argue that the magnitude of delay has been overestimated, and that it might even be unnecessary to compensate for such a delay.

Although time delay poses a difficult issue for visual motion perception in particular, there is nevertheless the benefit to allow large amounts of computation to take place unconsciously. Presumably, evolution must have set optimal delay lengths that are short enough to avoid too much delay, but long enough to allow sufficient amounts of computation. If visual delay is within such optimal lengths, compensation might be unnecessary. The key issue, then, is to determine the magnitude of delay inherent in the visual system. As mentioned by Nijhawan, the perceptual lag of a moving object due to neural delay can be calculated as $v \times t_{\max}$, where v is object velocity and $t_{\max}$ is the maximum cumulative delay in the vision pathway before conscious percepts emerge. The magnitude of delay is therefore determined by $t_{\max}$. This boils down to this question: To what extent can the visual process proceed in the visual hierarchy without awareness? The farther the unconscious visual process proceeds, the larger is the perceptual lag that occurs.

Because humans are not aware of the delay compensation process, it is tempting to speculate that the brain areas subserving it might be unrelated to awareness. Indeed, neural activity at the level of photoreceptors is not sufficient to yield visual awareness, and some researchers further argue that later stages after the retina, such as primary visual cortex (V1), are not involved in consciousness. This logic is problematic, however, and can be illustrated more clearly when we take a step back to examine a more general problem in visual perception – what Hermann von Helmholtz called unconscious inference (i.e., before percepts emerge, huge amounts of computational processes take place without consciousness). A key notion here is that unconscious inference does not imply that the neural substrates are unrelated to consciousness. For example, in the binding problem, a distributed hierarchical network in the brain processes different features of an object. The neural substrates that correctly select and integrate separate neural representations of features to form coherent object representations are believed to comprise both parietal and frontal cortex (Treisman 1996). Can we thus infer that activity in parietal and frontal cortex is unrelated to consciousness? Apparently this is not the case (Lin & He, in press).

Yet, the notion that observers are not directly aware of neural activity in the optic nerve, the lateral geniculate nucleus (LGN), and even V1 has been popular among many researchers (e.g., sect. 3 of the target article), especially after the appearance of the influential paper by Crick and Koch in 1995. Although this idea (Crick & Koch 1995) was initially based on behavioral studies that show equal adaptation strengths for both visible and invisible conditions (e.g., He et al. 1996), recent evidence demonstrates, instead, that V1 is directly involved in visual awareness (for a review, see Lin & He, in press). Monkey single-unit recordings, human electroencephalogram (EEG) recordings, and brain imaging converge to reveal robust awareness-dependent modulations in V1; neural events in V1 are actually attenuated in response to suppressed (vs. non-suppressed) visual stimuli. Convergent psychophysical evidence also comes from *binocular rivalry*, the phenomenon that when each of the two eyes views dissimilar images, the two images rival for visibility, with one temporarily dominating perception for several seconds before being replaced in dominance by the other. With adaptation, a recent study shows that rivalry suppression weakens the strengths of aftereffect for low-contrast adaptors (Blake et al. 2006), which is believed to transpire in V1. It appears that although unconscious inference per se is beyond consciousness, neural activity in V1 is actually correlated with consciousness.

Where, then, does consciousness enter into the information processing sequence? How can one distinguish conscious processing from unconscious processing in the brain? It is important to know which neural substrates of unconscious inference are conscious and which are unconscious, as measured by correlations

between neural activity and conscious experiences. This appears extremely difficult, given the unclear mechanisms of how the brain computes unconscious inference. In addition, it is almost impossible to isolate the inference processes (e.g., feature binding), since we only know that there are such processes after we are conscious of the objects. Moreover, although many studies focus on brain activation differences between conscious and unconscious experiences, it is an oversimplification to classify certain areas as conscious or unconscious based on such comparison. For example, a recent study shows that although V1 activity is closely correlated with conscious experience, even a considerable difference in cortical activation in many visual cortical areas does not necessarily lead to different conscious experiences (Jiang et al. 2007). It therefore appears that no single area, or even network, is sufficient for consciousness. A more dynamic approach to distinguish consciousness from unconsciousness is to go beyond isolated activity of certain brain areas, and, instead, focus on how different brain areas interact with each other. Lamme and Roelfsema (2000) suggest that when information first enters the early visual cortex in an feedforward sweep, it is unconscious; when such information is attentively grouped through recurrent processing (by means of horizontal connections and feedforward/feedback projections), it gives rise to consciousness. This might provide important clues as to what give rise to consciousness.

In sum, although the neural correlates of consciousness are still elusive, it is now well documented that V1 is directly related to awareness (Lin & He, in press). This has important implications for the magnitude of time delay by constraining $t_{\max}$ (the maximum cumulative delay in the vision pathway before conscious percepts emerge) to less than 72 msec on average in V1 neurons (Lamme & Roelfsema 2000). Such a short delay might be within optimal delay lengths to allow sufficient computations to take place, and might even make it unnecessary to compensate for the delay.

ACKNOWLEDGMENTS

The preparation of this commentary was supported by a University of Minnesota Graduate School Block Grant to Zhicheng Lin, and by the James S. McDonnell foundation and the U.S. National Institutes of Health (EY-015261-01) grants to Dr. Shang He (Lin's current advisor).

Neuronal adaptation: Delay compensation at the level of single neurons?

doi: 10.1017/S0140525X08003944

J. Patrick Mayo and Marc A. Sommer

Department of Neuroscience, Center for the Neural Basis of Cognition, Center for Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260.

jpm49@pitt.edu

masommer@pitt.edu

Abstract: Saccades divide visual input into rapid, discontinuous periods of stimulation on the retina. The response of single neurons to such sequential stimuli is neuronal adaptation; a robust first response followed by an interval-dependent diminished second response. Adaptation is pervasive in both early and late stages of visual processing. Given its inherent coding of brief time intervals, neuronal adaptation may play a fundamental role in compensating for visual delays.

Nijhawan presents a compelling argument in support of the existence of delay compensation in visual processing. Specifically, he draws a distinction between “discrete” and “continuous” stimuli and focuses on the flash-lag effect as a useful illustration of differential processing between the two. Our laboratory investigates an important class of stimuli that extends the discrete-continuous continuum. We use sequential stimuli, consisting of consecutive stimulation separated by

brief blank periods. This kind of visual stimulation is frequently encountered by the retina.

Humans make rapid eye movements, or saccades, approximately three times per second. These movements shift the position of the world on the retina. Thus, the seemingly simple act of looking at an object, such as a face, involves the successive repositioning of its image features on disparate locations of the retina. Importantly, we do not perceive these shifts of the visual world or the period of lost time during a saccade, because of a general phenomenon known as saccadic suppression. Natural vision is therefore largely comprised of stable visual input that is frequently interrupted by brief periods of saccadic suppression (although continuous stimulation resulting from object tracking is possible in the form of “smooth pursuit” movements; for review, see Krauzlis 2004). Saccadic suppression adds appreciable “dead time” to the normal delays that Nijhawan reviews; yet we fail to notice all of the lost perceptual time just as we fail to be disturbed by the afferent delays. Instead, we perceive a world that is spatially and temporally stable.

How do we maintain this two-dimensional visual stability in the face of the brain’s physical limitations? Nijhawan discusses feed-forward signaling as one way of coping with visual delays, and our earlier work demonstrated the existence of such signaling in a pathway from the midbrain to prefrontal cortex (Sommer & Wurtz 2002; 2006). This pathway seems to help keep the brain and retinotopic visual world in spatial register before and after a saccade. Our recent work has focused on how the brain maintains temporal stability (Mayo & Sommer, submitted). We sequentially present two visual probes in the receptive field of neurons in awake, fixating monkeys. We systematically vary the amount of time between probe presentations (interstimulus interval, or ISI). The purpose is to mimic the fixation-saccade-fixation timing of natural vision. We have recorded neuronal activity in two brain areas: the retinally recipient superficial layers of the superior colliculus (supSC) and the frontal eye field (FEF), a “higher” visuomotor area in prefrontal cortex.

Sequential stimulation elicited sequential visual responses in neurons, with each response separated by a period of baseline activity equivalent to the ISI. Therefore, the timing of neuronal responses accurately encodes the timing of real-world visual stimulation. However, the relative magnitude of the neuronal responses also varied significantly with ISI (Fig. 1). Second visual responses were only a fraction of the first; this is a well-known phenomenon called neuronal adaptation.

We suggest that neuronal adaptation may be useful for timing, because it carries information about the ISI during sequential visual stimulation. Extremely short ISIs (<50 msec) lead to negligible second neuronal responses (Fig. 1, left), whereas longer ISIs lead to monotonically increasing responses (Fig. 1, right) until ISI ≈ 400 msec, where second responses attain normal magnitudes (Fig. 2). We found that adaptation followed the same time-course in FEF and supSC, but was stronger overall in FEF. As a point of reference with previous studies, our supSC data matched those reported by other laboratories (Robinson & Kertzman 1995; Wurtz et al. 1980).

Figure 1 (Mayo & Sommers). Example visual responses from a small group of single neurons in superior colliculus. The results of using two probes presented at interstimulus intervals of 50 msec (left) and 200 msec (right) are shown.

Figure 2 (Mayo & Sommers). Normalized visual responses in the superior colliculus population. Vertical dotted line indicates time of first probe onset. Grayscale lines depict mean responses to the second probe after sequential, paired stimulation, at various interstimulus intervals. Dashed line represents typical response to first probe.

In light of the apparent need for a compensatory physiological mechanism as described by Nijhawan, we posit that neuronal adaptation may act as an alternative source of latency encoding. As the target article explains, the exact timing of neuronal responses in identical conditions can vary by tens of milliseconds from trial to trial. Neurons downstream of these time-varying sensory responses therefore receive an imprecise measurement of stimulus onset. However, for sequential stimuli presented within 400 msec of each other, downstream neurons could corroborate the raw timing of the neuronal responses with relative magnitude comparisons of the incoming neuronal responses. Neuronal adaptation therefore may function as a complementary latency code. The presence of this code in ecologically relevant scenarios – in response to sequential stimuli, and at presentation times that mimic typical periods of fixation and saccades – lends further support to our hypothesis.

Using the relative magnitudes of sequential visual responses as a clue to the timing of visual stimuli requires an assumption that the sequential visual stimuli themselves are nearly identical. Accepting this assumption means that differences in visual responses that occur within 400 msec of each other must be due to timing. The assumption seems valid for many real-world situations, such as making saccades across continuously textured images or when an object under inspection is briefly occluded. In the latter case, neuronal adaptation would provide a robust encoding of the duration of occlusion.

Although it is still unclear how exactly neuronal adaptation could account for delay compensation, it has two important traits that are central to Nijhawan’s proposal. First, neuronal adaptation is a sensory process that occurs only one synapse beyond the retina in cats and monkeys (Coenen & Eijkman 1972; Mayo & Sommer, submitted) and has been reported within the retina in lower species (Schwartz et al. 2007a). Thus, neuronal adaptation seems to be an immediate, automatic sensory mechanism. Second, neuronal adaptation is pervasive, seen at all stages of striate and extrastriate visual processing. This substantial concordance between adaptation and Nijhawan’s proposal calls for additional research and computational modeling. An improved understanding of the neurophysiology of timing would provide an important complement to Nijhawan’s thoughtful analysis.

ACKNOWLEDGMENTS

This research was supported by the Alfred P. Sloan foundation and NIH R01-EY017592.

Phase-alignment of delayed sensory signals by adaptive filters

doi: 10.1017/S0140525X08003956

Dennis J. McFarland

Wadsworth Laboratories, New York State Department of Health, Empire State Plaza, Albany, NY 12201.

mcfarlan@wadsworth.org

Abstract: Correction of sensory transmission delays is an intractable problem, since there is no absolute reference for calibration. Phase-alignment is a practical alternative solution and can be realized by adaptive filters that operate locally with simple error signals.

Nijhawan suggests that the visual system compensates for delays, with the perceived position of a moving object being based on a type of guess as to the object's actual position. This would be adaptive, as it is necessary that the output of the visual system be synchronized with the environment and behavior. However there is really no way that any part of the nervous system can be calibrated to correct for transmission delays, since there is no absolute reference available for calibration. As such, the calibration problem is intractable. This problem can be solved instead by aligning the relative phase of the signals in question. This is a common problem in engineering that has relatively simple and straightforward solutions.

Problems produced by transmission delays are frequently encountered in electrical engineering. Examples include cancellation of echoes produced in long-distance transmission lines and adaptive beam-forming from an array of sensors with delayed outputs (e.g., spatially separated sensors in a sonar system). These problems can be solved by means of adaptive filters with delayed inputs (Haykin 1996). The proper delay is selected by appropriate weighting on the delayed inputs. These weights are adjusted by some simple characteristic of the output, such as by minimizing the variance of the signal subject to the condition that the sum of the weights equals 1.

Adaptive filters capable of phase-aligning signals can be realized by simple operations that can be implemented either by algebraic equations or by a few elements of an artificial neural network. This means that the neural circuitry need not be complex. The cost function that tunes these filters can be very simple, so that it is not necessary to appeal to complex cognitive processes. Some characteristic of the fused signal, such as the variance, serves as an error signal to adjust the weights given to the delay elements. All that is required for phase-alignment is some sort of recursive feedback of the output of a perceptual subsystem that adjusts the connections of the delayed inputs. The simplest design is a tapped-delay line filter, although there is no unique solution to this problem.

The logical point for phase-alignment would be the step in processing that precedes the fusion of the to-be-aligned signals, as this would minimize the length of the path required for the recursive feedback and ensure the fidelity of the error signal. If filters are tuned by the output of areas far removed from the site of sensor fusion, the error signal will not closely reflect the characteristics of the fusion product. Hence, we would expect the site of phase-alignment to be the local network involving both the site of fusion and the immediate sources of its to-be-aligned inputs.

To the extent that the phase-alignment problem is solved with adaptive filters, we should expect that there would be no absolute delay correction employed by perceptual systems. Rather, the

phase correction would depend on the problems encountered during the recent history of the various perceptual subsystems. An understanding of how this process works might best be obtained by observing the plasticity in the calibration of perceptual subsystems (i.e., adaptation to the statistics of the observer's environment). Fusion of audiovisual speech is an example.

Vatakis et al. (2007) measured judgments of the temporal order of auditory and visual speech tokens. Participants were concurrently exposed to audiovisual speech that was presented either in synchrony or with the auditory stream lagging by 300 msec. They found that exposure to asynchronous audiovisual speech shifted the point of subjective simultaneity in the temporal order task. These results demonstrate that even the brief exposure encountered in a typical experimental session is capable of recalibrating the relative phase of perceptual subsystems. Thus, phase-alignment in this case is an adaptive process.

Phase-alignment by adaptive filters can operate locally within a relatively simple network using simple error signals. As a result, the top-down input to perceptual systems does not need to involve complex cognitive processes from remote higher cortical regions. Furthermore, the input is in the form of adjustments to the strength of synaptic weights that tune perceptual filters. This adjustment occurs in a slow post-hoc manner so that changes resulting from any current mismatch do not affect on-line processing. This slow change would occur on a timescale much longer than that of individual perceptual events. This is in contrast to the view expressed by Nijhawan, where prediction is based on an interaction of visual systems with information from areas much farther downstream.

Prediction is a difficult problem that is not explained by simply appealing to top-down input to perceptual processing streams. Phase-alignment is a tractable problem with simple solutions. Prediction is based on a Cartesian view of the problem that considers the perception of motion to be the instantaneous value of a continuous variable that starts and stops a counter somewhere in the brain. Phase-alignment involves a filter that integrates input from sensors across time to produce an output, the nature of which is adjusted by prior sensory experience.

ACKNOWLEDGMENT

My thanks to Jonathan Carp for comments on an earlier version.

Moving backward through perceptual compensation

doi: 10.1017/S0140525X08003968

Haluk Ögmen,^a Saamil S. Patel,^b Gopathy Purushothaman,^c and Harold E. Bedell^d

^aDepartment of Electrical and Computer Engineering, Center for Neuro-Engineering & Cognitive Science, University of Houston, Houston, TX 77204-4005; ^bDepartment of Neurobiology and Anatomy, University of Texas Medical School at Houston, Houston, TX 77030; ^cDepartment of Cell and Developmental Biology, Vanderbilt Vision Research Center, Vanderbilt University, T2304, Medical Center North, Nashville, TN 37232; ^dCollege of Optometry and Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-2020.

ogmen@uh.edu <http://www.egr.uh.edu/ece/faculty/ogmen/>

saamil@swbell.net

gopathy.purushothaman@vanderbilt.edu

Hbedell@optometry.uh.edu

Abstract: In the target article Nijhawan speculates that visual perceptual mechanisms compensate for neural delays so that moving objects may be perceived closer to their physical locations. However, the vast majority of published psychophysical data are inconsistent with this speculation.