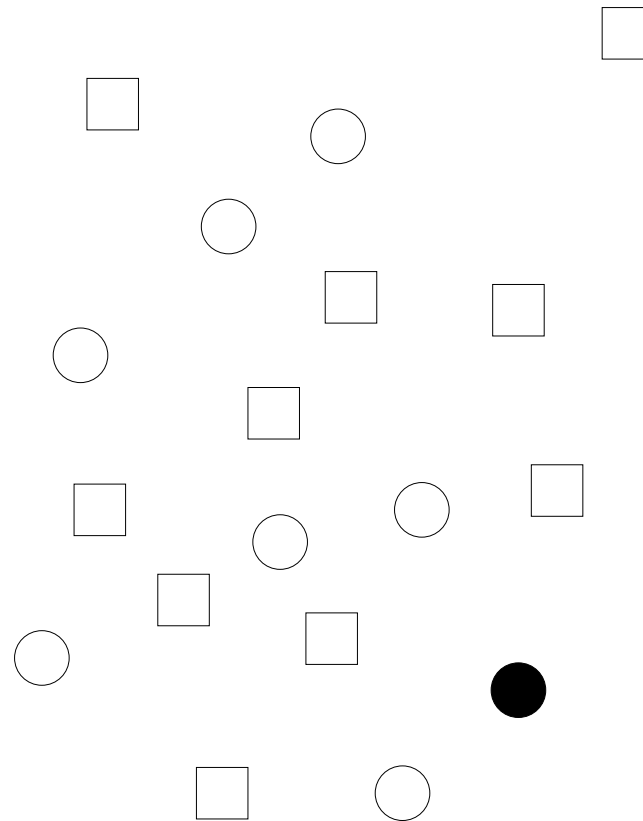


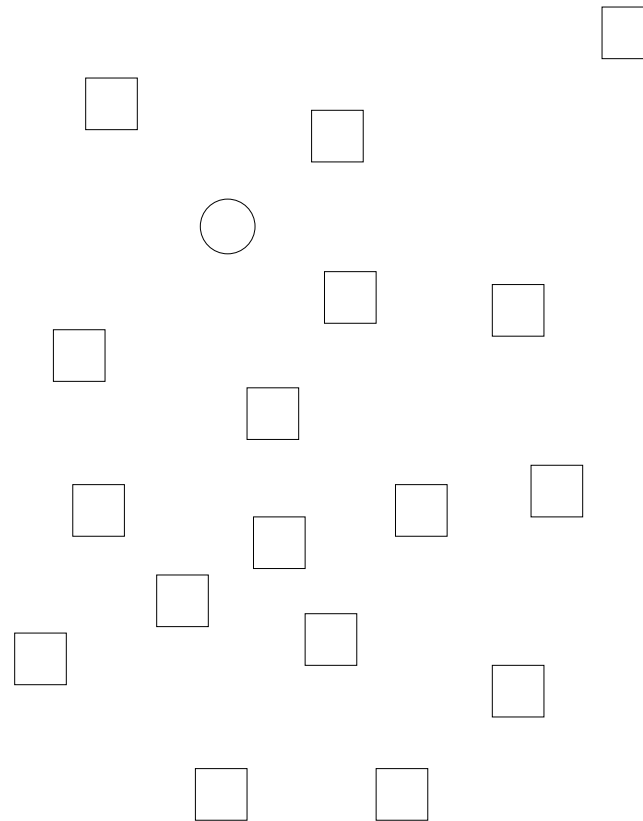
Visual Object Recognition and Pattern Recognition

Reinhard Moratz

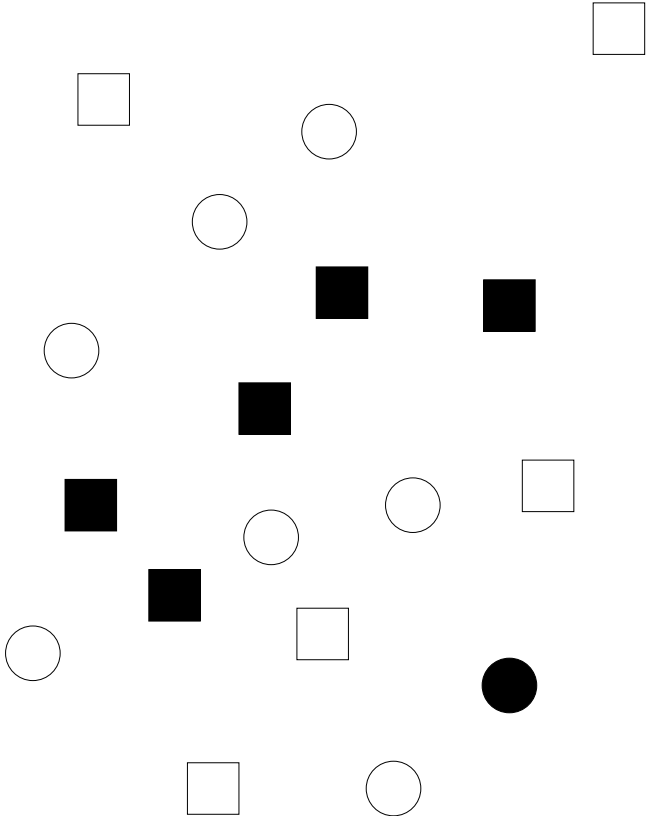
Treisman Experiments 1: Simple search for a black object is parallel



Treisman Experiments 2: Simple search for a circle object is parallel



Treisman Experiments 3:
Conjunctive search for a black circle is serial

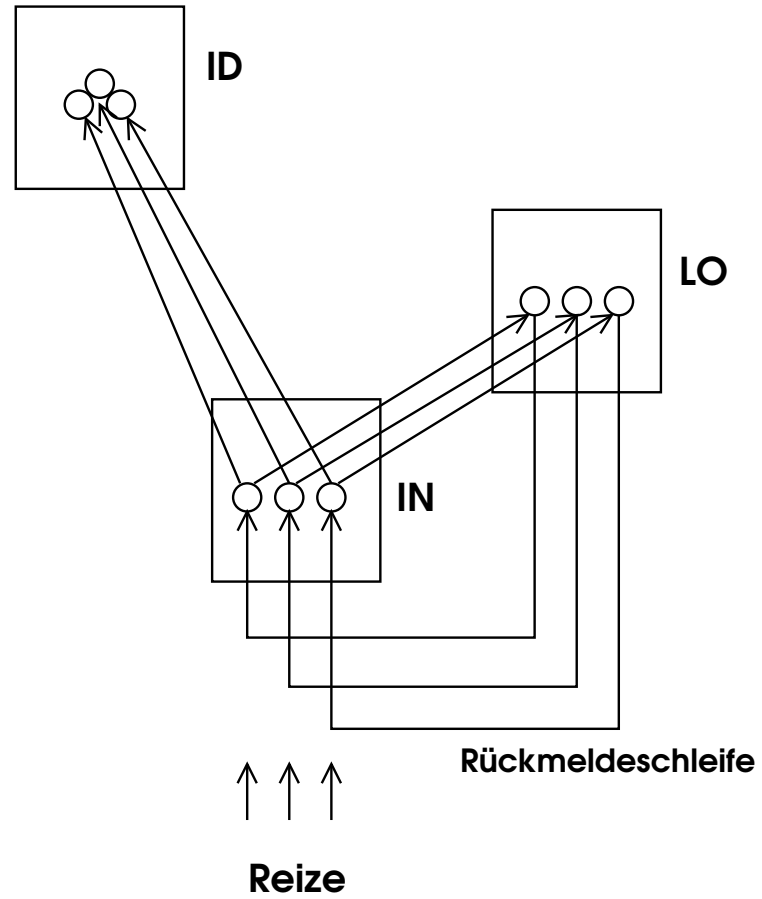


Visual attention

distinction into

- ↳ simple detection processes that run in parallel, and
- ↳ processes that need visual attention which is given sequentially to individual objects or spatial regions of interest of the scene.

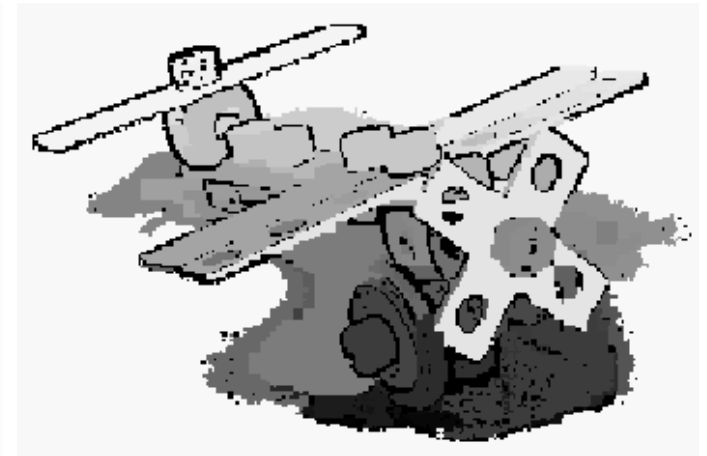
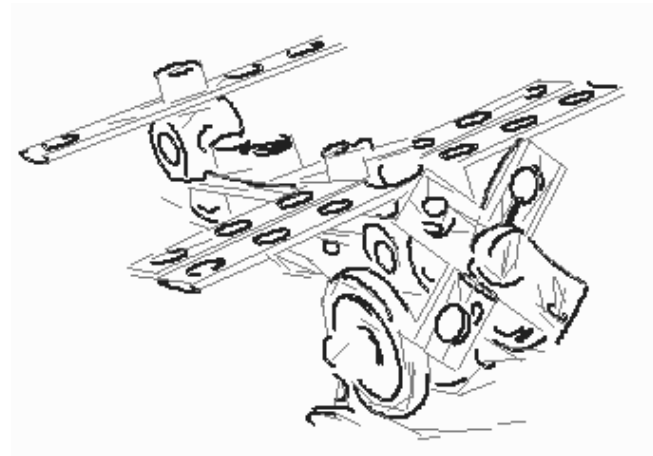
Model of visual attention from van der Heijden



Three levels of computer vision

- ➔ Low level image processing uses transformations which work on matrix structures which have coordinates corresponding to the sensor matrix. Results are mainly subsets of pixels which are put together into segments.
- ➔ Medium level image processing groups these segments into features. The resulting structures are numerical but do not reference to the image geometry any longer (e.g. feature vectors).
- ➔ High level image processing (image understanding) uses symbol-oriented and graph-based representations. This level produces qualitative descriptions of the objects and interprets the whole scene. It is also the interface to general world knowledge.

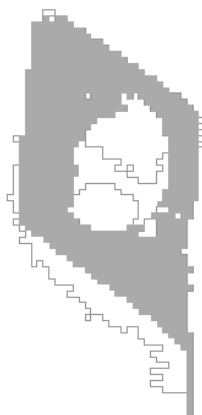
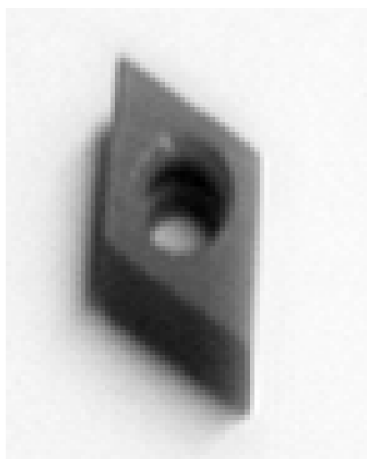
Segmentation: scene



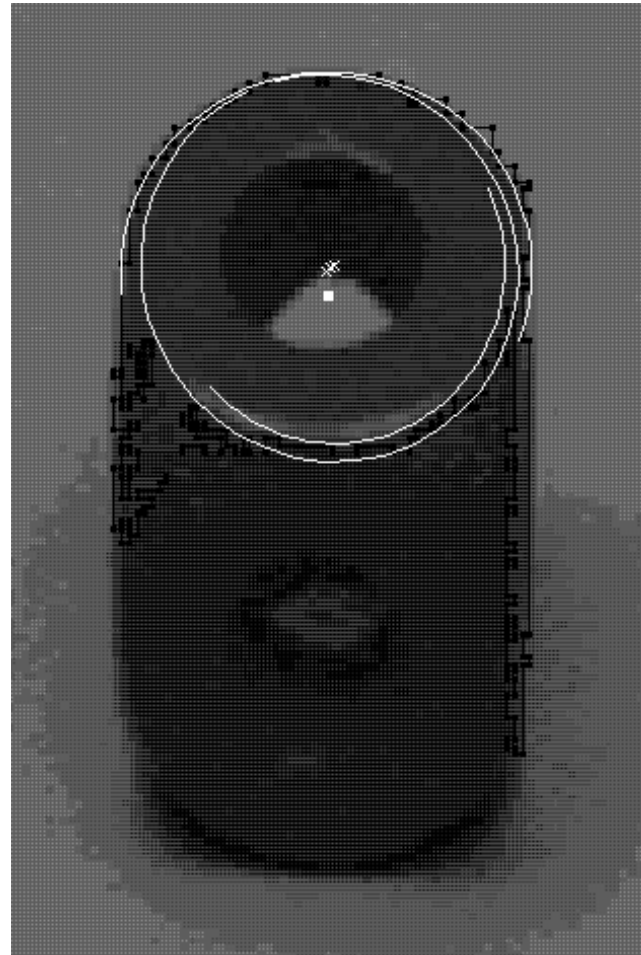
Segmentation errors

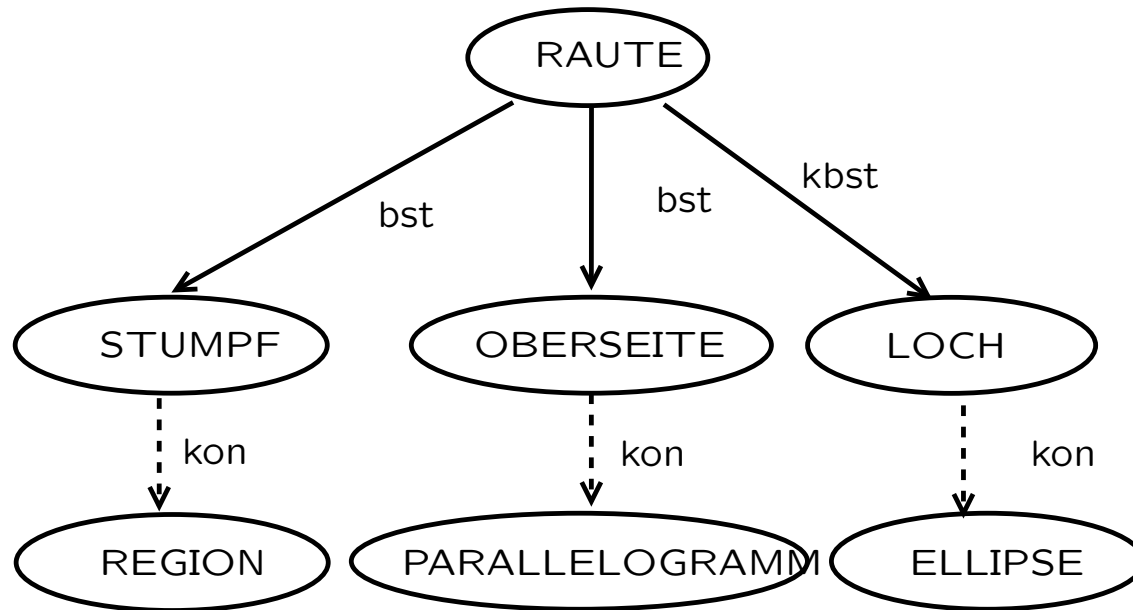
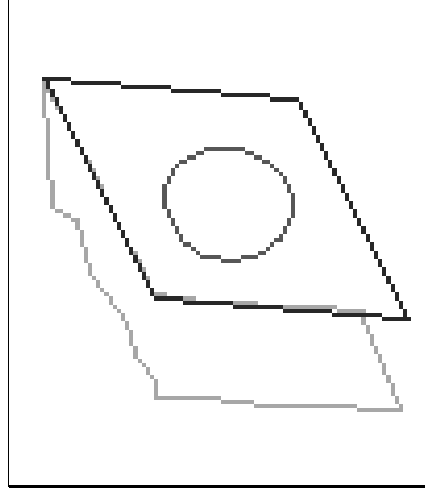
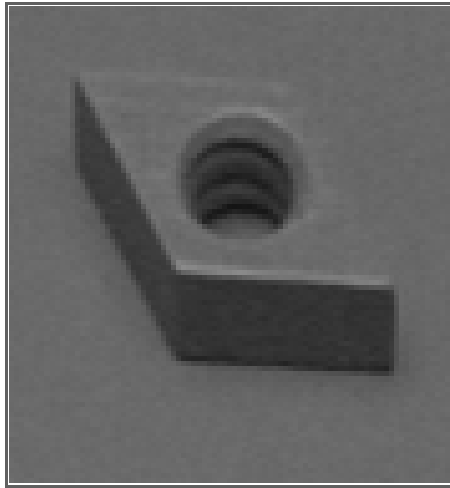
- ↳ Edges are often interrupted and can not be traced easily, but they detect object borders with high precision.
- ↳ Regions are closed by definition. Therefore they often can already be classified as meaningful subparts. The main drawback of regions are their imprecise borders. And in many cases different objects “glue together” to one region.

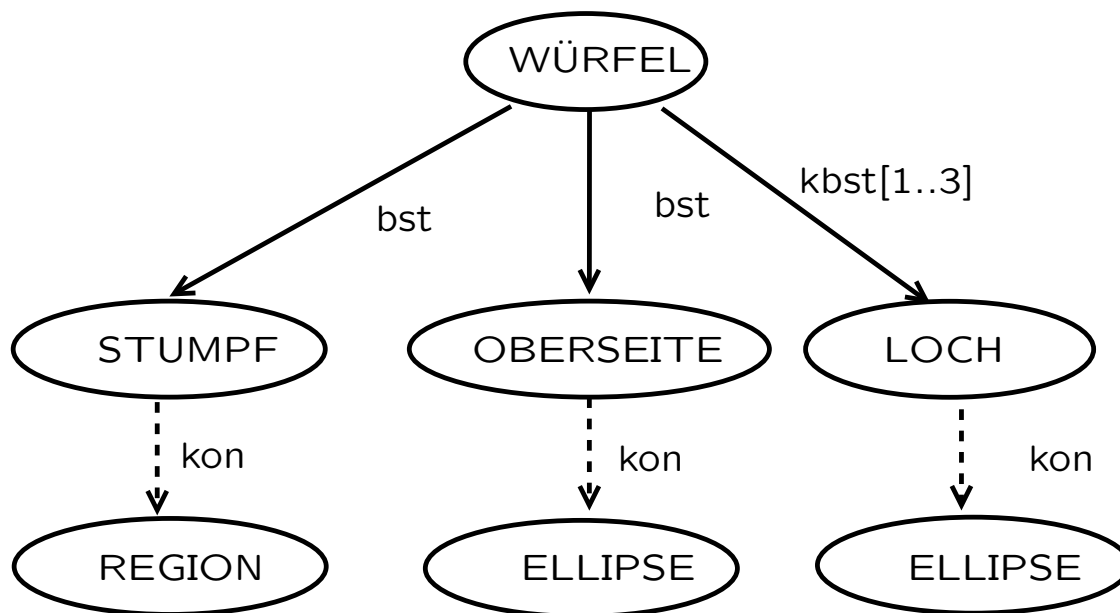
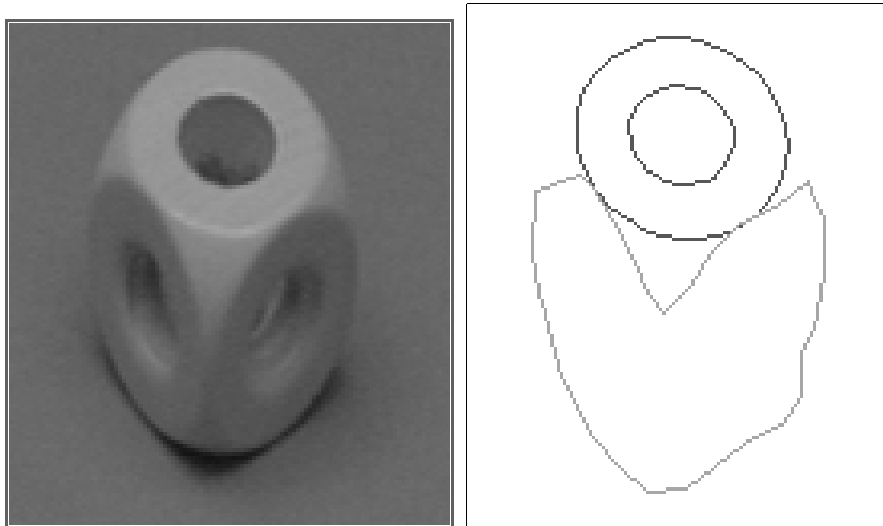
Segmentation: detail



Grouping of approximized edge and region segments







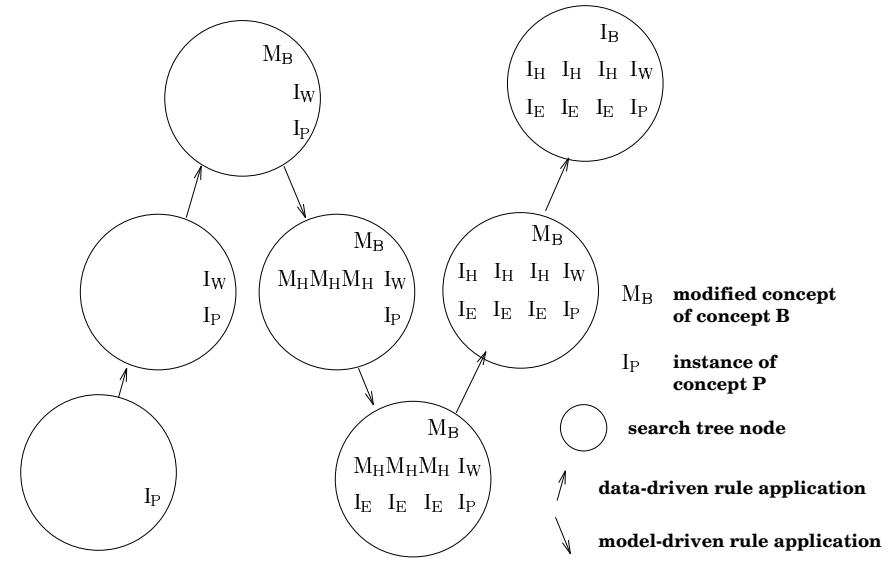
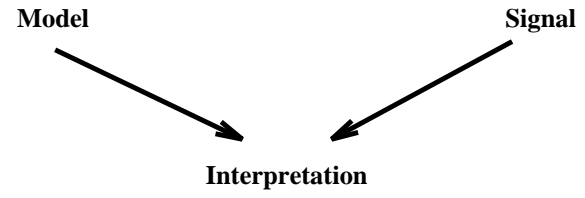
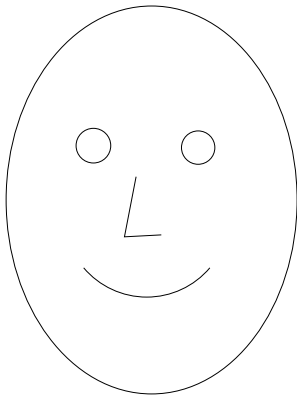
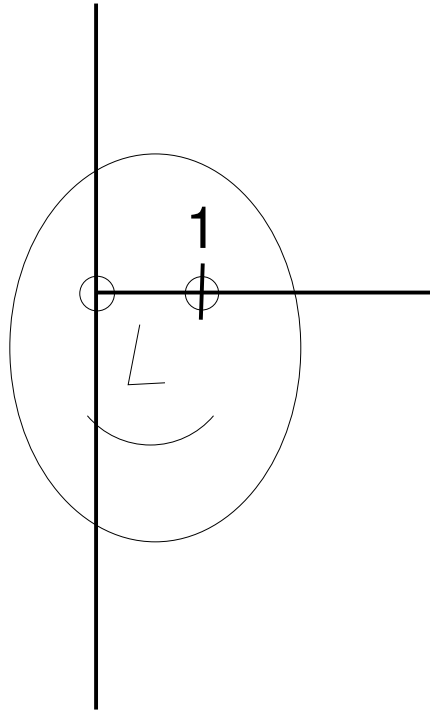


Figure 2: Signal Interpretation using a semantic Network, for the complete description of the network see figure 1.

Image



Normalisation



Feature Vector



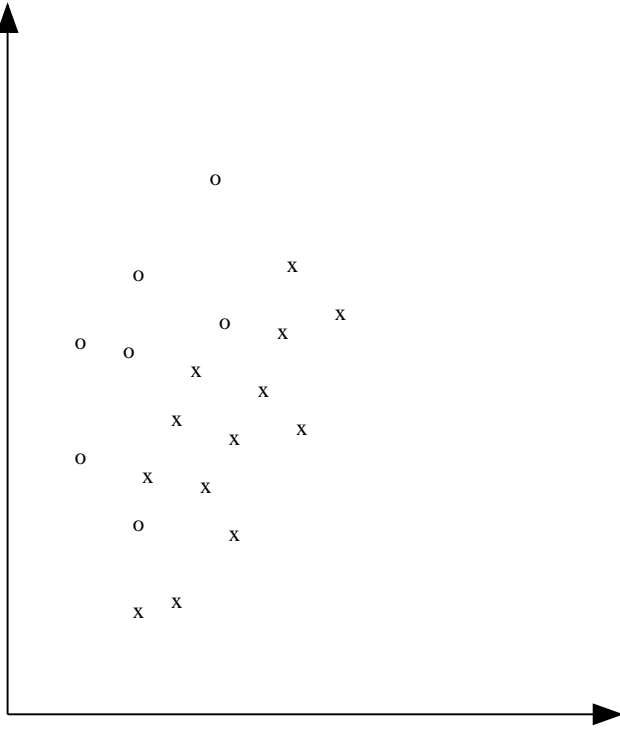
X-Coord. Nose

Y-Coord. Nose

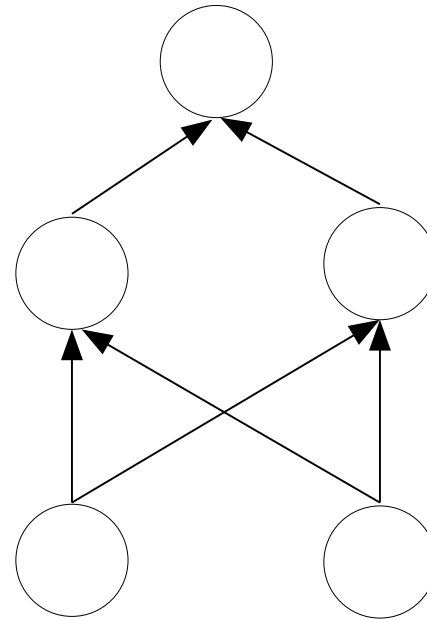
X-Coord. Mouth

Y-Coord. Mouth

Feature Space



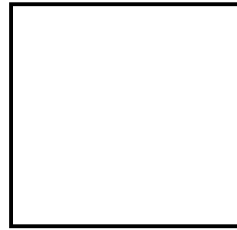
Artificial Neural Network



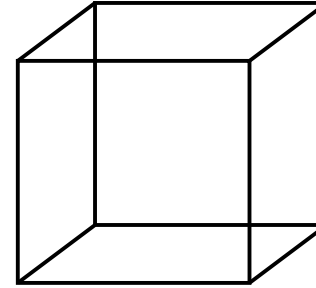
High-Dimensional Feature Spaces



2



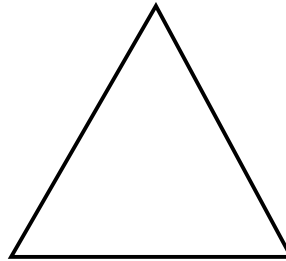
4



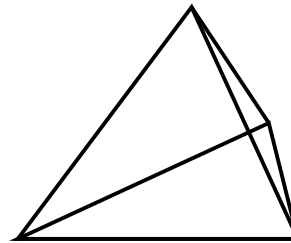
8



2



3



4