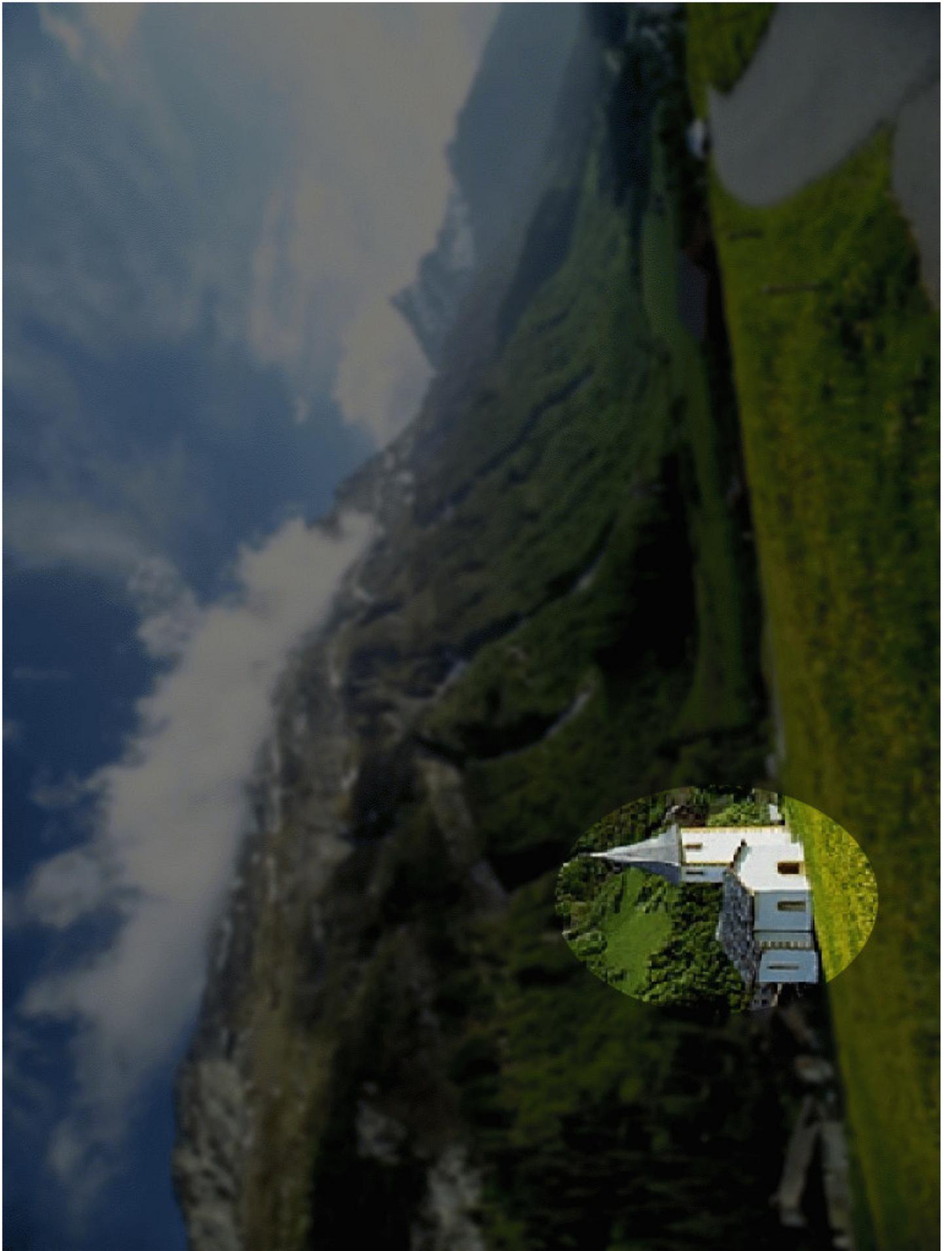




Digital Media Lab

Visual Attention







Digital Media Lab

Attention selects one modality over another



Cell phone use while driving



Where to look?

- Many visual processes are expensive
- Humans don't process the whole visual field
- How do we decide what to process?
- How can we use insights about this to make machine vision more efficient?



Selective Attention

- Perception and awareness are necessarily selective (mobile phone while driving)
- Perception is not (usually) passive, autonomous, feedforward and stimulus-driven--it is active, involves feedback, and goal-driven
- Behavioral goals determine states of perceptual selectivity, which in turn can modulate sensory input via cortical feedback
- Attention is the set of cognitive mechanisms that implement the selection of relevant perceptual input and the rejection of irrelevant input



Why can't we see and experience everything at once? That is, why is selective attention necessary? Here are two candidate answers

1. At a computational level, the brain has a limited capacity to process information (limited, for example, in storage capacity and in the number of computational units). The goal of perception—to accurately recognize or categorize an almost infinite variety of inputs very rapidly—requires suitable constraints to be tractable. Attention limits which inputs the system attempts to recognize or categorize.
2. Action—for example, locomotion, reaching movements, and eye movements—are also limited by biomechanical constraints to only one or a few at a time, in an appropriate sequence (cf. Karl Lashley, “The Problem of Serial Order in Behavior”), and because there is a need for perception-action correspondence, the scope of perception is limited in order to guide desired actions.



Three Modes of Attentional Control

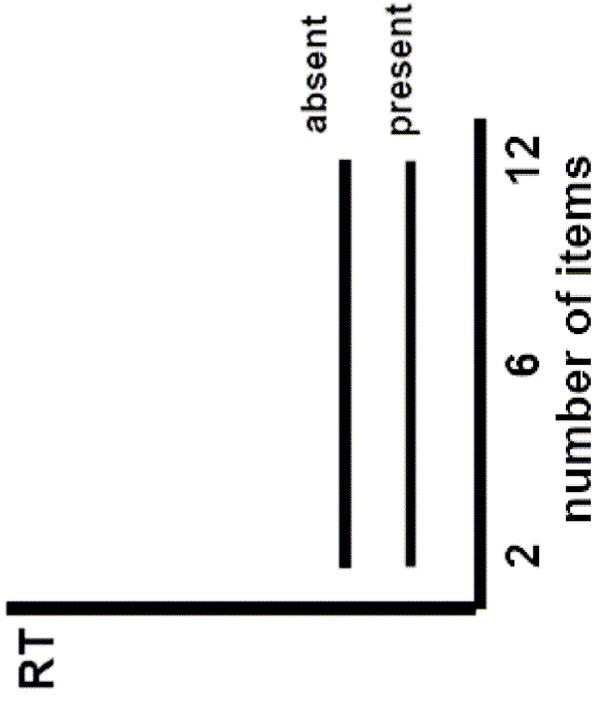
- Stimulus-driven, bottom-up, involuntary, automatic
- Goal-driven, top-down, voluntary, deliberate
- Hybrid: a combination of top-down and bottom-up control.



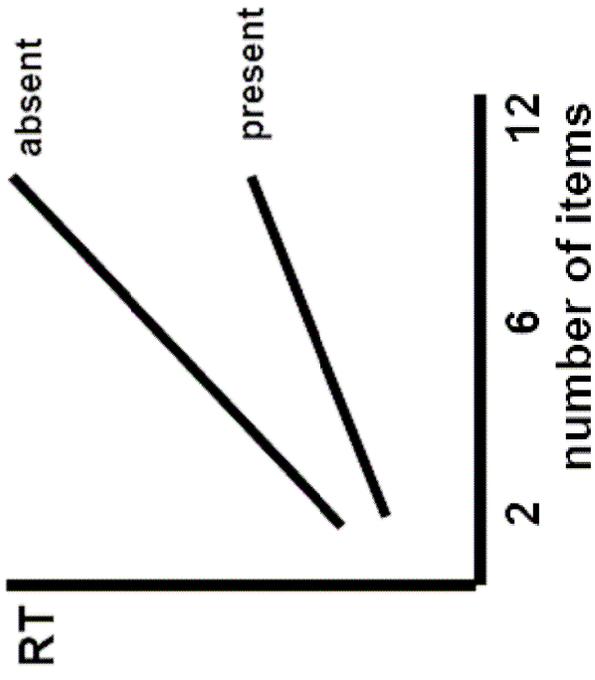
Digital Media Lab

Searching for a Single Target Amongst Distractors

PARALLEL SEARCH
(PREATTENTIVE)



SERIAL SEARCH
(FOCUSSED ATTENTION)





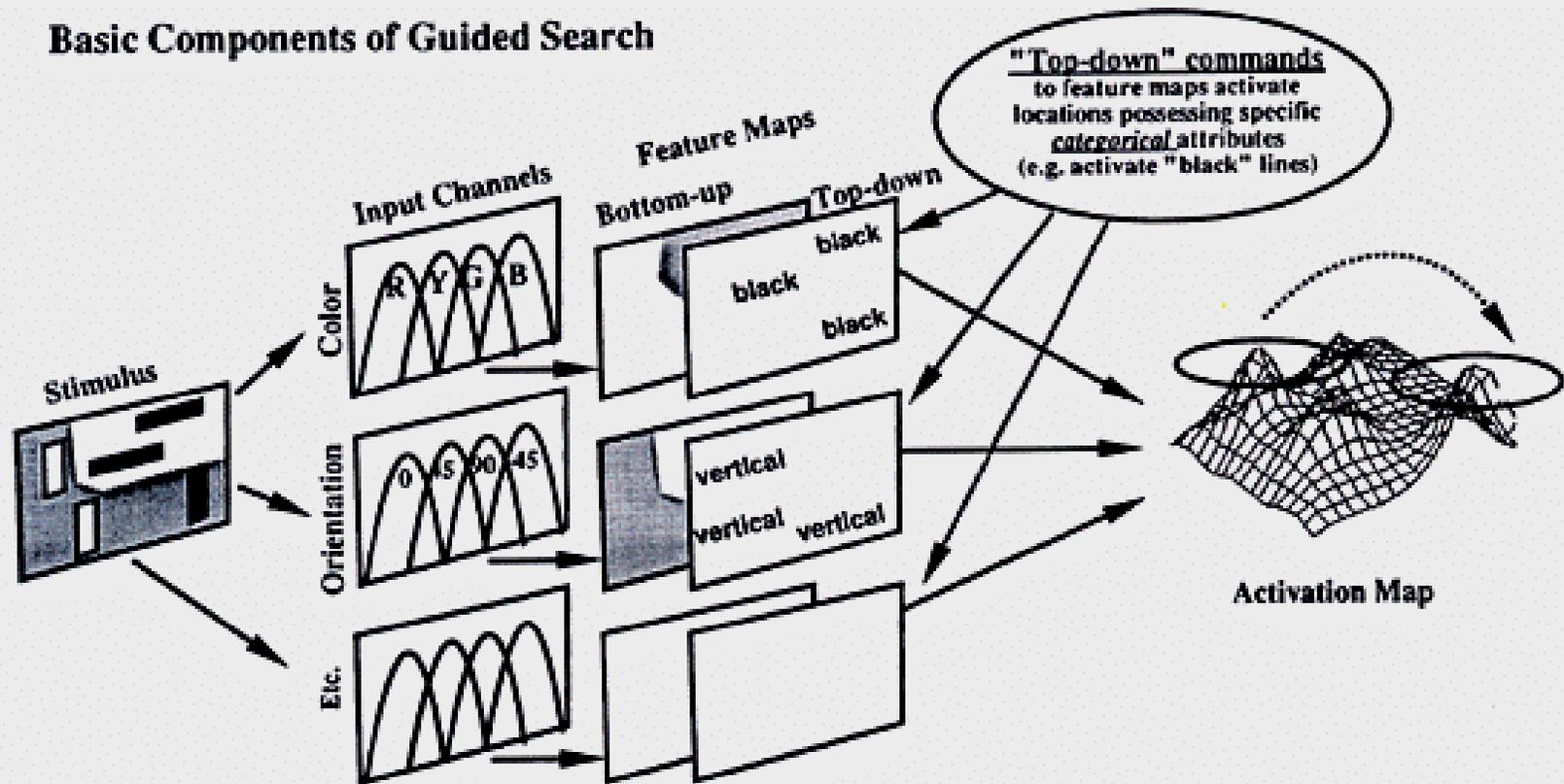
Four Domains of Perceptual Selectivity

- **Space:** the “spotlight of attention”
- **Features:** color, motion, orientation, etc.
- **Objects:** overlap, occlusion, transparency, grouping, segmentation, etc.
- **Sensory Modalities:**
 - Vision
 - Audition
 - Touch

Guided Search

Wolfe (1994)

Basic Components of Guided Search





Guided Search

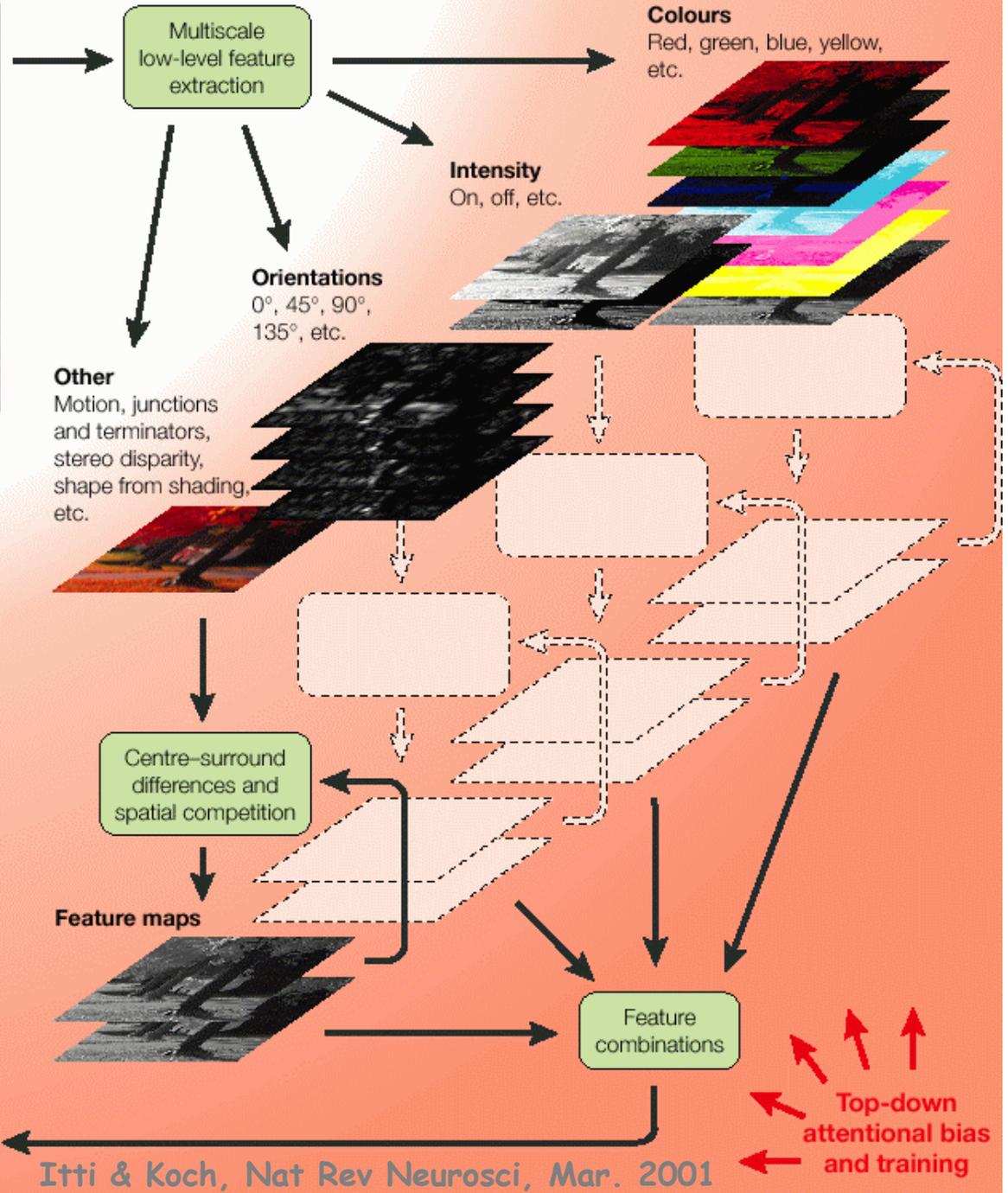
- Observed that some conjunction searches can be efficient.
- Suggested that these searches are not random, but can be guided by
 - salience as defined by local feature contrast
 - top-down goals as specified by the current search set



Visual salience

- Salience ~ visual prominence
- Must be cheap to calculate
- Related to features that we collect from very early stages of visual processing
- Colour, orientation, intensity change and motion are all important indicators of salience

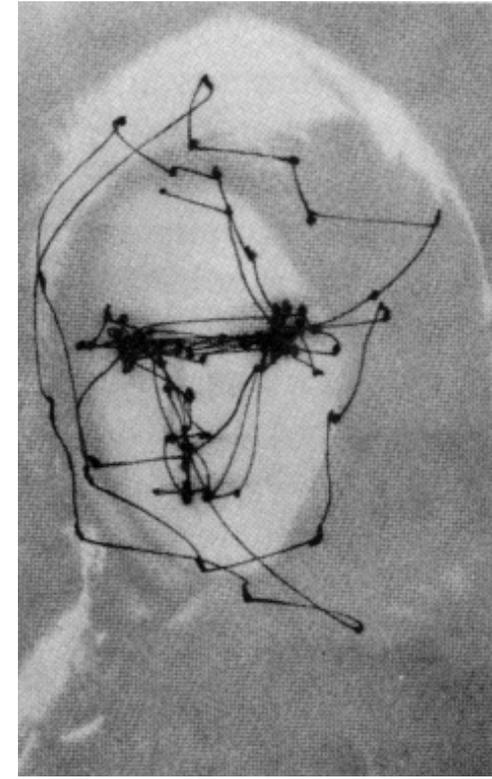
Input image



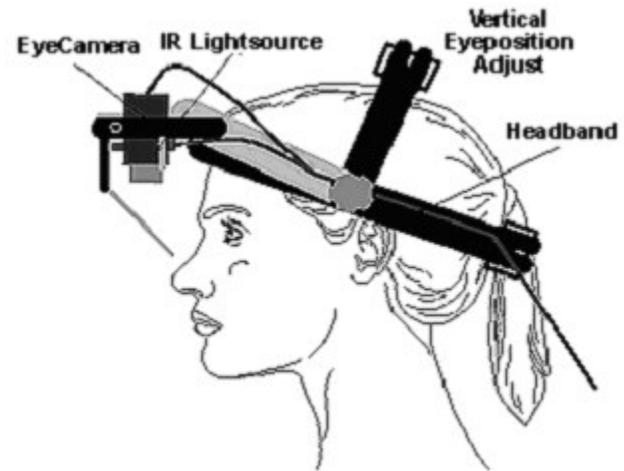
Itti & Koch, Nat Rev Neurosci, Mar. 2001

Eye Movement Experimental

Alfred L. Yarbus was a Russian psychologist (known there as Alfred Lukjanovic Jarbus) who studied eye movements in the 1950s and 1960s.

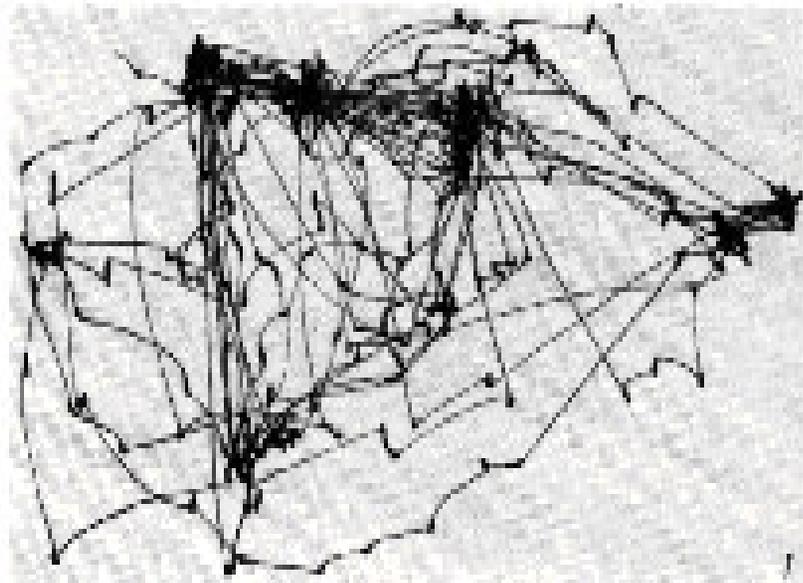


Eye Movement Experimental



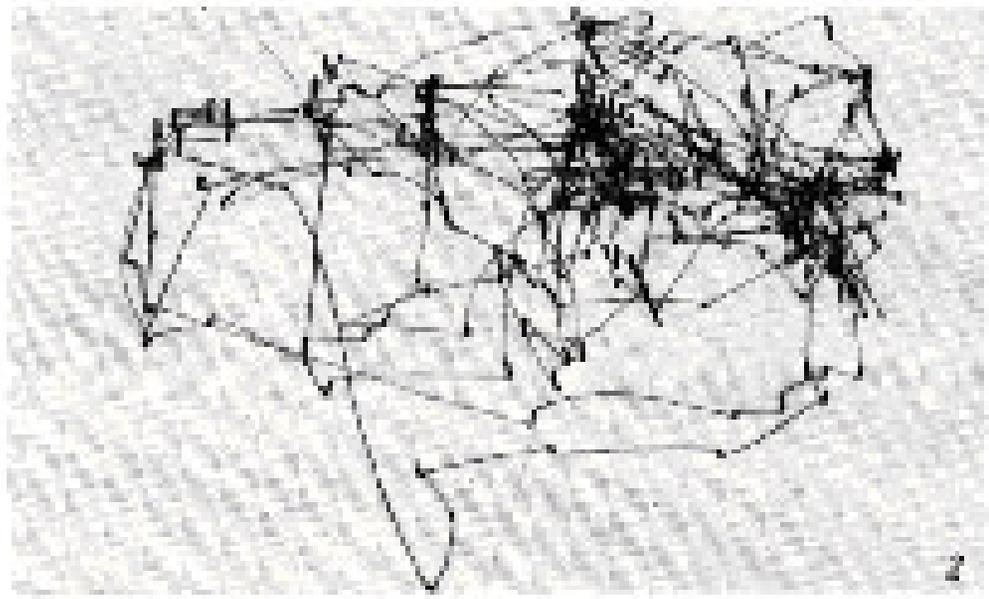
”The subject is asked to ...

” to have a free examination...”



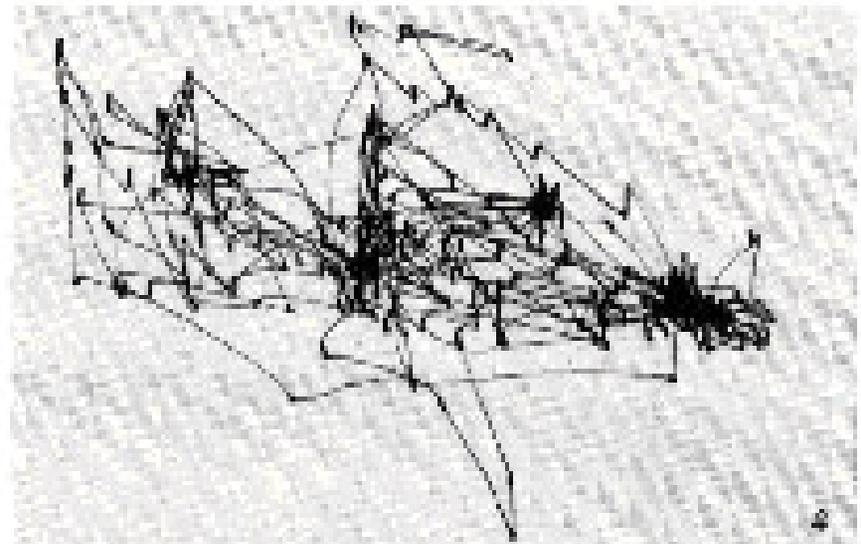


” to estimate the wealth of the family... ”





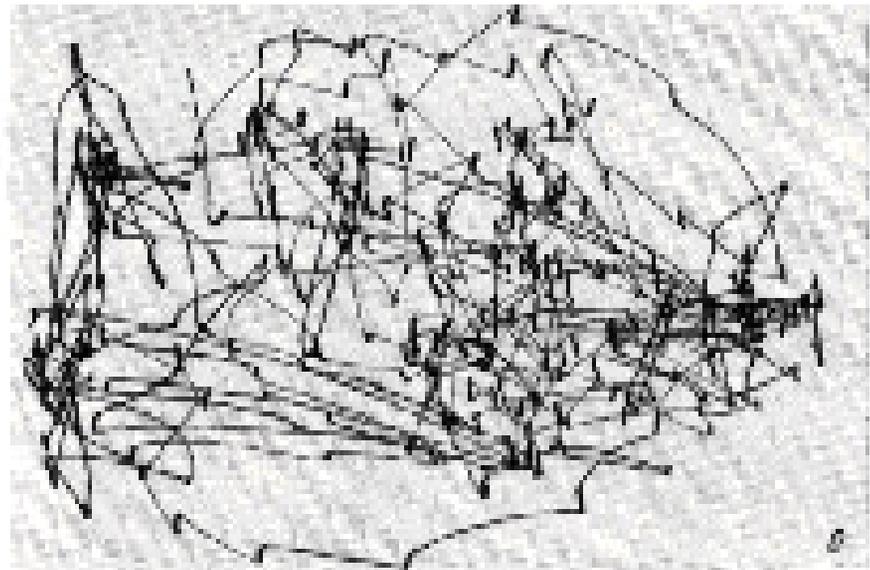
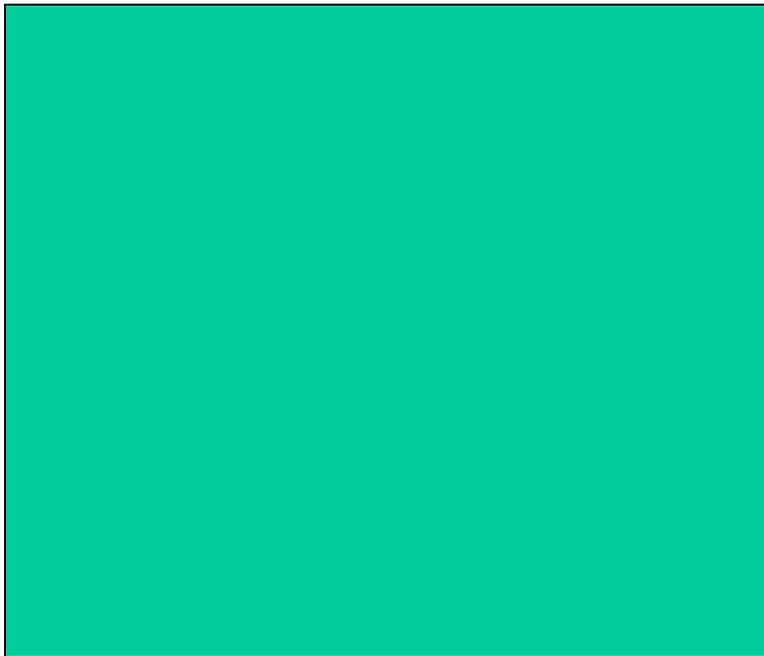
” to guess what the family had been doing before the arrival of the 'unexpected visitor'





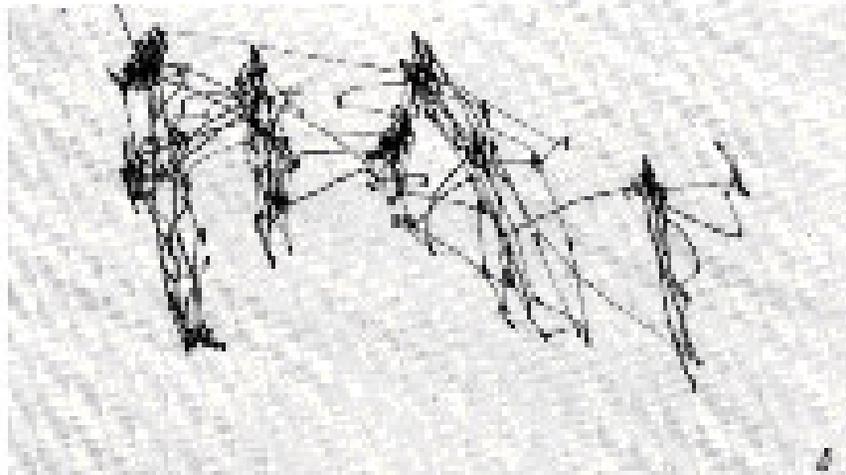
Digital Media Lab

” to memorize the location of the people and objects in the painting...



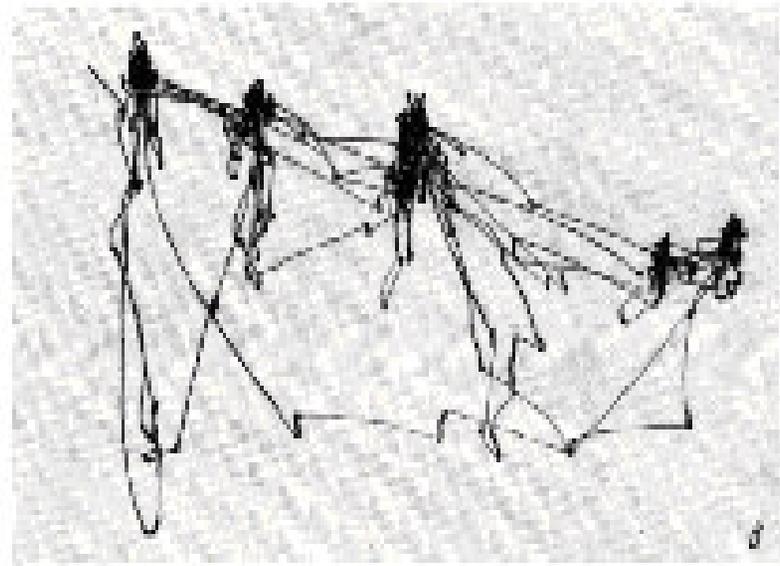


” to memorize the clothes worn by the position of the objects and people in the room.....

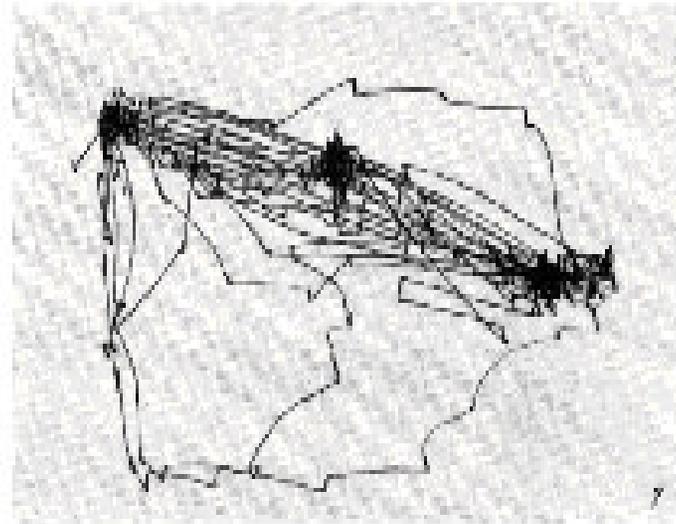




” to give the ages of the people ...”



” to estimate how long the 'unexpected visitor' had been away.”

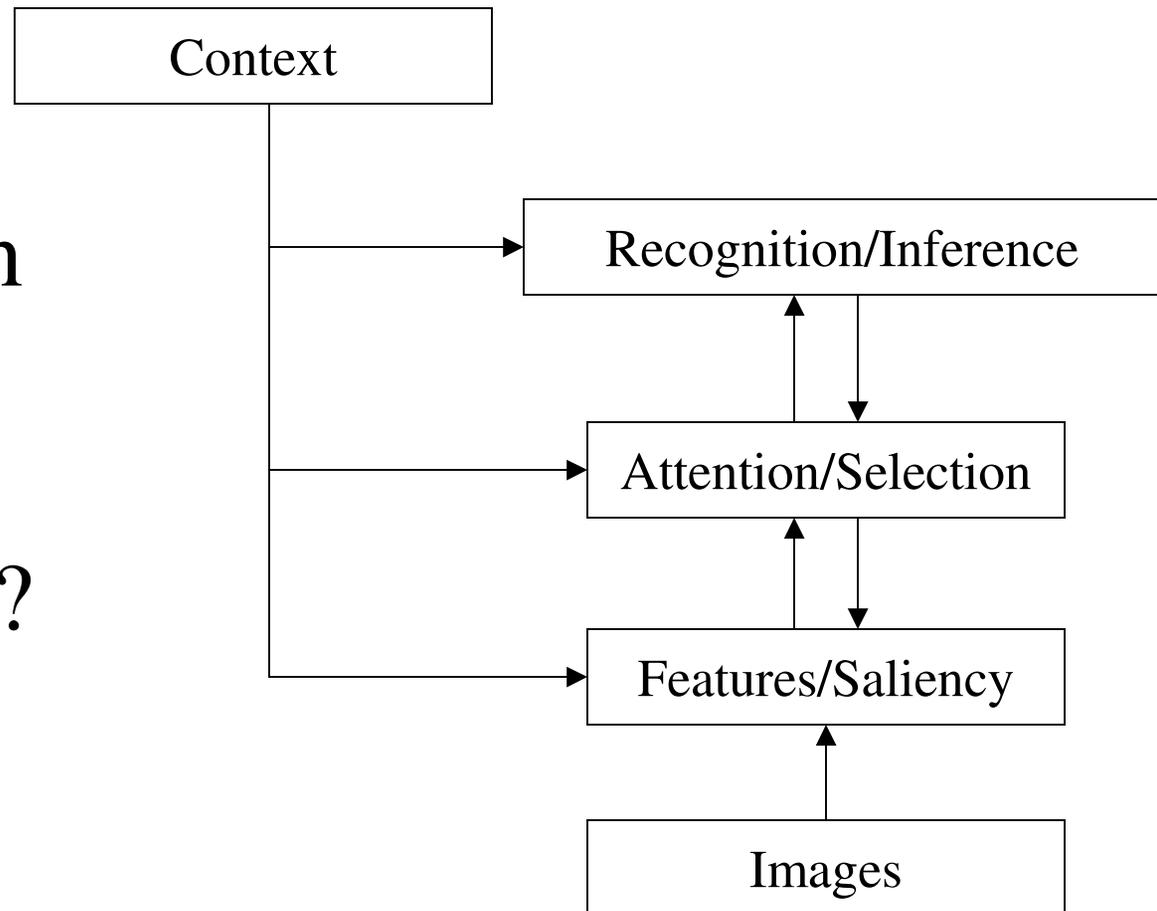


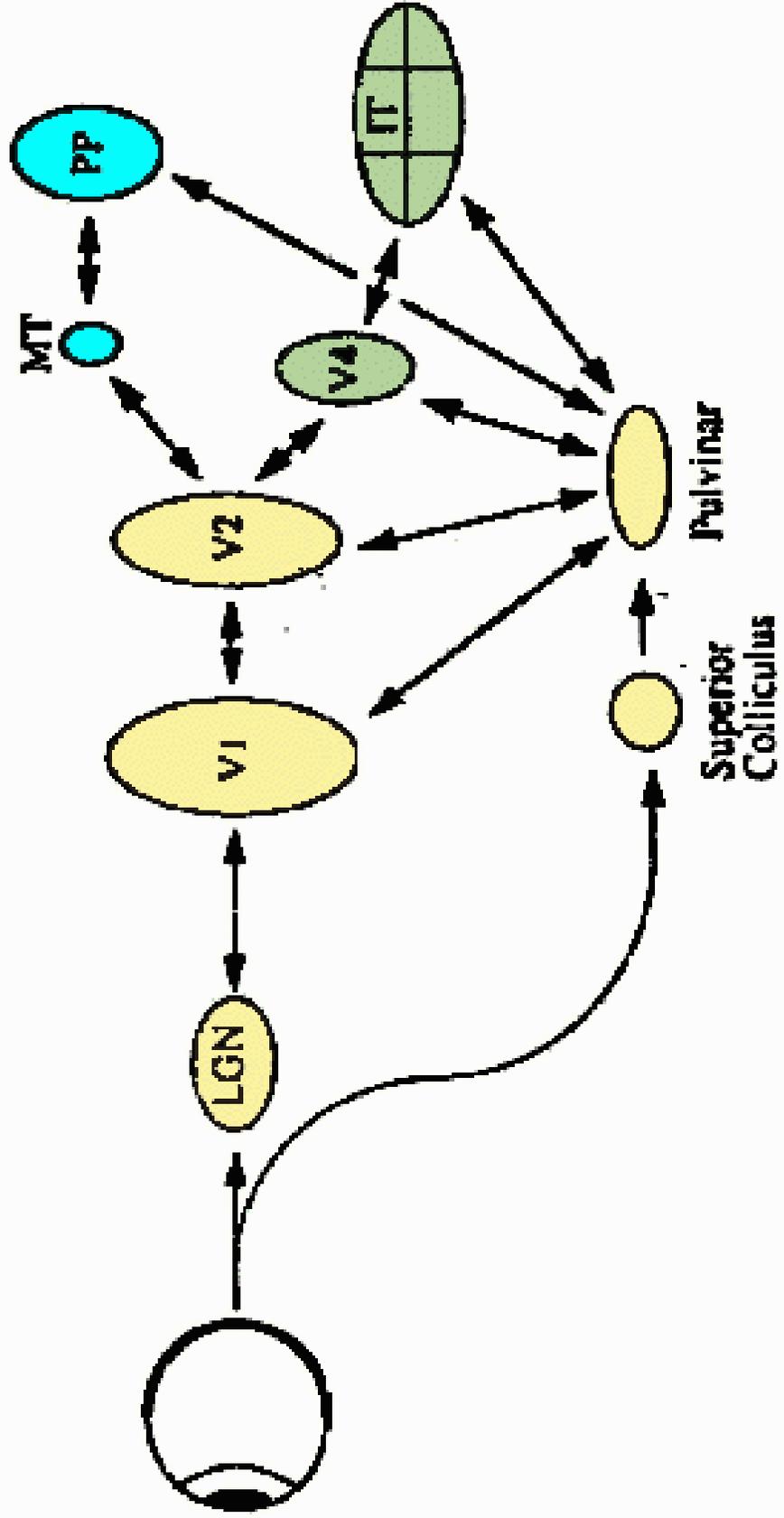


The General Vision Problem

(oversimplified)

Can all vision problems be described by this diagram ?





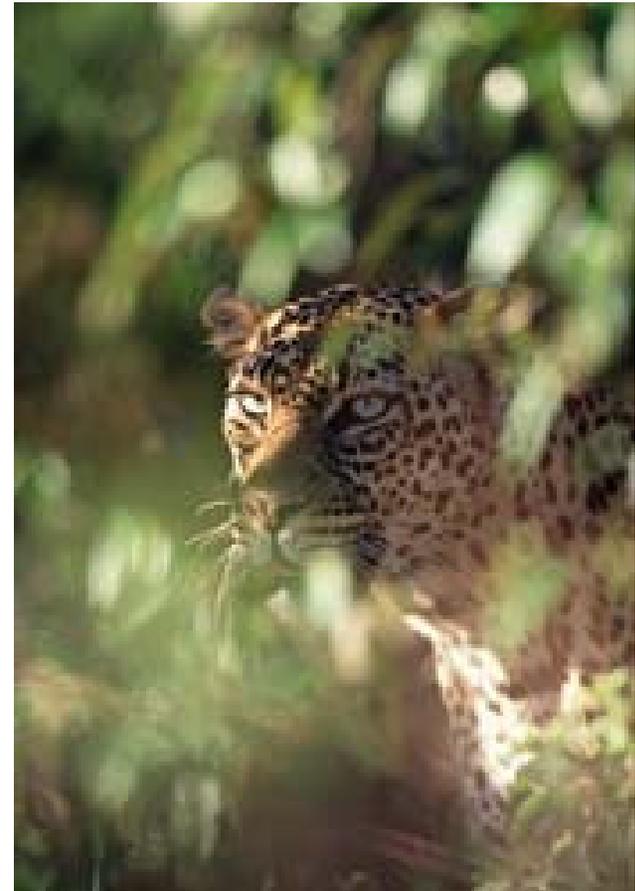


Web usability

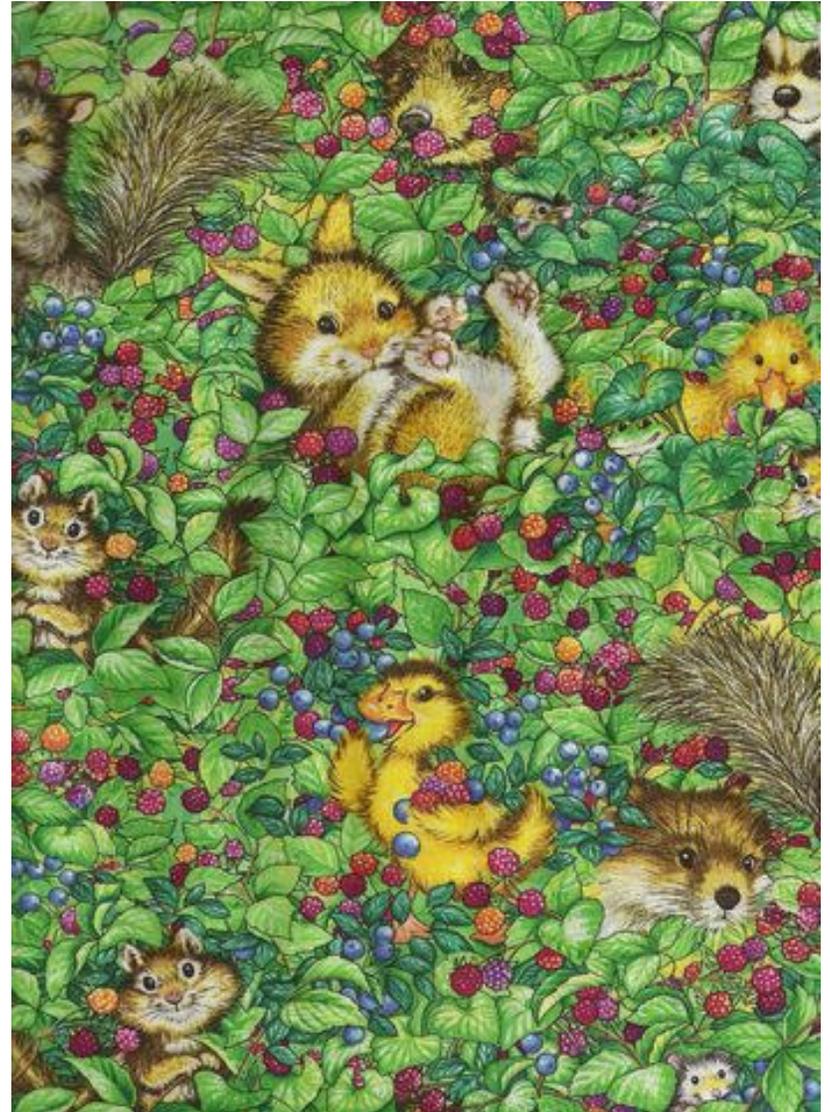


Direct Attention

Most higher animals in the world have an ability to sense danger by spotting anomalies in their environment and surviving by taking appropriate evasive action. Those organisms that have the benefit of vision are able to direct attention rapidly towards the unusual without any prior knowledge of the environment.



People and animals are able to spot anomalies in a scene no part of which they have seen before and attention is drawn in general to the anomalous Object in a scene, not to the common-or-garden or the familiar. It seems unlikely that such novelty can be sharply identified by processes which are totally dependent upon comparisons with a growing dictionary of independent recognition codes.



Why difference of Gaussians?

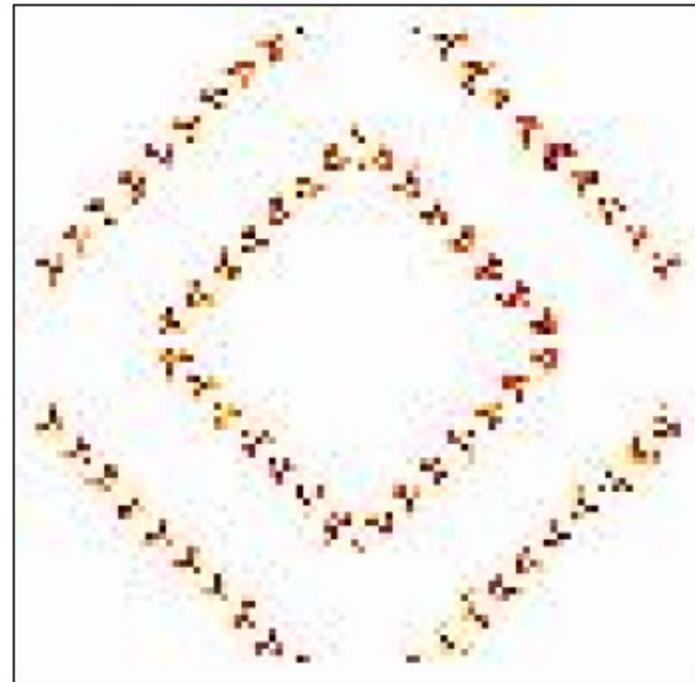
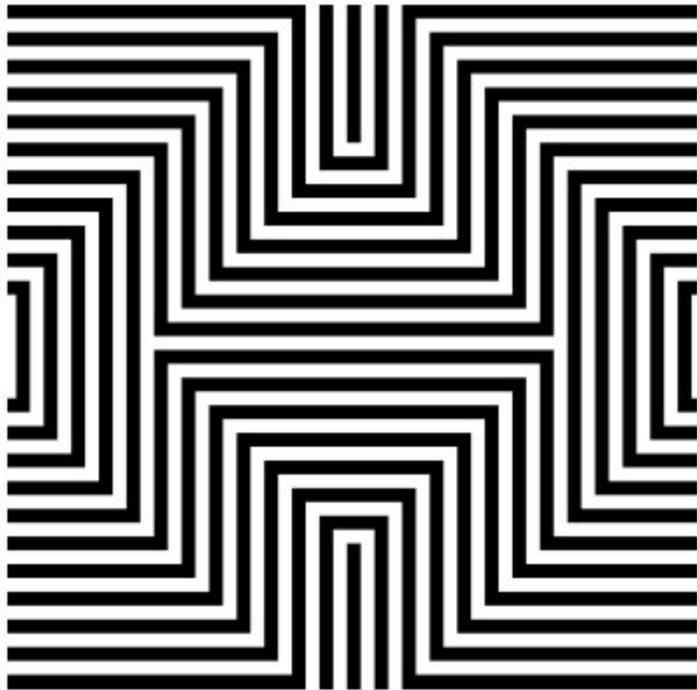




A Visual Attention Model

Visual attention is driven by the saliency. Saliency is normally strongest in a region if small neighborhoods in that region are dissimilar to neighbourhoods elsewhere in the image.

Visual Attention



Visual Attention



Figure 4a (21433 bytes)

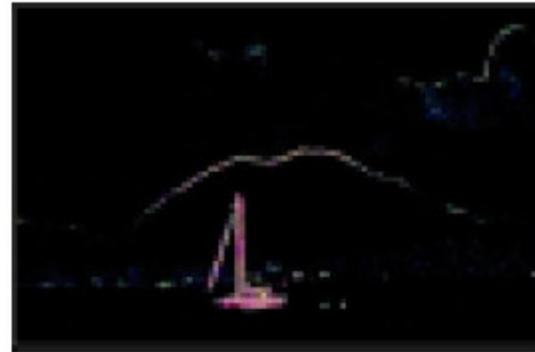


Figure 4b



Figure 5a (20672 bytes)

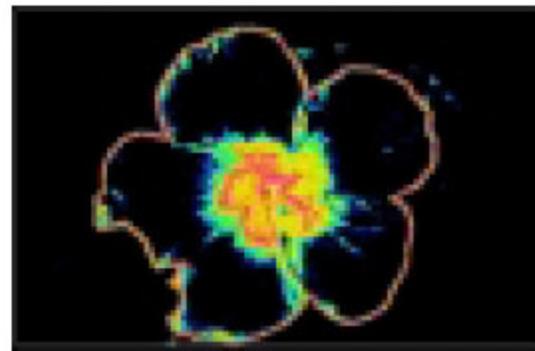
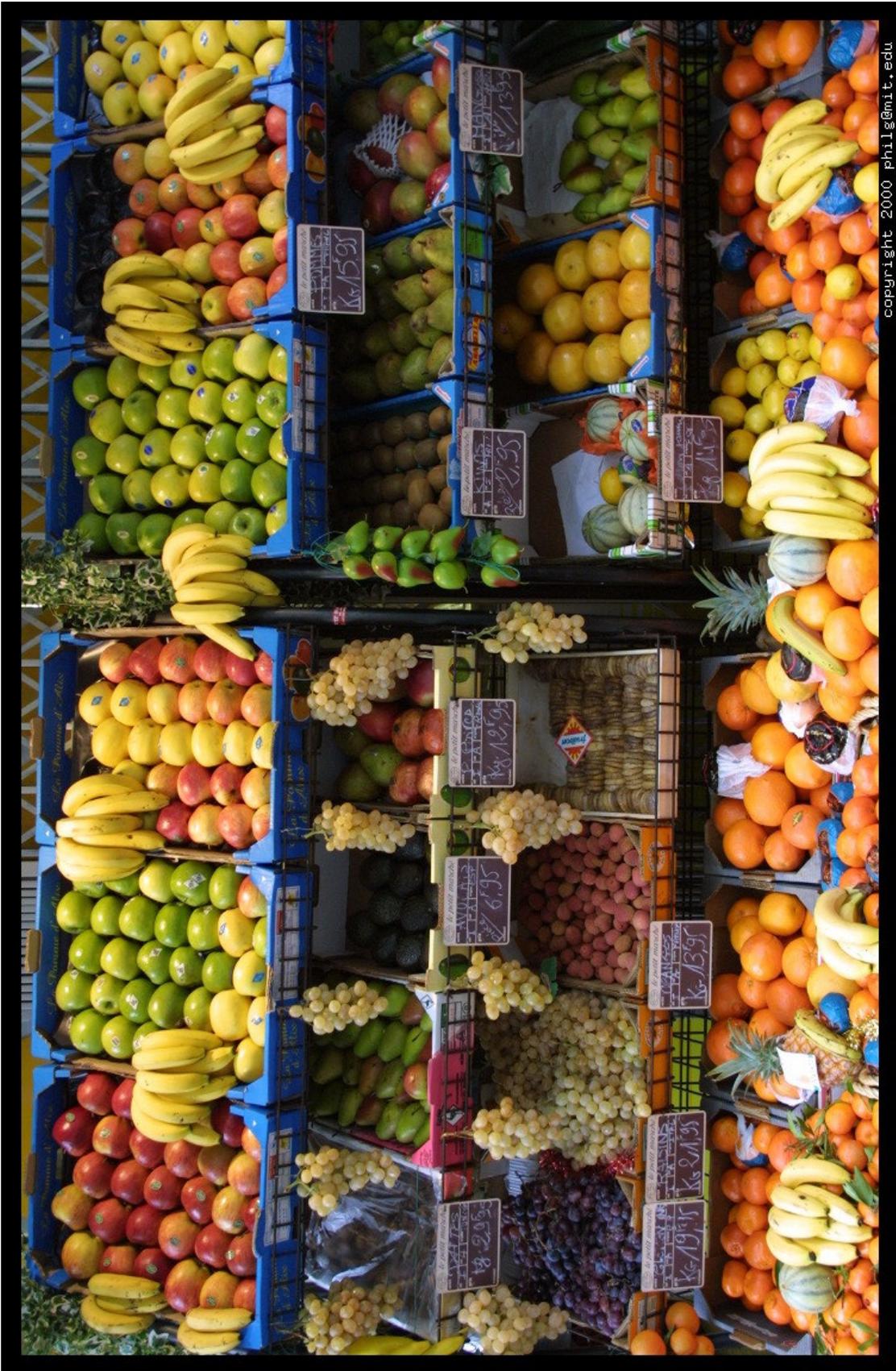


Figure 5b



Digital Media Lab



copyright 2000 phil@gmit.edu