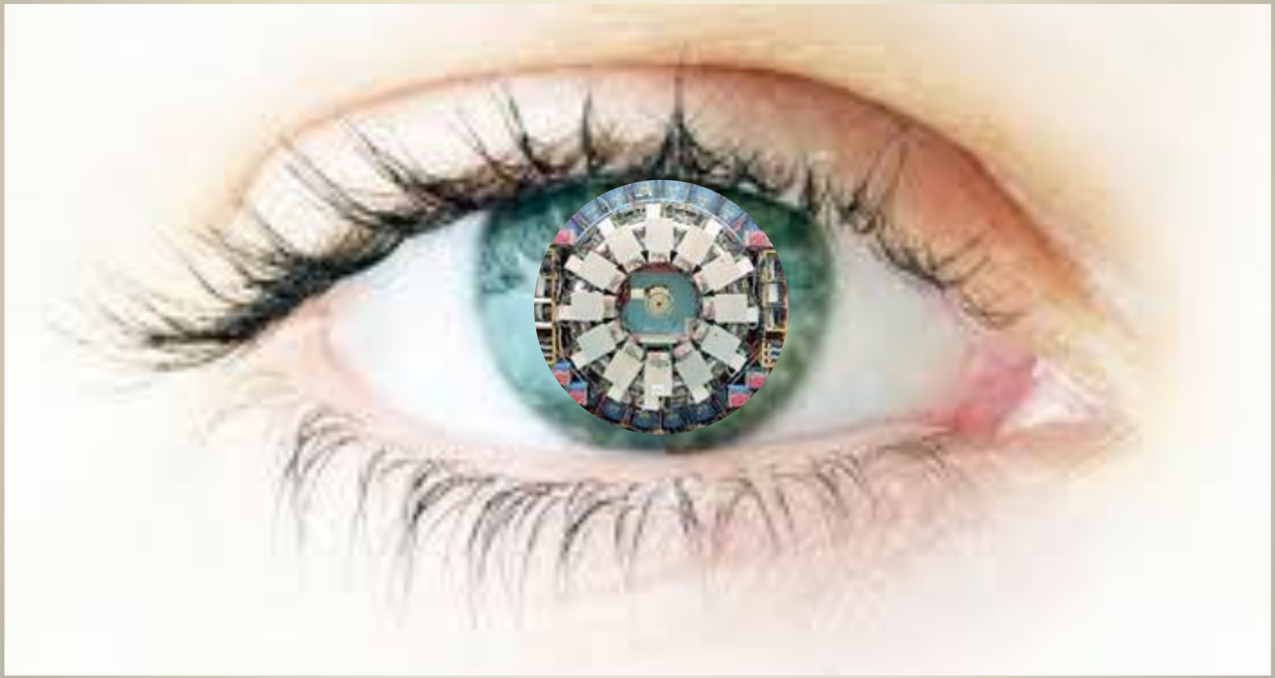
THE BRAIN AS A TRIGGER SYSTEM



Maria Michela Del Viva

Giovanni Punzi

UNIVERSITÀ DEGLI STUDI FIRENZE

UNIVERSITÀ DI PISA

PREFACE

- We spent some time exploring the analogies and differences between high-end trigger devices for experimental particle physics, and the mechanisms of natural vision
- We found that the analogy was more than superficial, and we learned a number of interesting things in both fields.[PloS one DOI:10.1371/journal.pone.0069154]

DATA ACQUISITION IN PARTICLE PHYSICS

- High-Energy Physics experiments produce particle collisions at high rates (MHz), and large event sizes – especially in large hadron colliders like the LHC, or the Tevatron before it.
- Cannot save all events for analysis (typical $<10^{-3}$): need to select interesting events in **real time – and to do it right.**
- Trigger is any device to select events, thus reducing data. However, we are interested in triggers that make complex decisions, based on a large fraction of the event data.
- To do such complex analysis in a short time, special techniques have been developed. In particular, extraction of meaningful information requires a strong data reduction internal to the trigger

NOT THE ONLY CASE OF LOTS OF DATA AND LITTLE TIME TO DECIDE

- The brain extracts biologically-relevant information from a large amount of input data.
- This must happen very quickly to initiate proper autonomic and motor response

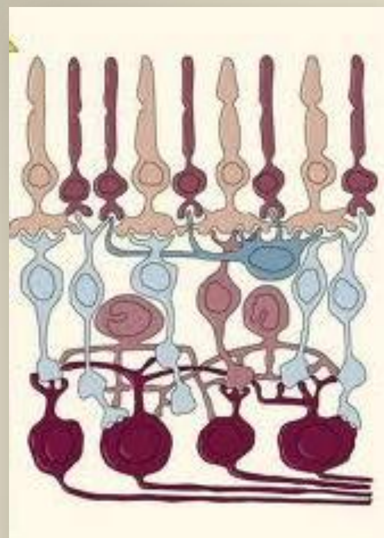
NOT THE ONLY CASE OF LOTS OF DATA AND LITTLE TIME TO DECIDE

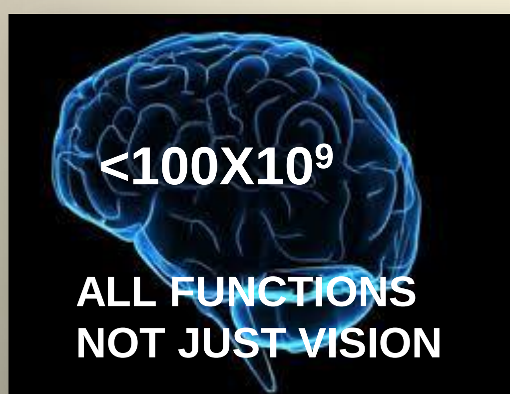
AMOUNT OF DATA

$9.2\text{X}10^7$ Rods +
$4.6\text{X}10^6$ Cones

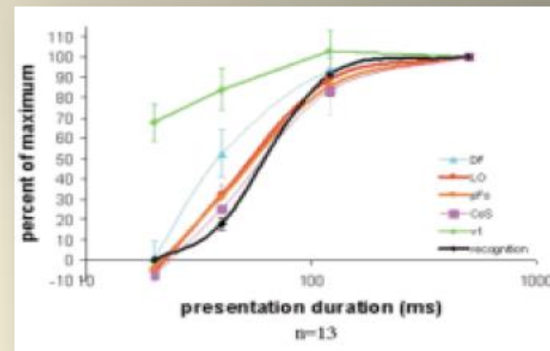
20.316 Gb/s

$1\text{X}10^6$ Optic Nerve fibers
0.8 Gb/s - 4 Gb/s





TIMING



- Early human visual areas can process images at 30-40Hz, with latencies <100ms
- Typical switching time of neurons ~1ms

1 KHZ FOR VISION = 1 GHZ FOR PP

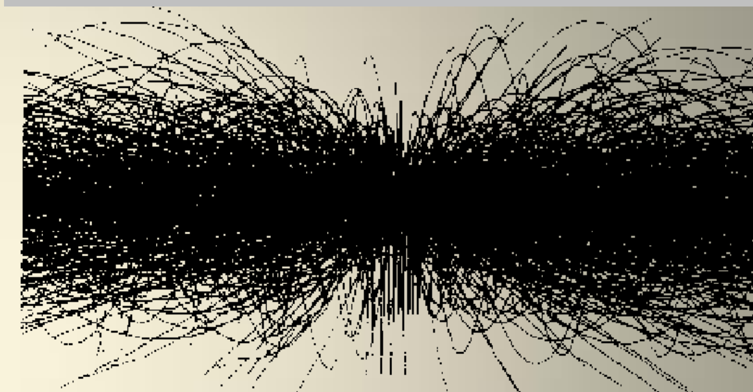
DATA REDUCTION PROBLEMS, COMPARED

VISION

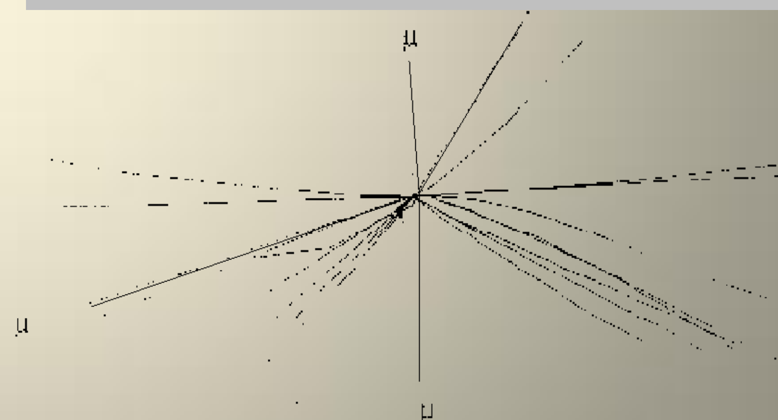
It is has been proposed that the visual system solves this problem by creating a compact summary of the image (“sketch”), based on few simple features. [Marr, 1982].

HEP

Full information



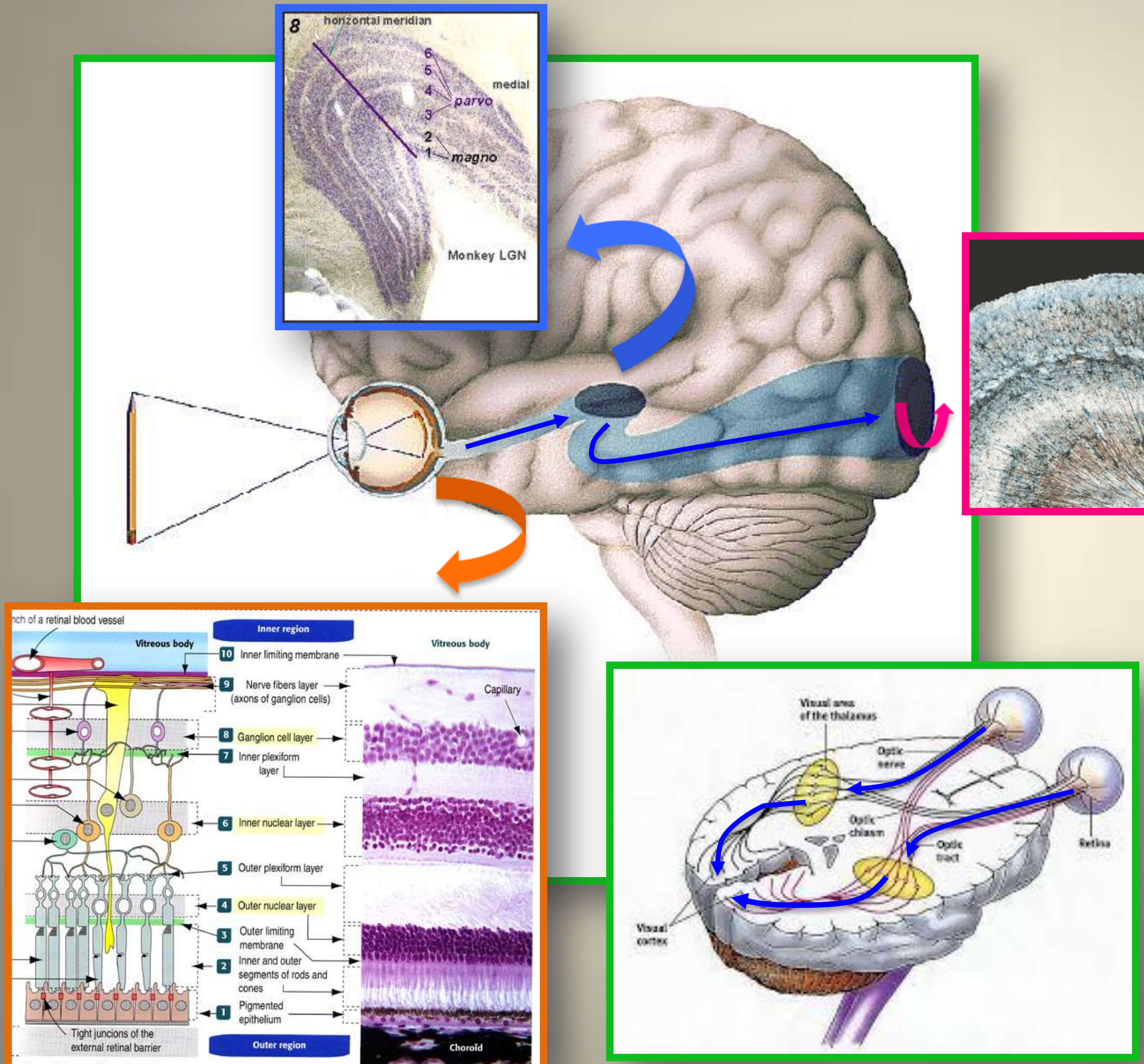
Extracted information

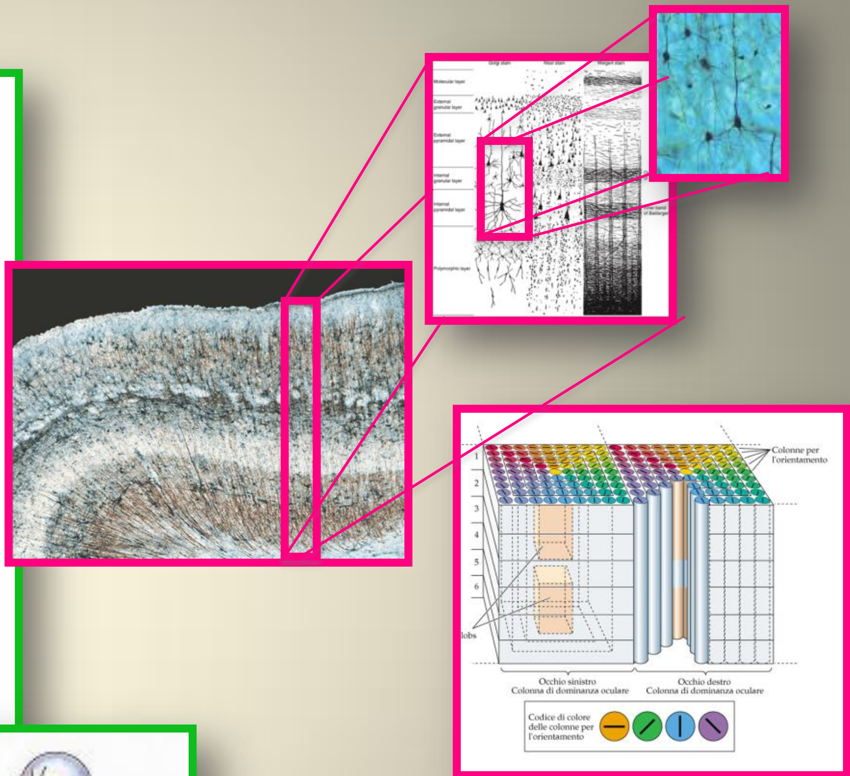


How similar are the two problems ?

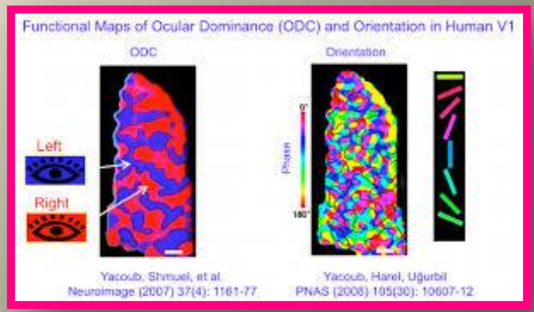
EARLY VISION AND DATA ACQUISITION IN HEP

A LOOK AT PHYSICAL IMPLEMENTATIONS



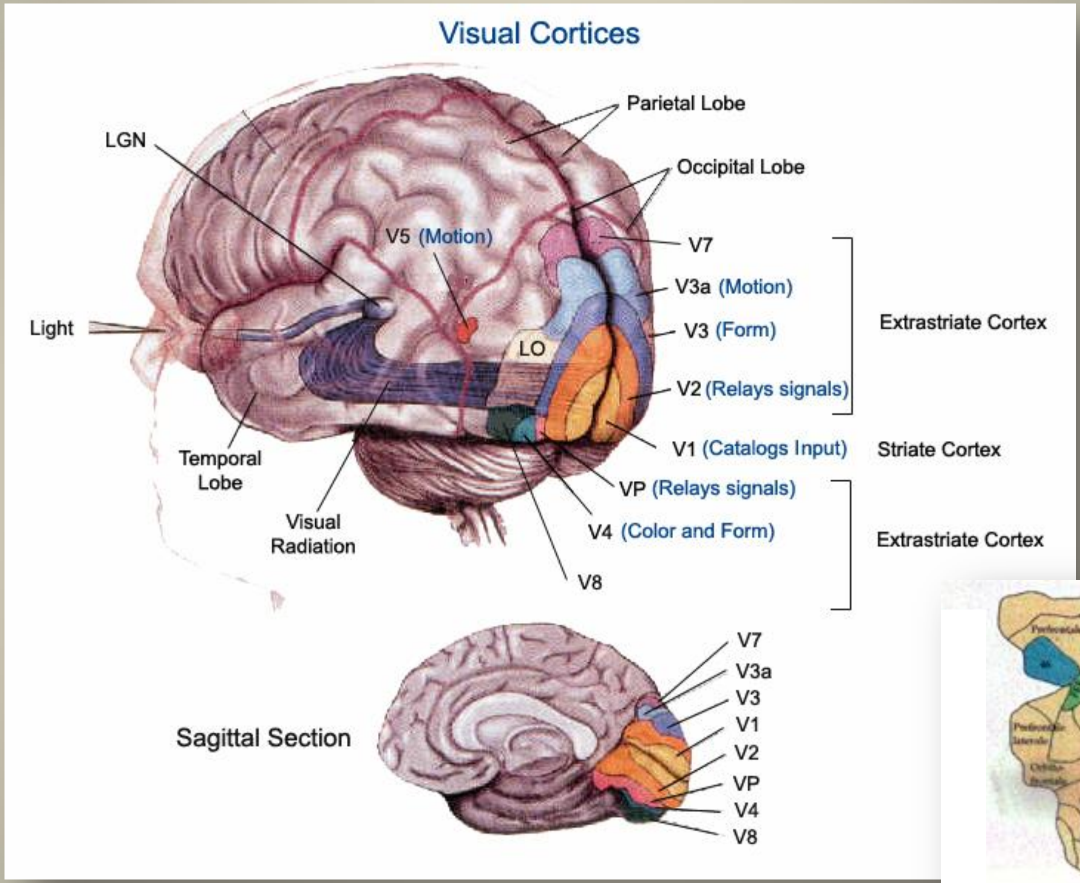


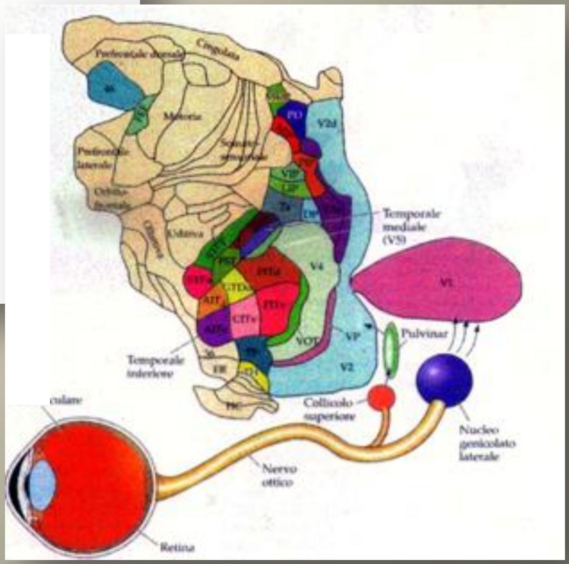
RETINOTOPIC ORGANIZATION



MANY DIFFERENT VISUAL AREAS

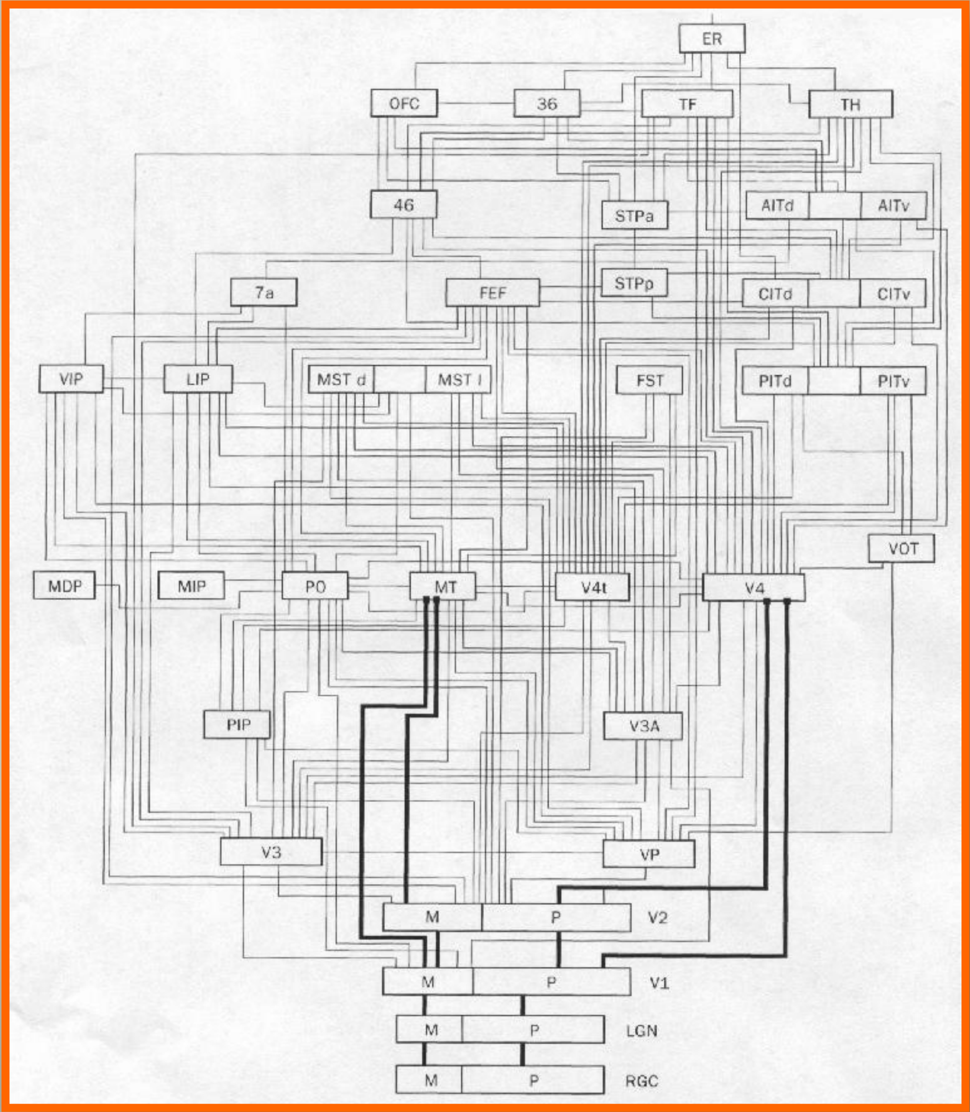
DIFFERENT REPRESENTATIONS OF VISUAL STIMULI IN MANY DIFFERENT AREAS



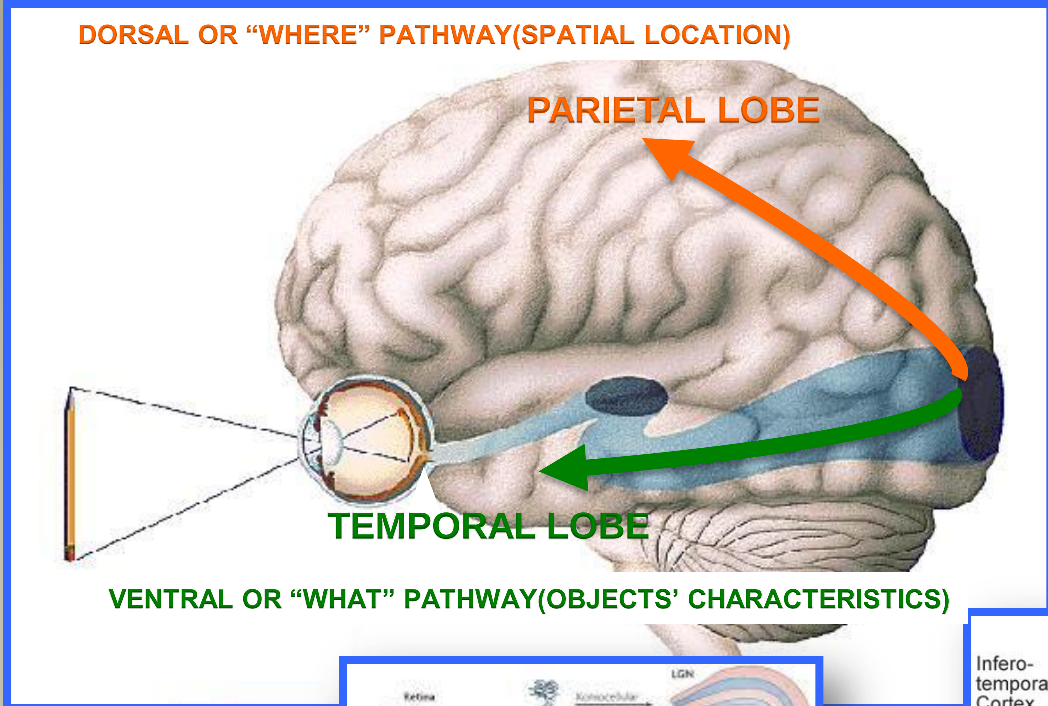


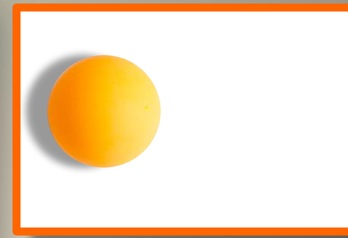
MACAQUE: OVER 30
RETINOTOPIC VISUAL AREAS

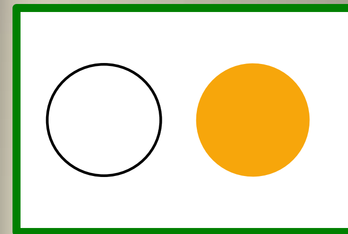
VERY COMPLEX NETWORK

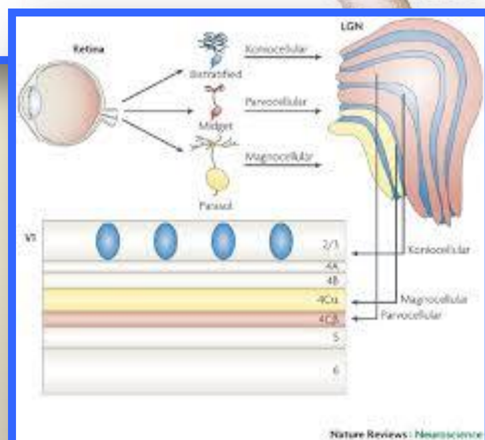


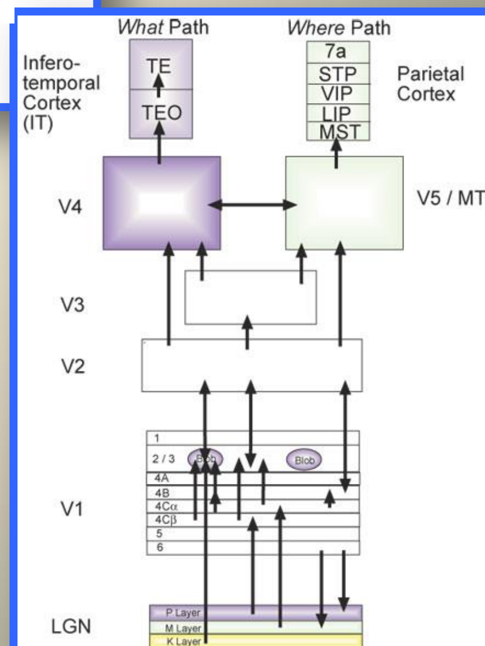
….SO FAR 2 MAIN PATHWAYS







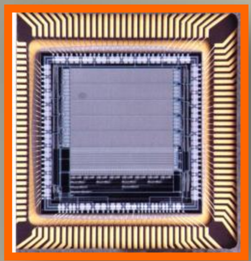


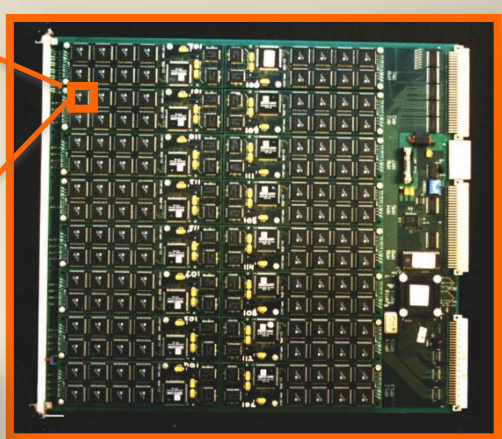


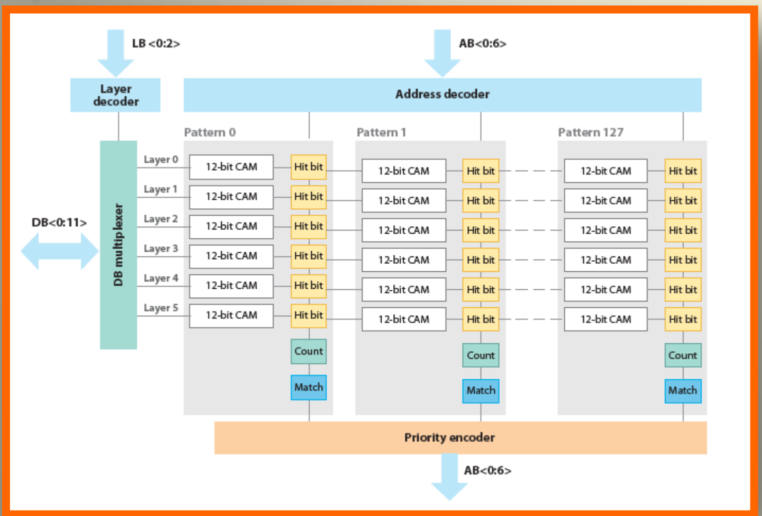
….an yet the networks and algorithms allowing us to see this are still largely unknown

EARLY VISION AND DATA ACQUISITION IN HEP

A LOOK AT PHYSICAL IMPLEMENTATIONS





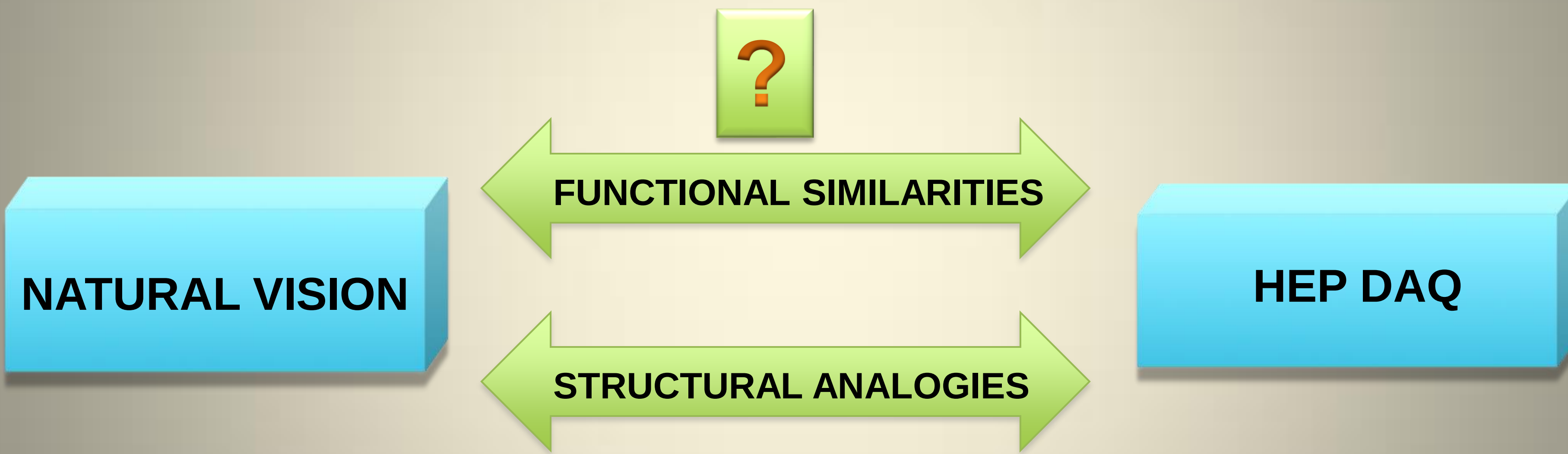


SVT TRIGGER@CDF was used to reconstruct tracks going through the detector in ~10µs

Based on custom parallel VLSI devices Associative Memory [NIM A278, (1989), 436-440]

Continuing in FTK for ATLAS.

EARLY VISION AND DATA ACQUISITION IN HEP



EARLY VISION AND DATA ACQUISITION IN HEP

BUT….SO SIMILAR IN COMPUTATIONAL FUNCTIONALITY

NATURAL VISION	HEP DAQ
Extensive data reduction must operate at an early stage [Attneave, 1954; Barlow, 1961]	Offline storage is limited
A strong, lossy data-reduction is highly likely [Zhaoping, 2006]	Use strongly reduced information
Size of brain limited	Size of device limited by cost
Limits to energy consumption [Attwell and Laughlin, 2001]	Limits to electrical power
Number of visual neurons and their discharge rate not sufficient to process all data. [Lennie, 2003; Echeverri, 2006]	Amount of bandwidth and computing power at higher trigger level cannot process the full rate.

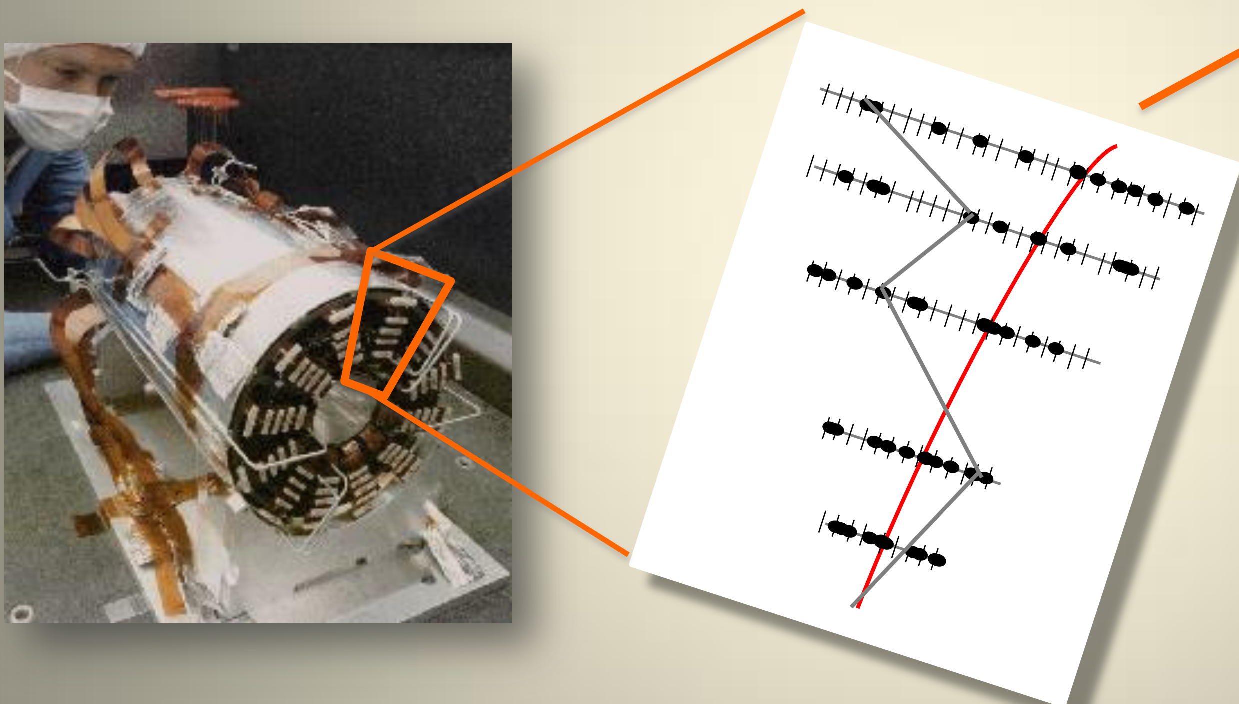
LOOKING FOR A SOLUTION

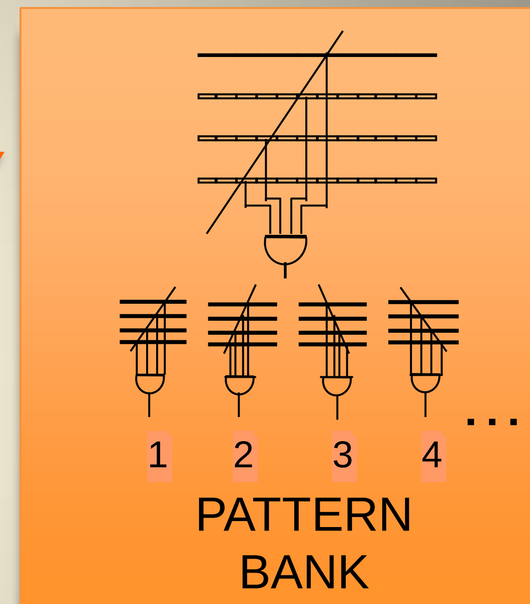
Could we learn something from known artificial systems?

SAME SOLUTION ?

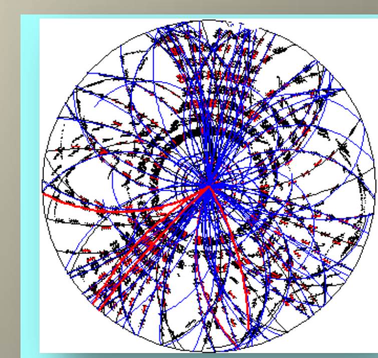
HEP EXAMPLE: PATTERN MATCHING

A pattern is a sequence of hits in the detector layers and it is represented by a set of coordinates. A particle trajectory is a specific sequence of hits.
Hit coordinates are read out sequentially into the AM, and compared in parallel to a set of pre-calculated “track patterns”, stored in a *pattern bank.*





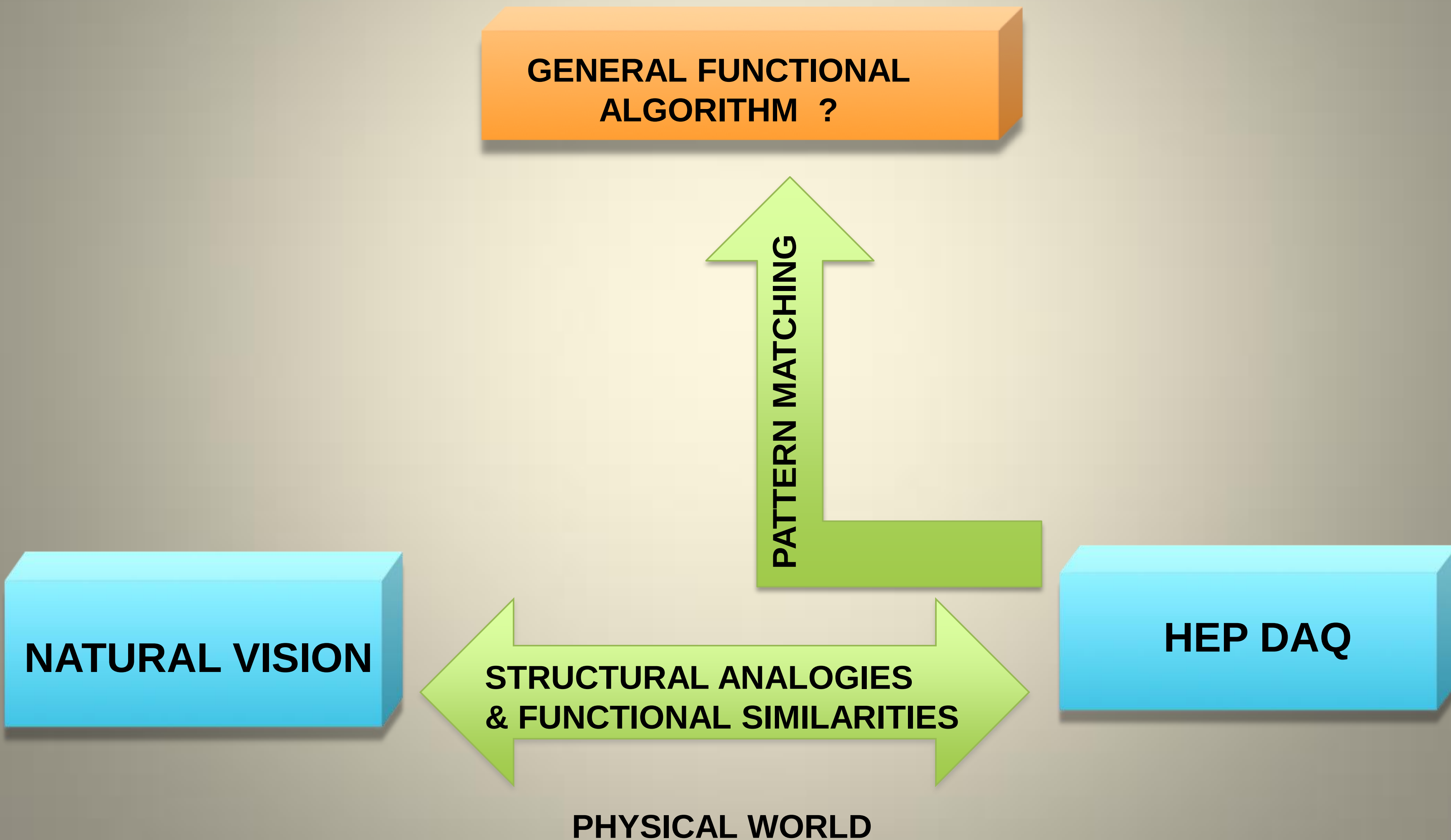
REAL EVENT



A specialized solution to a very specific problem.

Or is it ?

LET'S GENERALIZE AND OBSERVE THE TWO SYSTEMS FROM A FUNCTIONAL POINT OF VIEW

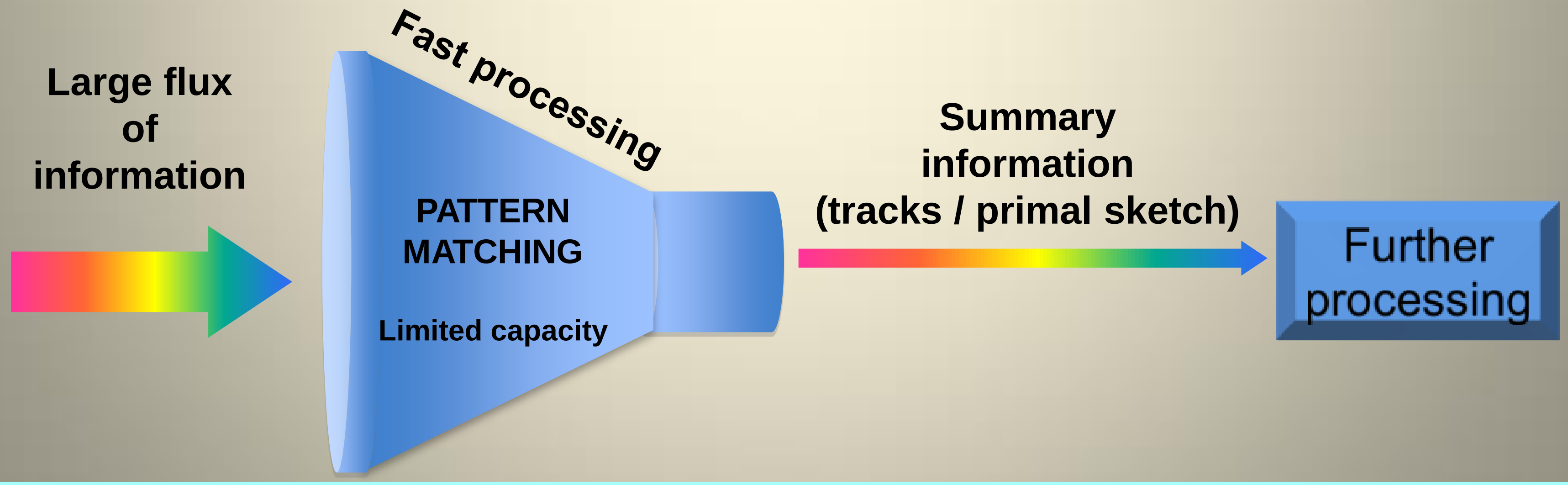


AN ABSTRACT MODEL FOR DATA REDUCTION

We have an "Information Processing System" receiving complex inputs that is expected to provide in output a "summary" of the information for another device to perform further processing.

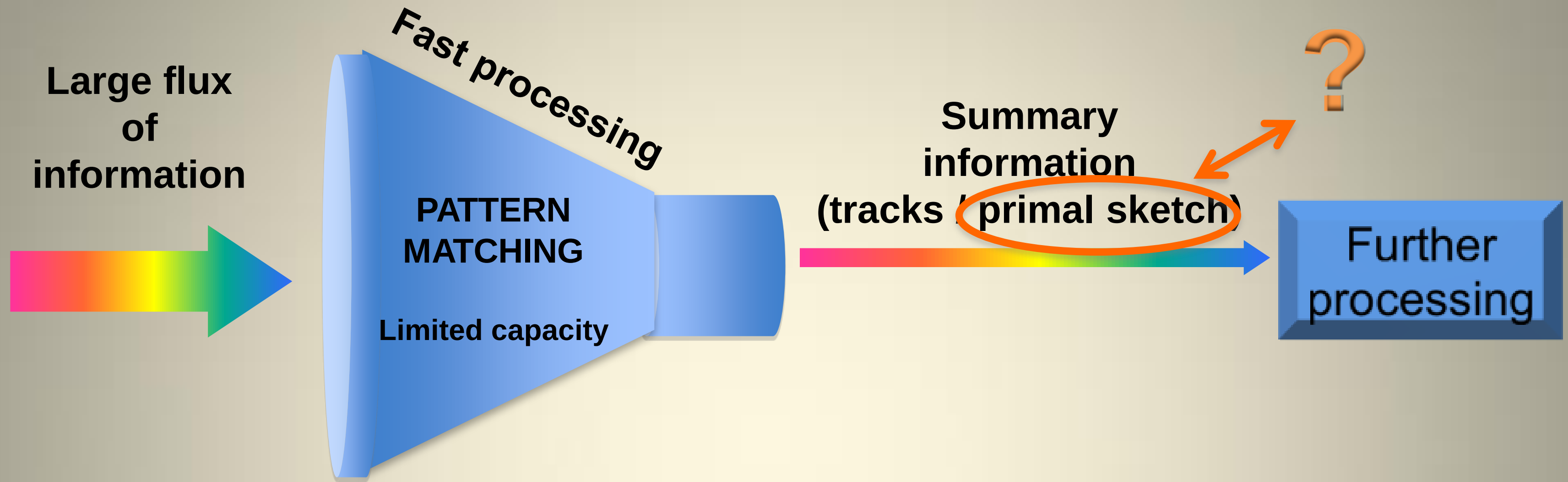
WE POSTULATE AN ALGORITHM BASED ON:

- The summary is based on **recognizing** a limited number of **meaningful patterns** of the input, dropping the remaining information (Pattern Matching)
- The device has finite **MEMORY**: fixed number of recognizable patterns
- The **OUTPUT BANDWIDTH is FIXED** by limitations of the next stage



THIS IS AN UNUSUAL WAY TO APPROACH THE PROBLEM IN VISION

AN ABSTRACT MODEL FOR DATA REDUCTION



For vision it is not obvious what is the information used for the sketch

We assume the system has been optimized by evolution

QUESTION:

WHAT IS THE OPTIMAL WAY TO SUMMARIZE INFORMATION ?

OPTIMAL SUMMARY OF INFORMATION FROM THE MAXIMUM ENTROPY PRINCIPLE

WE ASSUMED THE BEST STRATEGY IS TO CHOOSE THE PATTERN SET THAT MAXIMIZES OUTPUT ENTROPY H

Having assumed a discrete patterns representation makes it easy to calculate entropy as

$$H = \sum_{i}^{N} - p_i Log(p_i)$$

Where p_i is the probability of occurrence of each pattern in the input and N is number of patterns being considered

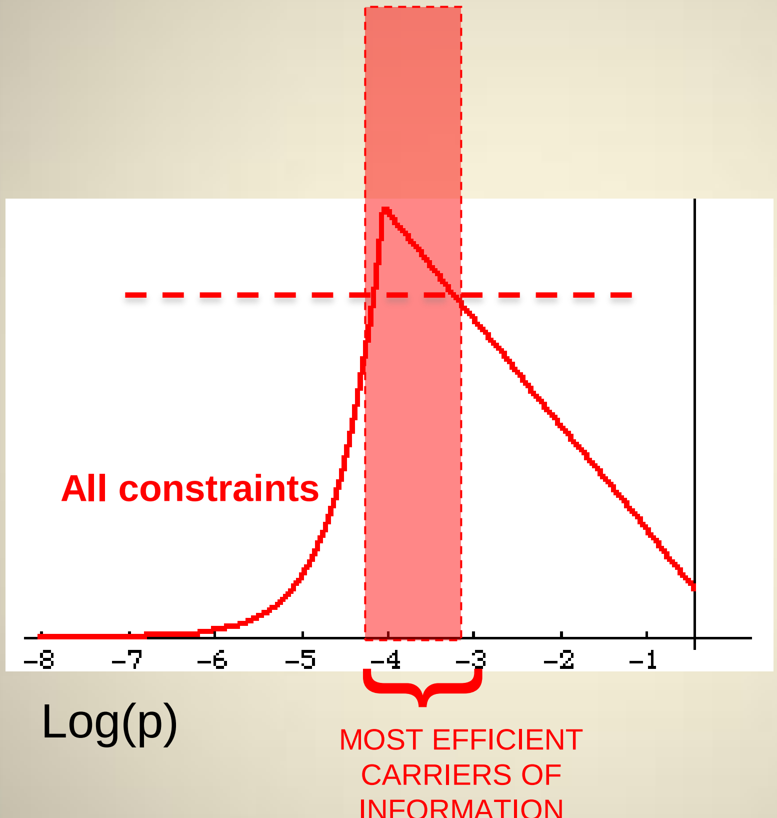
In absence of constraints, maximization is attained by including all possible input patterns => trivial solution: transfer to the output the whole input information

THE KEY TO A MEANINGFUL CHOICE IS THE EXPLICIT INCLUSION OF THE LIMITATIONS OF THE SYSTEM.

INCLUDE LIMITATIONS: CONSTRAINED MAXIMUM ENTROPY

Figure of Merit: ENTROPY YIELD PER UNIT COST:

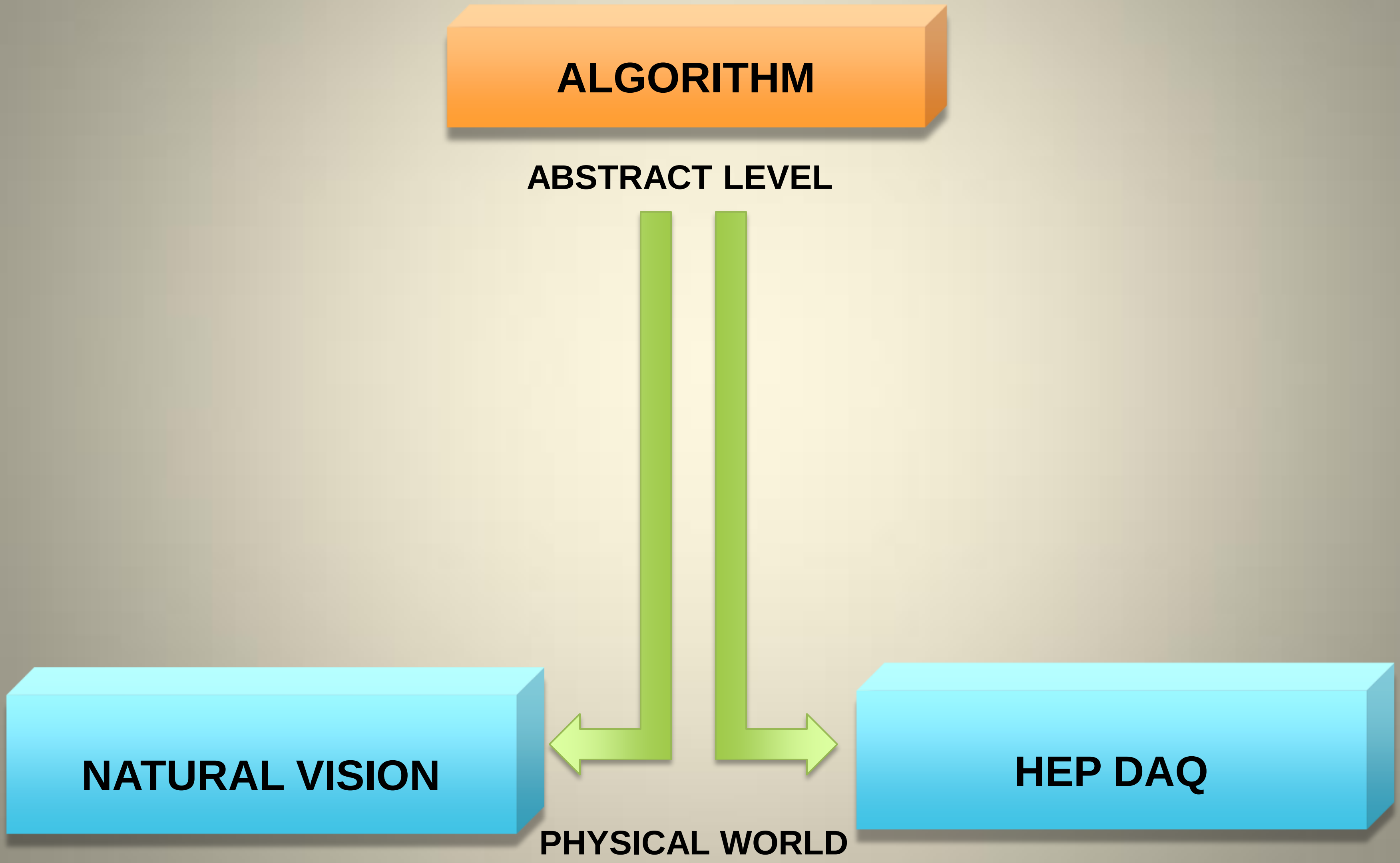
$$f(p) = \frac{-p\log(p)}{\max(1/N,\ p/W)}$$



Accounting for Output bandwidth and number of stored patterns

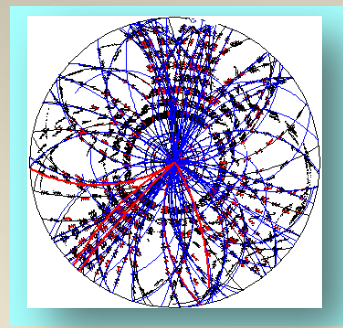
"worst-case" cost for each pattern: the larger of "storage cost" 1/N and "bandwidth cost" p_i/W, (W = maximum allowed total rate of pattern acceptance, $\Sigma\, p_i < W$.)

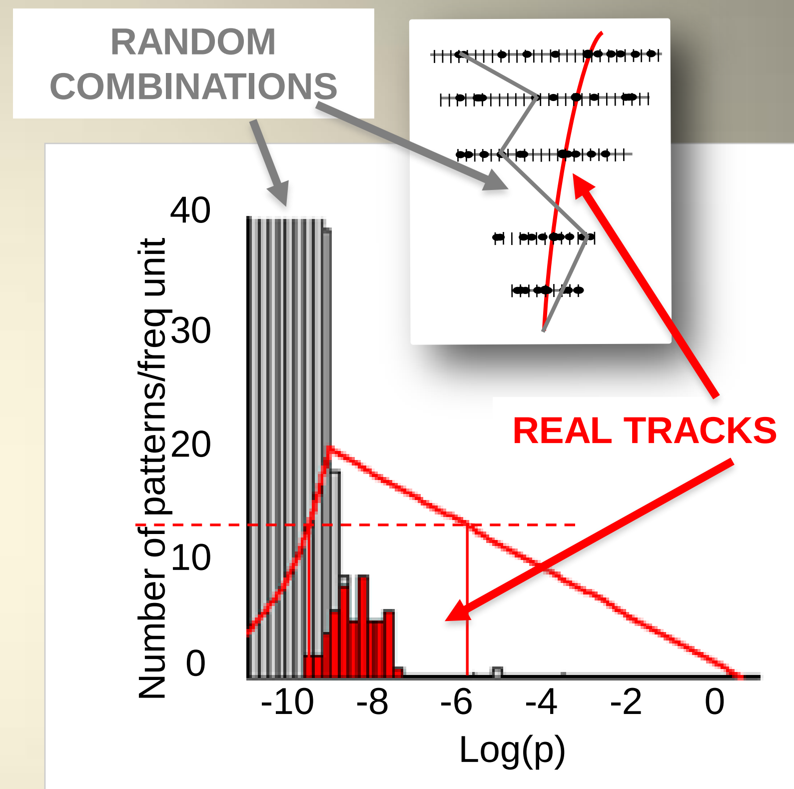
LET'S IMPLEMENT THE GENERAL ALGORITHM



TRACKING AND CONSTRAINED-MAXIMUM-ENTROPY

Realistic Montecarlo simulation of events detected by a wedge of a 5-layer detector (SVX-CDF)





REAL TRACKS CORRESPOND TO THE PEAK IN OUR FOM FUNCTION

NEW VIEWPOINT

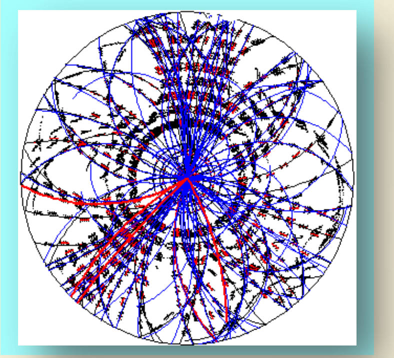
Tracks in a detector can be described as the piece of information that carries the maximum amount of information, under bandwidth and size constraints

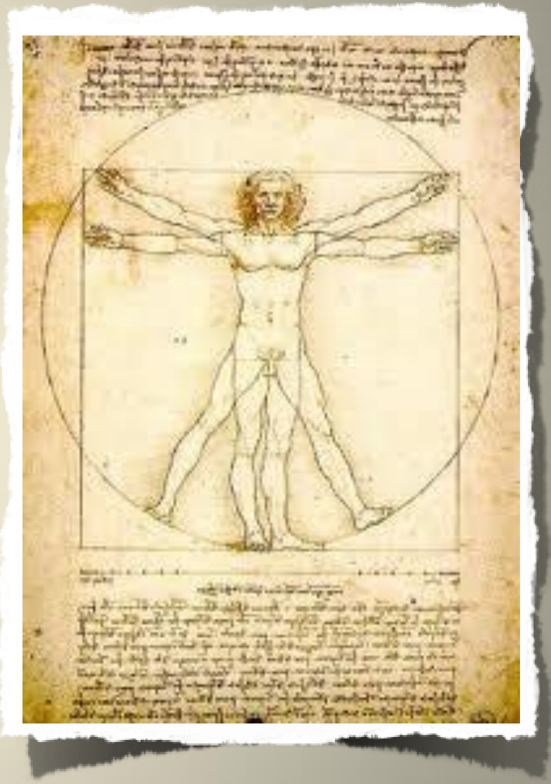
N.B. This implies that tracks can be recognized without prior knowledge of detector geometry. (May be a useful concept for alignment ?)

DATA REDUCTION IN EARLY VISUAL PROCESSING

THE CONSTRAINED-MAXIMUM-ENTROPY PRINCIPLE WORKS !

Having gained confidence about the method on a problem where the solution is known let’s apply it to vision where little is known about the nature of visual features composing these early compressed visual representations



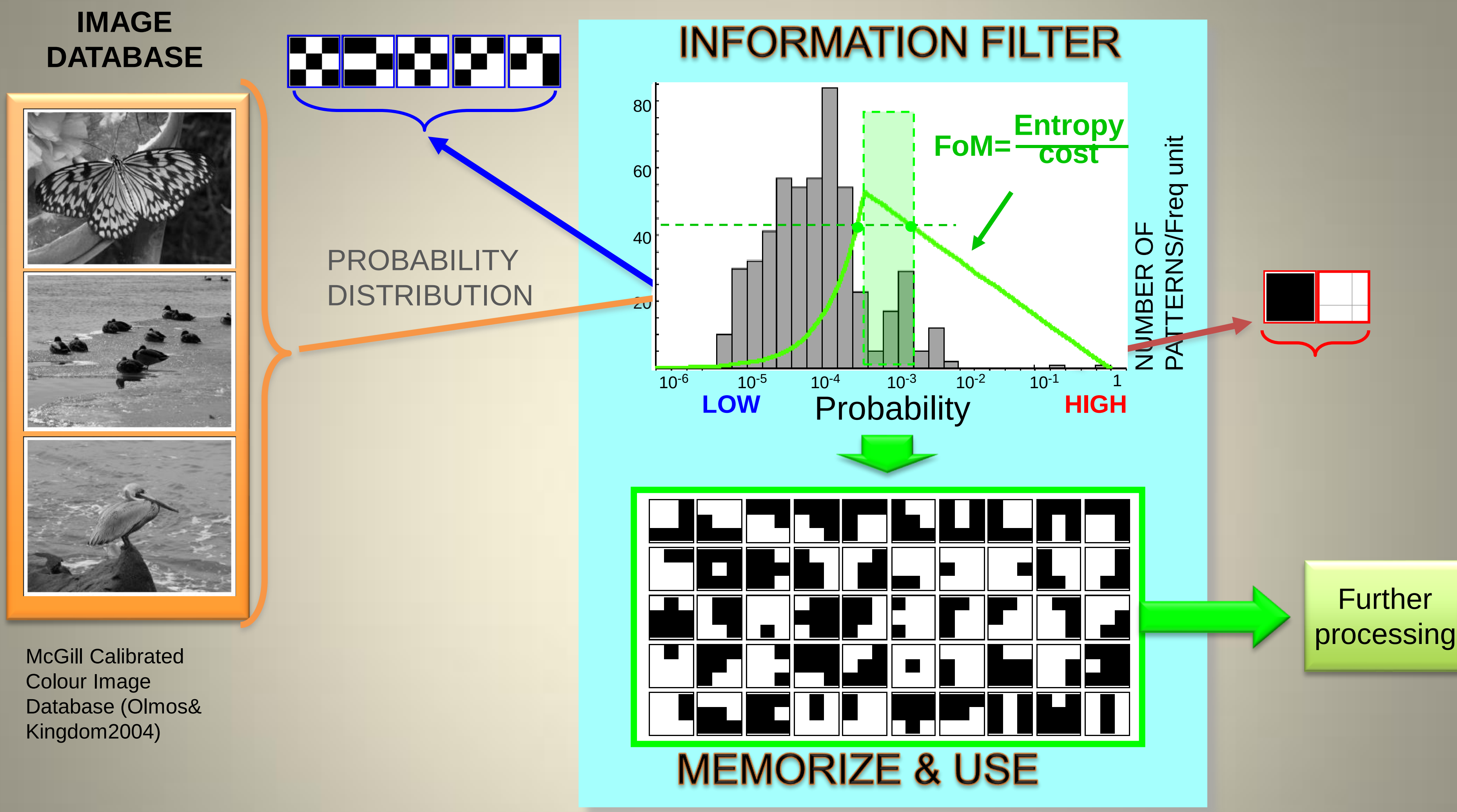


WHAT IS A PATTERN OR FEATURE FOR THE VISUAL SYSTEM?

For example any subset or patch of an image with a certain size

Example: B/W image
3x3 patches

EXTRACTION OF OPTIMAL PATTERNS



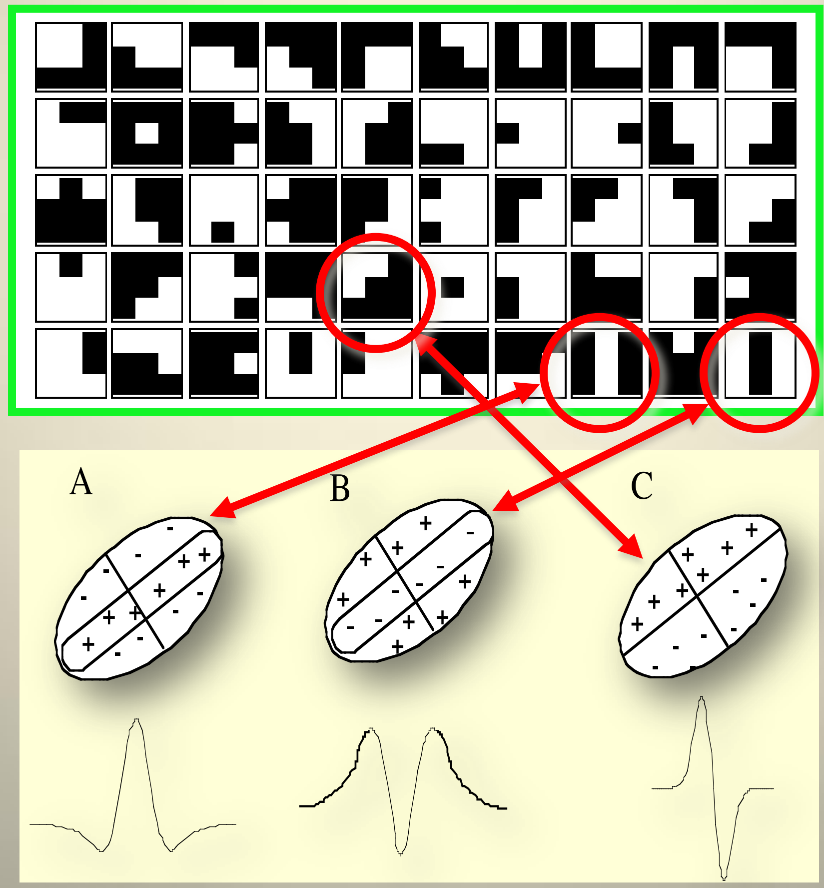
Count how many times a certain pattern occurs in the input (probability)

Keep only those with the predicted probability and discard the others

COMPARE WITH PHYSIOLOGY DATA

In the approximation of 1 bit and 3X3 pixel structure, selected pattern configurations are similar to well known spatial configurations of neurons (receptive fields) in primary visual areas (e.g. V1) : bars ,edges and corners with different spatial orientations (Hubel & Wiesel, 1962).

COMPARE RECEPTIVE FIELD STRUCTURE WITH SELECTED “OPTIMAL” PATTERNS

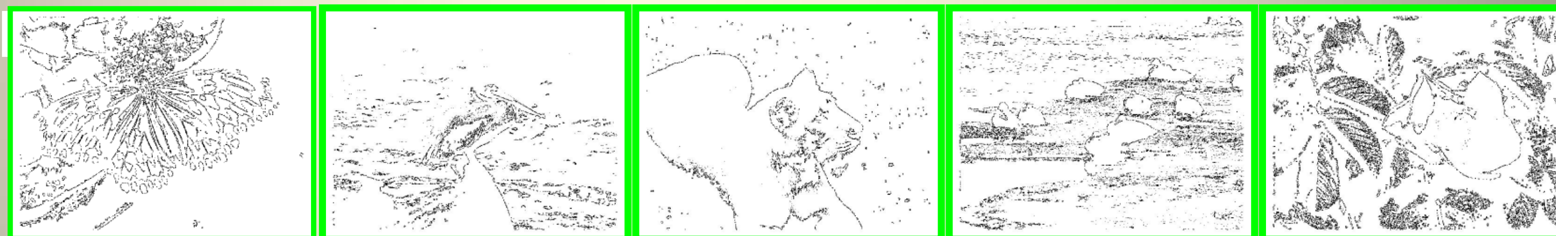


EXTRACTING SKETCHES

ORIGINAL FULL COLOR

DIGITIZED B/W
size 3.5 Mbit
H_{TOT}=271 kbit

FILTERED N=50
Size= 2.5 %
H = 9.8 % H_{TOT}



- SALIENT FEATURES ARE PRESERVED, DESPITE STRONG REDUCTION AND MASSIVE LOSS OF INFO
- RESULT DRIVEN ONLY BY STATISTICAL PROPERTIES OF INPUT AND CONSTRAINTS
- NO A-PRIORI “KNOWLEDGE OF THE WORLD”
- EDGE DETECTION IS DICTATED BY THE NEED FOR EFFICIENT INFORMATION TRANSMISSION, GIVEN CONSTRAINTS

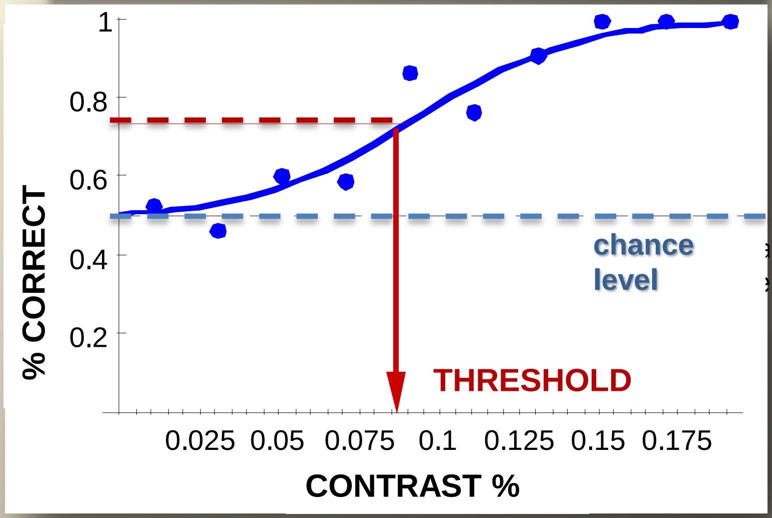
Simulation provides indirect evidence that the model works

But how can one possibly get direct evidence ?

MEASURING VISUAL SYSTEM (PERCEPTUAL) PERFORMANCE

- Perception is the result of probabilistic signaling of a population of neurons.
- While perception has a random component, its probability distribution is REPRODUCIBLE and can be ACCURATELY MEASURED**:** we call this **PSYCHOMETRIC CURVE.**
- Allows us to define THRESHOLDS AND SENSITIVITY.
- Techniques used are well established called(**PSYCHOPHYSICS**)
- Quantitative measurements of perception (eg. stimulus direction or orientation in space, etc) are done as a function of stimulus strength (S/N, luminance contrast..).

CONSTANT ACROSS SUBJECTS: TESTING LOW-LEVEL VISUAL PROPERTIES



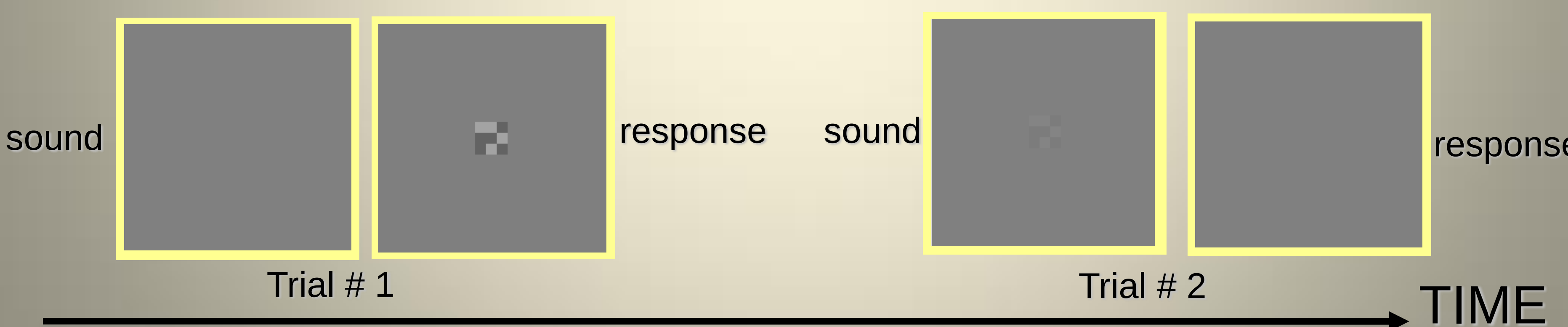
TWO ALTERNATIVE FORCED CHOICE

MEASURING CONTRAST SENSITIVITY TO VISUAL PATTERNS

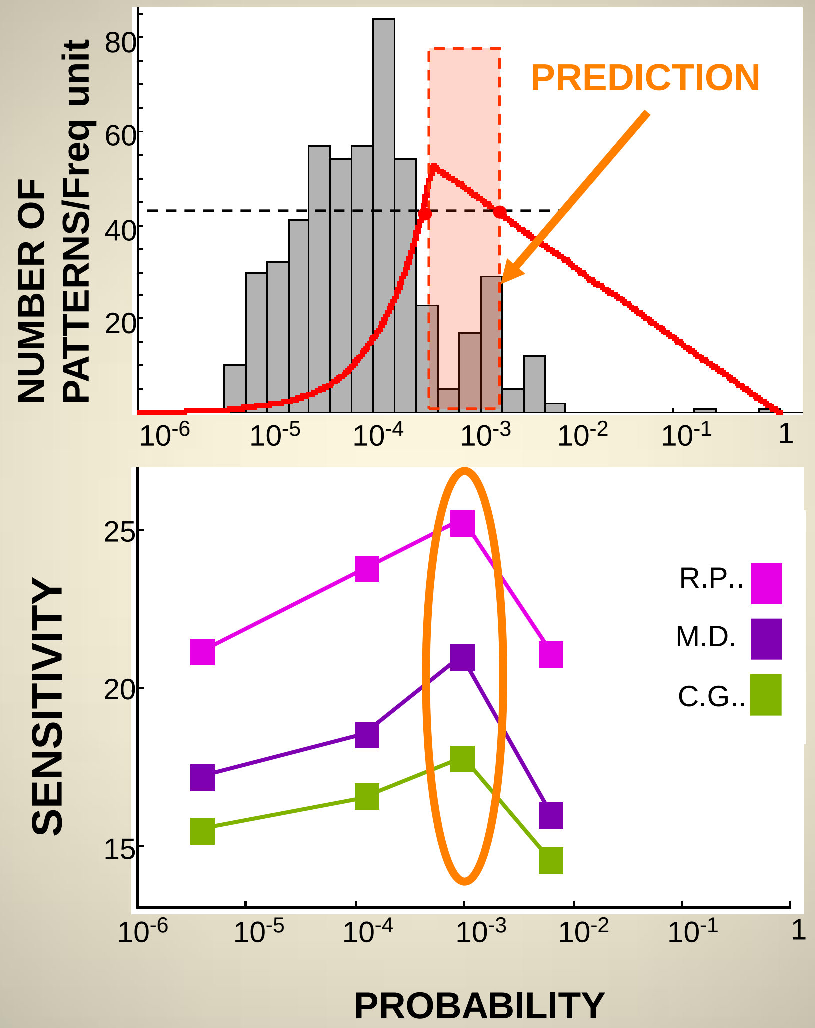
IF THIS IS THE WAY THE VISUAL SYSTEM SUMMARIZES IMAGES

HUMAN SENSITIVITY TO SELECTED PATTERNS SHOULD BE HIGHER THAN TO DISCARDED

We measured sensitivity for detection of several kinds of patterns as a function of luminance contrast in a 2IFC procedure



PSYCHOPHYSICAL EVIDENCE



SENSITIVITY TO DIFFERENT PATTERNS HAS THE SAME TREND AS OUR FoM FUNCTION

MAXIMUM SENSITIVITY FOR PREDICTED OPTIMAL PATTERNS

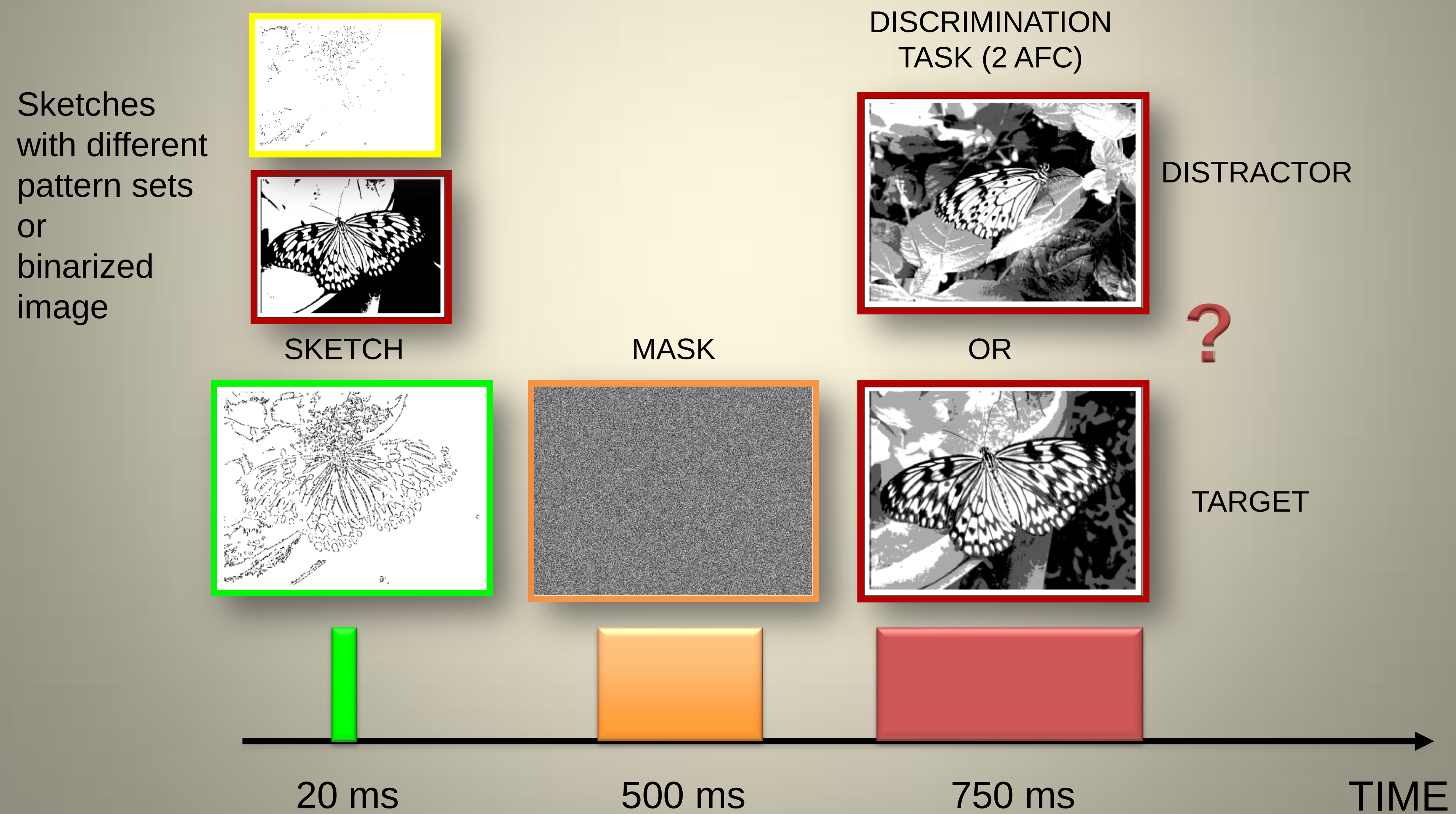
DOES THE VISUAL SYSTEM USE THE PATTERNS WE PREDICT IN EARLY (FAST) VISUAL ANALYSIS ?

If true

1. Our reduced-information sketches should be as easy to recognize as the original images
2. Sketches obtained with NON optimal pattern sets should be more difficult to recognize

DISCRIMINATION PERFORMANCE BASED ON IMAGE SKETCHES

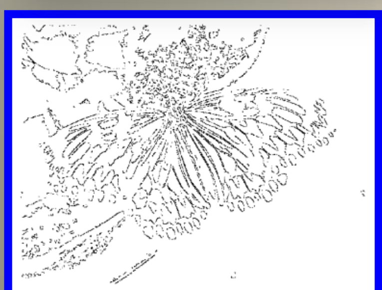
•Use very short presentation times (visual representation at early stages of visual analysis)



RESULTS

Original

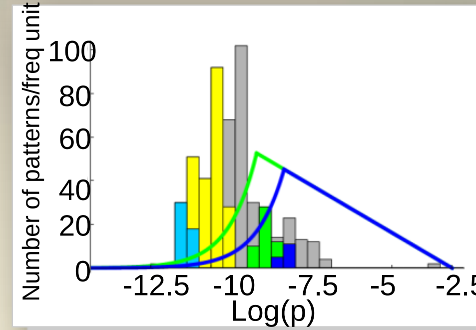
50 optimal patterns
Entropy=9.8%
Compression factor=40



16 optimal patterns
Entropy=5.5%
Compression factor= 67

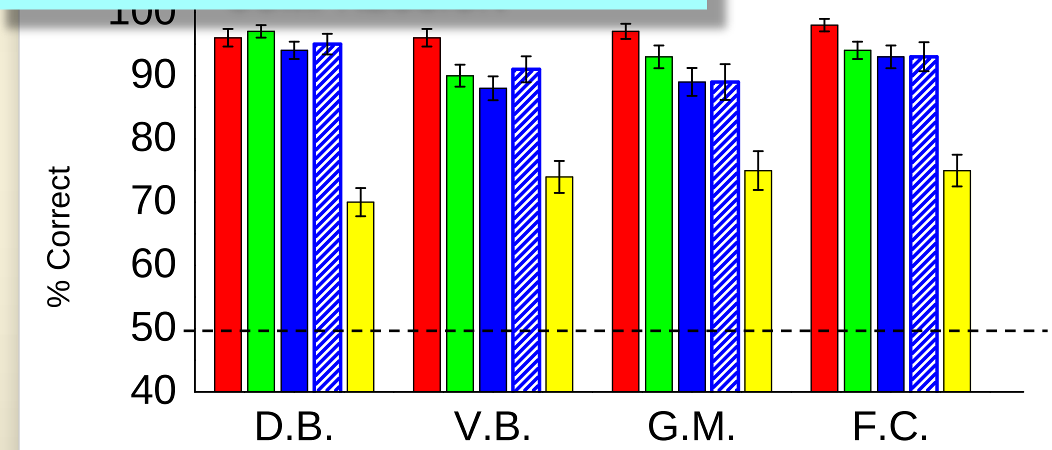
244 NON optimal patterns (rare)
Entropy=5.5%
Compression factor=90

$$\frac{1}{N_{tot}}\int_{f(p)>c} p f(p)\,dp < W$$



THE VISUAL SYSTEM WORKS BY SELECTING PATTERNS

SKETCH VISIBLE WITH HUGE DATA COMPRESSION



IN FAST VISION WE DO NOT SEE ALL THE IMAGE BUT JUST THE SKETCH

THE VISUAL SYSTEM SELECTS **MEANINGFUL** PATTERNS BY MAXIMIZING **CONSTRAINED-** ENTROPY

SUMMARY

- There are important similarities in the underlying functionalities of HEP triggers and the visual system. They are likely due to a process of "convergent evolution" towards a common set of solutions to difficult data handling issues
- Following this insight lead us to improve our understanding of vision

  ☞ Modeling the vision functionality using concepts borrowed from experimental physics led to a new model, different from what previously existed in the field.

  ☞ This new model, proved capable of explaining visual features from "first principles"

  ☞ Many paths for further study (plasticity, color, motion…)
- Comparative study of the two systems can reflect back on HEP experimental techniques. See for instance:
  - L. Ristori "An artificial retina for fast track finding" NIM A453 425-429;
  - G. Punzi "A specialized processor for track reconstruction at the LHC crossing rate" @ INSTR14 Novosibirsk
  - D.Tonelli, N. Neri talks @TIPP'14 (Thursday:Trigger & DAQ session 3)

THANKS!

BACKUP