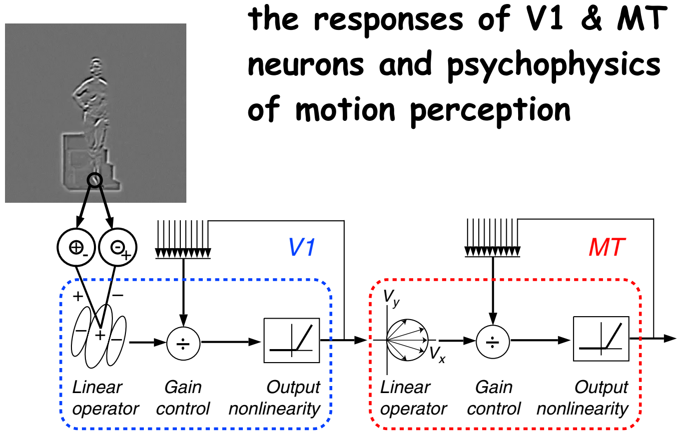
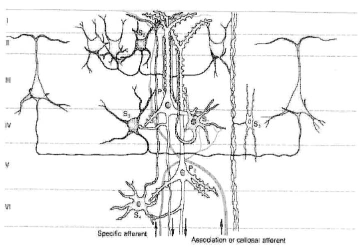
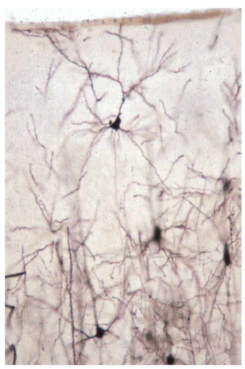
Computational theory of the responses of V1 & MT neurons and psychophysics of motion perception



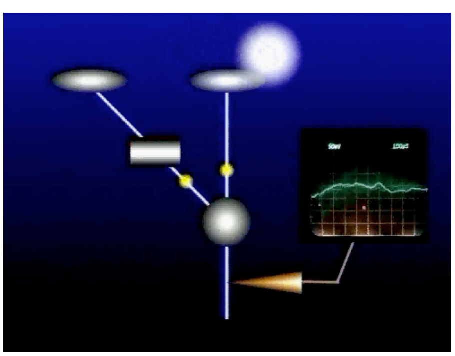
Neural circuits perform computations



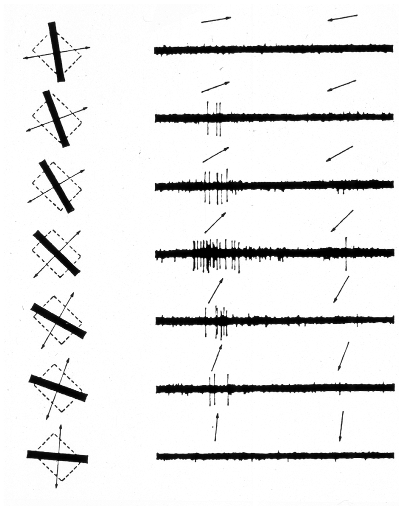


~50,000 neurons per cubic mm
~6,000 synapses per neuron
~10 billion neurons & ~60 trillion synapses in cortex

Computational theory: how do neurons compute motion?

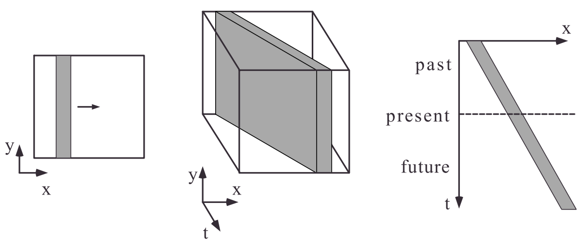


Direction selectivity



Hubel & Wiesel (1968)

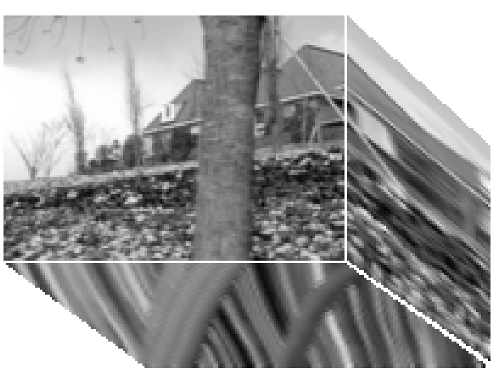
Orientation in space-time



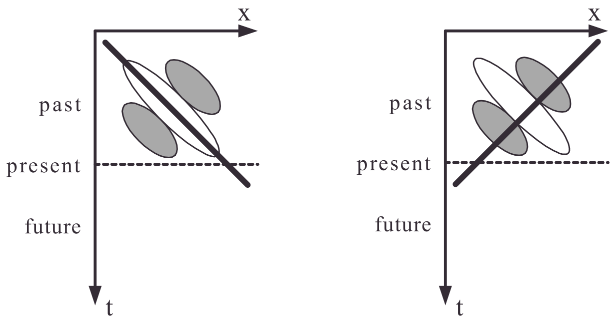
Motion is like orientation in space-time and spatiotemporally oriented filters can be used to detect and measure it.

Adelson & Bergen (1985)

Motion is orientation in space-time



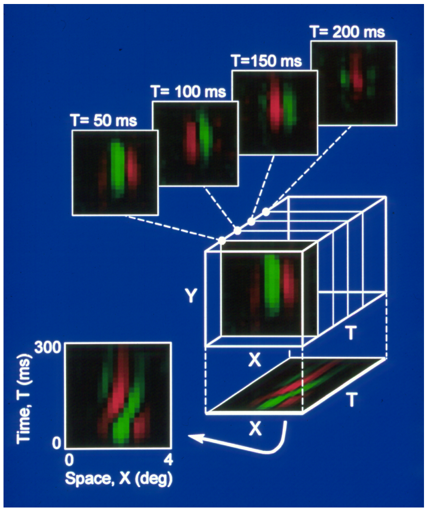
Direction selectivity model



Strong response for motion in preferred direction.

Weak response for motion in non-preferred direction.

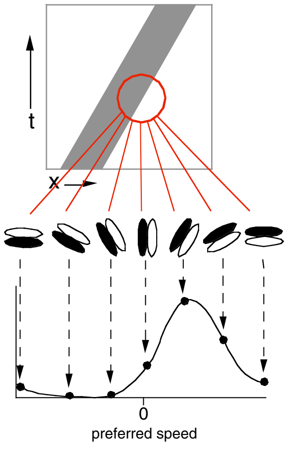
Space-time receptive field



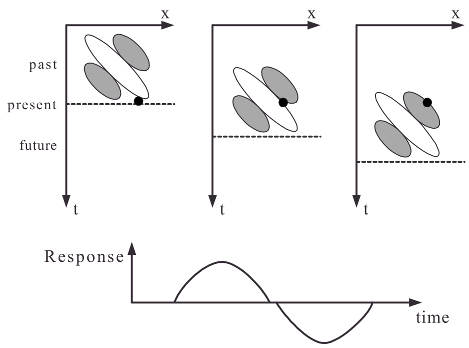
Ohzawa, DeAngelis, & Freeman

Distributed representation of speed

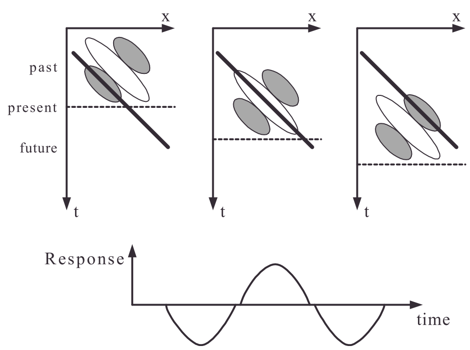
Each spatiotemporal filter computes something like a derivative of image intensity in space and/or time. "Perceived speed" is the orientation corresponding to the gradient in space-time (max response).



Impulse response

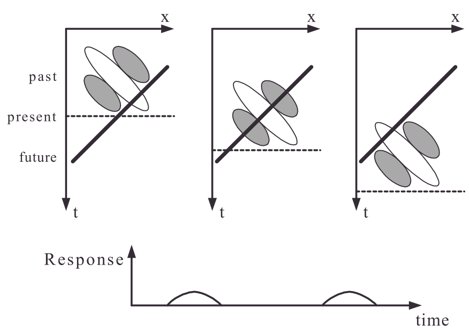


Strong response to preferred direction



Note: negative responses not
seen in neural firing rates

Weak response to opposite direction

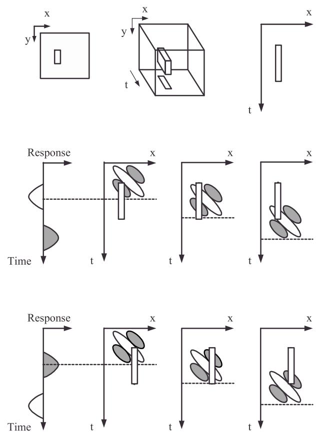


'On' and 'off' responses

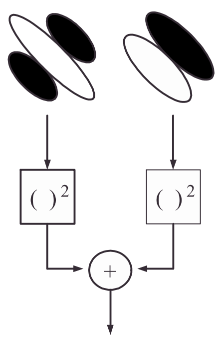
Stimulus

Off response

On response

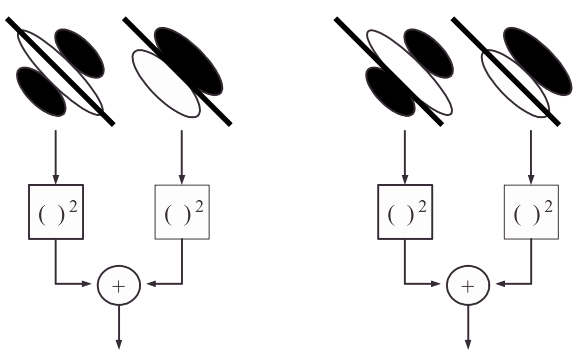


Complex cells: theory

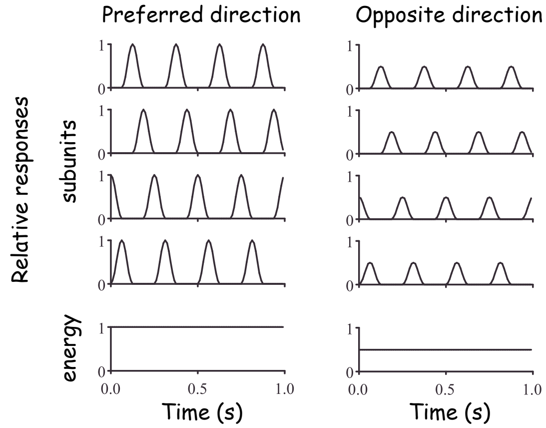


Complex cells & position invariance

Oriented stimulus as seen by both subunits at two different locations:



Motion energy responses to moving grating



Computing space-time RFs & motion energy

Cascade of recursive (streaming) low pass filters:

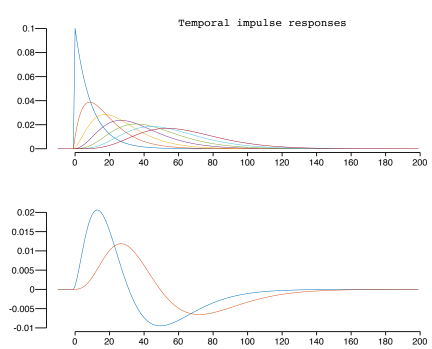
$\tau \, dy_1/dt = -y_1 + y_0$
$\tau \, dy_n/dt = -y_n + y_{n-1}$

$y_0(x,t)$: stimulus
$y_i(x,t)$: spatial array of temporally-filtered responses

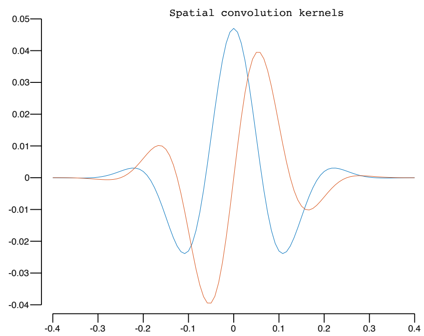
Biphasic temporal fliters:

$f_1 = y_3 - y_5$
$f_2 = y_5 - y_7$

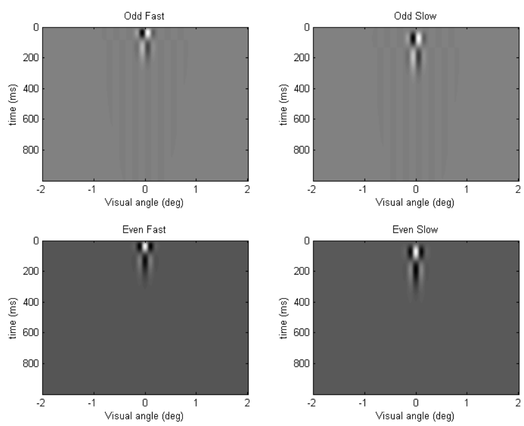
Cascade of temporal low-pass filters



Odd- and even-phase spatial weights

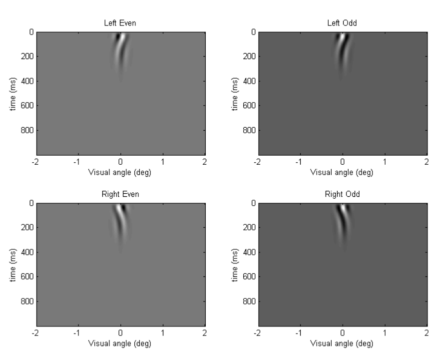


Space-time separable impulse responses



Space-time oriented impulse responses

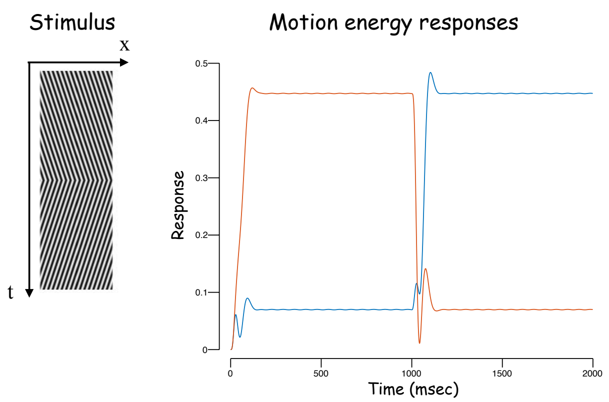
```
leftEven = oddFast + evenSlow;
leftOdd = -oddSlow + evenFast;
etc.
```



Matlab code

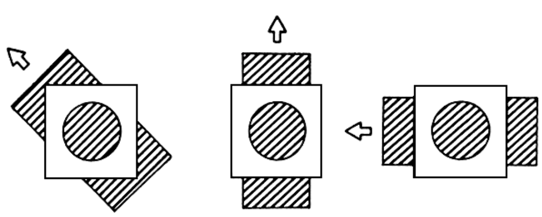
```
n = [3,5,5,7];
for tt = 1:size(input,1)

  % Temporal filters
  deltaY = (deltaT/tau) * (- y(1,:) + input(tt,:));
  y(1,:) = y(1,:) + deltaY;
  for nn = 2:max(n)
    deltaY = (deltaT/tau) * (-y(nn,:) + y(nn-1,:));
    y(nn,:) = y(nn,:) + deltaY;
  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);
  evenSlow = spatialConvolution(rtSlow,evenFilt);
  evenFast = spatialConvolution(rtFast,evenFilt);

  % Direction selective filters and motion energy
  leftEven = oddFast + evenSlow;
  leftOdd = -oddSlow + evenFast;
  leftEnergy = leftEven.^2 + leftOdd.^2;
  rightEven = -oddFast + evenSlow;
  rightOdd = oddSlow + evenFast;
  rightEnergy = rightEven.^2. + rightOdd.^2;

end
```

Direction-selective motion energy



The "aperture problem"

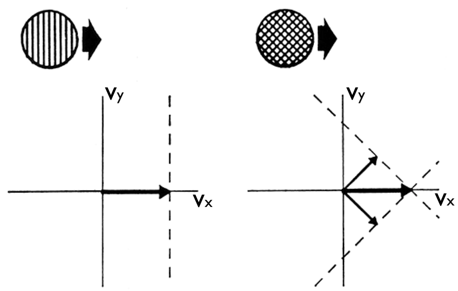
These three motions are different but look the same when viewed through a small aperture (i.e., that of a direction-selective receptive field).



Wallach (1935)

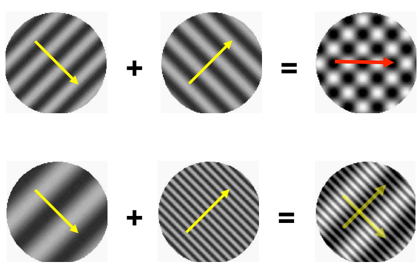
Intersection of constraints

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



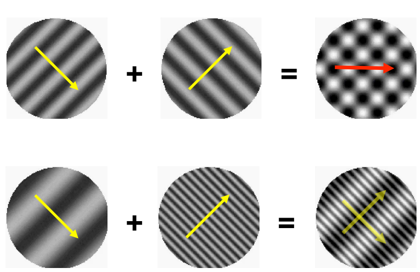
Adelson & Movshon (1981)

Component vs. pattern motion (perception)



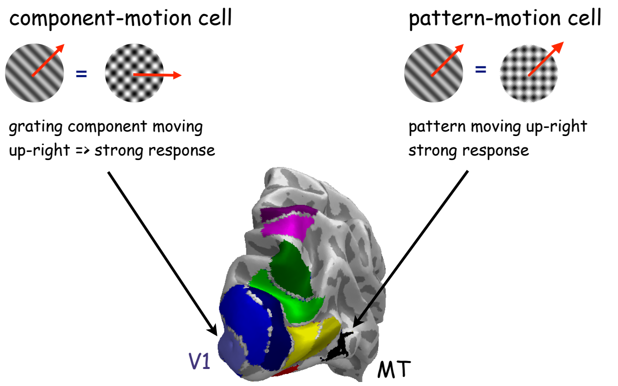
Adelson & Movshon (1981)

Component vs. pattern motion (perception)

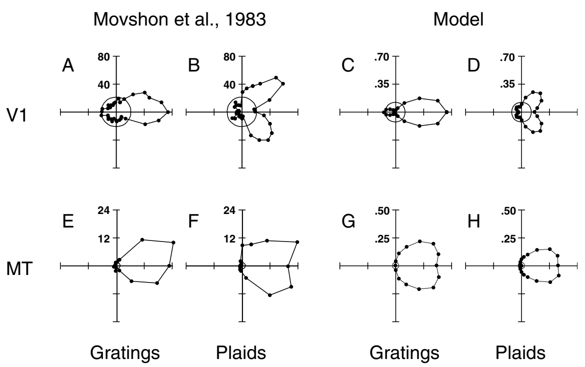


Adelson & Movshon (1981)

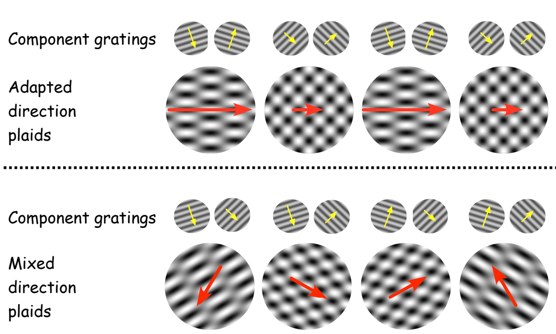
Component vs. pattern motion selectivity



Component vs. pattern motion: single neurons

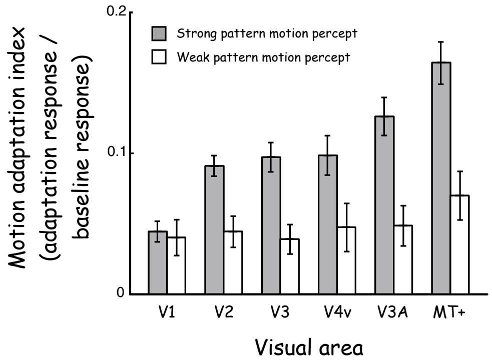


Component vs. pattern motion: fMRI adaptation



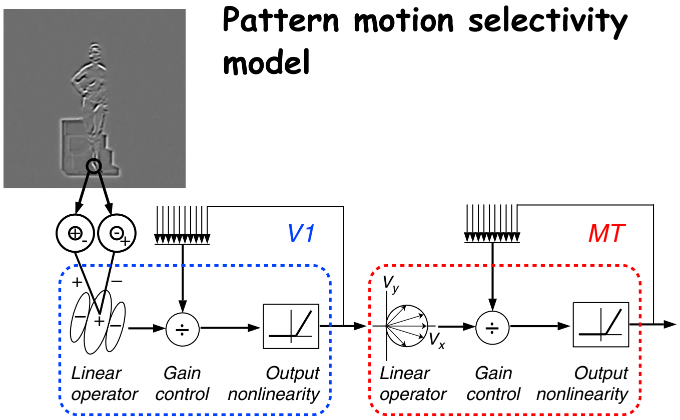
Huk & Heeger (2002)

Pattern motion selectivity across visual areas



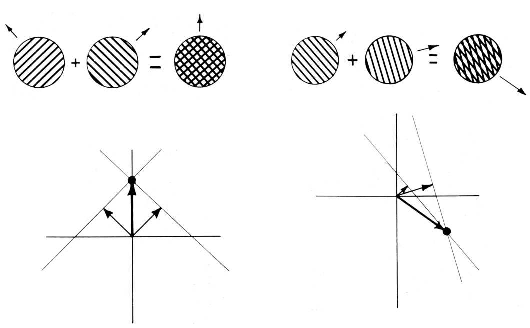
Huk & Heeger (2002)

Pattern motion selectivity model



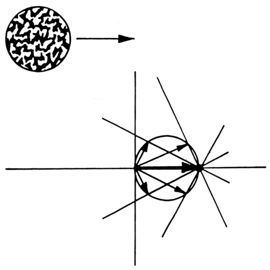
Simoncelli & Heeger (1998)

Intersection of constraints (two components)



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

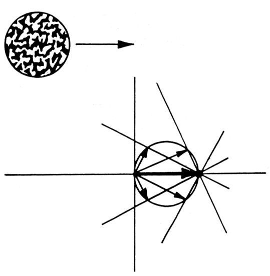
Intersection of constraints (many components)



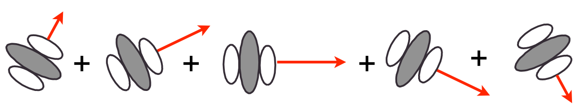
Each component activates a different V1 neuron, selective for a different orientation and speed.

How do you get selectivity for the moving pattern as a whole, not the individual components?

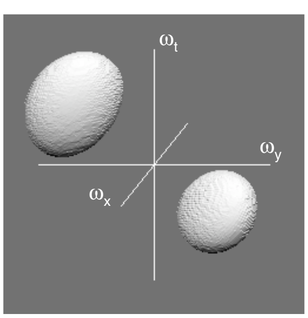
Neural implementation of IOC



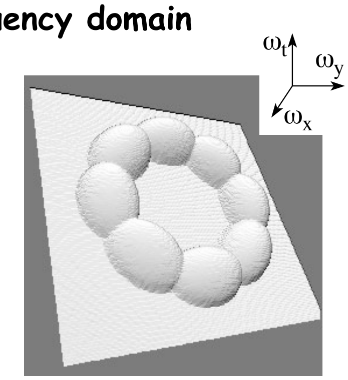
Answer: For each possible 2D velocity, add up the responses of those V1 neurons whose preferred orientation and speed is consistent with that 2D velocity.



Spatiotemporal frequency domain

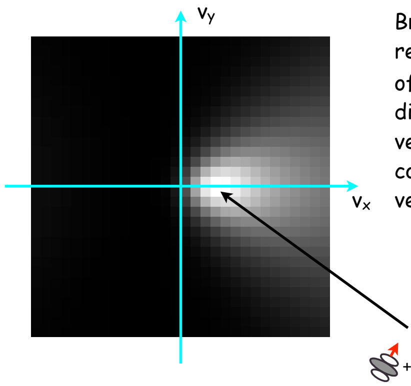


Spatiotemporal frequency response of space-time oriented linear filter.

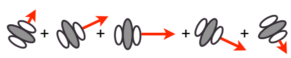


Frequency responses of filters that are all consistent with one velocity.

Distributed representation of 2D velocity

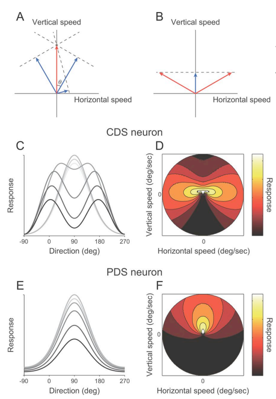


Brightness at each location represents the firing rate of a single MT neuron with a different preferred velocity. Location of peak corresponds to perceived velocity.



Predictions of the theory

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

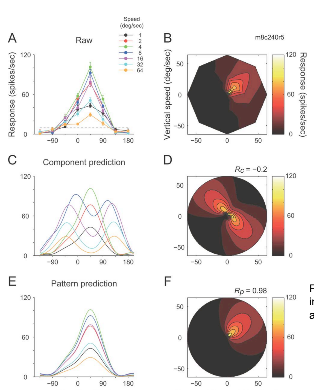


A single component velocity is consistent with two pattern velocities at faster speed.

Kuman & Uka (2013)

Testing the theory: pattern cell

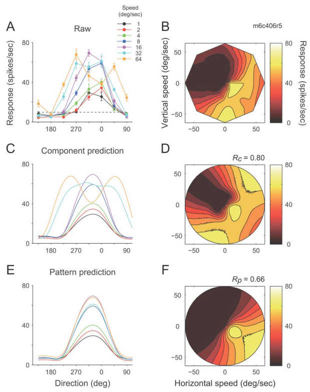
For CDS predictions, a periodic spline curve was interpolated to the direction-tuning data at the optimal speed. The direction tuning for a speed higher than the optimal speed was computed as the sum of two interpolated curves, each shifted by an amount determined from the ratio of the optimal speed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



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

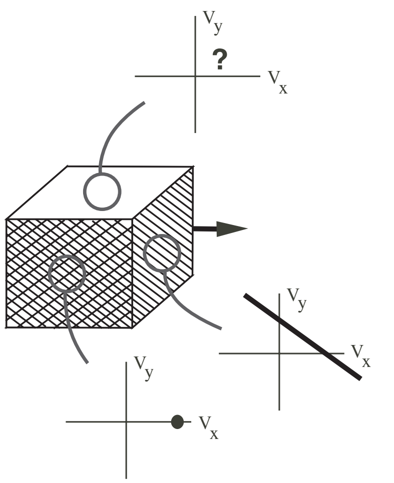
Kuman & Uka (2013)

Testing the theory: component cell

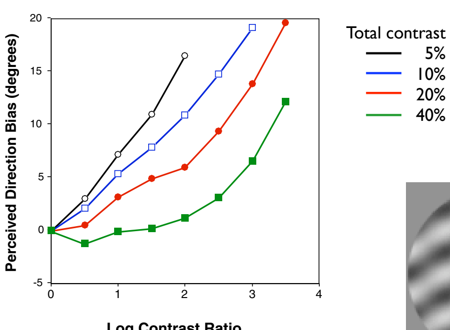


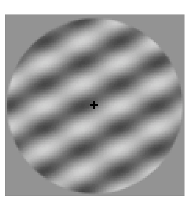
Kumano & Uka (2013)

Visual motion ambiguity



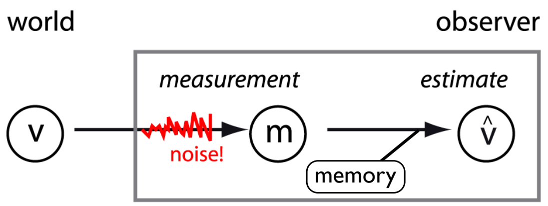
Bias in perceived velocity





Stone, Watson, & Mulligan (1990)

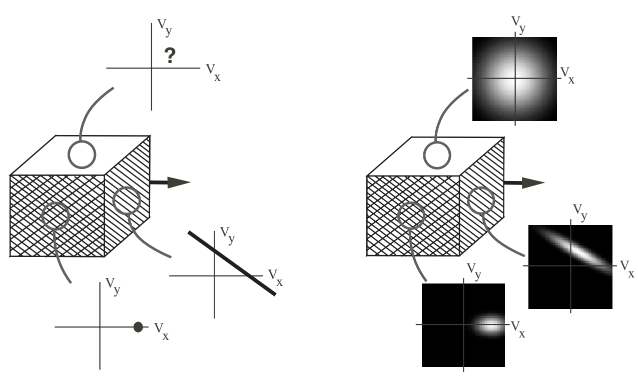
Bayesian models of perception



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

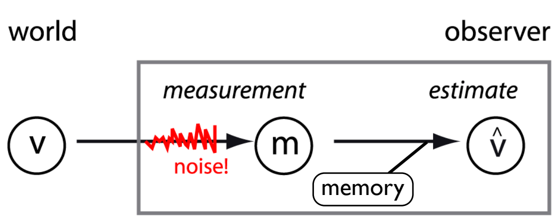
Goal: explain "mistakes" in perception as "optimal" solutions given the statistics of the environment.

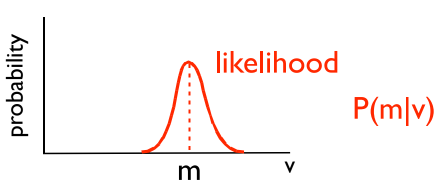
Prior bias for slower speeds



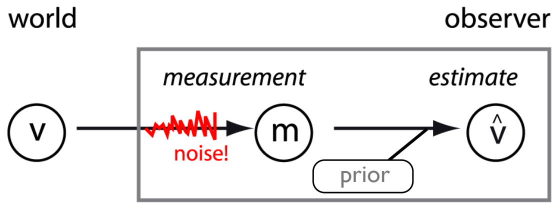
Simoncelli (1993)

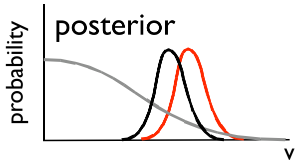
Bayesian estimation of velocity





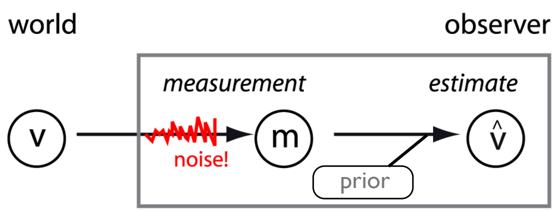
Bayesian estimation of velocity

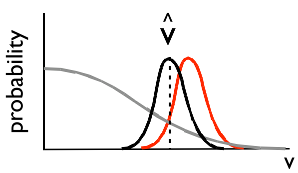




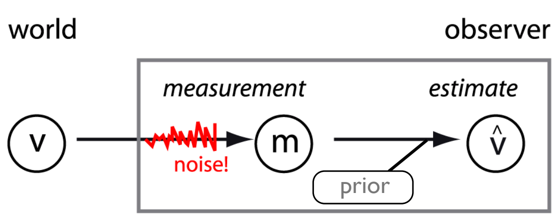
P(m|v) x P(v) ~ P(v|m)

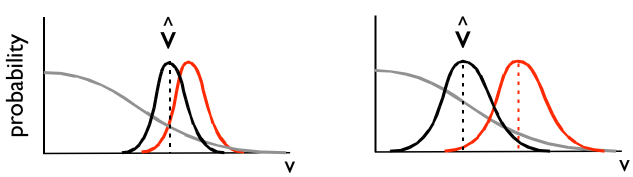
Bayesian estimation of velocity





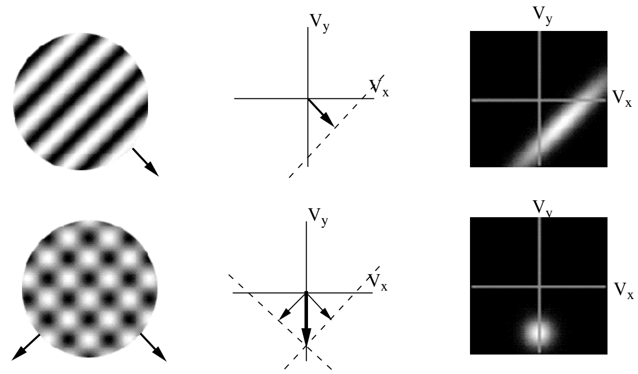
Bayesian estimation of velocity





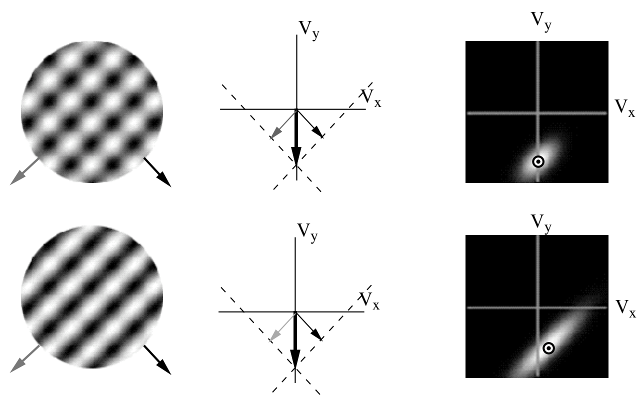
Bayesian model predictions

stimulus | idealization | model

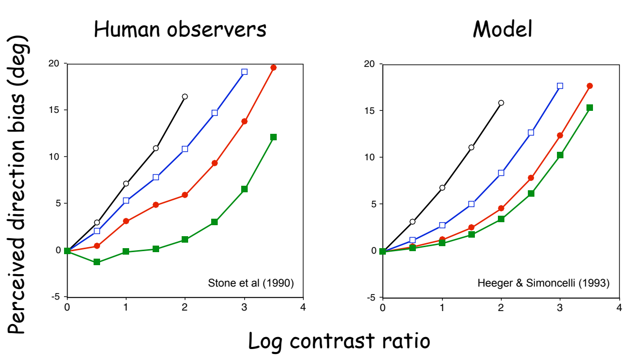


Bayesian model predictions

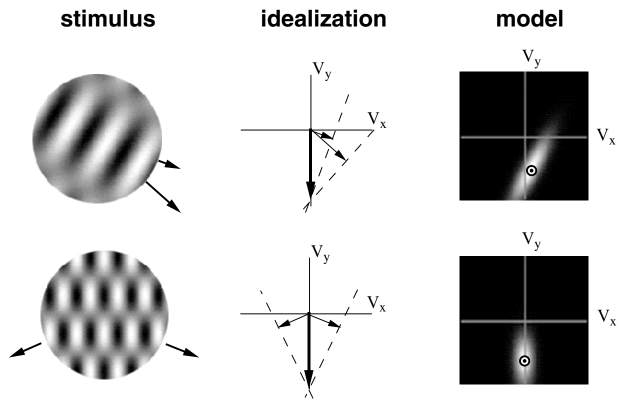
stimulus | idealization | model



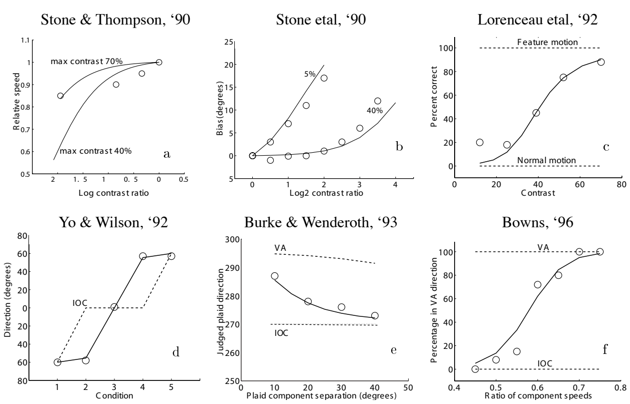
Prior for slow speeds explains bias in perceptual bias



Bayesian model predictions

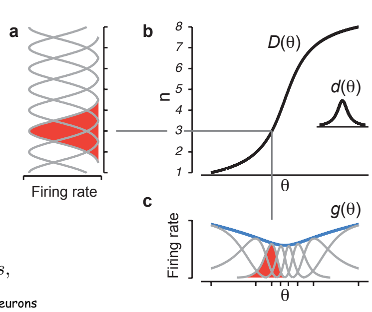


Theory fits lots of behavioral data



Weiss, Simoncelli, & Adelson (2002)
see also Stocker & Simoncelli (2006)

How does the brain represent the prior?



Fisher information

Prior

$$\underset{d(s),g(s)}{\operatorname{argmax}} \int p(s) \log\left(d^2(s) g(s)\right) \, \mathrm{d}s,$$

\# neurons

$$\int d(s)\, \mathrm{d}s = N,$$

Total spike rate

$$\int p(s) g(s)\, \mathrm{d}s = R$$

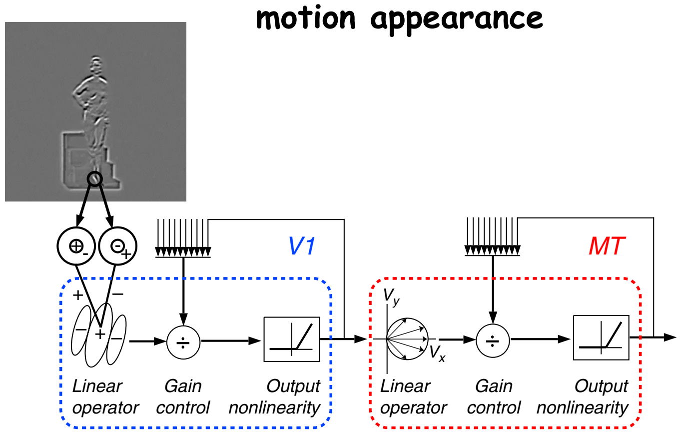
$$d(s) = Np(s), \quad g(s) = R.$$

Ganguli & Simoncelli (2014, 2016)

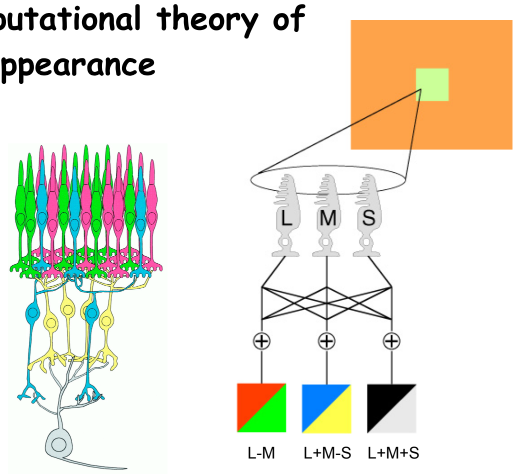
The "principles"

- 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!