Motion Perception

Lavanya Viswanathan

Week 12

April 14, 1998

Topics

• Optic flow
• Chromatic input to the motion system
• Motion models
  • Correlational models
    (Riechardt; Barlow and Levick; van Santen and Sperling)
  • Gradient models
    (Marr and Ullman)
  • Energy models
    (Adelson and Bergen)

Optic Flow

Motion flow field: the collective motion of points in an image from one moment to the next.

Important cue for the relative positions and motion of objects with respect to the observer.

Motion flow fields can be used to estimate

• depth map
• object motion
• observer motion
• image segmentation (common fate principle)

Components of the Motion Flow Field

Expansion due to translation of viewpoint
Shear due to rotation of viewpoint
Chromatic Input to the Motion System

Early studies, using equiluminant stimuli, suggested that chromatic edges are invisible to the motion system.

Equiluminant stimuli contain only chromatic edges, no luminance edges.

Ramachandran and Gregory, 1978:

No motion is observed in random dot kinematograms when the dots are equiluminant with the background.

Cavanagh, Tyler and Favreau, 1984:

Motion of chromatic gratings is either seen as stop-start motion or much slower than that of luminance gratings moving at the same speed. Relative slowing occurs only at low spatial frequencies. Chromatic gratings have higher velocity thresholds than luminance gratings.

Chromatic Input to the Motion System (contd.)

More recent work claims that the earlier results were based on the use of inappropriate scales for comparing color and luminance stimuli. They equated stimulus contrast for chromatic and luminance stimuli in terms of phosphor modulation and had not equated cone responses.

Chaparro et al. (1993) used detection thresholds to show that the best detected color stimulus is seen 3-8 times better than the best detected luminance stimulus. They conclude that “color is what the eye sees best” and suggest that this mechanism may compensate for the low chromatic contrasts typically found in natural scenes.

Using appropriate contrast scales, Cavanagh (1991) find that motion processes respond strongly to chromatic stimuli. Contrast thresholds for discriminating direction of motion can be similar for both color and luminance gratings.

Skeletal model: stages (Nakayama, 1985)

- early stage: sensitive to position and spatial frequency
- second stage: directionally selective subunits (e.g., Riechardt detectors)
- third stage: spatial and temporal integrator

Motion Models

Why a spatial frequency front-end?

Why not just use the change of surface intensity over time to encode the motion of the surface?

Because, for ramp luminances, motion of the source of light would be equivalent to motion of the surface itself.

So, the motion system would be contaminated by extraneous changes in ambient illumination.
**Issues unresolved by the Skeletal Model**

- Mechanism for directional selectivity unspecified. It could be
  - addition/thresholding,
  - multiplication, or
  - inhibition ("nulling" or "veto" mechanism)
- Contrast polarity: Must sign of contrast at edges match over time in order to stimulate detectors?
- What is the output of the system?
  - direction of motion?
  - metrical speed?
  - normal component of motion (ref., the aperture problem)?

---

**Correlational Models**

**Riechardt, 1961**

Consists of a pair of receptors separated by some physical distance such that the delayed output of one receptor is multiplied by the output of the other receptor.

---

**Barlow and Levick, 1965**

Two receptors connected to an AND-NOT gate, one directly and the other through a delay.

Active inhibition of the non-preferred direction. A spot arriving at $R_1$ before $R_2$ will cause the inhibition due to $R_1$ and the excitation due to $R_2$ to arrive at the gate simultaneously, thus shutting off the output.

---

**Barlow and Levick, 1965: Biological feasibility**

Originally proposed to explain mechanisms in the directionally selective cells of the rabbit’s retina.

Recent study shows that Meynert cells in the deep layers of the mammalian visual cortex may perform the same function. (Livingstone, M. S. (1998), Neuron; Ferster, D. (1998), Nature)

Meynert cells have asymmetrical dendrites that are believed to make them sensitive to motion in one direction.

**Inputs:**
- **Basal dendrites:** excitation from a different part of the visual field
- **Cell body:** inhibition from the current position in the visual field

**Velocity sensitivity:**
- Preferred velocity = \( \frac{\text{spatial displacement of two parts of receptive field}}{\text{relative delays of excitatory and inhibitory responses}} \)
Barlow and Levick, 1965: Biological feasibility (contd.)

van Santen and Sperling, 1985: Elaborated Riechardt Detectors (ERD)

Use a modified Riechardt detector with band-limited spatial frequency channels at the front end.

This eliminates the spatial aliasing of the original Riechardt model and corrects the wrong prediction of a reversed motion percept to any continuously moving object which is dominated by spatial frequencies whose half period is smaller than the inter-detector spacing.

Correlational Models

Issues with Correlational Models

Good news
Simple!
Provide speed sensitivity

Bad news
Identical outputs for fast motion in the preferred direction and for slow or stop-start motion in the non-preferred direction

But the human visual system is also subject to the correspondence problem of apparent motion!

Need to match features/objects at two spatial locations, i.e., subject to the correspondence problem of apparent motion
Contrast-polarity specific motion
Can’t explain second order motion

Correspondence Problem of Apparent Motion

Under strobed illumination, a fast clockwise turning wagon wheel seems to be turning slowly counter-clockwise.

Implicit Assumption: the nearest spoke in the next frame is the same one as in the last frame.
Gradient Models

Marr and Ullman, 1981

Convolve image with the Laplacian of a Gaussian ($\Delta^2 G$) to find the edges in the image (zero-crossings of the Laplacian) and then use spatial frequency filtering to detect motion.

Mathematical veridical velocity, given constant illumination, is given by

$$V_x = \left( \frac{dI}{dt} \right) \left( \frac{dI}{dx} \right)$$

Marr-Ullman version:

$$V_x = \frac{\partial (\nabla^2 G \bullet I)}{\partial t} \frac{\partial (\nabla^2 G \bullet I)}{\partial x}$$

Marr and Ullman, 1981: Biological feasibility

Claim:

$X$ cells signal $\nabla^2 G \bullet I$, i.e., presence of a zero-crossing and its direction of contrast.

$Y$ cells signal the time derivative, $\frac{\partial}{\partial t} (\nabla^2 G \bullet I)$.

However, $Y$ cells get rectified inputs and are unlikely to transmit the sign of the derivative.

Issues with Gradient Models

Good news

- Provide speed sensitivity
- Solve correspondence problem: output depends only on one point in the image, i.e.,
  - it can respond instantaneously
  - it is sensitive to very small displacements

Bad news

- Can’t explain long-range effects
- Can’t explain second order motion
- Not very biologically feasible

Spatio-temporal Energy Models

Adelson and Bergen, 1985

Feature-match models:

1) locate features in two frames
2) establish correspondences between features
3) determine $\Delta x$, $\Delta t$ between frames
4) compute $\Delta x/\Delta t$

Global-match models:

- Compute a global cross-correlation between frame regions, finding the $\Delta x$ that gives the best match between frames (whose time difference, $\Delta t$, is implicitly known).

Adelson and Bergen argue for global match models.
Adelson and Bergen, 1985 (contd.)

Correspondence problem again

Moving bar: few features to match

Moving random stripes: many features to match

Adelson and Bergen, 1985 (contd.)

Motion as “orientation in space-time”

a) Moving bar
b) Real motion
c) Sampled motion

Spatio-temporal receptive fields can respond to sampled or real motion, (equally well if sampling interval is below resolution thresholds).

Adelson and Bergen, 1985 (contd.)

Filters oriented in space-time seem to require “a different temporal impulse response correctly tailored to each spatial position within the receptive field.” (p. 288)

Such filters would not have separable responses in space and time, unlike the following one, that does:

Adelson and Bergen, 1985 (contd.)

Space-time separable filters can respond to motion, but directional selectivity is compromised.
Adelson and Bergen, 1985 (contd.)

Spatio-temporal energy filters are inherently *phase-sensitive*.

(I.e., their response fluctuates with the momentary alignment of a pattern with the spatial cross-section of the receptive field that is “operative” at that instant. In other words, their response depends on the *sign* of stimulus contrast.)

A & B recommend finding a *phase-insensitive* response by computing *motion energy*: by squaring and summing outputs of “quadrature pairs” of stimuli.

---

Correlational vs. Energy Models

In a 1984 paper on “elaborated Reichardt detectors” (ERDs), van Santen and Sperling recommend that we dispense with the classical conception motion as involving a *spatial object* occupying different *locations* at different *times*, and instead

“... think of motion as involving a temporal object (luminance modulation pattern that occurs at different points in time at different locations) (vS & S, 1984, p. 451)

In 1985 they prove that energy models and ERDs are *formally equivalent*, in that they compute identical outputs for any given input!

A & B dispute this assertion, claiming that the identity only holds for a special case of energy models, and later argue that physiological evidence concerning intermediate stages favors their approach.