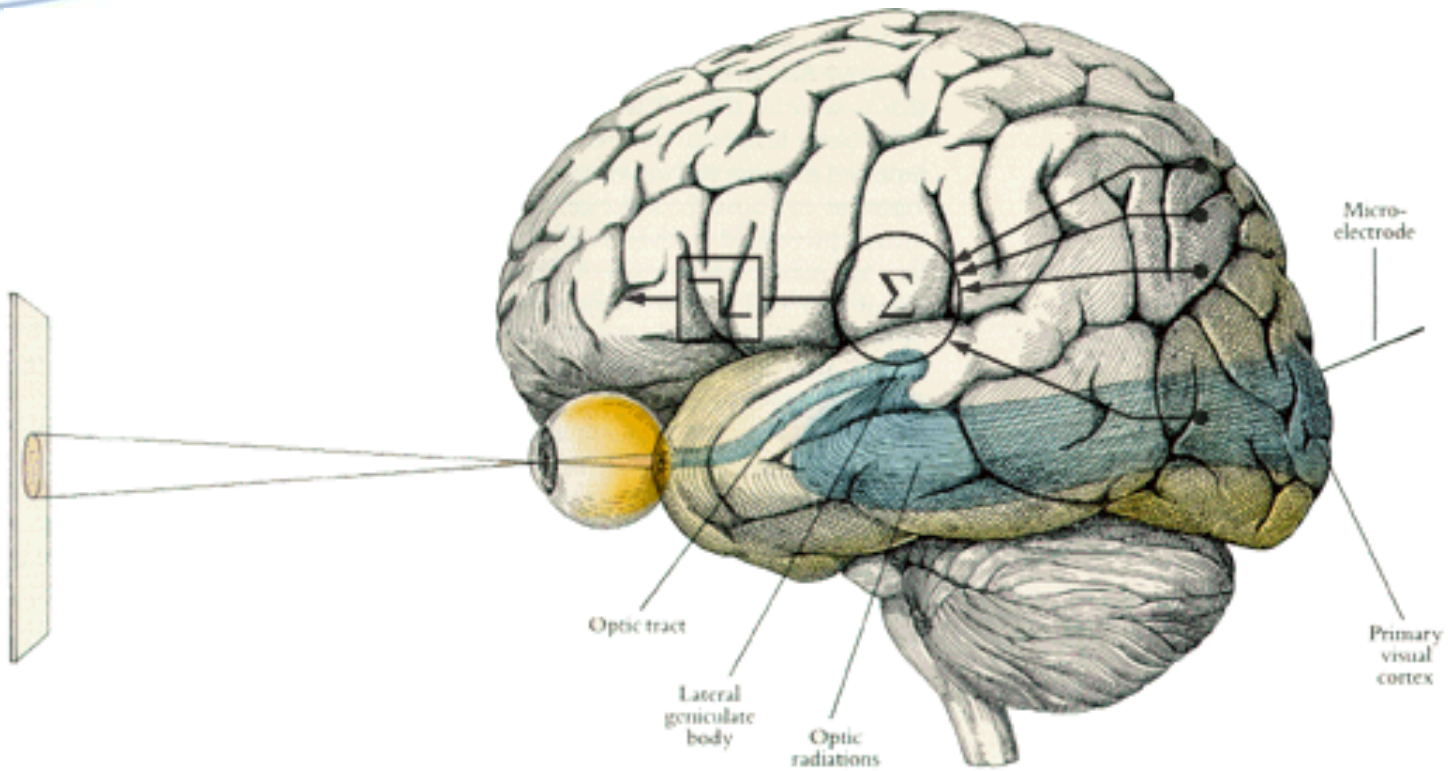
Reverse Engineering the Human Vision System

Reverse Engineering
the Human Vision System

Biologically Inspired Computer Vision Approaches

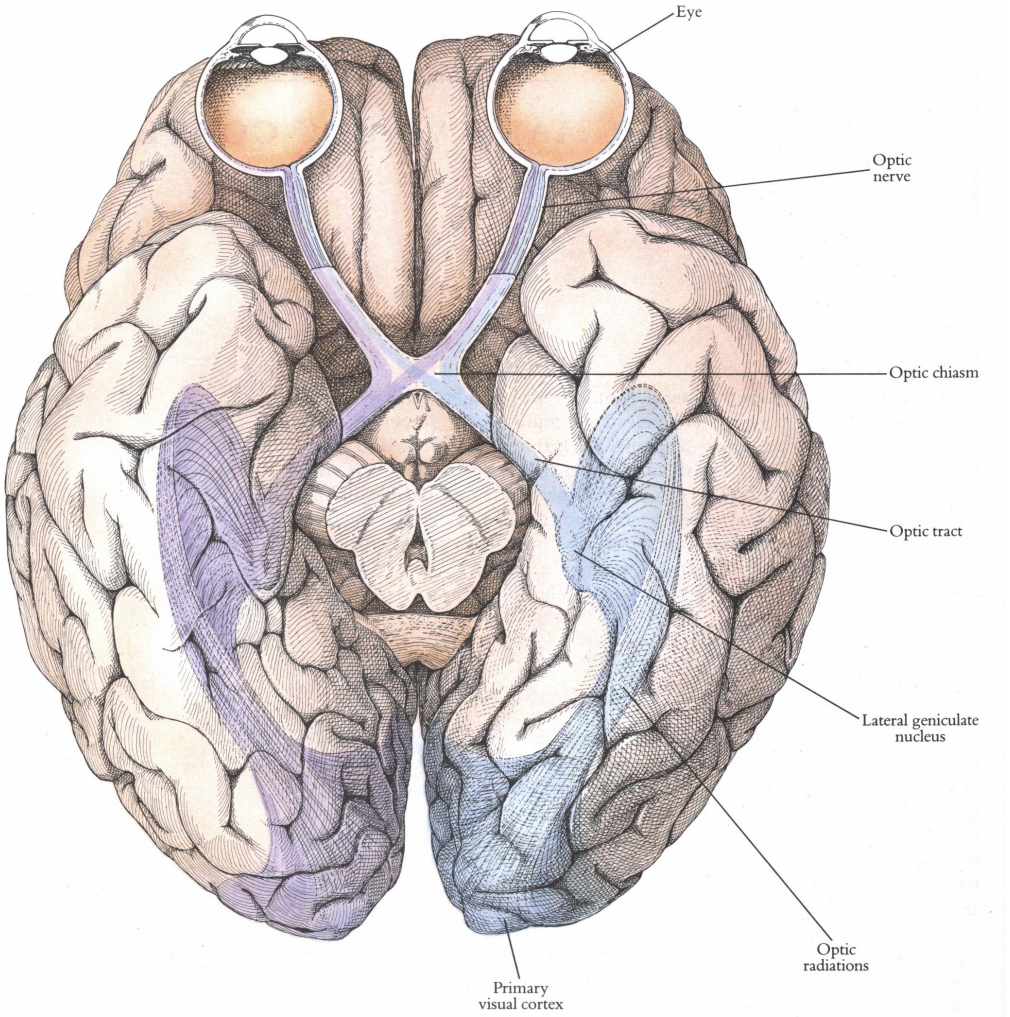
Maria Petrou
Imperial College London

Overview of the Human Visual System

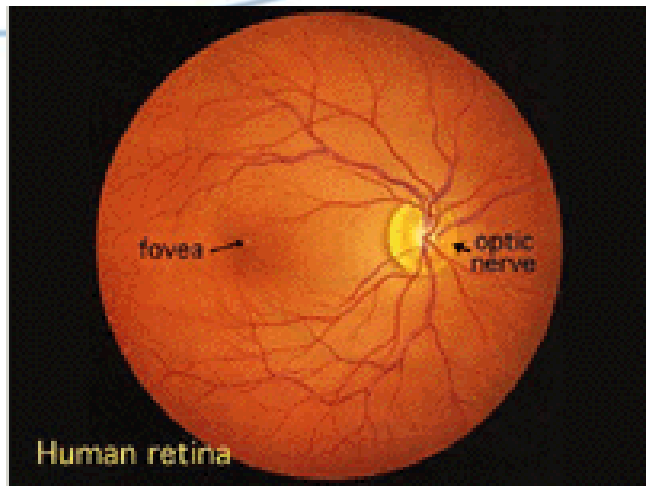


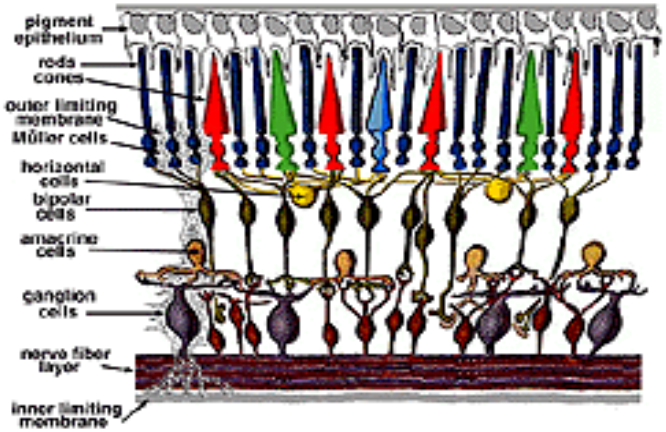
From eye to primary visual cortex

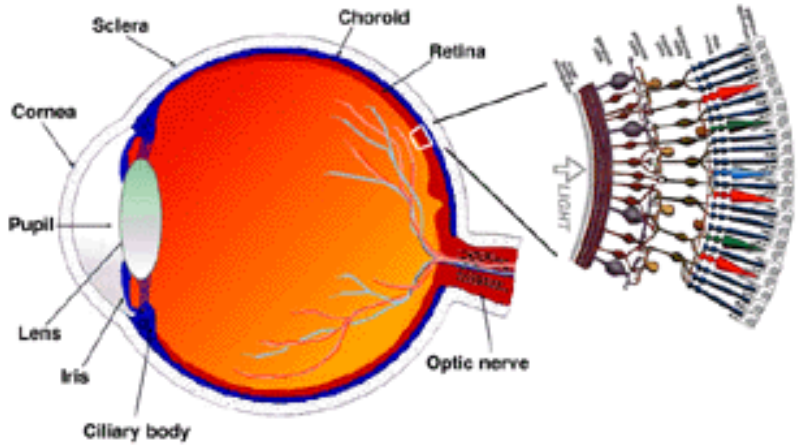
The visual path



The retina

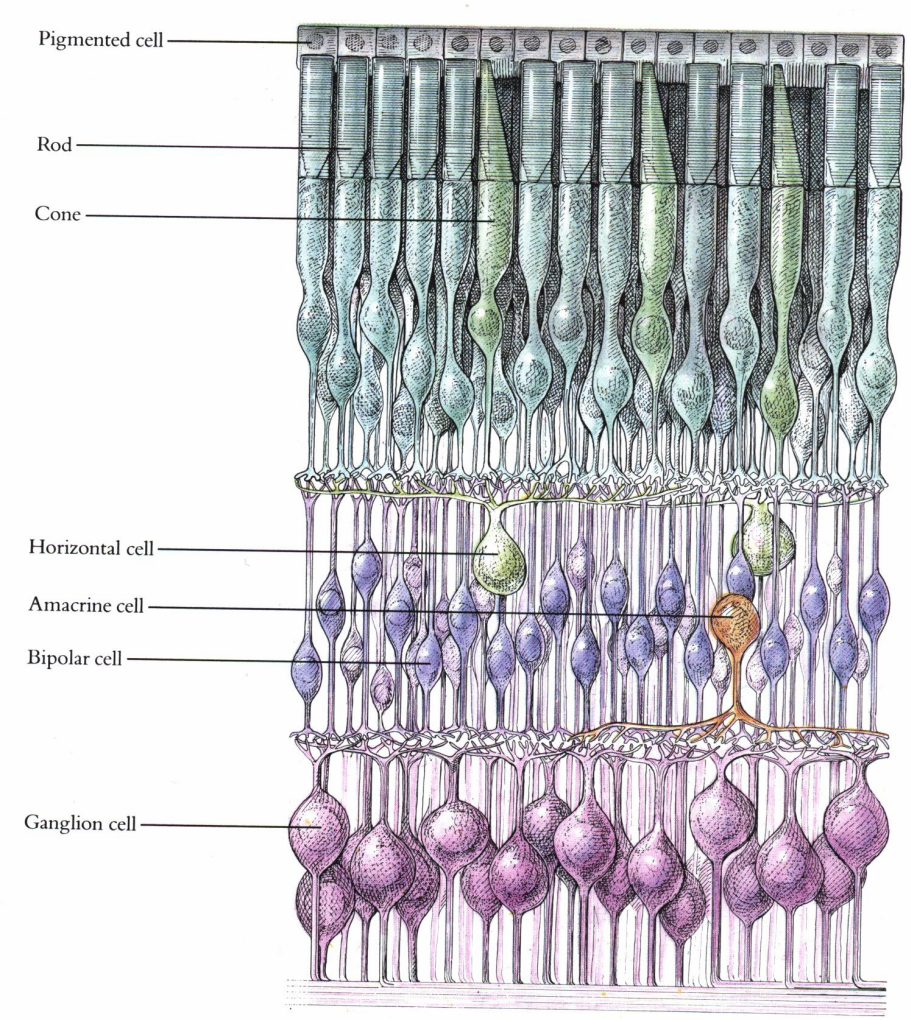






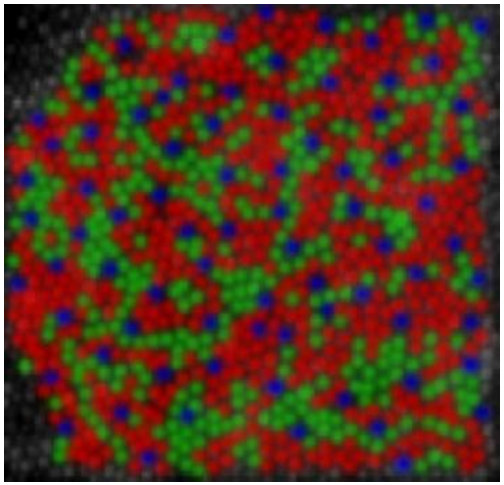
The Human retina is composed of five cellular layers.

The retina

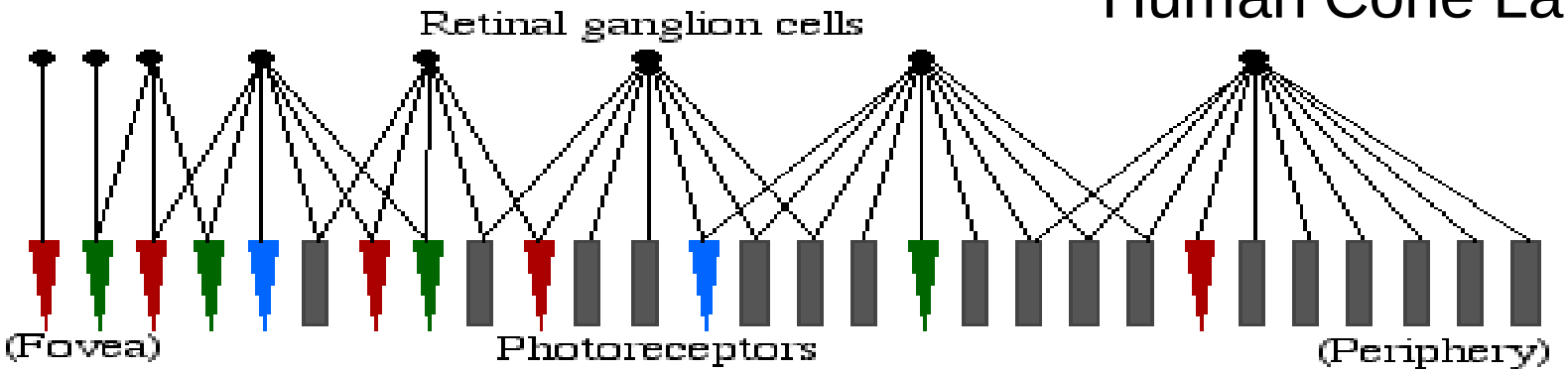

Pigmented cell
Rod
Cone
Horizontal cell
Amacrine cell
Bipolar cell
Ganglion cell

The retina

The first retinal layer is represented by photoreceptors (rods and cones). Their distribution is highly irregular.



Human Cone Lattice



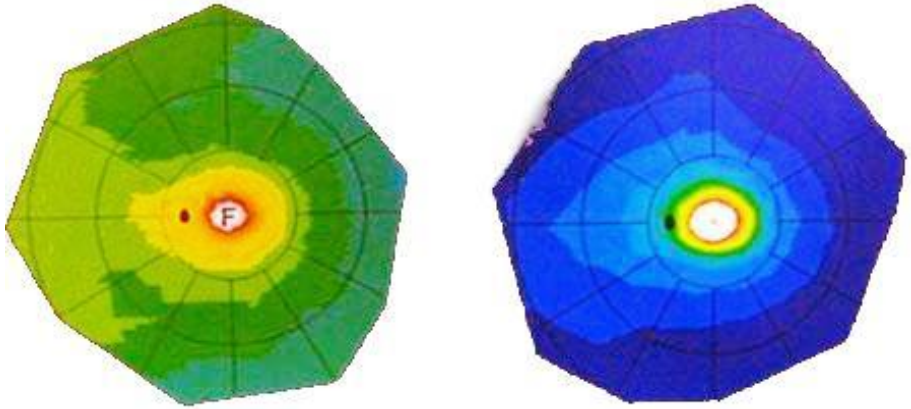
Retina sampling (From B. Olshausen)

Working with Irregular Sampling Grids

Sampling of photoreceptors in human retina has irregular topology

Human retina topographic map

Map shows the **cell density** in the human retina (F=fovea)

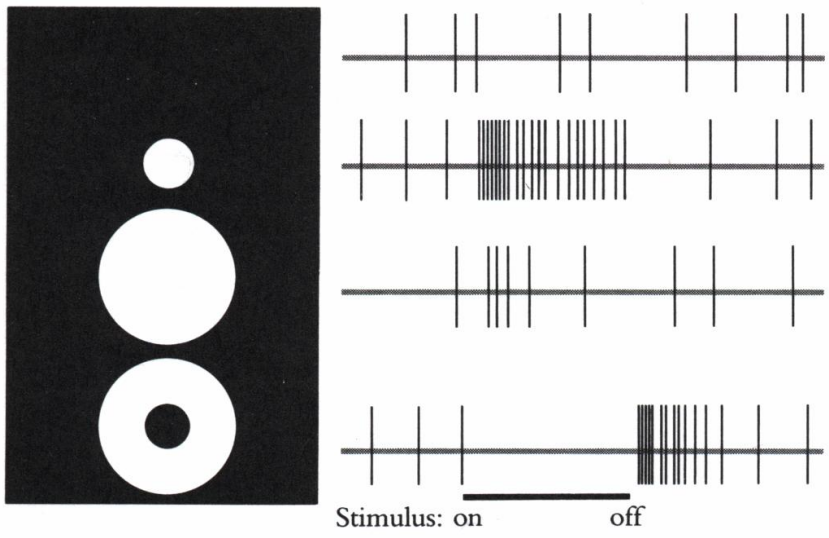


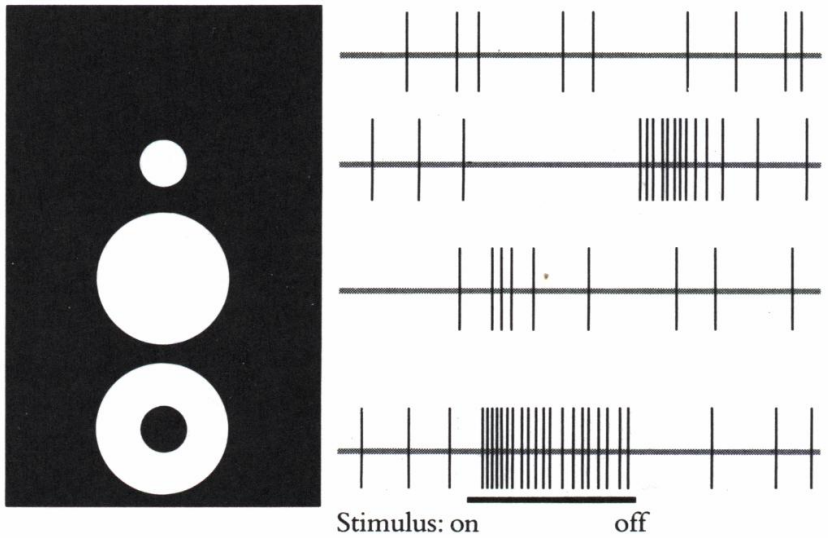
Dark red=16,000 cells/mm^2

Dark blue=1000 cells/mm^2

Cone photoreceptors Ganglion cells

The ganglion cells…





Issues to be investigated

The response curves of the photoreceptors to the various wavelengths are very different in the human eye and in the CCD sensors we currently use.

Issues to be investigated

The response curves of the photoreceptors to the various wavelengths are very different in the human eye and in the CCD sensors we currently use.

Can we construct sensors with response curves similar to those of the human eye?

Issues to be investigated

The response curves of the photoreceptors to the various wavelengths are very different in the human eye and in the CCD sensors we currently use.

Can we construct sensors with response curves similar to those of the human eye?

Why is this important from the technological point of view?

Issues to be investigated

The response curves of the photoreceptors to the various wavelengths are very different in the human eye and in the CCD sensors we currently use.

Can we construct sensors with response curves similar to those of the human eye?

Why is this important from the technological point of view?

In industrial inspection problems often the automatic inspection system has to see things the same way the human sees

Issues to be investigated

The topology of the photoreceptors is not that of a rectangular regular grid

Issues to be investigated

The topology of the photoreceptors is not that of a rectangular regular grid

Can we do image processing when the scene is sampled at irregularly spaced points?

Issues to be investigated

The topology of the photoreceptors is not that of a rectangular regular grid

Can we do image processing when the scene is sampled at irregularly spaced points?

Why is this important from the technological point of view?

Issues to be investigated

The topology of the photoreceptors is not that of a rectangular regular grid

Can we do image processing when the scene is sampled at irregularly spaced points?

Why is this important from the technological point of view?

In geoscience data are almost never sampled at regular points.
Occlusion and faults may damage the regularity of a grid.
Image compression may be achieved by reducing the redundancy of a regular grid.

Issues to be investigated

The structure of the retina is very complicated and very different from that of a CCD sensor array

Issues to be investigated

The structure of the retina is very complicated and very different from that of a CCD sensor array

Can we construct simulated and physical models of the retina?

Issues to be investigated

The structure of the retina is very complicated and very different from that of a CCD sensor array

Can we construct simulated and physical models of the retina?

Why is this important from the technological point of view?

Issues to be investigated

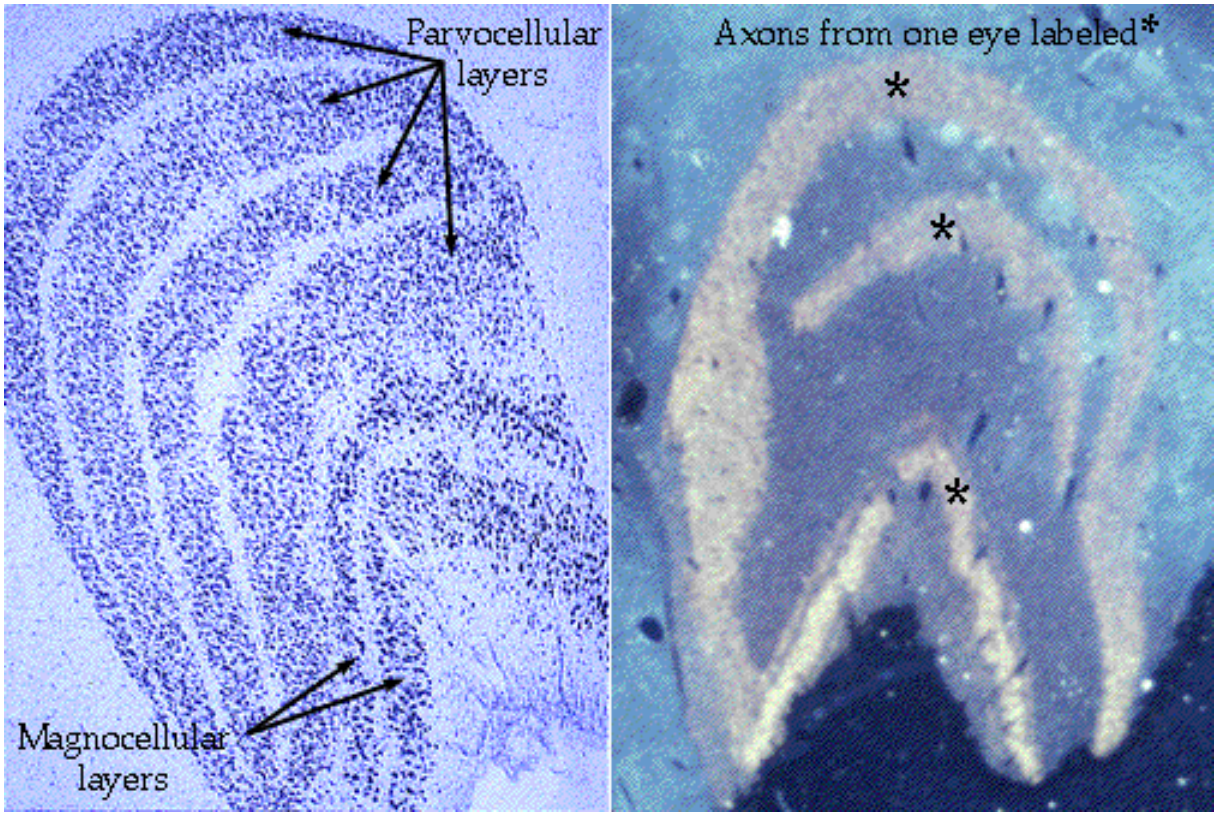
The structure of the retina is very complicated and very different from that of a CCD sensor array

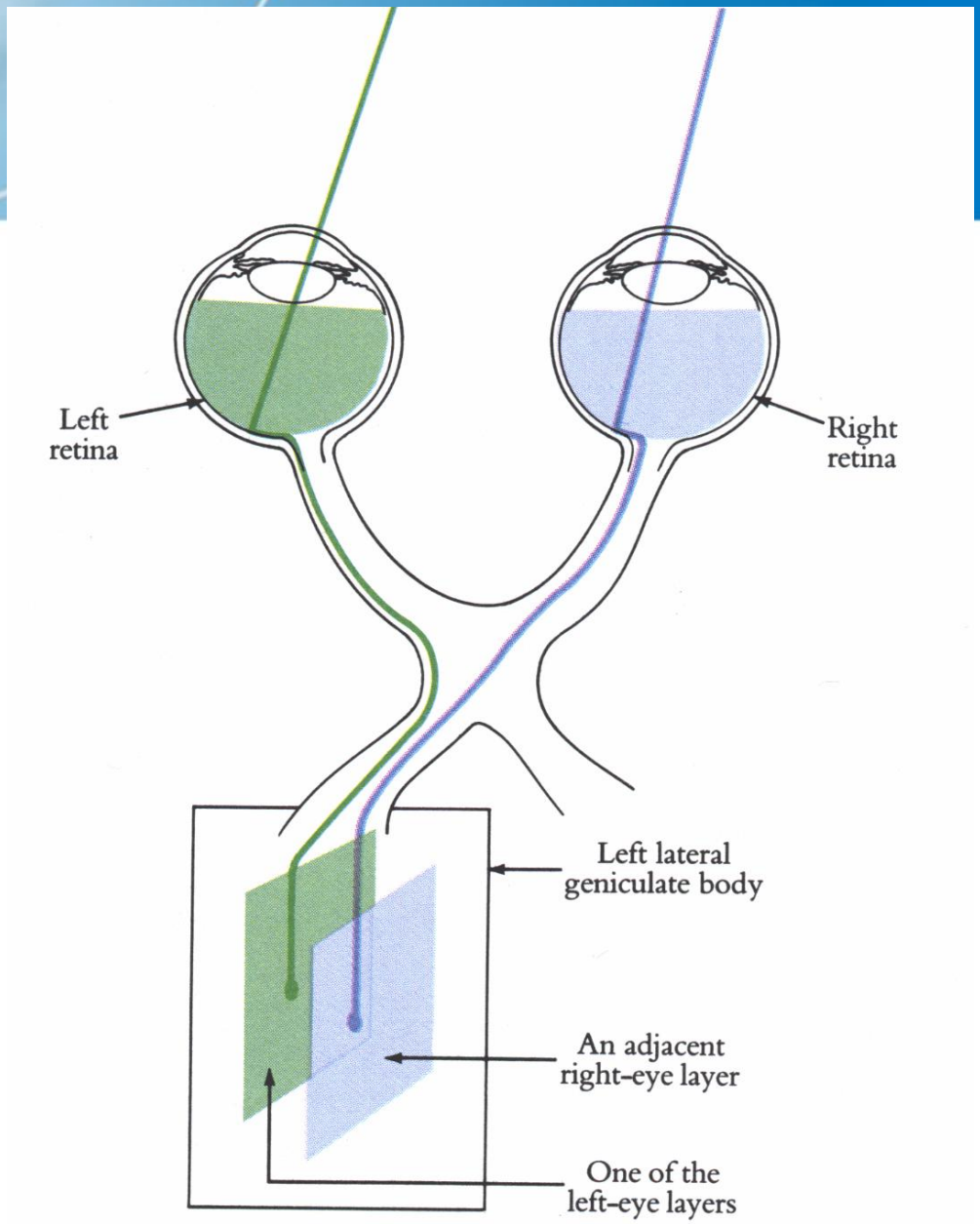
Can we construct simulated and physical models of the retina?

Why is this important from the technological point of view?

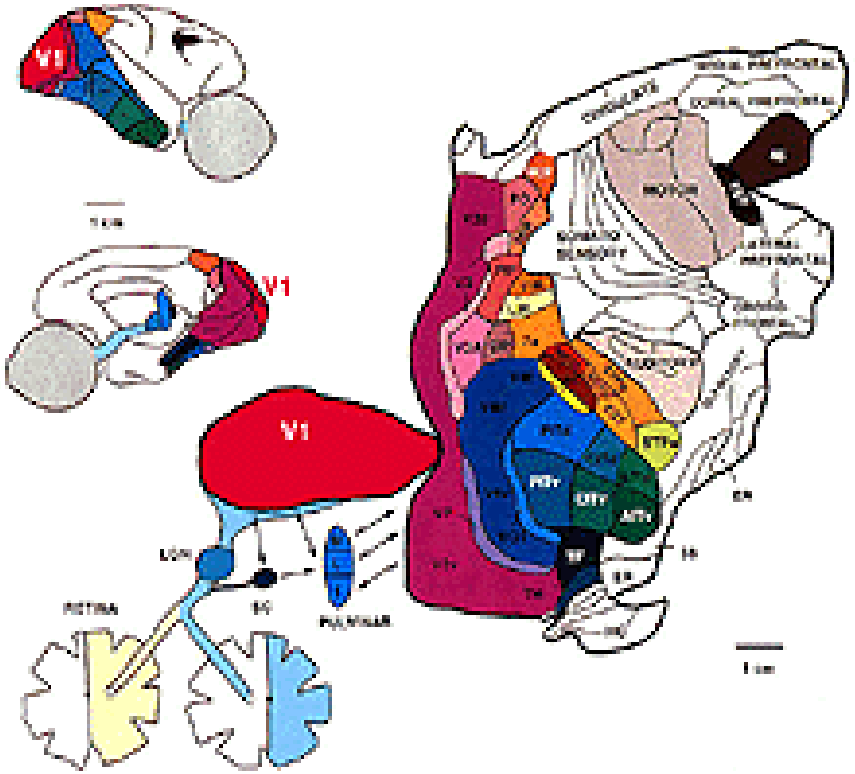
Models will help us understand, experiment with and even diagnose function and defects in the retina.
They may lead to the construction of implantable retinas!

The lateral geniculate nucleus

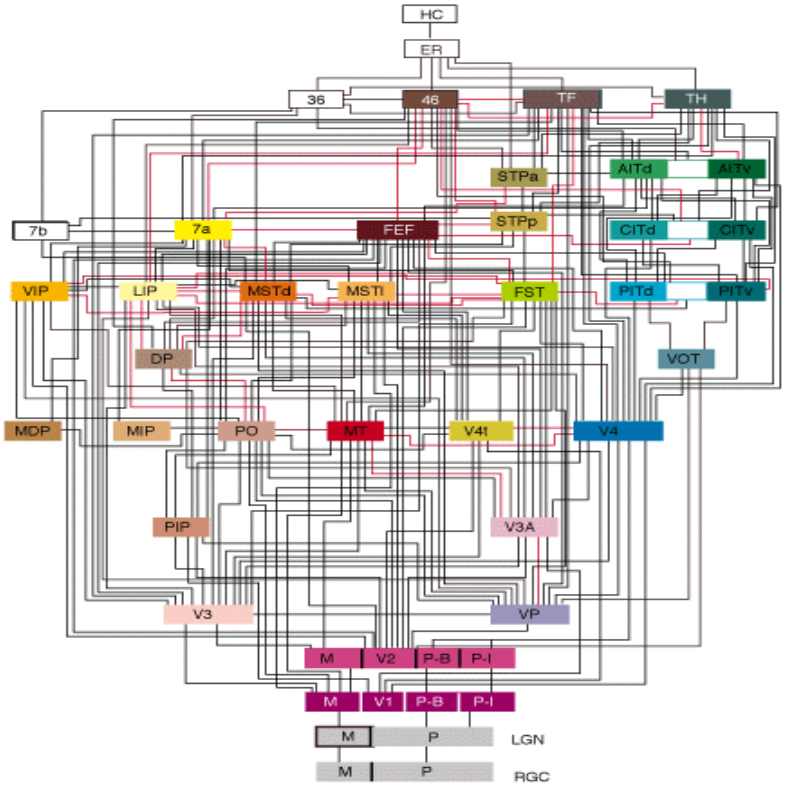



Left retina
Right retina
Left lateral geniculate body
An adjacent right-eye layer
One of the left-eye layers

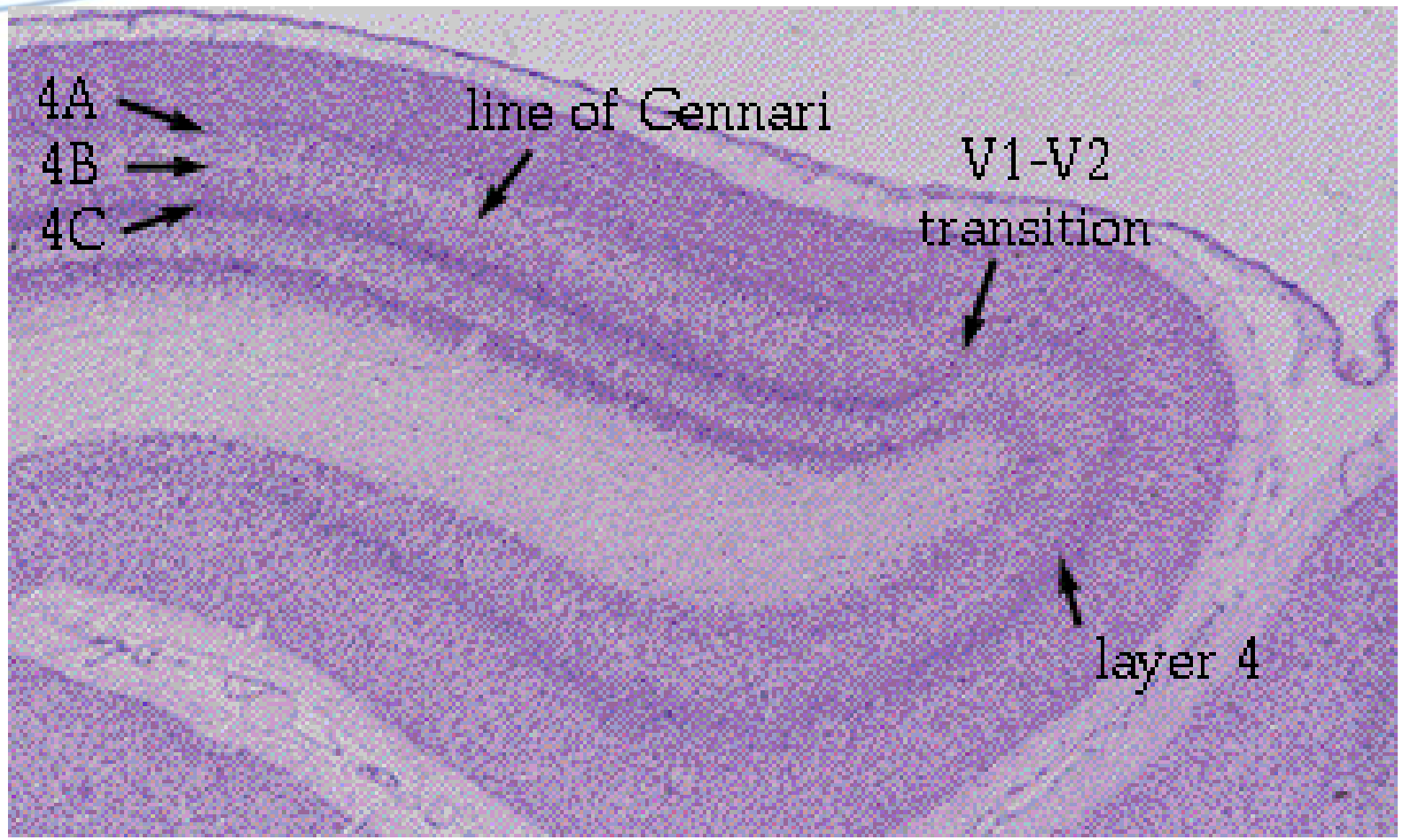
Overview of the Human Visual System

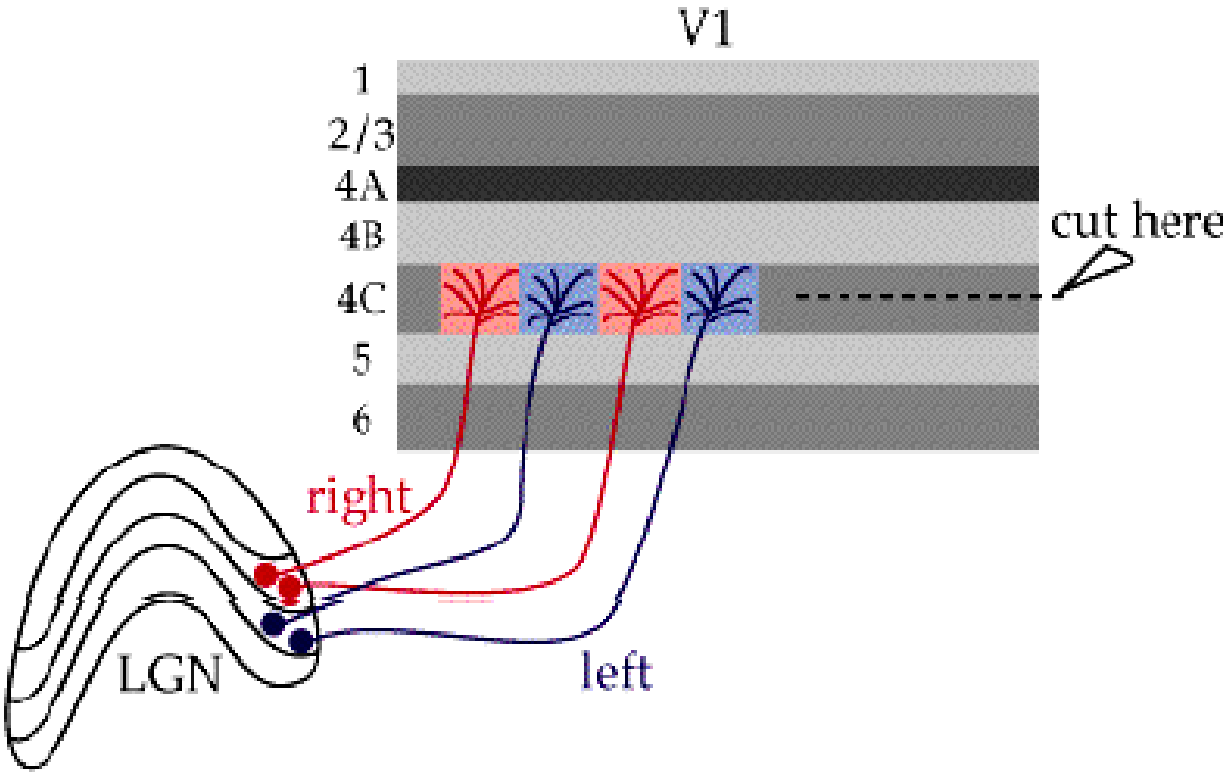


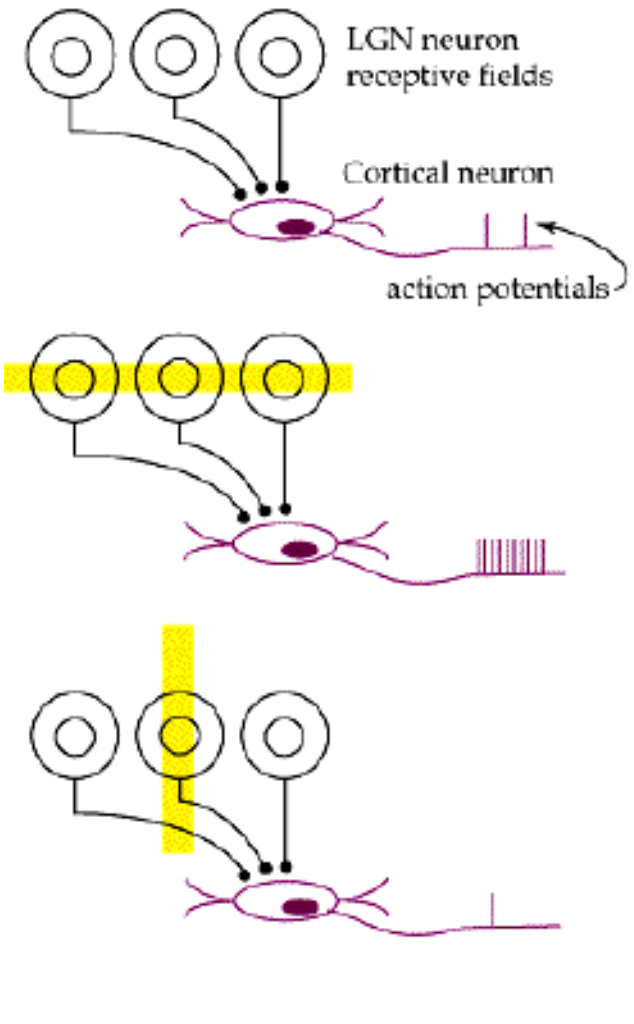
Visual areas in the brain



Visual Area wiring diagram


4A
4B
4C
line of Gennari
V1-V2
transition
layer 4


V1
1
2/3
4A
4B
4C
5
6
cut here
right
left
LGN


LGN neuron
receptive fields
Cortical neuron
action potentials

Issues to be investigated

What is the role of V1?

Issues to be investigated

What is the role of V1?

Can we learn anything from it for doing better edge detection, image segmentation, texture boundary detection artificial vision?

Issues to be investigated

What is the role of V1?

Can we learn anything from it for doing better edge detection, image segmentation, texture boundary detection artificial vision?

Why is this important from the technological point of view?

Issues to be investigated

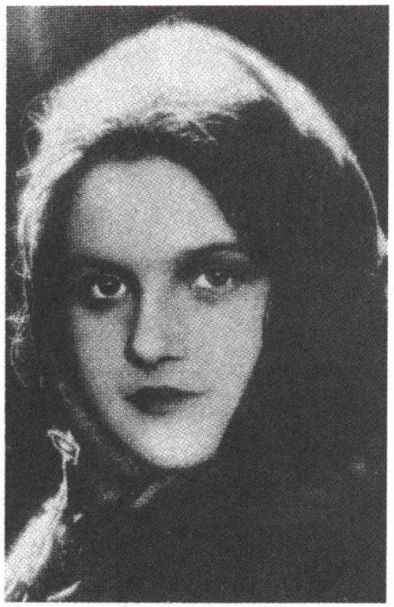
What is the role of V1?

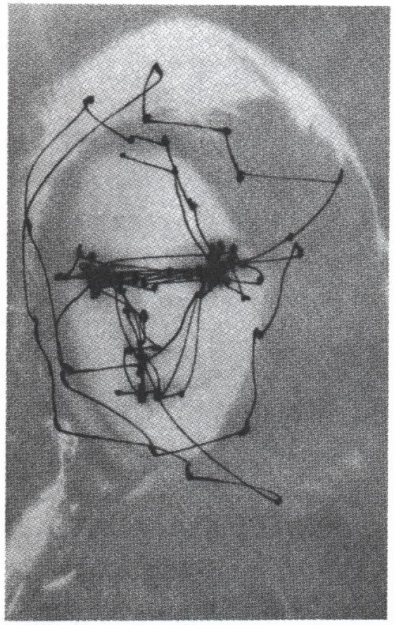
Can we learn anything from it for doing better edge detection, image segmentation, texture boundary detection artificial vision?

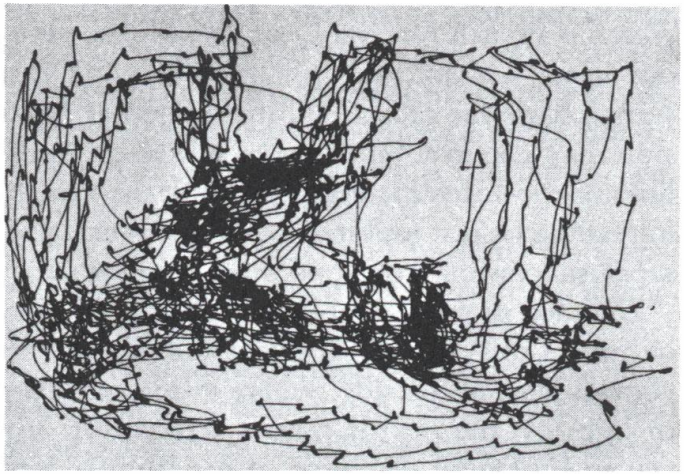
Why is this important from the technological point of view?

Artificial vision systems are still very far from acceptable performance in a generic environment. Understanding how our benchmark system works helps us improve them.

Vision and Perception







A picture is viewed by an observer while we monitor eye position and hence direction of gaze. The eyes jump, come to rest momentarily (producing a small dot on the record), then jump to a new locus of interest. It seems difficult to jump to a void—a place lacking abrupt luminance changes.

Issues to be investigated

How do we actually see?

Issues to be investigated

How do we actually see?

What is the relationship between perception and vision?

Issues to be investigated

How do we actually see?

What is the relationship between perception and vision?

Why is this important from the technological point of view?

Issues to be investigated

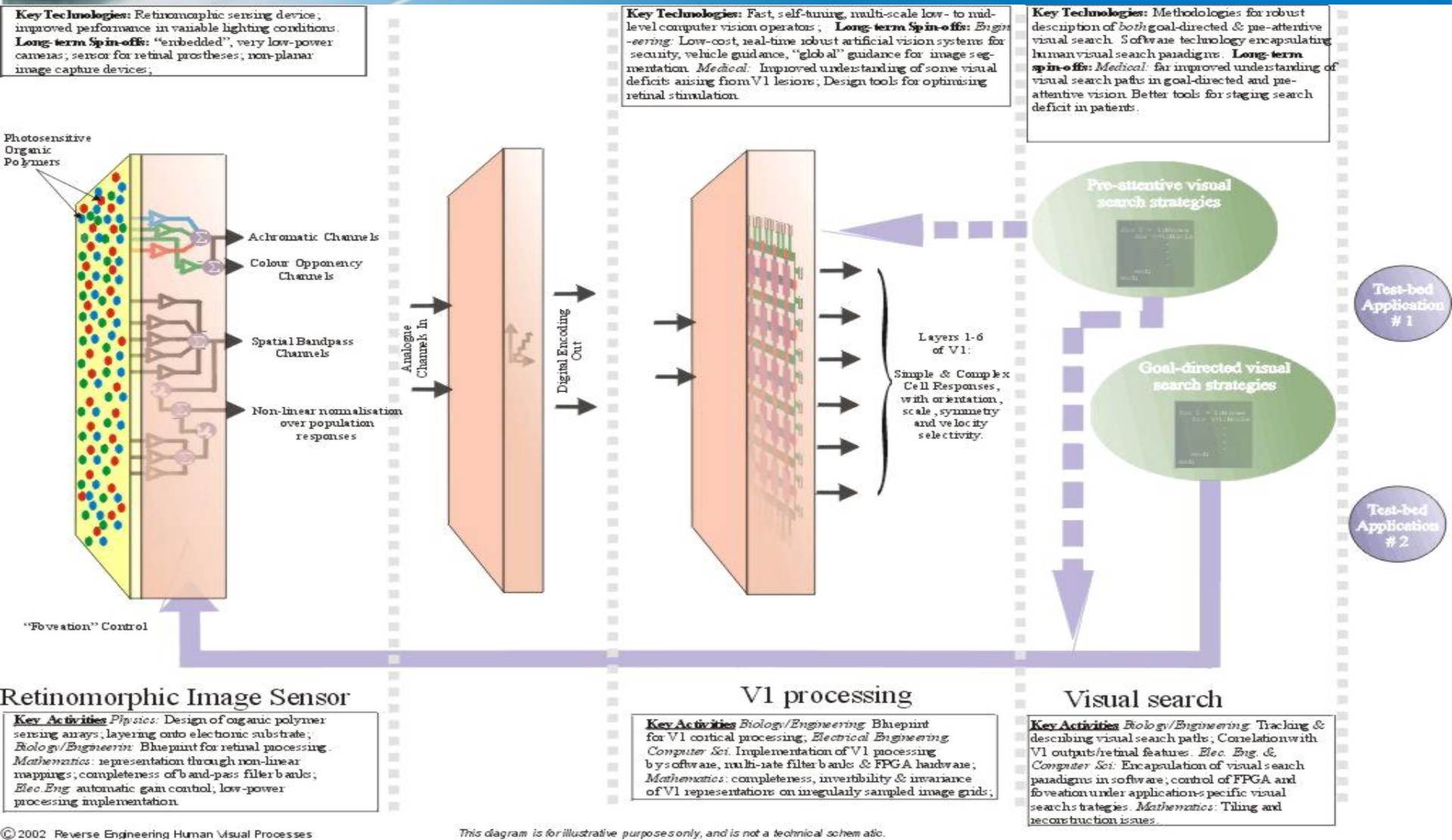
How do we actually see?

What is the relationship between perception and vision?

Why is this important from the technological point of view?

Improve the way we display information
Automatic driving/Robot navigation
Human/Computer communication

In summary…



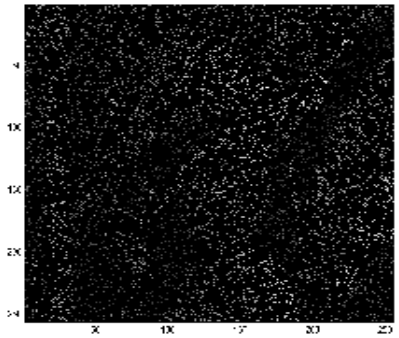
This diagram is for illustrative purposes only, and is not a technical schematic.

What we shall discuss in detail…

- Irregular sampling and how to deal with it
- Saliency model and how to generalise it for other problems
- Networks and how information may be organised in the brain

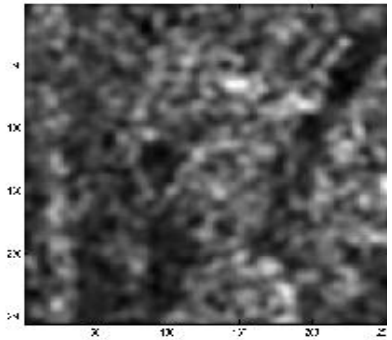
Image Reconstruction by using Normalized Convolution

original



Irregularly sampled (10%)

Conventional convolution



Normalized convolution

R Piroddi and M Petrou, 2004. “Analysis of irregularly sampled data: a Review”. Advances in Imaging and Electron Physics, Vol 132, pp 109-165.

Image Reconstruction by using Iterative Methods

original

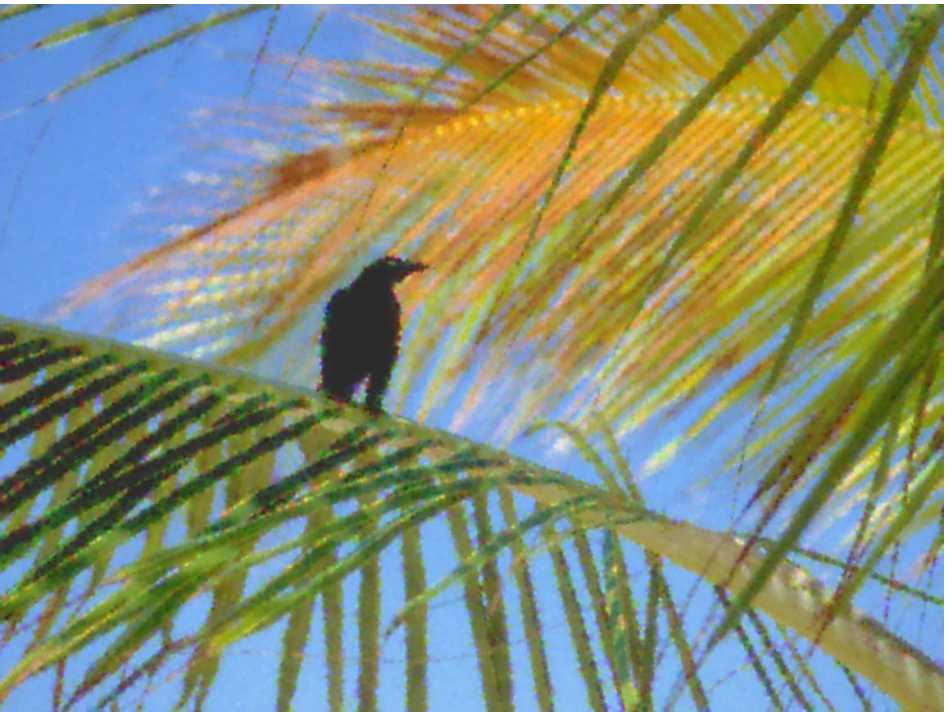
Irregularly sampled (5%)

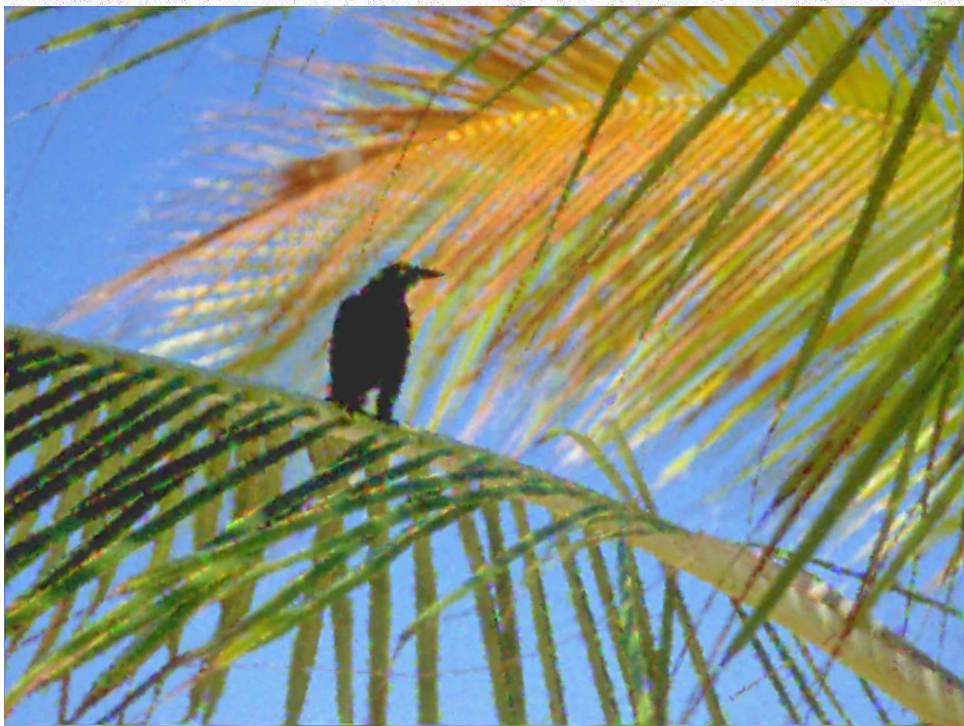
Reconstructed Normalized Convolution

Reconstructed Voronoi Iterative Method

from: Duijndam, A.J.W., M.A.~Schonewille and C.O.H. Hindriks, "Reconstruction of band-limited signals, irregularly sampled along one spatial direction," Geophysics, vol.64, no.2, 1999, pp. 524-538.

software: R.Piroddi and M. Petrou, CVSSP, UNIVERSITY OF SURREY






Texture reconstruction

S Chandra, M Petrou and R Piroddi, 2005. "Texture interpolation using ordinary Krigging". Pattern Recognition and Image Analysis, Second Iberian Conference, IbPRIA2005, Estoril, Portugal, June 7-9, J S Marques, N Perez de la Blanca and P Pina (eds), Springer LNCS 3523, Vol II, pp 183-190.

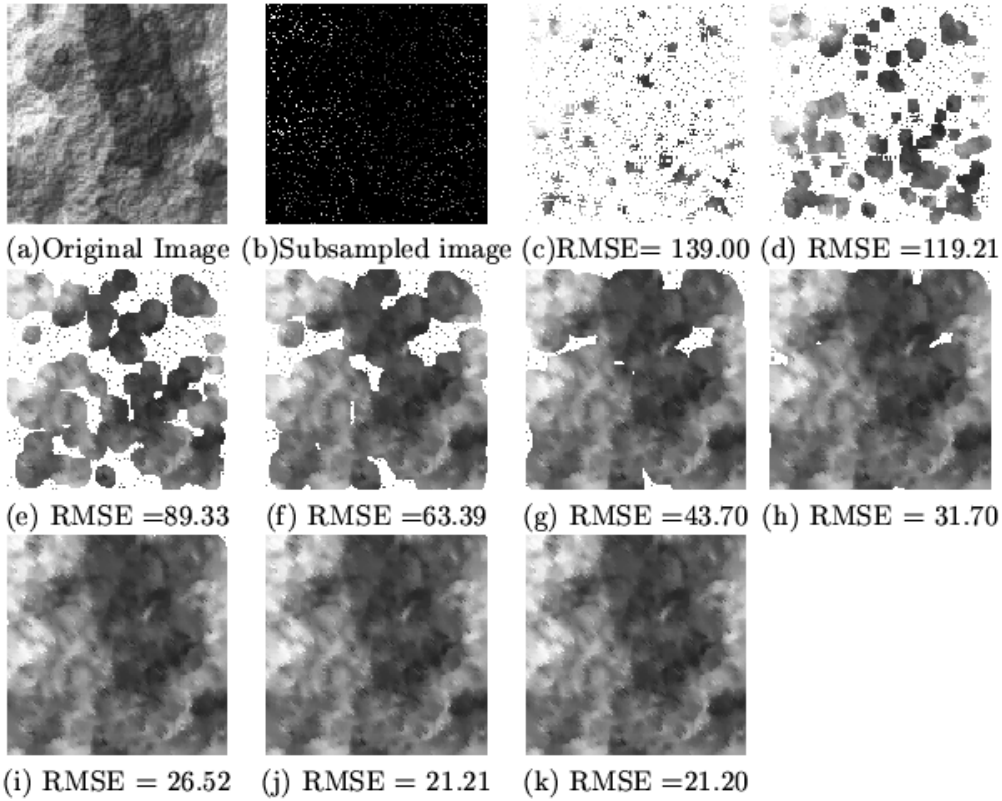


Fig. 4. Process of reconstruction for image III from 6% of its pixels, using the linear model

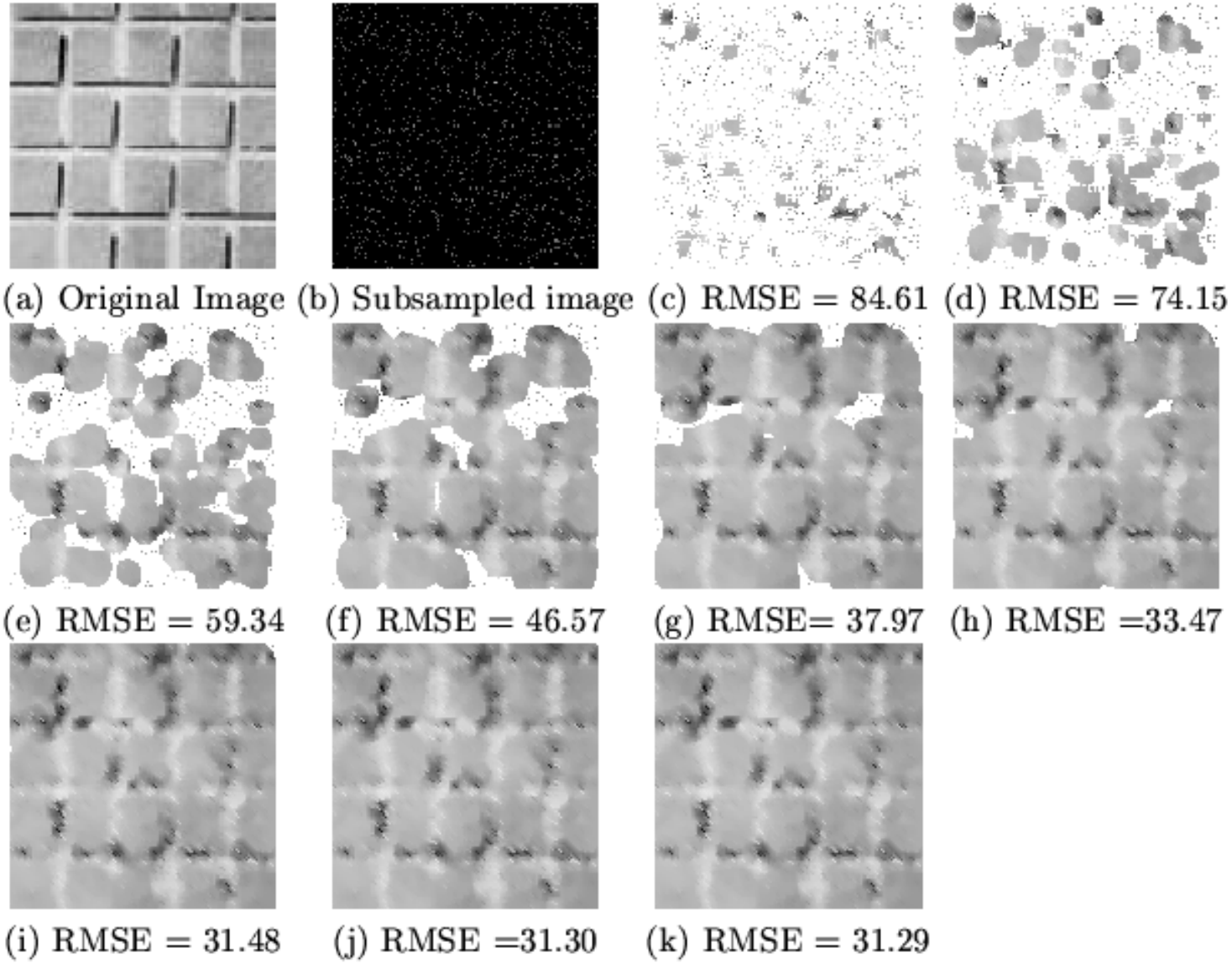


Fig. 3. Process of reconstruction for image II from 6% of its pixels, using the fractal model

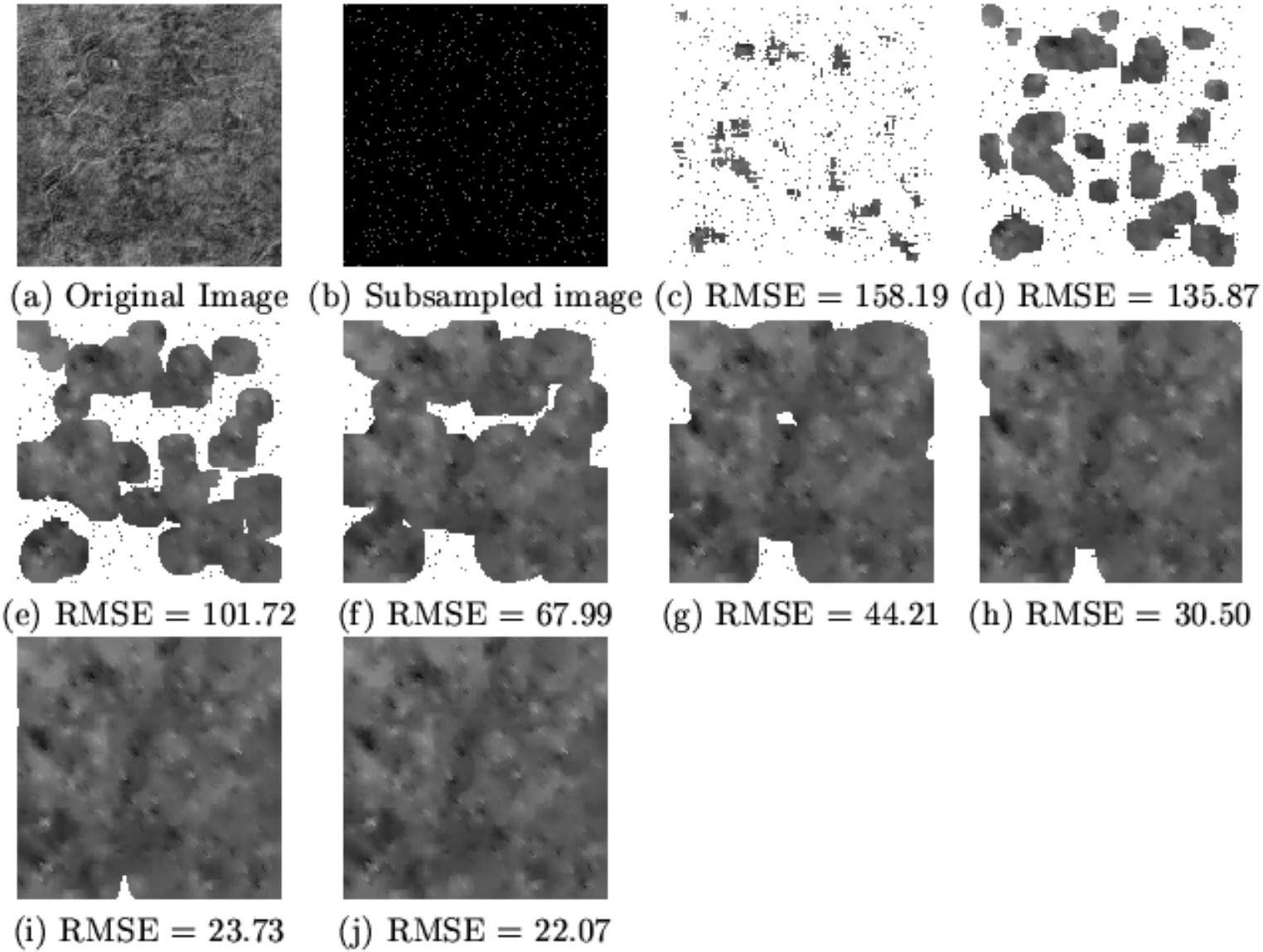
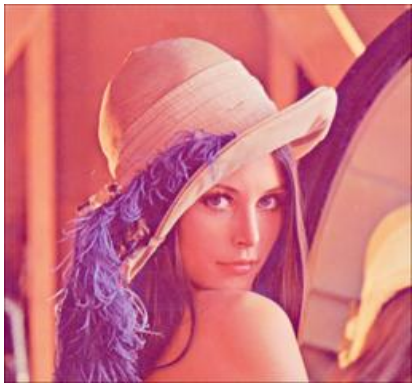
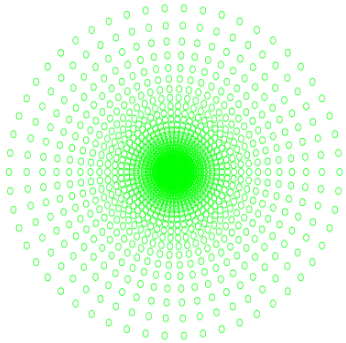


Fig. 5. Process of reconstruction for image IV from 4% of its pixels, using the exponential model

Image Reconstruction: Application to retinal sampling

original





Retinotopic sampling topology used to simulate retina (1% of original data).

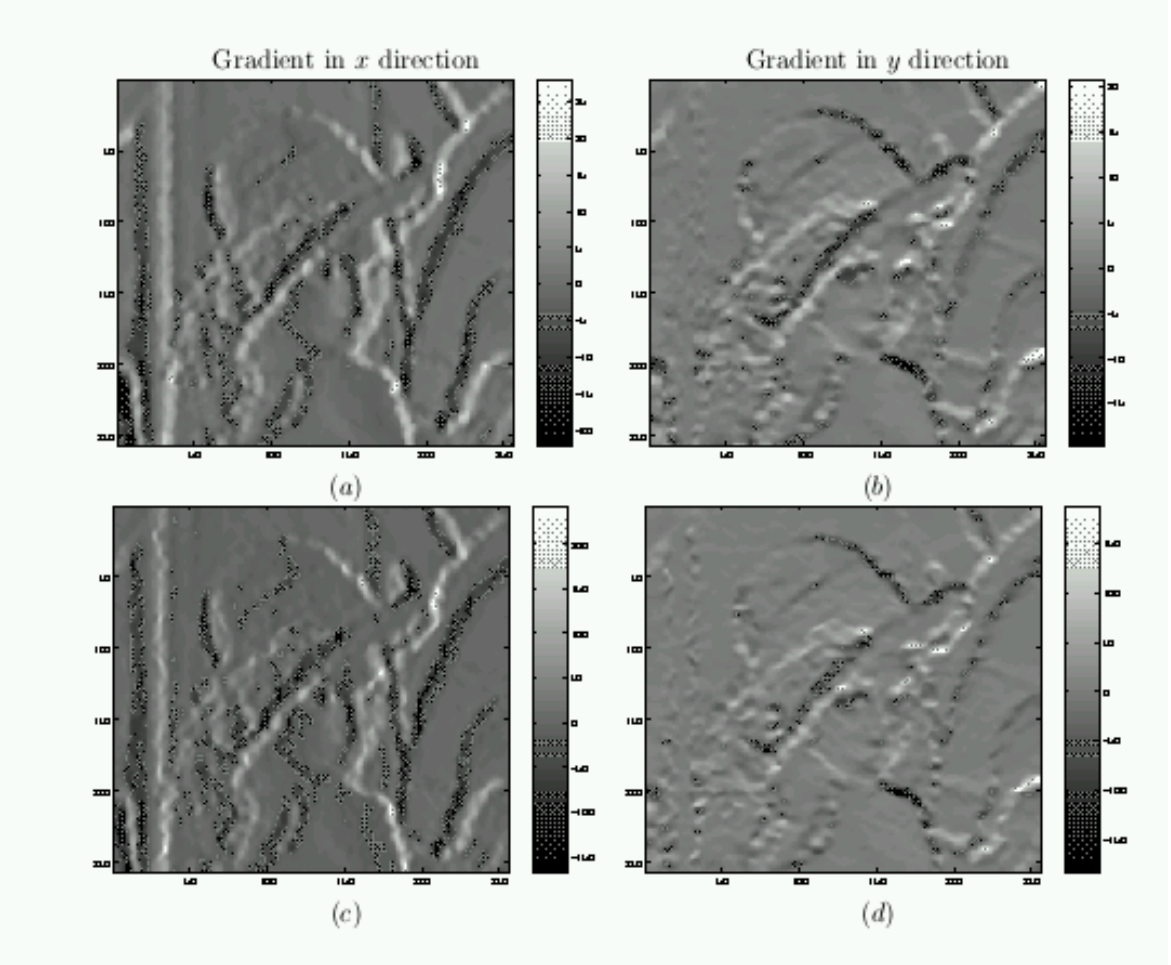
Reconstructed, foveated image

R Piroddi and M Petrou, 2004. "Analysis of irregularly sampled data: a Review". Advances in Imaging and Electron Physics, Vol 132, pp 109-165.

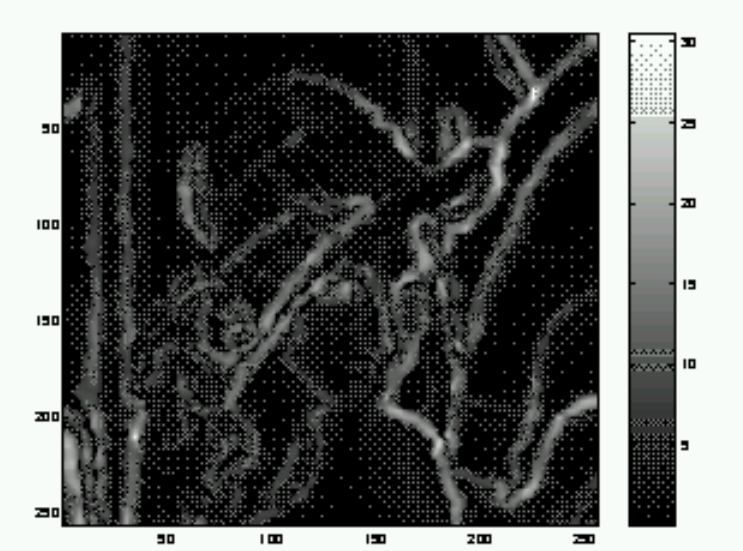
Gradient estimation on irregular samples

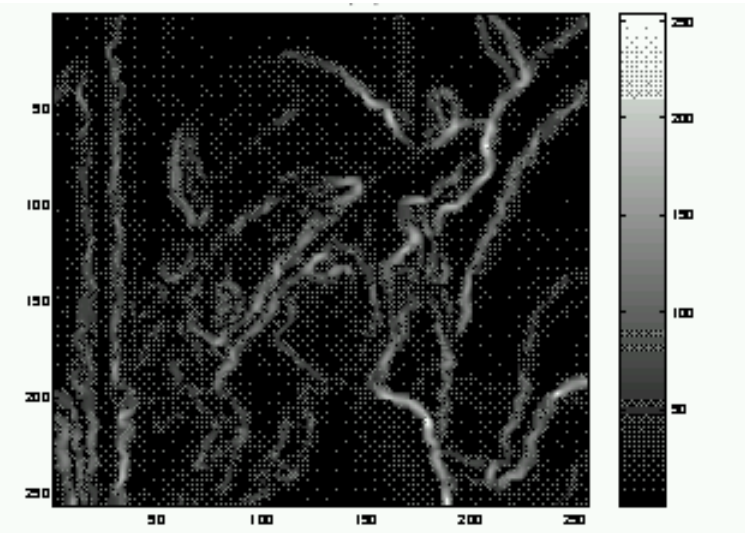
Normalized Differential Convolution (NDC)

Derivative of Normalized Differential Convolution (DoNC)



Gradient estimation on irregular samples: Magnitude

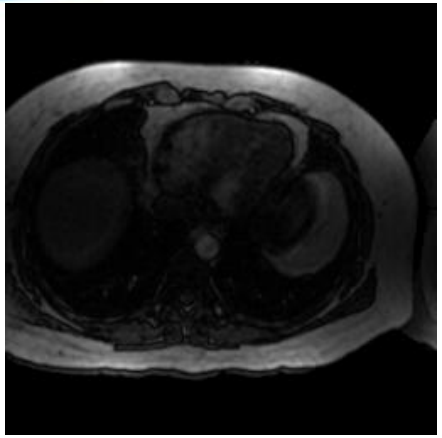




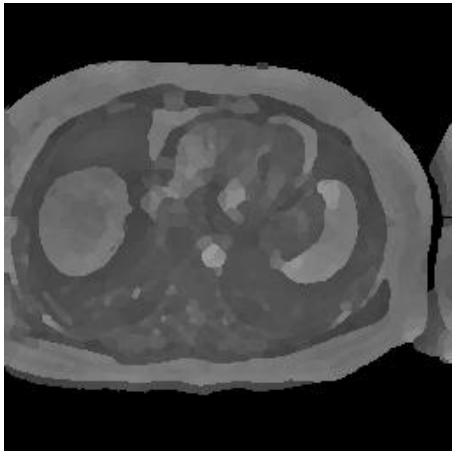
Normalized Differential Convolution (NDC)

Derivative of Normalized Differential Convolution (DoNC)

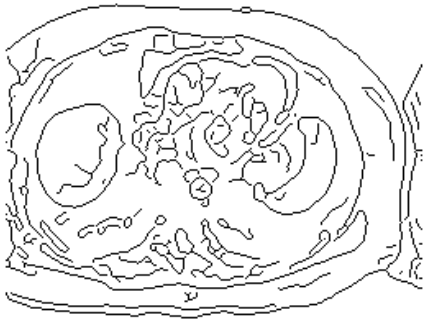
Modelling V1:
Application of V1 model



After pre-processing:



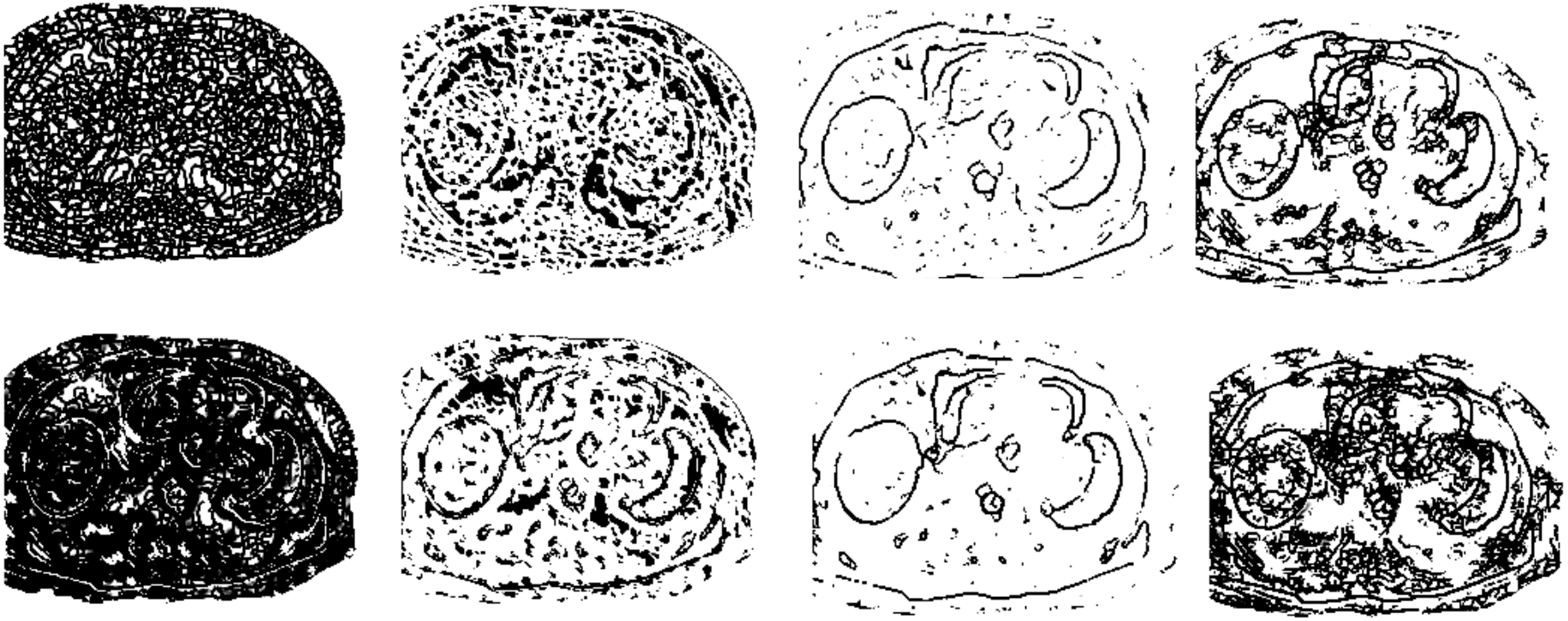
Canny



Zhaoping Li

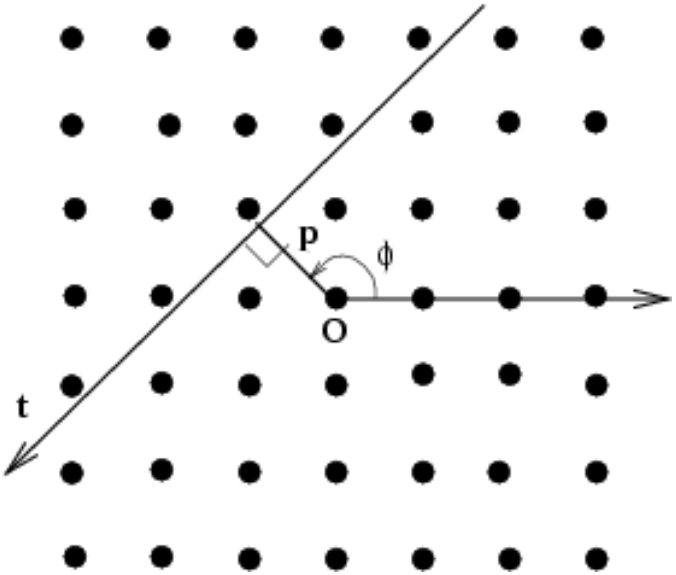
H Ibrahim, PhD thesis, Surrey University.

Modeling V1:
Application of V1 model



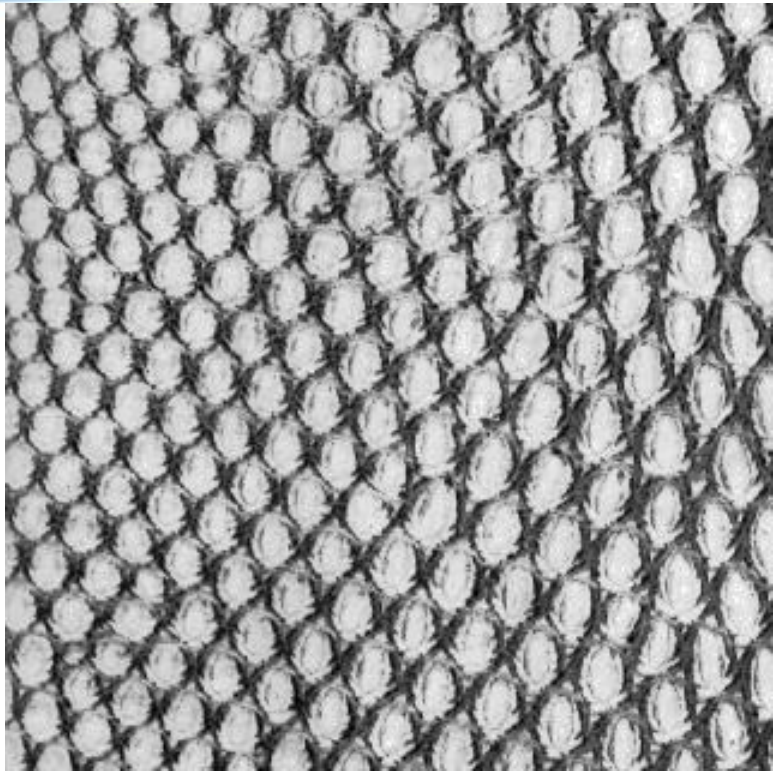
H Ibrahim, PhD thesis, Surrey University

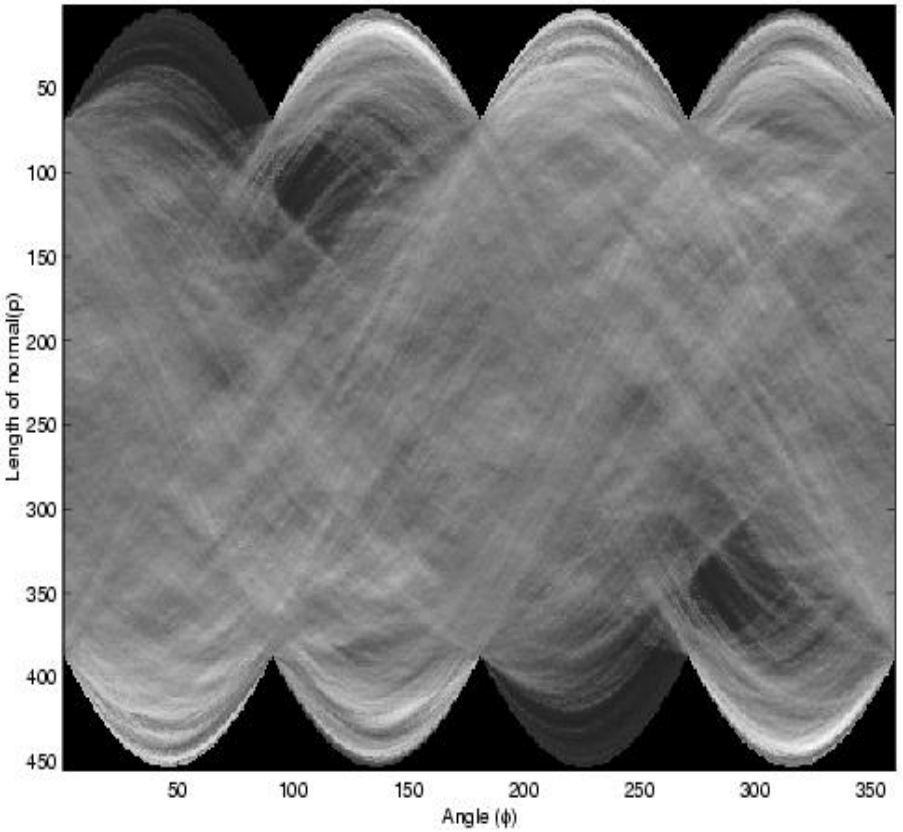
Pre-attentive texture ranking



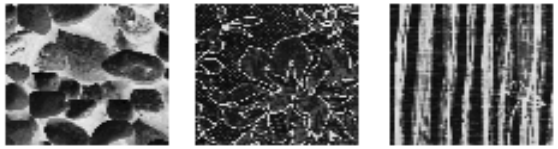
M Petrou, A Talebpour and A Kadyrov, 2007. "Reverse Engineering the way humans rank textures". Pattern Analysis and Applications, Vol 10 (2) pp 101-114.

Pre-attentive texture ranking

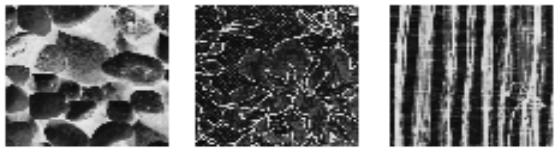




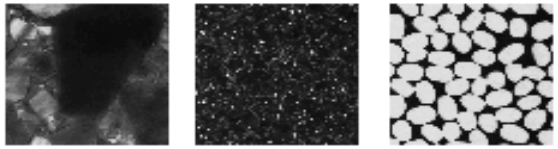
Pre-attentive texture ranking



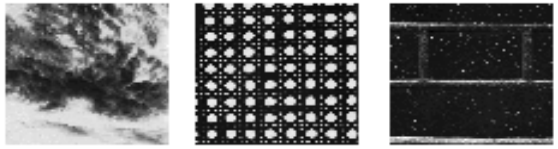
(a) T31 (b) T41 (c) T51



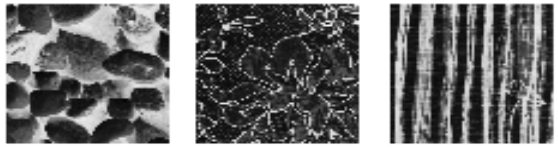
(d) T31 (e) T41 (f) T51



(g) T59 (h) T33 (i) T75



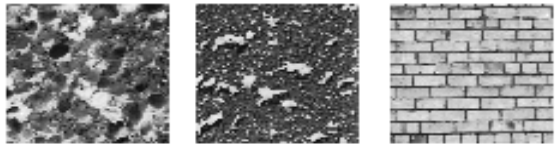
(j) T90 (k) T102 (l) T25



(a) T31 (b) T41 (c) T51

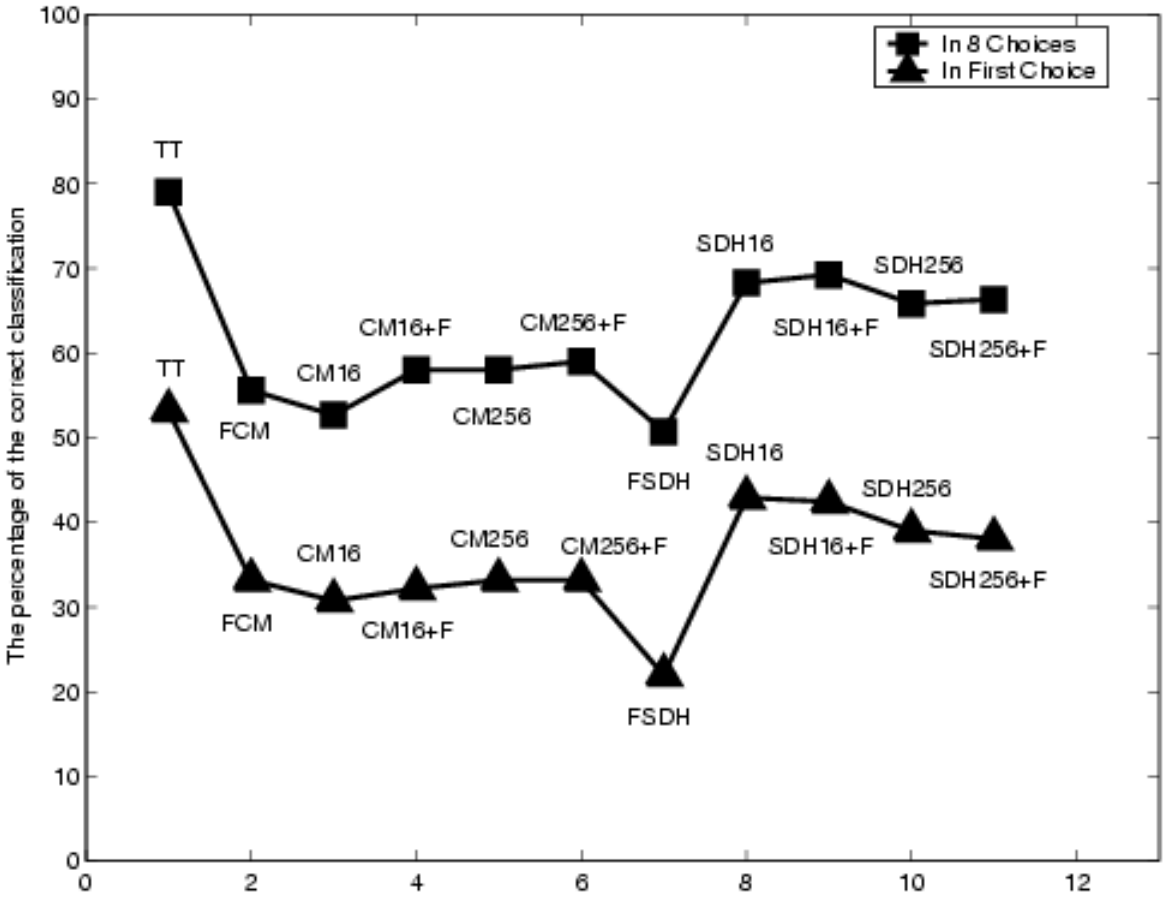
(d) T31 (e) T41 (f) T51

(g) T88 (h) T109 (i) T72



(j) T27 (k) T86 (l) T95

Pre-attentive texture ranking



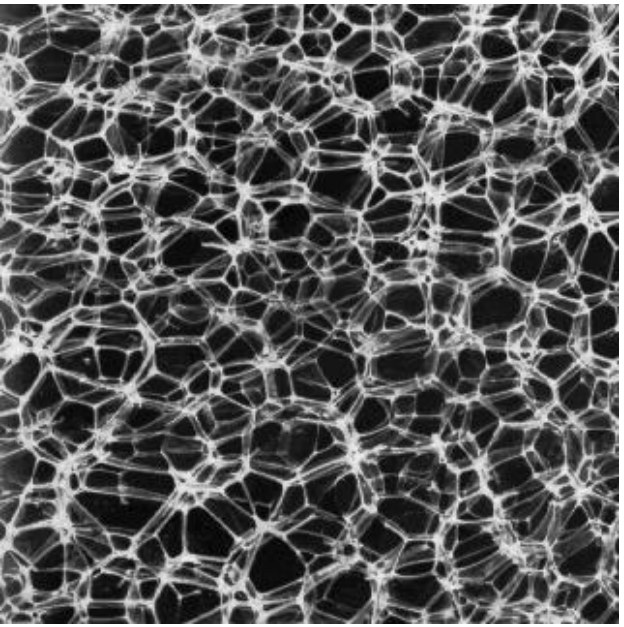
A. Talebpour and M Petrou, CVSSP, UNIVERSITY OF SURREY

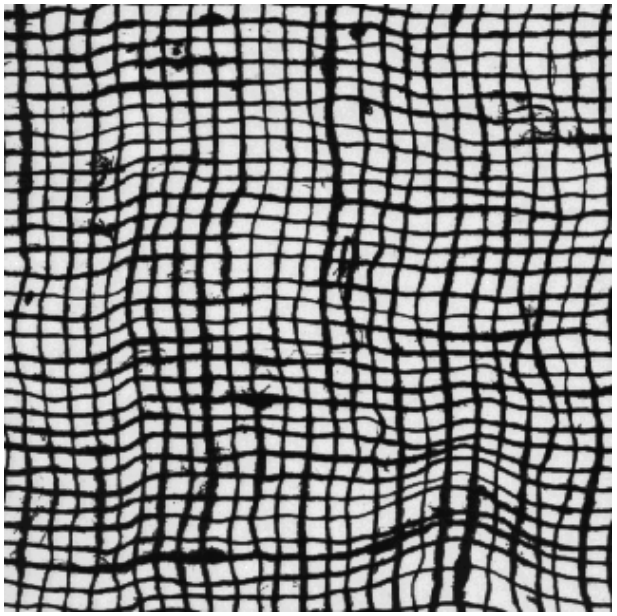
Pre-attentive texture ranking

	Functionals
T6	$\sum_{i=1}^{N-1} \lvert x_{i+1} - x_i \rvert$
T14	$\sum_{i=5}^{N-4} \sum_{k=1}^{4} \lvert x_{i-k} - x_{i+k} \rvert$
T24	$\sum_{i=1}^{N-2} \lvert x_i - 2x_{i+1} + x_{i+2} \rvert$
T25	$\sum_{i=1}^{N-3} \lvert x_i - 3x_{i+1} + 3x_{i+2} - x_{i+3} \rvert$
D1	$Max_{i=1}^{N} x_i$
D2	$Min_{i=1}^{N} x_i$
D5	$\sum_{i=1}^{N} i x_i$
D6	$\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2$
D10	$\sum_{i=1}^{N-4} \lvert x_i - 4x_{i+1} + 6x_{i+2} - 4x_{i+3} + x_{i+4} \rvert$
C1	$\sum_{i=1}^{N-1} \lvert x_{i+1} - x_i \rvert^p$
C13	Amplitude of the second harmonic
C17	Amplitude of the fourth harmonic

Table 1. The functionals of the best features used in the experiments. T, D and C in front of the numbers in the first column denote the Trace, Diametric and Circus functionals.

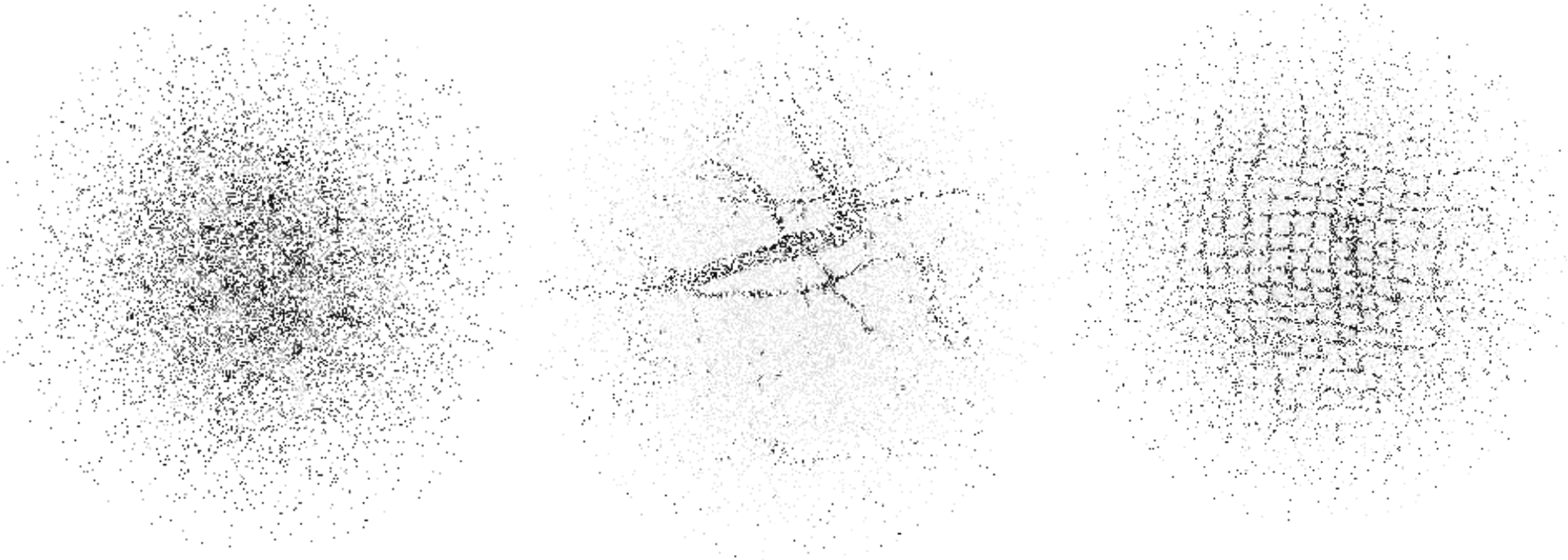
Texture recognition from Irregularly sampled data





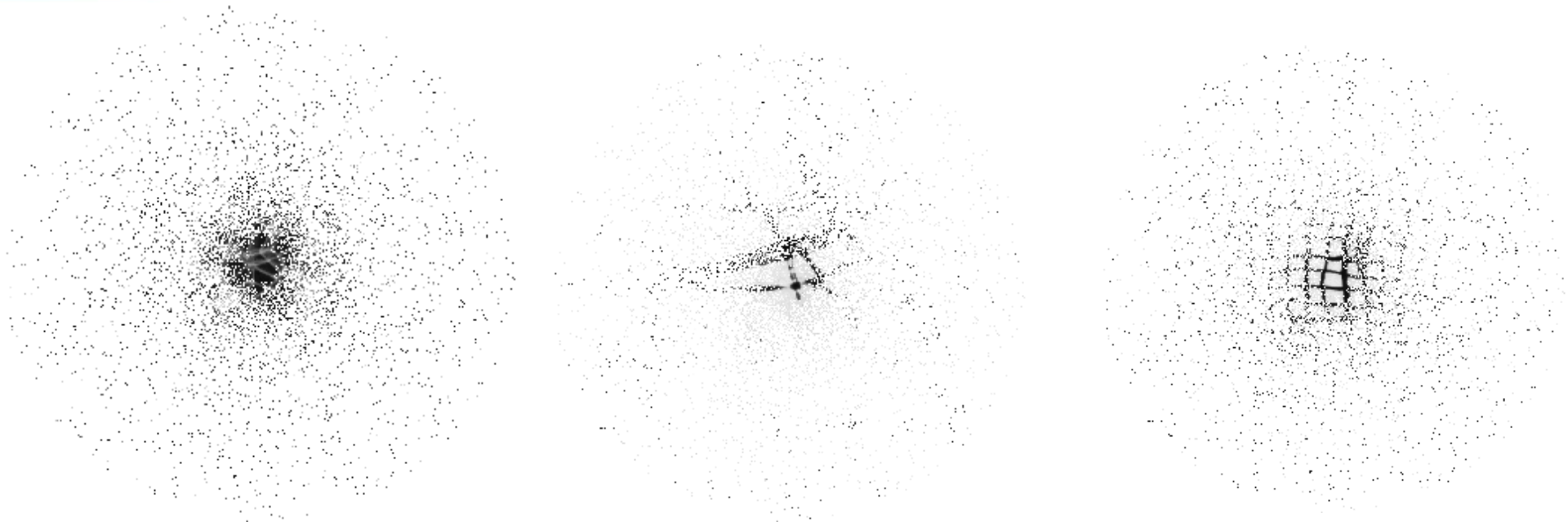
M Petrou, R Piroddi and A Talebpour, 2006. “Texture recognition from sparsely and irregularly sampled data”. Computer Vision and Image Understanding, Vol 102, pp 95-104.

Texture recognition from Irregularly sampled data



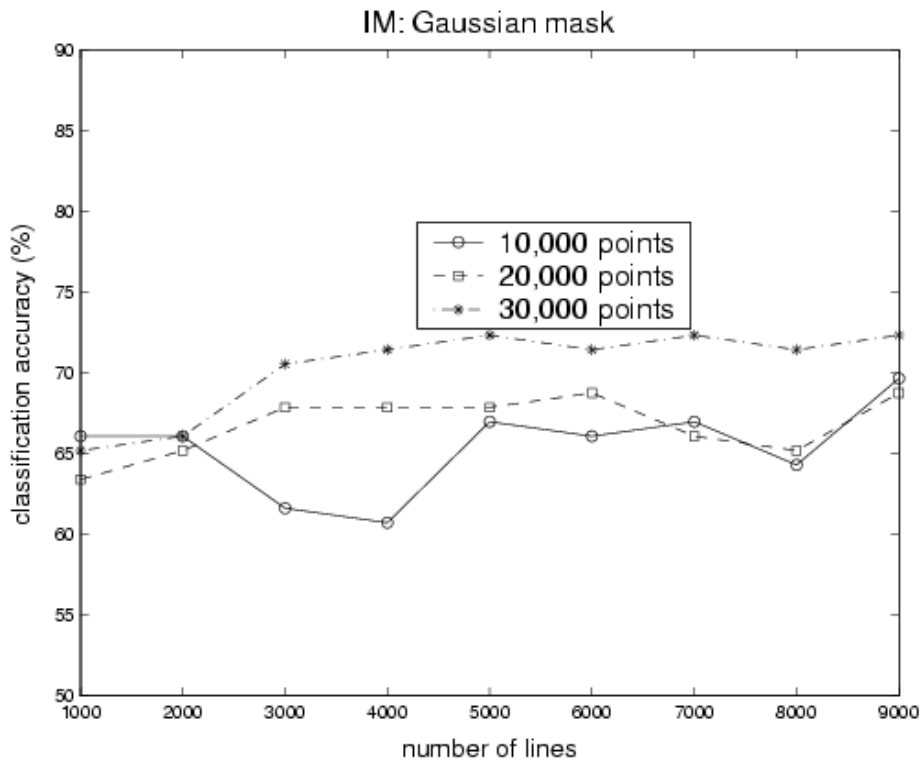
Gaussian masks 10,000 points

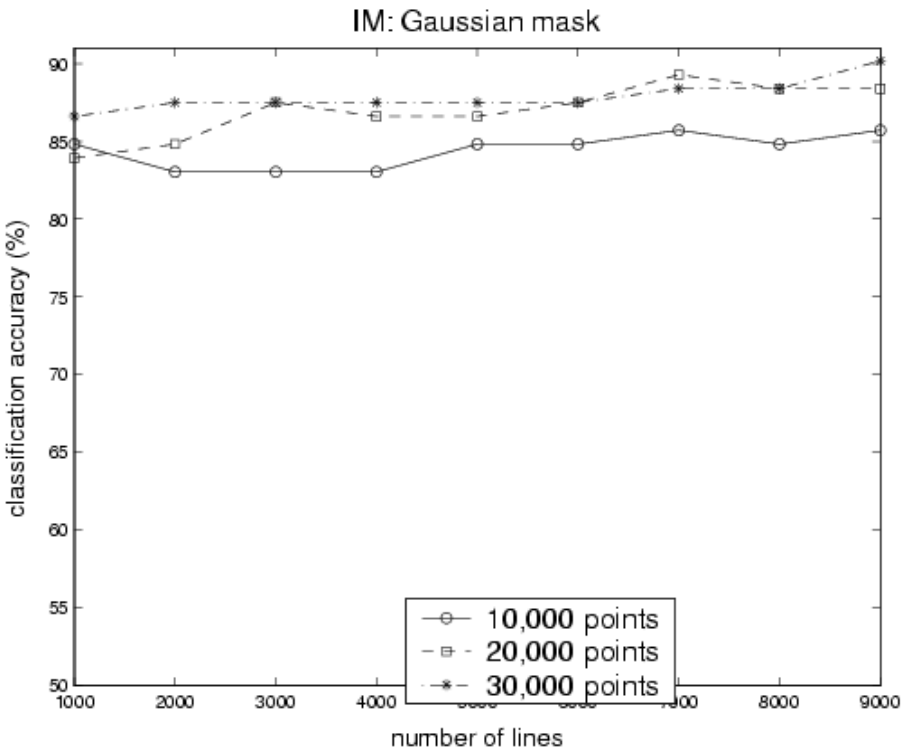
Texture recognition from Irregularly sampled data



Log-polar masks 10,000 points

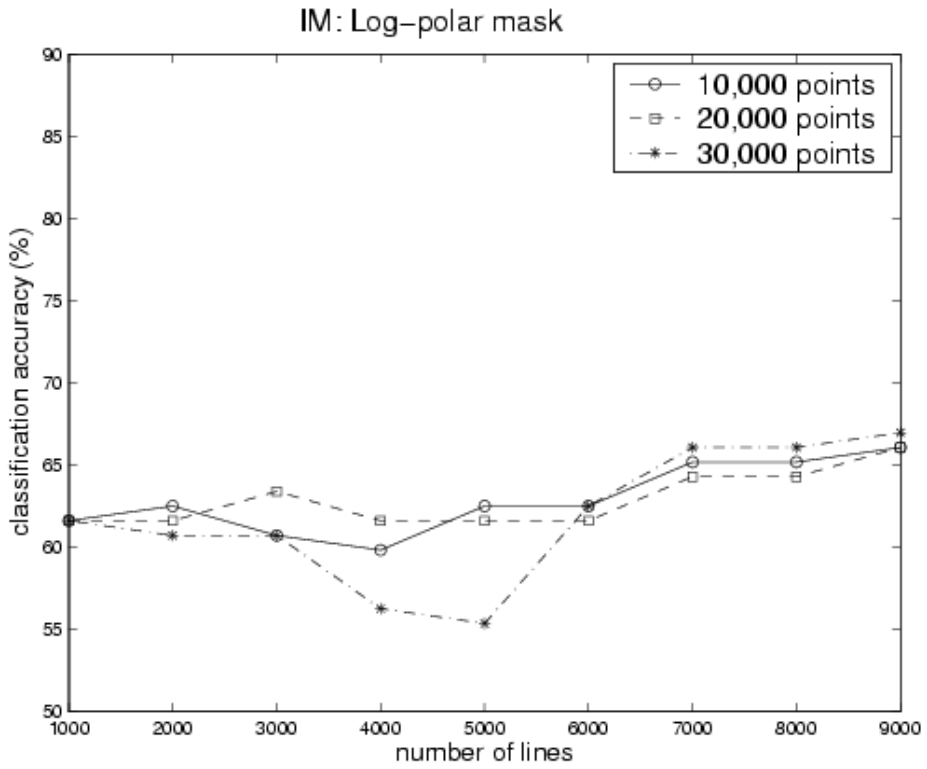
Texture recognition from Irregularly sampled data

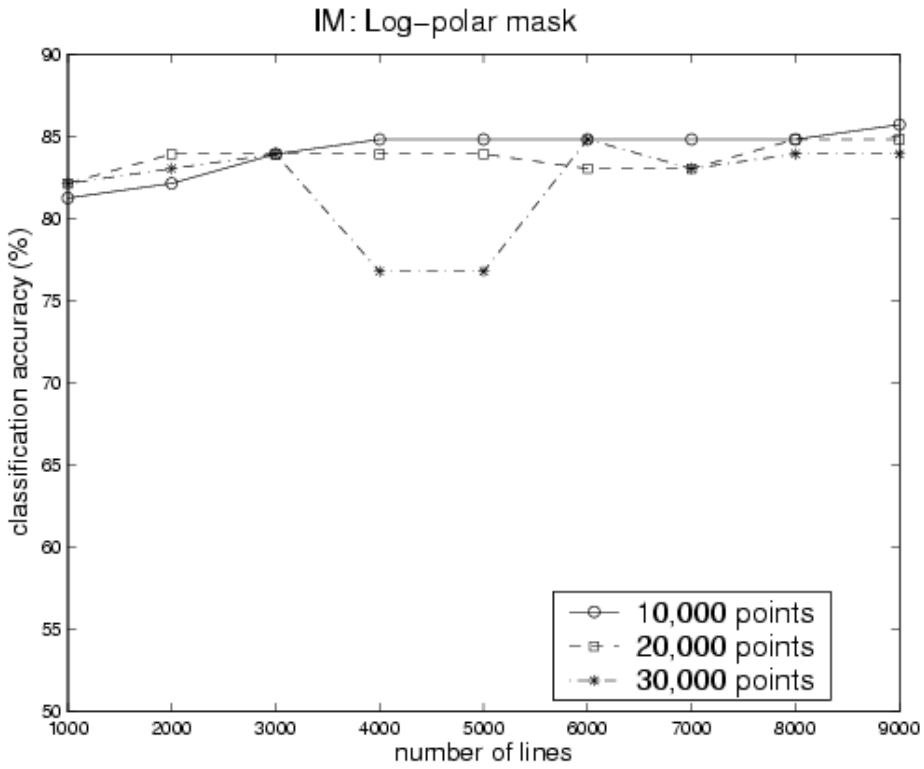




Left: Correct recall in the 1st position.
Right: Correct recall in the first 4 positions

Texture recognition from Irregularly sampled data





Left: Correct recall in the 1st position.
Right: Correct recall in the first 4 positions

How do visual cues work?

- How does information organise itself in the brain?
- How is it retrieved?
- What is the topology of the network of ideas?
- Is it different when the cues are visual from when the cues are verbal?

Designing an experiment

- People were shown 100 images of objects
- They were asked to find the most similar to a randomly picked one…
- People were shown the 100 names of the same objects.
- They were asked to pick the most similar one…

The conclusions

- The experiment was too limited to conclude on the topology of the network, BUT…
- It showed that both visual ideas and auditory ideas are organised in networks of similar characteristics, only different networks!

M Petrou and R Piroddi, 2006. “On the structure of the mind”, Proceedings of AISB'06: Adaptation in Artificial and Biological Systems, T Kovacs and J Marshall (eds), Vol 2, pp 60-63.

The future….

- A lot more to be done…
- Many aspects to be explored that may keep several PhD students going on for years and may lead to very exciting advances of technology…