

Temporal Codes and Computations for Sensory Representation and Scene Analysis

Peter A. Cariani

Abstract—This paper considers a space of possible temporal codes, surveys neurophysiological and psychological evidence for their use in nervous systems, and presents examples of neural timing networks that operate in the time-domain. Sensory qualities can be encoded temporally by means of two broad strategies: stimulus-driven temporal correlations (phase-locking) and stimulus-triggering of endogenous temporal response patterns. Evidence for stimulus-related spike timing patterns exists in nearly every sensory modality, and such information can be potentially utilized for representation of stimulus qualities, localization of sources, and perceptual grouping. Multiple strategies for temporal (time, frequency, and code-division) multiplexing of information for transmission and grouping are outlined. Using delays and multiplications (coincidences), neural timing networks perform time-domain signal processing operations to compare, extract and separate temporal patterns. Separation of synthetic double vowels by a recurrent neural timing network is used to illustrate how coherences in temporal fine structure can be exploited to build up and separate periodic signals with different fundamentals. Timing nets constitute a time-domain scene analysis strategy based on temporal pattern invariance rather than feature-based labeling, segregation and binding of channels. Further potential implications of temporal codes and computations for new kinds of neural networks are explored.

Index Terms—Autocorrelation, multiplexing, neural networks, neural timing nets, temporal coding.

I. INTRODUCTION

THE project of understanding how the brain works as an information processing device is a exercise in reverse-engineering: we seek to understand the functional principles by which an highly effective system of unknown design achieves its goals. In order to proceed, we need some grasp of its functional capabilities, its internal structural organization, the input–output behavior of its processing elements, the means by which internal signals represent information, and the nature of information-processing operations that are carried out on those signals [1]. Once functional principles are understood, artificial devices can be designed that rival and exceed their biological counterparts.

The neural coding problem involves identification of dimensions of neuronal activity that play particular functional roles: which aspects subserve which informational functions. We will consider basic types of temporal pulse codes and how they convey information. We define a neural code by its functional role in the system: which changes in neural activity constitute differences in information content that are used by the rest of

the system to guide behavior [2], [3]. Our aim is to consider representations and computations that different temporal codes afford, rather than respective information transmission capacities [4] or underlying biophysical mechanisms [5]. Temporal codes can be used for representation of stimulus properties, localization of sources, and perceptual grouping.

In Sections II–V, various types of temporal codes are outlined, followed by examples from sensory neurophysiology. Section VI discusses temporal codes for signal multiplexing and perceptual grouping. Latter sections present examples of neural timing nets for temporal processing, and contemplate future possibilities.

II. TYPES OF TEMPORAL CODES

Pulse codes afford many different means of encoding information. Many catalogues and taxonomies of possible neural pulse codes have been proposed [2], [3], [6]–[13]. Although many are extensive, none covers all possible codes. Some examples are depicted in Fig. 1(A)–(K). The most basic distinction to be made is between channel-based codes and temporal codes.

Channel-based codes convey sensory distinctions through patterns of neural channel activations. The neural “labeled line” channel determines the nature of the information conveyed. In rate-based codes [Fig. 1(A)], stimulus properties are encoded via differences in neural discharge rates. “Rate-place” schemes encode patterns and property-combinations through across-neuron profiles of average discharge rates (which neurons fire how frequently). Rate-place representations constitute the neural coding assumption that has historically predominated in the neurosciences and in neural networks.

In contrast to channel-codes, temporal codes [Fig. 1(B)–(G), (I)] convey distinctions through relative timings of spikes such that alterations of spike timings change functional meanings of messages conveyed. Temporal codes range from coarse temporal modulations of discharge rate [14]–[16] to temporal correlations between individual spikes. Time codes can serve as vehicles for representation of stimulus attributes [2], [17], [18] or for their grouping into perceptual objects [19]–[24].

Temporal codes can be divided by whether they rely on recurring temporal patterns of spikes in the same channel(s) or on relative timings spikes arriving through different channels. The simplest temporal pattern code is an interspike interval code [Fig. 1(C)] in which time durations between spikes, produced by the same neuron(s), convey information about stimulus properties [17]. More complex temporal patterns can represent multiple properties via multiple interspersed intervals [Fig. 1(D)], [25], [26] and interval sequences [Fig. 1(E)], [18], [27], [28].

Manuscript received June 9, 2003; revised December 30, 2003.

The author is with the Eaton Peabody Laboratory, Massachusetts Eye and Ear Infirmary, Boston, MA 02114 USA (e-mail: cariani@epl.meei.harvard.edu).

Digital Object Identifier 10.1109/TNN.2004.833305

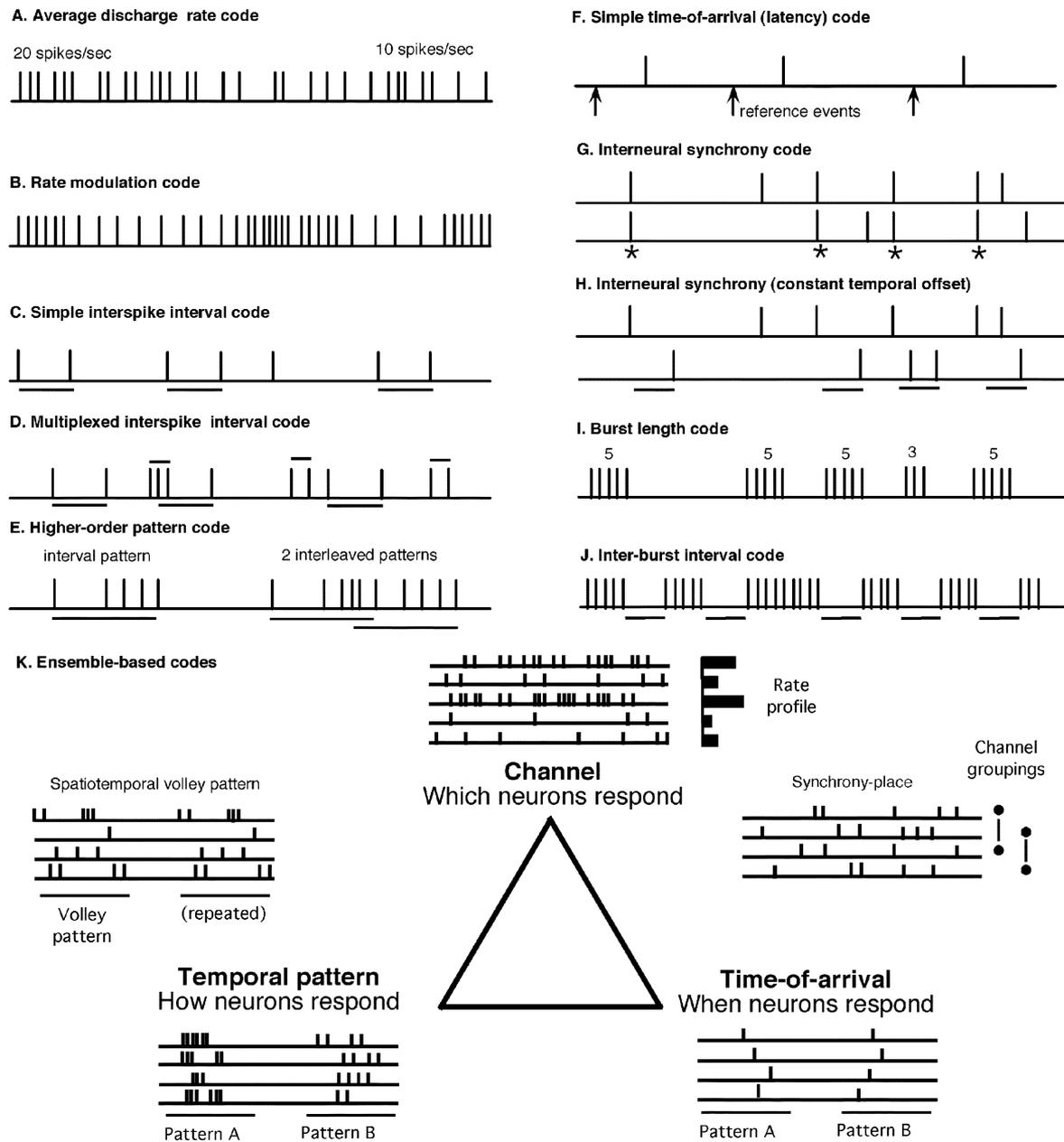


Fig. 1. Examples of neural pulse code schemes. (A) Rate-based channel code (# spikes in a given duration). (B) Rate modulation code (coarse temporal modulations of firing rate). (C) Simple interspike interval (ISI) code. (D) Multiplexing of two independent sets of intervals. (E) Complex temporal pattern code. (F) Simple time-of-arrival code (spike latency re: some reference event). (G) Interneuronal synchrony code (simultaneity). (H) Interneuronal synchrony (constant temporal offset). (I) Burst length code (nontemporal spike count code). (J) Interburst interval code (temporal pattern of burst events). (K) Combinations of neural ensemble-based channel, temporal pattern, and latency codes. Clockwise. Rate-channel scheme: rate profiles. Synchrony-place scheme: grouping of channels via common arrival time (spike synchrony). Time-of-arrival: two different latency-place patterns. Temporal pattern: two different ensemble-wide interval patterns. Spatiotemporal volley pattern (combination of all three code types): the same volley pattern shown twice.

Time-of-arrival or relative latency codes [Fig. 1(F)–(H)] use timing relations between spikes produced by different neurons. A well-known example [Fig. 1(H)] is found in binaural hearing, where time differences of sonic wavefronts reaching the two ears create different relative spike latencies in the two monaural auditory pathways. Here relative timing of spikes *across channels* rather than temporal structure *within channels* conveys information about source location. Since first spike latencies typically shorten with stimulus intensity [2], [29], [30], latency profiles (Fig. 1(K) lower right) can provide neuronal intensity cues. Time-of-arrival codes encompass interneuronal synchronies,

be they simultaneous [Fig. 1(G)] or delayed (Fig. 1(H), constant temporal offset).

While the taxonomy presented here is quite comprehensive in that it covers codes based on both spatial and temporal spike patterns, others stand outside its framework. For example, a burst length code [Fig. 1(I)] that counts spikes within a burst is neither a channel-code nor a temporal code. Similarly, an interburst interval code [Fig. 1(J)] relies on intervals between burst-events rather than between individual spikes. Still other schemes could rely on sequences of channel activations, spike times, and temporal patterns, i.e., ordinal rather than metrical

relations [31] Coarse temporal coding schemes intermediate between rate and spike correlation codes have also been proposed [10], [11]. Joint, statistical properties of neural ensembles, such as fractions of channels activated [32] or synchronized, variability of responses [2], [33], lie together with many other possible population-based codes [34], [35].

III. EXTRINSIC AND INTRINSIC TEMPORAL CODING MECHANISMS

Temporal codes are more common than is usually thought. Examination of the neurophysiology literature reveals stimulus-dependent temporal response patterns in early processing stages of virtually every sensory modality. Information-bearing temporal response patterns can be produced in two ways: extrinsically, through the temporal locking of neuronal discharges to stimulus waveforms (“phase-locking”), and intrinsically, through stimulus-specific triggering of endogenously generated time courses of response.

In many sensory systems, spike timings follow stimulus waveforms. Examples from auditory nerve fibers in the cat [36] and visual units in the monkey lateral geniculate nucleus [37] are shown in Fig. 2(A) and (B). Phase-locked responses also exist to mechanical vibrations of the skin [38], [39], to accelerations in the vestibular system, to inhalation cycles and odor pulses in olfaction [40], to self-produced electrical field oscillations and their external distortions in electroception [41], and to the movements of muscles via stretch receptors [42].

In the absence of stimulus locked responses, intrinsic temporal response patterns characteristic of particular stimuli, receptors, individual neurons, local neural circuits, and/or whole neural populations can bear information about the stimulus. Complex temporal response patterns are seen in the chemical senses. In early olfactory coding, many characteristic time courses of response depend upon both odorant and receptor [40], [43]–[45]. Time-delay neural nets for artificial noses [46] have been developed to effectively exploit relative latency patterns to discriminate odors. In later processing timings of discharge relative to oscillating field potentials are thought to play a role in organizing odor percepts [43]. In the rat gustatory system, different classes of tastants elicit characteristic temporal response patterns [Fig. 2(C)] that are associated with specific behavioral responses [47], [48]. Strikingly, playback of respective temporal patterns via electrical stimulation elicits appropriate behavioral responses in other rats, while scrambling of the patterns does not. In visual systems, intrinsic interspike interval patterns and coarse temporal modulations of firing rate result from differences in form, texture, and wavelength [14], [25], [49], [50]. The existence of subjective colors elicited by flicker-stimuli [51], [52] and temporally structured electrical stimulation [53] suggest the possibility of an intrinsic temporal code for color [12], [13].

IV. TIME-OF-ARRIVAL CODING OF STIMULUS LOCATION AND MOTION

Temporal patterns of activation across different sensory surfaces provide highly robust cues for stimulus direction. In audition, mechanoeception, and electroception, there may exist anal-

ogous mechanisms that make use of this cue to translate temporal differences into apparent location [2], [3], [41], [54], [55]. In all of these systems, receptors phase-lock to their respective adequate stimuli, such that the temporal structure of the stimulus is faithfully impressed on the timings of spikes produced by primary sensory neurons. By virtue of phase-locking, relative stimulus arrival times at different receptor sites are translated into relative spike latencies across sensory pathways. Relative spike timings can then in turn analyzed via neural delays and temporal coincidence detectors.

A well-known example is the localization of sounds in the azimuthal plane by means of interaural time differences (ITD) [56], [57]. Humans are able to distinguish differences of ITD as small as $\sim 20 \mu\text{s}$. Wavefronts from sound sources not directly in front of an observer arrive at the two ears at different times. These wavefronts produce phase-locked spikes in auditory nerve fibers whose relative timings reflect the interaural time differences. In the auditory brainstem, highly secure synapses, tapped delay lines, well-timed inhibitory inputs [5], and neural coincidence detectors in effect implement binaural cross-correlation operations that provide a readout of interaural time delays, and consequently, of azimuth estimates [41], [57], [58].

Analogous processing mechanisms may exist in mechanoeception, where relative delays of mechanical stimulation at different skin locations of a ms or more manifest themselves perceptually as changes in apparent location of stimulation [54]. As in audition, the apparent location moves toward the sensory surface that leads in time. Mechanoeceptive receptors phase-lock to skin deformations and hair displacements [38], [39], such that relative timings of neuronal discharges reflect relative timings of the stimulus arrival at different surfaces.

Active sensing involves observation of effects of actions on external environments. In bat and cetacean echolocation, acoustic signals are emitted and compared with echo patterns that contain time delays that provide information about distances and shapes of objects. In bat echolocation, time delays on the microsecond scale are registered by the relative timings between spikes produced by cries and their echoes [59], [60]. In some bat systems low-frequency modulations caused by interactions between the emitted signal and its Doppler shifted echo provide potential temporal cues for relative velocity (e.g., target flight speed and fluttering wingbeats) [59]. Another time-based active sensing strategy is found in electroception [41], [61]. Weakly electric fish produce sinusoidally-varying electrical fields around their bodies that are deformed by the presence of nearby external objects. These deformations alter the relative phases of the electric field at different body locations, which alter the relative latencies of spikes produced in afferent electroceptive pathways. As in the binaural example, these pathways have highly secure, low jitter connections, neural delays, and central coincidence detectors that permit use of sub-microsecond time-of-arrival differences.

Visual receptor arrays can be considered as collections of receptor surfaces. Phase-locking to temporal modulations of luminance produced by moving spatial patterns is ubiquitous in the visual systems of animals. As a consequence, temporal correlations between spikes produced in different visual channels

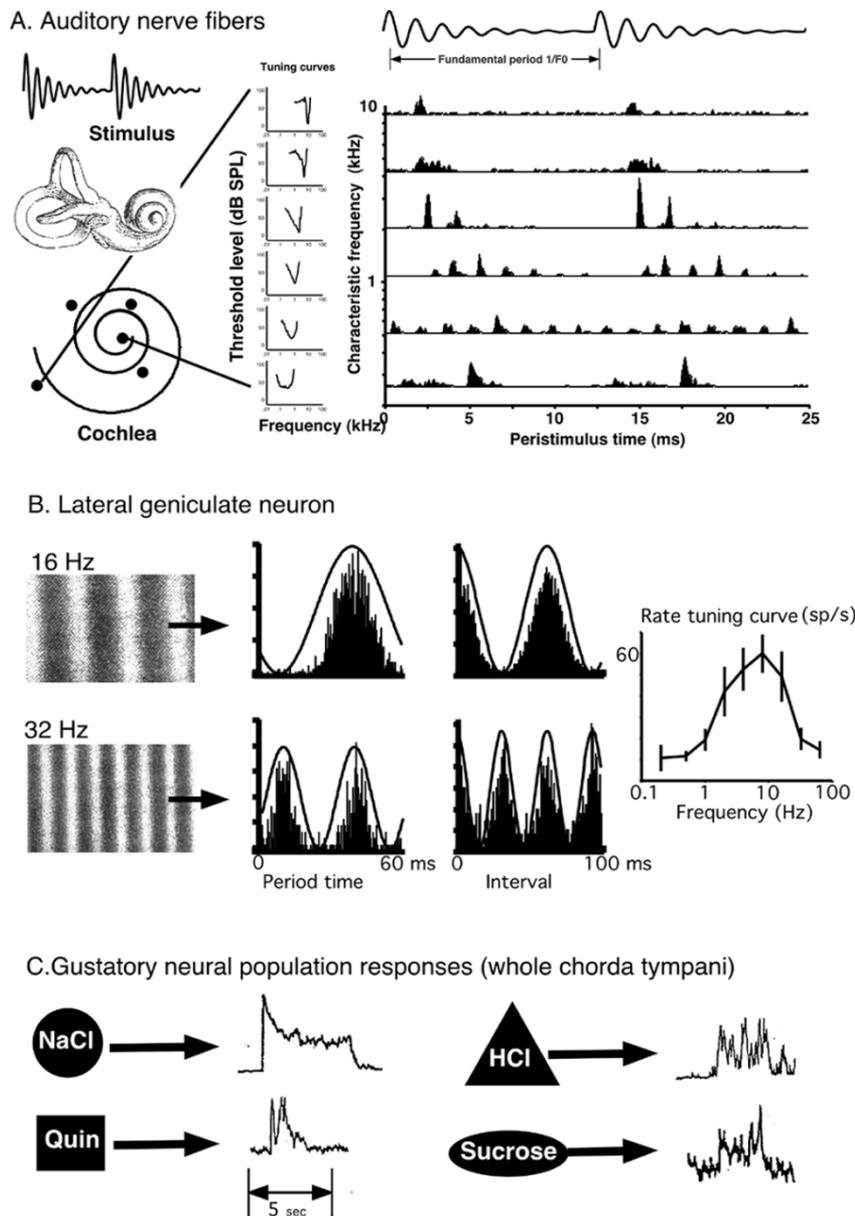


Fig. 2. Extrinsic and intrinsic temporal codes in early sensory processing. (A) Phase-locking of six cat auditory nerve fibers to a low-frequency harmonic complex tone. Rate-threshold tuning curves and poststimulus time histograms reflect effects of both cochlear filtering and the fine time structure of the stimulus. (B) Phase-locking of a visual neuron in the lateral geniculate nucleus of an anesthetized monkey in response to drifting gratings corresponding to sinusoidal temporal luminance modulations of 16 and 32 Hz. Period histograms and all-order interspike interval distributions are shown (16 Hz: 1233 spikes, 5897 intervals; 32 Hz: 495 spikes, 1102 intervals). Rate-based spatial tuning curve for the unit (mean \pm sd). Data courtesy of Przybylski & Pollen. (C) Intrinsic time courses of response in the gustatory system to four tastants of different classes: 0.1M NaCl (salty), 0.1M quinine (bitter), 0.1M HCl (sour), 0.5M sucrose (sweet). Waveforms are typical whole-nerve responses recorded from the chorda tympani of decerebrate rats. From Covey (1980).

potentially provide a general neurocomputational basis for the representation of visual motion. In the fly visual system, different spike timings in neighboring ommatidia are used for detection of motion [62]–[64]. Temporal crosscorrelation of small numbers of spikes permits rapid and precise motion estimations that inform flight course corrections in as little as 30 ms [63], [64].

Visual forms might also be encoded through temporal correlations between spikes produced across different retinal locations. Most vertebrate eyes are in constant drift even during fixation, and, many central visual neurons are known to discharge with relatively precise latency when contrast gradients (edges) cross their receptive fields. Moving images create

temporal modulations of luminance to which visual neurons phase-lock [Fig. 2(B)]. As in the auditory nerve, application of a temporally modulated stimulus converts a spike arrivals from a Poisson-like process to one dominated by the stimulus time structure. When sinusoidally-varying spatial gratings are drifted across receptive fields at constant velocities, at each retinal location spatial frequencies are converted to temporal frequencies of luminance modulation. Spike timings, as analyzed through peri-stimulus time (PST), period, and all-order interval histograms [Fig. 2(B)], faithfully replicate the temporal structure of the drifting image [65]. Temporal frequency can, thus, be accurately estimated from interspike interval statistics. For the lateral geniculate unit shown, temporal modulation

frequencies of 16 and 32 Hz were estimated to within 0.5% and 2%, respectively, of their true values on the basis of 5897 and 1102 all-order intervals. Given the coarseness of rate-based modulation tuning (rightmost plot), it is obvious that the spike timing information provides much finer precisions.

Similar stimulus-locking is seen for other visual stimuli. Many neurons in early stages of visual processing phase-lock to visual flickers of 50–100 Hz, but the relationship between perceptual flicker fusion limits and phase-locking is not simple [66]. Highest observed precisions of stimulus-driven spike timings in visual systems range from hundreds of microseconds to a few ms [63], [64], [67]–[69]. Stimulus-dependent synchronizations of spikes produced by different spatially remote neurons on the order of a few ms are also seen [70].

If the visual system used stimulus-driven spatial patterns of temporal correlations to register edges, then the disappearance of visual forms when an image is stabilized on the retina [71] could be readily explained. When images are stabilized, luminances are temporally unmodulated such that spike generation at each retinal position reverts to a Poisson-like process [72]. Temporal correlations of spikes across retinal positions disappear, as would spatial form information. The hypothesis is also consistent with recent psychophysical demonstrations that visual forms can be created through short-term spatiotemporal correlations that have no long-term spatial structure [73]. Such a hypothesis also potentially explains the hyperacuity problem in vision [64]—why visual acuity far exceeds the coarse spatial resolutions of individual receptive fields. It has been estimated from vernier acuity experiments with moving bars [74] that stimulus-driven spike timing with 1 ms jitter would be sufficient to account for observed performances, which lies within the realm of observed spike timings of visual units [75]. It also raises the possibility of spike-based spatiotemporal representations that couple to theories of visual form based on spatial auto and crosscorrelations [32], [76]–[78].

V. PHASE-LOCKED CODING OF STIMULUS WAVEFORMS

To the extent that sensory receptors follow the time structure of their adequate stimuli, that time structure is impressed on the timings of discharges of sensory neurons. As we have seen above, comparisons of the relative timings across receptors yields information about stimulus direction, movement, and perhaps, spatial form. For many kinds of stimuli, sensory qualities are related to the internal time structure of the stimulus waveform. Stimulus-locked temporal patterns of spikes in sensory neurons reflect this internal structure and, thus, provide information about the temporal form of the stimulus.

In the auditory system, phase-locked neural timing information is thought to be critical for most frequency discrimination [79]–[81]. Phase-locking of auditory nerve fibers to acoustic stimuli creates time intervals between spikes (interspike intervals) that are directly related to stimulus periodicities [Figs. 2(A) and 3(B)]. Distributions of interspike intervals from auditory neurons consequently contain information about the stimulus waveform and power spectrum for frequencies up to the limits of phase-locking. Phase-locking is maintained at virtually all sound pressure levels above response thresholds

such that interspike interval representations easily account for high precisions of auditory frequency discriminations over wide ranges of frequencies and sound pressure levels. As in binaural hearing, the time differences corresponding to the finest frequency discriminations (0.2% for pure tones near 1 kHz) are in the vicinity of 20 μ s. The progressively poorer ability of human listeners to distinguish higher frequencies from 2–10 kHz covaries with the progressive decline in phase-locking (and the quality of interval-based information) at those frequencies [79], [81]–[83]. Likewise, the frequency region for which there is good phase-locking, roughly up to 4 kHz, is coextensive with the existence region of musical tonality (octave matching, musical interval and melodic recognition).

Pitches of complex tones are explicable in similar terms. Harmonic complex tones produce pitches at their fundamentals (F0s), even in the absence of any spectral energy at that frequency (“pitch of the missing fundamental”). These low pitches of complex tones predominate in speech and music (they are the voice pitches we hear over telephones and cheap radios). In 1951, Licklider proposed a duplex time-delay neural network that operated on phase-locked auditory nerve responses to compute both rate-place and temporal autocorrelation representations of stimulus power spectra [81], [84], [85]. Decades later, the global interspike interval distribution of the auditory nerve was proposed as a candidate temporal representation for the pitch of complex tones. The general hypothesis has been confirmed in computer simulations [86] and in neurophysiological studies (Fig. 3) [36], [87]. With very few exceptions, the most common interval in the auditory nerve at any given time [Fig. 3(F)] corresponds to the pitch that is heard. On the basis of several thousand intervals, the fundamental pitch can be reliably estimated (within 1%). The fraction of pitch-related intervals *qualitatively* corresponds to the strength of the pitch that is heard. Such purely temporal, population-interval representations also account for many complex and subtle pitch phenomena: pitches of “missing fundamentals,” pitch equivalence classes, relative invariance of pitch and pitch salience with sound pressure level, pitches of stimuli having psychophysically-unresolved harmonics, the “nonspectral” pitches of amplitude modulated noise, as well as complex patterns of pitch shift that are heard for inharmonic stimuli. In effect, population-wide distributions of all-order interspike intervals form autocorrelation-like representations of the stimulus that in many (though not all) respects mirror stimulus autocorrelation functions. [87]; see [81] for recent developments and controversies. The close correspondences between the behavior of population-interval representations and psychophysically-observed patterns of pitch judgments provide very strong evidence that the interval information is utilized by the auditory system for pitch perception.

Purely temporal representations of vowel quality are also possible [80], [87]. Auditory nerve fibers phase-lock to frequency components in formant regions, consequently producing related intervals whose respective numbers depend on component intensities. Population-wide interval distributions consequently represent spectral energy distributions (formant patterns) through patterns of short intervals (0–4 ms) [85], [87]–[89]. These purely temporal representations of dominant

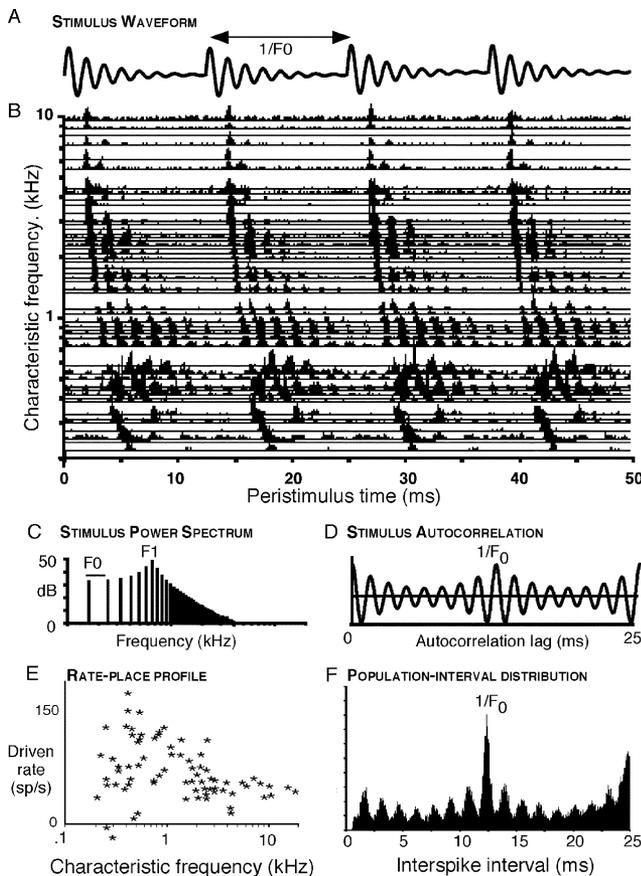


Fig. 3. Temporal coding of pitch in the auditory nerve. (A) Single-formant vowel waveform that produces a low pitch at its fundamental frequency ($F_0 = 80$ Hz). (B) Poststimulus time histograms of 52 cat auditory nerve fibers in response to 100 stimulus presentations at 60 dB SPL, arranged by characteristic frequency. (C) Stimulus power spectrum (log-frequency). (D) Stimulus autocorrelation function. (E) Population driven-rate profile (discharge rate—spontaneous rate). (F) Population-interval distribution (# all-order interspike intervals versus interval duration).

periodicity (pitch) and spectral energy distribution (timbre) have the high precision and robustness over large dynamic ranges that are characteristic of most auditory percepts. Their resilience to background noise [90] and sound separation capabilities [80] make them useful front-ends for artificial speech recognition systems.

The somatosensory analogue of auditory pitch is the sense of flutter-vibration [3], [38], [54]. Discrimination of frequencies up to 1 kHz appears to be based on interspike interval patterns produced through phase-locking of cutaneous afferents [39]. Complex tactile patterns produce corresponding temporal spike patterns [91] making an interval-based representations of tactile texture possible that parallel those for auditory pitch and timbre.

VI. TEMPORAL CODES IN SIGNAL MULTIPLEXING AND SCENE ANALYSIS

In addition to encoding sensory qualities themselves, temporal codes may be involved in perceptual grouping mechanisms. If object-properties are encoded via patterns of channel-activations, then the timing of those activations can be used to

sort them out. If object-properties themselves are encoded temporally, then commonalities of temporal pattern can be used as a basis for grouping.

Multiplexing involves the use of the same neural elements for concurrent representation and transmission of multiple signals. Multiplexing of information permits a given neuronal element to contribute to the representation of multiple perceptual objects, provided that there is a means of associating signals with particular objects. Three temporal strategies for binding ensembles of channels are time-division, frequency-division, and code-division multiplexing [Fig. 4(A)–(C)].

In time division multiplexing [Fig. 4(A)], the same channels participate in the encoding of multiple objects albeit at different times. If objects are represented by patterns of channels (e.g., feature detectors) then common spike timing in subsets of channels can be used as a principle for grouping [Fig. 4(B)]. The lines remain labeled to signify the features they encode, while the time domain is used to signify which channels are grouped or separated [23], [24]. The temporal label can involve either timing relative to a reference wave [92]–[94] or spike synchronization between channels [19]–[22], [95].

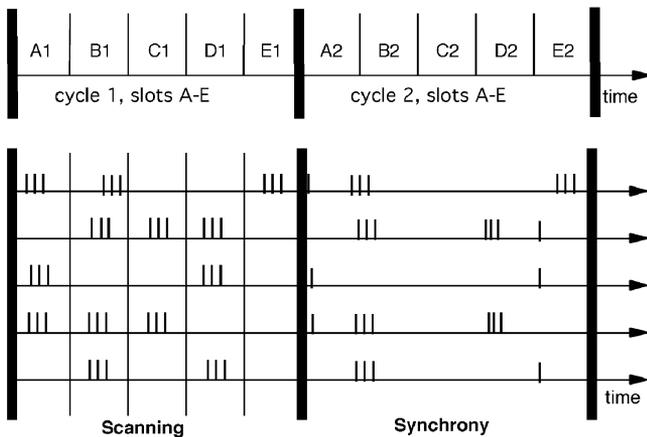
In frequency-division multiplexing [Fig. 4(B)], different signals utilize different frequency bands such that they can be mixed together in transmission channels and separated by receivers on the basis of their respective frequencies. In a pulse code, interpulse intervals encode different frequencies. The population-interval representation of low frequency sounds discussed above is a frequency-multiplexed system in that interspike intervals associated with different stimulus frequency components are conveyed by the same overlapping sets of auditory nerve fibers.

In code division multiplexing [Fig. 4(C)], patterns encoding features are not limited to any one frequency band. In the scheme shown in the figure, an initial header segment indicates signal-identity while a second segment conveys signal value. In a pulse code implementation, headers and value segments can be signaled by characteristic bursts consisting of different intra-burst intervals. The scheme here is similar to one proposed for multimodal encoding of cutaneous qualities [18]. Receivers selectively tuned to particular headers would process pulse trains segments with appropriate burst structure and ignore others. Multiple types of signals might then be asynchronously sent over each transmission line and demultiplexed by appropriate receivers.

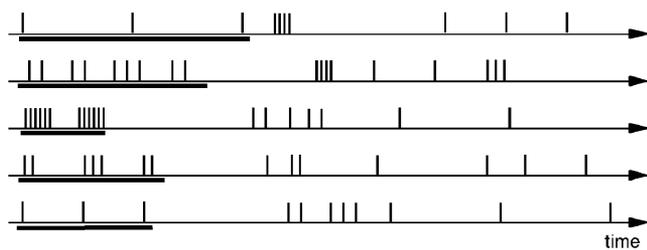
VII. NEURAL INFORMATION PROCESSING IN THE TIME DOMAIN

The near-ubiquity and robustness of spike timing information in early sensory representations begs questions concerning what neural temporal processing architectures are needed to fully exploit this information. The basic division between channel-based and temporal codes creates three broad classes of processing networks: connectionist networks, time-delay networks, and timing nets. Historically, time-delay networks have been used to convert analog, temporally-coded inputs into discrete patterns of channel-activations (e.g., [56], [84], [96]). However, neural networks can also be envisioned that carry out analog filtering and signal separation operations of sorts

A. Time-division multiplexing



B. Frequency-division multiplexing



C. Code-division multiplexing

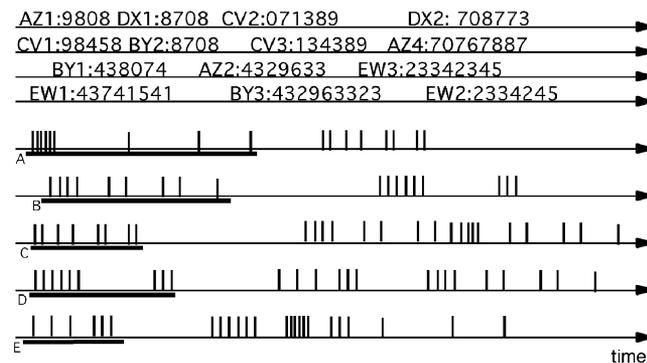


Fig. 4. Signal multiplexing: concurrent transmission of multiple signals over the same channels. (A) Time-division multiplexing (TDM) allocates particular time-slots to each signal (A-E). Neural pulse-coded TDM could use successive temporal windows or spike synchronies (B) Pulse-coded frequency-division multiplexing (FDM) maps signals onto different interspike intervals (bars). Multiple interval-types can then be transmitted over the same lines. (C) Code-division multiplexing encodes signal-type in the form of the signal, using header sequences that direct asynchronous and nonsequential transmission and reconstruction of the signal by each receiver. Code-division multiplexing using complex temporal pulse patterns as primitives. Initial burst encodes signal type, while subsequent spikes encode signal content. Patterns can be interleaved and transmitted asynchronously over the same channels.

familiar to signal processing engineers but relatively foreign to neuroscientists and neural network theorists. Using delay, multiplication (coincidence), and subtraction-cancellation [81] (inhibition) operations, neural timing networks carry out signal processing operations entirely in the time domain to compare, extract, and separate temporal patterns.

Feedforward timing nets (FFTNs) consist of arrays of idealized coincidence detectors that take inputs from multiple sets of tapped delay lines Fig. 5, [97]. Timing nets exploit the ability of axons to shift (delay) temporal patterns in time. Although structurally reminiscent of Jeffress and [56] and Braitenberg architectures [96], FFTNs have no subsequent “counting” or rate integration stage; their outputs are time rather than channel patterns. The computational properties of simple FFTNs with two sets of tapped delay input lines are straightforward. Spikes are fed in from either end of the array and propagate through their respective delay lines, crossing at different positions to produce output spikes. Each horizontal array position implements a particular relative delay between the input signals [Fig. 5(B)].

Several basic operations can be carried out [97]. First, the crosscorrelation function of the two inputs can be computed by counting the number of spikes in each output channel (vertical bar) as a function of relative delay (i.e., the Jeffress model). Convolutions can be computed by summing across relative delay channels for each time step [98], i.e., forming the collective peristimulus time histogram (PSTH) of the detector ensemble. The operation is similar to the common flip/shift/multiply method of computing convolution.

The “summary autocorrelation” of the outputs is the sum of the output channel autocorrelations (all-order interval distributions), analogous to the neural population-interval representations for pitch and timbre discussed above. Since the output summary autocorrelation of the FFTN is the product of the input autocorrelations, the coincidence array computes a cross spectral product in the time domain, from interval (lag) statistics. This multiplication of autocorrelations permits extraction of intervals related to common fundamentals (pitch) irrespective of the particular partials involved (timbre, vowel quality), and vice versa. Intervals associated with a common F0 shared by two signals can be extracted even if the signals have no harmonics in common.

A further consequence is that each interval or higher order spike arrival pattern (e.g., triplets) must be present in each of the inputs in order to appear in an output channel [Fig. 5(E)]. The array, thus, functions as a temporal sieve, passing those temporal patterns that are common to both sets of inputs, even when they are embedded in other spikes. These sifting operations are useful for computing intersections of sets of patterns, detecting embedded patterns, and for extracting patterns of interest from signals. Feedforward nets also separate temporal patterns on the basis of time-of-arrival disparities [Fig. 5(F)].

VIII. RECURRENT NEURAL TIMING NETS (RTNs)

Recurrent timing nets (RTNs) were developed to handle the detection and separation of periodic patterns. Periodic signals, such as those that produce strong pitches and rhythms, create strong temporal expectations. RTNs were inspired by time domain spike processing [98], stabilized auditory image architectures [99], reverberating circuits [100], neural loops [101], adaptive timing nets [102], Grossberg’s adaptive resonance circuits, and the psychology of temporal expectation. Concrete guidance has also come from auditory scene analysis: how perceptual attributes of sounds are organized into objects, voices, and streams

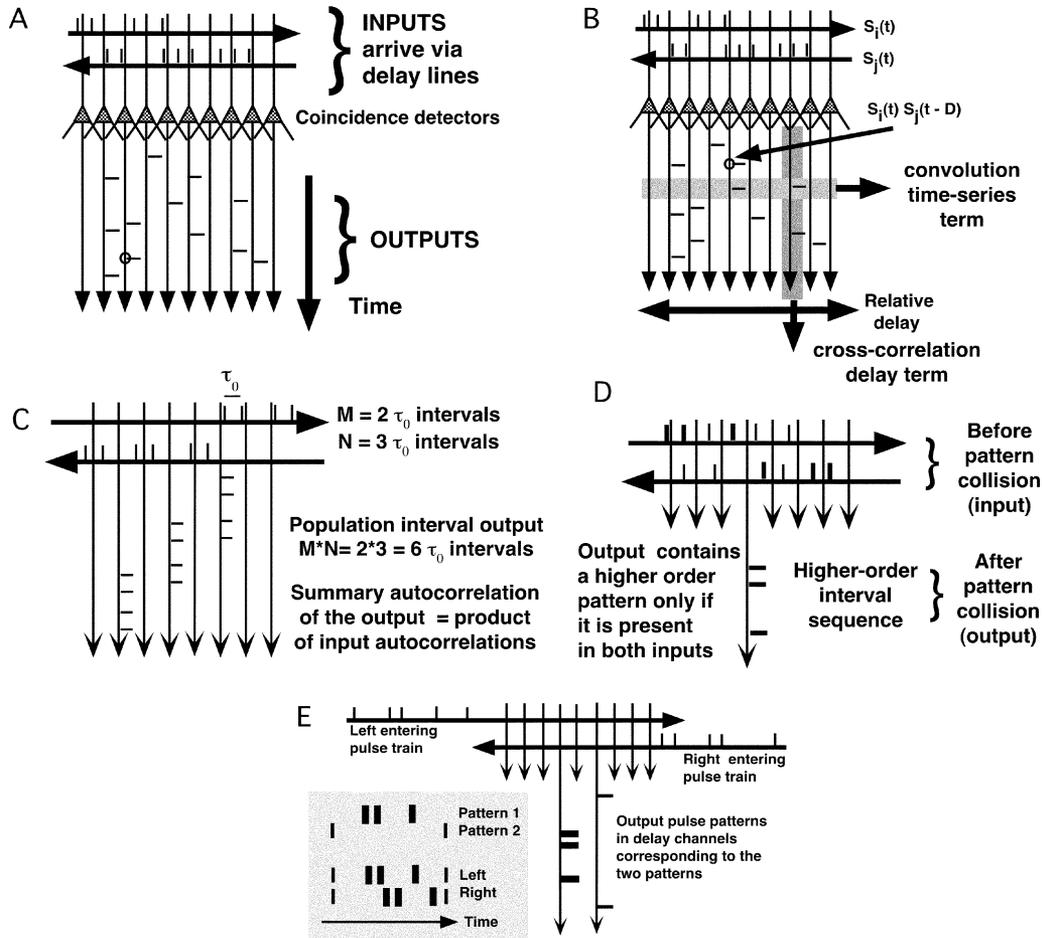


Fig. 5. Feedforward timing nets. (A) Basic structure. (B) Crosscorrelation by summing output pulses within channels at each relative delay. Convolution by summing across channels at each time step. (C) Summary autocorrelation of the output computes products of input autocorrelations. (D) Extraction of higher order temporal patterns containing interspersed spikes. (E) Separation of patterns (inset) with different relative delays or phases.

that can be separated from one another [103], [104]. Common periodicity (harmonicity) is a very strong grouping cue in audition. Although most low-frequency auditory perception is famously insensitive to the stationary phase-spectrum of the stimulus, transient phase and amplitude shifts are known to cause individual harmonics to separate from a complex. In the case of mistuned harmonics, which also readily separate, the phase of the mistuned component constantly precesses relative to that of the rest of the complex. In echolocation the neuronal mechanisms likely involve fine time comparisons between a signal (cry) and its repetition (echo) [60]. All of these examples suggest a process by which a signal waveform is compared with itself on a period-by-period basis.

In RTNs signals circulate in reverberating conduction loops, as temporal memory traces that interact with incoming temporal patterns. Neural representations then build up over time, with previous patterns dynamically creating sets of perceptual expectations that can either be confirmed and built up or violated. RTNs provide a basic mechanism for modeling echoic memory, separation of periodic sounds, rhythm induction, and creation of temporal expectancies.

Recurrent timing nets are in formative stages of development. The simplest recurrent timing networks consist of a 1-D array of coincidence detectors having common direct inputs [Fig. 6(A)].

The output of each coincidence element is fed into a recurrent delay line such that the output of the element at time t circulates through the line and arrives at the element at time $t + \tau$ [Fig. 6(B)]. In more recent implementations [105], a simple error-adjustment processing rule (1,2) governs the interaction of direct and circulating inputs

$$H(t) = H(t - \tau) + B_{\tau} [X(t) - H(t - \tau)] \quad (1)$$

$$B_{\tau} = \frac{\tau}{33 \text{ ms}}. \quad (2)$$

Here $X(t)$ is the direct input signal, $H(t - \tau)$ is the incoming circulating signal, and $H(t)$ is the outgoing circulating signal. τ is the recurrence time of the loop. B_{τ} determines the rate of adjustment and the length of the temporal processing window, which corresponds to the lower limit of pitch [81], [106].

In signal processing terms, RTNs most resemble neural time-domain implementations of comb filtering operations. Although the computations bear many similarities to correlogram-based segregation and oscillator-based channel allocations [95], by keeping processing entirely in the time domain, the RTN produces separated waveforms rather than ensembles of grouped frequency channels.

Some basic signal separation capabilities of RTNs have been tested on concurrently-presented vowel pairs (“double vowels”).

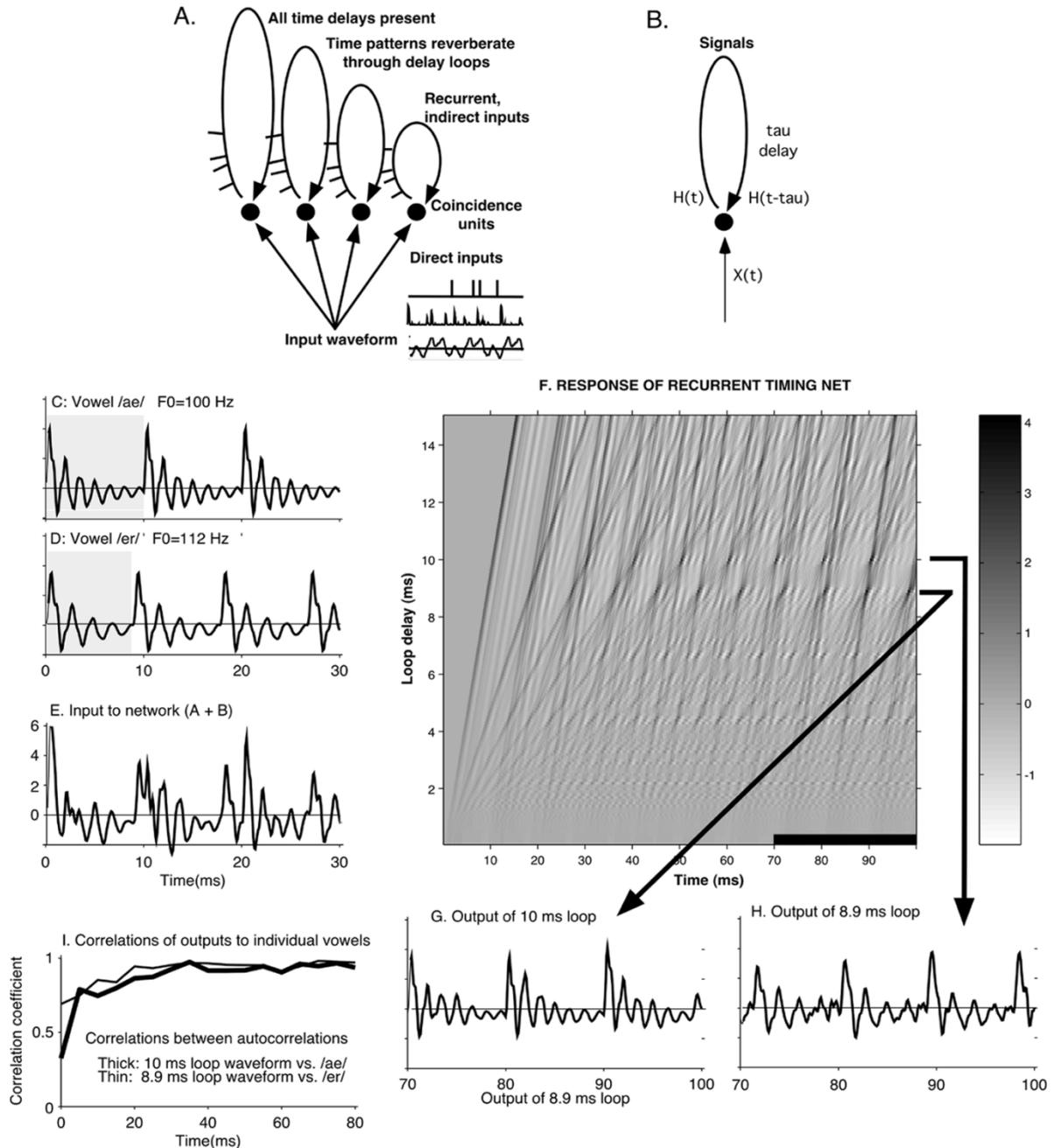


Fig. 6. Recurrent timing nets for F0-based separation of auditory objects. (A) Basic structure. (B) Processing nodes compare incoming and delayed (circulating) signals, and adjust output signal according to a linear error-adjustment rule. (C), (D) Waveforms of constituent single vowel waveforms. (E) Double vowel waveform. (F) Response of the network as a function of loop-delay and time. (G), (H) Circulating waveforms in the two most highly activated delay channels. (I) Resemblance of the autocorrelations (power spectra) of these two circulating waveforms to those of the individual constituent vowels.

A considerable body of psychophysical, neurophysiological, and neurocomputational work exists on their perception, [97]. When fundamentals are the same, the vowels are fused together, making the constituents harder to identify. When fundamentals are separated by a semitone (6%) or more, they are heard as two separate auditory objects, and can be identified with somewhat higher accuracy (15%–20% improvement).

A typical case is illustrated here (Fig. 6). Synthetic, three-formant double vowels (/ae/, /er/) with different fundamentals (100, 112 Hz, Fig. 8AB) were summed [Fig. 6(E)] and processed by a network with 150 delay loops that spanned 0–15 ms delays. Invariably, the two periodic patterns build up fastest in the delay loop whose recurrence time best matches their rep-

etition time [Fig. 6(F)]. Here the loops with the highest average signal strength correspond to the periods of the two vowels (8.9 and 10 ms). The signals circulating in these two delay channels after 70 ms of processing [Fig. 6(G) and (H)] highly resemble the two vowel constituents [Fig. 6(A) and (B)]. Correlations between the autocorrelations of separated signals and individual vowels [Fig. 6(I)] show how these resemblances increase as the signal separation unfolds over processing time. Multichannel RTNs with auditory nerve front-ends have also been explored, with similar results [97].

In effect, the loops dynamically create matched filters from repeating temporal patterns in the stimulus. Temporal-pattern invariances are then enhanced relative to uncorrelated patterns

in each loop, such that each loop functions as a pattern-amplifier for stimulus patterns whose period equals the loop's recurrence time. The network as a whole will amplify any periodic temporal patterns in its input and sort them out according to their periods. With longer sets of delays, such networks also separate multiple rhythm subpatterns [106].

RTNs illustrate how phase-locked fine timing information (and transient changes in phase) can be used to separate sounds. Most existing strategies for separating sounds on the basis of fundamental frequency attempt to group frequency channels together by finding F0-related features in each channel, e. g., [21], [22], [95], [107]. The RTN model demonstrates an alternate, time-domain strategy for auditory object separation that 1) exploits fine timing information and 2) uses no explicit feature detection (i.e., F0-detectors). Rather than labeling frequency channels for segregation, the delay loops amplify temporal pattern-invariances in the waveform. Auditory objects are, thus, separated on the basis of temporal pattern coherence [23]. Recurrent timing nets demonstrate how auditory images can be built up by comparing a signal with its immediate past [99]. They show how purely temporal representations and computations can effect separation and identification of auditory objects even when the information that constitutes them is temporally multiplexed in many of the same frequency channels. The nets provide a simple processing scheme in which object formation comes prior to analysis of auditory qualities (pitch, timbre) rather than as its consequence. They illustrate how auditory separation mechanisms might use transient changes in phase to form objects that are subsequently analyzed by phase-insensitive mechanisms for periodicity (pitch) and spectrum (timbre). In the context of speech separation and recognition, it may therefore make sense not to discard phase information in early stages of signal processing that may provide a basis for speaker separation.

Looking ahead, it is conceivable that judicious adjustment of synaptic weights and delays could result in networks capable of regenerating distributed temporal patterns—reverberating memories [100]. In randomly-connected nets, delay loops could be formed on the fly by short-term facilitations created by temporal correlations. The time structure of an incoming signal would dynamically organize central neural circuits so as to propagate and build up stable, reverberating patterns.

IX. POTENTIAL IMPLICATIONS FOR NEURAL NETWORKS

Temporal codes and computations may provide new avenues for artificial neural network development. Some of the general advantages of temporal codes stem from their ability to encode sensory information in a manner largely independent of stimulus intensity; this facilitates formation of intensity-invariant perceptual equivalence classes, which drastically simplifies pattern recognition. Signal multiplexing, mass statistical representations, and vectorial encodings may permit neural networks to handle information in more flexible ways that liberate signals from wires.

In a traditional connectionist network, all signals are scalars that are weighted and combined with many other signals at each node. The ability to temporally multiplex multiple independent

signals lessens their competitive interference. Networks can be envisioned in which neural assemblies tuned to respond to specific temporal patterns propagate both locally-relevant and irrelevant parts of neural signals. Inter-transparency of signals would then allow broadcast strategies of neural integration and coordination that would begin to resemble asynchronous, decentralized communications systems.

Temporal codes permit neural representations based on mass statistics of spike correlations. Population-interval representation of auditory qualities discussed above is a concrete example. If information can be encoded in temporal patterns of spikes, then processing can occur on population or ensemble-wide scales that obviate the need for particular elements, interconnections, synaptic efficacies, and transmission paths in order to function, a longstanding problem in theoretical neuroscience. Feedforward timing nets provide a simple processing strategy that obviates the need for precise point-to-point connectivities; although very primitive, they suggest strategies by which information might be processed on a mass-statistical basis without regard for which particular elements and connections are involved.

Finally, temporal codes support multiple, independent signal types. Independent signal types allow for vectorial, multidimensional representations in which temporal patterns related to different aspects of a situation or object might be flexibly superposed [20]. For example, characteristic temporal patterns might distinguish information from different sense modalities [18]. Tag systems based on orthogonal sets of temporal patterns might then permit representational compositionality and high-dimensional computations on them. Over the course of processing incoming information, neural assemblies would add specific annotational tags to circulating signals, such that the circulating temporal patterns could increase their effective dimensionality over time to create new categories.

REFERENCES

- [1] D. Marr, *Vision*. San Francisco, CA: W. H. Freeman, 1982.
- [2] D. H. Perkel and T. H. Bullock, "Neural coding," *Neurosci. Res. Program Bull.*, vol. 6, pp. 221–348, 1968.
- [3] W. R. Uttal, *The Psychobiology of Sensory Coding*. New York: Harper & Row, 1973.
- [4] D. M. MacKay and W. S. McCulloch, "The limiting information capacity of a neuronal link," *Bull. Math. Biophys.*, vol. 14, 1952.
- [5] A. Brand, O. Behrend, T. Marquardt, D. McAlpine, and B. Grothe, "Precise inhibition is essential for microsecond interaural time difference coding," *Nature*, vol. 417, pp. 543–7, 2002.
- [6] W. R. Uttal, *Sensory Coding: Selected Readings*. Boston, MA: Little-Brown, 1972, pp. 481–481.
- [7] F. Theunissen and J. P. Miller, "Temporal encoding in nervous systems: A rigorous definition," *J. Comput. Neurosci.*, vol. 2, pp. 149–62, 1995.
- [8] J. D. Victor and K. Purpura, "Metric-space analysis of spike trains: Theory, algorithms, and application," *Network*, vol. 8, pp. 127–164, 1997.
- [9] J. J. Eggermont, *The Correlative Brain: Theory and Experiment in Neural Interaction*. Berlin, Germany: Springer-Verlag, 1990, vol. 24.
- [10] W. Gerstner, "Spiking neurons," in *Pulsed Neural Networks*, W. Maas and C. M. Bishop, Eds. Cambridge, MA: MIT Press, 1999, pp. XIII–XXVI.
- [11] W. Maas, "Computing with spiking neurons," in *Pulsed Neural Networks*, W. Maas and C. M. Bishop, Eds, MA, Cambridge: MIT Press, 1999, pp. XIII–XXVI.
- [12] P. Cariani, "As if time really mattered: Temporal strategies for neural coding of sensory information," *Commun. Cogn. Artif. Intell.*, vol. 12, pp. 161–229, 1995.

- [13] ———, "Temporal coding of sensory information in the brain," *Acoust. Sci. Tech.*, vol. 22, pp. 77–84, 2001.
- [14] B. J. Richmond, L. M. Optican, and T. J. Gawne, "Neurons use multiple messages encoded intemporally modulated spike trains to represent pictures," in *Seeing Contour and Color*, J. J. Kulikowski and C. M. Dickenson, Eds. New York: Pergamon, 1989, pp. 705–713.
- [15] J. D. Victor and K. Purpura, "Nature and precision of temporal coding in visual cortex," *J. Neurophysiol.*, vol. 76, pp. 1310–1326, 1996.
- [16] B. J. Richmond and T. J. Gawne, "The relationship between neuronal codes and cortical organization," in *Neuronal Ensembles: Strategies for Recording and Decoding*, H. B. Eichenbaum and J. L. Davis, Eds. New York: Wiley, 1998, pp. 57–80.
- [17] V. Mountcastle, "The problem of sensing and the neural coding of sensory events," in *The Neurosciences: A Study Program*, G. C. Quarten, T. Melnechuk, and F. O. Schmitt, Eds. New York: Rockefeller Univ. Press, 1967.
- [18] R. Emmers, *Pain: A Spike-Interval Coded Message in the Brain*. New York: Raven, 1981.
- [19] C. M. Gray, "The temporal correlation hypothesis of visual feature integration: Still alive and well," *Neuron*, vol. 24, pp. 31–47, 1999.
- [20] W. Singer, "Neuronal synchrony: A versatile code for the definition of relations?," *Neuron*, vol. 24, pp. 49–65, 1999.
- [21] D. Wang, "The time dimension for neural computation," Dept. Comput. Inform. Sci., The Ohio State Univ., Columbus, OH, OSU-CISRC-09/02/TR21, 2002.
- [22] D. L. Wang, "An oscillatory correlation theory of temporal pattern segmentation," in *Neural Representation of Temporal Patterns*, E. Covey, H. L. Hawkins, and R. F. Port, Eds. New York: Plenum, 1995, pp. 53–76.
- [23] C. von der Malsburg, "The correlation theory of brain function," in *Models of Neural Networks II. Temporal Aspects of Coding and Information Processing in Biological Systems*, E. Doumany, J. L. van Hemmen, and K. Schulten, Eds. New York: Springer-Verlag, 1994, pp. 95–119.
- [24] L. Shastri and V. Ajanagadde, "From simple associations to systematic reasoning: A connectionist representation of rules, variables, and dynamic bindings using temporal synchrony," *Behav. Brain Sci.*, vol. 16, pp. 417–494, 1993.
- [25] G. S. Wasserman, "Isomorphism, task dependence, and the multiple meaning theory of neural coding," *Biol. Signals*, vol. 1, pp. 117–142, 1992.
- [26] S. H. Chung, S. A. Raymond, and J. Y. Lettvin, "Multiple meaning in single visual units," *Brain Behav. Evol.*, vol. 3, pp. 72–101, 1970.
- [27] R. Lestienne and B. L. Strehler, "Time structure and stimulus dependence of precise replicating patterns present in monkey cortical neuronal spike trains," *Brain Res.*, vol. 43, pp. 214–238, 1987.
- [28] M. Abeles, H. Bergman, E. Margalit, and E. Vaadia, "Spatiotemporal firing patterns in the frontal cortex of behaving monkeys," *J. Neurophysiol.*, vol. 70, pp. 1629–1638, 1993.
- [29] S. S. Stevens, "Sensory power functions and neural events," in *Principles of Receptor Physiology*, W. R. Loewenstein, Ed. Berlin, Germany: Springer-Verlag, 1971, pp. 226–242.
- [30] P. Heil, "Auditory cortical onset responses revisited. I. First-spike timing," *J. Neurophysiol.*, vol. 77, pp. 2616–2641, 1997.
- [31] J. P. Segundo, G. P. Moore, L. J. Stensaas, and T. H. Bullock, "Sensitivity of neurones in Aplysia to temporal pattern of arriving impulses," *J. Exp. Biol.*, vol. 40, pp. 643–667, 1963.
- [32] W. R. Uttal, *The Psychobiology of Mind*. Hillsdale, NJ: Lawrence Erlbaum, 1978.
- [33] G. Werner and V. B. Mountcastle, "The variability of central neural activity in a sensory system, and its implications for the central reflection of sensory events," *J. Neurophysiol.*, vol. 26, pp. 958–977, 1963.
- [34] L. Abbott and T. J. Sejnowski, "Introduction," in *Neural Codes and Distributed Representations: Foundations of Neural Computation*, L. Abbott and T. J. Sejnowski, Eds. Cambridge, MA: MIT Press, 1999, pp. 225–242.
- [35] S. A. Deadwyler and R. E. Hampson, "Ensemble activity and behavior: What's the code?," *Science*, vol. 270, pp. 1316–1318, 1995.
- [36] P. A. Cariani and B. Delgutte, "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. II. Pitch shift, pitch ambiguity, phase-invariance, pitch circularity, and the dominance region for pitch," *J. Neurophysiol.*, vol. 76, pp. 1698–1734, 1996.
- [37] A. Przybyszewski, J. P. Gasa, W. Foote, and D. A. Pollen, "Striate cortex increases contrast gain in macaque LGN neurons," *Vis. Neurosci.*, vol. 17, pp. 485–494, 2000.
- [38] W. Keidel, "The sensory detection of vibrations," in *Foundations of Sensory Science*, W. W. Dawson and J. M. Enoch, Eds. Berlin, Germany: Springer-Verlag, 1984, pp. 465–512.
- [39] V. Mountcastle, "Temporal order determinants in a somatosensory frequency discrimination: Sequential order coding," *Ann. New York Acad. Sci.*, vol. 682, pp. 151–170, 1993.
- [40] F. Macrides and S. L. Chorover, "Olfactory bulb units: Activity correlated with inhalation cycles and odor quality," *Science*, vol. 175, pp. 84–86, 1972.
- [41] C. E. Carr, "Processing of temporal information in the brain," *Annu. Rev. Neurosci.*, vol. 16, pp. 223–243, 1993.
- [42] M. A. Frye, "Encoding properties of the wing hinge stretch receptor in the hawkmoth *manduca sexta*," *J. Exp. Biol.*, vol. 204, pp. 3693–702, 2001.
- [43] G. Laurent, "A systems perspective on early olfactory coding," *Science*, vol. 286, pp. 723–8, 1999.
- [44] J. S. Kauer, "Response patterns of amphibian olfactory bulb neurones to odour stimulation," *J. Physiol.*, vol. 243, pp. 695–715, 1974.
- [45] G. Laurent and H. Davidowitz, "Encoding of olfactory information with oscillating neural assemblies," *Science*, vol. 265, pp. 1872–1875, 1994.
- [46] J. White, T. A. Dickinson, D. R. Walt, and J. S. Kauer, "An olfactory neuronal network for vapor recognition in an artificial nose," *Biol. Cybern.*, vol. 78, pp. 245–251, 1998.
- [47] E. Covey, *Temporal Neural Coding in Gustation*. Durham, NC: Duke Univ. Press, 1980.
- [48] P. M. Di Lorenzo and G. S. Hecht, "Perceptual consequences of electrical stimulation in the gustatory system," *Behav. Neurosci.*, vol. 107, pp. 130–138, 1993.
- [49] B. J. Richmond, L. M. Optican, M. Podell, and H. Spitzer, "Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics," *J. Neurophysiol.*, vol. 57, pp. 132–146, 1987.
- [50] W. M. Kozak and H. J. Reitboeck, "Color-dependent distribution of spikes in single optic tract fibers of the cat," *Vis. Res.*, vol. 14, pp. 405–419, 1974.
- [51] J. J. Sheppard, *Human Color Perception: A Critical Study of the Experimental Foundation*. New York: Elsevier, 1968.
- [52] L. Festinger, M. R. Allyn, and C. W. White, "The perception of color with achromatic stimulation," *Vis. Res.*, vol. 11, pp. 591–612, 1971.
- [53] R. A. Young, "Some observations on temporal coding of color vision: Psychophysical results," *Vis. Res.*, vol. 17, pp. 957–965, 1977.
- [54] G. von Békésy, *Sensory Inhibition*. Princeton, NJ: Princeton Univ. Press, 1967.
- [55] T. H. Bullock, "Signals and neural coding," in *The Neurosciences: A Study Program*, G. C. Quarten, T. Melnechuk, and F. O. Schmitt, Eds. New York: Rockefeller Univ. Press, 1967, pp. 347–352.
- [56] L. A. Jeffress, "A place theory of sound localization," *J. Comp. Physiol. Psychol.*, vol. 41, pp. 35–39, 1948.
- [57] S. Colburn, "Computational models of binaural processing," in *Auditory Computation*, H. Hawkins, T. McMullin, A. N. Popper, and R. R. Fay, Eds. New York: Springer-Verlag, 1996.
- [58] M. Konishi, "Deciphering the brain's codes," in *Neural Codes and Distributed Representations: Foundations of Neural Computation*, L. Abbott and T. J. Sejnowski, Eds. Cambridge, MA: MIT Press, 1999, pp. 1–18.
- [59] C. F. Moss and H.-U. Schnitzler, "Behavioral studies of auditory information processing," in *Hearing by Bats*, A. N. Popper and R. R. Fay, Eds. New York: Springer-Verlag, 1995, pp. 87–145.
- [60] J. A. Simmons, "Formation of perceptual objects from the timing of neural responses: Target-range images in bat sonar," in *The Mind-Brain Continuum*, R. Llinas and P. S. Churchland, Eds. Cambridge, MA: MIT Press, 1996, pp. 219–250.
- [61] W. Heiligenberg, "The coding and processing of temporal information in the electrosensory system of the fish," in *Temporal Coding in the Brain*, G. Buzaki, R. Llinas, W. Singer, A. Berthoz, and Y. Christen, Eds. Berlin, Germany: Springer-Verlag, 1994, pp. 1–12.
- [62] W. Reichardt, "Autocorrelation, a principle for the evaluation of sensory information by the central nervous system," in *Sensory Communication*, W. A. Rosenblith, Ed. New York: Wiley, 1961, pp. 303–317.
- [63] W. Bialek, F. Rieke, R. R. van Steveninck, and W. D. de Ruyter, "Reading a neural code," *Science*, vol. 252, pp. 1854–1856, 1991.
- [64] F. Rieke, D. Warland, R. de Ruyter van Steveninck, and W. Bialek, *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press, 1997.
- [65] D. A. Pollen, J. P. Gaska, and L. D. Jacobson, "Physiological constraints on models of visual cortical function," in *Models of Brain Function*, R. M. J. Cotterill, Ed. Cambridge, U.K.: Cambridge Univ. Press, 1989, pp. 115–136.
- [66] G. T. Plant, "Temporal properties of normal and abnormal spatial vision," in *Spatial Vision*, D. Regan, Ed. Boca Raton, FL: CRC Press, 1991, pp. 43–63.

- [67] S. P. Strong, R. Koberle, R. R. de Ruyter van Steveninck, and W. Bialek, "Entropy and information in neural spike trains," *Phys. Rev. Lett.*, vol. 80, pp. 197–200, 1998.
- [68] P. Reinagel and C. Reid, "Temporal coding of visual information in the thalamus," *J. Neurosci.*, vol. 20, pp. 5392–5400, 2000.
- [69] B. Buracas, A. M. Zador, M. R. DeWeese, and T. D. Albright, "Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex," *Neuron*, vol. 20, pp. 959–969, 1998.
- [70] A. Kreiter and W. Singer, "Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey," *J. Neurosci.*, vol. 16, pp. 2381–2396, 1996.
- [71] R. W. Ditchburn and B. L. Ginsborg, "Vision with a stabilized retinal image," *Nature*, vol. 170, pp. 178–94, 1952.
- [72] F. Mechler, J. D. Victor, K. P. Purpura, and R. Shapley, "Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli," *J. Neurosci.*, vol. 18, pp. 6583–6598, 1998.
- [73] S.-H. Lee and R. Blake, "Visual form created solely from temporal structure," *Science*, vol. 284, pp. 1165–1168, 1999.
- [74] T. Carney, D. A. Silverstein, and S. A. Klein, "Vernier acuity during image rotation and translation: Visual performance limits," *Vis. Res.*, vol. 35, pp. 1951–1964, 1995.
- [75] P. Reinagel and R. C. Reid, "Precise firing events are conserved across neurons," *J. Neurosci.*, vol. 22, pp. 6837–6841, 2002.
- [76] W. R. Uttal, *On Seeing Forms*. Hillsdale, NJ: Lawrence Erlbaum, 1988.
- [77] —, *An Autocorrelation Theory of Form Detection*. New York: Wiley, 1975.
- [78] H. J. Reitboeck, M. Pabst, and R. Eckhorn, "Texture description in the time domain," in *Computer Simulation in Brain Science*, R. M. J. Cotterill, Ed. Cambridge, U.K.: Cambridge Univ. Press, 1988.
- [79] B. C. J. Moore, *An Introduction to the Psychology of Hearing*, Fifth ed. San Diego, CA: Academic, 2003.
- [80] A. de Cheveigne, "Time-domain auditory processing of speech," *J. Phonetics*, vol. 31, pp. 547–561, 2003.
- [81] A. de Cheveigné, "Pitch perception models," in *Pitch*, C. J. Plack and A. J. Oxenham, Eds. New York: Springer-Verlag.
- [82] J. L. Goldstein and P. Srulovicz, "Auditory-nerve spike intervals as an adequate basis for aural frequency measurement," in *Psychophysics and Physiology of Hearing*, E. F. Evans and J. P. Wilson, Eds. San Diego, CA: Academic, 1977.
- [83] B. Delgutte, "Physiological models for basic auditory percepts," in *Auditory Computation*, H. Hawkins, T. McMullin, A. N. Popper, and R. R. Fay, Eds. New York: Springer-Verlag, 1996.
- [84] J. C. R. Licklider, "A duplex theory of pitch perception," *Experientia*, vol. VII, pp. 128–134, 1951.
- [85] R. Lyon and S. Shamma, "Auditory representations of timbre and pitch," in *Auditory Computation*, H. Hawkins, T. McMullin, A. N. Popper, and R. R. Fay, Eds. New York: Springer-Verlag, 1995, pp. 517–517.
- [86] R. Meddis and L. O'Mard, "A unitary model of pitch perception," *J. Acoust. Soc. Am.*, vol. 102, pp. 1811–1820, 1997.
- [87] P. Cariani, "Temporal coding of periodicity pitch in the auditory system: An overview," *Neural Plasticity*, vol. 6, pp. 147–172, 1999.
- [88] A. R. Palmer, "Segregation of the responses to paired vowels in the auditory nerve of the guinea pig using autocorrelation," in *The Auditory Processing of Speech*, S. M. E. H., Ed. Berlin: Springer-Verlag, 1992, pp. 115–124.
- [89] T. Hirahara, P. Cariani, and B. Delgutte, "Representation of low-frequency vowel formants in the auditory nerve," in *Proc. ESCA Research Workshop Auditory Basis of Speech Perception*, U.K., July 15–19, 1996.
- [90] O. Ghitza, "Temporal nonplace information in the auditory-nerve firing patterns as a front-end for speech recognition in a noisy environment," *J. Phonetics*, vol. 16, pp. 109–123, 1988.
- [91] J. W. Morley, J. S. Archer, D. G. Ferrington, M. J. Rowe, and A. B. Turman, "Neural coding of complex tactile vibration," in *Information Processing in Mammalian Auditory and Tactile Systems*: Alan R. Liss, 1990, pp. 127–140.
- [92] J. E. Lisman and M. A. Idiart, "Storage of 7 ± 2 short-term memories in oscillatory subcycles," in *Science*, 1995, vol. 267, pp. 1512–5.
- [93] G. Laurent, M. Wehr, and H. Davidowitz, "Temporal representations of odors in an olfactory network," *J. Neurosci.*, vol. 16, pp. 3837–3847, 1996.
- [94] R. Llinas, U. Ribary, M. Joliot, and X.-J. Wang, "Context and content in temporal thalamocortical binding," in *Temporal Coding in the Brain*, G. Buzsáki, R. Llinás, W. Singer, A. Berthoz, and Y. Christen, Eds. Berlin, Germany: Springer-Verlag, 1994, pp. 251–272.
- [95] D. L. Wang and G. J. Brown, "Separation of speech from interfering sounds based on oscillatory correlation," *IEEE Trans. Neural Networks*, vol. 10, pp. 684–697, July 1999.
- [96] V. Braitenberg, "Functional interpretation of cerebellar histology," *Nature*, vol. 190, pp. 539–540, 1961.
- [97] P. Cariani, "Neural timing nets," *Neural Netw.*, vol. 14, pp. 737–753, 2001.
- [98] H. C. Longuet-Higgins, "A mechanism for the storage of temporal correlations," in *The Computing Neuron*, R. Durbin, C. Miall, and G. Mitchison, Eds. Reading, MA: Addison-Wesley, 1989, pp. 99–104.
- [99] R. D. Patterson, M. H. Allerhand, and C. Giguere, "Time-domain modeling of peripheral auditory processing: A modular architecture and a software platform," *J. Acoust. Soc. Amer.*, vol. 98, pp. 1890–1894, 1995.
- [100] J. Orbach, *The Neuropsychological Theories of Lashley and Hebb*: Univ. Press America, 1998.
- [101] R. W. Thatcher and E. R. John, *Functional Neuroscience Vol. I. Foundations of Cognitive Processes*. Hillsdale, NJ: Lawrence Erlbaum, 1977.
- [102] D. M. MacKay, "Self-organization in the time domain," in *Self-Organizing Systems 1962*, M. C. Yovitts, G. T. Jacobi, and G. D. Goldstein, Eds. Washington, D.C.: Spartan, 1962, pp. 37–48.
- [103] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press, 1990.
- [104] C. J. Darwin and R. P. Carlyon, "Auditory grouping," in *Handbook of Perception and Cognition: Hearing*, B. C. J. Moore, Ed. New York: Academic, 1995, pp. 387–424.
- [105] P. Cariani, "Recurrent timing nets for auditory scene analysis," in *Proc. Int. Joint Conf. Neural Networks (IJCNN)*, Portland, OR, July 20–24, 2003, [Online] <http://cil.ece.uic.edu/ijcnn03/>.
- [106] —, "Temporal codes, timing nets, and music perception," *J. New Music Res.*, vol. 30, pp. 107–136, 2002.
- [107] R. Meddis and M. J. Hewitt, "Modeling the perception of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Amer.*, vol. 91, pp. 233–245, 1992.



Peter A. Cariani received the B.S. degree in life sciences from the Massachusetts Institute of Technology, Cambridge, in 1978 and the Ph.D. degree from Binghamton University, Binghamton, NY, in 1989. His thesis considered the epistemological implications of sensor evolution for artificial adaptive devices.

He is currently an Assistant Professor in Otolaryngology and Laryngology at Harvard Medical School, Cambridge, MA.