# FREE-FORM 3D OBJECT RECOGNITION

Christian Schütz, Heinz Hügli
Institute of Microtechnology, University of Neuchatel
Tivoli 28, CH–2003 Neuchatel, Switzerland
schutz@imt.unine.ch, hugli@imt.unine.ch

This paper investigates a new approach to the recognition of 3D objects of arbitrary shape. The proposed solution follows the principle of model-based recognition using geometric 3D models and geometric matching. It is an alternative to the classical segmentation and primitive extraction approach and provides a perspective to escape its difficulties to deal with free-form shapes. Using the iterative closest point matching at the heart of the recognition, we propose means to extend its use to the recognition of 3D objects obtained from range data. Examples demonstrate the feasibility of this approach to free-form recognition.

## 1 Introduction

The recognition of free-form 3D objects – i.e. objects of arbitrary shape – is one of the major problems in computer vision. The classical segmentation and

Christian Schütz & Heinz Hügli, "Free-form 3D object recognition", ISPRS, Third Conference on Optical 3-D Measurement Techniques, Vienna, Oct. 1995

primitive extraction approach cannot easily be extended to deal with free-form objects. As pointed out by other authors [5, 11], it is not clear, how object parts should be defined and how reliable segmentation may work. Therefore, we opted for a recognition principle based on geometric matching. It works directly on the measured 3D coordinates of the object surface. Hence the recognition is independent on assumptions of object primitives.

According to this geometric approach, the comparison of the test and model object is performed with an iterative closest point matching algorithm (ICP) [1]. It guarantees convergence but successful matching is obtained only for a limited range of orientation and translation differences between the test and model object [10]. The paper discusses means to extend the ICP algorithm to the recognition of 3D objects obtained from range data.

Our investigations refer to a recognition configuration used for the pose estimation of 3D industrial objects in automatic assembly [6]. The objects are described by range data. The range images are acquired with a range finder working on the principle of space coding with projected stripe pattern and triangulation.

Other researchers [2, 3, 4, 8, 9, 11] have used similar algorithms to track objects and register surfaces. In most of these systems the initial transformation, from where the iterative algorithm is launched, is entered by an operator or estimated, for example when tracking objects. As far as we know, the ICP algorithm has not yet been used to perform object recognition.

In a previous paper, we investigated the usefulness of the ICP algorithm to recognize of free-form 2D shapes and a simple 3D object [10]. We present now further experiments which refer to complex 3D objects. Section 3 presents rules to build a set of starting configurations for the ICP algorithm which helps to overcome its limited convergence and makes even subpart matching feasible. In section 4, we propose a measure of matching error best suited for the decision process. Successful application to real objects is shown in section 5.

## 2 ICP algorithm

The heart of the recognition is the ICP algorithm that compares two surfaces. One surface is the test which consists in a specific view of an object and the other a model containing the complete information of an object surface. First, the algorithm searches for every point of a test set, the point of a model set with the smallest Euclidean distance. These pairs of closest points between two surfaces to be matched are then used to calculate the translation and rotation, which minimize the mean square distance or error. The test object is then translated and rotated by the resulting transformation. This procedure is applied several times until the error falls below a threshold or the number of iteration exceeds a predefined constant. Fig. 1 gives an overview of the basic working principle of ICP.

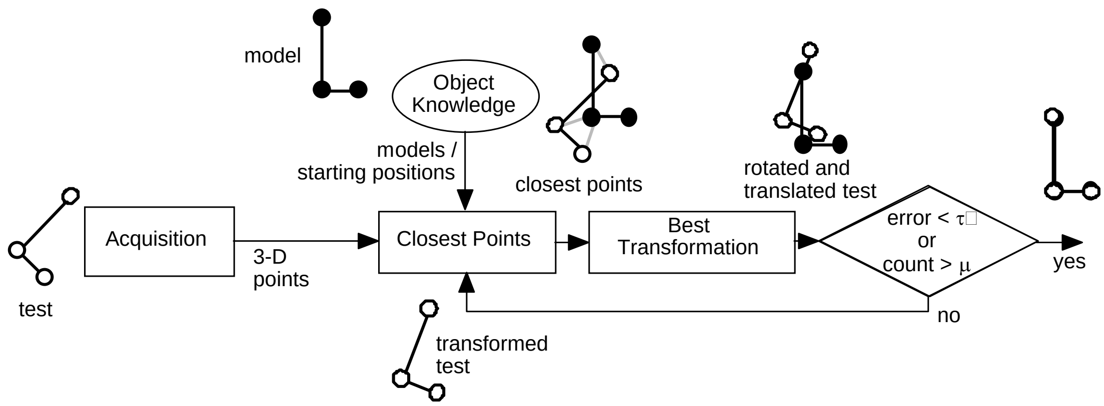


Fig. 1. Working principle of the iterative closest point algorithm (ICP)

The simplicity of this algorithm allows a fast implementation. Besides, it does not require any data pre-processing or local feature extraction, which makes it easily applicable to free-form objects. The algorithm converges after a few iterations to a solution which is not necessarily the best one.

## 3 Recognition algorithm

In a 3D object recognition system, the matching algorithm used to compare a test with the models in the database should allow the test to be any view of the corresponding model placed in any pose. Since the ICP algorithm converges only for a limited set of transformations between a test and its model, this algorithm has to be adapted to recognize reliably subparts of 3D objects. We propose a set of rules to select appropriate starting configurations, so that the ICP algorithm may converge towards the global minimum for at least one of them.

First, given a view axis defined by the camera position and the center of mass of the test, we place the model behind the test, as shown in Fig. 2. This placement ensures that the test surface not visible from the camera always faces the model and also excludes that test surfaces are compared with invisible model surfaces.

Second, we select N view points distributed uniformly on the sphere, circumscribing the model as drawn in Fig. 2. The model is now oriented so that every view point is lying once on the view axis. Furthermore, the model is rotated in M steps around the view axis for each of these configurations. The so defined N•M starting configurations for the ICP algorithm are selected in such a way, that their convergence zones cover the whole space around the model. This use of different starting configurations will ensure that the matching converges at least once towards a successful matching.

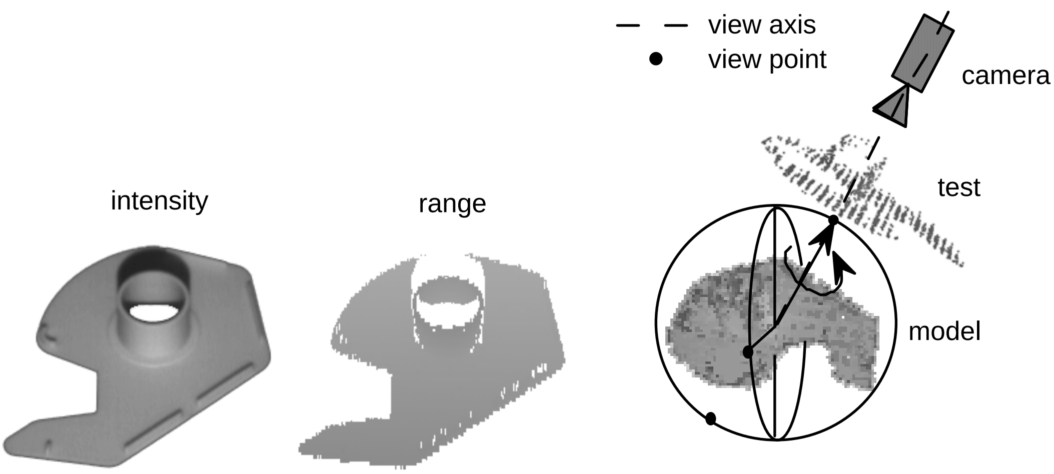


Fig. 2. Selection of starting configurations for a tape dispenser part

At a first glance the necessity of multiple starting configurations seems to introduce much overhead. But since the convergence zone is relatively large, the number of starting configurations can be kept low. Furthermore, the algorithm converges quickly and allows an estimation of the quality of a starting configuration after few iterations of ICP. This allows to prune the search tree. Starting configurations with low error are then used to do more iterations in order to find a complete matching. The definition of an adequate matching error measure will be presented now.

## 4 Matching error measure

A successful recognition needs a good error measurement which reflects the quality of the matching. In our recognition approach, we start the ICP algorithm from different configurations doing only few iterations to gain speed. Then only best configurations will be selected and used for further iterations to obtain final matching. This decision is based on the matching quality. Since ICP minimizes the mean square distance between the points of two objects, it seems obvious to use the minimized mean square distance as matching error measure. But experiments showed that this measure is insufficient to discriminate the quality of starting configurations. For example the mean square distances for the two cases shown in Fig. 3 differ by only 20%, which does not reflect the large difference between the two cases where the gray model object is in two completely different configurations.

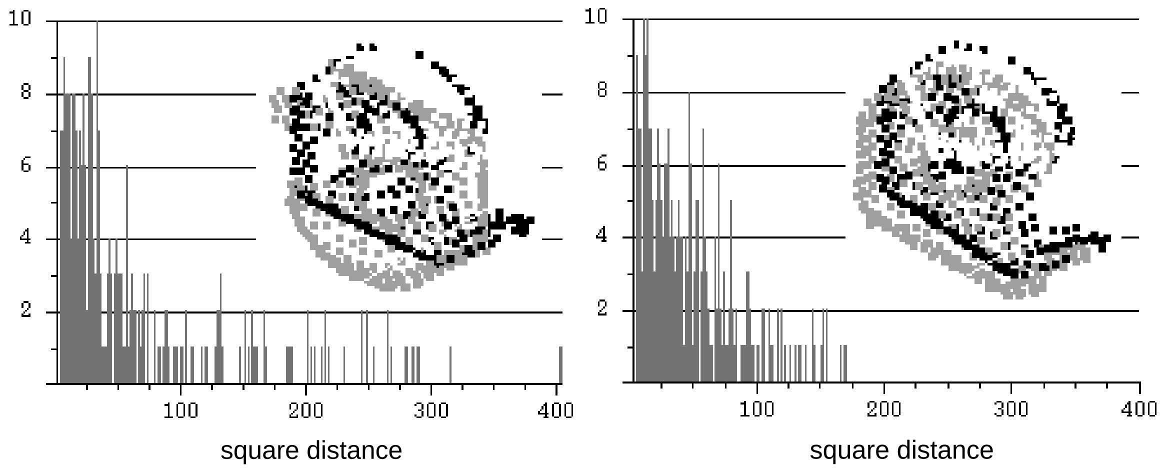


Fig. 3. Histogram of square distance between closest points

The square distance histograms corresponding to the two matches show that the distributions of the square distances differ even for similar means of square distance. In fact the deviation of the square distances is twice as large for the bad case (Fig. 3 left) compared to the good (Fig. 3 right). So, we finally define the matching error measure as the sum of the mean and the deviation of the square distances, similar as proposed previously by Zhang [11]. A low matching error will indicate cases with a promising match which are kept for further iterations.

## 5 Experimental results

As stated in the introduction, our experimental recognition system is used in an assembly environment. Our model database consists of the three parts of a tape dispenser. A typical scene is shown in Fig. 4, where the background has been removed for visualization convenience. Derived from the range image, the z-image gives the height above the working table for every image point. It is used to separate the objects, assumed not to touch each other. We simply set a threshold at the workplace height and obtain a binary image with the object zones (see Fig. 5 left). With this method we can extract objects even if the background has a complex texture, since texture does not appear in the z-image any more.

Every extracted test object is matched with all models. We start therefore the ICP algorithm for every model applying the rules defined in section 2. We select $N = 6$ view points uniformly distributed on the model sphere and do $M = 4$ rotations of 90 degrees for every view point aligned to the view axis. The ICP algorithm is expected to converge for one of these 24 configurations since its convergence zone is about 80 degrees [10].

The best model together with its two best configurations are selected. Several new configurations near to the two selected ones are used to start the ICP algorithm again calculating more iterations. Finally, we decide for the best configuration and do a final matching performing more than 20 iterations. The recognized model is then placed in the scene and projected as black dots on the intensity image (gray object) to verify the result, as presented in Fig. 5 right.

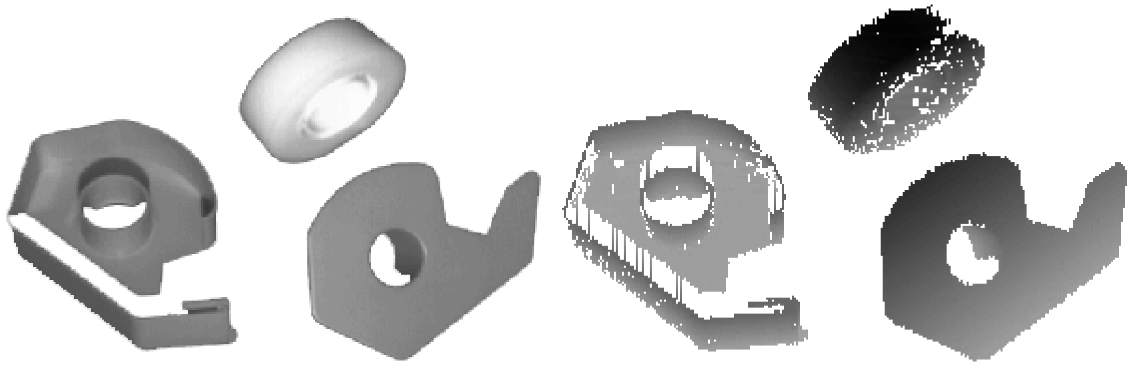

Fig. 4. Intensity and z-image of a typical robot workspace

Fig. 5. Thresholded z-image and recognized models superposed to the intensity image

We will now establish a comparison with the classical segmentation and primitive extraction approach. Typical for it, is the need to define primitives well suited to the objects to be recognized. If these primitives are very specific like planes, cylinders, superquadrics, they are not in a position to model free-form objects sufficiently well. If they are very general, they are usually not stable and lead to difficult segmentation.

To illustrate some of the difficulties found with the classical approach using general primitives, we present segmentation results for the same scene as in Fig. 4. We show in Fig. 6 the result of segmenting the range image into general primitives defined by the sole criterion of continuity [7]. Looking at the resulting segments, being the connected white regions of the presented image, we understand easily that they offer a very bad description of the tape parts. A symbolic matching based on such poor primitives will usually fail to recognize them.

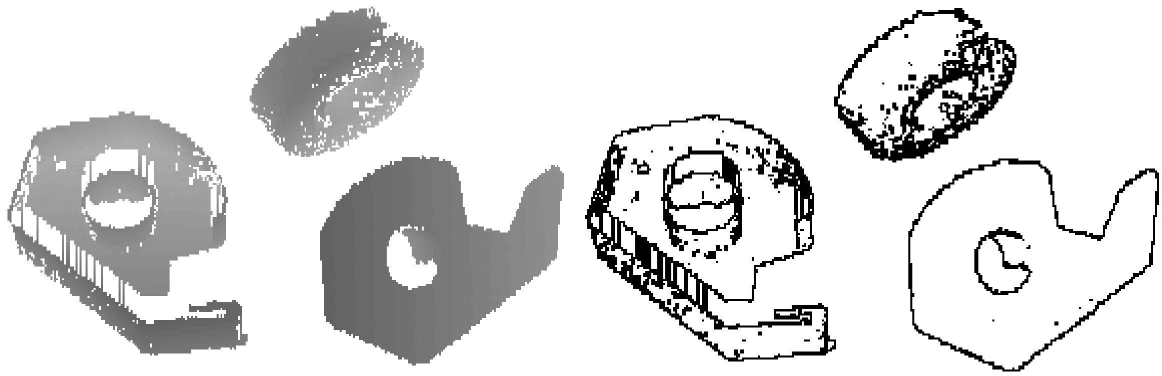

Fig. 6. Range image and its segmentation

Finally, these few examples are sufficient to show the feasibility and advantages of a geometric approach for modeling and recognizing free-form 3D objects.

## 6 Conclusions

The presented work is a contribution to a 3D object recognition approach which is easily applicable to free-form objects. The approach is based on geometric matching and applies to objects represented by sets of points or polygonal models. It differs from the classical approach which requires object segmentation and model construction in terms of geometric primitives.

Using the ICP algorithm at the heart of the recognition, we proposed a number of methods and techniques to extend its use to the recognition of 3D objects obtained from range images.

Since the ICP algorithm is not directly applicable to object recognition because of a limited convergence zone, we proposed a set of starting configurations using the knowledge of the camera position to overcome this handicap and to allow object or subpart matching.

We showed that the integration of the square distance deviation in the quality measure helps to extract reliably promising starting configurations for further observation.

Presented results show the successful recognition of the three parts of a tape dispenser and demonstrate the feasibility of the approach. Its intrinsic flexibility makes the approach applicable to any object form.

In the future, we will address possible limitations of the proposed approach which may arise when dealing with large number of models and having occlusion in the scene.

## 7 Acknowledgments

This research has been funded by the Swiss national Priority Program in Informatics, Knowledge-based systems, under project number 5003-344336, conducted in collaboration with the Institute of Microengineering at EPF Lausanne.

## 8 References

[1] P. Besl and N. McKay. "A method for registration of 3-D shapes," *IEEE Transactions On Pattern Analysis and Machine Intelligence,* Vol. 14, No. 2, pp. 239-256, 1992.

[2] Y. Chen and G. Medioni. "Object modelling by registration of multiple range images," *Image and Vision Computing,* Vol. 10, No. 3, pp. 145-155, 1992.

[3] J. Feldmar and N. Ayache. "Rigid and affine registration of smooth surfaces using differential properties," *Lecture Notes in Computer Science*, Vol. 801, No. 2, pp. 397-406, 1994.

[4] G. Godin and P. Boulanger. "Range image registration through viewpoint invariant computation of curvature," *From Pixels to Sequences, Proc. conf. ISPRS*, Zurich, Vol. 30, pp. 170-175, 1995.

[5] M. Hebert, J. Ponce, T. E. Boult, A. Gross and D. Forsyth. "Representation for computer vision," *Report of the NSF/ARPA workshop on 3D object*, CUNY Graduate School and University Center, 1994.

[6] H. Hügli, Ch. Schütz, Ch. Baur, E. Natonek. "Knowledge-based 3-D vision system," *Proceedings of Information Conference Module 2*, Knowledge based systems, Swiss National Science Fundation, pp. 46-53, 1994b.

[7] G. Maitre. *Segmentation et traitement preliminaire des images de profondeur*. These de doctorat a l'Universite de Neuchatel, 1994.

[8] G. Turk and M. Levoy. "Zippered polygon meshes from range images," *Computer Graphics*, ACM Siggraph 94 Proc, Orlando, Fla., pp. 311-318, 1994.

[9] D. A. Simon, M. Hebert and T. Kanade. "Techniques for Fast and Accurate Intra-Surgical Registration," *Journal of Image Guided Surgery*, Vol. 1, No. 1, pp.17-29, 1995.

[10] Ch. Schütz and H. Hügli. "Recognition of 3-D objects with a closest point matching algorithm," *From Pixels to Sequences,* Proc. conf. ISPRS, Zurich, Vol. 30, pp. 128-133, 1995.

[11] Z. Zhang. "Iterative point matching for registration of free-form curves and surfaces," *International Journal of Computer Vision*, Vol. 13, No.2, pp. 119-152, 1994.