Computational theory of the responses of V1 & MT neurons and psychophysics of motion perception

Neural circuits perform computations

~50,000 neurons per cubic mm
~6,000 synapses per neuron
~10 billion neurons & ~60 trillion synapses in cortex

Computational theory: how do neurons compute motion?

Hubel & Wiesel (1968)

Motion is like orientation in space-time and spatiotemporally oriented filters can be used to detect and measure it.

Adelson & Bergen (1985)
Strong response for motion in preferred direction.

Weak response for motion in non-preferred direction.

Direction selectivity model

Space-time receptive field

Distributed representation of speed

Each spatiotemporal filter computes something like a derivative of image intensity in space and/or time. "Perceived speed" is the orientation corresponding to the gradient in space-time (max response).

Impulse response

Strong response to preferred direction

Weak response to opposite direction

Note: negative responses not seen in neural firing rates
'On' and 'off' responses

Complex cells: theory

Complex cells & position invariance

Motion energy responses to moving grating

Computing space-time RFs & motion energy

Cascade of temporal low-pass filters
Odd- and even-phase spatial weights

Space-time oriented impulse responses

Odd- and even-phase spatial weights

Space-time separable impulse responses

Matlab code

```
n = [3,5,5,7];
for tt = 1:size(input,1)
    % Temporal filters
    deltaT = (deltaT/tau) * (-y(n(:,1)) + y(n(:,4))); y(n(:,1)) = y(n(:,1)) + deltaT;
    for nn = 1:size(n)
        if n(nn) == 1
            deltaT = (deltaT/tau) * (-y(n(n,1)) + y(n(n,1)-1,1)); y(n(n,1)) = y(n(n,1)) + deltaT;
    end
    rtFast = y(n(1,:)-y(n(2,:)); rtSlow = y(n(3,:)-y(n(4,:)));
    % Spatial filters
    oddFast = spatialConvolution(rtFast, oddFilt); oddSlow = spatialConvolution(rtSlow, oddFilt); evenFast = spatialConvolution(rtFast, evenFilt); evenSlow = spatialConvolution(rtSlow, evenFilt);
    % Direction selective filters and motion energy
    leftEven = oddFast + evenSlow;
    leftOdd = -oddSlow + evenFast;
    leftEnergy = leftEven.' + leftOdd.';
    rightEven = -oddFast + evenSlow;
    rightOdd = oddSlow + evenFast;
    rightEnergy = rightEven.' + rightOdd.';
end
```

Direction-selective motion energy

The “aperture problem”

These three motions are different but look the same when viewed through a small aperture (i.e., that of a direction-selective receptive field).

Wallach (1935)
**Intersection of constraints**

With two different motion components within the aperture, there is a unique solution:

\[
\begin{align*}
V_x & \uparrow \\
V_y & \uparrow
\end{align*}
\]

Adelson & Movshon (1981)

**Component vs. pattern motion (perception)**

\[
\begin{align*}
&\text{strong pattern-motion percept} \\
&\text{weak pattern-motion percept}
\end{align*}
\]

Adelson & Movshon (1981)

**Component vs. pattern motion selectivity**

\[
\begin{align*}
&\text{component-motion cell} \\
&\text{pattern-motion cell}
\end{align*}
\]

Adelson & Movshon (1981)

**Component vs. pattern motion: single neurons**

Movshon et al., 1983

**Component vs. pattern motion: fMRI adaptation**

Adapted =

Movshon et al., 1983

Model

Adapted =

Huk & Heeger (2002)
**Pattern motion selectivity across visual areas**

Huk & Heeger (2002)

**Intersection of constraints (two components)**

Each component activates a different V1 neuron, selective for a different orientation and speed.

**Intersection of constraints (many components)**

Each component activates a different V1 neuron, selective for a different orientation and speed.

**Neural implementation of IOC**

Answer: For each possible 2D velocity, add up the responses of those V1 neurons whose preferred orientation and speed is consistent with that 2D velocity.

**Spatiotemporal frequency domain**

Spatiotemporal frequency response of space-time oriented linear filter.

Frequency responses of filters that are all consistent with one velocity.
Linedar weighting

Distributed representation of 2D velocity

Brightness at each location represents the firing rate of a single MT neuron with a different preferred velocity. Location of peak corresponds to perceived velocity.

Predictions of the theory

Velocity of a random dot stimulus and the velocities of each oriented component.

A single component velocity is consistent with two pattern velocities at faster speed.

Kuman & Uka (2013)

Testing the theory: pattern cell

For CDS predictions, a periodic spline curve was interpolated to the direction-tuning data at the optimal speed. The direction tuning for a speed higher than the optimal speed was computed as the sum of two interpolated curves, each shifted by an amount determined from the ratio of the optimal speed to each speed.

Kuman & Uka (2013)

Testing the theory: component cell

For PDS predictions, the interpolated curve was used across all speeds.

Kumano & Uka (2013)

Visual motion ambiguity

Bias in perceived velocity

Stone, Watson, & Mulligan (1990)
Perception is our best guess as to what is in the world, given our current sensory input and our prior experience (Helmholtz, 1866).

Goal: explain "mistakes" in perception as "optimal" solutions given the statistics of the environment.

Bayesian models of perception

Prior bias for slower speeds

Bayesian estimation of velocity

Bayesian estimation of velocity

Bayesian estimation of velocity

Bayesian estimation of velocity
Perceived direction bias (deg).

The accuracy of the representation scales inversely with a good sensor system would allocate a higher proportion of neurons based on this population. A close form solutions are readily obtained using this density function to warp a homogenous population. If one assumes the environment is inhomogeneous, in that the frequency of occurrence are increased. This parameterization allows us to optimize the population directly. Specifically, the cell density is proportional to the stimulus value, stimulus duration, or intensity. As a result, the so-called allocation of neurons using a contextual parameterization allows us to optimize the population arrangement so as to maximize the information, which provides a bound on the variance of stimulus transmission for a variety of auditory attributes (acoustic frequency and modulation frequency), and three visual attributes. Consider a stimulus variable, heterogeneous in their frequency of occurrence, this should be the case, or how these heterogeneities might be accounted for.

Theory fits lots of behavioral data

How does the brain represent the prior?

Fisher information

Prior

\[
\int \frac{d(s)}{d(s)} g(s) \, ds, \quad \text{Firing rate} \quad \text{Neuronal total spike rate} \]

\[
\int d(s) \, ds = N \quad \text{and} \quad \int p(s) g(s) \, ds = R.
\]

The “principles”

- Perception is an inference that has evolved/developed to match the statistics of the environment (Bayesian estimation with priors that embody statistics of environment).
- Functional specialization. Each brain area (defined on the basis of physiology, architecture, connections, topography) performs a different function.
- Computational theory. Canonical computation (linear sum, threshold or sigmoid nonlinearity, adaptation) cascaded across a pathway of visual cortical areas. Selectivity and invariance.

A computational theory of motion appearance

A computational theory of color appearance

What distinguishes neural activity that underlies conscious visual appearance?

- Neural activity in certain brain areas.
- Activity of specific subtypes of neurons.
- Particular temporal patterns of neural activity (e.g., oscillations).
- Synchronous activity across groups of neurons in different brain areas.
- Neural activity that is driven by a coherent combination of bottom-up sensory information and top-down recurrent processing (e.g., linked to attention).
- Nothing. Once you know the computations, you’re done!