Research on pose point cloud matching error compensation method for confocal image assembly

Yijin Zhao, Xin Ye\textsuperscript{1}, Lei Wang, Xinhai Yu and Heng Zhang

Department of Mechanical Engineering, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing 100081, China;
\textsuperscript{1}E-mail: yexin@bit.edu.cn

Abstract. High-precision matching of spatial poses is the key to ensuring the assembly performance of complex three-dimensional devices. The human eye cannot directly judge the three-dimensional pose of mesoscale devices, and high-precision observation instruments are needed to assist accurate judgment, ensuring smooth assembly. In this paper, based on the three-dimensional assembly system of micro-devices of confocal microscope, it is difficult to realize the problem of assembly alignment of spatial poses in the existing methods, this focuses on solving the problem of how to perform high-precision pose matching by using point cloud data measured through the confocal microscope. A new method for directly performing alignment of three-dimensional poses is proposed. The principle of point cloud registration ICP algorithm is studied, and the theory of point cloud axis hole registration theory is carried out. A neural network about the mapping relationship between the angle calculated by the RT matrix obtained by ICP algorithm and the theoretical angle is constructed. The comprehensive assembly accuracy of the whole method is analyzed to be 2.41\textmu m. Finally, the actual experiment was carried out through the standard hole axis of 3\textmu m for the single-edge gap and assembled successfully.

1. Introduction

Difficulties in the whole micro-assembly system: the alignment detection, it refers to the matching of the end faces of two assembly objects. The common alignment detection method is difficult to achieve accurate assembly of the three-dimensional pose and an error in the assembly angle is generated. This paper proposes a method for spatial three-dimensional pose assembly based on point cloud data to obtain three-dimensional pose information.

At present, there are four detection alignment techniques most commonly used in micro-assembly, which are based on monocular micro-vision, based on multi-view micro-vision, and based on monocular vision multi-view detection of alignment and coaxial alignment detection methods.

Monocular microscopic visual inspection aligning method [1]: when the monocular microscopic vision method is used to detect the parts to be assembled, the optical system optical axis is generally perpendicular to the mounted mounting end surface to obtain the end face information of the target part and image processing to determine its position. It can only be used to detect the geometric position of a part.

Multi-objective microscopic visual inspection aligning method [2]: multi-objective microscopic visua inspection aligning method is traditionally using binocular, trinocular or multi-mesh microscopic vision to detect the alignment, but the positional information of the key matching features of the part cannot be obtained.
Monocular microscopic multi-view detection aligning method [3]: monocular microscopic multi-view detection aligning method is to use the monocular vision matching mirror group to simultaneously detect and align the image. Its shortcomings are similar to multi-view microscopic vision systems, but its spatial layout is slightly better than multi-view microscopic vision detection methods.

Coaxial alignment detection method [4]: in order to make up for the above-mentioned methods, the defects of the mating surface information cannot be seen for the purpose of the outer contour of the part. Beijing Institute of Technology proposed a coaxial detection method based on ordinary optical microscope. However, the defects of this method are: 1) unable to break the optical diffraction limit; 2) the accuracy of moving parts; 3) only the end face information can be detected and the coordination force and pose relationship mapping needs to be studied in order to ensure the coaxial alignment, the process is cumbersome and complicated and cannot be directly adjust the assembly of the three-dimensional pose of the micro device.

In order to solve the above problems, this paper proposes a point cloud pose error compensation method for three-dimensional micro-device confocal common image micro-assembly. The optical axis of the confocal microscope [5,6] is perpendicular to the axis of the component to be assembled. By the trapezoidal prism optical path transformation, the confocal microscope obtains the point cloud depth information of the matching end face. We use the ICP algorithm [7] commonly used in point cloud registration to register the point cloud scanned by the base part and the part to be assembled, calculating the angle to be rotated and the distance of translation from the obtained RT matrix.

This paper first introduces the composition and experimental principle of confocal image alignment micro-assembly system based on the laser confocal microscope. Secondly, the point cloud filtering algorithm, ICP algorithm, and ICP algorithm simulation experiment based point cloud registration are introduced. The generated ideal point cloud is used for registration, and the rotation angle translational displacement is calculated from the RT matrix obtained after registration. The neural network is then trained to derive the mapping between the angle calculated by the RT matrix obtained by the ICP algorithm and the true angle. Finally, the standard axial hole sample is used to carry out the axial hole three-dimensional pose assembly experiment based on the coaxial alignment system using laser confocal microscope to prove the correctness and feasibility of the method.

2. Overview of the confocal common image assembly

2.1. Confocal common image alignment assembly basic principle and assembly system

Laser confocal microscopes are generally used to measure the topographical features of a three-dimensional or quasi-three-dimensional object. Using the high-precision depth measurement capability of confocal microscopy, we designed and built a coaxial alignment micro-based laser confocal scanning microscope assembly system. The main idea is also derived from the coaxial alignment detection method, which the detection components only have changed from the original "microscope + CCD camera" to a laser confocal microscope. The laser confocal microscope used in this paper is Olympus LEXT OLS4100 with a planar resolution of 0.12μm and a longitudinal accuracy of 10nm. The OLS4100 uses a 405nm short-wavelength semiconductor laser with excellent planar resolution. For the high precision and low working distance of the laser confocal microscope, the alignment light path is designed by using the trapezoidal prism. The specific alignment principle is as follows:

First, as shown in Figure 1, the target part A1, A2 and the base part B1, B2 are placed on the opposite sides of the trapezoidal reflective prism. Since the two bottom angles of the trapezoidal prism are both 45°, the imaging light of the target part and the base part. The process can be kept equal, and after the oblique reflection of the trapezoidal prism, a clear image can be obtained on the image plane of the confocal scanning microscope. Then we use the precision module on the gantry to drive the confocal microscope head horizontally to the other side, moving a certain distance D to record the
relative positional relationship between the target part and the base part. Finally, the point cloud obtained by scanning the base part and the target part by laser confocal microscope is used for registration. The calculated RT matrix is used to obtain the pose relationship, and the actuator is started to complete the assembly process.

The three-dimensional micro-device confocal assembly micro-assembly system is mainly divided into three parts: a laser confocal scanning microscope part, a micro-motion stage part and a base support part, as shown in Figure 2 and Figure 3.

2.2. Assembly process

The main steps of the error compensation method for point cloud matching proposed in this paper can be divided into three main stages: laser confocal microscopy inspection, pose adjustment and assembly execution. First, the system is initialized and each module is reset; the base part is clamped on the 1# console and the target part is clamped on the 2# console. The confocal component is placed perpendicular to the 1# console and the axis of the console on the 2# vertical plane. Then we use a laser confocal scanning microscope to separately scan the three-dimensional point cloud for the collective part and target part. Using the 3D point cloud, the RT matrix of the collective part and the target part is calculated. Finally, the assemble the actuator will move and adjust the pose to complete the part assembly.

The overall experimental process of the pose-point cloud matching error compensation method for confocal image assembly proposed in this paper is shown in Figure 4:
3. **Point cloud error compensation principle and registration simulation experiment**

### 3.1. **Point cloud matching error compensation**

The main sources of error in confocal common image micro-assembly systems are: the assembly error, the motion error and the measurement error. The specific analysis is shown in Figure 5 below:

![Figure 5. Error source analysis.](image)

Pose point cloud matching error compensation method for confocal common image assembly proposed in this paper, because it mainly uses the three-dimensional data point cloud of the part and does not involve the alignment detection optical path error and image matching error, and neither does not need to consider the laser action. There is no need to consider the problems of scattering, diffraction, etc. caused by the laser acting on the edge of the part.

### 3.2. **Point cloud registration and ICP algorithm**

When a three-dimensional scanning device is used to scan an object to be measured, a three-dimensional discrete point set of the object can be easily obtained. The set of discrete points is a so-called point cloud data. When the point cloud is actually acquired, due to the influence of equipment accuracy and environment, some noise will inevitably appear in the point cloud data. Therefore, the filtering process in the point cloud processing process is an important step in the preprocessing, which has a great influence on the subsequent registration effect. Because the registration function is realized
in this paper, the edge requirements are higher. Meanwhile, the bilateral filtering has better effect of “guaranteeing the edge”. Therefore, the bilateral filtering is selected as the pre-registration.

Based on the predecessors, Besl and Mckay [8] proposed the exact registration ICP algorithm in 1992, which is the iterative nearest point algorithm. The registration of point cloud data is to solve the point set in two different coordinate systems, that is, the rigid body transformation matrix between the reference point set and the target point set. That is, knowing the point sets $P$ and $Q$ of the same object that are not in the same coordinate system, find the rigid body transformation between the two point sets, so that the distance between the two points corresponding to the pair of points and the minimum:

$$ F(R, T) = \sum_{i=1}^{N} ||Q_i - (RP_i + T)||^2 $$  \hspace{1cm} (1)

In the above formula, $P_i$ is the initial point set of the reference data, concentrating the nearest point $Q_i$ for the target data point, $R$ is a $3 \times 3$ Rotation matrix, $T$ is a $3 \times 1$ translation vector, $R$ and $T$ together constitute the entire rigid body transformation. $F(R, T)$ represents the sum of squared distances of corresponding points between two sets of points after the set of reference points has been rotated and translated. Before the micro-assembly, the assembly and the accessories to be assembled will be pre-assembled, so there will be no large offset. According to the actual application, the angular deviation is within $2^\circ$. Therefore, the focus of this paper is on accurate registration and point cloud registration simulation experiments within $2^\circ$ deviation.

3.3. Point cloud registration simulation experiment based on micro-assembly

We first study the rotation matrix obtained by the ICP algorithm, $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$. The following is the basic rotation matrix of the three-axis around the right-handed system. The rotation angles are respectively $\theta, \alpha, \beta$.

$$ R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} $$  \hspace{1cm} (2)

$$ R_y(\alpha) = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix} $$  \hspace{1cm} (3)

$$ R_z(\beta) = \begin{bmatrix} \cos \beta & -\sin \beta & 0 \\ \sin \beta & \cos \beta & 0 \\ 0 & 0 & 1 \end{bmatrix} $$  \hspace{1cm} (4)

The angle solution formula for the rotation matrix are:

$$ \theta = a \tan 2(r_{32}, r_{33}) $$  \hspace{1cm} (5)

$$ \alpha = a \tan 2(-r_{31}, \sqrt{r_{32}^2 + r_{33}^2}) $$  \hspace{1cm} (6)

$$ \beta = a \tan 2(r_{21}, r_{11}) $$  \hspace{1cm} (7)

And the rotation sequence is first to rotate the angle $\theta$ around the X axis, then to rotate the angle $\alpha$ around the Y axis, and finally to rotate the angle $\beta$ around the Z axis.

This paper studies the point cloud registration with an angular deviation of less than $2^\circ$. Due to the limitations of existing processing conditions and the design of standard shaft hole sizes, first, we use
the software Pro/E to draw a hole with an outer diameter of 6mm, an inner diameter of 1.5mm and a length of 2mm in its standard coordinate system, and save it as a point cloud with a point cloud of 100,000 points. Here, the number of point clouds is the result of experiments. The number of point clouds is 10,000, 100,000 and 1 million respectively. The experimental results show that the number of point clouds with a quantity of 10,000 is lower, the quantity is The precision of 100,000 point clouds is better and the matching speed is faster. The accuracy of the point cloud with a quantity of 1 million is not greatly improved, but the running speed is very slow. Therefore, the number of point clouds selected in this paper is 100,000. Then, we translate the center of the circle along the Z axis by 0.2mm, 0.4mm, and so on up to 2mm, drawing the axis with a diameter of 1.5mm and a length of 2mm, respectively, and using the above method to generate a point cloud of 100,000 points. Through the registration to obtain the transformation matrix, this paper gives the effect legend Figure 6 and RT matrix before and after the registration of 1mm translation. We sort out the real translation vector Z' and the translation vector Z obtained by the algorithm, and conclude that the translation vector calculated by the algorithm has higher precision and the error compensation is 2.6μm.

![Figure 6. Registration before and after rendering.](image)

\[
RT = \begin{bmatrix}
0.999991 & -0.000747948 & -0.0040643 & 0.00240126 \\
0.000739738 & 0.999998 & -0.00202143 & -0.00239826 \\
-0.0040658 & 0.00201841 & 0.99999 & -0.997576 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]

Then we rotate the original coordinate system, rotate the X, Y, Z axis by 2°, 1.98°, 1.96°, and so on, up to 0.02°, and draw the same size axis. The parameters are adjusted and registered with the previous registration method in the basic coordinate system to obtain the RT matrix. Since the shaft holes are all revolving bodies, it is only necessary to study the angle of rotation around the X and Y axes. After finishing the experiment, the experimental results are obtained. It can be seen that the accuracy of the rotation angle calculated by the RT matrix is not as good as the calculated translational displacement accuracy and there is a certain mapping relationship between the standard angle and the calculated angle.

### 3.4. Building a neural network

A neural network is a dynamic system with a directed graph as a topology. Feedforward neural network is the most common kind of network. Information is transmitted step by step. Each layer of neurons receives the input of the previous layer and continues to pass to the next layer. It can approximate any continuous function and square integrable function with arbitrary precision. Since the amount of data in this paper is not large and the mapping relationship is relatively simple, it is chosen to build a feedforward neural network. This article is based on Python and uses the Keras framework to build a feedforward neural network about the rotation angle around the X and Y axes.

The input is the rotation angle around the X and Y axes calculated by the RT matrix, and the output is the actual angle of rotation around the X and Y axes. Through the simulation experiment in Section 3.3, this paper uses Pro/E modeling to draw the base model and the target model of the known pose relationship, and uses the ICP algorithm to register. The corresponding RT matrix is obtained, the corresponding angle is calculated, and the training set is obtained by the above method. The selection of the number of hidden layer nodes is also a key issue affecting the performance of the network. At
this stage, there is no exact theoretical basis for the selection of the number of neurons in the hidden layer. Usually we use the trial and error method, according to the empirical formula:

$$Q = \sqrt{M + L + \alpha}$$  \hspace{1cm} (8)

In the above formula, M is the number of input layer neurons. Q is the number of neurons in the hidden layer. L is the number of neurons in the output layer. \( \alpha \) is a constant between one and ten.

Combined with the actual situation, the implicit layer excitation function is selected as the ReLU function:

$$f(x) = \max(0, x)$$  \hspace{1cm} (9)

The maximum number of training steps in the neural network set in this paper is 1000 steps. The loss function chosen in this paper is the mean square error (MSE). This loss function is commonly used in the regression problem of continuous variables in the real-value domain as follows:

$$MSE = \frac{1}{N} \sum (y_i - p_i)^2$$  \hspace{1cm} (10)

Finally, this paper trains a neural network with a number of single hidden layer neurons of 5.

4. Three-dimensional pose assembly experiment and result evaluation

Assembly accuracy throughout the assembly process depends on the accuracy of the inspection, the accuracy of the algorithm, and the accuracy of the execution system. The detection accuracy is the detection precision of the laser copolymer scanning focus microscope of 0.12\( \mu \)m, and the movement accuracy of the linear stage (PI stage) can be reached at 0.2\( \mu \)m. Therefore, we must focus on the accuracy of the entire algorithm. From the simulation experiment, the worst translation accuracy is 0.09\( \mu \)m. The 162 \( \mu \)rad rotation angle accuracy of the method is obtained by the neural network of the present invention and the assembly accuracy obtained by the actual range can be reached at 2.0\( \mu \)m. As can be seen from the above, the worst assembly accuracy of the method is 2.41\( \mu \)m. Therefore, this paper uses a standard axial hole with a single edge clearance of 3\( \mu \)m for actual verification experiments.

The three-dimensional axial hole solid figure obtained by scanning with a laser confocal microscope is shown in Figure 7 and Figure 8.

The test piece is a hole having an outer diameter of 6 mm, an inner diameter of 1.5 mm, a length of 2 mm, and a shaft having a diameter of 1.5 mm and a length of 2 mm. In this paper, three-dimensional point cloud data is acquired by confocal scanning microscopy, and pre-processing such as downsampling and bilateral filtering are performed. First, pre-assembly is performed by manually adjusting each platform of the system, and the initial RT matrix is obtained after registration:

$$RT = \begin{bmatrix}
0.999998 & -0.000548397 & 0.00179121 & -0.0471839 \\
0.000552598 & 0.999997 & -0.00234552 & 0.0411718 \\
-0.00178992 & 0.00234651 & 0.999995 & -0.54711 \\
0 & 0 & 0 & 1
\end{bmatrix}$$

From the calculation of the above RT matrix, the following calculation results are obtained, as shown in Figure 9. At this point start the actuator to perform the shaft hole assembly. The effect after
assembly is shown in Figure 10, which shows that the method can achieve the 3μm comprehensive assembly accuracy of the single-edge gap.

![Figure 9. Experimental result.](image)

![Figure 10. Shaft hole assembly.](image)

5. Conclusions

Based on the experimental principle of point cloud alignment assembly, this paper establishes a point cloud alignment micro-assembly system based on laser confocal microscope and describes the composition of the system. The method avoids the complicated and complicated process of studying the mapping of assembly force and pose relationship, and can directly perform assembly alignment of the three-dimensional pose of the micro device. Through the simulation of the axial hole registration of the point cloud, we built a neural network between the calculated angle and the theoretical angle, and analyzed the comprehensive assembly accuracy of the whole experiment which can reach at 2.41μm. Finally, we carried out the actual experiment and assembled successfully through the assembly experiment of the actual standard shaft hole through the standard shaft hole of the 3μm single side gap. The method can greatly improve the assembly efficiency of the micro assembly. Compared with the previous method of coaxial alignment, which only applies the two-dimensional end face information of the base part and the target part. The spatial three-dimensional pose of the base part and the target part is assembled. The assembly quality is improved, and the assembly process is also improved.

Acknowledgement

This work was financially supported by the National Natural Science Foundation of China (51575052).

References

[1] Janabi-Sharifi F, Deng L, Wilson W J 2010 Comparison of basic visual servoing methods IEEE/ASME Transactions on Mechatronics 16(5): 967-983
[2] Yun Y, Li Y 2011 Robotics and Computer-Integrated Manufacturing 27(6): 977-985
[3] Zhou Y 2017 Physics Reports 723: 1-160
[4] Ye X, Zhang ZH J, Jin X, Wang Q 2012 High precision visual coaxial optical alignment assembly system for micro-miniature structural parts China ZL201110007968
[5] Zurong Qiu, Ting Wang, Jinjiang Wang, Xinhua Li, Ying Zhong 2009 J.Vac.Sci.Technol.B 27(4): 1890–1896
[6] Zhang Z, Zhang J, Li H, et al 2012 2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE). IEEE 1: 471-475
[7] Rusinkiewicz S, Levoy M 2002 Efficient Variants of the ICP Algorithm International Conference on 3-d Digital Imaging & Modeling
[8] Besl P J, Mckay N D 2002 IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2):239-256