Global Hybrid Registration for 3D Constructed Surfaces Using Ray-Casting and Improved Self Adaptive Differential Evolution Algorithm

Linh Tao, Tinh Nguyen, and Hiroshi Hasegawa

Abstract—As a fundamental task in computer vision, registration has been a solution for many application such as: world modeling, part inspection and manufacturing, object recognition, pose estimation, robotic navigation, and reverse engineering. Given two images and set ones as the model, the aim is to find the best possible spatial transformation matrix causing 3D reconstruction of original object. The paper presents a new hybrid algorithm which improves both speed and convergence guarantee in comparison recently proposed methods of registering structured pointcloud surfaces by using a fast error calculation ray-casting based closest point method integrated with a new developed global optimization method Improve Self Adaptive Differential Evolution (ISADE). Ray-casting based $L_2$ error calculation method enables the algorithm to find the local minima error while ISADE exploit the searching boundary to find the global minima. The new algorithm is evaluated to show the significant improvement in quality and robustness to state-of-the-art methods.

Index Terms—3D registration, ISADE, hybrid global registration, Ray-casting.

I. INTRODUCTION

The introduction of commercial depth sensing devices such as Microsoft Kinect, Asus Xtion, etc. has shifted robotics, computer vision research areas from 2D based imaging and laser scanning toward 3D based depth scenes of environment processing. As a physical object or scenario cannot be completely captured with a single image, different images from different time and positions need to be aligned into a more completed view of the scenario; the process of alignment is called registration. Registration algorithms estimate the movement of the camera through calculating the transformation that optimally maps two point clouds. Various applications such as 3D object scanning, 3D mapping, and 3D localization use registration algorithms as backbone algorithms. According to how many views or images of the objects are processed at the same time, registration strategies are divided into multi-view registration (for all views case) and pair-wise registration (for two views case). Our paper focuses on the pair-wise registration of constructed range images taken by 3D cameras. As a consequence, starting from two views, i.e., the model and the data, the objective of our registration process is to find the best homogeneous transformation that, when applied to the data, aligns it with the model in a common coordinate system.

Iterative Closest Point (ICP) [1] and its variants such as non-linear ICP, generalized ICP and non-rigid ICP have been always indispensable tools in registration tasks. ICP's concept and implementation are easy to understand. ICP uses $L_2$ errors estimated from pair-wise point-clouds to derive a transformation which draws them closer to each other. Registration process finishes after many iterations of minimizing error and results in a homogeneous transformation.

However, ICP-class algorithms alone cannot solve problems for general registration tasks since they require a further assumption in which an initial near-optimal pose transformation is necessary for right convergences. Otherwise, the registration process would likely converge to local optimal solutions instead of the global optimal or near global optimal one. This result cannot be overcome merely by iteration procedure. In some mesh and point-cloud editor software such as Meshlab [2] registering tool for range data is available. It requires manually data pre-alignment from users before ICP comes into use.

To overcome the shortage of ICP-class methods, in general, registration processes are generally divided into two steps: coarse transformation or initialization and fine transformation. If two point-clouds are close enough, the first step could be omitted. Otherwise, the problem remains a big challenge for researchers. Coarse transformation, pre-alignment estimation or initialization solving has two approaches: local and global. Local methods use local descriptors (or signatures) such as PFH [3] and SIFT [4] which encode local shape variation in neighborhood points. If points with those descriptors appear in both registering point-clouds, initialization movement could be estimated by using sample consensus algorithms such as RANSAC [5]. The problem of local approaches is that those signatures are not always guarantied to appear on both registering point-clouds. On the other hand, global approaches take every point into account such as Go-ICP [6] and SAICP [7]. The biggest problem of those methods is computation cost in finding the corresponding points in point-clouds. If there are big numbers of point in point-clouds, the computation cost is going large. However, thanks to new algorithms especially heuristic optimal searching methods as well as the increasing in computer speed especially with parallel computing with multi-core CPU processor and Graphic Computation Unit (GPU) [8] it is possible to find solutions of global approaches of registration problem. After estimating coarse transformation, ICP algorithm is an efficient tool to find the fine transformation.

This paper proposes a new global registration method for...
3D constructed images without good initialization. It is called Global Hybrid Registration for 3D Constructed Surface Using Ray-casting and ISADE [9]. As other global registration methods, our method requires no local descriptors on works directly on raw scan surfaces. The method uses ray-casting based method for local minima searching together with ISADE as a search engine to find the global minima without using fine registration. Our method rapidly produces results at high rate convergence of the global optimization solution.

II. THREE DIMENSION REGISTRATION

This part summary some approaches for global range image registration task up to date. SVD, PCA [10] are integrated together with ICP as classical methods and global searching algorithms are integrated with ICP as in the most current methods.

A. ICP Algorithm

SVD and PCA have been used to find coarse transformation together with ICP as the fine transformation-estimating tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool. Original version of ICP algorithm relies on transformation together with ICP as the fine searching tool.

$$E(R, t) = \sum_{i=1}^{n} e_i(R, t) = \sum_{i=1}^{n} |R \cdot y_i + t - x_i|$$  \hspace{1cm} (1)

where $R$ and $t$ are rotation and translation matrix, $y_i$ is the corresponding point of $x_i$ denoted for its closest point in data point-cloud $Y$. There are some ICP variants, which rely on different categories to define closest points. Point-to-point and Point-to-plane are two popular examples. Equation 2 is used to search for closest point by Point-to-point category.

$$f^* = \arg\min_{j \in [1,...,n]} |R \cdot y_j + t - x_i|$$  \hspace{1cm} (2)

The iteration process is as following to archives the final transformation:

- Compute the closest model points for each data point as Equation 2.
- Compute the transformation $R$ and $t$ based on the error from Equation 1.
- Apply $R$ and $t$ to the data point-clouds.
- Repeat step 1, 2, 3 until error as Equation 1 smaller then a set tolerant or the procedure reaches its max iteration.

Step by step, ICP draws the data point-cloud closer to model point-cloud and the process stops at local minima. There are some variants of ICP algorithm based on different methods to calculate the transformation from error $E(R, t)$ and error itself as in LMICP [11] and SICP [12].

B. Global Hybrid Searching Algorithm

ICP algorithms are superior for registering close or pre-aligned point-cloud data; otherwise, it often converges wrongly. Global searching algorithms are solution to solve this problem since they are able to find the global minima instead of local one. To make the task of global searching algorithm less difficult, ICP are often applied to flatten the searching space. Fig. 1 and Fig. 2 show how ICP works as a flattening tool of objective functions. By using ICP, a complex fitness function in black turns into simpler one in red color. And with such a much more flatten fitness function, global searching method find a global minima more effectively.

The integration works well in case of point-cloud data with small point number. For large data case, ICP becomes slow and impossible for applying into real time applications. Our method integrates new global searching algorithm ISADE, which can handle complicated fitness functions without or few flattening process and fast error calculation method based on ray-casting corresponding searching algorithm which accelerates registration procedure.

III. METHOD OVERVIEW

A. Methodology Approach

The biggest disadvantage of ICP based registration methods in calculating cost function is runtime. In KinectFusion [13] a real-time scene reconstruction algorithm, ICP is used as an only method for registering two continuous frames. The method requires a powerful Graphic Card to fasten calculations and reduce runtime. However, in global registration algorithms with thousand times of error function calculation more than ICP through many iterations and populations, to make the algorithm can run real-time, we need a faster error calculation method.

The proposed algorithm takes the advantage of fast error calculating by using ray-casting based corresponding point
searching to apply for a new optimization algorithm ISADE with a purpose of getting a faster and global optimal convergence guaranty.

B. Ray-Casting Closest Point Method

ICP-class algorithms often use kd-tree [14] structure to speed up the process of finding \( j^* \) in Equation 2. The complexity of kd-tree searching closest algorithm is \( O(\log(n)) \) where \( n \) is number of searching point set. Fig. 3 shows an example of corresponding points of the data point-cloud in the model one.

Fig. 3. Kd-tree closest point in original ICP.

Since depth image or point cloud data are often obtained from 3D range camera in which the data could be consider as an 2D gray image \( G \) where value of each pixel show the depth of the point.

\[
z_{i,j} = G[i,j]
\]

Equations 4 is to convert from depth image and real 3D depth data \( \{x, y, z\} \).

\[
x_{i,j} = (i-cx)G[i,j]/f_x
\]

\[
y_{i,j} = (j-cy)G[i,j]/f_y
\]

\[
z_{i,j} = G[i,j]
\]

where \( f_x, f_y, cx, cy \) are intrinsics of the depth camera. In conversion, pixel position and structured expression of a point \( \{x, y, z\} \) can be calculated as Equation 5.

Equations 4 is to convert from depth image and real 3D depth data \( \{x, y, z\} \).

\[
z_{i,j} = G[i,j]
\]

where \( f_x, f_y, cx, cy \) are intrinsics of the depth camera. In conversion, pixel position and structured expression of a point \( \{x, y, z\} \) can be calculated as Equation 5.

\[
z_{i,j} = G[i,j]
\]

Those equations are to calculate \( i, j \) of data points which are also \( i, j \) of corresponding point in model point-clouds. The idea of the method is showed as Fig. 4, which reminds the ray-casting process in computer vision.

C. Objective Function

The fitness function need to provide an error score that is minimized when the best transformation matrix are applied. The paper uses fitness function as Equation 6.

\[
F(R, t) = f(n) \frac{1}{N^2} \sum_{i=1}^{N}(R \ast y_j + t - x_i)^2
\]

where \( f(n) \) is a function of inlier point number, \( n \). \( N \) is the number of points in the data point-cloud.

The error function should be smaller in bigger number of inlier point. Since that, searching algorithm would get rid of the case in which cost function is small for only small inlier points. Function \( f(n) \) is calculated as in Equation 7.

\[
f(n) = \begin{cases} 
\infty & \text{if } \frac{n}{N} < 0.1 \\
1 - \frac{n}{N} & \text{otherwise}
\end{cases}
\]

Instead of using ICP with iteration steps with meeting the condition of maximum iteration steps or error become smaller then a set error to flatten the cost function, the algorithm only do smothering by calculating transformation matrix to minimize error function with one step using SVD method. Equation 6 without \( F(R, t) \) can be rewrite as Equation 8 to find the one step better rotation and translation in term of cost from initial transformation matrix.

\[
F(\Delta R, \Delta t) = F(R + \Delta R, t + \Delta t)
\]

where \( R, t \) are initial rotation and translation matrix, \( \Delta R \) and \( \Delta t \) are smothering or fine matrix.

D. Translation Computing

We can find the optimal translation by taking derivative of \( F \) with respect to \( \Delta t \) and search for its roots.

\[
0 = \frac{\partial F}{\partial \Delta t} = \sum_{i=1}^{N} 2(\Delta R \ast y_j^* + \Delta t - x_i) = 2t \ast n + 2R(\sum_{i=1}^{N} y_j^*) - 2(\sum_{i=1}^{N} x_i)
\]

where \( y_j^* \) is new coordinate of \( y_j \) after rough transformation with \( R \) and \( t \).

Denote

\[
\bar{x} = (\sum_{i=1}^{N} x_i)/n \quad \text{and} \quad \bar{y} = (\sum_{i=1}^{N} y_j^*)/n
\]

The final results for translation:

\[
\Delta t = \bar{x} - \Delta R \bar{y}
\]

In other words, the translation of first movement draws two pointlouds close to each other so their weighted centroids coincide.

E. Translation Computing

Replacing \( \Delta t \) from Equation 10, \( F(\Delta R, \Delta t) \) is calculated as Equation 11.
\[ F(\Delta R, \Delta t) = \sum_{i=1}^{N} (R \ast y'_i + \Delta t - x_i)^2 \]

\[ = \sum_{i=1}^{N} (R \ast y'_i + (\bar{x} - R \bar{y}) - x_i)^2 \]

\[ = \sum_{i=1}^{N} (R \ast (y'_i - \bar{y}) - (x_i - \bar{x}))^2 \]

Denote \( x_i' = x_i - \bar{x} \) and \( y'_i = y'_i - \bar{y} \), rotation matrix is presented as Equation 12.

\[ \Delta R = \arg \min_{R} \sum_{i=1}^{N} (R \ast y'_i - x_i')^2 \]

Using SVD method for least square problem, covariance matrix is calculated as Equation 13.

\[ S = XY^T \]  

(13)

Decomposing \( S \) matrix \( S = U \sum V^T \), then rotation matrix is calculated as in Equation 14.

\[ \Delta R = V \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \det(VU^T) \end{bmatrix} U^T \]  

(14)

After having rotation matrix, translation matrix is recalculated as Equation 10.

\[ F_i = \frac{1}{1+\exp\left(-\frac{y_i}{a_i} \right)} \]

(ISADE give addition scaling \( F^\text{mean}_i \) as in Equation 10.

\[ F^\text{mean}_i = F_{\min} + (F_{\max} - F_{\min}) \left( \frac{i_{\text{iter}}}{i_{\text{max}}} \right) \]

where

\[ n_{\text{iter}} = n_{\min} + (n_{\max} - n_{\min}) \frac{i}{i_{\max}} \]

\[ F_i \] in Equation 9 is modified as in Equation 11.

\[ F_i = \frac{F_i + F_{\text{mean}}^i}{2} \]

Now scaling factor is set to be high in first iterations and after certain generations it become smaller for proper exploitation.

2) Crossover control parameter

ISADE algorithm is able to detect whether high values of \( C_r \) are useful and if a rotationally invariant crossover is required. A minimum base for \( C_r \) around its median value is incorporated to avoid stagnation around a single value. The control parameter \( C_r \) is assigned as Equation 13.

\[ C_{r_i}^{i+1} = \begin{cases} rand_2 & \text{if } rand_1 < \tau \\ C_r^i & \text{otherwise} \end{cases} \]

(13)

where \( rand_1 \) and \( rand_2 \) are random values \( \in [0,1] \). \( \tau \) presents probability to adjust \( C_r \). \( C_r \) is adjusted as in Equation 14.

\[ C_{r_i}^{i+1} = \begin{cases} C_{r_{\min}} & \text{if } C_{r_{\min}} \leq C_{r_i}^{i+1} \leq C_{r_{\text{medium}}} \\ C_{r_{\max}} & \text{if } C_{r_{\text{medium}}} \leq C_{r_i}^{i+1} \leq C_{r_{\max}} \end{cases} \]

(14)
where $C_{r_{\text{min}}}$, $C_{r_{\text{medium}}}$, $C_{r_{\text{max}}}$ denote low value, median value and high value of crossover parameter respectively. As in [12], we take $\tau = 0.1$, $C_{r_{\text{min}}} = 0.05$, $C_{r_{\text{medium}}} = 0.5$, $C_{r_{\text{max}}} = 0.95$.

All above ideas and theories are implemented as in flowchart in Fig. 5.

G. A New Combination

From initial position matrix, using one ICP iteration to gain a slightly better rotation and translation matrix, the algorithm recalculates the error as in Equation 6 and uses it in ISADE searching algorithm. Flowchart in Fig. 6 shows implementation of the whole algorithm.

IV. EXPERIMENT AND RESULTS

This section aims at presenting a number of experimental results to study how robust and accurate of ISADE results in comparison to other Global searching algorithm in using the same ray-casting based error function as well as comparison of result from new algorithm to KinectFusion in term of accuracy.

- De Falco et al.’s proposal (DE), Differential Evolution as a viable tool for satellite image registration [17].
- Valsecchi et al.’s proposal (GA), An Image Registration Approach using Genetic Algorithms [18].
- Talbi et al.’s proposal (PSO), Particle Swarm Optimization for Image Processing [19].
- Luck et al.’s proposal (SA), registration of range data using a hybrid simulated annealing and iterative closest point algorithm [20].

The proposed algorithm is implemented in C++ and compiled with GNU/g++ tool.

In order to perform a fair comparison between different optimization tools, in all methods, maximum iteration is set to 100 with population of 25 each generation. As SAICP is not a multi-agent method, its maximum iteration is set to 2500.

A. Range Image Datasets

Our experiments carried out number of pair-wise registration task using well-known Depth data taken from Kinect Microsoft Camera downloaded from website of Microsoft Research http://research.microsoft.com/en-us/projects/7-scenes/. Specifically, Fig. 7 shows all scenes: Chess, Fire, Heads, Office, Pumpkin, RedKitchen, and Stairs.

Those .PNG format depth images were sub-sampled into smaller solution of $128 \times 96$, which is 5 times smaller than original solution of $640 \times 480$ in each dimension. The reason for using smaller number of point dataset is to archive considerable suitable runtime while accuracy remains unchanged.

B. KinectFusion Error from Data Transpose

Accompany with depth datasets, 7 scenes database give us camera homogeneous transposes at each frame calculated from Kinect-Fusion algorithm. Using those transpose, we could calculate transformation matrix between two scenes as Equations 15.

$$T^i_j = T^{-1}_i \ast T_j \quad (15a)$$

$$T^i_j = \begin{bmatrix} R^i_j & t^i_j \\ 0 & 1 \end{bmatrix} \quad (15b)$$

where $T^i_j$ is transformation matrix to move frame $j$ to align with frame $i$, $T_i$ and $T_j$ are homogeneous transpose matrix for camera at frame $i$ and $j$ respectively, $R^i_j$, $t^i_j$ are rotation and translation matrix of $T^i_j$.

$R^i_j$, $t^i_j$ are applied into ray-casting error calculation methods for two frames as in Equation 6 to draw errors of KinectFusion algorithm for the next comparison step.

C. Parameter Setting

In each methods 30 runs were executed with two registration depth images are at distance of 20 frames in the sequence. The searching space is set so rotation and translation limitation at $[-2\pi/10, 2\pi/10]$ and $[-0.3, 0.3]$ separately. All methods are run on a PC of Intel core i7-4790 CPU 3.60 GHz ×8 processor and 8 GB of RAM memory.

D. Results Comparison between Algorithms

ISADE searching algorithm results are compared with other algorithms' results in three categories including convergent rate, mean and standard deviation, which are shown in Table I.
Image registration has been a very active research area. Recently, the approach of using evolutionary algorithms (EAs), especially new methods, proved their potential of tackling image registration problem based on their robustness and accuracy on searching for global optimal. With EAs algorithm as searching tools, it is not necessary to have good initials to avoid local minima and converge to near-global minima solutions. To do that, EAs algorithms need tuning carefully to gain best results.

We proposed the new registration algorithm by integrating a new self-adaptive optimization algorithm (ISADE) into a fast closest point searching method to tackle well-known challenging task of computer vision area. In the experiments, the results show that ISADE is able to find a robust and accurate transformation matrix of camera movement.

What is more important, accuracy and robustness results have been obtained in comparison with other state-of-the-art evolution based algorithms. ISADE shows its superior than GA, PSO, SA in searching for global minima solution. In comparison with DE, ISADE also show its much better in almost tested scenes. The robustness and accuracy is tested and proved in real 3D scenes captured by Microsoft Kinect camera.

In term of running time, by using fast searching closest point methods, proposed algorithms are considered fast in our sense. It shows potential of applying in real-time application if using parallel programing technique with multi-core processors.

In future work, ISADE algorithm can be implemented in parallel in GPU (Graphical Processor Unit) which can help algorithm reduces runtime to prove real-time implement possibility in 3D reconstruction, 3D mapping and 3D localization.

### REFERENCES


Tao Ngoc Linh received his B.E. in 2010 at Hanoi University of Science and Technology, Vietnam and M.E. in 2013 from Taiwan University of Science and Technology, Taiwan. He is pursuing Dr. Eng. degree at the Department of Functional Control Systems, Graduate School of Engineering and Science, Shibaura Institute of Technology, Japan. His research interests include 3D computer vision and intelligent algorithms.

Tinh Nguyen received the B.E. degrees in 2012 at Hanoi University of Science and Technology, currently, he is a second year Master student at Graduate School of Engineering, Shibaura Institute of Technology, Japan. His research interests include optimization system design, multi-body system, humanoid robot, and evolutionary algorithm.

Hiroshi Hasegawa received his B.E. in 1992 and M.E. in 1994 from Shibaura Institute of Technology, Japan. He has been working at Shibaura Institute of Technology, and currently is a Professor at the Department of Machinery and Control System, College of Systems Engineering and Science. He is a member of JSEE, JSME, ASME, JSCES, JSST and KES. His research interests include computer-aided exploration, creativity of design and systems engineering.