NEW APPROACH FOR THE DETERMINATION OF CORRESPONDING SURFACES IN SEGMENTED 3-D RANGE IMAGES

Daniel Fischer, Peter Kohlhepp

Forschungszentrum Karlsruhe, Institut für Angewandte Informatik
Postfach 3640, D-76021 Karlsruhe, Germany
Email: kohlhepp@iai.fzk.de, Phone: (49)-7247-825762, Fax: (49)-7247-825786

Abstract

Establishing feature correspondences in a fast and reliable way is a key task to many robotic vision applications, such as the recognition and localization of industrial objects, or the integration of partial sensor views in order to reconstruct a global geometry model of the environment. This paper presents a new solution of the correspondence problem for attributed surface graph descriptions derived from range images. It combines Neuro-Fuzzy based similarity measures at the single surface and relational level with tree search at the description level. The use of Neuro-Fuzzy rule bases makes the determination of surface and relational similarity robust with respect to measurement and segmentation noise, and tolerant of occlusion as well. Another novel aspect is the handling of relational costs. Conventional tree searching algorithms usually depend on neighborhood relations found between adjacent surface patches by the segmentation algorithm. Such relational information is sparse, and its detection a fragile and strongly view point dependent process. Our approach does not rely solely on these neighborhood relations. It rather exploits included angles, distance measures and similar relations between arbitrary, not necessarily adjacent patches. They are independent of the viewpoint and also largely occlusion invariant.

The implemented system has been tested so far at about 20 segmented images, mainly from several publicly available range image databases. In the paper we present preliminary performance results for a small number of images showing the influence of certain similarity thresholds and weight factors present in the system. By suitably adapting these experimental parameters, we may explicitly trade the desired efficiency of tree search against the discrimination accuracy required by the application.
1. Introduction

An autonomous mobile robot or vehicle operating in a previously unknown industrial environment has to solve a number of vision-related tasks: objects have to be recognized and gripped, and obstacles to be classified for a collision free path planning. A global environment map (world model) needs to be generated during an exploration trip, and the system needs to update its own position based on that map. Unpredictable light conditions, dust and shadows restrict the usage of video images in unknown environments. Range images collected by active optical sensors, e.g. laser radar sensors, are less dependent on these influences. 2-D or 3-D range images (point clouds) serve as a basis for the creation of geometric representations [Bes88]. 2-D representations, such as evidence grids, are sufficient if the robot’s sole purpose is to navigate on flat ground, but they are insufficient if the environment is more complex and objects have to be inspected or manipulated. In this case 3-D descriptions seem more appropriate despite their increased time for data capture and data processing.

The proposed approach is part of an experimental system for the reconstruction, recognition, and localization of objects called SOMBRERO (Surface Oriented Model Building for the REconstruction and Recognition of Objects). This system segments 3-D range images by a split and merge algorithm ([Koh95]) and produces 3-D surface representations which are more compact and meaningful than range images themselves. Some loss of information and additional segmentation noise have to be accepted and to be managed by the processing units performing object recognition and object reconstruction. A problem frequently arising in the context of recognition and reconstruction is to determine corresponding surfaces in different surface representations which may be given in different coordinate systems (“non-calibrated” case).

2. Surface Representation

The surface representation adopted by SOMBRERO consists of a list of surfaces and a list of neighborhood relations between these surfaces. A surface description $S$ consists of several attributes such as the surface type, area, center of gravity, direction vector, curvature histogram [Bes88], and the boundary which is approximated by a polygon. Relations are characterized by the two surface identifiers, the type of relation (occluding or occluded jump edge, convex or concave crease), the associated edge and the angle (only in case of convex or concave creases). An example of a surface representation is given in figure 1.

---

*Fig. 1: Surface Representation of a scene „chest and barrel“ (Photo of the scene: upper left window, 3-D range image: lower left window, wire frame surface representation: main window)*
3. Determining Corresponding Surfaces

The task of finding corresponding surfaces in two non-calibrated, partly overlapping, noisy and incomplete surface representations is addressed in this paper. The proposed method calculates surface similarities \( ssim \) between surfaces and relation similarities \( relsim \) between pairs of surfaces using Neuro-Fuzzy technology ([DHR96], [Kos92]). All the surface and relation attributes used for this purpose are position invariant. Both similarity measures are combined with an ordinary tree search algorithm [Fan90], [Gri91]. Figure 2 gives an overview of the correspondence finding process.

![Fig. 2: Determination of corresponding surfaces](image)

The output of the Neuro-Fuzzy rule base for \( ssim \) also serves as an input for the pre-selection of possible corresponding surfaces, and for the cost calculation. Relation similarities \( relsim \) are determined by the second Neuro-Fuzzy rule base and are used only for the cost calculation. Both rule bases were designed with the help of a commercially available fuzzy development environment, fuzzyTech® [NN97]. The Neuro-Fuzzy rule bases can be designed, adapted, and optimized using expert knowledge and/or training data. Small input differences have little effect on the similarity values. Therefore \( ssim \) and \( relsim \) are very robust and stable with respect to different kinds of noise, and occlusion can also be handled properly.

For constructing a correspondence between two surface descriptions, a tree search algorithm is started and a tree is created. Each node in the tree represents an assignment of a surface \( F_{1i} \) in description \( S_1 \) to a surface \( F_{2j} \) in description \( S_2 \). In order to restrict the number of nodes and to improve on the performance of the algorithm two conditions must be fulfilled for each new node to be created:

- The similarity between the two assigned surfaces must exceed a threshold value \( \alpha_{cut} \). If the similarity of a possible corresponding surface pair falls below \( \alpha_{cut} \) it is marked by the pre-selection task and will not be considered in the tree search.
- The cost for this node according to the cost calculation task must fall below a selectable threshold \( cost_{max} \).

Finally, a list of corresponding surfaces is generated from the path maximizing the total sum of surface and relation similarities. The determination of \( ssim \) and \( relsim \) is described in more detail in the following sections 3.1 and 3.2, while the pre-selection and cost calculation is found in section 3.3.

3.1 Determination of the surface similarity \( ssim \)

The similarity \( ssim \) \((F_{1i}, F_{2j})\) between two surfaces \( F_{1i} \) and \( F_{2j} \) is determined by a Neuro-Fuzzy rule base as indicated in fig. 3.

The following subset of the attributes introduced in chapter 2 was chosen to represent a surface: the area \( a \), the compactness \( c \), the curvature histogram \( ch \), and the degree of occlusion \( oc \) (table 1). These attributes are first scaled (normalized) to form a four-dimensional feature vector \( (a_D, c_D, ch_D, oc_D) \) (table 2). The scaled feature vector serves as the input vector to a Neuro-Fuzzy rule base, and the resulting output \( ssim \in [0,1] \) represents the similarity between the two surfaces \( F_{1i} \) and \( F_{2j} \).

A Neuro-Fuzzy rule base consists of the fuzzification, fuzzy rule, and defuzzification parts and can be seen as a non-linear function approximation transforming a n-dimensional input vector to a m-dimensional output vector. Each input variable is represented by a linguistic variable consisting of linguistic terms. The fuzzification part maps a real input value (e.g. \( a_c=0.3 \)) to a degree of membership to a linguistic term, such as occlusion is „high“, „medium“, or „low“. Rules have an IF (aggregation) and a THEN (conclusion) part, for example: IF [input1=high AND input2=low] THEN [output1=medium]. For the linguistic output term(s) of a rule, their degree of fulfillment depends on the degree to which the input terms are fulfilled. Real output values are finally generated in the defuzzification part from the output terms as received from different rules. Expert knowledge...
and/or training data is used for an optimal design of the membership functions, and for attaching proper weights to different rules. For details on fuzzy-logic and neuro-fuzzy rule bases see [DHR96] and [Kos92].

![Surface feature vectors](image)

**Fig. 3: Determination of surface similarity ssim**

<table>
<thead>
<tr>
<th>Surface feature</th>
<th>Description</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{ij} )</td>
<td>Area of surface ( F_{ij} ) in ( S_i )</td>
<td>( a &gt; 0 )</td>
</tr>
<tr>
<td>( c_{ij} )</td>
<td>Compactness of surface ( F_{ij} ) in ( S_i )</td>
<td>( c \in [0,1] )</td>
</tr>
<tr>
<td>( ch_{ij} )</td>
<td>Curvature histogram of surface ( F_{ij} ) in ( S_i ) contains the distribution of the eight possible curvature types (possible sign combinations of Mean and Gaussian curvature)</td>
<td>( 8 \sum_{m=1}^8 ch[m] = 1 ), ( ch[m] \in [0,1] )</td>
</tr>
<tr>
<td>( oc_{ij} )</td>
<td>Degree of occlusion</td>
<td>( oc \in [0,1] )</td>
</tr>
</tbody>
</table>

**Table 1: Surface features of a surface patch \( F_{ij} \) (equivalent for \( F_{2j} \))**

<table>
<thead>
<tr>
<th>Scaled surface feature</th>
<th>Description</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_D )</td>
<td>Min of ( a_{ij}, a_{2j} )</td>
<td>( a_D \in [0,1] )</td>
</tr>
<tr>
<td>( c_D )</td>
<td>Absolute difference between the compactness of ( F_{ij} ) and ( F_{2j} )</td>
<td>( c_D \in [0,1] )</td>
</tr>
<tr>
<td>( ch_D )</td>
<td>Norm of the curvature histogram difference vector of ( F_{ij} ) and ( F_{2j} )</td>
<td>( ch_D \in [0,2] )</td>
</tr>
<tr>
<td>( oc_D )</td>
<td>Minimum of both degrees of occlusion</td>
<td>( oc_D \in [0,1] )</td>
</tr>
</tbody>
</table>

**Table 2: Scaled surface features used for calculating the surface similarity**

If the boundary of \( F_{ij} \) is completely explained by neighborhood relations, \( oc_{ij} \) is 1. Roughly, a surface patch boundary is explained at those edges of the boundary polygon where there exists a nearby edge of a neighborhood relation (concave or convex crease, or curvature edge) or an occluding relation to an adjacent surface patch. Notice the smoothing influence of the feature \( oc_D \) on the remaining features: the higher the degree of occlusion of either surface patch, the less will be the influence that differences of their areas, their boundary shapes, or their curvature distributions will have on their similarity.

### 3.2 Determination of the relation similarity \( relsim \)

Similar to 3.1, a relation similarity \( relsim \) between two pairs of surface patches is determined (see figure 4).

\[
relsim = relsim \left((F_{ij}, F_{jk}), (F_{2j}, F_{2k})\right)
\]

As table 3 points out, a relation between two surfaces is represented by a three-dimensional feature vector, including the angle between the surface normal vectors \( n_i \), the distances between the center points \( cp_i \), and the minimum degree of flatness \( flat \). A scaled difference vector (see table 4) is calculated the components of which are used as input values for a Neuro-Fuzzy rule base. The more similar both relation vectors are, the more closely the \( t \) value \( relsime \in [0..1] \) will approach one.
3.3 Tree search algorithm

3.3.1 Pre-selection

Pre-selection is initiated before the tree search algorithm is started. Each surface $F_{1i}$ in $S_1$ is compared with each surface $F_{2j}$ in $S_2$. If their similarity $ssim$ is greater than a selectable threshold $\alpha_{cut} \in [0,1]$ they could be corresponding. Otherwise, the algorithm will not consider them as possible nodes in the tree. Therefore,

$$ssim( F_{1i}, F_{2j} ) > \alpha_{cut}$$

3.3.2 Cost calculation

During the tree search a cost function calculates individual costs for each generated node, representing a pair of corresponding surfaces. The cost calculation module uses both similarity values to determine two costs values. Assume that $F_{1i}$ is assigned to $F_{2j}$ at some node $N$ in the search tree. First, the surface costs of $N$, $cost_{surf}(N)$, are calculated directly from the surface similarity value $ssim$ of the corresponding surfaces at $N$:

$$cost_{surf}(N) := 1 - ssim( F_{1i}, F_{2j} )$$

Second, the relational costs, $cost_{rel}(N)$, are calculated by considering the nodes above $N$ in the search path. If $N$ is immediately below the root of the tree these costs are set to zero. Otherwise, a relational cost arises from each node $M$ above $N$ in the search path. More specifically, assume that such a node $M$ represents the assignment of surface $F_{1k}$ to $F_{2l}$. Then the relational cost for this pair of nodes $(N, M)$ is:
The total relational costs for \( N \) are determined by the maximum of the relational costs for each pair of nodes \((N, N_p)\) where \( N_p \) is any node along the path of length \( l_N \) from \( N \) to the root, excluding \( N \) and the root itself:

\[
\text{cost}_{\text{rel}}(N) := \max_{p \in \{1 \ldots l_N-1\}} \text{cost}_{\text{rel}}(N, N_p)
\]

The total cost per node can be calculated by using weighted surface and relational costs:

\[
\text{cost}(N) := \frac{k_1 \cdot \text{cost}_{\text{surf}}(N) + k_2 \cdot \text{cost}_{\text{rel}}(N)}{k_1 + k_2} \quad (k_1 + k_2 = 1)
\]

As soon as the total cost exceeds a selectable threshold \((\text{cost}(N) > \text{cost}_{\text{max}})\) at some node \( N \) the tree expansion will be stopped there. Finally, we define as the quality of a search path \( C \) - representing a correspondence between two descriptions \( S_1, S_2 \) - simply its length \( l \) (the number of surface pairs), decreased by the sum of costs for all nodes on that path. The best path (correspondence) \( C^* \) is the one maximizing quality:

\[
C^* = \arg \max_{\{N_1, \ldots, N_l\}} \left\{ l - \sum_{p=1}^{l} \text{cost}(N_p) \mid \text{cost}(N_{\text{max}}) \leq \text{cost}_{\text{max}} \right\}
\]

Each node \( N_p \) in the path \( C \) represents a pair of corresponding surfaces \( F_{1,p1}, F_{2,p2} \) in \( S_1, S_2 \), respectively.

4 Experimental Results

4.1 The influence of pre-selection

A real segmented view of an object is shown in fig. 5a whilst the complete synthetic model of this object is displayed in fig. 5b. Figure 6 points out the influence of \( \alpha_{\text{cut}} \) on the number of nodes and paths in the tree and, accordingly, on the processing time. The higher the value of \( \alpha_{\text{cut}} \) is chosen, the smaller will be the tree and the faster the algorithm. Tree search identified the correct surface correspondences (0-5, 1-8, 2-6, 3-0, 4-9, 5-1, 6-4) for any \( \alpha_{\text{cut}} < 0.7 \), but smaller values of \( \alpha_{\text{cut}} \) may be needed in case of significant occlusion. The example points out that the realized algorithm is robust to different kinds of noise and can be speeded up by changing the \( \alpha_{\text{cut}} \) parameter.

4.2 Tree search and cost calculation

Four segmented surface representations (see fig. 7a-d) serve as test objects for the validation of the proposed approach. Only \( S_{\text{Model}} \) and \( S_4 \) represent the same object but, due to slightly different viewing directions, not all surfaces of \( S_{\text{Model}} \) are available in \( S_4 \). \( S_{\text{Model}} \) constitutes a wrong object hypothesis for both \( S_1 \) and \( S_3 \). Only the object in \( S_3 \) has curved surfaces (0 and 2), whereas the other objects are described by flat surfaces.
The following figures 8, 9, and 10 (left column) summarize the matching results obtained, using a fixed and uniform parameter setting ($\alpha_{cut} = 0.5$, $cost_{max} = 0.2$, weighting factors $k_1 = 0.33$ and $k_2 = 0.67$). In the right column of each figure, the single surface correspondences available by pre-selection are visualized.

**$S_1$ and $S_{Model}$**

Corresponding surfaces (best path) found:

- (2-1), cost = 0.07
- (3-4), cost = 0.15
- (4-0), cost = 0.12

Three pairs of corresponding surfaces are found due to the similar dimensions and surface features present in both objects.

Three other surfaces in $S_1$ could not be assigned.

**$S_2$ and $S_{Model}$**

Corresponding surfaces (best path) found:

- (0-0), cost = 0.11
- (1-1), cost = 0.06
- (2-2), cost = 0.08
- (3-5), cost = 0.18

Four surfaces of $S_2$ were assigned correctly to their corresponding surfaces in $S_{Model}$. Due to noise the surfaces similarity between surface 4 in $S_2$ and surface 4 in $S_{Model}$ was lower than $\alpha_{cut}$ and was not considered any more during the tree search.

While the average costs per node in fig. 8 and 9 differ only little ($\approx 0.11$), the path cost (see in section 3.3.2) is lower for the correct correspondence in fig. 9 than for the wrong solution in fig. 8, due to the higher number of nodes ($4 \leftrightarrow 3$). Also, the thresholded similarity matrix is much sparser (banded structure) for the correct solution in fig. 9 than in fig. 8. A similar observation can be made for the distribution of path costs in the search tree: their variance is much higher for the correct object hypothesis (fig. 9) than for a wrong one.
Corresponding surfaces (best path) found:

- none -

No corresponding surfaces were found. Only the flat surface 1 in $S_3$ may possibly have corresponding surfaces in $S_{Model}$. But even the single surface dissimilarity, neglecting relational costs, is in all cases higher than $\text{cost}_{\text{max}}$.

5. Conclusion

We have presented a new solution of the correspondence problem for attributed surface graph descriptions, combining Neuro-Fuzzy similarity (or equivalently: cost) measures with tree search. The relational cost does not only consider neighborhood relations explicitly present in the segmented descriptions, but also view point invariant distance measures of arbitrary (non-adjacent) surface patches. Therefore the relational cost at a node strongly grows with its level in the search tree, i.e. the number of patches already assigned on the search path. At the tree level, the crucial experimental parameters are: the pre-selection parameter $\alpha_{\text{cut}}$, the node cost threshold $\text{cost}_{\text{max}}$, and the relative weights for surface and relational costs, $k_1$ and $k_2$. At the surface level, we have the parameters of the membership functions and the weights of different Neuro-Fuzzy rules. The number of nodes in the search tree, the number of entries in the thresholded surface similarity matrix, and the maximum, the mean, and the variance of achieved path quality are all important performance measures of the search. By using these measures and embedding the search tree in a learning loop, the Neuro-Fuzzy rule bases can optimally adapt themselves to an unknown operating environment, e.g. to different degrees of occlusion or to different object shapes.

In this work, we consider fine-grained matching based on surface correspondences. For greater efficiency at large model data bases we can hierarchically extend the pre-selection scheme from surface to object similarities, using object features like volume, center, number of patches, or distribution of surface types.

Acknowledgement

The segmented descriptions shown in figures 5 and 7a-c were derived from range images of the ABW structured light Range Image Database at the University of South Florida [http://marathon.csee.usf.edu/range/icons/ABW.html]. The range image in fig. 7d was taken from the MSU/WSU Range Image Database at Washington State University [http://www.eecs.wsu.edu:80/IRL/RID/RID.html].

References


