3-D Structures for Generic Object Recognition

Gérard G. Medioni and Alexandre R.J. François
Integrated Media Systems Center / Institute for Robotics and Intelligent Systems
University of Southern California, Los Angeles, CA 90089-0273
{medioni,afrancoi}@iris.usc.edu

Abstract

We discuss the issues and challenges of generic object recognition. We argue that high-level, volumetric part-based descriptions are essential in the process of recognizing objects that might never have been observed before, and for which no exact geometric model is available. We discuss the representation scheme and its relationships to the three main tasks to solve: extracting descriptions from real images, under a wide variety of viewing conditions; learning new objects by storing their description in a database; recognizing objects by matching their description to that of similar previously observed objects.

1. Introduction

Recognizing 3-D objects in an image is a complex problem along multiple dimensions. One has to *extract* relevant descriptions from the image, *match* them with one or more of the objects from a previously constructed database, and *learn* new objects for inclusion in the database. While these tasks are necessary, no consensus has emerged regarding the choice and level of features (2-D or 3-D), the type of indexing used, the matching strategy, and the order in which these tasks should be performed.

We argue that an appropriate way to solve *all* of the previous tasks is to generate and use high-level representations in terms of *volumetric parts*, their hierarchy and their arrangement. We further propose that part-based representations based on subsets of Generalized Cylinders constitute suitable means for shape abstraction of large classes of complex 3-D objects, and that these descriptions can indeed be generated from 2-D images under imperfect and realistic viewing conditions.

We start with an overview of the work appearing in the Computer Vision literature under the term *object recognition*. Then we outline the limitations of these schemes for solving the problem of *generic* object recognition. In section 3, we make a case for generic, high-level, volumetric part-based descriptions. We show how such descriptions can be computed from real images in section 4. In section 5, we show how to use the obtained descriptions for learning and recognition. We conclude the paper with a few remarks on what remains to be done to achieve generic object recognition.

2. Issues in object recognition

In the vision community, object recognition is often synonymous with model alignment. Most of the studies in object recognition are concerned with recognizing one object, for which an exact geometric model already exists, by finding its position and orientation. The approach is then a prediction/verification process, which has been used with great success in restricted environments. A review can be found in [11].

Approaches can be roughly classified into:

- *Appearance-based approaches*, in which no explicit 3D model is constructed, but multiple views are used instead. A very successful example is SLAM [20], which also takes lighting into consideration, and addresses indexing in a large space. This methodology has also been applied to faces [9].
- *3D model-based approaches*, in which a 3-D CAD-like model needs to be made available. The problem is then solved as a calibration/pose estimation one with respect to a reference 3-D [12]. Alternatively, one may compute, from the model, a set of view classes in the aspect graph approach [23][10].

A comparison between approaches is found in [19].

We are interested instead in the issues associated with *generic* object recognition, *i.e.* recognizing an object that might never have been observed before, and for which no exact model is available (see figure 1). In this context, the necessary components of an object recognition system are description, matching and learning.

Figure 1. *Generic* object recognition

- *Extraction of descriptions* consists of an interpretation of the image data into meaningful entities which could correspond to objects or portions of objects.
- *Matching* consists of assigning an identity to the extracted descriptions, a process which involves accessing stored models and comparing them with the image descriptions.
- *Learning* consists of acquiring objects not previously known to the system and describing similarities and differences with existing objects. This third component is generally ignored in many systems as it is performed by the user.

All these tasks rely strongly on the representation scheme used. It is this scheme which affects the choice of the strategies for the description process, the way models are accessed and compared, and the methods for learning. The nature of the representation also affects the complexity of each of the above tasks. Thus, choosing an adequate representation scheme is one of the most important design issues.

3. Description issues

As with any other problem, extracting generic descriptions requires an adequate representation scheme. Such a scheme becomes itself a generic model from which description methods are derived. As there is no universal representation scheme that can deal with arbitrary objects, we must derive generic shape representations that allow to capture a large class of objects.

3.1. Desirable properties

The desirable characteristics of a representation scheme for the purposes of generic shape analysis and recognition are well known [15][21]. They include local support, stability, discriminatory power and unambiguity. These characteristics are desirable not only for object space representation purposes, but also to recover descriptions from an image (or set of images) under realistic imaging effects.

The stability of a representation is with respect to changes in the scene properties including small irregularities on an object's surface, sensor noise, illumination changes and, especially, changes in viewpoint. We would like the descriptions recovered for the same object to be similar, even if its appearances are substantially different. This suggests that the shape representation scheme itself provides viewpoint tolerant image features from which final descriptions are built.

The search for such features has triggered much research in the community. An instance of which is the use of geometric invariants to characterize point-sets and certain curves [17]. Another instance is the derivation of higher-level viewpoint-tolerant properties of generic shape representations schemes, such as generalized cylinders [3]. To use point-sets requires visibility and selection of key points on an object. Even if certain invariant signatures for a number of points exist, they may not be stable or discriminatory enough. Building higher-level structures is necessary for stability and expressiveness. This constrains us to use a high-level representation scheme from which such structures can be identified in the image.

Besides the above characteristics, one would certainly wish to use a representation which allows the control of what is represented in an object. For each task, and at each of its steps, not all what can be represented in the object is needed. Rather, some tasks may require only coarse-level shape information whereas others some finer details. Representing an object's shape *hierarchically* is a natural way to control the representation scale.

3.2. Part-based representations

The complexity of viewed objects forces us to represent them in a way which makes their interpretation tractable. Man-made as well as natural objects are often made up of a number of components. This component (part-based) aspect stems from functional factors, as well as manufacturing constraints.

For some objects, a part-based representation may be difficult (such as bushes, terrain, face features), but for many others it is evident. For example, we have no hesitation in describing a teapot as consisting of a main body, a spout, a handle and a lid. For such objects, an explicit representation in terms of parts allows to capture their rich structure which, is important for the image interpretation process. It also allows to represent them even if some components are missing. A part-based representation is also useful for reasoning about function [25][27].

Such a representation is by design hierarchical, where parts can themselves be decomposed into sub-components, etc. This makes it convenient to control the description scale during image interpretation and during object recognition. Thus, it becomes possible to judge the similarity of objects in a coarse-to-fine fashion, a highly useful capability, especially in object learning. There exist, of course, many difficulties in generating such descriptions, as indicated later.

3.2.1. Volumes vs. surfaces. The two main representations for capturing the structured nature of an object are surfaces and volumes. Volumetric representations provide the correct level of abstraction for viewpoint tolerant descriptions. In fact, volumes lead themselves more naturally than surfaces to abstract an object's shape. For example, a stick figure using only the axes of generalized cylinders is often a good abstraction of the structure of a complex object (a horse for instance; see figure 2). Abstraction from a surface-based representation is unclear in this case. Surface attributes can be easily determined

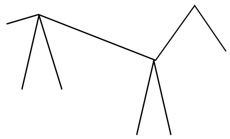


Figure 2. Volumetric part-based descriptions support natural shape abstraction (adapted from [31])

from a volumetric description, whereas the reverse is not as direct. The differences between two objects' shapes may not be directly accessible at a symbolic level when using surface-based descriptions. In fact, surface-based representations of otherwise similar objects can be drastically different.

A volumetric representation provides means to analyze shape similarities (or differences) along several criteria (for each of the intrinsic functions of the volumetric description). By decoupling the representation into a set of attributes, it becomes easier to compare objects and decide in which ways they are similar or different. Consider the two objects of figure 3. Despite their surface differences, we are still able to say that they have a similar structure, thanks to the volumetric parts composing them. Furthermore, the corresponding parts of the objects have quite different surface representations, a cylindrical and a piecewise polyhedral one for the heads, for example. But we are still able to judge their similarities (sizes along the parts axes and cross-sections), and differences (circular vs. polygonal cross-sections). This type of analysis is very desirable for recognition and even learning. In this case, the representation is symbolic and discriminative.

We find generalized cylinders (GCs) to be particularly suited in that they have many of the above-mentioned characteristics and they capture a large number of both natural and man-made objects. They have been used in many occasions, in past and on-going work, for the purpose of addressing one or more of the steps needed for generic object recognition [5][7][18][21][22][24][26][28][29]. However, without any restrictions, GCs are too general to be practical. Rather, subclasses of GCs are needed which are defined well enough to be practical yet general enough to capture a large set of objects. Examples of such subclasses of GCs are *straight homogeneous generalized cylinders* (SHGCs) [24][26][29], *planar right constant generalized cylinders* (PRCGCs) [28], and *circular planar right generalized cylinders* (circular PRGCs) [30].

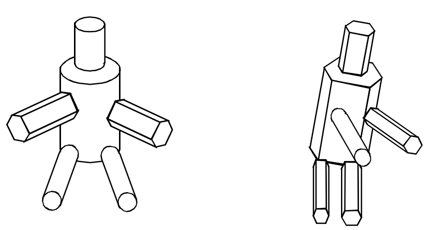


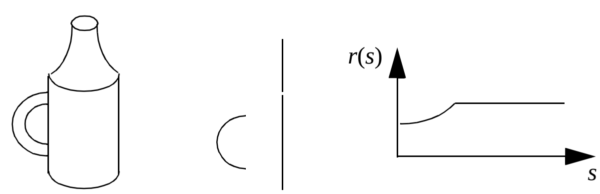
Figure 3. Volumetric representations are more suited to shape comparison (adapted from [31])

3.2.2. Segmentation into Parts. An issue that naturally arises at this point is the definition of a part. One could justifiably claim that the decomposition into parts is not a property of the object, but of our perception of it and, therefore, there might not be a rigorous way to define parts. However, in the case of GCs, a definition which was adopted since early work [21], consists of delimiting parts at discontinuities of any of the intrinsic functions, the axis, the sweep or the cross-section. An example is shown in figure 4. In this example, the segmentation of the handle (left-most part) is based on the discontinuity of its axis with the other axes (figure 4b), whereas the segmentation between the top part and the central one is based on the (first order) discontinuity of the sweep function (figure 4c), their axes being continuously connected.

This definition can be used both to partition a 3-D model and to segment the image of an object. The discontinuities in the intrinsic functions create discontinuities in the image, such as in symmetry relationships, and thus provide generic criteria by which to decompose shape. This captures, for example, the natural decomposition of parts at articulation joints where, due to the relative motion of parts, discontinuities in their shape are produced (in this case, axis orientation discontinuities, as between the upper and lower legs of a human).

Another issue is given a certain definition of a part, in terms of 3-D properties and image observables, to what extent can it faithfully capture the shape properties of an object. It is true that part models, such as GCs, are idealized in that they typically consist of smooth intrinsic functions and thus may not capture all shape details of an object. However, such part models, in the worst case, can still encode the coarse structure of the object's shape. This structure is often powerful enough to differentiate between objects and is thus very useful in recognition.

3.2.3. Psychological evidence. Part-based representations, based on volumetric primitives, appear to be psychologically plausible. Studies of human perception and recogni-



a. original object b. axis discontinuity c. sweep discontinuity

Figure 4. Segmentation into parts based on discontinuities in the intrinsic functions (adapted from [31])

tion of 3-D objects from their line drawings show a strong influence of the structured nature of objects on the perception performance of subjects [2]. The crux of the findings is the fact that simple geometric primitives (geons, analogous to qualitative GCs) and their interactions provide substantial viewpoint invariant object shape information. The geons are volumetric by nature and their perception has been found to be stable under rotations in depth, contour deletions and even occlusion.

This theory does not account for textured or highly irregular objects. Nonetheless it accounts for a large set of man-made and even natural ones (such as animals). It also accounts for new objects, not known *a priori* to the observer.

4. Computation of part-based descriptions

Deriving generic descriptions from an imperfect image (or images), in a data-driven fashion, is one of the most challenging problems in computer vision. A promising methodology is to adequately implement principles of perceptual grouping. The difficulty is then in the derivation of grouping (segmentation) and description methods that work not only on imperfect images but for large classes of objects as well.

4.1. Organization

We view the description module as consisting of a bottom-up process. However, this process makes use of expectations about image properties, which can be thought of as top-down generic knowledge.

An important aspect of this process is its organization. We do not believe it can be solved at a single feature level, for example by finding the best edge detector or boundary grouping method. There is a direct link between the hierarchical nature of the representations and the organization of the description process. The latter should proceed in successive stages where features are detected then grouped to form higher-level descriptions. The feature groups become themselves features of a higher level type which in turn can be grouped, and so on. Each level increases the scope of the interpretation process by building a geometric context based on previous feature groups. Thus, feedback loops can be used from higher levels to lower ones in order to refine the features or their groupings based on the information gathered from the increased scope.

This organization is illustrated in figure 5. The description module starts from an intensity image and ends at a level where structured representations in terms of parts (generalized cylinders) and their relationships are obtained. This representation includes information about the 3-D shape of each complex object so detected.

4.2. Tasks

The tasks that need to be performed at each feature level are feature *extraction*, *grouping* and *selection* (and refinement) of feature groups. Each of these tasks constitutes a real challenge in itself. Important issues here are which features to extract, what criteria to use in grouping them into meaningful clusters and how to know which ones are indeed meaningful and which are not.

The role of the generic shape representation scheme in this process is crucial. Knowledge of this scheme allows the derivation of generic constraints (in the form of expectations) in most of the above tasks, especially as we go higher in the hierarchy, to overcome noise and other image imperfections. Part of these constraints are image properties that any element of the class of objects captured by the shape scheme must satisfy if it is observed. In using these properties to extract object descriptions, the extraction process takes explicit account of the shape representation (the desired output of the system) including its dimensionality. In the case of 3-D representations, the expectations are in the form of projective properties which bear the 3-D attributes of the shape of the object. As illustrated in figure 6, the 3-D object projects onto the image object (a set of features to be extracted). The 3-D representation scheme (based on generalized cylinders) projects onto a set of projective properties. The segmentation and description process makes use of the projective properties to extract relevant descriptions from the image features (finds image features which satisfy these properties). This results in image (projective) descriptions from which 3-D attributes of the detected object's shape are inferred.

4.3. A hierarchy

The typical hierarchy used in past and current work for the description process includes *boundaries*, *surface patches*, *volumetric parts* and *compound objects*. The tasks at these levels are all inter-related as we believe a robust system should perform all levels of grouping in a concurrent manner. Unfortunately, in most systems the

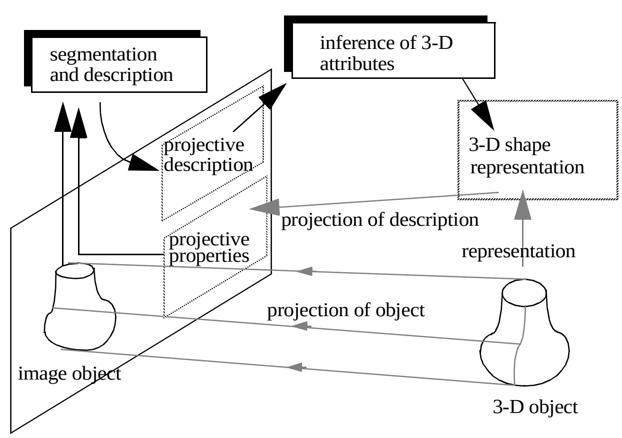


Figure 6. The description extraction process (adapted from [31])

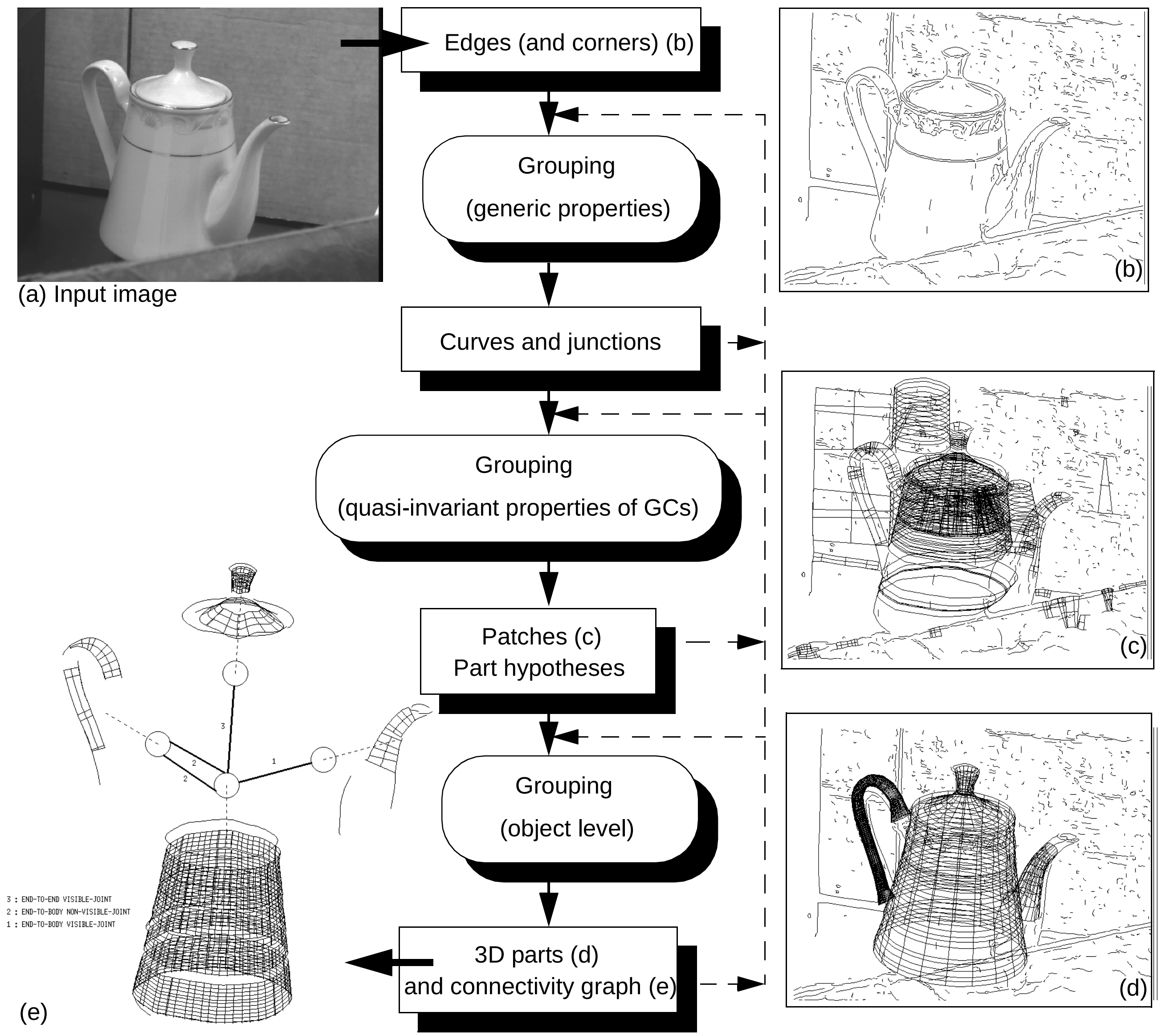


Figure 5. From image to 3D description (adapted from [31])

process is linear with little (or no) feedback, making higher levels sensitive to errors at lower levels.

For example, no working (robust and efficient) boundary grouping system has been demonstrated yet. Most of the work has either outlined some of the grouping criteria and/or implemented them in some prototype system demonstrated on few (real) images (e.g. [13][16]).

Some progress has been made in going beyond signal edges, detecting perceptual edges by enforcing continuity and co-curvilinearity generic constraints. While they significantly improve the detection of contours, they cannot be expected to solve the contour extraction problem, as they miss the importance of higher-level structures in the segmentation (and description) process. Thus, we believe it is important not to be over-ambitious in the design of continuity criteria solely based on boundary information. The more promising direction is to exploit feedback from higher-level processes.

Once partial contours are extracted, the next step is to infer surface patches from them. The most powerful constraint is that of symmetry. An important issue is which symmetry relationships to use. To serve the right purpose, they should have the characteristics outlined for the shape representation scheme. Symmetries which are chosen intuitively may not serve the purpose as their information contents, about the presence of relevant scene objects, is unknown. Rather, they should be viewpoint tolerant properties of the shape representation scheme itself. One would also expect these properties to be stable not only to imperfections in object shape but to approximations of the viewing geometry as well.

There have been several efforts on this issue, most notably on the derivation and use of *invariant* and *quasi-invariant properties* of certain generalized cylinders. For example, a perspective invariant property of SHGCs derived by Ponce *et al.* [24] is the intersection of tangent lines at corresponding (symmetric) points of the outline is a straight line which is the projection of the SHGC axis. Quasi-invariants are looser versions of invariants which have a small range of variation over a large portion of the viewing sphere [4]. For example, a quasi-invariant property of circular PRGCs derived by Zerroug and Nevatia [30] is that the segments joining the projection of co-cross-sectional points of the outline and the locus of their mid-points are very close to forming a right ribbon under most viewing directions. Other symmetry relationships used are parallel and skew symmetries [28]. These properties apply to certain subclasses of GCs and more research is needed to derive other (more general) properties. For this, we view quasi-invariance as a central element because invariant properties are rare or so general as to be useless.

The inferred patches need to be pieced together to form volumetric part hypotheses. This process involves the merging of surface fragments of the same part surface together as well as different surfaces of a part into a single part hypothesis. The grouping criteria must be rigorous and derived from the properties of the shape scheme.

Similarly, the object-level grouping should also exploit the projective properties of the shape representation scheme for grouping different parts which are likely to belong to the same object into a single object hypothesis and inferring the 3-D shape attributes (qualitative or quantitative) of each object.

5. Recognition using part-based descriptions

Once good descriptions (i.e. part-based, volumetric, hierarchical) are extracted from the image, we need to define a recognition engine to manipulate them. Each new object must be stored in a database in such a way that it can be retrieved efficiently when a candidate description is proposed for recognition. The key issues are the organization of the database (indexing based on high-level representations), and the retrieval methodology (matching high-level representations).

We rely on four principles to achieve our goal of generic object *re*-cognition, explicitly based on previous observation:

- Hierarchical object description
- Symbolic part description
- Hierarchical database organization
- Partial-match hypotheses generation, followed by interpretation and validation.

A system implementing these principles for 2-D shapes is described in [8]. The same principles apply to directed acyclic graph descriptions obtained for 3-D objects.

It should be noted that the described approach does not obviate the need for geometric reasoning and pose estimation, but rather proposes to invoke them once the object has been recognized. The purpose is then to estimate pose, but also to emphasize differences with existing stored models.

5.1. Symbolic, hierarchical descriptions

The recognition process uses symbolic, hierarchical object descriptions. The connectivity graph is transformed into a mono-rooted directed acyclic graph (that can be treated as a hierarchy for our purpose). Description hierarchies are constructed so that the few main parts of an object appear on the first level, while deeper levels add increasing levels of detail to the description (a good criterion for orienting the edges in the description graph is the size of the parts). Symbolic part descriptions are inferred from the quantitative descriptions extracted by the process described above. The intrinsic geometrical properties thus encoded are similar to that of geons [2]. The description also encodes geometrical relationships between connected parts.

5.2. Learning: dynamic database organization

Once the symbolic description is obtained for an object, it is stored in a database from which it can later be efficiently retrieved when another description is submitted for recognition. Efficient retrieval is achieved using an original indexing mechanism.

The most natural and efficient data structure for indexing hierarchical object descriptions for retrieval is a hierarchy. Organizing the descriptions in a hierarchy requires the definition of a partial order on the descriptions, which we infer from the description graph properties. Since the retrieval is data-driven, the indexing of a description hierarchy is based on its structure, complemented by a characteristic subset of qualitative part attributes. The qualitative partial description thus defined is called *I-Structure*.

Hierarchical descriptions are indexed as follows: to each node of the index is associated exactly one I-Structure. The actual descriptions referenced by the node are those that exactly match the node’s I-Structure. The nodes are themselves organized into a specialization hierarchy, based on a partial order on the I-Structures. An example of hierarchical indexing of hierarchical structures built from three types of parts is presented in figure 7.

Because of the definition of the partial order, the hierarchical index inherits interesting properties from the

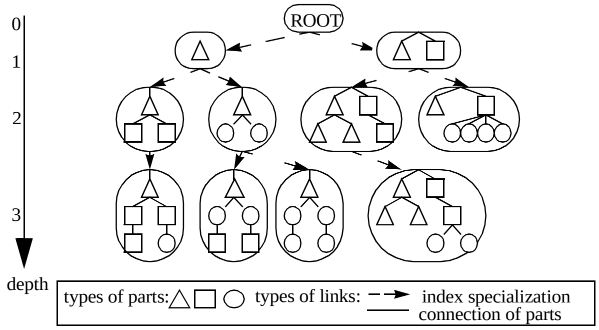


Figure 7. Hierarchical organization of the database for dynamic learning and efficient retrieval

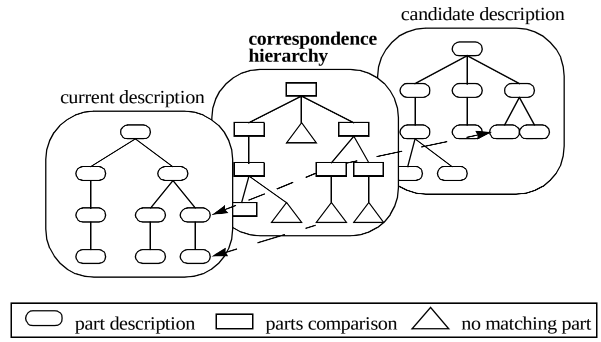


Figure 8. Comparison structure instantiated between two hierarchical descriptions. Only a few links are shown to preserve readability.

description hierarchies: the first level of the index hierarchy represents a coarse filter, that allows initial discrimination hypotheses based on the few main parts of the shapes, while the deeper levels allow to focus and refine the early recognition hypotheses.

One important feature of this indexing mechanism is that a description can be added to the database at anytime without any recomputation of the existing structure, which allows to start either with an important list of descriptions in the database or with a minimum database which is incrementally updated with new shapes (involving supervised learning).

5.3. Recognition: object description retrieval

Given an object description, recognition starts with the retrieval of similar descriptions in the database.

The retrieval step needs to proceed from partial matches to take into account possible occlusions and uncertainty. The partial matching retrieval produces hypotheses for the recognition. It is based on a metric to evaluate a similarity (or a dissimilarity) between shape descriptions, complemented by qualitative information about the comparison to allow high-level (symbolic) solution generation and explanation.

The dissimilarity between two shapes is based on the definition of *transition costs*, that represent the cost of the assumption that two different symbolic objects (parameter values, parts, descriptions) have been obtained from the same real object, because of noise in the original data, variation of the observation conditions, etc.

At the lowest level, transition costs are defined for each part description parameter. The transition cost between two part descriptions is defined as an aggregation of the transition costs between the description parameters.

In order to easily compute the transition costs between descriptions, we introduce a comparison structure which is a correspondence hierarchy instantiated between two shape descriptions (see figure 8). Each node of this graph points to two matching parts of the considered shapes, and its position in the correspondence hierarchy matches that of the parts it links in their respective description hierarchies. The transition cost between two paired parts is stored in the correspondence node. A transition cost between two descriptions is defined as the aggregation of transition costs between parts. The correspondence hierarchy allows to compute transition costs incrementally and allows to perform as well a symbolic, qualitative comparison of objects descriptions.

The retrieval process is a subgraph isomorphism problem, which is NP-complete. We have developed an algorithm which uses the index and the correspondence structures to avoid computing this expensive correspondence for all shapes. A partial match, based on the connection structure and the aggregation of dissimilarities between parts, is computed incrementally level by level between the new description and the possible candidates. The comparison structures are built level by level, keeping at each stage only the structures which point to a compatible description for the current level. This is made possible by the hierarchical indexing of the descriptions. At the beginning of the retrieval process, all compatible nodes on the first level of the index are selected, and the first level of the comparison structure is built for all the descriptions pointed by these nodes and the nodes in their sub-trees. The partially compatible subnodes of the previously selected nodes are then selected and the process is reiterated, increasing the precision of the matching. In the average case, the algorithm ends after the last level of the proposed shape has been processed, or after no candidate is left in the database for further investigation. The selected retrieved descriptions are used to give a classification for the new description.

The combination of the incremental process with the hierarchical indexing effectively decreases the number of shapes processed at each step, therefore dramatically reducing the average complexity of the retrieval, which we proved to be nearly independent of the number of descriptions in the database. Furthermore, the dynamic organization of the database provides a very powerful and flexible way of manipulating knowledge, especially with the open-world and adaptability assumptions, which are required for genericity.

6. Conclusion

We have articulated an approach to the inference of volumetric descriptions from images, have shown preliminary results on some images, and implemented a 2D recognition system. While this path shows promise, a number of deep issues remain.

Part-based representations using generalized cylinders are suited to capture the coarse structure of objects made up of basic primitives (a very large set). However, they are not suited to represent certain objects which are better described by statistical features, for instance bushes. Even for objects which can be approximated by generalized cylinders, the resulting descriptions may not be sufficient to capture fine aspects of the shape. Thus, generalized cylinders need to be complemented with other representation schemes which are better suited to capture these details.

The actual extraction of part-based descriptions from realistic images in unrestricted environments still remains a largely unsolved problem.

The recognition issues need to be explored in much greater detail. While everyone agrees that high-level descriptions reduce the complexity of the indexing and matching steps, the methods need not be *ad hoc*. More work is needed beyond the largely intuitive methods proposed so far.

7. References

[1] Bergevin R. and Levine M.D., Generic Object Recognition: Building and Matching Coarse Descriptions from Line Drawings. PAMI, 15:19-36, 1993.

[2] Biederman I. Recognition by Components: A Theory of Human Image Understanding. Psychological Review, 94(2):115-147, 1987.

[3] Binford T.O. Visual Perception by Computer. Proc. IEEE Systems and Controls, December 1971, Miami.

[4] Binford T.O., Levitt T.S. and Mann W.B. Bayesian Inference in Model-Based Machine Vision. Proc. AAAI Uncertainty Workshop, 1987.

[5] Brooks R.A. Model-Based Three Dimensional Interpretation of Two Dimensional Images. PAMI, 5(2):140-150, 1983.

[6] Clowes M.B. On Seeing Things. Artificial Intelligence, 2(1):79-116. 1971.

[7] Dickinson S.J., Bergevin R., Biederman I., Eklundh J.O., Munck-Fairwood R., Jain A.K., Pentland A.P. Panel Report: The Potential Of Geons For Generic 3-D Object Recognition. Image and Vision Computing, 15(4):277-292, April 1997.

[8] François A.R.J. and Medioni G.G. Generic Shape Learning and Recognition. Object Representation in Computer Vision II, pp. 287-320, Cambridge, U.K., April 1996.

[9] Georghiades A.S., Kriegman D.J., Belhumeur P.N. Illumination Cones for Recognition under Variable Lighting: Faces. CVPR, pp. 52-59, Santa Barbara, CA, 1998.

[10] Gigus Z. and Malik J. Computing the Aspect graph for Line Drawings of Polyhedral Objects. PAMI, 12(2):113-122, February 1990.

[11] Grimson W.E.L. Object Recognition by Computers: The Role of Geometric Constraints. MIT Press, Cambridge, MA, 1990.

[12] Kollnig H., Nagel H.H. 3D Pose Estimation by Directly Matching Polyhedral Models to Gray Value Gradients. IJCV, 23(3):283-302, Jun-Jul 1997.

[13] Lowe D.G. Perceptual Organization and Visual Recognition. Kluwer Academic Publishers, Hingham, MA. 1985.

[14] Mackworth A.K. Intrepreting Pictures of Polyhedral Scenes. Artificial Intelligence, 4:121-137. 1973.

[15] Marr D. Vision. W.H. Freeman and Co. Publishers, 1981.

[16] Medioni G., Lee M.S., Tang C.K. A Computational Framework for Segmentation and Grouping. Elsevier, 2000.

[17] Mundy J.L. and Zisserman A. editors. Geometric Invariance in Computer Vision. MIT Press, 1992.

[18] Mundy J.L., Huang C., Liu J., Hoffman W., Forsyth D.A., Rothwell C.A., Zisserman A., Utcke S. and Bournez O. MORSE: A 3-D object recognition system based on geometric invariants. IUW, pp. 1393-1402, Monterey CA, 1994.

[19] Mundy J., Liu A., Pillow N., Zisserman A., Abdallah S., Utcke S., Nayar S. and Rothwell C. An Experimental Comparison of Appearance and Geometric Model Based Recognition. Object Representation in Computer Vision II, pp. 242-269, Cambridge, U.K., April 1996.

[20] Murase H. and Nayar S.K. Visual Learning and Recognitioin of 3-D Objects from Appearance. IJCV, 14:5-24, 1995.

[21] Nevatia R. and Binford T.O. Description and Recognition of Complex Curved Objects. Artificial Intelligence, 8(1):77-98, 1977.

[22] Pentland A. Recognition by Parts. ICCV, pp. 612-620, 1987.

[23] Petitjean S., Ponce J., Kriegman D.J. Computing Exact Aspect Graphs of Curved Objects: Algebraic Surfaces. IJCV 9(3):231-255, 1992.

[24] Ponce J., Chelberg D. and Mann W.B. Invariant Properties of Straight Homogeneous Generalized Cylinders and their Contours. PAMI, 11(9):951-966, 1989.

[25] Rivlin E., Dickinson S.J. and Rosenfeld A. Recognition by Functional Parts. IUW, pp. 1531-1539, Monterey CA, 1994.

[26] Sato H. and Binford T.O. Finding and Recovering SHGC Objects in an Edge Image. CVGIP, 57(3):346-356, 1993.

[27] Stark L. and Bowyer K. Achieving generalized object recognition through reasoning about association of function to structure. PAMI, 13:1097-1104, 1991.

[28] Ulupinar F. and Nevatia R. Perception of 3-D Surfaces from 2-D Contours. PAMI, 15(1):3-18, January 1993.

[29] Zerroug M. and Nevatia R. From an Intensity Image to 3-D Segmented Descriptions. ICPR, 1:108-113, Jerusalem, Israel, October 1994.

[30] Zerroug M. and Nevatia R. Three-Dimensional Descriptions Based on the Analysis of the Invariant and Quasi-Invariant Properties of Some Curved-Axis Generalized Cylinders. PAMI, 18(3):237-253, March 1996.

[31] Zerroug M. and Medioni G. The Challenge of Generic Object Recognition. Object Representation in Computer Vision, pp. 217-232, New York City, NY, December 1994.