

Path Planning with Automatic Seam Extraction over Point Cloud Models for Robotic Arc Welding

Peng Zhou, Rui Peng, Maggie Xu, Victor Wu and David Navarro-Alarcon

Abstract—This paper presents a point cloud based robotic system for arc welding. Using hand gesture controls, the system scans partial point cloud views of workpiece and reconstructs them into a complete 3D model by a linear iterative closest point algorithm. Then, a bilateral filter is extended to denoise the workpiece model and preserve important geometrical information. To extract the welding seam from the model, a novel intensity-based algorithm is proposed that detects edge points and generates a smooth 6-DOF welding path. The methods are tested on multiple workpieces with different joint types and poses. Experimental results prove the robustness and efficiency of this robotic system on automatic path planning for welding applications.

I. INTRODUCTION

Industrial welding robots have been extensively studied with the aim to solve labor shortage problem and improve productivity. Currently, “teach and playback” mode or offline programming prevails in robotic welding workshops. As these conventional practices lack adaptation to accommodate changes in a working environment, many advanced techniques such as seam detection and tracking are gradually being employed in industry.

Solutions to intelligent robotic arc welding primarily require the assistance of computer vision. Previous studies have conducted in-depth research on image processing and established viable methods across several visual sensors, including mono camera, stereo camera, and CCD camera [1]–[6]. [7] developed an automatic image acquisition system for real-time weld pool analysis and seam tracking. The seam characteristics were measured by defining region of interest and implementing pattern learning algorithm. [8] proposed a shape-matching algorithm for an online laser-based seam detection. The algorithm enabled autonomous detection for different groove types and localizes the boundary of weld seam. [9] measured the cross-sectional areas of 2D profiles and finds joint features and volume of the workpiece.

While above mentioned studies can effectively locate the seam, their implementations are only validated on simple or flat welding workpieces with common joint types. In most cases, the groove problem is regarded as two-dimensional and image processing is sufficient to identify the target. However, it is important to discuss the generalization of these techniques to cope with complex geometries. The remaining challenge lies in capturing free-form grooves. Point cloud

processing is favored as it can comprehensively describe complex 3D shapes. A captured point cloud model can reveal useful geometric information but simultaneously introduce considerable noise. The complexity and computational cost also evolve as the information increases. Therefore, designing a robust algorithm that can efficiently utilise 3D information will be crucial for this application. Although, some recent studies using point cloud processing were able to detect the groove, they did not further exploit the features to improve system functionality, such as plan a welding path with well-defined orientation [10]–[14]. More experiments on complex and diverse seam patterns are needed for verification. In addition, the authors found that most systems need to manually set or program the initial poses for visual sensors to capture the workpiece. This guidance of scanning may be realized in a more intuitive way to facilitate human machine collaboration.

To address the above problems, we propose a novel point cloud based system for robotic welding planning. This system integrates a single RGB-D camera with a cooperative industrial manipulator to fulfill the generic requirements for seam detection. Guided by hand gestures, the system captures a set of partial point clouds of the workpiece from different perspectives. To reconstruct the surface, it adopts a linear iterative closest point algorithm that improves the rate of fitness and extends bilateral filter algorithm for 3D surface denoising. A novel algorithm is implemented to identify the groove based on edge density. After the localization, the system plans a smooth consecutive welding path with 6-DOF and executes it with the robot manipulator. To validate the proposed methods, we conduct detailed experiments on workpieces with different joint types.

The remaining parts of this article are as follows. Section II introduces the architecture of the proposed planning system and the key techniques applied. The experimental results of the proposed system are presented in Section III. Section IV gives final conclusions.

II. METHODOLOGY

A. System Overview

The robotic system architecture is depicted in Fig. 1. The complete surface model reconstruction for the target welding workpiece, which a hand gesture-based method, is designed to interact with the manipulator with assistance from a human operator. Subsequently, the depth camera, which is mounted on the end-effector of the robot, moves to several poses to capture different views of the point cloud. Then, the point cloud model of the workpiece is reconstructed by combining

All authors are with The Hong Kong Polytechnic University, Kowloon, Hong Kong. Corresponding author e-mail: dna@ieee.org

This work is supported in part by the Research Grants Council of HK (grant 14203917) and in part by the Chinese National Engineering Research Centre for Steel Construction Hong Kong Branch (grant BBV8).

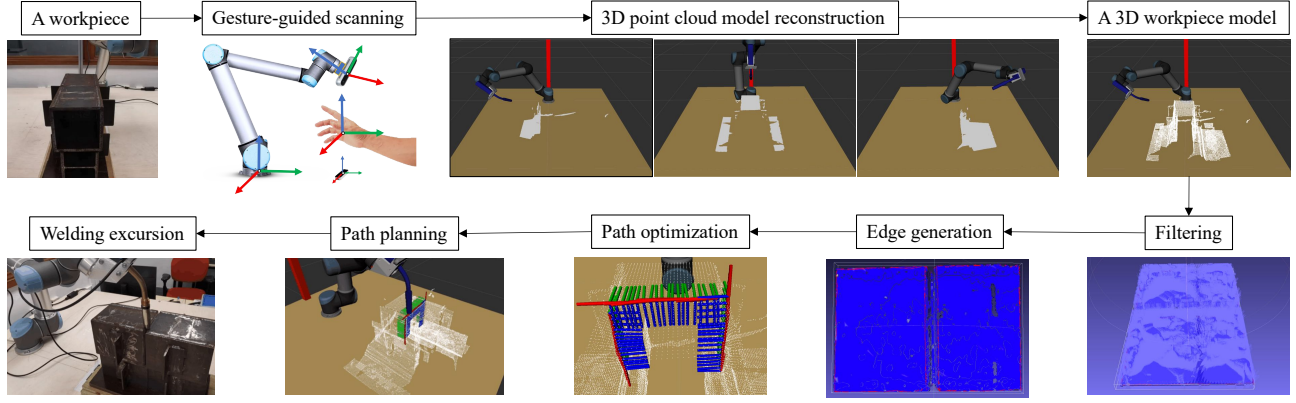


Fig. 1. Overview of the automatic robotic welding pipeline. Given a workpiece, a gesture-guided scanning is executed to merge captured point clouds into a completed 3D model, which is further improved with the bilateral filter. Subsequently, on basis of the edge points detected using an edge intensity gradient-based algorithm, a 6-DOF welding path is established.

all the individual frames of the point cloud for the workpiece. On the basis of the complete surface model, the gradient of the edge intensity-based welding seam detection algorithm is employed to locate the targeted weld path points. With these generated path points, an interpolation of both translation and rotation can be used to generate a 6-DOF welding path, which can then be sent to the robot manipulator. Finally, the robot executes the path tracking motion to finish the welding tasks on the targeted workpiece.

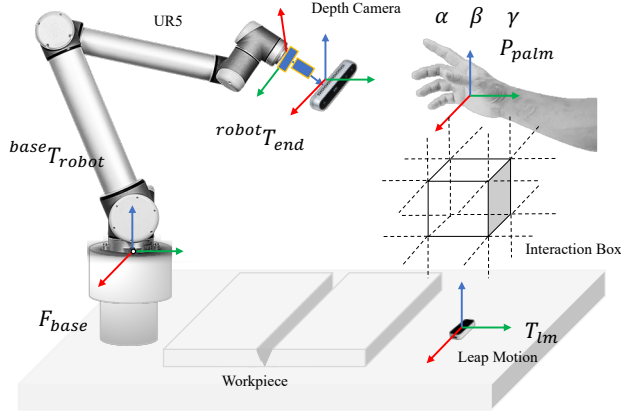


Fig. 2. Conceptual representation of the intuitively interactive scanning with hand gestures, in which an interaction box is built to define different mapping rules for robotic motions.

B. Hand Gesture Guided Scanning

To scan the welding workpiece, hand gestures are captured for generating control commands to move the camera. As shown in Fig. 2, we used a real-world setup to best represent the commonly used cases in this research. A “joystick” interaction mode is designed to interact with a robot, on which a depth camera is mounted. This interaction process is handled by the virtual interaction box above the hand gesture sensor. The moving direction of the scanning depth camera is determined according to the left hand’s relative position

with respect to the interaction box, and the rotation of the camera is mapped with the raw, pitch and yaw of left-hand gesture captured with leap motion. Meanwhile, the gesture of the right hand is used to control the point cloud capture and the pose of the robot end-effector at the current poses for the succeeding reconstruction process.

Algorithm 1: Hand Gesture Interaction

```

1 if True then
2    $(P_t^{left}, \alpha, \beta, \gamma) = \text{getPalmAttribute}(\mathcal{F}_t^{hand})$ 
3    $\mathbf{p}_t = (P_t^c, r_x, r_y, r_z) = \text{getPose}(t)$ 
4    $r_t = \text{getRegion}(P_t^{left})$ 
5   if  $r_t$  not in home region then
6      $P_{t+1}^{camera} = \mathbf{u}_d \cdot l$ 
7      $\mathbf{p}_{t+1}^{camera} = (P_{t+1}^{camera}, r_x, r_y, r_z)$ 
8     return pose  $\mathbf{p}_{t+1}^{camera}$ 
9   end
10  else
11     ${}^{lm}\mathbf{R}_{hand} = \text{Transform}(\alpha, \beta, \gamma)$ 
12     ${}^{base}\mathbf{R}_{hand} = {}^{base}\mathbf{R}_{lm} \cdot {}^{lm}\mathbf{R}_{hand}$ 
13     $(r_x^{t+1}, r_y^{t+1}, r_z^{t+1}) =$ 
14       $\text{InvRodrigues}({}^{base}\mathbf{R}_{hand})$ 
15     $\mathbf{p}_{t+1}^{camera} = (P_t^{camera}, r_x^{t+1}, r_y^{t+1}, r_z^{t+1})$ 
16    return  $\mathbf{p}_{t+1}^{camera}$ 
17  end
```

We present the entire interaction process in Alg. 1. In the frame of the UR manipulator, a standard pose is represented with $\mathbf{p} = (x, y, z, r_x, r_y, r_z)$. The algorithm takes the position of the left hand’s palm position P_{palm} , the hand’s Euler angles (α, β, γ) , and the current pose with respect to the camera ${}^{base}\mathbf{p}_{camera}^o$ as the inputs. The resulting desired output is target pose ${}^{base}\mathbf{p}_{camera}^*$ in the next time step. The entire interaction is in a loop of an infinite frame set $\{\mathcal{F}_i, i = t\}$ with respect to time step t captured by leap motion sensor. In each frame \mathcal{F}_i , the left hand’s palm position

P_{palm} and its hand's Euler angles (α, β, γ) are mapped into motions of the robotic manipulator based on the region in which the hand stays with respect to the interaction box above the sensor. According to the relative position with respect to the interaction box, the entire interaction space is separated into seven regions, namely, home, left, right, up, down, inward, and outward regions. When P_{palm} reaches other regions except the home region, a corresponding unit vector \mathbf{u}_d is computed to update the position part of the target pose ${}^{base}\mathbf{p}_{camera}^*$. However, if P_{palm} stays in the home region, the manipulator switches into rotation mode, then the Euler angles captured by the gesture sensor are to be transformed into rotation matrix ${}^{LM}\mathbf{R}_{hand}$, then this rotation matrix will be left multiplied by a rotation matrix from the gesture sensor to the robot base coordinate system ${}^{base}\mathbf{R}_{LM}$. With the resulting rotation matrix ${}^{base}\mathbf{R}_{hand}$, the inverse *Rodrigues* formula is used to transform it into the corresponding rotation part (r_x, r_y, r_z) for the target pose. Finally, the orientation part of the target pose ${}^{base}\mathbf{p}_{camera}^*$ is updated. During the process, the right hand's gesture is used to store the point cloud data \mathcal{P}_i and the corresponding pose of robot end-effector ${}^{base}\mathbf{p}_i$ corresponding to the base frame when a certain "ok" gesture is detected.

C. Welding Workpiece Reconstruction

In performing a high precision reconstruction for the welding workpieces, an initial reconstruction is executed on the basis of the scanned point cloud set, $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n\}$, and their calculated camera pose, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$, and then a linear iterative Closest point (ICP) algorithm is employed to further improve the fitness performance of the workpiece model. For each point cloud \mathcal{P} , the transformation from camera coordinate frame with respect to base the frame, ${}^b\mathbf{F}_p$, is given by:

$${}^b\mathbf{F}_p = {}^b\mathbf{M}_e \cdot {}^e\mathbf{M}_c \cdot {}^c\mathbf{F}_p \quad (1)$$

where ${}^e\mathbf{M}_c$ is a fixed rigid transformation obtained by calibrating the relative position of the robot end-effector with respect to the origin of the RGB-D camera by means of reverse engineering, and ${}^b\mathbf{M}_e$ can be easily computed using the inverse kinematics module. To simplify, the origin of the global coordinate system is set at the base of the robot.

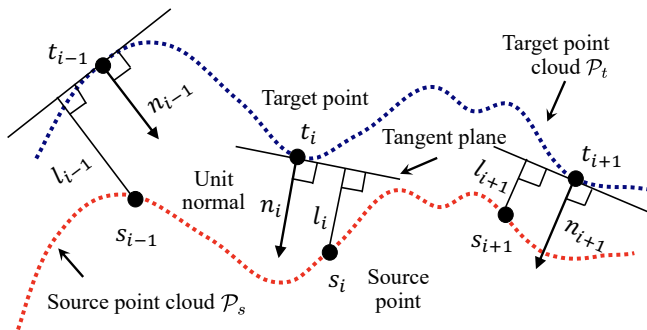


Fig. 3. Conceptual representation of the ICP process from a source point cloud \mathcal{P}_s to a target point cloud \mathcal{P}_t .

By performing the transformations illustrated in Eq. (1) on the dataset $\{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_n\}$, the entire point cloud data has been set in the base coordinate system, then they can be combined as a completed point cloud $\{{}^b\mathcal{P}_1 \cup {}^b\mathcal{P}_2 \cup \dots \cup {}^b\mathcal{P}_n\}$ for the welding workpiece. In further improving the fitness, an ICP process setting ${}^b\mathbf{M}_c$ as its initial configuration is needed. Specifically, the i -th source point is defined as $\mathbf{s}_i = (s_{ix}, s_{iy}, s_{iz}, 1)^T$, its corresponding target point is represented by $\mathbf{t}_i = (t_{ix}, t_{iy}, t_{iz}, 1)^T$, and the unit normal vector at \mathbf{t}_i can be denoted as $\mathbf{n}_i = (n_{ix}, n_{iy}, n_{iz}, 0)^T$. The goal of each ICP iteration is to obtain \mathbf{M}_* , that is:

$$\mathbf{M}_* = \arg \min_{\mathbf{M}} \sum_i ((\mathbf{M}_o \cdot \mathbf{s}_i - \mathbf{t}_i) \bullet \mathbf{n}_i)^2 \quad (2)$$

where \mathbf{M}_* and \mathbf{M}_o are 4×4 3D target and initial transformation matrices. In its first iteration, ${}^b\mathbf{M}_c$ is set as an initial transform matrix to accelerate the convergence. Usually, a rotation matrix $\mathbf{R}(\alpha, \beta, \gamma)$ and a translation matrix $\mathbf{T}(p_x, p_y, p_z)$ can be decomposed into a 3D transformation \mathbf{M} , which can be denoted as:

$$\mathbf{M} = \mathbf{T}(p_x, p_y, p_z) \cdot \mathbf{R}(\alpha, \beta, \gamma) \quad (3)$$

Eq. 2 is a essentially a least-square optimization problem whose solution normally requires the determination of the six parameters of $\alpha, \beta, \gamma, p_x, p_y,$ and p_z . However, in the rotation matrix \mathbf{R} , three of them (α, β and γ) correspond to the arguments of nonlinear trigonometric functions. Consequently, we cannot employ an efficient linear least-square technique to solve this problem. Nonetheless, the initial iteration ${}^b\mathbf{M}_c$ can ensure that three angles are $\alpha, \beta, \gamma \approx 0$. Therefore, the approximations $\sin \theta \approx \theta$ and $\cos \theta \approx 1$ can change the original nonlinear least-square problem into a linear one. In this manner, a linear least-square algorithm [15] can still be carried out as follows:

$$\hat{\mathbf{R}}(\alpha, \beta, \gamma) \approx \begin{pmatrix} 1 & -\gamma & \beta & 0 \\ \gamma & 1 & -\alpha & 0 \\ -\beta & \alpha & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

$$\hat{\mathbf{M}} = \mathbf{T}(p_x, p_y, p_z) \cdot \hat{\mathbf{R}}(\alpha, \beta, \gamma) \quad (5)$$

$$\hat{\mathbf{M}}_* = \arg \min_{\hat{\mathbf{M}}} \sum_i \left((\hat{\mathbf{M}} \cdot \mathbf{s}_i - \mathbf{t}_i) \bullet \mathbf{n}_i \right)^2 \quad (6)$$

where $(\hat{\mathbf{M}} \cdot \mathbf{s}_i - \mathbf{t}_i)$ can be written as a linear expression with six parameters ($\alpha, \beta, \gamma, p_x, p_y,$ and p_z). Then, N pairs of point correspondences can be arranged as a linear matrix expression as follows:

$$\min_{\hat{\mathbf{M}}} \sum_i \left((\hat{\mathbf{M}} \cdot \mathbf{s}_i - \mathbf{t}_i) \bullet \mathbf{n}_i \right)^2 = \min_{\mathbf{x}} |\mathbf{A}\mathbf{x} - \mathbf{b}|^2 \quad (7)$$

Thus, $\hat{\mathbf{M}}_*$ can be obtained by solving:

$$\mathbf{x}_* = \arg \min_{\mathbf{x}} |\mathbf{A}\mathbf{x} - \mathbf{b}|^2 \quad (8)$$

which can be further solved using singular value decomposition (SVD).

D. Point Cloud Denoising

In removing noise from the point cloud data without removing vital edge information in our succeeding procedure, the bilateral filter specific for point cloud data is considered. Normally, the bilateral filter is well defined for 2D pixels. Thus, we design and extend a similar version for point cloud data. Specifically, for a given point $p \in \mathcal{P}$, a unit normal vector \mathbf{n}_p is estimated for each point p via its neighbors $\mathcal{N}_r(p) = \{q \in \mathcal{P} \mid \|q - p\|_2 < r\}$. A regression plane is fitted by the neighbor point sets, then the normal of the plane is regarded the normal of the query point. The normal computation function is established as:

$$\mathbf{n}_p = \arg \min_{\mathbf{n}} \sum_{i=1}^n ((q_i - p_c)^T \mathbf{n})^2 \quad (9)$$

where \mathbf{n}_p is the targeted normal, p_c is the gravity center point of the current point's neighbor, and q_i is the i -th neighbor point within the neighbor. By computing the mean and covariance matrix of the neighbors, this least-square regression plane can be estimated.

With the estimated normal \mathbf{n}_p , $p' = p + \Delta p \cdot \mathbf{n}_p$ is used to update the point position to p' , and Δp is a normal distance weight defined as:

$$\Delta p = \frac{\sum_{q \in \mathcal{N}_r(p)} w_m(\|q - p\|) w_n(\langle \mathbf{n}_p, q - p \rangle) \langle \mathbf{n}_p, q - p \rangle}{\sum_{q \in \mathcal{N}_r(p)} w_m(\|q - p\|) w_n(\langle \mathbf{n}_p, q - p \rangle)} \quad (10)$$

where w_m and w_n are the weights generated by Gaussian distributions with variances σ_m and σ_n , respectively. Fig. 4 shows the conceptual representation of the process of filtering. In this scheme, the weights on the distances will be balanced by the weights of the normal distances, thus favoring the nearby points close to the tangent plane where the points lie on the same side of the edge around query point p . However, due to the uneven density of point surfaces caused by the process of ICP, a voxel downsampling is necessary to execute with a fixed voxel size given by multiple of the average distance of the point cloud.

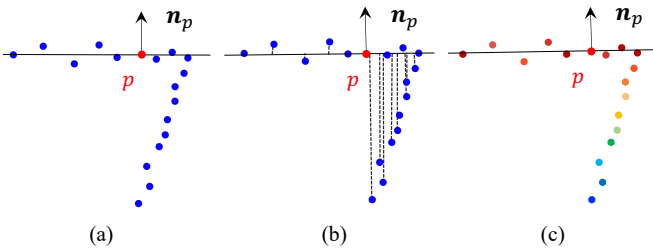


Fig. 4. Contributions of neighbors to the denoised point p in the bilateral filtering. (a) Regression plane estimation. (b) Distance weights. (c) Normal distance weights. (Blue: small weights; red: large weights)

E. Welding Seam Extraction

In extracting the welding seam for various workpieces, a generic seam detection algorithm is proposed on basis of the gradients of the edge intensity. As the point density of a reconstructed workpiece model varies, few points of a

neighborhood exist within a fixed range. The establishment of a minimum point m will provide a rational neighborhood to reduce the effect on points density. Therefore, for a given point $p \in \mathcal{P}$, its neighboring point q can be defined as:

$$\mathcal{N}_r(p) = \{q \in \mathcal{P} \mid \|q - p\|_2 < r \cup |\mathcal{N}| \geq m\} \quad (11)$$

where r is the fixed radius, which can adaptively adjust to guarantee that no less than m points will exist in the neighborhood of p . As shown in Fig. 5, the normalized displacement between original point p_o and its local geometric centroid c_o can be defined as the edge intensity of p_o , which is calculated by:

$$I_o = \frac{1}{r} \left\| p_o - \frac{1}{n} \sum_{i=0}^n q_i \right\| \quad (12)$$

where $\frac{1}{n} \sum_{i=0}^n q_i$ denotes the geometric centroid for all the neighboring points corresponding to original point p_o . Within the edge-free areas, the intensity remains low, but the value will increase when approaching edges and reaches a local maximum in the intensity space at edges. Fig. 5 shows a side view of two intersecting point cloud surfaces as an example, in which the edge intensity I_o of edge point p_o is much higher compared with the intensity I_t of its neighbor p_t in non-edge region, as the distance between the local geometric centroid c_o and p_o is larger than the distance for p_t .

For 2D images, the gradients typically represent the direction in which the features of a pixel mostly change. The way of defining gradients for 3D point clouds can represent the variations between edge intensities within its neighborhood. However, unlike the gradient decomposition in consecutive 2D pixels, gradients for a 3D point cloud can only be approximated due to the discrete points. Thus, we define the 3D gradient decompositions for a point p_o as:

$$I_k \approx \max_{i \in \mathcal{N}} \left[(I_i - I_o) \cdot \frac{d_{i,o}}{\delta k} \right], k = x, y, z \quad (13)$$

where δx , δy , and δz are the coordinate differences in three axes, and $d_{i,o}$ represents the distance between a current point p_o and its i -th neighboring point q_i . Fig. 5 shows an example of the point's gradient along Y axis.

A 3D offset function $F(x, y, z)$ for point clouds analogous with the decomposition model for a small offset between pixels is presented and approximated by a Taylor

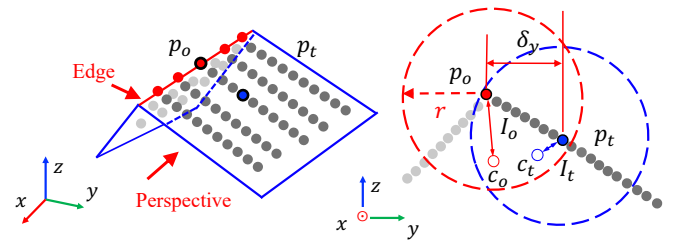


Fig. 5. Conceptual representation of the edge intensity and gradient along the Y-axis for current point p_o , in which edge intensity I_o with respect to p_o is higher than I_t with respect to its neighbor p_t . Gradient $I_{o,y}$ is defined as the difference between I_o and I_t divided by the coordinate difference along the Y-axis.

series expansion. For a given point p_o , the offset function $F(x_o, y_o, z_o)$ can be expressed as a multiplication of a small offset $(\delta x, \delta y, \delta z)$ and a symmetric matrix \mathbf{H} , in which I_x , I_y , and I_z can be calculated by:

$$\begin{aligned} F(x_o, y_o, z_o) &= \sum_{\delta x, \delta y, \delta z} [I_{x_o+\delta x, y_o+\delta y, z_o+\delta z} - I_{x, y, z}]^2 \\ &= \sum_{\delta x, \delta y, \delta z} \left[\begin{array}{c} \delta x \cdot I_x + \delta y \cdot I_y + \delta z \cdot I_z \\ + O(\delta x^2, \delta y^2, \delta z^2) \end{array} \right]^2 \\ &\approx (\delta x, \delta y, \delta z) \mathbf{H} (\delta x, \delta y, \delta z)^\top \end{aligned} \quad (14)$$

$$\mathbf{H} = \begin{bmatrix} I_x^2 & I_x I_y & I_x I_z \\ I_y I_x & I_y^2 & I_y I_z \\ I_z I_x & I_z I_y & I_z^2 \end{bmatrix} \quad (15)$$

where the distribution of the gradients for the local neighborhood is depicted by the eigenvalues of \mathbf{H} . Therefore, the intensity of gradient changes can be detected by comparing the difference of eigenvalues. For example, a high eigenvalue reveals a large shift in gradient along the related eigenvector, and this rule is widely used to classify the extracted features for 2D pixels. In extracting edges in point clouds, two observations will be detected, namely, 1) the co-existence of two high eigenvalues, which can indicate an edge lying on an intersection between two point cloud surfaces, and 2) a single existence of a high eigenvalue, which can reveal a boundary of the point cloud. In the context of welding applications, both of the situations will be considered because the welding seam only lies on the intersection between two surfaces or the boundary between different workpieces. Additionally, as "high" eigenvalues with a threshold are hard to describe, the ratios between them are measured to identify a threshold t instead of solely using the magnitude of eigenvalues. Most importantly, this ratio calculation scheme does not rely on a time-consuming SVD.

To avoid the calculation of eigenvalues, we use $\lambda_1 = \alpha \lambda_3$ and $\lambda_2 = \beta \lambda_3$ ($\alpha \geq \beta \geq 1$) to represent the three sorted eigenvalues ($\lambda_1 \geq \lambda_2 \geq \lambda_3$). Subsequently, the ratio between trace $\text{Tr}(\mathbf{H})$ and determinant $\text{Det}(\mathbf{H})$ can be calculated using Eq. 16. As the ratio increases with larger α and β , we can set a threshold t to guarantee the existence of large eigenvalues:

$$\frac{\text{Tr}(\mathbf{H})^3}{\text{Det}(\mathbf{H})} = \frac{(\lambda_1 + \lambda_2 + \lambda_3)^3}{\lambda_1 \lambda_2 \lambda_3} = \frac{(\alpha + \beta + 1)^3}{\alpha \beta} \geq t \quad (16)$$

$$G(o, i) = \frac{1}{\delta \sqrt{2\pi}} e^{-\frac{\|p_o - q_i\|^2}{2\delta^2}}, \quad q_i \in \mathcal{N} \quad (17)$$

Then, Eq. (17) is used to smoothen \mathbf{H} and ensure $\text{Det}(\mathbf{H})$ is a non-zero, in which σ is the standard deviation of the distance between current p_o and its neighborhood $\mathcal{N}_r(p_o)$. Finally, a set of edge points $\{p_1^{(e)}, p_2^{(e)}, \dots, p_n^{(e)}\}$ is generated.

F. Welding Path Optimization

On the basis of the extracted edge points, $p_i^{(w)} = p_i^{(e)} + d * \mathbf{n}_{p_i^{(e)}}$ is used to refine the edge point to welding path point for guiding the welding torch, where d is the optimal distance to weld a workpiece suggested by different sets of

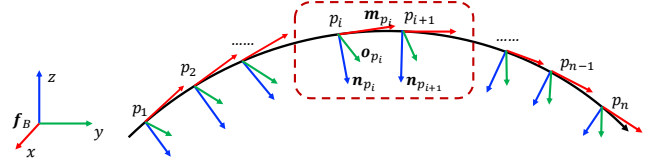


Fig. 6. Determination of three unit orthogonal basis vectors of the 3-DOF rotation for each welding seam pose.

welding equipment. Once the 3-DOF translation parameters $\{p_1^{(w)}, p_2^{(w)}, \dots, p_n^{(w)}\}$ for welding path points are determined, 3-DOF rotation parameters are necessary to form a complete welding path. The computation of rotation parameters needs to define three unit orthogonal basis vectors with respect to the robot base frame. As shown in Fig. 6, the first vector is defined as the pointing vector of the welding torch, which equals to the normal vector of the surrounding welding region point cloud. This approach aligns with the basic requirement of keeping the welding torch perpendicular to the welding region. The second vector is the moving direction of the welding torch, which is defined as the vector \mathbf{v}_m between two nearest path points. Furthermore, to reduce the disturbance of the moving direction, its final moving vector is defined as the average vector of the m nearest moving vectors, i.e., $\mathbf{m}_i = (\sum_{i-\frac{m}{2}}^{i+\frac{m}{2}} \mathbf{v}_{m_i})/m$. According to the right-hand rule, the third vector is computed via cross-multiplication, i.e., $\mathbf{o}_i = \mathbf{m}_i \times \mathbf{n}_i$. Therefore, each path point has its own coordinate that is formed by $(\mathbf{o}_i, \mathbf{m}_i, \mathbf{n}_i)^T$.

In view of achieving superior welding motion performance, increasing the density of path points is necessary to support the robotic manipulator in smoothly executing the path-tracking motion [16]. A linear path interpolation method is designed from two aspects, namely, 3D translation interpolation and 3D rotation interpolation. In particular, a number of ρ points is interpolated between $p_i^{(w)}$ and $p_{i+1}^{(w)}$ and the j -th interpolation point is defined as:

$$p_j^{(w_i)} = p_i^{(w)} + j \cdot \frac{p_{i+1}^{(w)} - p_i^{(w)}}{\rho} \quad (18)$$

Then, each initial path point $p_i^{(w)}$ owns its rotation vector $\mathbf{v}_i^{(w)}$ and the j -th rotation vector is interpolated between $\mathbf{v}_i^{(w)}$ and $\mathbf{v}_{i+1}^{(w)}$, which is computed with the spherical-linear rotation interpolation method:

$$\mathbf{v}_j^{(w_i)} = \frac{\sin((1-\alpha)\theta) \mathbf{v}_i^{(w)} + \sin(\alpha\theta) \mathbf{v}_{i+1}^{(w)}}{\sin \theta} \quad (19)$$

where $\alpha = j/\rho$ and $\theta = \arccos(\mathbf{v}_i^{(w)} \cdot \mathbf{v}_{i+1}^{(w)})$. Finally, a dense and even 6-DOF path for welding is planned.

III. RESULTS

In evaluating the performance of the proposed welding path planning method, we built a robotic system based on UR5 and tested it on several workpieces. To verify the robustness, two complicated types (cylinder and Y-shaped

workpieces) as shown in Fig. 10 are selected in each following step of the entire planning process for illustration. Fig. 7 has previously shown the experimental platform and the RGB-D sensor adopted in this work. The result of conventional teach–playback robotic path planning method, which is regarded the benchmark of welding path planning, is also obtained to compare with our proposed approach.

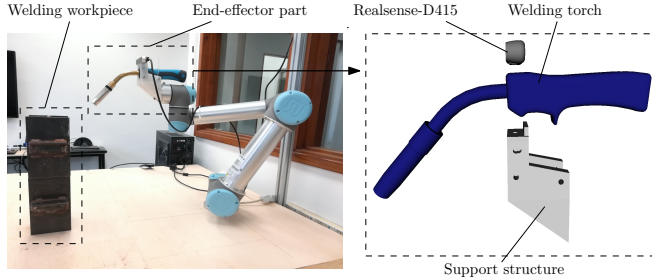


Fig. 7. (Left) Experimental welding platform. (Right) End-effector of the proposed welding robotic system comprising an RGB-D camera and welding torch. The camera is positioned above the welding torch.

A. Gesture-Guide Scanning

In intuitively scanning the workpiece with bare hands, a gesture interaction is implemented via leap motion Python SDK ver. 3.2 and Python library for UR robots, namely, URX. With the leap motion SDK, the hand’s palm position ${}^{LM}p_h$ with respect to the gesture sensor and angles of the raw, pitch and yaw (α, β, γ) for the left hand can be extracted in each captured frame. The interaction box is designed as a spatial cube with a length of 10cm edge, which is appropriate for the interactions in this study. The pose in UR5 is represented by (x, y, z, R_x, R_y, R_z) , while an API (*scipy.spatial.transform*) is used to convert R-P-Y angles into a rotation vector. During interactions, the robotic manipulator only executes six pure linear motions and one rotation which are both based on hand gestures. Precision can be ensured theoretically with a given sufficient time periods. As shown in Fig. 8, the approach direction of the RGB-D camera is fully aligned with the hand’s normal. Consequently, the interaction time is the only key metric

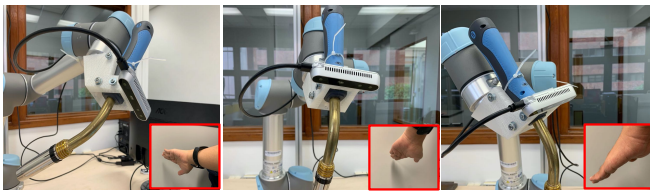


Fig. 8. Results of the robotic scanning process for the workpiece with differently performed hand gestures.

needed to measure this process. The interaction time is defined as the time period with respect to the time when the sensor receives the gesture to the camera, the robot starts to perform corresponding motions. The average interaction time of each motion is measured through 100 gesture actions. The results show that the rotation motion needs more execution

time, compared with linear motions. Linear interaction time is all approximately 0.05s, which is adequate for a normal interaction. By contrast, the average interaction time for rotations is approximately 0.9s, depending on the complexity of the transformation.

B. Workpiece Reconstruction

As shown in Fig. 9, partial point clouds from different perspectives are generated during the scanning process. The camera pose with respect to the world frame is calculated using a forward kinematics module provided by URX and a fixed transformation from the camera to the robot end-effector. On the basis of this information, a preliminary 3D model of the workpieces are reconstructed, as shown in Figs. 9(d) and (h). In measuring the reconstruction performance, a metric of the fitness score is computed as follows:

$$S(\mathbf{M}) = \frac{|\mathcal{P}_m|}{|\mathcal{P}_t|} \left\{ m_i \in \mathcal{P}_m \mid \|t_i - \mathbf{M}m_i\|^2 < d \right\} \quad (20)$$

where t_i and m_i are the target point and the model point, respectively, and \mathbf{M} is the iteratively optimized transformation. Therefore, this score can measure the ratio of successfully aligned points. As depicted by the trend in Fig. 11, the fitness score of the cylinder piece represented by the red spline is higher than the other Y-shaped piece labeled with the black spline throughout the entire ICP process. Before the ICP, the model of the cylinder piece was 86%, and that of the Y-shaped piece was 81%, but these values are not acceptable in high-precision welding. During the ICP process, the cylinder piece quickly reached the convergence when the iteration equalled 5, while the Y-shaped piece needed 15 iterations to reach the convergence. This trend can be explained by the Y-shaped piece that has a more complex geometrical surface than the cylinder piece.

C. Welding Path Detection

After rasterization, the harmonic mean of accuracy is taken to measure the overall performance. The mean is defined as:

$$F_1 = \frac{3}{a_{xy}^{-1} + a_{xz}^{-1} + a_{yz}^{-1}} \quad (21)$$

where a_{xy}^{-1} , a_{xz}^{-1} and a_{yz}^{-1} denote the accuracy calculated for the xy , xz , and yz planes after rasterization. The corresponding rasterized point clouds are shown in Figs. ??

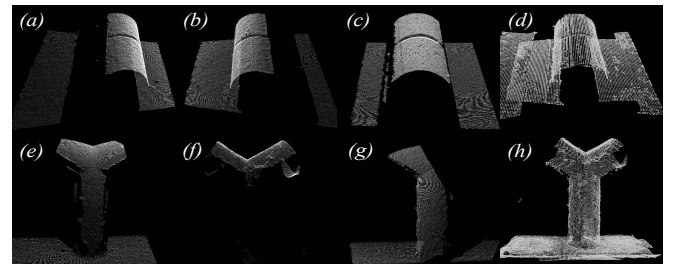


Fig. 9. Model reconstruction. (a)–(c), (e)–(g) Partial point clouds from different views. (d, h) Merge partial point clouds to form a completed model.

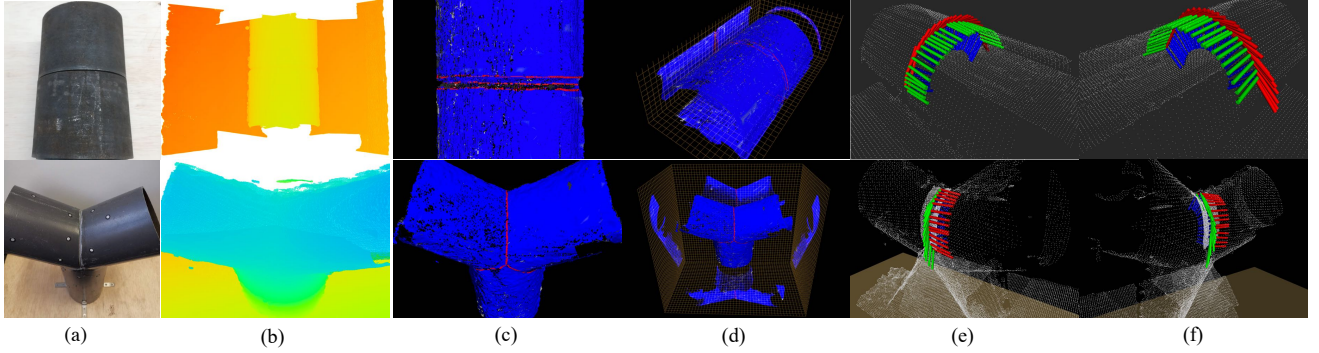


Fig. 10. Results of different processes in the proposed system. (a) Two actual welding metal workpieces, cylinder and Y-shape. (b) Reconstructed models of workpieces with a depth map. (c) Welding seam extracted with the edge intensity gradient-based algorithm. (d) Rasterized views of extracted edge points. (e, f) 6-DOF welding path after interpolation.

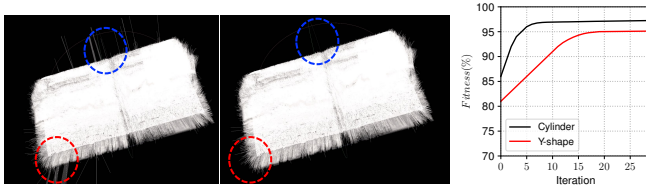


Fig. 11. (Left) Comparative results of pre and post-denoising with the bilateral filter. (Right) Fitness score of two workpieces with different iterations.

(e)-(h). Each detected edge point is displayed in red, and the ground truth path point is manually marked in green based on point cloud generated by a high-precision 3D scanner. The overall performance calculated on the basis of the coverage of the red edge point over the ground truth path rasterized on different planes. The effect of the bilateral filter on the welding path detection also needs to be determined. Figs. 12 (a) and (b) show the overall pre- and post-filtering performances of the two selected workpieces, respectively. The

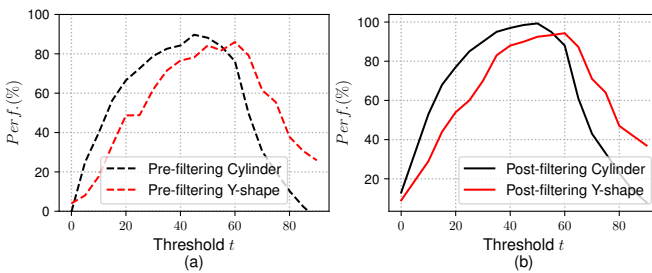


Fig. 12. Pre- and post-performance scores under different thresholds.

figures show that the performances of the two workpieces both have increased after filtering. This trend is due to the filtering, which reduces the noise and smoothens the surface for the point cloud data. Fig. 4 validates this difference based on the normal map of the pre- and post-filtering processes of the workpieces, especially on the corresponding regions surrounded by dashed circles. Besides, different predefined thresholds t have an impact on the overall performance. Figs. 13(a)-(c) show the extracted edges when t is equal to 10, 70, and 50, respectively. If the t is too small or too large, then

it will negatively affect the overall performance score. The cylinder has reached the peak when its threshold is near 53, whereas the Y-shaped piece needs a slightly larger threshold of approximately 60 to attain the peak performance.

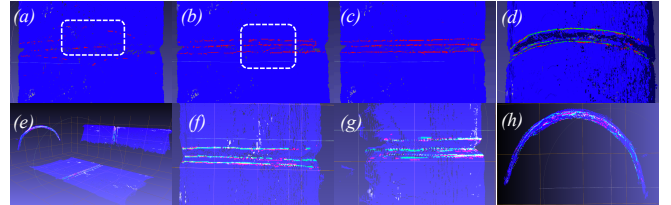


Fig. 13. (a)-(c) Overall performance with $t = 10, 70, 50$ (d) Extracted edge points and ground truth edge points. (e)-(h) Rasterization to measure the precision of the generated welding path.

D. Pose Optimization

After optimization, the welding path planned by the teach-playback method is selected as the standard path, which is regarded as our ground truth for comparing the welding path generated by our framework. For each workpiece, two individual paths are generated, and the root mean square error (RMSE) of the two paths is used to evaluate both the translation and rotation accuracy of the proposed method. The RMSE is defined as:

$$RMSE(m, g) = \sqrt{\frac{1}{n} \sum_{i=1}^n (m_i - g_i)^2} \quad (22)$$

where $\{m_i\}$ represents the generated welding path by our model, $\{g_i\}$ is the ground truth of the welding path, and n denotes the number of path points. The RMSE results of the two workpieces are shown in Table I, which reveals acceptable accuracy for both translation and rotation.

E. Comparison with Existing Methods

Compared with previous studies on groove detection, thresholding eigenvalue of gradient makes system performance less vulnerable to noise in complex environments and thus our system can be flexibly applied to a wide

TABLE I
RESULTS OF POSE ACCURACY FOR EACH WORKPIECE

Types	n_d	n_o	$RMSE_p$	$RMSE_\theta$
Cylinder	78	300	0.7mm	1.5°
Y-shaped	65	300	0.8mm	1.3°

n_d = The number of detected path points
 n_o = The number of points after optimization.

range of workpiece. This is supported by sufficient experiments from plane placed at various poses to complex 3d shapes. However, the output welding path is computed based on previous downsampled point cloud which may fail to maintain concrete local geometric information, especially for points at the corners. Accuracy is another weakness of the system as the standard value given by other studies [1]–[3], [7] is under 0.5 mm. Instead of using manual controls or pre-programming poses for positioning the camera, we adopt a gestured based interaction for scanning. Combined with a broad view vision sensor, the system can reconstruct free form workpieces and plan the path at a competitively fast speed. We measured the runtime for the whole system on a 500 mm straight V-type butt joint as this type is commonly presented in the literature. The full workpiece can be reconstructed from one capture. The average time for 20 runs is taken for comparison and the detection speed defined by length of seam over runtime is calculated. The results listed in Table II show that our system outputs consistent results across different settings and is significantly efficient for seam detection.

TABLE II
RUNTIME ON V-TYPE BUTT JOINT WELDING PATH PLANNING

System	Joint Pose	Method	Runtime(ms)	Speed(mm/s)
Ours	flat	Point cloud	3691	135
Ours	tilted	Point cloud	3891	129
Ours	horizontal	Point cloud	3600	139
Ours	vertical	Point cloud	4100	122
Ref [9]	flat	Laser vision	\	5
Ref [17]	flat	Point cloud	4913	\

IV. CONCLUSION

This article presents a sensor-guided welding robotic system. Specifically, an interactive algorithm based on gesture is first proposed to scan the workpiece in an intuitive manner. Then, the system reconstructs a 3D point cloud model of the target by a linear ICP algorithm. The precision and speed of the reconstruction is ensured by taking the pose obtained from robotic kinematics as the initial setting in ICP. Next, a bilateral filter is implemented to denoise point cloud while maintaining edge information and the welding seam joints is robustly identified using a gradient-based edge detection approach. Based on the detected seam, a smooth 6-DOF path is computed for execution. Experiments on various geometric shapes have validated the effectiveness of the proposed system. In the future, a higher precision

depth camera will be used to improve the accuracy of model reconstruction. Additionally, we may introduce other new interaction techniques to assist robotic welding tasks.

REFERENCES

- [1] Y. Xu, H. Yu, J. Zhong, T. Lin, and S. Chen, "Real-time seam tracking control technology during welding robot gtaw process based on passive vision sensor," *J Mate Process Tech.*, vol. 212, no. 8, pp. 1654–1662, 2012.
- [2] J. Fan, S. Deng, F. Jing, C. Zhou, L. Yang, T. Long, and M. Tan, "An initial point alignment and seam-tracking system for narrow weld," *IEEE Trans. Industr. Inform.*, vol. 16, no. 2, pp. 877–886, 2019.
- [3] C. Diao, J. Ding, S. Williams, Y. Zhao, *et al.*, "A passive imaging system for geometry measurement for the plasma arc welding process," *IEEE Trans. Ind. Electron.*, vol. 64, no. 9, pp. 7201–7209, 2017.
- [4] X. Li, X. Li, S. S. Ge, M. O. Khyam, and C. Luo, "Automatic welding seam tracking and identification," *IEEE Trans. Ind. Electron.*, vol. 64, no. 9, pp. 7261–7271, 2017.
- [5] H. N. M. Shah, M. Sulaiman, A. Z. Shukor, M. Rashid, and M. Jamaluddin, "Review paper on vision based identification, detection and tracking of weld seams path in welding robot environment," *Mod. App. Sci.*, vol. 10, no. 2, pp. 83–89, 2016.
- [6] A. Rout, B. Deepak, and B. Biswal, "Advances in weld seam tracking techniques for robotic welding: A review," *Robot. Cim-Int. Manuf.*, vol. 56, pp. 12–37, 2019.
- [7] L. Nele, E. Sarno, and A. Keshari, "An image acquisition system for real-time seam tracking," *Int. J. Adv. Manuf. Tech.*, vol. 69, no. 9-12, pp. 2099–2110, 2013.
- [8] Y. Ding, W. Huang, and R. Kovacevic, "An on-line shape-matching weld seam tracking system," *Robot. Cim-Int. Manuf.*, vol. 42, pp. 103–112, 2016.
- [9] R. Manorathna, P. Phairatt, P. Ogun, T. Widjanarko, M. Chamberlain, L. Justham, S. Marimuthu, and M. R. Jackson, "Feature extraction and tracking of a weld joint for adaptive robotic welding," in *Int. Conf. Cont. Auto. Robot & Vis.*, 2014, pp. 1368–1372.
- [10] L. Zhang, Y. Xu, S. Du, W. Zhao, Z. Hou, and S. Chen, "Point cloud based three-dimensional reconstruction and identification of initial welding position," in *Trans. Intel. Weld. Manuf.*, 2018, pp. 61–77.
- [11] S. M. Ahmed, Y. Z. Tan, *et al.*, "Edge and corner detection for unorganized 3d point clouds with application to robotic welding," in *IEEE/RSJ Int.Path Planning with Automatic Seam Extraction over Point Cloud Models for Robotic Arc Welding Peng Zhou, Rui Peng, Maggie Xu, Victor Wu and David Navarro-Alarcon Abstractâ This paper presents a point cloud based robotic system for arc welding. Using hand gesture controls, the system scans partial point cloud views of workpiece and reconstructs them into a complete 3D model by a linear iterative closest point algorithm. Then, a bilateral filter is extended to denoise the workpiece model and preserve important geometrical information. To extract the welding seam from the model, a novel intensity-based algorithm is proposed that detects edge points and generates a smooth 6-DOF welding path. The methods are tested on multiple workpieces with different joint types and poses. Experimental results prove the robustness and efficiency of this robotic system on automatic path planning for welding applications.* I. INTRODUCTION Industrial welding robots have been extensively studied with the aim to solve labor shortage problem and improve productivity. Currently, teach and playback mode or offline programming prevails in robotic welding workshops. As these conventional practices lack adaptation to accommodate changes in a working environment, many advanced techniques such as seam detection and tracking are gradually being employed in industry. Solutions to intelligent robotic arc welding primarily require the assistance of computer vision. Previous studies have conducted in-depth research on image processing and established viable methods across several visual sensors, including mono camera, stereo camera, and CCD camera [1]–[6]. [7] developed an automatic image acquisition system for real-time weld pool analysis and seam tracking. The seam characteristics were measured by defining region of interest and implementing pattern learning algorithm. [8] proposed a shape-matching algorithm for an online laser-based seam detection. The algorithm enabled autonomous detection for different groove types and localizes the boundary of weld seam. [9] measured the cross-sectional areas of 2D profiles and finds joint features and volume of the workpiece. While above

mentioned studies can effectively locate the seam, their implementations are only validated on simple or flat welding workpieces with common joint types. In most cases, the groove problem is regarded as two-dimensional and image processing is sufficient to identify the target. However, it is important to discuss the generalization of these techniques to cope with complex geometries. The remaining challenge lies in capturing free-form grooves. Point cloud All authors are with The Hong Kong Polytechnic University, Kowloon, Hong Kong. Corresponding author e-mail: dna@ieee.org This work is supported in part by the Research Grants Council of HK (grant 14203917) and in part by the Chinese National Engineering Research Centre for Steel Construction Hong Kong Branch (grant BBV8). processing is favored as it can comprehensively describe complex 3D shapes. A captured point cloud model can reveal useful geometric information but simultaneously introduce considerable noise. The complexity and computational cost also evolve as the information increases. Therefore, designing a robust algorithm that can efficiently utilize 3D information will be crucial for this application. Although, some recent studies using point cloud processing were able to detect the groove, they did not further exploit the features to improve system functionality, such as plan a welding path with well-defined orientation [10]–[14]. More experiments on complex and diverse seam patterns are needed for verification. In addition, the authors found that most systems need to manually set or program the initial poses for visual sensors to capture the workpiece. This guidance of scanning may be realized in a more intuitive way to facilitate human machine collaboration. To address the above problems, we propose a novel point cloud based system for robotic welding planning. This system integrates a single RGB-D camera with a cooperative industrial manipulator to fulfill the generic requirements for seam detection. Guided by hand gestures, the system captures a set of partial point clouds of the workpiece from different perspectives. To reconstruct the surface, it adopts a linear iterative closest point algorithm that improves the rate of fitness and extends bilateral filter algorithm for 3D surface denoising. A novel algorithm is implemented to identify the groove based on edge density. After the localization, the system plans a smooth consecutive welding path with 6-DOF and executes it with the robot manipulator. To validate the proposed methods, we conduct detailed experiments on workpieces with different joint types. The remaining parts of this article are as follows. Section II introduces the architecture of the proposed planning system and the key techniques applied. The experimental results of the proposed system are presented in Section III. Section IV gives final conclusions.

- II. METHODOLOGY
- A. System Overview
- The robotic system architecture is depicted in Fig. 1. The complete surface model reconstruction for the target welding workpiece, which a hand gesture-based method, is designed to interact with the manipulator with assistance from a human operator. Subsequently, the depth camera, which is mounted on the end-effector of the robot, moves to several poses to capture different views of the point cloud. Then, the point cloud model of the workpiece is reconstructed by combining Conf. on Robots and Intelligent Systems, 2018, pp. 7350–7355.
- [12] V. Patil, I. Patil, *et al.*, “Extraction of weld seam in 3d point clouds for real time welding using 5 dof robotic arm,” in *Int. Conf. Cont., Auto. and Robot*, 2019, pp. 727–733.
 - [13] L. Jing, J. Fengshui, and L. En, “Rgb-d sensor-based auto path generation method for arc welding robot,” in *Chinese Control and Decision Conf.*, 2016, pp. 4390–4395.
 - [14] K. Zhang, M. Yan, T. Huang, J. Zheng, and Z. Li, “3d reconstruction of complex spatial weld seam for autonomous welding by laser structured light scanning,” *J. Manuf. Process*, vol. 39, pp. 200–207, 2019.
 - [15] K.-L. Low, “Linear least-squares optimization for point-to-plane icp surface registration,” *Chapel Hill, Univ. North Carolina*, vol. 4, no. 10, pp. 1–3, 2004.
 - [16] R. Peng, D. Navarro-Alarcon, *et al.*, “A point cloud based method for automatic groove detection and trajectory planning of robotic arc welding tasks,” *IEEE Int. Conf. on Ubiquitous Robots*, pp. 1–7, 2020.
 - [17] L. Yang, Y. Liu, J. Peng, and Z. Liang, “A novel system for off-line 3d seam extraction and path planning based on point cloud segmentation for arc welding robot,” *Robot. and Cim-Int. Manuf.*, vol. 64, p. 101929, 2020.