Three-Dimensional Data Registration Based On Human Perception

Bruce Brendle, Ph.D.
US Army RDECOM-TARDEC
AMSRD-TAR-R (MS:205)
Warren, MI 48397-5000
bruce.brendle@us.army.mil

***Abstract*—Registration, the process of transforming different sets of data into a common coordinate system, is often required to allow the comparison or integration of the data sets. Techniques such as the widely used Iterative Closest Point (ICP) algorithm have limited effectiveness on data sets that require significant transformation or that have large degrees of inconsistencies. This paper describes a biologically inspired algorithm for data registration that is based on two theories of human perception. The use of macro level registration based on these theories combined with micro level registration using the ICP algorithm provides enhanced registration of these challenging data sets. The new algorithm was tested extensively on simulated sensor images in several scenarios key to successful application to autonomous ground navigation. The excellent performance of the biologically inspired algorithm in these cases makes it a promising candidate for this field.**

I. Introduction

Diverse applications ranging from medical imaging to computer vision make use of three-dimensional data. Often, multiple sets of data are acquired by sampling the same scene or object at different times, or from different perspectives, resulting in each data set having its own coordinate system. Registration, the process of transforming different sets of data into a common coordinate system, is then required to allow the comparison or integration of the data sets.

A basic rigid transformation between two coordinate systems, *i* and *k*, is composed of a translation and a rotation defined as:

$$ {}^{i}\mathbf{t}_k = {}^{i}\begin{bmatrix} x \\ y \\ z \end{bmatrix}_k = \begin{bmatrix} {}^{i}x_k \\ {}^{i}y_k \\ {}^{i}z_k \end{bmatrix} $$

$$ \begin{aligned} {}^{i}\mathbf{R}_k &= {}^{i}\mathbf{R}_{k,z(\phi)}\,{}^{i}\mathbf{R}_{k,y(\theta)}\,{}^{i}\mathbf{R}_{k,x(\psi)} \\ &= \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & -\sin\psi \\ 0 & \sin\psi & \cos\psi \end{bmatrix} \\ &= \begin{bmatrix} \cos\phi\cos\theta & -\sin\phi\cos\psi+\cos\phi\sin\theta\sin\psi & \sin\phi\sin\psi+c\phi s\theta\cos\psi \\ \sin\phi\cos\theta & \cos\phi\cos\psi+\sin\phi\sin\theta\sin\psi & -\cos\phi\sin\psi+\sin\phi\sin\theta\cos\psi \\ -\sin\theta & \cos\theta\sin\psi & \cos\theta\cos\psi \end{bmatrix} \end{aligned} $$

A data set comprised of *m* individual points is represented by a $m\times 3$ matrix, $\mathbf{D}$, with the following notation:

$$ \text{Frame of Reference}\,[\mathbf{D}]^{\text{Operation on the Data Set}}_{\text{Descriptor}} $$

Actual data viewed from viewpoint *i* is represented as:

$$ {}^{i}\mathbf{D}_A = \begin{bmatrix} {}^{i}\mathbf{p}_{a_1} & {}^{i}\mathbf{p}_{a_2} & \cdots & {}^{i}\mathbf{p}_{a_m} \end{bmatrix} = \begin{bmatrix} a_{x_1} & a_{x_2} & \cdots & a_{x_m} \\ a_{y_1} & a_{y_2} & \cdots & a_{y_m} \\ a_{z_1} & a_{z_2} & \cdots & a_{z_m} \end{bmatrix} $$

A key fact that will become important later is that ${}^{k}\mathbf{D}_A$ the actual data viewed from viewpoint *k*, is not necessarily equal to ${}^{i}\mathbf{D}_A$ for several reasons. If viewpoints *i* and *k* are separated spatially, data that was occluded at viewpoint *i* may become visible at viewpoint *k*. Also, if viewpoints *i* and *k* are separated temporally, i.e. the data is captured at different times, new data may appear.

The forward kinematics relationship between the data sets is given by:

$$ {}^{i}\mathbf{D}_A = {}^{i}\mathbf{R}_k\,{}^{k}\mathbf{D}_A + {}^{i}\mathbf{T}_k \qquad (1) $$

This relationship is the foundational equation for data registration. The goal of registration is to determine ${}^{i}\mathbf{R}_k$ and ${}^{i}\mathbf{T}_k$ that will transform the data sets into a common coordinate system so that the data sets can be compared or integrated. A byproduct of accurate registration having great benefit to the autonomous navigation application is that the vehicle's 3D movement can be calculated from ${}^{i}\mathbf{R}_k$ and ${}^{i}\mathbf{T}_k$, allowing localization of the vehicle with respect to a known starting position. This is similar to dead reckoning based upon heading and motion sensors without its drawback of error being proportional to the amount of distance traveled.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2006		

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Three-Dimensional Data Registration Based On Human Perception	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **US Army RDECOM-TARDEC,AMSRD-TAR-R (MS:205),Warren,MI,48397**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
See Report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**		**8**	

Several issues make direct calculation of ${}^{i}\mathbf{R}_{k}$ and ${}^{i}\mathbf{T}_{k}$ impractical for real world applications. Two of the most significant issues are *inconsistencies in the data sets* and *noise in the measurements of the data sets*.

Inconsistencies may exist due to new data appearing when data capture points are separated temporally or, in the case of spatially separated capture points, when data becomes visible that had previously been occluded. Also, actual data is not available in many applications and we are left to deal with noisy measurements of the data, the size of the errors being related to the quality of the sensors used to measure the data.

II. DATA REGISTRATION APPROACHES

A review of published literature reveals that it is rich in the area of data registration. There are many published approaches to application-independent image registration [1]. The majority of the literature comes from the medical imaging community, whose emphasis is on accuracy of registration over speed of registration [2]. The autonomous ground navigation community is also addressing the issue of range image registration [3]. The notable difference in the work performed by this community is the emphasis on speed over accuracy as registration speed is required for high-speed navigation. An optimal approach to registration should have both performance characteristics, accuracy and speed.

These approaches can be classified by several characteristics including rigid or elastic, extrinsic or intrinsic, and feature-based or voxel-based. For the application of autonomous ground navigation, techniques that are rigid and intrinsic are most relevant. Of the these approaches, voxel-based techniques are the most promising for applications requiring high-speed registration such as autonomous ground navigation as they avoid the processing time and error-potential of identifying features to be registered. The voxel-based technique most applicable to autonomous ground navigation is the Iterative Closest Point (ICP) algorithm.

III. ITERATIVE CLOSEST POINT ALGORITHM

Substituting measured or sensed data for the actual data in (1), the data sets for the image registration problem are the *sensed* data at viewpoints *i* and *k*, ${}^{i}\mathbf{D}_{S}$ and ${}^{k}\mathbf{D}_{S}$. The goal of registration is to find the three-dimensional transformation ${}^{k}\mathbf{D}_{S}$ (the combined rotation and translation) that minimizes the distance between the points in these sets, or more specifically:

$$\min_{\mathbf{T},\mathbf{R}} \sum_{n=1}^{m} \left\| {}^{i}\mathbf{D}_{n,S} - ({}^{i}\mathbf{R}_{k}\,{}^{k}\mathbf{D}_{n,S} + {}^{i}\mathbf{T}_{k}) \right\|^{2} \quad (2)$$

The ICP algorithm solves (2) iteratively in the following steps: **S**

1. Establish a set of *n* closest points between the data sets ${}^{i}\mathbf{D}_{S}$ and ${}^{k}\mathbf{D}_{S}$.
2. Compute the incremental transformation, and ${}^{i}\mathbf{T}_{k}$, using the set of closest points.
3. Apply the incremental transformation from Step 2 to ${}^{k}\mathbf{D}_{S}$
4. If relative changes in ${}^{i}\mathbf{R}_{k}$ and ${}^{i}\mathbf{T}_{k}$ are less than a threshold, terminate. Otherwise, repeat the procedure starting from Step 1.

The identification of closest points between two 3-D point sets required for the first step of the ICP algorithm can be accomplished by several methods including k-d trees and Voronoi polygons. The MATLAB function *dsearchn(X, XI)*, which returns the indices of the closest point in X for each point in XI was used for this research.

Likewise, the computation of ${}^{i}\mathbf{R}_{k}$ and ${}^{i}\mathbf{T}_{k}$ required for the second step of the algorithm can be performed by several methods including quaternions [4] and singular value decomposition [5].

Two major issues remain with the ICP approach. First, the amount of transformation possible with ICP is bounded and registration speed increases proportionally with the initial distance between images. The second issue is that inconsistencies in the data sets due to such real-world events as moving objects or occluded objects contribute directly to errors in registration accuracy.

IV. BIO-INSPIRED ALGORITHM

Fortunately, there exists an example of a system that handles these issues – the human perception system. This biological system, although not fully understood, has been modeled by physiologists and theories of its operation have been postulated and tested. This research project has applied recent theories in how humans register visual images during high speed eye movements and how they handle inconsistencies in the visual images to computer-based image registration. The resulting bio-inspired registration process has reduced dependency on iterative registration techniques such as ICP and is robust in cases of high degrees of inconsistency in the data sets. The process works well in all cases except in the extreme case in which there is no overlap between the data sets, making registration impossible by any method, and when there is not a sufficient difference in the data sets from which ${}^{i}\mathbf{R}_{k}$ and ${}^{i}\mathbf{T}_{k}$ can be calculated. An example of the latter case would be images from a sensor moving parallel to a flat surface, which are indistinguishable even by the human eye.

The proposed registration approach is based upon human use of extraretinal signals to estimate visual transformation and the assumption of stationarity. These concepts have been applied to macro level registration and then combined with the ICP algorithm for fine tuning in order to achieve accurate registration in cases where ICP alone does not perform well.

A. Saccadic Suppression

1) Perception Theory: A saccade is a rapid movement of the eye that results in the smearing of the image seen by the eye when the saccade occurs during the retinal integration time [6].

A similar phenomenon occurs while photographing high speed events. If the speed of the subject (relate this to a high-speed movement of the eye) is higher than the shutter speed of the camera (relate this to the integration time of the eye), the moving subject will appear blurry in the captured image.

To prevent this smearing, it has been suggested that the brain utilizes a *saccadic suppression* mechanism to shut off retinal processing during eye saccades [7] and that this suppression is triggered by extraretinal signals. The source of these signals is debatable and may come from extraocular muscles that measure actual movement of the eye [8], the efferent command that initiates the eye movement [9], or some combination of the two [10]. Regardless of the source, these signals can be used to estimate the required transformation between the pre- and post-saccadic images [11], allowing humans to accurately register the images. It is important to note that this mechanism accounts for transformation of images resulting solely from movement of the eye and does not account for inconsistencies in the images from such events as moving objects.

2) Registration Equivalent: Pre- and post-saccadic images can be equated to a series of medical images captured during a linear scan or to images captured from a sensor on a moving unmanned vehicle. In the same manner extraretinal signals are used to calculate the predicted transformation between the trans-saccadic images, positioning sensors on the medical imaging system or unmanned vehicle can be used to calculate an initial transformation between the sensor images.

Recall from (1) that the goal is to register two data sets, ${}^{i}\mathbf{D}_{S}$ and, ${}^{k}\mathbf{D}_{S}$ captured from viewpoints *i* and *k*, respectively. The hypothesis is that we can transform ${}^{i}\mathbf{D}_{S}$ to a viewpoint, *j*, that is very close to the viewpoint *k* using noisy measurements of the transformation between the viewpoints, ${}^{i}\mathbf{R}_{k,S}$ and ${}^{i}\mathbf{T}_{k,S}$. This results in ${}^{j}\mathbf{D}_{T}$, with the subscript indicating a *transformed* data set. Next, a registration technique such as ICP can be used on the resulting data sets to remove any error in the measurements of ${}^{i}\mathbf{R}_{k,S}$ and ${}^{i}\mathbf{T}_{k,S}$.

Applying the kinematics relationship from (1), to the data sets of interest, we have:

$$ {}^{i}\mathbf{D}_{S} = {}^{i}\mathbf{R}_{k,S}\,{}^{j}\mathbf{D}_{T} + {}^{i}\mathbf{T}_{k.S} \qquad (\cdot\,\cdot) $$

The transformed data set ${}^{j}\mathbf{D}_{T}$ can then be calculated as:

$$ {}^{j}\mathbf{D}_{T} = {}^{i}\mathbf{R}_{k,S}^{-1}({}^{i}\mathbf{D}_{S} - {}^{i}\mathbf{T}_{k,S}) \qquad (\cdot\,\cdot) $$

The registration of the resulting data set, ${}^{j}\mathbf{D}_{T}$, with ${}^{k}\mathbf{D}_{S}$ by means such as ICP provides the fine tuning rotation and translation parameters ${}^{j}\mathbf{R}_{k,I}$ and ${}^{j}\mathbf{T}_{k,I}$. Estimates of the desired total rotation and translation values can then be estimated as:

$$ {}^{i}\hat{\mathbf{R}}_{k,A} = {}^{i}\mathbf{R}_{j,S}\,{}^{j}\mathbf{R}_{k,I} \qquad (\cdot\,\cdot) $$

$$ {}^{i}\hat{\mathbf{T}}_{k,A} = {}^{i}\mathbf{T}_{j,S} + {}^{i}\mathbf{R}_{j,I}\,{}^{j}\mathbf{T}_{k,I} \qquad (\cdot\,\cdot) $$

Here the errors are limited to errors inherent in the registration technique. For accurate estimates of ${}^{j}\mathbf{R}_{k,A}$ and ${}^{j}\mathbf{T}_{k,A}$, the calculated values of the parameters ${}^{j}\mathbf{R}_{k,I}$ and ${}^{j}\mathbf{T}_{k,I}$ from the registration process must have negligible errors. However, *the registration of the data sets with negligible error is not trivial when there are inconsistencies in the data sets.*

B. Stationarity Assumption

1) Perception Theory: Registration of trans-saccadic images in the presence of moving objects or during the appearance of previously occluded objects is more complex. Although the extraretinal signals can be used to develop an estimate of the transformation between the images, the visual system must still account for unexpected changes in the images, which could be attributed to either moving objects or errors in the transformation estimate.

Image processing research in the area of Structure from Motion (SfM) has led to two hypotheses regarding the mechanism for registration in this case. A long-standing hypothesis, called the *rigidity assumption*, stated that the visual system would always choose the most rigid transformation, i.e. the one that required the least deformation [12]. An example of rigid transformation of a singular object would be a car moving parallel to an observer. Even though the car is moving, its size and shape remain the same.

Recent work has resulted in the postulation of a *stationarity assumption* [13] that would be considered along with rigidity in visual registration. Stationarity is a preference for objects to remain fixed in an allocentric, earth referenced coordinate system. An example of stationarity is the use of buildings as landmarks for navigation. As a person drives a car and these landmarks become occluded and then visible again, they assume that the buildings have remained stationary and they are able to estimate their position relative to these stationary objects. In weak stationarity, when multiple solutions are equally rigid, the visual system will select the one that is most stationary. In strong stationarity, the visual system will chose a stationary solution even if it is detectably non-rigid. Recent experiments support the hypothesis of strong stationarity [14].

2) Registration Equivalent: In the same manner that humans perceive visual images, an assumption of stationarity can be applied to the data sets ${}^{j}\mathbf{D}_{T}$, with ${}^{k}\mathbf{D}_{S}$ to compensate for any new data, changing data, or the appearance of previously occluded data. After translation of ${}^{i}\mathbf{D}_{S}$ from viewpoint *i* to viewpoint *j*, the data sets are within close enough proximity to each other that an assumption of, or preference for, stationarity would require corresponding data in the data sets that is not new, changed or previously occluded to be within some threshold, ε, of each other. To align the data sets using the assumption of stationarity, any points exceeding that

threshold should be excluded from the final registration process.

Prior to performing this distance check, however, it is first necessary to insure that the data sets are ordered by corresponding points, i.e. that each point ${}^{j}\mathbf{p}_{a_n,T}$ in ${}^{j}\mathbf{D}_T$, corresponds to ${}^{k}\mathbf{p}_{a_n,S}$ in ${}^{k}\mathbf{D}_S$ for all points n in the world coordinate system. This can be accomplished by identifying and removing any points in the data sets that are not within shared areas of the data sets, as defined by the their extreme boundaries in the allocentric coordinate system.

The removal of these outlying points from their respective data sets results in data sets ${}^{j}\mathbf{D}_N$ and ${}^{k}\mathbf{D}_N$ that contain only points in a common area of the world coordinate system and that are ordered by corresponding points. Here the subscript N indicates data sets corresponding to an intersection of space in world coordinates.

It is a computationally inexpensive process to identify those points in the data sets where $\left\| {}^{j}\mathbf{p}_{a_n,N} - {}^{k}\mathbf{p}_{a_n,N} \right\| > \varepsilon$ and to remove those points from each of the data sets. This results in data sets ${}^{j}\mathbf{D}_C$ and ${}^{k}\mathbf{D}_C$ containing only the *consistent* components of the data sets. These data sets can now be finely aligned using a standard technique such as ICP.

C. Bio-Inspired Algorithm

The use of macro level registration based on theories of human perception combined with micro level registration using the ICP algorithm, allows (2) to be solved iteratively in the following steps (representative data sets are shown for illustrative purposes):

1) Data sets (range images) are captured at viewpoints i and k.

2) ${}^{j}\mathbf{D}_T$ is constructed from estimates of the rotation and translation between images using (4).

3a) Areas of each image, ${}^{j}\mathbf{D}_T$ and ${}^{k}\mathbf{D}_S$, that are not within a common area of the world are identified.

3b) These areas of the images are removed, resulting in data sets ${}^{j}\mathbf{D}_N$ and ${}^{k}\mathbf{D}_N$.

4a) Corresponding points in ${}^{j}\mathbf{D}_N$ and ${}^{k}\mathbf{D}_N$ that have a large Euclidean separation are identified as inconsistencies.

4b) The stationarity assumption is applied by removing these points from both data sets to create consistent data sets ${}^{j}\mathbf{D}_C$ and ${}^{k}\mathbf{D}_C$.

5) ICP is performed on the resulting data sets to determine ${}^{j}\mathbf{R}_{k,I}$ and ${}^{j}\mathbf{T}_{k,I}$.

6) The overall estimate of the transformation from viewpoint i to viewpoint k is estimated as:

$${}^{i}\hat{\mathbf{R}}_{k,A} = {}^{i}\mathbf{R}_{j,S}\,{}^{j}\mathbf{R}_{k,I}$$

$${}^{i}\hat{\mathbf{T}}_{k,A} = {}^{i}\mathbf{T}_{j,S} + {}^{i}\mathbf{R}_{j,I}\,{}^{j}\mathbf{T}_{k,I}$$

V. Test Results

Three conditions were tested to verify the performance of the bio-inspired registration algorithms in comparison to the standard ICP algorithm in scenarios relevant to autonomous ground navigation. Experiments were run to simulate images from a 3-D laser detection and ranging sensor mounted on a vehicle driving through an urban environment, on a stationary vehicle scanning a scene, and finally on a stationary vehicle observing a scene while experiencing periods of high occlusion. These scenarios also represent the most challenging areas for ICP; images with large amounts of transformation and images with a high degree of inconsistency.

Three performance characteristics were measured to compare performance. Convergence speed was calculated by determining the iteration at which the algorithm had converged to within 95% of its final values for every transformation parameter. Translation error was calculated by subtracting the Euclidean distance between the translation calculated by the algorithm from the actual translation between the data sets. Rotation error was calculated by subtracting the rotation parameters calculated by the algorithm from the actual rotation values.

A. Moving Vehicle

Vehicle motion was simulated by moving the sensor viewpoint in half meter increments in the x and y dimensions for a total translation of ten meters in both directions. Figure 1 illustrates the motion of the vehicle during the experiment. Test cases for data collection consisted of the starting image and the image captured at the current sensor location, e.g. images for Case 10 consisted of the starting image and the image captured after moving the vehicle 5 meters in both the x and y dimensions.

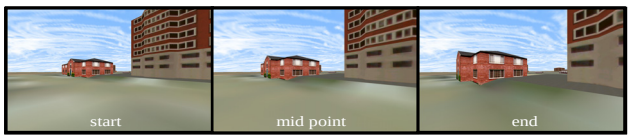


Fig. 1. Camera Images from Vehicle Motion Experiment

The biologically inspired registration algorithm outperformed the standard ICP algorithm for the moving vehicle experiment in all cases for registration speed and for registration accuracy as seen in Figures 2, 3, and 4. The major benefit of the new algorithm is seen when there is a large amount of transformation between images. The ICP algorithm stopped returning reasonable amounts of error when separation between the images exceeded 4.5 meters in both the x and y directions. The biologically inspired algorithm continued providing accurate results up to the maximum tested separation of 10 meters in both the x and y directions.

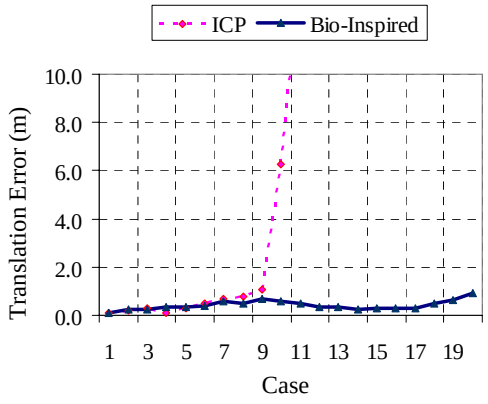


Fig. 2. Camera Images from Vehicle Motion Experiment

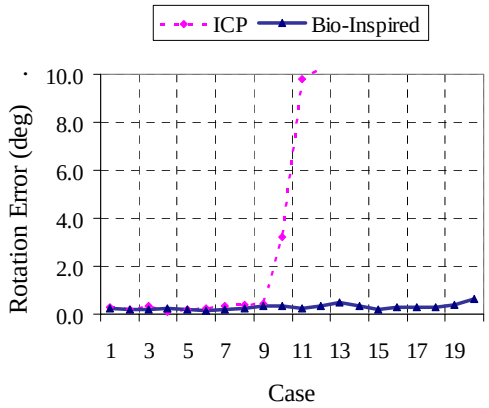


Fig. 3. Camera Images from Vehicle Motion Experiment

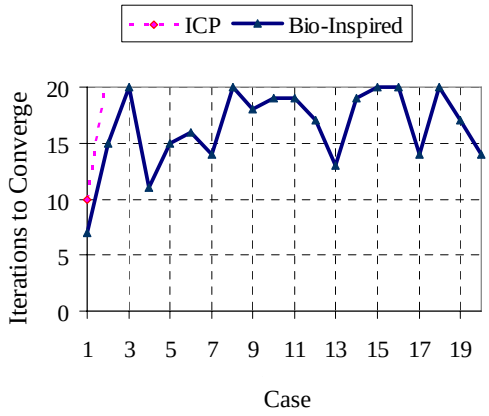


Fig. 4. Camera Images from Vehicle Motion Experiment

B. Scanning Sensor

Sensor scanning was simulated by fixing the sensor viewpoint and moving the sensor field of view in 0.5 degree increments for a total rotation of 10 degrees. Experiments were run with the sensor panning only, with the sensor tilting only, and then with the sensor panning and tilting. Figure 2 illustrates the sensor motion during each of these experiments.

.

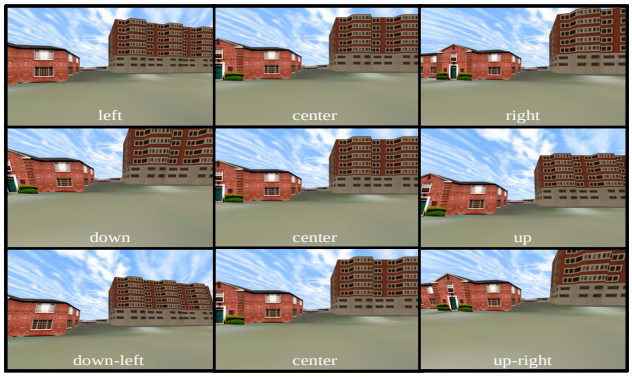
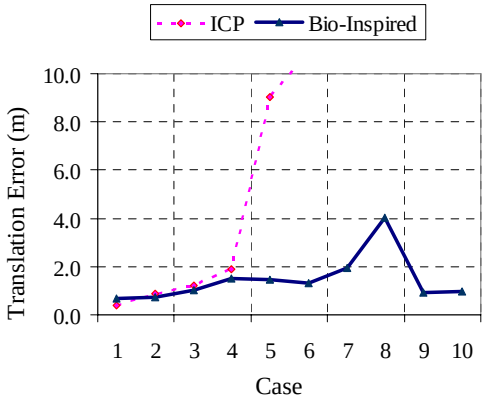


Fig. 5. Camera Images from Pan (top row), Tilt (second row) and Combined Pan and Tilt (bottom row) Experiments

Similar to the moving vehicle results, the biologically inspired algorithm proved much faster than the standard ICP algorithm during each of the sensor scanning experiments and outperformed the ICP algorithm for accuracy in all but two trials. For the experiments involving only panning of the sensor, the ICP algorithm was not able to accurately register the images for large amounts of panning, while the bio-inspired algorithm continued to perform well. Both algorithms were more sensitive to tilting than panning, which can be attributed to the comparatively smaller field of view of the sensor in elevation. Overall, the results were consistent, with the ICP algorithm experiencing significantly higher rates of error at relatively small amounts of sensor rotation, converging more slowly than the biologically inspired algorithm, and diverging at high amounts of rotation. Figures 6, 7, and 8 illustrate the results for the sensor pan and tilt experiment.

Fig. 6. Camera Images from Vehicle Motion Experiment



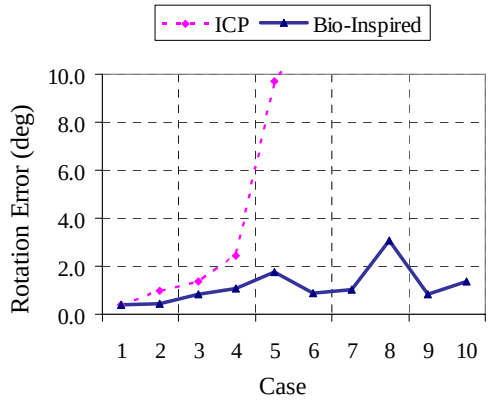


Fig. 7. Camera Images from Vehicle Motion Experiment

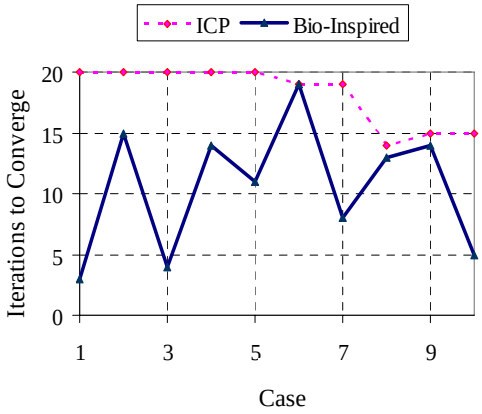


Fig. 8. Camera Images from Vehicle Motion Experiment

B. Sensor Occlusion

Sensor occlusion was simulated by fixing the sensor viewpoint and field of view and then moving a large object through the scene being observed by the sensor. Five images were captured of a truck moving through the scene in order to cover the range of cases from no occlusion to almost total occlusion. Figure 9 illustrates the images used during these experiments. Test cases for data collection consisted of the starting image, 'truck 1' and the image captured at the current truck location, e.g. images for Case 2 consisted of the starting image and the image labeled 'truck 2'.

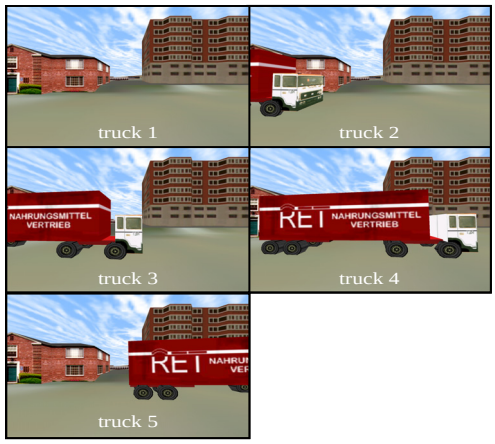


Fig. 9. Camera Images from Moving Object Experiment

Sensor occlusion is the most difficult test for voxel based registration techniques. In several of the images tested, 50% or more of the pixels from the base image were occluded in the second image, a scenario which proved impossible for the standard ICP algorithm to address. The ICP algorithm demonstrated poor accuracy and diverged, rather than converged, on a solution. On the contrary, the biologically inspired algorithm performed rapidly and with exceptional accuracy in all cases.

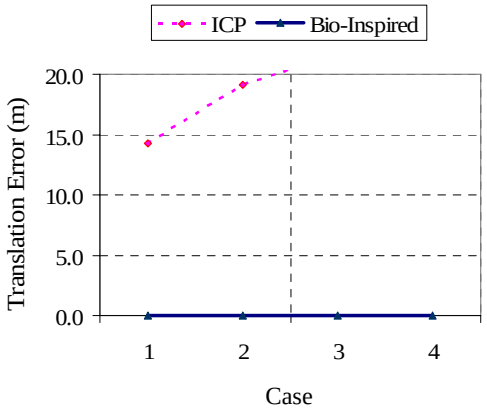


Fig. 10. Camera Images from Moving Object Experiment

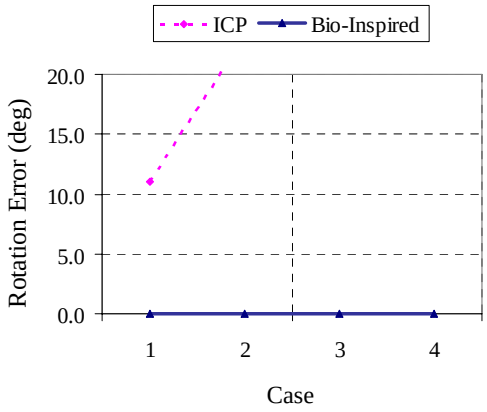


Fig. 11. Camera Images from Moving Object Experiment

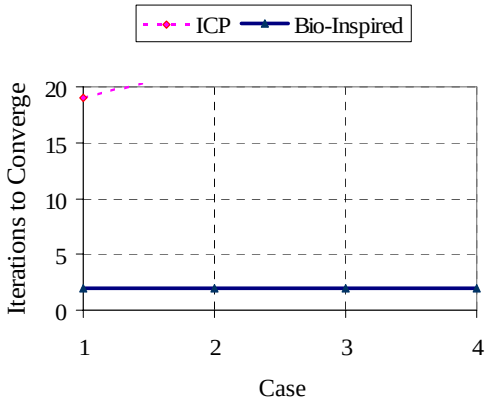


Fig. 12. Camera Images from Moving Object Experiment

D. Computational Efficiency

In addition to improved performance, the bio-inspired technique also improves computation time. On the computing platform used for this research, the combined computing time of the saccadic and staionarity corrections was only 10% of the computing time of a single ICP iteration. In addition, the ICP computing time on the resulting data sets was an average of 31% less per iteration than that required for data sets without the corrections. Together, this results in a decrease of the overall registration computation time.

VI. Conclusion

Registration of three-dimensional data sets, and image registration in particular, is a fundamental problem for a wide variety of applications ranging from medical imaging to autonomous ground navigation. Significant research has been conducted in this area, with the combined goal of achieving high-speed, precise registration of data sets. Particular difficulty persists in cases of noisy data sets requiring either significant transformation or those with a large degree of inconsistency.

This research developed a biologically inspired algorithm based on two theories of human perception that addresses these difficulties and performed extensive testing on data sets associated with the application of autonomous ground navigation.

The biologically inspired algorithm outperformed the ICP algorithm in all cases, with the most significant benefits being found in cases involving a large amounts of translation and/or rotation between the images, for which the ICP algorithm returned unrealistic results, and for cases where there were large amounts of occlusion between the images, for which the ICP algorithm actually diverged rather converged to a solution. Both of these cases are extremely important to autonomous ground navigation. The excellent performance of the biologically inspired algorithm in these cases makes it a promising candidate for this field.

ACKNOWLEDGMENT

The author conducted the work described in this paper as part of his graduate studies and would like to thank Dr. Ka C. Cheok, his advisor, Mr. Scott Pletz, who provided the simulated camera and ladar images that served as the data sets for the autonomous navigation case study, and Dr. Raj Madhaven from the National Institute of Standards and Technology, who shared insight into his work with the Iterative Closest Point algorithm.

REFERENCES

[1] L. G. Brown. “A Survey Of Image Registration Techniques”, ACM Computing Surveys, 24(4):325–375, December 1992.

[2] P. A. Van den Elsen, E.-J. D. Pol, and M. A. Viergever. ”Medical Image Matching – A Review With Classification”, IEEE Engineering Medical Biology Magazine, pages 26–39, March 1993.

[3] R. Madhavan and E. Messina, “Iterative Registration of 3D LADAR Data for Autonomous Navigation”, Proceedings of the IEEE Intelligent Vehicles Symposium, pages 186-191, June, 2003.

[4] B. K. P. Horn. “Closed-Form Solution Of Absolute Orientation Using Unit Quaternions”, Journal of the Optical Society of America, 4(4):629–642, 1987.

[5] K. Arun, T. Huang, and S. Bolstein. “Least-Squares Fitting of Two 3-D Point Sets”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 9(5):698–700, 1987.

[6] J. O’Regan, “Solving the ‘Real’ Mysteries of Visual Perception: The World as an Outside Memory”, Canadian Journal of Psychology, pages 461-488, 1992.

[7] F. Volkmann, A. M. L. Schick, L. A. Riggs. “Time Course of Visual Inhibition During Voluntary Saccades”, Journal of the Optical Society of America, 58, 562-569, 1968.

[8] L. Matin, E. Matin, D. G. Pearce, “Visual Perception Of Direction When Voluntary Saccades Occur”, Perception & Psychophysics, 5, 65-79, 1969.

[9] E. Von Holst, H. Mittelstaedt. “The Principle Of Reafference: Interactions between the Central Nervous System and the Peripheral Organs”, Perceptual processing: Stimulus equivalence and pattern recognition. New York: Appleton, 1971

[10] B. Bridgeman, A. H. C. Van der Hejiden, B.M. Velichkovsky, “A Theory Of Visual Stability Across Saccadic Eye Movements”, Behavioral and Brain Sciences 17 (2): 247-292, 1994.

[11] M. Wexler, “Anticipating The Three-Dimensional Consequences Of Eye Movements”, Proceedings of the National Academy of Science, U. S. A. 102, 1246–1251, 2005.

[12] H. Wallach, D. N. O'Connel. "The Kinetic Depth Effect", Journal of Experimental Psychology, 45: 205-217, 1953.
[13] M. Wexler, F. Panerai, I. Lamouret and J. Droulez, "Self-Motion and the Perception of Stationary Objects", Nature, 409: 85-88, 2001.
[14] M. Wexler, I. Lamouret and J. Droulez, "The Stationarity Hypothesis: An Allocentric Criterion in Visual Perception", Vision Research, 41:3023-3037, 2001.