# **Scan-Based Registration of Range Measurements**

C. Chen, P. Payeur Vision, Imaging, Video and Audio Research Laboratory School of Information Technology and Engineering University of Ottawa Ottawa, Ontario, Canada [chenadiu,ppayeur]@site.uottawa.ca

Abstract - This paper introduces an automatic approach for registration estimation between successive viewpoints of a laser range camera that takes advantage of the raw measurements and does not require any external device for pose estimation nor complex feature extraction or triangulation. Assuming only object rigidity and some overlap between the scan areas, the approach allows to estimate the six rotation and translation parameters that link 3-D scans gathered from different viewpoints. A compact modified Gaussian sphere representation is used to encode a simple planar patch approximation of the objects surface and to validate mapping between the measurements as the appropriate rotation and translation parameters are computed. This solution results in an important reduction of the computational workload and a sufficient accuracy for most robot navigation applications. The proposed approach is demonstrated in an experimental context using real range measurements collected from a series of viewpoints.

# I. INTRODUCTION

Building virtual representations of 3D environments from range measurements requires that data are gathered from a large number of viewpoints. This requirement results from the complexity of objects to be modelled, from the limited field of view of sensors and from occlusions that occur between objects. Each dataset gathered from a given point of view is defined with respect to a local sensor-based reference frame. As a result, the sensor position and orientation at each viewpoint must be precisely estimated to ensure that the information obtained from every source is merged in a consistent way to build a 3D model. The problem of registration consists of determining the geometric relationship that exists between different views provided by the sensor. An imprecise registration between viewpoints prevents from computing reliable models for collision avoidance or fine interaction between a robot and its environment [11].

The sensor pose can be measured with external means such as magnetic position and orientation trackers, robotic arms or even CCD cameras providing images from which the sensor position and orientation can be extracted. The latter solution implies very complex image processing and pattern recognition algorithms that are time consuming and rarely fully reliable. The first two approaches appear to be more realistic. A magnetic position and orientation tracking device, such as the Fastrak system commercialized by Polhemus Inc. has been tested in our robotic workcell. Unfortunately, the magnetic fields used by the device to track the pose appear to be very sensitive to the environment. In an experimental setup containing quite a large number of metallic parts such as computer boxes, power supplies and robotic equipments, such a device does not succeed in providing the required pose information except in very limited circumstances and under constrained displacements.

When a robotic arm is used to move the sensor from one viewpoint to another, the internal encoders of the robot also provide a good estimate of the sensor position and orientation. But our experiments revealed that there is still room for refinement on this information in order to enhance the quality of the virtual representation of the environment. Moreover, the sensor is then constrained to the robot physical workspace and cannot get an access to narrow areas of the environment. An interesting solution to estimate range sensor registration between successive viewpoints without any peripheral devices is to take advantage of the raw range data provided by the sensor. Assuming that there is an overlap between the areas of the scene that are measured from each viewpoint, it becomes possible to search for some matching characteristics in both sets of information and then compute the necessary registration information that would make the projections of those matching elements to superpose.

In spite of the fact that the registration problem between range measurements has been studied for a while in computer vision, no extensive and definitive solution has been found yet. Many variations to the widely known iterative closest point (ICP) algorithm [1] have been proposed to match characteristic point sets [3, 10], curves, meshes [2, 4] or parametric surfaces [8]. Some of them use both range and intensity data, also provided by most range sensors, to improve their selection of control points that are to be matched [7, 12]. These algorithms generally provide good results but the search for characteristic curves or surfaces is very complex and time consuming.

Moreover, research works on the topic of registration generally assume that full range images are directly available from the sensors. As a result, they search for matching characteristics between such full images and compute geometrical transformations from there. Such a framework does not correspond to the reality because the majority of range sensors currently available on the market or even prototypes found in laboratories do not provide such full images by themselves. They rather generate single points or scan lines of range measurements [6]. Those sensors that generate full images rely on an external mechanical device to translate the sensor or change its orientation [9]. This solution compares to the use of a robot to move the sensor and is sensitive in terms of registration errors. In this paper, we introduce an approach to estimate registration between range scan lines that makes use of a simple representation of raw data as they are provided by a laser range sensor. There is no need for any exhaustive search for features or sophisticated representations. In its normal operation mode, the technique does not require any help from an external position/orientation tracking device to provide an initial estimate of the translation and rotation between successive viewpoints. Computing registration from range data generated independently from any positioning device enhances the flexibility of sensing systems in a variety of applications. The following sections describe how scan lines are processed to create a compact representation of the surface shape. The estimation of the rotation and translation parameters is detailed. Finally, experimental results using real range measurements are presented and discussed.

#### II. PROPOSED APPROACH

The proposed algorithm is based on the fact that most manmade objects are composed of a set of planar surfaces. Even objects of a higher complexity can be approximated by a set of planar surfaces that are easy to represent. Assuming that the rigidity constraint is validated, these planar regions remain unchanged and the same reality is always projected on the image plane of a range sensor. Only the relative position and orientation of the sensor has an influence on the object representation. It is then possible to take advantage of these planar regions to estimate the registration between two or more sets of measurements gathered from different viewpoints. These assumptions are the same that are made in most research work on the problem of registration estimation. However, most of the proposed techniques invest a lot of efforts in building a sophisticated representation of object surfaces before they estimate the position and orientation parameters. Roth [12] provides a nice example of that by proposing a registration technique that is based on the match between surface representations obtained by means of a Delaunay triangulation. The conversion of range measurements into a triangular mesh mapping of the object surface is computationally intensive. As a result, much time is spent on modelling objects with respect to a camera-based reference frame while the actual goal is to estimate the motion of the sensor between viewpoints in order to eventually merge all range measurements in a comprehensive model defined with respect to a single reference frame that has nothing to do with the sensor reference frame. The strategy that is presented here rather relies on a simpler surface representation such that the emphasis is put on the estimation of the position and orientation variations of the sensor between viewpoints in 3-D space. This way, the main part of the effort is dedicated to the registration estimation rather than to an intermediate modelling technique or feature extraction.

A Jupiter laser range sensor able to provide scan lines of up to 256 range measurements is used. As for most practical systems, the resolution of these measurements depends on the depth and on the angle of the laser beam. Also, the spacing between the 256 points on one scan line is not constant. The sensor is moved in small constant increments along a straight line to collect a series of scan lines on the visible surface of the object. The sensor orientation is kept approximately constant. These small displacements are under the control of a F3 robotic arm operated in the world coordinate mode in order to ensure a precise control of the sensor position along the straight line. Then, the sensor is moved to a completely different viewpoint and the process is repeated. Figure 1 shows the experimental setup for two viewpoints.

Our goal is to estimate the translation, T, and rotation, R, parameters between two successive viewpoints from which range measurements are collected. The first step consists in segmenting each range profile to locate the linear sections and to approximate them with straight line segments. This allows to overcome the difficulties associated with irregular spacing between points on the same scan line that would preempt a point-based matching between two scans, even though they are collected on the exact same surface. These line segments provide an efficient way to merge neighbor profiles that show similar shapes into planar patches of various sizes. The normal vectors and the areas of these patches are then used to encode the surface representation as a modified Gaussian sphere which significantly simplifies the search for rotation parameters. Provided two sets of scan lines encoded as Gaussian spheres, the next step consists in estimating the rotation parameters by finding the appropriate rotation parameters that make similar vectors on the spheres to overlap. Finally, the translation parameters associated with the two sets of scan lines, and therefore with the displacement of the sensor between viewpoints, are estimated by computing the necessary shift of the patch centroids to make corresponding planar patches to match. Figure 2 illustrates the data flow of the proposed approach.

# **III. COMPACT REPRESENTATION OF SURFACES**

The proposed approach relies on a simple representation of surface shape that can be easily and quickly computed from the raw range measurements in order to estimate the rotation and translation parameters between different viewpoints. This section describes the two-step procedure that is used to obtain the modified Gaussian sphere representation of surfaces.



Figure 1: Range sensor scanning from two viewpoints.



Figure 2: Steps of data processing for registration estimation.

# III.1. Segmentation of range profiles

Starting from the raw range measurements, each scan line is segmented into a sequence of straight lines that best match the set of 256 points. Before a straight-line fitting technique is applied, outliers are removed from the scans and a median filter is applied on the range profiles to reduce the effect of noise. This eliminates large deviations that often occur in raw measurements. Next, the straight line segments are computed as shown in figure 3. The process starts at one extremity of the scan line. The first three measurements are initially used to determine the orientation of the straight line. The following points encountered along the scan are then successively checked for their proximity in terms of the distance with respect to the initial line estimate. When the deviation between a measurement point and the straight line estimate is below a given threshold (4 mm in our experiments), then the point is integrated into the straight-line segment. A segment is terminated when one measurement point is located too far away from the straight line estimate. In this case, a new straight-line segment is initialized and the remaining points in the scan line are successively checked for their proximity with this new segment. The process is repeated in the same way until the end of the profile is reached. Once a segment is terminated, it is revisited in order to refine its estimation. The parameters of each straight-line segment are computed on the basis of the entire set of measurement points that are associated with it rather than only with the estimate of the first three points.

The straight-line segments are next parameterized to obtain a compact representation of the profile of the surface along the scan line. The normal vector and the center point of each segment are computed. The length of the segment is used to define the length of its normal vector. As a result, a scan line of 256 measurements can be encoded as a set of 2aparameters, where *a* corresponds to the number of straight line segments associated with a given scan as shown in figure 4. This simplified notation allows both the compactness of the representation and an approximation of the surface structure along a single line in 3-D space.

#### III.2. Scan merging

The line segmentation previously obtained provides an efficient way to merge similar neighboring range profiles and to create a simple patch-based surface representation. Assuming that the gap between two successive scan lines is kept small enough to ensure a proper coverage of the object surfaces (5 mm in our experiments), the variation of the shape



Figure 3: Straight line fitting on a sequence of range measurements.

between successive profiles should be smooth. As a consequence, these profiles should have a similar normal vector distribution where the object surfaces are continuous, which is the case in most man-made environments. The vector representations for each scan obtained in the previous step are then compared with that of the first profile to measure the similarity between scans. If the angle and length of vectors associated with each straight-line segment are within a given deviation (0.025 rad for angle and 25% in length in our experiments), then neighbor scan lines can be merged to create patches associated with each of their segments.

On the other hand, transitions between objects appear as abrupt changes between two successive profiles. When such a transition occurs, the respective profiles are assigned to different neighbor patches as shown in figure 5. For example, scans 1 to 4 have been collected on a continuous surface and can therefore by merged. As these scans have been previously segmented, a different surface patch is created for each scan segment (patches 1 to 5 inclusively). The same idea applies to scans 5, 6 and 7 as they are similar. However, there is an important transition between scans 4 and 5 which are successive profiles in the raw data set. Locating this transition allows to delimitate the boundaries between patches associated with the first and the second group of scan lines respectively.

This way, planar patches are defined for each group of merged straight-line segments. Three points from each segment (one in the middle and two at the extremities) are considered to estimate the patch orientation. Figure 5 illustrates the process for the first patch as  $p_1$ ,  $p_2$  and  $p_3$  are selected from the first segment in scan 1, while  $p_4$ ,  $p_5$  and  $p_6$  are extracted from the same segment in the second profile. Knowing the distance between two successive profiles, the *X*, *Y* coordinates of  $p_4$  can be estimated. The corresponding *Z* coordinate of  $p_4$  is then extracted from the first segment of the second profile. The same steps also apply to estimate the coordinates of  $p_5$  and  $p_6$ . Eventually the process is repeated to locate  $p_7$  to  $p_9$  from  $p_4$  to  $p_6$  and so on for as many scans as necessary to fully define a given planar patch.

When the 3-D coordinates of all triplets are known for a given patch, the best fitting planar surface parameters are computed. For each planar patch previously defined, the center point coordinates,  $cp_i$ , the normal vector,  $\vec{n_i}$ , and the area are computed. The center point is defined as the average of the coordinates of all triplet points belonging to the patch.

$$cp_i = \frac{1}{N} \sum_{i=1}^{N} p_i \tag{1}$$



Figure 4: Compact representation of a segmented range profile.



Figure 5: Definition of patches as merged segmented scans.

The normal vector,  $\vec{n_i}$ , corresponds to the smallest eigen vector of the matrix  $\alpha^T \alpha$ , where:

$$\alpha = \begin{bmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ \dots & \dots & \dots & 1 \\ x_N & y_N & z_N & 1 \end{bmatrix}$$
(2)

with  $p_i = (x_i, y_i, z_i)$ .

The compact representation used for line segments is extended to 3-D space under the form of a modified Gaussian sphere representation to encode the distribution of patches that describe the surface as shown in figure 6. Here, the length of normal vectors defines the area of the planar patches.

Given the relatively large number of planar patches that might exist in a real 3-D scene, we observed that this representation can still be advantageously compacted to a limited number of normal vectors. Specifically, those normal vectors which have a small angle difference in between them can be merged into a single vector and their respective lengths added. Merging them is equivalent to the unification of the corresponding patches that must have a very similar orientation (within some tolerance level) with no consideration of their respective depth. Figure 6 shows the metric of co-normality that we introduced to select normal vectors that can be merged.  $\vec{n_1}$  and  $\vec{n_6}$  are two normal vectors corresponding to two patches.  $p_1$  and  $p_6$  are their projection points on the unit sphere surface. If the distance between  $p_I$ and  $p_6$  is less than a given threshold (0.14 mm in our experiments), then  $\vec{n_1}$  and  $\vec{n_6}$  can be merged. The resulting vector  $\vec{n}$  has the total length of  $\vec{n_1}$  and  $\vec{n_6}$ . In practice, the number of normal vectors is significantly reduced by this process. This opportunity to further reduce the complexity of the representation is provided by the fact that most objects can be approximated by a large number of separated patches, many of them having a similar orientation.

In comparison with the thousands of 3-D points collected by the range sensor, this simplified representation is much more compact. As a result, the computation time required to match the representations of scans decreases dramatically. In spite of some lost into the resolution of the representation, this strategy proves to be sufficient for many applications while it significantly speeds up the computation in comparison with a classical triangular mesh representation for which a large number of facets are required to approximate the object surface.



Figure 6: Compact representation of a surface shape.

# IV. REGISTRATION ESTIMATION

Using the modified Gaussian sphere representation for each planar patch, the three rotation parameters between successive viewpoints are directly estimated from the set of normal vectors. Then the three translation parameters can be computed from the center points coordinates,  $cp_i$ , while taking into account the rotation values previously obtained.

#### IV.1. Rotation parameters estimation

Three correspondences of normal vectors are required to uniquely determine the rotation matrix, R, from which the three rotation angles,  $(\theta, \phi, \psi)$ , can be computed [13]. Provided that  $\vec{n_1}$  and  $\vec{n_1}$ ,  $\vec{n_2}$  and  $\vec{n_2}$ ,  $\vec{n_3}$  and  $\vec{n_3}$  are three non-degenerated sets of corresponding normal vectors respectively extracted from the two sets of raw data associated with two viewpoints between which a rotation, R, exists, these vectors must satisfy the following constraint equations:

$$\overline{n_1} = R\overline{n_1'} \tag{3}$$

$$\vec{n_2} = R\vec{n_2}$$
 (4)

$$\vec{n_3} = R\vec{n_3}$$
 (5)

If we let:

$$\vec{n_1} = [x_1, y_1, z_1] \text{ and } \vec{n_1'} = [a_1, b_1, c_1]$$
 (6)

$$\vec{n_2} = [x_2, y_2, z_2] \text{ and } \vec{n_2} = [a_2, b_2, c_2]$$
(7)

$$\vec{n_3} = [x_3, y_3, z_3] \text{ and } \vec{n_3'} = [a_3, b_3, c_3]$$
 (8)

where the rotation matrix, *R*, is defined in accordance with the standard *RPY* (roll-pitch-yaw) convention [5]:

$$R = \begin{vmatrix} g & h & i \\ j & k & l \\ m & p & q \end{vmatrix}$$
(9)

The three rotation angles,  $(\theta, \phi, \psi)$ , can be extracted as follows from the *R* matrix for the *RPY* convention:

$$\theta = \operatorname{atan}(j, g) \tag{10}$$

$$\varphi = \operatorname{atan}(-m, j\sin\theta + g\cos\theta) \tag{11}$$

$$\Psi = \operatorname{atan}\left(p,q\right) \tag{12}$$

where  $\theta$  represents the rotation around the *Z* axis,  $\varphi$  is the rotation around the *Y* axis and  $\psi$  corresponds to the rotation around the *X* axis of the reference frame.

The values of the rotation matrix parameters can be computed by solving the following system of equations for the rightmost vector:

The resulting rotation matrix is then checked for orthogonality. Note here that the centroids of patches are not taken into account as only the orientation information is to be estimated at this step.

Applying the computed rotation matrix to all other normal vectors should result in the overlap of these vectors. However, because of errors in the measurements and the approximations made on the surface representation, the matching might not be exact. In order to refine the estimation of the rotation matrix, its computation is repeated for all possible triplets of corresponding normal vectors in the simplified Gaussian sphere representation. The rotation parameters that lead to a maximum overlap between normal vectors are considered as the best estimate for the rotation matrix.

# IV.2. Translation parameters estimation

Once the rotation parameters have been estimated, the application of the resulting rotation matrix to one Gaussian sphere representation leads to two sets of normal vectors having the same orientation but submitted to a shift in 3-D space. Therefore, the translation parameters, T=[dx, dy, dz], can be estimated as the translation along the reference axes between the centroids associated with each surface patch. As many surface patches need to be matched, a good estimate of the translation parameters is the average of the necessary displacement along each axis to align the two different sets of center points, that is:

$$T = \frac{1}{N} \sum_{i=1}^{N} Rcp_{i}' - cp_{i}$$
(14)

where *R* is the rotation matrix previously estimated,  $cp_i$  and  $cp'_i$  are the corresponding centroids of the *i*<sup>th</sup> matched planar

 $p_i$  are the corresponding centrolds of the i -matched planar patches.

The matched centroids correspond to patches having similar orientation. As many patches in the surface description might share the same orientation, this may result in some false matches of centroids. However, the correct translation parameter set is considered to be the one with the maximum number of correspondences. Figure 7 illustrates this idea with some experimental data. All possible coordinate shifts are shown as 3-D points. They tend to be spread everywhere in the working space. But if the area with the highest density of translation values is extracted for each axis, as depicted by the small cube, a good estimate of the translation parameters can be extracted as the average of those points within the high density area.

# V. EXPERIMENTAL RESULTS

The proposed algorithm has been tested on a real experimental testbed as illustrated in figure 1. The development and validation of the algorithm has been made on a set of range profiles collected on various objects with a Jupiter range finder mounted on a 7-DOFs robotic arm. However, the robotic arm is only used to move the sensor and to validate the results. It doesn't play any role in the estimation of the registration parameters. Figure 8 shows a set of raw range measurements from the right-side viewpoint along with its simplified patch representation. The results of the registration estimation procedure are illustrated in figure 9 for one of our experimental objects as a fitted superposition between the two sets of raw range measurements. The computation time for the whole process is about 30 seconds on a Pentium III - 933 Mhz processor running Matlab for two sets of scans having 44 scan lines each.

Our experiments revealed that for geometrically symmetric objects having only a few distinctive surfaces, an overlap of up to 50% between the scan areas might be required to achieve correct registration. For usual man-made objects, the required overlap area is much less. However it always depends on the complexity of the object. Increasing the overlap between scan areas increases the reliability but makes the scanning process longer.



Figure 7: Extracting translation parameters from all possible matches.



Figure 8: Range measurements and patch representation.



a) front view



Figure 9: Superposition of two sets of scan lines after registration.

The scan segmentation and surface representation as a modified Gaussian sphere revealed to significantly improve the computation efficiency. It also makes patch fitting simpler and faster. Using the overlap of normal vectors as an evaluation factor rather than the number of matched points between the two sets of data speeds up computation. It also gives more weight to those points that belong to larger patches. This strategy appears to be appropriate as large approximated patches appear to be a more reliable representation of the object surface for a given required accuracy in fitting.

The accuracy of the estimated registration parameters is sufficient for most applications in robot navigation with collision avoidance where computing time is a critical issue that makes more sophisticated algorithms based on triangular meshes or high-level feature extraction not tractable.

#### VI. CONCLUSION

We have introduced an automatic approach for registration estimation based on raw range measurements provided by a single line laser range sensor. The registration parameters are computed without any need for an external device to provide an initial estimate, nor any feature extraction or triangulation. Assuming only object rigidity and some overlap between the scan areas, the approach allows to estimate the six rotation and translation parameters that link 3-D scans gathered from different viewpoints. Taking advantage of a modified Gaussian sphere representation, the mapping of planar patches that result from the merge of similar range profiles is significantly compacted. This solution results in an important reduction of the computational workload while providing an efficient mean for matching estimation and validation all along the process. The proposed approach provided excellent experimental results on scans gathered with a Jupiter range finder with a sufficient accuracy for most robot navigation applications.

Further developments to this registration technique will examine some alternative ways to merge range profiles in the areas where only a small number of scan lines exhibit similar characteristics. Some improvements on the planar patch estimation step might also be introduced to be able to deal with generic objects that could advantageously be represented by non-rectangular patches. An implementation in C with optimization of the code will also result in an important reduction of the computation time.

# VII. REFERENCES

- P. Besl, N. McKay, "A Method of Registration of 3-D Shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no 2, Feb. 1992, pp. 239-256.
- [2] R. Bergevin, D. Laurendeau, D. Poussart, "Registering Range Views of Multipart Object", *Computer Vision and Image Understanding*, vol. 61, no 1, 1995, pp. 1-16.
- [3] Y. Chen, G. Medioni, "Object Modeling by Registration of Multiple Range Images", *Image and Vision Computing*, vol. 10, no 3, 1992, pp. 145-155.
- [4] C. S. Chua, "3D Free-Form Surface Registration and Object Recognition", International Journal of Computer Vision, vol. 17, 1996, pp. 77-99.
- [5] J. J. Craig, Introduction to Robotics, Mechanics and Control, 2nd edition, Addison-Wesley, 1995.
- [6] S. F. El-Hakim, J.-A. Beraldin, G. Godin, P. Boulanger, "Two 3-D Sensors for Environment Modeling and Virtual Reality: Calibration and Multi-View Registration", *ISPRS Congress, International Archives of Photogrammetry*, Vienna, 1996, pp. 140-146.
- [7] M. K. Elstrom, P. W. Smith, "Stereo-Based Registration of Multi-Sensor Imagery for Enhanced Visualization of Remote Environments", *Proceedings of the IEEE International Conference on Robotics and Automation*, Detroit, May 1999.
- [8] D. F. Huber, "Automatic 3D Modeling Using Range Images Obtained from Unknown Viewpoints", *Proceedings of the IEEE International Conference on 3-D Digital Imaging and Modeling*, Los Alamitos, CA, May 2001, pp. 153-160.
- [9] D. Laurendeau, C. Gosselin, F. Caron, S. Comtois, T. Laliberté, F. Loranger, "A Random Access 3D/2D Vision Sensor", *Proceedings of Vision Interface*, Vancouver, BC, June 1998, pp. 93-100.
- [10]T. Masuda, N. Yokoya, "A Robust Method for Registration and Segmentation of Multiple Range Images", *Computer Vision and Image Understanding*, vol. 61, no 3, 1995, pp. 295-307.
- [11]P. Payeur, C. M. Gosselin, D. Laurendeau, "Range Data Merging for Probabilistic Octree Modeling of 3-D Workspaces", *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 4, Leuven, Belgium, May 1998, pp. 3071-3078.
- [12]G. Roth, "Registering Two Overlapping Range Images", Proceedings of the 2nd International Conference on Recent Advances in 3-D Digital Imaging and Modeling, Ottawa, Canada, 1999, pp. 191-200.
- [13]E. Trucco, A. Verri, Introductory Techniques for 3-D Computer Vision, Prentice-Hall, 1998.