AN INEXACT GRAPH MATCHING ALGORITHM FOR THREE-DIMENSIONAL OBJECT RECOGNITION

HANS PETER AMANN AND HEINZ HUGLI
University of Neuchâtel, Institute for Microtechnology
Rue Tivoli 28, CH-2003 Neuchâtel-Serrières, Switzerland

ABSTRACT

This paper describes in detail the inexact matching algorithm which has been applied with success to 3-D object representations in a 3-D object recognition system for range data. The algorithm combines in a promising way several approaches proposed in the last couple of years: an extension of the search strategy to cope with inexact matching of attributed sub-graphs, the determination of a measure for attributal and structural dissimilarities, and the global transformation fitting, features which are used to perform efficient search-tree pruning. The algorithm was tested successfully in a series of experiments involving scenes with single and multiple objects.

1. Introduction

Object recognition is a key problem in machine vision. A central point of the recognition process is the match between model and observation at a given level of representation. We consider a system for the recognition of polyhedral objects that uses attributed relational graphs to represent the object. Since the object representations can be partial, inexact and even erroneous, the object recognition problem becomes an inexact (sub-)graph matching problem.

The complete recognition system that encompasses the various stages from data acquisition over object recognition to the verification of the resulting hypothesesis, has been described elsewhere[1,2,3]. This paper treats with more detail the algorithm for the efficient inexact matching of attributed relational graphs using a constrained search.

In the following, first we give a short overview over related work (§2). Next, we describe briefly the data acquisition system, preprocessing, matching and representation we apply to the acquired data matching (§3).

The inexact matching algorithm is presented in paragraph 4. First, we introduce the matching problem. Then, we present the graph matching as a search problem and propose to solve it by a constrained-based search strategy. We discuss the cost function which guides the search and the constrains used in order to keep the algorithm within reasonable limits of complexity. Paragraph 5 shows experimental results obtained with the described algorithm, paragraph 6 the conclusions and an outview on further subjects of interest.

2. Related work

The presented work is based on ideas gathered from many sources, the most important being presented in the following. For the graph matching approach, Eshra

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1992
(Proc. of Workshop on Structural and Syntactic Pattern Recognition, Aug. 26-28, Bern, Switzerland)

and Fu's backtrack-based inexact matching scheme using small subgraphs called Basic Attributed Relational Graphs (BARGs) has been a valuable base[4,5]. Shapiro presented a very helpful paper on the use of metrics for the determination of dissimilarities and their combination into a global cost measure[15]. Oshima and Shirai[13] showed the interesting feature of "root matches" that enables determination of interesting starting points in the search space. Furthermore, they proposed to introduce reliability measurements for graph attributes. Finally, the important concept of object rigidity - discussed in paragraph 4 - as well as the related mathematical tools that showed useful for its convenient implementation as a global search tree heuristic, are due to Faugeras and Hebert[8].

Recently, Flynn[16] and Fan[6] published on 3D object recognition approaches very similar to the one of the present paper, Grimson[18] with a larger view. While the main ideas of Flynn and Fan are almost identical, some differences can be found in the correspondence search strategy and in the determination of the "cost" criteria. Flynn focuses in particular on the order of application of the tree pruning criterion, Fan uses a two-stage depth-first tree search, extracting promising base-matches which then are extended further. Through their completeness, Grimson[18] and Fan[6] can serve as valuable references for the domain of 3D Object recognition, the former for a general approach, for the later in particular for approaches based on surface descriptions.

3. Data acquisition, data representation and preprocessing

For the present work, we decided to limit us to polyhedral objects that can be represented using planar surface patches. The corresponding high-level representation is an attributed relational graph called Planar Face Representation Graph PFRG where each node stands for a surface (attributes "surface area" and "orientation of the surface normal") and each arc for a border between two adjacent surfaces (attribute "border length").

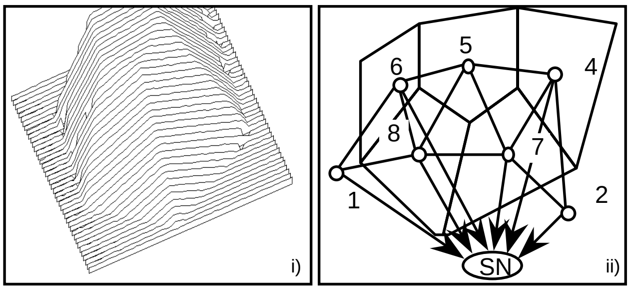


Figure 1: i) scanned test object pyramid and ii) corresponding Planar Face Representation Graph PFRG

By nature, acquired scene data will be neither complete (occlusions, …) nor precise (limited resolution, …) and the range image segmentation process can result in a cluttered PFRG representation. To deal with this kind of data, some useful intermediate level processing is helpful in order to remove elements which do not contain useful information and to merge elements which have been separated accidentally[2]. But also the subsequent matching algorithm must be conceived such that it is able to handle

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1992
(Proc. of Workshop on Structural and Syntactic Pattern Recognition, Aug. 26-28, Bern, Switzerland)

inexact information of this kind.

An important step preceding the correspondence search is the acquisition and preprocessing of raw input data. For our experiences, we used a laser range finder[14] delivering a dense set of points in the space. A subsequent range image segmentation process[3,11,12] delivers a more compact representation at a higher level of abstraction, i.e. in the form of regions and borders.

4. Algorithm for the inexact match of subgraphs

In the present paragraph, we treat inexact high level matching. First, we introduce the basic notations and define the problem. Then, we present the graph matching as a tree search problem, discuss the expansion rule of the search and the corresponding pruning constraints. Finally, we introduce the cost function which reflects the dissimilarity of the matched PFRGs.

Table 1: Notations

NRm	node m of reference graph R
NSn	node n of scene graph S
$A(N_{Sm}, N_{Sn})$	non-directed arc N_{Sm}, N_{Sn}
$A(N_{Rm},N_{Rn})$	non-directed arc N_{Rm}, N_{Rn}
$G_S = (N_S, A_S)$	scene graph (PFRG)
$G_R= (N_S, A_S)$	reference graph (PFRG)
$N_S = \{,N_{Sn},\}$	set of scene nodes

$N_R = \{,N_{Rm},\}$	set of reference nodes
$A_S = \{,A(N_{Sm}, N_{Sn}), \}$	set of scene arcs
$A_R =\{ , A(N_{Rm}, N_{Rn}), \}$	set of reference arcs
$M_i = \{ , (N_{Sn}, N_{Rm}), \}$	set of matched nodes
$H_i = (M_i, C_i, T_i)$	hypothesis on a match M_i
C_i	cost for match M_i
T_i	transformation for match M_i

Problem definition

Given two PFRGs, one for the scene G_S and one for the candidate reference G_R, find an ordered set of matching hypothesis $\{H_i\}$ describing partial matches between the two graphs. A matching hypothesis is a triple composed of the set of matched nodes $M_i = \{\ldots, (N_{Sn},N_{Rm}), \ldots\}$, the matching cost C_i and the geometrical transformation T_i the reference object has to undergo so that it maps the scene data the best (figure 2). The ordering is according to the cost C_i.

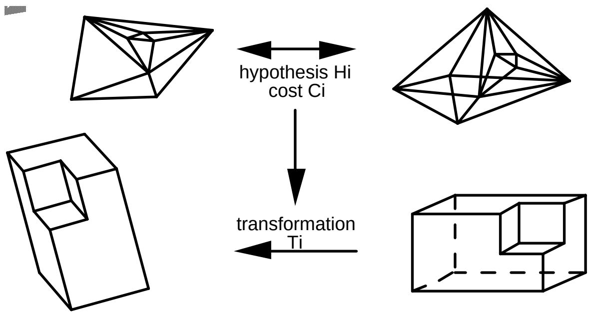


Figure 2: the matching process

A state space lattice for the search of hypotheses

The correspondence search can be seen as a search in a state space lattice: each state S_i corresponds to a match M_i and each extension of S_i to a new son state S_{i+1} to the extension of M_i to M_{i+1}. The new match extends the previous by the addition of a new

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1992
(Proc. of Workshop on Structural and Syntactic Pattern Recognition, Aug. 26-28, Bern, Switzerland)

node-to-node correspondence: $M_i = M_{i-1} \approx (N_{Sn}, N_{Rm})$.

As a new node of the scene is used at each state transition, the search space can be divided in hierarchical levels. Level 0 has the states with no match, level 1 has the states with one node-match and level i has the all states with i node-matches. As a full expansion of the search would lead to an unreasonable computational complexity, we apply a constrained search with following features:

- "Best-first" search: at each step of the search, the partial hypothesis $\{H_i\}$ are ordered according to their cost. The hypothesis with the minimal cost is expanded first to its son states.
- Pruning: not all expansions of a given state lead to valid hypothesis. If we reject a hypothesis, we also reject its further expansions. This is the essence of pruning.
- Expansion of a state is performed according to the graph structure of G_S and G_R; each hypothesis is expanded by a new node-match which brings together two not yet matched nodes and preserves the connectivity of G_S and G_R.
- Termination: heuristics are of various types and can be selected according to the application. Typical termination conditions are i) termination on found solution (solution is good enough), ii) termination on time or memory limits (best solutions so far).

Connectivity constraint

The state expansion is governed by the connectivity of both G_S and G_R. To formalize this statement, we consider figure 3 and define core sets C_{Si} and C_{Rj}, as the sets which contain the nodes already matched, for the scene as well as for the reference. The node candidates for further matches must satisfy two conditions. First, they are yet unmatched. Secondly, two of them form a candidate node-match if and only if they are respectively connected to core nodes which form a node-match. In other words, a node-match of core elements (N_{Sm},N_{Rn}) expands to all possible node-matches (N_{Si},N_{Rj}) such that both N_{Si} and N_{Rj} are yet unmatched and respectively connected N_{Sm} and N_{Rn}.

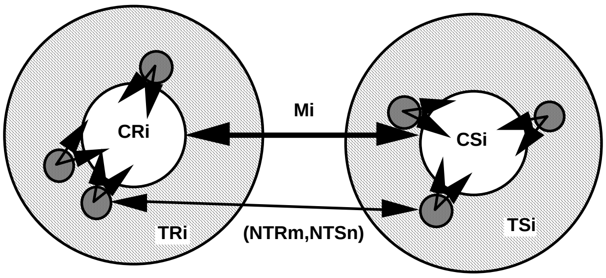
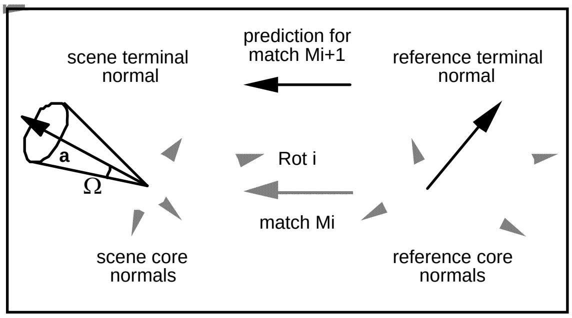


Figure 3 connectivity criterion (white: set of core nodes, dotted: candidate nodes) for further matches

Rigidity & visibility constraints

For pruning, Hebert and Faugeras[7] propose to use an interesting feature: the rigidity of objects. Given a certain number of matches between elements of the scene and the reference, one can find a hypothesis on the transformation T_i which maps the reference object to the scene.

In our case, as the PFRG carries no information on the surface position, we consider the transformation to be a pure rotation R_i. In principle, the rotation transforms the vector normal to the surface of node N_{Rm} to the one of N_{Sn} for each node-match (N_{Rm},N_{Sn}).



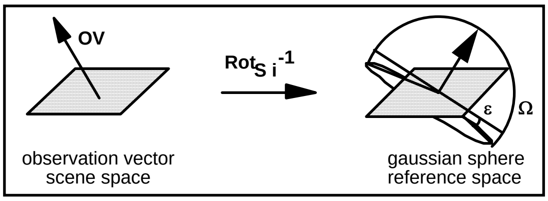


Figure 4: rigidity condition *Figure 5: observation space*

The criterion to accept further matches is that they can be done using a very similar rotation R_i' (identity can not be requested due to the non-idealities of the data acquisition process).

$$\left|\angle\left(normal(NT_{Sm}),\ normal(\ R_i\ (\ NT_{Rn}\ )\ )\right)\right| \leq TH_{\Omega} \tag{1}$$

Furthermore, the increase in the "rotation fitting" error due to the extension of match M_{k-1} to M_k must be bellow a given threshold:

$$\Delta e_{rotfit}(S_{i-1},\ S_i)\ \leq\ TH_{rotfit} \tag{2}$$

Note that for an empty match, the rotation R_i is not determined, for a match of a single node-match, it becomes underdetermined and for two and more node-matches generally overdetermined: the rotation R_i must be determined using a "best-fit" method (here least squared error fitting). Since R has to be recomputed after each successful extension of a partial match to a bigger one, its computation must be considered with the necessary attention. A very elegant approach enabling the computation of the transformations T_i in an non-iterative way while taking into account planes, lines and points has been shown in reference 8. The idea is based on the use of quaternions[10] and permits to determine in an extremely elegant way least square error fitted transformations taking even weighting factors into account. The idea has been developed in detail in two references[1,7].

The visibility condition can also be used. The observation of the scene space is usually mono-ocular and can be represented by an observation vector. The transformation of this vector back to the reference space enables to restrict the set of matchable planar faces of the reference object: their orientations must not be more then 90°±ε different from the transformed observation vector.

Attribute dissimilarity based constraints

Despite the fact that they are less reliable due to possible occlusions and shadowing, the attributes "surface area" (nodes) and "border length" (arcs) are also used for pruning in the two following ways.

A first constraint limits the increase of the respective cumulative dissimilarities for the transition from S_{i-1} to S_i. It is limited by a fixed threshold:

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1993

$$\textit{node dissimilarity} \qquad \Delta diss_{node}(S_{i-1},S_i) \leq TH_{node} \tag{3}$$

$$\textit{arc dissimilarity} \qquad \Delta diss_{arcatt}(S_{i-1},S_i) \leq TH_{arcatt} \tag{4}$$

A second constraint limits the increase of the global matching cost:

$$\textit{global cost} \qquad \Delta cost(S_{i-1}, S_i) \leq TH_{cost} \tag{5}$$

Determination of dissimilarities

We discuss now with more details the computation of the different dissimilarities which contribute to the global matching cost. The rotation fitting error $e_{rotfit}(S_q)$ is defined as the (weighted) sum of the squared Euclidean distances between the rotated reference normal vectors and the corresponding scene normal vectors, accumulated over all the q scene node / reference node correspondences (N_{Sn}, N_{Rn}) of a hypothesis H_q.

$$e_{rotfit}(S_q) = \frac{1}{2\sum_{root}^{q}\gamma_n} \sum_{root}^{q}\gamma_n|normal(N_{Sn})-R_q(normal(N_{Rn}))|^2) \tag{6}$$

$$\Delta e_{rotfit}(S_{q-1}, S_q) = e_{rotfit}(S_{q-1}) - e_{rotfit}(S_q) \tag{7}$$

The dissimilarity of surface area is a relative measure bounded to a range [0..1], enabling hence its integration in the global "cost" expression:

$$\Delta diss_{area}(S_{i-1}, S_i) =$$

$$diss_{area}(N_{Rm}, N_{Sn}) = \frac{|area(N_{Rm}) - area(N_{Sn})|}{max(area(N_{Rm}), area(N_{Sn}))} \tag{8}$$

In an identical way, the dissimilarity of the "border length" attributes is handled, the cumulative value is computed over all arcs matched for a given extension from hypothesis H_{q-1} to hypothesis H_q:

$$diss_{border}(A_{Rm},A_{Sn}) = \frac{|border(A_{Rm})-border(A_{Sn})|}{max(border(A_{Rm}),border(A_{Sn}))} \tag{9}$$

$$\Delta diss_{border}(S_i,S_k) = \sum diss_{border}(A_{Rm}, A_{Sn})\ ) \tag{10}$$

Structural dissimilarities

With the hypothesis that the reference graph contains reliable information, we apply some isomorphic error correction while extending the scene graph/reference graph matches, the number of deleted respectively added arcs being considered in the final cost computation.

Combining dissimilarities

In order to determine an overall cost function, dissimilarities determined for each type of attribute and for structural differences are combined so that search-tree states of different levels, representing matches of different sizes, can be compared: weighting must depend on the number of elements involved.

$$C(S_k) = \frac{\alpha\Sigma d_{area}(Sk) + \varepsilon e_{rotfit}(Sk)}{number\ of\ nodes(Sk)} + \frac{\beta\Sigma d_{border}(Sk) + \gamma A_{ins}(Sk) + \delta A_{del}(Sk)}{A_{match}(Sk) + A_{ins}(Sk) + A_{del}(Sk)} \quad (11)$$

The cost function is composed of two parts: the first related to the nodes, the second to the arcs. The first part is the weighted sum of accumulated dissimilarities of the area of the matched surface patches $\Sigma d_{area}(S_k)$ plus the mean squared rotation fitting error $e_{rotfit}(S_k)$ divided by the number of matched nodes. The second part consists of an arc-dissimilarity measure: the sum of arc attribute dissimilarities $\Sigma d_{border}(S_k)$, increased by a dissimilarity involving the number of inserted (A_{ins}) and deleted arcs (A_{del}) divided by the number of arcs involved (matched, inserted and deleted). α, β, γ, δ and ε are weighting parameters.

5. Experiments and results

Weighting parameters determination

In a first series of tests, we determined adequate weighting factors for the overall cost. The resulting set of parameters is $\{\alpha=0.05, \beta=0.04, \gamma=0.01, \delta=0.01, \varepsilon=0.9\}$. This same set was them used for all subsequent experiments.

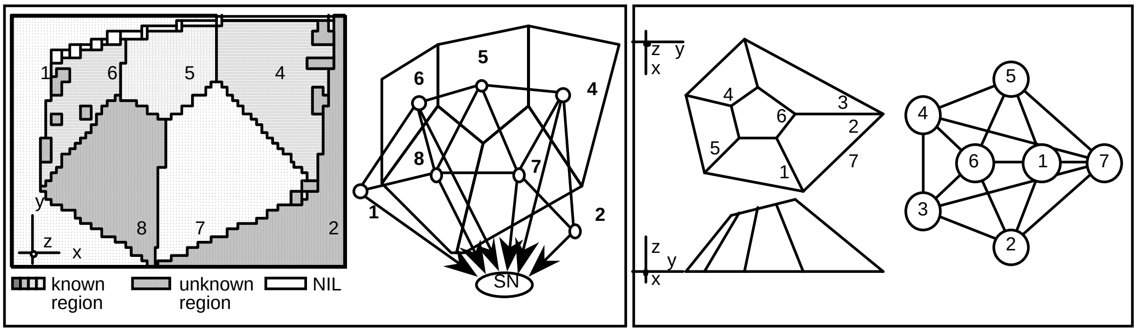


Figure 6: determination of weighting parameter (test object "P" left, reference object "Pyramid2" right)

Trying to interpret the influence of the weighting parameter on the search strategy, we see that the cost is dominated by the rotating fitting error (ε=0.9). The search takes therefore full advantage of the robustness of the rigidity constraint.

For the parts of the cost which are weighted by α, β, γ, δ and ε, they are discriminatory in cases showing fitting errors close to zero and in the case of a search state with a single node-match where rotating fitting does not apply.

Object recognition: single scene object / multiple reference objects

The second set of experiments deals with the 3D recognition problem where a single object in the scene is matched with multiple reference objects. Figure 7 shows a simple example: scene object P (b) was tested against the two reference objects, CUBE1 (a) and PYRAMID2 (c).

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1993

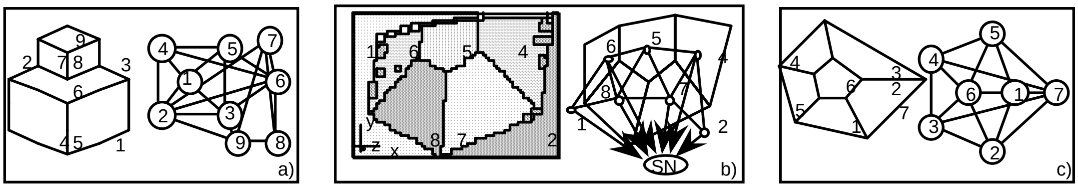


Figure 7: single-object recognition: reference objects CUBE1 (a) and PYRAMID2 (c); test object P (b)

For the first ten hypotheses found, the match between scene object P and reference object PYRAMID2 always delivers a lower global cost than the match with the second reference object CUBE1 : the algorithm behaves as expected.

Object recognition and object separation

The third set of experiments deals with the 3D recognition problem of occluded objects. The PFRG of a scene CP (Figure 8), showing a cube that partially occludes a pyramid, was tested vs. the reference PFRGs CUBE1 and PYRAMID2 (Figure 6, right). The algorithm described before, together with a simple test of hypothesis compatibility, we obtain the desired result, represented in figure 8 next page.

The "best" set of compatible solutions contains the correct match between reference object CUBE1 and the part of scene CP representing the cube and the reference object PYRAMID2 with the part of the scene representing the pyramid. The "best" set of compatible solutions contains the correct match between reference object CUBE1 and the part of scene CP representing the cube and the reference object PYRAMID2 with the part of the scene representing the cube and the reference object PYRAMID2 with the part of the scene representing the pyramid. The "best" set of compatible solutions contains the correct match between reference object CUBE1 and the part of scene CP representing the cube and the reference object PYRAMID2 with the part of the scene representing the pyramid.

Nevertheless, with our choice of high-level representation and correspondence-search criteria, the probability of erroneous correspondences increases, in particular for the partially occluded objects. Match results must be completed, either by subsequent (low-level) verification or by the use of richer data in both the high-level representation and the correspondence-search criteria.

H.P. Amann & H. Hügli, "An inexact graph matching algorithm for three-dimensional object recognition", in "Advances in structural and syntactic pattern recognition",Ed. H. Bunke, Series on Machine Perception and Artificial Intelligence - Vol. 5, pp 421-430, World Scientific Publ. Co., 1993

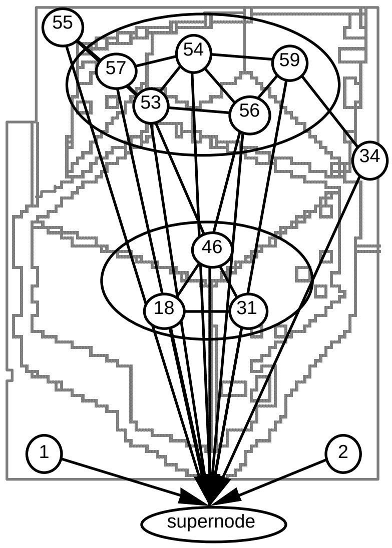


Figure 8: two compatible solutions: scene CP vs. references CUBE1 and PYRAMID2

6. Conclusions

In the present paper, we have described the algorithm used for the solution to the inexact high-level matching problem in the context of 3D-object recognition. Given two high-level object representations in the form of attributed relational graphs, we extract the most promising (partial) matches so that we can set up hypotheses on the correspondence between reference and scene objects. The match of the two subparts of reference and scene objects is characterised by a dissimilarity measure that accounts for matching cost and a transformation - the geometrical rotation the reference object has to undergo so that it fits the scene data best.

We treat the subject of graph matching as a tree-search problem. Since full search is computational complex, we use best first search together with constraints that enable efficient pruning of the search tree. The most important constraint we use is rigidity. Despite good results obtained with it, additional tree-pruning constraints based on the remaining graph attributes had to be added in order to deal with cases where the rigidity condition is not sufficient, e.g., for man-made objects where perpendicular and parallel faces are common.

A final match, consisting of a set of 1-1 node matches, is characterised by an overall similarity measure as well as by the geometrical transformation that best maps a reference object to the scene. Various experiments show sound results obtained with the described algorithm, i) for object recognition with single-object scenes and ii) for object recognition and object separation for multiple-object scenes, including objects which occlude themselves mutually.

Finally, the presented algorithm should not be considered a complete solution. Rather it should be considered as a building block for a more elaborate recognition system that also considers, among other things, subsequent verification of the generated hypotheses. A more elaborate system would also include object representations using additional information, e.g., descriptions of edges and surface shape.

7. Acknowledgements

We wish to thank Gilbert Maître who implemented the data-acquisition device as well as the range image segmentation algorithms used in the present experiments.

This work has been supported by the Swiss National Foundation for Scientific Research under project numbers FN2265, FN2000-5.138 and FN 20-25572.

References

1. H. P. Amann, 3D Object Recognition Based on Surface Representations, PhD thesis, University of Neuchâtel (Switzerland), Oct 1990
2. H. P. Amann, An algorithm for the inexact matching of high level 3D polyhedric representations, Proceedings SPIE Conf on Intelligent Robotics Syst., Boston, Nov 1991, Vol 1609, pp 143-140
3. P. J. Besl, R. C. Jain, Segmentation through variable-order surface fitting, IEEE PAMI Vol 10 No 2, March 88, pp167-192
4. M. A. Eshera, K. S. Fu, A similarity measure between Attributed Relational Graphs for Image Analysis, 7th Conf Pattern Recognition, Montreal 1984, pp75-77
5. M. A. Eshera, K. S. Fu, An Image Understanding System Using Attributed Symbolic Representation and Inexact Graph Matching, IEEE PAMI Vol 8 No 5, Sept 1986, pages
6. T.J. Fan, Describing and recognizing 3-D Objects Using Surface Properties, Springer, New York, 1990
7. O. Faugeras, M. Hebert, A 3D recognition and positioning algorithm using geometrical matching between primitive surfaces, Proc. 8th Int. Joint Conf. AI, Los Altos, Aug. 1983, pp 996-1002
8. O. D. Faugeras, M. Hebert, The Representation, Recognition and Positioning of 3D-Shapes from Range Data, in Three-Dimensional Machine Vision, ed T. Kanade, Kluwer Academic Press, 1987, pp 301-357
9. B.K. Horn, Extended Gaussian Images, Proc. IEEE, Vol 72 No 12, Dec 1984, pp 1656-1678
10. I. L. Kantor, Hypercomplex numbers: An elementary introduction to algebras, Springer, New York, 1989
11. G. Maître, H. Hugli, F. Tièche, H.P. Amann, Range Image Segmentation based on function approximation, Proceedings ISPRS90, Zürich, Switzerland, Sept 1990, Vol 1395, pp 275-282
12. G. Maître, H. Hugli, Range Image Segmentation by controlled-continuity spline approximation for parallel computing, Proc. SPIE Conf. on Intelligent Robotic Syst., Boston, Nov. 1991, Vol 1610, pp 238-249
13. M. Oshima, Y. Shirai, An object recognition system using three-dimensional information, in Three-Dimensional Machine Vision, ed T. Kanade, Kluwer Academic Press, 1987, pp355-397
14. H. Hügli & G. Maître, "3D by structured light: implementation and evaluation of a vision system for small parts", in Optical 3-D Measurement Techniques, Ed. A. Gruen & H. Kahmen, Wien, Wichmann, 1989, pp. 468-477
15. L. Shapiro, R. M. Haralick, "A Metric for Comparing Relational Descriptions", IEEE PAMI Vol 7, No 1, Jan 1985
16. P.J. Flynn, A.K. Jain, BONSAI: 3D Object Recognition Using Constrained Search, Journal IEEE, 1990, pp 263-267
17. E. Gmür, H. Bunke, 3D Object Recognition Based on Subgraph Matching in Polynomial Time, Workshop on Syntactical and Structural Pattern Recognition, World Scientific Publishing Co, 1989
18. W.E.L. Grimson, T. Lozano-Perez, D.P. Huttenlocher, Object Recognition by computer: the role of geometric constraints, The MIT Press, Cambridge MA, USA, 1990