



Figure I.3: The general three-level schema of visual information interpretation and representation.

The system modules and information structures shown in the schema of Fig. I.3 can be briefly characterized as follows.

External world contains physical objects whose pictures can be “seen” by the system, as well as physical objects affected by the system, like those “painted” by the system. It also may contain other similar systems, with which the system may communicate, exchanging information through the appropriate objects in the world, like pictures.

Picture input is a device capable of taking pictures of the external world and transferring them to the inside of the system.

Picture output is a device capable of producing pictures, generated within the system, as objects in the external world.

Encoded picture is an internal form of the picture as obtained by appropriate encoding of the external world picture, or as prepared by the system for output to the world. It is generally characterized by the fact that primitive elements of the picture correspond to elementary pictorial stimuli, like brightness and colour of otherwise undifferentiated picture points.

Image description is an internal partial interpretation of the contents of the encoded picture, or the structural description of visualized knowledge, in terms of some larger elements (like edges, lines, regions, and characteristic local configurations of them), their more complex attributes (like line direction, length, thickness, figure area and shape), and relations between them (like adjacency, relative size, position), cf. Section II.2.

Knowledge is the most abstract form of information, gathered and stored by the system as a result of interpretation and processing of external stimuli. It can be used to affect the external world in various ways, among others by producing pictures representing some chunks of the knowledge. At this level all data gathered by various sense organs of the system are ultimately integrated, although various links and influences between them can occur at lower levels as well.

The processes linking the modules and transforming the appropriate information structures comprise:

Picture coding produces an internal form of the raw external picture from the data provided by the picture input device.

Picture processing transforms the encoded picture in various ways, mostly in order to enhance it by emphasizing some its aspects or filter out the data considered to be noise. This may be an end in itself, or it may be a preparatory step to facilitate further analysis of the picture, or generation of its external form. Another popular kind of processing aims at size reduction of the picture encoding for its more effective storage or transmission.

Image analysis transforms the encoded picture into a higher-level description in terms of more meaningful parts, their attributes, and relations between them. In the context of traditional pattern recognition discipline, this process is usually called *feature extraction*.

Graphic processing transforms the image description in various ways, like calculating and adding to the description new, derivative attributes and relations, preparing the image for further analysis or picture synthesis.

Recognition may be said to give meaning to the information conveyed by the image description, producing usable knowledge about objects and relations depicted in the input picture. Often the result of the recognition process consists simply of a name (of a class) of the object found in the image.

Abstract processing, that can be also termed *abstract reasoning*, transforms the system knowledge producing new items of knowledge or restructuring it for various reasons, like facilitating compact storage or knowledge retrieval.

Visualization presents selected chunks of knowledge in a visual manner, producing a structural description of an appropriate image according to some selected visual language, cf. Section II.2.

Picture synthesis takes the image description and renders it in the form suitable for generation of the output picture, possibly after some amount of final processing. Because image descriptions have often the form of parameterised, continuous geometrical primitives, while encoded pictures are usually represented as raster images, i.e. digital arrays of picture points (pixels), this process is also known as *rasterization* (sometimes the term *digitization* is also used, not very properly in this context).

Picture generation feeds the picture output device with appropriate data produced from the internal encoded picture.