3D POINT CLOUD REGISTRATION WITH SHAPE CONSTRAINT

Swapna Agarwal, Brojeshwar Bhowmick

Embedded System and Robotics, TCS Research & Innovation, India

ABSTRACT

In this paper, a shape-constrained iterative algorithm is proposed to register a rigid template point-cloud to a given reference point-cloud. The algorithm embeds a shape-based similarity constraint into the principle of gravitation. The shape-constrained gravitation, as induced by the reference, controls the movement of the template such that at each iteration, the template better aligns with the reference in terms of shape. This constraint enables the alignment in difficult conditions indtroduced by change (presence of outliers and/or missing parts), translation, rotation and scaling. We discuss efficient implementation techniques with least manual intervention. The registration is shown to be important for change detection in the 3D point-cloud. The algorithm is compared with three state-of-the-art registration approaches. The experiments are done on both synthetic and real-world data. The proposed algorithm is shown to perform better in the presence of big rotation, structured and unstructured outliers and missing data.

Index Terms— Point-cloud registration, Shape registration, gravitational approach, change detection

1. INTRODUCTION

With the increasing availability of 3D data acquisition devices such as Kinect [1], [2], the researchers are inclining towards exploiting the 3D information for a variety of processes. Registration of 3D point-clouds is a necessary initial step for many such processes. Examples include change detection in a scene, industrial quality control, pose tracking [3] etc. Generally, two point-clouds (called reference and template respectively) of an object or scene are captured at two time instances, possibly from different camera positions. The scene may have gone under some changes during this time. The registration refers to the process of translating, rotating and scaling the template such that it optimally aligns with the reference. In this paper, we focus on the automatic rigid registration of 3D point-clouds. The proposed algorithm does not require any color, texture, point correspondence or point topology information of the point-clouds. Moreover, presence of noise, outliers and missing parts make the problem even more challenging. Now, we briefly review the related literatue and state our contributions (Fig. 1).

One of the earliest and most popular method for rigid

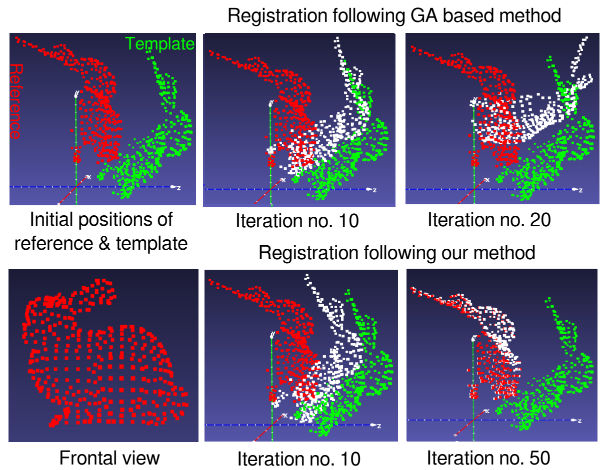


Fig. 1. Example comparing the registration results of the proposed method vs. that of GA [4]. The frontal view of the bunny data shows that the lower body part is heavier compared to the head. In GA, the heavier parts of the template (green) get attracted more by the reference (red), irrespective of shape, resulting in misalignment (first row). In our method, the parts of the template get attracted by the parts of the reference having similar shape. This gives better alignment (white) even in presence of large rotation. In the example, the template is given a $50°$ of rotation wrt the x-axis.

point set registration is Iterative Closest Point (ICP) [5], [6]. Using non-linear optimization algorithm, ICP minimizes the mean squared distances between the two point sets. This method is simple to implement but is sensitive to outliers and the performance depends on the initial alignment [4]. Some robust variants [7], [8], [9] of ICP are also there.

While ICP [5] assigns discrete correspondences between the points of the two sets, later methods such as Robust Point Matching (RPM) [10] and [11], [12], [13] assign soft correspondences. A set of methods [14], [15] treat the registration problem as maximum likelyhood estimation problem where the template points are assumed to be the centroids of Gaussian Mixture Model (GMM) and the reference points are seen as data points. Expectation-Maximization (EM) algorithm is generally used for optimization of likelyhood function. A closed form solution to the M-step of the EM algorithm for multidimensional case is proposed in [16]. Their method is called Coherent Point Drift (CPD). The GMM based meth-

978-1-5090-2175-8/17/$31.00 ©2017 IEEE

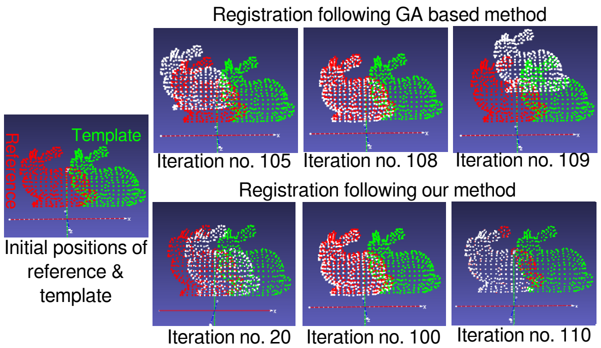


Fig. 2. Example comparing iterative alignment results of the proposed registration method with that of GA [4]. The white point cloud shows the registraion result. The proposed algorithm reaches local minima very fast (20 vs. 105 iterations in the second column). Near local minima, our approach takes smaller steps while GA overshoots (third to fourth columns).

ods perform better in the presence of noise. CPD was refined in [17] and [18] for structured outliers. Other recent mentionable approaches are [19], [20] and [21]. Recently in 2016, a new Gravitational Approach (GA) based registration algorithm is proposed [4]. Under GA algorithm, the template moves towards the reference under the influence of gravitaional force as induced by the reference. GA based method is shown to perform better compared to CPD under more than 50% uniform noise [4]. The performance of GA decreases with increasing rotation (more than 45^o) and change (extra or missing parts). We propose substantial modifications to the GA algorithm [4] to overcome the disadvantages of GA.

Our contributions are as follows. (1) GA method minimizes the distance between the center-of-mass of the two point-clouds but does not ensure proper shape alignment (see Fig. 1). We constrain the force of attraction with a measurement of local shape-similary for better alignment of the shapes (i.e., handles rotation and outliers better). (2) In GA, the force of attraction is made to be inversely proportional to the distance beetween the two points involved. As a result, the increased speed at local minima (where the distance is small) does not allow the algorithm to converge (see Fig. 2). To handle this situation, a number of free parameters are introduced in GA. In our algorithm, the force of attraction is proportional to a monotonically increasing function of distance. This modification allows for better convergence without the need for extra free parameters. (3) GA has limited capability to handle scale change [4]. In our method, to handle scale change, an orientation and translation invariant model is proposed that uses the spatial point distribution of the two point-clouds. (4) The new algorithm is evaluated extensively for fully automatic performance as opposed to a number of (seven) free, manually adjusted parameters of GA method [4].

The proposed algorithm is compared with ICP [5], CPD [16] and GA [4] in Section 5 and the conclusions are drawn in Section 6. Next, we briefly present the GA based registration method following which we elaborate the proposed shape-constrained registration algorithm.

2. GRAVITATIONAL APPROACH

In Gravitational Approach (GA) based registration [4], each point in one point set (called reference) attracts each point in the other point set (called template). The force of attraction is governed by the following formula.

$$\mathbf{f}^{Yi} = -Gm^{Yi} \sum_{j=1}^{N} \frac{m^{Xj}}{(\left\|r^{Yi} - r^{Xj}\right\|^2 + \epsilon^2)^{3/2}} \mathbf{n}_{ij} - \eta v^{Yi} \quad (1)$$

In (1) G represents the gravitational constant and $\mathbf{f}^{Yi}$ is the total force applied on a point Yi of the template. The symbols Xj, m^{Xj}, r^{Xj} and N rerpesent the jth point, mass of Xj, absolute coordinates of Xj and number of points in the reference respectively. The symbols $\mathbf{n}_{ij}$, v^{Yi} and η represent a unit vector with the same direction as force, the velocity of Yi in the previous iteration and a constant respectively. The template, as a rigid body, gets displaced in an iterative way under the influence of the cumulative force induced by all the points in the reference. The term ϵ^2 does not allow the force to increase beyond a certain small distance between the reference and the template (note that $\mathbf{f}^{Yi} \propto 1/(\left\|r^{Yi} - r^{Xj}\right\|^2)$). The term $-\eta v^{Yi}$ in (1) acts as friction that controls the velocity near local minima. Given the force, the displacement of the template is estimated following Newton's second law of motion. The scale and rotation, required for registration, are estimated using the new positions of the template points which are estimated following (1).

3. THE PROPOSED REGISTRATION APPROACH

The objective of the gravitational force is to minimize the distance between the center-of-mass of the two objects. On the other hand, the objective of registration is to align the two objects (point-clouds here) such that the parts of the objects having similar shapes map to each other. Therefore, we need to modifiy the principles of gravitation to accommodate this constraint. We explain these modifications next.

3.1. Modified Gravitational Approach

In GA, $\mathbf{f}^{Yi} \propto m^{Xj} m^{Yi}$. As a result, in registration following (1), the heavier parts of the point-clouds attract each other more irrespective of the shape, resulting in mis-alignment (e.g., Fig. 1). Therefore, we modify (1) such that,

$$\mathbf{f}^{Yi} \propto g(s^{Xj,Yi}), \quad (2)$$

where $s^{Xj,Yi}$ represents a measure of shape similarity of local neighborboods centered at Xj and Yi respectively. g is a monotonically increasing function of s. We have evaluated different features (e.g., histogram of normals, coefficients of polynomial approximating the local surface, curvature) that can represent shape in spatially local domain. The curvature

[22] seems to be the best representative of local shape. As a similarity measure of shape, we can use any Radial Basis Function (RBF) where the output decreases monotonically as the dissimilarity increases. In our implementation, we use

$$s^{Xj,Yi} = exp(-\frac{|a^{Xj} - a^{Yi}|^2}{\sigma^2}), \quad (3)$$

where, $|.|$ represents absolute value, σ and a^{Xj} represent the spread of the RBF funtion and the curvatue value of the neighborhood centered at Xj. In the proposed registration approach, we want the template to make large movement if the distance between the template and the reference is large. We want it to take tiny steps for fine-tuning when close to alignment. This is ensured by the following formula.

$$\mathbf{f}^{Yi} \propto h(\|r^{Yi} - r^{Xj}\|). \quad (4)$$

The value of the function g in (2) or the function h in (4) can be the value of the argument itself or some suitable monotonically increasing function as we shall discuss in Section 4. Therefore, we propose the following formula to estimate the total force applied on a point Yi of the template.

$$\mathbf{f}^{Yi} = -G \sum_{j=1}^{N} g(s^{Xj,Yi}) h(\|r^{Yi} - r^{Xj}\|) \mathbf{n}_{ij}. \quad (5)$$

Note that in (5) we could drop two free parameters ϵ and η as used in (1) in GA based registration [4].

3.2. Estimating Translation, Rotation and Scale

We estimate the displacement represented by say, d of the template following Newton's second law of motion.

$$d = (\frac{\mathbf{f}}{m^X} time + v^X) \times time. \quad (6)$$

In (6) $\mathbf{f}$, m^X and v^X represent the total force ($\sum_i^M \mathbf{f}^{Yi}$) applied on the template X, mass of X and velocity of X in a previous iteration repsectively.

The rotation matrix R is estimated following Kabsch algorithm [23]. Let Y and Y_D represent the mean subtracted $M \times 3$ matrices representing the M coordinates of the template before and after translation following (6). Let $\hat{C} = Y_D^T Y$ represent the corss-covariance matrix. The rotation matrix R is estimated using Singular Value Decomposition (SVD) of $\hat{C}$. After SVD $\hat{C} = \hat{U}\hat{S}\hat{V}^T$ and let $s' = sign(det(\hat{V}\hat{U}^T))$. Then,

$$R = \hat{V} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s' \end{bmatrix} \hat{U}^T. \quad (7)$$

In GA, the scale is estimated as the ratio of the positions estiamted using (1) and previous positions of the template cloud points. Any error occuring in estimating the translation using (1) gets propagated to estimating the scale at each iteration of GA based registration. Our model does not depend on the estimation of translation. We find the eigen-valules of the covariance matrix (say C) of each of the two point-clouds. The largest eigen-value represents a measure of the lengh of the point-cloud distribution in the direction of the largest variance. This measurement is independent of the orientation or relative position of the two point-clouds. Using SVD we have $C = USV^T$ where eigen-values are represented by the diagonal elements of S and U and V are orthogonal matrices. We estimate the scale as $c = \frac{eX}{eY}$ where eX and eY are the largest eigen-values of the reference and the template respectively. We scale the template (Y) as.

$$Y' = V diag(c) V^T Y^T(t). \quad (8)$$

In (8), t stands for iteration number and $diag(c)$ represents a square matrix with c in the diagonal and zeros elsewhere. In each iteration of our algorithm, the template is updated using

$$Y(t+1) = Y'R + \mu + d. \quad (9)$$

In (9) μ is a $M \times 3$ matrix where each row represents the mean position of the template.

4. IMPLEMENTATION DETAILS AND DISCUSSION

The value of shape similarity $g(s^{Xj,Yi})$ is the main factor controlling the translation and rotation of the template. If we give equal weightage to all the local shapes then we set $g(s^{Xj,Yi}) = s^{Xj,Yi}$. In indoor and some ourdoor scenes, non-planar local shapes play vital role in registration compared to planar shape. Therefore, we can set g as follows to give more weightage to non-planar local shapes.

$$g(s^{Xj,Yi}) = a^{Xj} a^{Yi} s^{Xj,Yi}. \quad (10)$$

For finding the curvature (a), setting the radius of the local neighborhood to $\sqrt{eX}$ and $\sqrt{eY}$ respectively yields good results. We design h of (4) as follows.

$$h(\|r^{Yi} - r^{Xj}\|) = h(\|r^{Yi} - r^{Xj}\|^p), \text{where } p \geq 1. \quad (11)$$

Higher value of p results in greater speed but oscillation about the local minimal. Lower value of p results in better registration but makes the process slow. The mass m of every point is assumed to be 1 and v^X is set to 0. m and v^X can be modified depending on prior information, if any. We also suggest that G of (5) should take larger value near convergence when the algorithm takes smaller steps. We modify G as follows.

$$G(t+1) = 0.1 \times G1 + 0.9 \times G1/(1 + exp[-\{t - E/2\}]) \quad (12)$$

where, we set $G1 = 10000$ and $E = 400$ is the number of iterations. For noisy real scenarios, pyramidal approach can be taken where the size of the local neighborhood and σ of (3) decreases with iteration.

5. EVALUATION

On Synthetic data: We evaluate the performance of the proposed method on data with different amounts and types of noise (structured and unstructured outliers, missing data). In Fig. 3 we compare our method with GA [4] in the presence of structured outliers. When the mass (no. of points) in the outlier is less than the reference (human scan), both the methods work well (Fig. 3, col 1). As the mass of the outlier increases, following GA, the outlier rather than the human shape moves towards the reference (Fig. 3, col 2, 3). Following our method, regardless of the mass of the outlier, the human shape (downloaded from [24], [25]) in the template registers with the human shape in the reference. Here comes into light the advantage of our method over GA [4]. Unlike GA where the movement of the template depends on the mass, in our method the movment depends on the shape.

Fig. 4a and b show the performance of our method on data with missing points. We delete 18% consecutive points from the head of the Armadillo [26] data. For both the cases when the full Armadillo or the one with missing points is used as the template and the other one as the reference, our method gives desired result (Fig. 4) whereas the competing methods give less accurate results.

Fig. 4c and d show the results when 50% Gausian and Uniform outliers respectively are added to the data. To quantitatively compare the proposed approach with that of GA [4], CPD [16] and ICP [5], we follow the same evalutaion protocol as used in [4] and use the same bunny point-cloud from Stanford 3D Scanning Repository [26] as used in [4] and [16]. We add 5%, 10%, 20%, 40% and 50% Gaussian and uniform noise to the template. For each of the 10 noise type-percentage combinations, 500 random tranformations are applied on the point-cloud to form the template. The non-transformed point-cloud acts as the reference. Fig. 5 compares the the amout of noise vs. Root Mean Square Error (RMSE) for the four competing methods. All the methods perform better for Gaussian compared to uniform noise. Overall CPD [16] performs better compared to ICP [5]. Our method outperforms all the three competing methods.

On Real Data: To evaluate the proposed method in real scenarios, we capture 3D point-cloud scans of a number of scenes, two captures for each scene using Kinect. Fig. 6 demonstrates one example. In the second capture, the position of the remote-controller has been changed and the front box has been opened. To make the problem more difficult, we translate the template by 0.2 meters along x-axis. The registration result shows proper alighment with the front box, cup and the background box. Notice that the cover of the front box and the remote-controller have not been aligned with anything as the postion of these items have been changed. Thus, our registration approach can be used for change detection in real scenes.

6. CONCLUSIONS

We have proposed a novel rigid point-set registration algorithm with special characteristics (e.g., shape constraint, translation proportional to distance, spatial point-set distribution model for handling scale) that outperforms other competing approaches [4], [16] and [5]. The proposed approach registers better in difficult conditions such as missing object part, rotation more than 50^o and different amounts of structured and unstructured outliers. We plan to employ the proposed registration approach for efficient change detection.

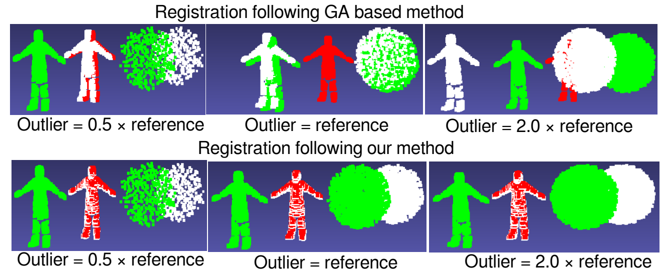
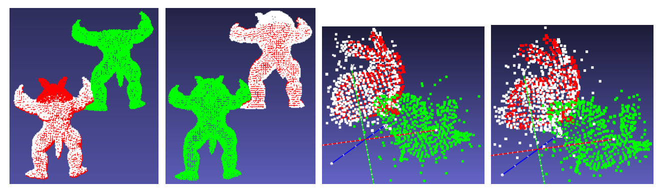


Fig. 3. Comparison of our method with GA [4] in the presence of structured outlier. With increasing mass of the outlier, the registration performance following GA degrades whereas our method consistently performs well. Red: reference, green: template, white: template after registraion.



(a) (b) (c) (d)

Fig. 4. Registration results of our method on point-clouds with 18% deleted points (a, b) and 50% outliers (c, d). In (a), (b), (c) and (d), the partial cloud, original full cloud, cloud with Gaussian outlier and Uniform outlier respectively are used as the template.

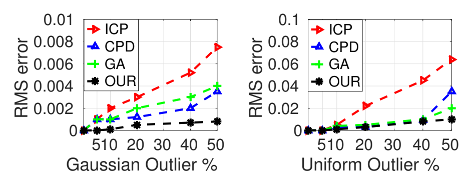


Fig. 5. Comparison of the proposed method with GA [4], ICP [5] and CPD [16] in the presence of unstructured outliers.

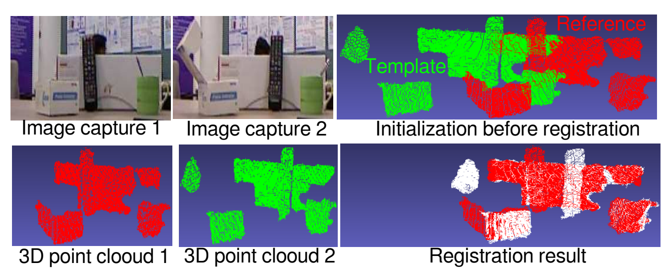


Fig. 6. Registration result on real scene.

7. REFERENCES

[1] Z. Zhang, "Microsoft kinect sensor and its effect," *IEEE multimedia*, vol. 19, no. 2, pp. 4–10, 2012.

[2] "Microsoft kinect," `https://developer.microsoft.com/en-us/windows/kinect`, Accessed: 4th Jan, 2017.

[3] S. Ge and G. Fan, "Sequential non-rigid point registration for 3d human pose tracking," in *Int. Conf. on Image Processing*. IEEE, 2015, pp. 1105–1109.

[4] V. Golyanik, Sk A. Ali, and D. Stricker, "Gravitational approach for point set registration," in *Computer Vision and Pattern Recognition*. IEEE, 2016, pp. 5802–5810.

[5] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[6] Z. Zhang, "Iterative point matching for registration of free-form curves and surfaces," *International journal of computer vision*, vol. 13, no. 2, pp. 119–152, 1994.

[7] A. W. Fitzgibbon, "Robust registration of 2d and 3d point sets," *Image and Vision Computing*, vol. 21, no. 13, pp. 1145–1153, 2003.

[8] Jeff M Phillips, Ran Liu, and Carlo Tomasi, "Outlier robust icp for minimizing fractional rmsd," in *3-D Digital Imaging and Modeling*. IEEE, 2007, pp. 427–434.

[9] Lei Zhang, Sung-In Choi, and Soon-Yong Park, "Robust icp registration using biunique correspondence," in *3D Imaging, Modeling, Processing, Visualization and Transmission*. IEEE, 2011, pp. 80–85.

[10] S. Gold, A. Rangarajan, C. Lu, S. Pappu, and E. Mjolsness, "New algorithms for 2d and 3d point matching: Pose estimation and correspondence," *Pattern recognition*, vol. 31, no. 8, pp. 1019–1031, 1998.

[11] B. Luo and E. R. Hancock, "A unified framework for alignment and correspondence," *Computer Vision and Image Understanding*, vol. 92, no. 1, pp. 26–55, 2003.

[12] H. Chui and A. Rangarajan, "A new algorithm for non-rigid point matching," in *Computer Vision and Pattern Recognition*. IEEE, 2000, vol. 2, pp. 44–51.

[13] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[14] G. McNeill and S. Vijayakumar, "A probabilistic approach to robust shape matching," in *Int. Conf. on Image Processing*. IEEE, 2006, pp. 937–940.

[15] M. Sofka, G. Yang, and C. V. Stewart, "Simultaneous covariance driven correspondence (cdc) and transformation estimation in the expectation maximization framework," in *Computer Vision and Pattern Recognition*. IEEE, 2007, pp. 1–8.

[16] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *Pattern analysis and machine intelligence*, vol. 32, no. 12, pp. 2262–2275, 2010.

[17] P. Wang, P. Wang, Z. Qu, Y. Gao, and Z. Shen, "A refined coherent point drift (cpd) algorithm for point set registration," *Science China Information Sciences*, vol. 54, no. 12, pp. 2639–2646, 2011.

[18] V. Golyanik, B. Taetz, G. Reis, and D. Stricker, "Extended coherent point drift algorithm with correspondence priors and optimal subsampling," in *Winter Conference on Applications of Computer Vision*. IEEE, 2016, pp. 1–9.

[19] S. Biswas, G. Aggarwal, and R. Chellappa, "Invariant geometric representation of 3d point clouds for registration and matching," in *Int. Conf. on Image Processing*. IEEE, 2006, pp. 1209–1212.

[20] M. Marques and J. Costeira, "Guided search consensus: Large scale point cloud registration by convex optimization," in *Int. Conf. on Image Processing*. IEEE, 2013, pp. 156–160.

[21] F. Peng, Q. Wu, L. Fan, J. Zhang, Y. You, J. Lu, and J. Yang, "Street view cross-sourced point cloud matching and registration," in *Int. Conf. on Image Processing*. IEEE, 2014, pp. 2026–2030.

[22] "Point cloud library," `http://pointclouds.org/documentation/tutorials/normal_estimation.php`, Accessed: 4th Jan, 2017.

[23] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics, Diffraction, Theoretical and General Crystallography*, vol. 32, no. 5, pp. 922–923, 1976.

[24] "Microsoft kinect," `https://sites.google.com/site/pgpapadakis/home/CAPOD`, Accessed: 11th Jan, 2017.

[25] P. Papadakis, "The canonically posed 3d objects dataset," in *Eurographics Workshop on 3D Object Retrieval*, 2014, pp. 33–36.

[26] "Stanford university computer graphics laboratory," `https://graphics.stanford.edu/data/3Dscanrep/`, Accessed: 11th Dec, 2016.