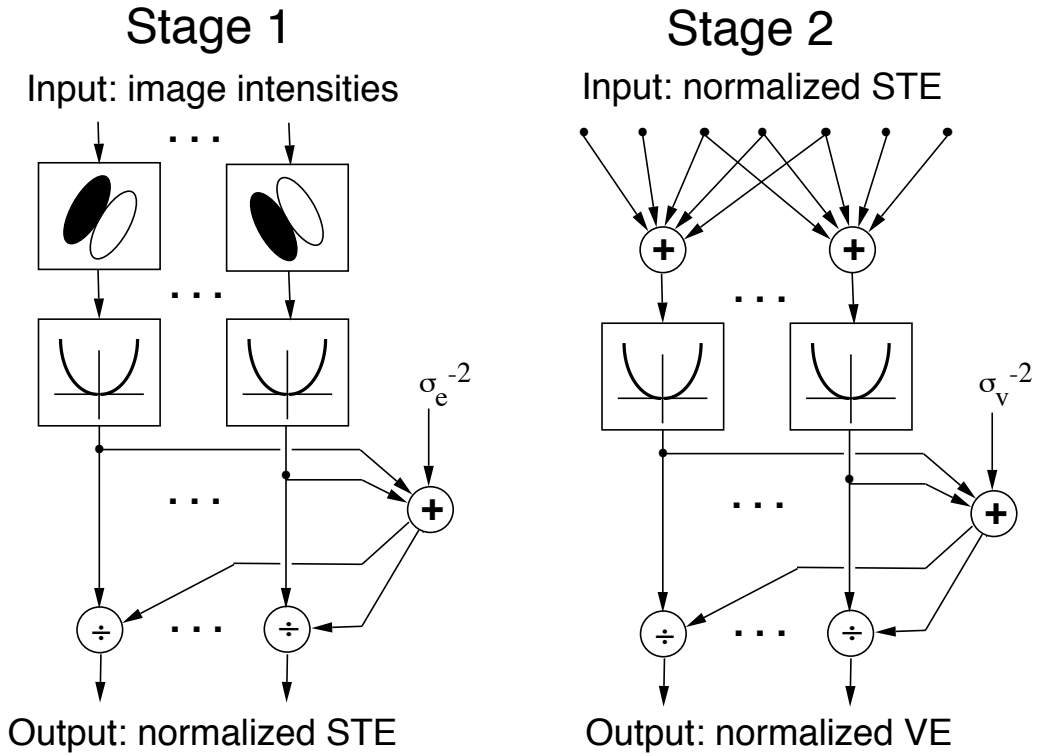
A Computational Model for Representation of Image Velocities

Eero Simoncelli
Computer and Information Science Department
University of Pennsylvania

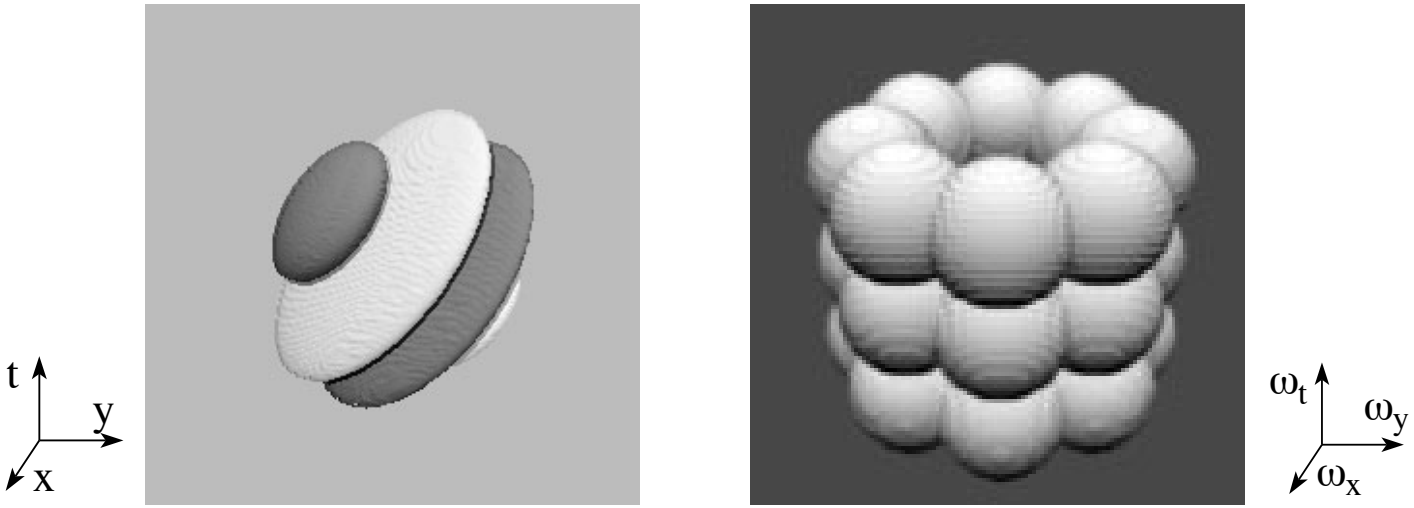
David Heeger
Psychology Department
Stanford University

The Model



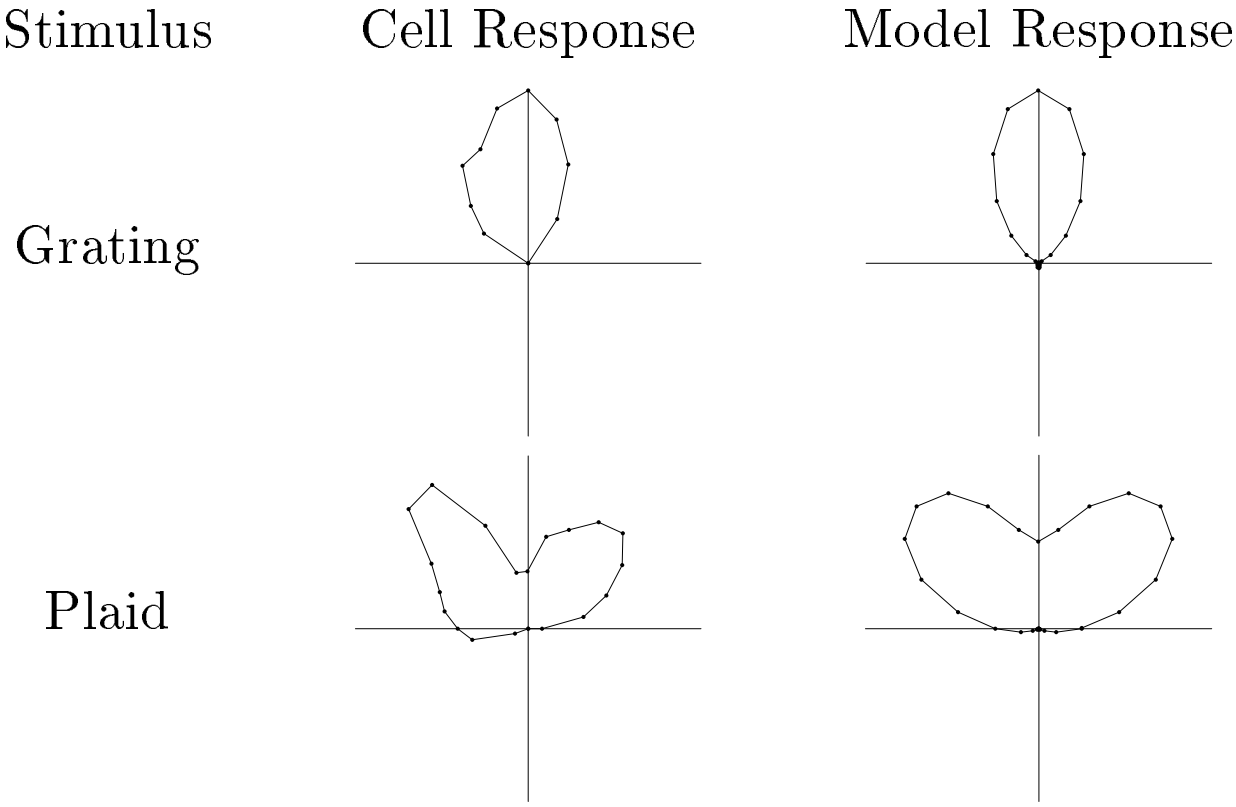
- First stage computes spatio-temporal energy (STE).
- Second stage computes "velocity energy" (VE).
- Simple computation: linear combination, squaring, normalization.

Spatio-temporal Energy (STE) Stage



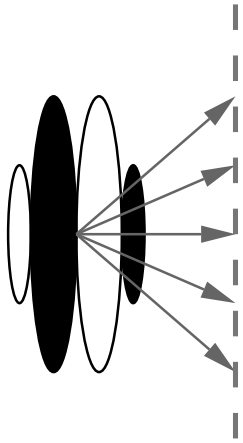
- Each unit is based on a linear combination of image intensities.
- Unit response is localized in spatio-temporal frequency.
- Frequency domain is "covered" by a minimal number of units. *Intermediate responses may be exactly interpolated.*

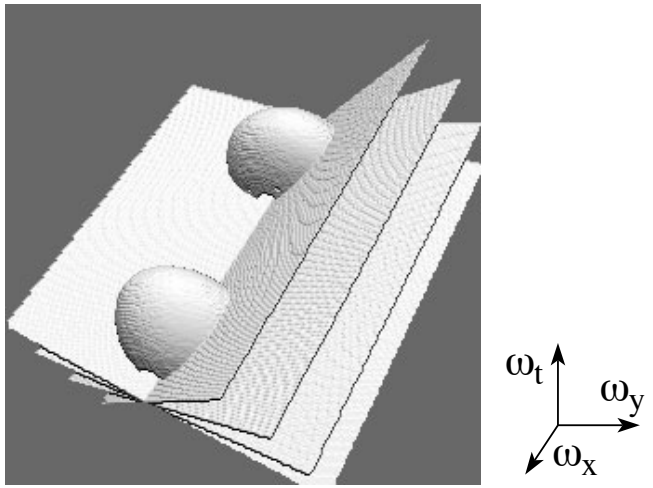
STE Unit vs. Complex Cell



- Polar plots of response vs. stimulus direction of movement.
- Single cell recordings replotted from Movshon et. al. (1986)

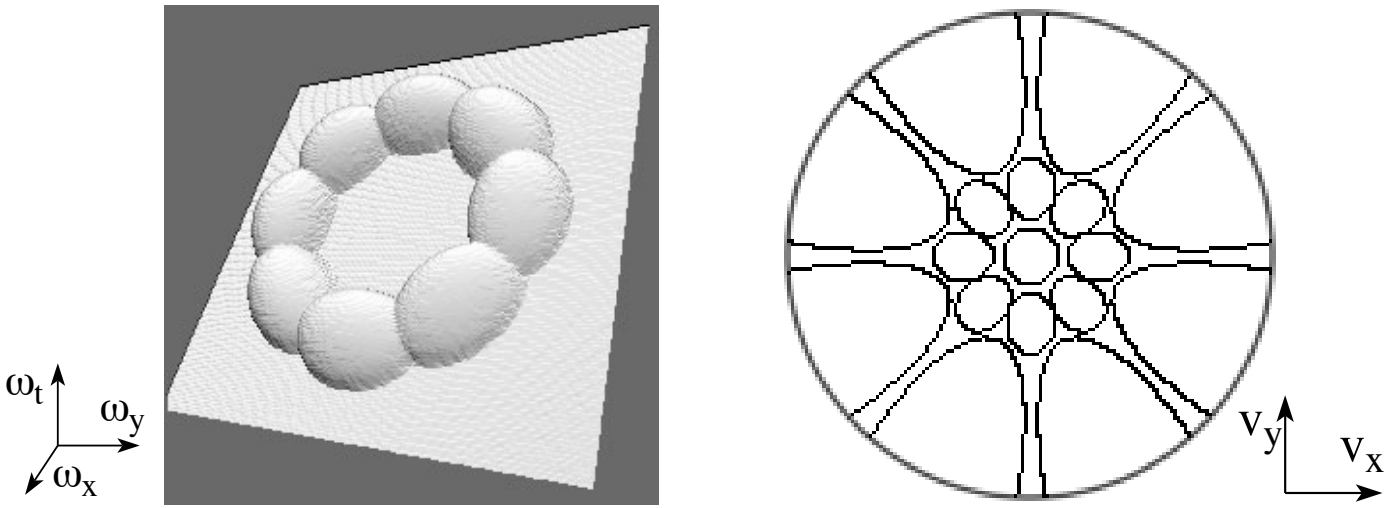
STE Unit Is Not Velocity-Tuned





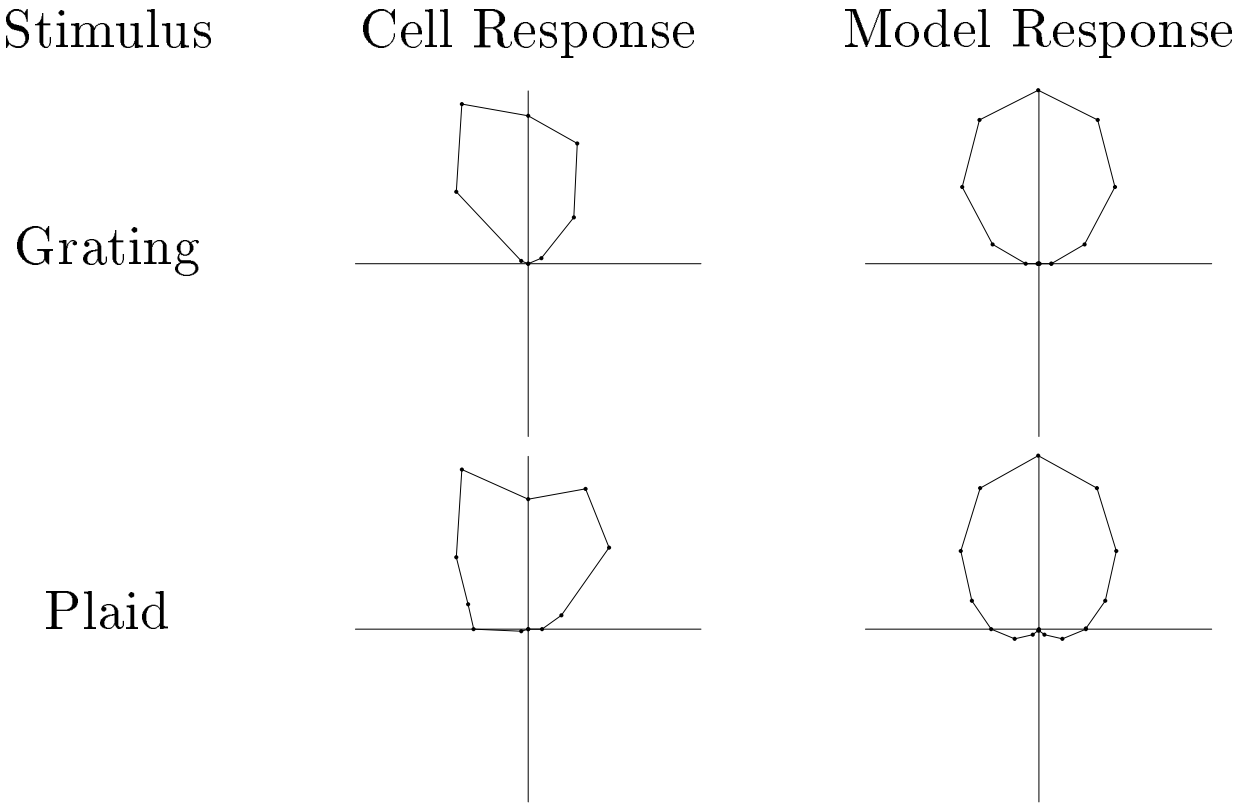
- Unit responds equally well to a whole family of velocities (aperture problem).
- Equivalently, in the Fourier domain, a family of planes cut through the unit.

Velocity Energy (VE) Stage



- Each unit is based on a linear combination of STE responses (as in Albright '84, Smith et al '92).
- STE responses are *interpolated* from previous stage.
- Unit response is localized in velocity space.
- Velocity domain is "covered" by a minimal number of units. *Intermediate responses may be exactly interpolated.*

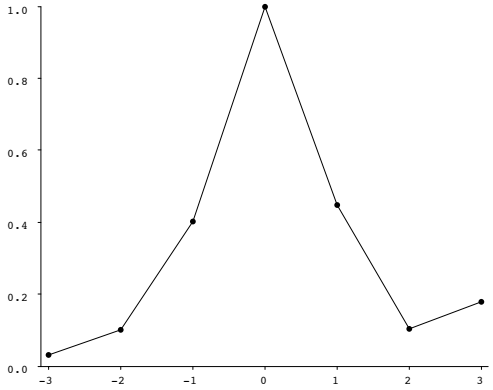
VE Unit vs. MT Cell



- Polar plots of response vs. stimulus direction of movement.
- MT cell recordings replotted from Movshon et. al. (1986)

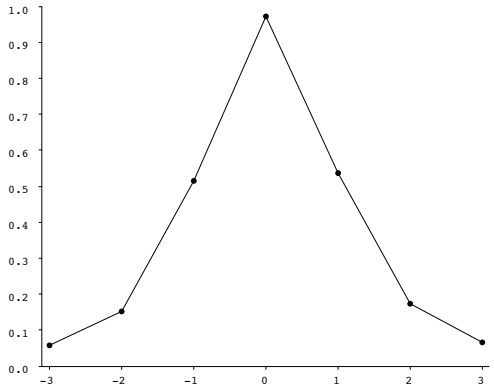
VE Unit vs. MT Cell

Cell Response



log speed

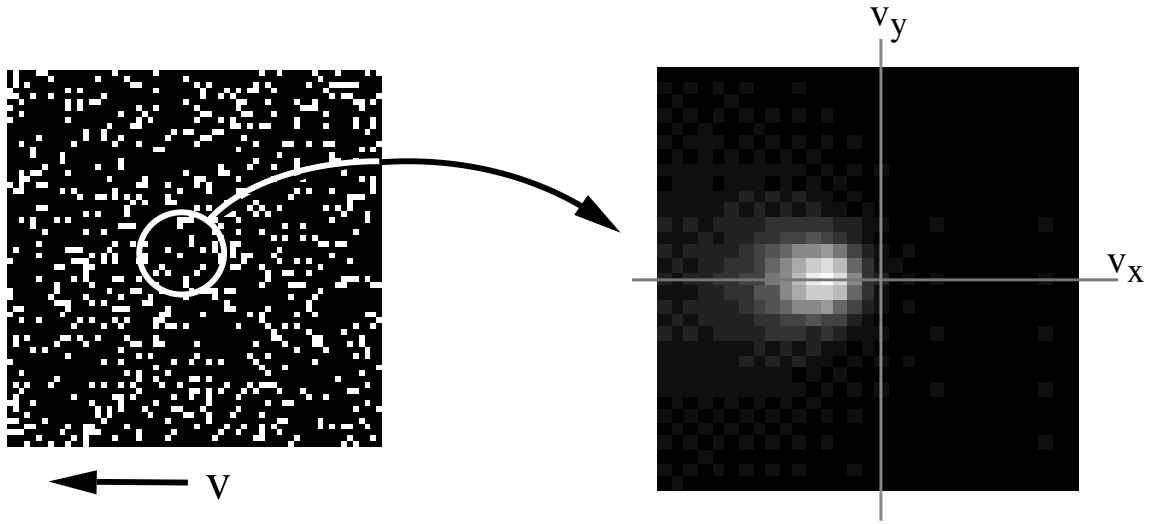
Model Response



log speed

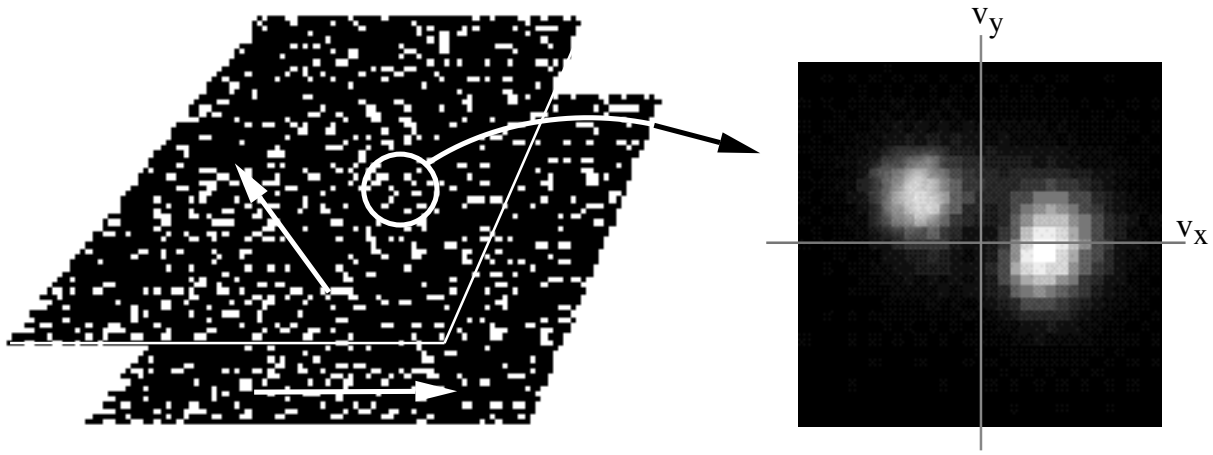
- Response to an oriented bar as a function of log speed.
- MT cell recordings replotted from Maunsell and Van Essen (1983).

VE Population Response: Single Motion



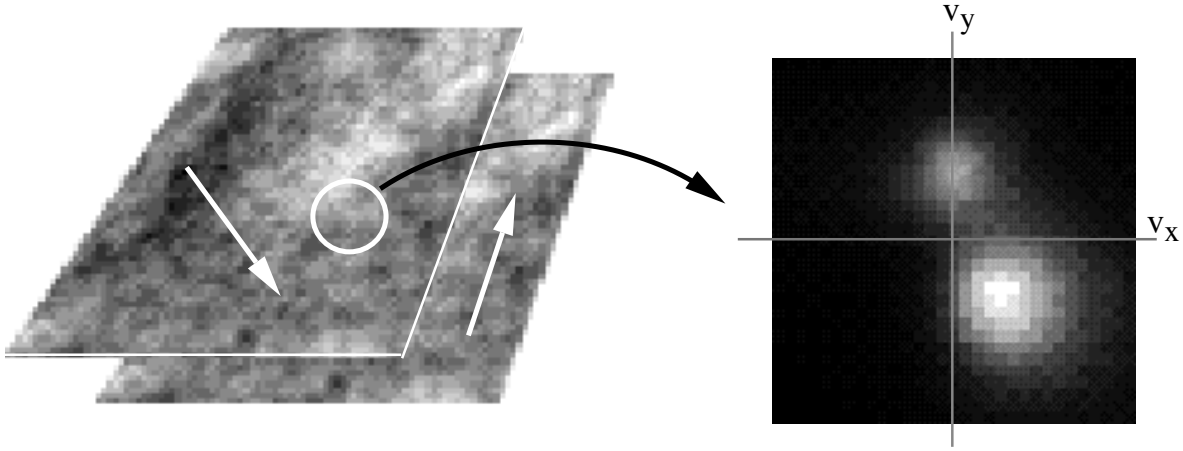
- Stimulus is a translating random dot pattern.
- Full velocity-space is interpolated from responses of a small number of units.
- Response is unimodal.

VE Population Response: Transparent Dots



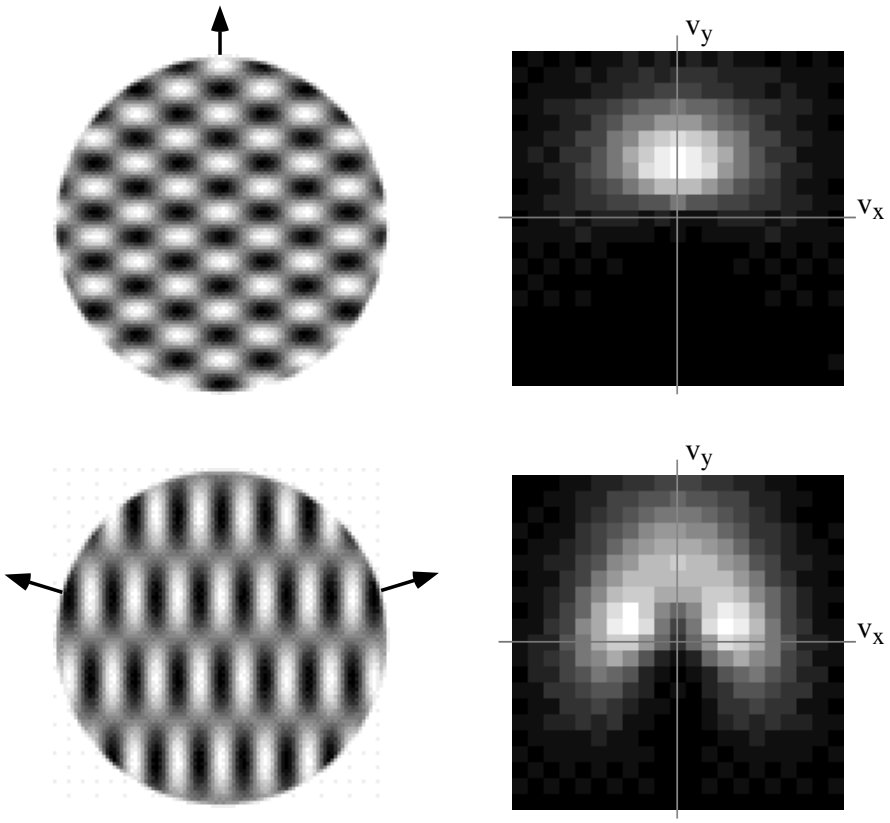
- Stimulus is two random dot patterns translating in different directions.
- VE Response is bimodal, indicating presence of two motions.

VE Population Response: Transparent Noise Patterns



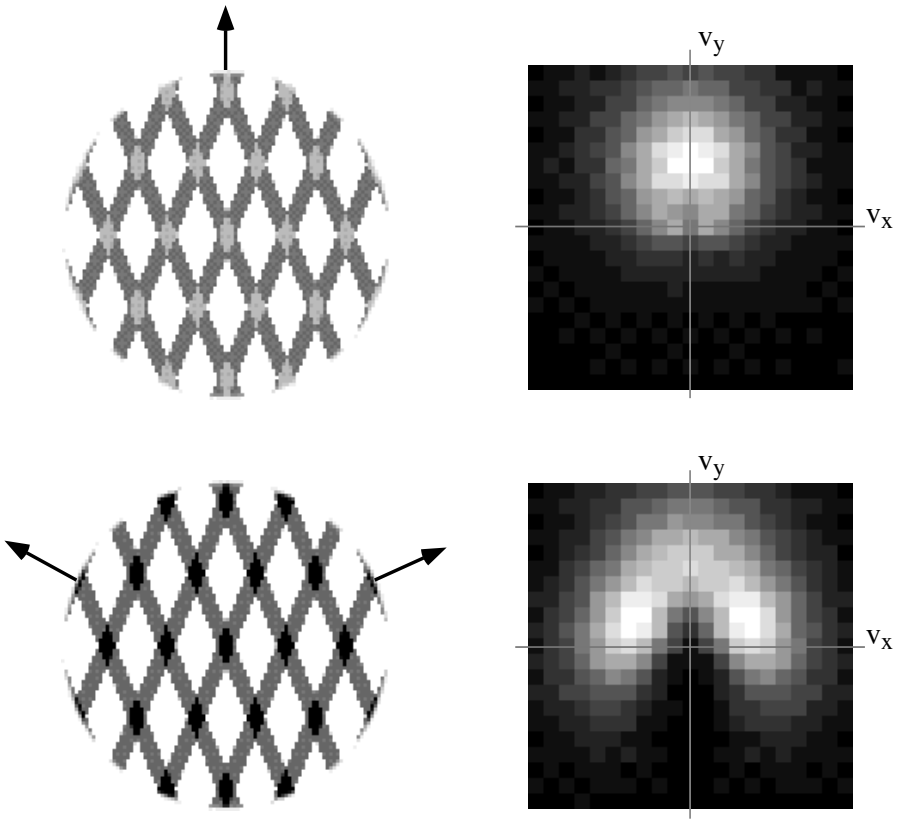
- Stimulus is two additively combined fractal noise patterns, moving in different directions.
- VE Response is bimodal, indicating presence of two motions.

VE Population Response: Sine Grating Plaids



- Steeper plaids look more transparent.
- Consistent with Adelson & Movshon (1982).

VE Population Response: Square Grating Plaids



- Transparency percept is influenced by luminance of intersections.
- Consistent with Stoner et al (1990).

Conclusions

- Simple two-stage distributed computation. For each stage:
  - Linear operators, squared and normalized.
  - Response space is minimally sampled.
  - Responses smoothly cover the space.
  - Intermediate responses may be exactly interpolated.
- Model is consistent with physiology of Complex & MT cells.
- Model is capable of representing multiple motions.
- Model is consistent with plaid transparency perception.