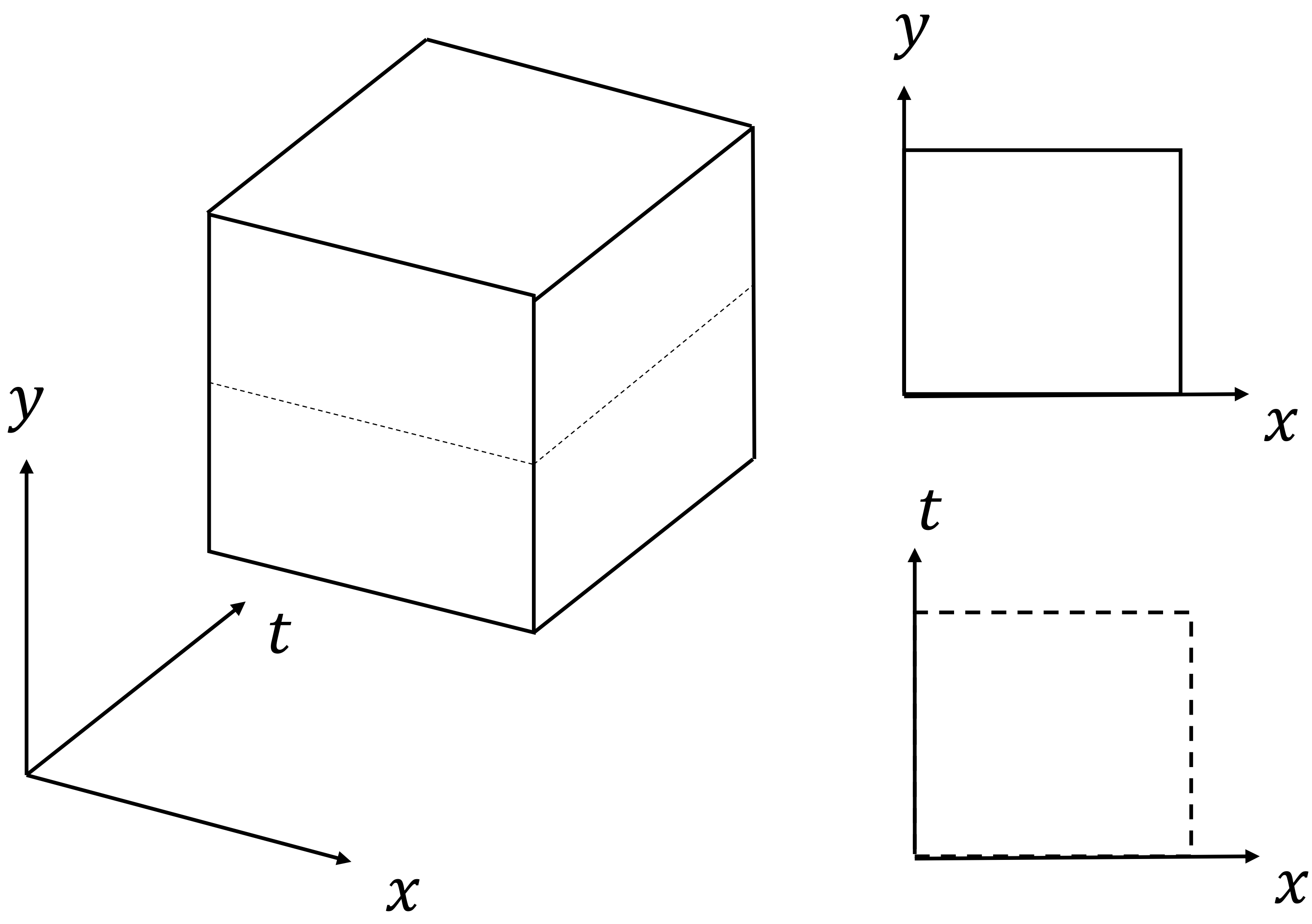
COMP 546

Lecture 7

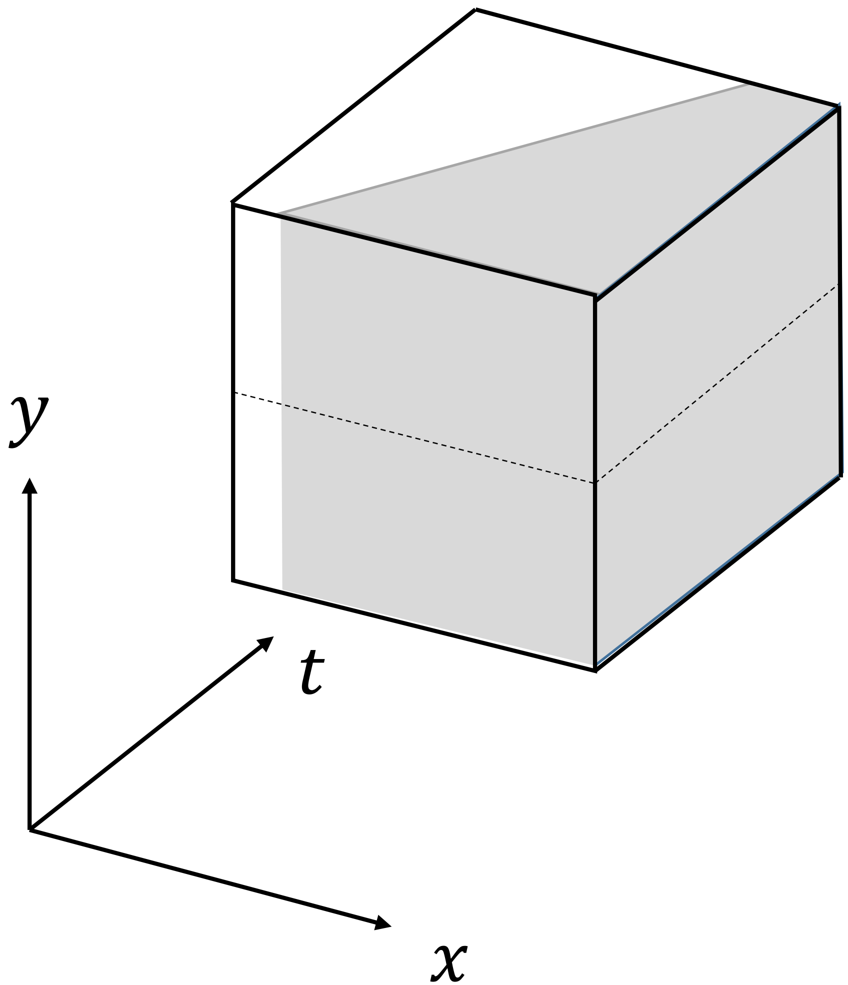
image motion

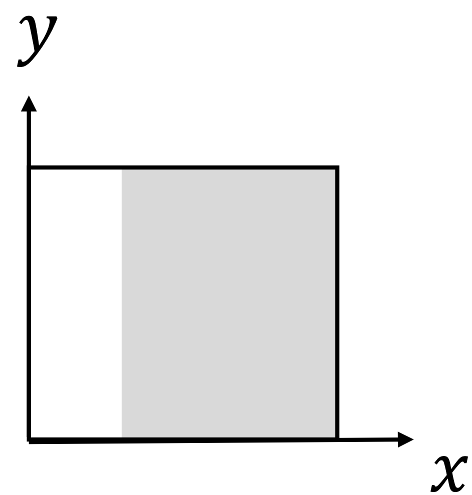
Thurs. Jan. 31, 2018

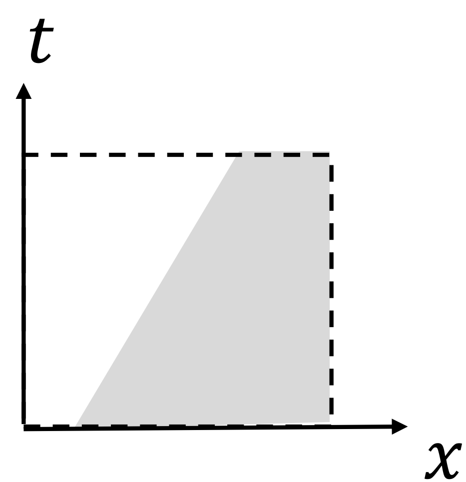
Time varying images (XYT)



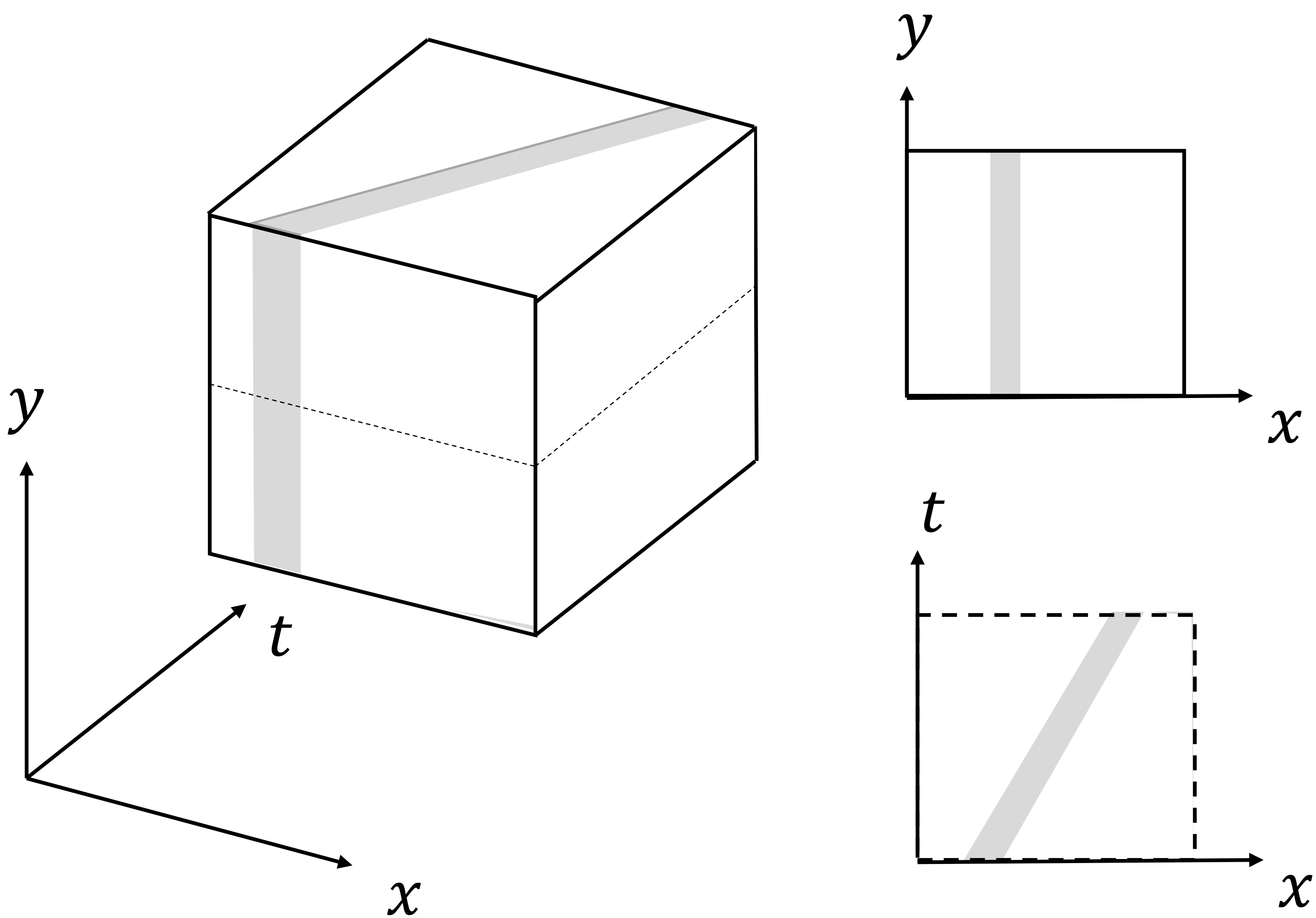
Motion in XYT
e.g. translating vertical edge

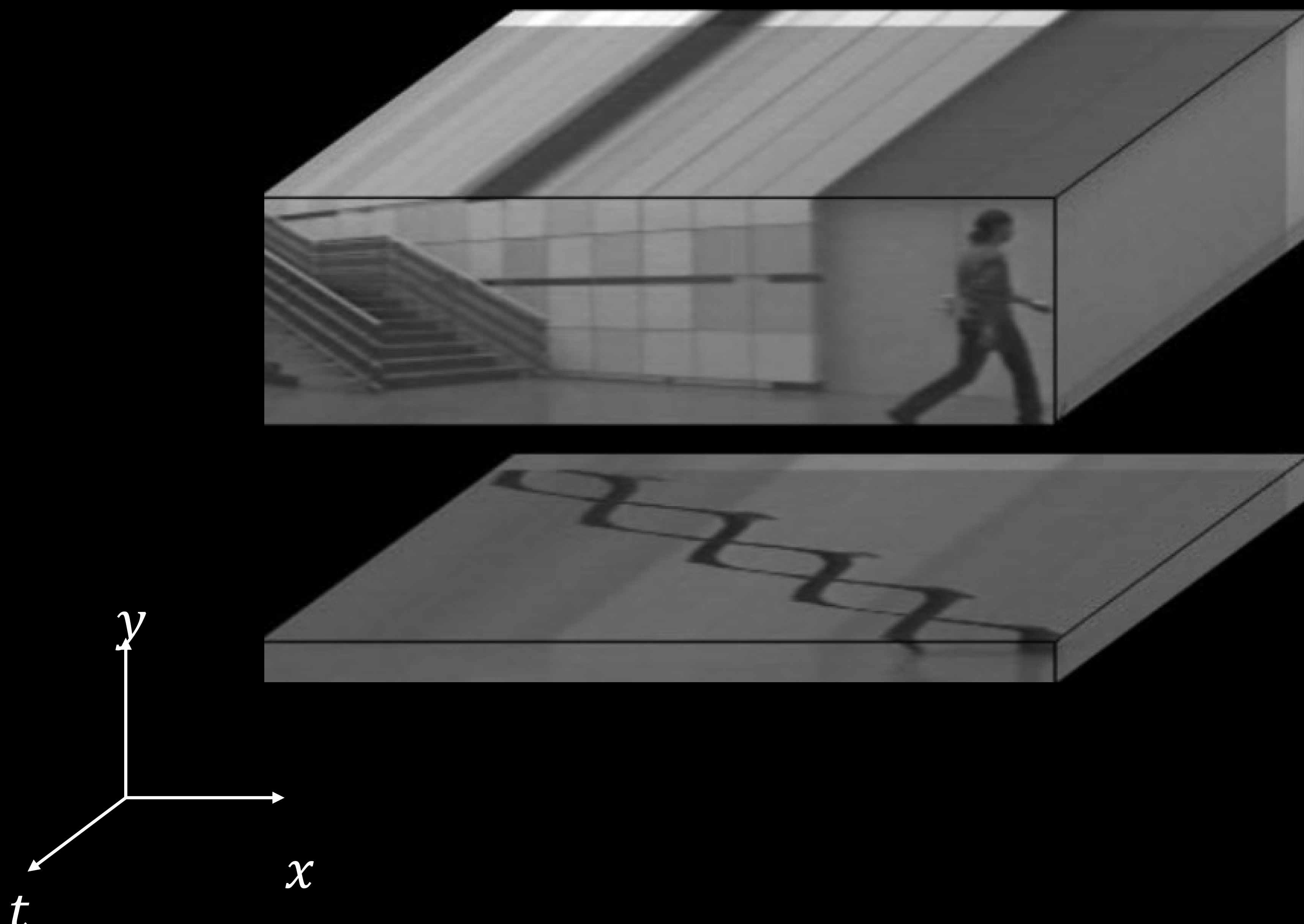






Motion in XYT
e.g. translating vertical bar

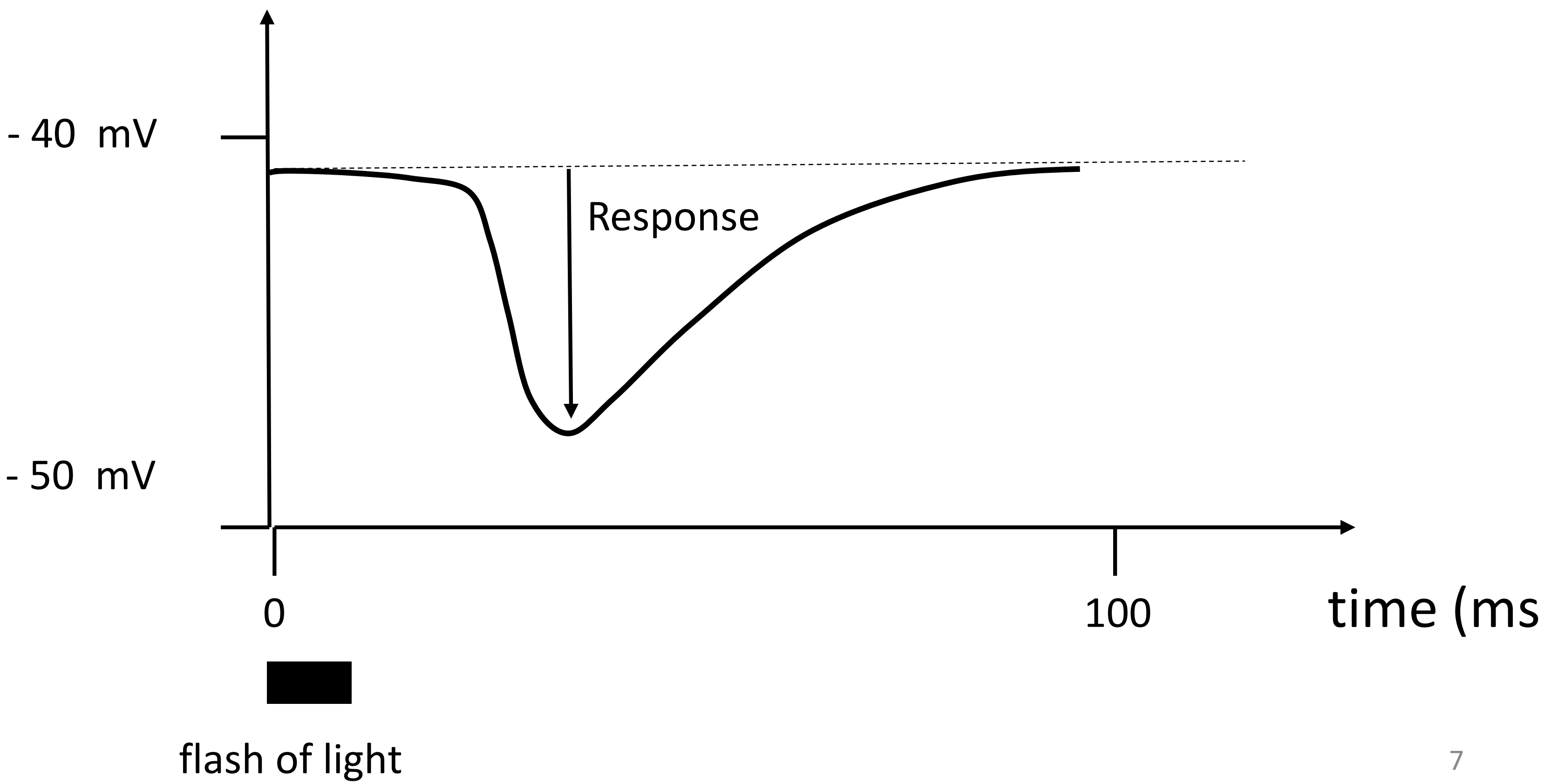



y
x
t

How do the eye and brain
(retina, LGN, V1)
measure image motion?

Photoreceptor response to a *brief flash of light*

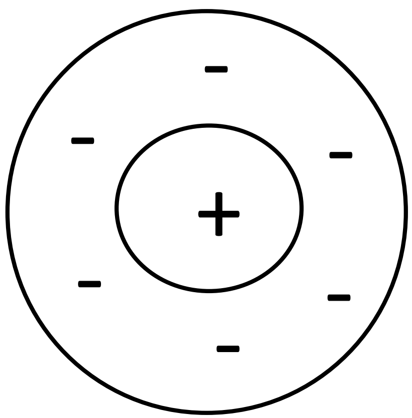
(recall from lecture 3)

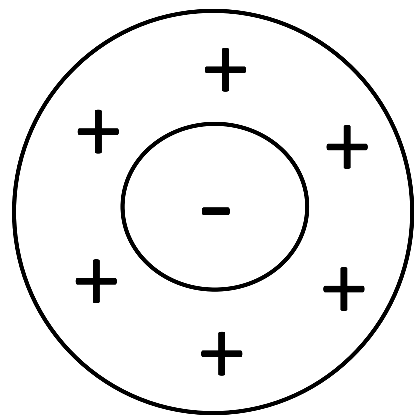


Retinal Ganglion and LGN cells

Our models up to now have been static only.

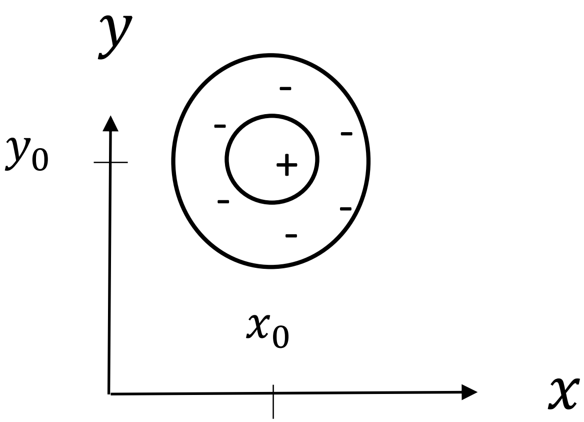
ON center,
OFF surround

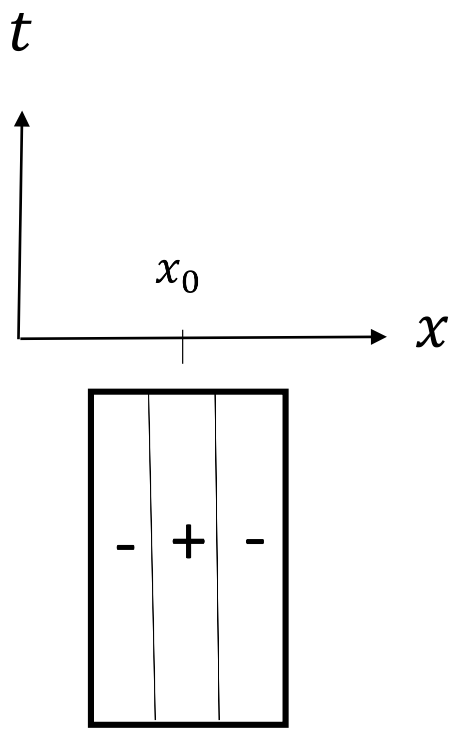




OFF center,
ON surround

XY and XT slices through DOG(x,y,t)

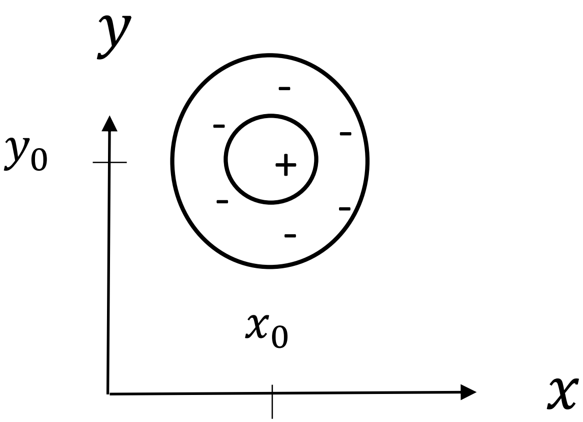


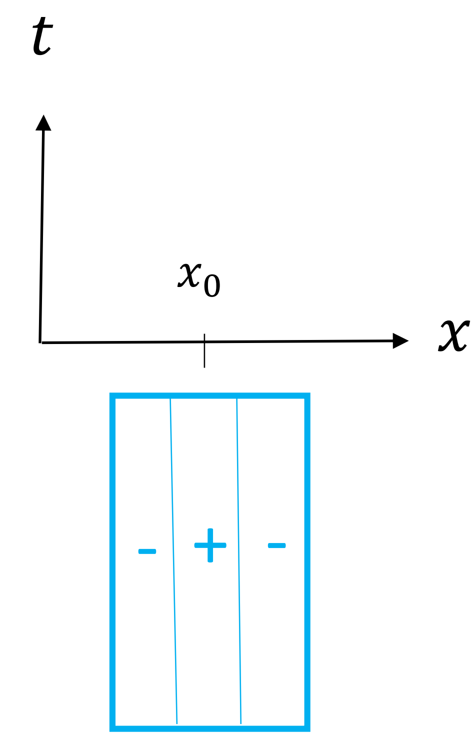


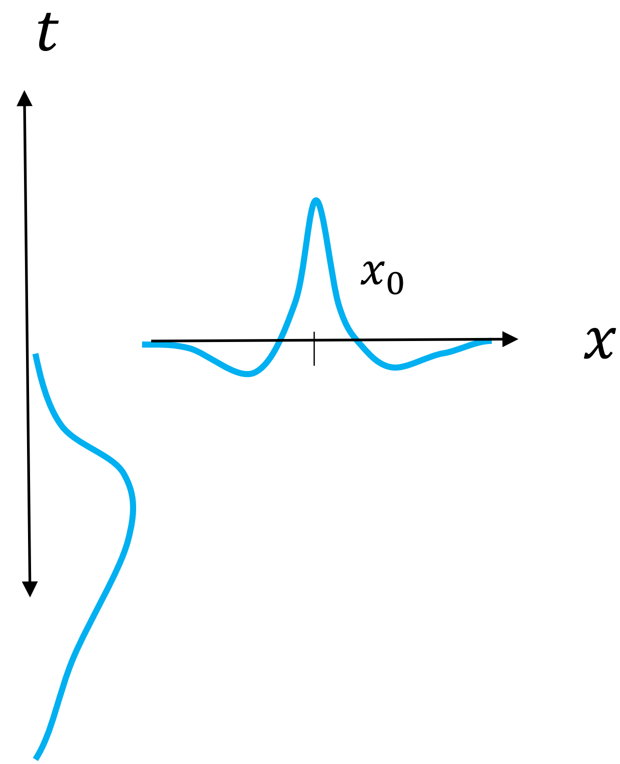
The response at t = 0 depends on the image intensities in the past (t < 0).

Space-time separable model

$g(x, y, t) = DOG(x, y)f(t)$

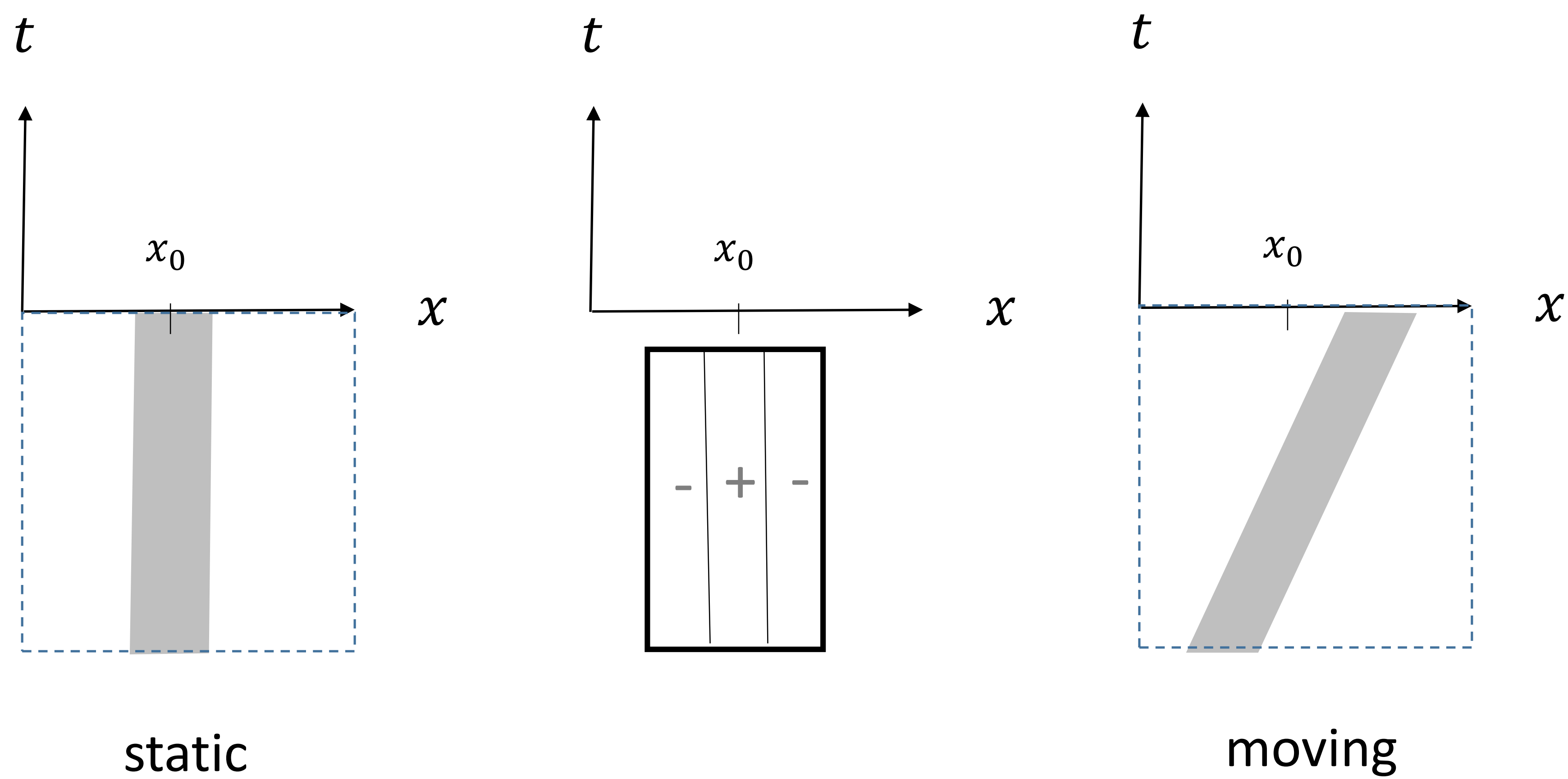




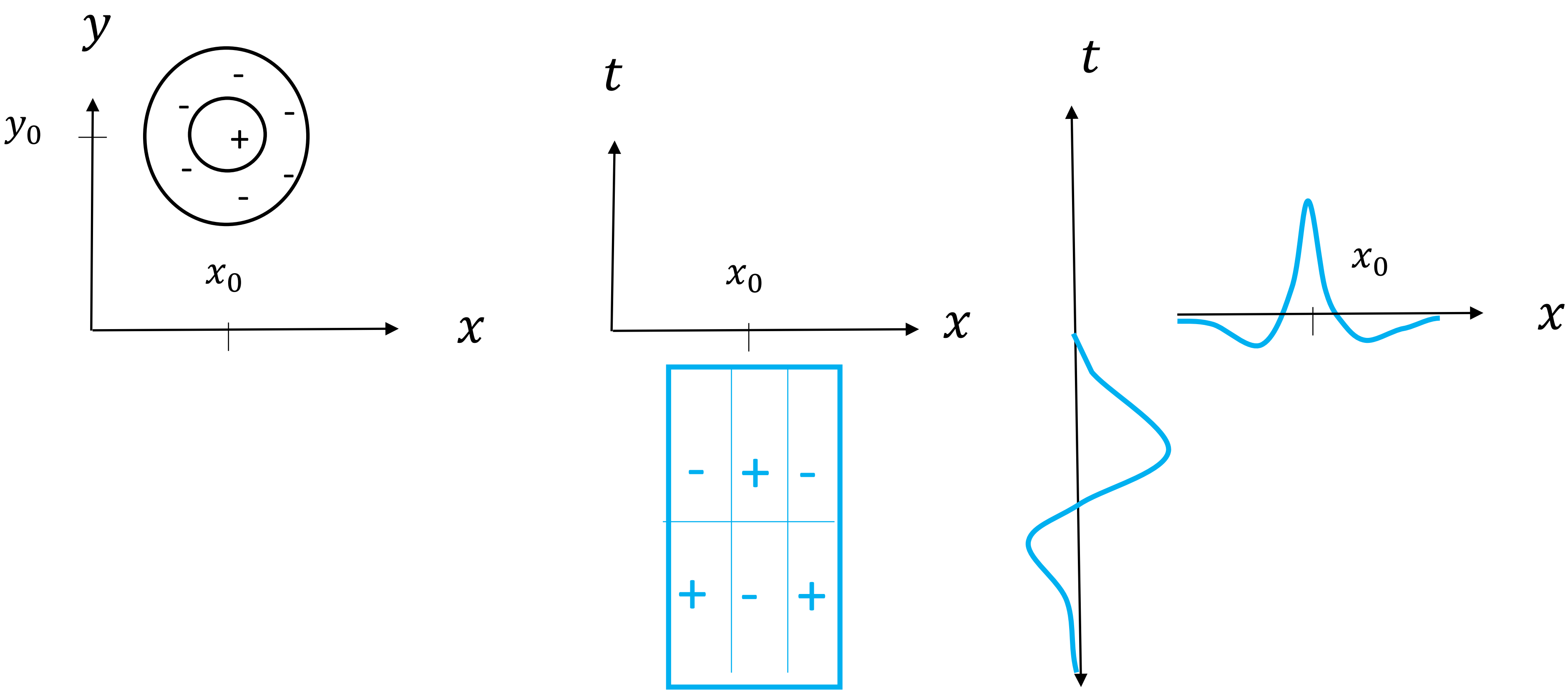


The response at t = 0
depends on the image
intensities in the past (t < 0).

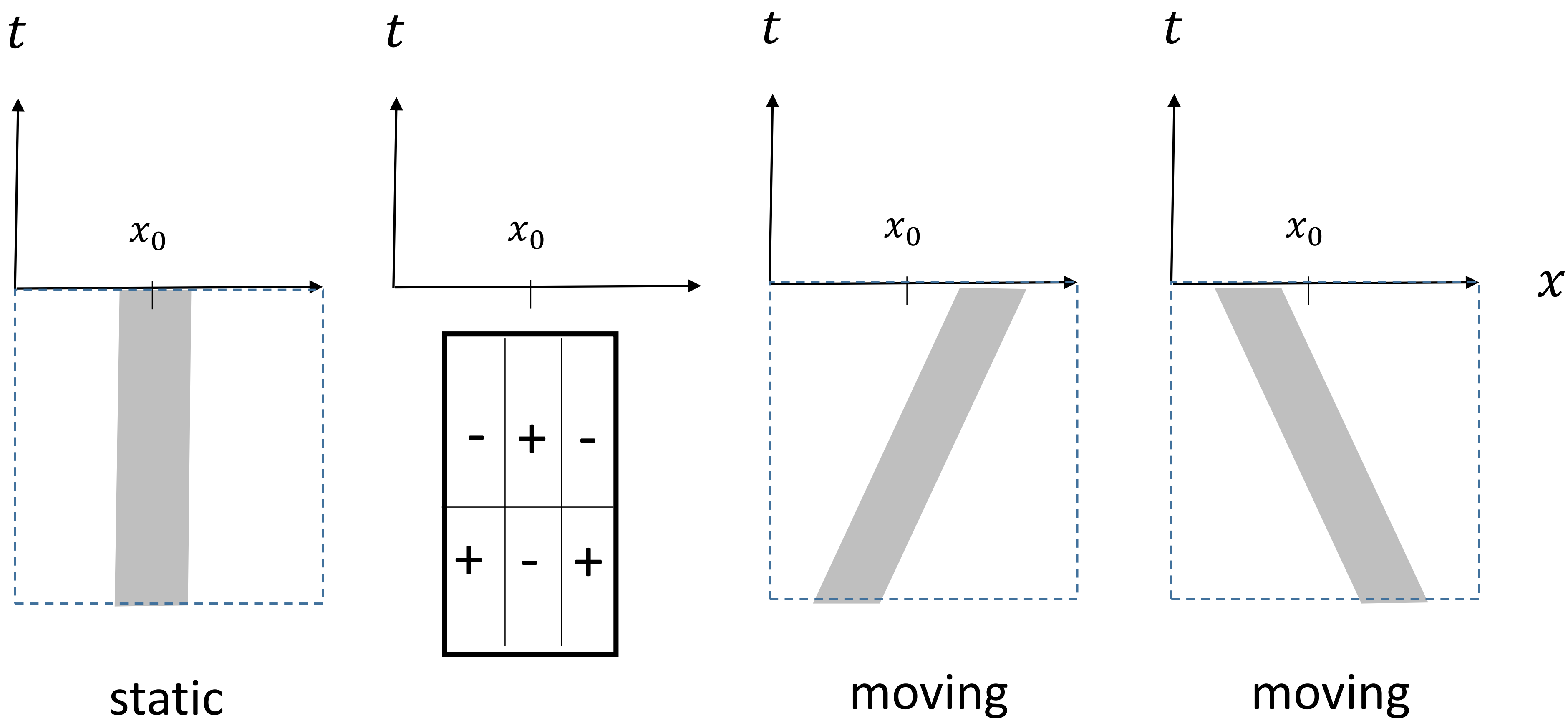
This cell would respond better to the static image.



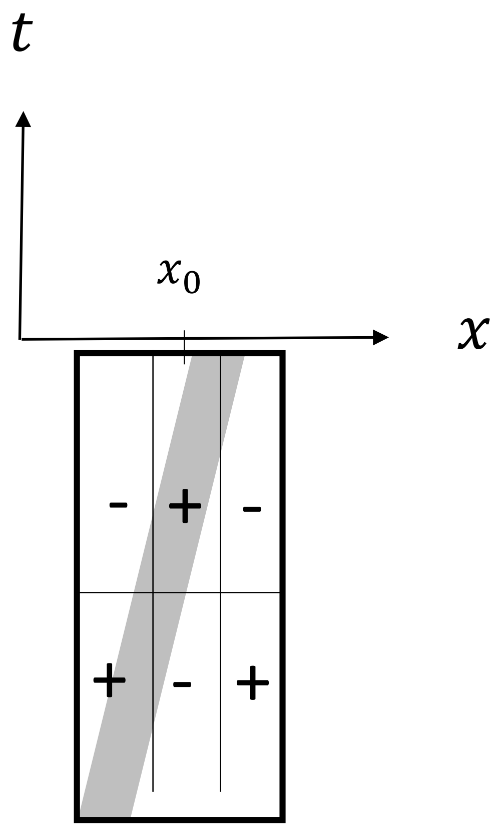
Space-time separable model (with temporal sensitivity)



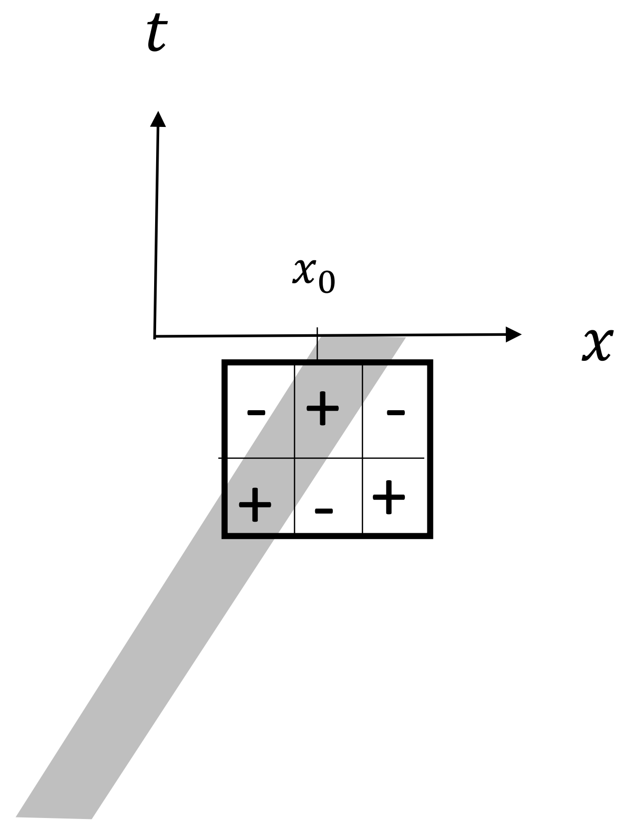
This cell would respond better to motion than to static image.



This cell would respond better to *slow* motion.



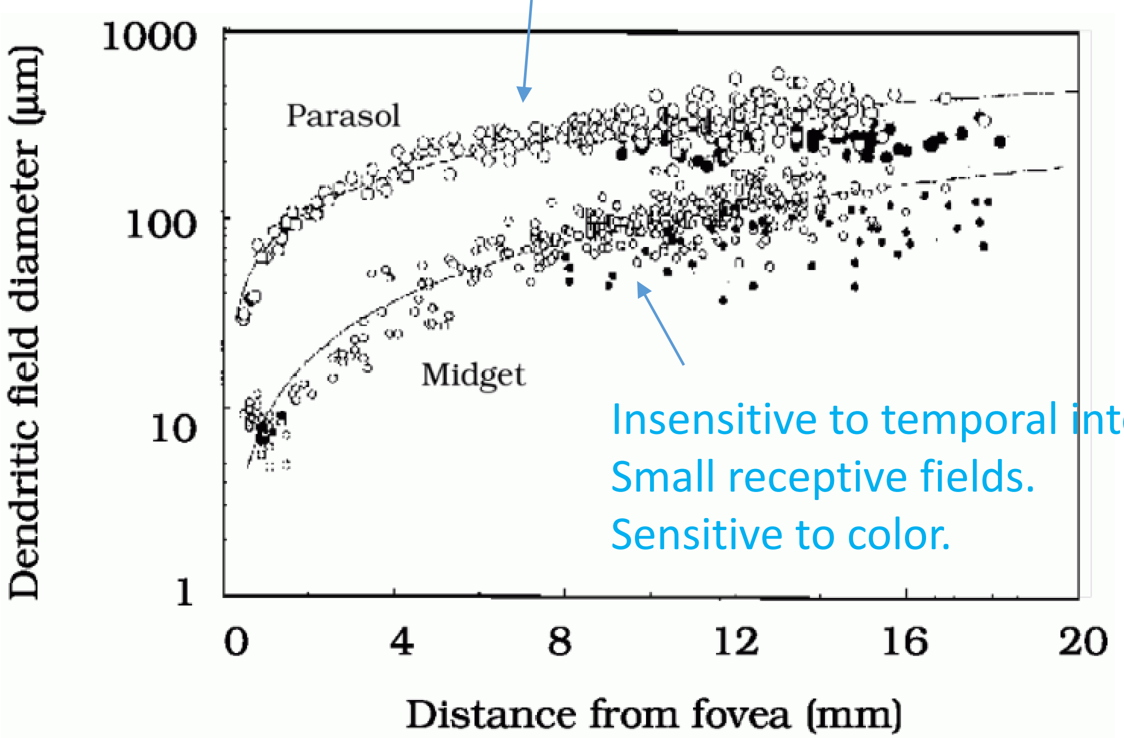
This cell would respond better to *fast* motion.



For your interest... (not on exam)

There are two classes of retinal ganglion cells...

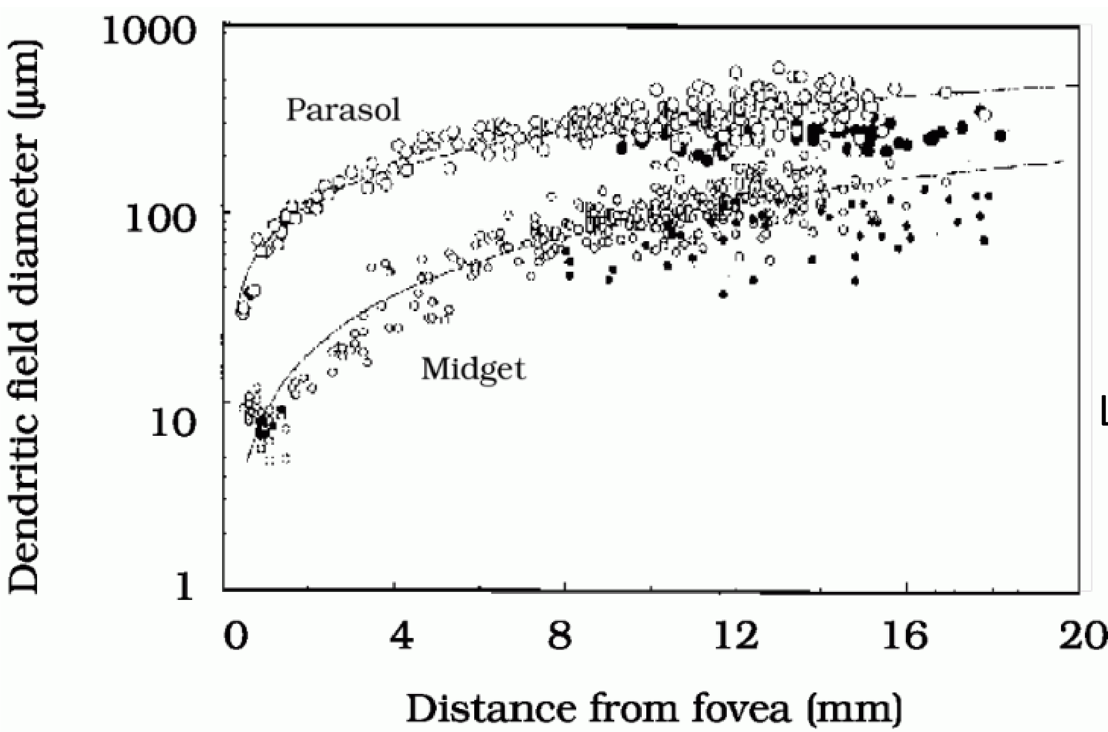
Sensitive to temporal intensity variations.
Large receptive fields.



Insensitive to temporal intensity variations.
Small receptive fields.
Sensitive to color.

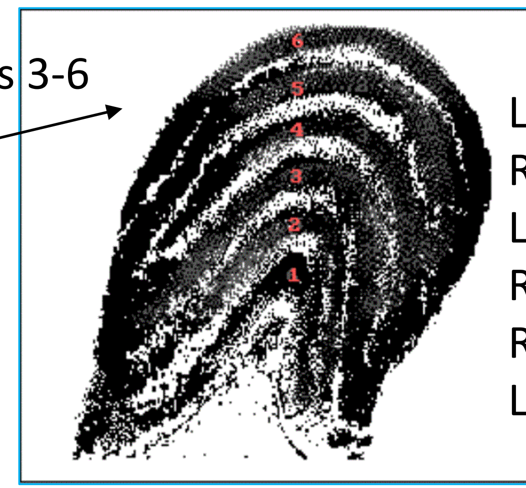
… and these two classes map to distinct layers in the LGN.

LGN



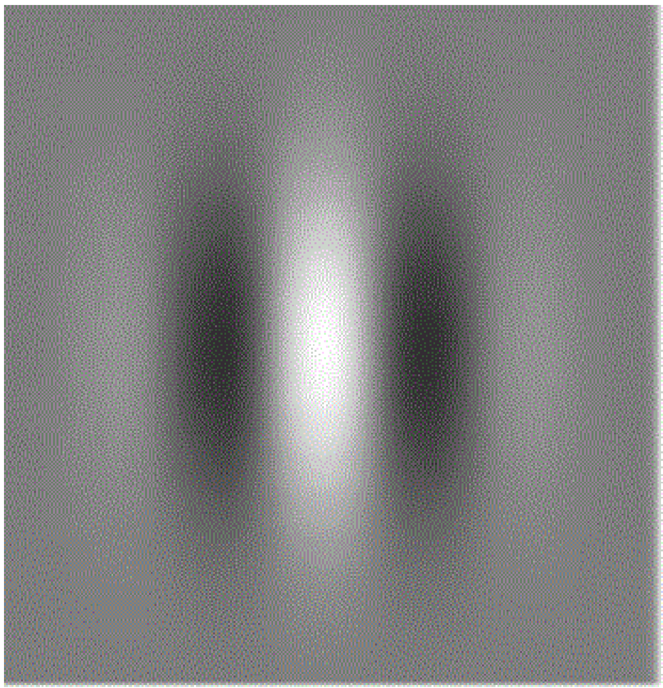
Layers 3-6

Layers 1, 2

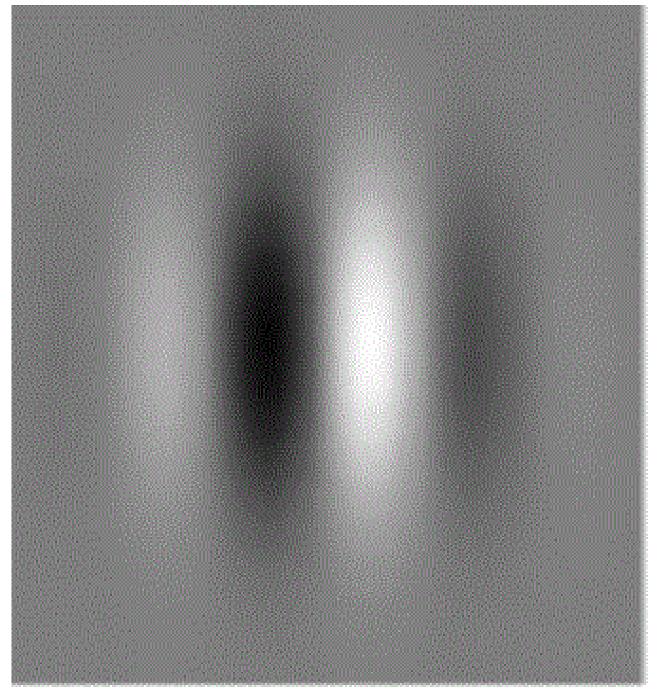


L
R
L
R
R
L

V1 cells can detect oriented structure in XY. What about XT and YT ?

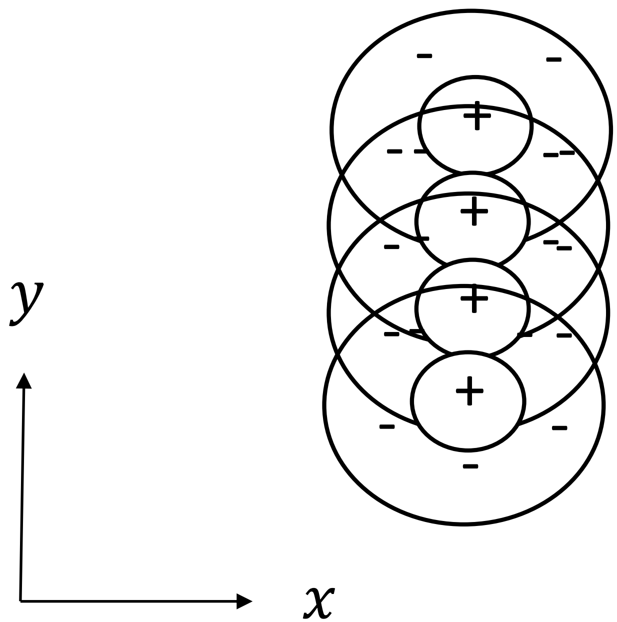


cosGabor(x,y)

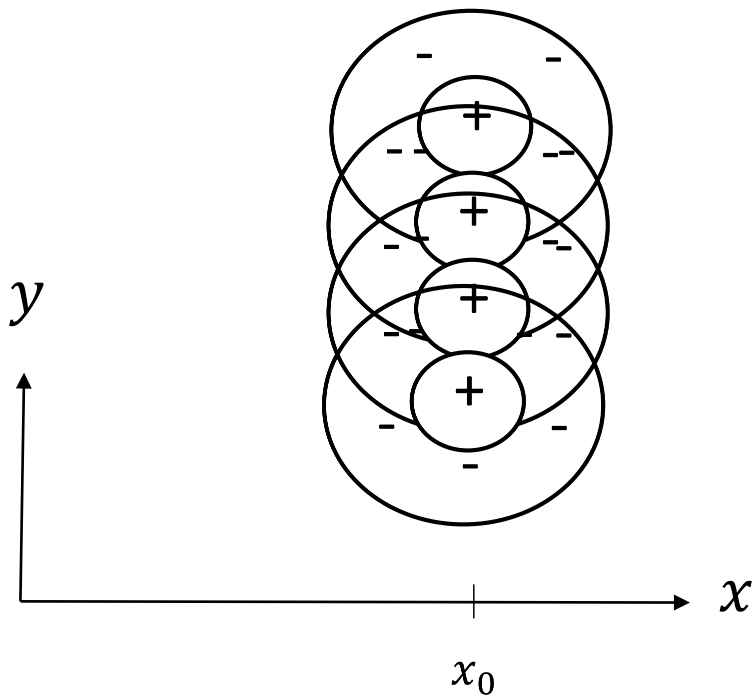


sinGabor(x,y)

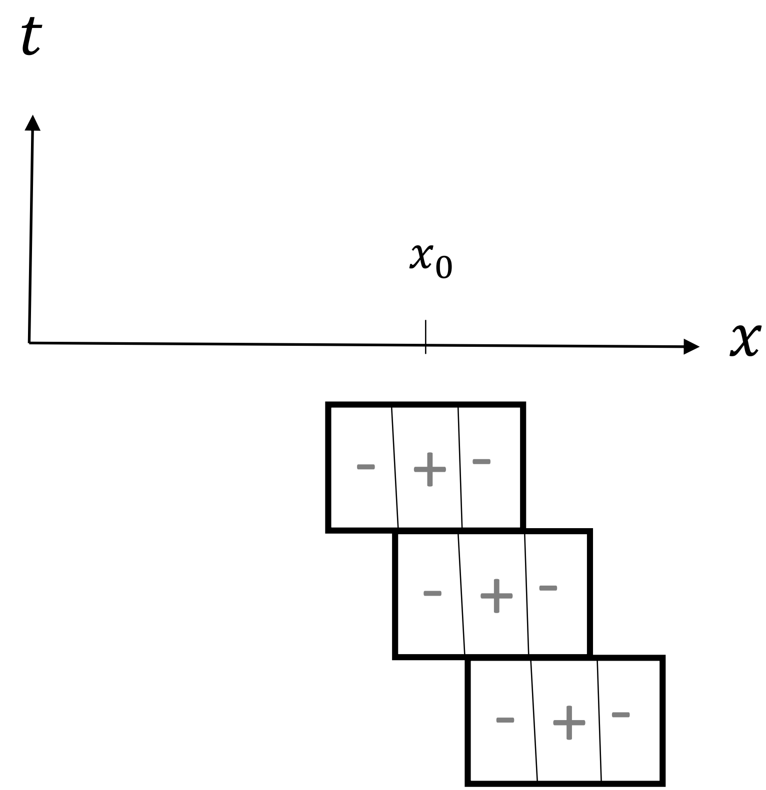
Hubel and Wiesel's idea for
a simple cell (line detector).

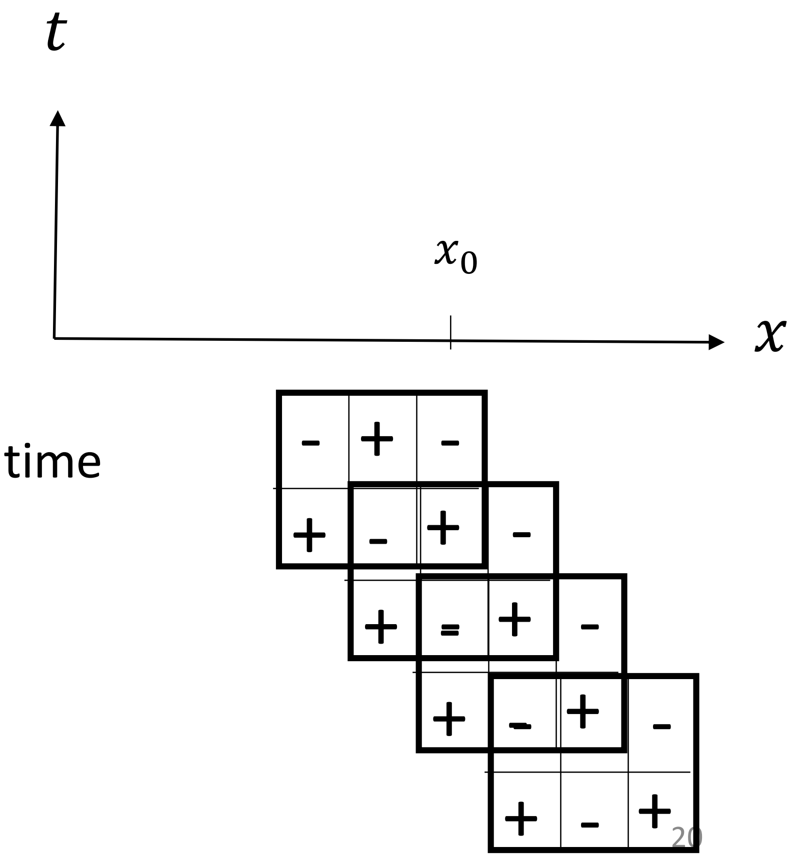


Hubel and Wiesel's idea for a simple cell (line detector).

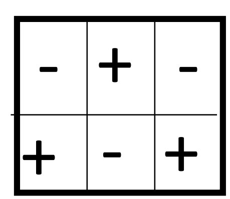


Reichardt (1950's): Temporal delays combined with spatial shifts could produce motion direction sensitivity.

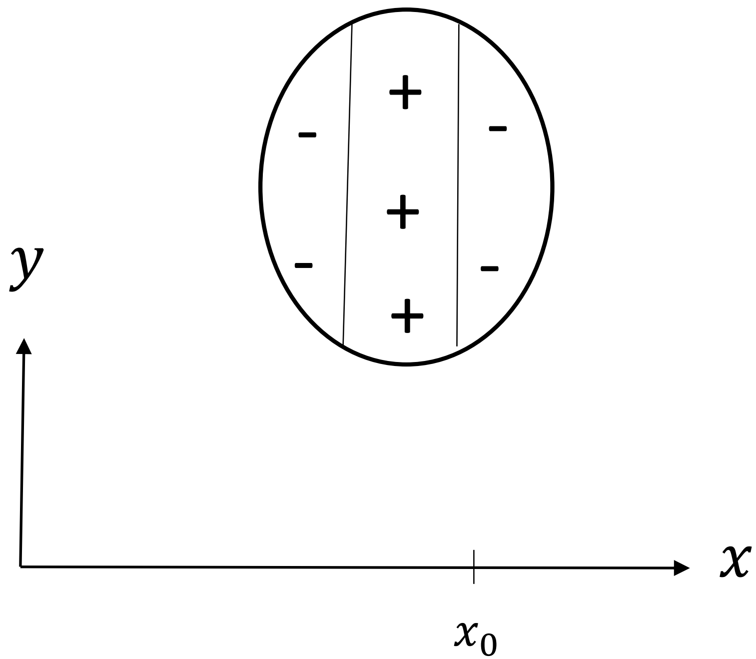


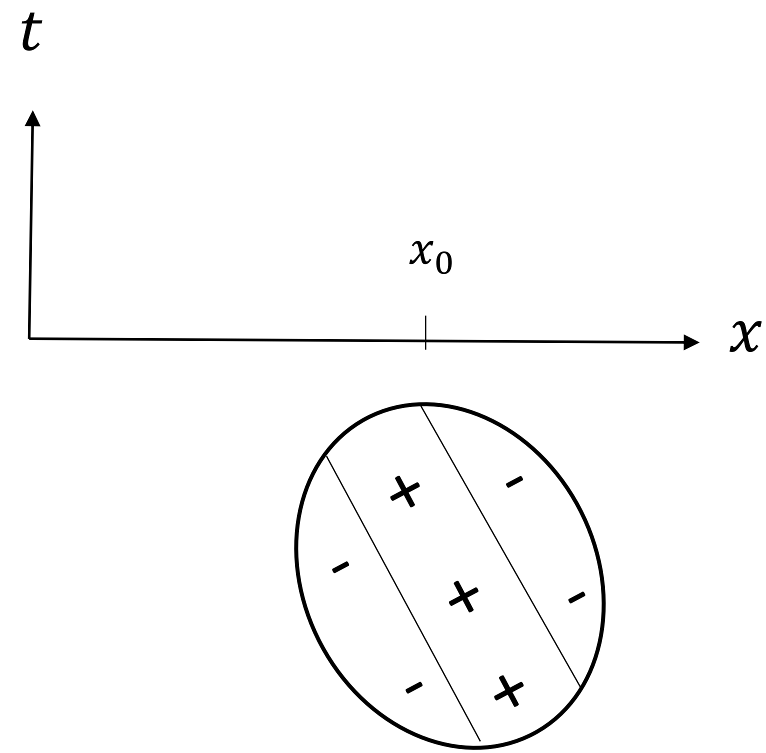


Alternatively, using time varying DOGs:



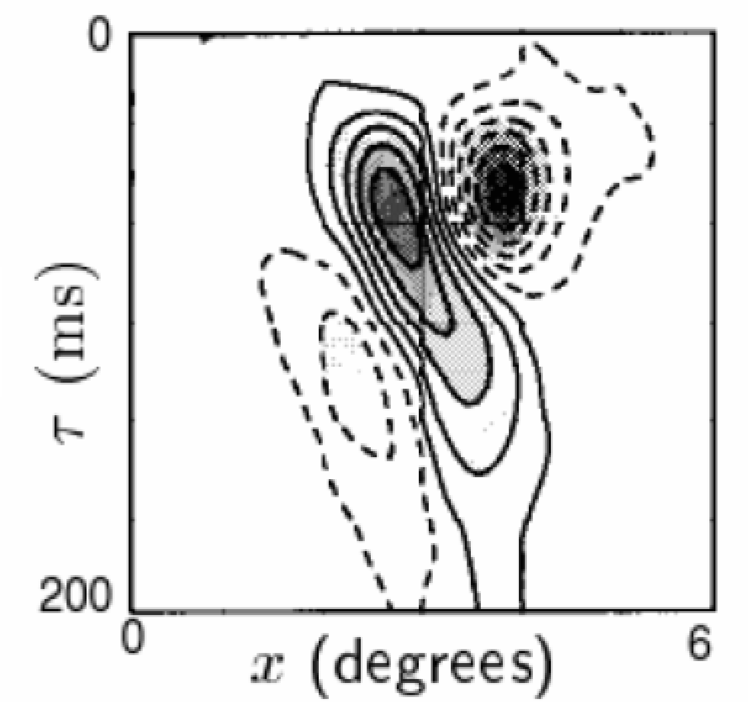
Orientation and direction tuning in V1



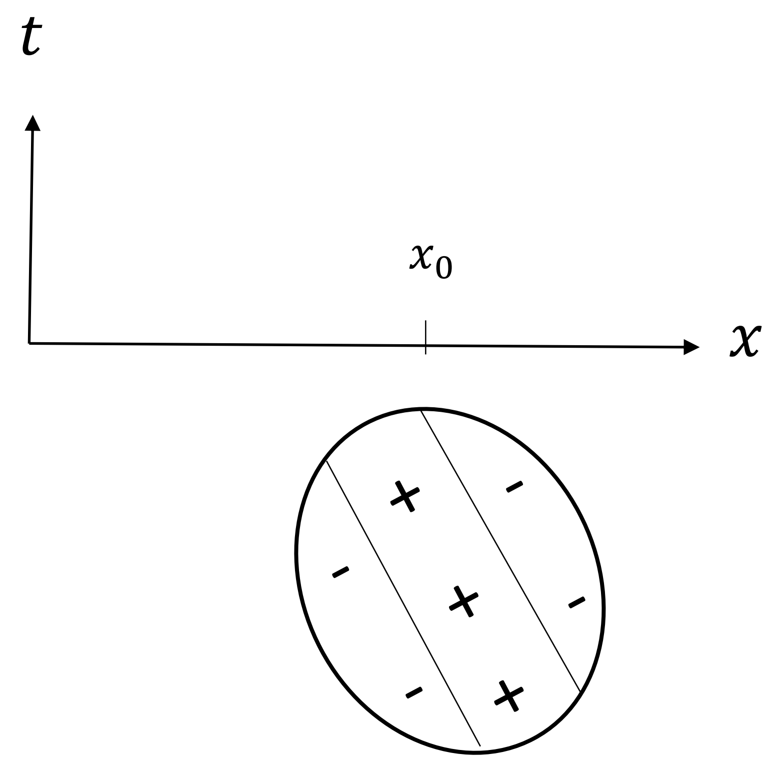


Cell is selective for vertically line moving to the left.

XT slice through receptive field profile of V1 cell

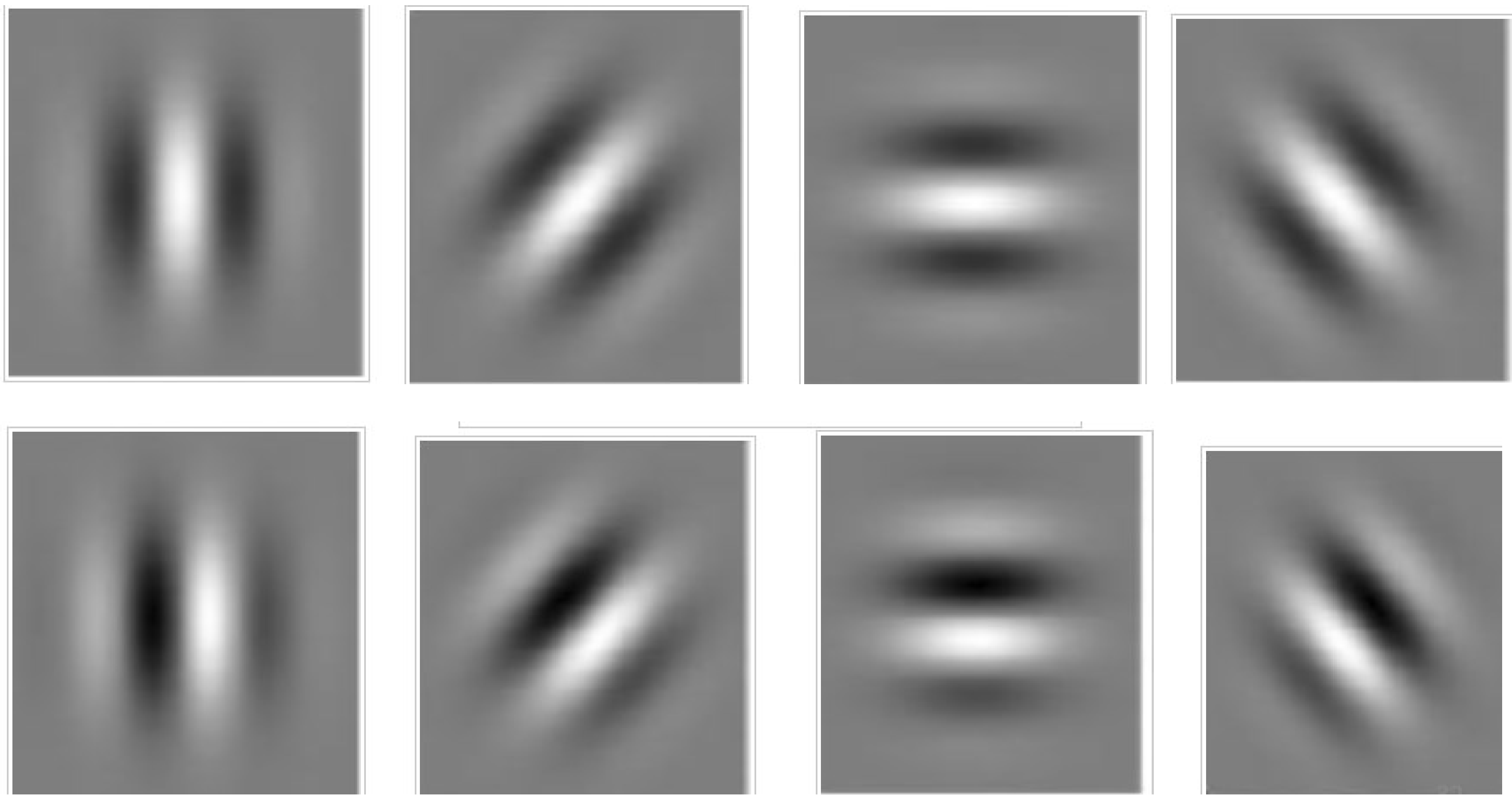


x

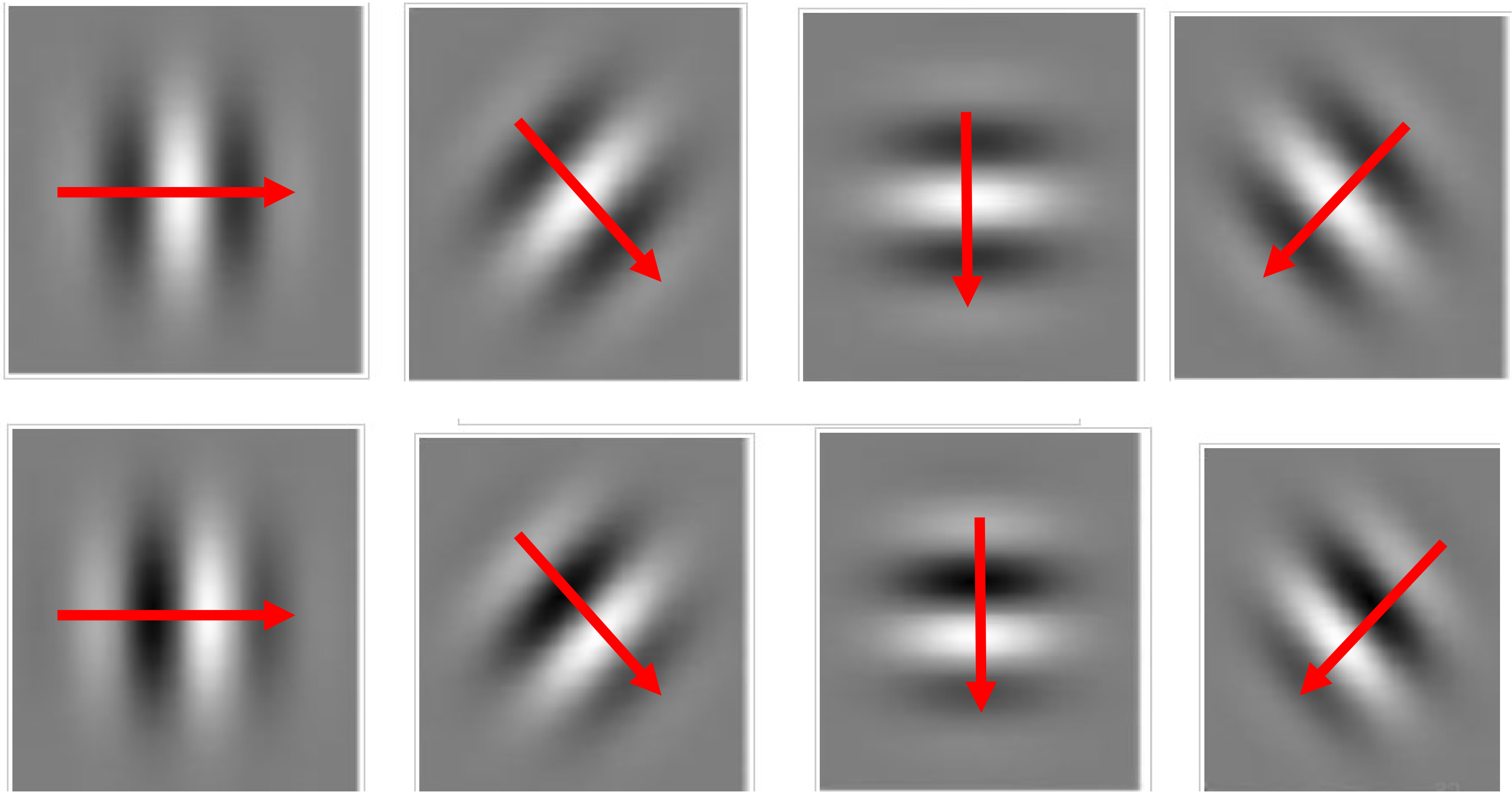


Simple and Complex Cells are *Orientation Tuned* (XY slice)

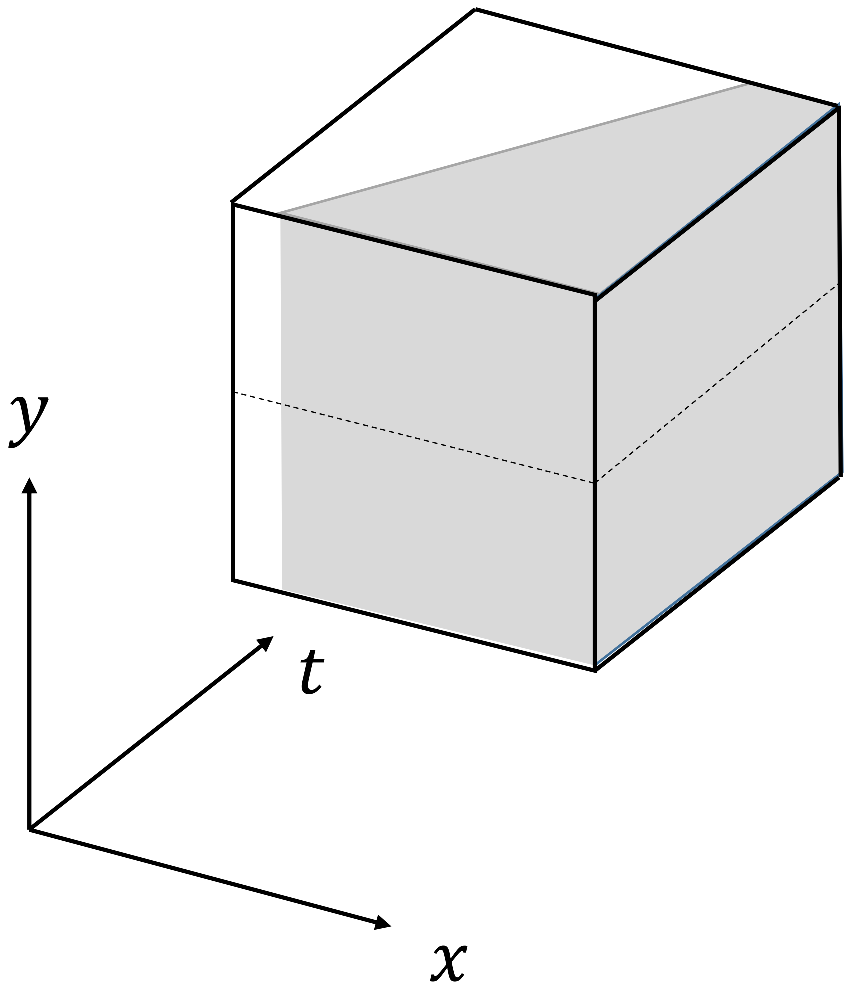
and many are binocularly disparity tuned.

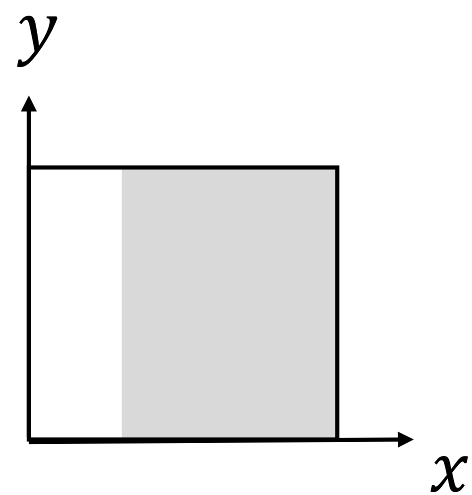


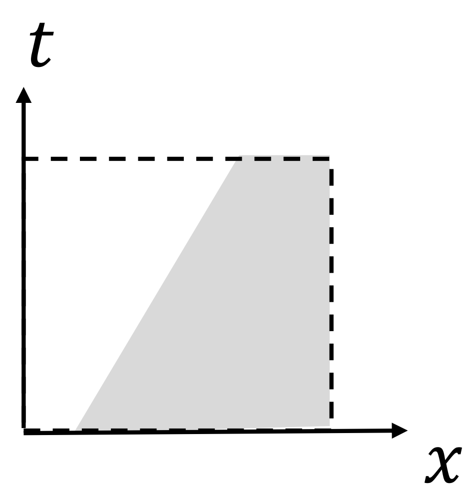
Many are also *motion direction and speed tuned.*



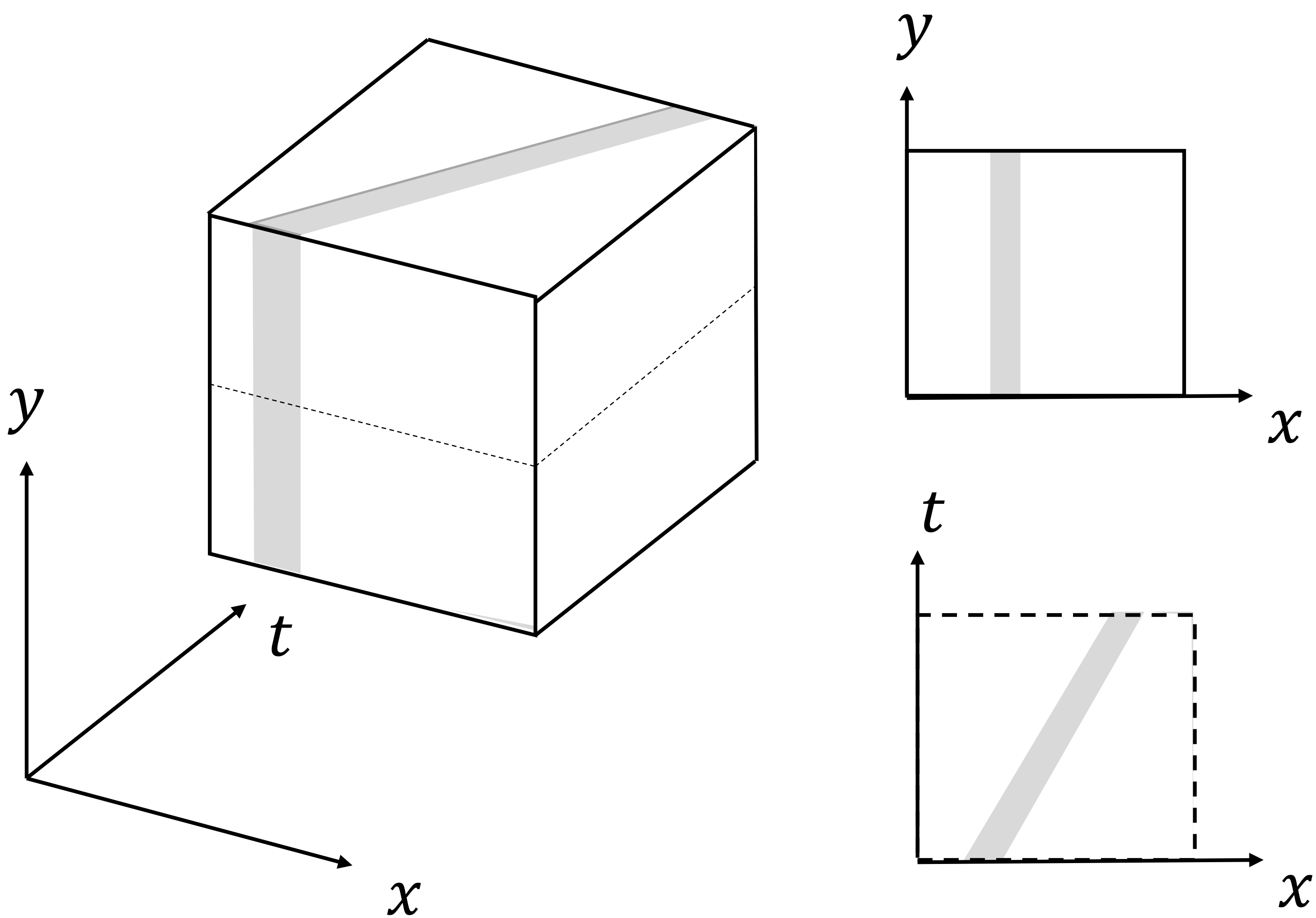
Vertical edge moving to the right



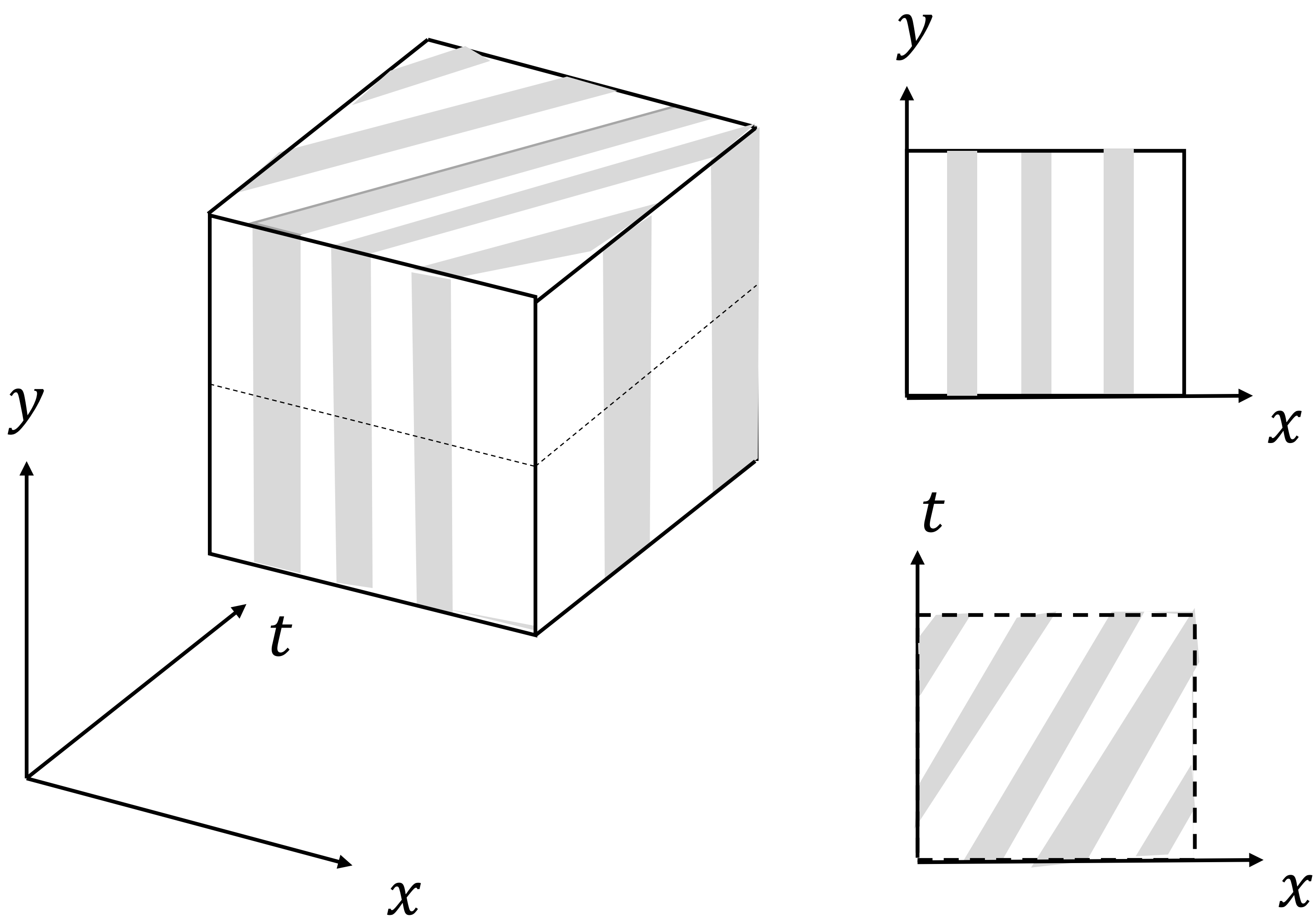




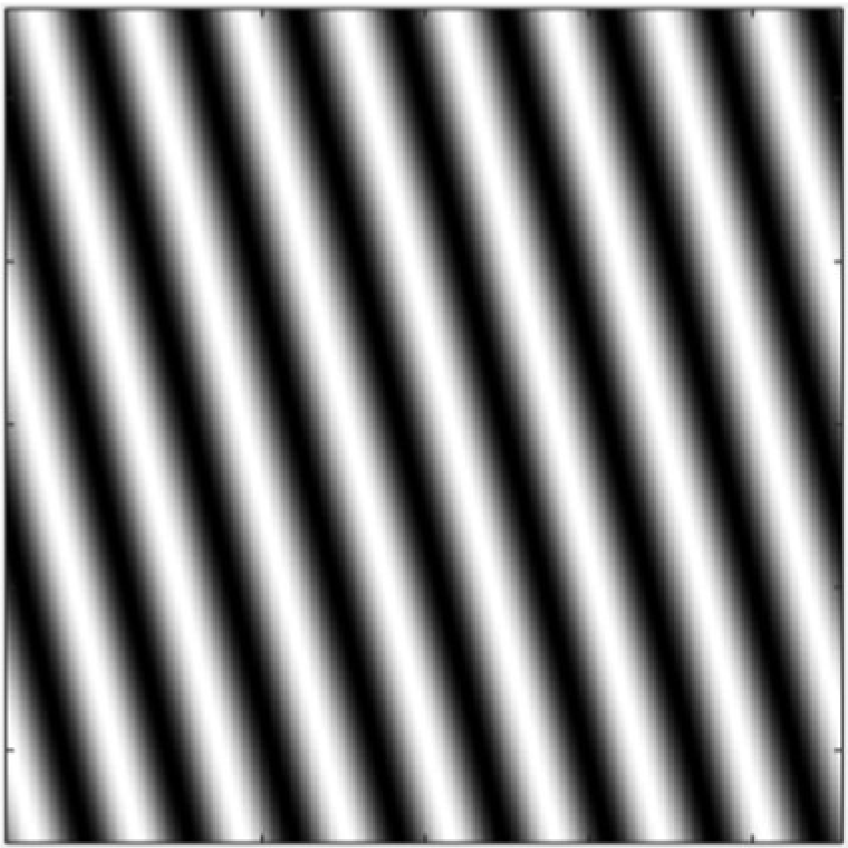
Vertical bar moving to the right



Vertically 2D sine moving to the right (wave)



Recall: 2D sine



$$\sin\left(\frac{2\pi}{N}(k_x\, x + k_y\, y)\right)$$

$$e.g.\quad k_x = 8$$

$$k_y = 2$$

$$N = 256$$

sine wave in XYT

$$\sin\left(\frac{2\pi}{N}(k_x\, x \;+\; k_y\, y) \;+\; \frac{2\pi}{T}\;\omega\, t\right)$$

Temporal frequency
(cycles per T frames)

Spatial frequency
(cycles per N pixels)

Exercise: what is the speed of the wave?

3D sine in XYT

$$\sin\left(\frac{2\pi}{N}(k_x\, x \;+\; k_y\, y) \;+\; \frac{2\pi}{T}\,\omega\, t\right)$$

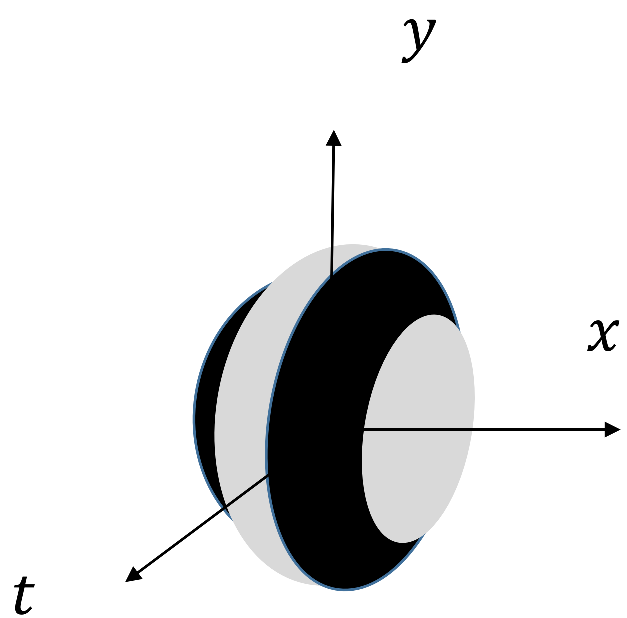
$$\frac{2\pi}{N}(k_x\, x \;+\; k_y\, y) \;+\; \frac{2\pi}{T}\,\omega\, t \;=\; \mathrm{c}$$

is the equation of a plane in XYT.

$$\left(\frac{2\pi}{N}k_x,\; \frac{2\pi}{N}k_y,\; \frac{2\pi}{T}\,\omega\right) \cdot \left(x, y, t\right) \;=\; \mathrm{c}$$

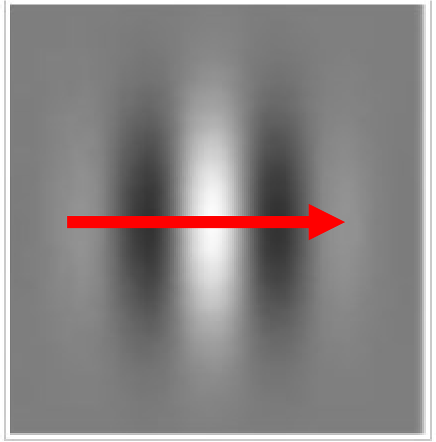
3D vector normal to the plane

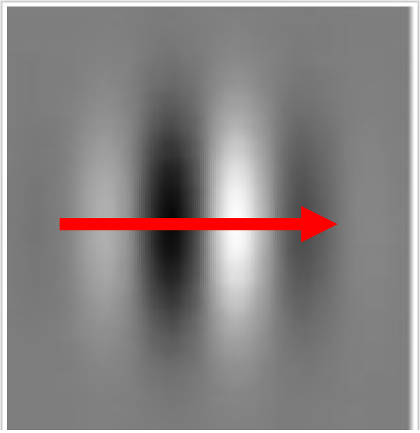
3D sine Gabor



$$\sin\left(\frac{2\pi}{N}(k_x\, x \;+\; k_y\, y) \;+\; \frac{2\pi}{T}\,\omega\, t\right)\; G(x, y, t, \sigma_x, \sigma_y, \sigma_t)$$

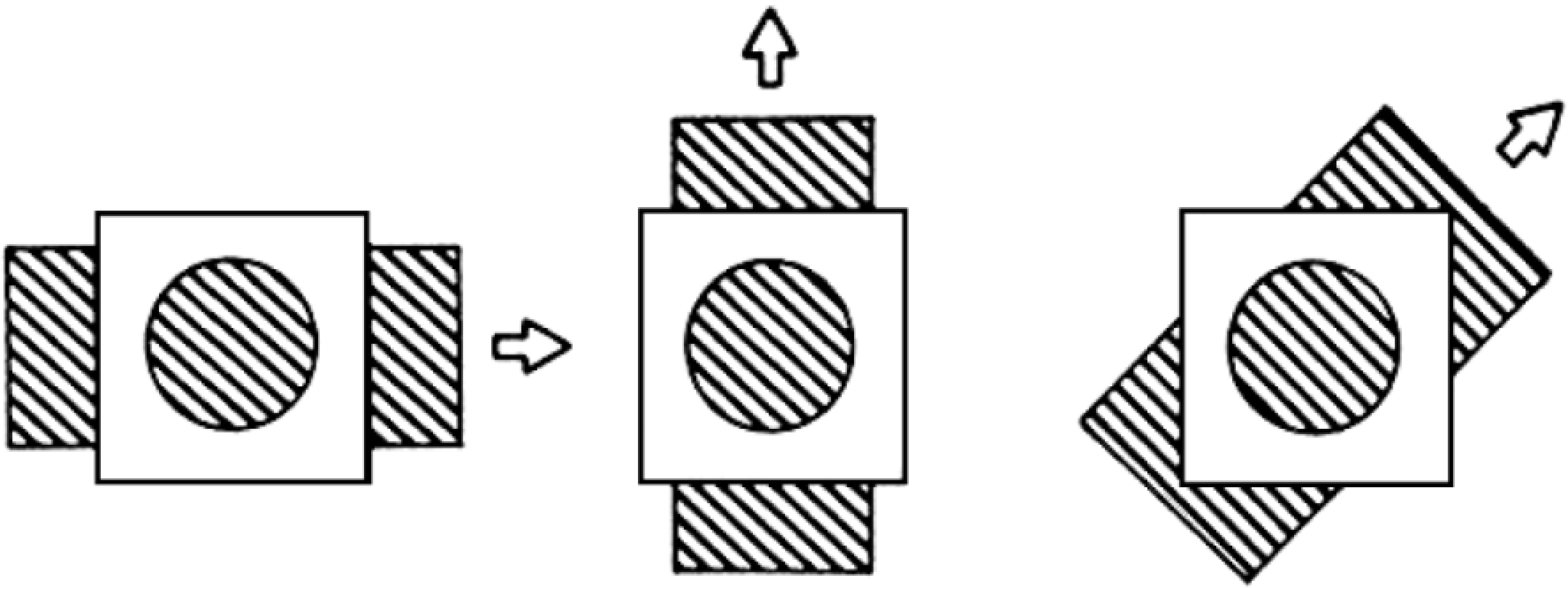
Normal Velocity





V1 cells only respond to the motion component that is normal (perpendicular) to their preferred orientation.

V1 : "Aperture Problem"



http://www.opticalillusion.net/optical-illusions/the-barber-pole-illusion/

The same issue arises with single bar, edge, or constant gradient.

To estimating image *velocity* $(v_x\,,\,v_y)$ at $(\,x, y)$, the visual system needs to combine the responses of many V1 (normal velocity) cells.