ABCM Symposium Series in Mechatronics - Vol. 2 - pp.81-88

IMPLEMENTATION OF 3D SHAPE RECONSTRUCTION FROM RANGE IMAGES FOR OBJECT DIGITAL MODELING

Landecir A. Albuquerque
Automation and Control Group (GRACO), Department of Mechanical Engineering, Faculty of Technology, University of Brasilia - 70910-900 – Brazil
lander@unb.br

José Mauricio S. T. Motta
Automation and Control Group (GRACO), Department of Mechanical Engineering, Faculty of Technology, University of Brasilia - 70910-900 – Brazil
jmmotta@unb.br

Abstract. *The necessity of obtaining geometric models in three-dimension that represent with precision a real world object is becoming common each day. For this, one has to recur to methods of 3D Modeling. Three-dimension models have application in several areas, amongst which one can cite photogrammetry, archaeology, reverse engineering, robotic guidance, virtual reality, medicine, cinema, game programming, and others. A current challenge is the construction of 3D models digitized with precision enough to be used in manufacturing systems or numerical simulation of the performance of machines and components in operation, such as turbines and flows in non-circular ducts when the geometric model is not available. The reconstruction of 3D shapes of objects or scenes from range images, also known as depth maps, is preferable than using intensity images or stereoscopy. These maps represent information of distances measured from an observer (optical sensor or camera) to the scene in a rectangular grid. Therefore, the 3D information is explicit and will not need to be recovered as in the case of intensity images. The reconstruction process presents three stages. The first one is sampling of the real world in depth maps. The second stage is the alignment of several views within the same coordinate system, known as image registration. The third stage is the integration of the views for the generation of surface meshes, named merging. The current challenges converge to searching methods that meet with the highest number of desirable properties, such as robustness to outliers, efficiency of time and space complexity and precision of results. This work consists in the discussion of different methods dealing with 3D shape reconstruction from range images found in the literature and in the implementation of the second phase of 3D reconstruction: range image registration.*

Keywords: *Image Registration, Range Images, Surface Reconstruction, 3D Shape*

1. Introduction

Commonly, in manufacture processes a 3D model of the product prototype is first constructed in CAD (Computer Aided Design) software. Next the computer model is imported to CAM (Computer Aided Manufacturing) software such that the product is manufactured using some manufacturing process as for example machining or rapid prototyping. Reverse Engineering works in the opposite way. The real, physical product, already exists and it is necessary to build up its computer model (Eggert et al., 1998, Cerrada et al., 1990). Three-dimension reconstruction can be used many times in this task, since a prototype can pass through several times between the real and virtual world, and vice versa, until the desired result is attained. The priority of 3D Reconstruction in Reverse Engineering is the precision of the obtained models.

The reconstruction process presents the following stages that are described shortly in this article (Dorai, Weng and Anil, 1997, Pulli, 1997, Chen and Medioni, 1992): i) data acquisition from multiple viewpoints (for short, these images are called views); ii) registration of range images, and iii) integration of views. This research has for focus the implementation of the second stage: registration of range images.

The method used to align the images in this work is the ICP (Iterative Closest Point) algorithm and the models obtained are mainly used for applications towards reverse engineering to CAD 3D models.

2. Modeling Based on 3D Reconstruction

Three-dimension models are an essential resource for several areas. As examples one can cite:

a) Modern industry: the availability of digital models used for design and improvement of prototypes has already become a decisive factor for productivity in companies and quality of products. Manufacture is one of the most benefited areas (Fig. 1).

b) Autonomous navigation: modeling by reconstruction has been studied intensively in recent times as a method to get three-dimensional environment maps for autonomous navigation. It is not uncommon that mobile robots are equipped with long distance range sensors (Ladars) for the specific task of environment 3D mapping;

c) Entertainment Industry: the entertainment industry has been stimulating the development of 3D modeling techniques. Since it has virtual reality as priority, creatures and objects to be used in films and 3D games are commonly generated by digitalization of a real model.

Figure 1: Reverse engineering for building a CAD model of a turbine rotor - GRACO/UnB. In the left, a vertical section of a Kaplan turbine. To the center, a single-blade turbine mockup. To the right, the computer model of the turbine rotor obtained from a contact profile digitizer.

d) Art and archaeology: the reconstruction of historical sculptures for study, preservation or creation of virtual museums;
e) Medicine: virtual models are largely used in medicine for surgical planning (e. g., plastic surgeries), prosthesis construction, etc.

The area of 3D Modeling is relatively young and as hardware and software technology advances and costs become lower new applications are thought.

Three-dimension models are traditionally constructed by using two methods: the first is related to Computer Graphics and consists of synthesizing the real world. This literally means to picture the three-dimension subject as an object or a real world scene. This approach is not recommended to all applications, since it is very laborious and it does not produce realistic results when the object has some shape complexity. The second method is named reconstruction. It consists in reconstructing the real world from images taken by sensors and matching them in order to build the model. This approach is traditionally the aim of research of Computer Vision.

The latter process consists of digitizing the object from several viewpoints using preferentially 3D digitizers (rangefinders) based on the principle of active sensing. These digitizers acquire samples of the object surface in the form of 2D matrices representing distances from the sensor to the object. These samples are commonly called depth maps or range images.

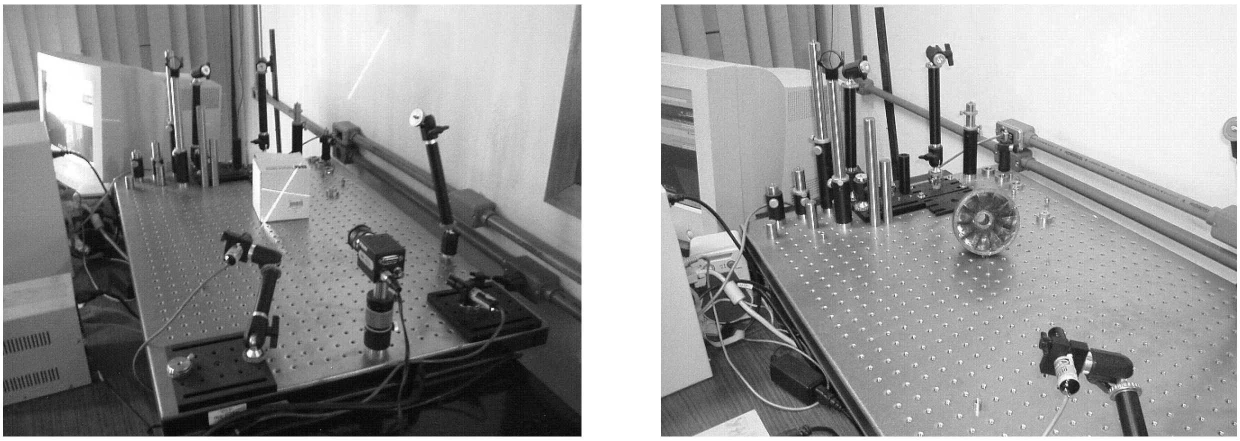
Following the image acquisition process it is necessary to align the images. Each image is acquired centered in the sensor coordinate system. The relative motion between the sensor and the object is a rigid body transformation recovering direction and magnitude of these motions (translations and rotations). These transformations can be applied successively to the image views to bring them together in a common coordinate system. Finally, when the maps are aligned it is possible to assemble them in order build a single model.

General solutions for the whole process do not exist, nor for each stage of reconstruction. The techniques to be employed in each stage depend on several factors that ranges from the cost of the sensoring system to object surface characteristics. All the stages are important and decisively affect the reconstructed model. However, the most important factor for choosing a technique is the type of model application.

3. Data Acquisition

One of the attractions for modeling using reconstruction is the large amount of 3D digitizers available in the market today (Petrov *et al.* , 1998) in comparison with some decades ago.

Three-dimension shape digitizers, also known as rangefinders, capture surfaces or even object volumes. Choosing the correct digitizer depends on some factors: type of application, object size and finally the cost. Currently, complete commercial packages already exist, including hardware and software to construct 3D models of free shape. However, in many of the cases it is necessary to construct a rangefinder from scratch, as in the acquisition setup showed in Fig. 2.



(a) (b)

Figure 2: Setup for range images acquisition - GRACO/UnB: Photograph of the experimental setup, showing camera, diodes and laser light planes on a (a) plain object and a (b) reduced turbine model.

Digitizers use different sensing techniques to interact with the scene and measure its distance from the sensor, such as contact devices, optical or acoustic probes, and others. Curless (1997) conceived a taxonomy that looks forward to organizing all these techniques (Fig. 3).

The techniques for 3D-shape data acquisition normally use passive and active optic sensors.

Passive techniques acquire surfaces from intensity images, while active techniques acquire a surface geometry projecting energy to the object surface. Due to this interaction with the object surface, active techniques are highly affected by the object surface properties (Curless, 1997).

A very used passive optical technique is stereopse. This technique involves two calibrated intensity cameras. The depth information is extracted by means of triangulation. The largest problem with this technique is the generation of sparse depth maps.

Amongst active optical techniques, the most popular are: Light Structure, also known as active laser triangulation, and Laser Imaging (Ladar).

A system of structured light is composed by calibrated cameras and a laser projector, i. e., the location relative to each other is known. Thus, it is possible to obtain depth information by triangulation.

For long distances such as in applications like autonomous navigation or photogrammetry the use of ladars is recommended. The great advantage of this type of rangefinder is its versatility to work embedded on a mobile vehicle. Ladars are expensive systems and use the TOF (Time of Flight) principle to recover depth.

Amongst active digitizers the most attractive ones use coherent illumination such as the low intensity laser. They produce depth maps with a rapid and high sampling tax in contrast to contact digitizers that are precise but also slow, disjointed and produce maps with a low sampling tax.

In applications such as mobile robotics ladars are the most sophisticated solution for environment 3D modeling. In the industry, rangefinders of active optical triangulation are the most used.

In practice, the cost of hardware of active methods is higher than passive methods. However, passive methods demand more complex software for reconstruction, which results also in expensive systems.

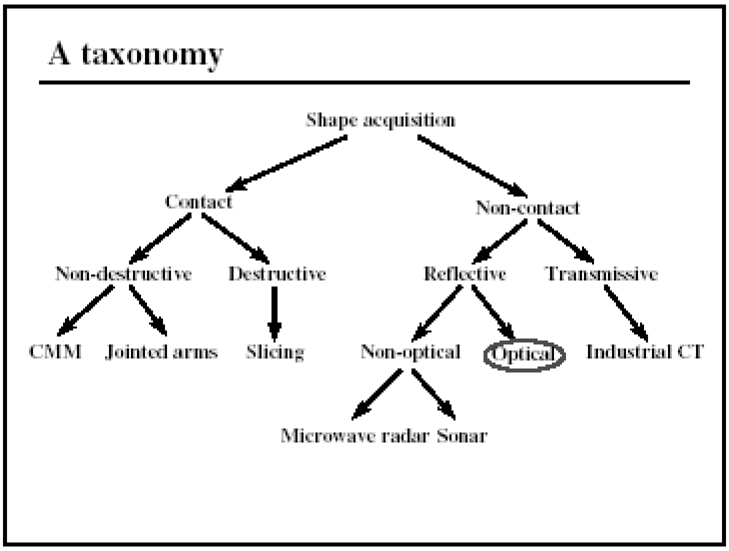
3. Image Registration

For the construction of 3-D models from range images it is highly desirable that the entire object surface is digitized. As each digitizer scan occurs in only one direction it is impossible that the entire object surface is digitized in only one pass. Therefore, more than one image has to be taken, coming up the need to align the set of acquired images. In the Computer Vision community this problem of alignment is known as Image Registration.

3.1 Previous Work

Image registration is a problem of crucial importance in computer vision and much research involves the subject. Up to date many methods had been developed and new methods are constantly being proposed, most of times, aiming at a solution for a specific application.

Image registration is commonly assorted as an optimization problem (Blais and Levine, 1993), since it aims at searching the parameters of an optimal rigid motion amongst a class of possible ones between two images. What in general distinguishes a registration method from another is the form they search for this optimal motion transformation.



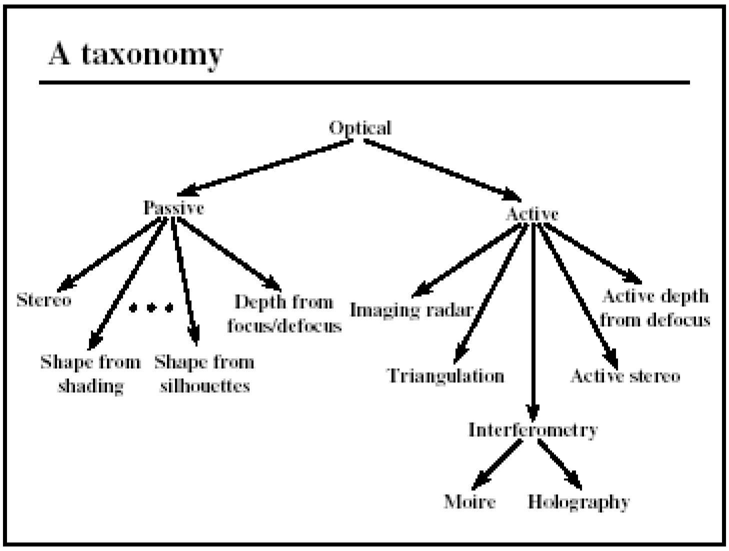


Figure 3. Curless Taxonomy for the 3D shape acquisition methods (Curless,1997).

Although it is difficult to sort out all the range of existing methods in the literature, many researchers divide the methods into two main categories (Blais & Levine, 1993, Dorai, Weng and Anil, 1997): a) the ones the data acquisition process is controlled in some way, either by means of a robotic manipulator or a revolving platform, and use the calibration parameters for the registration, and b) the ones that obtain the transformations straight from data. However, better results can be possibly obtained combining the two methods.

The first registration technique was based on matching of discrete features. This type of approach consists in the extraction of local image features invariant to rigid motion such as conical or polygons, for example. The motion estimation is obtained from the correspondence between features overlapping image areas. The problem with this approach is related to the extraction and sorting of these features.

Iterative approaches are more recent than the previous ones and usually search for the optimal transformation through iterative refining of an initial transformation. An example of those approaches is the ICP (Iterative Closest Point) algorithm, one of the most popular registration methods, which will be described in more details ahead. Approaches of this type are fast, easy to implement and are based on the minimization of a cost function. However, some assumptions need to be fulfilled for attaining some guarantee of convergence, such as having a good initial estimate.

Optimization approaches that search for the optimal transformation in the space of transformations also exist. Stochastic optimization are used with these methods such as Simulated Annealing (Blais & Levine, 1993), or techniques of robust computer vision such as RANSAC (Chen et al., 1999).

3.2 Pairwise and Simultaneous Registration

There are two distinct strategies to register images: a) Local or Pairwise Registration and b) Global, Simultaneous or Parallel Registration.

The idea of pairwise registration is to divide the entire process into stages. The direct strategy is to focus on two images each time, and to register one relative to the other. Next to an image pair is registered, a new pair, including a range image of the formed pair previously, is registered in the resulted coordinate system. This is repeated until all the images are used (Nishino and Ikeuchi, 2002).

The advantages of the pairwise registration are the low computational cost in relation to the global registration. The disadvantage is a lower accuracy in the final result. The errors of each stage are added to the errors of the previous one such that in the last stages there will be a considerable accumulation of registration errors.

The global registration of multiple range images solves the problem of error accumulation by registering all views at the same time. When registering all images at the same time the registration errors are spread among them.

The number of data sets on which the local and global strategies operate and how the correspondence problem is dealt with consist the main differences between the two approaches. Approaches in pairs operate on two sets of points, with correspondences one-to-one defined between these sets. Global registration, on the other hand, involves multiple sets of points with multiple sets of correspondence between them (Williams and Bennamoun, 2001).

The disadvantage of the global registration is the high computational cost, mainly the requirement of large spent of memory.

3.3 The ICP Algorithm

Introduced by Besl and McKay (1992), the algorithm is an iterative and fast approach and of easy implementation of the 3D data alignment problem.

The ICP presents several steps to which heuristics can be adjusted for turning it to be faster or more accurate, giving rise to a family of algorithms. Heuristics are approximation methods to solve problems in polynomial time complexity, and meta-heuristics are general purpose methods that give good solutions, but the optimal solution is not assured (Viana,1998).

Rusinkiewicz and Levoy (2001) proposed a classification system and compared different ICP variants based on the six stages of training identified in the algorithm: Selection, Matching, Weighting, Rejecting, error Metric and Minimizing.

The ICP algorithm carries out a 3D regression exploring the redundant points in the images to calculate the motion parameters and to approach the views (Arun, Huang and Blostein, 1987, Umeyama, 1991). That means, it is assumed that the images overlap. Another requirement is to supply an initial register estimation such that the algorithm refines it iteratively (Besl and McKay, 1992). The algorithm stops when a certain level of precision in the overlapping between images reaches a threshold.

The first stage of the algorithm is the selection of control points within an image. Next these points are matched to their nearest pair within the next image. This pair of points is to be associated to the same point on the object surface. This stage is the most challenging and needs a large span of time due to the corresponding problem complexity.

The errors associated to the cost function to be minimized are dependent on the precision by which each pair of points was corresponded. The minimization of this cost function is a nonlinear optimization problem. Depending on the precision of the initial register, this cost function can stick in a local minimum and results in a non-precise alignment.

Next to the minimization stage, the ICP recovers the 3D motion parameters between images and reuses them in one of the images to move them towards each other. The estimation of these parameters can be carried out using either quaternions (Horn, 1987) to represent rotation or the SVD (Singular Value Decomposition) (Arun, Huang and Blostein, 1987).

3.4 Mathematical Formulation of Registration

Consider two sets of points x_i e y_i, where $i = 1,...,N$.

Eq. (1) models the rigid motion between the two sets of points

$$x_i = Ry_i + T \quad (1)$$

where R is a rotation matrix and T is a translation vector.

In the least-square sense, the cost function for Eq. (1) is:

$$\sum^{2} = \sum_{i=1}^{N} \left\| x_i - \left(R y_i + T \right) \right\|^2 \quad (2)$$

Eq.(2) is a cost function based on the ordinary least-square. The optimal rotation and translation parameters are those that transform this cost function in a minimum. It is convenient to uncouple the rotation and translation components.

If the data points are aligned, then x_i e y_i have the same centroid, that means, $\overline{x} = \overline{y}$, e $T^{\bullet} = \overline{x} - R^{\bullet}\overline{y}$ where $T^{\bullet}$ and $R^{\bullet}$ are the optimal transformations.

To recover rotation, the errors p_i e q_i forms the cost function of Eq.(6).

$$p_i = x_i - \overline{x} \quad (4)$$

$$q_i = y_i - \overline{y} \quad (5)$$

$$\sum^{2} = \sum_{i=1}^{N} \left\| p_i - R q_i \right\|^2 \quad (6)$$

After some manipulation Eq.(6) yields Eq.(7) from which the rotation matrix can be uncoupled into two orthonormal basis V and U (Arun, Huang and Blostein, 1987) from the SVD (Singular Value Decomposition) of the covariance matrix H :

$$H = \sum_{i=1}^{N} q_i \, p_i^{\,T} \quad (7)$$

$$H = UDV^T \quad (8)$$

So,

$$R = VU^T \quad (8)$$

Eq.(8) returns a rotation if R is orthonormal ($R^T = R^{-1}$) , and det$|\mathrm{R}|$ =1. If det$|\mathrm{R}|$ =-1, R returns a reflection.

In the ICP algorithm this process is run iteratively until a threshold for the rotation matrix, R, is reached.

4. Integration of Views

After the range images are aligned they need to be integrated to shape the 3D model. The Integration stage consists of the generation of surface representations from the aligned data and the edition of the model. The integration is the process to create a representation of a unique surface from sampled points of two or more images (Turk and Levoy, 1994). This process can be considered a post-processing stage. Free-shape objects can have complex surfaces. Many parts may not be reached by the scanner such as regions with steep curvatures or regions that cannot be reached by the sensor. That may result in flaws in the final model that need to be mended, and sometimes some manual intervention is necessary to correct these problems.

5. Metodology and Experimental Results

Several test runs were carried out with the ICP algorithm with data from a cube (without noise), Fig. 4, and a range image pair of a sculpture (Buddha), Fig. 5.

In the tests with the cube, a regular object with well contrasted edge lines, the ICP algorithm reached the final alignment with few iterations. The estimated rotation angles showed a decreasing order as the iterations proceeded in the cases where there were no false pairs of matching points. In the presence of false pairs the convergence could not be guaranteed.

The Buddha range images used are available in an image database from OSU 3D Database, (Campbell and Flynn, 1998). These images were acquired with 200x200 laser scanning.

Several tests were planned with diverse initial starting positions. The results showed a clear relationship between the speed of convergence and the number of iterations with the proximity the range images were to each other, either results from the tests with the cube or with the Buddha.

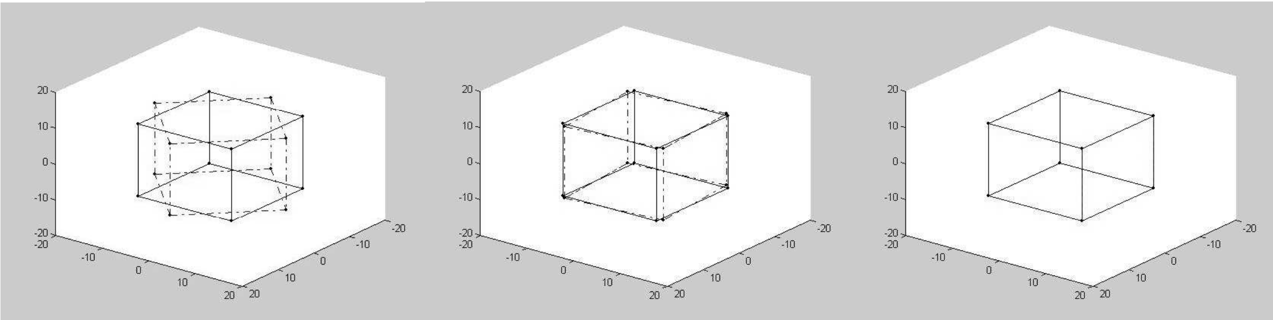


Figure. 4. Registration of two range images of a cube from two different viewpoints using the ICP algorithm. The images were initially rotated 45º around the *z* axis. The iterations stopped when the threshold of 1º was reached.

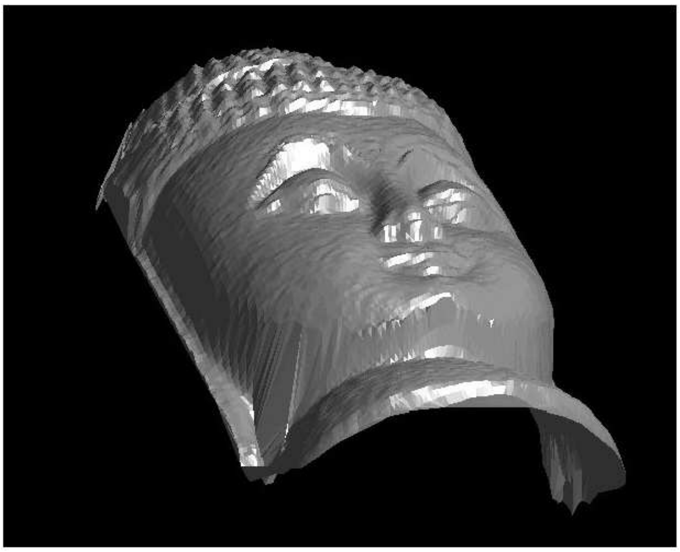


Figure 5. Rendering of a 3D model (Buddha) from a pair of range images acquired by laser scanning.

The experimental results had demonstrated that the convergence of the ICP algorithm is highly influenced by the presence of false pairs of correspondence. The influence of these false pairs decisively affects the convergence due to the error metric used to be based on ordinary least-squares (not robust). Another fact to point out is that, even in the cases without noise, the matching heuristic (in the case, the next point) can fail.

6. Conclusions

The biggest challenge of the ICP algorithm is the convergence. Convergence depends on several factors, such as the quality of the initial register, methods for choosing points and matching, methods for cost function minimization and methods to estimate the 3D transformations.

Speckles and occluded regions are natural in range images acquired by laser scanning. These problems can be tackled introducing some heuristics in the ICP stages turning it to be more robust.

This article presents a descriptive sequence of the main methods used for 3D reconstruction aiming at reverse engineering to produce precise 3D CAD models. The ICP algorithm was implemented and tests were carried out in several different conditions to show up the main factors that influence the final precision of the model.

Experimental results showed that convergence is a key problem with the ICP algorithm, which depends on several factors: initial starting registration point, the method used to choose the initial control points, method to minimize the cost function and method to obtain the 3D motion transformations.

To improve the algorithm accuracy, in future works one can adjust heuristics to avoid false correspondences or, alternatively, a more robust estimator to the error metric can be implemented.

7. References

Arun, K. S., Huang, T. S. and Blostein, S. D., 1987, "Least-Squares Fitting of Two 3-D Point Sets", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No. 5, pp. 698–700.

Besl, P. J. and McKay, N.D., 1992, "A Method for Registration of 3-D Shapes", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, pp. 239-256.

Blais, G. and Levine, M. D., 1993, "Registering Multiview Range Data to Create 3D Computer Objects", Technical Report TR-CIM-93-96, Centre for Intelligent Machines, McGill University.

Campbell, R., Flynn, P., "A WWW-Accessible 3D Image and Model Database for Computer Vision Research", Empirical Evaluation Methods in Computer Vision, Bowyer K.W. and Phillips P.J. (eds.), IEEE Computer Society Press, pp. 148-154, 1998.

Cerrada, C., Ikeuchi, K., Weiss, L., and Reddy, R., 1990, "A 3D-Object Reconstruction System Integrating Range-Image Processing and Rapid Prototyping", Technical Report CMU-RI-TR-90-32, Robotics Institute, Carnegie Mellon University.

Chen, C. S., Hung, Y. P. and Cheng, J. B., 1999, "RANSAC-based DARCES: A New Approach to Fast Automatic Registration of Partially Overlapping Range Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 11, pp. 1229-1234.

Chen, Y. and Medioni, G., 1992, "Object modeling by registration of multiple range images", Image and Vis. Computing, Vol. 10, No. 3, pp. 145-155.

Curless, B. L., 1997, "New Methods for Surface Reconstruction from Range Images", PhD Thesis, Stanford University.

Dorai, C., Weng, J. and Anil, K. J., 1997, "Optimal Registration of Object views using range data", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 10, pp. 1131-1138.

Eggert, D.W., Fitzgibbon, A.W. and Fisher, R.B., 1998, "Simultaneous Registration of Multiple Range Views for Use in Reverse Engineering of Cad Models", Computer Vision and Image Understanding, Vol. 69, No.3, pp. 253–272.

Horn, B. K. P., 1987, "Closed-Form Solution of Absolute Orientation Using Unit Quaternions", Journal of the Optical Society of America A, Vol. 4, No. 4, pp. 629–642.

Lorusso, A., Eggert, D. W. and Fisher, R. B., 1995, "A Comparison of Four Algorithms for Estimating 3-D Rigid Transformations", In British Machine Vision Conference, Birmingham, England, pp. 237–246.

Nishino, K. and Ikeuchi, K., 2002, "Robust Simultaneous registration of Multiple Range Images", In: The 5th Asian Conference on Computer Vision (ACCV2002), pp. 23-25.

Petrov, M. A. Talapov, T. Robertson, A, Lebedev, a. Zhilyaev, L. Polonskiy, 1998 "Optical 3D Digitizers: Bringing Life to the Virtual World", IEEE Computer Graphics and Applications, Vol. 18, No. 3, pp. 28-37.

Pulli, K., 1997, "Surface Reconstruction and Display from Range and Color Data", PhD Thesis, University of Washington.

Rusinkiewicz, S. and Levoy, M., 2001, "Efficient Variants of the ICP Algorithm", In Proc. Third Internat. Conf. 3D Digital Imaging and Modeling, pp. 145-152.

Turk, G. and Levoy, M., 1994, " Zippered Polygon Meshes from Range Images", In Proceedings of SIGGRAPH, pp. 311-318.

Umeyama, S., 1991, "Least-Squares Estimation of Transformation Parameters Between Two Point Patterns". IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 4, pp. 376-380.

Viana, G. V. R, 1998, "Meta-Heurísticas e Programação Paralela em Otimização Combinatória", Fortaleza, EUFC, Ceará, Brasil, 250 p.

Williams, J. and Bennamoun., M., 2001, "Simultaneous Registration of Multiple Corresponding Point Sets", Computer Vision and Image Understanding, Vol. 81, No. 1, pp. 117-142.