Visual Perception: Motion Perception

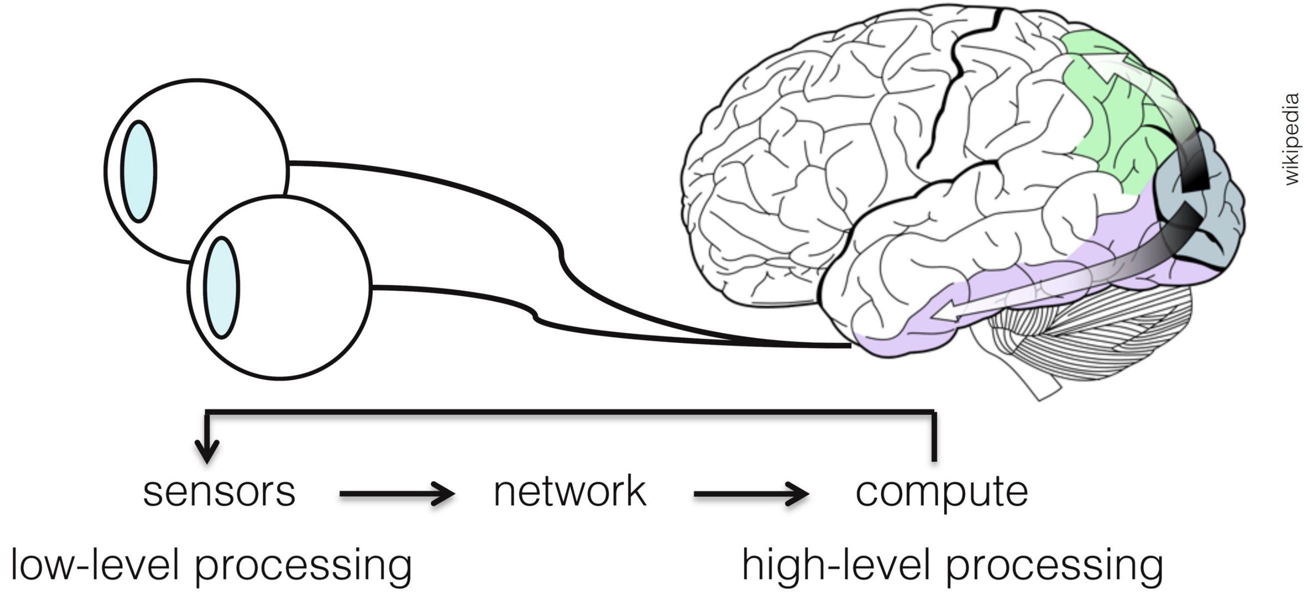
CS 6334 Virtual Reality

Professor Yapeng Tian

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

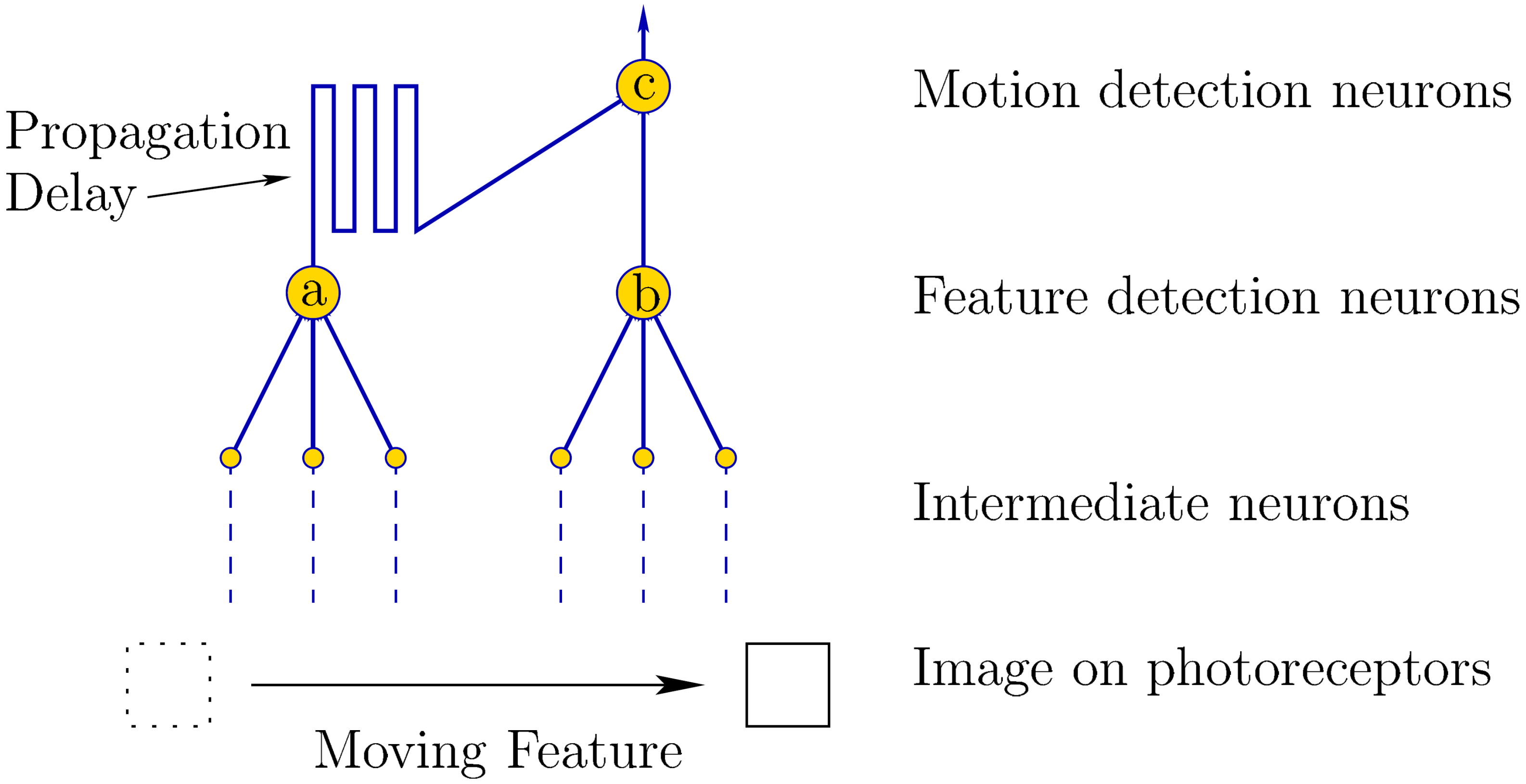
- Separate moving figure from a stationary background

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

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

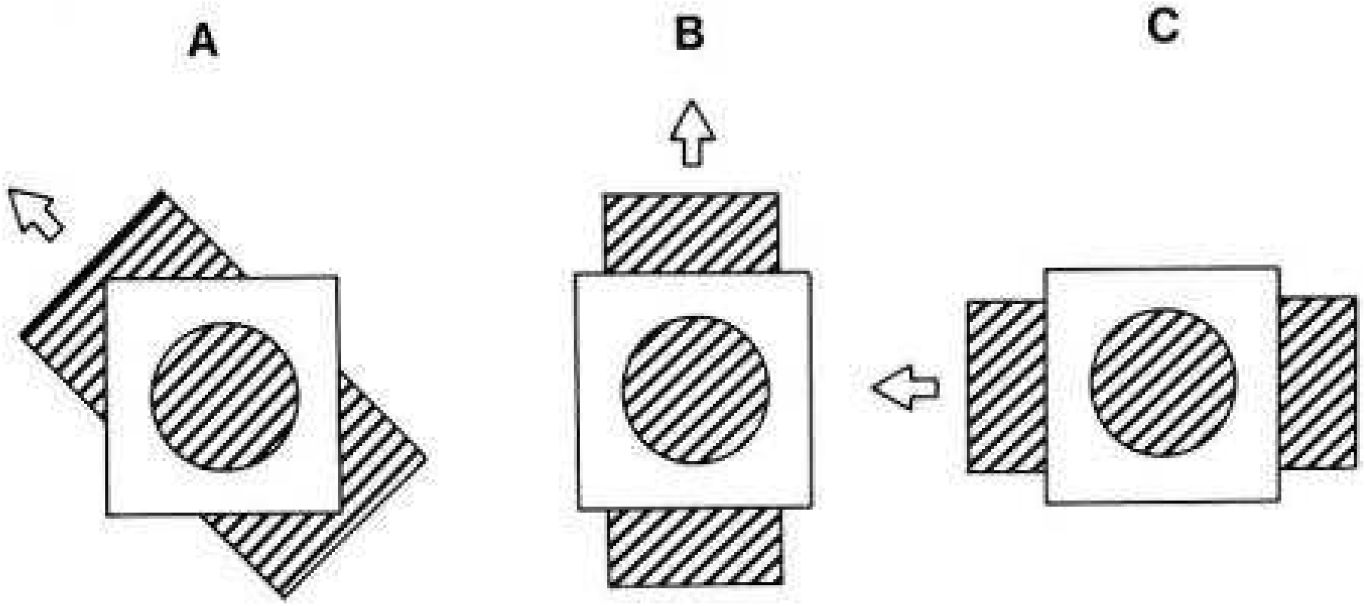
Reichardt Detector

- A neural circuitry model for motion perception

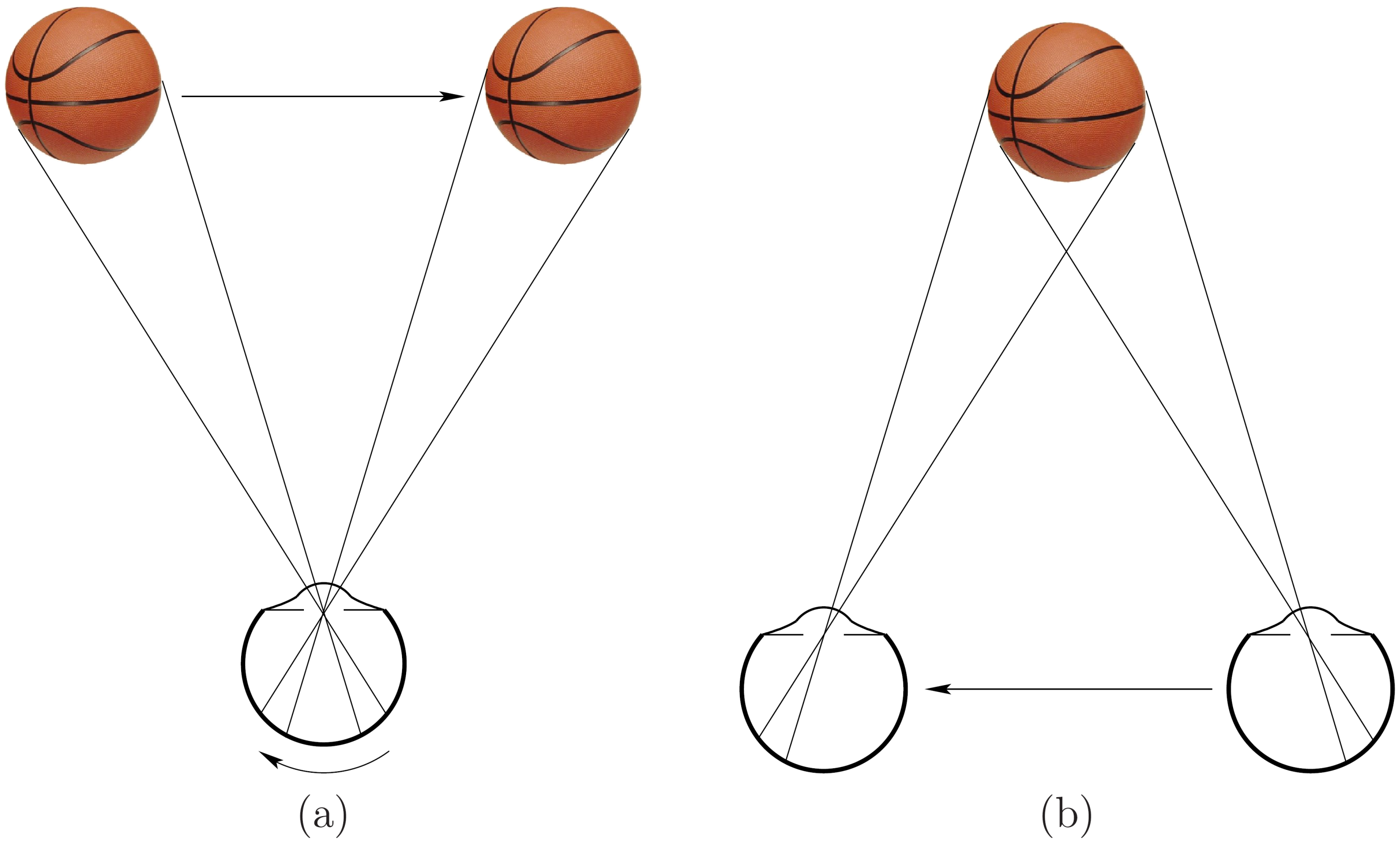


Local to Global

- 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

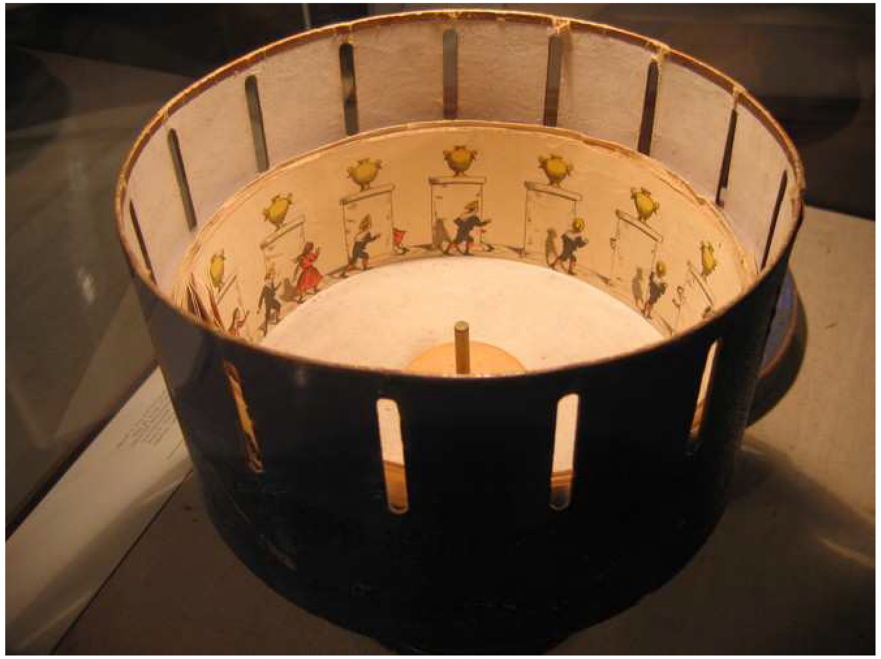
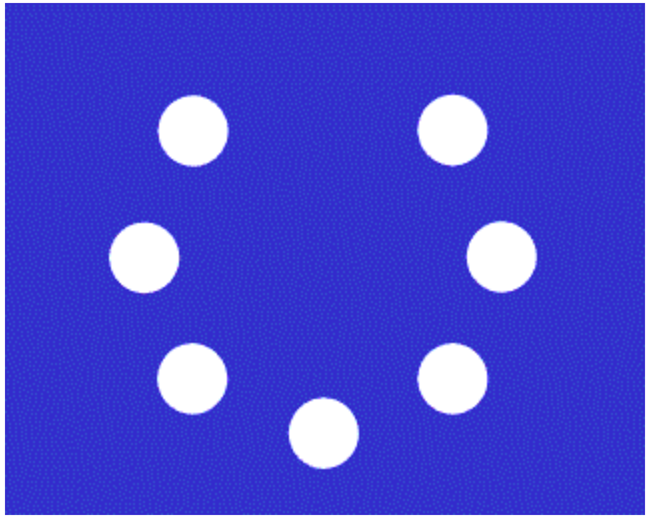


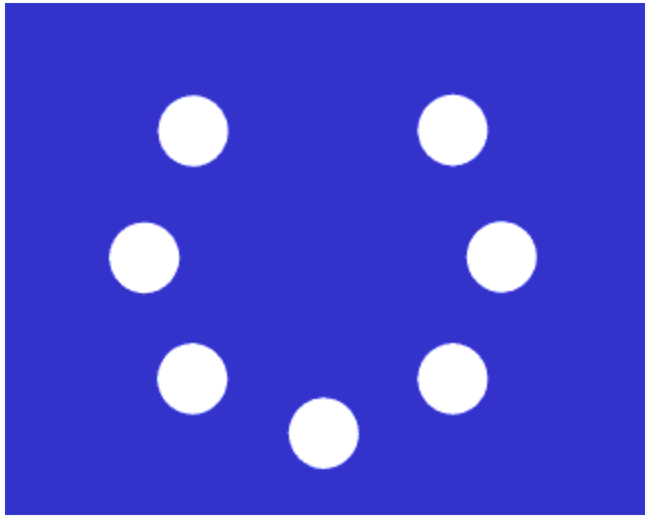
Figure 6.15: The *zoetrope* was developed in the 1830s and provided stroboscopic apparent motion as images became visible through slits in a rotating disc.

Beta Movement and Phi Phenomenon



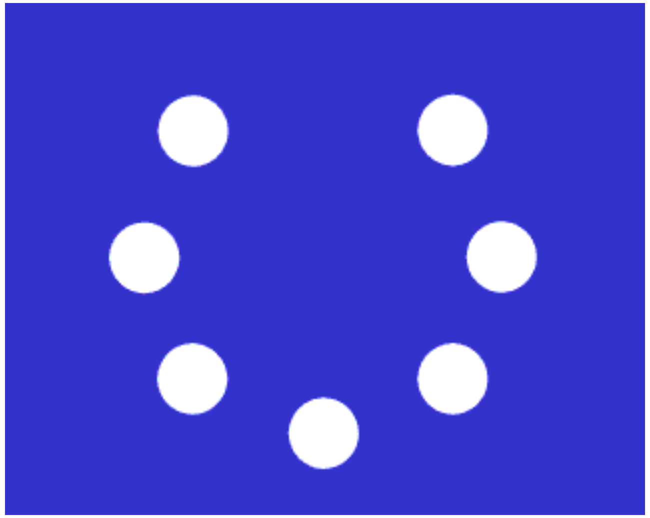
Still image

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

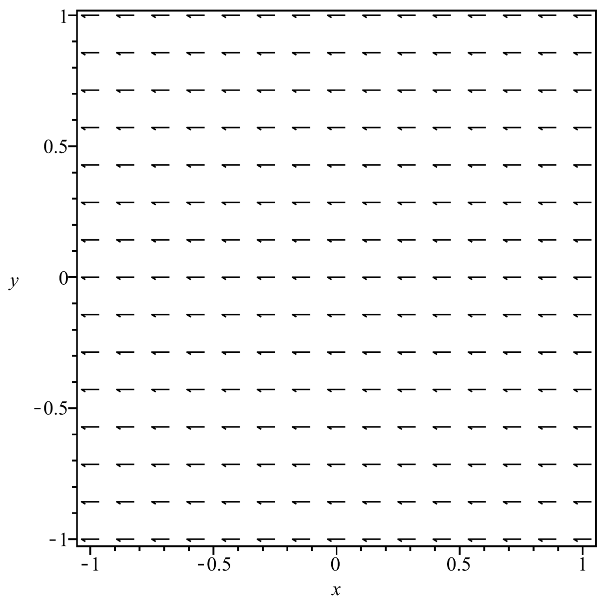
Frame Rates

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

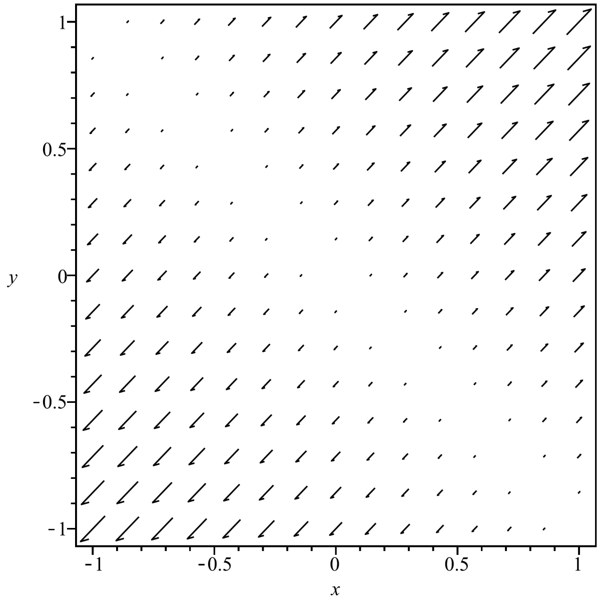
Optical Flow

- 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

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

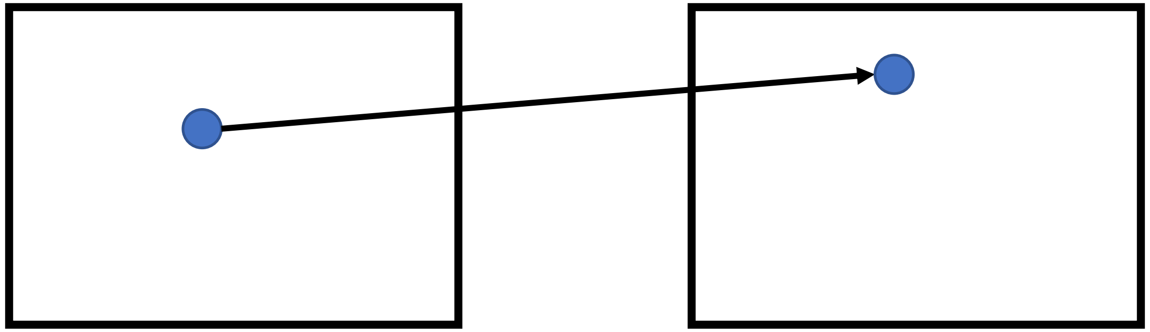


$$(x, y) \mapsto (-1, 0)$$



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

Brightness Constancy Constraint

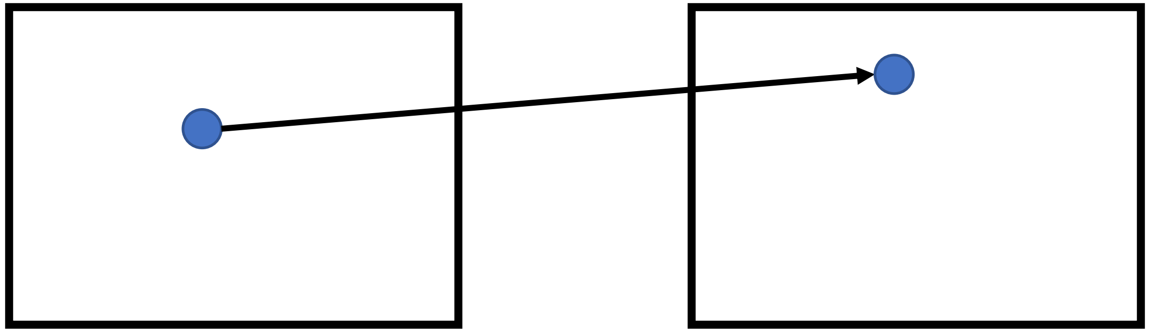


$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t)$$

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

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

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

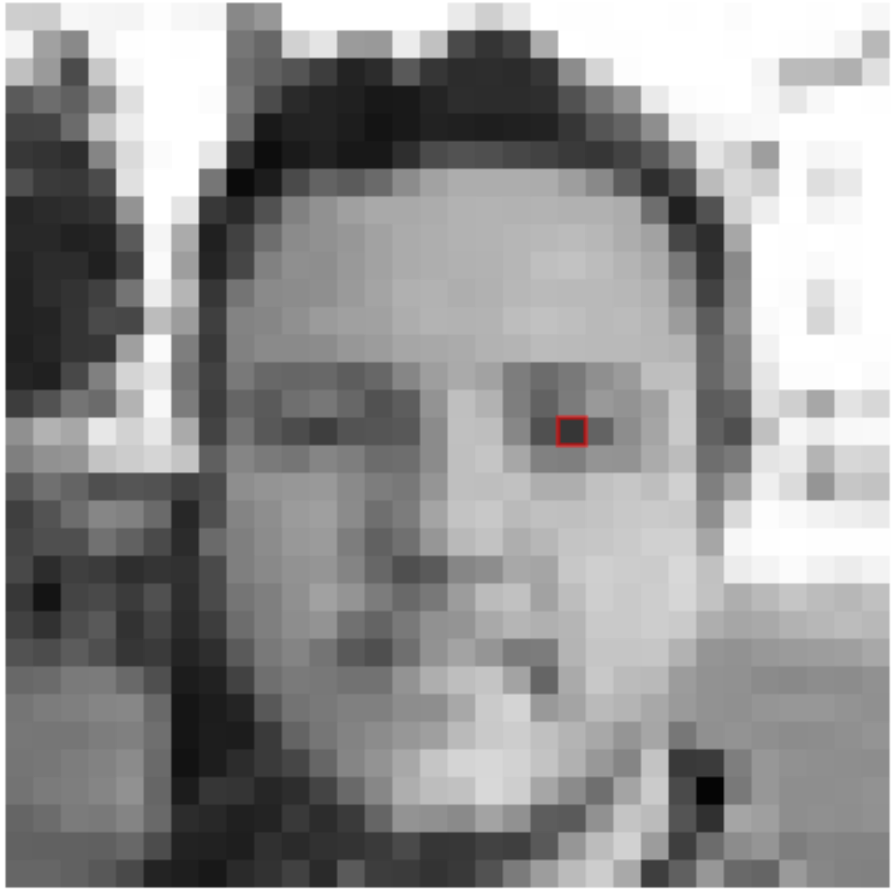
$$\boldsymbol{S}_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

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

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

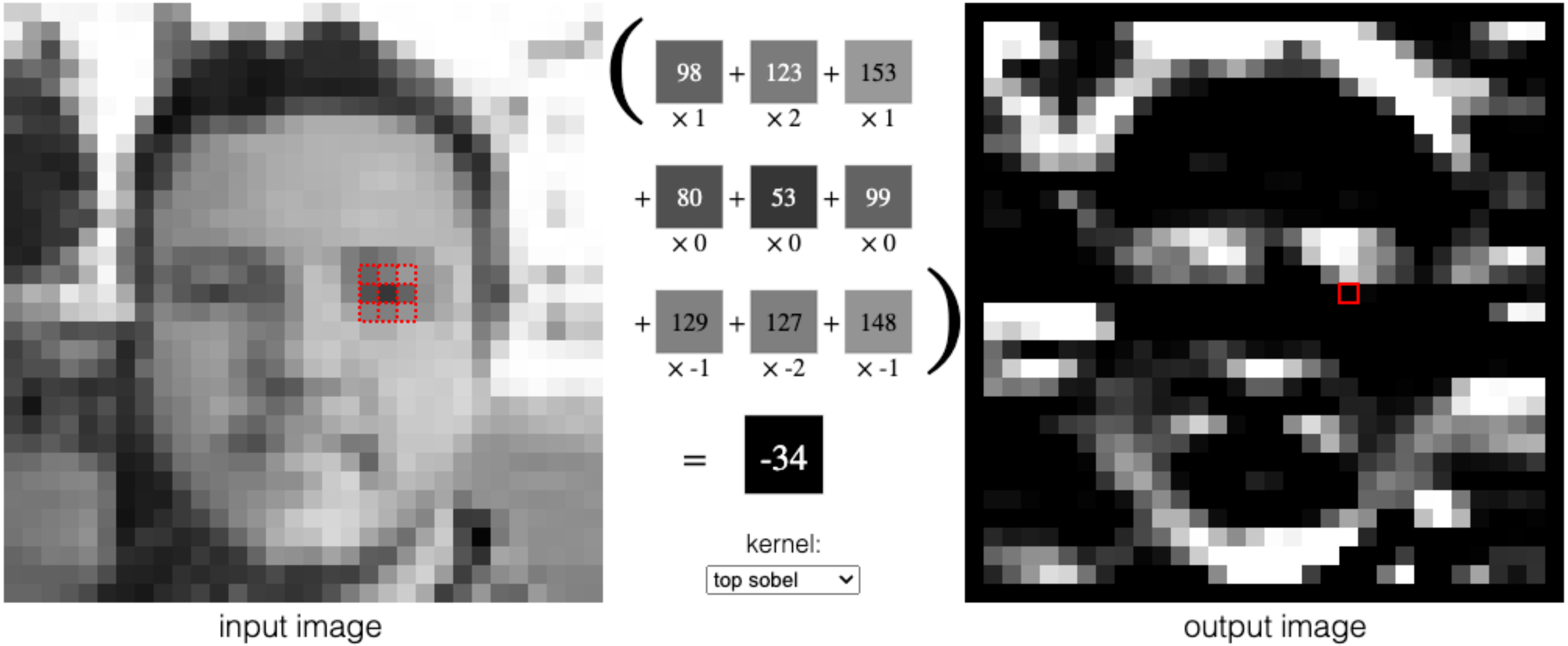
Example: Image Pixels

206	205	247	245	244	253	247	245	136	151	255	255	255	255	255	255	234	207	231	255	254	254	255	255	254	255	252	255	255	254	255	247
244	161	137	244	254	255	254	255	118	103	209	228	155	153	236	193	74	52	66	173	255	254	254	255	255	255	254	255	254	253	244	184
192	154	75	200	249	255	255	255	110	96	84	61	35	44	89	53	44	45	43	54	140	213	253	255	255	255	255	245	187	186	176	223
90	109	96	143	223	255	255	252	117	75	41	35	31	24	25	36	45	44	44	46	81	118	148	234	252	254	255	248	231	248	255	254
67	69	107	196	236	255	255	255	104	25	34	35	29	20	25	34	32	30	32	34	53	85	100	142	231	242	247	249	255	255	255	255
55	51	45	133	218	251	255	232	51	12	26	33	24	24	46	75	82	78	71	66	58	53	67	90	136	228	208	158	253	246	249	255
79	58	56	75	224	255	255	118	11	27	74	99	91	106	140	162	173	173	173	172	158	137	92	46	78	187	217	206	254	222	233	255
38	43	47	52	147	255	228	56	41	81	129	145	160	169	169	172	178	179	178	179	177	177	172	110	31	82	209	238	255	244	249	255
40	40	33	36	90	245	171	32	65	110	139	145	151	162	171	174	178	179	182	184	187	183	173	162	71	45	167	255	254	255	254	255
37	44	44	31	69	250	158	36	70	129	143	142	153	162	171	175	177	178	182	191	194	188	180	170	120	51	137	255	254	250	254	255
34	45	51	64	116	237	181	53	116	138	140	143	154	164	176	178	174	177	183	186	185	185	183	178	140	66	141	254	252	225	249	255
34	36	52	74	71	188	156	63	131	134	144	155	160	161	173	179	178	179	189	193	190	185	187	182	156	93	148	250	254	214	247	255
32	38	52	54	159	250	126	57	129	138	138	140	151	156	166	168	171	178	180	187	186	185	185	183	180	102	136	242	255	255	254	254
36	32	72	128	212	228	115	65	121	104	102	104	94	103	134	158	170	162	125	108	121	143	155	190	191	104	134	230	253	253	255	251
61	82	116	107	179	247	124	60	101	90	111	119	103	81	94	147	191	178	126	98	123	153	147	161	200	92	100	222	207	167	227	215
144	178	167	231	210	232	170	67	115	88	76	62	83	85	88	139	192	190	135	80	53	99	141	165	201	97	79	192	245	235	248	249
127	145	149	195	204	213	196	95	133	122	117	133	126	108	110	139	191	197	167	129	127	148	147	171	188	110	121	228	233	180	215	212
87	112	100	79	85	82	65	75	142	148	151	153	138	125	120	149	191	190	193	175	174	193	198	190	208	127	163	239	219	149	198	194
63	83	109	134	129	106	39	78	132	142	155	159	139	111	124	164	195	200	186	192	191	195	200	202	200	143	217	253	249	242	238	234
69	78	78	113	97	74	43	106	127	140	152	155	125	97	112	150	185	194	174	183	196	198	202	208	209	166	247	254	255	254	254	254
72	44	63	59	46	52	49	74	127	137	146	149	132	103	78	90	134	141	168	165	199	207	204	203	216	193	236	244	251	242	236	243
55	20	69	73	59	80	46	74	117	127	144	161	148	124	105	120	156	187	193	162	189	206	201	205	214	194	174	185	197	188	183	192
65	49	77	89	50	68	43	61	109	127	141	147	113	100	121	145	148	169	181	176	181	201	201	205	202	174	166	169	178	183	188	184
82	76	92	79	54	58	37	47	90	121	132	116	89	78	111	146	163	149	122	124	180	197	197	198	178	149	146	152	155	157	159	168
104	107	122	123	105	79	27	33	66	111	122	120	114	114	147	175	190	196	163	101	170	200	187	185	156	146	145	139	137	141	140	145
117	124	127	133	135	105	21	28	37	88	115	121	128	128	141	142	168	202	212	153	164	186	180	168	154	146	144	149	151	151	147	144
119	118	118	125	128	111	21	29	28	58	100	118	131	140	151	159	186	201	205	192	180	168	149	166	119	144	147	143	140	141	144	148
117	119	125	130	139	106	18	29	44	58	70	102	133	147	168	197	212	215	210	195	177	152	133	195	57	59	126	151	145	143	142	141
115	123	126	133	145	102	27	54	52	38	45	69	105	135	175	189	193	216	206	166	139	111	164	203	74	5	121	151	142	142	143	146
101	108	123	121	132	105	44	40	31	35	57	44	58	101	147	144	138	163	145	94	90	145	196	187	84	48	165	159	142	144	142	145
98	97	97	96	104	76	34	33	30	48	41	49	51	58	74	53	55	66	63	89	150	188	209	156	62	108	140	149	125	133	131	131
102	102	97	88	73	35	30	23	42	50	65	41	90	60	59	51	57	82	123	157	187	205	169	62	96	151	105	101	154	135	130	129



https://setosa.io/ev/image-kernels/

Example: Applying a Sobel Filter



https://setosa.io/ev/image-kernels/

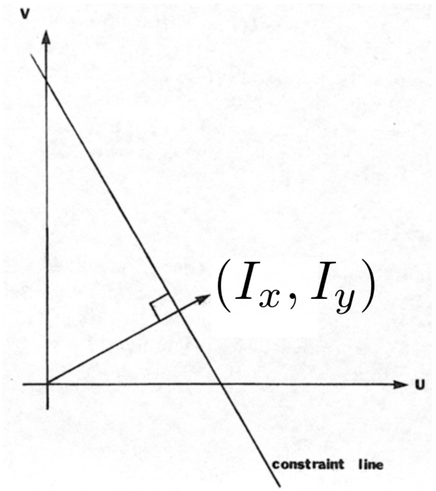
Brightness Constancy Constraint

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

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

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

Lucas-Kanade Method

- Solve the least squares problem

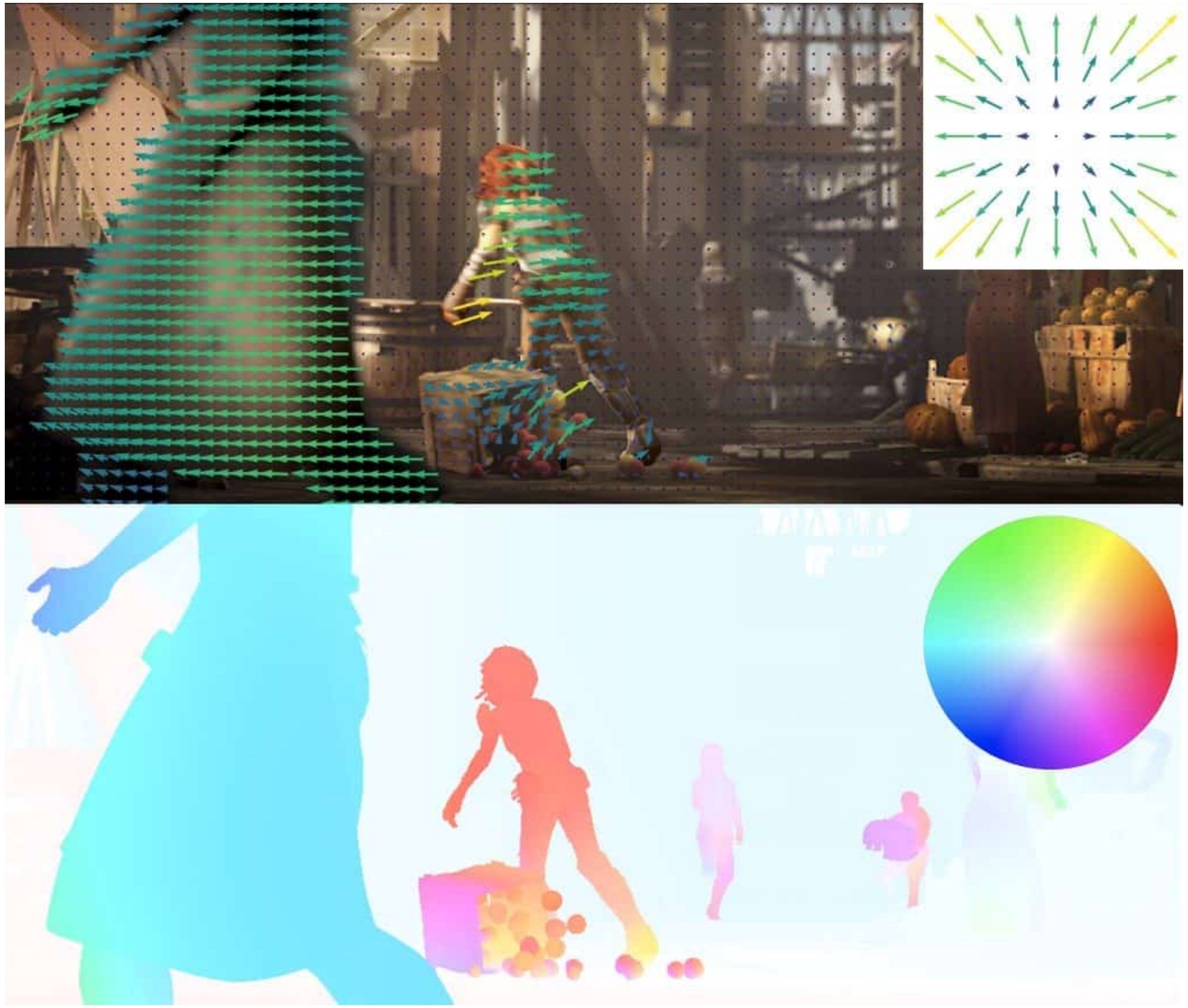
$$\underset{25\text{x}2}{A}\ \underset{2\text{x}1}{d} = \underset{25\text{x}1}{b} \longrightarrow \text{minimize } \|Ad - b\|^2$$

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

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

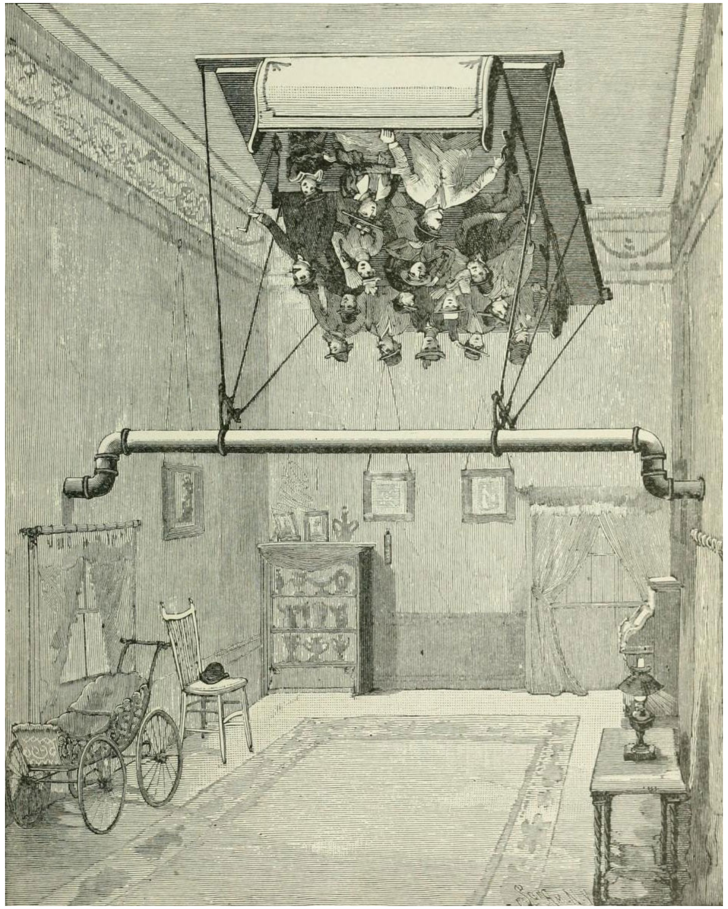
https://en.wikipedia.org/wiki/Proofs_involving_ordinary_least_squares#Least_squares_estimator_for_.CE.B2

Optical Flow Example

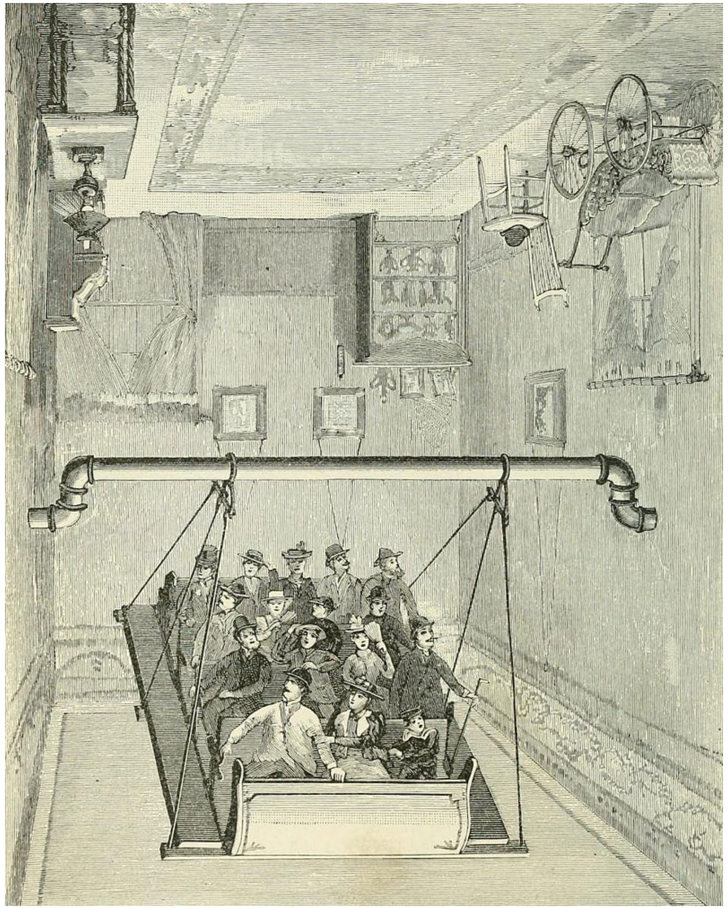


Vection

- 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



Perspective of riders



Actual swing position

VR motion sickness

- Industry leaders often proclaim that their latest VR headset has beaten the VR sickness problem.
- However, if a headset is better, the potential is higher for making people sick through vection and other mismatched cues.
- If the headset more accurately mimics reality, then the sensory cues are stronger, and our perceptual systems become more confident about mismatched cues.

Further Reading

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