Gregory's theory: Perception as hypothesis

Deals with classical philosophical problems associated with perceiving.

Concludes that perception is an activity resembling hypothesis formation and testing.

Theory of perception:

It is a chain of events -
Signals received by the sensory receptors trigger neural events. Appropriate knowledge interacts with these inputs to create psychological data. On the basis of the data, hypothesis are advanced to predict and make sense of events in the world.

Work of Richard Langton Gregory, CBE, MA, D.Sc., FRSE, FRS (born July 24, 1923) is a British psychologist and Emeritus Professor of Neuropsychology at the University of Bristol, UK.

- **Eye and Brain . Mind in Science**
- **The Encyclopedia of Ignorance**
- **The Oxford Companion to the Mind**

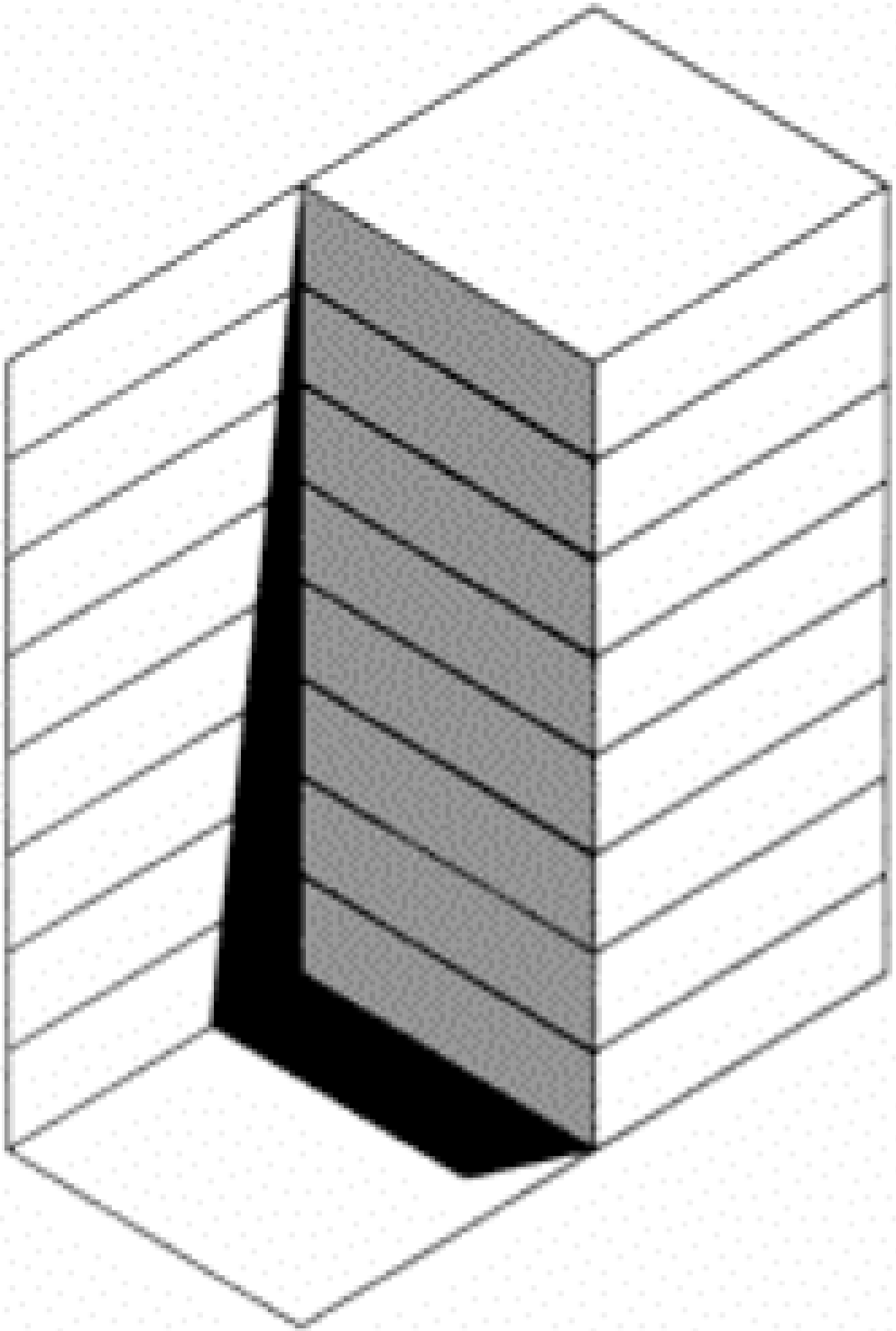
Sensations of consciousness—qualia—are created by the brain.

Qualia are psychologically projected into the external world of Objects (qualia are properties of sensory experiences).

Perceptions are predictive hypotheses, based on knowledge stored from the past.

As perceptions maybe 90% memory, how is the present moment distinguished from memory and anticipation?

Qualia may be evoked by afferent sensory stimuli to avoid confusion with the past and future by flagging the present moment.



Mach's corner.

When the corner flips in depth, the grey region may change quite dramatically in brightness. When the corner is in, it seems lighter than when out.

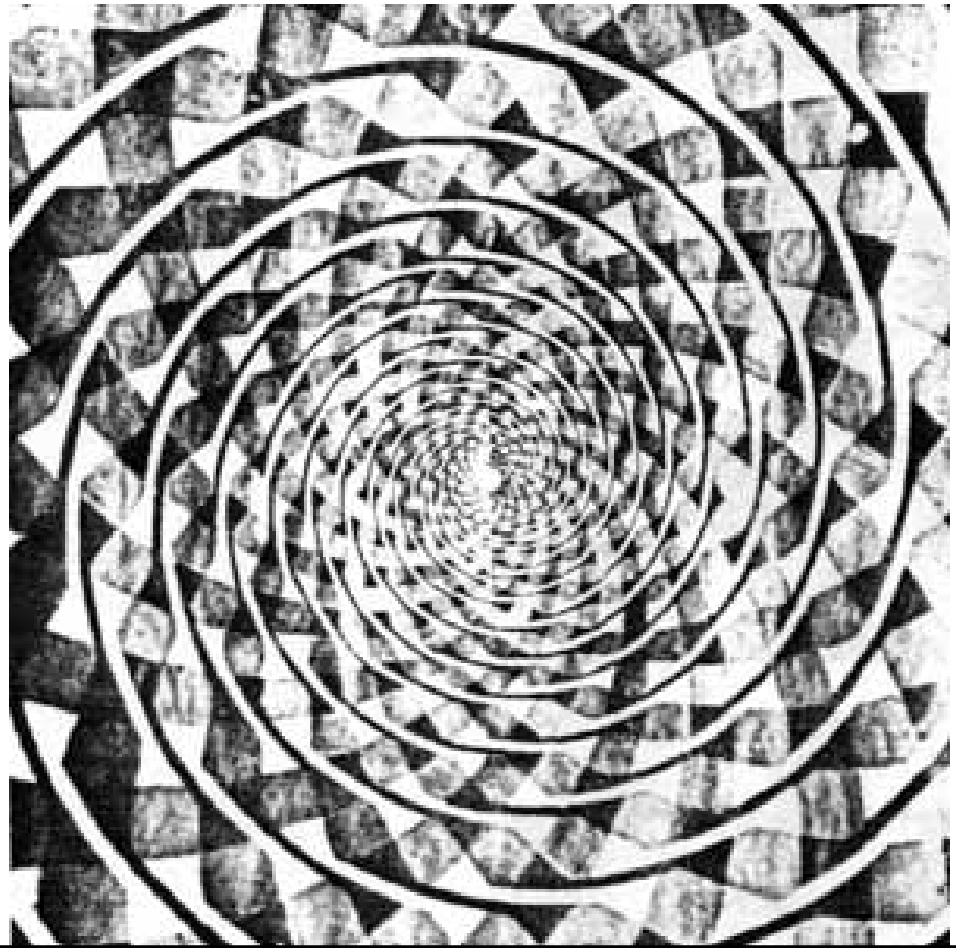
When the flipping corner is in, the grey region may look considerably lighter than when it sticks out. It looks lighter when seen as a shadow than as a mark on the surface, for shadows are normally minimised, as they are not objects for behaviour.

So seemingly simple qualia of consciousness are affected by meanings of perception.

A

12 13 14

C

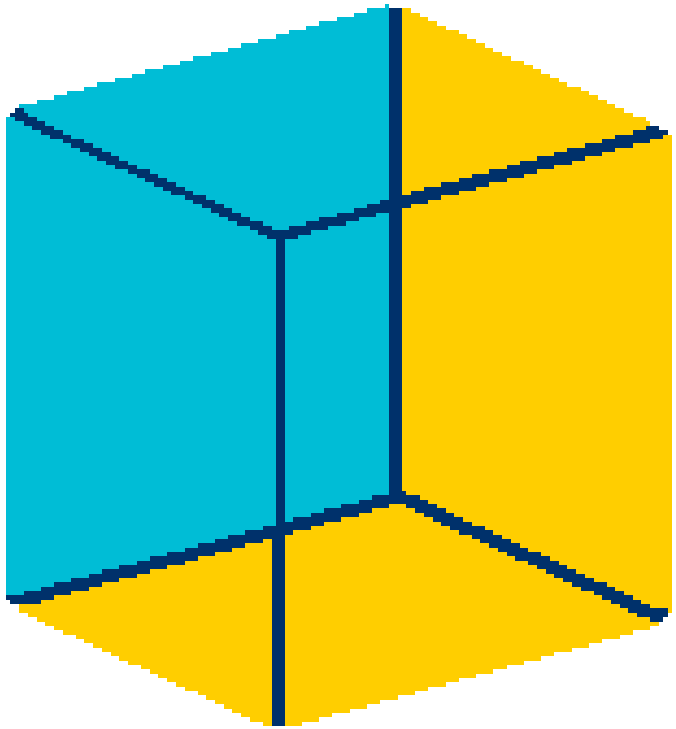


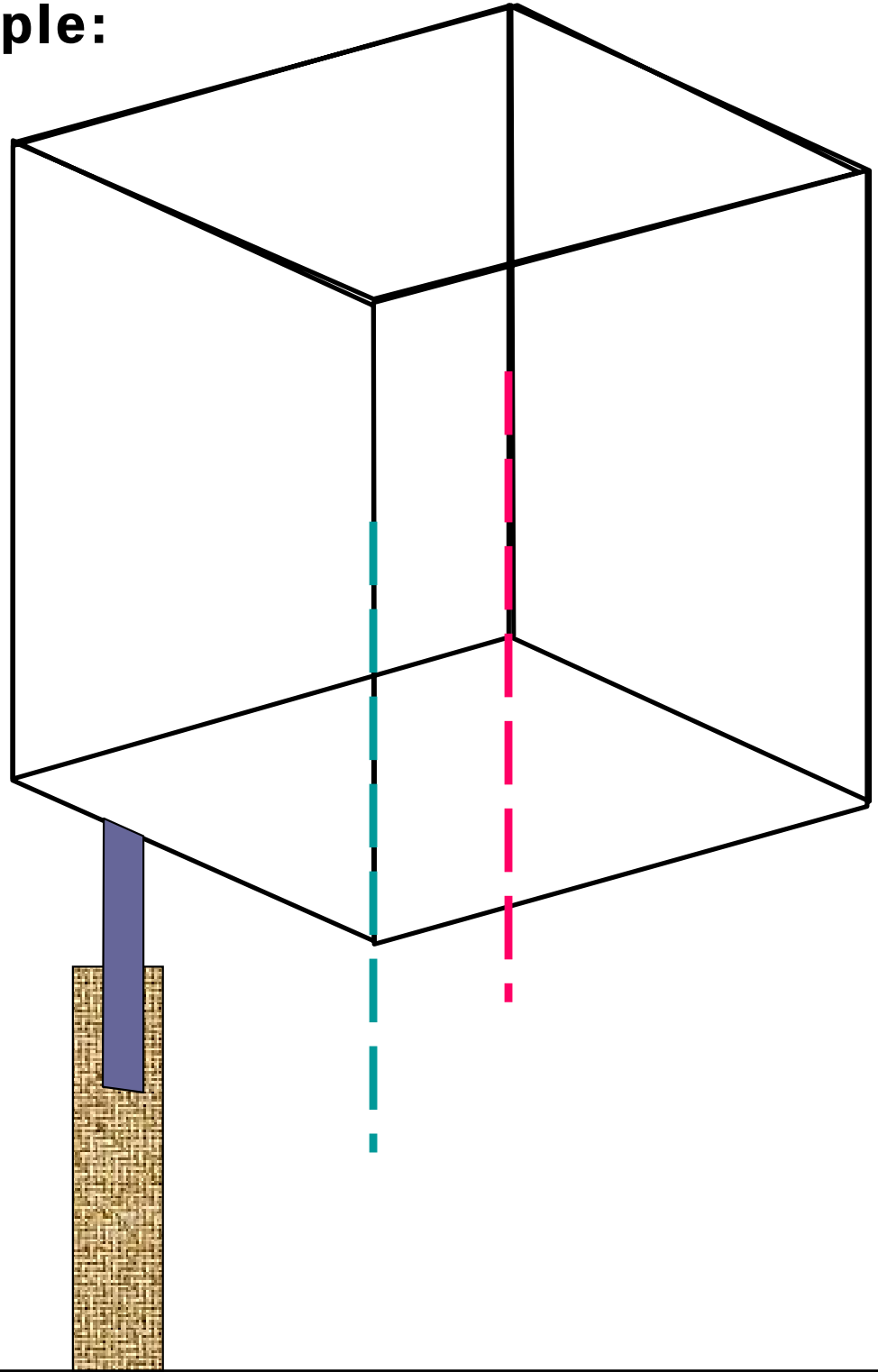
Necker's reversible cube example:

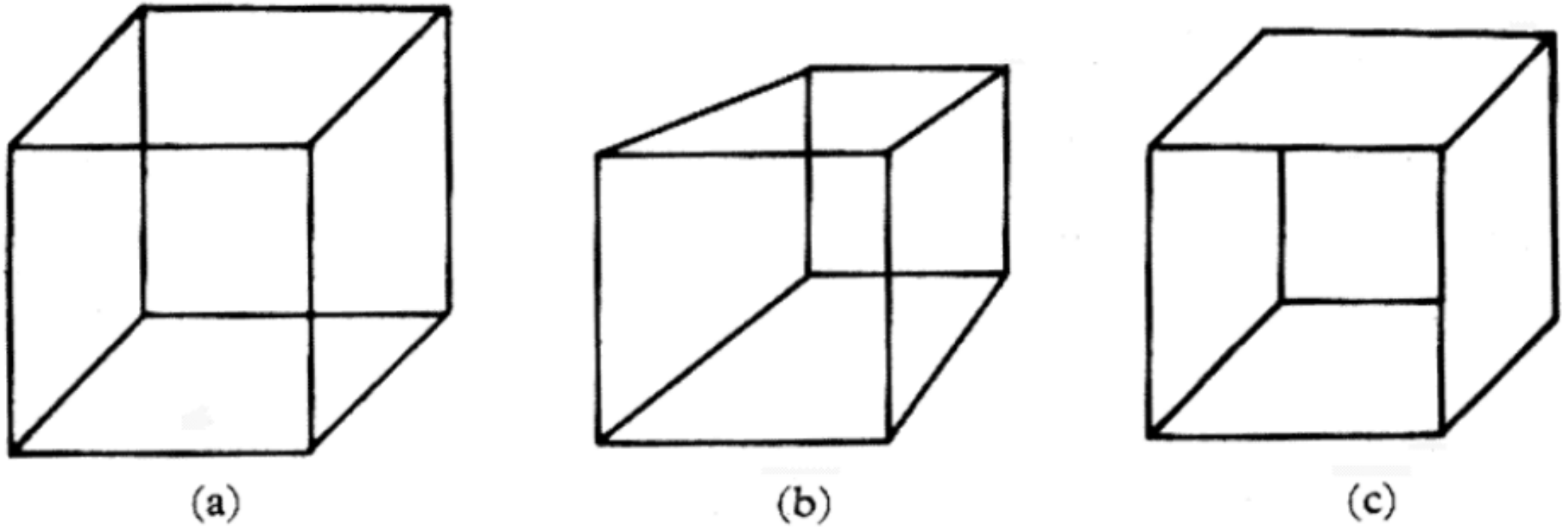
It cannot hang in air or space.

So give it a support.

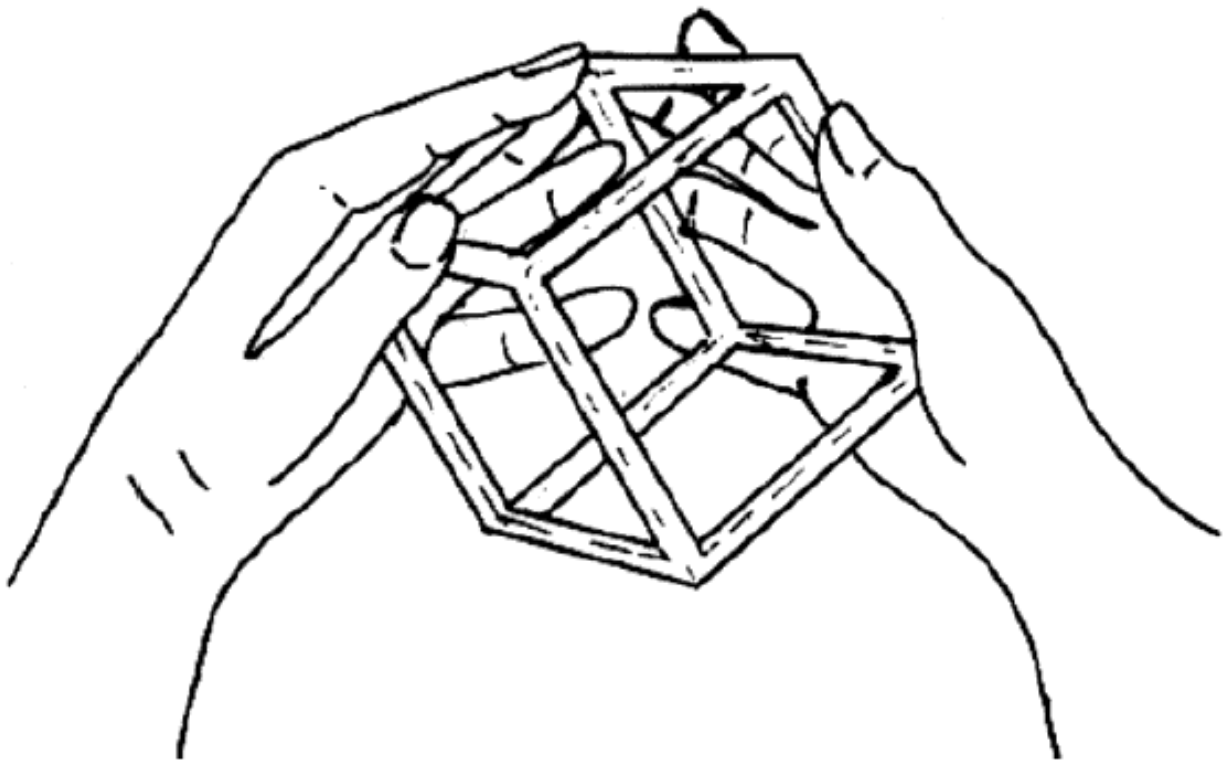
Feel anything different ?



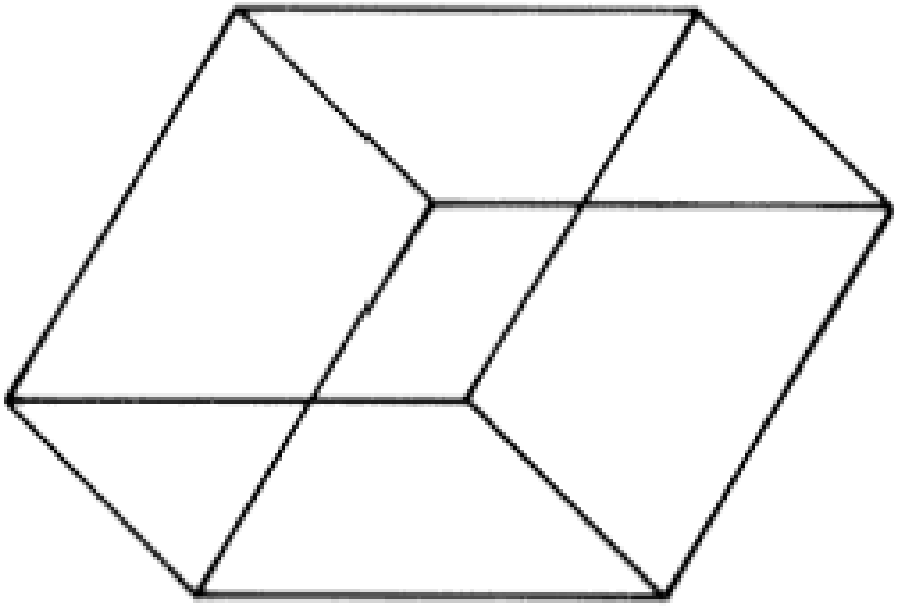


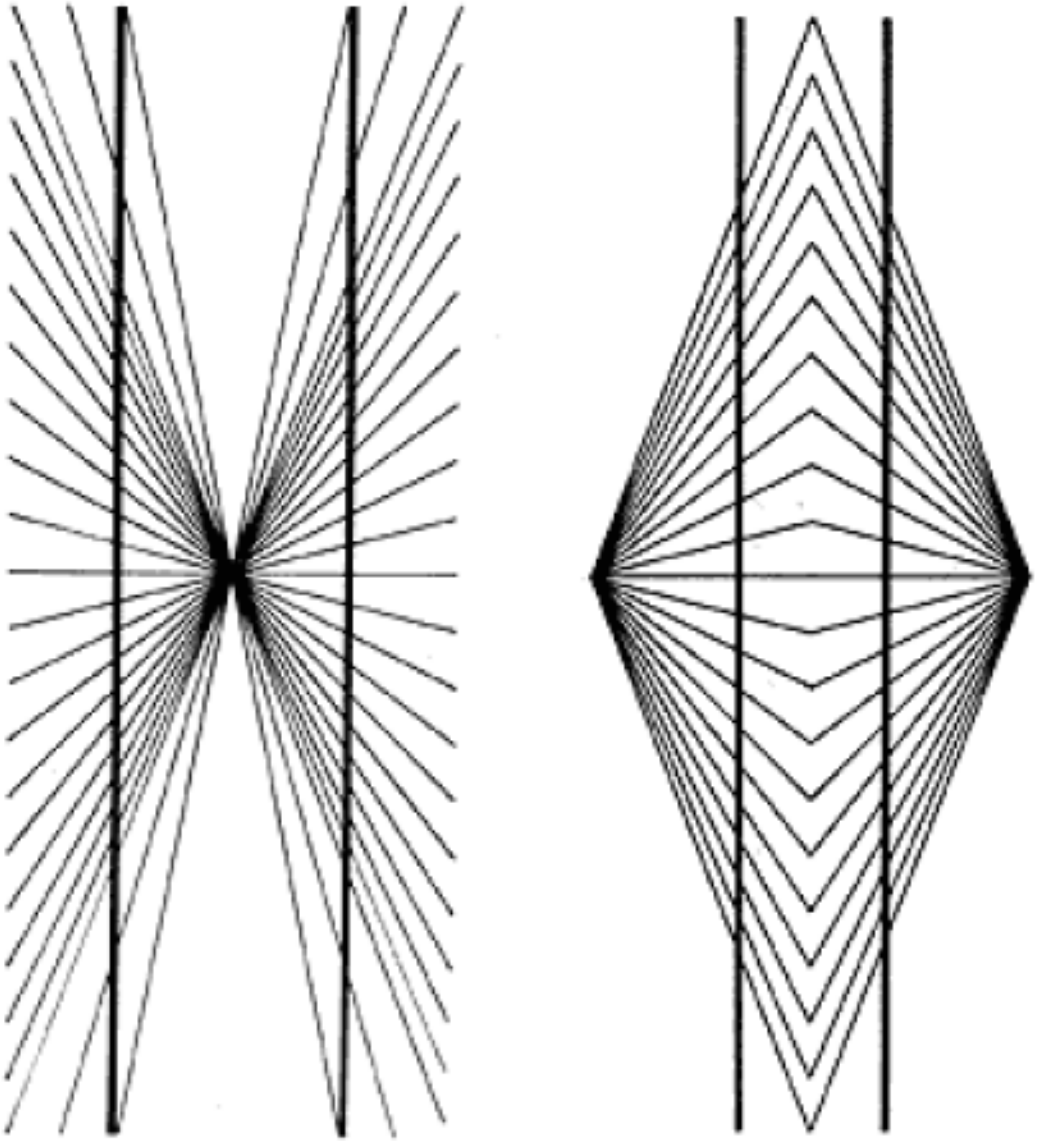


A simple Necker cube (a); and two examples: (b) and (c) of added depth cues (perspective and masking) generally tending to stabilise orientation.

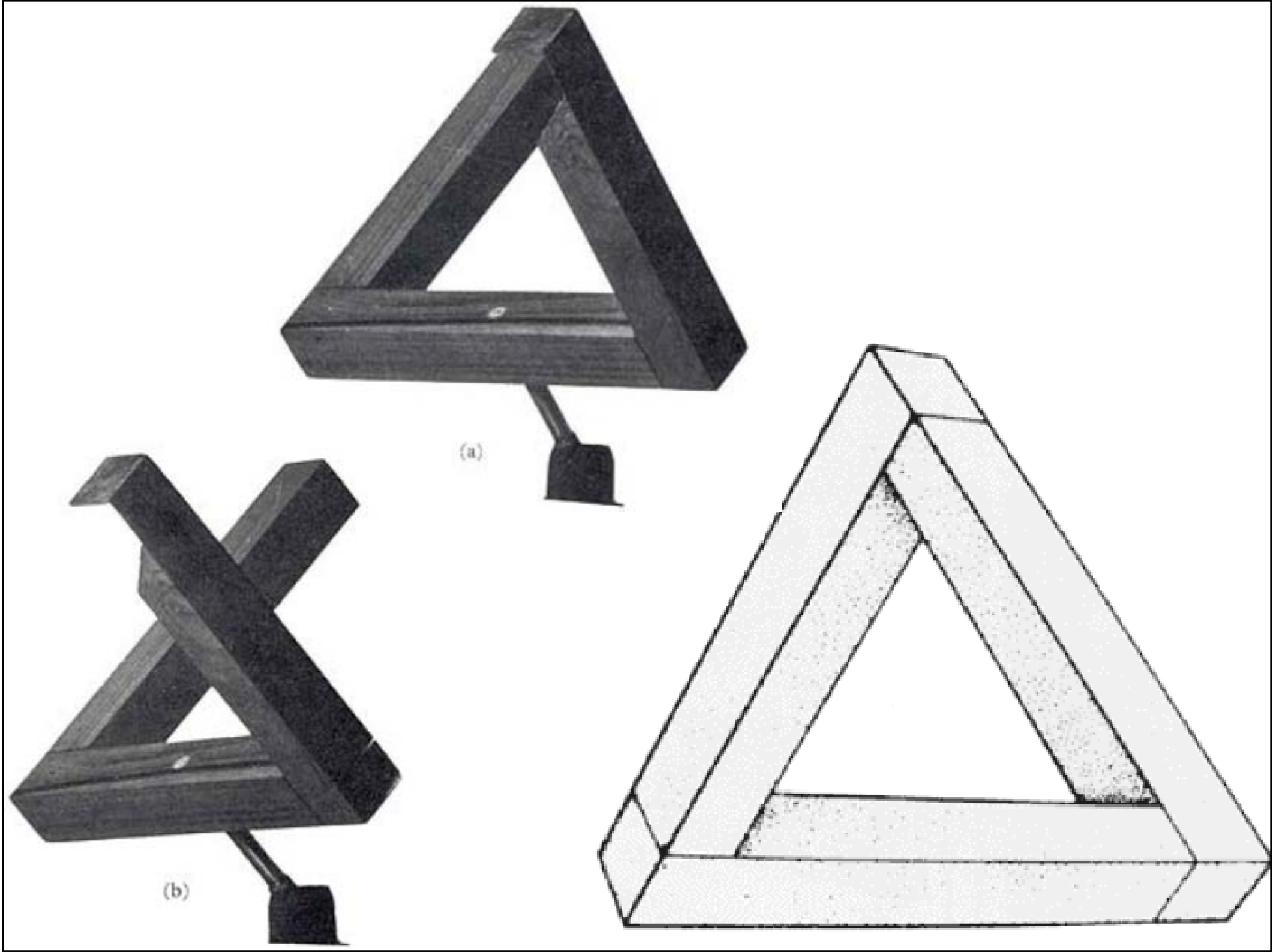


Necker rhomboid. This is the original form, presented by L. A. Necker in 1832.





Hering illusion


(a)
(b)

Significant points in Gregory's work:

- **Appropriate behavior for non-sensed object characteristics (table with 3 legs)**

- **Zero lag tracking time in case or regular and predictable path.**

- **Ambiguous illusions**

- **Extract familiar objects from a background clutter – even in speech (party in a large hall)**

- **Highly unlikely objects mistaken as likely objects**

- **Paradoxical phenomena (impossible figures: staircase chain, unsupported columns).**

- **Not always experience dictates perception**

Gregory provided a lot of evidences using experimentations to support his hypothesis.

Gibson's work on direct perception

and ecological optics.

Gibson: Pilots use information, such as, forms of patterns of movements of flowing textures (clouds, ground points etc.)

The importance of *movement* in perception and usefulness in considering perception in real-life conditions, as opposed to simple laboratory experiments.

What is indirect perception then ?

e.g. closure in patterns, illusory contours,

Go beyond simple stimulus and sensory evidence, to add details (as in IDWT?).

Indirect: Stages of model-based processing, with information flowing between them. Also involvement of memory in interpretation or perception ?

Is it really so ? Some theorists think not.

Gibson (cognitive theory) attempted to explain how we come to know of the world. Environment and inhabitants are integrated.

First let us understand “ecological optics”, and then understand more about the term “direct”.

LIGHT and Environment

An object does not hang in isolation. It is present in an environment. This is real-life, and HVS solves this problem, not like in CVS where we expect only a face or ball.

So in real-life, the retina is bathed in a sea of light. The visual world comprises of surfaces under illumination, some direct, some indirect rays (CG – ILLUM-SHADING).

However light carries information (known source) of the environment through which it has traveled, and from what it has been reflected. We do not live in a mad universe, where light moves erratically and is non-informative.

Light is structured – function may be complex. As the head and eyes sweeps through various solid angles during normal perceiving, we rely less on a momentary snap shot to extract information (ala, our poor CVS). Hence less of indirect information is also used.

Ecological optics

Take the case of size constancy, as object is drawn away from the viewer (perspective case). How do we still perceive the same object, although the stimulus changes.

Impose real-world environment factors:
Well, since the object should have a background (generally textured in a real-world case, or even otherwise any) the part of the background which it covers changes with time.

Any CVS which does this? Most systems do a segmentation to eliminate the background. We observe change in the fore-ground, but do not calculate the same for the background. In most other cases, we do not have or worry of the background.

Why this is so – overdose of laboratory experiments which overlook nature and environment.

In addition, Gibson talks about invariants in perception:

distal vs. proximal cues, transformational , structural, *affordances* (what is possible, what not).

Gibson suggests:
"Nervous system operated in circular loops. Information is never conveyed, but extracted by picking up invariants over time……
A perceptual system does not respond to stimuli but extracts invariants."

Gibson talks of resonance (like in a Radio) in perception:
Information pickup is the job of perception. In radio we do manual tuning, here it automatic, as well as simultaneously and coherently occurring over all bands (analogy with radio).

Also read about "realism".

So what is "direct perception" ?

What is a direct process:
say, implementation of Multiplication operator, *

Input: A, B; Output: C = A*B. But "*" function is not working.

Use a table. Row value and column value combination gives result. Direct approach, not decomposable.

Almost impossible with large real-life numbers. Decompose the numbers into units, tens, hundreds etc., and then superimpose. This is not direct (indirect). Rules are applied in stages. *Brain does not easily adapt to MATHS*.

Difficult to extrapolate in perception. This work was criticized by Marr – detection of physical invariants is an information-processing problem, and Gibson vastly underrated the difficulty in such detection.

Read about Norman's work

Importance of optical flow and movement during perception.

Difference between motor and sensory aspects of seeing.

Computational Theory of David Marr (MIT, 1970s)

Influenced by initial concepts in AI:

- **Information Theory (Shannon)**
- **Cybernetics (control systems with feedback)**
- **Digital Computer**

    - **Random Dot stereograms**
    - **Spatial frequency channels in receptive field**
    - **Early computations**

No concepts and use of neurons

Information (reduction of uncertainty)
vs. Information-processing

Representation vs. Description
- Use of symbols for representation (maps)

Feature vs. data

Almost the first computer for number crunching:

- MECHANICAl CALCULATOR or

CASH REGISTER

Marr's theory

Three levels of understanding information systems:

(i) Computational theory
(ii) The Algorithm
(iii) Hardware implementation.

Computational theory: What is the goal of the computation ? Why is it appropriate ? What is the strategy to carry out the computation ?

e.g. Add several sub-totals to get a SUM. Marr says, use a cash register, but the theory is simple rules of mathematics.

So we describe a machine, but do not bother how it operates. We also define the constraints on it, which allow processes within the machine to be defined.

The Algorithm:

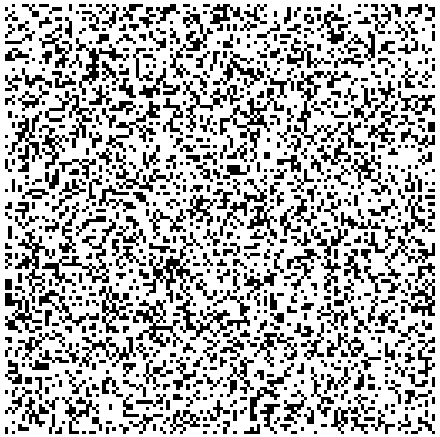
Mainly data representation - say for performing the SUM, BCD-Binary converter, formula used for binary ADDER, Binary-BCD converter.

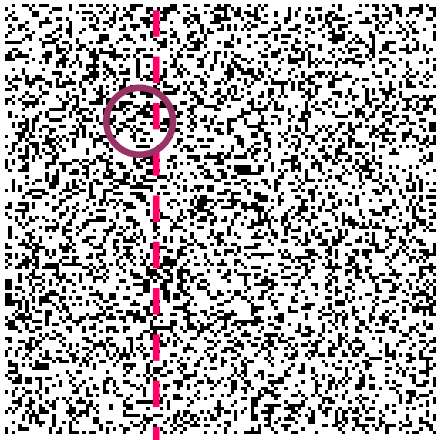
Knowledge of the representation dicates the algorithm to be used.

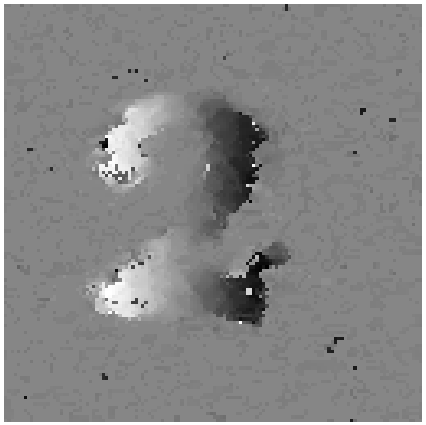
Hardware implementation:

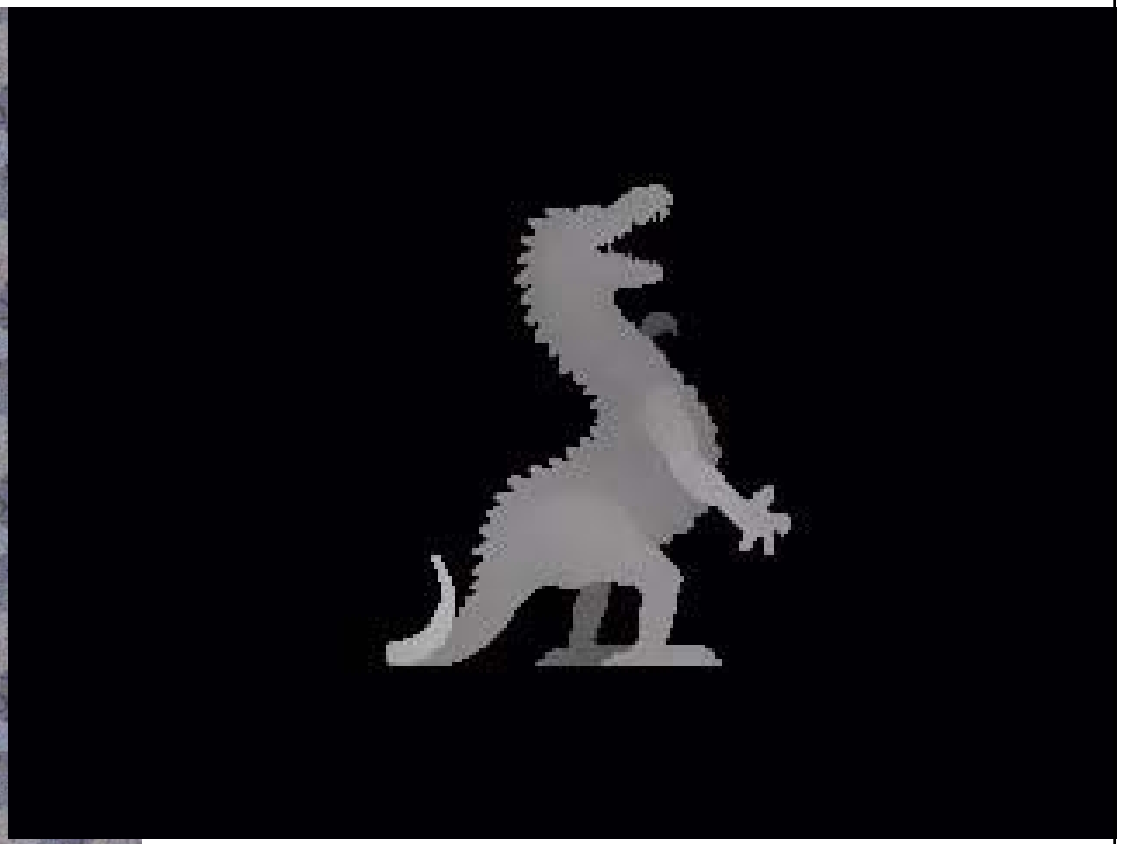
How does the machine work ? Series of electronic / mechanical functions. We know what is HARDWARE – this portion is for Non-Engg. (Arts, psychology, biologists, may be even purse science) students.

Random Dot Stereograms



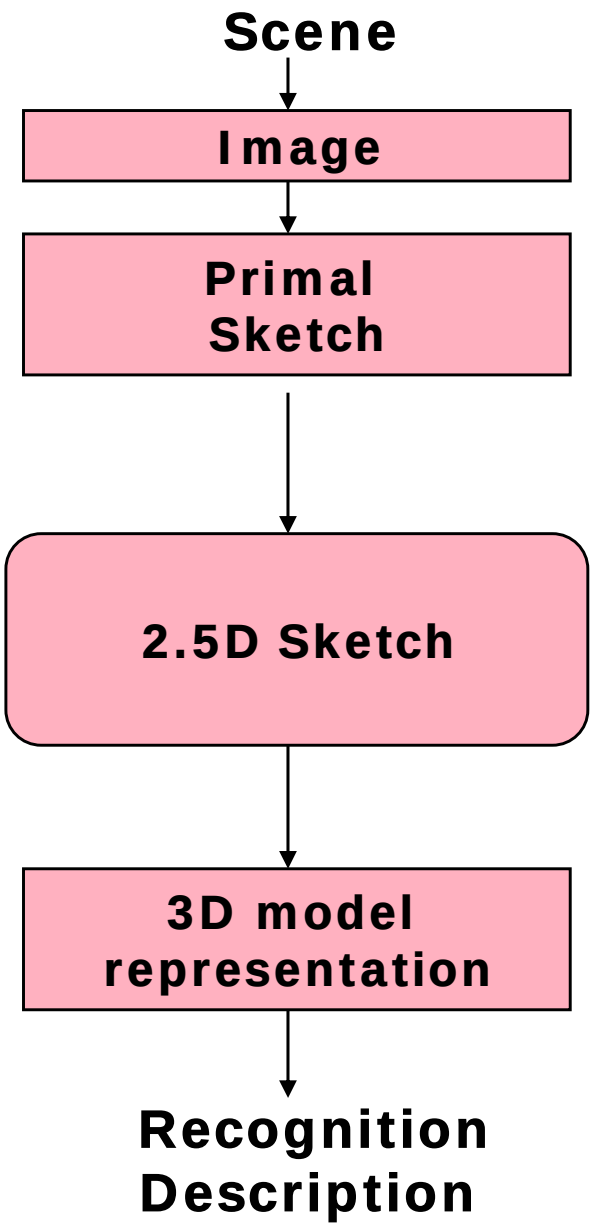






Now look at the corner of the room and visualize.

Marr's Object Recognition Framework

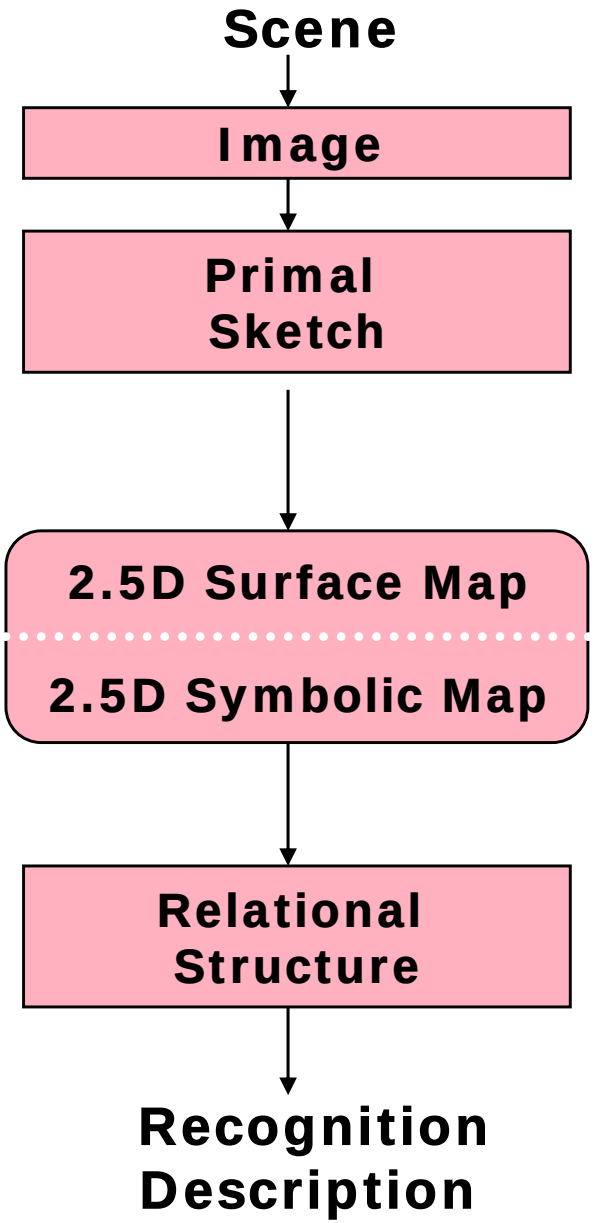


Primal Sketch : Extract significant intensity changes in an image, often in form of edges and blobs

$2^{1/2}$-D sketch – compute depth and orientation of visible parts of surfaces. "Picture" of the world w.r.t the viewer is beginning to emerge

3D model – Relate shape and orientation to 3D objects organized in an object-centered framework. Perceiver has attained a model of the external world. Position of orientation of the retina is irrelevant.

Marr's Object Recognition Framework



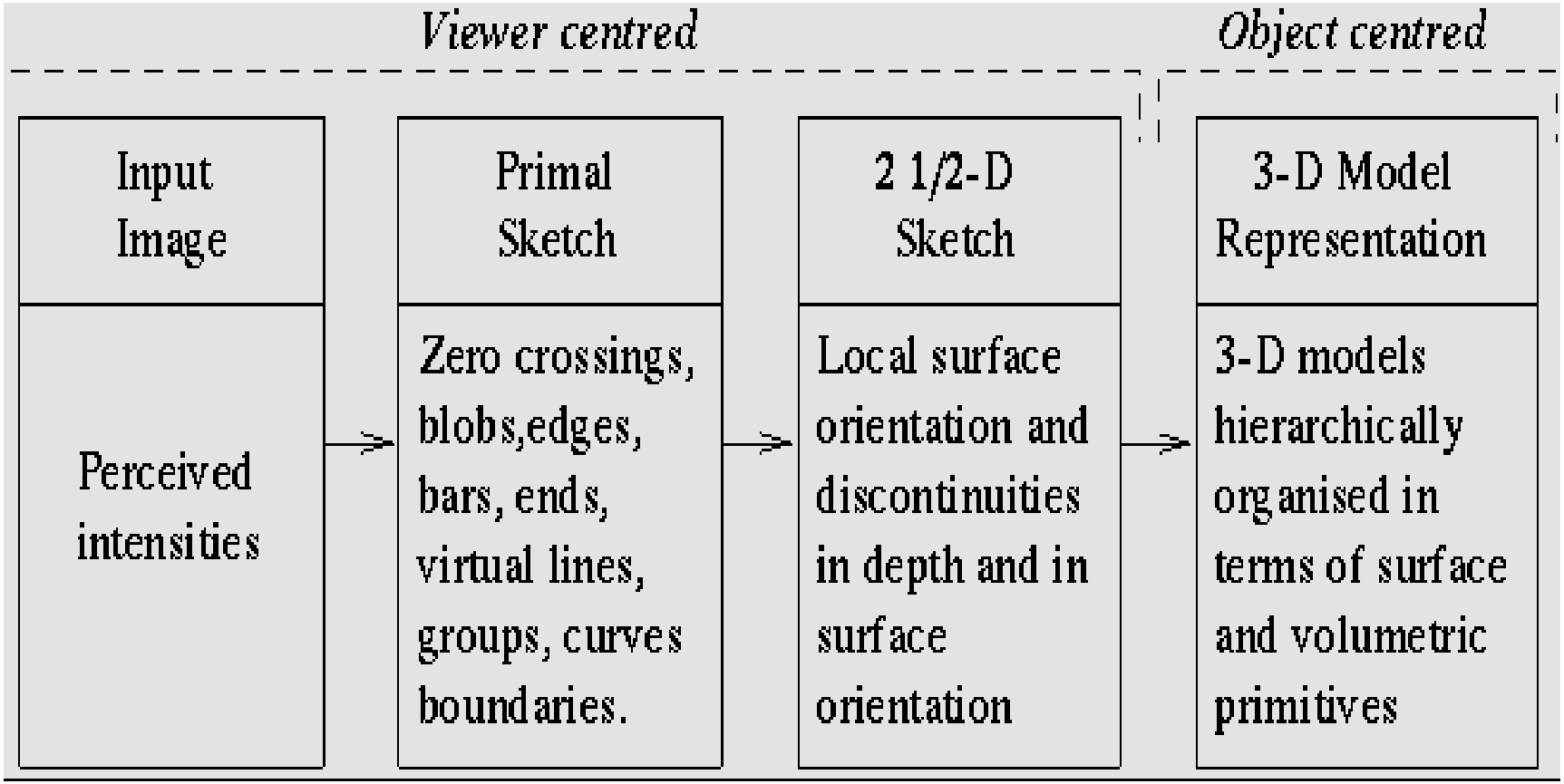
Primal Sketch : Extract significant intensity changes in an image, often in form of edges and blobs

The 2.5-D Sketch :

- **Transform image to a depth map, by calculating relative distance to features in the scene and infer surface orientations**
- **Use Shape form 'X' techniques**
- **Midway between 3D and Primal Sketch**

The Relational Structure : Transformation to an object-centered coordinate system and identification of object

Marr 1982 - Marr's representational framework



The detection of intensity changes, the representation and analysis of local geometric structures and the detection of illumination effects take place in the process of generation of the primal sketch.

One important principle of the primal sketch is that independent spatial organizations of the viewed intensities in a scene reflects the structure of the visible surfaces.

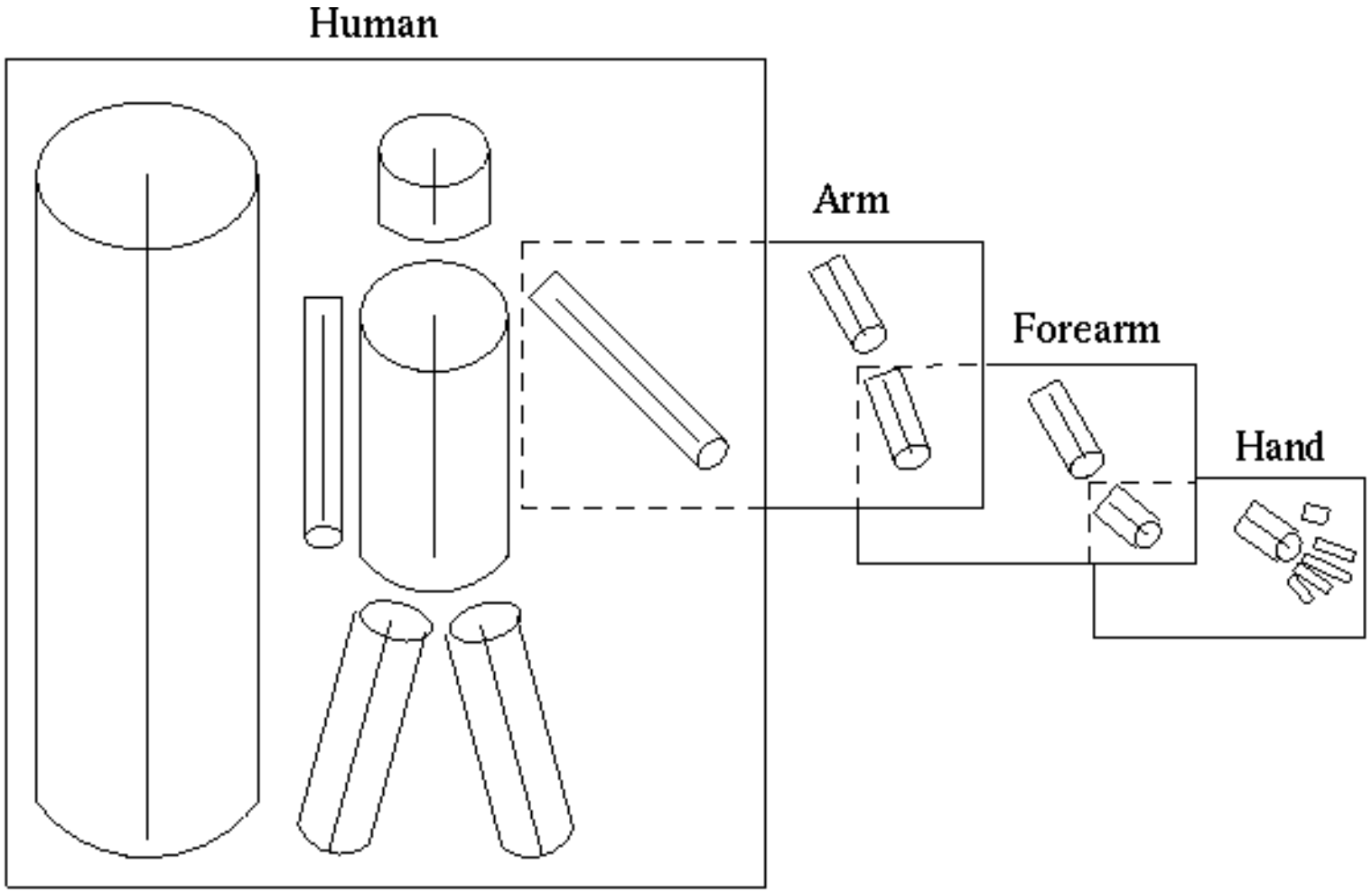
Marr proposed to capture these organizations by using a set of “place tokens'', or low level features, which correspond to oriented *edges*, *bars*, *ends* and *blobs*, which were represented by a 5-tuple: (*type*, *position*, *orientation*, *scale*, *contrast*).

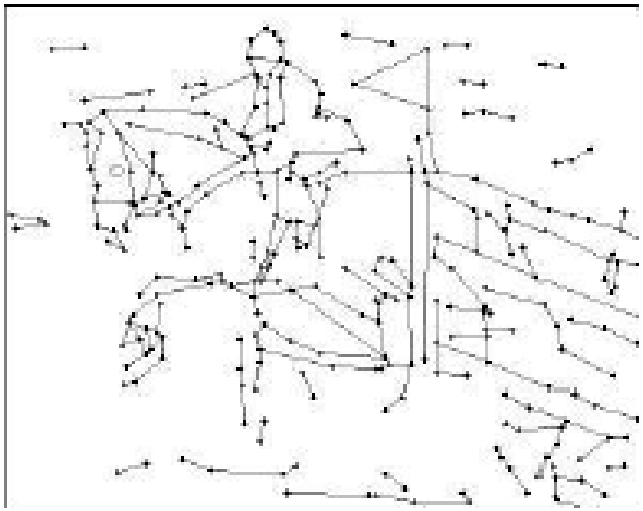
The 2.5 -D sketch is intended to represent the orientation and depth of the visible surfaces as well as discontinuities.

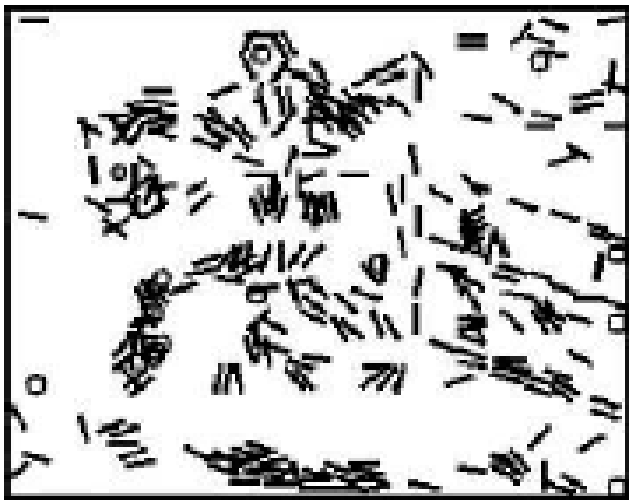
3-D model representation:

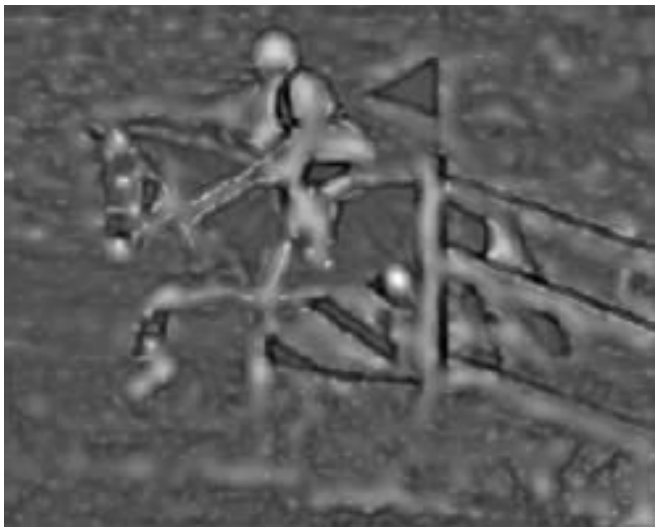
This representation is intended to describe shapes and their organization using a modular and hierarchical organization of volumetric and surface primitives (see human structure).

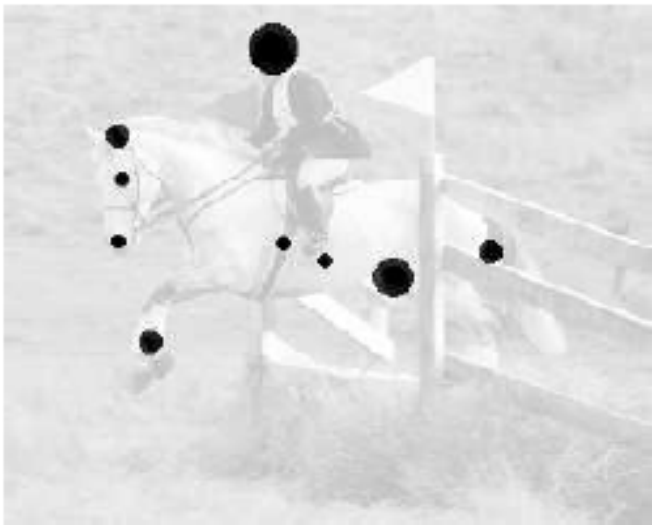
The recognition process uses a *catalogue of 3-D models* which is a collection of stored 3-D model descriptions and various indices into the collection that allow the association of a new description with the appropriate one in the collection.

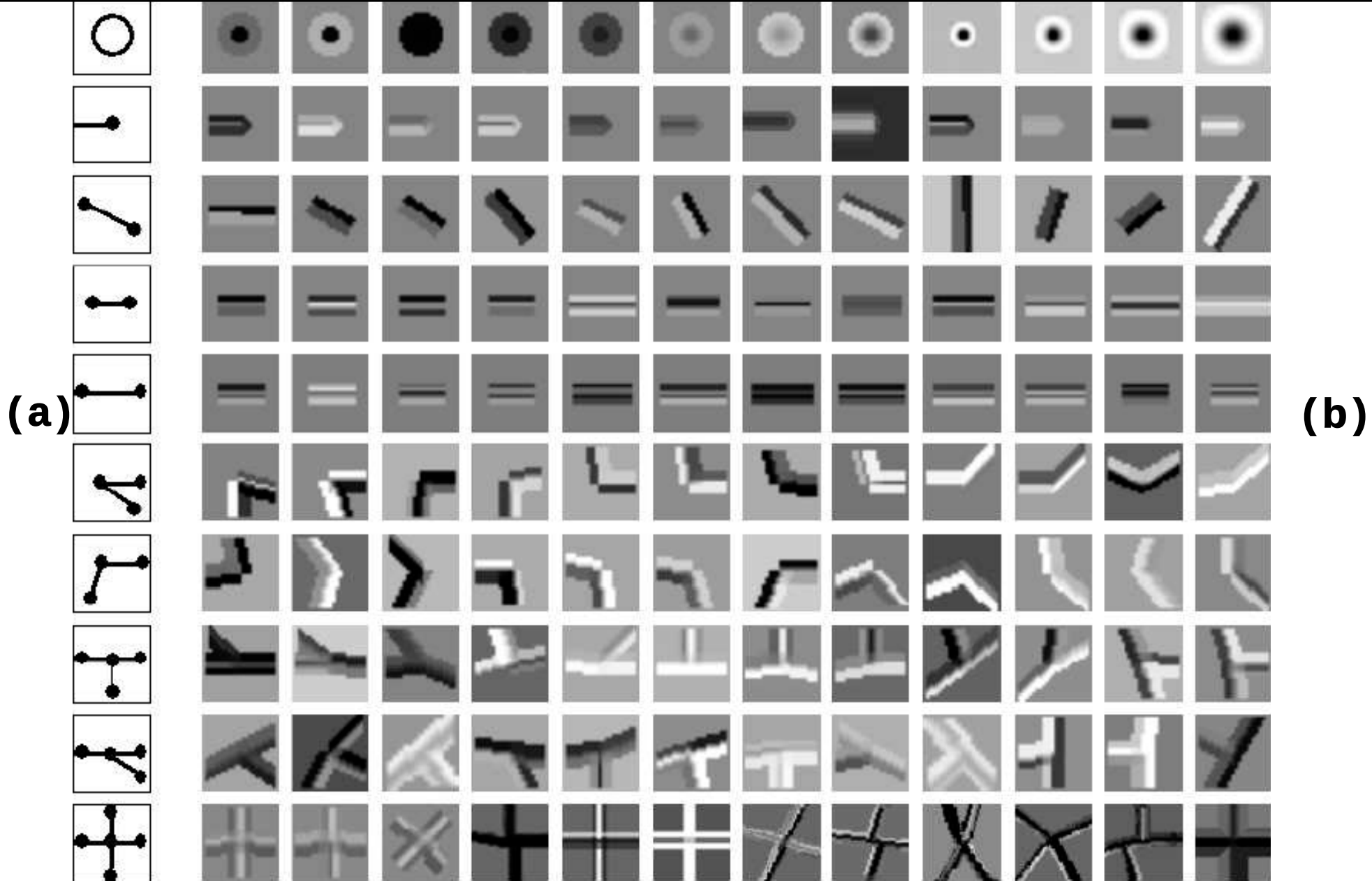

Human
Arm
Forearm
Hand



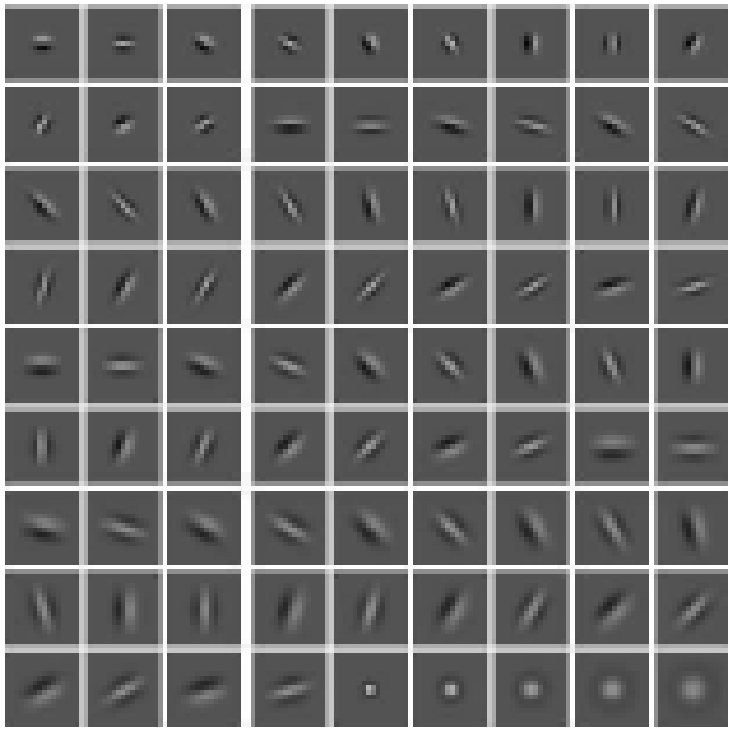


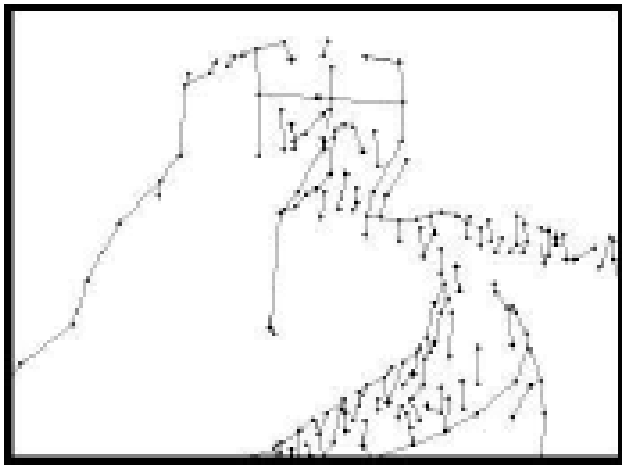


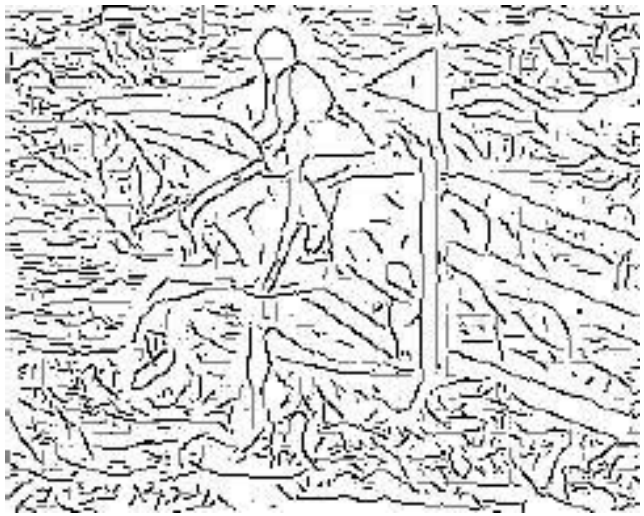




Samples from the visual primitive dictionary, consisting of eight types: blobs, end points, edges, ridges, multi-ridges, corners, junctions and crosses of different degrees. (a) The landmarks on the patches for topological and geometric attributes. (b) The photometric representation of the patches.







Accepted theory of LOG and DOG operators.

Rejected Primal Sketch to 3-D representation