COMP 546

Lecture 9

egomotion:

translation & rotation, eye movements

Thurs. Feb. 8, 2018

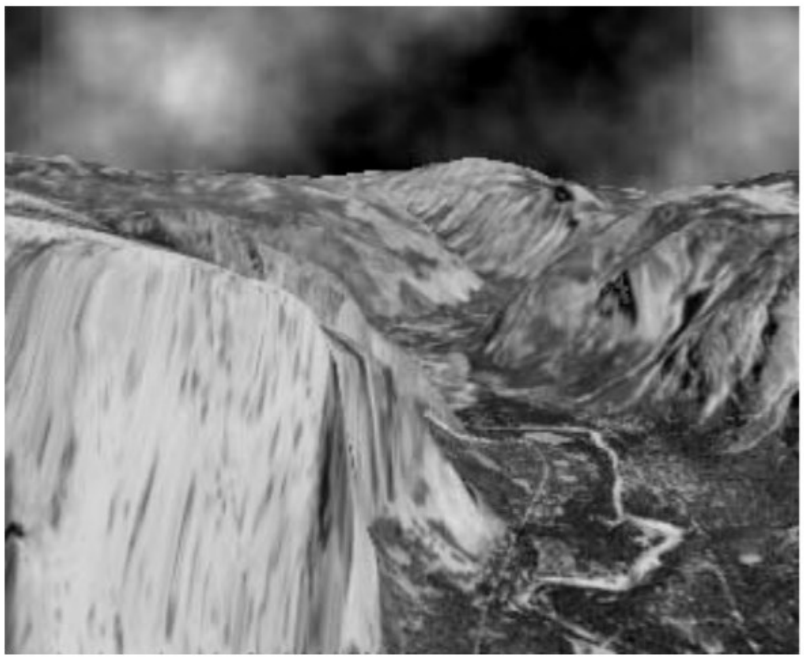
What is the image motion seen by a moving observer? ("egomotion")

Translation

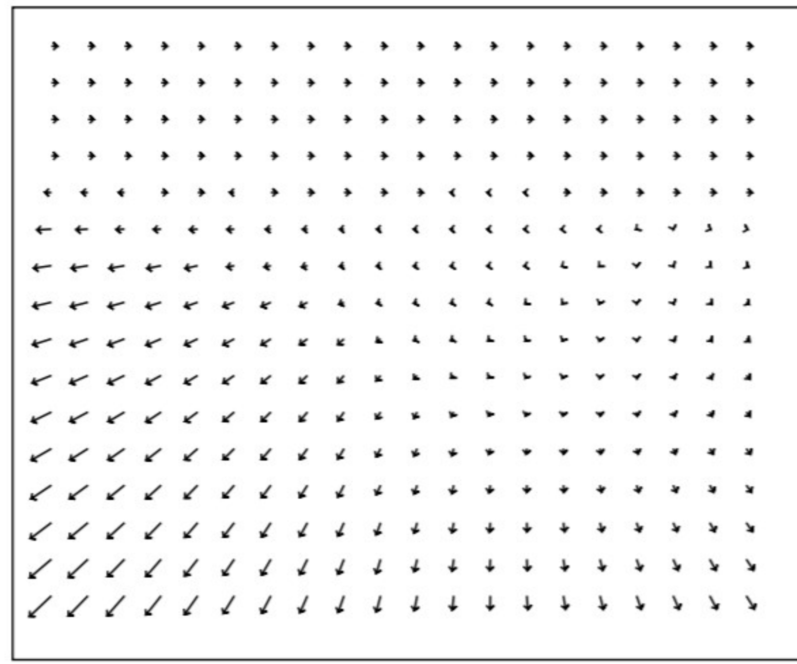
Rotation

Motion field seen by moving observer

For each image location (x, y), there is a velocity (v_x, v_y).

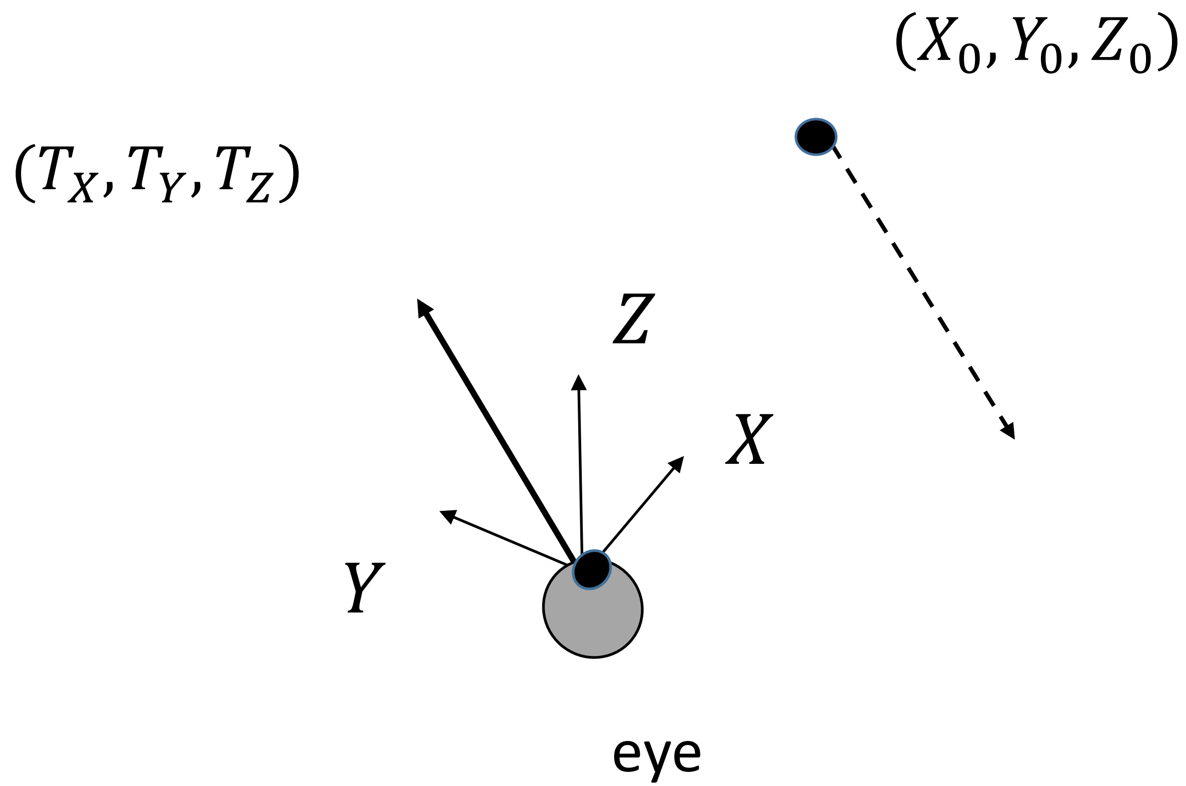


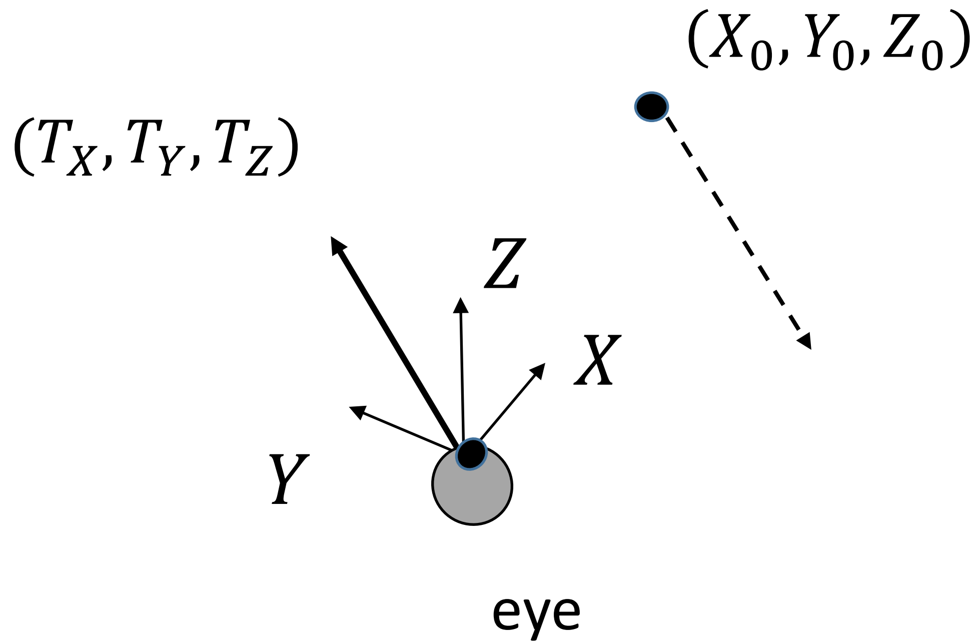
The Yosemite sequence



(Last lecture, I discussed how to estimate image velocities.)

Motion field seen by a *translating* observer



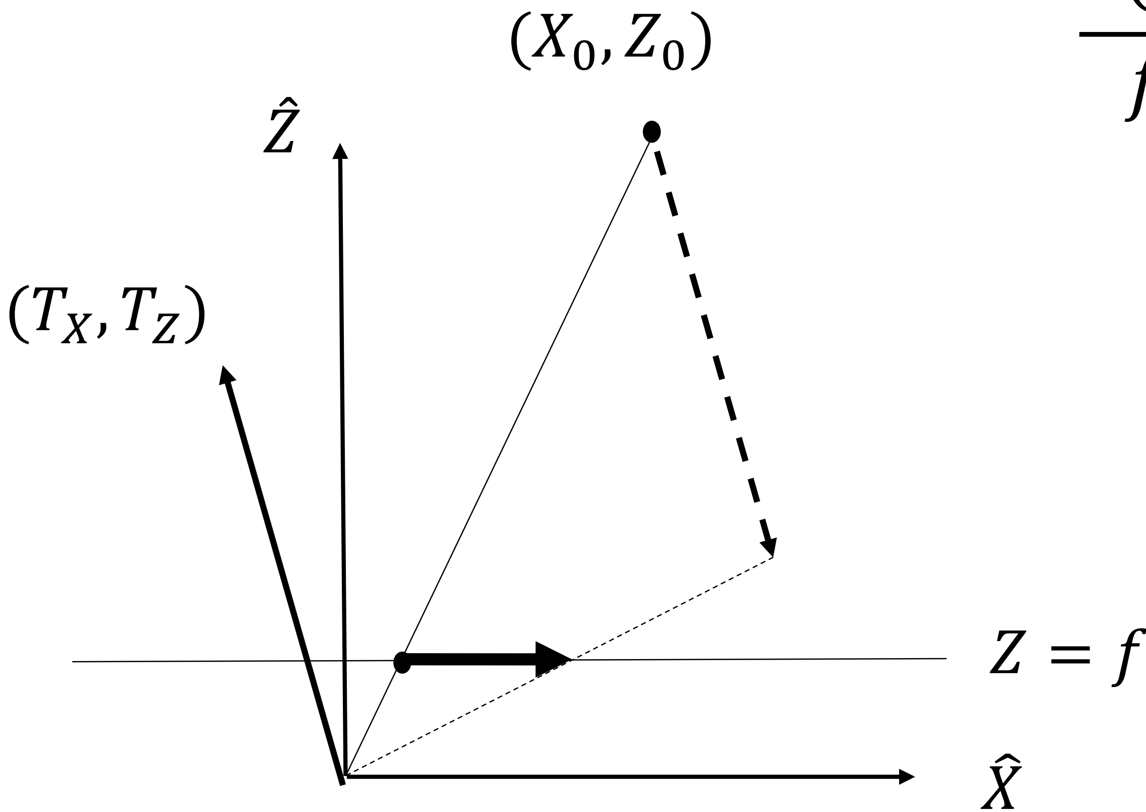


The path of the scene point in the eye's coordinate system is:

$$(X(t), Y(t), Z(t)) = \left(X_0 - T_x\, t, Y_0 - T_y\, t, Z_0 - T_z\, t\right)$$

The *relative* 3D velocity of the scene point $(-T_X, -T_Y, -T_Z)$

What is the **image path** of the scene point?

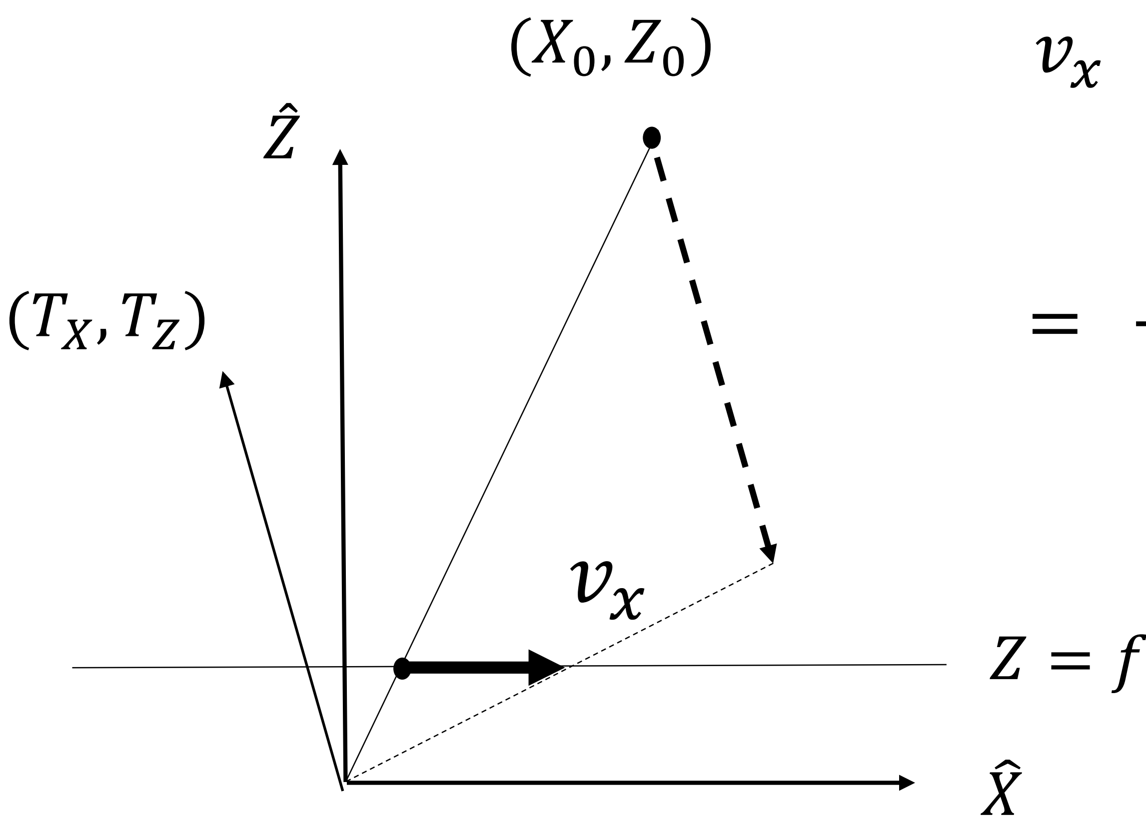


$$\frac{x(t)}{f} = \frac{X(t)}{Z(t)}$$

$$= \frac{X_o - T_X\, t}{Z_o - T_Z\, t}$$

Notation: here $x(t)$ is a position in the plane $Z = f$.

What is the **image velocity** of the scene point?



$$v_x = \frac{d}{dt}\frac{x(t)}{f} \Big|_{t=0}$$

$$= \frac{-T_X Z_0 + T_Z X_0}{{Z_o}^2}$$

Notation: here $v_x(t)$ is an angular velocity (radians/sec, assuming small angle approximation) rather than a velocity in the plane $Z = f$.

$$(v_x, v_y) = \left(\frac{-T_X Z_0 + T_Z X_0}{{Z_o}^2}, \frac{-T_Y Z_0 + T_Z Y_0}{{Z_o}^2} \right)$$

$$= \frac{1}{Z_o}(-T_X, -T_Y) \quad + \quad \frac{T_Z}{Z_o} \underbrace{\left(\frac{X_o}{Z_o}, \frac{Y_0}{Z_o} \right)}_{\left(\frac{x}{f}, \frac{y}{f} \right)}$$

Lateral component

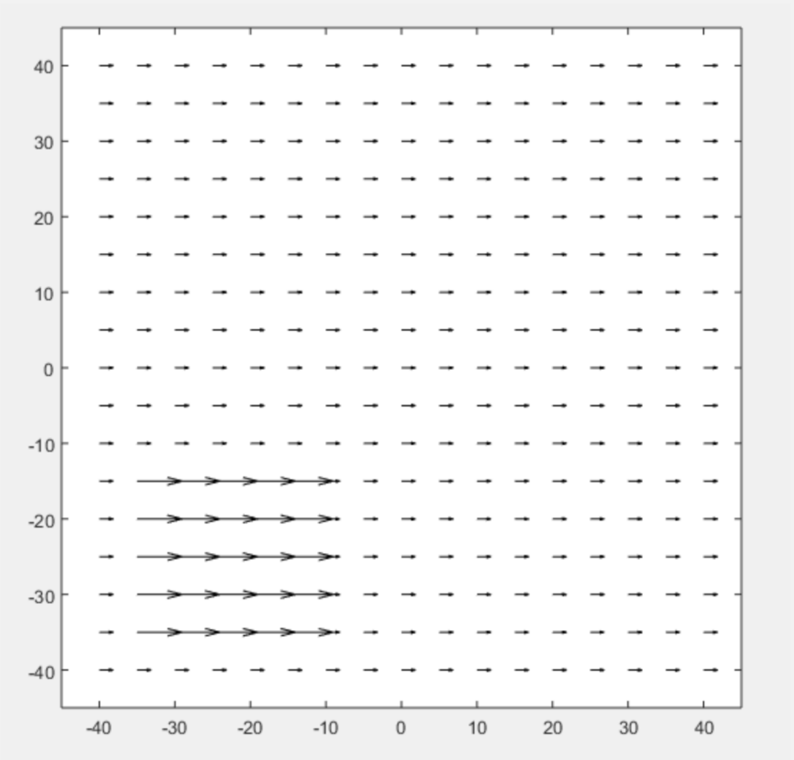
Forward component

Lateral Component $(T_Z = 0)$

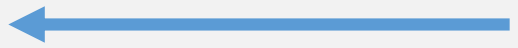
$$(v_x, v_y) = \frac{1}{Z_o}(-T_X, -T_Y)$$

Example:

wall $(Z = 10)$,
square $(Z = 3)$



eccentricity (deg.)



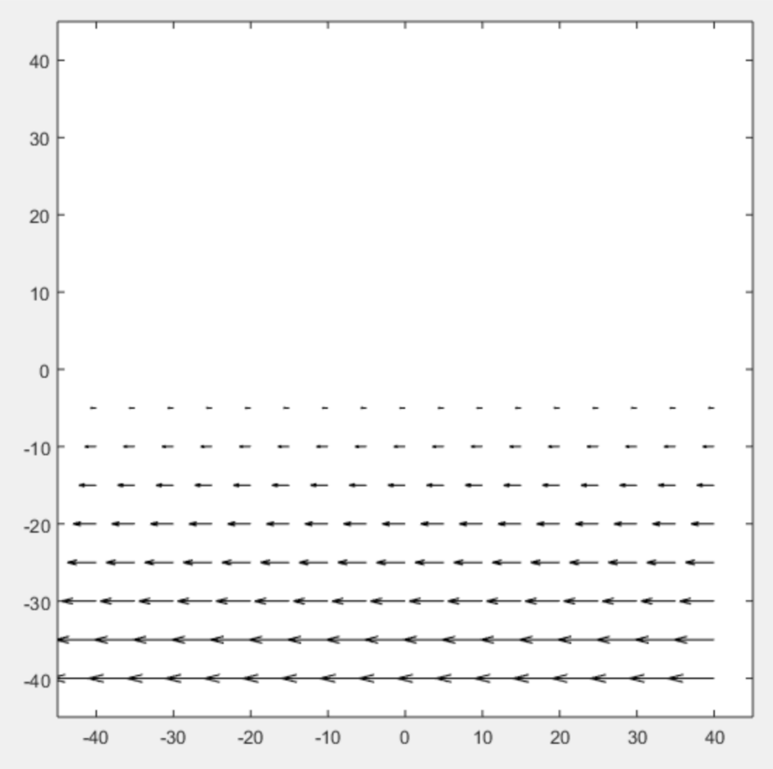
Lateral Component $(T_Z = 0)$

$$\left(v_x, v_y\right) = \frac{1}{Z_o}\left(-T_X, -T_Y\right)$$

Example:

ground plane

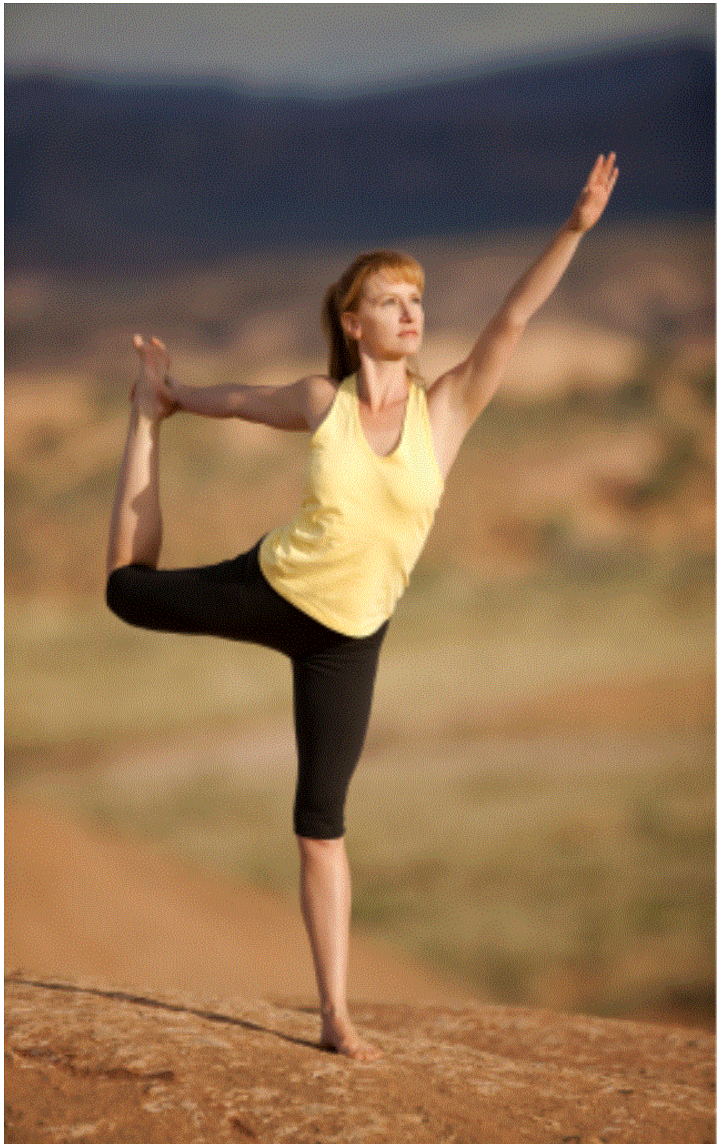
$$Z = -\frac{fh}{y}$$



eccentricity (deg.)

(T_X, T_Y)

Lateral Motion and Balance



Holding this pose is more difficult when looking up than when looking down.

Why?

Dizziness (‘height vertigo’)

Not to be confused with a more general ‘acrophobia’ (fear of heights)

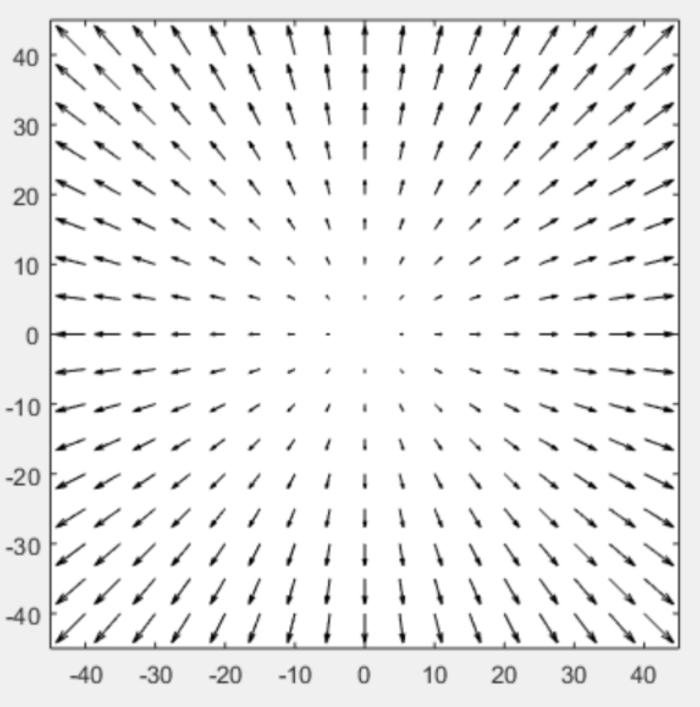
Forward Component $(T_X = T_Y = 0)$

$$(v_x, v_y) = \frac{T_Z}{Z_o}\left(\frac{x}{f}, \frac{y}{f}\right)$$

What does image flow depend on ?

Forward Component $(T_X = T_Y = 0)$

$$(v_x, v_y) = \frac{T_Z}{Z_o}\left(\frac{x}{f}, \frac{y}{f}\right)$$

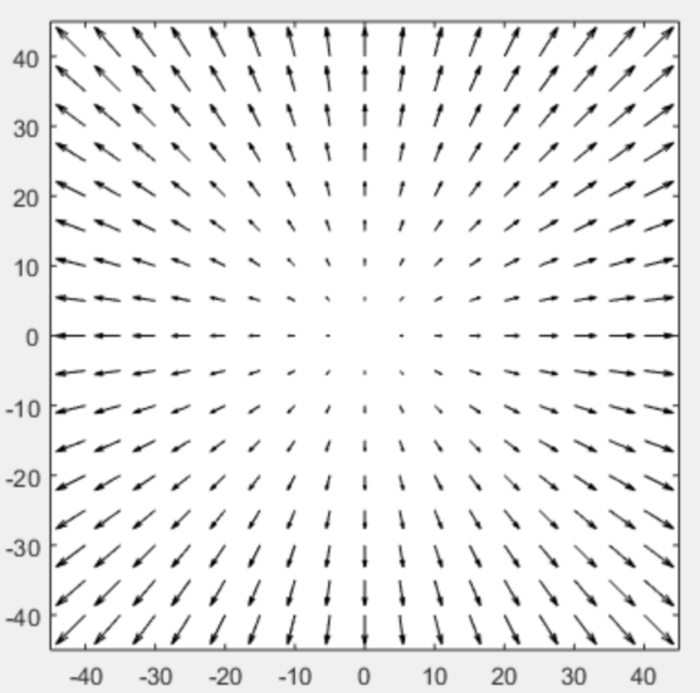


wall

Z_o = constant.

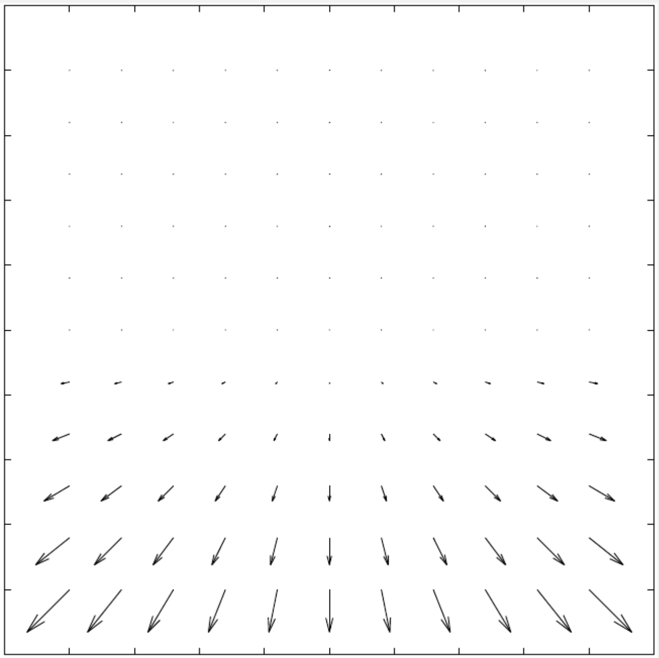
Forward Component $(T_X = T_Y = 0)$

$$(v_x, v_y) = \frac{T_Z}{Z_o}\left(\frac{x}{f}, \frac{y}{f}\right)$$



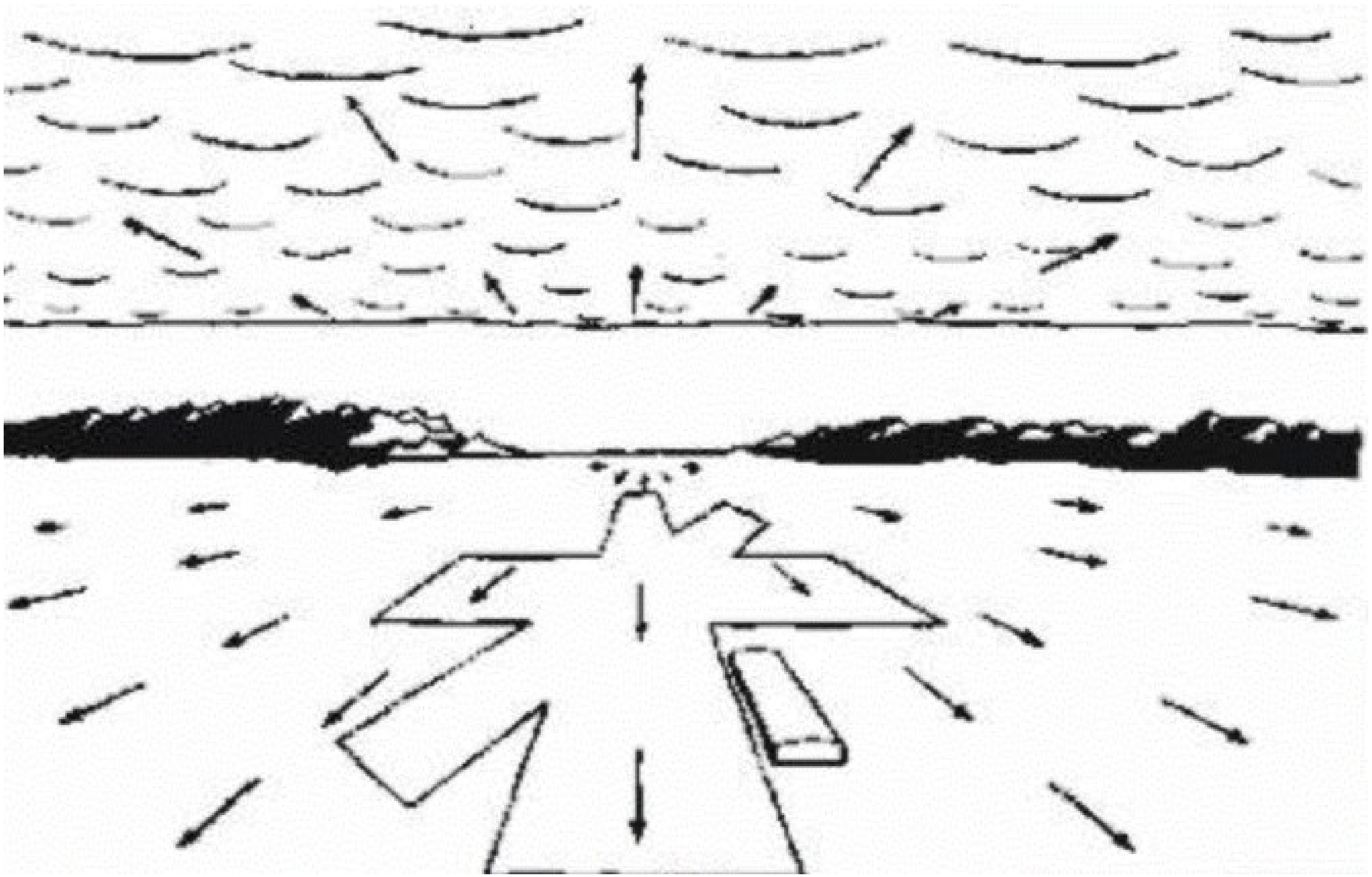
wall

Z_o = constant



ground plane

(see Exercises)



What does a pilot see when approaching the runway?

(from JJ Gibson 1950)

General Translation (when $T_z \neq \mathbf{0}$)

$$(v_x, v_y) = \frac{T_Z}{Z_o}\left(\frac{x}{f}, \frac{y}{f}\right) + \frac{1}{Z_o}(-T_X, -T_Y)$$

forward lateral

$$= \frac{T_Z}{Z_o}\left(\frac{x}{f} - \frac{T_x}{T_z}, \quad \frac{y}{f} - \frac{T_y}{T_z}\right)$$

General Translation (when $T_z \neq 0$)

$$(v_x, v_y) = \frac{T_Z}{Z_o}\left(\frac{x}{f}, \frac{y}{f}\right) + \frac{1}{Z_o}(-T_X, -T_Y)$$

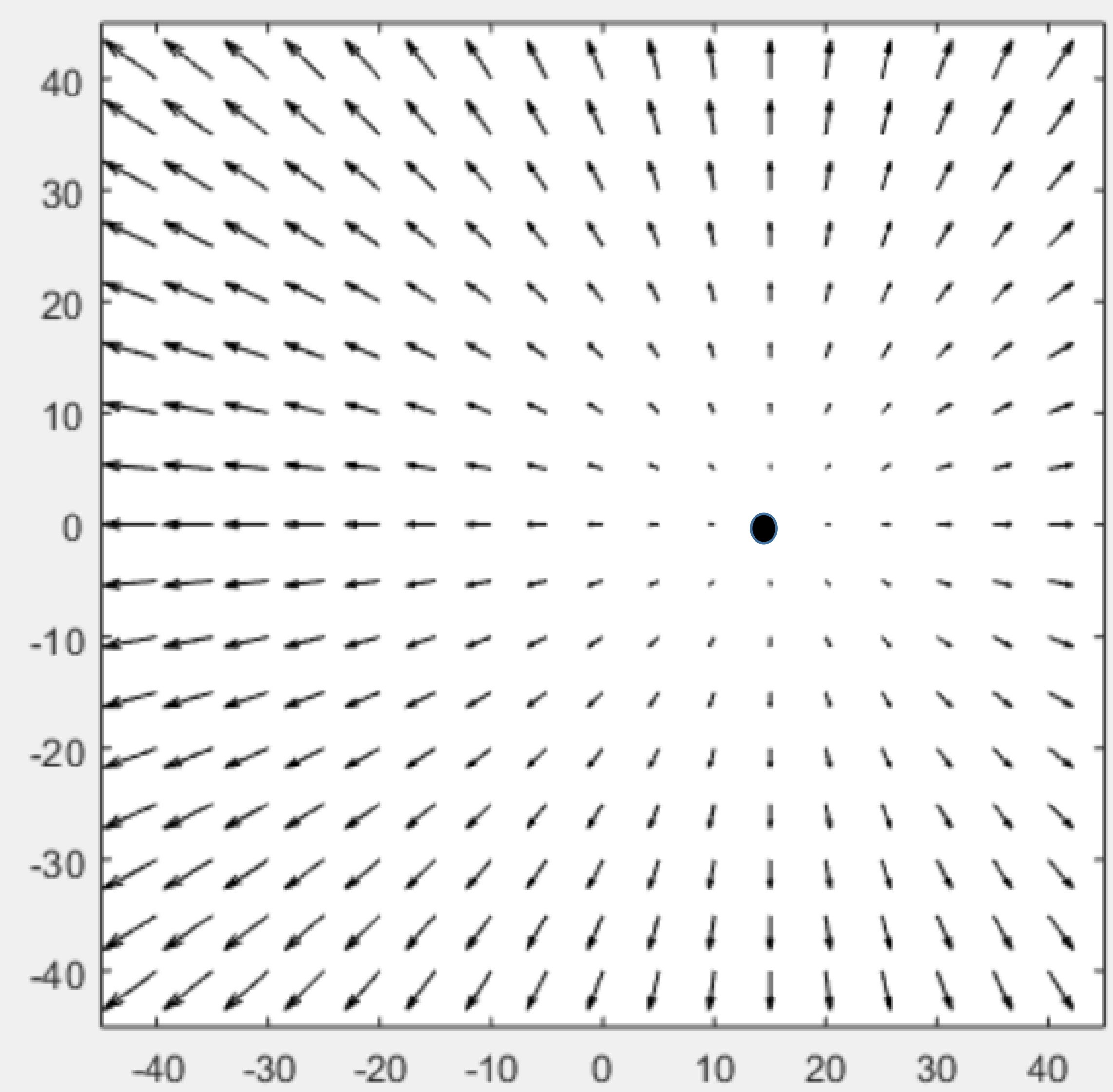
forward lateral

$$= \frac{T_Z}{Z_o}\left(\frac{x}{f} - \frac{T_x}{T_z}, \quad \frac{y}{f} - \frac{T_y}{T_z}\right)$$

$\left(\frac{T_x}{T_z}, \frac{T_y}{T_z}\right)$ is called the “direction of heading”.

Example: $(T_X, T_Y, T_Z) = (.3, 0, 1)$

Eccentricity (deg)



How can a *translating* observer estimate heading?

1) Estimate motion field (v_x, v_y).

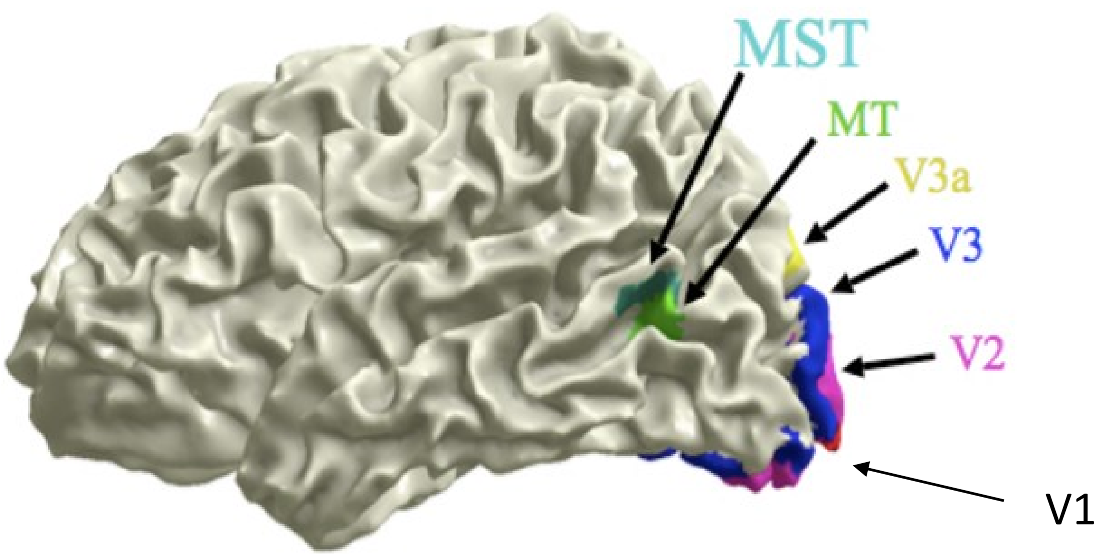
2) Estimate the direction of heading $\left(\frac{T_x}{T_z}, \frac{T_y}{T_z}\right)$ to be the image point where all motion vectors point away from that point.

How/where does the brain solve this problem?

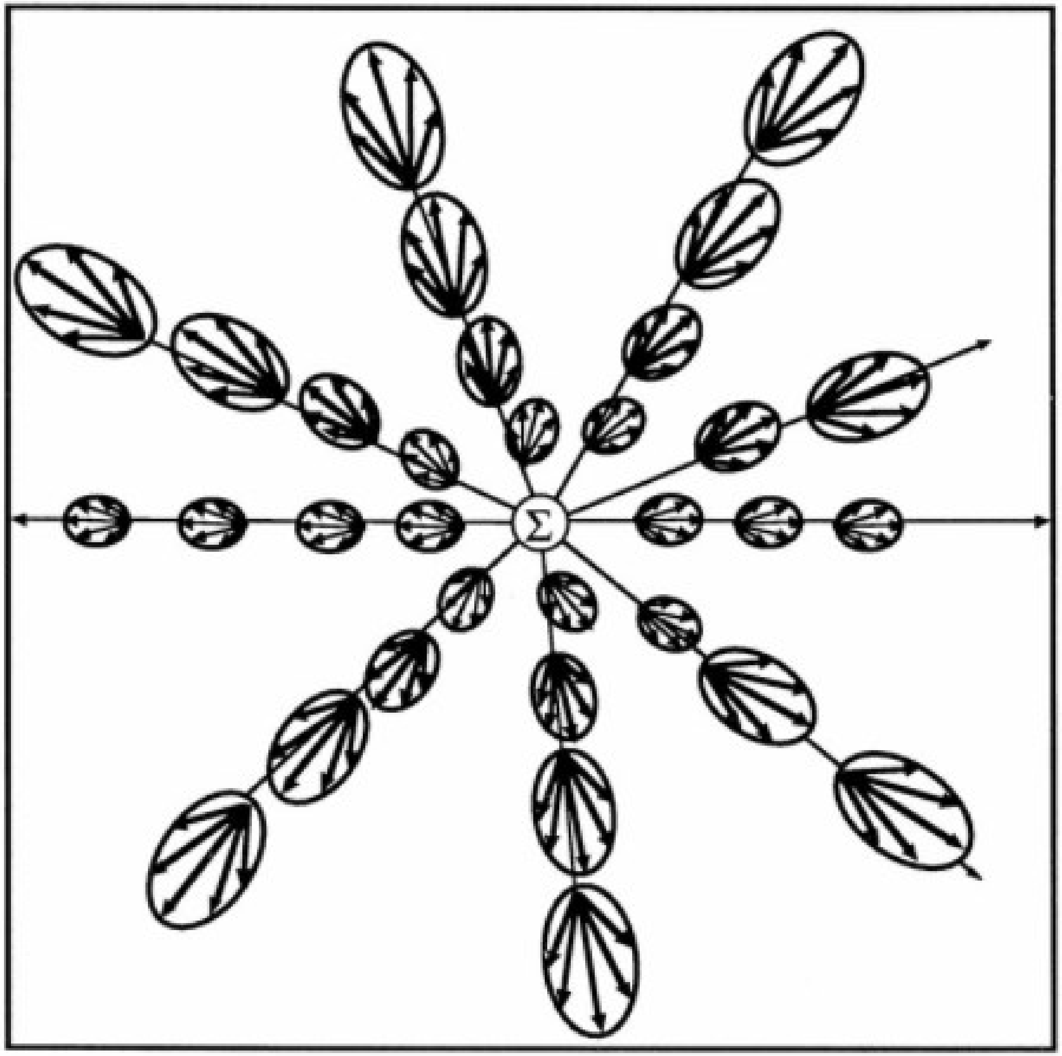
V1: measure normal velocity components with Gabors

→ MT (middle temporal lobe): estimate velocities (v_x, v_y)

→ **MST (medial superior temporal lobe): estimate global motion field**



MST cell 'templates'
(computation model)



Huge receptive fields

Each vector represents a normal component of velocity (V1 cell).

Each disk represents an MT cell (last lecture).

[Perrone & Stone, 1998]

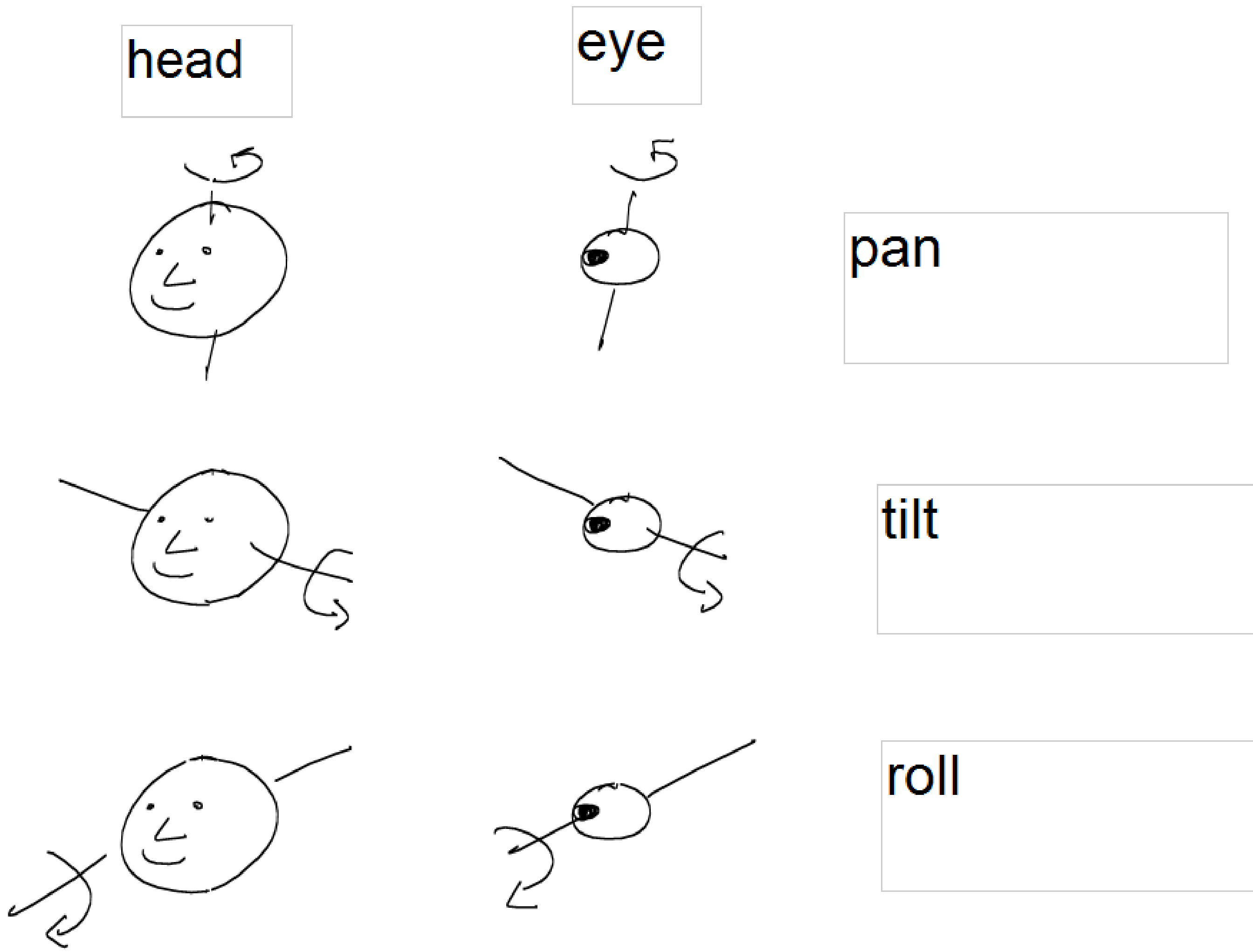
COMP 546

Lecture 9

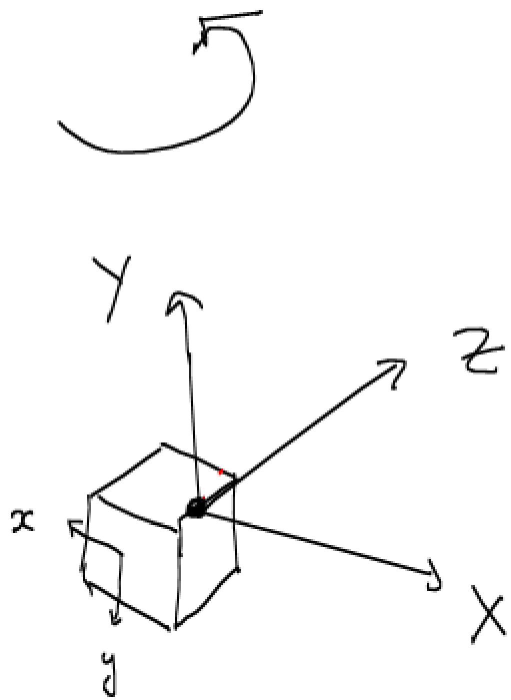
egomotion:
translation & rotation (eye movements)

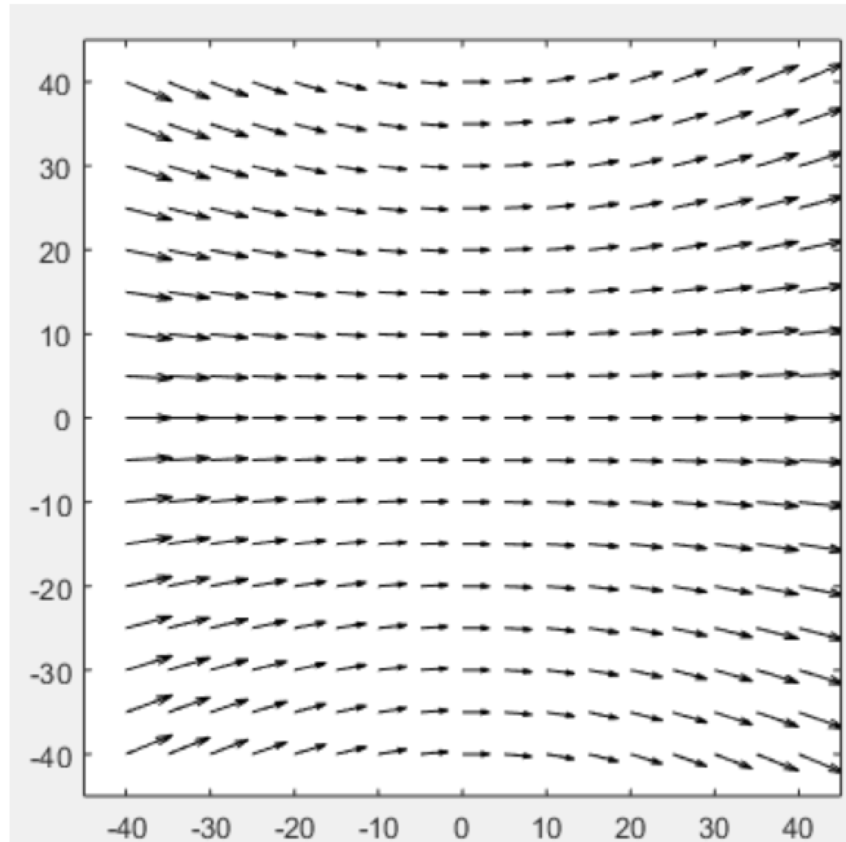
Thurs. Feb. 8, 2018

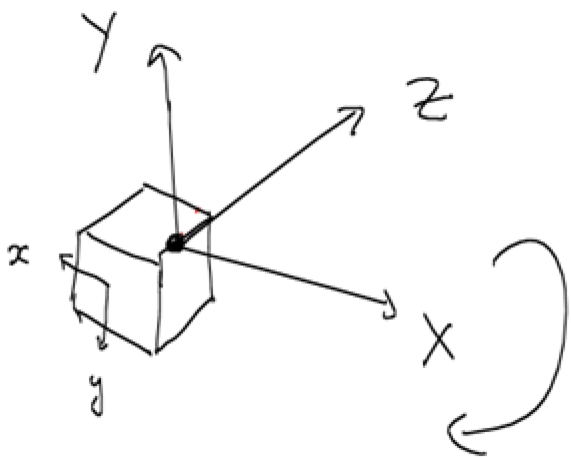
Motion field seen by a rotating observer ?



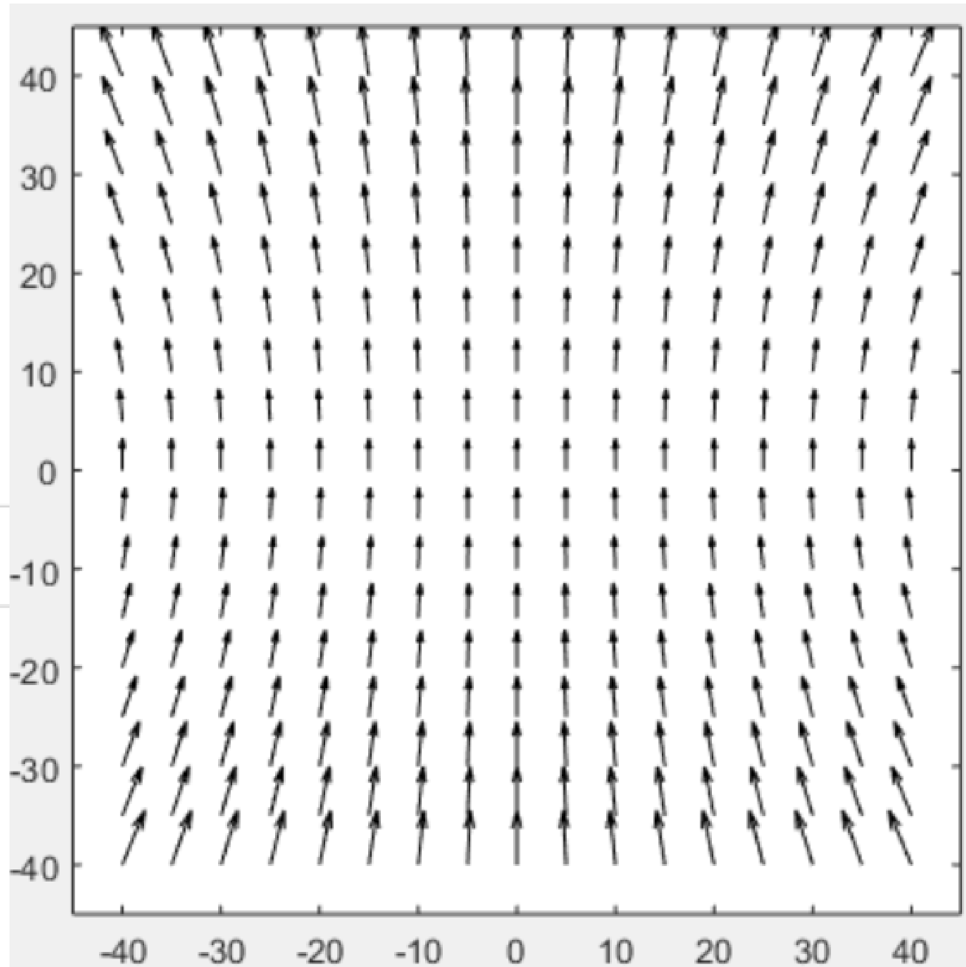
pan

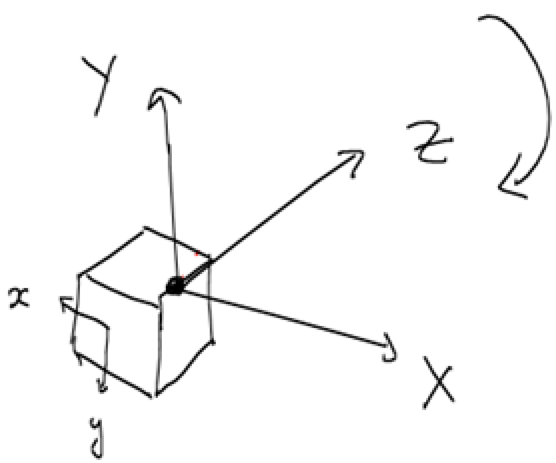




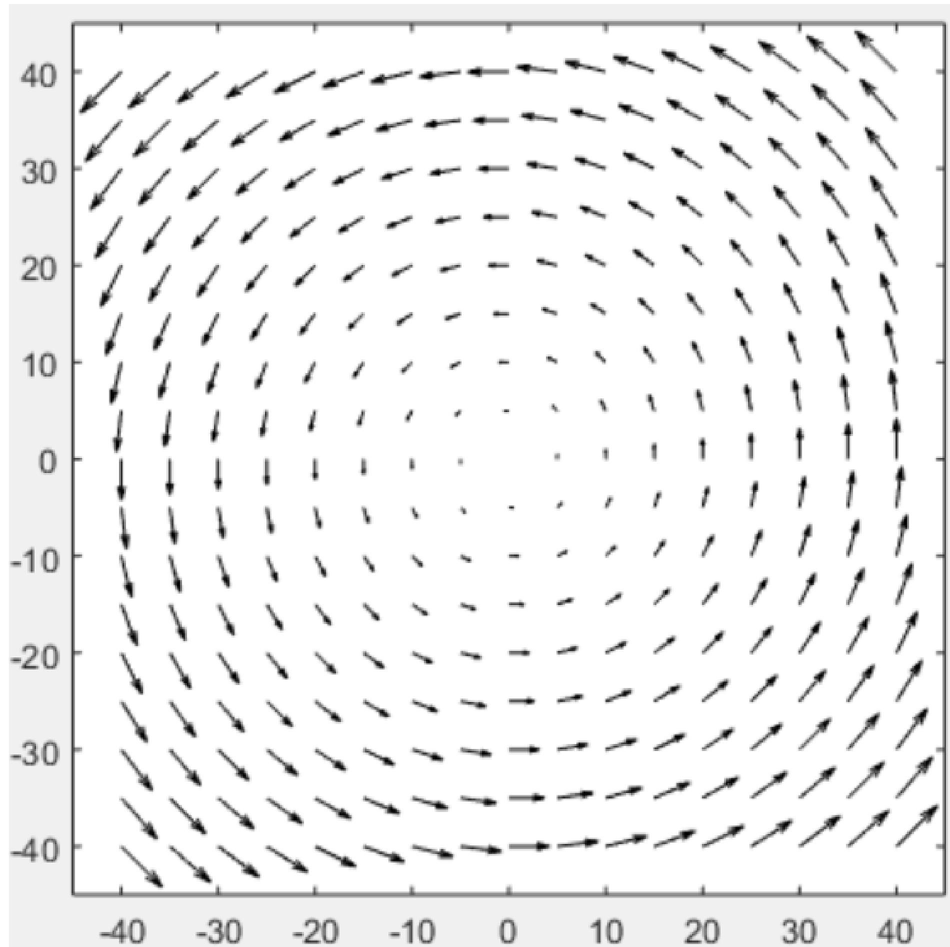


tilt





roll



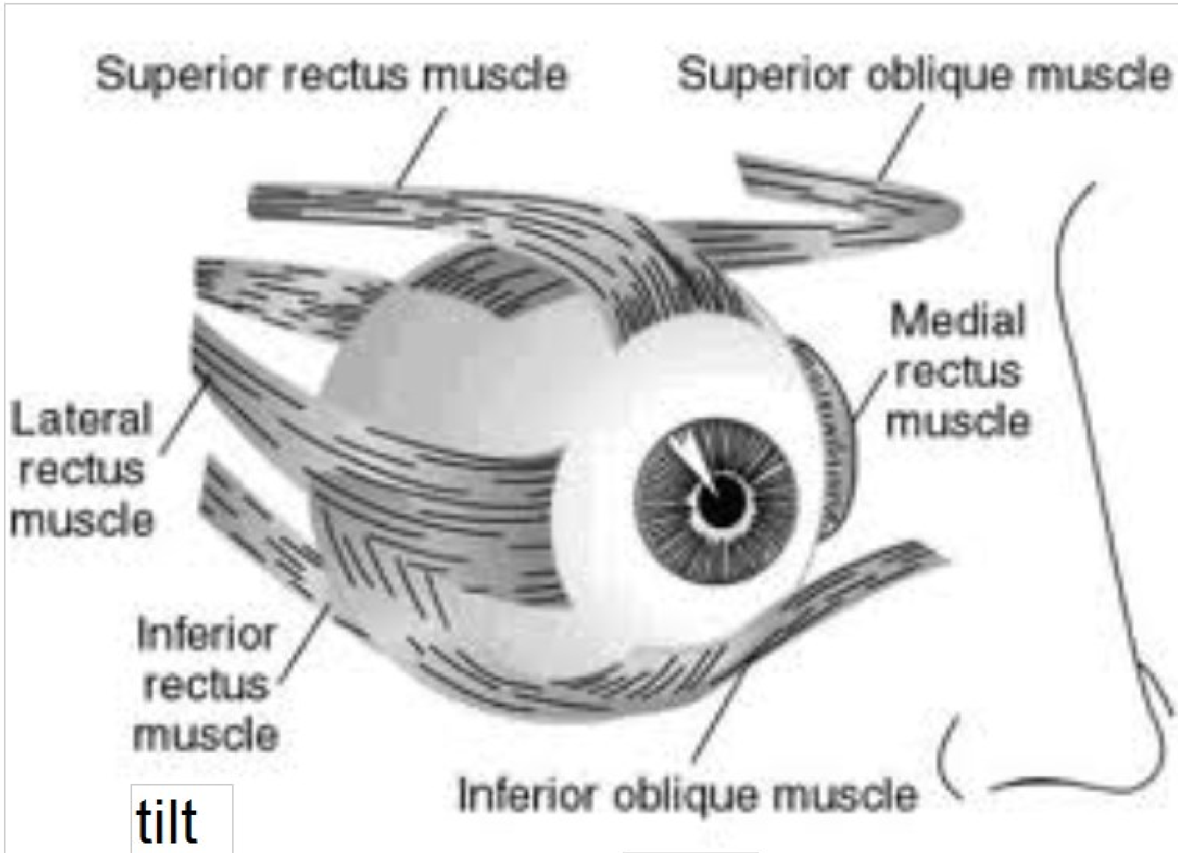
Eye Rotations (called "eye movements")

tilt

pan

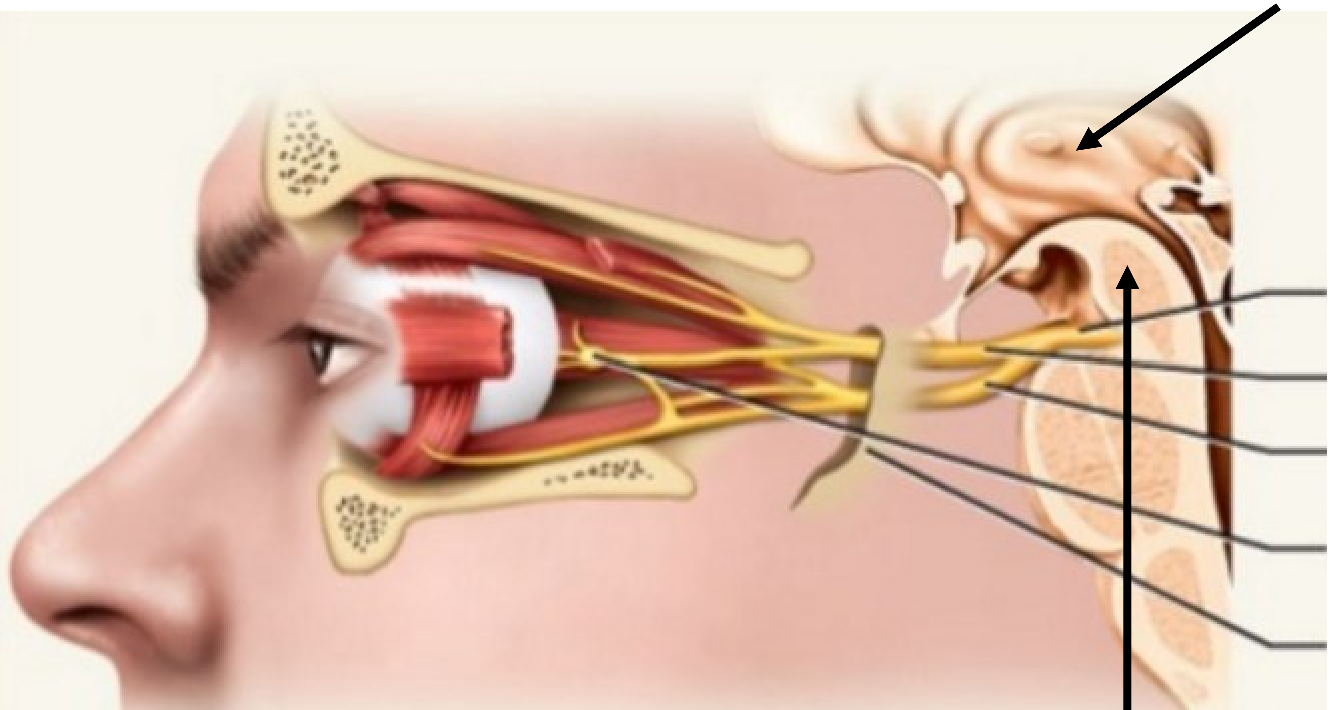
roll

pan



tilt

roll



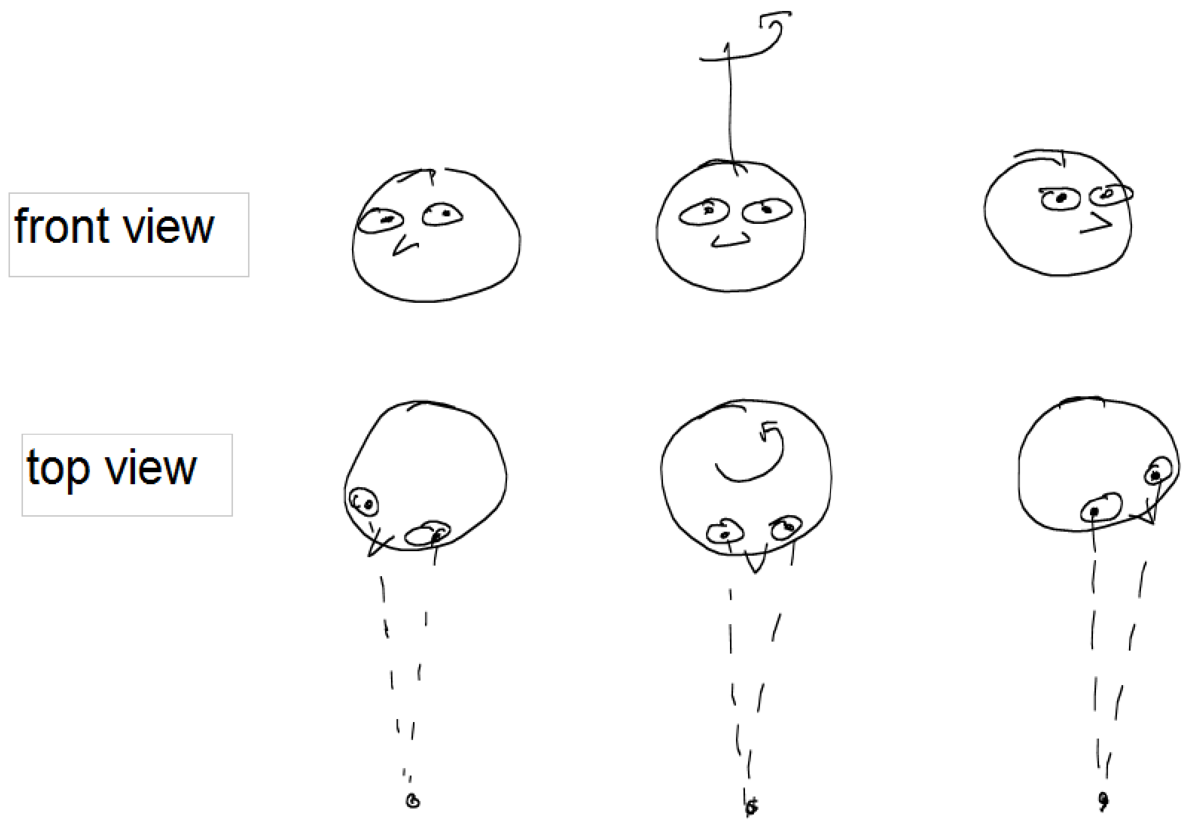
All eye movement *motor* (output) signals come from mid-brain e.g. oculomotor nerves.

These nerves also control accommodation, blinks, pupil contraction.

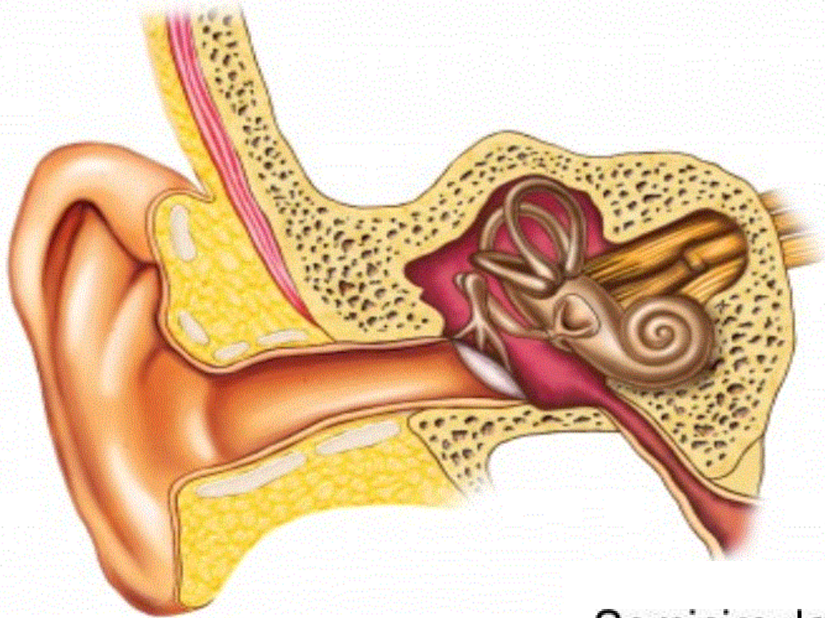
Types of eye movements

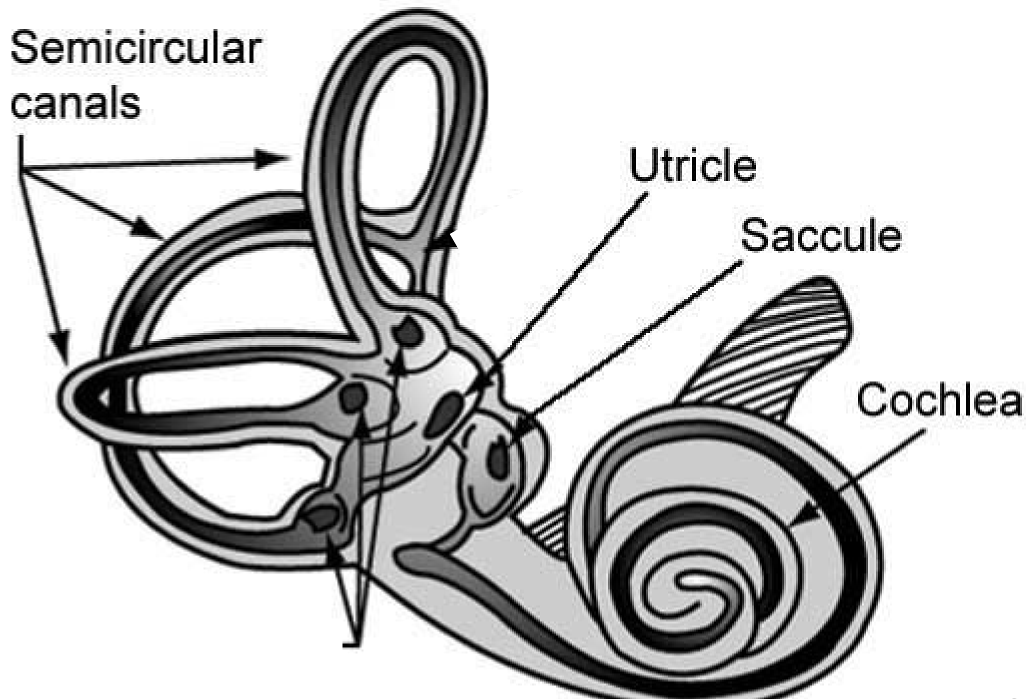
- vestibulo-ocular reflex (VOR)
- smooth pursuit
- vergence (next lecture)
- saccades (later in course)
- OKN (optokinetic nystagmus) OMIT

VOR
(eye rotations due to head movement)



Vestibular System
(in the inner ear)





Vestibular system: the brain's IMU

(inertial measurement unit – a term used in robotics)

It measures:

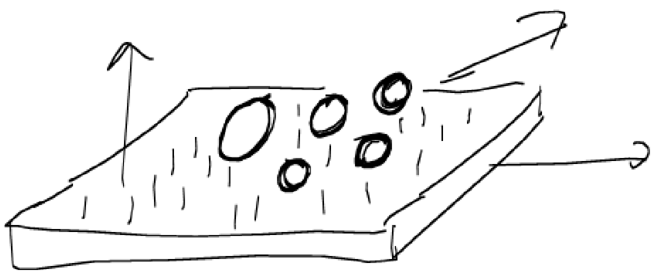
- linear acceleration of head

$$\frac{d}{dt} \quad (T_X, T_Y, T_Z)$$

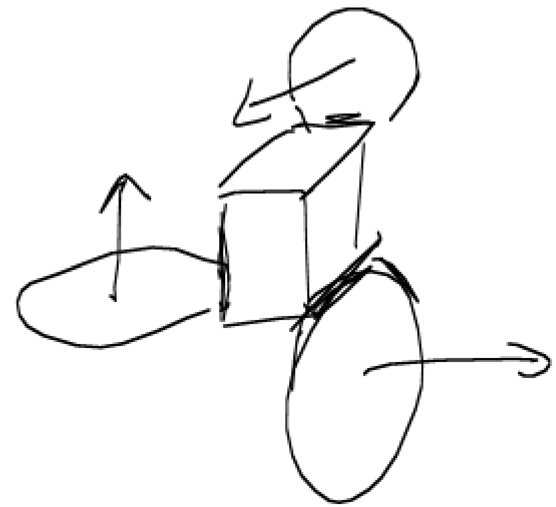
- angular acceleration of head

$\frac{d}{dt}$ (pan, tilt, roll)

otoliths

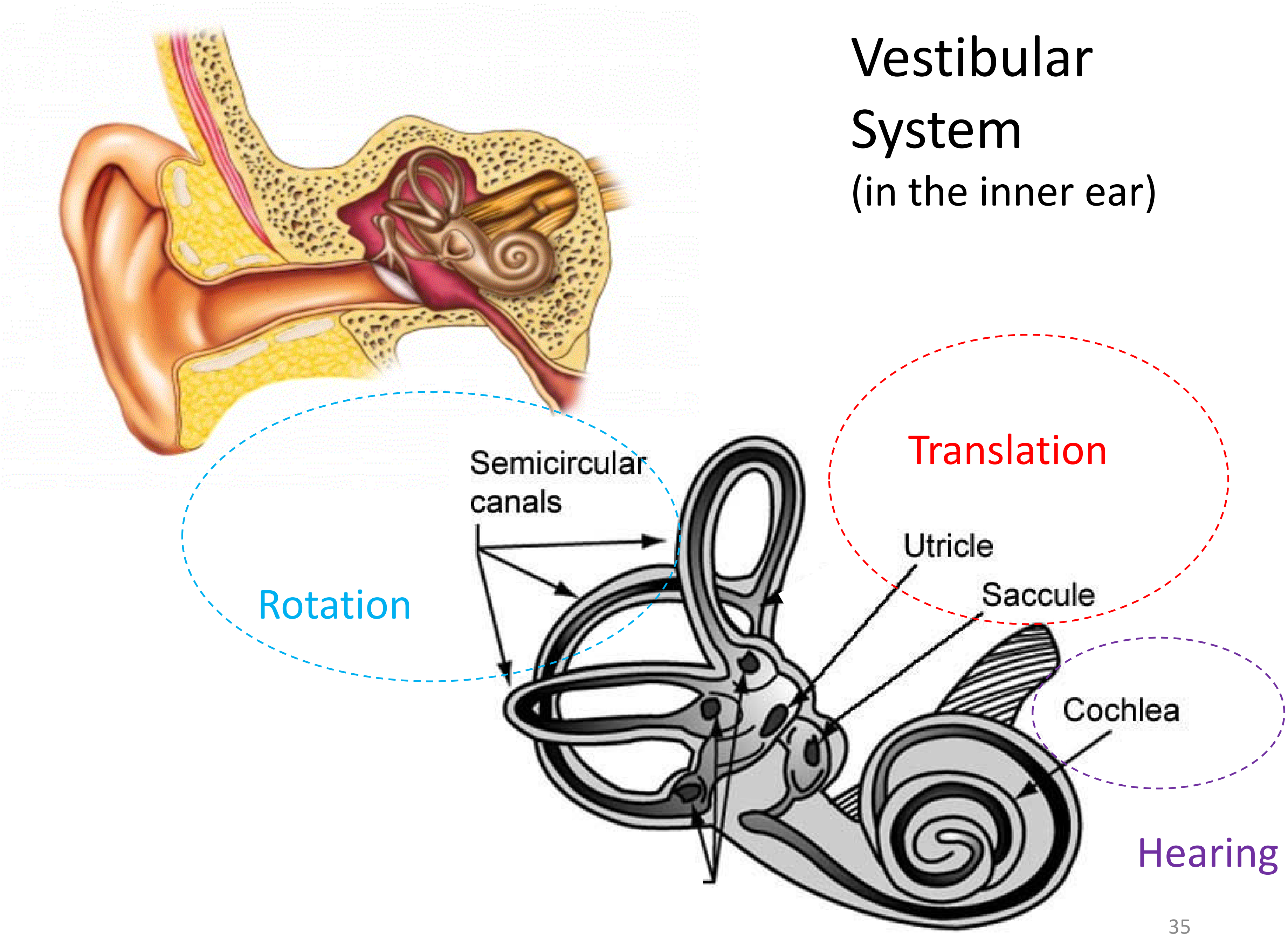


Translation
(linear acceleration)



Rotation
(angular acceleration)

Vestibular System
(in the inner ear)



Smooth Pursuit Eye Movements

Tracking a static object as the observer moves

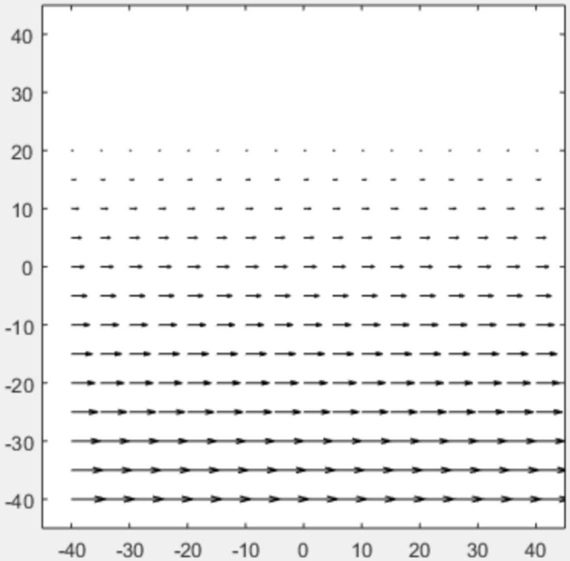
(or smoothly tracking a moving object)

Reduces retinal motion of object to 0.

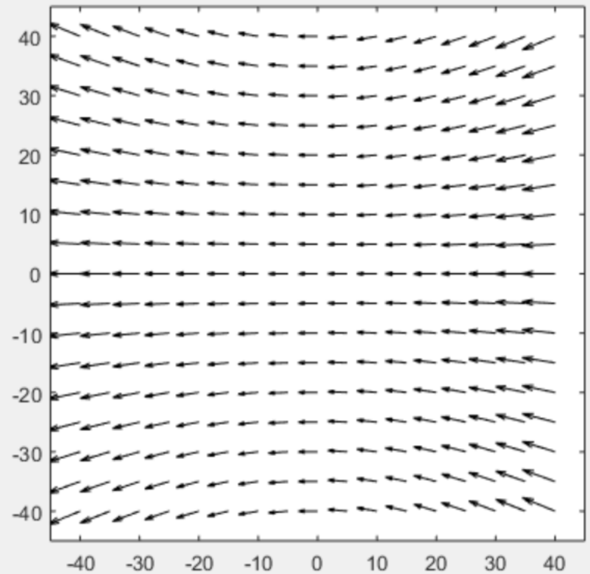
Combined translation and rotation

Q: where is direction of heading?

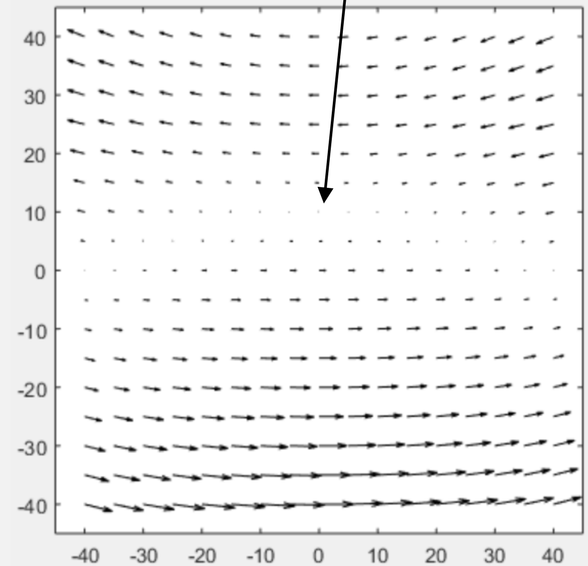
Zero image velocity (tracking)



translation only



rotation only

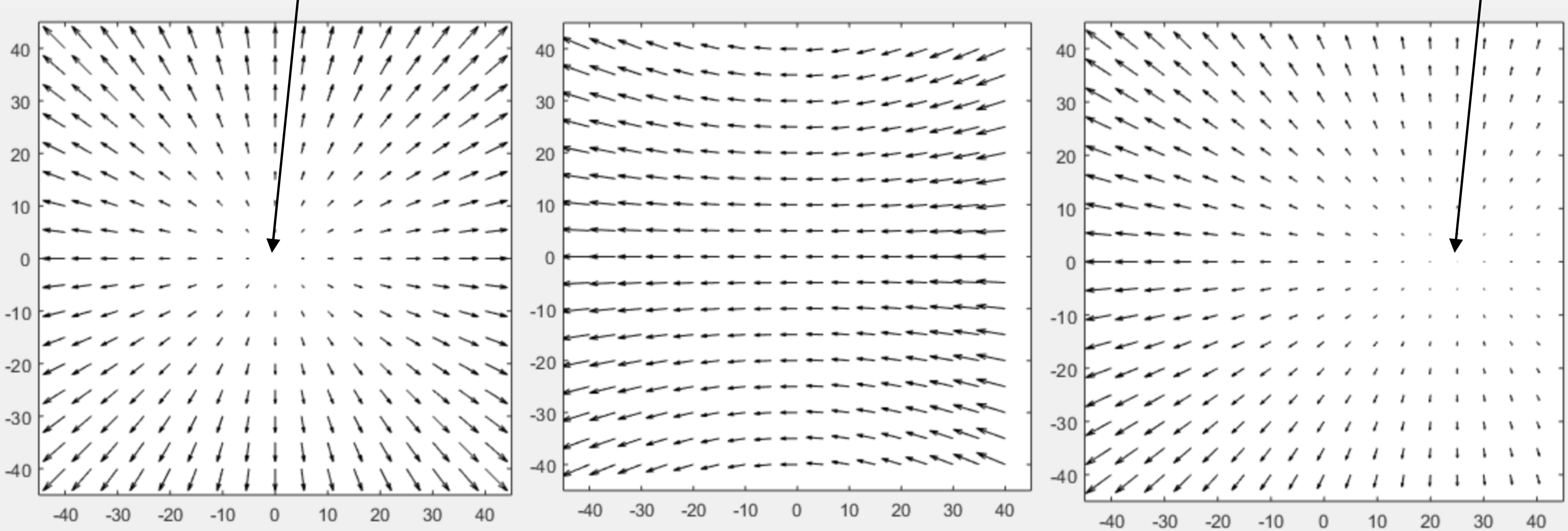


translation + rotation

Combined translation and rotation

Direction of heading

Zero image velocity (tracking)



translation only

rotation only

translation + rotation

Discussion/Summary

- Motion field seen by moving observer is the sum of translation & rotation fields

- Cells in MST are sensitive to global motion fields (huge receptive fields) and are believed to be involved in estimating egomotion. [It is not entirely clear what exactly the computational role of these cells is. But there is a lot of evidence that these cells exist, and that they do play a role in estimating heading direction.]

- VOR adds eye rotation to cancel the image motion that is due to head motion. Smooth pursuit adds eye rotation that reduces the retinal velocity of certain "interest points" to zero, allowing a detailed "still" image analysis. In both cases, these rotations are known (controlled by the brain) so their effects can be accounted for, i.e. the translation and rotation components of the motion field can be disentangled.