Invariant representations of visual patterns in a temporal population code

Reto Wyss*, Peter König, and Paul F. M. J. Verschure

Institute of Neuroinformatics, University of Zürich and Swiss Federal Institute of Technology, CH-8057 Zürich, Switzerland

Mammalian visual systems are characterized by their ability to recognize stimuli invariant to various transformations. Here, we investigate the hypothesis that this ability is achieved by the temporal encoding of visual stimuli. By using a model of a cortical network, we show that this encoding is invariant to several transformations and robust with respect to stimulus variability. Furthermore, we show that the proposed model provides a rapid encoding, in accordance with recent physiological results. Taking into account properties of primary visual cortex, the application of the encoding scheme to an enhanced network demonstrates favorable scaling and high performance in a task humans excel at.

Mammals demonstrate highly evolved visual object recognition skills, tolerating considerable changes in images caused by, for instance, different viewing angles and deformations. Elucidating the mechanisms of such invariant pattern recognition is an active field of research in neuroscience (1–5). However, very little is known about the underlying algorithms and mechanisms. A number of models have been proposed which aim to reproduce capabilities of the biological visual system, such as invariance to shifts in position, rotation, and scaling (6–8). Most of these models are based on the "Neocognitron" (9), a hierarchical multilayer network of spatial feature detectors. As a result of a gradual increase of receptive field sizes, translation-invariant representations emerge in the form of activity patterns at the highest level. These models do not consider time as a coding dimension for neural representations. Recently, however, the importance of the temporal dynamics of neuronal activity in representing visual stimuli has gained increased attention (10, 11). Hence, it seems timely to consider the role of temporal coding in the context of tasks like invariant object recognition. In recent years, several modeling studies have addressed properties of temporal codes (12–14). For instance, Buonomano and Merzenich (15) proposed a model for position-invariant pattern recognition which uses temporal coding. In this model, feed-forward inhibition modulates the spike-timing such that stimuli are represented by the response latencies of the neurons in the network. This architecture naturally leads to translation invariant representations. This model assigns a critical role to inhibitory interactions in the feed-forward path of the visual system (retina-LGN-V1), whereas anatomical studies suggest that these connections are predominantly excitatory (16). Furthermore, the majority of inputs to cortical neurons are excitatory and of cortical origin (17). Indeed, a recent theoretical study has shown that lateral excitatory coupling has pronounced effects on the global network dynamics (18). In particular, the combination of intracortical connectivity and dendritic processing allowed context-dependent representations of different stimuli to be expressed in the temporal dynamics of the network. Here, we build on these previous proposals and concepts (15) and investigate the formation of invariant representations by the dynamics of activity of neuronal populations.

In this study, we investigate a model of primary visual cortex consisting of a map of integrate-and-fire neurons with lateral excitatory interactions. A central feature of this model is the monotonic relationship between transmission delays in this lateral coupling and the distance between pre- and postsynaptic neurons. We hypothesize that this network property induces dynamics of neuronal activity (Fig. 1a) that are specific to the geometry of a stimulus and invariant with respect to several transformations. The advantage of such a representation is that it emerges naturally without the need of training the network repeatedly for different stimulus positions or orientations. To investigate the validity of this hypothesis, we determined the amount of information contained in the temporal population responses of this network for different parameters and stimulus sets. Furthermore, we investigate the speed of encoding and its robustness to synaptic noise. If the proposed encoding scheme reflects relevant properties of the mammalian visual system, it should be well suited for those tasks at which mammals excel. Therefore, we tested it in two demanding tasks taken from the domain of pattern recognition. The results suggest that invariant pattern recognition can be achieved by using temporal coding at the population level.

Methods

The basic network we investigated consisted of a two-dimensional array of 40×40 conductance-based leaky integrate-and-fire neurons, which included a spike-triggered potassium conductance-yielding frequency adaptation (see *Appendix*). Under constant excitation and after adaptation, these neurons spiked regularly. Each neuron connected to a circular neighborhood of fixed size, such that neurons with Euclidean distance ≤ 9 cells were connected. The synapses were of equal strength (ν) and were modeled as instantaneous excitatory conductances, whereas transmission delays were related to the Euclidean distance between the positions of the pre- and postsynaptic neurons with a proportionality factor of $\tau = 1$ ms per cell. Stimuli were presented continuously to the network and first passed through an edge-detection stage (LGN), and the resulting contours were projected topographically onto the array of neurons (V1) by using a tonic excitatory input conductance (Fig. 1a). After frequency adaptation, the stimulated neurons spiked at a frequency of ≈ 42 Hz.

For some experiments, a more sophisticated network was investigated, incorporating orientation and spatial frequency selectivity. The input was preprocessed at three different resolutions (40×40, 20×20, and 10×10 pixels). Subsequently, orientation selective filters of four different orientations (0°, 45°, 90°, and 135°) were applied, and the resulting preprocessed stimuli were fed into three networks of corresponding size, i.e., $40 \times 40 \times 4$, $20 \times 20 \times 4$, and $10 \times 10 \times 4$ neurons. Thus, the different networks consisted of four different layers, each of which was responsible for the processing of visual information in one of the four orientations. Within each layer, lateral connectivity is long-range and asymmetric. Each neuron projects within a segment of a circle along the preferred orientation with an apex angle of 30° and radii of 13, 10, and 6 cells for the three resolutions, respectively. Between layers, connectivity is short-range and symmetric. Neurons project within a circle with a

---

*To whom correspondence should be addressed at: Institute of Neuroinformatics, University/ETH Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland. E-mail: rwyss@ini.unizh.ch.
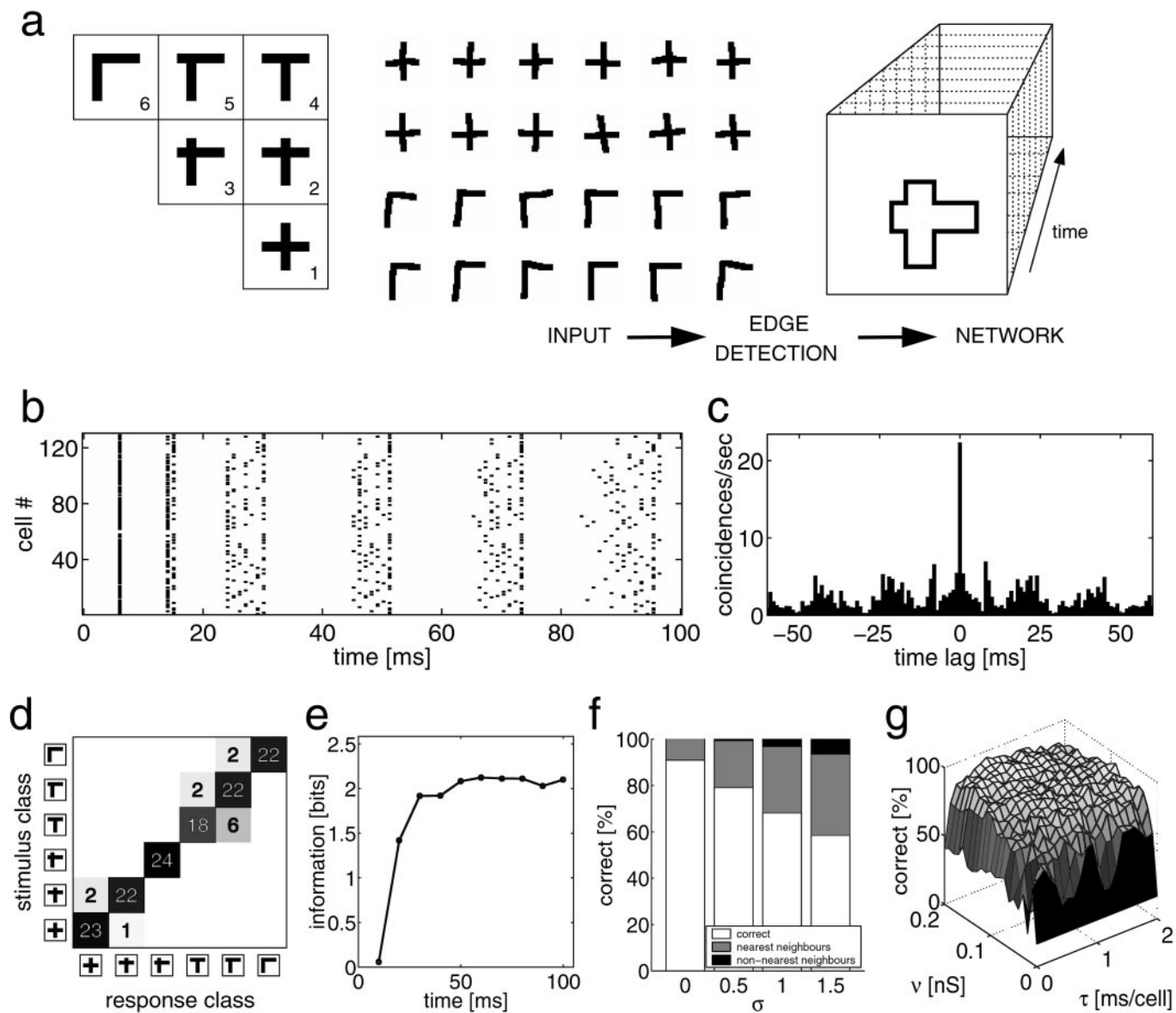
Fig. 1. (a) Schematic of the encoding paradigm. The stimulus set contains six stimulus classes, each of which consists of a horizontal and a vertical bar of equal length, but intersecting at different locations for the individual classes. From each stimulus class, 24 hand-drawn samples are presented to the network (12 samples for classes 1 and 6 are shown in the middle). These "solid" input patterns pass through an edge-detection stage and the resulting contour is projected topographically onto a map of cortical neurons. Because of the lateral intracortical interactions, the stimulus becomes encoded in the network's activity trace. (b) Spike raster of 130 neurons of the network while presenting a stimulus sample of class 1. (c) The population cross-correlation function for the neurons shown in b. (d) After response clustering, the entries of the hit matrix represent the number of times a stimulus class is assigned to a response class ($\nu = 0.13$ nS). (e) Speed of encoding. Average information encoded in the network's activity traces as a function of time with $\nu = 0.13$ nS. (f) Influence of synaptic noise on encoding. Percentage of stimuli being classified either correctly, to a nearest neighbor, or to a nonnearest neighbor for different noise levels σ, with $\nu = 0.13$ nS. (g) Correctly classified stimuli as a function of the inverse of transmission speed and synaptic strength of the lateral coupling.

radius of two cells for all resolutions. There are no connections between networks of different resolutions.

Results

In our first experiment, we investigated the concept and tested the basic network's performance in the invariant encoding of hand-drawn stimuli. The arrangement of the six stimulus classes (Fig. 1a *Left*) reflects an intuitive notion of topology; i.e., class 1 is visually more similar to class 2 or 3 than to class 5 or 6. For each stimulus class, 24 samples were presented to the network. The spike raster plot for a sample of class 1 illustrates the network activity (Fig. 1b). An initial synchronous phase is followed by a dispersion of activity. Such raster plots give a good

description of network activity. However, because activities of such large numbers of neurons are rarely recorded simultaneously in visual cortex, they do not allow a direct comparison to physiological data. Therefore, population-averaged histograms are not used to describe physiological data but rather describe time-averaged histograms (19). For comparison, we compute the multiunit cross-correlation function (Fig. 1c). The central peak signifies the population bursts, and the satellite peaks are a sign of the temporally structured activity. Overall, this cross-correlation function resembles those acquired in the visual cortex of mammals (20, 21).

For a quantitative analysis, the responses of the network are clustered into six classes by using the temporal correlation of the

population response as a similarity measure. For each stimulus, the correlation is computed with all other responses over an interval of 100 ms. It is assigned to the response class which maximizes the average of this measure (see *Appendix*). The resulting hit matrix (Fig. 1*d*) reveals that 91% of the stimuli are classified correctly (diagonal entries). The misclassifications (9%, nonzero off-diagonal entries), however, do show some regularity. They only occur among visually similar stimuli. Indeed, more than half of the misclassifications result from a confusion of classes 4 and 5, which only differ in the precise position of the vertical bar. To investigate the relationship between the strength of the lateral coupling, ν, and the encoding performed by the network, we vary this parameter in the range of $0 < \nu < 0.25$ nS. For each of these conditions, we calculate the mutual information from the hit matrix (see *Appendix*). We find that for $\nu < 0.01$ nS, the activity traces contain no information about the input stimuli. For $0.01 < \nu < 0.25$ nS, however, mutual information reaches 1.77 ± 0.21 bits (mean $\pm$ SD, of $\log_2(6) = 2.6$ bits possible), and all hit-matrices qualitatively resemble the one shown in Fig. 1*d*. In summary, the network successfully and reliably encodes the six stimulus classes in a large range of lateral coupling strengths.

In the next step, we investigated the classification of rotated stimuli. A stimulus set was constructed by taking one sample from each stimulus class (Fig. 1*a Left*) and generating the complete sets by rotating each sample by 23 evenly spaced angles between 0° and 360°. As in the previous experiment, no information about the input is conveyed by the network's activity traces for $\nu < 0.01$ nS. For $0.01 < \nu < 0.25$ nS, mutual information reaches 1.54 ± 0.31 bits. Maximal information of 2.2 bits is attained for $\nu = 0.075$ nS, where 92% of the stimuli are classified correctly, 7% are confused with nearest neighbors, and only 1% with nonnearest neighbors. Hence, the population-activity trace reliably encodes the six stimulus classes invariant to rotation.

To analyze the speed of encoding, we determined the amount of information encoded in the network's activity trace at different times after stimulus onset. By varying the length of the interval used to compute the correlation between different responses between 2 and 100 ms, we observed that 66% of the information is available after 20 ms (Fig. 1*e*). This property of our encoding scheme is compatible with the impressive speed of processing found in the mammalian visual system (22).

It has been argued that a potential problem of encoding information in the timing of action potentials is the highly unreliable transmission of signals across synapses (23). Thus, we investigated the robustness of the proposed encoding scheme with respect to synaptic noise in the lateral coupling. Synaptic noise is modeled by perturbing the individual synaptic conductances dynamically: each conductance is multiplied by a random factor f, drawn from a normal distribution with a mean of one and variance σ^2, i.e., $f \in N(1, \sigma^2)$. To prevent negative conductances, the normal distribution was clipped at zero. The system's performance in encoding the input stimuli decreases linearly with increasing noise (Fig. 1*f*). The number of correctly classified stimuli decreases by not more than 25% for $\sigma = 1$, which corresponds to a signal-to-noise ratio of 1. Furthermore, misclassifications are, to a substantial degree (90% for $\sigma = 1$), caused by the confusion of nearest neighbors (Fig. 1*f*). Thus, the information encoded at the population level is robust with respect to synaptic noise.

In the following experiment, we investigated the network's encoding performance for different speeds of tangential interactions. For this purpose, the temporal resolution of the simulation was increased to $\Delta t = 0.1$ ms, allowing for the delay slopes τ to take values from 0 to 2 ms per cell in 0.1 ms per cell steps. For the clustering, we binned the network responses with a temporal resolution of 1 ms. In addition, the synaptic strength ν

was varied simultaneously in the range from $0 < \nu < 0.2$ nS to detect combined effects of τ and ν on encoding performance. As it is shown in Fig. 1*g*, the performance does not significantly depend on τ, as long as the latter is above ≈ 0.3 ms per cell. Furthermore, apart from weak synaptic strengths (i.e., $\nu < 0.04$ nS), we do not observe any systematic relationship between ν and τ. Thus, the encoding scheme proposed here is remarkably invariant to the choice of two defining parameters of the lateral coupling, which constitutes the central component of the proposed network.

As a further control, we studied the encoding of random stimuli. We presented random dot patterns which had the same pixel density as the bar-stimuli investigated before. The random stimuli are compared with the clusters of the bar stimuli from the first experiment. We find that the distance of the random stimuli from these clusters, in units of the width of the distribution of the bar-stimuli, is large (29.8 ± 2.4, mean $\pm$ SD, $n = 500.000$). Hence, with a reasonable choice of a classification threshold, e.g., a distance of at most 3 SDs resulting in $\approx 0.5\%$ false rejections, the number of false positives is virtually zero. Thus, it is unlikely that in the proposed encoding scheme random stimuli are confused with structured stimuli.

So far, the network we have investigated only incorporates excitatory coupling between neurons, whereas inhibitory interactions, which are ubiquitous throughout the cortex, have not been considered. Hence, as a control, we add lateral inhibitory connections to our network model. Each cortical neuron inhibits its direct neighbors in the map with a synaptic conductance of 10 nS and a delay of 2 ms. We find that for all stimulus sets presented above, there is no significant difference in the system's performance; i.e., base set, 1.78 ± 0.19 bits; rotation set, 1.5 ± 0.27 bits. These results indicate that the proposed encoding scheme is not affected by incorporating inhibitory interactions in the neural network.

The effect of the presence of multiple stimuli was addressed as a further control. We investigated a four times wider network. The activity induced by a target stimulus was evaluated by pooling neuronal activity within a large region (readout region), which, however, does not encompass the whole network. Two randomly chosen stimuli serve as distractors on either side (Fig. 2*a*). When the readout region of the stimulus includes the distractors, a small decrease in performance is observed (Fig. 2*c*). If, by virtue of the tangential connections in the network, the distractors interact with the representation of the target stimulus, the encoded information is more reduced (Fig. 2*b*). A combination of both effects leads to a more severe interference (Fig. 2*d*). Note, however, that this situation creates a continuous stimulus pattern that poses severe problems for any recognition system. Furthermore, given the scaling properties of the network (see below), it is possible to encode pairs of neighboring stimuli as one compound stimulus. Hence, the presence of multiple stimuli does not pose a fundamental problem to the encoding scheme investigated here.

To test the basic encoding principles in more realistic and demanding tasks, we studied an enhanced version of the network. In this case, the network units are endowed with orientation- and spatial frequency-selective responses similar to neurons in primary visual cortex. The lateral connections depend on the neural response preferences (see *Methods*). This network was tested on handwritten digits from 250 writers contained in the modified database of the National Institute of Standards and Technology (MNIST; Fig. 3*a*; ref. 24), which is a standard benchmark in the domain of character recognition. To facilitate comparison with other systems tested on this database, we used the original partitioning into a training set and a test set, with 5,400 samples and 950 samples per digit, respectively. Our model is capable of classifying 94.8% of the stimuli correctly (Fig. 3*b*).

Fig. 2. Multiple stimuli. (*a–d*) Target stimulus (black), the distracting stimuli (gray), and the readout region (dashed rectangle) in four characteristic situations. (*a*) Control condition: stimuli do not interact, and readout only captures the target stimulus. (*b*) Representations of the stimuli interact through the lateral coupling within the network. (*c*) Readout captures activity from the distracting stimuli. (*d*) Conditions *b* and *c* are combined. (*e*) The performance of the system in encoding and classifying stimuli in the four situations, respectively. (Bars = ±SD over 20 trials.)

In a final control, we studied scaling of the model to very large numbers of synthetic stimulus classes. Stimulus classes were defined by randomly choosing five points in the input field and connecting each pair of points by a straight line with a probability of 0.3. Individual samples of each class were constructed by jittering the position of the vertices and the thickness of lines to varying degrees (Fig. 4 *a–c*). We investigated classification by the network for up to 800 stimulus classes (Fig. 4*d*). For an increasing number of stimulus classes, the encoded information stayed close to the theoretical limit. Even in the case of high variability and the maximal number of stimulus classes, 80% of the information is represented in the temporal structure of neuronal activity. Thus, the proposed encoding scheme scales to problems of interesting complexity.

Discussion

We have shown that, in a model of a cortical network, the interaction of network and stimulus topology induced stimulus-specific but transformation-invariant temporal dynamics. Thus, stimuli are represented by a temporal population code. This representation is position- and rotation-invariant and reasonably robust to stimulus variability. The stimulus encoding preserves the intuitive notion of visual similarity. Furthermore, it is robust

Fig. 3. The MNIST database contains handwritten digits from 250 different writers. (*a*) Five samples for each digit, as they were presented to the network. The stimulus set is split into a training set (5,400 samples per class) and a test set (950 samples per class). (*b*) The hit matrix. Please note that some stimuli, such as 0 or 1, reach a higher number of hits, whereas others, such as 4 or 9, are likely to be confused because of the writing habit of some writers.

Fig. 4. Performance vs. number of stimulus classes. (*a–c*) Samples of the random stimuli generated synthetically (vertical, different classes; horizontal, different samples). Each class is defined by five randomly chosen points; i.e., $P_i(x_i, y_i)$ for $i = 1 \ldots 5$ with $x_i, y_i \in U(0, 1)$ (uniformly distributed, the plane $[0,1] \times [0,1]$ corresponds to the input to the network). Each pair of points was connected with a probability of 0.3 by a bar of thickness $d = 0.12$. The different samples were created by varying the positions of the points and the thickness of the bars randomly: $P_i'(x_i + \delta_x, y_i + \delta_y)$ for $i = 1 \ldots 5$ with $\delta_{x,y} \in N(0, \sigma_{pos}^2)$ and $d' = d + \delta_d$ with $\delta_d \in N(0, \sigma_{thick}^2)$. a: $\sigma_{pos}^2 = 0.03$, $\sigma_{thick}^2 = 0.018$; b: $\sigma_{pos}^2 = 0.04$, $\sigma_{thick}^2 = 0.0216$; c: $\sigma_{pos}^2 = 0.05$, $\sigma_{thick}^2 = 0.0252$. (*d*) Information encoded by the network as a function of the number of stimulus classes. Optimal performance [information = $\log_2$(no. of classes)] is indicated by the thin dashed line.

with respect to synaptic noise. These properties apply in a range of the strength and speed of the lateral coupling, each spanning more than one order of magnitude. As these results demonstrate only basic properties of the encoding scheme and leave many questions open, we investigated an extension of the network which incorporates orientation- and spatial frequency-selective neurons. We show that this leads to invariant representations and reliable classification of large stimulus sets of increased complexity.

In this study, we tried to understand how key functional properties of the visual system can be accounted for by a temporal population code. Compared with area 17, however, our model contains several assumptions. It is composed of conductance-based integrate-and-fire neurons, which are coupled through excitatory connections. The key assumption of our model is that the transduction latencies of these connections depend systematically on the distance between pre- and postsynaptic neurons. Indeed, indirect evidence for such a relationship has been found (25–26). These transduction delays could result from dendritic delays caused by the distance a signal has to travel between the synapse and the soma of the postsynaptic neuron (27). Thus, our study predicts that a positive correlation exists between the separation of receptive fields and the distance between the dendritic location of the respective synapses and the soma.

The speed of information flow found in biological visual systems poses severe constraints upon computational models of pattern recognition. Experimental studies (28) have shown that the analysis and classification of complex visual patterns can be carried out by humans in not more than 200 ms. Considering that a minimum of 10 areas are involved to reach the relevant processing areas from the retina, little time is left for intra-areal processing. This view is supported by studies in macaque monkeys (29–30), which show that single visual areas can process significant amounts of information in just 20–30 ms. This result is compatible with our finding that 66% of information about a stimulus becomes encoded within 20 ms. The gradual increase of encoded information over the time period we observed indicates that subsequent processing stages can be engaged before all of the information about a stimulus has been encoded. Indeed, experimental data on signal timing in the macaque visual system show that the distribution of onset latencies of different visual areas overlap substantially (31). This property suggests that the dominant mode of processing in the visual system is concurrent rather than strictly feed-forward.

A number of studies have proposed a role for temporal dynamics in the formation of visual representations. One of the first proposals (32) suggested that local features of visual stimuli are coded in the temporal patterns of spikes of single neurons. Based on principal component analysis of the response patterns of single neurons, a significant amount of information carried by the first ($\cong$firing rate) as well as the higher components ($\cong$temporal pattern) was found (29). An important difference with the present study is that we consider temporal coding at the population level, rather than at the level of a single neuron. Recently, Buonomano and Merzenich (15) presented a model on the generation of temporal population codes that contained orientation-selective feature detectors and strong feed-forward inhibition. It was shown that the latencies between stimulus onset and the first spike of the neurons in the network constitute a representation that is invariant to the position of presented stimuli. In contrast, by relying on lateral interactions, the present model accumulates information over time, which results in reliable encoding.

In a widely used benchmark on the recognition of hand written digits, nearly 95% correct classification was achieved. This falls short of the performance of the very best specialized character recognition systems (24). However, it compares well to an approach in which a k-nearest neighbor clustering is applied directly to the spatial representation of the stimuli centered at a common position (24). Thus, the encoding process performed by the network discards part of the information for the generation of invariant representations, while preserving the relevant information for classification of the stimuli.

As a particular example of a temporal population code, synchronized neuronal activity has been intensively studied (10, 11, 33). A large number of experiments have reported synchronized activity in a variety of species and cortical structures (20, 34). In particular, it has been argued that the synchronization of neural activity provides a substrate for the binding and segmentation of visual patterns (35, 36). Furthermore, it has been shown that the synchronization of neuronal activity is mediated by intracortical connections without changing receptive-field properties of the postsynaptic neurons (37, 38). Thus, synchronous activity does not contain information about stimulus features as such and can be seen as a binary signal about global stimulus properties. Our model shows that synchronous activity at a broad time scale, combined with dispersion on a fine time scale, provides a high-dimensional signal in a temporal population code. This signal contains detailed information about a stimulus, including local features and their global relationship. Indeed, experimental evidence is available which shows that cortical neurons can produce feature-specific phase lags in their activity (39). Theoretical studies argue that synchronous and dispersed activity are different functional modes of the same basic network structure (40). The transition between these modes depends on the transmission delays in the lateral coupling. Moreover, Maass (41) has shown that neurons that encode information in their spike-timing have interesting computational properties. Therefore, we believe that the temporal population code we present here provides a promising approach toward both invariant pattern recognition and the understanding of the encoding of information by the nervous system.

Appendix

The Neuron. The time course of a leaky integrate-and-fire neuron's membrane voltage $V(t)$ is described by the differential equation:

$$C_m \frac{dV}{dt} = -(I_{exc}(t) + I_{inh}(t) + I_K(t) + I_{leak}(t)), \quad \textbf{[1]}$$

where C_m is the membrane capacitance ($C_m = 0.2$ nF), and I represents the transmembrane current, i.e., excitatory input (I_{exc}), inhibitory input (I_{inh}), spike-triggered potassium current (I_K), and leak current (I_{leak}). These currents are computed by multiplying a conductance g with the driving force: $I(t) = g(t)(V(t) - V^{rev})$, where V^{rev} is the reversal potential of the conductance ($V_{exc}^{rev} = 60$ mV, $V_{inh}^{rev} = -70$ mV, $V_K^{rev} = -90$ mV, $V_{leak}^{rev} = -70$ mV). The neuron's activity at time t, $A(t)$, is given by $A(t) = H(V(t) - \theta)$, where H is the Heaviside function, and θ is the firing threshold ($\theta = -55$ mV). Each time a spike is emitted, the neuron's potential is reset to $V_{rest} = V_{leak}^{rev}$. The constant leak conductance g_{leak} is 20 nS. The time course of the potassium conductance is given by $\tau_K dg_K/dt = -(g_K(t) - g_K^{peak} A(t))$, where $A(t) \in \{0,1\}$, with a time constant τ_K and a peak conductance g_K^{peak} ($\tau_K = 40$ ms, $g_K^{peak} = 200$ nS). The synaptic interactions are "instantaneous," such that the total synaptic conductance at time t is the linear sum over all active conductances derived from the individual synapses at time t. In the discrete-time simulations, the equations above are integrated with Euler's method and a temporal resolution Δt of 1 ms.

Clustering Algorithm and Mutual Information. The algorithm for clustering the responses of the network was adapted from Victor and Purpura (42). The network's responses to stimuli from C

stimulus classes $S_1, S_2, \ldots, S_C$ are assigned to C response classes $R_1, R_2, \ldots, R_C$, yielding a $C \times C$ hit matrix $N(S_\alpha, R_\beta)$, whose entries denote the number of times that a stimulus from class S_α elicits a response in class R_β. Initially, the matrix $N(S_\alpha, R_\beta)$ is set to zero. For each response, $r \in S_\alpha$, we calculate the average temporal correlation of r to the responses $r' \neq r$ elicited by stimuli of class S_γ:

$$\bar{\rho}(r, S_\gamma) = Z^{-1}\left(\left\langle Z\big(\rho(r, r')\big)\right\rangle_{r' \text{ elicited by } S_\gamma}\right), \qquad [2]$$

where $\rho(r, r')$ is the temporal correlation between r and r', $\langle \cdot \rangle$ denotes the average. Z is the Fisher Z-transform given by $Z(\rho) = \frac{1}{2} \ln((1 + \rho)/(1 - \rho))$, which transforms a distribution of correlation coefficients ρ into an approximately normal distribution of coefficients, $Z(\rho)$. The average correlation is also computed for the stimulus class S_α which elicited r, but because $r \neq r'$, the term $\rho(r, r)$ is excluded from (2). The response r is classified into the response-class R_β, for which $\bar{\rho}(r, S_\beta)$ is maximal, and $N(S_\alpha, R_\beta)$ is incremented by one. If $k \ \bar{\rho}$ share the maximum, each corresponding matrix element is increased by $1/k$.

An information-theoretic measure, the mutual information I, quantifies the extent to which this clustering is random. For stimuli that are drawn from discrete classes $S_1, S_2, \ldots$, and responses that have been grouped into discrete classes $R_1, R_2, \ldots$, the mutual information I is given by

$$I = \frac{1}{N_{\text{tot}}} \sum_{\alpha, \beta} N(S_\alpha, R_\beta) \left[\log_2 N(S_\alpha, R_\beta) + \log_2 N_{\text{tot}} \right.$$

$$\left. - \log_2 \sum_\alpha N(S_\alpha, R_\beta) - \log_2 \sum_\beta N(S_\alpha, R_\beta) \right], \qquad [3]$$

where N_{tot} is the total number of stimuli. For C equally probable stimulus classes, random classification corresponds to $N(S_\alpha, R_\beta) = N_{\text{tot}}/C^2$ for $\forall \alpha, \beta \in \{1, \ldots, C\}$, where I becomes zero. For perfect classification, where each diagonal element of $N(S_\alpha, R_\beta)$ is equal to N_{tot}/C, the mutual information becomes maximal, i.e., $I = \log_2 C$.

1. Fujita, I., Tanaka, K., Ito, M. & Cheng, K. (1992) *Nature* **360,** 343–346.
2. Ito, M., Tamura, H., Fujita, I. & Tanaka, K. (1995) *J. Neurophysiol.* **73,** 218–226.
3. Logothetis, N. & Sheinberg, D. (1996) *Annu. Rev. Neurosci.* **19,** 577–621.
4. Tanaka, K. (1997) *Curr. Opin. Neurobiol.* **7,** 523–529.
5. Rolls, E. (2000) *Neuron* **27,** 205–218.
6. Perrett, D. & Oram, M. (1993) *Imaging Vis. Comput.* **11,** 317–333.
7. Wallis, G. & Rolls, E. (1997) *Prog. Neurobiol.* **51,** 167–194.
8. Riesenhuber, M. & Poggio, T. (1999) *Nat. Neurosci.* **2,** 1019–1025.
9. Fukushima, K. (1980) *Biol. Cybern.* **36,** 193–202.
10. Bair, W. (1999) *Curr. Opin. Neurobiol.* **9,** 447–453.
11. Singer, W. (1999) *Curr. Opin. Neurobiol.* **9,** 189–194.
12. Hopfield, J. (1995) *Nature* **376,** 33–36.
13. van Rullen, R., Gautrais, J., Delmore, A. & Thorpe, S. (1998) *BioSystems* **48,** 229–239.
14. Arnoldi, H., Englmeier, K. & Brauer, W. (1999) *Biol. Cybern.* **80,** 433–447.
15. Buonomano, D. & Merzenich, M. (1999) *Neural Comput.* **11,** 103–116.
16. Kandel, E., Schwartz, J. & Jessell, T. (2000) *Principles of Neural Science* (McGraw–Hill, New York), 4th Ed.
17. Douglas, R. & Martin, K. (1990) in *The Synaptic Organization of the Brain*, ed. Shepherd, G. (Oxford Univ. Press, New York), pp. 389–438.
18. Verschure, P. & König, P. (1999) *Neural Comput.* **11,** 1113–1138.
19. Gerstner, W. (1998) in *Pulsed Neural Networks*, eds. Maas, W. & Bishop, C. (MIT Press, Cambridge, MA), pp. 3–53.
20. Singer, W. (1999) *Neuron* **24,** 49–125.
21. König, P. & Engel, A. (1995) *Curr. Opin. Neurobiol.* **5,** 511–519.
22. Thorpe, S., Fize, D. & Marlot, C. (1996) *Nature* **381,** 520–522.
23. Allen, C. & Stevens, C. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 943–952.
24. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. (1998) *Proc. IEEE* **86,** 2278–2324.
25. Mitzdorf, U. & Singer, W. (1978) *Exp. Brain. Res.* **33,** 371–394.
26. Thomson, A. & Deuchars, J. (1994) *Trends Neurosci.* **17,** 119–126.
27. Agmon-Snir, H. & Segev, I. (1993) *J. Neurophysiol.* **70,** 2066–2085.
28. Thorpe, S. & Imbert, M. (1989) in *Connectionism in Perspective*, eds. Pfeifer, R., Schreter, Z. & Fogelman-Soulié, F. (Elsevier, New York), pp. 63–92.
29. Tovee, M., Rolls, E., Treves, A. & Bellis, R. (1993) *J. Neurophysiol.* **70,** 640–654.
30. Tovee, M. & Rolls, E. (1995) *Visual Cognit.* **2,** 25–58.
31. Schmolesky, M., Wang, Y., Hanes, D., Thompson, K., Leutgeb, S., Schall, J. & Leventhal, A. (1998) *J. Neurophysiol.* **79,** 3272–3278.
32. Richmond, B. & Optican, L. (1987) *J. Neurophysiol.* **57,** 147–161.
33. Abeles, M. (1982) *Local Cortical Circuits: An Electrophysiological Study* (Springer, Berlin).
34. Singer, W. & Gray, C. (1995) *Annu. Rev. Neurosci.* **18,** 555–586.
35. Milner, P. (1974) *Psychol. Rev.* **81,** 521–535.
36. von der Malsburg, C. (1995) *Curr. Opin. Neurobiol.* **5,** 520–526.
37. Engel, A., König, P., Kreiter, A. & Singer, W. (1991) *Science* **252,** 1177–1179.
38. Nowak, L., Munk, M., Nelson, J. & Bullier, A. (1995) *J. Neurophysiol.* **74,** 2379–2400.
39. König, P., Engel, A., Roelfsema, P. & Singer, W. (1995) *Neural Comput.* **7,** 469–485.
40. Crook, S., Ermentrout, G., Vanier, M. & Bower, J. (1997) *J. Comput. Neurosci.* **4,** 161–172.
41. Maass, W. (1997) *Neural Comput.* **9,** 279–304.
42. Victor, J. & Purpura, K. (1997) *Network Comput. Neural Syst.* **8,** 127–164.