REGISTRATION OF LIDAR POINT CLOUDS USING IMAGE FEATURES

Manoranjan Majji, Research Associate
Brein Flewelling, Graduate Student
Brent Macomber, Graduate Student
John L. Junkins, Distinguished Professor
Aerospace Engineering Department
Texas A&M University
College Station, TX 77843-3141
majji@tamu.edu
junkins@tamu.edu

Anup B. Katake, Scientist
Starvision Technologies
College Station, TX, 77840

Hyochung Bang, Associate Professor
Department of Aerospace Engineering
KAIST, Taejon, Korea

ABSTRACT

An optimal linear translation and attitude estimation (OLTAE) algorithm is proposed to register 3dimensional point clouds based on the image features associated with the individual data sets. In LIDAR applications, such images are created by projecting the point cloud data on to an image plane. Physically, this image is the return light intensity observed by the LIDAR imager that is usually available to the analyst for post processing. Associated image features are extracted from the corresponding images by utilizing the recent advances in computational vision and image processing. Features thus obtained have unique descriptors that automate the matching process and ease the solution of the so-called correspondence problem. Corresponding matched features from the images are then used as vector measurements for the 3 dimensional point cloud registration algorithm. As a byproduct the algorithm is shown to provide the uncertainties associated with the translation vector and the pose orientation estimate. The methods developed in the paper are subsequently applied to measurement sets obtained from a Light Detection and Ranging (LIDAR) system for spacecraft proximity operation emulation applications.

INTRODUCTION

Recent advances in micro electronics and MEMS technologies are working towards the development of more cost effective and reliable LIDAR systems. This mini-revolution in hardware has caused explosive growth in the development of ranging devices and we anticipate novel sensing solutions in increasingly compact form factors in the near future. These devices are all set to stream large volumes of data for processing and the compaction will enable portability and hence the challenging problem of registration of the data streams and models generated by the LIDAR systems ensues.

Such problems have been addressed previously in the robotics community. Mobile robots typically are required to scan their environment for model building and simultaneous localization (SLAM, for short, cf. (Besl and McKay 1992), (Surmann, Nuchter and Hertzberg 2003), (Andreasson and Lilienthal 2007)). Enabling LIDAR sensors to the mobile platform enhances the model accuracy and eliminates the scale factor ambiguity associated with traditional modeling methods such as stereo vision sensing solutions (cf. (Morodohoi, et al. 2007), (Se, Lowe and Little 2002)). Iterative Closest Point (ICP) algorithm implementations are commonly used to solve the 3 dimensional point cloud registration problem (cf. (Rusinkiewicz and Levoy 2001)) . From a mapping perspective this is rather useful information since geo-referencing is possible. In the context of aerospace applications, this is rather attractive for us since we are interested in proximity guidance, navigation and control applications that can be possible only by precise knowledge of the scene geometry.

In the areas of photogrammetry and remote sensing, the LIDAR data processing algorithms are playing an increasingly key role (cf. (Gerlek 2009), (Cekada, Crosilla and Kosmatin-Fras 2009) among other recent papers). Data registration has been a key problem in this community and a lot of work has been done by several researchers (cf. (Wilkinson, et al. 2009),). With increasing use of LIDAR data, algorithms for the calibration and auto-calibration of the LIDAR scanning systems are also being actively developed (cf. (Lichti 2010),(Amiri and Armin 2010)).

Our paper is presented as follows. We first outline the novel OLTAE algorithm for relative translation and rotational motion estimation. Subsequent section gives a brief over view of the feature extraction algorithms and our implementation of the SURF algorithm. We then discuss the details of our scanner and experimental setup. Experimental results and validation details are presented subsequently. We conclude the paper by discussing our current and future work on the improvement of the current data processing pipeline for efficient reconstructions and accurate model building.

OPTIMAL LINEAR TRANSLATION AND ATTITUDE ESTIMATION ALGORITHM

We now present the technical details associated with an algorithm that is central to the point cloud registration process, the main subject matter of this paper. It is convenient to first assume that the correspondence problem in two consecutive point cloud data sets has been solved and the analyst at this stage has a list of the matching "tie-points" to stitch the point clouds together. We note that this is one of the most challenging steps in the 3D point cloud registration process and is addressed in the subsequent developments of this paper. With this assumption, the coordinates of the matched 3dimensional point as viewed by the sensor in two different coordinate systems (observation stations) is depicted in the schematic of Fig. 1.
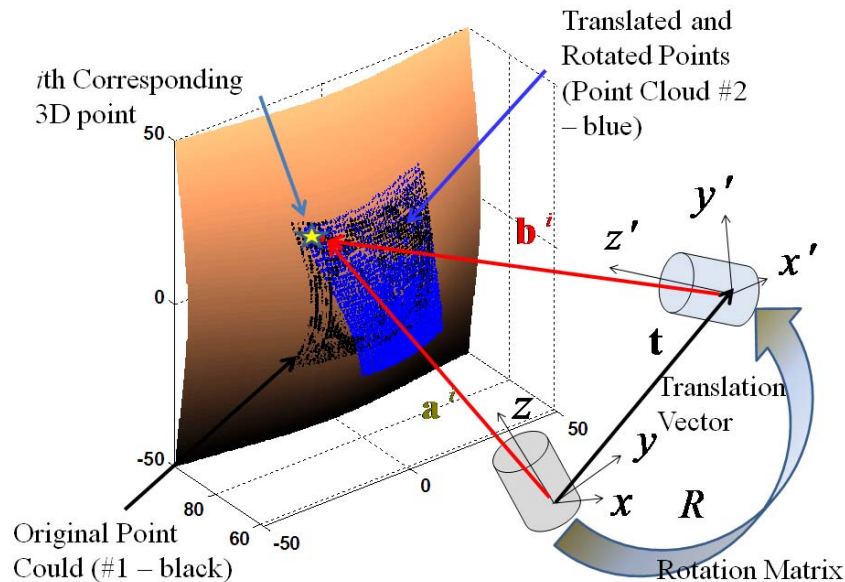
Figure 1. Schematic depicting the LIDAR point cloud registration problem.

Mathematically, the dependence of the matched 3D coordinates (as observed in the new coordinate system) on the translation and rotation of the scanning station and the corresponding coordinates in the "old" coordinate system is given by the following vector relation

.
$$\mathbf{b}^i = R\mathbf{a}^i + \mathbf{t} \tag{1}$$

where $\mathbf{b}^i$, $\mathbf{a}^i$ are the ith point as observed in the subsequent frames, R is the proper orthogonal rotational matrix ($R^T R = I$) and $\mathbf{t}$ denotes the translation vector between the two successive frames in the experiment.

Now let us parameterize the direction cosine matrix in terms of the classical Rodrigues parameters (Gibbs vector). This parameterization of the orthogonal matrices in 3D is quite conveniently accomplished by the Cayley transform given by

$$\mathbf{b}_i = (I + Q)^{-1} (I - Q) \mathbf{a}_i + \mathbf{t} \tag{2}$$

where $Q = [\tilde{\mathbf{q}}]$, with $\mathbf{q}$ being the vector of classical Rodrigues parameters (Schaub and Junkins 2004) (Gibbs vector) and the notation $[\cdot]$ is used to denote the cross product operation, creating a skew-symmetric matrix from the corresponding vector.

Now the same equation can be written as

$$\begin{aligned} (I + Q)\mathbf{b}_i &= (I - Q)\mathbf{a}_i + (I + Q)\mathbf{t} \\ &= (I - Q)\mathbf{a}_i + \mathbf{t}^* \end{aligned} \tag{3}$$

where the redefinition of $\mathbf{t}^* = (I + Q)\mathbf{t}$ has been used.

The relationship above can be simplified further by writing

$$\mathbf{b}_i - \mathbf{a}_i = -Q(\mathbf{b}_i + \mathbf{a}_i) + \mathbf{t}^* \tag{4}$$

Defining $\mathbf{a}_i + \mathbf{b}_i = \mathbf{v}_i, \mathbf{b}_i - \mathbf{a}_i = \boldsymbol{\varepsilon}_i$ we have that

$$\begin{aligned} \boldsymbol{\varepsilon}_i &= -Q\mathbf{v}_i + \mathbf{t}^* \\ &= [\tilde{\mathbf{v}}_i]\mathbf{q} + \mathbf{t}^* \\ &= \begin{bmatrix} [\tilde{\mathbf{v}}_i] & I_3 \end{bmatrix} \begin{bmatrix} \mathbf{q} \\ \mathbf{t}^* \end{bmatrix} \end{aligned} \tag{5}$$

The determination of the attitude parameter and the direction cosine matrix then becomes rigorously linear – a linear least squares problem ensues quite elegantly. By stacking the system of (3 element) vectors from each measurement, we obtain the least squares problem to be given by

$$\mathbf{y} = \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_m \end{bmatrix} = \begin{bmatrix} [\tilde{\mathbf{v}}_1] & I_3 \\ \vdots & \\ [\tilde{\mathbf{v}}_m] & I_3 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}} \\ \hat{\mathbf{t}}^* \end{bmatrix} = \mathbf{H} \begin{bmatrix} \hat{\mathbf{q}} \\ \hat{\mathbf{t}}^* \end{bmatrix} \tag{6}$$

The best estimate (in the least squares sense) is consequently obtained by the solution to the normal equations as

$$\begin{bmatrix} \hat{\mathbf{q}} \\ \hat{\mathbf{t}}^* \end{bmatrix} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y} \tag{7}$$

Please note however that although we obtain the estimates of the translation vector and rotation matrix directly from this linear least squares solution directly, the covariance calculations cannot be obtained reliably from the expression $(\mathbf{H}^T \mathbf{H})^{-1}$. This is because of two assumptions carried out in the above calculation. One is that the data in the

coefficient matrix was assumed to be error free. In fact this is far from the truth. Since the coefficient matrix is derived from measurements, it is noisy typically. To account for the noise, one typically uses a total least squares error criterion for the estimation problem (cf. (Crassidis 2004), (Majji and Junkins 2007)). A second complication arises from the fact that the parameters being estimated have been "lumped" in a nonlinear fashion by the redefinition of the variables (definition of $\mathbf{t}^*$).

FEATURE EXTRACTION AND MATCHING

Feature extraction and matching are one of the most important image processing contributions of modern times. While Harris corner detector (cf. (Moravec 1980)(Harris and Stephens 1988)) instigated the revolution in image processing and artificial intelligence applications, the robustness in feature extraction and matching really started off from the Scale Invariant Feature Transform (SIFT) (cf. (Lowe 2004)). A reliable, repeatable and identifiable feature extraction process was developed by Lowe based on the assignment of keypoint descriptors to each feature extracted by a corner detector. Introduction of scale space ideas in this paper made the keypoint localization and extraction process robust to small degrees of illumination variations, image rotations and compression artifacts and image features are now able to be reliably identified across several textured regions.

Computationally efficient alternatives soon followed. Some researchers proposed implementation of this algorithm on graphical processing units (GPUs) utilizing the parallelization inherently present in the graphics hardware (cf. (Wu 2004)). One of the most computationally attractive alternatives that arrived later is a modification of the feature extraction and description process known as the Speeded-Up Robust Features (SURF) (cf. (Bay, et al. 2008), (Evans 2009)). It is quite attractive because of the fact that the image operations are highly speeded up by using box filters using an integral image concept (cf. (Viola 2001)). Interest point selection and scale space ideas are borrowed from the SIFT and approximated appropriately for the desired speed-ups. We wish to direct interested readers to the appropriate literature. In subsequent developments of the paper, we use our own implementation of the SURF algorithm for feature extraction that was found to perform similar to other implementations depending on the tuning parameters. A sample feature extraction result and the associated feature scale is plotted in Fig. 2.

**Figure 2.** Image showing TAMU SURF implementation.

In the SURF paradigm, one typically obtains an identification tag associated with each image feature that has been isolated from the corner detection algorithm. This is known as the feature descriptor. For our implementation, owing to computational considerations, we use a 64 dimensional feature descriptor for each feature that is isolated from the image. Normalizing the descriptor is an elegant way to achieve slight invariance of the extraction and matching process to illumination variations in different scenes. When the analyst intends to produce matches

between sets of features, one typically tries to look at several characteristics. Here we used a dot product measure between features of the successive scans to isolate the greatest potential matches. This also provides us with robustness pre-filters in the data set and usually was found to give dependable results without much tuning by the analyst.

We now proceed to the discussions associated with our experimental setup and associated data and image processing details.

EXPERIMENTAL SETUP AND RESULTS

At the Land, Air and Space Robotics (LASR) Laboratory, Texas A&M University, we have a Delta-Sphere 3000 scanning LIDAR instrument. The instrument consists of a laser ranging device (Acuity 1D scanner) mounted on a pan mechanism that can be precisely measured by means of an encoder embedded in the system. Data processing pipeline is provided by 3rd tech corporation. More details on the scanning system are provided on their webpage(England 2009). Since the test data is available directly from the scanner, we have been able to exercise our algorithms with some modest data processing conversions. Sample scans of a representative scene are shown in Fig 3. Please note that the data has been projected on to a plane through origin situated at the imager. Also note that the scene color has been added to the light return data using a camera system attached to the scanner.

Scan # 1   Scan # 2
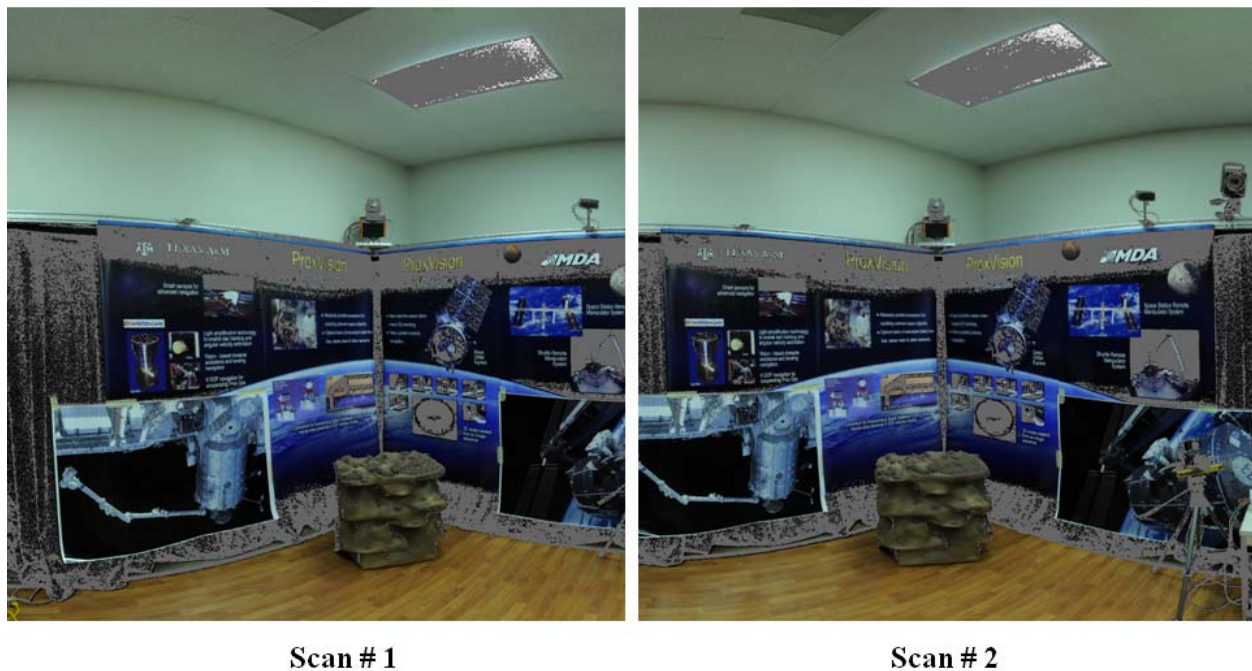
Figure 3. Representative scans (second scan is at a translated and rotated location with respect to the first).

Gray dots in each scan denote insufficient data (consequently invalid range measurement) by the photodetector of the light return measurement system. The experimental setup represents a typical scenario in the 3D point cloud registration problem. The first step is to use our feature extraction and image processing tools to extract features that are common to the image sets in this representative situation. The matching procedure outlined previously was employed in this situation and we obtain 42 reliable matches as shown in Fig. 4. It was observed that the correspondence of the features detected and matched was correct without much tuning by the analyst. While we understand that an efficient procedure should produce correct matches without any intervention by the analyst (a scheme based on residual errors and variance calculations is used by us for removal of feature matches) this observation is to demonstrate the clear reliability of the feature extraction and matching algorithms developed by our researchers. The matched features are shown in Fig. 4.

Feature Matching using SURF

Figure 4. Feature matching using SURF (TAMU implementation).

Corresponding models associated with the individual scans are created as shown in the Fig.5. A texture mapping process included in the vendor suite is used to produce the models being shown in Fig. 5. We are investigating alternative and precise calibration alternatives to understand this transformation using stereo vision concepts.

Model #1   Model #2

Figure 5. Models from individual scans.

Using the feature matches as indicated in the main developments of the paper, we are able to estimate the rotational motion of the frames to be estimated as

$$\hat{\theta} = \begin{bmatrix} -0.3108, 0.1877, 0.0868 \end{bmatrix} \tag{8}$$

degrees (Yaw, Pitch and Roll). Translation estimate was found to be

$$\hat{t} = \begin{bmatrix} -20.9258, 1.6007, 0.2099 \end{bmatrix} \text{in.} \tag{9}$$

The true displacement along the positive x axis of the initial frame was 21 inches (with no rotational motion).

CONCLUSION

A novel algorithm, called OLTAE for image based 3D point cloud registration is presented in the paper using recent advances in computational vision. An implementation of the speeded-up robust features (SURF) is used in the feature extraction matching process to identify corresponding features in the rigid-body displaced scanned data. The novel algorithm is demonstrated on a representative scan data taken by the sensor in our laboratory. Current work includes the development of more accurate calibration methods (manual as well as automated options are being investigated) to align the image data with the scan data. Efficient computational implementations (hardware and software solutions) are also under study. Investigation in to the applicability of the present strategy to problems in photogrammetry and remote sensing, especially for air-borne scanning and satellite sensing data is in progress. Error characterization of the current sensor with respect to inertial measurement units present in our laboratory is also being carried out.

REFERENCES

Amiri, P. J., and G. Armin. "Sensor modeling, self-calibration and accuracy testing of panoramic cameras and laser scanners." *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010: 60-76.

Andreasson, H., and A Lilienthal. "Vision Aided 3D Laser Scanner Based Registration." *IEEE International Conference on Autonomous Robots and Agents (ICARA).* Palmerson North, New Zealand, 2007. 1-6.

Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. "SURF: Speeded Up Robust Features." *Computer Vision and Image Understanding*, 2008: 346-359.

Besl, P. J. "A Method for Registration of 3-D Shapes." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992: 239-256.

Besl, P.J., and N. D McKay. "A Method for Registration of 3-D Shapes." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1992: 239-256.

Cekada, M.T, F. Crosilla, and M. Kosmatin-Fras. "A Simplified Analytical Model for a-priori Lidar POoint-Positioning Error Estimation and a Review of Lidar Error Sources." *Photogrammetric Engineering and Remote Sensing*, 2009: 1425-1440.

Crassidis, J., Junkins, J.L.,. *Optimal Estimation of Dynamical Systems.* Boca Raton, FL: Chapman - Hall/ CRC Press, 2004.

England, Nick. *Delta Sphere 3D Laser Scanner.* 2009. http://www.deltasphere.com/ (accessed March 4, 2010).

Evans, C. *Notes on the OpenSURF Librar.* Internal Report No. CSTR-09-001, Bristol, UK: University of Bristol, 2009.

Gerlek, M. P. "Compressing LIDAR Data." *Photogrammetric Engineering and Remote Sensing*, 2009: 1253-1255.

Harris, C., and M. Stephens. "A Combined Edge and Corner Detector." *Proceedings of the 4th Alvey Vision Conference.* 1988. 147-151.

Lichti, D. D. "Terrestrial laser scanner self-calibration: Correlation sources and their mitigation." *ISPRS Journal of Photogrammetry and Remote Sensing* 65, no. 1 (2010): 93-102.

Lowe, D. G. "Distinctive Image Features from Scale Invariant Keypoints." *International Journal of Computer Vision*, 2004: 91-110.

Majji, M., and J. L. Junkins. "Total Least Squares Estimation of Dynamical Systems." *AIAA Guidance, Navigation and Control Conference and Exhibit.* Savannah, GA: AIAA, 2007. 1-15.

Moravec, H. *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover.* Institute Reports, Pittsburgh, PA: The Robotics Institute, Carnegie Mellon University, 1980.

Morodohoi, P., J. M. Frahm, A. Akbarzadeh, D. Gallup, Clipp. B., and D. Engles. "Real-time video-based reconstruction of urban environments." *ISPRS Working Group V/4 Workshop 3D-ARCH 2007: 3D Virtual Reconstruction and Visualization of Complex Architectures.* Zurich, Switzerland: ISPRS, 2007.

Rusinkiewicz, S., and M. Levoy. "Efficient Variants of the ICP Algorithm." *International Conference on 3D Digital Imaging and Modeling 3DIM.* 2001.

Schaub, H., and J.L. Junkins. *Analytical Mechanics of Space Systems.* Washington, D. C., : AIAA Press, 2004.

Se, S., D. Lowe, and J. Little. "Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks." *The International Journal of Robotics Research*, 2002: 735-758.

Surmann, H., A. Nuchter, and J Hertzberg. "An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitization of indoor environments." *Robotics and Autonomous Systems*, 2003: 181-198.

Viola, P., and Jones, M. "Rapid object detection using a boosted cascade of simple features." *IEEE Computer Vision and Pattern Recognition (CVPR).* IEEE, 2001.

Wilkinson, B.E., B.A. Dewitt, A.C. Watts, A. H. Mohamed, and M.A Burgess. "A New Approach for Pass-point Generation from Aerial Video Imagery." *Photogrammetric Engineering and Remote Sensing*, 2009: 1415-1423.

Wu, C. C. *SIFT on GPU (siftgpu).* 2004. http://www.cs.unc.edu/~ccwu/siftgpu/ (accessed March 1, 2010).