Published in Photonics East (Proceedings of SPIE), 2409, issue 9, 66-74, 1996
which should be used for any reference to this work

HOW WELL PERFORMS FREE-FORM 3D OBJECT RECOGNITION FROM RANGE IMAGES?

Heinz Hügli, Christian Schütz

University of Neuchâtel, Institute for Microtechnology
Rue Breguet 2, CH-2000 Neuchâtel, Switzerland
hugli@imt.unine.ch, schutz@imt.unine.ch

ABSTRACT

This paper investigates the recognition performance of a geometric matching approach to the recognition of free-form objects obtained from range images. The heart of this approach is a closest point matching algorithm which, starting from an initial configuration of two rigid objects, iteratively finds their best correspondence. While the effective performance of this algorithm is known to depend largely on the chosen set of initial configurations, this paper investigates the quantitative nature of this dependence. In essence, we experimentally measure the range of successful configurations for a set of test objects and derive quantitative rules for the recognition strategy. These results show the conditions under which the closest point matching algorithm can be successfully applied to free-form 3D object recognition and help to design a reliable and cost-effective recognition system.

KEYWORDS: 3D vision, free-form object recognition, pose estimation, range imaging

1. INTRODUCTION

The recognition of free-form 3D objects is still a major problem in computer vision. Although successful recognition has been obtained with simple objects, it is difficult to extend the traditional segmentation and primitive extraction approach to deal with free-form objects. It is not clear, how object parts should be defined and how reliable segmentation may work, as pointed out by other authors [8, 13]. Therefore, a recognition principle based on geometric matching is used as presented in [12, 13]. This approach works directly on the 3D coordinates of the object surfaces as obtained from range data. Hence, the recognition is independent on assumptions related to object features. Furthermore, since a range finder senses the 3D object information, perspective and scaling effects are not present.

According to this geometric approach to object recognition, the comparison of the test and model object is performed with an iterative closest point matching algorithm (ICP) [2]. In this present work the test surface is represented by a range image whereas the model surfaces may be represented by sets of 3D points, polygons or other geometric primitives. For this reason, even CAD models can be easily used as models.

Although the ICP algorithm is proven to converge [2] the found solution is often only a local minimum. Convergence to the global minimum is not guaranteed generally. It is known to depend on a well chosen initial configuration of test and model [10].

In previous works using geometric matching methods, initial configurations are derived from contextual knowledge about the range finder setup [3, 4, 13] or found automatically by extracting some features as curvature or edges [1, 5, 6, 7]. In absence of a priori knowledge or robust features, the ICP algorithm is started with multiple different initial configurations [9, 12]. In any case, the recognition performance depends on the choice of the initial configurations and it is the first objective of this paper to investigate the quantitative nature of this dependence.

To do so, we propose a methodology for measuring and representing the recognition performance as a function of the initial configurations and apply it to the recognition of free-form 3D objects from range imaging, applying the setup discussed in [12]. In essence, the notion of successful initial configuration (SIC) is introduced, then the SIC-range of example objects is measured experimentally and displayed in the form of an original SIC-map. We present SIC-maps obtained from a series of experiments performed on injected plastic free-form objects.

The second objective aims at using the information from the SIC-maps in order to select a set of required initial configurations (RIC) to design a reliable and cost effective recognition system. Recognition experiments on a database of several object scenes assess the performance of the chosen RIC.

Section 2 introduces the ICP algorithm and its application to free-form object recognition. The SIC-range is investigated experimentally in section 3 for three toy objects. Section 4 develops the RIC estimation using the obtained SIC-maps. Finally in section 5, a recognition experiment using a database with different scenes representing toy objects gives recognition performances for several RIC setups.

2. OBJECT RECOGNITION WITH GEOMETRIC MATCHING

The geometric matching of two object surfaces is performed with the iterative closest point (ICP) algorithm [2]. Fig. 1 gives an overview of the basic principle of ICP. One surface is the test which consists in a specific view of the object; the other is a model containing the complete information of the object surface. The algorithm proceeds iteratively. First, it pairs every point of a test set with the closest point of a model set. 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 repeated until a predefined error falls below a threshold τ or the number of iterations exceeds a chosen constant N.

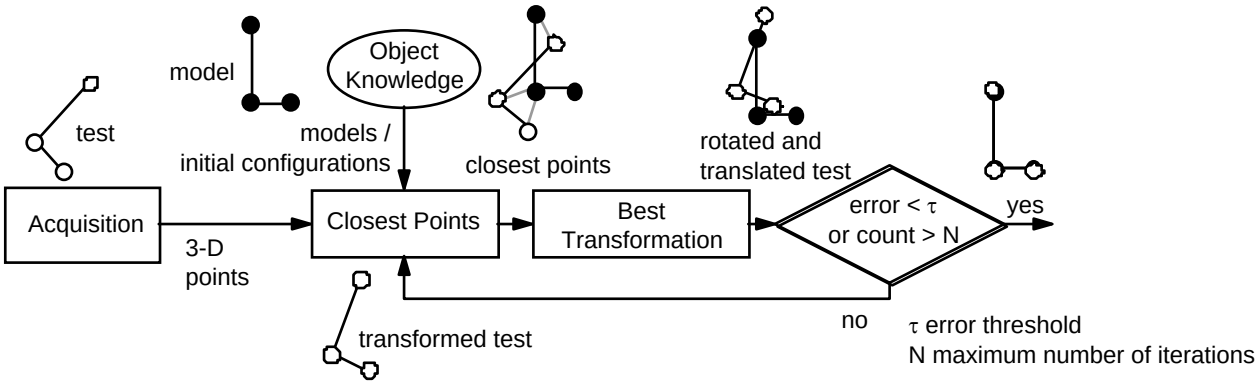

Fig. 1. Matching two objects with the iterative closest point algorithm (ICP)

The algorithm uses directly the sensed range data points and does not require any data pre-processing or local feature extraction, which makes it easily applicable to free-form objects and several types of object representations. Note that in our application the test represents a single view of the object and is obtained from a single range image while the model represents the complete object and is obtained by the registration of several range images representing different object views.

The algorithm converges after few iterations to a solution which is characterized by its error value. Successful matching occurs if the matching error, which is calculated as the sum of the mean and the deviation of the square coupling distances, is below a predefined threshold [11]. Otherwise, the algorithm is trapped in a local minimum.

Considering the space of possible initial configurations, successful matching is obtained only from a limited range of it, named successful initial configuration range or SIC-range. To practically measure the SIC-range, we first reduce the dimension of the 6D space of initial configurations to the 3D space proposed in [12] and corresponding to the setup in Fig. 2. The test object is placed on the view axis defined by the camera pointing towards the model object at a fixed distance between model and camera. The space of initial configurations can then be defined by the triple (ϕ, θ, ω) where ϕ and θ are respectively zenith and azimuth angles of the view axis in the model spherical reference system and ω designates the camera rotation angle around this axis. From now on, we name the pair (ϕ, θ) view point.

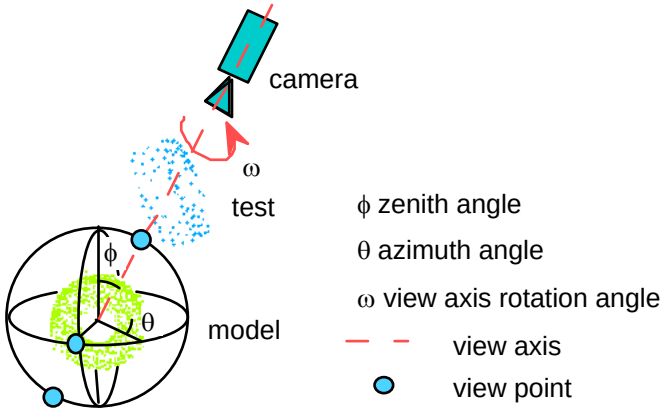

Fig. 2. Initial configuration setup of test and model

3. ASSESSMENT OF SUCCESSFUL INITIAL CONFIGURATIONS

We thus measure the SIC-range in the (ϕ, θ, ω) space. In order to do so, the (ϕ, θ, ω) parameter space of all possible initial configurations has to be inspected for successful matching. Successful configurations form the SIC-range. We estimate the SIC-range using the three free-form objects shown in Fig. 3. They represent injected plastic toys and are called duck, fish and swan.

Fig. 3. Free-form objects: duck, fish and swan

For each toy one view is selected and its SIC-range elaborated as follows. First, the test is directly placed at its pole initial configuration (see Fig. 4), which is its original orientation and a translation according to the setup presented in Fig. 2.

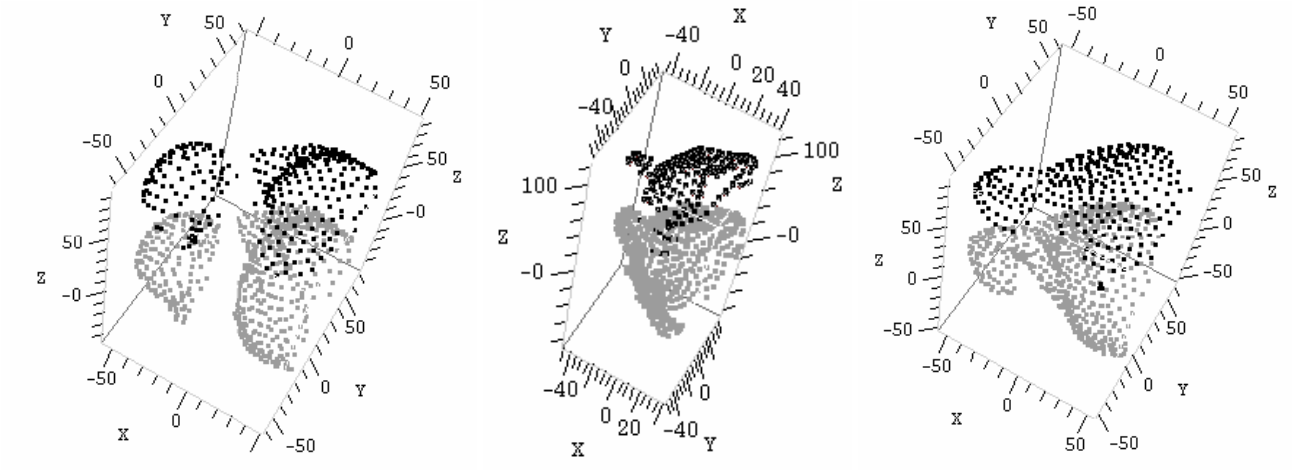

Fig. 4. Pole initial configuration of the three toy objects showing the model in gray and the test in black

Starting from this pole configuration, the complete (ϕ, θ, ω) parameter space is traversed in steps of 10 degrees varying azimuth angle θ from 0 to 360, zenith angle ϕ from 0 to 180 and view axis rotation angle ω from 0 to 360 degrees. This procedure orientates the test at different view points and then rotates it around the view axis. More formally, the

transformation, which orientates the test on the view point, rotates the test object by zenith angle ϕ around the rotation axis which is perpendicular to the great circle of azimuth angle θ as illustrated in Fig. 5. Then, follows the rotation ω around the view axis.

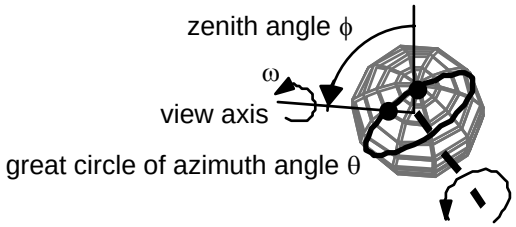

Fig. 5. Test transformation from pole to view point

For every initial configuration the ICP matching algorithm is launched, running enough (40) iterations to ensure convergence. An initial configuration is labeled as successful if the final matching error falls below a fixed threshold. An appropriate threshold is simply derived from the matching behavior of a number of cases.

The matching results are presented for every view point as a small circle where the black segment represents the SIC-range for ω (Fig. 6). The view points on the sphere are projected on a plane tangential to the pole for visualization in 2D giving rise to the SIC-map as defined in Fig. 6. View points having the same zenith angle lie on a circle around the pole. Note that view points from the lower hemisphere are omitted since there is no successful matching at all.

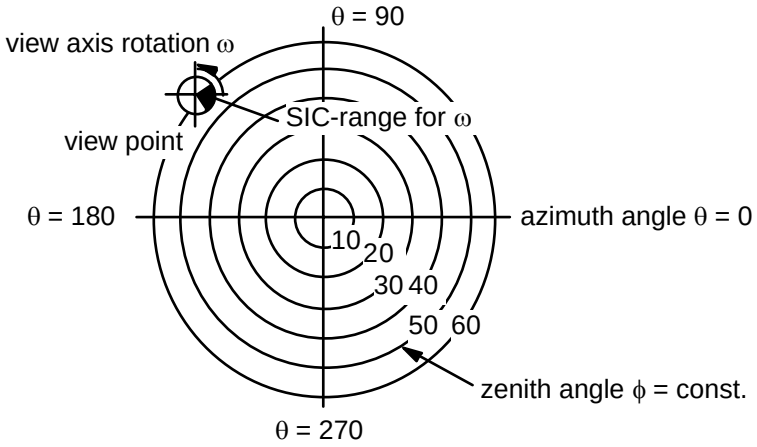

Fig. 6. The SIC-map illustrates the SIC-range in the (ϕ, θ, ω) space

Fig. 7 shows the SIC-maps of the three toy objects obtained experimentally according to the above explained method.

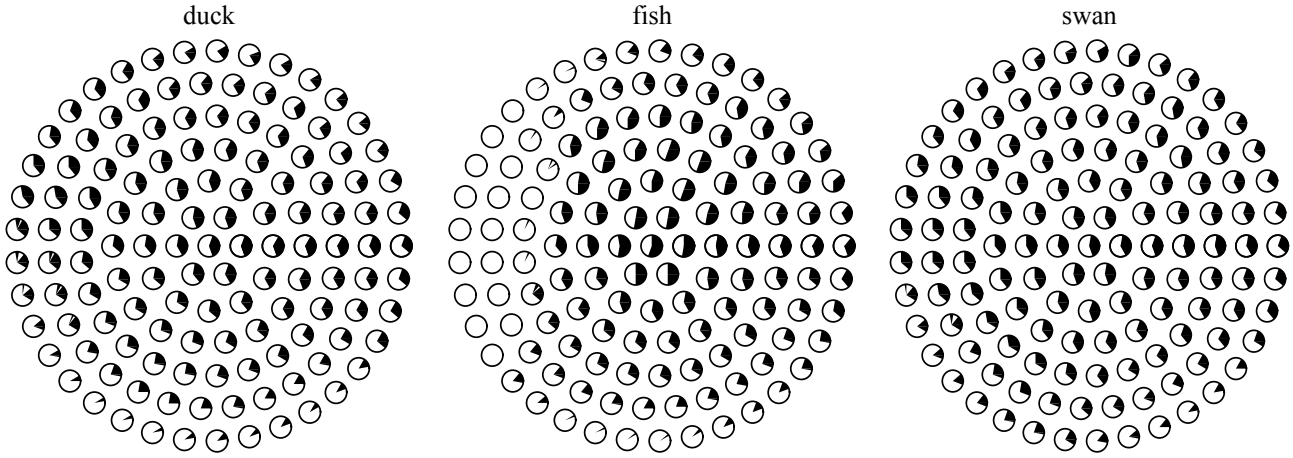

Fig. 7. SIC-map for three free-form objects (scale as in Fig. 6)

We observe that the SIC-range for ω decreases for view points with growing zenith angle. While this decrease is azimuth dependent for the fish it is nearly azimuth independent for the duck and swan.

In any case, it is interesting to analyze how the size of the SIC-range varies with increasing zenith angle. Let us define mean size $\Delta\overline{\omega}$ and minimum size $\Delta\omega_0$, the size of the black section of the SIC range for ω. We observe the decreasing functions reported in Fig. 8. A $\Delta\omega_0$ of 150 falls to 0 at a zenith angle ϕ_0 of 60 for duck and swan. For the fish, a $\Delta\omega_0$ of 200 falls to 0 at a zenith angle ϕ_0 of 40.

Considering these functions and defining now the simple rectangular range

$$R(\phi_0, \Delta\omega_0) = [0\ldots\phi_0, 0\ldots\Delta\omega_0], \quad \phi_0, \Delta\omega_0 \geq 0$$

an approximate and conservative modeling of above objects could be for example:

SIC-range of duck and swan R(40, 90)
SIC-range of fish R(30, 120)

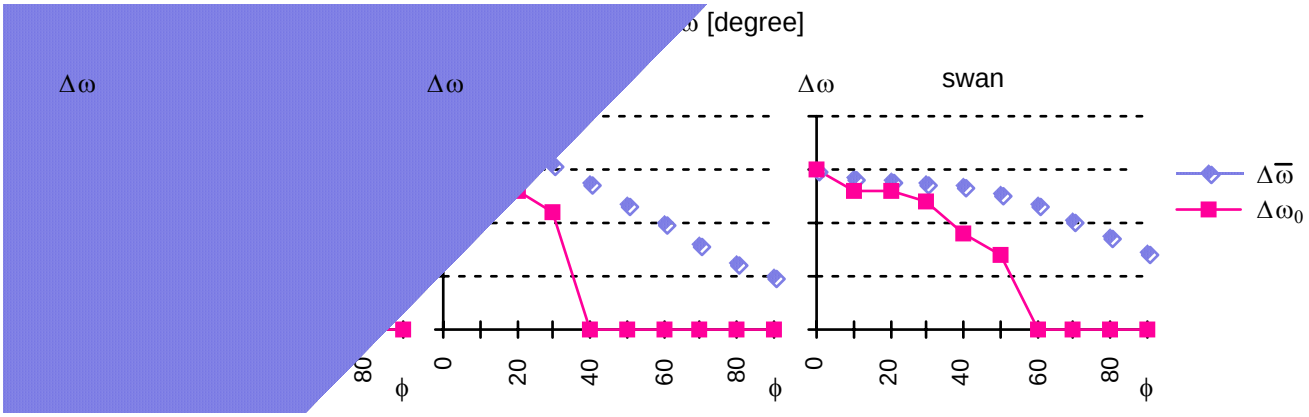

Fig. 8. Zenithal variation of the size of SIC-range for ω

In the next section we will use this object SIC-range modeling for the design of a reliable recognition system.

4. ASSESSMENT OF REQUIRED INITIAL CONFIGURATIONS

During recognition when no a priori knowledge is available, the matching algorithm is started with multiple different initial configurations. A simple choice of these required initial configurations (RIC) considers N_v view points distributed evenly on the observation sphere and N_r possible view axis rotations, giving rise to a number of $N_v \bullet N_r$ RIC [12]. The choice of N_v and N_r is discussed in this section. It is constrained by two factors. On one hand, the computational cost of recognition [9] is proportional to $N_v \bullet N_r$, calling for small values of N_v and N_r.

On the other hand, N_v and N_r must be sufficiently large to ensure reliable recognition. It means that at least one of the multiple initial configurations must belong to the SIC-range of the test. Considering an object with an hypothetical uniform SIC-range given by $R(\phi_0, \Delta\omega_0)$, above condition is satisfied if the covering of the space with the set of RIC is dense enough to fulfill the conditions of a maximum spacing in rotation angle

$$\frac{360}{N_r} < \Delta\omega_0$$

and a maximum spacing in zenith angle

$$\phi^*(N_v) < \phi_0$$

Function $\phi^*(N)$ is a basic function that relates the number of evenly distributed points on a sphere with their maximum spacing. More formally, it expresses the maximum angle distance of any point on the sphere to the set of N points distributed on it. Some values of $\phi^*(N)$, computed from a distribution obtained by a point repulsion algorithm, are reported in Fig. 9.

Fig. 9. Maximum angular distance $\phi^*(N)$ to N points evenly distributed on a sphere

Applying now above condition to the object models of previous section, we obtain example choices of N_v and N_r as follows:

duck and swan	R(40, 90)	$N_v = 10$, $N_r = 4$
fish	R(30, 120)	$N_v = 12$, $N_r = 3$

In a further step, we now vary the parameter $\Delta\omega_0$ and ϕ_0 of the SIC-modeling and minimize computational costs. On one hand, with above two conditions, the cost expresses as function of $\Delta\omega_0$ and ϕ_0 as

$$C(\phi_0, \Delta\omega_0) \geq \text{const} \cdot \frac{1}{\Delta\omega_0} \cdot \phi^{*-1}(\phi_0)$$

On the other hand, $\Delta\omega_0$ and ϕ_0 are bound by the object specific functions $\Delta\omega_0(\phi_0)$ of Fig. 8. Under this constraint, optimizing the costs leads to an optimal solution $(\phi_0, \Delta\omega_0)_{opt}$. Fig. 10 represents the cost function $C(\phi_0, \Delta\omega_0)$ together with the subset of values (black line) corresponding to the $\Delta\omega_0(\phi_0)$ dependence for the duck object. The optimum cost is obtained

for $(\phi_0, \Delta\omega_0)_{opt}$ = (50, 60) corresponding to N_v = 8, N_r = 6.

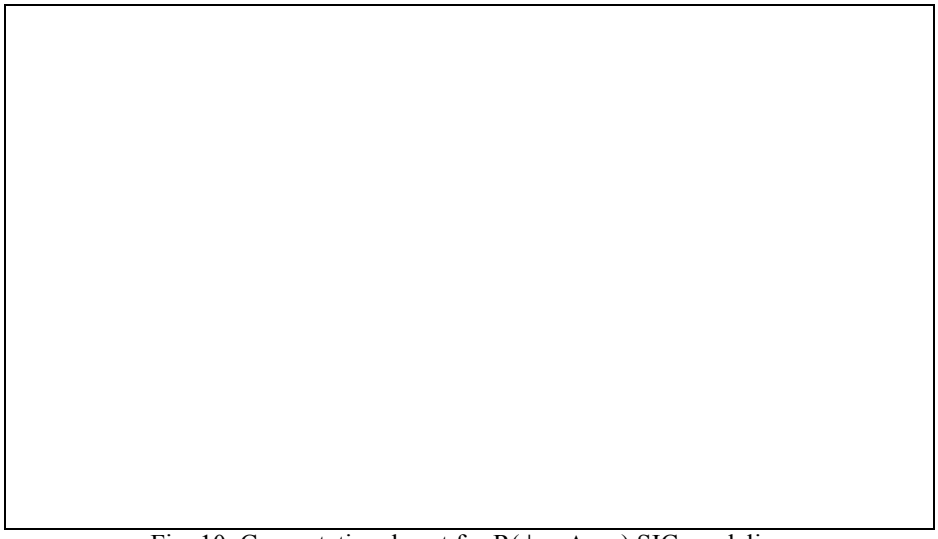

Fig. 10. Computational cost for R(ϕ_0, $\Delta\omega_0$) SIC-modeling

The $(\phi_0, \Delta\omega_0)_{opt}$ results for all objects are combined by selecting the smallest value of ϕ_0 which is 35 for the fish. Knowing ϕ_0, again the smallest of the $\Delta\omega_0(\phi_0)$ angles is chosen which equals to 90. Now all SIC-models are satisfied and we obtain a common set of $(\phi_0, \Delta\omega_0)_{opt}$ = (12, 4).

These numbers are only estimates since the SIC-range elaboration has been performed only with one view per object. The next section presents recognition experiments on a scene database in order to verify the performances of the designed setup.

5. RECOGNITION PERFORMANCE FOR DIFFERENT SETUPS

To verify the validity of the selection of a set of RIC characterized by numbers N_v and N_r according to the previous section, a series of recognition experiments is performed. It involves the three object models known as duck, fish and swan and ninety tests from scenes acquired with a range finder, each containing one of the toys in a arbitrary pose.

The recognition experiment is used to evaluate the recognition performance as a function of sets of RIC characterized by different numbers of view points N_v and view axis rotation steps N_r. A specific set is runned three times and the mean of the recognition results is calculated since the view point distribution is initialized randomly.

The decision of correct recognition is made by selecting a matching error threshold τ of 40 mm^2. This is an appropriate level to identify good correspondence since the number of sensed points is reduced by grouping local neighbors [12] and the resulting mean square distance is about 36 mm^2.

Under these conditions, there are no false classifications (confusions) at all. The result is thus either correct recognition or rejection. In the following, we report the percentage of correct recognition.

Fig. 11 compares RIC setups characterized by N_v/N_r values. Solid bars represent the recognition rate while the gray curve reports computational cost.

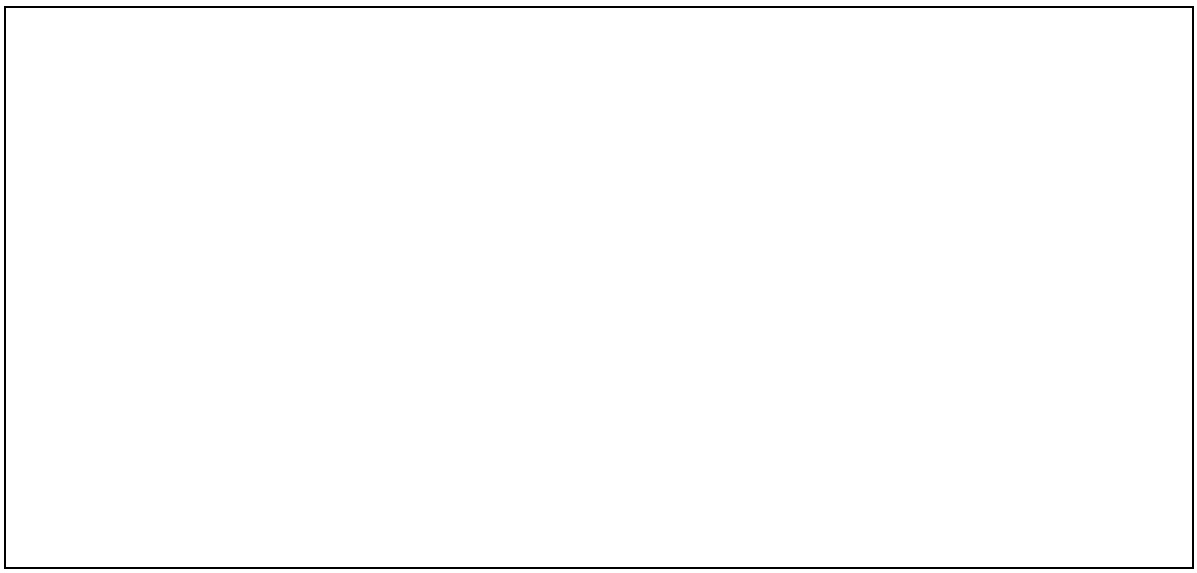

Fig. 11. Recognition performance for different RIC setups

Maximal recognition rate of about 97% is reached for a set of RIC N_v/N_r = 12/8. Almost the same recognition rate is obtained with less cost for N_v/N_r = 12/4.

By looking at the setups N_v/N_r = 8/8 and N_v/N_r = 12/4, we also observe that a higher number of view points needs a lower number of rotations to obtain similar performance.

The experimental results thus confirm the design rules of the previous section.

6. CONCLUSIONS

In this paper, we investigated the quantitative nature of the dependence of recognition performance with initial configurations used in the ICP matching algorithm. We introduced the notion of SIC-range and proposed a suitable representation named SIC-map to describe this dependencies. SIC-maps of three objects were presented and analyzed.

In a second part, we used the knowledge of object SIC-rnaps to design a recognition system with an optimal number of view points N_v and view axis rotations N_r for the recognition configuration (RIC).

By recognition experiments on a scene database, we measured computational costs and recognition performances which confirm the validity of the design rules.

Globally, the presented results confirm the successful application of range images and geometric matching to free-form 3D object recognition. The applied setup allows a recognition of objects lying in arbitrary pose.

7. ACKNOWLEDGMENTS

This research has been funded by the Swiss national Priority Program in Informatics, Knowledge-based systems, under project number 5003-344336 and by Swiss national science foundation under project number 2100-43530.

8. REFERENCES

1. R. Bergevin, D. Laurendeau and D. Poussart, “Estimating the 3D Rigid Transformation Between Two Range Views of a Complex Object,” IEEE Proceedings of Int. Conf. on Pattern Recognition, pp. 478-482, 1992.

2. P.J. Besl and N.D. McKay, “A Method for Registration of 3-D Shapes,” Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 14, no. 2, pp. 239-256, 1992.

3. G. Blais and M. Levine, “Registering multiview range data to create 3D computer objects,” Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 17, no. 8, pp. 820-824, 1995.

4. Y. Chen and G. Medioni, “Object modelling by registration of multiple range images,” International Journal of Image and Vision Computing (IVC), vol. 10, no. 3, pp. 145-155, 1992.

5. J.-L. Chen and G.C. Stockman, “Determining pose of 3D objects with curved surfaces,” Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 18, no. 1, pp. 52-57, 1996.

6. C.S. Chua, R. Jarvis, “3D Free-Form Surface Registration and Object Recognition,” International Journal of Computer Vision, Kluwer Academic Publishers, vol. 17, pp. 77-99, 1996.

7. 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.

8. 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.

9. H. Hugli, Ch. Schutz, D. Semitekos, “Geometric matching for free-form 3D object recognition,” ACCV, Singapore, vol. 3, pp. 819-823, 1995.

10. Ch. Schutz and H. Hugli, “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. Ch. Schutz, H. Hugli, “Free-form 3D object recognition,” Optical 3-D Measurement Techniques, Wichmann, Heidelberg, vol. 3, pp. 516-525, 1995.

12. Ch. Schutz, H. Hugli, “Towards the recognition of 3D free-form objects,” Intelligent Robots and Computer Vision XIV, Algorithms, Techniques, Active Vision and Materials Handling, SPIE, Philadelphia, vol. 2588, pp. 476-484, 1995.

13. 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.