

A review of recent range image registration methods with accuracy evaluation [☆]

Joaquim Salvi ^{a,*}, Carles Matabosch ^a, David Fofi ^b, Josep Forest ^a

^a *Institute of Informatics and Applications, University of Girona, Av. Lluís Santalo s/n, 17071 Girona, Spain*

^b *Laboratoire Electronique, Informatique et Image, University of Burgundy, rue de la Fonderie 12, 71200 Le Creusot, France*

Received 10 August 2005; received in revised form 8 February 2006; accepted 16 May 2006

Abstract

The three-dimensional reconstruction of real objects is an important topic in computer vision. Most of the acquisition systems are limited to reconstruct a partial view of the object obtaining in blind areas and occlusions, while in most applications a full reconstruction is required. Many authors have proposed techniques to fuse 3D surfaces by determining the motion between the different views. The first problem is related to obtaining a rough registration when such motion is not available. The second one is focused on obtaining a fine registration from an initial approximation. In this paper, a survey of the most common techniques is presented. Furthermore, a sample of the techniques has been programmed and experimental results are reported to determine the best method in the presence of noise and outliers, providing a useful guide for an interested reader including a Matlab toolbox available at the webpage of the authors.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Computer vision; 3D reconstruction; Range image; Registration

1. Introduction

Surface registration is an intermediate but crucial step inside the whole computer vision complexity. Registration benefits by previous steps such as basically surface reconstruction and view pose estimation without ignoring image processing and camera calibration, and it is the input for further steps such as object model computation. The variety of applications is worthwhile: reverse engineering and mould fabrication in the manufacturing process, artifact reproduction and 3D modeling of carving pieces and sculptures both with applications in the souvenir industry and virtual museums, and many others such as augmented reality in graphics and map building in robotics.

There are many sorts of range finders involved in surface reconstruction which are based on time-of-flight lasers [1], laser scanning [2], stereovision [3], and pattern projection [4]. All of these range finders have their pros and cons and different fields of applications but all of them provide a more or less accurate reconstruction of a surface. In all the cases, the reconstruction does not represent the entire object due to occlusions and the limited field of view of the sensor. In order to solve this problem a set of range images taken from different positions are acquired and then must be fused. Many devices to overcome this task exist in the market, which are basically based on calibrated mechanics to compute the geometry between the views such as rotating tables and robot arms. However, these devices are still limited due to the fact that the object must be located inside the device or close to the working area. Many applications involved with large surfaces or simply with the complexity of moving the object to the scanning system demand a process of accurate registration without a previous knowledge of the pose of the views. This is the aim of this paper, which presents a comprehensive survey

[☆] This research has been partly supported by Spanish Project MCYT TIC2003-08106-C02-02.

* Corresponding author. Tel.: +34 972419812; fax: +34 972418976.
E-mail addresses: qsalvi@eia.udg.es (J. Salvi), cmatabos@eia.udg.es (C. Matabosch), d.fofi@iutlecreusot.u-bourgogne.fr (D. Fofi), forest@eia.udg.es (J. Forest).

of surface registration including a new classification, the description of every category and comparative experimental results with both synthetic and real data.

Other interesting overviews are already available in the literature. For instance, Rusinkiewicz and Levoy [5] published a paper involved on comparing the ICP techniques specially centered in the point-to-point and point-to-plane correspondences and different methods involved in sampling the control points. Although some coarse registration techniques are explained, the paper is incomplete because multi-view registration is neglected. Furthermore, only synthetic data is used in comparing the different techniques without bringing in some of the problems involved using real images in the tests and all the views used in the experiments have overlapping regions which eases the registration. Besides, Dalley and Flynn [6] presented another study centered in pair-wise registration and especially in the different techniques involved in removing the outliers such as basically the Schtz's distance threshold [7] and the Zhang's statistical outlier [8]. Our paper is a completing survey and an extension of these previous works because it analyses the different techniques in both pair-wise and multi-view registration and contributes with experimental results with real range images.

The remainder is organized as follows. First, a new classification is presented in Section 2. Second, coarse registration techniques are presented and compared in Section 3, while fine registration is discussed in Section 4. Next, experimental results compare the surveyed techniques in the presence of both synthetic data and real data acquired some of them from real scanners and others provided by other authors. The article ends with the conclusions, which includes a discussion of the advantages and drawbacks of every group of techniques sort out in the new classification proposal.

2. Classification of registration methods

The goal of registration is to find the Euclidean motion between a set of range images of a given object taken from different positions in order to represent them all with respect to a reference frame. The proposed techniques differ as to whether initial information is required, so that a rough registration can only be estimated without an initial guess. If an estimated motion between views is available, a fine registration can then be computed. The classification of the surveyed methods is revealed in Table 1.

In coarse registration, the main goal is to compute an initial estimation of the rigid motion between two clouds of 3D points using correspondences between both surfaces, as explained in Section 3. These methods can be classified in terms of (a) the kind of correspondences used; (b) the method used to compute the motion; (c) the robustness of the method; (d) the registration strategy (see Table 1). In general, the most common correspondence method used is point-to-point, such as the *Point Signature* method (see Section 3.1) and the method of *Spin Image* (see Section 3.2). However, there are other methods that align lines, like

methods of *bitangent curves* (see Section 3.6) and others that match the surfaces directly, like the *algebraic surface model* (see Section 3.5). Another important aspect of coarse registration is the way of computing the motion when correspondences are found. Robustness in the presence of noise is another important property, because there are usually no corresponding regions between views. Most methods are robust, looking for the best combination of correspondences [9–11]. Other methods may converge to a local solution [12], and in theory this increases the speed of the method but the solution is not always the best, and in some cases it is far from the right solution. In general, coarse registration methods are iterative, usually maximizing the rate of overlapping points. However, a few provides linear solutions, like the methods based on *Principal Component Analysis* or the *Algebraic Surface model* (see Sections 3.3 and 3.5, respectively).

In fine registration, the goal is to obtain the most accurate solution as possible. These methods use an initial estimate of the motion to first represent all range images with respect to a reference system, and then refine the transformation matrix by minimizing the distances between temporal correspondences, known as closest points. Table 1 also classifies fine registration methods in terms of: (a) the registration strategy; (b) the use of an efficient search method, such as k-d trees in order to speed up the algorithm; (c) the way of computing the minimization distance, either point-to-point or point-to-plane; (d) the way of computing the motion in each iteration; (e) the robustness of the method.

The registration strategy can differ according to whether all range views of the object are registered at the same time (*multi-view registration*) or the method registers only a pair of range images in every execution (*pair-wise registration*). Moreover, fine registration methods need a lot of processing time to decide which is the closest point. In order to deal with this problem, several proposals to increase the searching speed have been presented, such as the use of k-d trees to alleviate the problem of searching neighbors.

Another important parameter is the distance to minimize. Most methods use the distance between point-correspondences, while others use the distance between a given point in the first range image and the corresponding tangent plane in the second. The problem of point-to-point distance is that the correspondence of a given point in the first view may not exist in the second view because of the limited number of points acquired by the sensor, especially on low resolution surfaces. To address this problem, some authors use the point-to-plane distance. In this case, a tangent plane in the second view is computed at the position pointed by the given point in the first view. The distance between the point in the first view and that tangent plane in the second is the minimization distance. Theoretically, point-to-plane converges in less iterations than point-to-point.

Finally, robust methods can cope with noise and false correspondences due to the presence of non-overlapping regions. In real images, the robustness is very important, especially when only a small part of the first view has a

Table 1
Classification of the registration methods

Registration strategy	Linear						
	Iterative						
Robustness							
Motion Estimation	Not defined/Other						
	Eigenvectors						
	Least Squares						
Kind of Correspondence	Surface						
	Curves						
	Points						

Coarse Registration	Point Signature	Chua97	√		√		√	√
		Ho99	√			√	√	√
		Chua00				√		√
	Spin Image	Johnson99	√		√		√	√
		Carmichael99	√		√		√	√
		Huber99	√		√		√	√
		Huber02	√		√		√	√
	PCA	Chung		√		√		√
		kim02		√		√		√
		kim03		√		√		√
	RANSAC-Based Darces	Chen98	√		√		√	√
		Hung99	√		√		√	√
	Algebraic Surface Model	Tarel98		√		√		√
	Line-based	Wyngaerd02		√		√	√	√
		Stamos03		√		√	√	√
Chen05			√		√	√	√	
Genetic Alg	Brunnstrom96	√		√		√	√	
Principal Curvature	Feldmar94	√			√		√	

Fine Registration	ICP	Besl92	√		√		
		Kapoutsis98	√		√	√	
		Yamany98	√		√		
		Trucco99	√		√		√
		Greenspan01	√		√		√
		Jost02	√		√	√	
		Sharp02	√		√		
		Zinsser03	√		√	√	√
	Chen	Chen91	√			√	
		Gagnon94		√		√	
		Pulli99		√		√	
		Rusinkiewicz01	√			√	
	Signed Distance Fiel	Masuda01		√		√	√
		Masuda02		√		√	√
	Genetic Algorithms	Chow03	√		√	√	√
Silva04			√		√	√	

Registration strategy	Pair-wise registration				
	Multi-view registration				
Efficient search	k-d trees				
Minimization distance	distance point-point				
	distance point-plane				
	Robustness				

correspondence in the second, that is in the presence of a reduced overlapping region.

3. Coarse registration methods

Coarse registration methods search for an initial estimate of the motion between pairs of consecutive 3D views

leading to the complete registration of the surface. In order to compute this motion, distances between correspondences in different views are minimized. Features from both surfaces are usually extracted with the aim of matching them to obtain the set of correspondences, whereas other techniques find such correspondences without any feature extraction but with some Euclidean invar-

iants. The most common correspondences are points, curves and surfaces.

In some situations, coarse registration techniques can be classified on *shape features* or *matching methods*. The first group searches for characteristics of points, using usually neighborhood information, in order to search for correspondences. Examples of this group are *Point Signature*, *Spin Image*, etc. Matching methods are based on the process of matching points from both surfaces, as *Ransac* or *Genetic Algorithm*. In some situations both techniques can be combined to find correspondences, as Brunnström [13], who used the normal vectors at every point to define the fitness function of the genetic algorithm. On the other hand, techniques of both groups can be used independently as *Ransac* which do not use features in the matching process or *Point Signature* that when points are characterized only a comparison between features from both surfaces is required to detect correspondences.

3.1. Point signature

Point signature is a point descriptor introduced by Chua [9] and used to search for correspondences. Given a point p , the curve of the surface that intersects with a sphere of radius r centered to p gives the contour of points (C). These points are then represented in a new coordinate frame centered at p . The orientation axes are given by the normal vector (n_1) at p , a reference vector (n_2) and the vector obtained by the cross-product. All points on C are projected to the tangent plane giving a curve C' . The vector n_2 is computed as the unit vector from p to a point on C' which gives the largest distance. Thus, every point on C can be characterized by: (a) the signed distance between its own correspondence in C' ; and (b) a clockwise rotation angle θ from the reference vector n_2 . Depending on the resolution, different $\Delta\theta$ s are chosen. Then, the point signature can be expressed as a set of distances in each θ from 0° to 360° . Finally point signatures from two views are compared to determine potential correspondences. The matching process is very fast and efficient.

The main drawback of the algorithm is the process to compute the point signature. The intersection of a sphere to the surface is not very easy, especially when the surface is represented as a cloud of points or a triangulated surface. In this situation interpolation is required, incrementing the computing time and decrementing the quality of the point signature. Moreover the computation of the reference vector is very sensible to noise, and errors in this computation effects the point signature descriptor obtained considerably.

3.2. Spin image

Spin image is a 2D image characterization of a point belonging to a surface [14]. Like point signature, spin image was initially proposed for image recognition. However, it has been used in several registration applications since then.

Consider a given point at which a tangent plane is computed by using the position of its neighboring points. Then, a region around the given point is considered in which two distances are computed to determine the spin image: (a) the distance α between each point to the normal vector defined by the tangent plane; (b) the distance β between this point to the tangent plane; obtaining:

$$\alpha = \sqrt{\|x - p\|^2 - (n(x - p))^2} \quad (1)$$

$$\beta = n(x - p) \quad (2)$$

where p is the given point, n is the normal vector at this point, and x is the set of neighboring points used to generate the spin image. Using these distances, a table is generated representing α on the x -axis and β on the y -axis. Each cell of this table contains the number of points that belong to the corresponding region. In order to choose the size of the table that determines the resolution of the image, the double length of the triangle mesh is selected.

Some spin images are computed in the first view and then, for each one, the best correspondences are searched for in the second view. When the point-correspondences are found, outliers are removed by using the mean and the standard deviation of the residual as a threshold. The rigid transformation is finally computed from the best correspondence found.

The main problem of this method is that the spin image strongly depends on the resolution of the method. In order to solve this problem, Carmichael et al. [15] proposed the *face-based spin image* in which a set of points are interpolated inside every triangular mesh with the aim of uniforming the number of points in every spin image computation. In addition, other approaches have been presented to solve the problem of false mesh triangles given by surface boundaries and occlusions [16]. In this case, the method is used as a filter to remove such false triangles before registration.

Finally, using the variants of spin image, good results can be found in Range Image Registration. The spin image feature is very robust, except in case of symmetries or repeated regions in the object. However this is a problem present in most part of coarse registration techniques.

3.3. Principal component analysis

This method is based on using the direction of the main axis of the volume given by the cloud of points of the range image to align the sequence of range images between them. If the overlapping region is large enough, both main axes should be almost coincident and related to a rigid motion so that registration may succeed. Therefore, this transformation matrix is found to be the one that aligns both axes by only applying a simple product (see Eq. (5)). This method is very fast with respect to others that identify point or curve correspondences. However, the overlapping region must be a very important part of the view in order to obtain good results. Chung and Lee [17] proposed a registration algorithm using the direction vectors of a cloud of

points (a similar approach was used by Kim et al. [18]). The method involves calculating the covariance matrix of each range image as follows:

$$\text{Cov} = \frac{1}{N} \sum_{i=0}^{N-1} (p_i - \bar{p})(p_i - \bar{p})^T \quad (3)$$

where N is the number of points, \bar{p} is the center of mass of the cloud of points, and p_i is the i th point of the surface. Then, the direction U_i of the main axis can be computed by singular value decomposition:

$$\text{Cov}_i = U_i D_i U_i^T \quad (4)$$

The rotation is determined by the product of the eigenvector matrices:

$$R = U_1 U_2^{-1} \quad (5)$$

and the translation is determined by the distance between the centers of mass of both clouds of points, expressed with respect to the same axis:

$$t = \bar{\mu}_2 - R\bar{\mu}_1 \quad (6)$$

Principal component analysis is very fast. However, it can only be used with effectiveness when there is a sufficient number of points. In addition, this method obtains accurate solutions when most part of the points are common. Results are less accurate when the overlapping region constitutes a smaller part of the image. In practice, a 50% overlapping of the region is critical. However, the solution obtained can be used as an initial guess in a further fine registration. The main problem of principal component analysis is its limitation in coping with surfaces that contain symmetries. Thus, if the eigenvalues obtained representing two axes are similar, the order of these axes can change in the matrix U_i , and the final result obtained is completely different from the correct solution. Although PCA provides a fast solution, in most cases this one is far from the expected.

3.4. Ransac-based darces

This method is based on finding the best three point-correspondences between two range images to obtain an estimation of the Euclidean motion. Three points are the minimum required to compute the motion between both surfaces if no other information is used [10]. As will be commented in Section 3.8, Feldmar used only a single point but also considered the normal vector and the principal curvature to obtain enough information to compute the rigid motion [12].

Three points (primary, secondary and auxiliary) in the first view are characterized by the three distances between them (d_{ps} , d_{pa} and d_{sa}). Each point in the second view is hypothesized to be the correspondence of the primary point (p'). Next, the secondary point is searched for among the points located at a distance d_{ps} from p' . If there are not any points in that position, another primary point is tested. Otherwise, a third point in the second view that satisfies the

distances defined in the triplet is searched. Once a triplet is identified, the rigid transformation between both points can be determined. This search is repeated exhaustively for every satisfied triplet between both views and a set of potential Euclidean motions is obtained. The correct transformation is the one that obtains the largest number of corresponding points between both views.

A modification of this method focused on decreasing the computing time related to the search of correspondences was proposed [19]. The results obtained were very good because of its robustness even in the presence of outliers. However, it can only be used when the number of points in each view is relatively small. Theoretically it is a good method. However, the precision depends on the resolution of the surface and the time increases considerably with the number of points, so that it can only be used in applications where time is not critical.

3.5. Algebraic surface model

Tarel et al. [20] proposed a method to estimate the motion between surfaces represented as a polynomial model. First, two implicit polynomial models are determined from all the points of both range images using *3L Fitting*, a linear algorithm based on Least Squares. In general, the algorithms used to obtain a model are iterative, and require a lot of processing to compute the polynomial function. However, the linear algorithm does not require so much computational time and offers better repeatability compared to other implicit polynomial fitting methods.

This method is based on obtaining a function of the distance between the polynomial model and the points, where these distances are nearly zero. In order to improve the accuracy of this method, fictional points are added to the range image located at distances of $+c$ and $-c$ from the surface.

As this method does not need points or curve correspondences, the computation time is faster compared to others. However, a normal vector at each point is required to estimate the model, which it is not easy to compute when only points are available. If the range scanner gives this information, the computing time decreases considerably. The principal drawback of this method is the requirement that a large part of both images must belong to the overlapping region. The author reports good results with less than 15% of non-overlapping region, which is quite unusual in range image registration.

3.6. Line-based algorithm

Some authors proposed to use lines to find pairs of correspondences. Examples are the straight line-based method proposed by Stamos and Leordeanu [21] and the curved line-based method proposed by Wyngaerd [22].

The former is based on the extraction of straight segments directly in the range images which are further registered with the aim of computing the motion between the

different views. The algorithm is applied to large and structured environments such as buildings in which planar regions and straight lines can be easily found. The segmentation algorithm determines a set of border lines and their corresponding planes. First, a robust algorithm is used to efficiently search pairs of lines based on line length and plane area. Then, the rotation and translation among potential pairs is computed. Finally, the one that maximizes the number of planes is taken as the solution.

Some years later, the same authors changed the approach used in computing the motion between straight lines [23]. As most part of lines in a structured environment is contained in the three planes of a coordinate system, they proposed to compute first the three main directions of every view. Hence, 24 combinations arise to potential align both views. Then, the rotation matrix is computed for every combination and, finally, the one that maximizes the number of diagonal elements is selected as the rotation solution. The rest of rotation matrices are kept because final results are supervised by an operator. Translation vectors are computed as the one that connect midpoints of two pair of segments. The more repeated vector is selected to become the solution. Finally, the registration is refined by using an ICP-based method.

The algorithm obtains good results even considering that it is classified into coarse registration. The main drawback is the difficulty to segment the straight segments as well as the supervisor required to check the final results given by the method. Both drawbacks decrease the number of applications but the method has performed very well in the registration of buildings.

The general case of line-based matching is the consideration of curved lines in order to register free-form surfaces

Vanden Wyngaerd [22] proposed a rough estimation of motion by matching bitangent curves. A bitangent curve is a pair of curves composed by the union of bitangent points, which are simultaneously defined as a pair of points tangent to the same plane. The bitangent curves are found by means of a search in the dual space.

The main idea is that all bitangent points are coincident in the dual space. In order to do the search, it is necessary to represent the four parameters of any plane using only three components or coordinates. So, the normal vectors at each point of the range image are computed and their norms are set to one. Using these vectors and the coordinates of their points, it is easy to compute the four parameters of the plane (a , b , c and d) tangent to that point.

Since the norms of the normal vectors are set to one, it is possible to represent this vector using just two parameters. The author used a and b to parameterize the normal vector. In theory, it is possible to construct the dual space using a , b and d . However, it is necessary to normalize the parameter d between -1 and $+1$ to scale the values.

Once all the bitangent curves present in a range image are extracted from the dual space, the matching between these curves with the curves in the next range image starts. In this way, an invariant description of a pair of bitangent curves is

used with the goal of matching only the most representative curves, i.e. the 15 longest ones. The invariant used is defined as the set of distances between bitangent points.

In order to increase efficiency, the curve is divided into segments of equal length. Once a correspondence is found, four corresponding points, that is the two end-points of both bitangent segments, are obtained. With these four correspondences, the Euclidean transformation can be computed, and then the error can be analyzed by transforming all the points with respect to the reference system. The matching of bitangent segments that correspond to the minimum error is selected as the best one among all the potential matches.

Compared to other methods in which the correspondence is based on points, this method has the advantage that the range image is previously transformed into the dual space before the search for possible matches starts. This transformation decreases the computing time and increases the robustness. However, depending on the shape of the object, the number of bitangent points can be insufficient to ensure good results.

3.7. Genetic algorithm

Brunnström and Stoddart [13] used a genetic algorithm to solve the problem of searching for correspondences between two range images. The interest in this method is centered on defining the vector that contains the n index of correspondences between both range images, where the size of the vector is set to n , i.e. the number of points in the second range image (the image that is matched with respect to the first). Genetic algorithms require a fitness function to measure the quality of each potential solution. In order to determine this fitness function, four invariants between the two pairs of correspondences are used:

$$\|\vec{v}_{ij}\| = \|\vec{r}_j - \vec{r}_i\| \tag{7}$$

$$\cos(\theta_{ij}) = \frac{\vec{n}_j \cdot \vec{v}_{ij}}{\|\vec{n}_j\| \cdot \|\vec{v}_{ij}\|} \tag{8}$$

$$\cos(\theta_{ji}) = \frac{\vec{n}_i \cdot \vec{v}_{ji}}{\|\vec{n}_i\| \cdot \|\vec{v}_{ji}\|} \tag{9}$$

$$\cos(\beta_{ji}) = \frac{(\vec{n}_j \times \vec{v}_{ij}) \cdot (\vec{n}_i \times \vec{v}_{ij})}{\|\vec{n}_i\| \cdot \|\vec{n}_j\| \cdot \|\vec{v}_{ij}\|^2} \tag{10}$$

where r_i and r_j are the position of two points belonging to the same surface and n_i and n_j are the normal vectors at both points, respectively.

Using these invariants, the quality of the correspondences is computed analyzing the distance error and the error in the normal parameters as follows:

$$q_d(\alpha_i, \alpha_j) = e^{-\frac{\|\vec{v}_{\alpha_i \alpha_j} - \vec{v}_{ij}\|^2}{2\sigma^2}} \tag{11}$$

$$q_n(\alpha_i, \alpha_j) = e^{-\frac{(\theta_{\alpha_i, \alpha_j} - \theta_{ij})^2 + (\theta_{\alpha_j, \alpha_i} - \theta_{ji})^2 + (\beta_{\alpha_i, \alpha_j} - \beta_{ij})^2}{2\mu^2}} \tag{12}$$

$$Q(\alpha_i, \alpha_j) = q_d(\alpha_i, \alpha_j) q_n(\alpha_i, \alpha_j) \tag{13}$$

where σ and μ are experimental parameters that must be estimated; i and j are the indexes that determine two points in the second range image; and α_i and α_j represent the indexes of two other points in the first range image. Then, the quality of a correspondence α_i can be computed as the sum of the qualities of every pair of correspondences between α_i and the rest of the points.

$$Q(\alpha_i) = \sum_{j \neq i} Q(\alpha_i, \alpha_j) \quad (14)$$

The previous function indicates the quality of a pair of correspondences, while the fitness function indicates the global matching quality, which is expressed as a function of $Q(\alpha_i)$ as follows,

$$f(\vec{\alpha}) = e^{\sqrt{\sum_i Q(\alpha_i)}} \quad (15)$$

When the fitness function is defined, the cross-over and mutation probabilities are fixed to characterize the algorithm. The mutation is not very important when searching for correspondences because nearby indexes do not imply nearby points in the space. Therefore, the author set the probability of mutation at 1%, with a cross-over of 90%.

Finally, the stopping criteria, which it is not a very well-studied problem in genetic algorithms, must be defined. Three different approaches were presented: (a) setting a % of good correspondences (computed starting from the closest points); (b) supervising the fitness function so that it does not increase after a certain number of iterations; (c) counting iterations until a certain number. When the algorithm finishes, the Euclidean motion might be computed because the chromosome that maximises the fitness function contains the point-correspondences. However, some correspondences in the chromosome might be wrong, which means that these bad correspondences must be previously removed in order to guarantee the computation of a good Euclidean Motion, using SVD, for instance. So, only the 30% of the correspondences that maximize $Q(\alpha_i)$ are used in the computation. As in most genetic approaches, the results obtained are quite good but the computing time is expensive, specially in the presence of a large number of points, where exists lots of potential correspondences. As Ransac-based algorithm, it is not appropriate when time is critical.

3.8. Principal curvature

Feldmar and Ayache [12] proposed that use of the differential of points characterized by the principal curvature in the matching process. This method only needs a single correspondence to compute the Euclidean motion between two range images. It characterizes a point by its principal curvatures (k_1, k_2) . Principal curvatures are the maximum and the minimum curvature of the surface at a point. Additionally, the normal vector and the principal direction corresponding to the principal curvature are also considered. In order to facilitate the search, points of the second range image are organized in a table indexed by their curvature

values. Then, considering a point M in the first view, whose curvatures are (k_1, k_2) , the set of points in the second view, whose curvatures are close to (k_1, k_2) , can be quickly found and evaluated as the corresponding point. For every potential matching, the Euclidean motion that aligns $P_1 = (M_1, \vec{e}_{11}, \vec{e}_{21}, \vec{n}_1)$ with $P_2 = (M_2, \vec{e}_{12}, \vec{e}_{22}, \vec{n}_2)$ is computed, where M_i is a 3D point in the i image, $\vec{e}_{1i}, \vec{e}_{2i}$ are the principal directions of the curvature at the M_i point, and \vec{n}_i is the normal vector at that point. Thus, two rigid displacements, D and D' , are computed, where D corresponds to align P_1 to P_2 , while D' aligns P_1 to $P'_2 = (M_2, -\vec{e}_{12}, -\vec{e}_{22}, \vec{n}_2)$. Finally, the transformation matrix (R, t) defining the Euclidean motion between both views can be easily computed as follows:

$$D = (BA', M_2 - BA'M_1) \quad (16)$$

$$D' = (B'A', M_2 - B'A'M_1) \quad (17)$$

where A is the 3x3 matrix whose columns are $(\vec{e}_{11}, \vec{e}_{21}, \vec{n}_1)$, B is $(\vec{e}_{12}, \vec{e}_{22}, \vec{n}_2)$ and B' is $(-\vec{e}_{12}, -\vec{e}_{22}, \vec{n}_2)$. Then, every transformation matrix is evaluated. A ratio is computed by considering the number of points in the transformed surface (computed from the Euclidean motion) which have a corresponding point in the second surface at a distance smaller than a threshold related to the total number of points of the range image. If both D and D' do not reach the termination criteria, the algorithm is repeated using an alternative initial point. Otherwise, the transformation matrix computed is considered to be a good estimation.

Although the curvatures and principal directions are not always known, the author described a method to determine them based on the use of a parameterization of the surface similar to the polynomial fitting of Tarel (see Section 3.5).

The main problem of this method is that it is not robust. Only a good correspondence is search for, and when a possible solution is found, the algorithm stops. However, as the correspondence is not validated, it might be a false matching, obtaining a bad initial registration that may satisfies the rate of overlapping.

4. Fine registration methods

The term *fine registration* is used when an estimation of the motion is previously known and used as an initial guess to iterate and converge to a more accurate solution. In order to solve this problem, a distance function is minimized. Some authors used the distance between point-correspondences, while others used the distance of a given point in the first view to a plane tangent to its closest point in the second view. In recent years, some methods have been presented: (a) iterative closest point; (b) Chen's method; (c) signed distance fields; and (d) genetic algorithms, among others.

In the registration process, different methodology can be used independently of the technique chosen, which are hereafter briefly related:

4.1. Control points

Although some authors use all points in the registration step [24,25], others use Uniform subsampling [26,27], Random Sampling [28] or Normal Sampling [5]. In the common case, the better solution is the Normal Sampling for two reasons. As a sampling is applied, the time required is smaller. Furthermore, as points from significant parts of the surfaces are more important compared to parts of uniform surfaces, usually better results are obtained, specially in cases where the surfaces are not very shaped.

4.2. Points weight

Although in some situations all points have the same weight [24], in other situations weights are introduced depending on: distance between point-correspondences [29], compatibility of normals [29] and the uncertainty of the covariance matrix [30], among others.

4.3. Rejecting pairs

Some authors used all point-correspondences to determine the motion because they usually work with image-to-model registration. However, when an image is not a subset of the following in the sequence, some correspondences are outliers and must be rejected.

Now, the surveyed fine registration techniques are discussed in the following sections.

4.4. Iterative closest point (ICP)

The ICP method was presented by Besl and McKay [24]. The goal of this method is to obtain an accurate solution by minimizing the distance between point-correspondences, known as closest point. When an initial estimation is known, all the points are transformed to a reference system applying the Euclidean motion. Then, every point in the first image (p_i) is taken into consideration to search for its closest point in the second image (m_i), so that the distance between these correspondences is minimized, and the process is iterated until convergence (see Eq. (18)).

$$f = \frac{1}{N_p} \sum_{i=1}^{N_p} \|\vec{m}_i - R(\vec{q}_R) \vec{p}_i - \vec{t}\|^2 \quad (18)$$

In order to minimize the previous equation, a symmetric 4×4 matrix $Q(\Sigma_{pm})$ is defined as follows,

$$Q(\Sigma_{pm}) = \begin{bmatrix} \text{tr}(\Sigma_{pm}) & \Delta^T \\ \Delta & \Sigma_{pm} + \Sigma_{pm}^T - \text{tr}(\Sigma_{pm})I_3 \end{bmatrix} \quad (19)$$

where tr is the trace, $\Delta = [A_{23}A_{31}A_{12}]^T$ is computed from the anti-symmetric matrix $A_{ij} = (\Sigma_{px} - \Sigma_{px}^T)_{ij}$, Δ^T is the transpose of Δ , I_3 is the identity matrix and Σ_{pm} is the cross-covariance matrix of the points P and M given by:

$$\Sigma_{px} = \frac{1}{N_p} \sum_{i=1}^{N_p} [\vec{p}_i \vec{m}_i] - \vec{\mu}_p \vec{\mu}_m \quad (20)$$

The unit eigenvector $\vec{q}_R = [q_0 \ q_1 \ q_2 \ q_3]^T$ corresponding to the maximum eigenvalue of the matrix Q is selected as the optimal rotation expressed in quaternions. Once R is computed (see Eq. (21)), the translation vector can be easily computed (see Eq. (22)) and the motion determined by R and t.

$$R = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 - q_0q_3) & 2(q_1q_3 + q_0q_2) \\ 2(q_1q_2 + q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (21)$$

$$t = \vec{\mu}_m - R\vec{\mu}_p \quad (22)$$

The motion is applied to the first surface and the process is repeated until distances between corresponding points decrease below a threshold. ICP obtains good results even in the presence of Gaussian noise. However, the main drawback is that the method can not cope with non-overlapping regions because the outliers are never removed. Moreover, when starting from a rough estimation of the motion, the convergence is not guaranteed.

Some modifications of ICP have been presented in recent years. In 2001, Greenspan and Godin [31] applied the Nearest Neighbor Problem to facilitate the search of closest points. The first range image is considered as a reference set of points, which is preprocessed in order to find, for every point, the neighborhood of points in the second view located at a certain distance. The points of the neighborhood are sorted according to that distance. The use of this pretreatment leads to consider the closest point of the previous iteration as an estimation of the correspondence in the current iteration. If this estimation satisfies the spherical constraint, the current closest point is considered to belong to the neighborhood of the estimate. A property of the spherical constraint holds that any point that is closer to \vec{q} (a point in the second range image) than the current estimate \vec{p}_c (the closest point obtained in the last iteration) must be located inside a sphere centered at \vec{p}_c with a radius of $2\|\vec{q} - \vec{p}_c\|$. This pretreatment decreases the computing time drastically. A year later, Jost and Hugli [32] presented the Multi-resolution Scheme ICP algorithm, which is a modification of ICP for fast registration. The main idea of the algorithm is to solve the first few iterations using down sampled points and to progressively increase the resolution by increasing the number of points considered. The author divides the number of points by a factor in each resolution step. The number of iterations in each resolution step is not fixed, so that the algorithm goes to the next resolution when the distance between correspondences falls below a threshold.

In the same year, Sharp et al. [33] proposed the ICP using invariant features ICPIF. In this case, points are matched using a weighted feature distance as follows,

$$d_x(p, m) = d_e(p, m) + \alpha^2 d_f(p, m) \quad (23)$$

where d_e is the Euclidean distance, d_f is the distance in the feature space between each correspondence points and α

controls the relative contribution of the features. Different invariant features were proposed: (a) curvature; (b) moment; (c) spherical harmonics. Experiments reported by the author pointed out that the spherical harmonics provided the best convergence rate, while the traditional ICP provided the worst one.

One year before, Godin et al. [34] presented a similar work, where color and curvature were used as feature information. As color is used as a matching constraint, symmetric objects can be well registered. However the author only presents results with simple objects without shadows, occlusions or luminance changes.

In addition, other authors proposed some improvements to increase the robustness of ICP. For instance, Trucco et al. [35] implemented the RICP method making use of the Least Median of Squares approach. The method is based on executing the registration with just a sample of random points (m points), computing this operation a sufficient number of times with the aim of finding a registration without outliers. The Monte Carlo algorithm was used to estimate the number of executions. Once all the potential registrations were computed, the one that minimizes the median of the residuals is chosen as the solution. Finally, the correspondences with a residual larger than 2.5σ were removed and the transformation between both views was computed using only the remaining points (inliers). Note that σ was estimated by using a robust standard deviation [36].

Moreover, Zinsser et al. [37] proposed a robust method based on outlier thresholding known as the Picky ICP algorithm. The main difference with respects to the previous methods is that at every iteration only the pairs of correspondences with the smallest distances are used in the motion computation. The threshold was fixed at a given multiple of the standard deviation.

Overall, ICP is the most common registration method used and the results provided by authors are very good. However, this method usually presents problems of convergence, lots of iterations are required, and in some cases the algorithm converges to a local minimum. Moreover, unless a robust implementation is used, the algorithm only can be used in surface-to-model registration.

4.5. Method of Chen

In 1991, Chen and Medioni [25] proposed an alternative to the ICP algorithm, which was based on minimizing the distance between points and planes. The minimization function was selected to be the distances between points in the first image with respect to tangent planes in the second. That is, considering a point in the first image, the intersection of the normal vector at this point with the second surface determines a second point in which the tangent plane is computed.

Despite other authors that considered this algorithm just an improvement of ICP, we have considered it a new technique for several reasons. First, the paper was presented at

the same time of Besl's approach. Second, the search of correspondences is an important aspect of the algorithm, consequently it should be considered a different method.

A new algorithm to find these intersections between lines and range images was proposed, bearing in mind that the intersection of a given line with a cloud of points requires a high computational cost. A more accurate technique to find such a distance was proposed by Gagnon et al. [38].

Finally, once the distances between points and planes were obtained, the motion that minimizes these distances was estimated by least squares. The process was repeated until convergence was attained.

In 2001, Rusinkiewicz and Levoy [5] presented several variations of this algorithm to improvement the precision of the algorithm. The author proposed the *Normal Space Sampling*. The main idea is to select more points in the region where the normal is different from the other parts of the surface. Using this sampling technique better results are obtain in low shaped surfaces.

Point-to-plane distance is normally more difficult to compute compare to point-to-point. When no normal information is given, the plane must be computed using neighborhood information, which requires a lot of time to detect this neighborhood, and not always with sufficient accuracy in the estimation of the normal vector. However, nowadays this estimation can be obtained directly from most part of range finders. This method is more robust to local minimum and, in general, results are better. Despite no robust method are applied, this method is less sensible in the presence of non-overlapping regions. The reason is that only the control points that their normal vector intersects the second view are considered in the matching. Moreover, point-to-point distance ensures a correspondence unless a distance threshold or other robust identifier is used. Chen's approach usually requires less iterations than the ICP approach.

4.6. Matching signed distance fields

In 2001, Masuda [27,39] presented a new registration algorithm, which was based on the matching signed distance fields. The method was a robust one so the outliers were removed, and all the views of a given object were registered at the same time, which means a *multi-view registration*.

First, all views (α) were transformed to a reference coordinate system using the initial estimation of the motion. A set of key points was generated on a 3D grid of buckets with a fixed size δ . Then, the closest point from every key point was searched for in order to establish the correspondences, limiting the distance between points to $\sqrt{3}\delta$.

The process was composed by the following steps:

- (1) Initialization: Compute the initial values of the motion parameters, $T_0^\alpha = [R_0^\alpha t_0^\alpha]$.
- (2) Determine the closest point p_i^α to every key point p .
- (3) Compute the new motion parameters T_{i+1}^α using the correspondences between p and p_i^α .

Steps 2 and 3 are repeated until there is convergence and an integrated model of the object made by the p points is obtained.

At every iteration the closest point was computed as follows:

$$p_i^\alpha = (T_i^\alpha)^{-1}p - (n_p^{\alpha T}(T_i^\alpha)^{-1}p - k_p^\alpha)n_p^\alpha \quad (24)$$

where n_p^α represents the normal vector of p and k_p^α represents the signed distance, which is computed as:

$$k_p^\alpha = n_p^{\alpha T}c_p^\alpha - d_p \quad (25)$$

where c_p^α is the closest point to each key point for every range image α , and d_p is the average distance between the point p and each c_p^α .

$$d_p = \frac{1}{\sum_{1 \leq \alpha \leq N_R} w_p^\alpha} \sum_{1 \leq \alpha \leq N_R} w_p^\alpha d_p^\alpha \quad (26)$$

where d_p^α represents the distance between point p and the nearest key point and w_p^α is the weight of point p .

Finally, once the solution has converged, it is possible to compute the points (signed distance field) and their normal vectors, obtaining the final model as follows:

$$SDF_p = p + n_p d_p \quad (27)$$

$$n_p = \frac{\sum_{1 \leq \alpha \leq N_R} w_p^\alpha R^\alpha n_p^\alpha}{\|\sum_{1 \leq \alpha \leq N_R} w_p^\alpha R^\alpha n_p^\alpha\|} \quad (28)$$

The advantage of this method is that it registers all the views at the same time, which implies a more robust solution and avoids the error propagation problem present in *pair-wise registration* methods. On the other hand, this algorithm can not be used in real time as *localization and mapping* because not all views are already available when the motion must be computed.

Another advantage of this algorithm is that the final result is directly an integrated model, compared to pair-wise registration techniques that require an off-line algorithm to eliminate the overlapping region and triangulate all the registered surfaces.

4.7. Genetic algorithms

Chow et al. [40] presented a dynamic genetic algorithm to solve the registration problem. The goal of this method is to find a chromosome composed of the 6 parameters of the motion that aligns a pair of range images accurately. The chromosome is composed of the three components of the translation vector and the three angles of the rotation. In order to minimize the registration error, the median is chosen as the fitness function, as follows:

$$F(T) = \text{Median}(E_i) \quad (29)$$

$$E_i = \min|Tp_i - m_j| \quad (30)$$

where T is the transformation matrix composed by 6 motion parameters and p_i and m_i the points of both surfaces.

Therefore, only a sample of points of the first image were used to compute the error with the aim of decreasing the computing time. The cross-over operation consisted of combining genes made by two chromosomes to create a new chromosome. The author randomly selected the number of genes to be swapped. The cross-over operation works well when the chromosome is far from the final solution but it is useless for improving the solution in a situation close to convergence. Therefore, the mutation operation was defined as follows: a gene is randomly selected and a value randomly obtained between the limits $[-MV, +MV]$ is added. The limits are very wide at the beginning and become narrower at every step in order to guarantee the convergence in the final steps.

A similar method was proposed the same year by Silva et al. [41]. The main advantage of this work is that a more robust fitness function is used and the initial guess is not required. The author defined the Surface Interpenetration Measure (SIM) as a new robust measurement that quantifies visual registration errors. Another advantage of Chow's method is the multi-view registration approach. Finally, the hillclimbing strategy was used to speed up the convergence.

Overall, the use of genetic algorithms has the advantage of avoiding local minima which is a common problem in registration, especially when the initial motion is not provided or it is given with low precision. This algorithm also works well in the presence of noise and outliers given by non-overlapping regions. The main problem of this algorithm is the time required to converge. Additionally, fitness function must be theoretically computed using real correspondences, however, as they are unknown, temporally correspondences are used using the estimated motion. However, as the motion change iteratively, correspondences must be searched several times (for example, every time a better result is found), resulting again in a lot of computing time.

5. Experimental results

Although many authors evaluated the accuracy of their methods, very few provided a comparison between already published approaches [5,6,42]. In order to provide such a useful comparison, some methods to evaluate the accuracy have been tested which allow us to decide the best method in every situation. The computing time has also been considered, which might be critical in some applications. The measures used to determine the accuracy are the following: *rotation error*, *translation error* and *RMS (root mean square)*.

In order to evaluate the accuracy in rotation, it is represented as a directional vector and a rotating angle around such a vector, which both can be easily extracted from any rotating matrix. Then, the estimated rotation is compared to a real one. So, the error is determined as the angle between both directional vectors (γ) and the discrepancy between both angles of rotation ($\alpha - \beta$). The translation

error is defined as the distance between the origin of the coordinate system estimated by registration with respect to the real one, that is, the norm of the difference between both translation vectors. The RMS error is defined as the mean of the distances between point-correspondences considering all views in the same reference frame.

In order to compare the main algorithms, synthetic data is used to evaluate the registration errors. First of all, synthetic test scenes are generated. Additionally more realistic synthetic data is used, and finally experiments with real data are realized.

The first test scenes used are *Wave* and *Fractal landscape* (see Fig. 1). These surfaces have been used by other authors [37,5] to test their algorithms. These surfaces are contained in a range of 1 in *X* and *Y*-axis and approximately 0.5 in *Z*-axis. Although these scenes certainly do not cover all sort of scanned objects, they are quite representative of many real scannings. Several experiments are performed to evaluate the precision in front of different error sources, as shape measurement error (noise, surface sampling, etc.), and correspondence error (occlusion, outliers, etc.), among others. Although it is very difficult to distinguish these error sources, because they are usually related, these experiments are organized to show basically the effects of each source of error.

As final results can change depending on the implementation of the algorithm, details of this implementation are briefly commented. Ransac-based approach is programmed using a subset of 5% of the points of the first view to search for correspondences using all the points in the second view. Moreover, as distances between primary, secondary and auxiliary points can not be exactly found, a variation of a 2% of such a distance is permitted. Furthermore, the distance between the 3 control points is limited to a 15% of the maximum distance to speed up the process. In the spin image algorithm, the number of correspondences searched is fixed to the 5% of the number of points, and 25% of points are used to compute the spin image. In the implementation of the genetic algorithm of Brunnström,

only 5% points of one surface are searched in the second. In this situation, the second surface have all points, differing from the author that used only 50 points in every surface. However, it is still difficult to work with accuracy using such large number of points.

In fine registration algorithms, a threshold of 0.00001 is fixed to guarantee convergence. However, a maximum number of iterations is fixed to 20 in all the algorithms, except in the case of Chow which number of iterations is fixed to 600 due to the fact that the method is based on genetic algorithms.

In order to evaluate the effects of shape measurement error, two experiments have been done. First of all, Gaussian noise is independently added into the test scenes. Experiments are performed using different values of standard deviation error; the first with an error of 1% of the range dimension in each angle direction, and in the second one with 2% of standard deviation.

In Coarse registration algorithm fractal surface and wave surface are subsampled to 512 points. The motion between both surfaces is a rotation of 45° in each angle and a displacement of 1 unity in each axis.

The results obtained shown that the final result is not very influenced by the noise, because this noise is negligible with respects to the errors in coarse registration techniques.

In order to test the fine registration algorithms, a sequence of surfaces of 2000 points are registered. Every surface is related to the previous by a rotation of 5° around every axis and a displacement of 0.2 unities in the *Z*-axis. This initial motion is similar to the pre-alignment used by Rusinkiewicz and Levoy [5]. The same percentage of noise is independently added in each surface.

Results obtained in Fig. 2 shown that while the error obtained by Chen algorithm is directly related to the noise introduced, Besl and Chow have always a residual error not depending on the noise introduced. This error is usually consequence of the convergence to a local minima. As Chen approach solves better the problem of local minima, excellent results are obtained without noise.

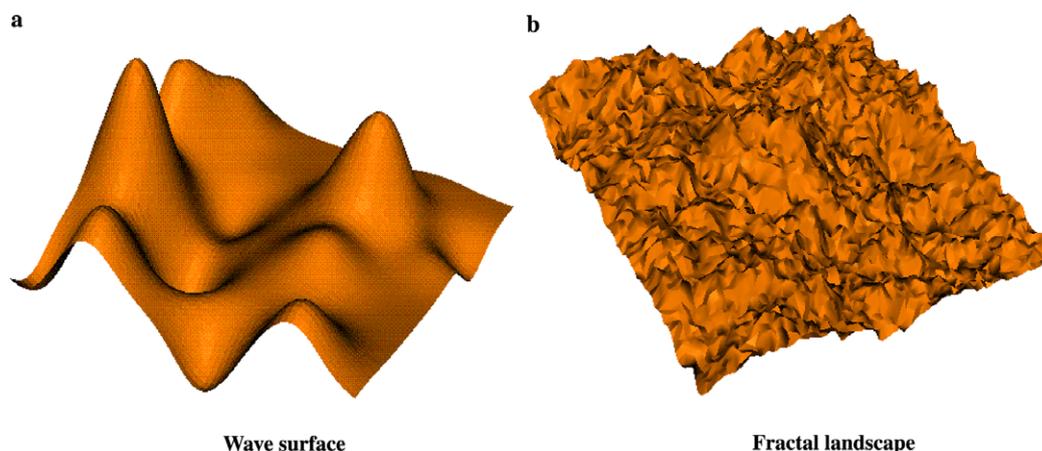


Fig. 1. Test scenes used throughout this paper.

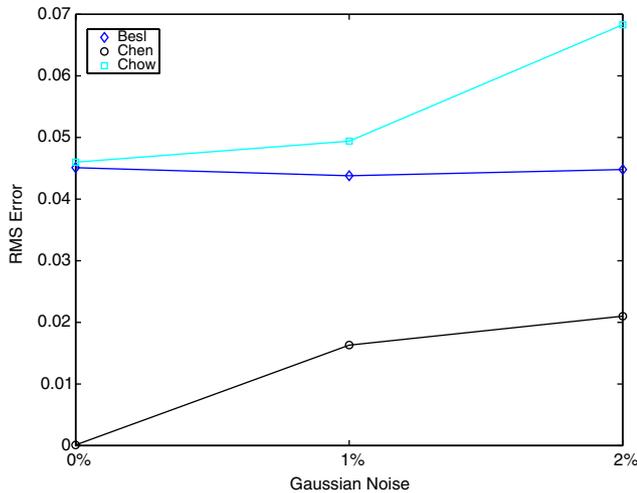


Fig. 2. Effects of Gaussian noise in fine registration algorithms.

Another case of misalignments is the surface sampling. Theoretically when surfaces are scanned with low resolution, worse results are obtained. As surfaces have less points, points correspondences are found with low precision. In order to test the effects, original surfaces are independently subsampled into different resolutions: (a) (12.5% and 25% of the initial size (2048 points) in coarse registration; and (b) 12.5% and 25% of the initial size (4096 points) in fine registration).

As is shown in Fig. 3, errors in Fine registration algorithms extremely depends on the sampling surface. While the variations in Trucco, Besl and Chow are not significant, in Chen and Zinsser algorithms, the errors using surfaces of 2000 points are approximately the half than using a subsampling of 500 points.

Furthermore, in the case of low sampling (500 points), Zinsser and Chen algorithms presented the best solutions. As the Chen approach use a point-to-plane distance, the accuracy is better compared to point-to-point because the

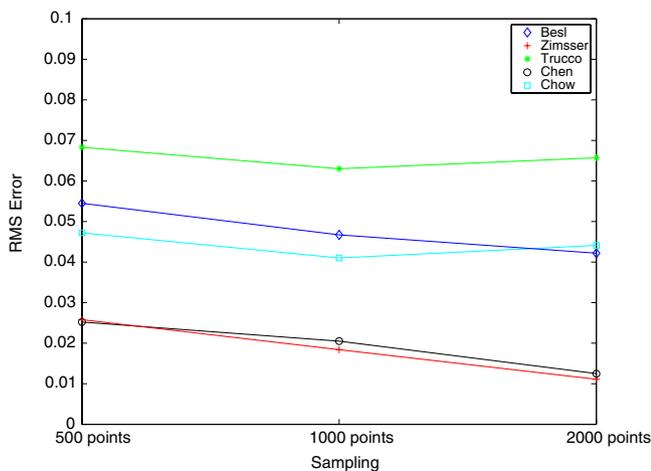


Fig. 3. Evolution of the RMS error in fine registration algorithms in wave surface with respects to the surface sampling.

Table 2
Fine registration errors in subsampled fractal surface (500 points)

Method	RMS error	Translation error	$\alpha-\beta$ (rad)	Time (s)
Besl	0.052	0.029	-0.0297	5.47
Zinsser	0.006	0.28	-0.013	2.08
Trucco	0.053	0.056	-0.028	3.48
Chen	0.022	0.006	-0.001	2.96
Chow	0.043	0.033	-0.146	473.47

plane interpolation decrements the effects of the subsampling. Although Zinsser algorithms presents similar results in RMS errors, the rotation errors of Zinsser’s are considerably larger compared to Chen’s (see Table 2). The fact is that not always RMS errors are representative of a correct registration, especially in robust variant when correspondences related with large distances are removed, and only the correspondences with small distances remains.

On the other hand, in most part of results obtained with Coarse Registration methods, sampling does not effect considerably the final results, and low resolution surfaces can be used in this situation. In this experiment only Ransac-based is clearly effected by the sampling. This is because it is difficult to find 3 points in both surfaces that corresponds exactly to the same position. Besides, PCA does not use directly the points to computed the motion, and spin image and the genetic algorithm use more points reducing the error in the motion estimation. However, experiments realized with more complex objects show that in this situation, results are more influenced by the sampling.

In both experiments realized, 100% of overlapping area is used. However, in most part of real applications both surfaces contains non-overlapping regions. In these situations, false correspondences are very common, and they affect considerably the final resolution. In this experiment, several percentages (5%, 10%, 20% and 50% of the surface) of non-overlapping region are introduced.

Results obtained shown that the presence of non-overlapping regions do not effect significantly coarse registration techniques. In general, coarse registration techniques only use a few part of points to obtain the motion, and, if these points are correctly selected, the result is similar to the one obtained without outliers. On the other hand, Principal Component Analysis obtained similar eigenvalues from the surfaces with or without non-overlapping regions because they are almost planar.

In fine registration techniques, errors in Besl and Chow algorithms increase directly proportional to the percentage of non-overlapping region. This change is specially significant with 50% of outliers (see Fig. 4). On the other hand, Zinsser algorithm can cope with 5% and 10% of outliers, however, the accuracy decreases. Finally, Chen approach is robust to outliers. This is because point-correspondences are computed using the projection of points in a grid Z space, and points whose projection belongs to an empty

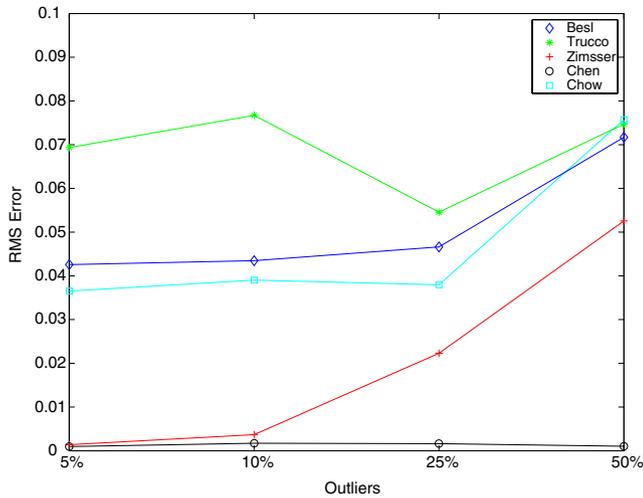


Fig. 4. Effects of outliers in the quality of the fine registrations algorithms.

Table 3
Experimental results using synthetic data obtained by coarse registration methods

Points	Method	Translation error (mm)	$\alpha-\beta$ (rad)	Time (s)
200	Ransac-based	50.613	0.98573	1.062
	PCA	21.229	2.9533	0.016
	Spin image	149.84	2.8911	7.969
	GenAlg	81.359	0	4.281
400	Ransac-based	21.394	0.30569	106.08
	PCA	18.777	3.0037	0.016
	Spin image	81.19	2.3465	50.703
	GenAlg	33.537	1.6721	122.56
700	Ransac-based	27.557	0.096427	32820
	PCA	16.566	0.21871	0.016
	Spin image	64.254	0.90339	169.76
	GenAlg	18.017	2.9687	249.89

cell are not considered. Most part of outliers are removed and motion is only computed using points belonging to both surfaces.

Other experiments are realized using more realistic objects. Tables 3 and 4 show the experimental results obtained by using synthetic clouds of 3D points of Fig. 6, taken from the database of RapidForm2004.

Finally, real scanned objects are used to take into account the effects of the shape measurement error. Results are shown in Tables 5 and 6. It can be shown in Fig. 5 that the executing time of Ransac algorithm is surprisingly smaller in the case of 500 points compared to the case of 250 points. This is because in the case of 250 points, the distance between three points is larger compare to the case of 500 and no good correspondences might be found without increasing the searching distance.

Although the time required is very important, all the methods have been programmed using Matlab 6.5 in a Pentium IV 2,6 GHz because we are just interested in the comparison among the methods and Matlab guarantees an

Table 4
Experimental results using synthetic data obtained by fine registration methods

Points	Method	Translation error (mm)	$\alpha-\beta$ (rad)	RMS error (mm)	Time (s)
700	Besl	0.37	0.01	1.59	0.91
	Zinsser	0.25	0.003	2.64	3.28
	Jost	0.82	0.006	4.85	0.38
	Trucco	0.87	0.011	2.27	4.13
	Chen	0	0	1.49	25.03
1500	Besl	0.41	0.007	1.48	4.77
	Zinsser	0.093	0.002	1.75	9.28
	Jost	0.257	0.009	4.22	1.30
	Trucco	0.768	0.016	5.44	112.8
	Chen	0	0	1.68	145.06
6000	Besl	0.166	0.002	0.71	47.86
	Zinsser	0.169	0.002	1.86	406.05
	Jost	0.801	0.019	3.01	8.51
	Trucco	0.458	0.015	2.91	198.28
	Chen	0	0	1.36	1217.3

Table 5
Experimental results using real range images obtained by coarse registration methods

Points	Method	Translation error (mm)	$\alpha-\beta$ (rad)	γ (rad)	Time (s)
250	Ransac-based	69.406	2.3562	1.3855	2.86
	PCA	67.464	1.9923	1.2985	0.016
	Spin image	53.861	0.91301	1.9844	24.156
	GenAlg	72.094	2.1475	1.5708	44.766
500	Ransac-based	127.03	2.3562	1.5708	2.453
	PCA	69.87	1.9703	1.3306	0.016
	Spin image	40.147	0.54215	0.68824	101.59
	GenAlg	82.698	2.1988	1.5708	243.94
1000	Ransac-based	40.53	0.17204	1.5236	1657.7
	PCA	69.893	1.9236	1.3262	0.203
	Spin image	23.133	0.54339	1.5708	565.45
	GenAlg	38.059	1.7571	0.85619	1051.6

easy implementation. Furthermore kd-tree or other velocity improvements are not used in these experiments to easy again the comparison.

The 3D clouds of points have been acquired using a Minolta Vivid 700 laser scanner, obtaining up to 8 views of a given object. The registration results obtained are shown in Fig. 10.

The real positions of the object are known and so the accuracy of the methods compared because the real objects have been scanned on a turnable table with a the precision of less than 1°.

Several objects have been registered and we have observed that the shape of the object affects the accuracy of the registration considerably and independently of the registration method used. Overall, when the views have poor details, planarity or symmetry, the registration does

Table 6
Experimental results using real range images obtained by fine registration methods

Points	Method	Translation error (mm)	α - β (rad)	RMS error (mm)	Time (s)
500	Besl	3.34	-0.09	4.18	1.41
	Zinsser	0.68	-0.06	2.35	1.95
	Jost	2.01	-0.07	4.54	0.36
	Trucco	2.87	-0.05	3.93	19.33
	Chen	1.3373	-0.008	2.1608	18.391
	Chow	0.27	0.0	4.60	154.05
1000	Besl	3.47	-0.09	3.78	5.22
	Zinsser	0.47	-0.02	1.42	7.97
	Jost	2.38	-0.08	3.70	1.38
	Trucco	2.71	-0.05	5.13	33.95
	Chen	0.29957	0.003	1.7305	67.954
	Chow	0.12	-0.01	4.51	281.61
5000	Besl	3.12	-0.08	3.20	57.36
	Zinsser	0.26	-0.02	0.85	281.58
	Jost	2.24	-0.09	2.83	9.14
	Trucco	2.57	-0.05	1.77	149.59
	Chen	1.2535	-0.015	1.2543	993.09
	Chow	0.06	0.0	3.57	1776.00



Fig. 6. 3D Model from RapidForm2004 used in the synthetic experiments.

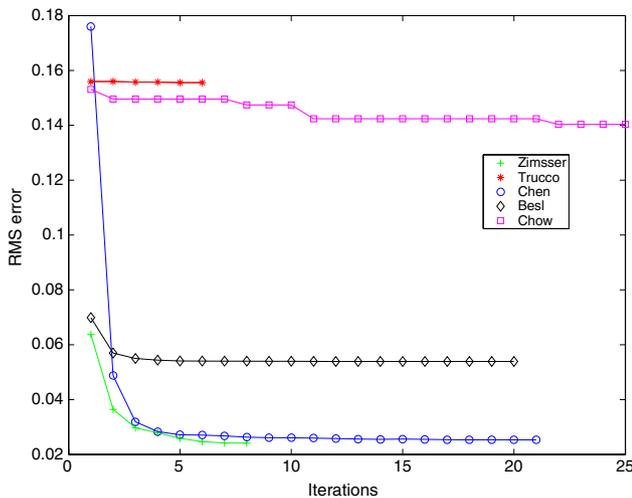


Fig. 5. Evolution of the fine registration RMS error in each iterations. Note that Trucco algorithm stops because the error difference between consecutive iterations is very small. On the other hand, Chow algorithm needs a lot of iterations (a maximum of 600 is fixed), however only the first ones are represented.

not converge (see Fig. 9). Moreover, a robust method is essential to register real range images. See, for instance, Fig. 10a compared to the rest of registrations (b, c and d) in which the registrations are more accurate because only the points belonging to the overlapping region have been used in the computation.

Although this paper is focused on pair-wise registration so that the interest is centered in the alignment of two views, the same algorithms may be used to register large

surfaces. The topic of large surface registration is focused on the registration of several views among them with the aim of obtaining a complete model, but then we have to make additional considerations besides of just a change of resolution. First, information of vicinity is required to know which pairs of views contain enough overlapping area. Second, the drift or error propagation problem has to be considered to reduce the residual in the registration. With the aim of solving both difficulties, multi-view registration techniques appeared. Basically, these techniques register a set of views simultaneously and the drift is distributed between the views [27,43]. However, classic multi-view algorithms are not the best option because in large surfaces there are lots of views without overlapping area. So, some authors proposed to build adjacent graphs to determine loops which are further register using a multi-view algorithm [44]. Fig. 11 shows the registration of a large object composed of 27 different views (see Fig. 12). The algorithm implemented is based on the graph of adjacencies which is built from consecutive pair-wise registration. Then, a multi-view algorithm is applied to minimize the drift of every loop in the graph Fig. 13.

6. Conclusions

In this paper, most representative Range Image Registration algorithms have been surveyed, programmed and tested with different surfaces: synthetic images, test scenes used by other authors and real images. Experiments performed shown the main characteristics of each method.

Coarse registration techniques are used when an initial estimation of the Euclidean motion is unknown. The main interest in these techniques is the algorithm used to find the correspondences between clouds of points, which can be based on points, curves, surfaces and directional vectors. Besides, fine registration methods are based on converging to a solution from an initial estimation of the rigid motion. Depending on the method used, a quite accurate initial guess is required because some methods have convergence problems due to the presence of local minima.

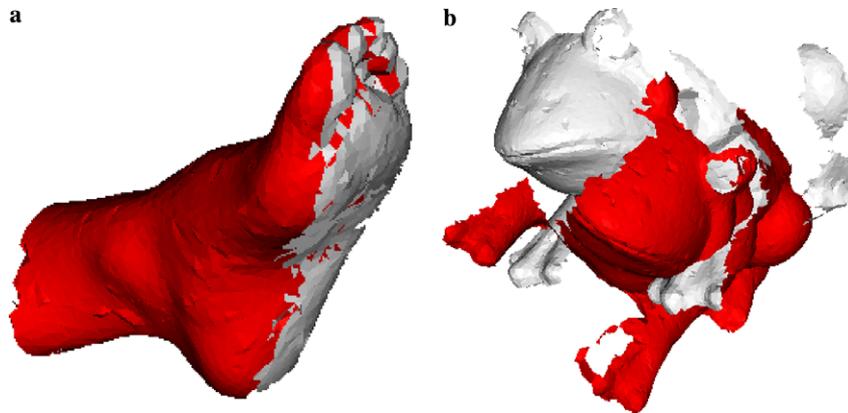


Fig. 7. Coarse registration obtained by the Ransac-based method of Chen: (a) Synthetic data; (b) Real data.

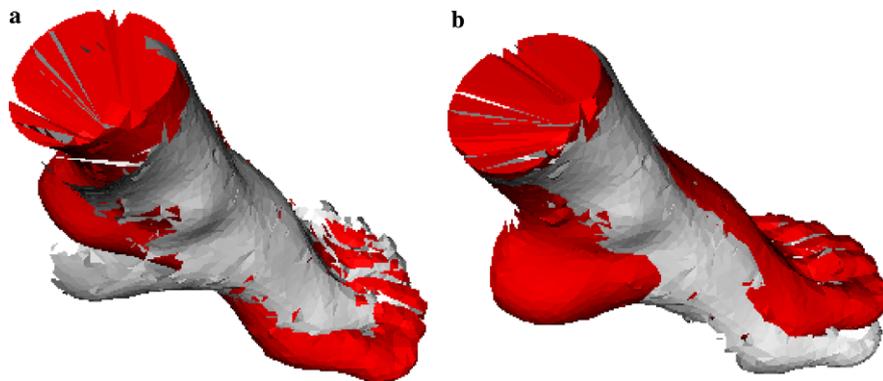


Fig. 8. Influence of the number of points in the accuracy provided by the spin image method: (a) using 400 points; (b) using 1500 points.

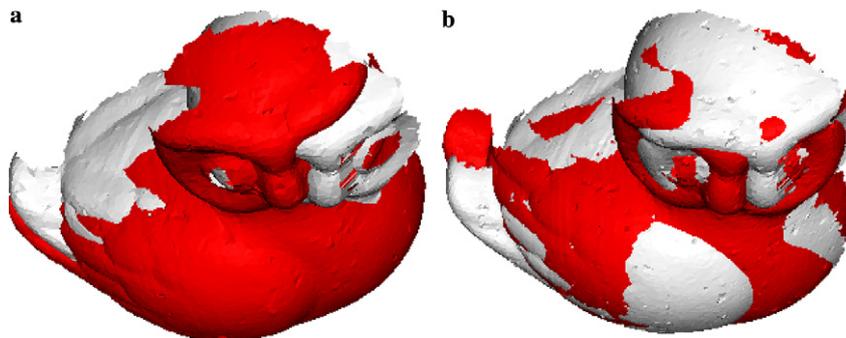


Fig. 9. Inaccurate pair-wise registration of two real range images in the presence of surfaces with few shape details: (a) method of Besl; (b) method of Zinsser.

In this paper, the surveyed methods are explained in detail focusing the analysis on the main contribution of each technique and especially detailing the pros and cons between them. Experimental results using both synthetic and real data are reported providing a useful comparison of the methods, which is a worthwhile information that is not usually provided by authors.

Analyzing the results of coarse registration methods, we conclude that in the presence of low resolution views

Chen's Ransac-based method is the one that obtains better results, especially in synthetic models where the points are coincident and only related to a transformation matrix. Besides, in real images where points might not coincide, the results obtained are less accurate (see Fig. 7). Moreover, when the amount of points used in the registration increases, the computing time required to get the estimation is really a problem. In this case the best method becomes spin image, whose accuracy

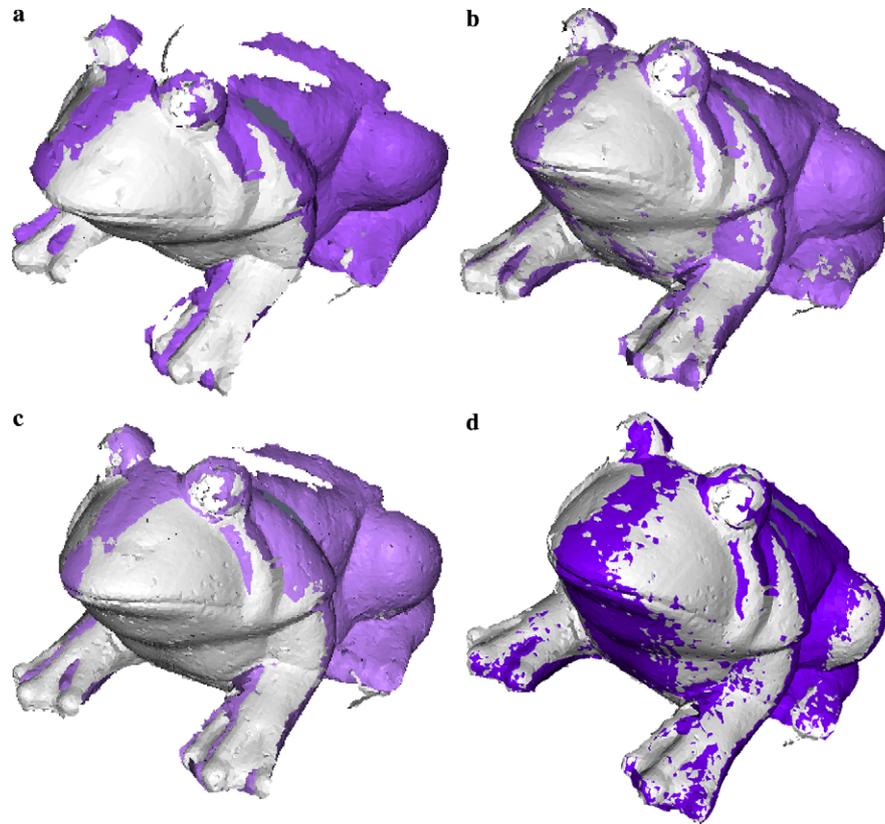


Fig. 10. Pair-wise registration of two real range images using some of the most frequently used fine registration methods: (a) method of Besl; (b) method of Zinsser; (c) method of Trucco; (d) method of Chow.

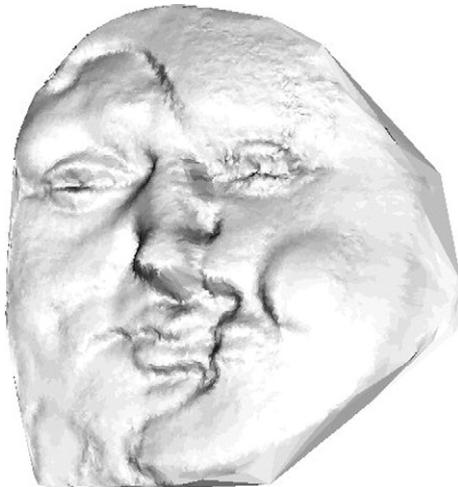


Fig. 11. Registration of a 3D object by means of the alignment of 27 range images.

depends on the number of points used (see Fig. 8) while the computing time remains important. If computing time is critical, the Principal Component Analysis is the best fast method. The problem of this method is that it is based on aligning the three principal axes of the cloud of points, so that the given solution may suffer from symmetries of 90° or 180° . This problem can be solved mod-

ifying the method to find the solution that maximizes the number of overlapping points. However, in this case, the solution obtained is very bad when the overlapping region is not significant. Finally, a genetic algorithm is robust to noise and the results are quite good, but again a lot of time is required to reach a solution.

In addition, considering all the fine registration methods, the Chen method is the best one from our point of view. This method solves the local minima problem presented by ICP variants. Although the point-to-plane distance is theoretically difficult to compute, the iterative algorithm proposed by Chen is very fast and efficient. Furthermore, this method can work with non-overlapping regions, because points whose projection is far from the correspondences are not considered.

On the other hand, results provided by the method of Besl have the problem of local minima, presenting errors even without the presence of noise. Moreover, Besl's approach can not work with non-overlapping regions. In order to solve this problem, it is necessary to use a robust method like the ones proposed by Zinsser or Trucco. The main difference between the two of them is the way of computing the overlapping region. Although both strategies are good, the results obtained indicate that Zinsser's method is more accurate than the one proposed by Trucco's. However, Trucco's method is faster in the presence of high

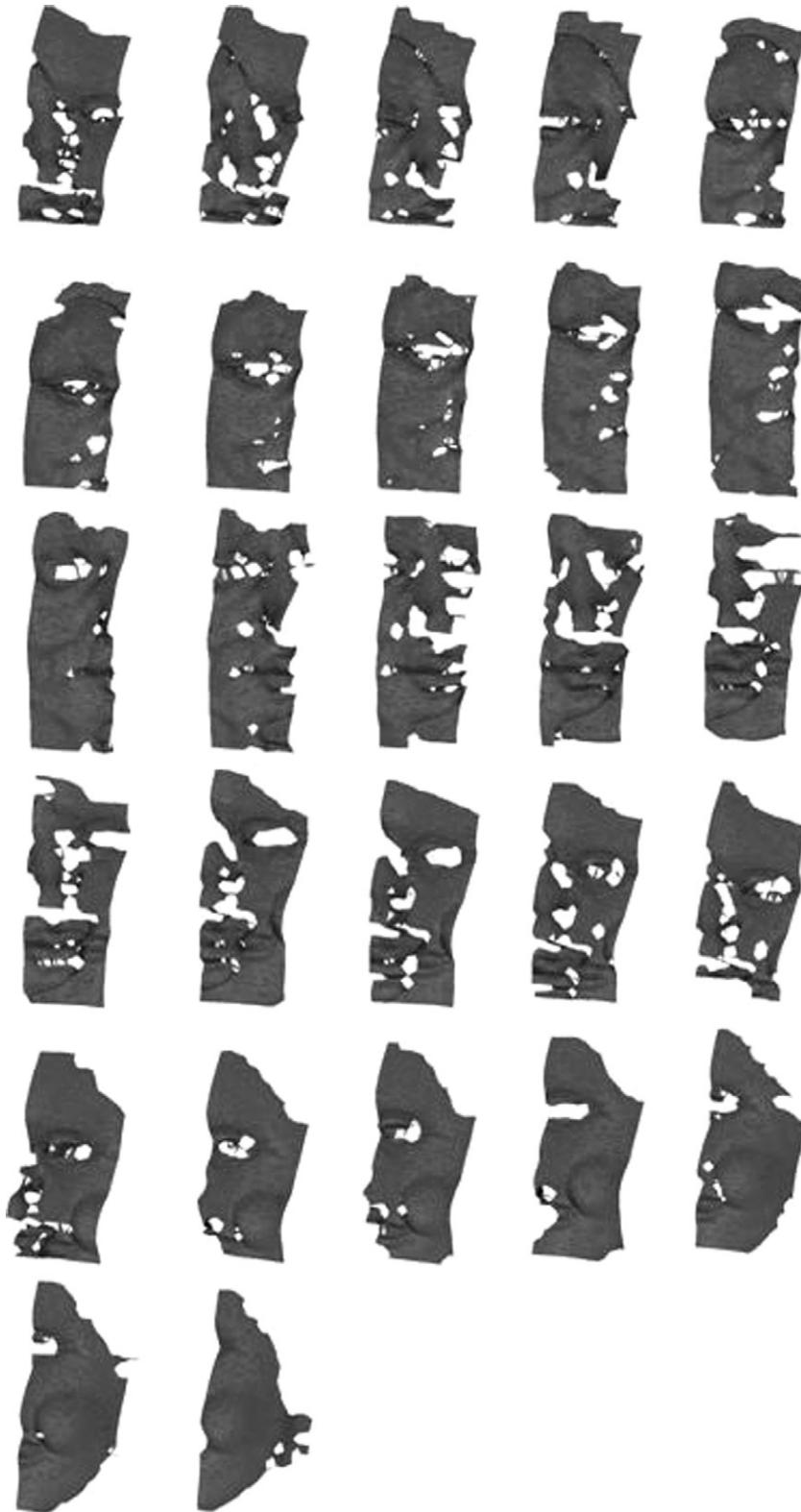


Fig. 12. Set of 27 range images used in the registration.

resolution surfaces. The variant proposed by Jost obtains good results considering that the process has been sped up considerably.

A useful Matlab toolbox including all the programmed methods and test algorithms is available at <http://eia.udg.es/~cmatabos/research.htm>.



Fig. 13. Picture of the large scale object used in multi-view registration.

Acknowledgments

The authors thank Shariar Negahdaripour of the University of Miami for the comments and advices provided during a sabbatic in the University of Girona. The authors thank also Timo Zinsser and Szymon Rusinkiewicz for providing the synthetic images used in their papers.

References

- [1] A. Ullrich, N. Studnicka, J. Riegl, S. Orlandini, Long-range high-performance time-of-flight-based 3d imaging sensors, in: 3D Data Processing Visualization and Transmission, Padova, Italy, 2002, pp. 852–855.
- [2] J. Forest, J. Salvi, An overview of laser slit 3d digitasers, in: International Conference on Robots and Systems, Lausanne, 2002, pp. 73–78.
- [3] C. Matabosch, J. Salvi, J. Forest, Stereo rig geometry determination by fundamental matrix decomposition, in: Workshop on European Scientific and Industrial Collaboration, 2003, pp. 405–412.
- [4] J. Salvi, J. Pagès, J. Batlle, Pattern codification strategies in structured light systems, *Pattern Recogn.* 37 (4) (2004) 827–849.
- [5] S. Rusinkiewicz, M. Levoy, Efficient variant of the ICP algorithm, in: 3rd International Conference on 3-D Digital Imaging and Modeling, 2001, pp. 145–152.
- [6] G. Dalley, P. Flynn, Pair-wise range image registration: a study in outlier classification, *Comput. Vis. Image Und.* 87 (1–3) (2002) 104–115.
- [7] C. Schütz, T. Jost, H. Hügli, Semi-automatic 3d object digitizing system using range images, in: Proceedings of the Third Asian Conference on Computer Vision, vol. 1, 1998, pp. 490–497.
- [8] Z. Zhang, Iterative point matching for registration of free-form curves and surfaces, *Int. J. Comput. Vision* 13 (2) (1994) 111–152.
- [9] C.J.R. Chua, Point signatures: a new representation for 3d object recognition, *Int. J. Comput. Vision* 25 (1) (1997) 63–85.
- [10] C.-S. Chen, Y.-P. Hung, J.-B. Cheng, A fast automatic method for registration of partially overlapping range images, in: International Conference on Computer Vision, Bombay, 1998, pp. 242–248.
- [11] A. Johnson, M. Hebert, Using spin images for efficient object recognition in cluttered 3d scenes, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (5) (1999) 433–449.
- [12] J. Feldmar, N. Ayache, Rigid, affine and locally affine registration of free-form surfaces, Tech. rep., Technical Report of INRIA, Sophia Antipolis (March 1994).
- [13] K. Brunnström, A. Stoddart, Genetic algorithms for free-form surface matching, in: International Conference of Pattern Recognition, Vienna, 1996, pp. 689–693.
- [14] A. Johnson, Spin-images: a representation for 3-d surface matching, Ph.D. thesis, Carnegie Mellon University, USA, 1997.
- [15] O. Carmichael, D. Huber, M. Hebert, Large data sets and confusing scenes in 3-d surface matching and recognition, in: Second International Conference on 3-D Digital Imaging and Modeling, Ottawa, Ont., Canada, 1999, pp. 258–367.
- [16] D. Huber, M. Hebert, A new approach to 3-d terrain mapping, in: International Conference on Intelligent Robotics and Systems, 1999, pp. 1121–1127.
- [17] D. Chung, Y.D.S. Lee, Registration of multiple-range views using the reverse-calibration technique, *Pattern Recogn.* 31 (4) (1998) 457–464.
- [18] S. Kim, C. Jho, H. Hong, Automatic registration of 3d data sets from unknown viewpoints, in: Workshop on Frontiers of Computer Vision, 2003, pp. 155–159.
- [19] C.-S. Chen, Y.-P. Hung, J.-B. Cheng, Ransac-based darces: a new approach to fast automatic registration of partially overlapping range images, *IEEE Trans. Pattern Anal. Mach. Intell.* 21 (11) (1999) 1229–1234.
- [20] J. Tarel, H. Civi, D. Cooper, Pose estimation of free-form 3d objects without point matching using algebraic surface models, in: Proceedings of IEEE Workshop on Model-Based 3D, 1998, pp. 13–21.
- [21] I. Stamos, M. Leordeanu, Automated feature-based range registration of urban scenes of large scale, in: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, 2003, pp. 555–561.
- [22] J.V. Wyngaerd, Automatic crude patch registration: Toward automatic 3d model building, *Comput. Vis. Image Und.* 87 (2002) 8–26.
- [23] C. Chen, I. Stamos, Semi-automatic range to range registration: a feature-based method, in: International Conference on 3-D Digital Imaging and Modeling, 2005, pp. 254–261.
- [24] P. Besl, N. McKay, A method for registration of 3-d shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (2) (1992) 239–256.
- [25] Y. Chen, G. Medioni, Object modeling by registration of multiple range images, in: IEEE International Conference on Robotics and Automation, 1991, pp. 2724–2729.
- [26] G. Turk, M. Levoy, Zippered polygon meshes from range images, in: SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques, Orlando, Florida, 1996, pp. 311–318.
- [27] T. Masuda, Generation of geometric model by registration and integration of multiple range images, in: Third International Conference on 3-D Digital Imaging and Modeling, 2001, pp. 254–261.
- [28] T. Masuda, K. Sakaue, N. Yokoya, Registration and integration of multiple range images for 3-d model construction, in: Proceedings of the 13th International Conference on Pattern Recognition, vol. 1, 1996, pp. 879–883.
- [29] G. Godin, M. Rioux, R. Baribeau, Three-dimensional registration using range and intensity information, in: Proceedings of SPIE Videometric III, vol. 2350, 1994, pp. 279–290.
- [30] Z. Zhang, Iterative point matching for registration of free-form curves, Tech. rep., Technical Report of INRIA, Sophia Antipolis (March 1992).
- [31] M. Greenspan, G. Godin, A nearest neighbor method for efficient ICP, in: Third International Conference on 3-D Digital Imaging and Modeling, Quebec City, Canada, 2001, pp. 161–168.
- [32] T. Jost, H. Hügli, A multi-resolution scheme ICP algorithm for fast shape registration, in: First International Symposium on 3D Data Processing Visualization and Transmission, 2002, pp. 540–543.
- [33] G. Sharp, S. Lee, D. Wehe, ICP registration using invariant features, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (1) (2002) 90–102.
- [34] G. Godin, D. Laurendeau, R. Bergevin, A method for the registration of attributed range images, in: 3DIM01 Third International Conference on 3D Digital Imaging and Modeling, Québec, Canada, 2001, pp. 179–186.
- [35] E. Trucco, A. Fusiello, V. Roberto, Robust motion and correspondences of noisy 3-d point sets with missing data, *Pattern Recogn. Lett.* 20 (9) (1999) 889–898.

- [36] P. Rousseeuw, A. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 1987.
- [37] T. Zinsser, H. Schnidt, J. Niermann, A refined ICP algorithm for robust 3-d correspondences estimation, in: *International Conference on Image Processing*, 2003, pp. 695–698.
- [38] H. Gagnon, M. Soucy, R. Bergevin, D. Laurendeau, Registration of multiple range views for automatic 3-d model building, in: *Computer Vision and Pattern Recognition*, 1994, pp. 581–586.
- [39] T. Masuda, Object shape modelling from multiple range images by matching signed distance fields, in: *First International Symposium on 3D Data Processing Visualization and Transmission*, 2002, pp. 439–448.
- [40] C. Chow, H. Tsui, T. Lee, Surface registration using a dynamic genetic algorithm, *Pattern Recogn.* 37 (1) (2004) 105–117.
- [41] L. Silva, O. Bellon, K. Boyer, Enhanced, robust genetic algorithms for multiview range image registration, in: *3DIM03. Fourth International Conference on 3-D Digital Imaging and Modeling*, 2003, pp. 268–275.
- [42] C. Matabosch, J. Salvi, X. Pinsach, R. García, Surface registration from range image fusion, in: *IEEE International Conference on Robotics and Automation*, New Orleans, 2004, pp. 678 – 683.
- [43] K. Pulli, Multiview registration for large data sets, in: *Second International Conference of 3-D Digital Imaging and Modeling*, 1999, pp. 160–168.
- [44] G. Sharp, S. Lee, D. Wehe, Multiview registration of 3d scenes by minimizing error between coordinate frames, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (8) (2004) 1037–1050.