Lecture Notes - 07/13/99
Alex Huk

# Seeing *Motion*: Lecture Notes

## 0. Introduction

Motion perception can be broken into 2 main categories:

- perceiving motion of objects in the world ("Object motion").
- perceiving motion of ourselves through the world ("Observer motion").

The organization of this lecture reflects this split; note that the first half does not correspond to an assigned reading (but see Chapter 8 of the course text for an advanced treatment, if you'd like), the second half will review a selection of the material from Chapter 13.

## 1. What is motion?

- Motion is often defined as a change in position over time. From high school, you may remember the equation speed = distance/time.
- Motion can be broken into two main components: direction and speed.
- However, for studying visual motion, there are several other distinctions that are helpful. Consider the variety of types of motion information that you are capable of using in everyday life:
    1. **Simple translation**: watching a ball thrown across your field of view, you can easily perceive that the ball is moving relative to the world.
    2. **Complex motions**: overlooking a crowded street, with cars, pedestrians, and cyclists all moving in different directions at once, you can detect the individual directions and speeds of each person or vehicle, but can also detect the general "flow" of traffic.
    3. **Apparent motion**: looking at a neon street sign, where a series of lights flash one after another, it appears to move.
    4. **Stroboscopic** motion: you see movement on TV or at the movies, although you're actually watching a series of static images.
    5. **Motion aftereffects**: after watching a waterfall for a few minutes, you look at the nearby rocks, and they appear to move upward, even though they're obviously not actually changing their position over time.
    6. **Structure from motion**: a well-camouflaged animal is impossible to see as long as it doesn't move. When it does, you can easily identify it.
    7. **Optic flow**: as you walk down the street, objects in the world change their position in your field of view, and you are easily able to use this information to navigate the world.

8. **Induced movement**: sitting on a stationary train and looking out the window, the train on the rail next to yours begins to move. You at first feel like *your* train is moving, even though it is not.
9. **Eye-movements**: your eyes are constantly moving, making small shaking motions or long, smooth motions. In either case, the image falling on your retina changes, but the world seems stationary.

---

# 2. Object motion

- Let's call examples 1-6 "Object Motion", since they all involve objects changing their positions over time.
- Although objects are really moving out in the world, your visual system *infers* motion. Often this inference is correct, but sometimes it is not. Illusions of movement are therefore often used to uncover the basic mechanisms that we use to perceive real motion:
    - **Motion (movement) aftereffect** ("MAE"). After viewing the attached movie of continuous motion in the same direction, stationary objects appear to move in the *opposite* direction. Think about how the adaptation of direction-selective neurons in V1 might contribute to this illusory percept of movement.
    - **Movement-without-motion**. (film viewed in class). In these demonstrations, the stimulus also appears moving but never actually changes position. This stimulus separates the contributions of "local" and "global" motion-detection mechanisms: the local (or short-range) detectors register motion (eg, each contour seems to be moving), but the more global (long-range) detectors know that the larger stimulus is not actually changing position. This suggests a hierarchy of motion-detectors in the visual system.
- In the case of **apparent motion**, a stimulus is flashed briefly in one location and then another (usually identical) stimulus is flashed in another (nearby) location. Under certain circumstances, observers perceive the first stimulus moving to the position of the second (as if there were only 1 stimulus, actually moving from the first position to the second). This is the well-known principle underlying neon marquee lights that you see at old theatres and casinos. The perception of apparent motion depends crucially on the interval of time between the flashing of the two stimulus (the "ISI": inter-stimulus interval):
    - If the ISI is < 30 msec (.03 sec), the two stimuli appear to be simultaneous.
    - If the ISI is between 30 and 60 msec, the first stimulus will appear to move partially toward the second stimulus.
    - If the ISI is > 60 msec, the first stimulus appears to move smoothly and continuous to the location of the second stimulus (apparent motion is achieved).
    - When the ISI is longer (> 200 msec), you (veridically) perceive the first stimulus being flashed, followed by the second stimulus being flashed at a different location.
- **Stroboscopic motion**: Apparent motion is 1 type of stroboscopic motion. A more familiar instance of stroboscopic motion occurs in movies and on television. As you know, film is made of a series of individual, static frames (snapshots), presented rapidly. Early movies showed 16 frames per second; although fast enough for stroboscopic motion, people were still able to detect the flicker between frames (hence the name

"flicks"). To solve this problem, film engineers had to play upon not just apparent motion but also "visual persistence". To see this phenomenon, wave a pencil quickly in front of your face. You'll see the pencil blur, leaving a brief streak behind it at fast speeds. The early frame rate of movies was too slow to utilize this persistence to build a more coherent, smooth percept, but a change to 24 frames per second was adequate, and the "flicking" went away. Current films still show 24 frames per second, but now flash each frame 3 times, artificially increasing the flicker rate to 72 flashes/sec, which puts modern film well above our ability to detect flicker.

# 3. Simple motion detectors in V1

How does our visual system infer motion from the variety of cases described above?

- We already know from the experiments of Hubel and Weisel that there are directionally-selective neurons in V1.
- Building a **directionally-selective** receptor: when an object moves, its representation in the retinal image moves, and therefore stimulates a consecutive series of photoreceptors. See the simple circuit proposed in the figure based on simple inhibitory connections. See the responses to 2 different directions of motion ((a)-(b) for rightward motion, (c)-(d) for leftward).
- **Direction tuning** curves: direction-selective complex cells fire maximally when a bar of light is moved through their receptive field in a certain ("preferred") direction, and fire less and less as the direction is changed more and more from the preferred. [see tuning curves].
- Note that the direction tuning curves measured for direction-selective neurons in V1 are measured by passing a bar through the receptive field. The direction of movement of the bar is assumed to be perpendicular to the orientation of the bar. However, if the bar passes through the entire receptive field, the motion of the bar could actually be in a number of directions. This is known as the *aperture problem*. The aperture problem reminds us that the direction and speed of motion of a bar (or a "grating" pattern of bars) is ambiguous when edges or texture are not present. Also note that we, as people, rarely see motion through an aperture that is ambiguous-- *but* the receptive field of each motion-responsive neuron is essentially an aperature, so the aperture problem is something that our motion pathway must always solve.
- When two grating patterns with different orientations are superimposed, the resulting **plaid** pattern(usually) appears to move as a whole. What is the perceived direction?
- When a plaid grating pattern is passed through the receptive field of direction-selective V1 neurons, the neurons respond to the *individual components*: i.e., if one of the two superimposed gratings is moving in the preferred direction, the cell will fire maximally. These neurons are therefore often called *component cells*. However, our percept is of a coherent plaid, drifting in a direction that is not the same as the direction of motion of either of the component gratings.

# 4. Motion detectors in area MT

- To see the neural mechanisms which underlie our actual percept of plaid-pattern motion, we need to leave V1 and head to extrastriate area MT. MT can be defined according to the F-A-C-T criteria we have discussed earlier:
  - **F**unction. MT is notoriously responsive to motion. Compare the fMRI responses of various brain areas to motion (left) and flicker (right). While most areas respond to both, MT responds much more strongly to motion vs stationary stimuli than flicker vs mean-gray stimuli. This reflects the fact that approximately 90% of MT neurons are direction-sensitive.
  - **A**rchitecture. Staining for the presence of cytochrome oxidase, an enzyme used in metabolism, is often used to help identify different visual areas. See the distinctive patch (quite dark and dense) that is evident in MT. Perhaps more interestingly, MT has a columnar architecture, no longer based on eye-of-origin (ocular dominance columns) or orientation preference (orientation columns/"pinwheels"), but instead is systematically organized with regard to the preferrred direction of motion of neurons. [see the direction-columns in MT]
  - **C**onnections. The direction-selective neurons in V1 project (connect) to MT neurons.
  - **T**opography. Indeed, MT has its own retinotopic map of the world (first demonstrated in humans by R Khan and R Dougherty, a graduate student and a postdoc here in the Psych department at Stanford). However, the receptive fields of MT neurons are much larger than V1 neurons (i.e., MT-neuron receptive fields cover or "see" a patch of visual space 100 times the size of V1 neuron receptive fields), suggesting that they are able to integrate more global motion signals.
- When a plaid pattern is passed through the receptive field of most MT neurons, the cell responds to the direction of the plaid, e.g., the perceived direction of motion, and not to the directions of motion of the component gratings. These cells are called *pattern cells* [see the different responses of component and pattern cells to a plaid pattern]
- Building a **pattern cell**: a pattern cell can receive inputs from a series of component cells, each of which has a different preferred direction, but all of which have receptive fields covering the same part of visual space.
- The responses of MT neurons have additional correspondences to our conscious percepts of motion. Bill Newsome and colleagues (here at Stanford) trained monkeys to perform a difficult motion discrimination task, judging the direction of motion of noisy, not-very-coherent moving dot stimuli. Monkeys view the dots and decide which direction the dots moved. When the experimenters would electrically stimulate an MT column selective for upward motion, the monkey would be more likely to respond that the stimulus moved upward.[see the Newsome experimental setup, stimulus, and results].
- **Human MT and the motion aftereffect**: The time course of activity in human MT seems to match the time course of the perceptual MAE (as measured by Roger Tootell at Harvard/MGH). [see the fMRI data].

# 5. Observer motion

- Information derived from vision is a major guide to our actions; the perception of motion is especially central to our own movements through the world. (see examples 7 and 8 at top).
- **Optic flow and Heading perception**: an idea originally described by JJ Gibson, the *optic flow field* is a representation of the direction and speed of motion of the visual field, relative to the observer. [see an optic flow field for a pilot landing a plane]. Note that near points move fast (large arrows) while faraway points move slower (small arrows). All of the arrows point away from a central spot, the *focus of expansion*, which corresponds to where the observer is headed.
- **J.J. Gibson** hypothesized that the visual system estimates the optic flow from the chaning pattern of light on the retina, and uses it to estimate the 3D motions of observers and objects. Note that as observer motion becomes more complex (e.g., eye and head movements while walking) and objects in the scene move relative to one another (e.g., a bird flies across the pilot's field of view), the optic flow field becomes increasing complex. No matter how complex, the optic flow field provides information that underlies many aspects of our ability to navigate the physical world.
- **Avoiding collisions**: We are quite good at detecting objects that are on a path that will make contact with us, even as early as 8 days old. Such "looming" objects become increasingly large on the retina as they approach us, and we appear to make reflexive defensive responses ("flinching") that are based on a (wisely-conservative) estimate of "tau", the "time to impact" of the object.
- **Maintaining balance**: We use optic flow information, in addition to proprioceptive (= "where we are in space") information from joints and muscles, to compensate for slight misbalances. In a "swinging room" apparatus, the optic flow can be artificially swayed, and people make unconscious compensatory leanings in the opposite direction. Children are especially sensitive to this manipulation. Adults can confirm the importance of optic flow information to balance by closing their eyes while standing on one leg.
- **Induced motion**: Although the train example at top (motion example #8) often produces an illusory feeling of motion, this usually only happens when the other (actually-moving) train is close. Why doesn't induced motion occur as often in cars?
- See the text for more research applying these concepts to athletics and driving.

---

# 6. Eye movements

- Our eyes are constantly making small, fast movements, although your perception of the world isn't nearly so "jumpy". However, there are ways to make your eye movements more obvious; see this demo using afterimages.
- In fact, as we already learned, we know that images stabilized on the retina fade. The quick eye-movements that you are constantly making keep "refreshing" the visual system, so that it keeps detecting change, and nothing fades.
- There are 2 main types of larger eye-movements: *saccades* and *smooth pursuit*. Saccades are fast, point-A to point-B, motions. Smooth pursuits, meanwhile, follow moving objects

gradually. Interestingly, smooth pursuit can only be performed when there's something to pursue; try moving your eyes gradually and evenly from left to right. You'll notice that they jump, and are difficult to control. Now, move your finger across your field of view from left to right, and follow with your eyes. You should have no trouble, now.

- **Corollary discharge**: Despite the nearly-constant motions of our eyes (as well as our heads), we perceive a steady world. This is likely due to an integration of visual information with information about eye movements/position and body movements. Usually, this compensatory process works. However, it can be misled, especially when you don't control your eye movements in the traditional way. Try *gently*pushing on the side of one eye (keeping the other closed): the world will appear to jiggle around, because there isn't an eye-muscle (motor) signal available to compensate for the motion of the eye and subsequent motion of the retinal image.

# 7. Motion processing after MT: areas MST and STS

- Area MT sends projections to area MST, which appears to process more complex motion patterns. It includes cells that respond to expansion/contraction (optic flow?), and other combinations of rotation and translation. Additionally, there is evidence that MST is the "comparator" of corollary discharge, as it receives inputs from vestibular and movement areas of the brain. [see sample MST responses to expansion and rotation].
- Area STS, meanwhile, has neurons that are responsive to *biological motion*: see demo in class of "point-light walkers". This may underlie our surprisingly-acute abilities to recognize human-like figures from very impoverished information (remember the Peter Gabriel video for "Sledgehammer"? You can identify the dance steps merely from a few light bulbs stuck on the dancers).

# 8. Where to now?

Next lecture, we'll examine another way we perceive the "where" in the visual world: seeing the shape and location of objects in three dimensions.