

Economy of scale: A motion sensor with variable speed tuning

John A. Perrone

Department of Psychology, The University of Waikato,
Hamilton, New Zealand



We have previously presented a model of how neurons in the primate middle temporal (MT/V5) area can develop selectivity for image speed by using common properties of the V1 neurons that precede them in the visual motion pathway (J. A. Perrone & A. Thiele, 2002). The motion sensor developed in this model is based on two broad classes of V1 complex neurons (sustained and transient). The S-type neuron has low-pass temporal frequency tuning, $p(\omega)$, and the T-type has band-pass temporal frequency tuning, $m(\omega)$. The outputs from the S and T neurons are combined in a special way (weighted intersection mechanism [WIM]) to generate a sensor tuned to a particular speed, v . Here I go on to show that if the S and T temporal frequency tuning functions have a particular form (i.e., $p(\omega)/m(\omega) = k/\omega$), then a motion sensor with variable speed tuning can be generated from just two V1 neurons. A simple scaling of the S- or T-type neuron output before it is incorporated into the WIM model produces a motion sensor that can be tuned to a wide continuous range of optimal speeds.

Keywords: speed tuning, temporal frequency tuning, V1, MT, motion model

Introduction

Understanding how the brain processes visual speed information is integral to the question of how we gather information about the environment from retinal image motion. Our knowledge of how this process occurs would improve if we could deduce the mechanisms underlying the properties of the neurons that respond selectively to image speed. We know that for a neuron to be tuned to a particular image velocity (speed), v , it needs to respond maximally to combinations of spatial (u) and temporal (ω) frequencies that are related by the equation $\omega = -vu$ (Watson & Ahumada, 1983). It is well established that neurons in the MT area respond best to a particular edge or bar speed (Felleman & Kaas, 1984; Maunsell & Van Essen, 1983) and that some of them are capable of coding image speed independently of changes to the stimulus pattern (i.e., they follow the $\omega = -vu$ rule) (Perrone & Thiele, 2001; Priebe, Cassanello, & Lisberger, 2003). However, until recently, it was not clear how MT neurons could have acquired these abilities from the V1 neurons that provide their inputs. The V1 neurons are not speed tuned; their responses are dependent on the spatial frequency content of the stimulus, and they are broadly tuned for temporal frequency (Foster, Gaska, Nagler, & Pollen, 1985).

We have recently shown that despite these limited V1 properties, it is possible to generate the type of speed tuning found in MT neurons (Perrone, 2004; Perrone & Thiele, 2002). We referred to the mechanism by which speed tuning could be generated from V1 neurons as the weighted intersection mechanism (WIM) model.

The WIM model

The building blocks for a WIM speed tuned sensor are two types of commonly occurring V1 complex neurons: a sustained type (S), which has low-pass temporal frequency tuning, $p(\omega)$, and a transient type (T) with band-pass temporal frequency tuning, $m(\omega)$, (see red and blue lines in Figure 1a).

The spatial frequency (sf) tuning of the S- and T-type V1 neurons in the WIM model also differ from each other in a special way. The S-type sf tuning function, $f(u)$, used in the model is based on actual V1 neuron data (Hawken & Parker, 1987) (see dashed red line in Figure 1b). The T sf function (blue line in Figure 1b), $f'(u)$, differs from the S type by an amount determined by the shape of the temporal frequency tuning functions (see Equation 1 below). Let $S(u, \omega)$ represent the combined spatiotemporal frequency sensitivity function of the sustained V1 neuron (or equivalently, its spatiotemporal energy output) and $T(u, \omega)$ represent the transient neuron sensitivity [i.e., $S(u, \omega) = f(u)p(\omega)$ and $T(u, \omega) = f'(u)m(\omega)$]. Note that this multiplication operation (and the steps that follow) assumes that the temporal function retains its shape as the spatial frequency changes and vice versa. There is evidence to support this “separability” assumption in V1 monkey (Foster et al., 1985) and cat (Tolhurst & Movshon, 1975) neurons. The issue of separability will be raised again in the Discussion.

Let v be the optimal speed (velocity) that elicits a maximal response from a sensor made up from an S- and T-type V1 neuron. We have previously demonstrated that if

$$f'(u) = f(u) \frac{p(vu)}{m(vu)}, \quad (1)$$

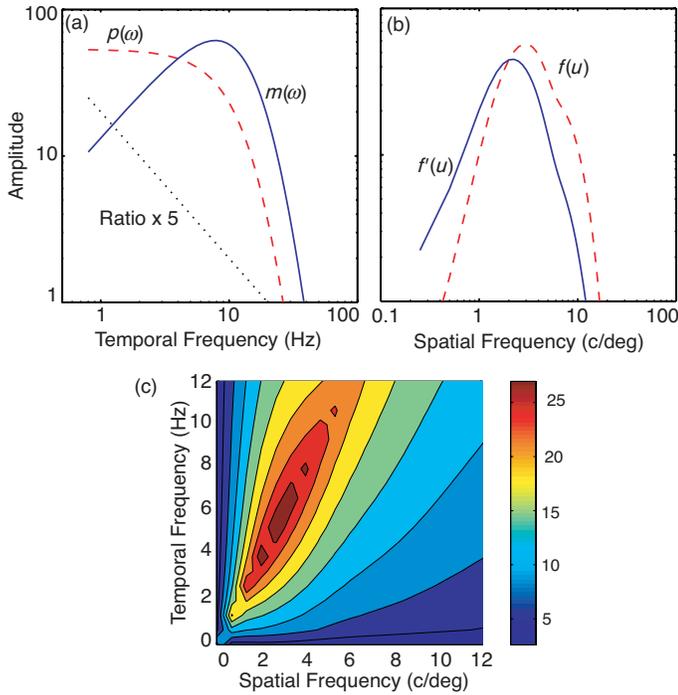


Figure 1. Creating a speed tuned sensor from V1 neurons. (a). V1 neuron temporal frequency tuning curves. (b). V1 neuron spatial frequency tuning curves. (c). Spectral receptive field of a model sensor tuned to 2 deg/s.

then $S(u_i, \omega_i) = T(u_i, \omega_i)$ for all u_i, ω_i , such that $\omega_i/u_i = -v$ (Perrone & Thiele, 2002). In other words, if the sf tuning of the transient-type V1 neuron differs from the sustained sf tuning in the manner specified by Equation 1, then the two V1 neurons (S and T) will respond equally to a particular set of spatial and temporal frequencies corresponding to a stimulus speed v .

In previous presentations of the WIM model (Perrone, 2004; Perrone & Thiele, 2002), we have adopted the arbitrary convention of first setting the spatial frequency tuning of the sustained neuron and then generating the transient neuron tuning from Equation 1. We have no particular reason (e.g., developmental or evolutionary) to favor this particular ordering, and one could just as easily rewrite Equation 1 so that the sustained tuning is derived from the transient tuning. For consistency, in the derivation of the variable speed tuning mechanism outlined below, I have retained the convention of fixing the S neuron tuning properties and modifying the T neuron properties.

The next stage of the model is to introduce a mechanism that produces a large output whenever the S and T neurons are responding equally. The algorithm we adopted was

$$WIM(u, \omega) = \frac{\log(S + T + \alpha)}{|\log T - \log S| + \delta} \quad (2)$$

where α and δ are constants that control the gain and tuning bandwidth of the sensor.

This mechanism produces a motion sensor with a spatiotemporal frequency sensitivity profile (the spectral receptive field) that is oriented in (u, ω) frequency space and which is maximally sensitive to a particular edge speed, v (see Figure 1c). This is because, in frequency space, a moving edge has a Fourier spectrum that is oriented relative to the (u, ω) axes and which passes through the origin (i.e., the equation for the spectral line is given by $\omega = -vu$) (Watson & Ahumada, 1983). For the particular temporal and spatial functions chosen in Figure 1a and 1b, the WIM sensor (Equation 2) has a spectral receptive field with a slope that is maximally responsive to edges moving at 2 deg/s to the left.

We have shown that the spatiotemporal frequency sensitivity profile generated by Equation 2 closely matches those commonly found in MT neurons (Perrone & Thiele, 2001) and have argued that the WIM mechanism could form the basis of MT speed tuning (Perrone, 2004; Perrone & Thiele, 2002). The requirements for setting up a speed tuned sensor using a WIM-type scheme are actually quite modest. A broad range of temporal frequency tuning functions will work, as long as one is slightly more band-pass than the other (see Figure 6, Perrone & Thiele, 2002). The scheme is also tolerant of a broad range of spatial frequency tuning functions as long as they can be adjusted sufficiently to meet the requirements of Equation 1.

While it is an efficient means of generating speed tuning from V1 neurons compared to alternative schemes (e.g., Simoncelli & Heeger, 1998), the current configuration of the WIM model still requires a new transient-type V1 neuron to be used for each new speed tuning value, v_i . For each optimum speed required in a WIM sensor tuned to a particular spatial frequency, u_0 , separate matched pairs of S and T inputs are required: (S_1, T_1) , (S_1, T_2) , (S_1, T_3) , etc. Given the multitude of speeds that need to be registered in a typical retinal image sequence, this is a resource intensive mechanism for achieving speed tuning. It would be more efficient if we could use the same S-T pair for a range of speed tunings. It turns out that a judicious selection of the V1 temporal frequency tuning functions enables this economy to be achieved.

V1 neuron temporal frequency tuning

Figure 2 shows a sample of temporal frequency tuning data derived from V1 neurons. They range from low-pass through to band-pass in their temporal frequency tuning.

Previously, in the WIM model (Perrone, 2004; Perrone & Thiele, 2002), we have used functions developed by Watson (1986) to simulate the temporal frequency tuning of V1 neurons. For sustained (lowpass) tuning, the function used was

$$p(\omega) = \sqrt{a^2 + b^2} \quad (3)$$

where

$$a = \left((2\pi\omega\tau_1)^2 + 1 \right)^{\frac{9}{2}} \text{ and } b = \left((2\pi\omega\tau_2)^2 + 1 \right)^{\frac{10}{2}} .$$

The parameters τ_1 and τ_2 are time constants, measured in seconds. As can be seen from Figure 1a (red dashed line), a good fit to data such as those shown in Figure 2a can be obtained by setting (τ_1, τ_2) in Equation 3 to (0.0072, 0.0043).

To simulate the temporal frequency tuning of transient (band-pass)-type V1 neurons (e.g., see Figure 2c), a more complex version of the Watson function has been used up till now, which includes a “transience factor” (ζ) that increases the degree of band-pass tuning (Perrone & Thiele, 2002). However, I have since discovered that a more useful function for the transient V1 neuron temporal frequency tuning is one given by the following equation:

$$m(\omega) = \frac{\omega}{k} p(\omega) , \tag{4}$$

where k is a constant (set to 4.0 for Figure 1a).

The two functions given by Equations 3 and 4 are shown in Figure 1a, and they easily fall within the family of tuning curves found in V1 neurons (Figure 2). Besides providing a good fit to typical V1 temporal frequency tuning data, these two particular temporal frequency functions offer a special benefit when it comes to setting up speed tuning in the WIM model.

Variable speed tuning

If the transient V1 neuron has temporal frequency tuning based on Equation 4, then the ratio of the S and T functions is given by

$$\frac{p(\omega)}{m(\omega)} = R(\omega) = \frac{k}{\omega} . \tag{5}$$

On a log-log plot, this ratio is represented by a straight line defined by $\log R = -\log \omega + \log k$ (see dotted line in Figure 1a, but note that it has been shifted upwards for clarity). This ratio function possesses a unique property: If ϕ is any real number, then from Equation 5,

$$R(\phi\omega) = \frac{k}{\phi\omega} = \frac{1}{\phi} R(\omega)$$

i.e.,

$$\frac{1}{\phi} R(\omega) = R(\phi\omega) . \tag{6}$$

This property turns out to be very useful in the new speed tuning mechanism. Using Equation 5 again, we can rewrite Equation 6 as

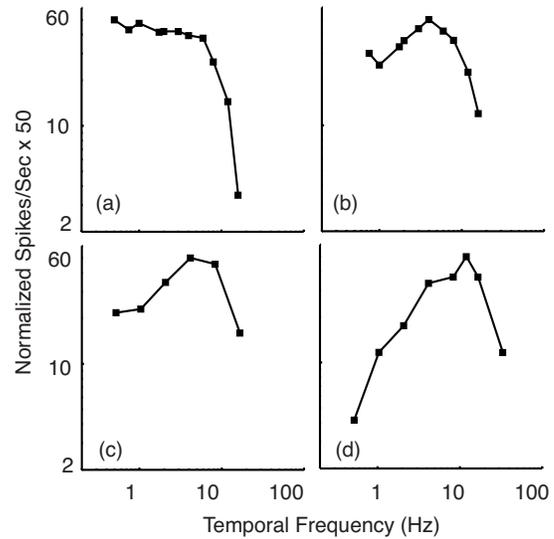


Figure 2. Replotted temporal frequency tuning data from V1 neurons. (a) and (b). Foster et al. (1985). Type unknown, moving gratings. (c). M. J. Hawken (personal communication, 1999). Complex type, moving gratings. (d). Hawken, Shapley, and Grosof (1996). Complex type, alternating gratings.

$$\frac{1}{\phi} \frac{p(\omega)}{m(\omega)} = \frac{p(\phi\omega)}{m(\phi\omega)} . \tag{7}$$

For a WIM sensor tuned to speed v_1 , we require the following relationship to exist between the different spatial and temporal frequency functions (see Equation 1):

$$f_1'(u) = f_1(u) \frac{p(v_1 u)}{m(v_1 u)} . \tag{8}$$

To generate a WIM sensor tuned to a new speed v_2 using the current version of the WIM model (Perrone, 2004; Perrone & Thiele, 2002), it is necessary to incorporate a new transient-type V1 neuron (T_2) with new spatial frequency tuning, $f_2'(u)$, also controlled by Equation 1, i.e.,

$$f_2'(u) = f_1(u) \frac{p(v_2 u)}{m(v_2 u)} , \tag{9}$$

where $f_1(u)$ is the sustained spatial frequency tuning function of the original WIM sensor, tuned to speed v_1 . If we let $v_2 = \phi v_1$, Equation 9 can be rewritten as

$$f_2'(u) = f_1(u) \frac{p(\phi v_1 u)}{m(\phi v_1 u)} . \tag{10}$$

Using the result from Equation 7 gives

$$f_2'(u) = \frac{1}{\phi} f_1(u) \frac{p(v_1 u)}{m(v_1 u)} . \tag{11}$$

Combining this result with Equation 8 gives

$$f'_2(u) = \frac{1}{\phi} f'_1(u). \tag{12}$$

In other words, we do not need to use a new transient spatial frequency tuning function, $f'_2(u)$, to generate speed tuning v_2 . We can simply scale the original transient neuron spatial frequency function. This is a powerful result and it enables a great saving in the number of V1 neurons required to generate different speed tunings. Equation 12 shows that if we start with a single pair of complex V1 neurons, S and T, and scale the T output by a factor = $1/\phi$ prior to the WIM algorithm (Equation 2), we will produce a sensor tuned to speed $v_2 = \phi v_1$.

Note that the same result could be derived with the S and T neurons inter-changed in the above treatment, such that a scaling factor is applied to the S neuron rather than the T neuron. In fact, both the S and T outputs could be scaled to keep the overall gain of the WIM sensor constant. For simplicity, only the T scaling option has been presented here and this choice is based on convention (see the WIM model section above).

Results

Figure 3 shows examples of the theory being put into practice. The same S and T units that generated the spectral receptive field in Figure 1c (tuned to 2 deg/s) were used to generate units tuned to 1 deg/s (Figure 3a) and 4 deg/s (Figure 3b), simply by scaling the T sensitivity by $\phi = 2$ and 0.5 for Figures 3a and 3b, respectively. Figure 3c shows the speed tuning curves for the three different sensors. These were generated using a moving bar (20 pixels wide) and two-dimensional image-based versions of the WIM sensors (Perrone, 2004). By changing the size of the scaling parameter, ϕ , a wide continuous range of speed tuning values can be generated.

Figure 4 is a still frame from an animated movie (Figure 5) demonstrating the variable speed tuning mechanism. For clarity, the actual movie does not contain all of the text labels. The left hand part of the figure shows the sustained and transient V1 neuron spectral receptive fields in perspective plot form. The sustained amplitude plot is shown in red, and it is rendered slightly transparent to make the locus of intersection of the two functions more apparent. Note that the axes in this plot are linear, and so the spatial and temporal frequency contrast sensitivity profiles will differ from those in Figure 1.

In the movie shown in Figure 5, the amplitude of the transient unit is being scaled up and down using values of ϕ that range from 0.3 to 4. Note how the two surfaces of the S and T functions intersect on a straight line in the (u, ω) plane. This is the basis of the WIM model, and it comes about because of the special way the transient spatial function, $f'(u)$, is constructed (Equation 1). No other spatial function will generate a locus of intersection that is exactly

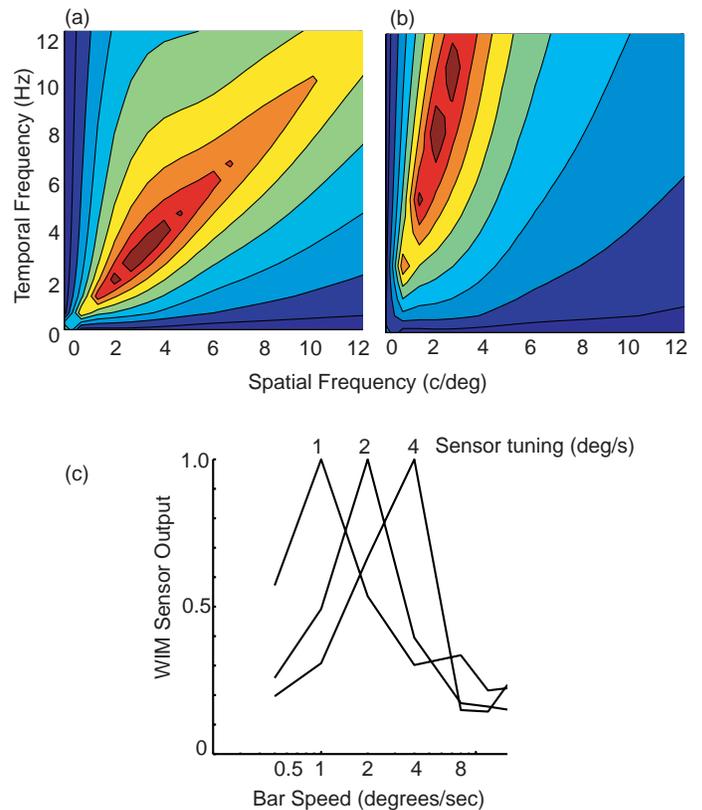


Figure 3. Speed tuning changes brought about by a simple weighting of the V1 neuron inputs. The sensors shown here use the same two V1 components as in Figure 1c. (a). Sensor tuned to 1 deg/s. (b). Sensor tuned to 4 deg/s. (c). Speed tuning curves for the two new sensors and the original sensor. All outputs have been normalized to the maximum.

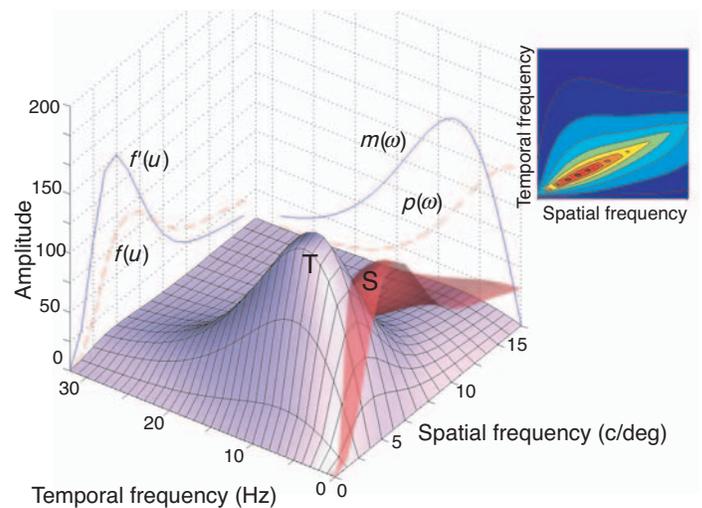


Figure 4. Explanation of the movie sequence demonstrating the variable speed tuning mechanism. The amplitude scale has been stretched in this figure to help reveal the S and T spatiotemporal frequency perspective plots.

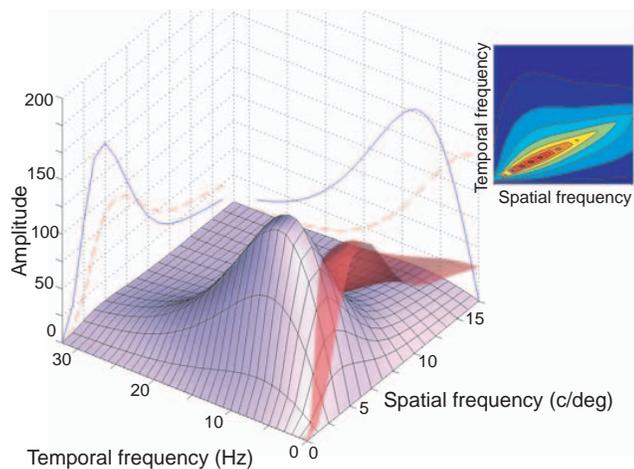


Figure 5. Animated movie sequence demonstrating the variable speed tuning mechanism. It is best viewed frame by frame using the slider control in QuickTime.

straight and oriented in this manner. Notice also how the slope of this line changes with different values of ϕ . The locus of intersection remains straight only for different values of ϕ because of the special relationship between the $p(\omega)$ and $m(\omega)$ temporal functions (Equation 5). Other temporal functions without this property will not retain the exact linear intersection as ϕ changes.

The spectral receptive field of the WIM sensor generated by the S and T neurons is shown as an inset in the upper right part of the movie. Locations along the point of intersection of the S and T surfaces correspond to maximally sensitive regions of the WIM sensor spectral receptive field (see Equation 2).

Discussion

I have shown that a single S-T pair of V1 neurons can generate a huge number of speed tunings simply by adjusting the strength of the connections between the V1 and WIM stages. The variable speed tuning mechanism is an amazingly economical strategy that relies on a special relationship between the temporal frequency tuning curves of the different V1 neuron classes used in the WIM model. Whether or not the primate brain has actually capitalized on this source of economy will be difficult to establish. One would need to find the appropriate matched pairs of V1 neurons that feed into the putative WIM stage and test their temporal frequency tuning. The ratio of the two responses at each tested temporal frequency should follow the k/ω rule (Equation 5). The currently available physiological data from V1 neurons (e.g., Figure 2) can certainly accommodate the functions required for the variable speed tuning mechanism to work.

As mentioned in the section on the WIM model, the speed tuning mechanism relies to some extent on the fact that the S and T neuron spatiotemporal response functions are separable (within a single quadrant). The development

of the variable speed tuning mechanism presented above is certainly simplified by assuming that $S(u, \omega) = f(u)p(\omega)$ and $T(u, \omega) = f'(u)m(\omega)$. The data from some V1 neurons show that this assumption is not unreasonable (Foster et al., 1985; Tolhurst & Movshon, 1975). However, mathematical convenience should not be mistaken for biological practicality. In the end, the basic WIM mechanism requires only that the S and T neuron spatiotemporal frequency functions overlap along a line given by $v = -\omega_i/u_i$. One way of achieving this is to assume separability and to use Equation 1, but there are other options. Two inseparable functions $S'(u, \omega)$ and $T'(u, \omega)$ could also be made to intersect along the $v = -\omega_i/u_i$ line by changing their overall shape. Similarly, the primate brain may have evolved S' and T' (non-separable) spatiotemporal frequency functions for its V1 neurons that enable the variable speed tuning mechanism to work. I have simply shown that if separability is a property of these neurons, then the theoretical ideal temporal frequency tuning curves for variable speed tuning will be ones based on the k/ω relationship (Equation 5). The WIM model and the variable speed tuning concept are not invalidated if further physiological studies reveal that the majority of V1 complex neurons are (one quadrant) inseparable in (u, ω) space.

The animated sequence in Figure 5 was intended to convey the continuous nature of the tuning mechanism and to demonstrate that very fine adjustments can be made to the optimum speed tuning value of the WIM sensor. As currently conceived, the different WIM sensors are assumed to be set at some optimum speed tuning value using a fixed weight (ϕ value). However, the animated sequence does raise the possibility of a dynamical system in which the speed tuning of the sensor could be altered rapidly in response to events occurring in other parts of the visual field or from extraretinal sources, such as eye movements.

Conclusion

Many accounts of motion processing in the brain tend to rely on the idea that neurons exist that are able to deliver a signal proportional to the speed of patterns moving over their receptive fields (see Perrone, 2001; Perrone, 2004). Neurons with this property have yet to be found. Instead, neurons in one of the key motion processing areas of the primate brain (MT) tend to be speed tuned. The fact that their responses fall off when a sub-optimal speed occurs is advantageous to global motion processing schemes that are based on “template matching” (e.g., Perrone, 1992; Perrone & Stone, 1994). However, speed tuning is an inefficient way of coding image speed compared to systems that directly output the speed value. Many neurons are required to register the wide range of possible speeds encountered during normal behaviors. The variable speed tuning mechanism outlined in this work overcomes this issue of resource intensiveness and shows that speed tuning can be both useful and economical.

Acknowledgments

Thanks to Rich Krauzlis for his helpful comments on previous drafts and to Frederic Glanois for his input into the early stages of this project. Thanks again to Michael Hawken for providing some of the V1 temporal frequency tuning data.

Commercial relationships: none.

Corresponding author: John A. Perrone.

Email: jpnz@waikato.ac.nz.

Address: Psychology Dept., The University of Waikato, Private Bag 3105, Hamilton, New Zealand.

References

- Felleman, D. J., & Kaas, J. H. (1984). Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, *52*(3), 488-513. [[PubMed](#)]
- Foster, K. H., Gaska, J. P., Nagler, M., & Pollen, D. A. (1985). Spatial and temporal frequency selectivity of neurones in visual cortical areas V1 and V2 of the macaque monkey. *Journal of Physiology*, *365*, 331-363. [[PubMed](#)]
- Hawken, M. J., & Parker, A. J. (1987). Spatial properties of neurons in the monkey striate cortex. *Proceedings of the Royal Society of London B*, *231*(1263), 251-288. [[PubMed](#)]
- Hawken, M. J., Shapley, R. M., & Gross, D. H. (1996). Temporal frequency selectivity in monkey visual cortex. *Visual Neuroscience*, *13*, 477-492. [[PubMed](#)]
- Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, *49*(5), 1127-1147. [[PubMed](#)]
- Perrone, J. A. (1992). Model for the computation of self-motion in biological systems. *Journal of the Optical Society of America A*, *9*(2), 177-194. [[PubMed](#)]
- Perrone, J. A. (2001). A closer look at the visual input to self-motion estimation. In J. M. Zanker & J. Zeil (Eds.), *Motion vision: Computational, neural, and ecological constraints* (pp. 169-179). Heidelberg: Springer-Verlag.
- Perrone, J. A. (2004). A visual motion sensor based on the properties of V1 and MT neurons. *Vision Research*, *44*(15), 1733-1755. [[PubMed](#)]
- Perrone, J. A., & Stone, L. S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research*, *34*(21), 2917-2938. [[PubMed](#)]
- Perrone, J. A., & Thiele, A. (2001). Speed skills: Measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience*, *4*(5), 526-532. [[PubMed](#)]
- Perrone, J. A., & Thiele, A. (2002). A model of speed tuning in MT neurons. *Vision Research*, *42*(8), 1035-1051. [[PubMed](#)]
- Priebe, N. J., Cassanello, C. R., & Lisberger, S. G. (2003). The neural representation of speed in macaque area MT/V5. *Journal of Neuroscience*, *23*(13), 5650-5661. [[PubMed](#)]
- Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743-761. [[PubMed](#)]
- Tolhurst, D. J., & Movshon, J. A. (1975). Spatial and temporal contrast sensitivity of striate cortical neurones. *Nature*, *257*(5528), 674-675. [[PubMed](#)]
- Watson, A. B. (1986). Temporal sensitivity. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 6.1-6.42). New York: Wiley.
- Watson, A. B., & Ahumada, A. J. (1983). A look at motion in the frequency domain. In J. K. Tsotsos (Ed.), *Motion: Perception and representation* (pp. 1-10). New York: Association for Computing Machinery.