Motion Perception

Lavanya Viswanathan
Week 13
April 20, 1999

Topics
- Traveling wave model (Grossberg and Rudd)
- Aperture problem
- Motion grouping
  - Extrinsic vs. intrinsic terminators
  - Barber-pole illusion
  - Plaids (vector sum vs. IOC vs. feature tracking)
  - Motion integration across apertures (Shiffrar et al.)
  - Motion capture (Ramachandran)
  - Cooperativity and hysteresis
- Motion segmentation
- Motion transparency
- Velocity decomposition
- Optic Flow

Some kinds of Apparent Motion

**Beta:** Object motion across empty intervening space. A well-defined object is perceived moving continuously.

**Phi:** Sense of motion without a concurrent perception of moving object (e.g., motion in periphery).

**Gamma:** Apparent expansion at onset and contraction at offset of a single flash of light.

**Delta:** Beta or phi motion directed toward the first flash, found when the intensity of the second flash is sufficiently greater than the first.

Two-flash, Long-range Apparent Motion

Key facts:
- A *single flash* (stationary) does *not* generate a motion percept (gamma motion aside).
- Two or more properly timed and positioned flashes *do*.
- A *change of ISI* (only) can alter the global motion percept qualitatively (e.g., Ternus effect)

Issues:
- Why is the long-range influence of a single flash -- which must exist -- not perceived as motion?
- How do individual influences interact to bridge *variable distances* without a loss of sharpness in the percept?
- How does varying ISI cause a speed-up or slowdown of the smooth motion signal?
**Traveling Wave Model (Grossberg and Rudd, 1989)**

2. Forget them.
3. Remember the “level-less” long-range Gaussian filter!

---

**SINGLE FLASH: Spatial Patterns**

If a single peak is produced at Level 4, the Gaussian filter produces a Gaussian profile, which is then sharpened back to a single peak by a “winner-take-all” competition at Level 5.

Note 1: The width of the Gaussian varies, depending on the (size of) the “scale” considered.

Note 2: A Gaussian of any scale changes amplitude over time.

---

**SINGLE FLASH: Temporal Dynamics**

A single flash creates a characteristic “rapid increase toward saturation level” while ON and an equally characteristic exponential temporal decay soon after it shuts OFF.

This temporal profile modulates the amplitude of the Gaussian signal.

Since the peak of the Gaussian stays in place regardless of amplitude, the model’s “percept” is that nothing moves.

---

**TWO FLASHES: Temporal Dynamics**

If two flashes occur in rapid succession, the “waning” of the first signal may overlap in time with the “waxing” of the second.

Note: To produce apparent motion, it is not necessary that the temporal profiles of the stimulus “abut” (i.e. zero ISI). ISI could be positive, or the flashes themselves could overlap in time somewhat.
Long-range Apparent Motion

If the Gaussian activity profiles of two flashes overlap sufficiently in space and time, the sum of Gaussians produced by the waning of the first-flash Gaussian, combined with the waxing of the second-flash Gaussian, can produce a single-peaked traveling wave:

The resulting motion percept is thus both long-range and sharp.

“Just in time delivery” of motion signals

For a given ISI: How does perceived velocity increase with distance between flashes?

For Gaussian filter: \[ G_{ji} = H \exp \left( \frac{(j-i)^2}{2K} \right) \], the largest separation for which sufficient spatial overlap between two Gaussians centered at locations \( i \) and \( j \) will exist to support a traveling wave of summed peak activity is: \( L_{\text{crit}} = 2K \)

Interactions between flash separation and filter scale

Multiple spatial scales!

To upper left of dashed lines: NO MOTION

Why no motion?

The waxing and waning Gaussian profiles never produce a single peak of summed activation.

To lower right of dashed lines: MOTION

Equal half-time property

Not only does the traveling wave for a given Gaussian scale bridge variable distances in (almost) equal times, but also:

Traveling waves from Gaussian filters of a variety of sizes bridge the same distance in comparable times; in fact, the time needed to bridge half the distance between flashes is precisely the same:
Ternus effect

Simulate the Ternus effect, using Gaussian filters, combined with gating interactions among sustained and transient model cells.

Variation of a single parameter, ISI, yields a qualitative change in percept, from group to element motion.

Adaptation of motion detectors

Adaptation to 30 motion

Repulsion: 0 motion seen as --10 motion

The ability to “selectively” adapt for a specific stimulus feature (i.e. direction of motion) is generally considered indicative of the presence of mechanisms tuned to that specific feature.

Motion Aftereffect (MAE)

Waterfall Illusion:

“Having steadfastly looked for a few seconds at a particular part of the cascade, ................. I saw the rocky surface as if in motion upwards, and with an apparent velocity equal to that of the descending water”

Addams (1834)
Falls of Foyers

Short-range vs. Long-range Motion Processes

Table 1 from Cavanagh & Mather (1989) -- See Syllabus
(principally after Anstis (1980) and Braddick (1980), with some additions).

<table>
<thead>
<tr>
<th>Short-range</th>
<th>Long-range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short spatial range (&lt; 15 arc min)</td>
<td>Operates over many degrees</td>
</tr>
<tr>
<td>Brief temporal range (80-100 ms ISI)</td>
<td>ISI up to 500 ms</td>
</tr>
<tr>
<td>Motion aftereffect</td>
<td>No motion aftereffect</td>
</tr>
<tr>
<td>Not dichoptic</td>
<td>Dichoptic</td>
</tr>
<tr>
<td>Braddick, 1974</td>
<td>Shipley et al. (1945)</td>
</tr>
<tr>
<td>No response to colour</td>
<td>Response to colour</td>
</tr>
<tr>
<td>Low-level neural comparator</td>
<td>Responsive to higher-level correspondences</td>
</tr>
<tr>
<td>Passive motion response, velocity space computations</td>
<td>do not activate motion detectors</td>
</tr>
<tr>
<td>Adelson and Movshon, 1982</td>
<td>Cooperative processes, inference</td>
</tr>
</tbody>
</table>
First-order vs. Second-order Processes
(Cavanagh and Mather, 1989)

<table>
<thead>
<tr>
<th>First-order stimuli</th>
<th>Second-order stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td>characterize responses to visual stimuli at the earliest level (retinal ganglion cells)</td>
<td>frequency with which specific combinations of intensity or color values occur for pairs of points same mean luminance and color</td>
</tr>
<tr>
<td>luminance</td>
<td>texture</td>
</tr>
<tr>
<td>color</td>
<td>spatial separation of pairs in the same image</td>
</tr>
<tr>
<td></td>
<td>motion</td>
</tr>
<tr>
<td></td>
<td>pairs in different time frames of motion sequence</td>
</tr>
<tr>
<td></td>
<td>binocular disparity (or depth)</td>
</tr>
<tr>
<td></td>
<td>pairs belong to a stereo-pair of images</td>
</tr>
</tbody>
</table>

Aperture Problem
First described by Wallach (1935):
When viewed through a circular aperture, a line is seen to move in the direction perpendicular to its orientation, regardless of its real direction of motion.
This occurs because any motion parallel to the line causes no change in the stimulus.

Extrinsic vs. Intrinsic terminators

Not all line ends are created equal!

**Intrinsic**: belongs to the line (or object)
**Extrinsic**: does not belong to the line (or object). The object is seen to extend behind some occluder.

Extrinsic terminators are useful cues for occlusion within a scene. Motion of such terminators tells us NOTHING about the direction of motion of the occluded object, though it may be used to determine the form of the occluding object (ref., Michotte’s rabbit-hole effect).

Motion Grouping
Assume that you’ve (somehow!) managed to figure out how to do early motion detection.
Now what?
Consider the leopard jumping at you through a rain forest canopy.
Motion Grouping

How do you . . . survive?!

Locally ambiguous, contradictory, disconnected signals must be transformed into a coherent motion percept.

First pass: Getting agreement among motion directions
Second pass: Trade-off of possible directions and possible speeds
Big picture: Using motion signals for figure/ground segmentation

1. How do motion signals that code for different (or ambiguous) direction-of-motion (in 2D) “choose” a coherent direction?
   
   Barber-pole illusion
   Plaids

2. How (and under what circumstances) do motion signals that are far apart get combined (pooled)?
   
   Motion integration across apertures
   Motion capture

Barber-pole Illusion (Wallach, 1935)

Viewed through a rectangular horizontal aperture, a diagonal line appears to move horizontally (i.e. in the direction of greatest elongation of the aperture) most of the time, at least while both of its visible endpoints are in contact with the horizontal boundaries of the aperture.
Barber-pole Illusion (contd.)

What happens when a line of the same orientation is just beginning to become visible through the aperture?

A “cognitive” theory might predict that we still perceive horizontal motion . . .

Barber-pole Illusion: Role of Terminators

Line terminators play a crucial role in determining perceived line motion.

Note: Motion direction signals from points (such as line ends) are unambiguous, unlike signals from line interiors, for which the aperture problem restricts information to the normal component of velocity (i.e., true direction is ambiguous).

Barber-pole Illusion: Role of Terminators (contd.)

Line terminators, which have unambiguous motion signals, are used to disambiguate motion signals from the interiors of lines.

CLAIM: To be useful, unambiguous information from line ends has to be “propagated” to line interiors, to help select consistent direction signals and suppress others.

Note: This claim is analogous to those concerning the need for “filling-in” of brightness signals.

One alternative to “filling-in” or “propagation” of motion signals from line-endings to the interior: one could concede that line ends get a stronger, even determining “vote” in some labeling scheme, without granting that motion direction signals must actually propagate across a retinotopic map.

Barber-pole Illusion: Extrinsic and Intrinsic Terminators

The lines in the barber-pole illusion have intrinsic terminators (i.e., the terminators belong to the line).

When these terminators are made extrinsic, say by presenting the tilted lines at a further depth than the occluding rectangle (e.g. through stereoscopic viewing), the illusion is abolished.

Line then appears to move in a direction perpendicular to its orientation, as predicted by the aperture problem at line interiors.
Revisit the Aperture Problem

All single neurons are subject to the aperture problem for edges that are longer than the extent of their receptive fields.

The detector picks up only the normal component of velocity.

Constraint Line

The locus in velocity space of possible positions of the leading edge of the bar (or line) after some time interval ($\Delta t$) is called the constraint line.

Plaids

The superposition of two drifting sinusoidal gratings can result in the percept of coherent (or rigid) plaid motion.

Coherent plaid motion is seen if contrasts, velocities, and spatial frequencies are similar. Otherwise, transparent overlapping motions can be seen, i.e., the component gratings are seen to glide over each other.

Question: When the gratings do cohere into a rigid plaid, how do we compute the velocity of the resultant plaid?

One solution: Vector Sum or Average

The resultant velocity could be the sum or average of the perpendicular velocity components of the individual gratings.

Not always true!
Second solution: Intersection of Constraints (IOC)

Or, the resultant velocity of the coherent plaid could be determined by the intersection of constraint lines for the component gratings. This is called the IOC solution (Adelson and Movshon, 1982).

So, for the gratings given on the last page, the IOC solution would be

\[ v_x \quad v_y \]

Third solution: Feature Tracking

We could also be computing the unambiguous velocity of features in the display from one frame to the next. These unambiguous signals would then veto the ambiguous signals from the interiors of objects.

Some unambiguous features are

- Line endings (as in the barber-pole illusion)
- Line intersections (as in plaid motion)
- Object corners (as in the displays of Shiffrar et al.)

IOC vs. Feature Tracking

At a real corner, the IOC computation and the trajectory of the corner feature are always identical!

**Question:** Does IOC always give the same solution as feature tracking?

**Answer:** NO! (more later)

Motion Integration across Apertures

Consider the following displays (Mingolla, Todd and Norman, 1992):

Line intersections absent  Line intersections present
Motion Integration: (vector average vs IOC vs feature-tracking)

In both cases, observers’ perception of motion was dominated by the motions of the component contours. Observers nearly always perceived motion in the direction of the vector average solution and away from the IOC or the feature-tracking solution (i.e., the veridical solution).

Vector averaging is an efficient technique for noise suppression, especially when the feature points themselves provide ambiguous information (as in the case of features formed by occlusion).

The visual system probably uses a compromise between vector-average of the component motions and feature-tracking.

Motion Integration: (IOC vs Feature-tracking)

IOC predicts human perception only when coherent motion is perceived (and not always even then, as seen above)!

Feature-tracking can explain human perception even when globally coherent motion is not perceived, as in the displays of Shiffrar and Lorenceau, 1996.

The saliency of the features in the display could help explain whether locally ambiguous motions will be integrated into a global percept.

Motion Capture (Ramachandran, 1981)

**Stimulus**: A square is displaced, while randomly distributed dots are replaced by uncorrelated dots.

**Percept**: The frame and dots move coherently.

Motion Capture with Barber-pole Gratings

(Ramachandran and Inada, 1985)

**Stimulus**

random dots superimposed on the grating

**Percept**

motion of the dots is captured by grating motion
**Spotted Barber-pole**
(Shiffrar, Li and Lorenceau, 1995)

**Stimulus**
- random dots superimposed on the grating
- all dots move coherently down

**Percept**
- motion of the grating is captured by unambiguous motion signals of the dots

**Cooperativity and Hysteresis**

Imagine a display of randomly positioned dots.

Define a motion direction in the plane, $\beta$ and some tolerance range about that direction, $\partial$. That is, for any two frames of a multiframe sequence, each dot displaces by a fixed step size in a direction in the range $(\beta - \partial, \beta + \partial)$. If $\partial$ is “small enough,” coherent motion will be perceived in direction $\beta$, (along with “jittery motion” of individual dots).

**Stimulus**

**Percept**

**Cooperativity and Hysteresis (contd.)**

If, while a continuous “movie” of such dot motions is displayed, the parameter $\partial$ is varied continuously in time, a hysteresis effect is observed for the transition from coherent to incoherent motion:

Cartoon of data from Williams, Phillips, and Sekuler (1986).

**Motion Segmentation**

This is the problem of extracting form structure from motion cues.

Two types of algorithms:

**Type A**
- Motion is a property of material points in space, not the positions those points occupy at a given time
- Structure from motion is achieved by taking multiple measures of a given material point at different moments in time. This requires the ability to reliably track the identity of such points. (correspondence problem)

**Type B**
- Motion is a property of positions in space, not the material points that occupy those positions at a given time
- Structure from motion is achieved by taking multiple measures at different positions within a local spatiotemporal neighborhood. This requires a smoothly varying flow field.
Motion Transparency

**Stimulus:** A field of random dots with two subsets moving in different directions.

**Percept:** The field is decomposed into two superimposed transparent surfaces seen in depth and sliding across each other.

*Note:* In the actual display, all dots have the same color and can be differentiated only on the basis of motion cues.

Velocity Decomposition

A physically unidirectional motion of random dots is perceived to be decomposed into two component motion directions. Observers see two transparent square regions overlapping with the dots in each square moving in a distinct direction.

Velocity and surface decomposition is strongest when velocity of overlapping region is equal to the *vector sum* of the velocities in the adjacent regions.

*Note:* The square contours shown in this figure are not present in the actual stimulus.

Induced Motion: Frames of Reference

Duncker (1929):

**Stimulus**

**Percept**

The difficulty of the perceptual "frame of reference" problem can scarcely be exaggerated, and its importance to motor control (e.g. posture, grasping, etc.) is paramount.

What moves with respect to what?

Use: Retinal coordinates? Body coordinates? World coordinates?

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

- direction of heading
- 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.