1

THE CORRELATIVE BRAIN

The human brain is a hugely complex information processing system. In this chapter, it is our intention neither to review the brain anatomy and structures in detail nor to discuss every aspect of brain functions. Instead, we try to present an overview of the correlative brain at both the microscopic and macroscopic levels. Before discussing various correlative neural mechanisms, we first provide a brief background of some fundamental concepts of the human brain.

1.1 BACKGROUND

1.1.1 Spiking Neurons

The human brain consists of about 10^{11} (a hundred billion) neurons and $10^{15}-10^{16}$ (quadrillion) synapses. Each neuron is connected via synapses to about 1000–10,000 other neurons. It is the vast amounts of neurons and synapses that empower the brain with a high capacity for memory and "computing power" in a way that is quite different from the Turing machine or von Neumann–type computer.

A neuron is the basic functioning unit in the nervous system; it is responsible for receiving, integrating, and transmitting information. Despite the fact that there are many different types of neurons in terms of shape or size, most of them share a similar structure, as illustrated in Figure 1.1. Typically, a single cortical neuron receives thousands of inputs from other connecting neurons and sends its output spikes to about the same number of other neurons.

Correlative Learning: A Basis for Brain and Adaptive Systems, by Zhe Chen, Simon Haykin, Jos J. Eggermont, and Suzanna Becker

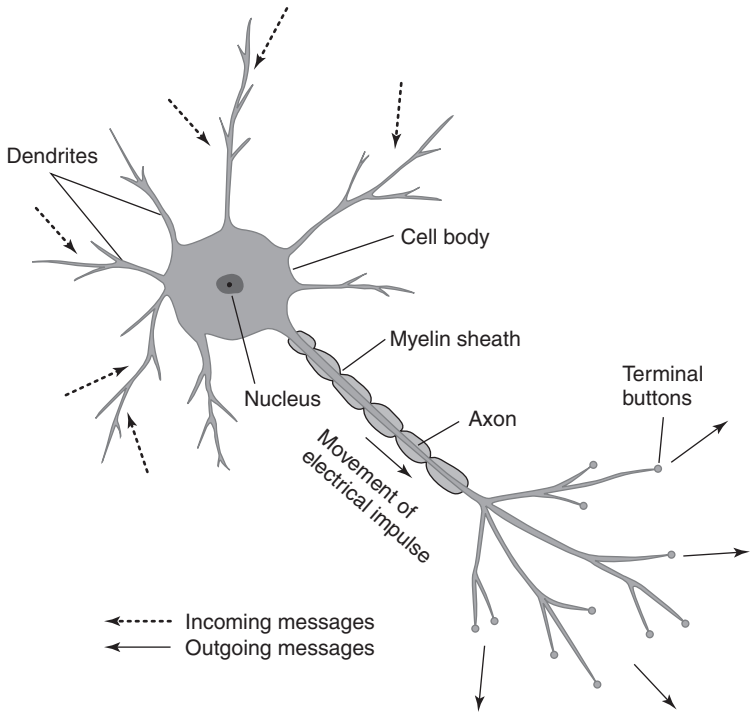


Figure 1.1 Schematic of neuron structure.

In Figure 1.1, there are several distinct components inside or outside the neuron:

Soma (cell body): Soma (Latin, meaning "body") is the cell body of the neuron and contains the nucleus and other structures that support the chemical processing.

Dendrite: Dendrites (Greek, meaning "tree") are the branching fibers that connect the soma; the fibers are the site of the synapses that are responsible for receiving incoming information from other neurons.

Axon: Axon is a singular fiber that carries information away from the soma to the synaptic sites of other neurons (dendrites and somas).

Synapse: Synapse (Greek, meaning "association")[1] is the connection that bridges two neurons or the connection between a neuron and a muscle. The synapse consists of three elements: (i) the presynaptic membrane, which is formed by the terminal button of an axon; (ii) the postsynaptic membrane consisting of a segment of dendrite or soma; and (iii) the space between these two structures, which is called the synaptic cleft.

Terminal buttons (boutons) are the small knobs at the end of an axon that release chemicals called neurotransmitters; the terminal buttons (boutons) form the presynaptic side of the synapse.

Myelin sheath consists of fat-containing cells that insulate the axon from electrical activity and increase the rate of transmission of signals. Axons that carry information over long distances, for example, from the periphery to the brain or between the two hemispheres of the cortex, tend to be myelinated while short-range axons do not.

Synapses are commonly believed to be the initial places where information is gained and stored. The massive number of synapses connecting the neurons across the brain constitutes a distributed memory system for storing the knowledge learned from experience. Depending on their electrical and chemical properties, synapses can be either *excitatory* or *inhibitory*. For the excitatory synapse, the neurotransmitters "depolarize" the postsynaptic membrane, that is, make the inside of the cell less negative with respect to its resting value (about $-70\,\text{mV}$). The change in membrane potential due to depolarization (i.e., electrical discharge) is called the *excitatory postsynaptic potential* (EPSP). If the depolarization of the postsynaptic membrane reaches a threshold (about $-55\,\text{mV}$), an action potential (i.e., spike) is generated in the postsynaptic neuron. In contrast, at the inhibitory synapse, the neurotransmitters "hyperpolarize" the postsynaptic membrane, that is, make the membrane potential more negative. The change in membrane potential due to hyperpolarization (i.e., electrical charge) is called the *inhibitory postsynaptic potential* (IPSP). The IPSP will make the neuron much less likely to spike when simultaneously receiving excitatory input.

The action potential generated at the postsynaptic neuron is a pulse of electrical activity that is created by a depolarizing current that exceeds the critical threshold level. This occurs because the exchange of ions across the membrane causes more sodium ions to enter the neuron; the spiking process often occurs over a time course of $2-100\,\text{ms}$, depending on the specific neuron. As a function of time, the spike trains can be observed at the location of a specific postsynaptic neuron, and these spike trains produce the spiking neural codes (see Figure 1.2). The spike train sequences can be roughly modeled as a homogenous Poisson process with the average firing rate as a rate parameter [201]. Specifically, let k denote the number of the spikes in the interval $(0, T]$, and let $r = k/T$ denote the average firing rate; by letting k and T approach infinity in the limit while keeping the ratio r constant, it follows that the probability of N spikes falling within an interval of time of bin

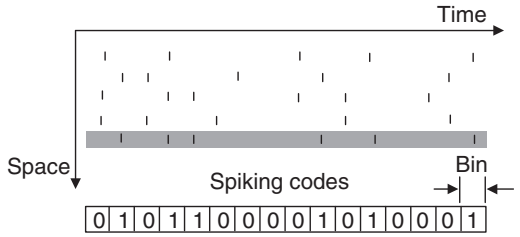


Figure 1.2 Graphical illustration of spiking neural codes.

size Δt is equal to

$$\Pr(N \text{ spikes in } \Delta t) = e^{-r\,\Delta t}\frac{(r\,\Delta t)^N}{N!}, \tag{1.1}$$

which defines a Poisson probability density function (pdf). Calculating the mean and variance of spike counts with respect to the Poisson probability would yield

$$\langle N \rangle = r\,\Delta t, \qquad \text{var}\,[N] = r\,\Delta t. \tag{1.2}$$

Additionally, given a spike at the present time, the waiting time (denoted by τ) between the current spike and the next spike follows an exponential distribution that has the pdf form

$$p(\tau) = re^{-r\tau}. \tag{1.3}$$

Calculating the mean and variance of τ with respect to $p(\tau)$ would yield

$$\langle \tau \rangle = \frac{1}{r}, \qquad \text{var}[\tau] = \frac{1}{r^2}. \tag{1.4}$$

A graphical illustration of simulated Poisson distributed spike trains is given in Figure 1.3. Figure 1.4 also presents an illustration of measuring firing rate via spike counting.

To understand brain function, we have to look into the "code" that neurons use. Action potentials (or spikes) are the primary way in which neurons communicate with each other; hence neural spikes are the unique "language" used inside the brain. In addition to the rate code (i.e., the number of spikes in a specific time interval), neurons may also use spike *timing* to code information (therefore referred to as temporal code). It appears that spike timing is important, at least in some neural systems such as the auditory regions, in that specific times between action potentials may carry information that is not available from the rate code. Experiments in vivo suggest that firing rates and synchrony are often simultaneously relevant. However, how firing rate and synchrony comodulate and which aspects of inputs are effectively encoded have yet remained elusive.

Functionally, a neuron is often simplified as an *integrate-and-fire* unit: The input x_i to a neuron i is generated by the firing rates x_j of other neurons j subject to a gain function

$$x_i = f\left(\sum_{j\in N_i}\theta_{ij}x_j - b_i\right), \tag{1.5}$$

where θ_{ij} denotes the synaptic efficacy and $f(\cdot)$ is a *gain function* which can be linear, nonlinear, or binary (all or none). Biologically speaking, equation (1.5) has the following interpretation:

- The parameter N_i defines the neighborhood region where neurons are connected to neuron i.

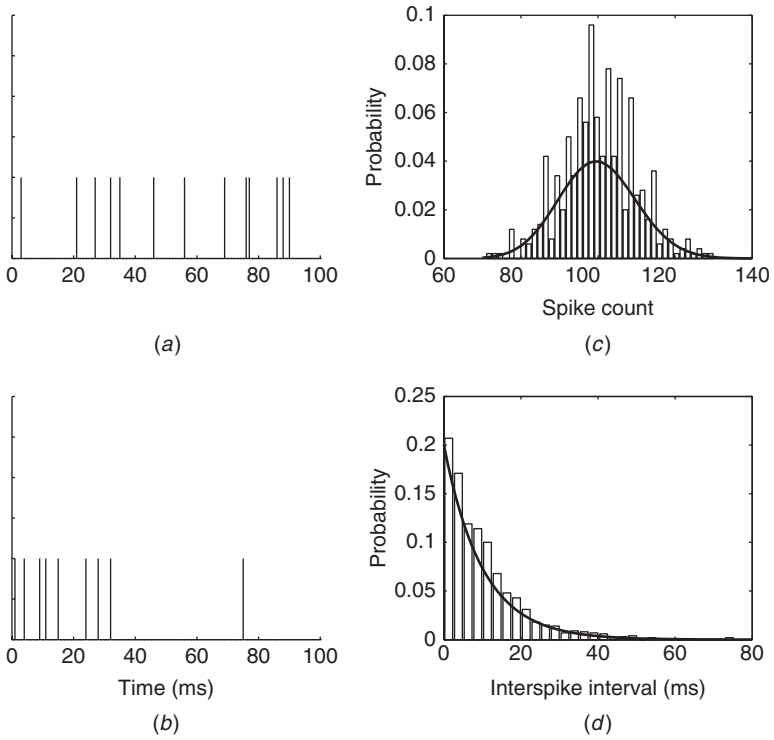


Figure 1.3 A graphical illustration of the Poisson spike trains. (*a, b*) Simulations of two Poisson spike trains with $r = 100$ and $\Delta t = 1$ ms. (*c*) Spike count histogram calculated from 1000 Poisson trains simulated within 1s duration; the solid curve is the Poisson spike count density. (*d*) Interspike interval (waiting time) histogram calculated from the simulations; the solid curve is the exponential interspike interval density.

- The weighted summed current $I_i = \sum_{j \in N_i} \theta_{ij} x_j$ is often called the *postsynaptic potential* (PSP) of neuron i.
- The voltage x_i is viewed as the firing rate of neuron i.
- The threshold bias b_i is viewed as a baseline current.
- The function f can be viewed as an operation that is implemented via dendritic integration.

It is this "integrate-and-fire" mechanism described in (1.5) that motivated Warren McCulloch and Walter Pitts [606] to first develop the computational neuron model. The McCulloch–Pitts neuron is a static model; despite its simplicity, the McCulloch–Pitts neuron model has been widely used in the neural network literature. In the meantime, more biologically accurate neuron models, such as Caianiello's neuron model [132] and the Hodgkin–Huxley model [395], also have been developed to analyze neuronal dynamics.

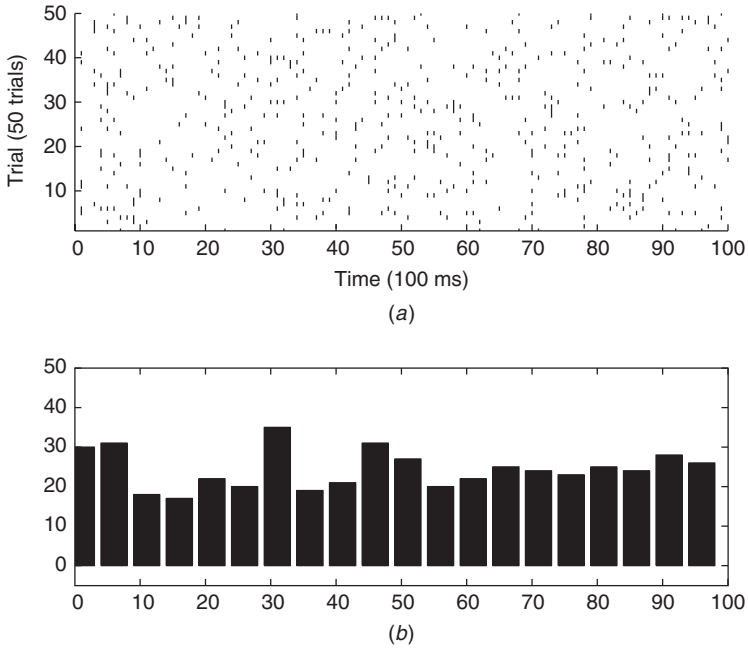


Figure 1.4 (*a*) The spike trains observed within 100 ms over 50 independent trials. (*b*) The total number of spike counts per 5 ms within 50 trials, from which we can calculate the mean firing rate as about 100 spikes/s.

Specifically, Caianiello [132] introduced the *time delay* into the model of a neuron's temporal dynamics,

$$x_i(t) = f\left(\sum_j \sum_k \theta_{ij} x_j(t - k\tau) - b_i(t)\right). \tag{1.6}$$

The above so-called *neuronic equation* essentially states that neuron j can influence the firing of neuron i up to $k\tau$ time steps in the future, and the dynamics can be modeled as a Markov process.[2]

To model the single neuron's firing rate, a simple way is to link the Poisson rate to the membrane potential from a biophysical viewpoint:

$$r(t) \approx \alpha[V(t) - V_{\text{th}}], \tag{1.7}$$

where V_{th} (in millivots) denotes a potential threshold value, α (in spikes per second per millivolt) denotes the slope parameter, and $V(t)$ denotes the instantaneous membrane potential. Taking the time average of (1.7) would yield the mean firing rate expression

$$\langle r(t) \rangle \approx \alpha[V_0(t) - V_{\text{th}}], \tag{1.8}$$

where $V_0(t) = \langle V(t) \rangle$ denotes the time-averaged membrane potential. Nevertheless, the neural firing of a single cell is known to be very noisy. If we measure the firing rate in different trials by presenting the same or correlated stimuli, a significantly different firing pattern can be observed. Such random firing effects can be overcome by averaging an ensemble of neurons or a population of cells; by doing that the firing rate function appears more deterministic. In practice, the firing rate is modeled as a filtered version of a known stimulus signal

$$r(t) = r_0 g\left(\int_{-\infty}^{\infty} d\tau\ f(\tau) s(t-\tau)\right), \tag{1.9}$$

where r_0 denotes the background firing rate when no stimulus occurs (i.e., $s = 0$), $f(t)$ denotes a filter, and $g(\cdot)$ denotes a memoryless nonlinear function whose argument is a *reverse correlation function*. Note that if the stimulus signal $s(t)$ is close in shape to that of the filter $f(t)$, specifically $s(t) = f(-t)$, then the rate function $r(t)$ will increase its value considerably, thereby achieving the maximum modulation.

1.1.2 Neocortex

The brain of vertebrates consists of the forebrain, brainstem, and spinal cord. In the forebrain the most recently evolved component, and the most prominent component in higher vertebrates, is the neocortex. In addition, the forebrain includes phylogenetically older cortical areas (allocortex) such as the olfactory cortex and hippocampus as well as many nuclei important for emotion (e.g., the amygdala), motor control (the basal ganglia), and numerous other functions. The brain is divided into left and right hemispheres. Different sides of the brain are responsible for controlling their *opposite* sides of the body. While the precise role of each hemisphere is still under debate, it is generally agreed that the left hemisphere plays a greater role in language and object recognition while the right plays a greater role in spatial cognition. The hemispheres of the cerebral cortex are also divided into four divisions, or lobes, the *frontal, parietal, occipital*, and *temporal* lobes. The gray matter volume within a given region of the brain often correlates positively with specific skills associated with that region.

In different cortical areas, there are specialized *functional* cortices responsible for specific tasks of sensory perception, cognition, or motor control. The neurons in those specific cortical areas often form specific topographic maps; the neurons within the same cortical region also have similar functional roles and structures. In particular, five important cortices of the neocortex are described here:

Visual cortex is specialized for vision; it is located at the back of brain in the occipital lobe. There are also numerous visual areas within the temporal and parietal lobes. The neurons within the visual cortex receive and process the information from the eyes (namely, their retinae) and complete the visual tasks. In monkeys nearly half of the cerebral cortex is related to visual processing.

Auditory cortex is specialized for audition or hearing; it is located in the temporal lobe. The neurons in the auditory cortex process the information received at the auditory nerves from the inner ear (cochlea) and further propagated through the auditory brainstem and the ascending auditory system.

Somatosensory cortex is mainly specialized for haptic sensations; it is located in the parietal lobe.

Motor cortex is specialized for movement; it is located in the back portion of the frontal lobe.

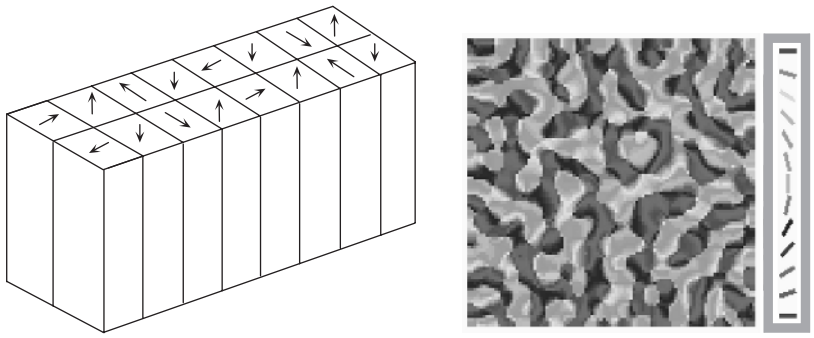
Association cortex refers to the areas of the lobes that are multimodal, receiving converging inputs from multiple sensory modalities. Different association cortices may be specialized for different functions, such as language comprehension, spatial imagery, memory, or sensorimotor transformations.

Within the motor or sensory cortices, there are also *primary* and *secondary* motor or sensory areas. The primary motor or sensory areas are those where motor or sensory information first arrives at the cortex. These primary areas are responsible for processing the primitive motor command or low-level sensory stimuli. For representing cortical areas of neocortex, Table 1.1 lists some abbreviated terms commonly used in neuroscience.

The neocortex is thought to be a self-organizing system[3] in the sense that a larger degree of order emerges from the system as time progresses. The neocortex is structurally ordered at many levels, including the layered and columnar structure, groupings of columns into hypercolumns, and at a larger scale into topographically organized feature maps. A central and long-standing theme in neuroscience has been to study *why* and *how* these ordered structures and maps are formed in the neocortex. Information arriving at the neocortex, in the form of spatiotemporal spike patterns, is structured, redundant, high dimensional, and somewhat random. In terms of their roles, there are two categories of maps: *functional* and *topographic*. Topographic maps are by definition functionally structured, but functional maps

Table 1.1 Common Terminology for Areas in Sensory and Motor Cortices

Term	Description
V1	Primary visual cortex, striate cortex
V2	Secondary visual cortex
MT	Medial temporal, V5
IT	Inferior temporal
A1	Primary auditory cortex
A2	Secondary auditory cortex
S1	Primary somatosensory cortex
S2	Secondary somatosensory cortex
M1	Primary motor cortex
M2	Secondary motor cortex



(*a*) (*b*)

Figure 1.5 (*a*) Graphical illustration of three-dimensional columnar structure with two arrays of orientation selective cells. (*b*) Computer simulation of two-dimensional orientation maps of visual cortex.

might not be topographically organized. Different cortical areas have their own specific functional maps, for example:

- Visual maps can represent the distance to an object, line orientation, movement direction, binocular disparity, and so on.
- Auditory maps can represent the object in terms of azimuth, elevation, and distance by synthesizing the maps of time and intensity disparity.
- Motor maps can, for instance, represent gaze direction; variations in motor commands are represented topographically into spatiotemporal patterns within the motor maps.

Topographic (such as retinotopic, somatotopic, or tonotopic) maps arise as a result of the anatomical structure of the sensory receptor surface and the innervating nerve fibers preserving this orderliness in the fiber tracts and in each interposed nucleus. Although the roles of topographic maps vary, a commonly accepted view is that the maps provide a low-dimensional representation of complex stimuli in the cortices. Topographic map formation has been widely studied using correlation-based neural models and learning rules (to be discussed in Chapter 3). As an example, the orientation-selective columnar cells in the visual cortex are illustrated in Figure 1.5.

1.1.3 Receptive Fields

Another important notion for understanding how neurons process and respond the sensory stimuli is the so-called *receptive field* (RF).[4] Each neuron has its own RF. Although the size and property of different neurons may vary, their common goals are to detect, match, and encode the (primitive or abstract) features of the information flow. By appropriate tuning of the synaptic strengths of inputs within a neuron's RF, that neuron can be viewed as a feature detector whose task is to extract a set

of information-bearing features to represent (with maximum information retention) the complex sensory stimuli.

Within the neural maps, neighboring cells often have *similar* and *overlapping* RFs, which enable them to cooperate with each other in processing the incoming stimuli. For instance, the neurons in the visual orientation maps have RFs that cause them to respond only to a small subset of visual stimuli that are strongly localized in the retinal space as well as the orientation angle space. Computationally, Daugman [199] used two-dimensional Gabor filters to model the spatial RFs of simple cells in the visual cortex,

$$\mathrm{RF}(x, y) = \exp\left(-\frac{\tilde{x}^2 + \gamma^2 \tilde{y}^2}{2\sigma^2}\right) \cos\left(2\pi \frac{\tilde{x}}{\lambda} + \varphi\right), \tag{1.10}$$

where

$$\tilde{x} = x \cos\ \Theta + y\ \sin\ \Theta, \qquad \tilde{y} = -x\ \sin\ \Theta + y\ \cos\ \Theta,$$

where the arguments x and y define the spatial position of the visual RF; parameter γ is the aspect ratio that specifies the support of the Gabor filter; parameter λ defines the *wavelength*, and $1/\lambda$ defines the *spatial frequency*; parameter σ defines the *size* of the RF, and the ratio σ/λ determines the spatial frequency bandwidth of the cells; the angle parameter $\Theta = 2\pi/k$ ($k \in \mathbb{N}$) specifies the *orientation* of the impulse response, and φ is a *phase offset* parameter (when $\varphi = 0$, the RF function is symmetric; when $\varphi = \pi$, the function is antisymmetric). It is well believed that the Gabor filter provides a good approximation of the response properties of visual cells [276]. Figure 1.6 depicts some computer simulations of visual RFs using a Gabor filter with varying parameters (γ, λ, σ, Θ, φ).

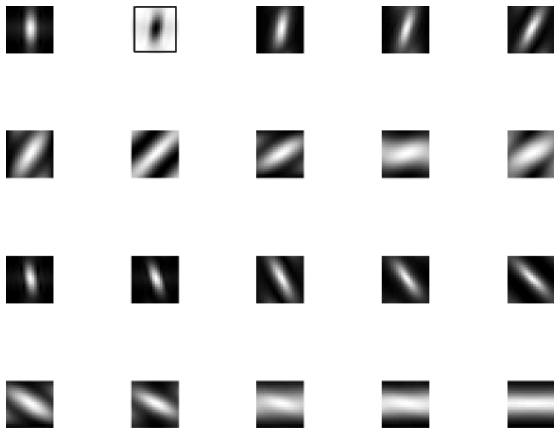


Figure 1.6 Illustration of visual receptive fields. The orientation-selective receptive fields are simulated by two-dimensional Gabor filters. The first two correspond to the ''ON-center-OFF-surround'' and ''OFF-center-ON-surround'' cells, respectively.

Likewise, the neurons in the auditory maps have similar and overlapping spectrotemporal receptive fields (STRFs) in terms of either the amplitude (modulation) or frequency (tone) of the sound stimuli. In a similar vein, we can define the STRF with a two-dimensional complex Gabor filter,

$$\mathrm{STRF}(t, f) = \frac{1}{2\pi\sigma_t\sigma_f}\exp\left(-\frac{(t-t_0)^2}{2\sigma_t^2} - \frac{(f-f_0)^2}{2\sigma_f^2}\right) \times \exp\left[j\omega_t(t-t_0) + j\omega_f(f-f_0)\right] \qquad (j=\sqrt{-1}) \qquad (1.11)$$

which is modeled by the product of a Gaussian envelope and a complex-valued Euler function. The Gaussian envelope is specified by the mean parameters t_0 and f_0 (central frequency) and the standard deviation parameters σ_t and σ_f. The periodicity is defined by the radian frequencies ω_t and ω_f. The scaling factors σ_t and σ_f at time and frequency make the Gabor filter act like a wavelet function for multiresolution analysis. Therefore, the auditory neurons with a waveletlike STRF can tune their auditory responses according to varying auditory stimuli.

1.1.4 Thalamus

Most sensory input to the cortex (including visual, auditory, and somatosensory but not olfactory) project to the cortex primarily via the thalamus, although there are also nonthalamic pathways. Thus the thalamus is the last region in the primary processing chain between sensory receptors and the cortex. Despite its relatively compact volume, its role in information processing is extremely important. It is now widely believed that the thalamus is more than a relay station between the received sensory stimuli and sensory cortices. Indeed, surprisingly it has been found that the number of feedback connections in the corticothalamic loop is about 10 times as many as that of feedforward connections in the thalamocortical loop.[5] In the visual pathway, the thalamic structure is known as the lateral geniculate nucleus (LGN); whereas in the auditory pathway, it is referred to as the medial geniculate nucleus (MGN) or medial geniculate body (MGB). The motor information generated by the cerebellum or basal ganglia also passes through thalamus to motor cortex. The feedback projections are believed to play a crucial role for selective attention, top-down expectation, or prediction (given the contextual prior). See Figure 1.7 for an illustration of thalamocortical and corticothalamic loops in the visual system.

1.1.5 Hippocampus

The hippocampus,[6] an older part of cerebral cortex, is located inside the temporal lobe of the brain. The perforant path constitutes the predominant input pathway to the hippocampus and it projects mainly to the superficial layers of the entorhinal cortex (EC), which in turn projects to the dentate gyrus and CA fields (CA stands for cornu ammonis—so called because the whole structure looks like rams' horns). There are also connections from the dentate gyrus to CA3, from CA3 to CA1, and

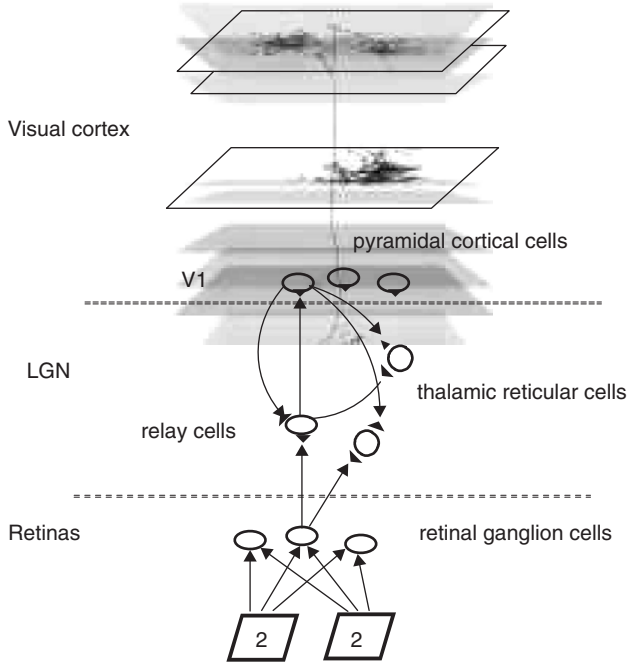


Figure 1.7 Schematic of thalamocortical and corticothalamic loops between the LGN and primary visual cortex (V1).

CA1 back to the EC (as shown later in Figure 1.15). Studies in rats have shown that neurons in the hippocampus have spatial firing fields, for which these cells are known as the *place cells*. The discovery of place cells has led to the idea that the hippocampus might act like a cognitive map [682].

1.2 CORRELATION DETECTION IN SINGLE NEURONS

The most important characteristic of a well-functioning brain is that it learns by experience. Learning starts with modifiable synapses, which are considered more and more as important computational systems of the brain [2]. The idea of synapse involvement in memory and thus implicitly that of modifiable synapses has a rather long history [747].

The Law of Neural Habit and Correlative Synapses. An early idea of the correlative synapse can be traced back to William James. In his classic work on psychology [436] (excerpted in [39]), James proposed the laws of association ([39], p. 225):

> How does a man come, after having the thought of A, to have the thought of B the next moment? or how does he come to think of A and B always together? These

> were the phenomena which Hartley undertook to explain by cerebral physiology. I believe he was in essentially respects, on the right track, and I propose simply to revise his conclusions by the aid of distributions which he did not make.

In James's theory, he claimed that ([39], p. 566; also in [122])

> there is no other elementary causal law of association than the law of neural habit: When two elementary brain-processes have been active together or in immediate succession, one of them, on reoccurring, tends to propagate its excitement into the other.

Essentially, James's *law of neural habit* indicates the basic conditions ("being coactive" and "reoccurring") for the modification of neural synapses, although he did not restrict himself to the synapses; instead, he used the term "elementary brain processes." However, James's theory clearly bears a resemblance with the theory on synaptic plasticity established later.[7]

Herbert Spencer, in *The Principles of Psychology* [844], has also described similar concepts of correlation-based modification of synaptic connections; he also indicated the fundamental connection between nervous changes and psychological states and discussed the psychological aspects of intelligence. In his words ([844], p. 408)

> when any state a occurs, the tendency of some other state d to follow it, must be strong or weak according to the degree of persistence with which A and D (the objects or attributes that produce a and d) occur together in the environment.

Basically, this law of connection states that if two external events occur in a correlative fashion, the associated internal states will also be correlated correspondingly; it is the "strengths of the connection" between the internal states and external events that are important to encode the information or knowledge within the brain [844].

Following the early research studies in psychology, Young [990] also suggested that repeated excitation leads to a permanent facilitation, that is, stronger and more efficacious synapses between neurons. McCulloch and Pitts [606] were among the first to phrase the properties of what later would be called Hebb's synapse in the following words:

> The phenomena of learning, which are of a character persisting over most physiological changes in nervous activity, seem to require the possibility of permanent alterations in the structure of [neural] nets. The simplest such alteration is the formation of new synapses or equivalent local depressions of threshold. We suppose that some axonal termination cannot at first excite the succeeding neuron; but if at any time the neuron fires, and the axonal terminations are simultaneously excited, they become synapses of the ordinary kind, henceforth capable of exciting the neuron. The loss of inhibitory synapses gives an entirely equivalent result.

According to Changeux and Heidmann [155], the first mention of changes in strength or number of connections in neural networks can already be found in

Descartes' *Traité de l'homme* (1677). In this case, we have to convert several aspects of Descartes' concept of a hydraulic nervous system to those fitting the present electrochemical one.

Postulate of Hebbian Learning. The most influential proponent of learning as a correlative process was Donald Hebb, who postulated the following, now referred to as *Hebb's postulate* ([377], p. 62)[8]:

> When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic changes take place in one or both cells such that A's efficiency as one of the cells firing B, is increased.

The clause "*takes part in firing it*" indicates the causality condition and implies both temporal specificity, that is, the spikes from cell A occur prior to and within a short time window of the firings in cell B, and spatial specificity so that only the synapse involved in firing cell B gets strengthened. Stated mathematically, Hebb's postulate can be formulated as

$$\Delta\theta_{AB}(t) = \eta x_A(t) y_B(t), \tag{1.12}$$

where x_A and y_B represent the pre- and postsynaptic activities (i.e., firing rates), respectively, between the synapse connecting neurons A and B; $\Delta\theta_{AB}$ denotes the change of synaptic strength; and η is a small step-size (also known as learning-rate) parameter. Namely, the change of the synaptic weight $\theta_{AB}(t)$ is proportional to the product of input $x_A(t)$ and output $y_B(t)$. The learning rule is local, since the information for modifying the synapse is easily available at the location of the synapse. Averaged over many time steps, the synaptic weight becomes proportional to the *correlation* between pre- and postsynaptic firing [320].

Although Hebb's postulate became well known in 1949, it is not until nearly a quarter of a century later that physiological experiments first offered the validated evidence of Hebb's proposal. In 1973, Bliss and Lomo [100] published a paper describing a form of activation-induced synaptic modification in the hippocampus of the brain. In their experiments, they applied pulses of electrical simulation to the major pathway entering the hippocampus while recording the synaptically evoked responses, and they reported the *long-term potentiation* (LTP) phenomenon.[9] Long-term potentiation shows a number of associative properties in that there are interaction effects between *coactive* pathways. Specifically, if a weak input that would not normally cause a strong postsynaptic response is paired with a strong input, the weak input can be potentiated. Such an associative property can find its links with Pavlov's conditioning experiments and Hebb's postulate, and it is believed to form the cellular basis of memory.

Hence, the "Hebb-like effect" can be long lasting. In Hebb's original words, this consequence is described as ([377], p. 70)

> any two cells or systems of cells that are repeatedly active at the same time will tend to become associated, so that the activity in one facilitates activity in the other ... such that a reverberation in the structure might be possible.

In the literature, synapses that follow Hebb's postulate, when using the standard LTP protocol described above, are called *Hebbian synapses*. The important features of Hebb's rule include (i) a time-dependent mechanism, (ii) a local mechanism, (iii) an associative mechanism, and (iv) a correlational mechanism for which the Hebbian synapses are often referred to as *correlational synapses* [36].

Nowadays, Hebb's postulate has been widely accepted and supported by numerous neurophysiological data. It is believed that Hebbian correlation between presynaptic and postsynaptic neurons, which leads to synaptic plasticity, is mediated by backpropagating action potentials that are actively or passively transmitted to the synapse.

Experience-Dependent Synaptic Plasticity in Neocortex. The formulation of the Hebb rule $\Delta\theta_{AB}(t) = \eta\langle x_A(t)y_B(t)\rangle$, that is, the change in synaptic weight is proportional to the correlation of presynaptic and postsynaptic activities, appears to lead to untenable predictions [10, 11]. These authors recorded from pairs of neurons that either directly excited or directly inhibited each other in the auditory cortex of behaving monkeys. They found that functional plasticity is a function of the change in correlation (or covariance) and not of correlation or covariance per se. They also found that the size of the plasticity effect was increased approximately sixfold during appropriate behavior. The spike activity of the presynaptic cell was considered as the conditioned stimulus (CS), the response from the postsynaptic cell the conditioned response (CR), and the auditory stimulus the unconditioned stimulus (US) when presented 2–4 ms after a spike of the presynaptic cell. The monkey was trained to respond to the US. Specifically, they suggested a modified Hebbian learning rule as follows:

$$\Delta\theta_{AB}(t) = \eta\left[x_A(t)y_B(t+\tau) - \langle x_A y_B\rangle\right], \tag{1.13}$$

where the time interval τ is only a few tens of milliseconds after the time of a CS spike at time t and the average correlation $\langle x_A y_B\rangle$ is taken over at least several minutes. Thus the changes in synaptic weights are proportional to the changes in correlation. Appropriate behavior increases the modification factor by about a factor 6, as more or less required by *Thorndike's law of effect* [882]. Ahissar and colleagues [10, 11] also suggested that, alternatively, fractional changes in synaptic weights could be proportional to fractional changes in the correlation.

Spike-Timing-Dependent Plasticity. There are two main problems with the classical Hebbian synapse; one is that under the standard formulation the synaptic strength can only increase. Such a system, when linear, is inherently unstable and results in unlimited growth of excitatory synapse strength. The system can be kept stable through nonlinear saturation or by imposing normalization conditions. One could, for instance, keep the total summed weight of all synapses to a given neuron constant, that is, when one synapse increases in strength the others have to decrease collectively by the same amount. This mechanism contradicts with the supposed spatial selectivity of synaptic strengthening or weakening. However, numerous

reports about the occurrence of heterosynaptic LTP and LTD have surfaced in recent years [89], so this is a feasible solution. Another problem with the firing-rate-based Hebb synapse is the way the association between the firings of the input and output neuron is supposed to occur. This can be assessed much more effectively on the basis of a spike-timing-based correlation procedure compared to a rate-based one.

Recently several investigators [80, 584, 589] presented evidence that the precise timing difference between pre- and postsynaptic action potentials determines whether LTP or LTD will occur. Long-term potentiation occurs when the presynaptic spikes precede the postsynaptic ones, whereas LTD occurs when the postsynaptic spikes precede the presynaptic ones. The time window for these phenomena is rather short (Figure 1.8), of the order of tens of milliseconds, and the phenomenon is called *spike-timing-dependent plasticity* (STDP). Essentially, STDP imposes a *temporally asymmetric* time window on Hebbian learning [89]; that is, if a presynaptic neuron fires a short time before the postsynaptic neuron, positive Hebbian learning occurs, whereas if the postsynaptic neuron fires a short time before the presynaptic neuron, anti-Hebbian learning occurs. This form of spiking-time-dependent Hebbian learning is more realistic in that it captures the causal relationship that exists between presynaptic and postsynaptic firing [317, 320, 484]. Specifically, the STDP learning rule has several distinct features [195]: (i) the bidirectionality of synaptic modification with approximately balanced LTP and LTD, which helps the neural circuit maintain its net synaptic excitation at a stable level; (ii) the spike sequence dependence of synaptic modification, which allows the circuit to learn sequences and to encode causality of external events; and (iii) the narrow

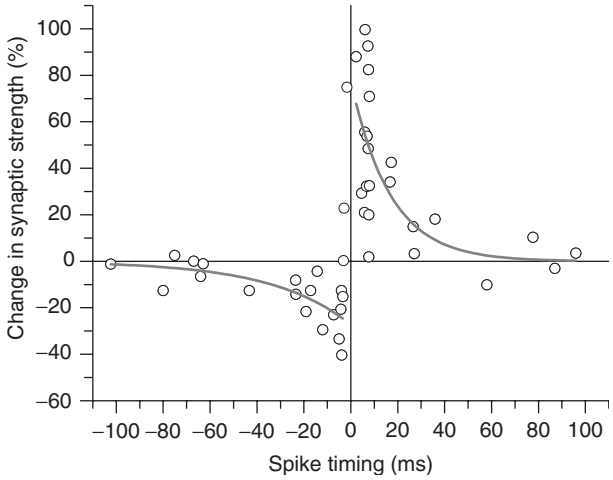


Figure 1.8 Illustration of temporally asymmetric spiking-time-dependent Hebbian synaptic plasticity. The synaptic modifications (LTP or LTD) are induced by correlated pre- and postsynaptic spiking. (Reprinted, with permission, from the *Annual Review of Neuroscience,* Vol. 24. Copyright © 2001 by Annual Reviews.)

temporal window, which allows the system to select inputs based on its response latency with a millisecond precision, thus shaping the temporal dynamics of the circuit.

The biphasic learning window of STDP overcomes the instability problem inherent in the rate-based Hebbian learning rule if there is slightly more depression than potentiation. The temporal window length arises naturally in a model where backpropagation of action potentials from the cell soma, where they are initiated, into the dendrites toward the synapses is considered. This makes the timing of the postsynaptic spikes available at the synapse, and the backpropagated signal functions as an associative signal for synapse modification. The conduction velocity of these backpropagated action potentials is of the order of 0.5 m/s in cortical pyramidal cells [130, 863], and with a typical dendritic length of 0.5 mm this translates in a delay of about 1 ms between the initiation of the action potential and its availability at the dendritic synapse. Recently several investigators [80, 584, 589] presented evidence that the precise timing difference between pre- and postsynaptic action potentials determines whether activity-dependent LTP or LTD will occur.

Depolarization of the postsynaptic membrane (e.g., by a backpropagating action potential) can remove a Mg^{2+} ion from the pore of an NMDA (*N*-methyl-D-aspartate) receptor channel, thereby allowing an influx of Ca^{2+} when the presynaptic terminal releases glutamate. This mechanism allows an NMDA receptor channel to function as a molecular detector of the coincidence of presynaptic activity and postsynaptic depolarization [106]. The resulting influx of Ca^{2+} may lead to synaptic potentiation.

The STDP is dependent not only on the timing interval between pre- and postsynaptic spikes but also on the timing of preceding presynaptic spikes. Such spikes can depress the efficacy of following spikes in producing STDP. Therefore the first spike of a burst in the presynaptic neuron is the dominant one in causing synaptic modification [297]. Recent studies [298] suggest that STDP is also locationdependent; specifically, the activity-dependent synaptic modification depends on dendritic location according to the temporal characteristics of presynaptic spikes.

In experimental studies, STDP was shown to be instrumental in eliciting changes in orientation columns in cat visual cortex, thereby demonstrating the link between *synaptic plasticity* and *representational plasticity*. Schuett et al. [803] paired brief flickering gratings of low spatial frequency and with a particular orientation with one 60-μA electrical pulse in about 300 μm below the cortex surface for 3–4 h. The timing of the pairing was critical; a shift in orientation preference toward the paired orientation occurred at the site of electrical stimulation if cortex was activated first visually and then electrically. A similar result was found by repetitive pairing of two visual stimuli with different orientations for 3–6 min [988]. A shift in orientation tuning of cortical neurons was found with the direction of shift determined by the order of presentation. An effect was found when the time difference of the presentation was about 40 ms. They also demonstrated that this stimulation paradigm in humans produced a shift in perceived orientation, thereby demonstrating a link between synaptic plasticity, representational plasticity, and perception. Song and Abbott [842] in a modeling study demonstrated that the formation of orientation

columns during development as well as their remapping in adulthood follows the timescales and biphasic shape of STDP.

1.3 CORRELATION IN ENSEMBLES OF NEURONS: SYNCHRONY AND POPULATION CODING

Correlative Firing. In neuroscience, correlative firing refers to two or more neurons (or ensembles of neurons) that tend to be activated at the same time [786]. According to Cook [183], correlated firing occurs at two levels. In the short term, since few neurons can be driven reliably by a single axon, the relative timing of multiple inputs is crucial to their influence. For a population of neurons, a "window of opportunity" focuses on the moment at which a strong volley of afferent impulses shifts the membrane potential toward the firing threshold; within that window the effect of another input on the neuron's output may be enhanced. In the longer term, for some neurons and synapses, the relative timing of multiple inputs can modulate synaptic efficacy in long-lasting ways and thus change the functional properties of the circuit.

Correlated activities are widely witnessed in various sensory (visual, auditory, olfactory, or somatosensory) systems (e.g., [336, 337, 531, 582, 583]) and motor system (e.g., [501]). Although there remain some distinctions between different systems, the basic functional principles are similar. For example, in the visual system, neighboring neurons, in areas from retina to cortex, tend to fire synchronously more often than would be expected by chance; correlated firing among neural assemblies abounds at cortical and subcortical (e.g., thalamic) levels [16, 833].

For the auditory system, Eggermont [243] reviewed the role of correlation and synchrony in auditory cortex. Specifically, in the auditory brainstem and midbrain, inhibitory interactions between neurons further add to the highly nonlinear nature of the coding of sound whereby the firings of individual cells become highly interdependent and their firing times may become correlated. The way sound is represented at various levels of the auditory system forms the basis for its neural coding. A *neural code* is considered here as a vocabulary of the firings represented at a subcortical and/or cortical level on which perceptual discrimination is based. This vocabulary, an N-dimensional vector (with N the number of participating neurons, i.e., the size of the assembly), contains all the information needed for the perceptual decision process. Examples of such vocabularies are those based on instantaneous firing rates, integrated firing rates, and mean interspike interval duration of a group of specialized neurons [248]. How a neural code is constructed out of neural representations depends on (i) the sensitivity of the neurons to detect changes in the stimulus, (ii) the variability in the individual neurons' responses to the same stimulus, and (iii) the correlation between the responses of the individual neurons. If a neural code was based on firing rate, then independence of the firings in neighboring neurons would allow more information to be transmitted and correlations between the firings of individual neurons would generally diminish the information capacity of a neuronal population [1002]; however, it can improve the accuracy of the neural code [1, 770].

Population Coding in Motor and Sensory Systems. Animals extract information in parallel from an initially unknown, usually time-varying stimulus on the basis of short segments of a large number of spike trains to allow real-time estimation of some aspects of the stimulus [761]. Potential examples of pseudo-real-time estimation procedures are found in the *population vector coding* method applied to motor cortex [315] and the superior colliculus [694]. In these models, assuming independence of neuronal firing, the firing rates of neurons were weighted by their preferred hand-pointing or saccadic eye-movement directions and added up to provide a movement vector that predicted the motor output in strength and direction. If the motor neurons are assumed to be tuned in cosine fashion to a particular angle-of-motion direction (d, in radians), that is, the individual neuronal firing rate r_n that depends on d and achieves its maximum $r_{n,\max}$ in the preferred angle of direction d_n satisfies a cosine tuning function

$$r_n(d) = r_{n,\max} \cos(d - d_n), \tag{1.14}$$

where only positive cosine values are taken into account, then the weight of each individual contribution to the final compound saccade vector will be given by the correlation of its preferred firing direction and the desired direction of motion, which is proportional to the cosine of the angle between the two vectors. The population vector model then states that the direction of motion induced by the population activity is given by

$$d_{\text{pop}} = \frac{1}{N} \sum_{n=1}^{N} \frac{r_n}{r_{n,\max}} d_n., \tag{1.15}$$

where N denotes the total number of motor neurons. This model is equally applicable to encoding of a stimulus direction, for example, orientation of a visual object, but its assumption about independence of individual neuron activity and its sensitivity to noise (i.e., the spontaneous firing activity) make it less than ideal [736, 785].

Place cells in the hippocampus that code the position of the animal in reference to its environment, cells in visual cortical field MT that detect direction of motion, and cells in visual cortical field V1 that are tuned to the orientation of a stimulus are also prime examples of population coding on the basis of firing rate that can produce adequate stimulus reconstruction [203, 206, 735]. Recently the importance of dedicated subgroups of neurons in the hippocampus ("cliques") that can initiate various startle responses has been highlighted [556]. Dedicated subgroups of neurons ("clusters") have been identified for representation of auditory space in the midbrain and forebrain [179]. These clusters are not part of topographic maps because neighboring clusters may be coding for completely different sound location cues.

Examples of population coding in auditory cortex based on the firing rate of (presumably independently firing) neurons are found in the panoramic code of sound location [299, 619], in the population vector model of sound azimuth coding [252], and in the coding of vocalizations [312, 797] or periodic sounds [574].

The sampling of the neuronal population in all these studies was done sequentially, thereby making their activities in fact independent. The coding of the sound direction features by firing rate was much better than those of the vocalizations or periodic sounds. Thus, better representational codes must exist for aspects of sensory stimuli other than those related to direction or location.

Role of Correlated Firing in Neural Coding. Sensory systems often represent distinct features of the environment by spatially distinct sets of neurons. For instance, in the visual system, color, texture, and size are encoded in different visual areas. Thus, a yellow, fuzzy tennis ball and a red, smooth pool ball would be coded in one area as yellow versus red, in another area as fuzzy versus smooth, and in the third one as slightly different sizes. Somehow, the relationships of the properties belonging to the tennis ball and the pool ball need to be tagged to prevent us from seeing a fuzzy, red pool ball. This may require a mechanism, such as enhanced neural synchrony between cortical areas [335], to group the extracted features belonging to a specific object. It is also possible that the common spatial location of [yellow, fuzzy] for the tennis ball is a sufficient tag that could be accomplished by connections of the color and texture areas to the retinal maps in V1.

In the auditory system, important sound features are "components of an auditory scene [that] appear to be perceptually grouped if they are harmonically related, start and end at the same time, share a common rate of amplitude modulation or if they are proximate in time and frequency" [184]. Thus, important sound features allow correlations in the temporal domain and spectral domain that signal sufficient overlap to be grouped into one percept or assigned to one sound source (Figure 1.9). Sounds can be meaningfully decomposed into contours (e.g., temporal envelopes) and texture (e.g., frequency content), as is common for visual images [248]. The most meaningful aspects of speech are likely the sound envelopes as these play a crucial role in speech recognition as demonstrated by replacing the detailed frequency information by octavewide bands of noise without affecting recognition to an appreciable extent [824]. These sound envelopes also produce the largest changes in the correlation of neural activity, compared to a nonstimulus condition, in auditory cortex [248]. The correlated activity across a neural population may emphasize these stimulus contours above their texture, despite the fact that STRF overlap accounts for up to 40% of the variance in pairwise neural correlation [250]. This suggests that the fraction of shared inputs from the auditory thalamus by cortical cells represents those that potentially take part in a correlated neural assembly but the firing times of a neuron are codetermined by the sound envelopes as filtered by the neuron's STRF.

Coding of complex sounds requires a population of neurons. In response to complex sounds, cortical neurons typically show a correlation in their time-varying firing rates and even in their spike-firing times. Thus, the coding mechanism utilized by a cell population to extract stimulus information cannot be inferred from the activities of different neurons recorded at different times. The role of these correlated firings in the coding of complex sound is not fully known. Coincident

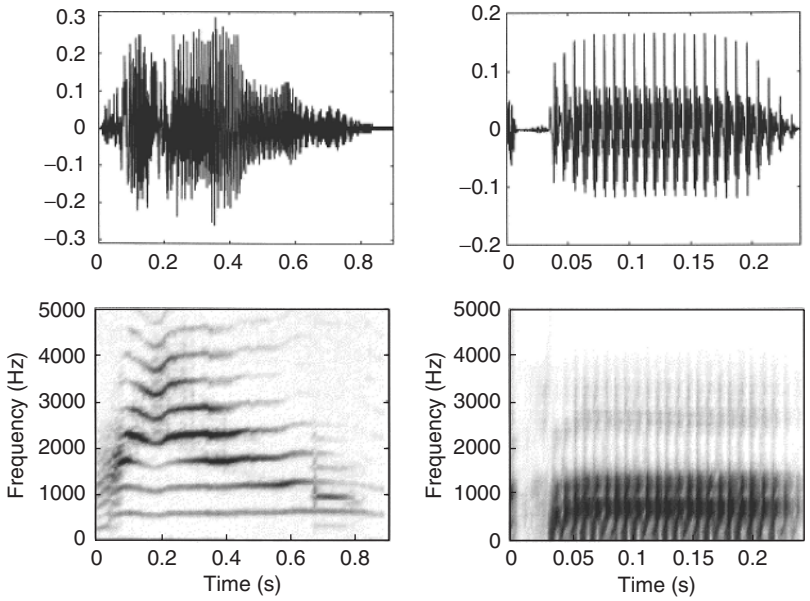
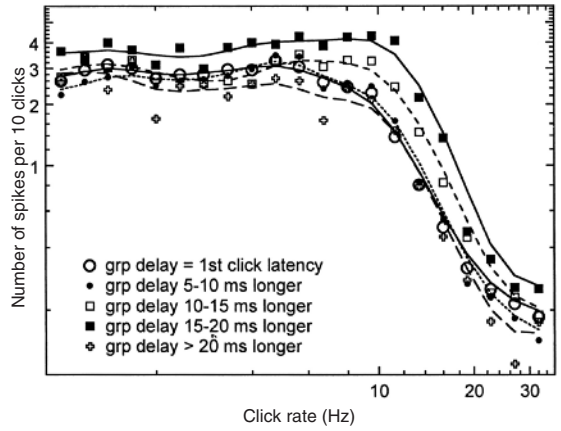
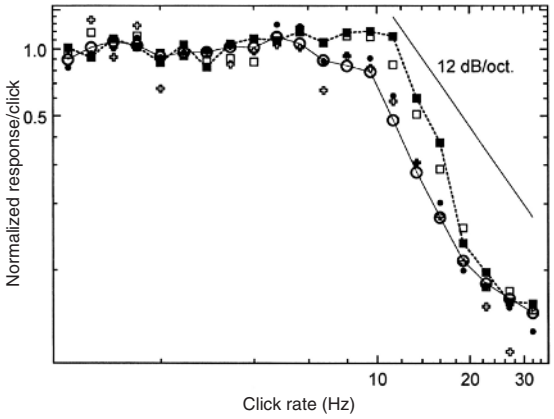


Figure 1.9 Two vocalization sounds that illustrate similarities and differences in binding features. In the left-hand column, the waveform and spectrogram of a kitten meow are presented. The average fundamental frequency is 550 Hz, and the highest frequency component (not shown) is 5.2 kHz. Distinct downward and upward frequency modulations occur simultaneously in all formants between 100 and 200 ms after onset. The meow has a slow amplitude modulation. In the right-hand column, the waveforms of a /pa/ syllable with a 30-ms voice-onset time (VOT) and its spectrogram are shown. The periodicity of the vowel and the VOT are evident from the waveform. The fundamental frequency (i.e., the periodicity) started at 125 Hz and remained at that value for 100 ms and dropped from there to 100 Hz at the end of the vowel. The first formant started at 512 Hz and increased in 25 ms to 700 Hz, the second formant started at 1019 Hz and increased in 25 ms to 1200 Hz, and the third formant changed in the same time span from 2153 to 2600 Hz. The dominant role of the periodicity in binding of frequency components is noted. (Reprinted from *Hearing Research*, Vol. 157, J. J. Eggermont, Between sound and perception: Reviewing the search for a neural code, pp. 1–42. Copyright © 2001, with permission from Elsevier.)

firings that frequently occur without concomitant firing rate changes (such as in the neural response to the steady-state portion of a pure tone, which can show the same firing rate as under silence but with increased neural synchrony between pairs of neurons [205]) can in principle be detected by depressing cortical synapses [819]. These synapses have an initial high probability of transmitter release and act as low-pass filters that are most effective at the onset of presynaptic activity and respond most vigorously to transient stimuli and to slow modulation envelopes. These synapses are responsible for the low-pass properties of temporal modulation transfer functions (Figure 1.10) as measured electrophysiologically in primary auditory cortex (A1) [246, 248].



(*a*)



(*b*)

Figure 1.10 Low-pass filtering in auditory cortex neurons. Stimuli presented were 1-s-long periodic click trains and the number of synchronized spikes per click is shown here as a function of the click repetition rate. (*a*) Group averages are distinguished by group delay as determined from the phase repetition rate dependence. This plays only a modest role, except that neurons with large group delays show a slightly higher cutoff rate compared to those with group delays below 15 ms. (*b*) Various curves are normalized to their mean response between 1 and 4 Hz. (Reprinted from (246), with permission. Copyright © 1999, *Journal of Neuroscience*, by the Society for Neuroscience.)

It has been predicted [485], and shown recently in the avian forebrain [481] and in vitro [759], that correlated neural activity is capable of propagating through cortical structures without diminishing in strength and with preserved temporal precision. This would facilitate grouping across distinct cortical fields and the formation of interarea neural codes. This is reminiscent of the theory of synfire chains [4, 920], which require this property.

Observations That Favor Role of Coincident Firings in Neural Coding. In the primary motor cortex (M1) of behaving macaque monkeys, correlated neural firings play a significant role in coding movement direction [601]. The information carried by neural interactions using a simultaneous recording from 12–16 neurons during an arm-reaching task was investigated. Pairs of simultaneously recorded cells revealed significant correlations in firing rate variation when estimated over 600-ms time intervals. This covariation was only weakly related to the preferred directions of the individual M1 neurons estimated from their maximal firing rate. Interelectrode distance had no significant effect either. In some of the cell pairs, the strength of the neural correlation varied with the direction of the arm movement. Prediction of the direction was consistently better when correlations were incorporated as compared to one based on the average firing rate of presumably independent neurons. Thus, neural interactions quantified by correlated activity carried additional information about movement direction beyond that based on the firing rates of the individual neurons. The correlated neural activity was also much higher for a planned sequence of movements compared to the same movements when executed independently by the monkey, although the firing rates were the same in the two conditions [360]. Simultaneously recorded activities of neurons in M1 of monkeys during performance of a delayed-pointing task showed that accurate spike time synchronization occurred in relation to stimuli and movements and was commonly accompanied by discharge rate modulations but without precise time locking of the spikes to these external events [760].

In primary somatosensory cortex (S1) of the anesthetized cat, stimulation of the front paw with an air jet resulted in neuron pair correlograms (see examples in Figure 0.2) with much sharper peaks than observed without stimulation [776]. The incidence and rate of stimulus-induced synchronization decreased with the distance between the recording sites. These results suggest that neuronal synchronization measures may supplement the changes in firing rate that code intensity and other attributes of a tactile stimulus.

The synchronous firing in the secondary somatosensory cortex (S2) of three monkeys trained to switch attention between a visual task and a tactile discrimination task increased in up to 35% of the pairs tested and so did the firing rates, however without a significant correlation between the changes in firing rate and changes in synchrony [854].

Cells in cat primary visual cortex showed enhanced orientation discrimination by including the synchronization of the firings between two to six cells in addition to their firing rates [787, 788].

Pairs of neurons recorded with electrodes in different auditory cortical areas showed a fourfold increase in firing synchrony during stimulation with tones or noise compared to silence combined with modest increases in firing rate [247]. Neural synchrony in rat auditory cortex also increased in a delayed go/no-go task, a task where one stimulus required a behavioral response after some prescribed time and the other one did not, but specifically in the waiting period [916].

Observations That Argue Against Role of Coincident Firings in Cortical Neural Coding. In V1 of the awake monkey, neural synchrony was observed between neurons with distant RFs in response to textured "figure–ground" stimuli. However, there was no difference in synchrony between pairs with both RFs overlapping the "figure" part and pairs in which one or both units had RFs within the "background" part of the stimulus. Thus, no evidence was found for a role of neural synchrony in the binding of those features that lead to texture segregation [521]. In a coherent motion detection task, the neural synchrony in awake monkey visual field MT was actually lower than for noncoherent conditions [879] and thus not likely to play a role in binding of motion by synchrony. Pairwise correlation strength for units recorded on the same electrode in MT of the behaving monkey was independent of the presence of visual stimulation and the behavioral choice of the animal [53]. Rolls et al. [770] and Aggelopoulos et al. [9] also found little gain of stimulus-dependent synchronization on the information available about the stimulus in the neuronal firing rate in inferior temporal visual cortex.

Simultaneously recorded firings from 30–40 neurons from three somatosensory cortical areas were able to predict the type of stimulus regardless of whether the trials were shuffled for each single neuron [659]. This suggests that precise timing information between those neurons was irrelevant. In secondary somatosensory cortex (S2) of anesthetized cats, Alloway et al. [15] found no evidence that synchrony played a role in the coding of the direction of movement of a tactile stimulus. Similarly, in rat barrel cortex, synchronized firing did not contribute to coding the stimulated whiskers [714]; coding was instead solidly based on first-spike latency.

A similar absence of change in correlation strength with increased auditory stimulation level was reported for units recorded on separate electrodes in A1 of the anesthetized cat [247]. Thus neural synchrony likely does not code for stimulus level.

Hence, it appears that in the early stages of motor and sensory cortical processing (M1, S1, V1, A1) neural synchrony may play a greater role than in later stages (S2, MT, IT). We return to this issue later in our discussion of the role of synchrony in feature binding via bottom-up versus top-down attentional processes.

1.4 CORRELATION IS THE BASIS OF NOVELTY DETECTION AND LEARNING

It has been already suggested that coincidence detection by rather broadly tuned neurons may result in sharper tuning or greater specificity for particular stimuli [57]. This can be obtained either by a simple convergence of two neural activity patterns on a coincidence detecting neuron [250, 886] or by strengthening the direct connections between simultaneously active neurons. The latter mechanism has been postulated for the creation of sharply tuned neural assemblies [377], secondary repertoires [238], and synfire chains [3].

Neural Assemblies. Hebb [377] has pointed out that there are two extreme views of neural assembly action. One was called the *switchboard theory*: The cortex is considered as an elaborate kind of telephone exchange with precise connections; the other was called the *field theory*, which regards the cortex as an aggregate of cells forming a statistically homogeneous medium with mostly random connections. An example of a switchboard theory was presented by Ballard [55]. Examples of field theories are those by Beurle [86], Cowan [190], Griffith [339], and Hopfield [399], to mention a few. Hebb's own assembly model was somewhat intermediate in assuming that precise connections existed but with modifiable synapses that could be changed by experience. An elaboration of such an assembly theory was presented by John [444] in what he called a "statistical configuration theory." In this theory, learning and memory are envisioned as the establishment of a representational system of a large number of neurons in different parts of the brain. The activity of these neurons will be affected in a coordinated way by the spatiotemporal characteristics of the stimuli presented during the learning task. This was assumed to initiate a common mode of activity in various brain regions specific for that stimulus. Information about an event is represented by the average behavior of such a responsive neural ensemble. Another event can be represented by the same ensemble but with a different correlation pattern.

A big leap in the concept of neural assemblies was made by von der Malsburg [922] by proposing the following description: "a cell assembly is a set of neurons cross-connected such that the whole set is brought to become simultaneously active upon activation of appropriate subsets which have to be sufficiently similar to the assembly to single it out from overlapping others." Thus, given suitable input, the assembly can be ignited and then acts as a logical unit by going through a spatiotemporal activity pattern characteristic for that assembly. The ignition character of an assembly is also evident in the concept of the synfire chain [3, 4]: "the activity of the neurons that transmit information is organized along a chain of sets of neurons. Each link in the chain is made of a set of neurons that fire in exact synchrony whenever the chain becomes active." The concept of neural assembly also includes the necessarily hierarchical character of the organization and is related to the concept of repertoires [238] defined such that "the main unit of function and selection in the higher brain is a group of cells connected in various ways. Groups of cells build repertoires."

Neural assemblies have more recently been defined as "a group of neurons [that are] at least transiently working together as indicated by correlation of unit activity" [316]. In visual cortex, cells with approximately 0.5-mm separation showed the highest correlation among cells with similar RFs and similar connectivity from the LGN [522]. This suggests that overlapping or shared connectivity is a dominant factor in neural assembly formation. It is common to think about a neural assembly as widely distributed in cortical space, potentially extending over various subdivisions of cortex [838]. For instance, connections over large spatial divisions of auditory cortex are provided by the thalamic cell axonal divergence and convergence, often estimated to be between 2 and 5 mm at the cortical level [536] and intracortically through horizontal fibers [932] that can range up to 8 mm. In visual

cortex, the spatially periodic effects of the patchy connections of these horizontal fibers have been shown by cross-correlation [893]. These cortico-cortical connections are for a sizable part heterotopic. In auditory cortex, they connect cell groups with characteristic frequencies (CFs) differing by more than one octave [537]. In visual cortex, the horizontal fibers connect cell groups without spatial RF overlap but with similar orientation tuning.

Neural assembly membership is expected to be stimulus dependent and context specific and may reflect the number and functional strength of its common inputs under different conditions [903]. It is however likely that at any point in time several spatially overlapping neural assemblies are active. In response to external events, a group of neurons forming a dynamical cell assembly may spontaneously organize itself temporarily by correlated firing of their spiking activity.

Neural assemblies, thus defined, may potentially be probed using microelectrode arrays that allow recording from a set of relatively widely spaced neurons. These neurons could participate in one or more neural assemblies. The quantification of the correlation in spiking activity occurring between pairs of such widely spaced neurons thus becomes crucial in defining membership of neural assemblies. The stimulus will be one of the dominant sources of neural correlation because of its common input character. Although it is common to correct for stimulus-induced correlations by using shift predictors or joint peristimulus time histogram (JPSTH) techniques [244], the brain does not have that luxury but may exploit this stimulus-dependent correlation to change the extent and structure of the neural assemblies.

Secondary Repertoires. The selection theory of brain function [238–240] assumes that after ontogeny and early development the brain contains cellular configurations (groups) that can already respond in a discriminatory way to sensory stimuli (e.g., the orientation selectivity in the visual system of newborn monkeys), because of their genetically determined structures or because of epigenetic alterations that have occurred independently of the structure of these sensory signals. This prespecified collection of neuronal groups is called a primary repertoire and consists of a large number of groups (of the order of 10^6), each with a modest number (50–10,000) of cells. The primary repertoire is degenerate, that is, it contains multiple neuronal groups, with different internal structures, that are capable of carrying out the same function. The primary repertoire should contain enough neuronal groups such that sensory signals have a high probability to find matching groups; and finally it must have provisions for amplifying a selective recognition event, probably by synaptic alterations, either through the formation of new synapses or through changes in already existing contacts. All these properties are very much the same as in the classical perceptron [773].

In addition, the neuronal group selection theory requires a secondary repertoire as a collection of different, higher order neuronal groups whose internal and external synaptic connectivity can be altered by selection during experience. This cell group selection can occur in two stages, first by filtering—selecting all groups that react more or less well to the spatiotemporal input pattern—and second by an inhibition process (a threshold mechanism) that eliminates those selected groups from stage

1 that have an insufficient response. An important aspect of the theory is the reentrance of signals at the level of the secondary repertoire. The dominant cell type, the pyramidal cells in cortex receive far more collaterals from other pyramidal cells (>99%) than from specific afferents (<1%). Thus, the cortex is to a large extent a thinking machine working on its own output [110]. For the pyramidal cells there is probably no way to distinguish these inputs. Thus, internally generated signals are reentered as if they were external signals; this reentrance might be able to guarantee a continuity in the neural construct as well as a succession of temporal order of associated memory events. Moreover, reentrant signals in the secondary repertoire might be of a different modality from the more direct input from first repertoire groups. Thereby the secondary repertoire might be able to relate multimodal activity patterns. Reentrance is also a mechanism that allows ongoing cross-correlations between inputs from first repertoire groups and second repertoire groups, thereby providing the possibilities of association and classification [845]. The role of STDP in the neuronal group selection theory has been elucidated in [435].

Learning. Learning is a process that generates a "brain" that is different from that prior to the learning process. The results of learning are memories. Memories are likely laid down in spatial patterns of synaptic connectivity. Novelty detection is related to memory recall; the evidence for the existence of novelty detectors comes among others from evoked potential studies where an "oddball" stimulus creates large activity in the latency region beyond 100 ms following the stimulus, called the *mismatch negativity* (MMN) [646]. A novelty detector is most likely a neuronal group and not an individual neuron. Whenever information is processed through various assemblies, familiar input will activate already formed functional connections in a neuronal group. An acceptable match between incoming information and stored information is based on correlation and will generate either an inhibitory signal to arousal centers or no signal at all. Whenever there is a mismatch, a signal is sent to the arousal centers whose nervous activity, or activity induced by the arousal center in the sensory cortex, may result in detectable evoked potentials. Novelty detection then in this view is the result of a "template-matching" procedure that has to be carried out in parallel for a very large number of potential templates: In principle all information that does not match with the templates has a novelty character.

In monkeys, MMN-like activity corresponded to spikes generated in superficial layers of primary auditory cortex. When NMDA receptors in the auditory cortex were blocked, the MMN disappeared [439]. In analogy, administering the anesthetic ketamine, which is an NMDA blocker, to human volunteers also abolished the MMN [508]. This suggests that the MMN requires NMDA receptors involved in the formation of memory traces used in the novelty detection. This is not surprising as NMDA channels have also been linked to LTP and working memory [564].

Topographic and Functional Brain Maps. Braitenberg [110] came to the conclusion, on anatomical grounds, that the probability of two neighboring pyramidal cells in cortex being interconnected by one synapse was about 10%. Experimentally, this was confirmed much later. Dual intracellular recordings from

neighboring pyramidal cells were used to assess the probability of synaptic contact between randomly selected neuron pairs within 250 μm distance from each other [880] as 1 : 11 (layer III to layer V), 1 : 21 (layer III to layer III), 1 : 41 (layer V to layer V), and 1 : 86 (layer V to layer III). Thus, most of the pyramidal cells are not interconnected, and if they are, the connection is very weak (by only one synapse out of the 5000–10,000 located on each pyramidal cell) compared to the estimated firing threshold of a pyramidal cell, which is about 10–30 for time-correlated input and more than 300 for asynchronously arriving inputs [4]. Thus, correlated input activity keeps the brain going. In rat somatosensory cortex, Bruno and Sakmann [124] measured in vivo the excitatory postsynaptic potential evoked by a single thalamocortical synaptic connection and confirmed its low efficacy. They also demonstrated that the thalamocortical inputs are numerous and synchronous, thereby confirming the suggested synchrony mechanism by which thalamic inputs alone can drive cortex.

Most correlations of spiking activity in the cortex are through shared specific afferent (i.e., thalamocortical) input and therefore serve to functionally link neurons with overlapping STRFs. What the experimenter observes as topographic maps (e.g., retinal position, sound frequency, body surface) does not have meaning for the brain itself based on this spatial organization but only through the correlations in the neuronal firings. Thus, activity in one cell can just be spontaneous ("noise") and will be very difficult for other neurons to distinguish from stimulus-induced activity. Correlated activity in two or more neurons that project to another neuron potentially signifies outside world activity if their firings are correlated. Furthermore, as described above, correlated inputs are more likely to fire the receiving cell.

As discussed earlier, two types of neuronal maps are found in the cortex: *topographic maps* and *functional maps*. Topographic maps arise from the anatomical structure of the receptor surface and from the afferent nerve fibers preserving this orderliness in the nerve fiber tracts and in each interposed nucleus. Receptor surface maps may undergo quite complex transformations, resulting in topographic map deformation. An example is the complex logarithmic transformation between retinal surface and visual cortex surface (e.g., [807, 808]) or superior colliculus map (e.g., [694]). Topographic maps can also result through computations carried out in the brain. A prime example is the map of auditory space in the midbrain [492], which is based on a combination of interaural differences in arrival time, phase differences in continuous sounds, and intensity differences at the two ears as represented centrally through differences in excitation and inhibition [345], but in reference to a topographic map of visual space in the same structure (e.g., [491]). To the experimenter there is no difference between the computational map of auditory space and the topographic map of visual space: Both show a spatially ordered structure. To the animal there is no difference either: The internal order of both is present in the correlation structure. Thus, in the brain all maps are *correlation maps* and there is no real way to distinguish between the two types. Correlation maps represent signal-associated aspects, both sensory and motor, as encoded by individual neurons or small neuronal groups. Topographic maps are only obvious to the investigator, who relates neural activity to the stimuli presented.

Topographic maps are organized on the basis of their sensory receptor or effector surfaces and are therefore likely to be found at the sensory and motor sides of the brain. Functional maps have mainly an organization or ordering through their correlation structure [494, 495]. Neural unit pairs with strong correlations can be considered close in the neural organization; neural units with weak correlations have a larger functional distance. Functional maps may differ for different stimulus conditions [250]. Topographic maps also have a cross-correlation structure, but this is mainly the result of the anatomical ordering and spectrotemporal or spatiotemporal overlap of RFs (Figure 1.11). Here, the principle of obtaining coincident-spike STRFs is illustrated. Typically spikes in two neurons that contribute to their potentially overlapping STRF (i.e., occur within the time–frequency domain of their STRF) are more strongly correlated than those that do not contribute. As can be

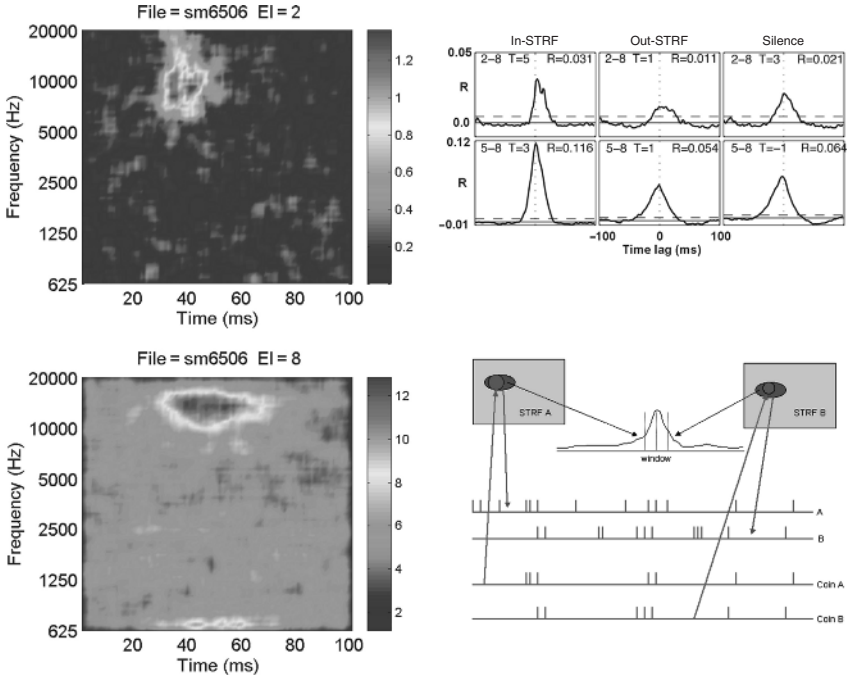


Figure 1.11 Correlation structure of overlapping STRFs in auditory cortex. Left column shows two simultaneously recorded STRFs with partial overlap in the time–frequency domain. The upper STRF shows a responsive area between about 6 and 20 kHz with a CF slightly below 10 kHz and extending in time between 30 and 50 ms after stimulus onset. The lower STRF is generally between 10 and 17 kHz and 30–70 ms and a CF at about 13 kHz. Upper right panels show cross-correlograms between spikes that contributed to the STRFs (In-STRF), between spikes that did not contribute to the STRFs (Out-STRF), and for comparison from the same neurons under spontaneous firing (silence). The right bottom panel shows in cartoon form how coincident-spike STRFs are constructed. Original spike trains in red, only spikes that occur within a given time window from those in the other neuron, and considered as coincident spikes, are shown in green.

seen from the cross-correlograms, spontaneous activity during silence in turn is more strongly correlated than "spontaneous" spikes, that is, those not contributing to the excitatory part of the STRF, during continuous stimulation. Selection of a certain coincidence window allows the construction of STRFs for coincident spikes only.

In Figure 1.12, we show comparisons of STRFs for four sets of recordings comprising activity from four to seven separate electrodes (indicated in the bottom panels) under the conditions of only coincident spikes (within a 10-ms window, top row) and all spikes merged (bottom row). Here, N is the number of recording sites (typically out of 16) that belong to the same cluster and whose spike trains are used in constructing the all-spike and coincident-spike STRFs. It is clear that only in some cases the coincident-spike and all-spike STRFs are similar. For instance, in the third column, where the two STRFs show the highest similarity, the contour of the excitatory part of the coincident-spike STRF overlaps with the all-spike STRF, but differences still exist for the inhibitory parts. The largest differences between coincident-spike and all-spike STRFs are shown in the second column. Here, the coincident-spike STRF forms a small subset of the all-spike STRF but overlaps with its most responsive part. Note the large differences in the inhibitory parts of the two STRFs.

This suggests that different computations can be carried out by the cortex depending on the mode of action of the downstream cells: coincidence detection or temporal integration. Whether neocortical pyramidal cells are acting as coincidence detectors or temporal integrators depends on the degree of synchrony among the synaptic inputs [333, 779]. High-input synchrony leads to the more efficient

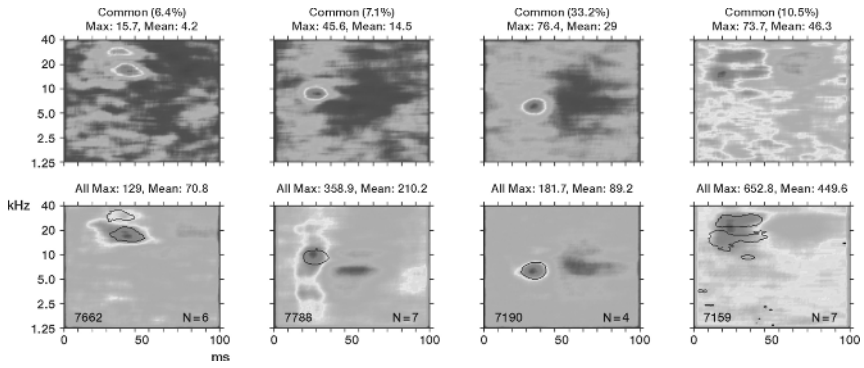


Figure 1.12 Four examples of coincident-spike STRFs (top row) compared, by superimposing its contours, to the all-spike STRFs (bottom row). In the bottom row the number of STRFs that contributed to both types of STRF are indicated (e.g., $N = 6$). Peak STRF values and mean level of activity are indicated at the top of each panel, which give an indication of the different signal-to-noise ratios (SNRs) of the two types of STRF. The percentage of the spikes that contributed to the coincident-spike STRFs is displayed above each set of panels. Warm colors (yellow and red) indicate excitation, cold color (blue) indicates inhibition. Green is typically neutral. (Reprinted from (250), with permission. Copyright © 2006, by the American Physiological Society.)

coincidence detection, whereas low-input synchrony leads to temporal integration. The asynchronous background spikes from other cortical cells could also play an important role in setting the processing mode of the cell [8].

Functional maps are likely to be found in the association areas between sensory and motor areas. However, Vaadia et al. [904] found that about 15% of neurons in primary and secondary auditory cortex of behaving monkeys showed sensorimotor association properties. Besides a strict sensory component in their response, these units showed task-dependent activity as well. Since these units were found widely dispersed throughout the auditory cortex, they suggested that association cortex might overlap with sensory cortex. If this conjecture is correct, it suggests the coexistence of topographic and functional maps in the same cortical structures. Neurons from one type of map may simultaneously take part in another type of map.

Synchronous or correlated activity is the only way in which sensory events propagate through the brain and manifest themselves to the brain; it will also be the way in which such events are remembered. Thus, in a more general sense, simultaneous activity of neurons is also related to memory recall. Activity in representations of different sensory modalities may be correlated on a larger timescale and can be an indication of events in a more general context. The relation between topographic maps of different modalities can be maintained in a functional map. This functional map then constitutes a memory of a situation in which such a multimodal stimulus was present. Later, this map can be activated again and be used to complete a certain pattern of activity or to predict the temporal sequence of a stimulus pattern.

How is such a functional map formed? One may assume that the correlations between the occurrences of events that together form a situation can be stored in terms of connections between the neural units that represent these events. The question can be raised whether there is a sufficient number of connections to store all such items. Since the neocortex has about 10^{11} neurons and about 10^{15} synapses, there seems to be ample space to do so. Functional maps are only characterized by a "correlation structure" among firings from different neurons. They have been attributed as resulting from neural assemblies [316], neural clusters [179], and neural cliques [556]. Functional maps have not yet been visualized by electrophysiological techniques [250].

1.5 CORRELATION IN SENSORY SYSTEMS: CODING, PERCEPTION, AND DEVELOPMENT

In many layered brain structures, spontaneous and highly synchronized neural activity can be found. One of these areas is the hippocampus, where theta wave (4–8-Hz) activity is found that seems to reflect the initiation of purposive movement patterns such as self-generated motion through space [98, 128]. There seems to be less and more localized correlated firing of neurons in the neocortex during active processing compared to the sleep state, where many single units show a high degree of correlation in their firings [213, 663]. Thus, correlated activity does not always accompany active behavior. It could be postulated that during sleep cells in the

brain are recruited into large assemblies that generate reverberant activity in the slow-wave or spindle frequency region. Its purpose could be to stabilize memories [856]. We return to these issues in our discussion of correlation in memory systems.

Correlation in Perceptual Coding. Barlow [60] proposed that a major goal of perceptual coding is to produce a minimally redundant neural code, which should facilitate subsequent learning. The information about an underlying signal of interest, such as the visual form or the sound of a predator, may be distributed across many input channels, making it difficult to associate particular stimulus values with distinct responses. A neural code having minimal redundancy should alleviate this problem. If the encoding of the sensory input vector into an N-element feature vector has the property that the N elements are statistically independent, then all that is required to form new associations with some event V (assuming the features are also approximately independent conditioned on V) is knowledge of the conditional probabilities $p(V|y_i)$ for each feature y_i, rather than complete knowledge of the probabilities of events conditional upon each of all combinatorially possible sensory inputs. Atick and Redlich [47] proposed a cost function for Barlow's principle that minimizes the power (redundancy) in the outputs subject to a minimal information loss constraint. Under conditions of high noise (low redundancy), the RFs that emerged were Gaussian-shaped spatial smoothing filters, while at low noise levels (high redundancy) "ON-center OFF-surround" RFs resembling second-order spatial derivative filters emerged. In fact, cells in the mammalian retina and LGN of the thalamus dynamically adjust their filtering characteristics as light levels fluctuate between these two extremes under conditions of low versus high contrast [825, 918]. Moreover, this strategy of adaptive rescaling of neural responses has been shown to be optimal with respect to information transmission [111].

While on the one hand the sensory periphery may strive to remove correlations by lowering redundancy in the neural code, on the other hand there is ample evidence that the brain generates correlated signals for many different purposes, for example, in order to synchronize large-scale networks involved in attention and memory and to provide meaningful commands to the motor system.

Correlation and Its Role in Sensory Development. Genetic information is generally insufficient to provide for more than a crude topographic wiring of the receptor surface onto the cortex in the form of retinotopic maps in visual cortex, tonotopic maps in auditory cortex, and body surface somatotopic maps in the somatosensory cortex. More elaborate topographic orderings, such as those for orientation sensitivity in visual cortex or auditory space, which need to be computed from information at the two ears, in midbrain structures require self-organizing processes based on spontaneous firing [826], lateral inhibition [133], and the ability to change the connection strengths (i.e., that of synapses) between neurons based on correlation of incoming and existing activity [125]. Most rules for modifications in synaptic strength assume that the change is proportional to the change in the correlation or covariance of two neural activity patterns [241]. Evidence for the necessity of temporally correlated activity and the formation of

and changes in somatosensory maps [174], tonotopic maps [56, 57, 998], and visual maps [573] is now provided in abundance. In many of these cases the maps and changes therein are guided by experience.

A similar phenomenon may serve to align topographic maps of space for different sensory modalities. Knudsen [491] has shown that auditory and visual maps of young barn owls raised with one ear plugged are in register although the binaural input is distorted. Removing the earplug resulted in a shift of the auditory map of space with respect to the visual one. Apparently, an alignment of the representation of the auditory and visual space occurs in the midbrain (optic tectum) during the early periods of life. Since auditory responses are mainly found in the deep layers of the optic tectum, that is, the visual midbrain, where visuomotor units are also found, it is suggested that localizing sound (auditory) together with the sound sources (visual) results in the alignment of auditory and visual spatial maps. This requires either multimodal neurons [615] or separate sets of auditory and visual neurons with some form of neural interaction between them [778].

The appraisal of the outside world by our senses is done in parallel fashion and results in neural activity patterns organized in topographic maps of the brain. Correlation of nervous activity takes place at many levels. First of all there is the single neuron level. Neurons that have overlapping RFs, defined as the set of sensory receptors that significantly affect its rate of firing, will show a covariance in overall firing rate as well as a coincidence in the occurrence of spikes [250, 886]. Thus, a visual RF is an area on the retina, a somatosensory RF is an area on the body surface, and an auditory RF is a range of sound frequencies.

In the auditory system, overlap of STRFs or a difference in CF, the most sensitive frequency in the RF, is a strong indicator for neural correlation [117, 244, 250] with STRF overlap explaining nearly 40% in the variance of the peak correlation (Figure 1.13). The STRFs represent both the frequency range of sensitivity and the temporal window in which neural activity is elicited. The visual equivalent is the spatiotemporal RF combining the spatial area of sensitivity with the temporal window of evoked neural activity. Usually STRFs are measured using a broadband continuous stimulus such as dynamic ripple noise [486] or multifrequency stimuli consisting of randomly presented gamma tone pips [250]. In the example shown in Figure 1.13, tone pips for each of 81 frequencies over 5 octaves were randomly presented according to a Poisson process, with similar average rate but different realization for each frequency. The topographic mappings will assure that the strength of coincident firing will decrease with distance. Coincident firings are a subset of the firing of the neurons involved and may function to extract relevant information from the neuronal "noise" and improve the SNR [250]. In the visual cortex, neurons in different areas or even different hemispheres show correlated neural activity when they have similar orientation sensitivity or originate from the same eye [260, 502].

Another example, from the field of auditory disorders, where neural synchrony plays an important role is the phenomenon of tinnitus [380], a sensation of a hissing or ringing sound in the ear, in the absence of stimulation. The auditory system has the highest spontaneous activity of all sensory systems, yet we normally do not

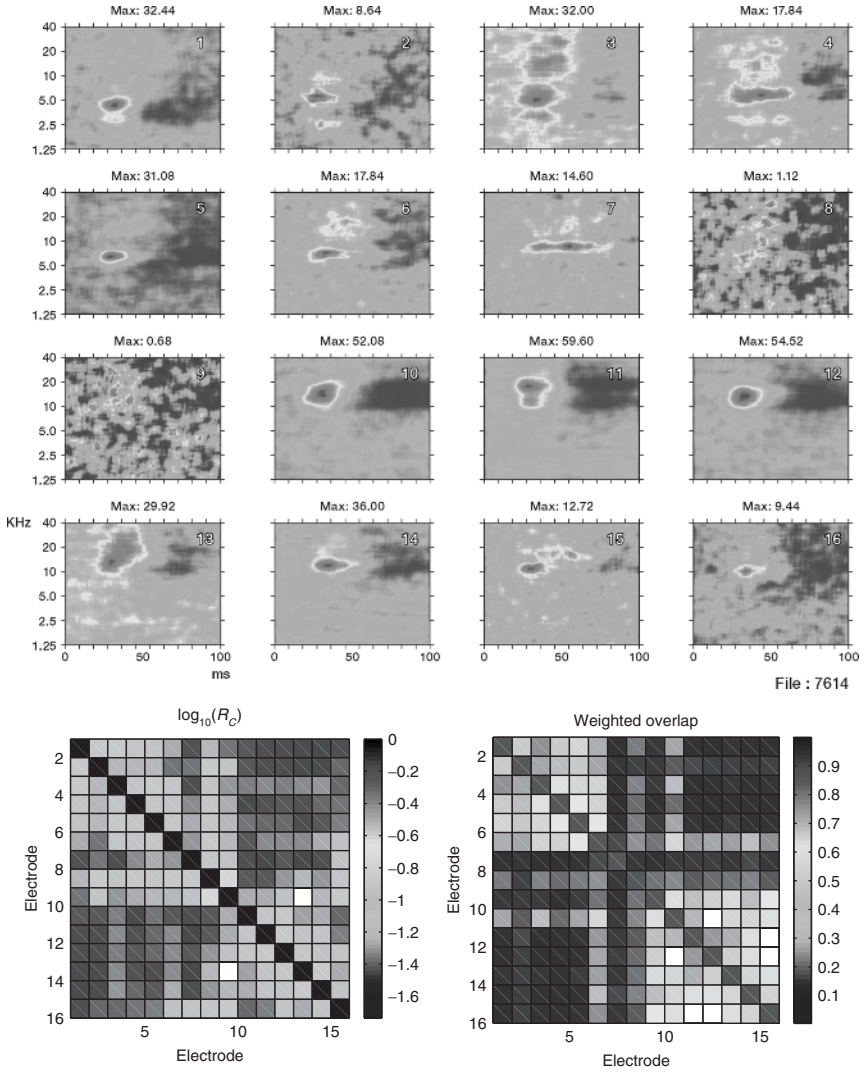


Figure 1.13 The top 16 panels indicate individual STRFs recorded simultaneously with 16 electrodes in the auditory cortex. The left bottom panel shows the $\log_{10}$ of the peak cross-correlation coefficients (R_C) between the spike trains of all electrode pairs. The right bottom panel shows the same for weighted STRF overlap between the STRFs of all electrode pairs.

hear our spontaneous activity, presumably because its firings are not correlated across auditory nerve fibers [445] and very little among nonneighbor auditory cortex neurons [242]. Tinnitus is often present following noise-induced hearing loss, something that is more and more prevalent due to the addiction of a substantial part of the population to overly loud recreational sounds. In contrast to industrial noise

exposure, the recreational noise levels are not regulated by any legislation. Animal models of tinnitus usually show a specific reduction in the spontaneous firing rate of auditory nerve fibers tuned to the frequency range of the hearing loss. Tinnitus can be understood on the basis of increased synchrony among neurons [253] in the central auditory system accompanied by increased spontaneous firing rates [463]. Computational models [220, 794] suggest that tinnitus is a byproduct of the action of relatively slow homeostatic mechanisms that tend to stabilize firing rates among neurons [896, 897]. When these homeostatic mechanisms operate at the network level, by upregulating lateral excitation and downregulating lateral inhibition, such models then generate not only increased spontaneous activity but also increased synchrony between neuronal firings [220].

In addition, at the onset of vision, mammals show ocular dominance columns in visual cortex and a segregation of input from the two eyes in different layers in the visual thalamus (i.e., LGN). However, long before onset of vision this layer separation does not exist and cells in the LGN receive input from both eyes. By blocking the formation of action potentials in the retina using tetrodoxin, which blocks sodium channels needed for action potential initiation, the segregation of individual eye input in individual LGN layers is prevented. Similarly, by electrically stimulating both optic nerves synchronously, the segregation is also prevented. Finally, after blocking activity initiation in the retina by tetrodoxin but electrically stimulating the individual optic nerves asynchronously, the segregation of eye input does occur. This all suggests that neural activity that is correlated in individual retinas but asynchronous between retinas is what drives the formation of ocular dominance layers in the LGN, and similarly the formation of ocular dominance columns in visual cortex. This correlated neural activity has to be generated in the retina before the onset of vision (reviewed in [826]).

Correlative Activity in Prevision Retina Shapes Ocular Dominance and Orientation Columns. In a series of studies by Shatz and colleagues, by using a hexagonal sized multielectrode array with 61 electrodes spaced 50 μm apart, recordings were made from up to 100 individual ganglion cells showing bursting periods of about 5 s long and interspersed with 1–2 min of silence. For neighboring electrodes the bursting occurred synchronously, that is, the cross-correlogram of the action potential firings was centered around zero and showed a peak with a half time of about 2.5 s. For more distant electrode pairs the correlogram peak was displaced by a value commensurate with diffusion of an excitable substance at a speed of about 0.2–0.6 mm/s [127, 613]. This results in a wavelike activity pattern that spreads across the retinal surface; the pattern can start anywhere and propagate in any direction. The determining factor is in what direction the local percentage of recruitable cells, that is, those that are not refractory, is large enough to participate in a wavelike activation. This required percentage for wave propagation turned out to be about 30% in modeling studies [127]. The other crucial aspect is a fairly long refractory period (here about 1–2 min). Crucially, these propagating waves disappeared at the onset of vision and the strength of the correlated activity also decreased over the 30-day period that the waves were observed in

ferrets by more than a factor 10. In adult ferrets no correlation was detected using these retinal electrode arrays. Therefore, ocular dominance columns are dependent on highly synchronized activity within but not between retinas. This bursting pattern in developing animals may also act to reinforce the topographic accuracy of retinal projections based on the exponentially decreasing peak correlation with distance between ganglion cells. A correlation-based mechanism (i.e., a Hebbian one) in the LGN and visual cortex would favor proximate ganglion cells to connect to the same area in LGN. The initially rough retinotopic projection might thus become gradually more refined with age [973]. This story suggests that correlation of neighboring ganglion cells in the retina ceases to exist once the developmental period is finished. This is not so (reviewed in [612]) because at least three types of correlations can be found between retinal ganglion cells in adult animals. Very narrow correlogram peaks (<1 ms wide) are found between neighboring ganglion cells that share gap junctions (electrical synapses). Intermediate-width correlogram peaks (2–10 ms) are found for those ganglion cells that share gap junction input from the same amacrine cell. This is the most prevalent type. Finally, ganglion cells might share common input from the same bipolar cell through standard chemical synapses and this is reflected in correlogram peaks that are 40–50 ms wide. The shared activity from amacrine cells may encompass ganglion cell clusters over a distance of up to 0.4 mm. In this case amacrine cells provide most of the overlap of the spatiotemporal RFs of the ganglion cells. Each ganglion cell will receive input from many amacrine cells, which typically have small RFs. Coincident firings between two ganglion cells require precise timing, which likely only happens when they get input from the same amacrine cell. For a population of ganglion cells to fire in coincident fashion, they need to have at least one amacrine cell in common. A coincident-spike RF then could reflect that of this particular amacrine cell [796]. Thus, it is likely that it is only the wave pattern of synchronous activity that disappears after eye opening. Orientation columns and directional sensitivity are also formed before eye opening but require subsequent visual activity to maintain these higher order RF properties. That the driving force again is locally synchronized activity in the retina or optic nerve will not be a surprise. Both rearing in darkness and rearing in the absence of visual contrast result in weakening of the individual cortical neuron's selectivity for stimulus orientation and movement direction. Inducing a global synchrony, by electrically stimulating the optic nerve periodically does the same: The individual neuron's orientation and directional responses are greatly reduced. And although orientation columns can still be demonstrated using optical imaging, they lack the normal salience [947].

Formation of Tonotopic and Retinotopic Maps. Higher than normal pair correlations thus appear to be present during map development. When they are local, their effect is beneficial, whereas when they are global, it generally disrupts normal map formation and neuronal properties. Another example where local pair correlations decrease in strength with age is found in area A1 of the cat. Here, the peak strength for neighboring unit pairs recorded on the same electrode remained high up to postnatal day 50 (P50) and then decreased with age [242]. This decrease

starts around the same time that frequency-tuning curve bandwidth starts to increase in dorsal and ventral parts of A1 [103] and when the percentage of γ-amino butyric acid (γABA)−ergic neurons in A1 starts to decrease [306]. The tonotopic map in A1 does not appreciably change from P14 until P50 albeit that its gradient (in kilohertz per millimeter) decreases, likely because of cortical growth in that time period [103]. Thus, taking into account that the cat does not hear external sounds of less than 90 dB sound pressure level (SPL) before P9 [245], it appears that there is not much refinement in the tonotopic gradient except for the increases in the bandwidth of frequency tuning in noncentral parts of A1. In subcortical structures such as the inferior colliculus the frequency-tuning curves bandwidth decreases by half over the same period. Thus these changes in frequency-tuning curve bandwidth are occurring in the auditory thalamus and/or the auditory cortex. In visual cortex, a reasonably accurate retinotopy is present at birth in monkey and at the time of eye opening in the kitten, again suggesting that visually evoked activity does not play a major role here [958]. Topographic maps, such as the tonotopic map in A1, can be affected by abnormal neural synchrony during development. Zhang et al. [998] introduced not normally occurring synchronous input to many neurons in the auditory pathway by exposing rat pups to pulsed white noise during the period between P9 and P28. They found a disruption of the individual cortical neuron's frequency selectivity and a disruption of the tonotopicity. A critical period for this effect was suggested by the fact that the same stimulation after P30 did not produce any effect.

Models of Columnar Development Based on Correlated Firings. Columnar organization in visual cortex comes in many shapes. Roughly a column is a three-dimensional structure that cuts perpendicularly through the cortical layers and keeps the same cross section in each layer. The simplest column is cylindrical in shape and is usually called a minicolumn (see Figure 1.5). It is thought to reflect the retinotopic organization in cortex so that a small region in the retina is mapped onto a small region in cortex. The cross section of such a column is also called the RF. Neighboring areas in the retina have neighboring minicolumns in visual cortex giving rise to a retinotopic organization. Spatial RFs like the retinotopic ones typically undergo a characteristic distortion when mapped upon, for example, the superior colliculus and the striate cortex. Schwartz [807–810] has described the mapping on V1 as a logarithmic conformal mapping:

$$w = B \ln(z + A), \tag{1.16}$$

where $w = u + jv$ and $z = x + jy$ represent the cortical and retinal complex coordinates, respectively. The logarithmic conformal mapping maps foveal parts of the RFs such that they are overrepresented at the expense of more peripheral parts of the retina. One of the properties of the logarithmic conformal mapping is that a multiplication of the size of the image plane—the retina—results in a translation in the projection plane—the cortex. This mapping therefore is size invariant, and angles are preserved (Figure 1.14). Thus lines intersecting at right angles on the

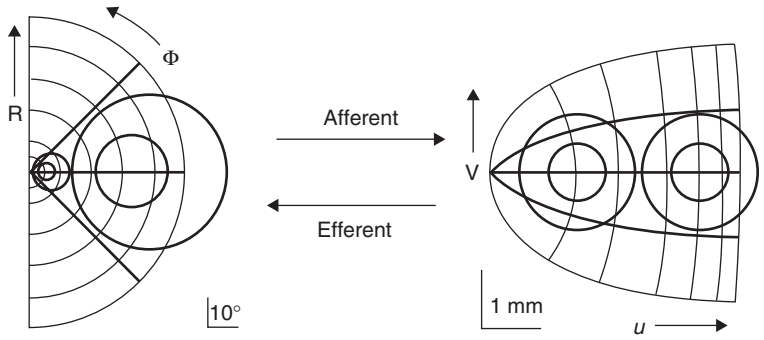


Figure 1.14 The point-to-point mapping between the external world (see half field on the left) and the superior colliculus (on the right). The mapping function used is isotropic ($B_u = 1.5$ mm; $B_v = 1.5$ mm/deg; $A = 3°$; see equations in the text). As a result, the distance from 0 to 80° eccentricity along the horizontal meridian in the colliculus map is about 5 mm. When replotted in the R, Φ coordinates of visual motor space, the corresponding circular visual (movement) fields in the superior colliculus have strikingly different sizes and noticeable skewness along the R dimension. The polar coordinate grid on the left has meridians every 45° and isoeccentricity hemicircles of 5°, 10°, 20°, 30°, 40°, and 50°. (Reprinted from *Vision Research*, Vol. 26, F. P. Ottes et al., Visuomotor fields of the superior colliculus: A quantitative model, pp. 857–873. © 1986, with permission from Elsevier.)

retina will do so in the cortical map as well. Such a mapping has also been shown to apply to the superior colliculus in rhesus monkey [694]. The mapping rule in some detail reads:

$$u = B_u \ln \sqrt{R^2 + A^2 + 2AR\cos\Phi} - B_u \ln A, \tag{1.17a}$$

$$v = B_v \arctan \frac{R\sin\Phi}{R\cos\Phi + A}, \tag{1.17b}$$

where A is the offset in degrees and R is the eccentricity in the visual motor space; B_u and B_v are two scaling constants; parameter A together with B_u/B_v determines the shape of the mapping. For numerical values see the caption of Figure 1.14.

More complex cortical columns are the ocular dominance columns, regions where input from the two eyes alternates (as introduced above). Orientation columns are defined as regions that respond optimally to a narrow range of stimulus orientations and are related to the ocular dominance columns in such a way that regions with a high spatial rate of change of orientation tend to either be aligned along centers of ocular dominance columns or intersect them at right angles.

Nearly all the neural net models for visual cortical map development, and ideally for the retinotopic, ocular dominance and orientation map combined, are based on four common assumptions: (i) Hebbian synapses, (ii) correlated and/or spatially patterned activity in the afferent nerve fibers to the cortex, (iii) fixed synaptic connections between cortical neurons that are excitatory at short distances and inhibitory at longer distances, and (iv) normalization of synaptic strength [869].

The most successful model class is that of the *dimension-reduction models*, such as elastic nets, which make use of competitive Hebbian synapses and are generally based on Kohonen's learning rules for self-organizing maps [497]. Most cortical maps are a combination of a receptor surface map (e.g., retinotopic) with a mapping of other response properties such as orientation selectivity and ocular dominance. Thus, the visual cortex maps several stimulus dimensions onto a two-dimensional surface under the constraint that shape and position of the RFs vary smoothly over the cortical surface. The stimulus dimensions are five: two for retinal position, two for orientation selectivity, and one for ocular dominance. It appears that modeling the formation of retinotopic maps requires quite different learning rules compared to modeling both the orientation and ocular dominance columns [726].

Two types of activity-driven Hebbian learning hypotheses have been used to explain topographic map formation. One is correlation-based learning, which assumes that lateral connections in cortex act linearly such that the activity of cortical neurons is related to their total afferent input. The other is the competitive neural network model where only a small localized region of cortex is active for a given afferent input for a given time and located at the region which receives the strongest total input. This is a "winner-take-all" type of model. Correlation-based learning models cannot predict the emergence of localized RFs but can predict the formation of orientation and ocular dominance columns. The competitive neural network can model all three topographic maps [673].

Specifically, in the correlation-based learning model, the activity of a neuron in layer M is a linear combination of the output of the previous layer L:

$$\mathbf{x}^M = \sum_{i \in L} \mathbf{w}_i x_i^L + \mathbf{a}, \tag{1.18}$$

and the learning rule is

$$\Delta \mathbf{w}_i = \sum_{j \in L} (Q_{ij}^L + b)\mathbf{w}_i + \mathbf{c}, \tag{1.19}$$

where Q_{ij}^L denotes the covariance matrix of points in layer L and $\mathbf{a}, b$, and $\mathbf{c}$ are constants. The competitive neural network model uses feature vectors $\mathbf{x}$ as stimuli, and a point on the cortical surface after learning represents such a feature vector. Other points j on the cortical surface have their connection weights updated according to the following learning rule:

$$\Delta \mathbf{w}_j = \eta h(r)(\mathbf{x} - \mathbf{w}_j), \tag{1.20}$$

where η is a constant and $h(r) = \exp(-r^2/2\sigma^2)$ is a circular symmetric (or elliptic shape) Gaussian function with $r = \|\mathbf{x} - \mathbf{w}_j\|$ denoting the distance between cortical point j and the cortical point closest to the ideal representation of stimulus $\mathbf{x}$. Here the rate of change is proportional to the current level of presynaptic activity as reflected in $(\mathbf{x} - \mathbf{w}_j)$ but conditional on the synapse being close [as reflected in $h(r)$] to the region of cortex that responds most strongly to the stimulus.

It appears that the correlation-based learning model and the competitive neural network model are two extremes of a model with variable intracortical competition. Recapitulating, a Hebbian learning rule can be formulated under the assumption of synaptic competition between cortical neurons. In the limit of weak competition the system becomes linear, no topographic projection emerges, and only the correlations of the input patterns, together with the shape of the interaction function, determine the developing RFs and neural maps. For strong competition a topographic map emerges. In other words, the correlation-based learning model and the competitive neural network model are two extremes of a model with variable intracortical synaptic competition [726].

1.6 CORRELATION IN MEMORY SYSTEMS

Correlation in neural coding and processing is likely to be used in rather different ways by different memory systems. Posterior neocortical areas organize and represent sensory information roughly into hierarchies of topographically organized feature maps, as discussed in the preceding sections. This type of representation may be optimal for extracting and representing the overall statistical structure of the sensory world averaged over a long timescale. The medial temporal lobe (MTL) memory system, on the other hand, exhibits strikingly different properties and plays a unique role in episodic memory. In spite of these differences in timescale, all forms of long-term memory including episodic memory have a common neural mechanism, namely, synaptic plasticity. In contrast, short-term or working memory represents transient information that is stored on a timescale of seconds to minutes and may not involve synaptic changes.

Working memory allows information to be maintained in an active state over a period of seconds to minutes in the form of persistent neural activations. It is thought to result from dynamic interactions between frontal and posterior cortical memory sites. For example, single neurons exhibiting similar, persistent firing during the delay period in spatial working memory tasks have been recorded in both prefrontal and parietal cortices [153]. Functional imaging studies implicate the coactivation of frontoparietal and frontotemporal networks in working memory for auditory [44, 598] and visual [175, 188], information with the specific subregions within these areas dependent on the specific task. Correlated, synchronous firing may play a role in the maintenance of information in working memory. For example, gamma oscillations recorded in electroencephalography (EEG) increase in power in proportion to memory load in a working memory task [407]. Magnetoencephalography (MEG) studies reveal long-range phase synchronization in a network of prefrontal, parietal, and temporal areas implicated in the maintenance of working memory, with the degree of synchronization varying with the attentional demands of the task [341].

Episodic memory refers to memory for specific events situated in particular places and times. In contrast to posterior cortical areas, which recode the sensory signal so as to extract general features and abstract away the specific details of

each learning episode, the episodic memory system codes and retrieves the specific details of particular episodes. Thus, the key properties of an associative memory system can be observed: It binds together the elements of a complex event into a single memory trace and subsequently can perform pattern completion, whereby a subset of the elements of the original memory trace can cue the retrieval of the entire episode. The MTL memory system, comprised of the hippocampus and surrounding MTL structures (parahippocampal region, subiculum, entorhinal and perirhinal cortices), is crucial for episodic memory. Hence, people with MTL lesions show severe retrograde and anterograde amnesia for episodic details, while semantic, perceptual and procedural memory, and simple forms of conditioning are spared (e.g., [637]).

Whereas stimulus correlations on a long timescale appear to drive the learning of long-term memory representations in neocortical systems, the hippocampal system must be able to ignore correlations between different events, so as to create unique memory traces for otherwise similar events. For example, the act of driving to work and parking one's car in the same parking lot is extremely similar from one day to the next, and yet finding one's car at the end of the day requires the retrieval of a unique event memory for where the car was parked most recently. Thus the goal of learning in an episodic memory system should be to capture instantaneous correlations between inputs as accurately as possible while decorrelating separate events into nonoverlapping memory traces. Three unique features of the hippocampal memory system account for its ability to carry out this goal: (i) *Anatomy:* Being reciprocally connected to most cortical and subcortical regions, the hippocampal system is well positioned to bind together features from the various cortical maps as well as emotional and other contextual information. It is a massive convergence zone, in terms of both its incoming and outgoing connections. Moreover, the circuitry of the hippocampus, illustrated in Figure 1.15, is strikingly different from that of neocortex. Most notably, the extensive associational fiber pathways, including the CA3 recurrent collaterals and CA3-to-CA1 Shaffer collaterals, may help the hippocampus to perform associative memory and retrieval functions. (ii) *High plasticity:* The ability to encode complex associations ultrarapidly or even with one-shot learning necessitates a very high level of plasticity. Electrophysiological support for this comes from studies of LTP. Long-term potentiation can be induced in the hippocampus within a single session of high-frequency electrical stimulation, whereas neocortical LTP in the awake behaving animal occurs on a much slower timescale, requiring multiple sessions spaced over several days and taking about 15 days to reach asymptotic levels [741]. (iii) *Sparse coding:* Neurons within the hippocampus, under the tight control of inhibitory interneurons, fire at extremely low rates [63, 453]. Using sparse codes, on average, decreases the amount of correlation between neurons and results in highly specific coding at the single-neuron level. For example, hippocampal "place cells" each fire when a rat is in a particular location within a given environment and tend not to respond in other locations in the same environment or in similar locations in different environments. Place-selective neurons have also been recorded in the nonhuman primate [600, 689] and in the human hippocampus [256].

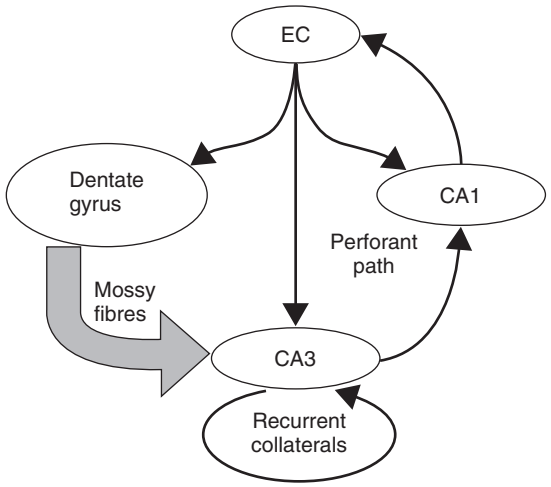


Figure 1.15 The major regions and pathways within the hippocampus. The entorhinal cortex (EC) is reciprocally connected to most cortical areas and in turn projects to all regions within the hippocampus. In addition to receiving direct input from the EC, each region is connected in series in what is known as the trisynaptic circuit, from the EC to the dentate gyrus, CA3, CA1, and back to the EC, thus completing the loop. Some of the most striking and unique anatomical features of the hippocampus are the presence of the mossy fibers, which have the largest and most potent synaptic connections in the mammalian brain, and the massive system of recurrent collateral connections within the CA3 region.

Sparse Coding and Pattern Separation. In 1971 Marr [591] put forward a highly influential computational theory of hippocampal coding. The central ideas in this theory included a rapid, temporary memory store mediated by sparse activations and Hebbian learning, an associative retrieval system mediated by recurrent connections, as well as a gradual consolidation process by which new memories would be transferred into a long-term neocortical store. Many subsequent models have built upon these same basic principles (e.g., [71, 358–360, 460, 550, 604, 605, 610, 693, 769, 892]). Computer simulations have demonstrated that sparse coding serves to remap the input from a space in which many correlated features are present to a new space in which feature correlation is minimized, a process that has come to be known as "pattern separation" [692]. Hence sparse codes are optimal for creating unique event memories with minimal overlap between different memories [591]. When overlap is minimal, interference between different memories is reduced. This allows the hippocampus to employ a very high level of plasticity without suffering unduly from interference [605].

Associative Learning and Pattern Completion. Largely based on predictions from computational models, it is now widely assumed that the associational pathways within the hippocampus, especially recurrent connections within the CA3 region, explain its capacity for pattern completion, that is, the cued retrieval of

a complete memory. In support of this notion, place cells maintain their place selectivity even in total darkness [609, 740]. Moreover, knockout mice having selective loss of NMDA receptors in CA3 show normal acquisition and normal place fields in spatial memory tasks but a loss of place-selective firing after partial cue removal [652].

Temporal Sequence Learning and Consolidation. Hippocampal neurons not only encode static "snapshot" memories, they also encode temporal sequences. Strong evidence for this claim comes from cellular recordings suggesting replay of recently experienced event sequences. When a rat runs through a series of locations in an environment, a corresponding series of place cells fire in sequence. Interestingly, during the next few hundred milliseconds, the same set of place cells tend to fire in the same sequence, albeit on a compressed timescale [840]. Similar patterns of very brief temporal sequence replay have also been recorded during slow-wave sleep [535, 648, 712, 964], and very long sequences of up to 2 min have been recorded during rapid-eye-movement (REM) sleep [572], while replay of sequences in reverse order has been reported during periods of inactivity immediately following periods of locomotion in the awake state [287]. One hypothesis as to the functional significance of this sequence replay is that the hippocampus may be involved in long-term memory consolidation, consistent with Marr's theory of hippocampal function as a temporary memory system. The hippocampus, being ideally suited for rapid memory acquisition, could act as a cache for memorizing recently experienced sequences. The repeated replay of these sequences and their corresponding correlated firing patterns in the neocortex via hippocampal–cortical back projections could allow slower, correlation-based learning mechanisms in the neocortex to process the newly learned information. The consolidation hypothesis is controversial, particularly with respect to human episodic memory (e.g., [638]) and animal spatial memory [593], but is supported by evidence from animal lesion studies for at least some types of learning [848].

Oscillatory Firing within the Hippocampus. The hippocampus exhibits at least two distinct modes of firing that appear to have very different behavioral correlates (for a review, see [170]). One mode is highly synchronized to the so-called theta rhythm, a slow, regular rhythm of about 4–8 Hz associated with active wakeful periods and REM sleep. A second predominant mode is characterized by irregular, high-intensity "sharp-wave" events and is seen during quiet wakeful activities as well as in slow-wave sleep. Yet a third mode, the "hippocampal slow oscillation" ($\leq$1 Hz) has also been reported [967]. These different rhythms may help to organize the activities of different subregions of the hippocampus into synchronously firing assemblies to promote learning and memory operations. They may also serve a similar function with regard to coordinating hippocampal–cortical interactions.

During theta mode, hippocampal neurons fire in gamma frequency volleys (40–100 Hz) strongly modulated by the theta rhythm [170]. A number of lines of evidence suggest that the theta oscillation may serve to lock the hippocampus into

a memory acquisition/retrieval mode. For example, in many species, theta occurs during alert attentive and exploratory behaviors and REM sleep, but not during quiet waking or consummatory behaviors [965]. Hippocampal theta oscillations have also been observed in human intracranial field recordings during exploration and goal-seeking behavior in a virtual environment [140] as well as during REM sleep [139]. The neuromodulator acetylcholine regulates theta oscillations, and cholinergic agonists enhance both theta oscillations and plasticity [417]. Further, it is the superficial layer neurons of the entorhinal cortex, which convey cortical inputs into the hippocampus, that fire predominantly during theta, whereas the deep-layer neurons responsible for conveying hippocampally generated signals back to the cortex fire more weakly during theta mode [169]. Additionally, electrical stimulation in theta frequency bursts is optimal for LTP induction [526]. Finally, the direction of plasticity is linked to the theta phase: Neurons that fire in phase with theta undergo LTP while neurons firing in antiphase undergo LTD [398, 418, 423, 711]. Thus, it appears that encoding of new information is performed optimally when the incoming information is maximally correlated through the synchronizing effect of the theta oscillations. A related hypothesis is that the hippocampus rapidly switches between encoding and retrieval modes within each theta cycle [357]. In support of this hypothesis, human intracranial recordings have revealed that neurons in numerous brain locations exhibit theta-phase reset during both encoding and retrieval (in response to presentation of study items and test items, respectively), but firing during retrieval is nearly 180° out of phase with that seen during encoding [763]. Whether there is rapid alternation between retrieval and encoding during each theta cycle or switching between these states at a longer timescale remains to be seen.

The theta oscillation may also serve to coordinate assemblies of neurons involved in temporal coding of information. This could explain the phenomenon of theta-phase precession [683]: When a rat first enters the place field of a given place cell, the cell fires at a relatively late phase in the theta cycle, and as the rat moves through the place field, the cell fires at progressively earlier phases. One explanation for phase precession is that strong lateral synaptic connections are formed between neurons coding for nearby places, causing place cells whose fields are about to be entered to receive some lateral activation from place cells whose fields are already occupied [128]. An alternative explanation for phase precession is that it reflects a precise use of spike timing to encode for spatial location, and it is caused by temporal properties of the hippocampal input [681] rather than by lateral interactions. In this view, the theta rhythm would serve as a clock for coordinating the relative timing of neurons, with phase offsets coding for specific place-related information while firing rates correlate with other variables such as running speed [422].

In addition to coordinating circuits within the hippocampus, the theta oscillation may function to coordinate multiple brain regions into functional networks for different purposes. For example, individual neurons in medial prefrontal cortex fire in synchrony with hippocampal theta during foraging and running [424], while hippocampal neurons fire in synchrony with amygdala neurons at theta frequency during retrieval of fearful memories [813]. Increased theta synchrony has been

observed in human EEG recordings across prefrontal, MTL, and visual areas during recollection of images compared to object recognition [346]. Thus, correlated firing across multiple brain regions by synchronization to the theta rhythm may be a general principle by which selective regions are recruited into specific tasks.

In the absence of theta activity, during silent wakeful and consummatory behaviors and slow-wave sleep (SWS), hippocampal neurons generate "sharp waves" coinciding with high-frequency "ripples," causing the output region of the hippocampal–cortical interface (deep layers of entorhinal cortex) to fire in synchrony with these sharp wave/ripple events [129, 167, 168]. During sharp waves, the hippocampus appears to act in "top-down" mode, sending information back out to cortex. Sharp waves are generated and propagated via the associational pathways within the hippocampal circuit—the CA3 recurrent collaterals and CA3–CA1 Shaffer collaterals; they arise from highly synchronized population bursts generated in the CA3 triggering aperiodic, high-intensity dendritic field potentials (sharp waves) lasting 40–100 ms in CA1 [167]. Coinciding with the sharp wave events, a 200-Hz oscillatory field potential synchronizes spiking in CA1. The deep layers of the entorhinal cortex, subiculum, and parasubiculum fire strongly during sharp-wave events, propagating activity back out to cortex, while the superficial layers are relatively silent [167]. Thus, during each sharp-wave event, there is a powerful synchronization of the neural circuits connecting the hippocampus to the neocortex [167, 168]. It is widely believed that sharp-wave events during SWS represent hippocampal replay of recently acquired memories for the purpose of memory consolidation. In support of this notion, acetylcholine levels are naturally lowest during SWS, and consolidation of recently learned word pairs is blocked when a cholinergic agonist is injected during SWS but not when the injection occurs during waking periods [305].

As a computational neural network model, the Boltzmann machine, which was first proposed by Hinton and Sejnowski [6, 390], operates in alternating bottom-up and top-down modes reminiscent of the hippocampal–cortical interactions during theta versus sharp waves, suggesting a computational explanation for these two modes. During the "waking" phase, the circuit operates in a data-driven mode and learns to represent the current input pattern via Hebbian learning. During the "sleep" phase, the circuit operates in a generative mode, triggering activity states that may be similar to recently learned memories but that should be "unlearned" via anti-Hebbian learning. Recently, Kali and Dayan [460] developed this idea further in a Boltzmann machine model that demonstrates the feasibility of a rapid memory store within the hippocampus promoting memory consolidation in the neocortex. Moreover, Kali and Dayan proposed that the periodic replay of hippocampally stored memories is required to keep the neocortical and hippocampal memory representations in register with one another.

1.7 CORRELATION IN SENSORIMOTOR LEARNING

Sensory and motor systems are closely connected and form an intrinsic sensorimotor loop/cycle. On the one hand, the motor primitive commands are driven by the

sensory events, and on the other hand, the motor events influence the forthcoming sensory inputs. Learning occurs at many levels of the brain from simple conditioning to the learning of complex sensorimotor response mappings. Correlations between sensory inputs and outcomes underlie all types of sensorimotor learning. Specifically, temporal difference learning theory (e.g., see [201]) has proven to be a powerful model for capturing these many levels of learning.

Temporal-Difference (TD) Models of Classical Conditioning and Their Relation to STDP. We remind the reader that a CS is the originally neutral stimulus that, after training, comes to elicit the learned behavior (the CR). This change results from the experimenter-enforced arrangement that the CS temporally precedes and thus predicts reinforcement of the US. Unit firing patterns that develop during conditioning and are specific to the CS (i.e., are absent in response to stimuli that do not predict the US) represent learning-relevant neuronal plasticity. Thus, a CS is a convenient probe for detecting neuronal plasticity, thereby permitting identification of regions and circuits involved in learning processes. In the classical Pavlovian experiment, the CS is the sound of a bell, the US is a plate of food, and the CR is the salivation by the dog involved. Previously, in Section 1.2, we presented an example where the CS was the firings of a presynaptic neuron, the CR was the response of the postsynaptic neuron, and a sound that activated the presynaptic neurons served as the US. The *Rescorla–Wagner rule* [758] for classical conditioning states that "organisms can only learn when events violate their expectations," namely, learning only occurs when the information content of the message is high. In the conditioning experiment, the sound of a bell (although not uncommon to call people to natural or spiritual food) is initially highly unlikely to indicate food for a dog, so it violates expectations. Specifically, the learning rule can be written as

$$\Delta w_{ij}(t) = \eta[r_j(t) - v_j(t)]x_i(t), \tag{1.21}$$

where $r_j(t)$ is the value of a specialized training signal that induces the desired response of the unit and $v_j(t) = \sum_j w_{ij}(t)x_i(t)$ is the neuron's standard response to an input $x_i(t)$. However, the above model used in animal learning is not a TD model, but all of its predictions can be obtained from TD models.

Specifically, TD models relate to the effect of interstimulus interval between a CS and US and thus cover the class of STDP synaptic plasticity models in principle. In order for an association between the CS and US to occur, they must occur roughly at the same time. This is reflected in an update of association strength, Δw, according to the product of the levels of CS and US processing. Sutton and Barto [867] introduce "reinforcement" as the level of US processing and "eligibility" as the level of CS processing; thus, the learning rule is written as

$$\Delta w = \text{reinforcement} \times \text{eligibility}.$$

Eligibility is always positive, but reinforcement can be either positive or negative. In this context, the Rescorla–Wagner rule can be rewritten as

$$\Delta w_i = \eta(r - V)\beta_i x_i, \tag{1.22}$$

where $r - V$ denotes the reinforcement and $V = \sum_i w_i x_i$ is the predicted signal and $\beta_i x_i$ denotes the eligibility. When $x_i = 0$, then the ith CS, denoted as CS_i, is absent; when $x_i = 1$, CS_i is present. A subsequent improvement over the Rescorla–Wagner model is that all CS in a given trial can produce associative strengths with total sum that is given by $Y = \sum_t r(t) - V(t)$; then the reinforcement is postulated to be equal to dY/dt, which further leads to

$$\Delta w_i = \eta \frac{dY}{dt} \beta_i x_i. \tag{1.23}$$

However, a drawback of this model is that it still does not perform well for long interstimulus interval. As we have seen, classical conditioning can be viewed as a manifestation of the subject's attempt to predict the arrival of the US. In terms of the Rescorla–Wagner model, V is the predicted US level on the trial and r is the actual US level. The difference $r - V$ drives the learning process as the model's reinforcement term. How is this to be extended to a time-dependent form? Here r will change with time within a trial and each successive delayed US within the same trial will have a forgetting effect γr $(0 \leq \gamma < 1)$, and the total prediction becomes nonlinear, $V(t) = r(t + 1) + \gamma r(t + 2) + \gamma^2 r(t + 3) + \ldots$. This results in a modification of the reinforcement term as follows:

$$\Delta w_i = \eta \left[r(t + 1) + \gamma V(t + 1) - V(t) \right] \beta_i x_i(t + 1). \tag{1.24}$$

The reinforcement term can be interpreted as a difference between two predictions, one at time $t + 1$ and the other at time t, and can be considered as a discrete-time analog of the prediction dV/dt. This TD model performs well, but its relationship to the STDP synaptic plasticity model is only qualitative. The STDP model works over timescales of the order of 100 ms whereas classical conditioning operates over timescales of one order of magnitude higher. It is believed that in classical conditioning many brain areas are likely involved, with polysynaptic and recurrent circuitry that can effectively extend the timescale of computation without sacrificing the requirement of a very short STDP window at each synapse [88]. Drew and Abbott [229] suggested that cortical up and down states or changes in the spontaneous background activity might generate the required extension of the correlations and the asymmetry in the potentiation and depression part on the behavioral timescale.

Neural Adaptive Information Processing. How does learning alter neuronal representation of sensory events, that is, when animals acquire information about their behavioral significance and meaning? This approach entails an analysis of the responses of neurons to the CS within the appropriate sensory system.

Classical conditioning in the auditory system [194] produces highly specific modification of RFs in auditory cortex. In this case, responses to frequencies near the boundaries but still within the original RF are strengthened at the expense of those to frequencies in the center of the RF. The strengths of the lemniscal input to pyramidal cells are continuously adjusted during learning, whereas the nonlemniscal auditory influence produces a diffuse increase in the excitability of these neurons throughout auditory cortex. Here we note that lemniscal inputs to the cortex synapse generally in layer IV, whereas nonlemniscal inputs synapse with pyramidal cells in layer I. The latter effect is greatly modulated by acetylcholine released diffusely in cortex by electrical stimulation of the basal forebrain. Specificity of RF plasticity (or representational plasticity) is said to result from a modified Hebbian rule, so that increased excitability strengthens some synapses and weakens others within the same neuron.

Effects of Modulatory Neural Systems. Direct neural transmission uses glutamate and GABA as transmitter substances to excite and inhibit, respectively, the receiving neurons. In addition to the direct-acting neurotransmitters, there are numerous modulatory substances including, among others, acetylcholine, dopamine, and serotonin. Modulatory transmitter substances are released in cortex diffusely by systems originating in the forebrain or brainstem. Acetylcholine is released by the nucleus basalis in the basal forebrain, whereas dopamine is released by neurons in the ventral tegmentum area and substantia nigra. Natural release of these modulatory substances during learning or behavior can be mimicked in experimental conditions by electrical stimulation of the releasing structures. The modulatory effects are studied by pairing electrical stimulation of the relevant brain structure with an acoustic stimulus. Pairing nucleus basalis stimulation with narrow-band stimuli for several weeks enhances cortical representation of the sound frequencies [475]. Specifically, pairing nucleus basalis stimulation with periodic tone-pip trains with a rate (15 Hz) above the normal range in cortex enhances responsiveness at this rate; and pairing with periodic tone-pip trains at 5 Hz makes the cortex less responsive at rates it would normally process. This depended on the tone frequency; when it was randomly changed, the effect occurred as described; when the tone frequency was fixed, there was no effect [476]. The effects of the cholinergic and the dopaminergic systems on cortical plasticity are different. Dopaminergic activity may enhance or reduce cortical representation depending on the stimulus contingency [56]. Thus ventral tegmental dopaminergic activity may be essential in contingency-based associative learning. Dopamine has long been thought to play a role in reinforcement learning. Phasic firing of dopamine neurons exhibits a striking correlation with TD error (a reward prediction error signal): A first exposure to a US evokes strong dopamine firing, whereas after repeated CS–US pairings the dopamine response transfers to the CS and is much weaker upon arrival of the US, signaling a near-zero prediction error [805]. On the other hand, when an anticipated reward fails to arrive, dopamine neurons fire below baseline levels, signaling a negative prediction error. Thus the actions of dopamine may constitute the brain's implementation of the TD learning algorithm [64, 865, 866].

Cholinergic effects, unlike dopamine, are largely determined by the spectral and temporal characteristics of the paired sensory stimulus [477, 478]. The cholinergic system could thus be more engaged in stimulus feature-directed perceptual learning [57]. In addition to direct effects on plasticity, acetylcholine suppresses synaptic transmission in cortical feedback pathways [322, 359, 411, 482]. It has been suggested that under conditions of uncertainty, such as when an unexpected stimulus arrives or under high noise levels, acetylcholine upregulates bottom-up, thalamocortical transmission of information, leading to optimal inference and learning [992].

Behavioral Training-Induced STRF Changes. Auditory tasks can influence RF properties of cortical neurons. In a series of such experiments, ferrets were trained on target tone detection and two-tone discrimination tasks and on gap detection and click-rate discrimination. STRF changes were measured online during task performance and facilitative changes occurred within minutes of task onset. During frequency detection or discrimination tasks, there were only spectral changes in the STRF. However, during and following temporal tasks, the STRF showed sharpened temporal aspects [295]. The fact that RF plasticity occurs during very different tasks and learning situations suggested that it represents a general process of information storage and representation [945]. Cortical RF plasticity can be induced within a few trials and the changes paralleled the appearance of the first behavioral signs of learning [237]. Receptive field plasticity has a short-term, fast-learning component which is only demonstrable in a behavioral context [217]. Polley et al. [734] tested whether topographic map plasticity in the adult auditory cortex was controlled by sensory inputs alone and/or by task dependence. Rats trained to attend to intensity cues had an increased proportion of units that were tuned to the target intensity range but showed no change in tonotopic map organization. The degree of topographic map plasticity within the task-relevant stimulus dimension was correlated with the degree of perceptual learning for rats in both tasks.

Learning More Complex Sensorimotor Mappings. According to computational models, learning to associate together a CS and a US can be accomplished by a single neuron. In contrast, learning to perform a complex motor task such as reaching for an object requires the coordination of many muscle groups and numerous brain regions. One way to study how animals accomplish such tasks is to perturb the system. Under such perturbations, the brain must learn to compensate by recalibrating the sensorimotor mapping. For example, it is well known that one can adapt to distorting prisms, even when they completely reverse the visual image, and that the adaptation persists for some period of time after removal of the glasses. Interestingly, wearing laterally displacing prisms during walking results in plasticity, which generalizes to reaching movements [636]. In the auditory domain, when people are fitted with an artificial pinna (outer ear) that perturbs the binaural cues for sound localization, they adapt and eventually regain normal auditory sound localization [396]. However, in contrast to the case of the visual system, upon removal of the artificial pinna the system immediately reverts to its original

configuration, demonstrating a capability for maintaining two sets of mappings simultaneously.

It is thought that the difference between the predicted and actual consequences of one's actions is the error signal driving such learning processes. For example, when one puts on distorting prisms and tries to reach for an object under visual guidance, misreaching occurs. The predicted consequence of one's action does not match up with the visual percept. This error is a potent signal for driving plasticity and has been capitalized on in treatment of several neurological disorders. For example, prism adaptation has been used to treat hemispatial neglect [775]. Neglect patients, usually as a result of a right parietal stroke, fail to attend to objects in the left side of space even though they are able to process all of the visual information. Adapting to prisms that shift the visual field in the direction of the "good" side of space results in improvement in neglect symptoms, not only while the prisms are worn but up to 2 h after removal of the glasses. The same principle of altered visual feedback has been used in the treatment of phantom limb pain. The patient is shown a projected, mirrored image of his or her good arm in the place where the paralyzed arm should be and observes the limb movements during an adaptation period, thus providing altered visual feedback that can recalibrate the sensorimotor circuits to some degree; this treatment has been reported to alleviate phantom pain in some patients [746]. Numerous computational models of learning are also based on prediction error. In the case of the TD model [64, 865, 866], the error signal is the difference between received and predicted reinforcement. The TD learning model has been extended to model learning motor actions in more complex domains. The Q-learning algorithm [942] involves learning to predict future reinforcement obtained when choosing among a set of alternative actions. In this case, the agent must learn to predict how much reinforcement can be expected for each combination of actions and states in the environment.

1.8 CORRELATION, FEATURE BINDING, AND ATTENTION

It has been widely believed that very strong correlations in the sensory signals are likely to be learned early in development, effectively becoming hard-wired into neural circuits at early stages of processing. Within spatially localized regions of the visual field, features such as color, orientation, and texture are strongly correlated across space; within the auditory domain, spectral properties are correlated across time. It appears that some neurons are tuned to combinations of such correlated features. For example, people are about 10 times faster at detecting combinations of orientation and color if they are superimposed within the same visuospatial region relative to when they are spatially separated [397]. These findings suggest that correlated features within the same location are coded in combination explicitly at an early stage of processing. However, employing a different neuron to encode every possible combination of stimulus features is clearly infeasible. Thus combinatorial coding cannot be the only solution employed by the brain for detecting stimulus correlations.

How are correlated features detected when there is no neuron already tuned to a given feature combination? The *temporal correlation theory* [922] posits that when features activate a population of neurons coincidentally this temporal coincidence in firing leads to the emergence of a synchronously firing cell assembly. Moreover, von der Malsburg proposed that such transient correlation patterns may become temporarily strengthened by rapid reversible synaptic plasticity. Further, it has been proposed that multiple synchronously oscillating cell assemblies could represent coherent groupings of parts of the same object, offering a solution to the binding problem, a central problem in perception [235, 337] A substantial body of evidence supports a role for synchronization in bottom-up feature integration (for a review, see [871], although there is also evidence to the contrary, as reviewed in Section 1.3). For example, in cat visual cortex, unit recordings revealed periods of oscillatory spike synchronization between neuron pairs in two different visual areas (striate and extrastriate cortices); moreover, the synchrony was greatest for coherent motion of a single stimulus that activated both RFs, weaker for a pair of coherently moving stimuli each within one of the pair's RFs, and weakest for independently moving stimuli [260]. It is unclear how a pattern of synchronous firing could give rise to a categorical object percept, and thus the theory has been much debated (e.g., [820]). Nonetheless, there is broad agreement that the detection of multiple features represented across different neurons is greatly facilitated when those neurons are active more or less synchronously to within a narrow temporal margin.

The temporal correlation theory implies that the synchronous firing of cell assemblies arises via self-organizing, bottom-up processes. An alternative or perhaps complementary view (if both processes take place) is that object perception involves sequential, top-down attention to each feature or object part. Treisman's *feature integration theory* [889, 891] posits that features represented in separate maps must be attended to sequentially in order to be detected conjunctively. An implication of this theory is that individual features may be detected in parallel, whereas a search for a conjunction of features requires a serial search. In support of this notion, Treisman showed that reaction time in visual search for a single feature such as color or orientation is unaffected by the number of distractors in a display, whereas searching for a conjunction of features takes time proportional to the number of distractors. Although the notion of a strict dichotomy between parallel and serial attention has been called into question (e.g., [970]), as has the necessity of attention for object recognition [883], nonetheless, there is substantial evidence for a role for top-down attention in object perception. Perhaps the strongest evidence comes from the study of illusory conjunctions. When people are shown displays of multiple objects, with insufficient time for sequential attention to operate, they frequently make binding errors, for example, reporting having seen a blue "O" in a display that actually contained a red "O" and a blue "H" [889]. Similarly, in the auditory modality, when presented with overlapping streams of high- and low-pitched tones, if subjects were not attending to either stream, they failed to notice deviant tones and did not exhibit EEG mismatch negativity (MMN), whereas attention to the high-pitched tones resulted in MMN for

either stream [864]; this suggests that stream segregation, a form of auditory object segmentation, depends upon attention. The top-down application of selective attention to different parts of the same object may serve to synchronize the firing of the corresponding feature detectors and thereby increase the correlations in their outputs. There is ample evidence from EEG recordings that attention during object processing boosts synchronous activity in the gamma-band range (40–130 Hz) (e.g., [204, 768, 872, 884]). Intracranial recordings from awake animals provide further evidence for the synchronizing effects of attention. For example, when macaque V4 neurons are presented simultaneously with stimuli within and outside their RFs, they synchronize much more when the animal's attention is directed to a region within the neurons' RF relative to attention outside the RF [293]. Likewise, in somatosensory cortex, attentional switching between visual and tactile tasks modulates neuronal synchrony, with the degree of synchrony increasing with task difficulty [854]. What might be the role for attentional modulation of neural synchrony? As discussed in Section 1.3, an increase in correlated firing increases the efficiency within which that neuronal representation can be detected at higher levels of processing. Thus, when multiple feature detectors are persistently firing in a correlated manner, they may be more likely to jointly drive a higher level process such as object detection, learning and memory, or a motor response. Another possible role for attention is to increase discriminability. McAdams and Maunsell [603] found that attentional modulation improved the reliability with which V4 neurons could discriminate their preferred orientation. An increase in neuronal reliability could explain the increase in correlated firing resulting from attentional modulation.

Interestingly, recent evidence from human intracranial recordings suggests that attention may affect synchrony in different brain regions in very different ways. For example, attention to a visual stimulus was found to increase stimulus-driven gamma-band oscillations in the fusiform gyrus, whereas attention increased baseline gamma-band oscillations preceding stimulus onset in the lateral occipital sulcus [872]. These authors speculated that the effect of attention during the preparatory phase in the lateral occipital area may have been to put neurons into a state of readiness, thereby allowing their earliest stimulus-driven responses to be synchronized within a narrow temporal window. Many open questions remain regarding the role of correlation and synchrony in feature binding and object perception, and it is currently an active area of research.

1.9 CORRELATION AND CORTICAL MAP CHANGES AFTER PERIPHERAL LESIONS AND BRAIN STIMULATION

As we have seen, activity-dependent reorganization of the cerebral cortex is found during development and learning. In each sensory system, the cortical map can also be induced to undergo large-scale rearrangement after amputation or other surgical or damaging manipulations of the sensory periphery [133, 323, 448]. The best known is the sensation of a phantom limb and in more serious forms that of phantom pain following amputation or deafferentation of a limb in humans. Another less known example is tinnitus that follows partial deafferentation of the

output of the cochlea through hearing loss. Although in these cases the effect of cortical reorganization is not beneficial, the neural mechanisms that underlie such deprivation-dependent cortical reorganization may be the same as those responsible for the improvements in perceptual skills that accompany learning. Thus, they warrant a discussion. There are several phases to the reorganization process: an immediate phase of expansion of the representations of parts with intact innervation adjacent to the (partially) deafferented region; a phase lasting weeks or months in which the new representation is consolidated and topographic order at least partly restored; and a late phase of further expansion and use-dependent refinement of internal topography. These common changes can happen in various ways, and we will see that the three major sensory systems, *touch, audition*, and *vision*, differ in how much of the observed topographic map changes in cortex are due to subcortical changes. To appreciate this fully, we have to know the differences in the pathways from receptor to cortex. These are illustrated in Figure 1.16 and discussed below. The same approximate levels in hierarchy are indicated with similar shapes.

Comparative Anatomy of Major Sensory Systems. The sensory system that is confined to the brain is the visual one. The eye with its sensory surface, the retina, is typically considered an extracranial extension of the brain. Briefly,

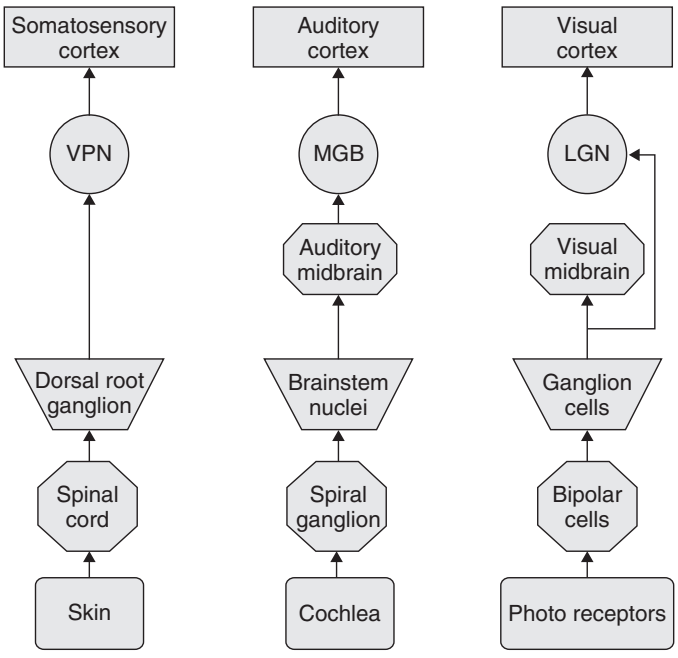


Figure 1.16 Schematic of sensory pathways in somatosensory, auditory, and visual systems (VPN: ventral posterior nucleus; MGB: medial geniculate body; LGN: lateral geniculate nucleus).

the photoreceptors are connected to the bipolar cells which input to the ganglion cells that form the optic nerve. But this is not a one-to-one relay; small groups of neighboring photoreceptors provide common input to individual horizontal cells that mediate lateral inhibition. The bipolar cells that collect input from several neighboring photoreceptors activate several ganglion cells; the amacrine cells connect several ganglion cells together. Thus the retina is a broadly interconnected processing network with the ganglion cells two synapses removed from the receptors. The optic nerve partly activates the superior colliculus in the midbrain but mostly bypasses this to directly innervate the visual thalamic nucleus (i.e., LGN). In this nucleus, input from each eye activates different and mutually exclusive cell layers. The LGN projects mainly to layer IV of V1. The projection of the external world on the retina is topographically mapped onto the thalamus and visual cortex. On the other extreme we find the somatosensory system that has most of its receptors at a long distance from the brain and where a large amount of preprocessing is done outside the brain. Different receptor types reside in the skin of the body surface, information about their activity is transferred to the spinal cord by a large number of nerves that each are responsible for a subset of receptors. Output from the spinal cord ends in the dorsal column nuclei, the cuneate and gracile nuclei, in the brainstem and via the medial lemniscus reaches the ventral posterior nucleus of the thalamus and finally arrives at S1. What sets the somatosensory system apart from the visual one is the extent of divergence of its projections to the dorsal column nuclei and to the thalamus. Thus, a very large number of cells in the dorsal column nuclei and the thalamus come to represent a particular body part, but there is no clear segregation of the outputs of individual nerves representing that body part. In between there is the auditory system with its receptor surfaces clearly outside the brain but close by in the inner ear [248]. Here, there is no map of external space such as in the visual and somatosensory systems, instead sound frequency is mapped "one dimensionally" onto the receptor surface in the cochlea. Thus, the topographic maps of the auditory system are tonotopic or frequency maps and these are in the form of a series of isofrequency sheets, columns that transect the entire cortical area more or less perpendicular to the tonotopic axis. Maps of auditory space are constructed by comparison of sound level and arrival times at the two ears; consequently there are two separate systems, a monaural one and a binaural one, that process information independently. The output of the cochlea is via the auditory nerve, of which each nerve fiber trifurcates at the level of the cochlear nucleus in the brainstem to branch in a tonotopic fashion in each of its three subnuclei. The anterior and posterior ventral cochlear nuclei (VCNs) are purely auditory, but the dorsal cochlear nucleus (DCN) also receives input from the trigeminal nerve among others about the position of the external ear. The output from the DCN terminates in the central nucleus of the inferior colliculus (ICC) of the midbrain. The output from the VCN either goes, via the lateral lemniscus, directly to the ICC (the monaural pathway) or goes to the superior olivary complex in the brainstem where neural activity from the two ears is compared (the binaural pathway) and then to the ICC. The left and right ICC share information and their outputs go to the MGB in the thalamus and then on to A1. Thus, segregation

of individual ear output is not as complete as individual eye output in the visual system.

Cortical and Subcortical Topographic Map Reorganization after Peripheral Lesions. For the somatosensory system, Wall et al. [931] reviewed evidence that peripheral injuries cause widespread neurochemical/molecular, functional, and structural alterations in subcortical and cortical substrates of the brain and that cortical changes are but one reflection of global mechanisms that, beginning from the moments after injury, operate at multiple subcortical levels of the somatosensory core. Faggin et al. [270] also suggest both cortical and subcortical reorganization in the somatosensory system. They recorded simultaneously from up to 135 neurons in S1, ventral posterior medial nucleus of the thalamus, and trigeminal brainstem complex of adult rats before and after reversible sensory deactivations by subcutaneous injections of lidocaine. Immediate and simultaneous sensory reorganization was observed at all levels. Thus, peripheral sensory deafferentation triggers a systemwide reorganization. In the visual system of both adult cats and adult macaque monkeys after circumscribed retinal lesions were made, there was no significant sprouting of the retinogeniculate terminals in the lesion projection zone (LPZ). First, the LGN neurons in the LPZ cannot be activated by visual stimuli presented through the lesioned eye and the silent zone closely approximates in size to the normal representation of the lesioned retinal area [268]. Second, the displacements of the RFs observed in the projection zones of retinal lesions in area 17 of cats and macaque monkeys ($\sim$6–8 mm) exceeded the expected limits of the lateral spread of geniculocortical afferents ($\sim$2 mm). The LPZ in area 17 received their geniculate inputs from the LPZ in the LGN and not from parts of the LGN in which cells responsive to visual stimuli were located. This implicates a cortical mechanism rather than a thalamic mechanism in the topographic reorganization of area 17 [228]. In cat's A1 area, the effect of unilateral localized cochlear lesions in adult cats on the topographic maps of the lesioned cochleas showed large reorganizations [744, 767]. Two to eleven months after a unilateral cochlear lesion affecting the high frequencies, the map of the lesioned cochlea in the contralateral A1 was altered so that the region normally representing the hearing loss frequencies was now occupied by an enlarged representation of lesion-edge frequencies. Along the tonotopic axis the total representation of lesion-edge frequencies could extend up to $\sim$2.6 mm into the hearing loss area, that is, about what one can expect from the thalamocortical divergence. On the basis of threshold sensitivity at the CF, the changes in the map reflect a plastic reorganization rather than simply the residue of prelesion input. In contrast, the map of the unlesioned ipsilateral cochlea, obtained from recordings of the same binaurally sensitive cortical cells, did not differ from those in normal animals. The difference between the ipsilateral and contralateral maps in the region of contralateral map reorganization suggested, in light of the physiology of binaural interactions in the auditory pathway, that the cortical reorganization reflected, at least partly, subcortical changes. To investigate those potential subcortical contributions to cortical reorganization, the frequency organization of the ventral nucleus of the medial geniculate body (MGBv) was

investigated 40–186 days following lesioning [464]. In the lesioned animals it was found that, in the region where mid-to-high frequencies are normally represented, there was an "expanded representation" of lesion-edge frequencies. Neuron clusters within these regions of enlarged representation that had "new" characteristic frequencies displayed response properties (latency, bandwidth) very similar to those in normal animals. The tonotopic reorganization observed in MGBv was similar to that seen in A1 and suggested that the auditory thalamus played an important role in cortical plasticity. To additionally examine the contribution of subthalamic changes to the thalamic and cortical map reorganization, the effects of unilateral mechanical cochlear lesions on the frequency organization of the central nucleus of the ICC were examined in adult cats [431]. After recovery periods of 2.5–18 months, the frequency organization of ICC contralateral to the lesioned cochlea was determined separately for the onset and late components of multiunit responses to toneburst stimuli. For the late-response component in all but one penetration through the ICC and for the onset response component in more than half of the penetrations, changes in frequency organization in the lesion projection zone were explicable as the residue of prelesion responses. In half of the penetrations exhibiting nonresidue-type changes in onset response frequency organization, the changes appeared to reflect the unmasking of normally inhibited inputs. In the other half it was unclear whether the changes reflected unmasking or a dynamic process of reorganization. Thus, most of the observed changes were explicable as passive consequences of the lesion, and there was limited evidence for plasticity in the ICC. Immediate unmasking of subthreshold inputs to the ICC was also noted after minute lesions in the spiral ganglion of the auditory nerve [841]. So most of the changes in ICC might never evolve beyond this initial unmasking of excitatory inputs. No evidence was found for reorganization of the topographic maps in the dorsal cochlear nucleus after partial destruction of the cochlea similar to those that cause massive map reorganization in auditory cortex [614, 743], although small regions of the DCN that were deprived of their normal, most sensitive frequency (or CF) input by the cochlear lesion appeared to have acquired new CFs at frequencies at or near the edge of the cochlear lesion. However, because of the elevated thresholds at the new CFs, the changes simply reflected the residue of prelesion input to those sites. The results suggest that the DCN does not exhibit the type of plasticity that has been found in the auditory cortex and even the midbrain. Combined, this suggests that in the somatosensory system topographic map changes may already occur in the spinal cord and definitely in the brainstem nuclei, in the auditory system potential changes are seen in the midbrain and definitely in the thalamus, whereas in the visual system topographic map changes are likely confined to the cortex.

Time Course of Cortical Topographic Map Changes. Immediately after the lesion, neurons in the LPZ area of cortex either fall silent or show dramatically enlarged RFs that extend far outside the prelesioned area. This may represent unmasking of subthreshold excitatory inputs to a suprathreshold driving control by either suppressing inhibitory or potentiating excitatory connections. The second phase, marked by the gradual expansion of the representation of a

perilesion receptor surface into the initially silenced region of cortex, takes place over weeks and months. The cortex then develops a new, piecewise-continuous map, continuous up to the border of the lesion and then jumping across. It is during this extended period that axonal growth and synaptogenesis occur [196]. Calford et al. [135] monitored topographic reorganization in the V1 of the cat by recording extracellular activity of cells over the 11 h following the circumscribed outer layer monocular lesion in the retina. In the first hour following the lesion, no neural responses could be elicited within the LPZ by photic stimulation of the lesioned eye. However, in the next 1–3 h after the lesion only 39% of the recording sites within the LPZ remained unresponsive. This percentage further decreased to 31% (3–7 h) and 27% (7–11 h). In these cats, the ectopic RFs recorded from the LPZ within hours after lesioning were up to 10-fold larger than their normal counterparts revealed by stimulating the same cells via the nonlesioned eye. In LPZ regions, 1–2 weeks after bilateral retinal lesions, both spontaneous activity and driven activity were significantly reduced. At the same time, both spontaneous and driven activity significantly increased in cortical regions immediately adjacent to the LPZ (associated with a sharp increase in glutamate immunoreactivity [269]). There are several phases to the reorganization process in somatosensory cortex that are very similar to those in visual cortex: an immediate phase of expansion of the representations of parts with intact innervation adjacent to the deafferented region; a phase lasting weeks or months in which the new representation is consolidated and topographic order restored, and a late phase of further expansion and use-dependent refinement of internal topography. In detail, the time course of changes in somatosensory cortex following median nerve section [616] revealed that large cortical areas were silenced by median nerve transection. Inputs from fragments of dorsal skin were immediately unmasked and had greater than normal RF overlap as a function of distance across the cortical surface. They were transformed over time into very large highly topographic and complete representations of dorsal skin surfaces. Representations of bordering glabrous skin surfaces progressively expanded to occupy larger and larger portions of the former median nerve cortical representation zone. These expanded representations of ulnar nerve–innervated skin surfaces sometimes moved, in entirety, into the former median nerve representational zone. Most of the former median nerve zone was driven by new inputs in a map derived 22 days after nerve section. At 11 days reoccupation was still incomplete. Immediately after amputation of a single exposed digit on the forelimb of the flying fox [134], neurons in the area of cortex receiving inputs from the missing digit were *not* silent but responded to stimulation of adjoining regions of the digit, hand, arm, and wing. In the week following amputation, the enlarged RFs shrank until they covered only the skin around the amputation wound. The immediate response can be interpreted as a removal of inhibition and the subsequent shrinking of the RF might be due to reestablishment of the inhibitory balance in the affected cortex and its inputs. In auditory cortex, immediately after a noise trauma unmasking of excitatory inputs was observed [670]. Initially thresholds were elevated (by about 40 dB) and CFs of units recorded before and immediately after the trauma were shifted to lower values, that is, to the edge of the hearing

loss range. Gradually, over a few hours thresholds recovered but CFs did not change, further and essentially all units recorded from acquired new CFs. This confirmed previous findings by Robertson and Irvine [767] where the responses of neuron clusters were examined within hours of making small mechanical cochlear lesions. It was found that shifts in CF toward frequencies spared by the lesions could occur, but thresholds were greatly elevated compared to normal (mean difference was 31.7 dB in five animals). The emergence of driven activity in such regions after prolonged recovery periods in lesioned animals thus suggests that the auditory cortical frequency map undergoes reorganization in cases of partial deafness. Typically, after about 3 weeks reorganization of the cortical tonotopic map is observed [251]. Some features of this reorganization are similar to changes reported in somatosensory cortex after peripheral nerve injury and in visual cortex after retinal lesions, and this form of plasticity may therefore be a feature of all adult sensory systems.

Mechanism of Cortical Topographic Map Changes. In somatosensory cortex, because of the limited extent of the lesion ($<2-3$ mm), it was proposed that thalamocortical axonal divergence within the cortex represented the neural substrate of reorganization. In auditory cortex, the results by Rajan et al. [744] also pointed to the same extent of reorganization. In visual cortex, the extent of reorganization, $\sim 6-8$ mm, was much larger than the lateral spread of thalamocortical afferents. Pyramidal neurons in all cortical areas receive some of their input from afferent connections to the same vertical column. But they also receive inputs from a wide-ranging intracortical network of axons (the horizontal fibers) from more remote pyramidal cells. These horizontal fibers extend about 6-8 mm in cortex, about $2-4$ mm radiating outward from each source neuron. The synapse strength of these horizontal fibers can be altered by appropriate patterns of stimulation, that is, by stimulating the horizontal inputs while simultaneously depolarizing the recorded neuron with injected current [394]. The degree of synaptic strengthening depended on the degree of inhibition present: The greater the inhibition, the less effective the synaptic potentiation. The fact that synapses made by horizontal collaterals can be potentiated suggests that synapses that normally play a modulatory role can, under the proper conditions, be strengthened so as to drive their target neurons above the threshold for spiking. After reorganization, the cortex becomes visually responsive again, but via signals conducted horizontally, intracortically, and in an orientation-selective manner from columns of neurons in unaffected regions of cortex outside the original LPZ. For the RFs of cells in the LPZ to shift, the horizontal connections must be strengthened. The way this could be done is by sprouting axon collaterals and by synaptogenesis [323]. The potential synaptic mechanisms that play a role in cortical reorganization are strengthening of existing but weak or subthreshold synaptic inputs and depression or otherwise weakening of existing strong synaptic inputs. Likely bases of map plasticity may lie in the cortical neurons' ability to compensate for changes in excitatory input by regulating turnover of postsynaptic α-amino-3-hydroxy-5-methylisoxazole-4-propionic acid (AMPA) receptors, thereby scaling the size of EPSP amplitudes and thus the overall responses of a

neuron to stimulation [896], an effect that is mediated by brain-derived neurotropic factor. There are numerous observations that implicate the inhibitory neurotransmitter GABA in map plasticity. Cortical cells released from inhibition commonly increase the sizes of their RFs, and removal of inhibition is thought to underlie the immediate expansions of RFs of somatosensory cortical neurons after loss of peripheral input by amputation or local anesthesia of a digit, an effect that may depend on loss of tonic control by C-fiber inputs over central inhibitory mechanisms [448].

Correlated Neural Activity and Topographic Map Organizations. The specific patterns of reorganization in cortex were driven by the patterns of correlation in natural sensory stimulation. The experiment reported in [174] fusing digits showed that imposing synchrony in the activation of receptors on both digits drove the reorganization and fusion of cortical RFs. At the cellular level, reorganization was being driven by an activity-dependent process of synaptic strengthening (reminiscent of processes during development). In the visual system RF expansion was accompanied by a selective increase in the strengths of intracortical connections [197]. By monitoring the strength of cross-correlations between pairs of neurons through different stages of RF expansion, it was found that populations of neurons increased the strength of their effective synaptic interconnections in the expanded regions of their RFs. Most of the cross-correlograms, recorded with electrodes separated by 0.1–5 mm, had widths of 5–15 ms and were asymmetric, typical for interactions observed between pairs of V1 neurons in different cortical columns. The increase in effective connection strength was also orientation selective and only occurred between pairs of neurons with similar orientation preferences. Neurons whose orientations differed by more than 30° did not show any increase in their mutual connection strength despite substantial increases and overlap in RF area. All the evidence is consistent with the idea that dynamic RF expansion is mediated through horizontal intracortical connections that link RFs of similar orientation preference. This was corroborated by lesioning studies [136]. Increases in neural synchrony between neurons in the lesion zone were found in auditory cortex immediately after a noise trauma [670] and several weeks to months after the trauma in reorganized cortex [668]. The increased synchrony was specific to reorganization because when reorganization was prevented by targeted acoustic stimulation of the frequency range of the hearing loss, but with hearing loss still present, there was no change in neural synchrony [667, 668]. On the other hand, inducing cortical reorganization in the absence of hearing loss both increased neural synchrony and strengthened horizontal fiber synapses with pyramidal cells in the reorganized area [669].

Behavioral Consequences of Topographic Map Changes Resulting from Receptor Injury. Adult cortex, and potentially also the thalamus, is highly plastic and can change as a result of learning (adaptive plasticity) and as a result of receptor injuries. The map changes that occur may or may not be related to the increased performance after learning [120, 757], and the map changes that

occur after peripheral injury may have mostly maladaptive consequences, albeit that some increased performance has been noted. Specifically this was after putative reorganization in auditory cortex where the overrepresentation of the edge frequency was related to improved frequency discrimination [607]. The maladaptive consequences appear to be dominating, as suggested by the correlation of phantom limb sensations and tinnitus with cortical map changes [95, 282, 642]. Assuming that reorganized adult cortex remains plastic, it should be possible to reverse the changes. Because of potential subcortical influences, this should, likely have been done by interference at or close to the receptor level. One successful approach has been to restore the balance of output from the cochlea resulting from high-frequency hearing loss [667]. This was accomplished by continuous stimulation in the frequency region of the noise-induced hearing loss, to compensate for the downregulated spontaneous and driven firing rates. After such a hearing loss, one typically observes reorganization of the cortical tonotopic map, accompanied by increased spontaneous firing rates and increased neural synchrony [666, 668]. The observations after applying the enriched acoustic environment for at least 3 weeks and starting immediately after the trauma included not only absence of cortical reorganization, normal spontaneous firing rates, and unaltered (i.e., normal) neural synchrony but also an improvement of the presumably neurotoxicity-induced hearing loss at frequencies above the hair cell loss frequency range. The sound was composed out of 32 series of tonepips (each for one of the 32 frequencies) with each series an independent realization of a Poisson process with rate of 1.5 Hz, and the co-occurrence of any combination of frequencies was thus completely random. Another take on the driving force of the changes in the cortical activation, besides restoration of the balance between the output from the normal and hearing loss regions in the cochlea, could be that the enriched acoustic environment produced a desynchronization of the inputs to the (to be) reorganized cortex and thus of the activity of that cortical region. There is some doubt about the possibility to use this sound treatment to reverse long-standing reorganizations, that is, those involving axonal sprouting and the formation of new connections. This would suggest that any treatment has to start well before this sprouting process starts.

1.10 DISCUSSION

In this chapter we have reviewed many correlative neural mechanisms and the important roles of correlation in perception, detection, memory, and sensorimotor coordination. We discussed how correlation is detected in sensory inputs in both the single neuron and neural populations and how it is employed in neural systems to encode and transmit information. Specifically, temporally correlated neuronal spike activities emerge from various interactions, such as correlation of firing of neighboring neurons or neural assemblies or synchronous firing of neurons from different subcortical or cortical areas (e.g., retinal ganglion cells, LGN cells, and V1 cells) [16, 833]. We conclude by the observation that correlation is the brain's "basic mechanism to evaluate, to control, and to learn" and "the basis of learning,

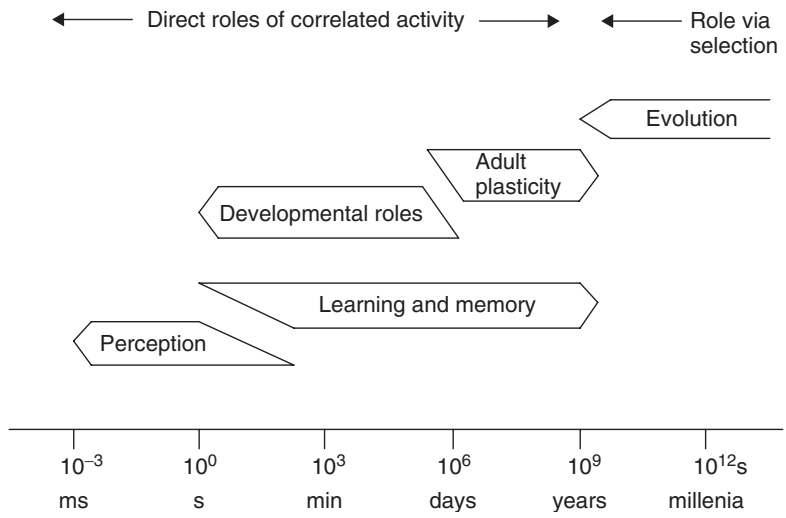


Figure 1.17 An illustrative diagram of timescales of the various roles of correlated activities within the central nervous system of human being. (Reprinted from *Trends in Neuroscience,* Vol. 14, J. E. Cook, Correlated activity in the CNS: A role on every timescale? pp. 397–401. Copyright © 1991, with permission from Elsevier.)

association, pattern recognition, and memory recall in the human nervous system" [241]. In addition, correlation also occurs at both macroscopic and microscopic timescales. According to Cook's categorization [183], correlated activities emerge in the central nervous system on every timescale, ranging from the momentary to the evolutionary (see Figure 1.17).

Learning and synaptic plasticity are the keys to the brain's ability to adapt to the changing environment. Motivated by this neurobiological evidence, we have mentioned several of the most biologically plausible learning algorithms, including Hebbian learning, competitive learning, self-organizing feature maps, STDP, Rescorla–Wagner rule, and TD learning. Despite their different motivational roots, all of these learning rules are based on the principle of correlation. The competitive learning and self-organizing map models share with Hebbian learning the principle that the synaptic strength should change in proportion to presynaptic and postsynaptic correlation. In the case of STDP, the synapse increases its efficacy whenever its input activity correlates with a subsequent change in neuronal output firing. The Rescorla–Wagner and TD learning rules are based on a prediction error signal, specifically, the difference between received and expected reinforcement. When a neural input is correlated with a prediction error, synaptic strengths are adjusted so as to reduce the prediction error. This idea can be generalized to allow for learning based on other prediction error signals. For example, the Kalman filter [461] is a standard method in engineering for training a system to predict a temporally varying signal, and the prediction error drives the parameter adjustment process. Computational neural models based on Kalman filtering are described in

greater detail in Chapter 7 of this book. Still more generally, the correlation-based ALOPEX learning algorithm [355, 902] is based on the correlation between an error signal and the parameter changes. Here, the error signal could be any analytically computable cost function. The basic idea is that if a change in a parameter value correlates with an increase in error, then the parameter is more likely to be changed in the opposite direction in the future, whereas a parameter change correlated with a decrease in error is more likely to be repeated in the future. Detailed discussions of the ALOPEX paradigm are presented in Chapter 6 of this book.

Finally, to conclude this chapter, we note that it is the correlative nature of the brain that motivates us to study the notion of correlative learning; moreover, we suggest that the principle of correlation might serve the role of encompassing most (if not all) learning paradigms, many of which will be reviewed in details in Chapter 3.

BIBLIOGRAPHICAL NOTES

For a general background on spiking neurons, textbook treatments can be found in [320, 493]. The classic books [171, 201] serve as excellent sources and references on neuroscience and computational brain models.

Synaptic plasticity is referred to as an adaptive response of the neuron to specific external (stimulus) signals; it is closely related to the notion of "learning." The notion of synaptic plasticity has a long history. The first mention of phenomena that can now be related to modifications in synaptic strengths or number of connections in neural networks can be found in René Descartes' *Traité de l'homme* (*Treatise on Man* (1664) [212], in which he described the concept of a hydraulic nervous system. William James, in his classic work on psychology [436], proposes the law of neural habit, which is also reminiscent of the ideas of Herbert Spencer [844] and Young [990]. In 1949, Donald Hebb published his famous learning postulate in *Organization of Behavior: A Neuropsychological Theory* [377], in which he first hypothesized the correlative mechanism for synaptic plasticity. A general review of Hebb's work can be found in [628, 816]. Hebbian synaptic mechanisms have been reviewed from biological [89, 472], biophysical [122], and physiological [66, 855] perspectives. The discussion of Hebbian synaptic plasticity in hippocampus is found in [472, 584, 817].

The milestone of Hebbian plasticity was further established by the discovery of the phenomenon of LTP [100]. The review of LTP and its relation to Hebbian synaptic plasticity can be found in [121, 586]. The review of spike-timing-dependent Hebbian plasticity can be found in [89].

The notion of the correlative brain was first proposed by [922] and reviewed in more detail in [241]. While it is impossible to include all the bibliographical references regarding the omnipresent importance of correlation and synchronization in population coding, topographic map formation, perception, memory, sensorimotor learning, attention, and feature binding, we will refer the reader to relevant references when discussing specific contents throughout the book.

NOTES

1. The term synapse was first used by Charles S. Sherrington [829] to designate the functional junction between nerve cells.
2. Caianiello [132] also proposed a learning rule known as the *mnemonic equation* for describing the synaptic plasticity. The learning rule, appearing in the form of a differential equation, is given as

$$\frac{d\theta_{ij}}{dt} = \theta_{ij}(t)\left[\alpha x_j(t - k\tau)x_i(t) - \beta\,\mathrm{sgn}(\theta_{ij}(t) - \theta_{ij}(0))\right]\mathrm{sgn}(a_{ij}(t) - \theta_{ij}(0)),$$

   where α and β are constants.
3. The term *self-organization* was first used by Farley and Clark [272] of MIT Lincoln Laboratory in 1954 while they studied the behavior of networks with nonlinear elements. Self-organization was used also by analogy to physics. In physical systems of densely packed units, where the activity of adjacent units is mutually influencing, several self-organizing principles of self-organization are particularly important [929]:

   - Local interactions tend to self-amplify. Synchronous or correlated interactions among coupled units strengthen and spread across ensembles of units, creating coherent activity patterns.
   - Developing activity patterns compete. The strongest (most coherent) patterns vigorously grow at the expense of others. This leads to the formation of activity "domains" of different self-amplifying patterns.
   - Domains of activity tend to cooperate. In spite of the overall competition in the system, domains of correlated activity will tend to coalesce to form larger, coherent activity patterns. If there are no outside influences acting on the system, the activity patterns with the most internal cooperativity and the least competition will win out.

   Von der Malsburg and Singer [929] noted that a fundamental correlate of these three basic principles is that "global order can arise from local interactions . . . ultimately leading to coherent behavior" (p. 71). In a large self-organizing network, a number of competing local domains can coexist, but the tendency of the network is generally toward attaining a globally ordered state. Because of this, even a relatively weak stimulus toward global organization can decisively influence developing local patterns.
4. The notion of RF is a cornerstone in visual physiology. According to H. K. Hartline (1938). The response of single optic nerve fibers of vertebrate eye to illumination of the retina. American Journal of Physiology 121. 400–415. It is "the region of the retina that must be illuminated in order to obtain a response in any given fiber." Nowadays, this notion has been widely extended and generally understood as "the area in which stimulation leads to response of a particular sensory neuron."
5. Sherman and Koch [828] have found that in the cat there are roughly 10^6 fibers from LGN in the thalamus to the visual cortex and about 10^7 fibers in the reverse direction.
6. The term *hippocampus* derives from Greek mythology with the meaning of "horselike sea monster." In anatomy it was named because of its curved "seahorse shape"; the "hippocampus" was first used by the anatomist Giulio Cesare Aranzi (circa 1564) for describing this brain region.

7. William James [436] also emphasized the principle of association that governs activation:

   > The amount of activity at any given point in the brain-cortex is the sum of the tendencies of all other points to discharge into it, such tendencies being proportionate (i) to the number of times the excitement of each other point may have accompanied that of the point in question; (ii) to the intensity of such excitements; and (iii) to the absence of any rival point functionally disconnected with the first point, into which the discharges might be diverted.

   These words clearly capture the essence of (i) Hebbian synapse, (ii) presynaptic neural activations, and (iii) the role of inhibition and anti-Hebbian synapse.
8. In Donald Hebb's book, *The Organization of Behavior*, there were three pivotal postulates in the context of synaptic plasticity and learning [377]:

   - The first one is the most celebrated Hebb's postulate formulated as Hebb's rule (p. 62), which provides the basis for adjusting connection weights in biological or artificial neural networks.
   - The second postulate speculates that groups of neurons that tend to fire together form a cell assembly whose activity can persist after the triggering event and serves to represent it (illustrated in Figure 10 of Hebb's book, p. 72).
   - The third postulate states that thinking is the sequential activation of sets of cell assemblies.

9. The phenomenon of LTP was first observed experimentally by Terje Lømo in 1966. Subsequent physiological experiments also validated another phenomenon called long-term depression (LTD), which refers to the weakening of a synapse that lasts from hours to days, as a counterpart of the LTP.