & Poggio, T. *Nature* **333**, 364–367 (1988).

Computing texture boundaries from images

Harry Voorhees* & Tomaso Poggio

Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Recent computational and psychological theories of human texture vision[1–3] assert that texture discrimination is based on first-order differences in geometric and luminance attributes of texture elements, called 'textons'[4]. Significant differences in the density, orientation, size, or contrast of line segments or other small features in an image have been shown to cause immediate perception of texture boundaries. However, the psychological theories, which are based on the perception of synthetic images composed of lines and symbols, neglect two important issues. First, how can textons be computed from grey-level images of natural scenes? And second, how, exactly, can texture boundaries be found? Our analysis of these two issues has led to an algorithm that is fully implemented and which successfully detects boundaries in natural images[5]. We propose that blobs computed by a centre-surround operator are useful as texture elements, and that a simple non-parametric statistic can be used to compare local distributions of blob attributes to locate texture boundaries. Although designed for natural images, our computation agrees with some psychophysical findings, in particular, those of Adelson and Bergen (described in the preceding article[6]), which cast doubt on the hypothesis that line segment crossings or termination points are textons.

How can texture elements be computed from images of natural scenes? From the grey-level image our algorithm extracts 'blobs'—small compact and elongated linear regions (also known as 'bars') which are darker or lighter than their surround. To detect relatively dark blobs, the algorithm works as follows. First, the image is convolved with a centre-surround filter at a fine scale (we typically use a laplacian of gaussian of standard deviation 1.5 pixels). By using the logarithm of intensity, the resulting blob regions are fairly insensitive to shadows or other large-scale changes in illumination. Positive values of the filtered image indicate regions which are relatively darker than their surround. Hence, blobs can be regarded as the duals of intensity edges, which are often computed as the zero crossings of the filtered image. To remove some spurious connections between blob regions, the filtered image is thresholded about a small positive value. The threshold is proportional to an estimate of noise in the image, which is automatically computed from a histogram of the intensity gradient magnitude[5].

This computation yields a map of dark blob pixels. Using morphological operations, these areas are segmented on the basis of local shape into small compact and thin elongated components. For each of these blobs, five attributes are computed—contrast, orientation, width, length, area and area density. Finally, spurious or noisy blobs are removed by requiring that the intensity gradient at most points along a blob's perimeter are above a noise threshold and pointing outwards.

We have found that in natural images, the computed blobs and bars capture physically meaningful information about the surface—small impressions and markings—and that attributes

* Present address: Thinking Machines Corporation, 245 First Street, Cambridge, Massachusetts 02142, USA.

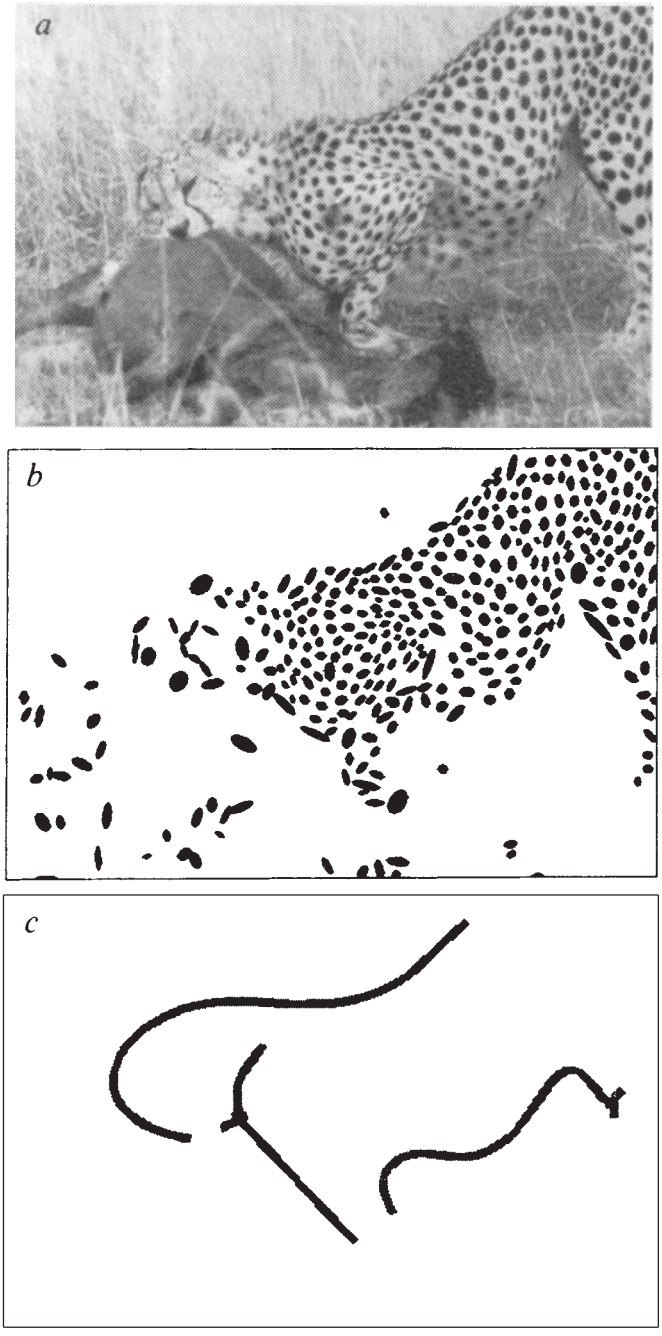

Fig. 1 *a*, Image of a leopard (from *Time*; 23 February 1987, page 40). *b*, The computed textons (blobs and bars). *c*, Texture boundaries due to density as found by our algorithm.

of these blobs characterize the texture. The algorithm, therefore, provides a method of computing discrete 'textons' from natural images in the spirit of Marr's raw primal sketch[1]. Figures 1 and 2 show the result of applying our algorithm to natural images of textured surfaces.

Once blobs are computed, how can their attributes be compared to locate texture boundaries? Specifically, what statistic is used to compare two distributions of attributes from nearby neighbourhoods? We define a non-parametric statistic which reliably predicts under which circumstances a texture boundary will be perceived between two regions.

Mimicking the inability of humans to immediately discriminate small changes in attribute values, we use histograms rather than the exact sample distributions themselves. The histogram bucket size equals the minimum difference which is immediately discriminable (for example, 15° for orientation). From the standpoint of detecting physical boundaries, implementing this 'weakness' makes sense: if we did not ignore small differences in attribute values, a graded texture gradient, perhaps formed by the projection of a curved surface, would yield undesirably significant texture boundaries across its face. Discretizing the distribution (with overlapping buckets) avoids this problem.

To compare two sample attribute distributions, we compute the maximum difference between corresponding histogram buckets. This maximum frequency difference (MFD) statistic is better suited than most standard statistics for this purpose, because it is less sensitive to outliers, which are common in natural images. The statistic also agrees with several human discrimination capabilities. For example, a region of randomly oriented bars is readily distinguished from a region of equally oriented bars, and the two distributions yield a high MFD. Conversely, we were surprised to discover that a region of bars randomly oriented between 0 and 90° is hard to discriminate from a region of bars randomly oriented between 90 and 180°. Any standard statistical test would regard these distributions as very significantly different. The MFD, however, is small, again in accord with perception[5].

A ubiquitous problem in texture discrimination is the choice of neighbourhood size. It has been said that large neighbourhoods are needed to robustly represent texton distributions in the presence of noise and outliers, but that small neighbourhoods are needed to localize or even to avoid missing certain texture boundaries. Our statistic seems to offer a way out of this dilemma. Because it is quite insensitive to outliers, small distributions (of about ten samples) often suffice to compute reliable difference measures. Such neighbourhoods are small enough to detect boundaries between regions which are as narrow as two or three textons across. In our implementation, the neighbourhood diameter is between 5 and 10 times the average blob width. These figures are based on some psychophysical results of Julesz[7], and are designed so that any neighbourhood of textons dense enough to be perceived as a texture is of sufficient size.

When applied to the synthetic images tested by Bergen and Adelson[6] in the previous article, our computation provides an alternative to the conjecture[3] that crossings are textons (see Fig. 3). In the left image, originally offered as evidence of crossings as textons, the crosses give smaller blobs of higher contrast than do the *L*s (Fig. 3*b*). This demonstrates that blob attributes offer an alternative to crossings as an explanation of human texture discrimination.

In the right image, where the size of the *L*s and the intensity of crosses are decreased, the computed blobs from each region are similar. Bergen and Adelson[6] reported in this case that discrimination is harder. It therefore appears that simple blob attributes offer not only an alternative, but a more accurate explanation of human discrimination capabilities. At the very least, the psychophysical results demonstrate that such spatial filtering cannot be neglected in a satisfactory explanation of human texture perception. Recent psychophysical experiments[8] also support this view.

We go on to illustrate how our algorithm computes texture boundaries in this example. Figure 3*c* shows histograms of blob contrast values from neighbourhoods in different regions in each of the two images, together with the computed MFD statistic. Figure 3*d* shows the MFD computed at a grid of points across each image array. Significantly high and long ridges in each array identify blob contrast boundaries, as shown in the left image of Fig. 3*e*. In the right image, no significant boundaries are detected. Similar results are obtained on this example using blob width as an attribute instead of blob contrast.

The boundaries computed by our algorithm are consistent with pre-attentive texture perception, without the need for crossings as textons. From the standpoint of analysing natural images, this result is not surprising. Seldom would a natural scene give texture elements shaped like characters, each having a distinct number of crossings or terminators. Although we have not

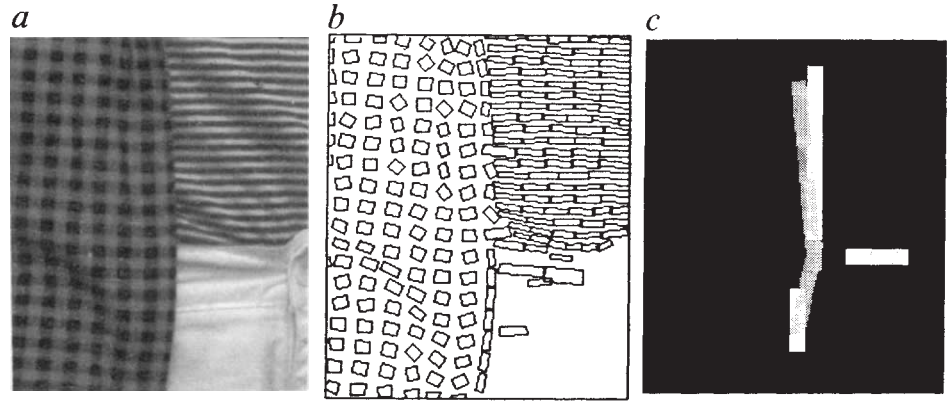

Fig. 2 *a*, Natural image of clothing. *b*, The computed textons (blobs and bars). *c*, Texture boundaries found by our algorithm. Boundaries due to differences blob width are shown in white with boundaries due to differences in blob contrast shown in grey. The algorithm has been applied to several other examples of natural texture[5].

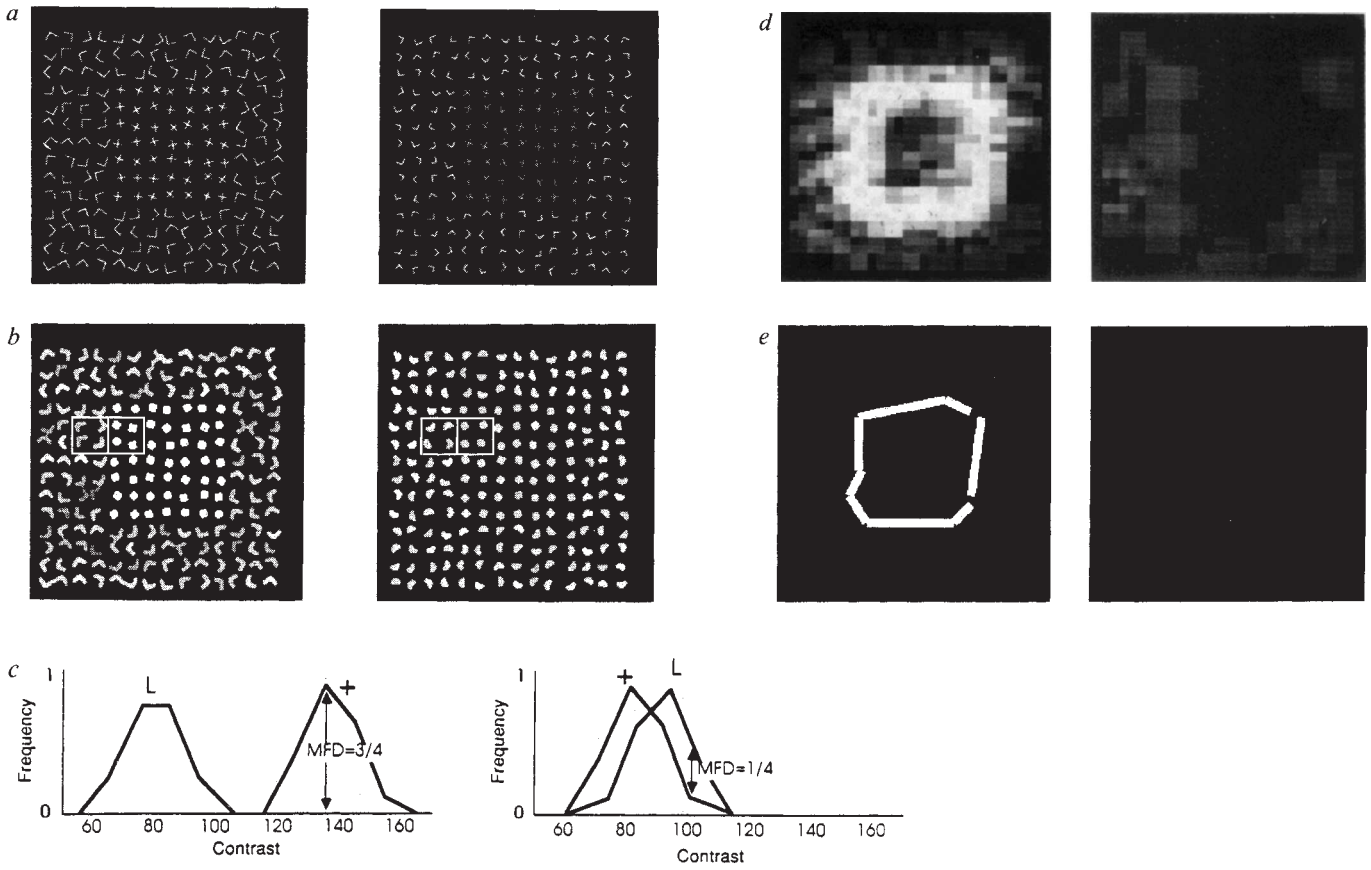

Fig. 3 Demonstration that crossings are not textons. *a*, (Left) Demonstration offered as evidence that crossings are textons. (Right) Bergen and Adelson's example[6] where discrimination is decreased. *b*, Blobs obtained by $\nabla^2 G$ filtering, shown here with intensity proportional to blob contrast. At left, the crosses give smaller blobs of higher contrast than do the *L*s. The gaussian filter, *G*, has standard deviation 2.5 pixels, and the image size is 256 pixels square. *c*, Superimposed histograms of blob contrast from adjacent neighbourhoods shown in *b*, with maximum frequency difference (MFD) statistic. In our implementation, the histograms are smoothed slightly to reduce the effect of discretization. *d*, MFD of contrast, computed from adjacent neighbourhoods over each entire array. *e*, Significantly high and long ridges of *d* identify contrast boundaries. Contrast accounts for the discrimination of the left image, without the explicit computation of crossings. No boundaries are detected in the right image, which humans also find harder to discriminate according to Bergen and Adelson (see preceding article[6]).

proven that blobs are textons any more than Julesz and Bergen[3] proved that crossings are textons, we believe that our explanation is preferable, because the computation works for natural images as well as for synthetic ones. Also, in natural images, blob attributes provide information about the physical cause of a texture boundary. Changes only in blob orientation, for example, suggest the presence of a surface orientation discontinuity, while large changes in size or density can only be due to occlusion or to a change in material[9].

Our experience with natural images suggests that most texture boundaries are probably due to differences in texton density. At a particular scale of analysis, two surfaces adjacent in the image are unlikely to give textons at the same scale of analysis (as in Fig. 1). It is possible that texture boundaries due only to changes in texton orientation or size (as in Fig. 2) are less common in images of natural scenes than synthetic, psychophysical images would lead us to believe.

We suspect that terminators may not be textons either. We have already shown how the discrimination of Fig. 3, which contains terminator density ratio of 4:2, can be explained by blob attributes. We also found, using images of squarish randomly oriented 6s and *E*s (similar to those of Fig. 6 of Julesz[4]), that even a terminator density ratio of 3:1 is not sufficient for discrimination. Breaks in lines (pairs of terminators) seem more

salient than ends of lines, and may be more relevant to natural textures (see ref. 5, Fig. 6.5). Breaks in dark lines may be computed as small-scale light blobs.

Our computation leaves many questions unanswered. We have not dealt with multiple scales of blob detection or the integration of various attribute boundaries. Nevertheless, we believe that we have extended perceptually based theories of texture vision to natural images by implementing an algorithm which computes textons, and which compares textons to locate texture boundaries, in accordance with human texture capabilities. We have successfully applied the algorithm to natural images, and have shown how it accounts for previous psychophysical results as well. Our algorithm demonstrates the feasibility of a first-order, symbolic approach to texture discrimination in the spirit of Marr's primal sketch[1] and Julesz's texton theory[4]. It is possible that our statistical test applied to the outputs of a sufficient number of arrays of linear filters would also provide acceptable texture segmentation. Such a scheme, similar to Bergen and Adelson's suggestion[6], remains to be elaborated, implemented and tested. It is possible that the main steps of the algorithm could be implemented in a relatively natural way by known neural structures in the retina and the visual cortex.

The authors thank Fero Simoncelli and Mike Sokolov for their help in producing the images.

Received 21 January; accepted 7 April 1988.

1. Marr, D. *Cold Spring Harb. Symp. quant. Biol.* Vol. XL, 647-662 (1976).
2. Beck, J., Prazdny, K. & Rosenfeld, A. in *Human and Machine Vision* (eds Beck, J., Hope, B. & Rosenfeld, A.) 1-38 (Academic, New York, 1983).
3. Julesz, B. & Bergen, J. R. *The Bell System Technical Journal* **62**, 1619-1645 (1983).
4. Julesz, B. *Nature* **290**, 91-97 ((1981).
5. Voorhees, H. *M.I.T. Artificial Intelligence Lab, Technical Report number 968*, 1987.
6. Bergen, J. P. & Adelson, E. H. *Nature* **333**, 363-364 (1988).
7. Julesz, B. *Biol. Cybern.* **54**, 245-251 (1986).
8. Beck, J., Sutter, A. & Ivry, R. *Computer Vision Graphics Image Proc.* **37**(2), 299-325 (1987).
9. Riley, M. D. *M.I.T. Artificial Intelligence Lab, Technical Report* 649 (1981).

A cellular analogue of visual cortical plasticity

Yves Frégnac, Daniel Shulz, Simon Thorpe* & Elie Bienenstock

Laboratoire de Neurobiologie et Neuropharmacologie du Développement, Bâtiment 440, Université Paris XI, F-91405, Orsay Cedex France

Neuronal activity plays an important role in the development of the visual pathway. The modulation of synaptic transmission by temporal correlation between pre- and postsynaptic activity is one mechanism which could underly visual cortical plasticity[1-4]. We report here that functional changes in single neurons of area 17, analogous to those known to take place during epigenesis of visual cortex[5,6], can be induced experimentally during the time of recording. This was done by a differential pairing procedure, during which iontophoresis was used to artificially increase the visual response for a given stimulus, and to decrease (or block) the response for a second stimulus which differed in ocularity or orientation. Long-term modifications in ocular dominance and orientation selectivity were produced in 33% and 43% of recorded cells respectively. Neuronal selectivity was nearly always displaced towards the stimulus paired with the reinforced visual response. The largest changes were obtained at the peak of the critical period in normally reared and visually deprived kittens, but changes were also observed in adults. Our findings support the role of temporal correlation between pre- and postsynaptic activity in the induction of long-lasting modifications of synaptic transmission during development, and in associative learning.

* Present address: Institut des Neurosciences, University Paris VI, F-75230, Paris, France.

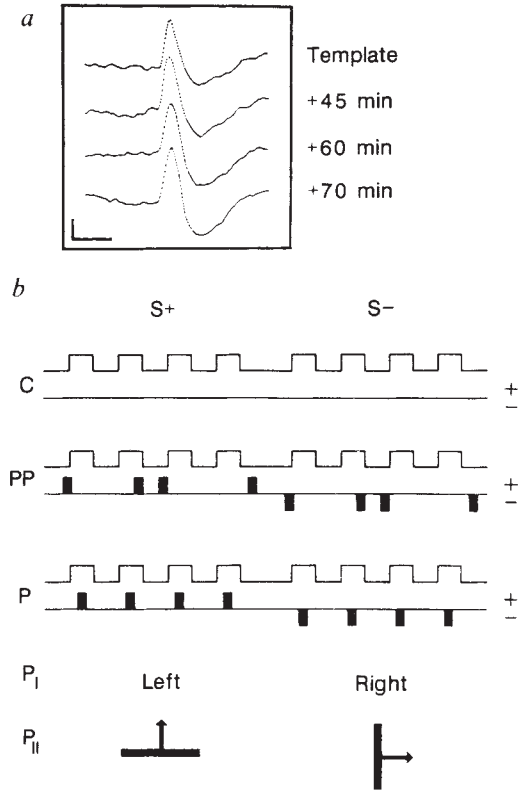

Fig. 1 Imposed temporal correlation between visual afferent activity and postsynaptic firing induced iontophoretically. *a*, For each cell, the shape of the action potential (calibration bars, 1 mV and 1 ms) was continuously monitored on a digital scope to ensure that the same neuron was recorded throughout the experiment. *b*, Each indentation indicates the temporal occurrence of a single visual stimulation, and each filled rectangle that of an iontophoretic pulse of a given polarity (shown at right). During control (C, upper row), two stimuli (S^+ and S^-) were presented by blocks of four trials in succession (upper line), without iontophoretic current (lower line). During pseudo-pairing (PP, middle row), iontophoretic pulses were uncorrelated with visual stimulation. During pairing (P, lower row) iontophoretic pulses were concomitant with the visual response, in such a way as to impose a significant increase (S^+) or decrease (S^-) of the visual response. Cases where no modulation of activity could be produced by iontophoresis are noted S^0 (see right eye stimulation in Fig. 2). The two test stimuli differed in either ocular dominance (P_I), or orientation (P_{II}). The relative preference response between the two test stimuli, given by the normalized ratio of visual responses $S^+/(S^+ + S^-)$, was measured during control periods in which the sequence of stimulation shown in row C was repeated 10-50 times. The two series of values taken from this ratio before and after pairing (or pseudo-pairing) were compared using both parametric (unpaired Student's *t*-test; significance level of $P < 0.005$) and non-parametric tests (Kolmogorov-Smirnov [K.S.]; significance level of $P < 0.05$). To assess non-associative effects, pseudo-pairing procedures (PP) were interposed between controls prior to pairing in some experiments. In the case of the orientation protocol (P_{II}), the full orientation tuning curves were analysed before and after pairing, to reveal possible generalization effects for stimuli other than those used during pairing.

Methods. In addition to standard anaesthesia and electrophysiological procedures detailed elsewhere[16], we used a simple method to artificially control the level of postsynaptic activity by varying the retention/ejection current of the 1-3 M potassium acetate or chloride extracellular recording electrode (2-20 MΩ). Positive current (average value of +4 nA) and increase in the concentration of potassium in the extracellular medium resulted (in 79% of the cases) in a significant increase in spontaneous and/or evoked activity. Negative current (average value of −9 nA) reduced the cells' activity through a field effect[17] (in 76% of the cases) which sometimes led to a total blockade of the response for the preferred stimulus (in 16% of visual cells).