A Generic Shape Matching with Anchoring of Knowledge Primitives of Object Ontology

Dongil Han[1], Bum-Jae You[2], Yong Se Kim[3], and Il Hong Suh[4]

[1] Vision and Image Processing Lab., Sejong University
Seoul, Korea
dihan@sejong.ac.kr
[2] Intelligent Robotics Research Center
Korea Institute of Science and Technology
Seoul, Korea
ybj@kist.re.kr
[3] Creative Design and Intelligent Tutoring Systems Research Center
Sungkyunkwan University
Suwon, Korea
yskim@skku.edu
[4] Intelligence and Interaction Lab.
Hanyang University
Seoul, Korea
ihsuh@hanyang.ac.kr

Abstract. We have developed a generic ontology of objects, and a knowledge base of everyday physical objects. Objects are represented as assemblies of functional features and their spatial relations. Generic shape information of objects and features is stored using a partial boundary representation. Form-function reasoning is applied to deduce geometric shape elements from a feature's functions. We have also developed a generic geometric shape based object recognition method which uses many local features. The proposed recognition method considers the concept of ontology for representation of generic functions of objects. And the use of a general shape-function reasoning with context understanding enhances the performance of object recognition.

1 Introduction

To support a robot's interaction with a typical human environment requires a machine-understandable representation of objects, including their shapes, functions, and usages. Object recognition is supported by reasoning from each object's generic shape information. An object may have internal degrees of freedom, which means that its appearance and detailed geometry are highly variable, even though it fulfils the same function. Hence, many objects which have different shapes and geometry structures may be commonly known by the same name.

This condition can make model-based object recognition [1][2] extremely difficult because one may require either a classifier with a flexible boundary, or many different object models. Thus, for capturing and recognizing the object shape, function-based approach is introduced in [3]. The function models would capture a broad

M. Kamel and A. Campilho (Eds.): ICIAR 2005, LNCS 3656, pp. 473 – 480, 2005.

variation in allowed shape without reference to any specific geometric or structural plan. For this reason, function-based models seem to provide better support for "purposive" and "task-oriented" vision.

Previous research has explored the relationship between form and function for object recognition. The Generic Recognition Using Form and Function (GRUFF) system [4] represents objects as a set of functional elements (mostly planar surfaces), and spatial relations between elements. It performs generic object recognition by matching functional surfaces in the sensor input data to objects' definitions. It uses the Object Plus Unseen Space (OPUS) method to construct a partial 3D model from image and rangefinder data, but this has the drawback of sensitivity to varying image conditions.

Neumann *et al.* [5] performs context-based scene interpretation by modeling scenes as *aggregates*, where an aggregate is a set of entities and their spatial and temporal relations. They represent aggregates of scenes in description logic (DL), and match input models to scene definitions using the RACER DL reasoner [6]. However, their scene interpretation capability is beyond the current state-of-the-art in description logics, because a complete representation of the relations between entities exceeds the allowed expressiveness of RACER's DL.

2 Object Ontology

We adopt the ontology formalism in developing a generic ontology of objects. We use the standard OWL web ontology language, and the de facto standard Protégé ontology editor with OWL plugin [7]. Using this ontology, we have instantiated a knowledge base of ~300 objects for a typical indoor environment.

A. Representation of objects

Manufactured objects are typically assembled from multiple components, where each component contributes some specific functionality. Reflecting this, we adopt a hierarchical feature-based representation. An object is decomposed into a set of features and their spatial relationships, where a *feature* is a functionally significant subset of an object or another feature. Features are characterized by the functions they provide. Each feature can be further decomposed into more features.

B. Spatial relations

We define several spatial relations that frequently occur in everyday objects. For each spatial relation, we provide a definition that can be implemented as a (geometric) algorithm. For example, the ***above****(A, B)* relation is defined as: A is above B iff A's highest point is higher than B's highest point (with respect to the gravity direction), and A's lowest point is not lower than B's highest point.

C. Form-function reasoning

We characterize features using generic functions taken from function-based taxonomies for design [8][9]. While a feature is a 3D component, its functional elements, or *organs* [10], may correspond to subsets of its 3D shape. By applying form-function reasoning, we deduce geometric shape requirements for each functional element.

For example, a table's primary function is to limit the downward motion of many objects of any shape. The key feature for a table is a counter, which is typically a thin, rigid 3D slab. A counter's key organ is its top surface. To contact many objects implies many contact points, from which we deduce a planar surface. A table should also minimize the energy required to translate objects to different positions, which implies a horizontal orientation. Hence, we deduce a shape requirement of a *horizontal planar surface* for a counter's top surface.

D. Geometric shape elements

We define a qualitative representation of geometric shape elements. A shape element has a geometric datum (usually a surface), which represents a generalized portion of a solid's boundary. Other constraints on the allowable orientation, curvature, and tolerance of a shape element are specified using a phrase structure.

E. Representation of solids

A boundary representation (B-rep) is a 3D model that rigorously describes a solid by enumerating the topological elements of its boundary, including its faces, edges, and vertices. Other solid representations can be converted to B-rep, so a B-rep is a good candidate to be a generic solid representation.

On the other hand, an ontology of objects should also support generic representations of object families. This requires a capability to tolerate wide variations in specific geometry, while capturing the critical geometric relations only.

We adopt a *partial B-rep* scheme, in which a subset of a solid's boundary is fully specified, representing the critical geometric and topological relations only. Remaining portions of the boundary are abstracted away. Each solid has a bounding box data field, reflecting the principle that all real solid objects are bounded. Each feature class's shape information is then represented as a partial B-rep with 1 or more geometric shape elements.

3 Ontology-Based Object Recognition

A goal of this work is to design a vision-based context understanding system to enable a mobile robot to look for an object that it never seen before, in a place of first visit, where the object may be partially or completely obscured by other object. Such a visual context understanding system usually requires us to recognize place, objects, spatial and temporal relations, activities, and intentions.

In this paper, we describe the 2D object extractor. This module recognizes objects using two approaches. In the model-based approach, SIFT features [11]– [13] and edge features are directly matched to pre-computed vision feature-based models of objects. In the case that no vision feature models exist for an object, ontology-based object recognition proceeds as shown in Fig. 1.

- Local feature extraction obtains low-level vision feature information such as edges, lines, arcs, etc.
- The object ontology is queried for the object's feature decomposition and generic shape information, which includes geometric shape elements such as surfaces and curves, and spatial relations between features and shape elements.

- Low-level edge vision features are further processed to obtain mid-level vision features, such as rectangles. These are matched to the geometric shape elements to identify a set of candidate object features.
- For each object in the object ontology, check if all of its required features exist, and whether all spatial relations between its features are satisfied. This groups a set of features and spatial relations into a new instance of that object class.
- Repeat using only the unassigned shape elements in the scene data, until all input elements have been assigned to some object.

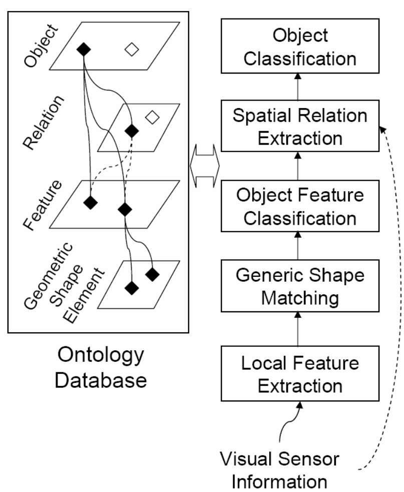


Fig. 1. Ontology-based object recognition scheme

A simplified subset of the ontology representation of a beam projector object is shown in Fig. 2.

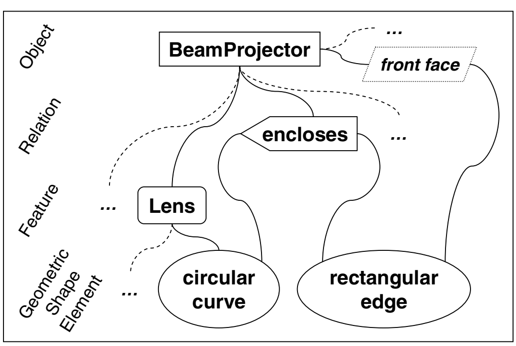
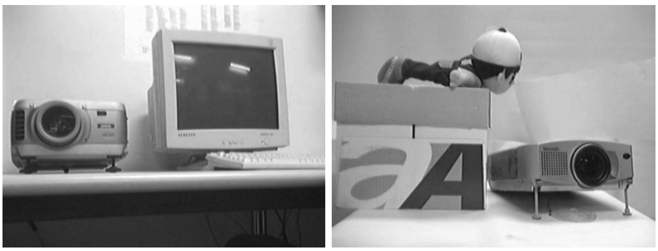


Fig. 2. Object decomposition of a beam projector (partial)

4 Experimental Results

To test this object recognition scheme, experiments were conducted on two kinds of beam projectors with different shapes and orientations, as shown in Fig. 3. First, edge information is extracted by using the canny edge detector, and it is further processed to generate the low-level image features such as connected line, arc, etc., as shown in Fig. 4. Circular edges are identified as shown in Fig. 5, and these are matched to the *circular curve* geometric shape element for a beam projector's lens feature. Similarly, rectangular edges are identified as shown in Fig. 6, and these are matched to the *rectangular edge* shape element for one face of a beam projector. In addition, the *encloses* spatial relation is checked, which rejects all rectangular edges that do not enclose any circular curve. The result of successful recognition of both beam projector objects is shown in Fig. 7.



(a) Beam Projector-A (b) Beam Projector-B

Fig. 3. Two test images

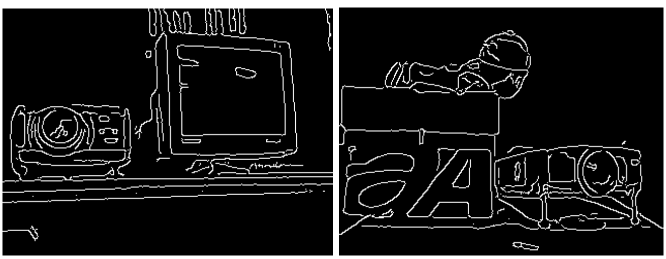


Fig. 4. Edge extraction and low-level image features

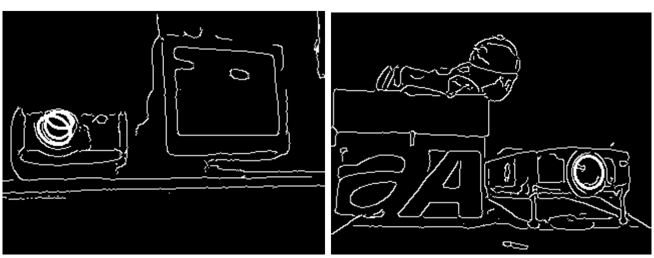


Fig. 5. Matching the *circular curve* geometric shape element

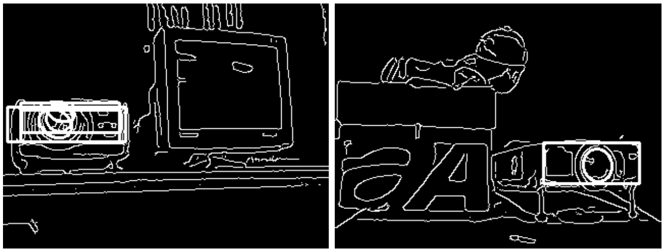


Fig. 6. Matching the *rectangular edge* geometric shape elementand the *encloses* relation

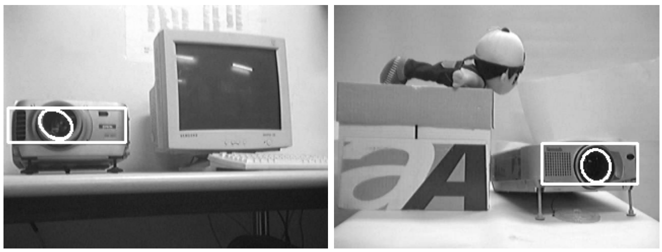


Fig. 7. Successful recognition of both beam projectors

The performance of proposed ontology based recognition method is tested by comparing the model-based recognition system (Matrox MIL 7.5). As shown in Fig. 8 to Fig. 10, the several projector images are captured that have different size and orientation. The left images show the results of proposed recognition and the right images

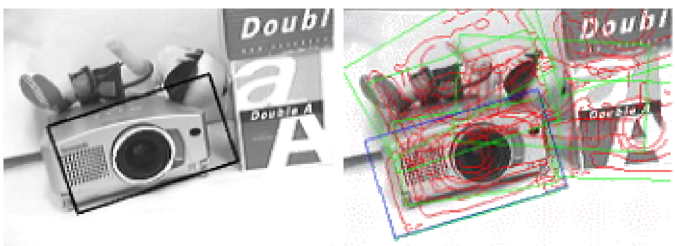


Fig. 8. ontology (left) vs. model-based (right) recognition example

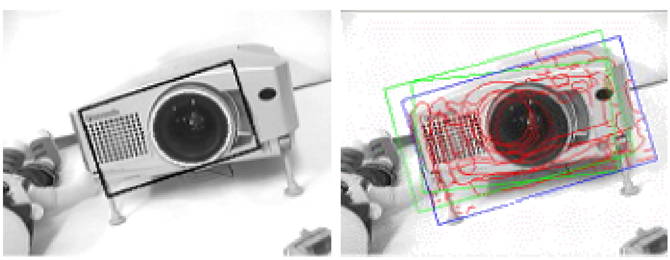


Fig. 9. ontology (left) vs. model-based (right) recognition example

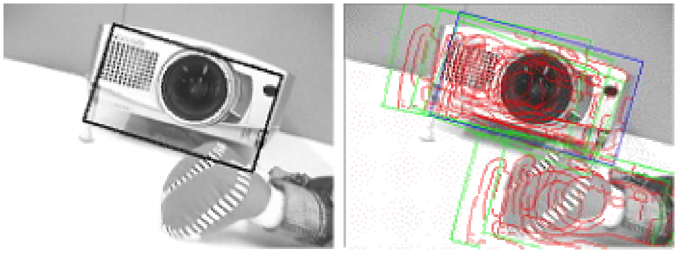


Fig. 10. ontology (left) vs. model-based (right) recognition example

Table 1. The comparison of maximum probability of presence of beam projector

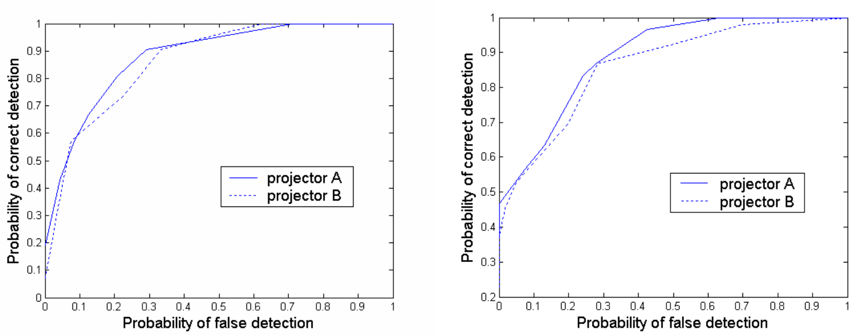
Max. Probability	Fig. 8.	Fig. 9.	Fig. 10.
Model based	24.82%	29.30%	33.17%
Ontology based	58.33%	60.36%	53.37%

show the results of Matrox MIL 7.5. The Matrox MIL 7.5 shows the several matched results that the probability of recognition results exceeds the certain threshold level.

Table 1 shows the maximum probability of recognizing projector-B using the model-based recognition method and proposed method. In most cases, the proposed method shows better results and average performance of recognition result also shows better result.

The receiver operating characteristic (ROC) curves for the beam projector detectors are shown in Fig. 11. More than 60 test sample images are used in this experiment. The result of model-based beam projector detection method with Matrox MIL 7.5 is shown in Fig. 11-(b). The curve A in this figure shows the projector-A detection result with the model of projector-A. The curve B shows the projector-B detection result with the model of projector-A.

The curve A and B of Fig. 11-(a). show the projector-A and projector-B detection results with the proposed object recognition scheme.



(a) Proposed detection scheme (b) Model-based detection scheme

Fig. 11. The ROC curve comparison

5 Conclusion

We have developed a new object recognition scheme that combines generic shape information extraction and reasoning with function ontology for effective object recognition. The results of our research show that ontology based object recognition concept can be used to create a powerful object recognition scheme.

As a future work, we will include 3-D features such as surface patches, surface normal vectors for enhanced objection performance with more complex objects.

Acknowledgments

This work is supported by KIST and Intelligent Robotics Development Program, one of the 21st Century Frontier R&D Programs funded by the Korea Ministry of Commerce, Industry and Energy.

References

1. David A. Forsyth and J. Ponce: Computer Vision: A Modern Approach. in Prentice Hall, August, 2002.
2. R. Chin and C. Dyer: Model-based recognition in robot vision. in ACM Computing Surveys, vol. 18, no. 1, pp. 67-108. 1986.
3. L. Stark and K. Bowyer: Generic recognition through qualitative reasoning about 3-D shape and object function. in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 251-256. 1991.
4. Stark, L., and Bowyer, K.: Function-Based Generic Recognition for Multiple Object Categories. in Computer Vision, Graphics and Image Processing, Vol. 59, No. 1, pp. 1–21, Jan. 1994.
5. Neumann, B., and Möller, R.: On Scene Interpretation with Description Logics. in FBI-B-257/04 (Technical Report), Fachbereich Informatik, Universität Hamburg, 2004.
6. Haarslev, V., and Möller, R.: RACER System Description. in Proc. Int'l Joint Conf. on Automated Reasoning (IJCAR 2001), LNAI Vol. 2083, Springer, 2001, pp. 701–705.
7. Protégé, http://protege.stanford.edu.
8. Kirschman, C. F. and Fadel, G. M.: Classifying Functions for Mechanical Design. in Journal of Mechanical Design, Vol. 120, pp. 475–482, 1998.
9. Stone, R. B., and Wood, K. L.: Development of a Functional Basis for Design. in Proc. ASME Conf. on Design Theory and Methodology, Las Vegas, 1999.
10. Haudrum, J.: Creating the Basis for Process Selection in the Design Stage. in Ph.D. Thesis, Institute of Manufacturing Engineering, Technical University of Denmark, 1994.
11. D. Lowe: Distinctive image features from scale-invariant keypoints. in International Journal of Computer Vision, vol. 60, pp. 91-110, 2004.
12. S. Se, D. Lowe, and I. Little: Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks. in International Journal of Robotics Research, vol. 21, pp. 735-758, 2002.
13. S. Se, D. Lowe, and J. Little: Global localization using distinctive visual features. in International Conference on Intelligent Robots and Systems, pp. 226-231, 2002.