

Fig. 1. Different models for glutamate spillover in the hippocampus. (a) 'Local' spillover involving only a couple of synapses. Diffusion is fast enough that the time course of NMDA receptor activation (right) may be similar at the site of release (red) and in adjacent synapses (green). (b) More broadly distributed spillover, in which glutamate released from one synapse reaches a larger number of synapses, though at a much lower concentration. As a result, very few receptors are activated in any adjacent synapse (right). (c) Cooperation between active synapses could enhance activation of receptors at quiescent synapses in between, resulting in a larger postsynaptic response (right).

a center-surround arrangement that may even enhance synapse specificity. Other factors, such as alternate routes of calcium entry into the postsynaptic dendrite or a possible role for postsynaptic neuronal glutamate transporters⁸, are likely to complicate this simplified scenario. Although the evidence for spillover continues to mount, it remains to be seen whether the phenomenon, contrary to its unfortunate, accidental moniker, proves beneficial to neuronal information processing.

the induction of LTP without violating Donald Hebb's postulated requirement for coincident pre- and postsynaptic activity¹³. If nearest neighbors were to experience an intermediate influx of calcium, these synapses could undergo depression¹⁴, resulting in

1. Arnth-Jensen, N., Jaboudon, D. & Scanziani, M. *Nature Neurosci.* 5, 325-331 (2002).
2. Asztely, F., Erdemli, G. & Kullmann, D. M. *Neuron* 18, 281-293 (1997).
3. Malenka, R. C. & Nicoll, R. A. *Neuron* 19, 473-476 (1997).

4. Choi, S., Klingauf, J. & Tsien, R. W. *Nature Neurosci.* 3, 330-336 (2000).
5. Rusakov, D. A. & Kullmann, D. M. *J. Neurosci.* 18, 3158-3170 (1998).
6. Barbour, B. J. *Neurosci.* 21, 7969-7984 (2001).
7. Lozovaya, N. A., Kopanitsa, M. V., Boychuk, Y. A. & Krishtal, O. A. *Neuroscience* 91, 1321-1330 (1999).
8. Diamond, J. S. *J. Neurosci.* 21, 8328-8338 (2001).
9. Lester, R. A., Clements, J. D., Westbrook, G. L. & Jahr, C. E. *Nature* 346, 565-567 (1990).
10. Bergles, D. E. & Jahr, C. E. *J. Neurosci.* 18, 7709-7716 (1998).
11. Clements, J. D. & Westbrook, G. L. *Neuron* 7, 605-613 (1991).
12. Denk, W., Yuste, R., Svoboda, K. & Tank, D. W. *Curr. Opin. Neurobiol.* 6, 372-378 (1996).
13. Hebb, D. O. *The Organization of Behavior: A Neuropsychological Theory* (Wiley, New York, 1949).
14. Cummings, J. A., Mulkey, R. M., Nicoll, R. A. & Malenka, R. C. *Neuron* 16, 825-833 (1996).

A new window on sound

Bruno A. Olshausen and Kevin N. O'Connor

Auditory filters must trade off frequency tuning against temporal precision. The compromise achieved by the mammalian cochlea seems well matched to the sounds of the natural environment.

In the vertebrate cochlea, sound is detected by an array of several thousand hair cells that transduce mechanical vibrations into electrical activity. The individual hair cells, and the auditory nerve fibers to which they are connected, are tuned to specific frequencies. The population of auditory nerve fibers thus provides us with a frequency analysis of sound waveforms in the environment, but there are as many ways to perform a frequency analysis as there are to build a house. Which one is the most appropriate for the auditory system? An article by

Lewicki in this issue sheds new light on this question¹. Using powerful new statistical methods, the author shows that the frequency analysis performed by the mammalian cochlea is well matched to the range of sounds encountered in the natural environment.

Each auditory nerve fiber may be considered as a filter that signals information about the temporal structure of stimuli within its preferred frequency range. As engineers have understood for years, the design of a filter involves an inevitable trade-off between the precision of frequency tuning and temporal tuning. A tone consists of cyclical fluctuations of air pressure, and to obtain an accurate frequency estimate, many cycles must be integrated. But a longer integration period means a decrease in the temporal accuracy of the filter—in other words, a filter cannot signal both the frequency and the timing of a sound with

arbitrary precision. Yet discrimination of real-world sounds often requires accurate measurements of both frequency and timing. In human speech, for example, the difference between vowel sounds depends on the relative strengths of different frequencies (harmonics), whereas the distinction between certain consonant sounds ('ta' and 'da' or 'ba' and 'pa') is a matter of a difference in voice onset time of a few tens of milliseconds. Precise temporal information is also important for sound localization, which in many cases depends on time-of-arrival differences between the two ears. The challenge for the auditory system, then, is to find the right trade-off between timing and frequency analysis.

One way to think about the time-frequency tradeoff is in terms of a 'tiling' of the time-frequency plane (Fig. 1). At one extreme, frequency discrimination is sacrificed completely to maximize temporal discrimination. At the other extreme is the Fourier transform, in which temporal discrimination is ignored to extract the maximum information about the frequency composition of the stimulus. Between these extremes are many other possibilities. Two schemes that are well known to

The authors are at the Center for Neuroscience and the Department of Psychology (B.A.O.) and Section of Neurobiology, Physiology and Behavior (K.N.O.), University of California at Davis, 1544 Newton Court, Davis, California 95616, USA.
e-mail: baolshausen@ucdavis.edu

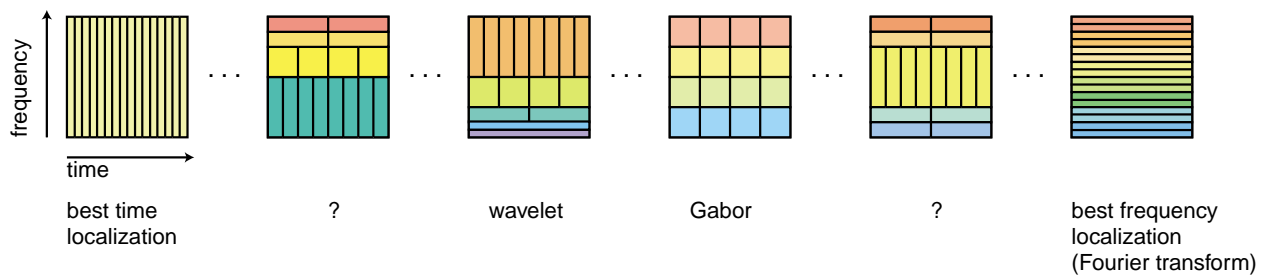


Fig. 1. The time–frequency plane can be tiled in multiple ways, any of which provides a complete representation of a signal. Some of these possibilities have been named by engineers (Gabor, wavelet, Fourier transform). Each row within a tiling represents one filter. The vertical dimension of each row represents the frequency specificity of the filter. The width of each box within a row represents the temporal resolution of the filter. Note that although the shapes vary, the area of each box is the same; this area represents the lower bound imposed by the fact that it is impossible to achieve arbitrarily precise resolution of both timing and frequency. How the boxes within a tiling are shaped reflects the chosen trade-off between time and frequency.

engineers are the Gabor filter and the wavelet filter. Gabor filters, which consist of Gaussian-windowed sinusoids, provide the optimal joint resolution in both time and frequency, and they tile the time–frequency plane with windows of equal temporal width at all frequencies. Wavelet filters can be viewed as variants of Gabor filters in which the temporal windows become narrower as frequency increases, giving the filters the property of self-similarity. They are popular for many signal-processing applications because they capture the self-similar structure that is present in many natural signals. But these are just examples, and there is no end to the variety of tiling schemes that can be imagined.

Which scheme is ‘optimal’ would seem to depend on a number of factors: the relative behavioral importance of different types of information, neurobiological or biophysical constraints, and the statistical properties of signals present in the environment. Lewicki focuses on the last of these, attempting to find a time–frequency tiling scheme that maximizes statistical independence among the filters. The motivation for maximizing independence is related to ideas proposed long ago by Attneave² and Barlow³, who argued that the nervous system should try to exploit the redundancies present in signals in order to form representations of the structure present in the environment. It is also related to principles of efficient coding, which aim to make the most use of limited neural resources⁴.

Lewicki’s method¹ for deriving a set of optimal filters draws on a recent advance in signal analysis called ‘independent component analysis’ (ICA). ICA provides a method for extracting a lin-

ear decomposition of signals that minimizes not just correlations but many higher-order statistical dependencies as well⁵. Lewicki shows that when ICA is applied to different ensembles of natural sounds (using short samples of 8 ms duration), the time–frequency tiling patterns that emerge are strikingly different. For environmental sounds (such as crackling twigs), one obtains time–frequency windows similar to a wavelet, whereas for animal vocalizations (monkey coos), one obtains a tiling pattern similar to the Fourier transform. Speech, which (as noted above) contains a mixture of temporal and frequency cues, gives rise to an intermediate tiling pattern, somewhere between a Gabor and a wavelet; the temporal accuracy increases with frequency, but to a lesser extent than with wavelet filters.

The tiling pattern that is optimal for speech is thus intermediate between those optimized for environmental sounds and animal vocalizations. As Lewicki shows, a close match is obtained with a 2:1 mixture of environmental to animal sounds. Interestingly, this pattern is similar to what has been observed physiologically in cat auditory nerve fibers, and it also bears similarity to the auditory filters that have been characterized psychophysically in humans and other animals. Although auditory filters measured behaviorally are not necessarily determined by the cochlea⁶ (they could theoretically arise anywhere within the auditory system), these results, taken together, suggest that the cochlea and auditory nerve may be optimized to transmit a wide range of naturally occurring sounds to the brain. It is even possible, as the author suggests, that the acoustic properties of human speech have evolved to make efficient use of the

pre-existing properties of the peripheral auditory system.

Lewicki’s analysis does not attempt to provide a comprehensive account of auditory coding. For example, it does not consider the effect of changing sound intensity. The tuning of Lewicki’s filters is independent of sound intensity, but this is not true of real auditory nerve fibers. Most fibers reach the limit of their dynamic range roughly 30–40 dB above threshold, meaning their firing rates saturate at even moderate intensities⁷. At these intensities, their frequency tuning also becomes considerably broader⁸. But these facts are not necessarily inconsistent with Lewicki’s results, because most physiological measurements are made using isolated pure tones. Less is known about how auditory nerve fibers behave in response to more ecologically realistic broadband stimuli, and it is possible that gain control mechanisms maintain frequency selectivity even with high-intensity stimuli⁹.

There are also a few peculiarities to Lewicki’s filters that arise from the particular way in which ICA was implemented. For example, the filters learned by the algorithm are fairly symmetric in time (the attack and decay occur at about the same rate), whereas the ‘gamma-tone’ filters that have been characterized physiologically are asymmetric in time (they rise more steeply than they decay). In addition, the algorithm produces filters with the same frequency response but shifted in time, whereas auditory nerve fibers do not show such delays. But it would be fairly straightforward to modify the algorithm so that the filters are constrained to be causal (that is, filter outputs are determined from present and past values of the input), in which case their temporal

envelopes would most likely become asymmetric, like gamma-tone filters. And if the filters were also assumed to be time-invariant, so that they are actually convolved with the input signal, then there would be no need for time-shifted filters. However, such modifications are unlikely to dramatically affect the time-frequency tiling scheme learned by the algorithm, which is really the main point of the paper.

Lewicki's results share an intriguing similarity to recent work in vision. Neurons in the visual cortex encode both the location and the spatial frequency of visual stimuli, and the trade-off between these two variables is analogous to that between timing and frequency in auditory coding. It has been shown that maximizing statistical independence or 'sparseness', of visual representations yields spatial receptive field properties similar to those of cortical neurons^{10–12}. Curiously, the space–frequency tiling scheme of both the derived filters and those measured physiologically deviates from a wavelet in much the same way as Lewicki finds for the auditory system; bandwidth at high spatial frequencies is narrower than one would expect¹³. It is not yet clear whether this similarity is profound or simply coincidental.

Perhaps an even deeper question is why ICA accounts for neural response properties at the very earliest stage of analysis in the auditory system, whereas in the visual system ICA accounts for the response properties of cortical neurons, which are many synapses removed from photoreceptors. It seems likely that structural or neurobiological constraints

are crucial in determining the stage of analysis appropriate for an independent component analysis of sensory signals. For example, the visual system is faced with an early bottleneck, where information from more than 100 million photoreceptors is funneled into 1 million optic nerve fibers. The representation is then expanded by a factor of 50 in the cortex. By contrast, in the auditory system, there is no early bottleneck, and the 3000 inner hair cells of the cochlea immediately fan out onto 30,000 auditory nerve fibers. Thus it seems that in both systems, ICA is applied at the point of expansion in the representation.

It is tempting to speculate about what additional insights may be gained regarding neural mechanisms higher up in the auditory stream—the midbrain, thalamus and cortex—by considering higher-order forms of structure over longer time scales and across multiple frequency bands. But this can not be done by simply applying the same analysis yet again. To add descriptive power, any additional stage of analysis would have to be nonlinear. Divisive normalization⁹ or a signal power representation (as in a spectrogram) seem like obvious choices, and some preliminary work along these lines has produced promising results¹⁴. Conceivably, this type of approach could begin to account for more complex properties of auditory neurons such as multiple excitatory–inhibitory band structure or frequency modulation sensitivity¹⁵, or perhaps even predict heretofore unnoticed response properties of cortical neurons. One gets the feeling that these findings

are really just the tip of the iceberg, and that ahead lies a vast territory ripe for investigation.

1. Lewicki, M. *Nature Neurosci.* 5, 356–363 (2002).
2. Attneave, F. *Psychol. Rev.* 61, 183–193 (1954).
3. Barlow, H. B. in *Sensory Communication* (ed. Rosenbluth, W. A.) 217–234 (MIT Press, Cambridge, MA, 1961).
4. Simoncelli, E. P. & Olshausen, B. A. *Annu. Rev. Neurosci.* 24, 1193–1215 (2001).
5. Hyvarinen, A., Karhunen, J. & Oja, E. *Independent Component Analysis* (Wiley, New York, 2001).
6. Ehret, G. in *Advances in Hearing Research. Proceedings of the 10th International Symposium on Hearing* (eds. Manley, G. A., Klump, G. M., Koppl, C., Fastl, H. & Oeckinghaus, H.) 387–400 (World Scientific, London, 1995).
7. Evans, E. F. & Palmer, A. R. *Exp. Brain Res.* 40, 115–118 (1980).
8. Kiang, N. Y.-S., Watanabe, T., Thomas, E. C. & Clark, L. F. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, Massachusetts, 1965).
9. Schwartz, O. & Simoncelli, E. P. *Nature Neurosci.* 4, 819–825 (2001).
10. Olshausen, B. A. & Field, D. J. *Nature* 381, 607–609 (1996).
11. Bell, A. J. & Sejnowski, T. J. *Vision Res.* 37, 3327–3338 (1997).
12. van Hateren, J. H. & van der Schaafi, A. *Proc. Royal Soc. Lond. B* 265, 359–366 (1998).
13. De Valois, R. L., Albrecht, D. G. & Thorell, L. G. *Vision Res.* 22, 545–559 (1982).
14. Kording, K.P., Konig, P., Klein, D.J. (2002) Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN'2002) (IEEE Press, New York, in press)
15. Schreiner, C. E., Read, H. L. & Sutter, M. L. *Annu. Rev. Neurosci.* 23, 501–529 (2000).

Color visions in the brain



A vivid perception of color is evoked by spoken words ("seven" is blue, for instance) in people with a condition called 'colored-hearing synesthesia'. This percept is associated with activity in an area of the brain that responds to color vision, report Julia Nunn and colleagues on page 371 of this issue. Because other visual areas are not activated, these results suggest that a conscious perception of color can be created by activation of the brain's 'color center' alone.

In a control experiment, normal subjects did not show activity in the color center in response to spoken words, even after they had been extensively trained to visualize particular colors in association with those words. The authors conclude that synesthesia is much more like a color hallucination than color imagery. Synesthesia may also have a genetic basis, as it runs in families and is strongly sex-linked (six times more common in women). The authors speculate that this condition may result from developmental errors in the formation or retraction of connections between auditory cortex and visual cortex.

Sandra Aamodt