Vision in Space-Time:

Basic Mechanisms for Seeing Motion

Joe Lappin

17 March 2004

***General theme*: Visual images are moving images. The eyes move, and objects move. Vision is very sensitive to image motion and to information it provides about spatial structure.**

Mechanisms:
- dipole image changes
- autocorrelation & the bi-local Reichardt detector
- directional selectivity
- cortical area MT / V1

Computational problems: ***local*** — the aperture problem

global — integrating multiple local directions

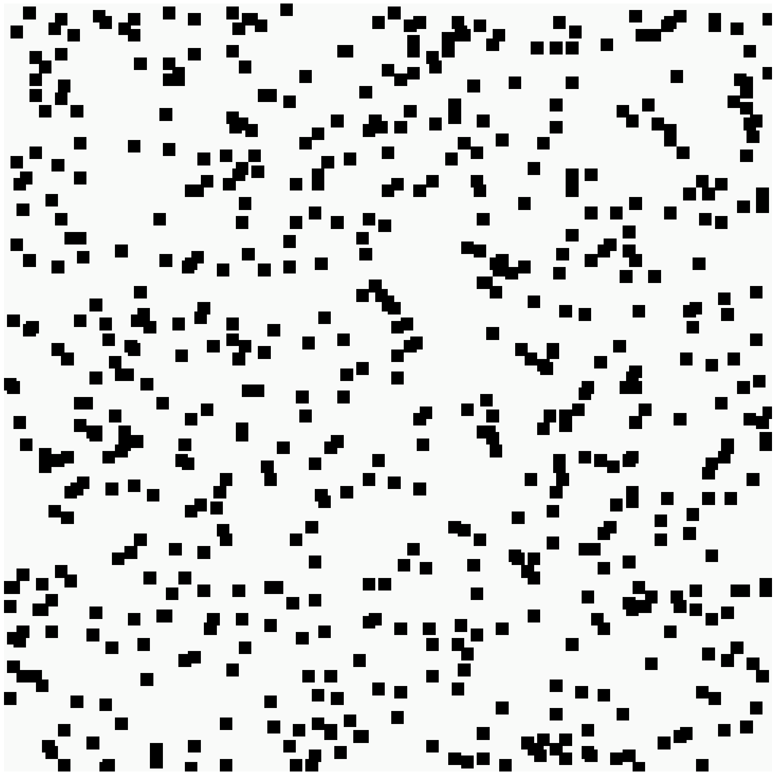
Readings:

• Borst, E. (2000). Models of motion perception. *Nature Neuroscience Supplement,* **3**, 1168. [a brief history of the "Reichardt detector"]

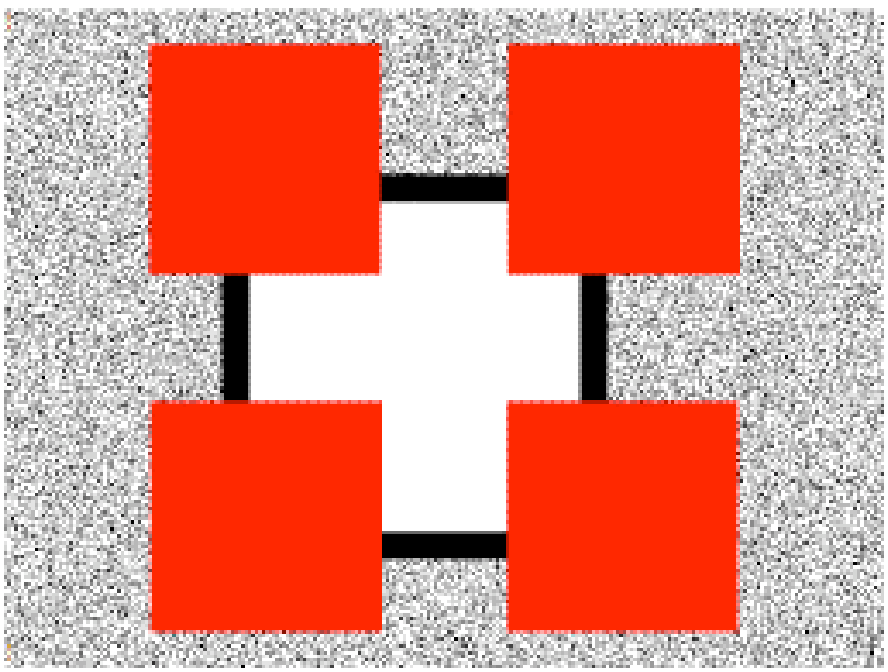
• Newsome, W.T., Britten, K.H., & Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. *Nature*, **341** , 52-54.

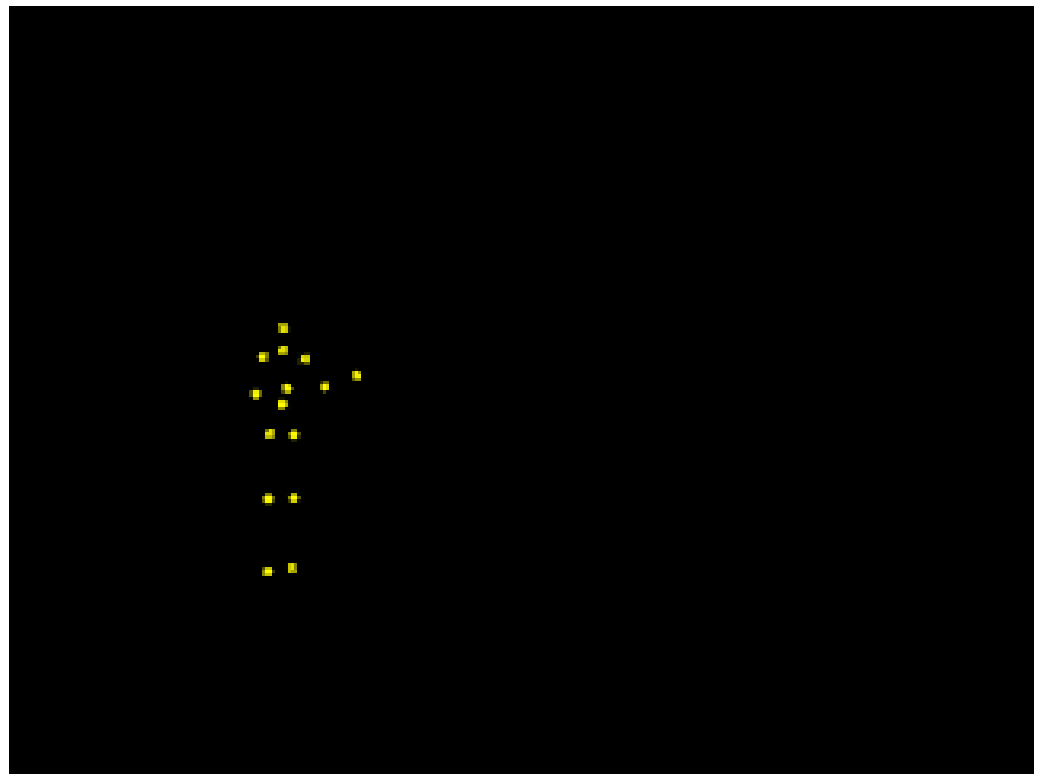
• Saltzman, C.D., Britten, K.H., & Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgments of motion direction. *Nature,* **346**, 174-177.

• also relevant: Tovée, M.J., (1996). *An introduction to the visual system.* Ch. 10, "Motion perception". New York: Cambridge U. Press.



Motion provides spatial information



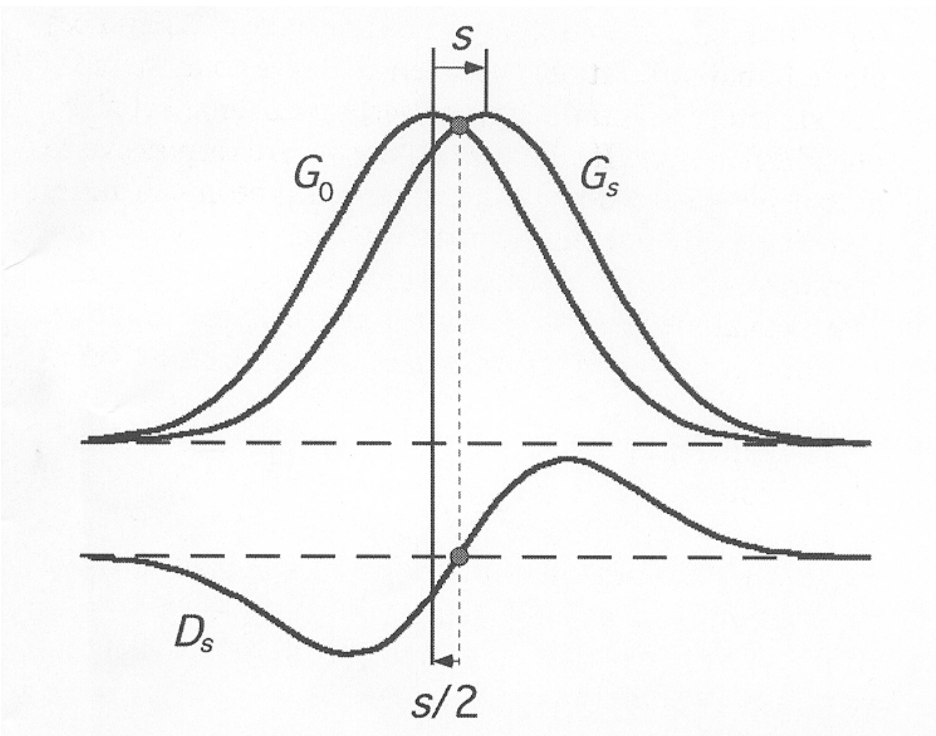


Motion Detection

The dipole image change produced by motion:

<u>Visual neurons are virtually all highly responsive to motion</u>!

• Image motion produces a simultaneous increase and decrease in optical intensity. The dipole nature of this image change shifts the balance between excitatory and inhibitory effects — analogous to adding weight to one side of a balance scale and subtracting it simultaneously from the other side.

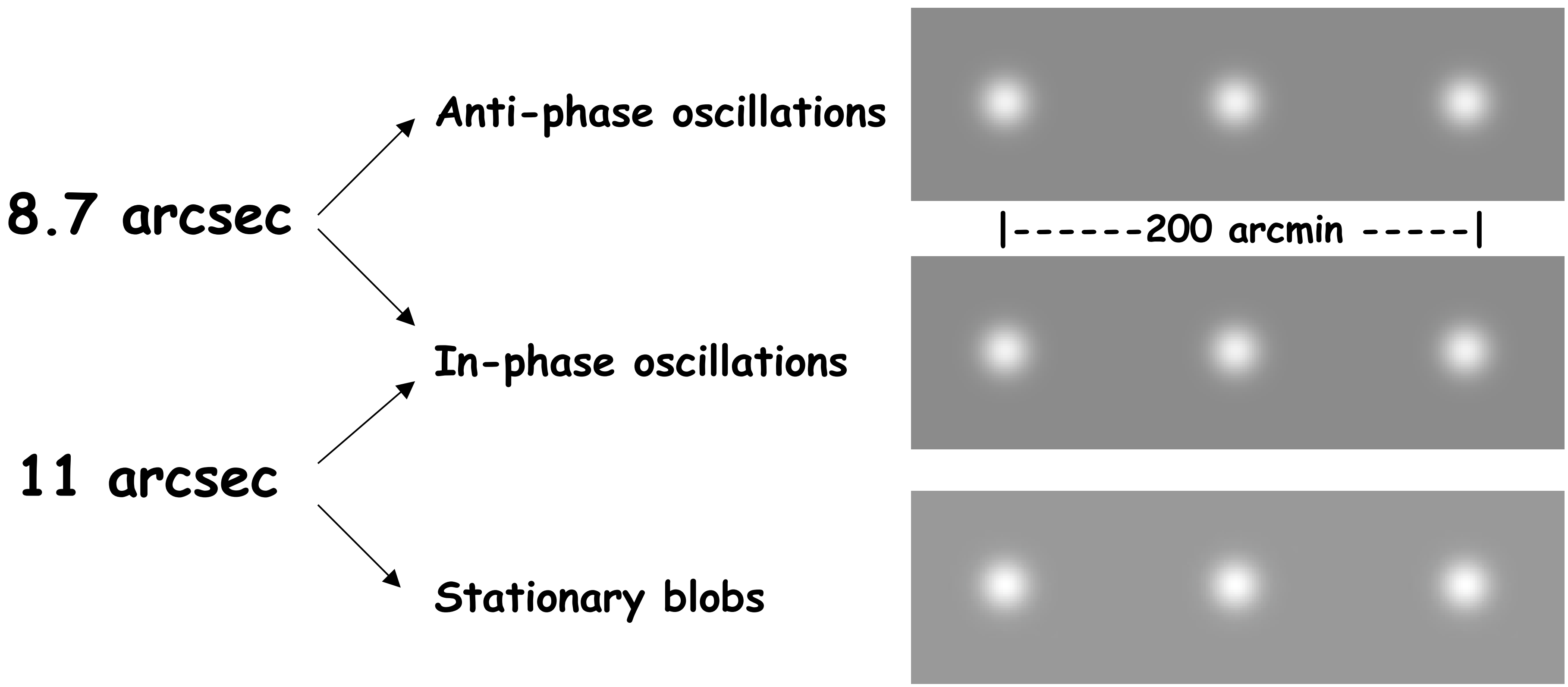


• Both P and M retinal ganglion cells are very responsive to image motion; and both cell classes carry motion information (contrary to what textbooks sometimes say). In humans and other primates, retinal ganglion cells and LGN cells are not directionally selective. For this reason, motion information is sometimes regarded as not encoded by ganglion cells, but this is misleading.

• Moreover, the temporal patterns of spike trains in neighboring ganglion cells are <u>coherent</u> (correlated) with one another. Responses of neighboring ganglion cells carry information about relative motion, even if the two responses do not influence each other.

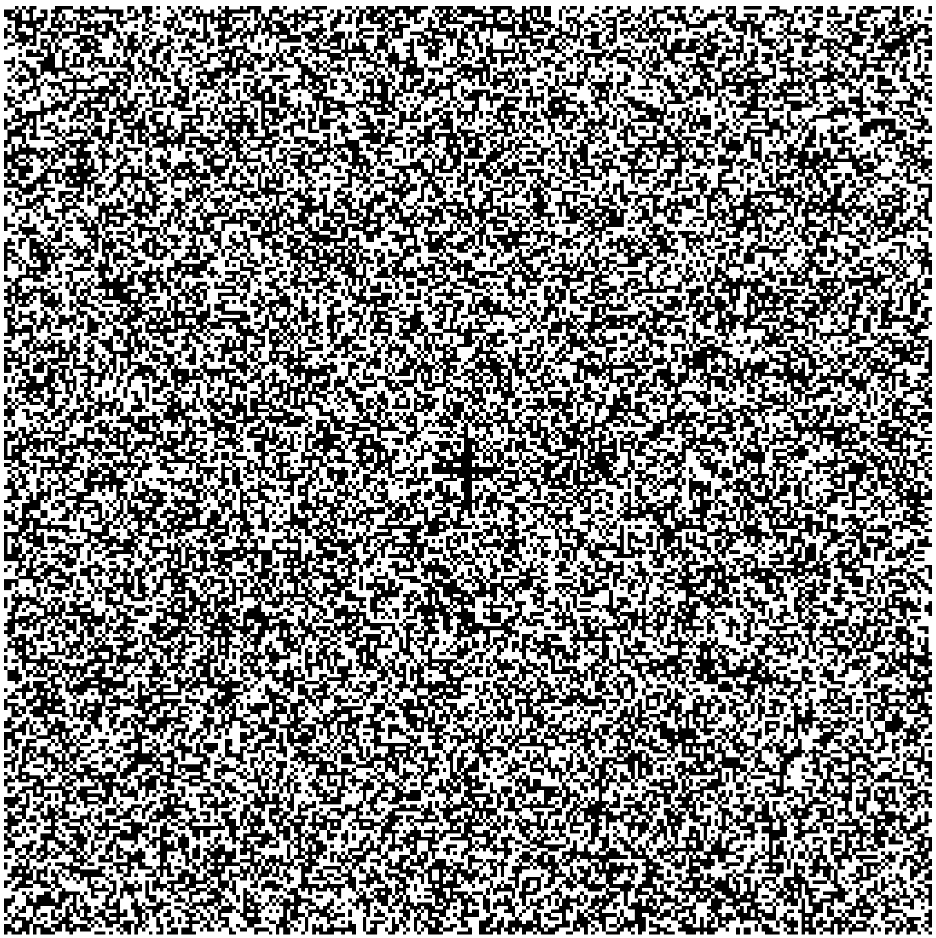
Lappin, Donnelly, & Kojima (2001): "Coherence of early motion signals"

- The acuity for *relative* motion is better than acuity for *detecting* motion !!



Common fate

Vision is highly sensitive to the "common fate" (a classical Gestalt organizing principle) of changing images. Two ways of illustrating this are shown below.



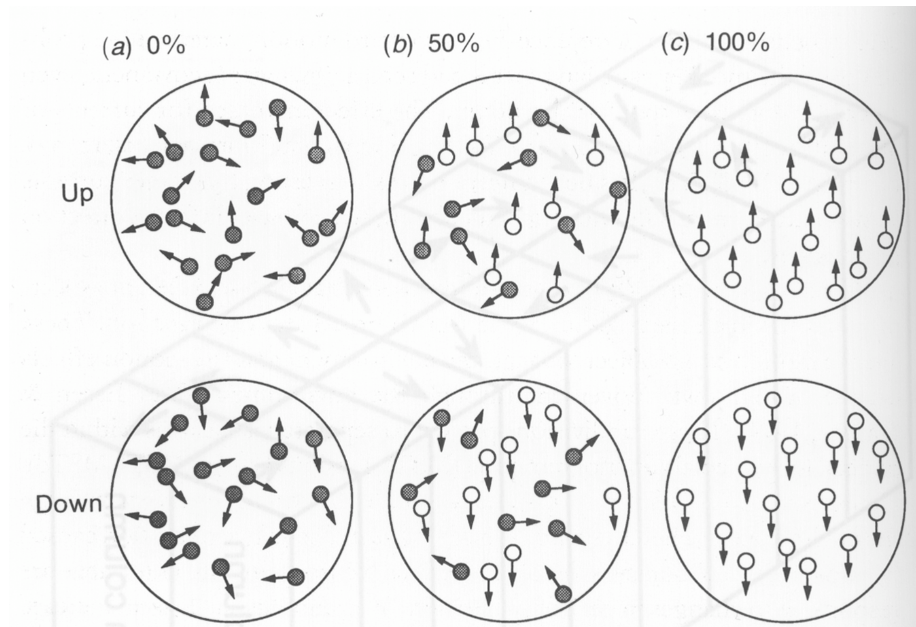


Fig. 10.7. Schematic representation of a random-dot stimulus that can be used to motion thresholds. In each panel, the arrow indicates the direction in which the attached dot moves. The dots moving in a random direction are shown in grey and those moving in the same direction are shown in white. (*a*) The dots are moving randomly and there is 0% correlation between their movements. (*c*) All the dots are moving together in the same direction and there is 100% correlation of movement. (*b*) Half the dots are moving in a random manner and half the dots are moving in the same direction. In this case, the dots are 50% correlated. (Redrawn with permission from Sekuler & Blake, 1994. Copyright (1994) McGraw-Hill.)

Autocorrelation: Structure and motion (even in 3D) may be described by the motion that maps a pattern onto itself.

3D motions can be described by autocorrelations of 2D images. Autocorrelation is defined on a transformation that maps a pattern onto itself. Such transformations could, for example, consist of groups of motions in 3D. For a discrete 2D pattern, $f(x, y)$, the autocorrelation function for translations, say $M(a, b)$, is usually written as

$$A(a, b) = \sum\sum [f(x, y) \cdot f(x+a, y+b)]$$

(with summation over the space coordinates x and y).

It is important to understand that the autocorrelation is defined on the transformation parameters a and b rather than on the initial spatial coordinates of the pattern.

Thus, we can generalize this notion of autocorrelation by the formula

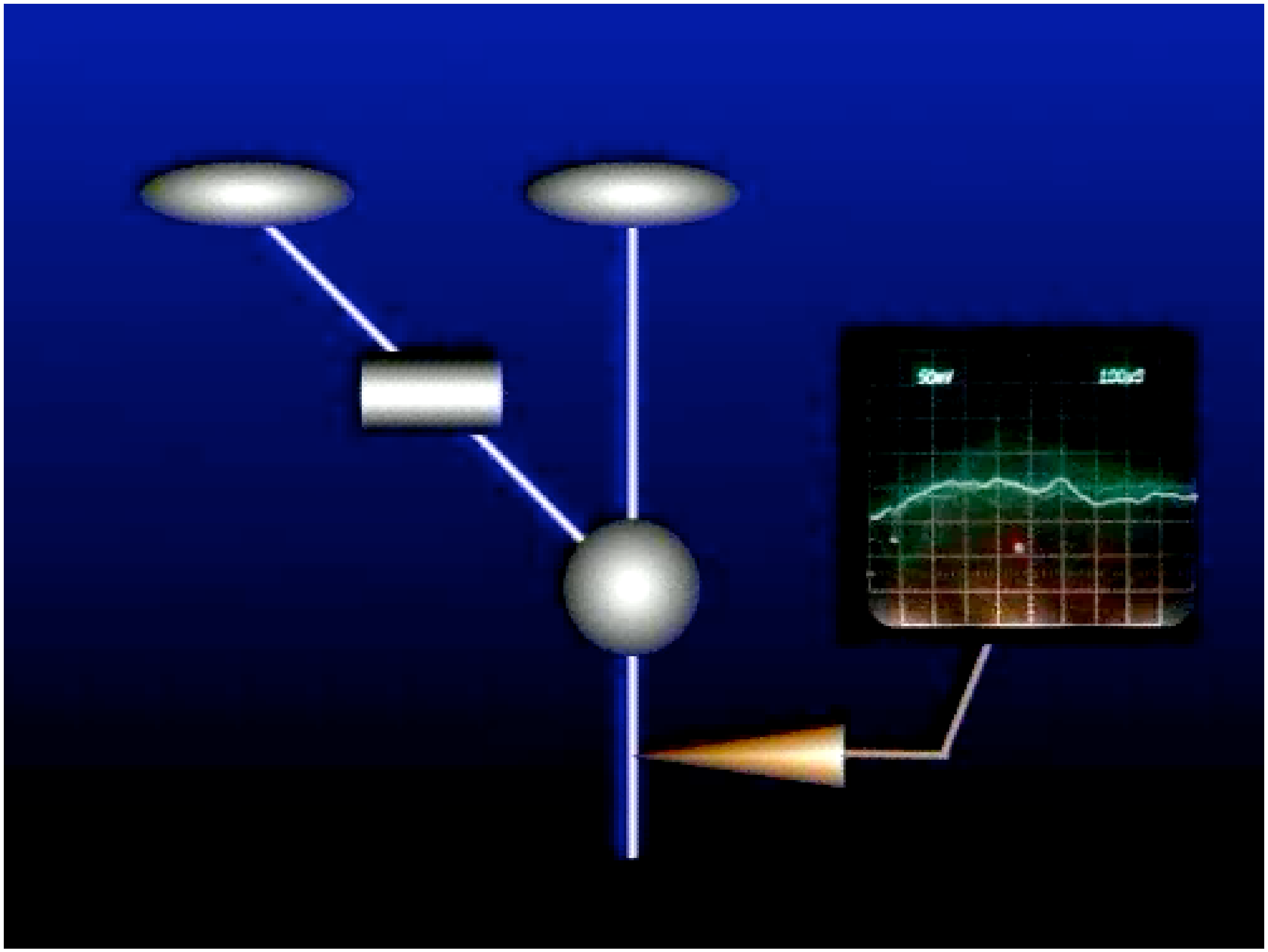
$$A(M) = \sum\sum \{f(x, y) \cdot M[f(x, y)]\}$$

where M is any transformation that maps a pattern onto itself — *e.g.*, a motion (translations and rotations) even in 3D space.

The basic idea is that space and motion can be specified by transformations of images over time. As James Gibson once pointed out, the transformations that a form undergoes may be as important as the form that undergoes the transformations.

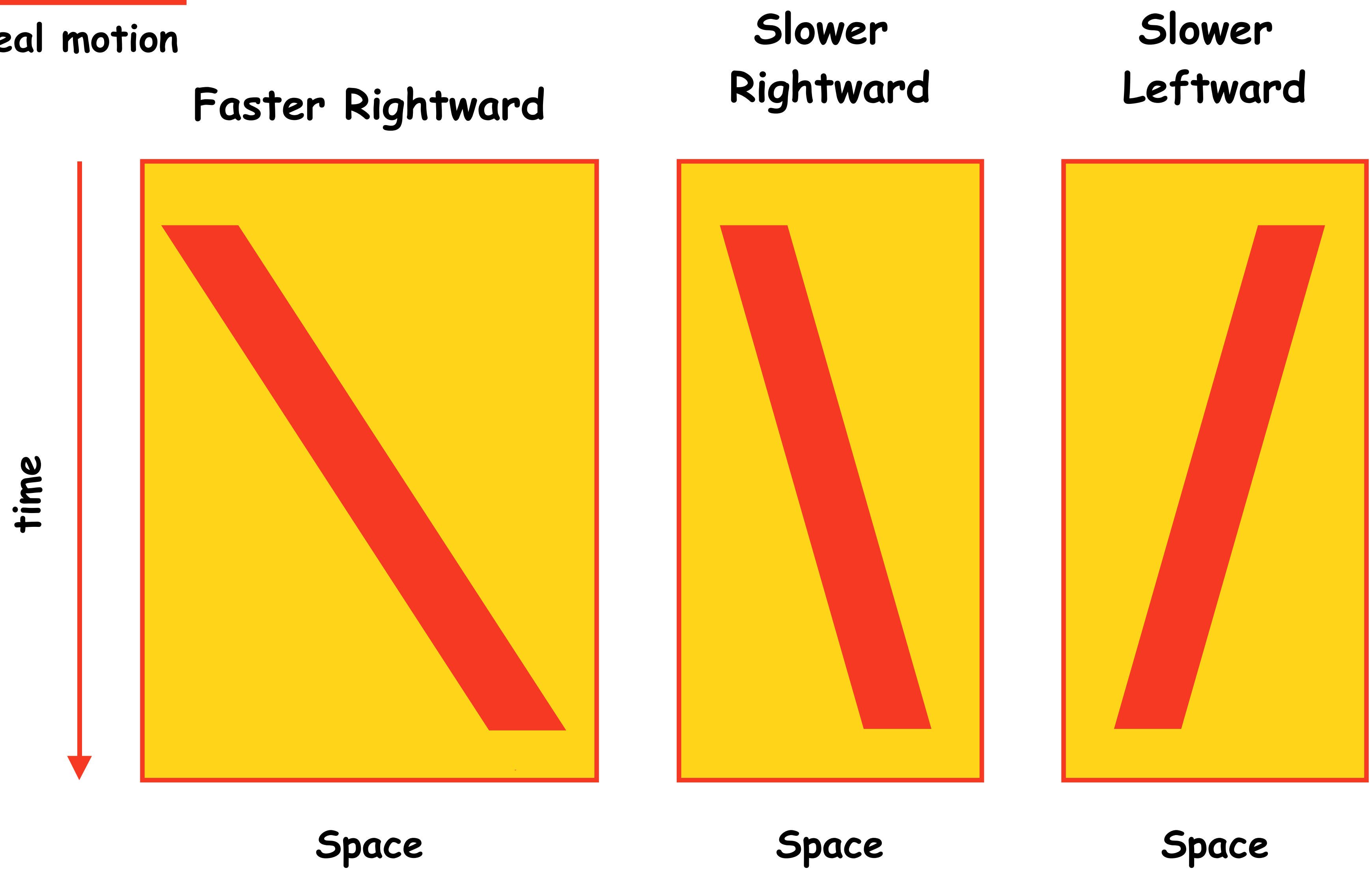
One mechanism that might accomplish such autocorrelations is known as the "*Reichardt detector*".

Reichardt's bi-local autocorrelation-detector

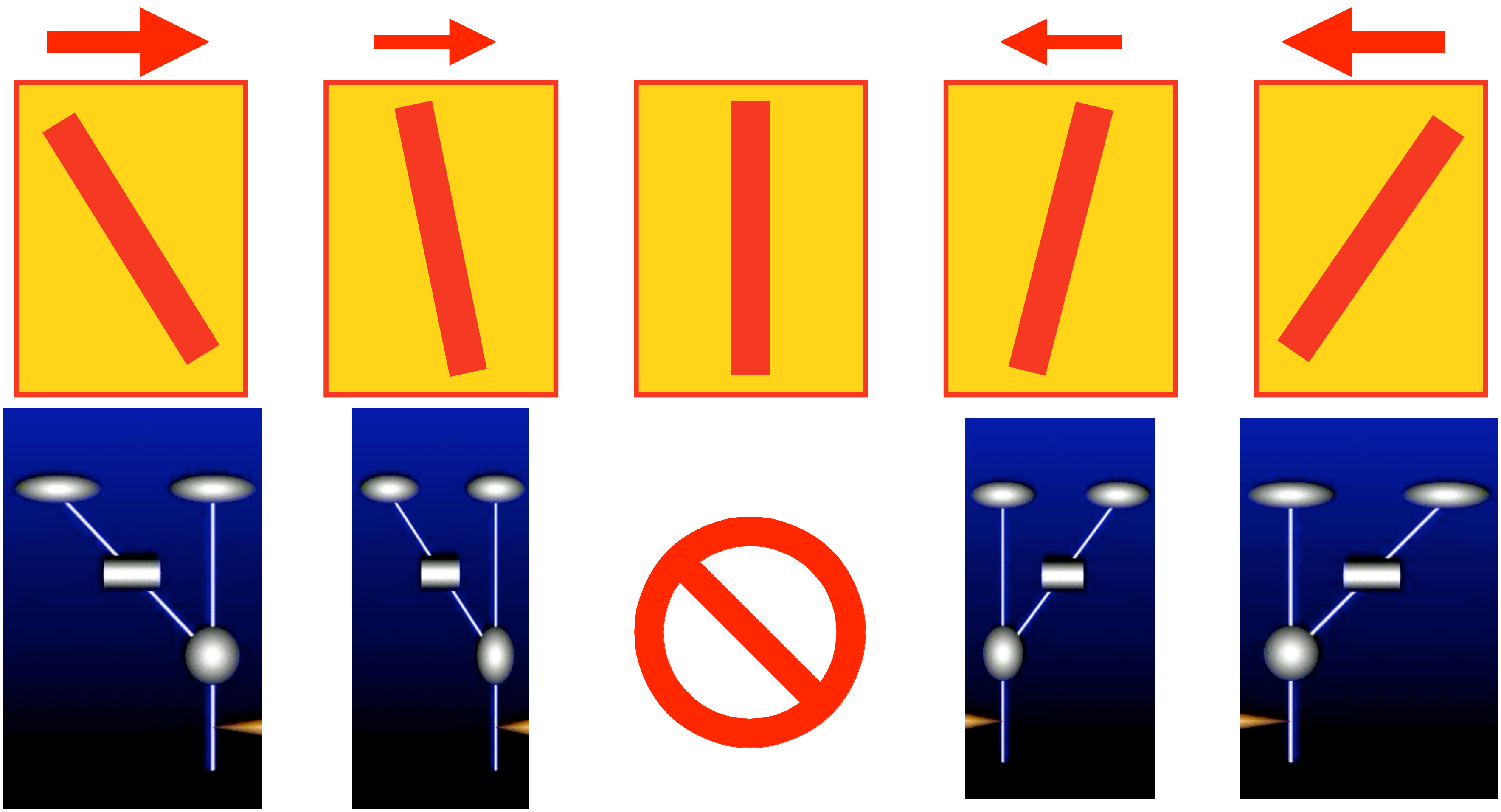


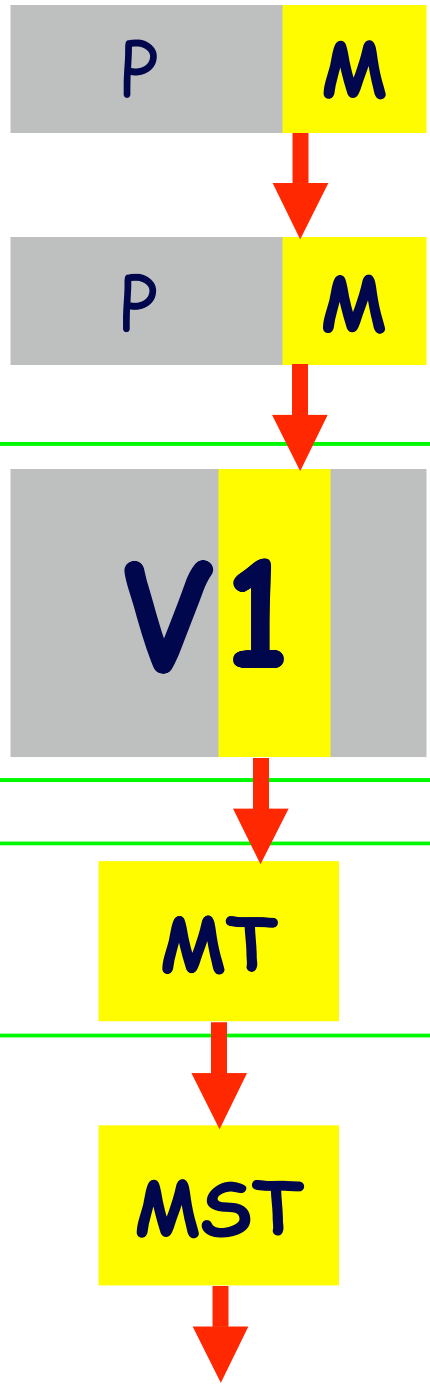
Space/time Plot

Real motion



How Can the Visual System Detect the "Orientation" of Motion Energy in Space/Time?





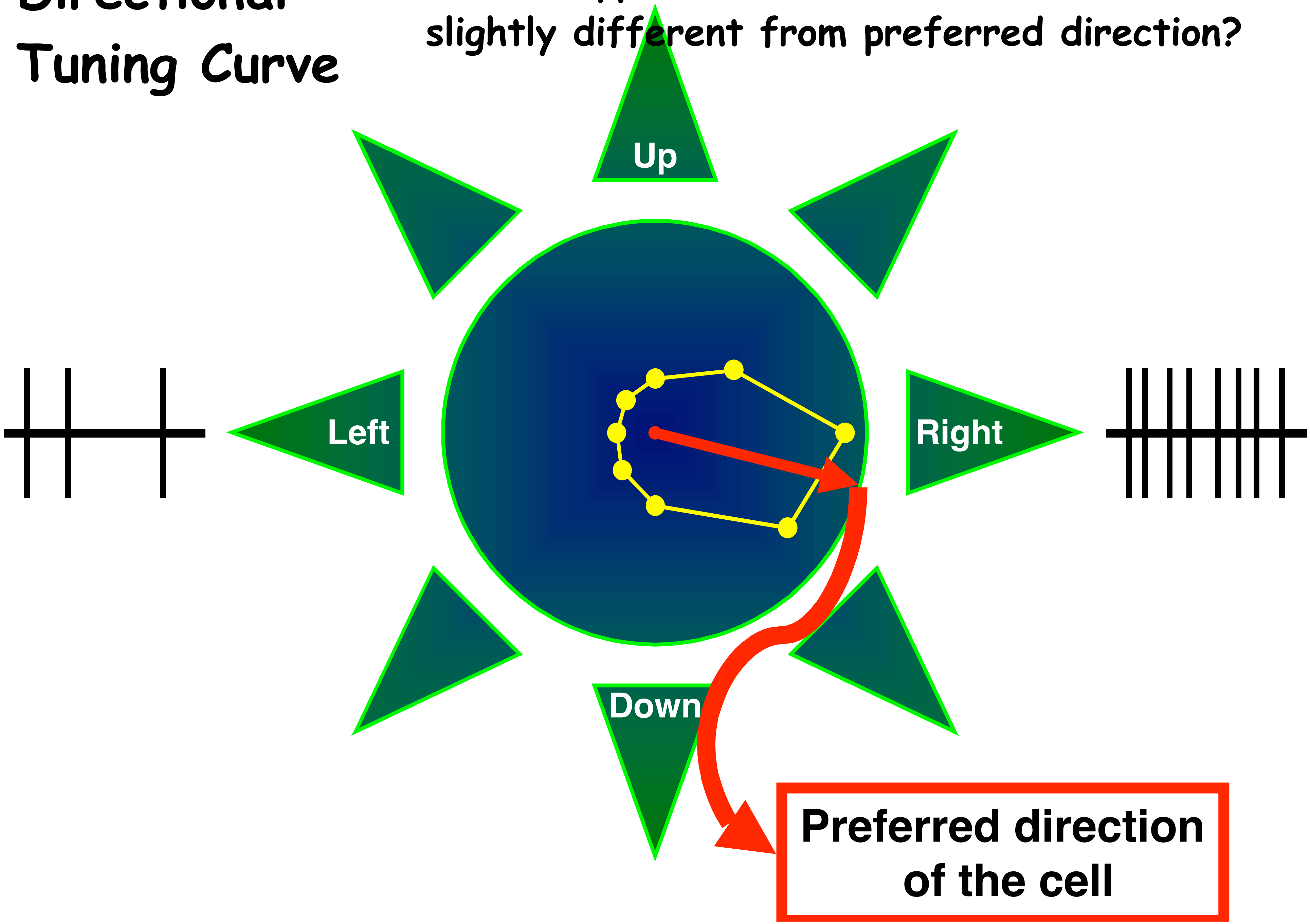
M cells respond strongly to rapid changes in intensity, ideal for conveying motion info to the brain.

V1 : Appearance of directionally selective (DS) cells (~15%)

MT : “The motion area.” Crucial for motion processing (~90% DS)

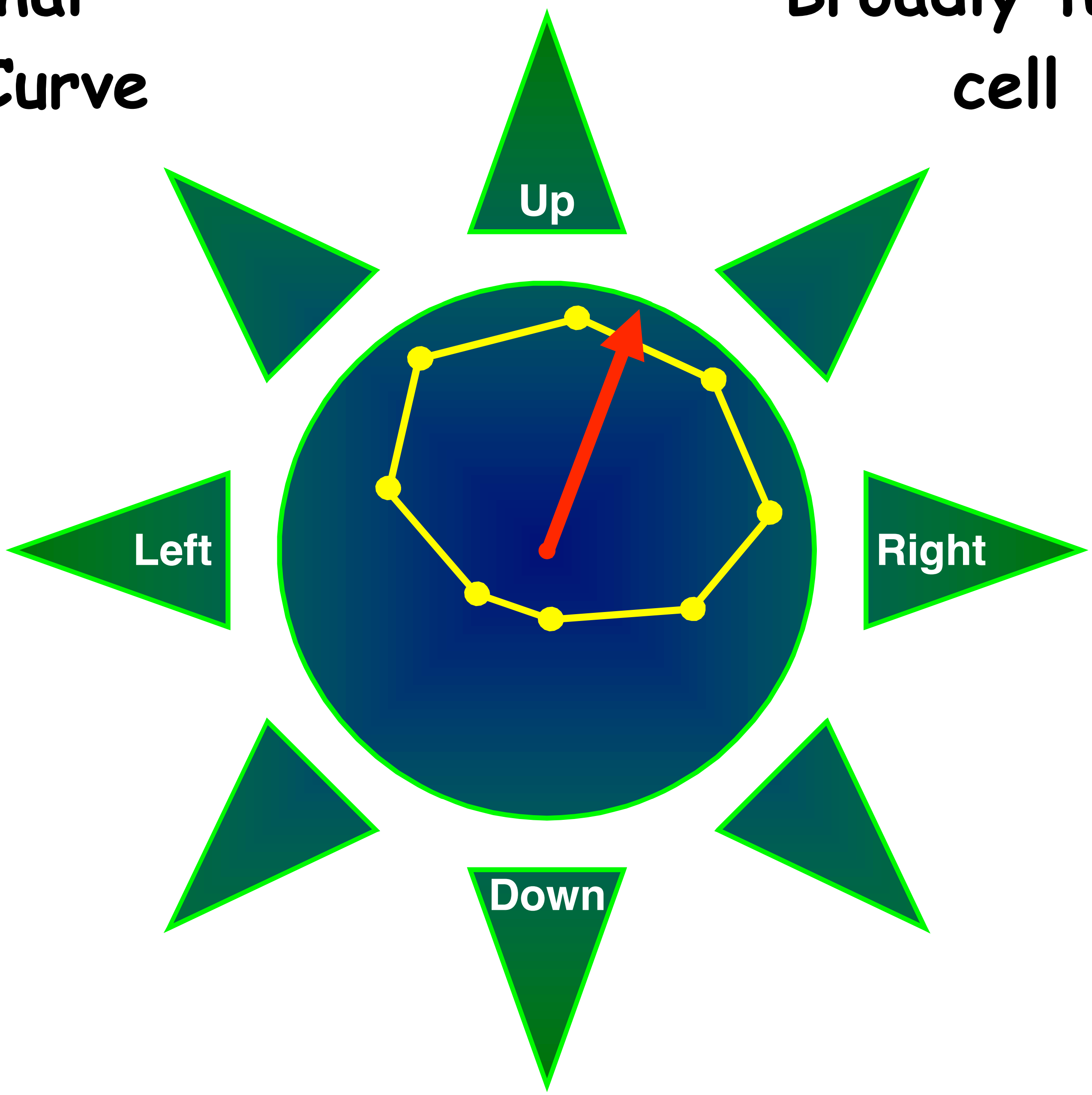
Directional Tuning Curve

What happens if the direction of motion is slightly different from preferred direction?



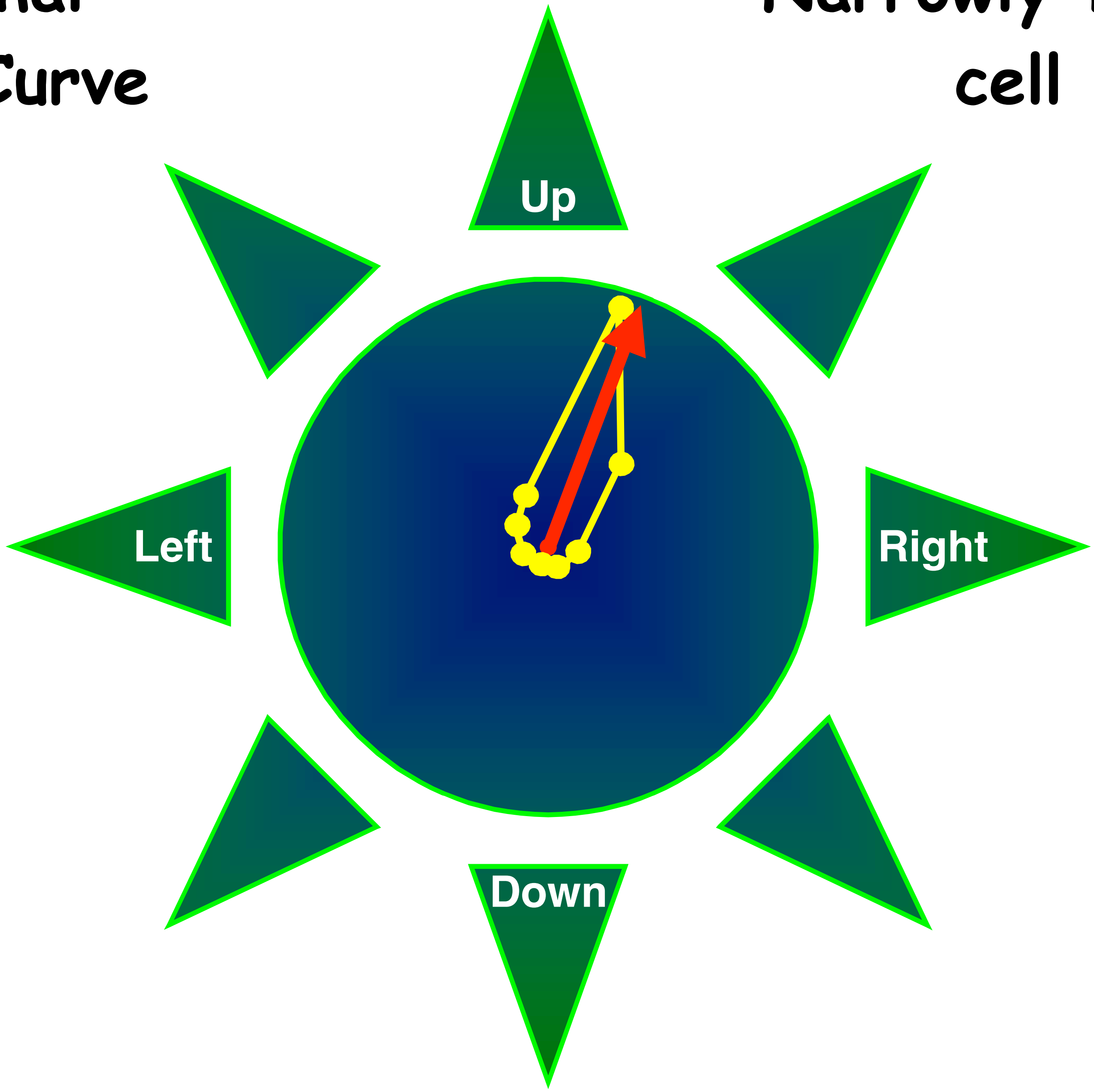
Directional Tuning Curve

Broadly tuned cell

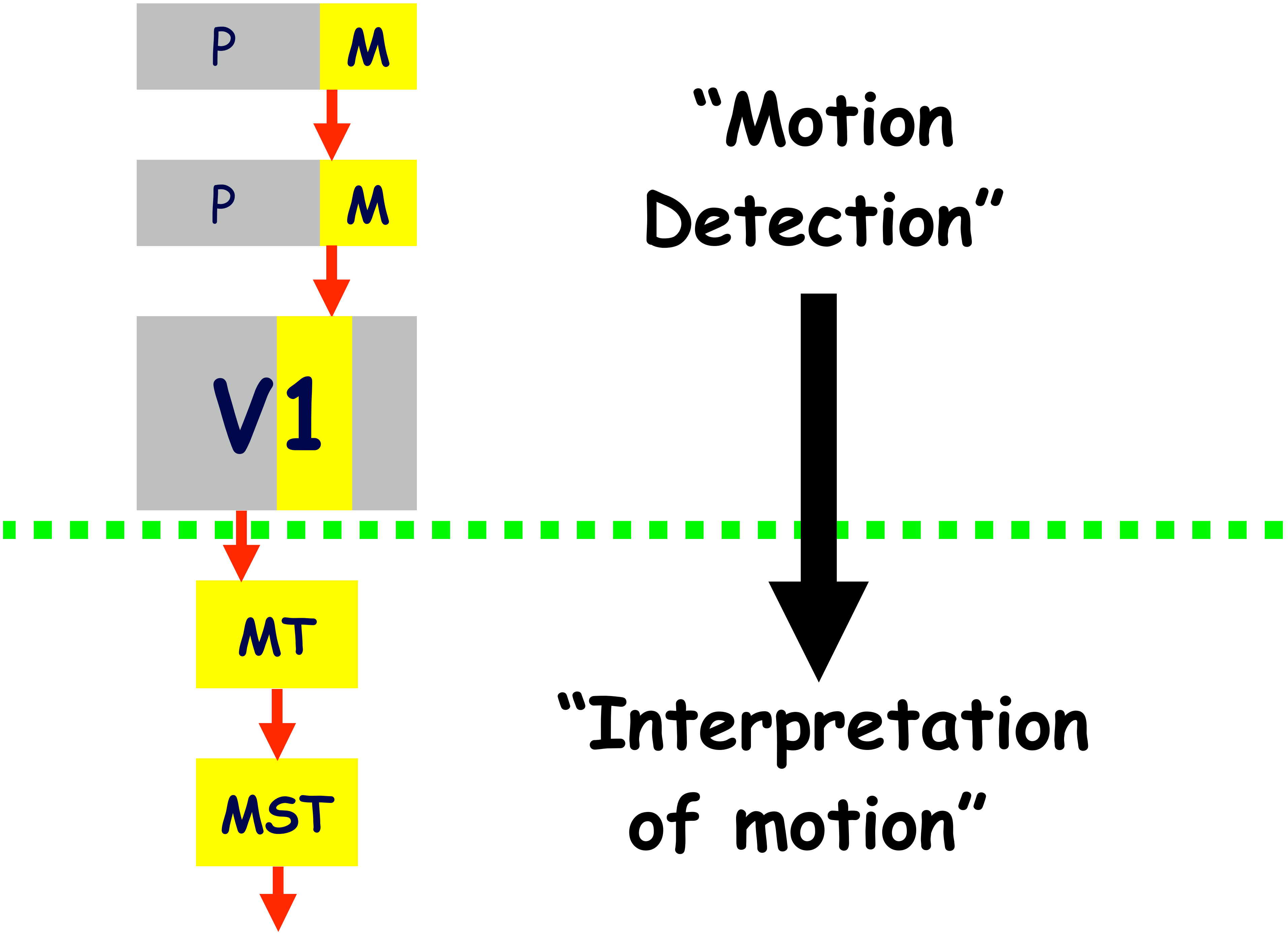


Directional Tuning Curve

Narrowly tuned cell

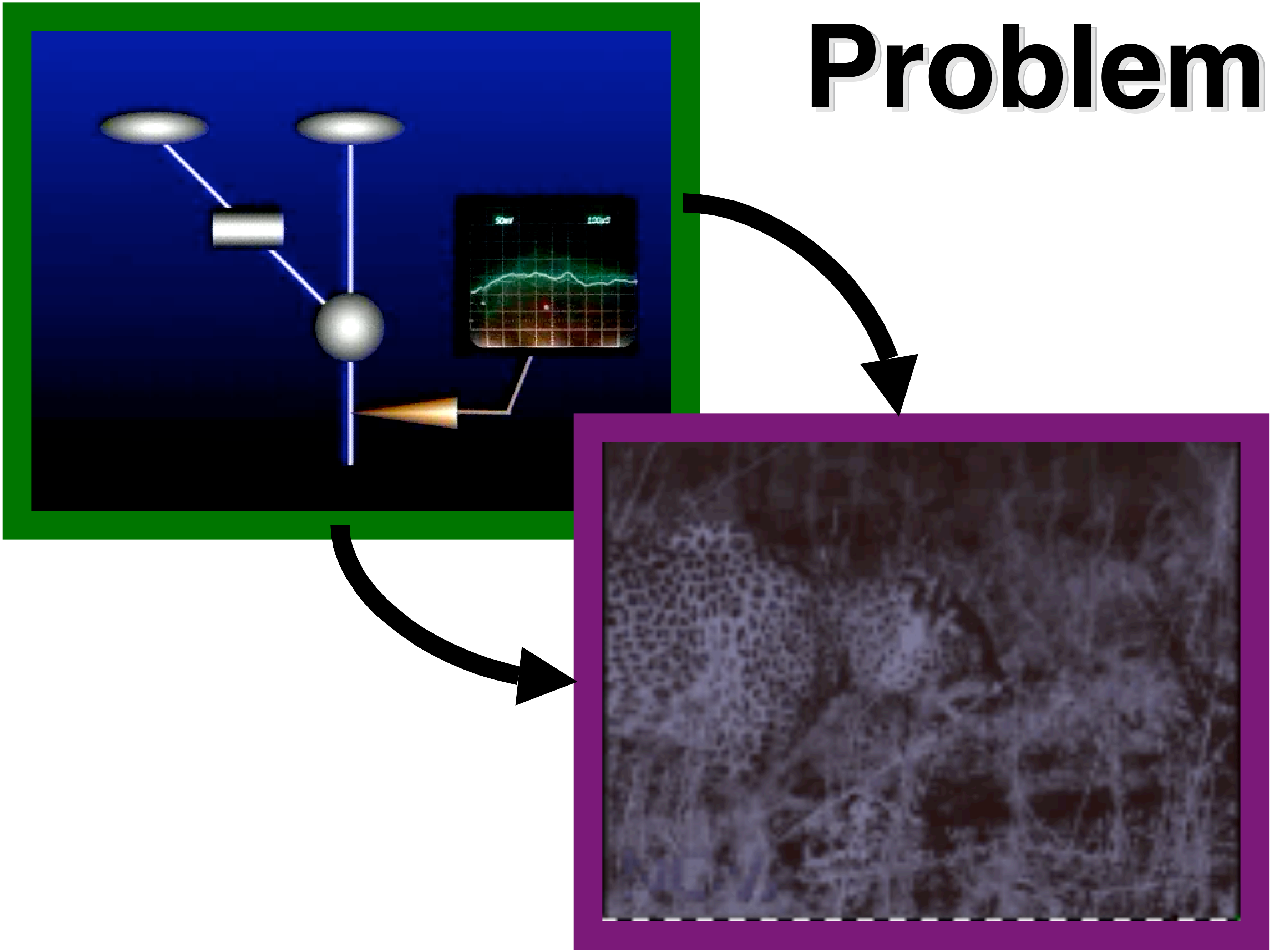


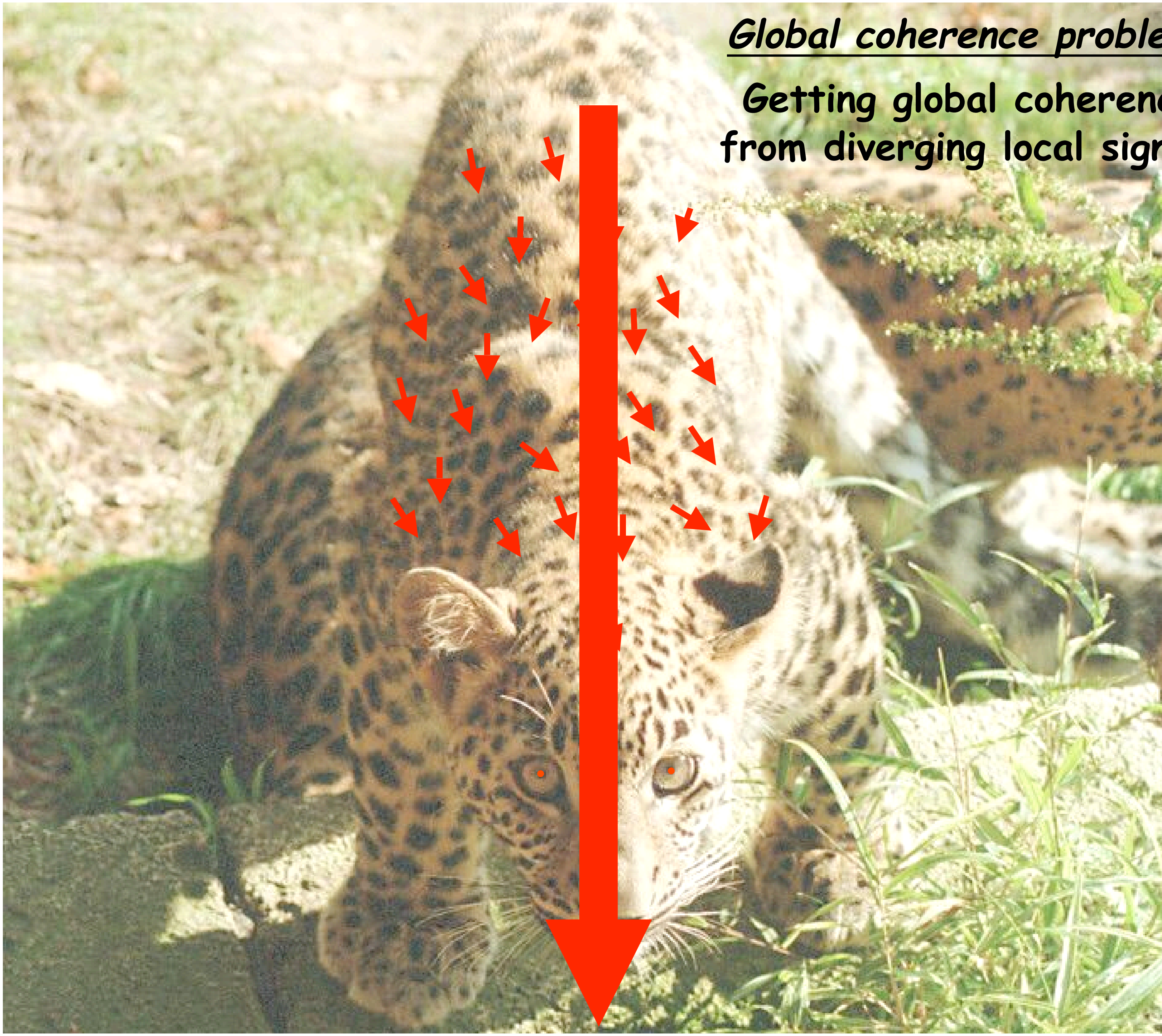
Interpretation of motion signals

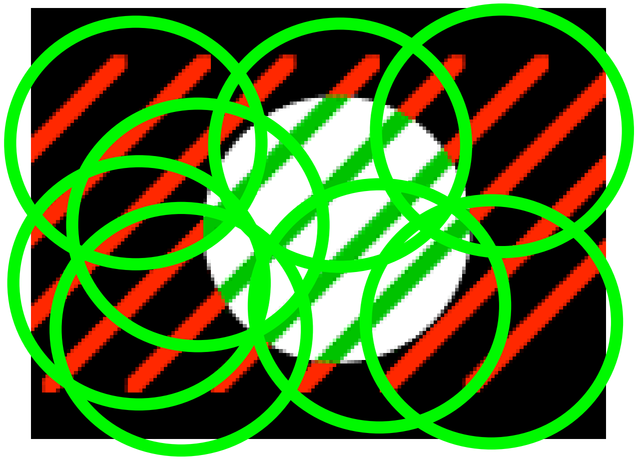


P
M
P
M
V1
MT
MST
"Motion Detection"
"Interpretation of motion"

Problem

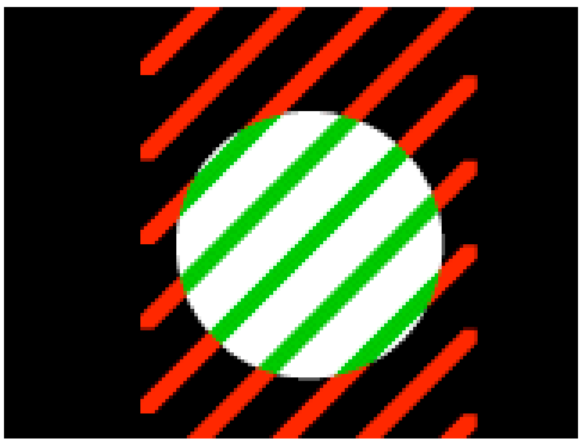



Global coherence problem:
Getting global coherence
from diverging local signals



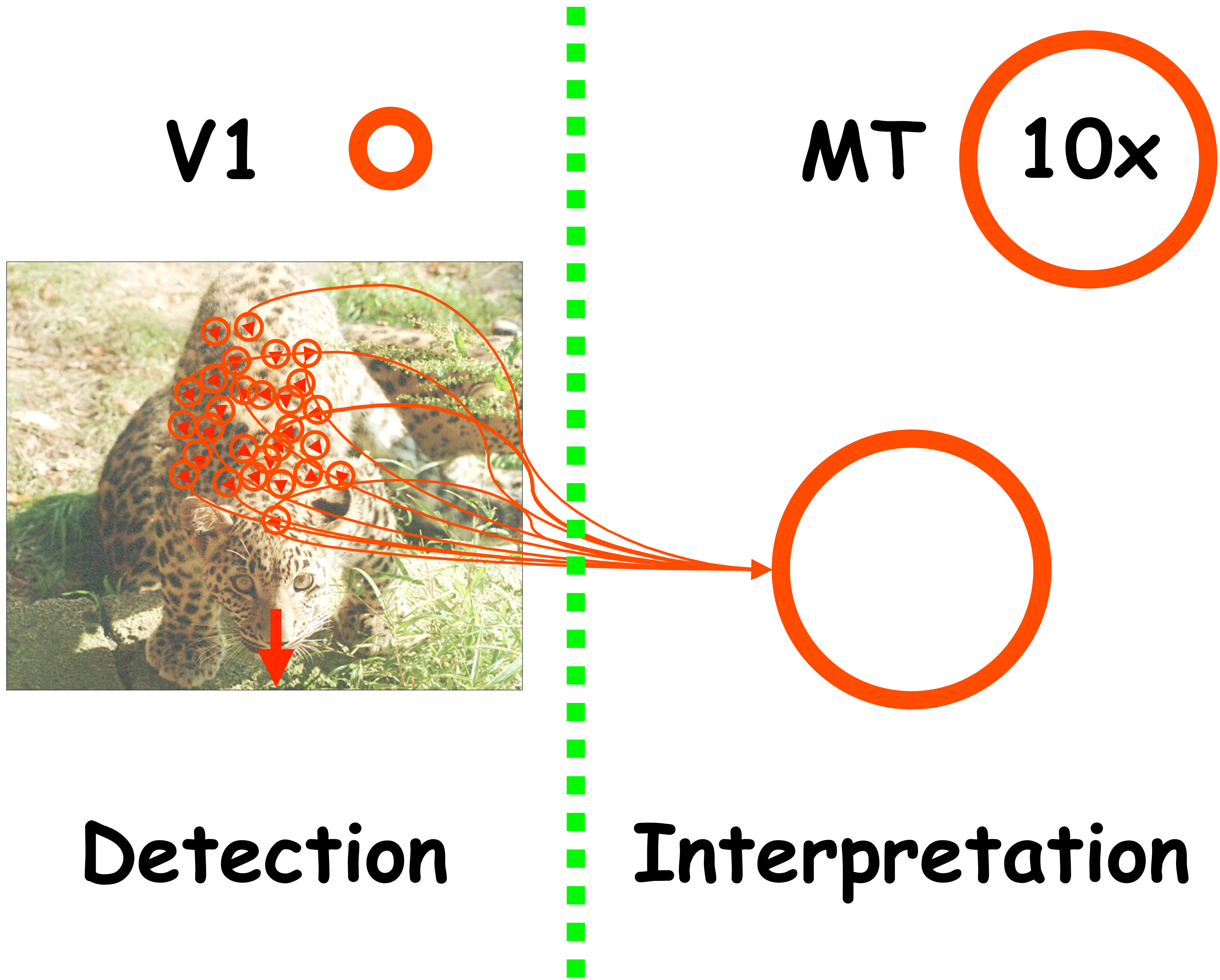
Local Ambiguity:

the "aperture problem"



Moral:

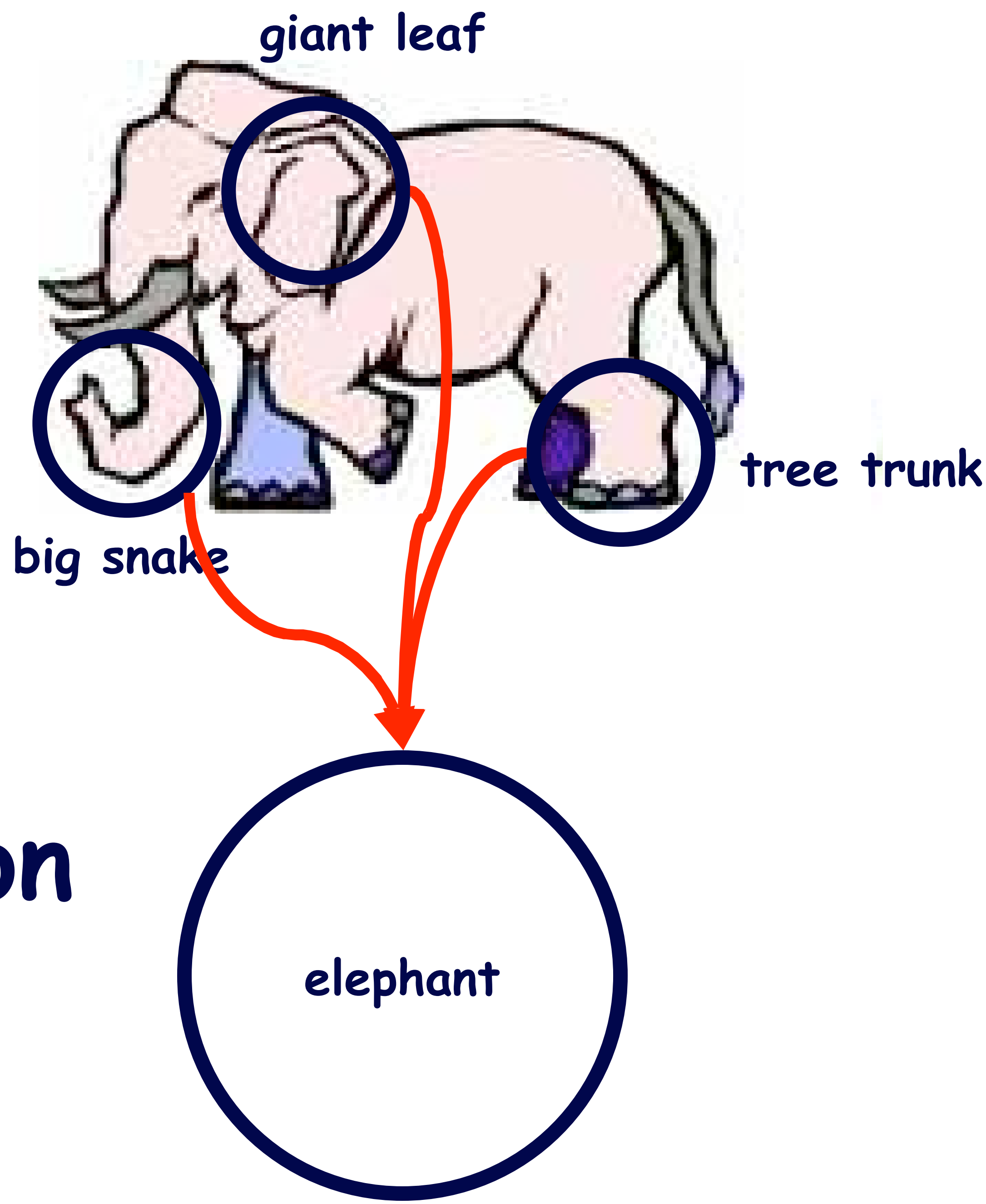
The response of one neuron is ambiguous. Combination of outputs among many neurons is necessary to resolve the ambiguity.



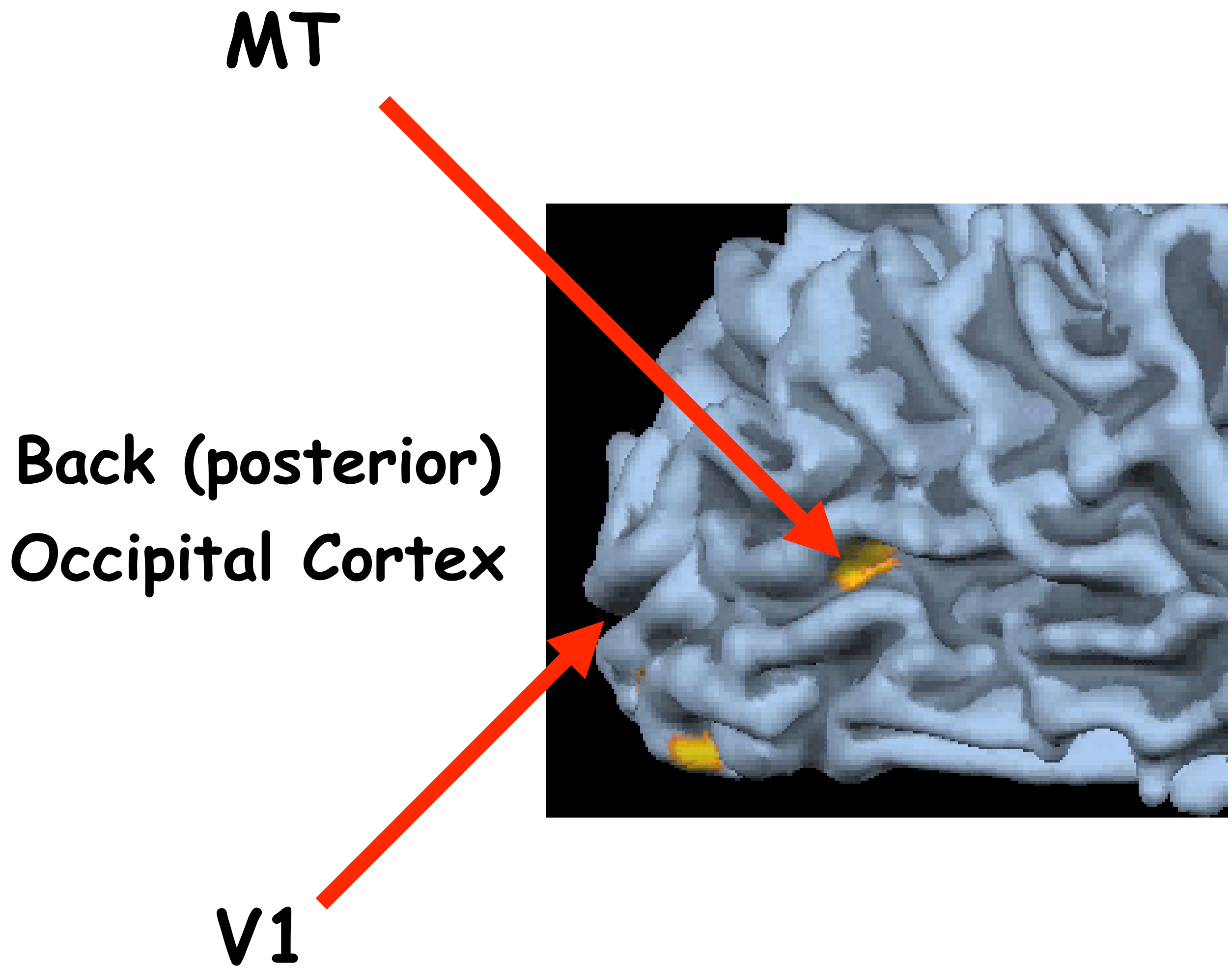
V1
MT
10x
Detection
Interpretation

Detection
V1

Interpretation
MT



Human Brain



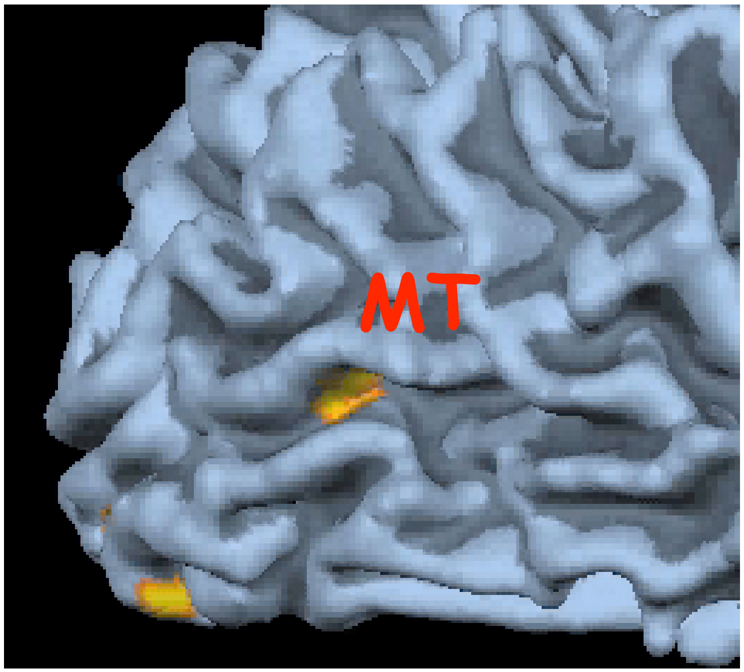
Intermediate Conclusions

~ Integration of motion signals is necessary to resolve the ambiguities in the responses of V1 neurons

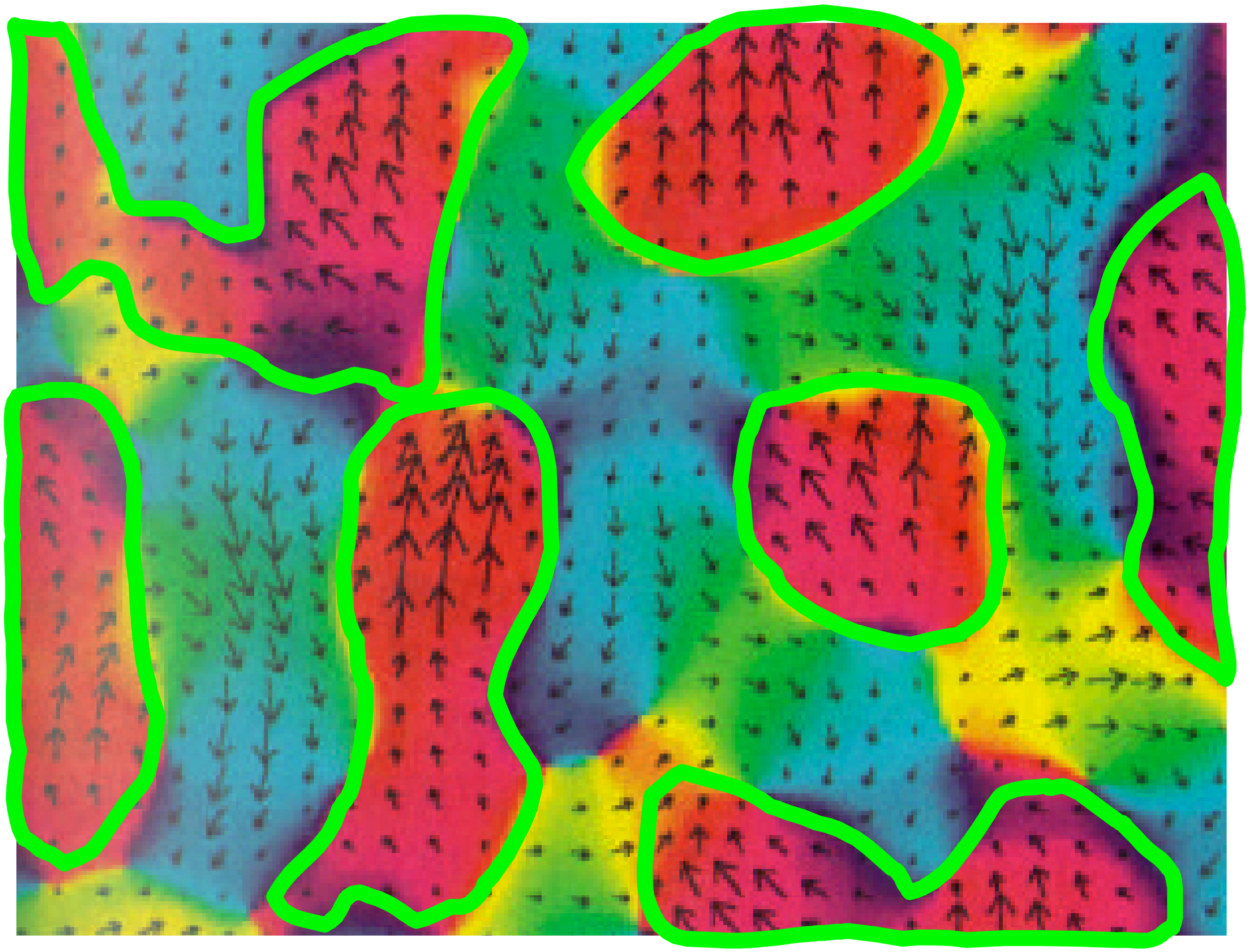
~ This integration is thought to occur in MT - the main motion area

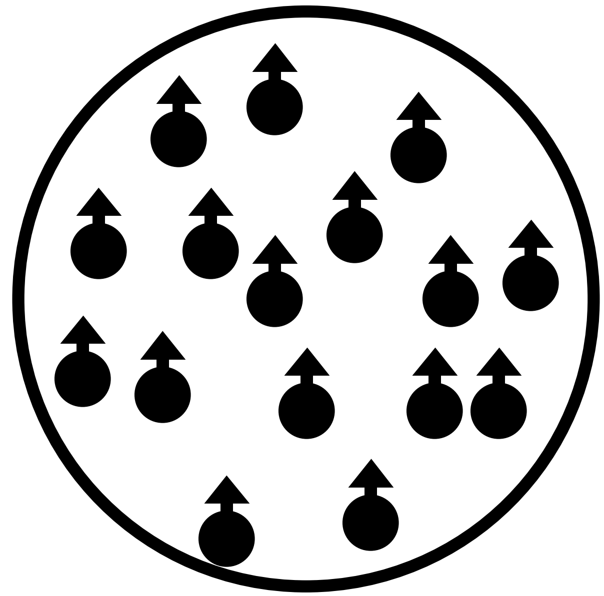
~ Why do we think MT is the main motion area?

Why do we think MT is the main motion area?



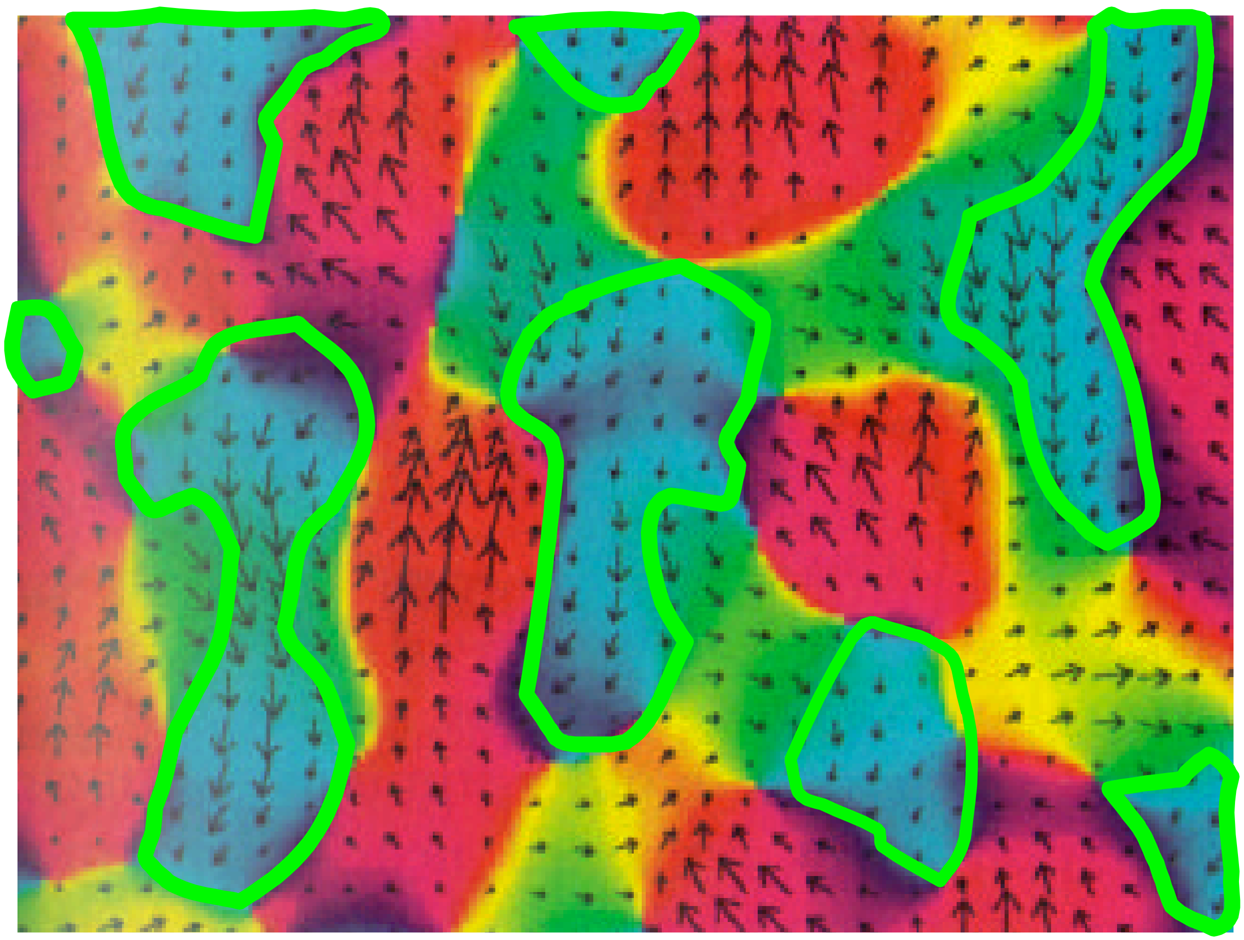
Directional maps in MT (using optical imaging)

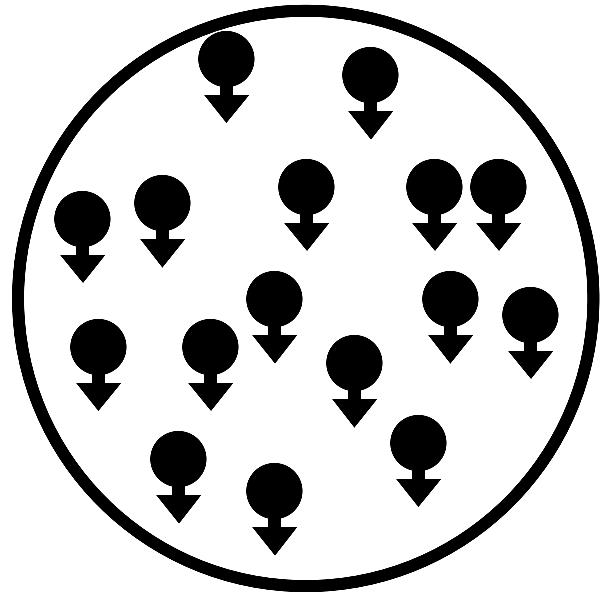




Upward motion

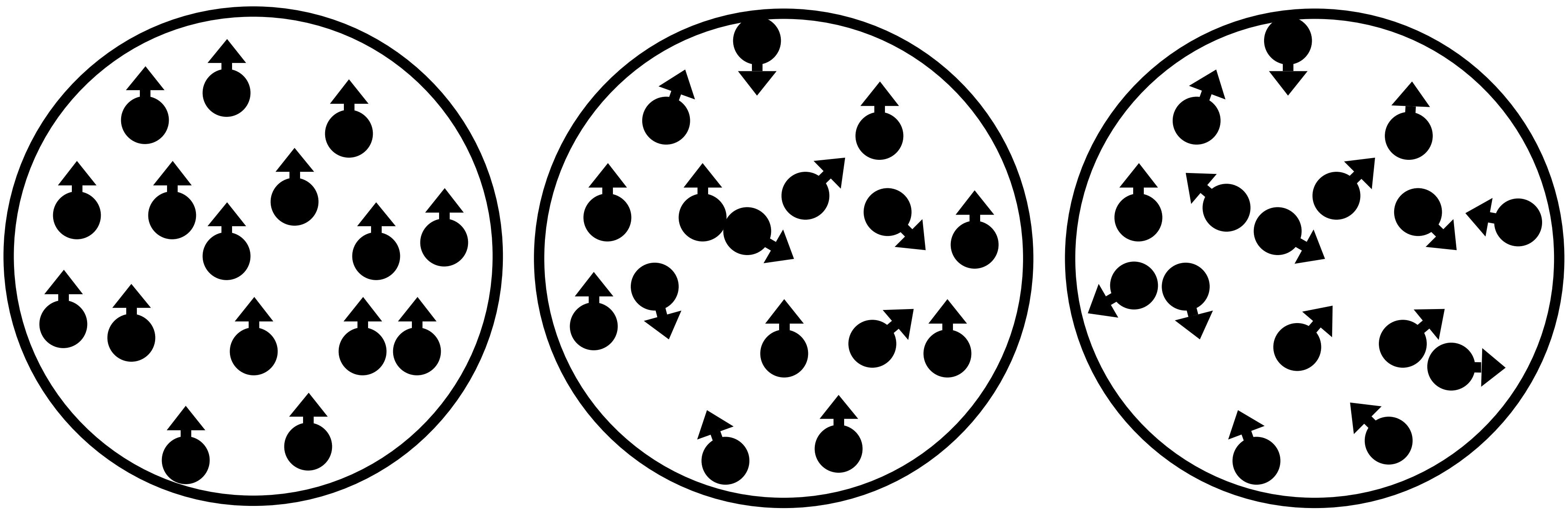
Directional maps in MT (using optical imaging)





Downward motion

Random-dot stimuli
(thresholds are around 5-10%)

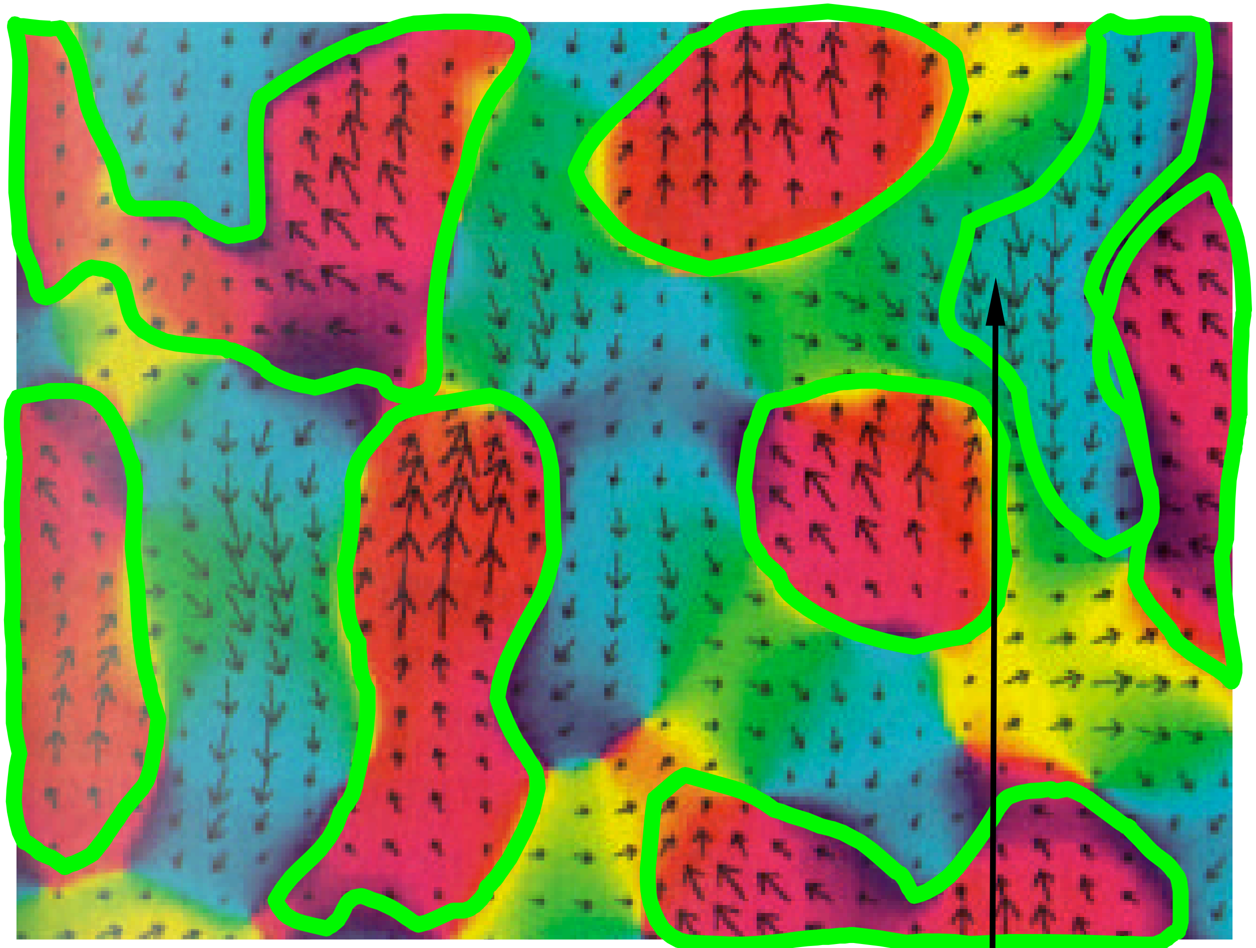


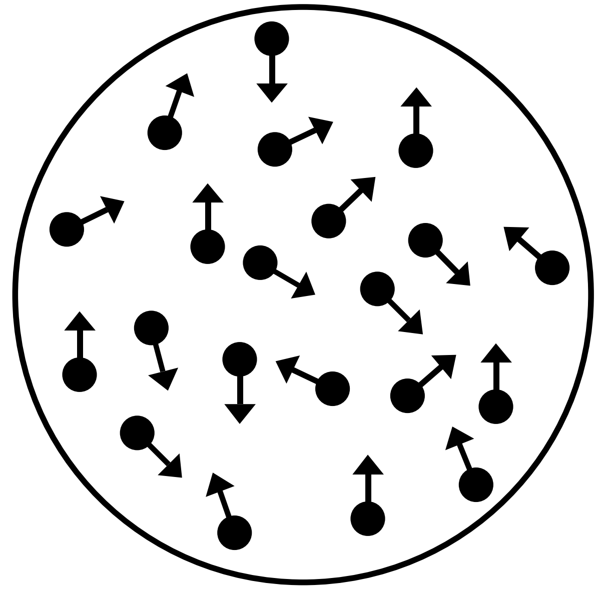
100% correlation

50% correlation

0% correlation

Salzman, Britten & Newsome, 1990





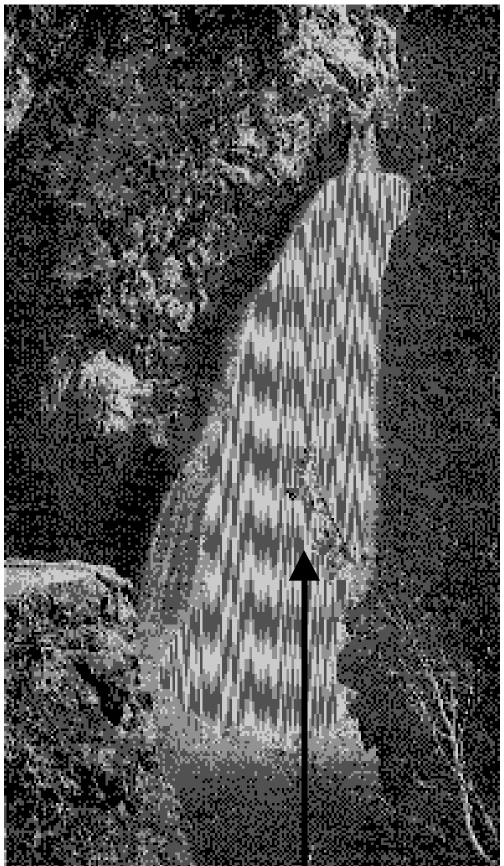
weak correlation - Up

Monkey signals perception of upward motion

Stimulate with microelectrode

Monkey signals perception of downward motion

Motion-after effect (MAE)

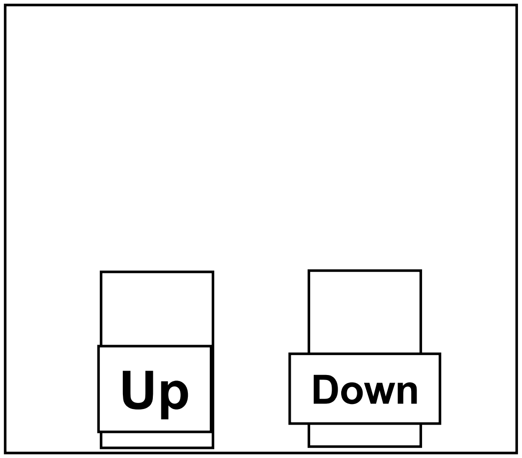


The Falls of Foyers

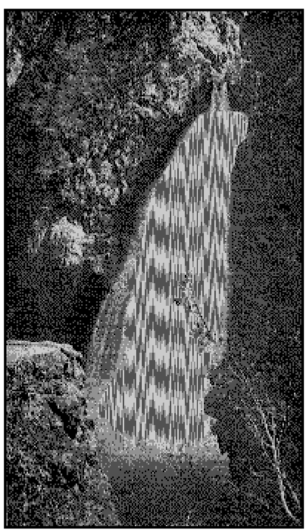
~ Initial reports by Ancient Greeks

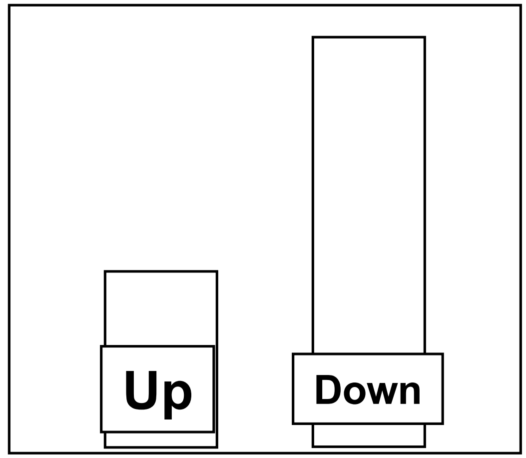
~ First modern report by Robert Adams (1834) while viewing a waterfall at Foyers, Scotland.

MAE Explanation

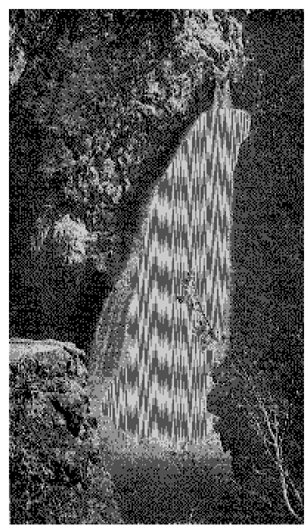


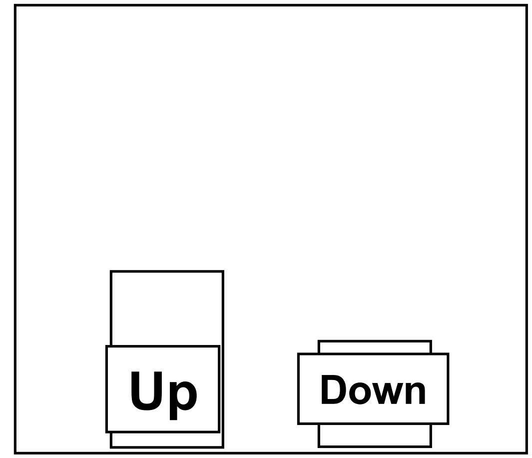
Before adaptation



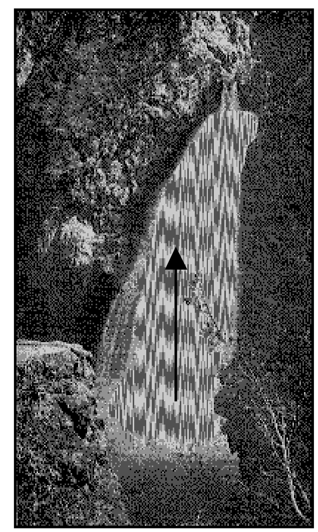


During adaptation





After adaptation



Thank you for your attention!