Journal of Computing and Information Technology - CIT 3, 1995, 4, 215–227

Looking for Visual Primitives*

C. Arcelli, L. P. Cordella and L. De Floriani

Istituto di Cibernetica del CNR, Napoli, Italy
Dipartimento di Informatica e Sistemistica, Universita' di Napoli "Federico II", Napoli, Italy
Dipartimento di Informatica e Scienze dell'Informazione, Universita' di Genova, Genova, Italy

Visual primitives can be considered as abstractions of those informative subsets of an image which are of interest in a given vision task. After discussing their nature and some problems related to their extraction, pattern description in terms of primitives is considered. Eventually, models relating 3-D visual primitives in high level vision are discussed.

1. Introduction

When referring to human vision, visual primitives may be defined as abstractions of some informative subset of what is seen. They originate at perceptual level as a response to visual stimuli detected at retinal level. In the early times of computer vision research, hints on the types of primitives to look for in the digital pictures to be analyzed, came from the investigations on some of the nervous systems of various species (e.g., frog, squirrel, rabbit, cat and monkey). These investigations [1] showed that some specific triggers were events such as convex edges, oriented slits or bars (possibly moving), ends of lines, line segments and corners.

Machine vision is concerned with extracting information from visual sensors to enable machines to make "intelligent" decisions. In this respect, redundancy reduction is of basic importance in the design of information processing systems that can perceive and interact with the external world, since it gives evidence to the singularities of the images which will constitute the primitives to be utilized for higher level processing. In machine vision, visual primitives are the smallest image subsets with specific geometrical and/or structural properties, which are of interest in a given analysis and description task. They define the basic level of description, in a hierachy of descriptions, used to identify a given pattern.

In the framework of computer vision, the term "visual primitive", or better "primitive component", has a clear meaning for people familiar with the structural approach to visual pattern recognition [2], [3]. This approach assumes that a complex pattern can be decomposed into simpler subpatterns, possibly in a recursive way, and then characterized (i.e., described) in terms of simple components, which are called "primitives", and of their relations. In this way the structure of a pattern can be outlined. Primitive components should not be further structured, or better their structure should not be of interest for the considered purpose.

In the decision-theoretic approach to recognition [4], [5], the term "feature" is certainly more familiar than the term "primitive". In this case, a pattern is characterized by a set of features (e.g., simply a set of measurements performed on the raw data) and a feature vector is assumed to represent the pattern to be recognized. Recognition implies the partition of a feature space into regions each pertaining to a different class. The elements of a feature vector should be such as to characterize a pattern so that it can be attributed to a specific class with high reliability. If this aim is achieved, it can be said that the used features effectively represent the pattern or that its description has been given in terms of essential features which include information about the primitives.

In pattern recognition, the most general mean-

* Reprint from book *Human and Machine Vision: Analogies and Divergencies* edited by Plenum Press, 1994

ing of the term feature could be: any pattern property that can be parameterized. The term could thus be applied uniformly to specify any type of entity assumed to characterize a pattern, from the simplest to the most structured one. However, a terminology for a feature hierarchy can be suggested: starting from the raw data, "local features" can be generated directly from them (absence of any structure); "intermediate features" are partial aggregates of the former ones, while a structured set of them, playing a special descriptive role for a pattern family, is a "primitive feature". Semantically significant assemblies of primitive features can be defined, and an object can be seen as a terminal (primitive) feature assembly [6]. This terminology may also suggest a strategy for achieving pattern description.

The determination of a set of primitives, in terms of which the patterns of interest can be described, is influenced by a number of circumstances, such as the nature of the data, the specific application considered and the technology available for implementing the recognition system.

Selecting primitive features is a complex problem that has no general solution. There isn't a "universal picture element" nor a simple alphabet of primitives generally valid. According to K. S. Fu [3], a general selection criterion is that the primitives should serve as basic pattern elements to provide a compact but adequate description of the data in terms of the specified structural relations; moreover they should be easily extracted.

In Section 2, visual primitives and problems related to their extraction are discussed, while in Section 3, pattern description in terms of primitives is considered. 3-D visual primitives and models relating such primitives in high-level vision are discussed in Section 4, where aspects regarding the use of CAD models and relational representations are emphasized.

2. Visual primitives

Although features useful to form visual primitives can generally be evaluated from texture, shading, motion, depth, etc., for a large class of images they are mainly extracted from boundaries which, in their turn, consist of edges. Edges represent small areas of high local contrast in correspondence with discontinuities in intensity, color, texture and so on. Accurate detection of edges is therefore crucial to automatic feature detection and object recognition.

In general, images may contain a number of edges occurring at any orientation and with sizes varying from very short (in this case the edges may be regarded as dots) to rather long. Edges may constitute the boundaries of closed regions, or may originate other primitives such as crossings, junctions, corners and bends. The significance of these primitives was shown in an early paper [7], where they were taken as the lowest level units in a recognition procedure for line patterns. The usefulness of a set of similar features (horizontal, vertical and oblique straight lines; curves which are closed, open vertically, open horizontally) was also discussed [8].

The edge detectors usually form the basis for complex digital image processing and analysis. For instance, to extract arbitrary curves such as objects boundaries from a noisy background one can use line detection operations which characterize line-like features as a succession of short edges, aligned in a given direction, which are brighter (or darker) than the points on either side of them in the orthogonal direction.

Edges provide an indication of the shape of the objects in a picture. However, since in machine vision edge detection is concerned with ideal edges corrupted by various forms of noise, it is often difficult to design edge detectors which manifest good performance characteristics when the edges are incomplete and degraded.

It has been suggested that edges should be found by using a symbolic data representation rather than the original numerical output from an edge detector. To this end, a vector description is associated with each edge element in the two-dimensional image. This vector contains information about the type of edge, the degree of contrast as given by the gradient, the position of the edge, its orientation. After applying an elementary edge detector, the edges are aggregated into line segments by using rules based on the examination of the compatibility of the symbolic edge descriptions in a local region. The new data array results from a purely data driven procedure, not determined by the observer's knowledge of

Fig. 1. Examples of figures with equal global organization, but different constituent parts.

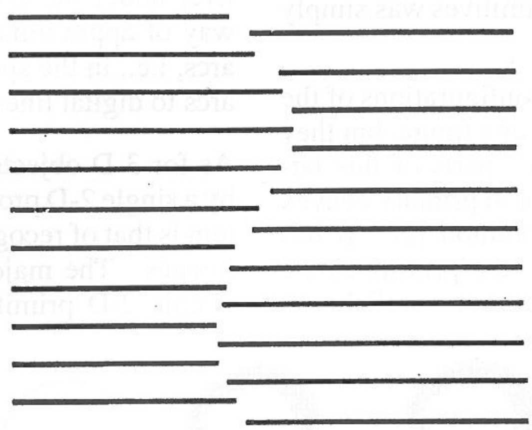


Fig. 2. Boundary line defined by the good continuation of the end points of the right and left gratings.

the semantic aspects of the visual input, and is termed the primal sketch [9].

According to Gestalt theory [10], [11], human vision is designed to structure the spacial features of an object in such a way that they are perceived as properties of the object, all together in the same construction, and not merely as individual parts. Gestalt psychologists attempted to define the organizational principles by which the global aspects of a visual scene are abstracted from the details. For instance, while figures may differ considerably regarding the nature of their constituent parts, the proximity of the parts produces the same perceptual interpretation, and such figures are understood as equivalent patterns (see Figure 1). Similarly, the role of good continuation is crucial to force the end points of the lines to be perceived as linked so as to form a boundary line (see Figure 2).

Principles of good continuation, similarity and proximity play a major role in the perception of closed regions. Accordingly, in image analysis the closed regions may be identified i) by linking a suitable sequence of edges; ii) by grouping pixels on the basis of gray-level homogeneity (region growing).

Segmentation is the assignment of the pixels in a given picture to one of many disjoint sets such that the pixels in each set share a common property. The criterion used to assign each pixel to a set and the number of sets are largely dependent on the desired task, the desired description, and the scene. Thresholds are used to assess property significance. The processes of edge finding and region growing may cooperate to build a segmented image, producing better results than those that would be obtained by either technique alone [12].

Texture analysis may also be considered as a problem of region segmentation, where the comparison has to be accomplished between neighboring areas or patches of pixels. Comparison is still performed on the basis of similarity and proximity, but it takes into account local patterns, i.e., a whole set of primitives. A major source of difficulty in the study of texture has been how to describe these patterns. Features as coarseness, contrast, and edge orientation appear to be the primary factors which influence the aggregation processes originating textured regions.

From the applicative point of view, among the oldest attempts to primitive feature based description and recognition of 2-D patterns, are the works of Shaw [13] for bubble chamber tracks and of Ledley [14] for chromosome analysis (late 60's). Pieces of lines or edges of different curvature, orientation and size were typically considered as primitives in these cases and the relation between primitives was simply concatenation.

2-D primitives can be line configurations of the boundary or of the skeleton of a figure, but they can also be two-dimensional parts of this latter. Figure decompositions into primary convex subsets have been quite common [2]. It has been claimed that the more the primitives are simple, the more inefficient and unreliable is the recognition process. However, extracting more complex primitives implies the difficulty of grouping simple features to form more complex features. For instance, the variability of the descriptions possible for the P's shown in Figure 3 depends on the way they are represented and on the primitives that are selected. If characters are approximated with polygonal lines, thus a possible primitive is the straight segment. The circular arc would be a more powerful primitive, under the condition to have an effective way of approximating characters with circular arcs, i.e., in the specific case, of fitting circular arcs to digital lines (see Figure 3).

As for 3-D objects, they are often represented by a single 2-D projection. In this case the problem is that of recognizing 3-D objects from 2-D images. The majority of the approaches uses simple 2-D primitives such as line segments,

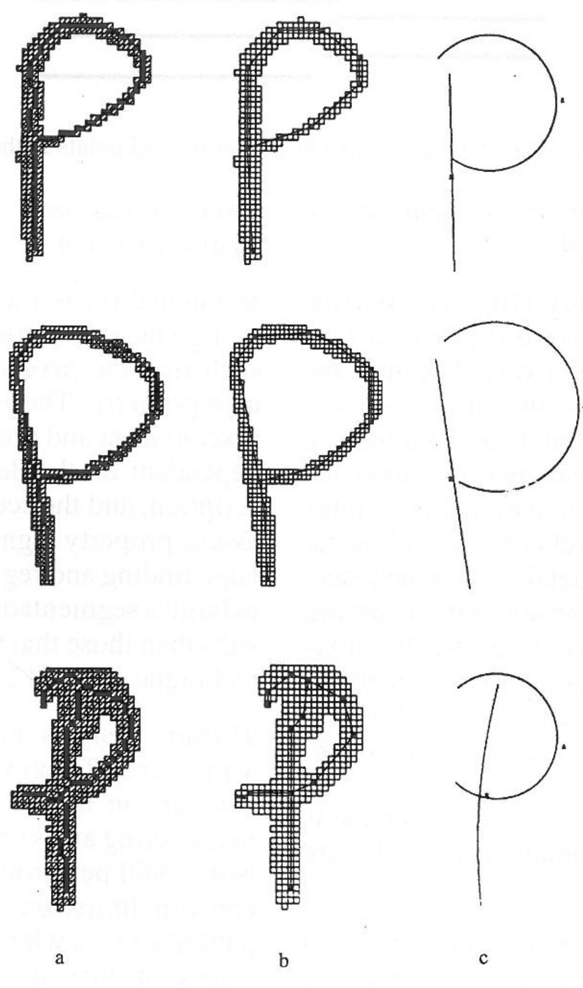


Fig. 3. Some samples of handwritten characters (letter "P"); a) bit map and skeleton (superimposed), b) polygonal approximation of the skeleton (superimposed), c) approximation of the skeleton with circular arcs. For different representations (b or c), different primitives can be convenient (see text).

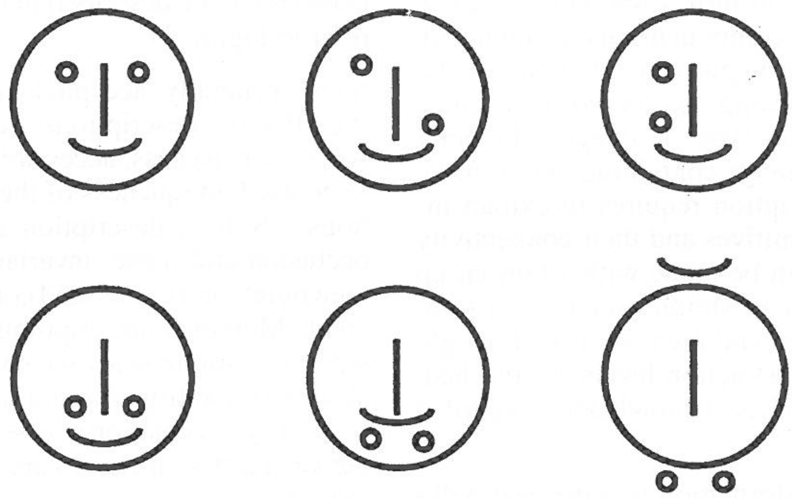


Fig. 4. Schematic drawing of a face and some of its scrambled versions. When the relations among the constituent parts are changed, the figure may be no longer perceived as a face.

corners, inflections and 2-D perceptual structures.

3-D volumetric primitives that assembled together in different ways can serve to build up objects have been proposed and employed (polyedra, generalized cylinders, geons) [15], [16]. They are preferably specified in a qualitative way, so as their use can be more general. 3-D objects must be recognized independently of the particular view in which they may appear: this implies that a number of possible views per primitive has to be taken into account.

Most computer vision systems, which are designed to recognize three-dimensional objects, compare a scene model, constructed by processing images obtained from one or more sensors, against entities in a model database containing a description of each object the system is expected to recognize [17]. The development of such model-based recognition techniques has occupied the attention of many researchers in the computer vision community for years [18], [19], [20]. Three-dimensional model-based computer vision uses geometric models of objects and sensor data to recognize objects in a scene. Likewise, CAD systems are used to interactively generate 3D models during the design process. Despite this similarity, there has been a dichotomy between the two fields. In recent years, the unification of CAD and vision systems has become the focus of research in the context of manufacturing automation [21]. The term CAD-based vision has been coined for research in vision employing CAD models for various visual tasks [20], [21]. Other approaches to model-based recognition use object representations derived from a model of human vision [22].

3. Image Description

Image description is a parameterization process according to which a suitable representation of an image is transformed into a data structure made up of features and of relations among them. In fact, a listing of features does not generally provide a sufficient description. One cannot distinguish between T and L if features such as "horizontal segment" and "vertical segment" are just listed. It is also necessary to know how the segments join together. Thus, in addition to a set of features, one must specify a set of relations and a set of rules describing the patterns in terms of features and relations. Figure 4 exemplifies this problem. Examples of relational features are: right of, left of, above, below, inside, outside, at the center of, surrounded by, near, far, next to, attached to, visible from, overlapping, occluding, isolated, grouped, larger, smaller, longer, shorter.

Thus, for complex images, the most interesting approach seems to be the structural one according to which the image is divided into parts (high level primitives for that image) and described in terms of those primitives and their interrelationships; in turn, each part is seen as made of primitive components and analogously described. In principle, the process can be iterated untill reaching the feature level most convenient for that type of image. In other words, given an image containing one or more objects, the description requires to extract instances of the primitives and their connectivity relations. This can be made with a bottom up approach, combining simpler features to form complex primitives and then objects, although, as far as higher abstraction levels are reached, the need for model based knowledge is regarded unavoidable.

In early vision, information is processed without any prior knowledge of the viewed image. The goal is to extract relevant information useful for further analysis (e.g., sharp changes in image luminosity are a priori relevant information). The result of early processing is generally still an image, which is a (possible) representation of the original one. Further processing can lead, through successive stages, to obtain more convenient representations.

The aims of the representation phase are manyfold: it serves to reduce noise, to compact information, to decompose the pattern into "meaningful" primitive components, to represent components in such a way that insignificant shape variations are hidden. In fact, features actually invariant within a class and peculiar of that class may not be directly detectable from the raw data. Suitable processing is generally needed. The representation phase should be intended as a first step of a process of abstraction that leads from the specimen to the prototype (model).

Although the decomposition process outlines specific features of an object, it leads to a representation that is still susceptible of different descriptions depending on the scope, the wideness of the image class, and so on. It can be said [23] that a meaningful decomposition and description implies to single out the primitives (features, component parts) as much as possible invariant with respect to the differences existing among the specimens of a same class, and to describe the components and the structure they form (i.e., what we have called the representation) in such an essential way as not to display the differences between members of the same class, but rather to put in evidence the similarities. Ideally, all the members of a same class should be described in the same way (still refer to Figure 3).

Some generally accepted criteria for obtaining effective descriptions are: attention to the scope, conciseness, accessibility of the features to be used, uniqueness of the obtained descriptions. Still, a description must be robust to occlusion and noise, invariant over a range of viewpoints and scales, and computationally efficient. Moreover, descriptions should be stable, while remaining sensitive enough; however, to absorb variability within a same class and not to destroy information necessary to discriminate between different classes are generally conflicting aims.

Indeed, all the above issues are often competitive, in the sense that using certain features may maximally satisfy some criterion but preclude to maximize the others. This could suggest to introduce a redundancy of feature types.

As mentioned above, the number and the nature of the features to be used for describing a pattern depend on the task. If a description serves to the purpose of faithfully reproducing the described object, then it has to be quantitative and very detailed. On the contrary, for recognition purposes, few qualitative features are more convenient especially if, for the application at hand, the number of classes is restricted. In human vision, the dependency of the used features on the task, has been outlined since long time ago by recording eye movements during constrained and unconstrained observation of pictures [24]. These experiments also outline the role of knowledge when looking for primitive features.

One of the image features that is most commonly used for description is shape. However, shape is generally a very complex feature, whose description is not trivial. Recent studies on anatomical connections have shown that in non-human primates the cortex processing visual information can be divided into several areas, and that the visual pathways appear separated into two streams, running dorsally and ventrally [25]. Studies of the effects of brain lesions have also indicated that this ventral stream of processing is associated with shape perception and object recognition [26]. As for studying the human brain mechanism involved in recognition, a recent technique has concentrated on monitoring local blood flow by using positron emission tomography within the brain during shape processing tasks [27]. All the previous studies, however, did not provide information about mechanisms by which shape is computed; for that information single cell recording techniques can provide useful insights. In this respect, key neurophysiological studies along the posterior anterior axis of the ventral pathway have recently been reviewed [28].

The term "shape" usually refers to a configuration consisting of several elements in some kind of mutual relationship. The elements are interchangeably called features, attributes, cues, or components. Examples of shape features are: boundary features (dominant points, curvature extrema, circular arcs, ...), region features (convex and concave pattern components, maximal discs, ...), boundary sinuosity, symmetry, compactness, angular variability, direction of the dominating axis, elongation. Each feature can assume values that may vary along a continuous or a discrete scale.

The quantitative study of object shape in terms of edges, angles, and contours began with Attneave [29]. Namely, he suggested that a boundary curve could be segmented by means of critical points which coincide with the points of maximum inflection. An alternative study of object shape in terms of regions was carried out by Blum [30], who proposed a new geometry for shapes based on the primitive notion of symmetric point. In this case, the shape of the object can be described in terms of the symmetric points and of the maximal discs, obtained as growth of these points and completely contained inside the object.

Also suitable parts of the object may be regarded as shape primitives [31] (e.g., see Figure 5). In this respect, a major problem in machine perception is to develop procedures that distinguish functional parts from purely nominal parts that lack psychological reality. In any case, a shape measure must be used to quantify the similarity or dissimilarity between shape descriptions.

4. Model-based object recognition.

Three issues dominate the design of a modeling scheme for vision: (i) the method adopted for model acquisition; (ii) the choice of a reference frame for the model, and (iii) the choice of a representation [32].

Methods for acquiring models have received relatively little attention. Usually, the models used for recognition have been provided manually. Manual construction of object descriptions is obviously time consuming and requires detailed knowledge of the object recognition system. This approach to model construction is impractical in applications where the set of objects to be recognized is large or changes frequently. An alternative is to construct models from examples using prototypical features extracted from images taken from a number of viewpoints. Models built by learning are often limited in their precision by the quality of the sensor. Much interest centers on using CAD models to construct models for object recognition. Object-centered CAD models provide a natural way to define an object and can be the source of the information necessary for its recognition, although they are often organized in ways that are not appropriate for vision.

The attributes and the geometric relations among component parts of an object must be defined with respect to a coordinate system, which in vision research is either object- or viewer-centered. Most work in object recognition has adopted object-centerd models, since they provide a natural way to express objects independently of the view [33]. Recently, there has been a lot of interest on 3D multiview representations, which model objects by a finite set of viewer-centered descriptions [34]. Each member of the set models the object by its 2D projection as seen from one viewpoint on the view sphere. By using such representations, features extracted from images can be directly matched with features associated with each member of the multiview model set. The shift from object-centered models to viewer-centered ones shifts processing requirements from object recognition time to object modeling time.

Computer vision researchers have used a variety of model types, which can be broadly classified as either descriptive (or quantitative), i.e., the model can be used to generate a synthetic image of the object, or discriminatory (or qualitative), i.e., the model information can be used to distinguish between different objects, but not to generate synthetic imaginery. Solid representations used in CAD are quantitative by nature:

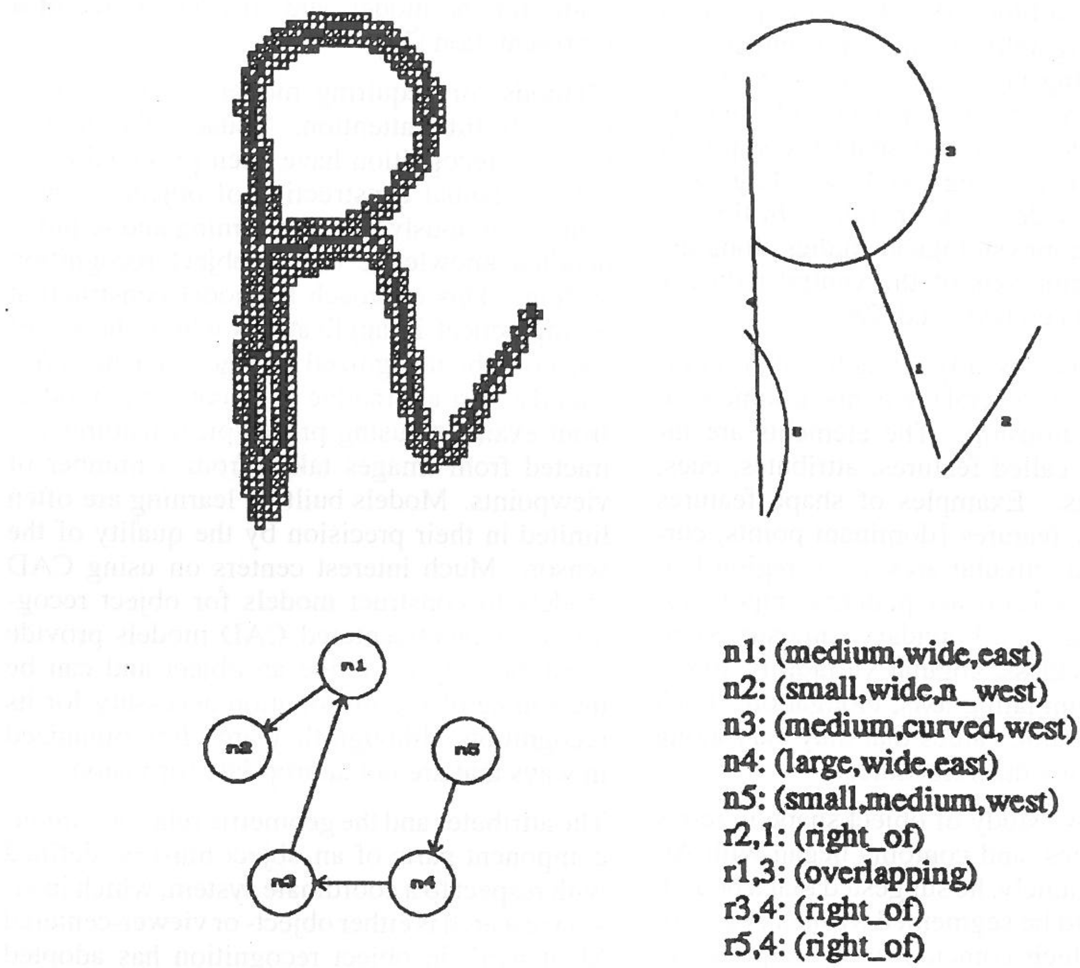


Fig. 5. Character shape description: a) the bit map of a character and its skeleton (superimposed); b) a possible representation and decomposition of the character in terms of circular arcs: c) description of the above representation with an Attributed Relational Graph whose nodes correspond to character components.

primitives (surfaces and volumes) are specified in terms of numerical parameters. If we are performing a visual recognition task and the objects to be recognized can be distinguished by examining some qualitative features of the segmented primitives, representations, which capture only those variations, might offer real advantages in processing [17].

The idea of qualitative representation was proposed by Biederman [22] as a model of human vision, but it also offers interesting properties for a computer model. In human vision, the retinal image is transformed at different levels of the visual pathway into various data representations as a precursor to possible object recognition. At the highest levels in this process, we have only a very sketchy knowledge of the exact details [23]. The main idea behind Biederman's approach is to coarsely reconstruct 3D objects using generic primitives, called geons.

While the geon representation has an intuitive appeal, the lack of quantitative information limits its usefulness in environments where discrimination between qualitatively similar, but quantitatively different objects is performed. If an assembly line is making a" family" of parts which differ only in scale, they would have identical geon representations, making discrimination between the differently sized items impossible without additional (quantitative) information [17].

Quantitative modeling consists of using" classical" CAD models, i.e., a boundary model or a volumetric one. A boundary model describes a solid object as a collection of boundary entities (i.e., vertices, edges, faces) and their mutual adjacency/incidence relations. Volumetric models (like, Constructive Solid Geometry) describe a solid object as the boolean combination of predefined volumetric primitives. Thus, they are typical of a design environment, since they somehow simulate the process of object construction by a designer. Moreover, CAD systems based on a CSG representation always store a boundary description of the object to speed up rendering as well as analysis operations. On the other hand, a boundary representation of the solid should be augmented with geometrical information to speed up the matching process. CAD models contain information for the local design operations such as what shape to extrude or what the profile curve for a sweep operation is. Features used in construction of models are implicitly rather than explicitly used in the CAD representation. For example, a dihedral edge, comprised within its adjoining surfaces, is not modeled as an edge per se but as two surfaces with adjacency information.

Object recognition techniques are based, for the most part on geometric features of the objects to be recognized. These include corners, edges and planar faces for polyhedra, as well as points, arcs of distinct curvature and regions of constant curvature for sculptured surfaces. Other features, such as axes of inertia, profile curves, surface textures properties, reflectance, etc. can also be used. Thus, several authors have recently proposed the use of a boundary model, produced by a commercial CAD system, augmented with additional features. For instance, in a boundary model, a line is characterized by its endpoints; an additional feature associated with a line is its length. Similarly, a planar surface is characterized in a boundary model by (i) coefficients of its plane equation and (ii) its bounding curve. Such description can be augmented, for instance, with a list of visible areas corresponding to the viewpoints on the uniformly sampled viewsphere.

Often, not only geometric features enhancing the object recognition task are added to boundary models, but relations other than the adjacency/incidence ones commonly encoded in a boundary data structures are stored in the representation. Examples of such relations are orientation, proximity, containment, covisibility, etc. Such relations are usually described in the form of a graph.

A graph representation has the flexibility of representing different types of attributes of, and relations among primitives composing an object model [32], [25]. Not only binary relations can be represented by a graph, but also higher-order relations can be encoded as hyperarcs connecting three or more nodes in a hypergraph [36]. Graph structures are the basis of classical boundary models. An example is given by the incidence graph, in which nodes describe the three basic topological entities (vertices, edges and faces) with their geometric attributes, while the arcs correspond to four incidence relations, i.e., Vertex-Edge, Edge-Vertex, Edge-Face and Face-Edge relations (see Figure 6).

An example of the use of a graph structure as a view-independent representation of a solid is given by the work of Zhang, Sullivan and Baker [32]. A hypergraph model of an object is computed from its boundary description produced through a CAD system. In such a hypergraph, nodes correspond to extended model features, including their shapes and the types of 2D image features that might match, while arcs and hyperarcs describe covisibility of model features. Covisibility of model features is generally view-dependent, in the sense that two features, that are covisible from one viewpoint, may not be covisible from another. Features are considered covisible if the probability of their co-occurence in images is high.

Graphs are also used in vision as conceptual representation of the visibility structure of a scene. An example is provided by the aspect graph introduced by Koenderink and Van Doorn [37]. An aspect graph is a representation of an object's topology; thus, it captures all viewpoints of an object. The aspect is the topological appearence of the object from a particular viewpoint. Slight changes in the viewpoint change the size of features, edges and faces, but do not cause them to appear or disappear. When a slight change in viewpoint causes a feature to appear or disappear, an event takes place. An aspect graph, or visual potential graph, is obtained by representing aspects as nodes and events between aspects as paths between corresponding nodes [21].

We have seen that graph representations are the basis for both "classical" boundary representations used in CAD and for "enriched" boundary ones used in vision systems. Recently, a lot of attention in CAD/CAM has been devoted to the development of the so-called feature-based

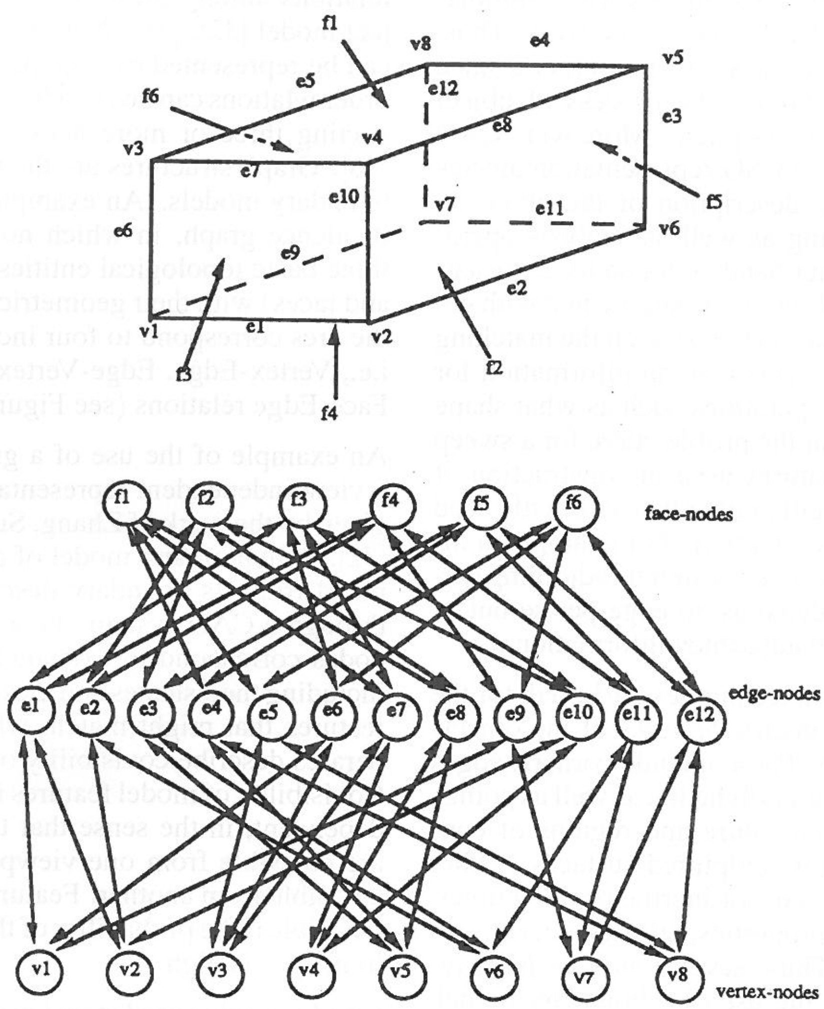
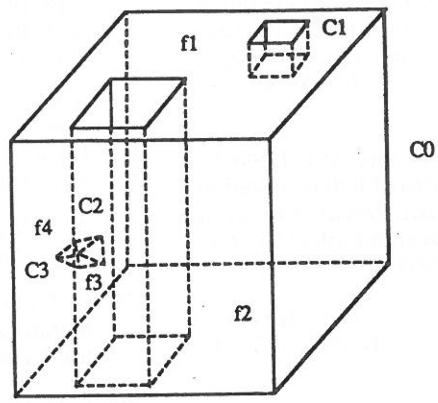


Fig. 6. An example of incidence graph of an object.

models. The term feature (shape or form feature) is used in the computer vision and in CAD/CAM literature with a different meaning, and sometimes a lot of confusion arises. Form features are a compact way of describing subparts of an object, which have a specific meaning in the context of the design or production process. There have been several attempts to give a truly satisfactory general definition of form feature. Pratt [38] defines a form feature as" a related set of elements of a geometric model conforming to characteristic rules allowing its recognition and classification and that, regarded as an independent entity, has some significance during the life cycle of the modeled product". The elements might be volumetric primitives in a CSG representation or geometric and topological entities in a boundary model. The basic idea is that such elements occur in a recognizable pattern when we consider the model from the point of view of a specific application, such as process planning for machining, or assembly planning. Examples of form features are protrusions, blind holes, slots, pockets, through holes, etc.

Relational descriptions of the decomposition of a solid object into its form features have been developed in the form of hypergraphs called feature graphs, which describe an object as a structured aggregation of face-adjacent object parts called components (see Figure 7). Such face-adjacencies are defined by subsets of the faces of each feature, shared by two or more components. Each node of the feature graph is again a graph describing the topology and the geometry of the feature boundary.



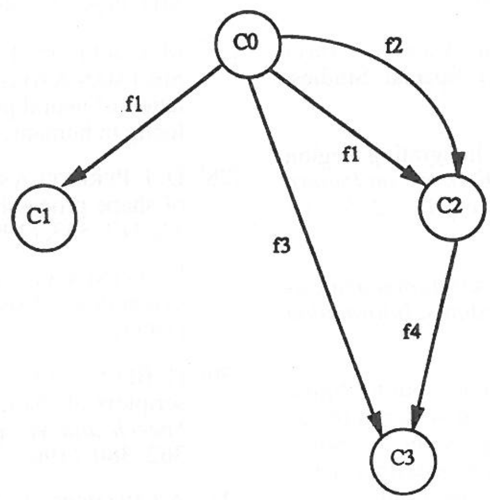


Fig. 7. Decomposition of an object in terms of form features.

Although feature-based modeling is still a research issue in CAD/CAM, the development of feature-based models as well as of form feature recognition algorithms will make feature-based models available as input representations for vision systems. In fact, feature-based models could be enhanced with geometric information as boundary models to allow the representation of features as geometric characteristics of a solid easily detectable from an image (i.e., visual features). Moreover, if an object is decomposed into shape features, we can use the aspects to represent a small set of primitives rather than the entire object. This could lead to a possible application of a "recognition-by-part" procedure. An open issue is whether the process knowledge embedded in a feature-based model can facilitate the object recognition process.

References

[1] H. B. Barlow, R. Narasimhan and A. Rosenfeld, "Visual pattern analysis in machines and animals", *Science*, 177, 567–575, (1972).

[2] T. Pavlidis, *Structural Pattern Recognition*, Springer–Verlag, 1977.

[3] K. S. Fu, *Syntactic Methods in Pattern Recognition*, Academic Press, (1974).

[4] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, Academic Press, N. Y., (1972).

[5] P. Devijver and J. Kittler, *Pattern Recognition: a Statistical Approach*, Prentice–Hall, Englewood Cliffs, N. J., (1982).

[6] R. M. Bolle, A. Califano, R. Kjeldsen, A complete and extendable approach to visual recognition, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14, 534–548, (1992).

[7] R. NARASIMHAN, On the description, generation, and recognition of classes of pictures, in *Automatic Interpretation and Classification of Images*, A. Grasselli Ed., Academic Press, New York, 1–42, (1969).

[8] E. J. GIBSON, H. OSSER, W. SCHIFF AND J. SMITH, "An analysis of cirtical features of letters, tested by a confusion matrix", in *A Basic Research Program on Reading*, Cooperative Research Project No. 630, U.S. Office of Education, (1963).

[9] D. MARR AND E. HILDRETH, Theory of edge detection, *Proc. of the Royal Society*, B, 207, 187–217, (1980).

[10] K. KOFFKA, *Principles of Gestalt Psychology*, Harcourt–Brace, New York, (1963).

[11] G. KANIZSA, *Organization in Vision, Essays on Gestalt Perception*, Praeger Special Studies, Praeger, New York, (1979).

[12] T. PAVLIDIS AND Y. T. LIOW, Integrating region growing and edge detectio, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12 , 225–233, (1990).

[13] A. C. SHAW, A formal picture description scheme as a basic for picture processing systems, *Information and Control*, 14, 9–52, (1962).

[14] R. S. LEDLEY ET AL., FIDAC: film input to digital automatic computer and associated syntax-directed pattern recognition programming system, in *Optical and Electro-Optical Information Processing*, chapt. 33, 591–614, MIT Press, Cambridge, (1965).

[15] S. J. DICKINSON, A. ROSENFELD, A. P. PENTLAND, Primitive-based shape modeling and recognition, in *Visual Form: Analysis and Recognition*, Plenum Press, New York, 213–229, (1992).

[16] M. D. LEVINE, R. BERGEVIN, Q. L. NGUYEN, Shape description using geons as 3D primitives, in *Visual Form: Analysis and Recognition*, Plenum Press, New York, 363–377, (1992).

[17] P. J. FLYNN, A. K. JAIN, CAD-based Computer Vision: From CAD Models to Relational Graphs, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13, 2, 114–132, (1991).

[18] P. J. BESL, R. C. JAIN, Three-dimensional object recognition, *ACM Comput. Surveys*, 17, 1, 75–145, (1985).

[19] R. T. CHIN, C. R. DYER, Model-based recognition in robot vision, *ACM Comput. Surveys*, 18, 1, 67–108, (1986).

[20] J. BRADY, N. NANDHAKUMAR, J. AGGARWAL, Recent progress in the recognition of objects from range data, *Proceedings 9th Int. Conf. Pattern Recognition*, 85–92, (1988).

[21] C. HANSEN, T. C. HENDERSON, "CAGD-based Computer Vision", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11, 11, 1181–1193, (1989).

[22] I. BIEDERMAN, Human Image Understanding: Recent Research and a Theory, *Computer Vision, Graphics and Image Processing*, 32, 29–73, (1985).

[23] A CHIANESE, L. P. CORDELLA, M. DE SANTO, M. VENTO, Classifying character shapes, *in Visual Form: Analysis and Recognition*, Plenum Press, New York, 155-164, (1992).

[24] A. L. YARBUS, *Eye Movements and Vision*, Plenum Press, New York, (1967).

[25] D. J. FELLEMAN AND D. C. VAN ESSEN, Distributed hierarchical processing in the primate cerebral cortex, *Cerebral Cortex*, 1, 1–47, (1991).

[26] L. G. UNGERLEIDER AND M. MISHKIN, Two cortical visual systems, in *Analysis of Visual Behaviour*, MIT Press, Cambridge, 549–586, (1982).

[27] M. CORBETTA, F. M. MIEZIN, S. DOBMEYER, G. L. SHULMAN AND S. E. PETERSEN, Attentional modulation of neural processing of shape, colour and velocity in humans, *Science*, 248, 1556–1559, (1990).

[28] D. I. PERRETT AND M. W. ORAM, Neurophysiology of shape processing, *Image and Vision Computing*, 11, 317–333, (1993).

[29] F. ATTNEAVE, "Informational aspects of visual perception", *Psychological Review*, 61, 183–193, (1954).

[30] H. BLUM, "A transformation for extracting new descriptors of shape", in *Models for the Perception of Speech and Visual Form*, MIT Press, Cambridge, 362–380, (1967).

[31] A CHIANESE, L. P. CORDELLA, M. DE SANTO, M. VENTO, Decomposition of ribbon-like shapes, *Proc. 6th Scandinavian Conf. on Image Analysis*, Oulu, Finland, 416–423, (1989).

[32] S. ZHANG, G. D. SULLIVAN, K. D. BAKER, The Automatic Construction of a View-Independent Relational Model for 3-D Object Recognition, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15, 6, pp. 531–544, (1993).

[33] D. MARR, K. H. NISHIHARA, Representation and recognition of the spatial organization of three-dimensional shape, *Proc. R. Soc. Lond.*, B, 200, 269–294, (1978).

[34] M. R. KORN, C. R. DYER, 3-D multiview object representations for model-based recognition, *Pattern Recognition*, 20, 1, 91–103, (1987).

[35] L. DE FLORIANI, Feature extraction from boundary models of three dimensional objects, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11, 8, 785–798, (1989).

[36] A. K. C. WONG, S. W. LU, M. RIOUX, Recognition and shape synthesis of 3D objects based on attributed hypergraphs, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11, 3, 279–290, (1990).

[37] J. J. KOENDERINK, A. J. VAN DOORN, The singularities of the visual mapping, *Biol. Cybern.*, 24, 51–59, (1976).

[38] M. PRATT, Solid Modelling — Survey and Current Research Issues, (1990).

Received: July, 1995
Accepted: November, 1995

Contact address:

C. Arcelli
Istituto di Cibernetica del CNR
Via Toiano 6
80072 Arco Felice
Napoli, Italy

L. P. Cordella
Dipartimento di Informatica e Sistemistica
Universita' di Napoli "Federico II"
Via Claudio 21
80125 Napoli, Italy
tel: +39 81 7683185
fax: +39 81 7683186
e-mail: cordel@nadis.dis.unina.it

L. De Floriani
Dipartimento di Informatica
e Scienze dell'Informazione
Universita' di Genova
Viale Benedetto XV 3
16132 Genova, Italy

CARLO ARCELLI received the doctoral degree in Physics from University of Bologna, Italy, in 1969. Since 1970, he has been working at the Istituto di Cibernetica of the National Research Council of Italy, Naples, where he has done research in the fields of picture processing and digital geometry. His current interests are mainly in shape analysis and description.

LUIGI P. CORDELLA received the degree of Doctor in Physics from the University of Rome, Italy, in 1964. He is Full Professor of Computer Science at the Faculty of Engineering of the University of Naples, Italy. He has written more than 100 papers in his fields of interest and is co-editor of four books. He has been active in the field of image processing and pattern recognition since 1972. His present research interests include shape description and recognition, document processing, optical character recognition and parallel computer architectures for vision.

LEILA DE FLORIANI is Professor of Computer Science at the University of Genova, Italy, since 1990. From 1982 to 1992 she has been a member of the research staff at the Institute for Applied Mathematics of the National Research Council in Genova. Leila De Floriani has written over 90 technical publications on the subjects of algorithms and data structures, geometric modeling, and computational geometry. She has been a member of the programme committee of several international conferences and one of the general chairpersons of the Eighth International Conference on Image Analysis and Processing. She is a member of the Editorial Board of the journals "Computer Aided Design" and "International Journal of Geographic Information Systems".