Direct physiological evidence for scene segmentation by temporal coding

(cat/visual cortex/correlation analysis/synchronization/cell assembly)

ANDREAS K. ENGEL, PETER KÖNIG, AND WOLF SINGER

Max-Planck-Institut für Hirnforschung, Deutschordenstrasse 46, 6000 Frankfurt 71, Federal Republic of Germany

ABSTRACT Theoretical studies have suggested that scene segmentation may be accomplished by a temporal coding mechanism using synchronization of neuronal responses. Here we report a direct experimental test of this hypothesis. Neuronal responses were recorded simultaneously from two to four sites with overlapping receptive fields in cat visual cortex. Correlation analysis revealed that all cells synchronized their responses irrespective of their orientation preference when they were activated by a single light bar. However, when stimulated with two superimposed light bars of different orientations, the same cells segregated into distinct assemblies according to their orientation preferences. Within each of these assemblies responses were synchronized, but correlation was absent between the two assemblies. These results are compatible with the hypothesis that responses to individual objects in a scene are distinguished by synchrony, whereas responses to different objects show no temporal correlation, thus allowing for the segmentation of superimposed stimuli. We conclude that stimulus-specific synchronization of spatially distributed neuronal responses may provide a physiological mechanism for scene segmentation.

Since the early work of the Gestalt psychologists (1, 2) psychophysical research has established that scene segmentation and perceptual grouping represent important aspects of early vision. It is generally agreed (3) that, as a prerequisite for object recognition, a visual scene has to be parsed, and coherence between elementary features has to be evaluated to permit the discrimination of different objects within the scene. The physiological correlate of this segmentation process has not yet been identified with certainty. According to a now-classic proposal, segmentation should be accomplished by single neurons dedicated to certain feature constellations (4). However, such a model is not supported by experimental evidence and leads to a "combinatorial explosion," because every new pattern requires an additional dedicated unit. Therefore, alternative mechanisms have been suggested which imply that feature constellations are represented by cell assemblies rather than by the activity of single cells (5–10). A presumed advantage of assembly coding is that the combinatorial exhaustion of the system can be avoided, because one particular cell may participate in different assemblies (9). In the classical model introduced by Hebb (5), assemblies are defined by amplitude coding—i.e., the concurrent elevation of the average firing frequency of the participating cells for several hundred milliseconds. However, the Hebb mechanism does not permit the coactivation of two spatially superimposed assemblies in the same cortical region (9, 11). Because all cells raise their average firing frequencies, information about the partition into distinct assemblies is lost. This "superposition catastrophe" will result in false conjunctions of features and render object

recognition impossible (9, 11). Therefore, it has been suggested that precise synchronization of neuronal discharges in the millisecond range should be used to define functional assemblies and to achieve scene segmentation (7–9). According to these proposals, perceptual coherence within a segment of a scene should be expressed by response synchronization of feature-detecting neurons. In contrast, cells responding to different objects in a scene should be uncorrelated. This mechanism of assembly formation will in the following be referred to as temporal coding. As shown by simulation studies (11, 12), such a mechanism permits the coexistence of several assemblies in the same network and, thus, seems capable of solving the superposition problem.

Increasing physiological evidence suggests the presence of a temporal-coding mechanism in cortical networks that uses neuronal responses with a periodic temporal structure. As shown in recent studies by us and others (13–17), neurons in striate and extrastriate areas of cat visual cortex tend to discharge in an oscillatory manner. Because adjacent cells within a single orientation column have a strong tendency to oscillate synchronously (15, 16), we have adopted the term neuronal group for the designation of such coherently active cell clusters (6, 10). Subsequent studies (14, 17–19) have provided evidence that spatially separate groups within one visual area, as well as cell groups located in different areas, can synchronize their oscillatory responses. In addition, we have recently observed response synchronization across the cerebral hemispheres (20). These results suggest that the phase-relationship of neuronal oscillations may be used to define cortical assemblies—i.e., families of spatially distributed neuronal groups that represent visual objects. In this manner, temporal coding may be used to achieve perceptual grouping and scene segmentation (17–23). If this holds true, the following predictions should be confirmed experimentally: The synchronization between spatially separate cell groups should be stimulus dependent. Neuronal groups should oscillate in synchrony when they respond to the same perceptual object, whereas cell groups responding to two unrelated stimuli should oscillate in an uncorrelated manner. In particular, it should be demonstrated that this mechanism is capable of segregating stimuli that are superimposed in the visual field. In the present study, we used a simple experimental paradigm to test these predictions. Essentially, the experiment involves simultaneous recording from several cell groups with overlapping receptive fields. When activated with a single light bar, these groups will fire synchronously, even when they differ in their preferred orientation (18, 19). We have now tested whether the synchronization is affected by using two superimposed light bars as stimuli. These "conflicting stimuli" constitute a simple visual scene in which two figures have to be discriminated. A preliminary report of the results (21) has appeared.

MATERIALS AND METHODS

The data were collected from eight anesthetized and paralyzed adult cats. The animals were prepared and maintained

Neurobiology: Engel *et al.*

Proc. Natl. Acad. Sci. USA 88 (1991)   9137

as described in detail elsewhere (18, 19). Multiunit activity was recorded from area 17 of cat visual cortex by using an array of four or five platinum-iridium electrodes with a spacing of 400–500 μm. Thus, we were able to monitor the activity of coherently active cell groups within different orientation columns (10, 15). The array was centered on the representation of the area centralis and was oriented along the anterior–posterior axis. Due to this electrode arrangement, the groups of cells recorded by different electrodes usually had overlapping receptive fields but differed in their preferred orientation. At the onset of a recording session, size and location of the receptive field, preferred orientation, and width of orientation tuning were assessed for each group of cells with light stimuli projected manually onto a tangent screen in front of the cat. Subsequently, the receptive fields of the two eyes were aligned with a combination of prisms. For correlation measurements, light bar stimuli were generated with computer-controlled optical benches.

The cases included in this study were selected in the following way: Simultaneous recordings from four sites (quadruplets) were included when the respective orientation preferences matched pairwise but differed by 45° or more between these pairs. Recordings from three sites (triplets) were accepted when two sites had similar orientation preferences and when these differed from the third site by at least 45°. Finally, we also evaluated simultaneous recordings from two sites (doublets) when the respective orientation preferences differed by at least 45°. Cases were only included if cell groups with different orientation preference had sufficiently broad tuning to be coactivated with a single stimulus of intermediate orientation.

The cells were activated with either single moving light bars of various orientations or two bars moving simultaneously across the receptive fields (conflicting stimuli). In the latter case, orientations of the two light bars were adjusted to match orientation preferences of the cell groups (see Figs. 1–3). The length of the bar stimuli always corresponded to the extension of the receptive fields. The bars were moved orthogonal to their orientation, and their velocity was adjusted to elicit strong and sustained responses. For each trial, light bars were moved forward and backward across the receptive fields within 10 sec, and 10 trials to the same stimulus were run in succession. For both single as well as conflicting stimuli, several blocks of 10 trials were recorded in an alternating manner to control for the consistency of the effects observed.

To record multiunit activity, electrode signals were amplified, bandpass-filtered between 1 and 3 kHz, and then fed through a Schmitt trigger, the threshold of which was set to at least twice the noise level. The resulting pulse train was digitized with a time resolution of 1 ms. As described previously, peri-stimulus-time histograms as well as auto- and cross-correlation functions were computed for the recorded spike trains (18, 19). As a control, we computed shift-predictors of all correlograms (19). The oscillatory modulation of neuronal responses is reflected by a periodic auto-correlogram. A periodic cross-correlogram indicates the synchronization of two oscillatory responses. In both cases, the significance of the modulation can be estimated by fitting a damped sine wave (Gabor) function to the correlograms (19) (see Figs. 1–3). The strength of the modulation can then be quantified by using the amplitude, decay, and offset of the respective Gabor function. To obtain a measure that is independent of the absolute response strength, we determined the relative modulation amplitude by computing the ratio of the Gabor function amplitude over its offset (19). Cross-correlograms were considered as reflecting a significant synchronization when the Gabor function amplitude was significantly different from zero (at the 5% level), when the relative modulation amplitude exceeded a value of 0.1, and

when the Gabor function displayed at least three distinct peaks (19) (see Figs. 1 and 2). The same criteria were applied to auto-correlograms to quantify the periodic modulation of the responses. It should be noted that cross-correlograms with only one center peak, which reflect the interaction of two nonoscillatory responses, can be described by a Gabor function with a rapid decay. Therefore, we used the same quantification procedure in cases where we studied the influence of conflicting stimuli on the interaction of nonoscillating cells (see Fig. 3).

RESULTS

We studied the effect of single and conflicting stimuli on the cross-columnar interaction in 22 cases, corresponding to 4 quadruplets, 7 triplets, and 11 doublets. As described previously, cells in different orientation columns tend to synchronize their oscillatory responses when their receptive fields are overlapping and when they are activated with a single light bar (19). We have now tested whether conflicting stimuli eliminate the temporal correlation between cells of different orientation preference. The constellation of quadruplets and triplets allowed us to investigate whether within pairs of recordings with matching orientation preference the correlation is maintained when conflicting stimuli are applied. This result should be expected because in such pairs both cell groups respond to the same component of the stimulus configuration.

One of the quadruplets is illustrated in Fig. 1. We recorded from four sites with overlapping receptive fields and roughly alternating orientation preferences. Stimulation with single light bars of different orientations yielded a synchronization of oscillatory responses between all sites activated by the respective orientation. Thus, cells at sites 1 and 3 responded synchronously to a vertical (0°) light bar (Fig. 1A), cells at sites 2 and 4 responded synchronously to a light bar of 112° orientation (Fig. 1B), and cells at sites 2 and 3 responded synchronously to an intermediate stimulus (Fig. 1C). When a 0° and a 112° light bar were presented together to form a pair of conflicting stimuli, the cells at all four recording sites were simultaneously activated (Fig. 1D). In this case, response synchronization was still seen between sites 1 and 3 and between sites 2 and 4 and was as good as with single bar stimuli. However, no correlation occurred between the two synchronously active pairs. For instance, responses were uncorrelated between sites 2 and 3 (Fig. 1D), although at both sites the responses were even stronger than with a single stimulus of intermediate orientation. Thus, the conflicting stimuli segregate the cells in two distinct assemblies according to their orientation preferences. Cells responding to the same stimulus synchronize their responses, but they do not correlate with cells preferring the other stimulus. Importantly, this result cannot be explained by assuming selective anatomical connections between cells with similar orientation preferences (24) because with single stimuli a temporal correlation can be observed between cells preferring different orientations (Fig. 1C).

Similar observations were made in triplets and doublets, which represent fragments of the quadruplet constellation. A doublet of oscillatory responses is illustrated in Fig. 2. As in the previous example, synchronized responses were evoked with a single stimulus of orientation intermediate between those preferred by the two groups of cells (Fig. 2 A and B). Because the cells exhibited a rather broad orientation tuning, they could also be activated simultaneously with a light bar oriented optimally for one of the recording sites. In this case, the responses were also synchronized (Fig. 2 C and D). However, combined presentation of two light bars with orientations optimal for the two cell groups still eliminated the synchronization (Fig. 2 E and F). This case demonstrates
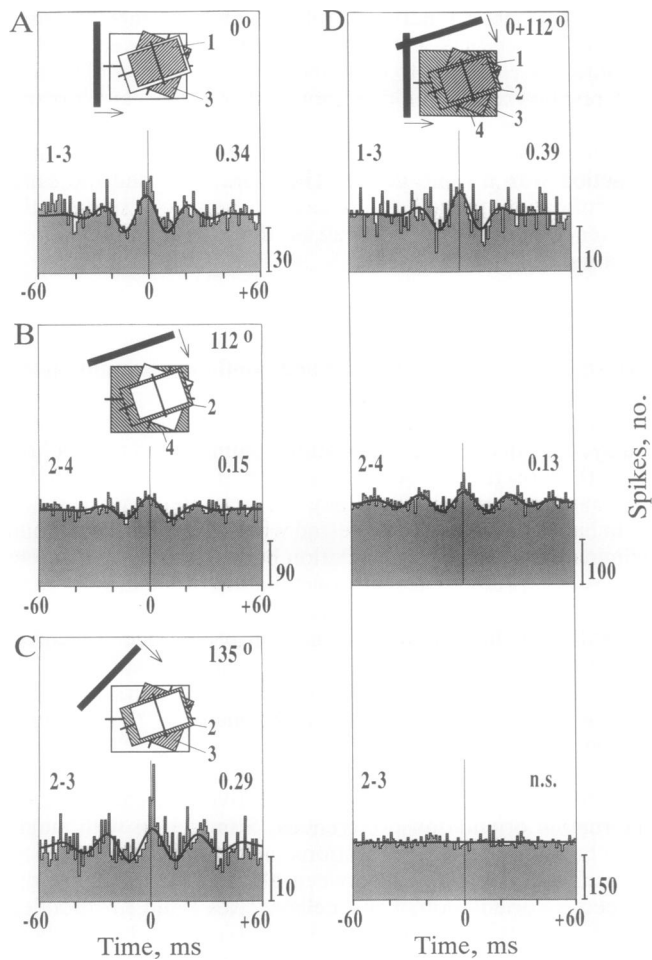
FIG. 1. Conflicting stimuli alter the cross-columnar interaction. We recorded simultaneously from four different sites separated by 400 μm. Cells at recording sites 1 and 3 preferred stimulus orientations near vertical. In contrast, cells at sites 2 and 4 had orientation preferences near horizontal (indicated by the thick line drawn across each receptive field in A–D). Stimulation with single light bars of 0° (A), 112° (B), and 135° (C) orientation yielded a synchronization between all responding sites (hatched receptive fields). (D) With combined presentation of both 0° and 112° light bars, responses were synchronized between sites 1 and 3 and between sites 2 and 4. However, no significant (n.s.) synchronization occurred between these pairs—e.g., between responses 2 and 3. The cross-correlograms between sites 1 and 2, sites 1 and 4, and sites 3 and 4 were also flat (data not shown). The graph superimposed on each of the correlograms represents the Gabor function used to assess strength of the modulation. The number at the upper right of each correlogram indicates relative modulation amplitude. Note that, for both pairs 1, 3 and 2, 4, relative modulation amplitude was undiminished with the conflicting stimuli. Scale bars indicate number of spikes. (This figure is adapted, with modification, from ref. 21.)

FIG. 2. Effect on conflicting stimuli on cell pairs with broadly overlapping orientation tuning. Recorded cells had a spatial separation of 400 μm and differed in their orientation preference by 45°. (A and B) Stimulation with a light bar of intermediate orientation elicited oscillatory responses at both sites that were synchronized to a significant degree. (C and D) Due to their broad orientation tuning, the cells at site 2 responded also to a horizontally oriented stimulus, which was optimal for the cells at site 1. With this stimulus, the two responses were also synchronized. (E and F) Adding the optimal stimulus for cell group 2 to the stimulus configuration eliminated the synchronization between the two responses (n.s.). Based on the evidence presented in Fig. 1, it can be assumed that cell groups 1 and 2 were still synchronized with other cells having the same respective orientation preference. Thus, these cell groups participated in two large assemblies desynchronized with respect to each other. In B, D, and F, the thick continuous line represents the Gabor function that was fitted to the correlogram. The number in the upper right corner indicates relative modulation amplitude. Vertical scale bars indicate numbers of spikes.

that even cells with largely overlapping orientation tuning can be desynchronized by conflicting stimuli. The observation that the cells were synchronized with one stimulus of optimal orientation but desynchronized by just adding the second optimal stimulus was confirmed in five additional cases.

In a number of cases (5 of 22), we recorded from nonoscillatory cells. In all of these cases, conflicting stimuli had a desynchronizing effect. Fig. 3 shows an example of a doublet in which the two recording sites differed in their orientation preference by 45°. Stimulation with a single light bar of intermediate orientation yielded a strong temporal correlation (Fig. 3 A and B). However, when the cells were activated by two stimuli adjusted to the respective preferred orientations, the responses were uncorrelated (Fig. 3 C and D). This
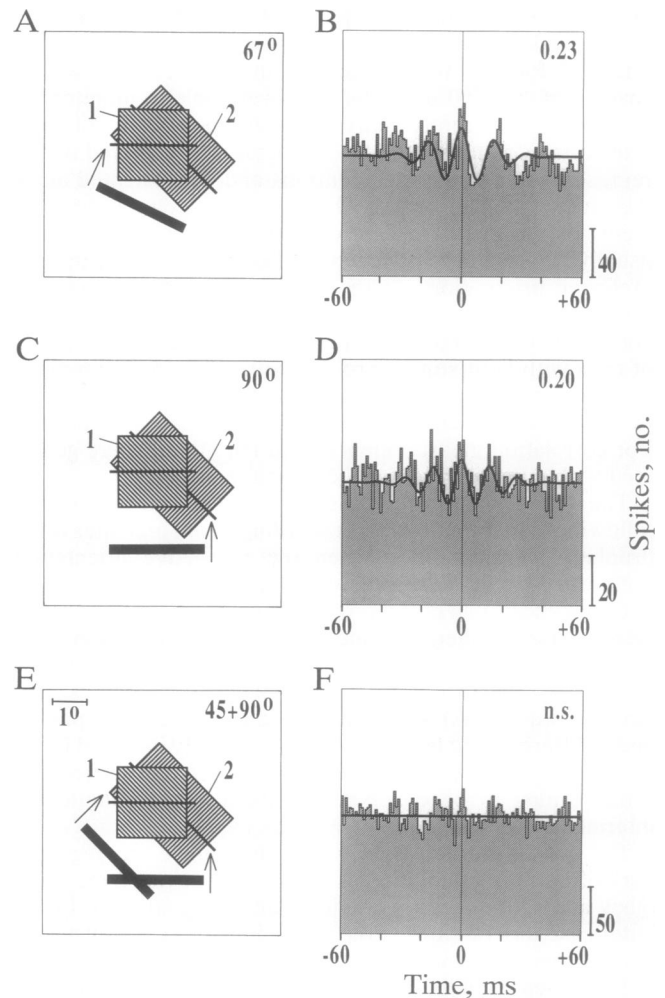
experiment shows that the stimulus dependence of temporal correlations, as demonstrated by the conflicting-stimulus paradigm, is not confined to cells with oscillatory responses.

The results illustrated in Figs. 1–3 were confirmed by analysis of the whole data sample. In 12 of the 22 cases studied, conflicting stimuli eliminated the temporal correlation between groups of cells with different orientation preferences that had been observed with single light bars. In 2 additional cases, the temporal correlation was reduced but not completely abolished by the conflicting stimuli. In these 14 cases, the average strength of correlation between cells with different orientation preferences, as obtained with single light bar stimuli, was 0.41 (SEM = ±0.08; average of the relative modulation amplitudes, see *Materials and Methods*).
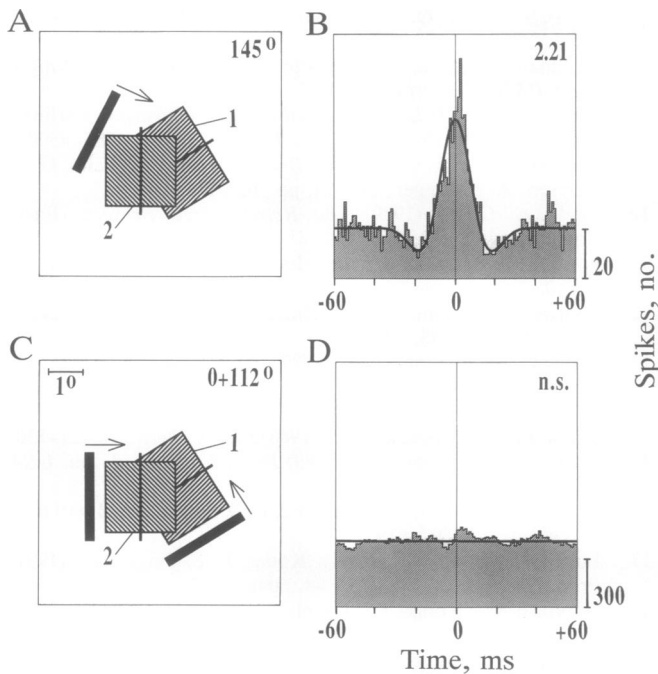
Neurobiology: Engel *et al.*

Proc. Natl. Acad. Sci. USA 88 (1991)   9139

FIG. 3. Effect of conflicting stimuli on temporal correlation of nonoscillatory responses. Recording sites were separated by 400 μm. (*A* and *B*) When coactivated with a light bar of intermediate orientation, the two cell groups fired in a strongly correlated manner. (*C* and *D*) Activation with two independent light bars eliminated the temporal correlation. In *B*, the number in the upper right corner represents relative modulation amplitude. The correlogram in *D* cannot significantly be described by a Gabor function (n.s.). Vertical scale bars indicate numbers of spikes.

Altogether, these 14 positive cases comprised 2 quadruplets, 4 triplets, and 8 doublets. In the quadruplets and triplets, there was no evidence that the synchronization within pairs of recordings with like orientation preference was reduced by the conflicting stimuli (5 out of 6). In the remaining 8 cases, the conflicting stimuli had no segregating effect on cells with different orientation preference. However, there was also no evidence for enhanced synchrony, which might have been expected as a consequence of activating all cells optimally with their preferred stimuli. We were unable to determine why conflicting stimuli did not always desynchronize the responses. However, the data indicate that the segregating effect of conflicting stimuli did not depend on the extent to which the cells differed in their preferred orientation.

In a previous study, we have reported that conflicting stimuli can reduce the rhythmicity of oscillatory responses (16). The results of this study are consistent with this previous observation. At 42 out of 59 recording sites investigated, oscillatory responses were evoked with single light bars. At 17 of these sites, the conflicting stimuli reduced the modulation of the auto-correlograms and at 8 sites they abolished the oscillation. At the remaining 17 sites, the periodicity of the auto-correlograms was unattenuated.

DISCUSSION

The results of this study support the hypothesis that temporal coherence of neuronal firing patterns can be used for the segmentation of visual scenes. This was demonstrated by challenging cell assemblies with a simple "visual scene" consisting of two light bars moving in different directions. In psychophysical terms, a segmentation of this scene into two independent figures is expected because the two superimposed light bars were relatively short and, in addition, the combined figure fulfilled the transparency condition (25)—i.e., the moving intersection was brighter than the remaining

part of the light bars. Thus, the conflicting stimuli would not appear as a single moving object, even if viewed through a relatively small aperture.

The observation that conflicting stimuli can desynchronize cells that respond synchronously to a single coherent object has several implications. The data demonstrate that assemblies of synchronously oscillating cell groups are, indeed, formed in a stimulus-dependent manner. When stimulated with a single object, all responding cells join the same assembly. However, the same cells are segregated into different assemblies by desynchronization when the stimulus has several components with different orientations and directions of motion. Thus, cells can switch between assemblies by changing the phase-relationship to neurons in neighboring columns. These results demonstrate the parsimony of temporal coding: Formation of new representations does not require new units. Interestingly, we obtained the same results for cells the responses of which did not yield periodically modulated auto-correlograms and, hence, did not oscillate at a fixed frequency. This result shows that dynamic coupling and response synchronization are not confined to response epochs characterized by regular oscillations. We consider this an important issue because it implies that temporal coding can also serve for the formation of functional cell assemblies in cortical areas where regular oscillatory responses are less prominent than in cat visual cortex. Altogether, the present results confirm the predictions about assembly formation by synchronous neuronal discharges that have been formulated by von der Malsburg (9, 11). Moreover, the results are fully compatible with the theory of neuronal group selection as proposed by Edelman (6, 10).

Evidence for assembly formation by temporal coding has already been obtained in a recent study (18). There, we could demonstrate that cells in area 17 with nonoverlapping receptive fields fire synchronously when they are activated by a single long light bar and, thus, respond to a single coherent object, but they are segregated into two assemblies when two independent objects are used (18). In the latter case, the two assemblies formed in area 17 are spatially disjunct and separated by an inactive region of cortex. Segregation of these assemblies can be achieved by amplitude coding and does not require a temporal-coding mechanism. Therefore, segmentation of spatially superimposed stimuli provides a crucial test of our working hypothesis. These stimuli are represented by spatially overlapping assemblies in the same cortical region that cannot be segregated by amplitude coding (9, 11). The present results demonstrate that such superimposed assemblies can, indeed, be distinguished by the temporal coherence of neuronal firing patterns and, thus, provide direct evidence for the existence of a temporal-coding mechanism capable of solving the superposition problem (9, 11).

Recently, we have obtained direct experimental evidence for the hypothesis that synchronization of neuronal responses is achieved by reciprocal connections at the cortical level (20). The results presented here suggest that quite different states of functional connectivity must be realized within the network of horizontal connections. The cortical microcircuitry must permit the synchronization of neurons with different feature-selectivity when highly coherent stimuli are used to activate the network. On the other hand, the same circuitry must permit desynchronization when incoherent stimuli are applied. Recent simulation studies by Schillen and König (26, 27) demonstrate how both requirements can be met in a network of reciprocally coupled oscillators. These studies show that the segregating effect of conflicting stimuli can be modeled without fast changes of synaptic efficacy, which had been considered as a prerequisite for scene segmentation in previous studies (11, 12).

The observation that the functional coupling of cortical cells is dynamic and subject to rapid changes has particular

implications for the interpretation of cross-correlograms (19). Our data demonstrate that no direct inferences can be made from correlation data on the underlying microcircuitry and that, in particular, the absence of correlogram peaks does not prove a lack of anatomical connectivity. In a recent cross-correlation study of cat visual cortex, Ts'o and coworkers (24) reported preferential correlation between cells with similar orientation preference and concluded that only cells with like receptive-field properties are linked via horizontal connections. At least with respect to short-range interactions, our results (18, 19) are not consistent with this interpretation because with single light bar stimuli we readily find temporal correlations between cells in adjacent columns with dissimilar orientation preferences. The apparent discrepancy may, in fact, be explained by the stimulus dependence of temporal correlations as observed in this study. Ts'o *et al.* (24) always stimulated the cells with their preferred orientation. Therefore, when recording from cells with overlapping receptive fields and different orientation preferences, these authors used a configuration similar to our conflicting stimuli. Thus, they may have decreased the probability of observing correlated firing of cells with dissimilar orientation preferences.

In conclusion, the present results provide further evidence that cortical representations may be created in a highly dynamic manner by transient synchronization of neuronal responses (17–23). As demonstrated here, superposition of two stimuli in the visual field induces the formation of two assemblies of synchronously firing cells that are desynchronized with respect to each other. This result suggests the existence of a temporal-coding mechanism for scene segmentation, which involves synchronization of neurons responding to the same object, but desynchronization of cells coding for different objects of the scene.

1. Köhler, W. (1930) *Gestalt Psychology* (Bell & Sons, London).
2. Koffka, K. (1935) *Principles of Gestalt Psychology* (Routledge & Kegan Paul, London).
3. Treisman, A. (1986) in *Handbook of Perception and Human Performance*, eds. Boff, K., Kaufman, L. & Thomas, I. (Wiley, New York), pp. 35.1–35.70.
4. Barlow, H. B. (1972) *Perception* 1, 371–394.
5. Hebb, D. O. (1949) *The Organization of Behaviour* (Wiley, New York).
6. Edelman, G. M. & Mountcastle, V. B. (1978) *The Mindful Brain* (MIT, Cambridge, MA).
7. Abeles, M. (1982) *Local Cortical Circuits* (Springer, Berlin).
8. Crick, F. (1984) *Proc. Natl. Acad. Sci. USA* 81, 4586–4590.
9. von der Malsburg, C. (1986) in *Brain Theory*, eds. Palm, G. & Aertsen, A. (Springer, Berlin), pp. 161–176.
10. Edelman, G. M. (1989) *The Remembered Present* (Basic Books, New York).
11. von der Malsburg, C. & Schneider, W. (1986) *Biol. Cybern.* 54, 29–40.
12. Sporns, O., Tononi, G. & Edelman, G. M. (1991) *Proc. Natl. Acad. Sci. USA* 88, 129–133.
13. Gray, C. M. & Singer, W. (1987) *Soc. Neurosci. Abstr.* 13, 404.3.
14. Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M. & Reitboeck, H. J. (1988) *Biol. Cybern.* 60, 121–130.
15. Gray, C. M. & Singer, W. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1698–1702.
16. Gray, C. M., Engel, A. K., König, P. & Singer, W. (1990) *Eur. J. Neurosci.* 2, 607–619.
17. Engel, A. K., Kreiter, A. K., König, P. & Singer, W. (1991) *Proc. Natl. Acad. Sci. USA* 88, 6048–6052.
18. Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989) *Nature (London)* 338, 334–337.
19. Engel, A. K., König, P., Gray, C. M. & Singer, W. (1990) *Eur. J. Neurosci.* 2, 588–606.
20. Engel, A. K., König, P., Kreiter, A. K. & Singer, W. (1991) *Science* 252, 1177–1179.
21. Engel, A. K., König, P., Gray, C. M. & Singer, W. (1990) in *Parallel Processing in Neural Systems and Computers*, eds. Eckmiller, R., Hartmann, G. & Hauske, G. (Elsevier, Amsterdam), pp. 105–108.
22. Singer, W., Gray, C. M., Engel, A. K., König, P., Artola, A. & Bröcher, S. (1990) *Cold Spring Harbor Symp. Quant. Biol.* 55, 939–952.
23. Engel, A. K., König, P., Kreiter, A. K., Gray, C. M. & Singer, W. (1991) in *Nonlinear Dynamics and Neural Networks*, eds. Schuster, H. G. & Singer, W. (VCH, Weinheim, F.R.G.), in press.
24. Ts'o, D. Y., Gilbert, C. D. & Wiesel, T. N. (1986) *J. Neurosci.* 6, 1160–1170.
25. Stoner, G. R., Albright, T. D. & Ramachandran, V. S. (1990) *Nature (London)* 344, 153–155.
26. König, P. & Schillen, T. B. (1990) in *Parallel Processing in Neural Systems and Computers*, eds. Eckmiller, R., Hartmann, G. & Hauske, G. (Elsevier, Amsterdam), pp. 117–120.
27. Schillen, T. B. & König, P. (1991) *Neural Comput.* 3, 167–177.