Visual Perception: Motion Perception

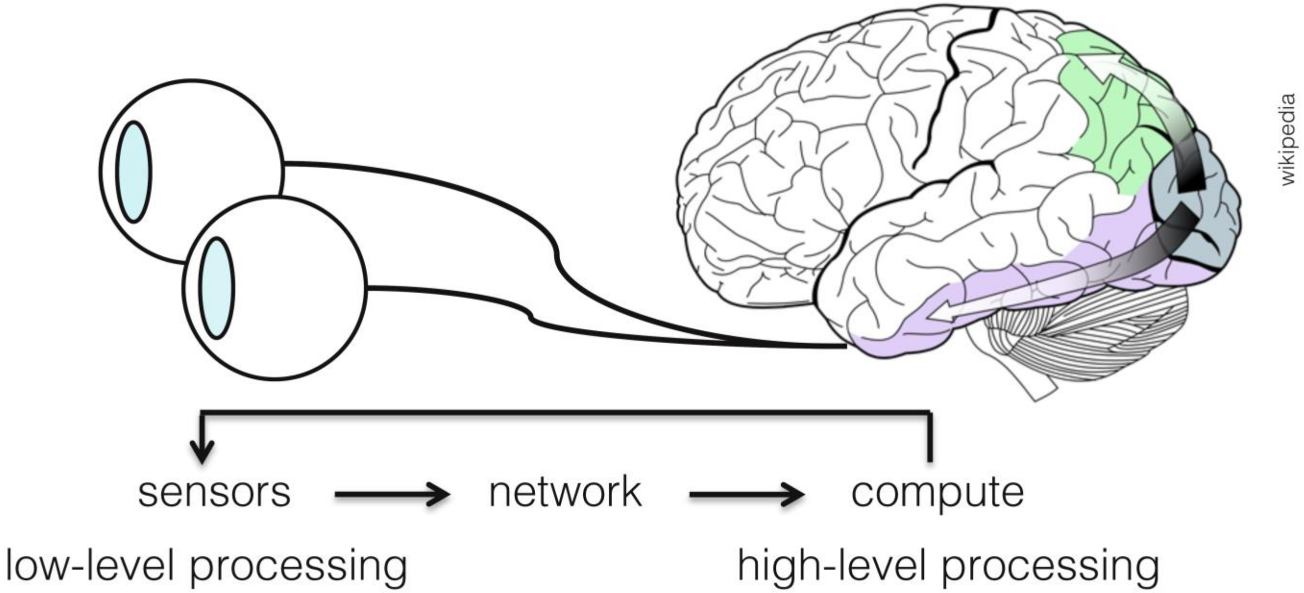
CS 6334 Virtual Reality

Professor Yu Xiang

The University of Texas at Dallas

Visual Perception

- How humans perceive or interpret the real world using vision?



- We need to understand visual perception to achieve visual unawareness in VR systems

Motion Perception

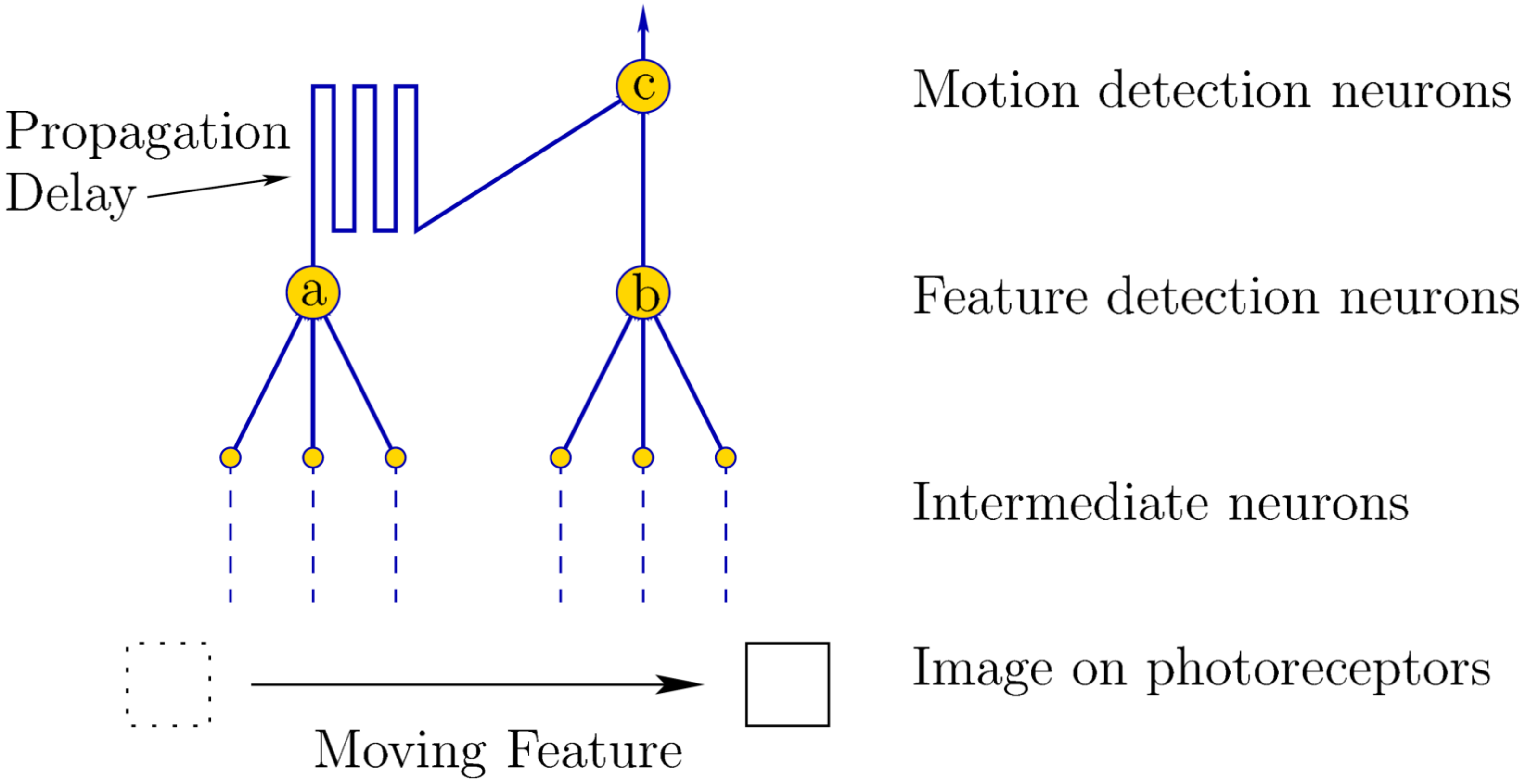
- Separate moving figure from a stationary background

- Motion for 3D perception
  - Look at a fruit by rotating it around

- Guide actions
  - Walking down the street or hammering a nail

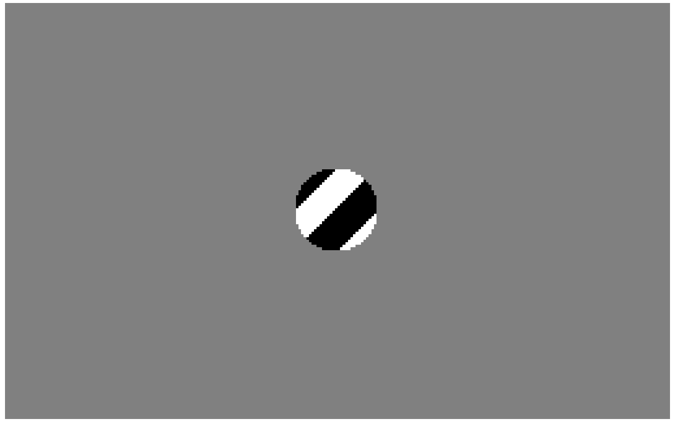
Reichardt Detector

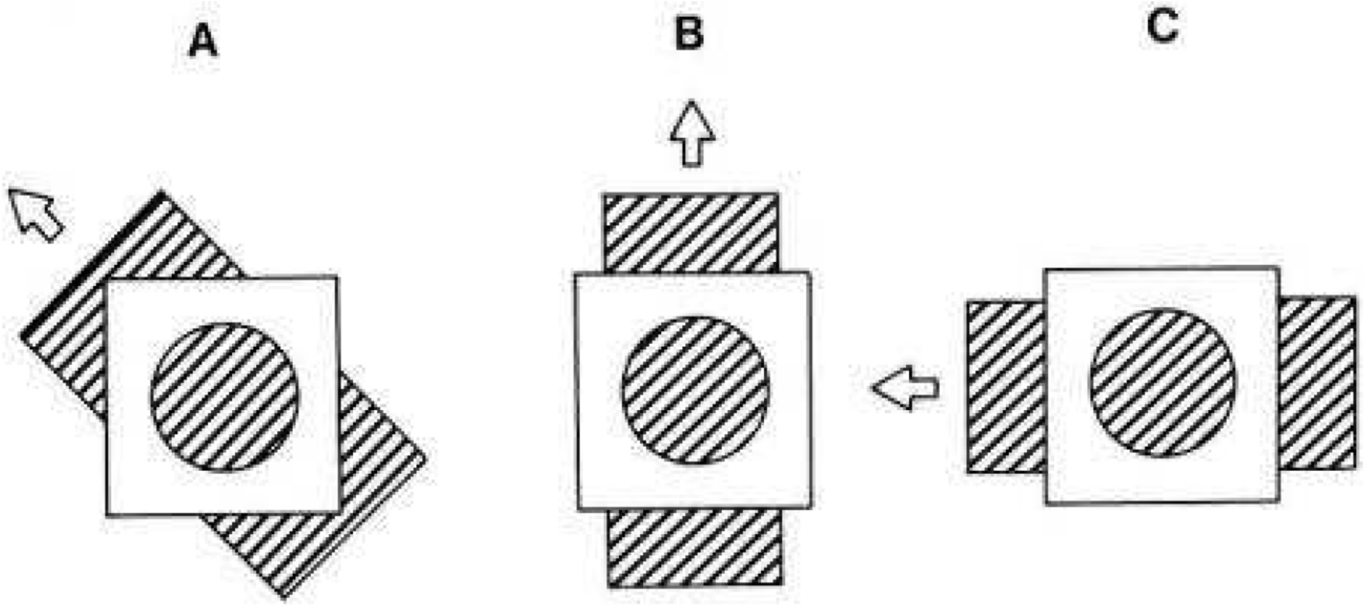
- A neural circuitry model for motion perception



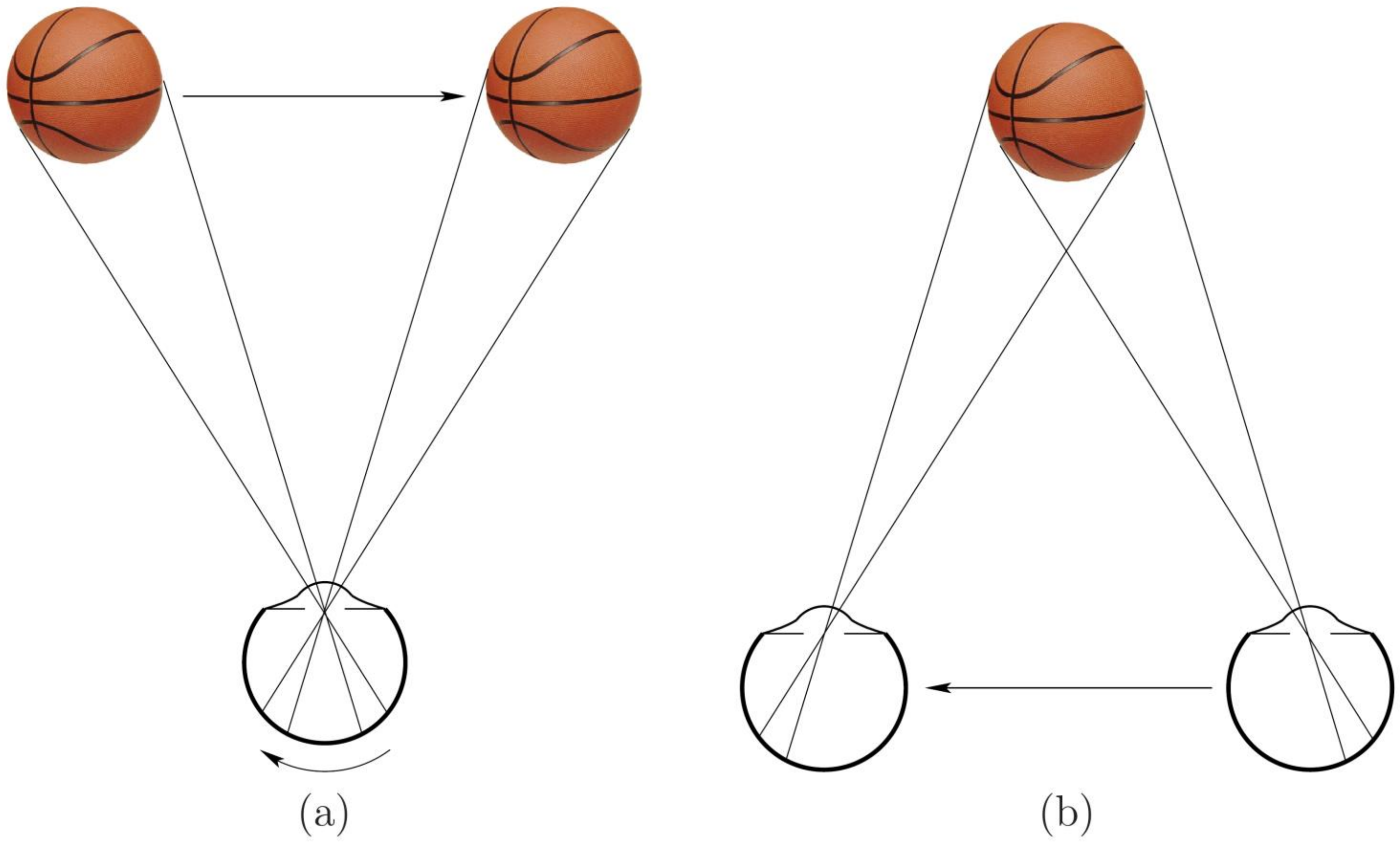
Local to Global

- Motion detectors are local
- Our visual system infers the global motion
- The aperture problem





Object Motion vs. Eye Movement



Two motions that cause equivalent movement of the image on the retina

- **Saccadic masking (saccadic suppression)**: the brain selectively block visual processing during eye movements, suppress motion detectors in the second case
- **Proprioception**: the body's ability to estimation its own motions due to motor commands (i.e., use of eye muscles)
- **Information is provided by large-scale motion**: if the entire scene is moving, the brain interprets the user must be moving

Stroboscopic Apparent Motion

- Motion from a sequence of still images being flashed onto the screen
  - TV, small phone, movie screen

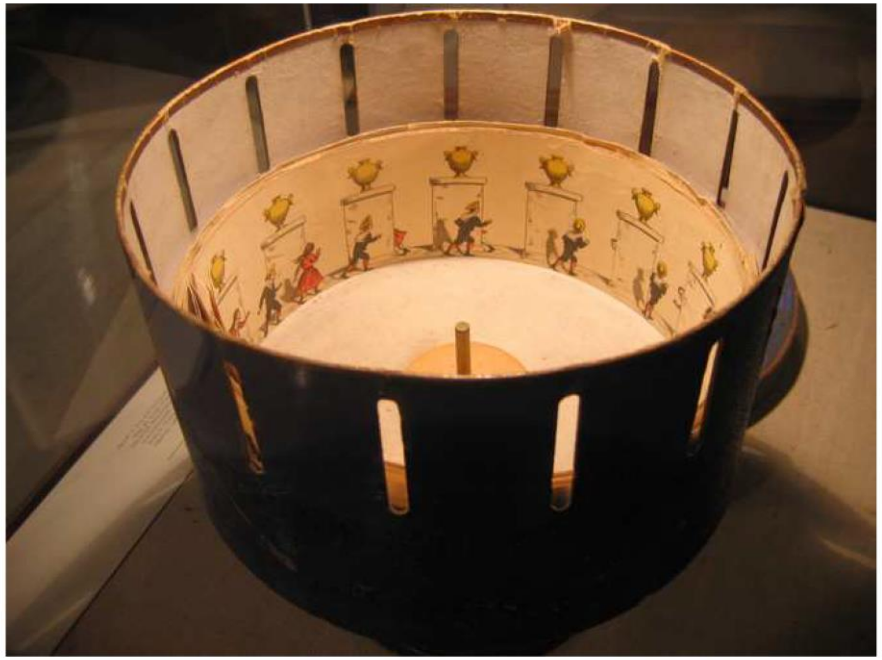
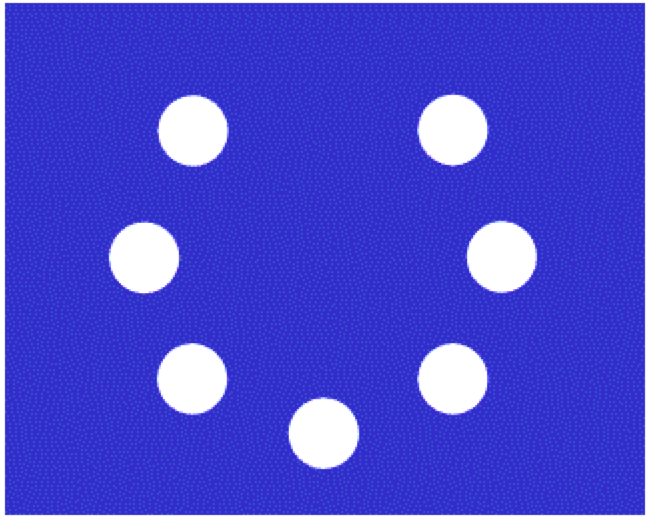


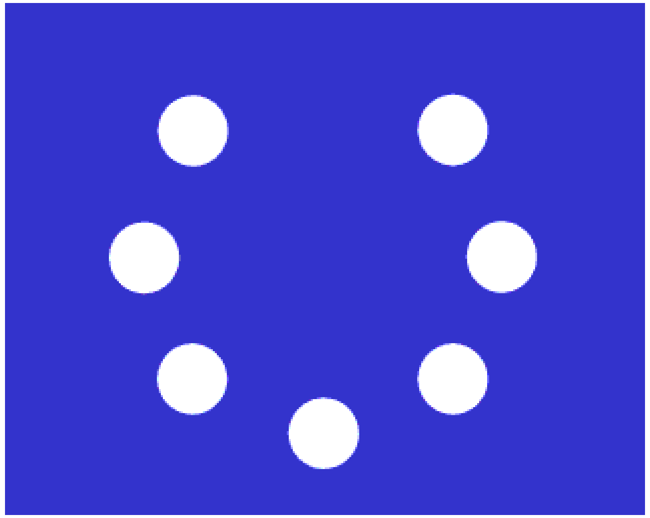
Figure 6.15: The *zoetrope* was developed in the 1830s and provided stroboscopic apparent motion as images became visible through slits in a rotating disc.

Beta Movement and Phi Phenomenon



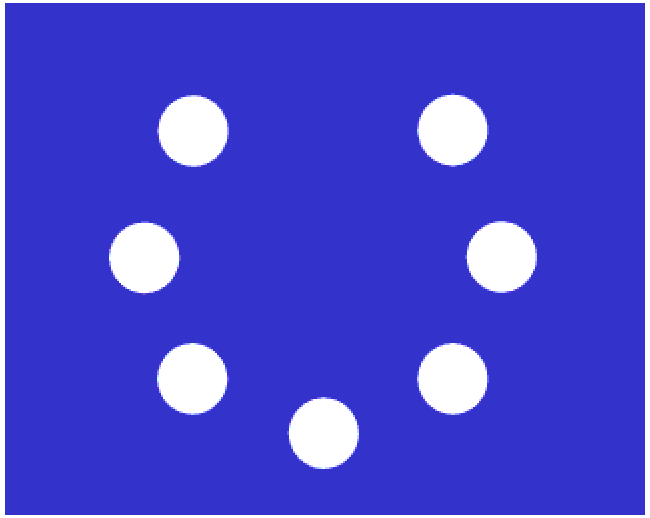
Still image

Beta Movement



Low frequency (2fps)
Jumping dot

Phi Phenomenon



High frequency (15fps)
Moving hole

We can perceive motion at 2fps!

- Stroboscopic apparent motion triggers the neural motion detection circuitry

Frame Rates

FPS	Occurrence
2	Stroboscopic apparent motion starts
10	Ability to distinguish individual frames is lost
16	Old home movies; early silent films
24	Hollywood classic standard
25	PAL television before interlacing
30	NTSC television before interlacing
48	Two-blade shutter; proposed new Hollywood standard
50	Interlaced PAL television
60	Interlaced NTSC television; perceived flicker in some displays
72	Three-blade shutter; minimum CRT refresh rate for comfort
90	Modern VR headsets; no more discomfort from flicker
1000	Ability to see zipper effect for fast, blinking LED
5000	Cannot perceive zipper effect

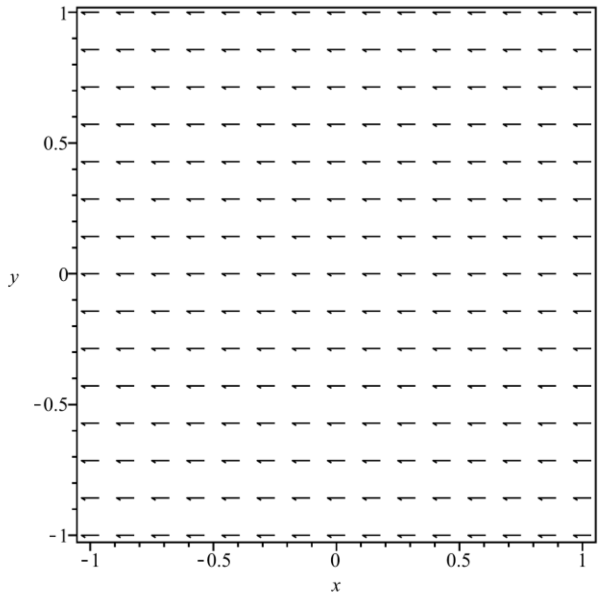
Interlacing: draw half the image in one frame time, and then half in the other

Blade shutter: show each frame two or three times

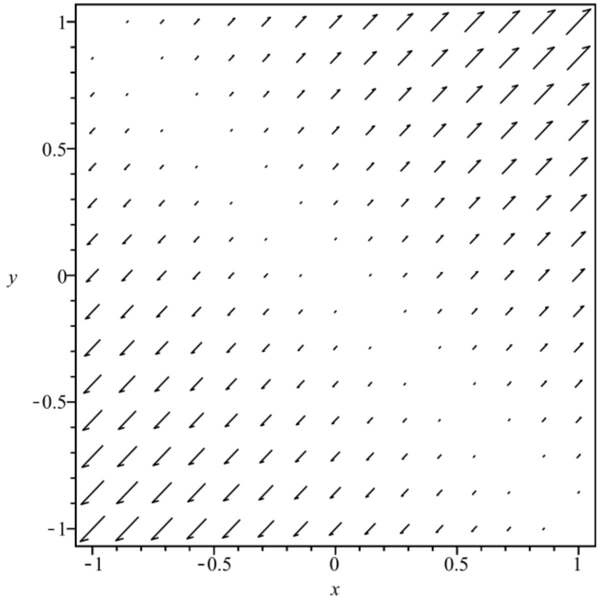
Optical Flow

- The pattern of apparent motion of objects, surfaces and edges in a visual scene caused by the relative motion between an observer and a scene
- Velocity field

$$(v_x, v_y) = \left(\frac{dx}{dt}, \frac{dy}{dt}\right)$$

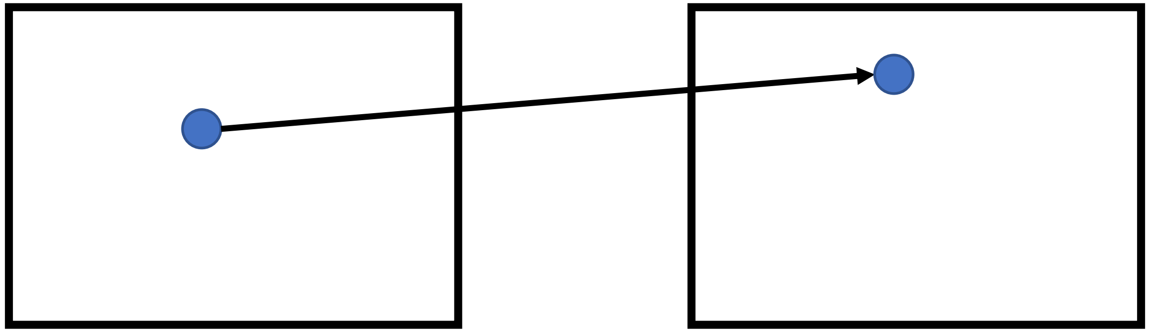


$$(x, y) \mapsto (-1, 0)$$



$$(x, y) \mapsto (x + y, x + y)$$

Brightness Constancy Constraint

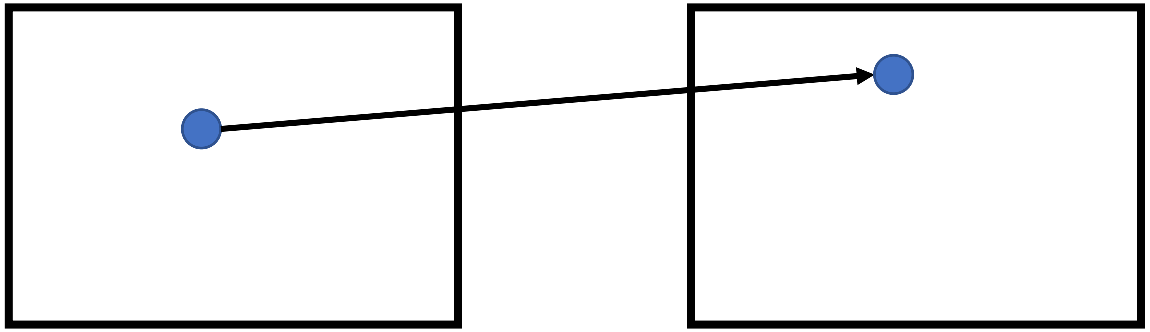


$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

Taylor series

$$I(x + \Delta x, y + \Delta y, t + \Delta t) = I(x, y, t) + \frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t + \text{higher-order terms}$$

Brightness Constancy Constraint



$$\frac{\partial I}{\partial x}\Delta x + \frac{\partial I}{\partial y}\Delta y + \frac{\partial I}{\partial t}\Delta t = 0$$

$$\frac{\partial I}{\partial x}\frac{\Delta x}{\Delta t} + \frac{\partial I}{\partial y}\frac{\Delta y}{\Delta t} + \frac{\partial I}{\partial t}\frac{\Delta t}{\Delta t} = 0$$

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$$

Brightness Constancy Constraint

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0$$

$\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}$ (spatial gradient; we can compute this!)

$\frac{dx}{dt}, \frac{dy}{dt} = (u, v)$ (optical flow, what we want to find)

$\frac{\partial I}{\partial t}$ (derivative across frames. Also known, e.g. frame difference)

Image Gradient

- Derivative of a function

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

- Central difference is more accurate

$$f'(x) = \lim_{h \to 0} \frac{f(x+0.5h) - f(x-0.5h)}{h}$$

- Image gradient with central difference
  - Applying a filter

1	0	-1

X derivative

1
0
-1

Y derivative

Image Gradient

- Sobel Filter

1	0	-1
2	0	-2
1	0	-1

Sobel

=

1
2
1

weighted average and scaling

1	0	-1

x-derivative

$$\boldsymbol{S}_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

$$\frac{\partial \boldsymbol{f}}{\partial x} = \boldsymbol{S}_x \otimes \boldsymbol{f}$$

$$\boldsymbol{S}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

$$\frac{\partial \boldsymbol{f}}{\partial y} = \boldsymbol{S}_y \otimes \boldsymbol{f}$$

$$\nabla \boldsymbol{f} = \left[\frac{\partial \boldsymbol{f}}{\partial x}, \frac{\partial \boldsymbol{f}}{\partial y} \right]$$

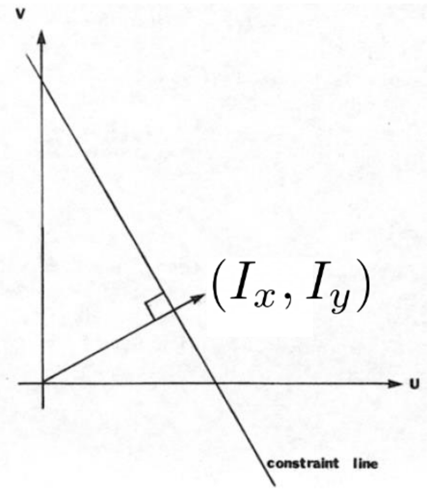
Brightness Constancy Constraint

$$I_x u + I_y v + I_t = 0$$

Known (spatial and temporal gradients)

Unknown (optical flow)

- For each pixel, there are two unknowns



Brightness Constancy Constraint

$$I_x u + I_y v + I_t = 0$$

- The component of the flow vector in the gradient direction is determined (called normal flow) (Recall vector projection geometry)

$$\frac{1}{\sqrt{I_x^2 + I_y^2}}(I_x, I_y) \cdot (u, v) = \frac{-I_t}{\sqrt{I_x^2 + I_y^2}}$$

- The component of the flow vector orthogonal to this direction cannot be determined.

https://en.wikipedia.org/wiki/Dot_product

Lucas-Kanade Method

$$I_x u + I_y v + I_t = 0$$

- Assumption: the flow is constant in a local neighborhood of a pixel under consideration
- Use two or more pixels to compute optical flow

5x5 window

$$\underbrace{\begin{bmatrix} I_x(\mathbf{p}_1) & I_y(\mathbf{p}_1) \\ I_x(\mathbf{p}_2) & I_y(\mathbf{p}_2) \\ \vdots & \vdots \\ I_x(\mathbf{p}_{25}) & I_y(\mathbf{p}_{25}) \end{bmatrix}}_{\substack{A \\ 25\times 2}} \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_{\substack{d \\ 2\times 1}} = - \underbrace{\begin{bmatrix} I_t(\mathbf{p}_1) \\ I_t(\mathbf{p}_2) \\ \vdots \\ I_t(\mathbf{p}_{25}) \end{bmatrix}}_{\substack{b \\ 25\times 1}}$$

Lucas-Kanade Method

- Solve the least squares problem

$$\underset{25\text{x}2}{A}\ \underset{2\text{x}1}{d} = \underset{25\text{x}1}{b} \longrightarrow \text{minimize } \|Ad - b\|^2$$

$$\overset{2\text{x}2}{(A^T A)}\ \overset{2\text{x}1}{d} = \overset{2\text{x}1}{A^T b}$$

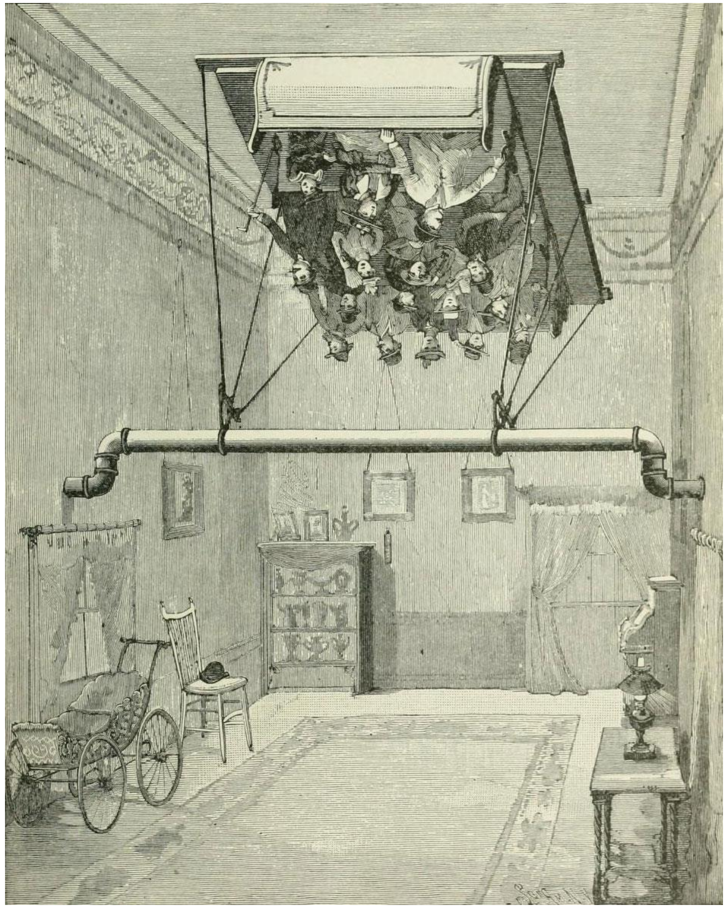
$$\underbrace{\begin{bmatrix} \sum I_x I_x & \sum I_x I_y \\ \sum I_x I_y & \sum I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} u \\ v \end{bmatrix} = - \underbrace{\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix}}_{A^T b}$$

https://en.wikipedia.org/wiki/Proofs_involving_ordinary_least_squares#Least_squares_estimator_for_.CE.B2

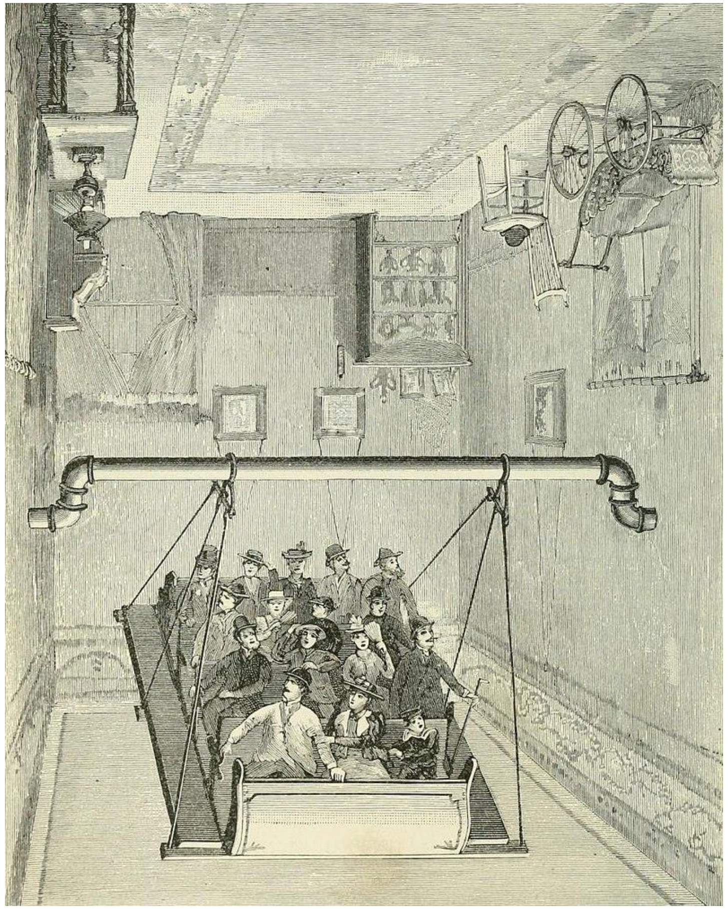
Optical Flow Example

Vection

- Illusions of self-motion
  - The brain is tricked into believing that the head is moving based on what is seen, even though no motion occurs.
- The haunted swing illusion
  - The room is rotating, and the persons are stationary
- Vection is commonly induced in VR
  - Moving the user's viewpoint
  - Leads to VR sickness, such as dizziness, nausea or even vomiting



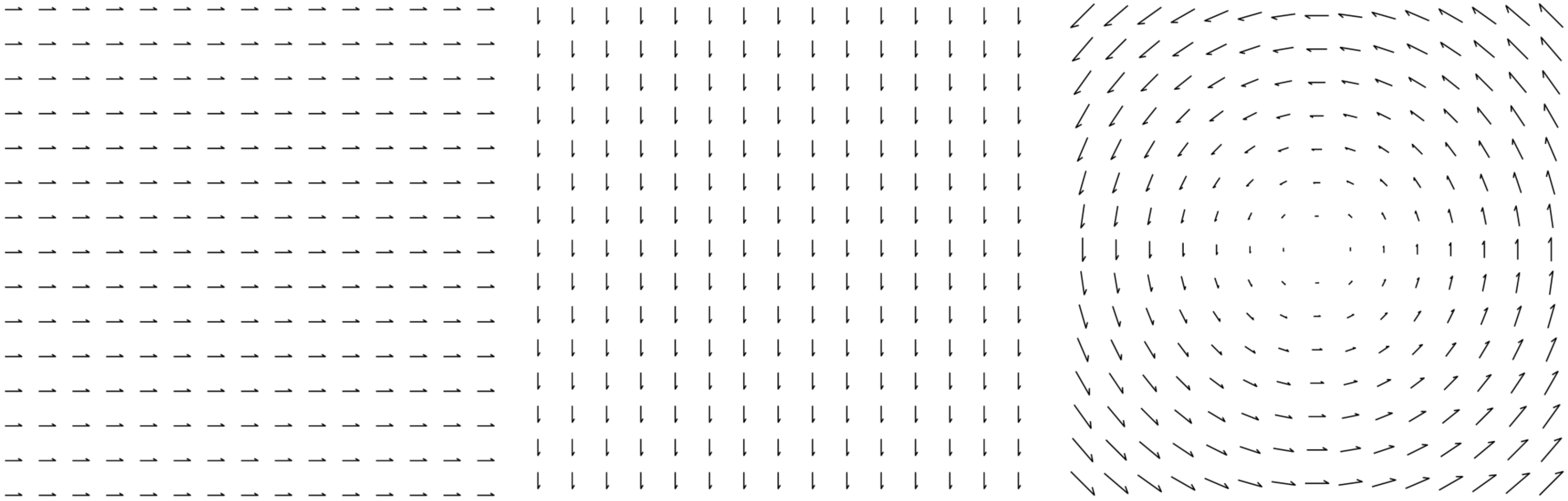
Perspective of riders



Actual swing position

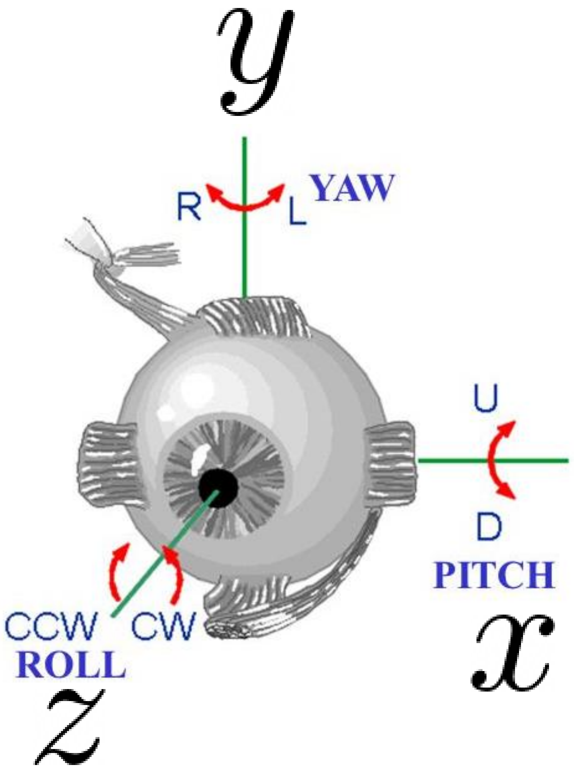
Vection Types

Changing viewpoints (rotation)



(a) yaw   (b) pitch   (c) roll

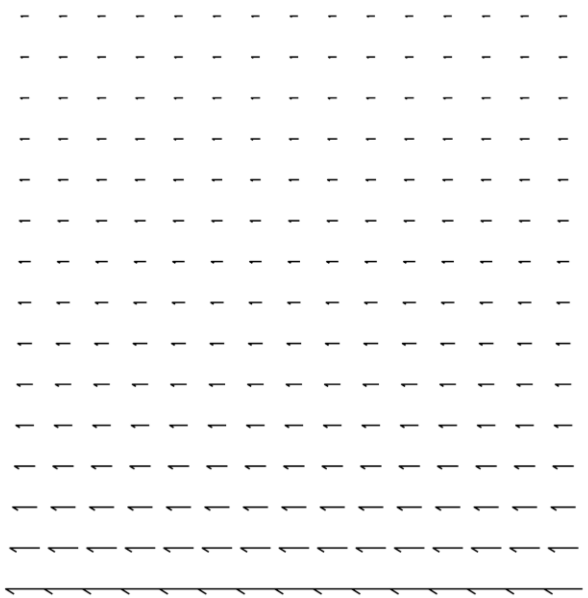
ω_y ω_x ω_z



Vection Types

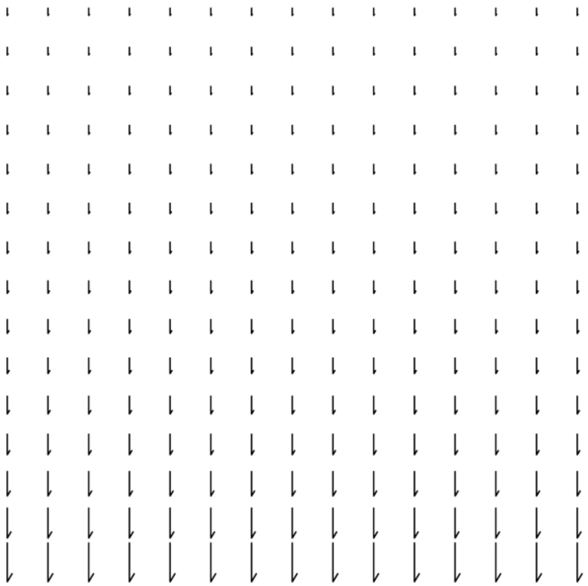
Translation

- Features with closer distances move faster (motion parallax)



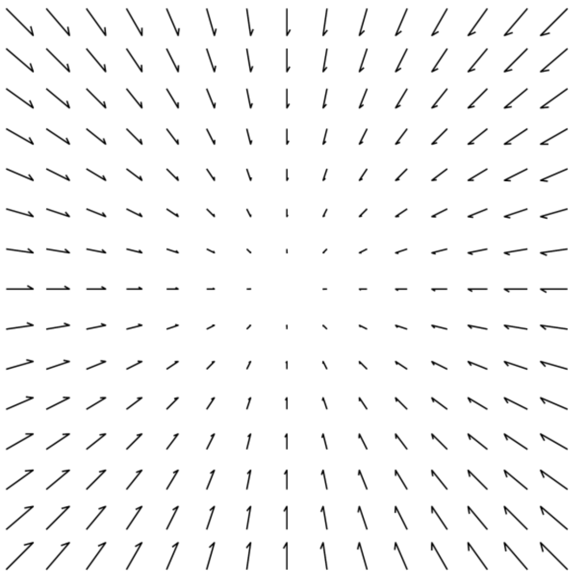
(d) lateral

v_x



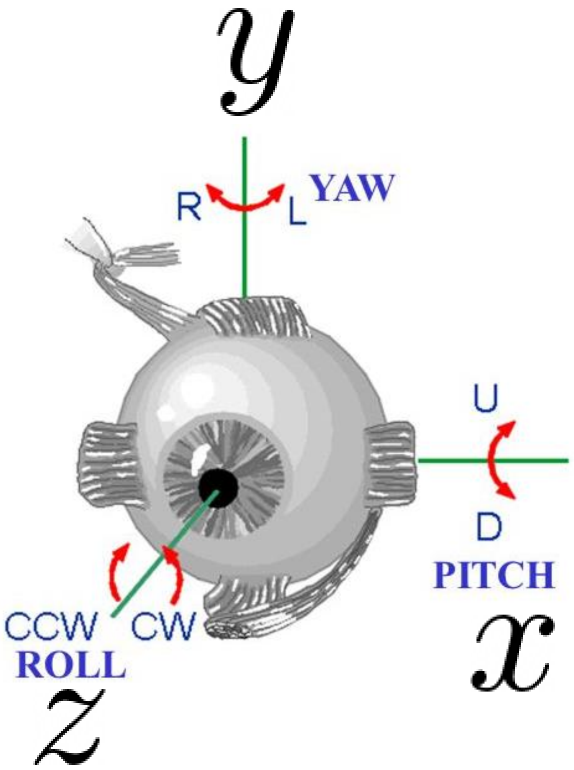
(e) vertical

v_y



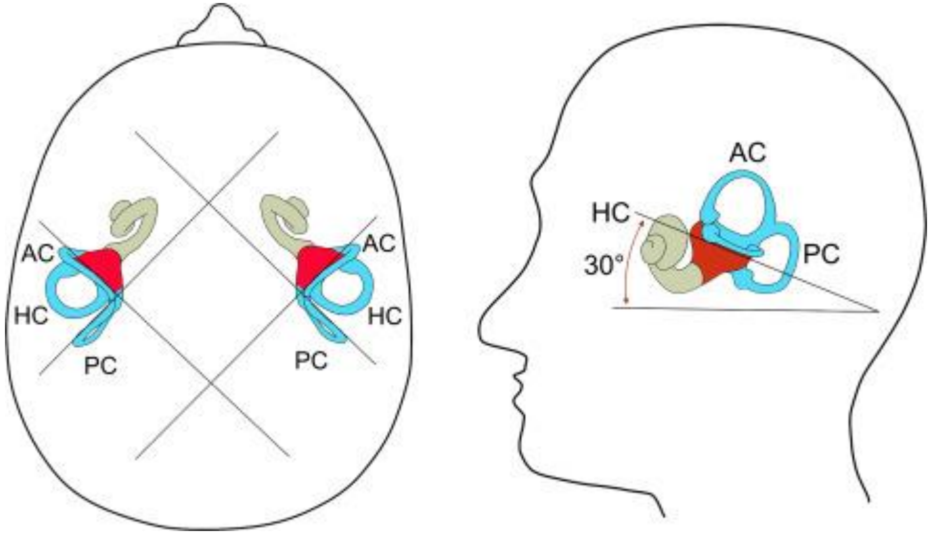
(f) forward/backward

v_z



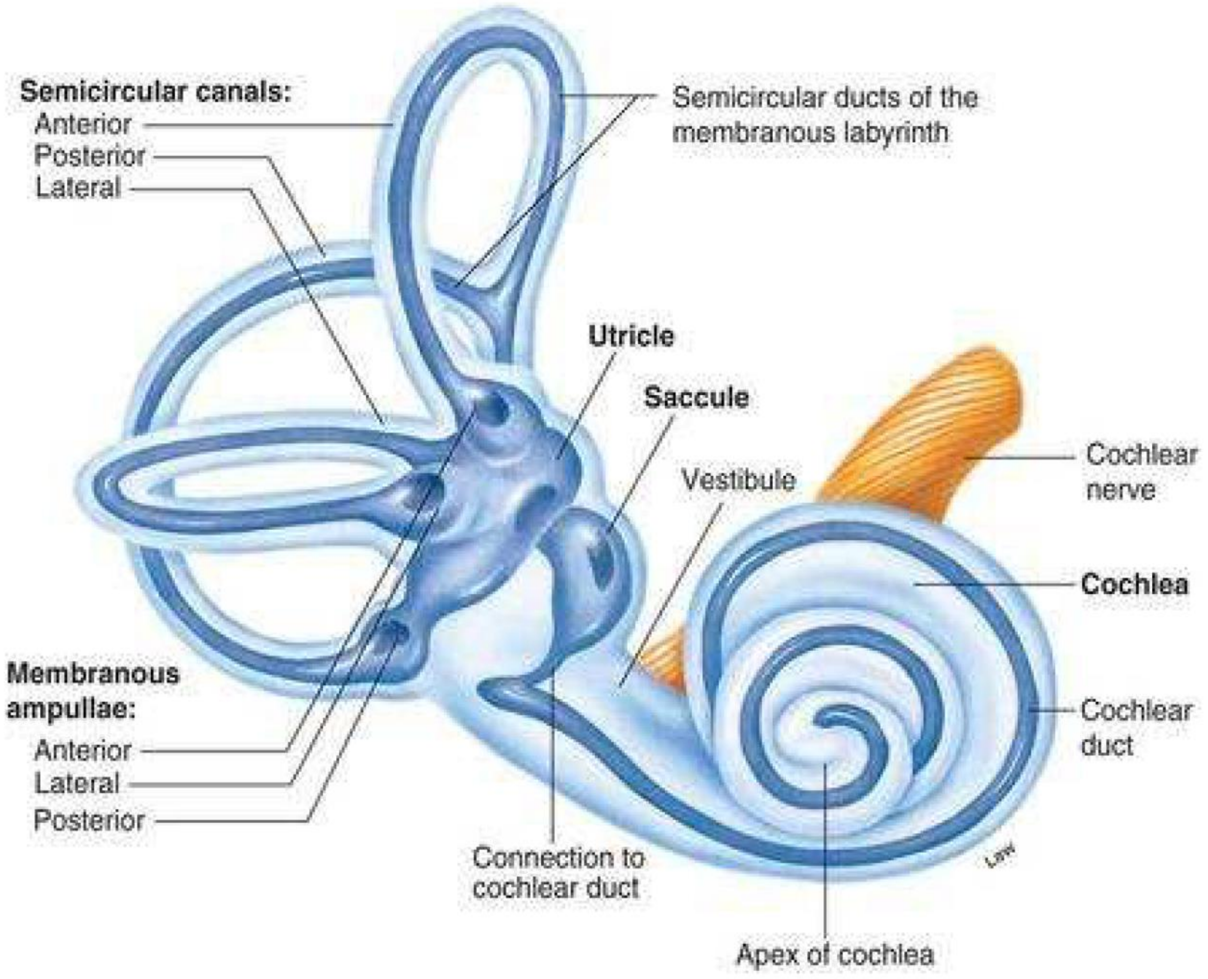
The Vestibular System

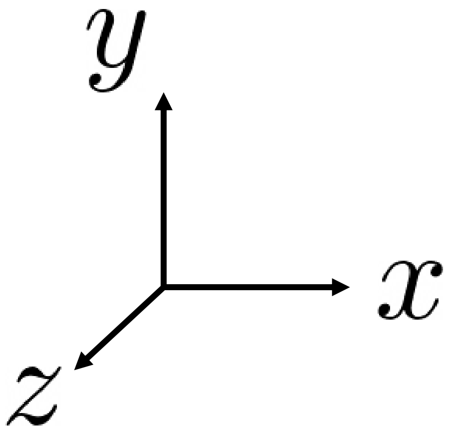
- The vestibular organs measure both linear and angular velocities of the head
- The vestibular system: vestibular organs and the associated neural pathways
  - Plays a crucial role for bodily functions that involve motion



The vestibular organs are behind the ears

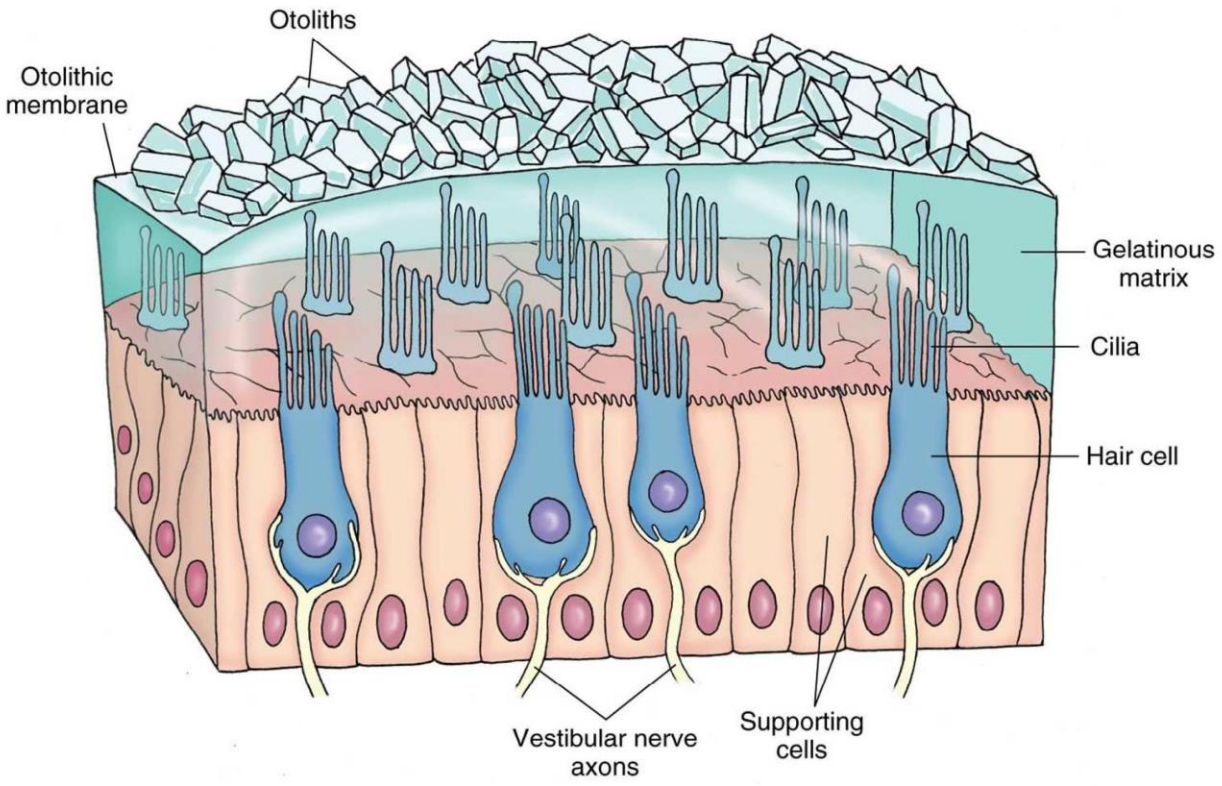
The Vestibular Organ



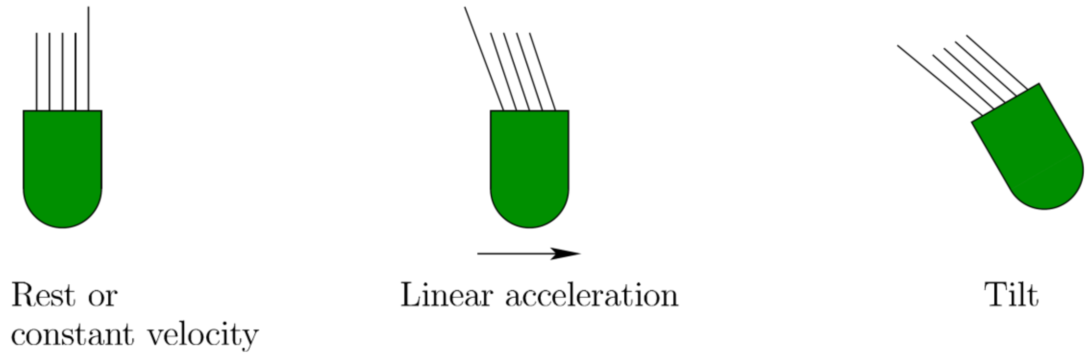


- Utricle and Saccule: measure linear acceleration
  - Utricle a_x, a_z
  - Saccule a_y, a_z
- Semicircular canals: measure angular acceleration
  - Each canal diameter 0.2~0.3mm
  - Circular arc diameter 2~3mm
  - Three canals are rough perpendicular

Sensing Linear Acceleration



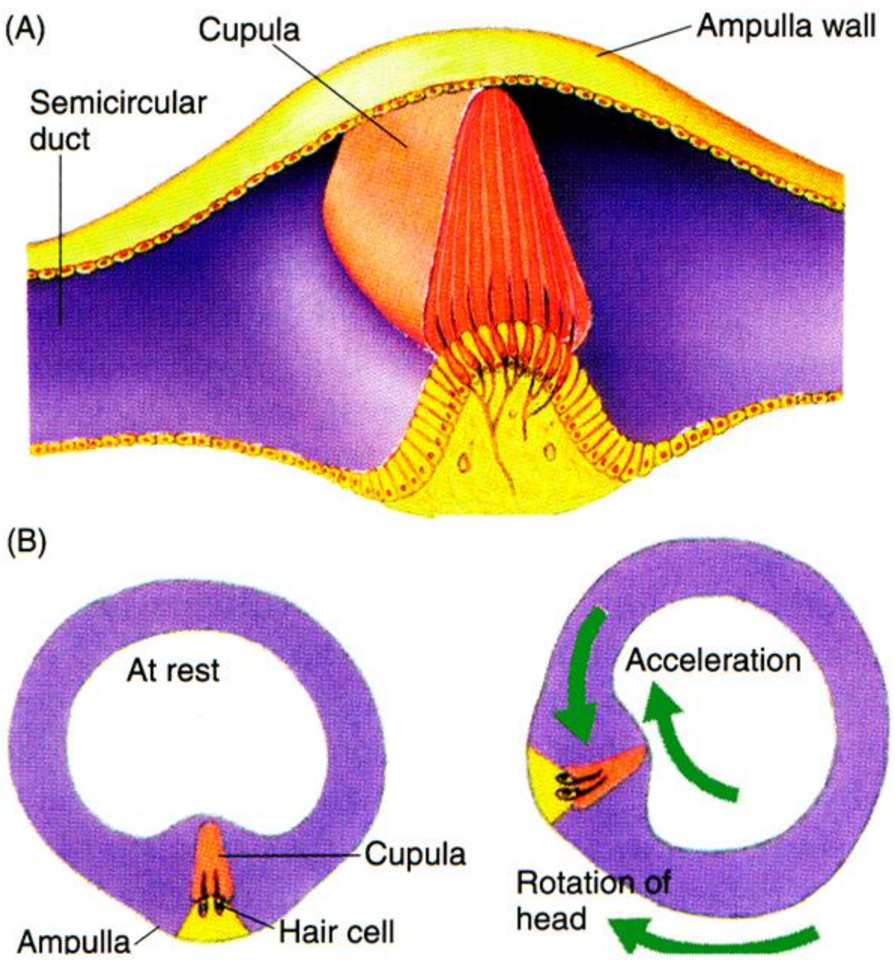
An otolith organ (utricle or saccule)

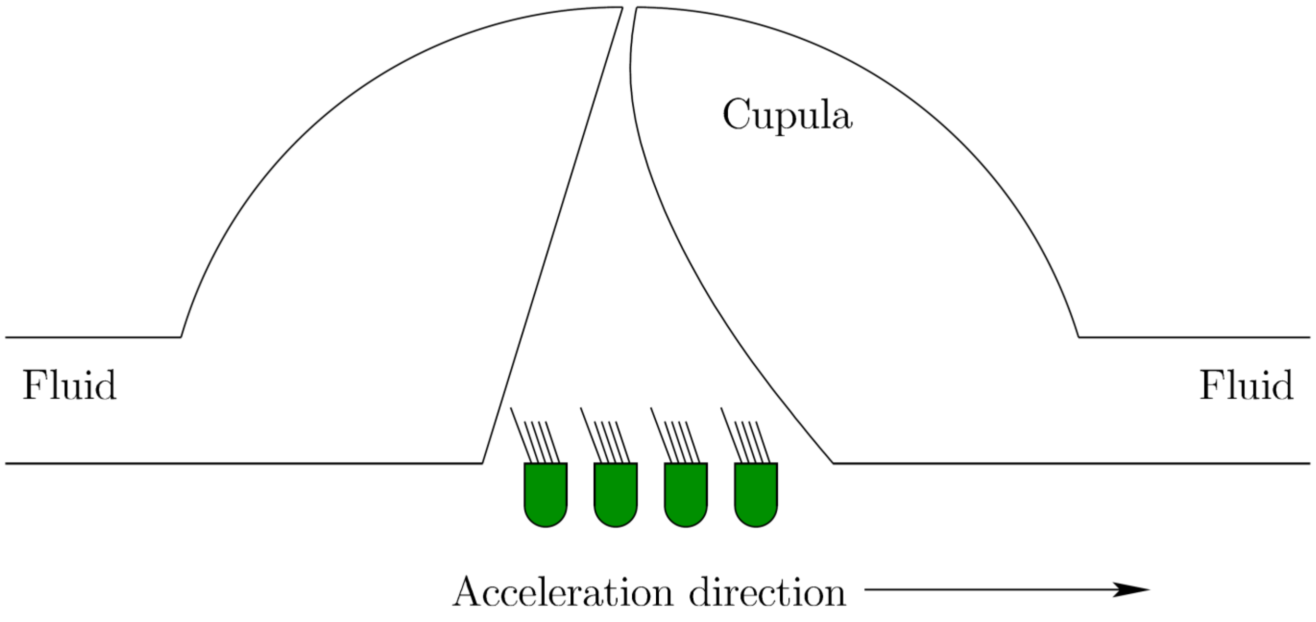


Einstein equivalent principle: the effects of gravity and true linear accelerations on a body are indistinguishable

- The same signals are sent to the brain whether the head is tilted or it is linearly accelerating
- Can use other stimuli such as vision

Sensing Angular Acceleration





https://www.d.umn.edu/~jfitzake/Lectures/DMED/InnerEar/VestibPhysiol/SemicircCanals.html

Vestibular Mismatch

- If the head is not moving, but the viewpoint is changing. The perceived acceleration from vection cause a mismatch with vestibular cues.
- Vection leads to sickness symptoms such as dizziness, nausea, and even vomiting
- We need to consider VR sickness in VR design

Further Reading

- Section 6.2, 8.2, 8.4, Virtual Reality, Steven LaValle
- Determine Constant Optical Flow, Berthold K.P. Horn
  https://people.csail.mit.edu/bkph/articles/Fixed_Flow.pdf