

ORIGINAL ARTICLE

Open Access



# Three-Dimensional Reconstruction of Welding Pool Surface by Binocular Vision

Zunan Gu, Ji Chen\*  and Chuansong Wu

## Abstract

Current research of binocular vision systems mainly need to resolve the camera's intrinsic parameters before the reconstruction of three-dimensional (3D) objects. The classical Zhang' calibration is hardly to calculate all errors caused by perspective distortion and lens distortion. Also, the image-matching algorithm of the binocular vision system still needs to be improved to accelerate the reconstruction speed of welding pool surfaces. In this paper, a preset coordinate system was utilized for camera calibration instead of Zhang' calibration. The binocular vision system was modified to capture images of welding pool surfaces by suppressing the strong arc interference during gas metal arc welding. Combining and improving the algorithms of speeded up robust features, binary robust invariant scalable keypoints, and KAZE, the feature information of points (i.e., RGB values, pixel coordinates) was extracted as the feature vector of the welding pool surface. Based on the characteristics of the welding images, a mismatch-elimination algorithm was developed to increase the accuracy of image-matching algorithms. The world coordinates of matching feature points were calculated to reconstruct the 3D shape of the welding pool surface. The effectiveness and accuracy of the reconstruction of welding pool surfaces were verified by experimental results. This research proposes the development of binocular vision algorithms that can reconstruct the surface of welding pools accurately to realize intelligent welding control systems in the future.

**Keywords:** Surface reconstruction, Welding pool, Binocular imaging, Gas metal arc welding, Feature points matching

## 1 Introduction

Gas metal arc welding (GMAW) is widely applied in modern manufacturing industries. To improve the weld quality of manual GMAW, welders correct either the welding parameters or the position of welding gun based on information from the welding pool surface acquired by sight and their expertise [1]. This means that the welding pool surface contains important visual information that allows skilled welders to control GMAW process. To decrease the workload and occupational risks for welders, much research has been done in the last few decades to transfer this inspection work to intelligent control systems [2].

Caprio [3] utilized the oscillation of the welding pool to estimate its penetration during the Laser Powder Bed Fusion process. Thomas [4] reconstructed the thermal images of cold metal transfer process to monitor the porosity defects and improper weld beads. Vasilev [5] utilized an ultrasonic thickness measurement system to control the welding current and welding speed. These methods have mostly focused on improving the welding process with information from one-dimensional or two-dimensional welding pool surface data. However, it is clear that three-dimensional (3D) information of the welding pool surface can better reflect the weld quality.

The 3D reconstruction method for welding pool can be categorized into the 3D structured light method [6], shape from shading method [7], and binocular stereo vision method [8]. To avoid arc interference during measurement, Zhang et al. [9] proposed a novel-structured laser vision system for reconstructing 3D welding

\*Correspondence: chenji@sdu.edu.cn

Key Laboratory for Liquid-Solid Structural Evolution & Processing of Materials, Ministry of Education, Institute of Materials Joining, Shandong University, Jinan 250061, China

pool surface. In their experiments, a structured laser was projected onto the welding pool surface and was specularly reflected onto a preset imaging plane. The images of the laser points on the plate were captured by camera. Finally, the geometry of the welding pool surface was reconstructed by the Edge-point algorithm and One-point algorithm [10]. However, the deviation of this method can reach a maximum of 1.22 mm. Chen [11] proposed a new image processing framework that could extract key characteristics of 3D welding pool geometry through its two-dimensional passive vision images. It was found that the support vector machine classification model was accurate enough to obtain the width and height of the welding pool. The camera's exposure time had to be carefully adjusted for different welding currents, as it remarkably affected the image quality of the welding pool and the accuracy of the results. Zhong et al. [12] put forward measures to increase the effectiveness and accuracy of shape from shading (SFS) method. An improved algorithm was used to reconstruct the weld surface of aluminum alloy for gas tungsten arc welding (GTAW). However, the SFS method is difficult to be used in GMAW process as its spatters, arc interference, and welding stability is much worse than that in GTAW. An accurate and anti-jamming 3D reconstruction method for GMAW is still needed for the suppression of strong electromagnetic interference, high heat radiation interference, weld fumes, and other conditions.

The binocular stereo vision method is more accurate and efficient in acquiring reconstruction results. Mnich [13] attempted to use this method to reconstruct a GMAW pool, yet results showed that the binocular vision system needed to be improved, as extra-bright areas on pool surface could not be reconstructed. Richardson [14] used a particle image velocimeter (PIV) to track the movement of oxide particles to determine the velocity field on welding pool and then transferred the above data into a 3D fluid velocity by the stereo vision method. The calibration process was the main factor that influenced measurement precision during 3D reconstruction; the average distance error between the tested corner points was 0.14 mm. Liang [15] established a biprism stereo vision system to characterize weld penetration based on the geometry of the welding pool surface under pulsed GMAW. A two-step stereo matching algorithm was proposed to reconstruct the 3D shape of the welding pool surface. However, some regions on the surface were discontinuous, and the accuracy of this algorithm was not discussed. Xiong [16] also developed a biprism stereo vision system to reconstruct welding pool geometry in GMAW. A global-based iterative matching algorithm and triangle measuring method were optimized. The results were validated by a reconstructed standard cylinder

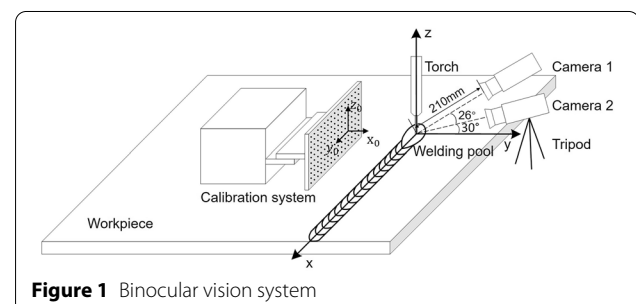
with clear checkboard, yet the maximum height error was 4.15%, showing that the accuracy and usability of a global-based iterative matching algorithm for GMAW process without checkboard still needs to be verified and improved.

In this paper, an improved binocular vision system was developed to observe the welding pool surface. An automatic 3D reconstruction method was proposed to suppress arc interference on welding pool information. The mathematical models including detection, description and matching of feature points were established to effectively and robustly calculate the world coordinates of welding pool surface against different welding conditions. This study was foundation of on-line intelligent control of welding pool behavior for GMAW process.

## 2 Experimental System

The binocular vision system is given in Figure 1. Two Basler acA2000-165uc cameras (color) were used, and their CMOS sizes were both 2/3". The focal length of the camera lens was 35 mm with a 650-nm cut-off filter and 7# welding goggle. The aperture was F/5.6. Two cameras were connected with a synchronizer trigger to generate two synchronous images. The frequency of sampling was 200 frames/s with a resolution of  $600 \times 500$  pixels and exposure time of 40  $\mu$ s.

The weld workpiece was fixed on a moveable workbench. Bead-on-plate welding tests were carried out on Q235 mild steel plates with dimensions of 250.0 mm  $\times$  70.0 mm  $\times$  5.0 mm. The welding wire material was H08Mn2Si with a diameter of 1.2 mm. The distance between the welding wire tip and workpiece surface was 18.0 mm. The chemical compositions of the welding wire and base metal are presented in Table 1. The direct current electrode negative mode was used, and the distance from the wire tip to the workpiece was 18.0 mm. Other welding parameters are listed in Table 2. The torch and cameras were stationary during the welding process. The workpiece and workbench were moved at a preset welding speed.



**Figure 1** Binocular vision system

**Table 1** Chemical compositions of welding wire and base metal (wt%)

Composition (%)	C	Si	Mn	P	S
Wire (H08Mn2Si)	0.11	0.65	1.80	0.030	0.030
Q235	0.17	0.15	0.30	0.015	0.035

**Table 2** Welding parameters

Welding current (A)	Arc voltage (V)	Welding speed (m/min)	Shielding gas (L/min)
240–280	33.0	0.5–0.8	20(Ar)

### 3 Calibration System

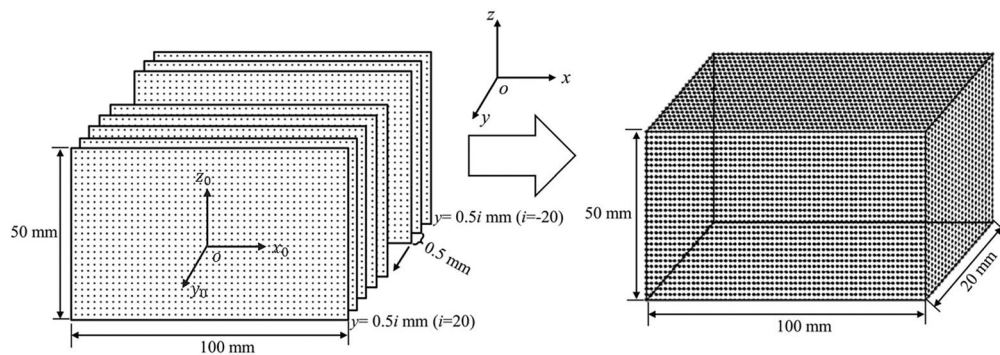
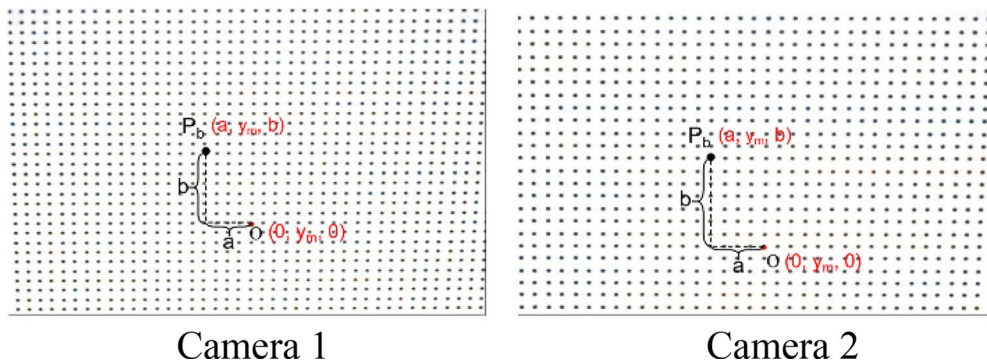
In order to effectively avoid errors caused by perspective distortion and lens distortion (including radial distortion, centrifugal distortion, and prism distortion), a preset coordinate system for camera calibration [17] was used to establish the corresponding relationship between the world coordinates and the image coordinates via a calibration target with a point matrix at different positions. For example, Figure 1 illustrates how

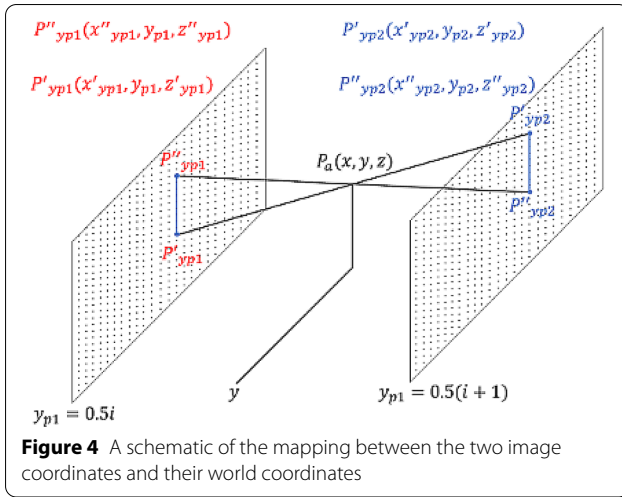
two cameras were fixed on a tripod. Figure 2 shows how the target paper was moved along the  $y$  direction from  $y=0.0$  mm to  $y=20.0$  mm with a moving step length of 0.5 mm. The target paper was captured by the two cameras at each position. The red dot on the target paper represents the coordinate origin of the point matrix. Finally,  $O_{xyz}$  and  $O_{x_0y_0z_0}$  were the world coordinate system and pixel coordinate system, respectively.

The world coordinates of black dots in each target image were calculated by their distance from the red dot. That is, when the black dot  $P_b$  was located in the  $a$  row to the left of the red dot, its world coordinate was  $x_d = a \times 1$  mm =  $a$  mm; when the black dot was located in the  $b$  row above the red dot, its world coordinate was  $z_d = b \times 1$  mm =  $b$  mm, as shown in Figure 3. It can be seen that the world coordinates of the black dots are  $[x_d, y_d, z_d] = [a, y_m, b]$ . When the world coordinates of all grid points on the graph were assigned, the world coordinate system of the calibration system was successfully constructed.

It is assumed that a point  $P_a$  was located in the area of calibration that was captured by two cameras, as shown in Figure 4.

The world coordinates of  $P_a$  could be reconstructed from the following procedure:

**Figure 2** The 3D point array**Figure 3** Captured images of the calibration target by two cameras at  $y_m$



(1) The image coordinates of  $P_a$  were  $P_1(X_1, Y_1)$  in camera 1 and  $P_2(X_2, Y_2)$  in camera 2. The four nearest neighbors of  $P_1$  in the captured image were  $A(x_A, y_A)$ ,  $B(x_B, y_B)$ ,  $C(x_C, y_C)$ , and  $D(x_D, y_D)$ . Their world coordinates, which were calibrated prior, were  $A'(x_A, y_p, z_A)$ ,  $B'(x_B, y_p, z_B)$ ,  $C'(x_C, y_p, z_C)$ , and  $D'(x_D, y_p, z_D)$ .  $P'(x'_{yp}, y_p, z'_{yp})$  could be calculated by the four nearest neighbors of  $P_1$  as follows:

$$\begin{bmatrix} x_{A'} & x_{B'} & x_{C'} & x_{D'} \\ z_{A'} & z_{B'} & z_{C'} & z_{D'} \end{bmatrix} \times \begin{bmatrix} X_A & X_B & X_C & X_D \\ Y_A & Y_B & Y_C & Y_D \end{bmatrix}^{-1} \times \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix} = \begin{bmatrix} x'_{yp} \\ z'_{yp} \end{bmatrix}. \quad (1)$$

Similarly, all the corresponding world coordinates of  $P_2$  at 41 planes can be obtained as  $P''(x''_{yp}, y_p, z''_{yp})$  ( $y_p = 0.5i$ ,  $i = -20, -19, \dots, 19, 20$ ).

(2) The distance  $D_{yp}$  between  $P'(x'_{yp}, y_p, z'_{yp})$  and  $P''(x''_{yp}, y_p, z''_{yp})$  at all the  $y$ -planes was calculated as follows:

$$D_{yp} = \sqrt{(x'_{yp} - x''_{yp})^2 + (z'_{yp} - z''_{yp})^2}, \quad (y_p = 0.5i, i = -20, -19, \dots, 19, 20). \quad (2)$$

When  $P_a$  was in a certain  $y$ -plane ( $y = 0.5i$ ),  $P'$  and  $P''$  overlapped, and  $D_{yp}$  was 0. Then  $x'_{yp} = x''_{yp}$  and  $z'_{yp} = z''_{yp}$ ; the world coordinates of  $P_a$  were  $(x'_{yp}, y_p, z'_{yp})$ . Otherwise,  $P_a$  was located between two adjacent plane matrices where the sum of  $D_{yp1}$  and  $D_{yp2}$  represented the minimum. If  $y_{p1} = 0.5i$  and  $y_{p2} = 0.5(i+1)$ , the world coordinates were as follows:

$$x = \frac{1}{2}(x''_{yp1} + x'_{yp1}), \quad (3)$$

$$y = \frac{(z'_{yp2} - z''_{yp2}) \times y_{p1} + (z''_{yp1} - z'_{yp2}) \times y_{p2}}{(z'_{yp2} - z''_{yp2}) + (z''_{yp1} - z'_{yp2})}, \quad (4)$$

$$z = \frac{(z'_{yp2} - z''_{yp2}) \times z'_{yp1} + (z''_{yp1} - z'_{yp2}) \times z''_{yp2}}{(z'_{yp2} - z''_{yp2}) + (z''_{yp1} - z'_{yp2})}, \quad (5)$$

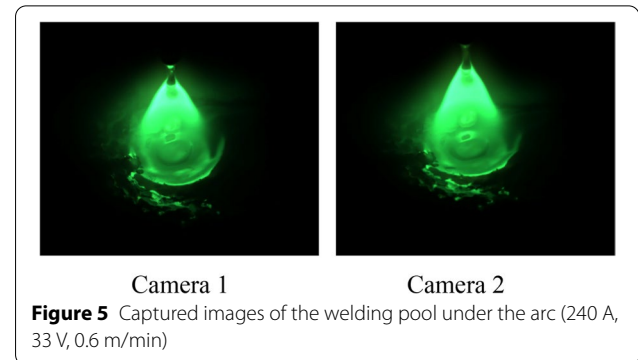
where  $P'_{yp1}$  and  $P'_{yp2}$  were points captured by camera 1, and their world coordinates are  $(x'_{yp1}, y_{p1}, z'_{yp1})$  and  $(x'_{yp2}, y_{p2}, z'_{yp2})$ , respectively. Similarly,  $P''_{yp1}$  and  $P''_{yp2}$  were points captured by camera 2, and their world coordinates were  $(x''_{yp1}, y_{p1}, z''_{yp1})$  and  $(x''_{yp2}, y_{p2}, z''_{yp2})$ , respectively. Thus, the world coordinates of point  $P_a$  were successfully achieved. Verification experiments showed that the maximum relative error between the recovered and measured results was less than 0.6% [17].

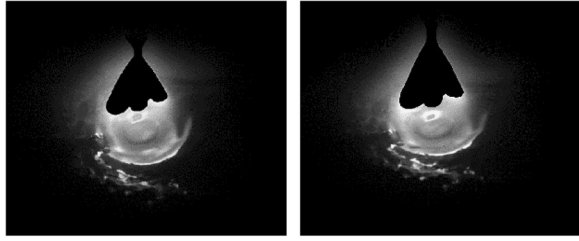
#### 4 Pre-preprocessing of Welding Pool Image

The size of the welding pool under the arc is small, and there are various noises in the process of image acquisition and quantification. In addition, the existence of a strong arc will also affect the extracting and matching processes of feature points, which decreases the accuracy of 3D reconstruction. Therefore, image processing algorithms, which mainly include gray-level transformations, image filtering, and edge detection, are needed to pre-process the welding pool images [18]. Figure 5 shows the original images of the welding pool under the arc. Figure 6 shows the processed images, which can be used for feature-point detection.

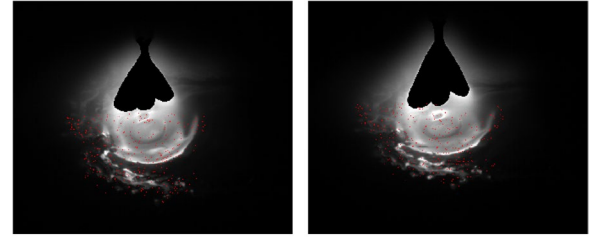
#### 5 SURF-BRISK-KAZE Feature Point Matching Algorithm

Finding the matching feature points accurately and efficiently is a high priority for welding pool surface reconstruction. The algorithms need to adapt to various requirements, such as the small scale of the welding pool, affine transformation of camera angles, noise interference





**Figure 6** Pre-processed images of the welding pool under the arc (240 A, 33 V, 0.6 m/min)



**Figure 7** SURF feature points of the welding pool image under the arc (240 A, 33 V, 0.6 m/min)

of spatter and arc electromagnetism, severe illumination of arc, and nonlinear image distortion.

### 5.1 SURF Feature Point Detection

For some welding pool images captured by cameras 1 and 2, box filters [19] were used to approximately replace the Gaussian second order derivatives. The different box filters were selected to establish the scale space for feature-point detection. The feature points were extracted by a Hessian matrix and a threshold on each scale space layer. For an image  $I(x, y)$ , a Hessian matrix with scale  $\sigma$  at point  $(x, y)$  was defined as follows:

$$H = \begin{bmatrix} L_{xx}(x, y, \lambda) & L_{xy}(x, y, \lambda) \\ L_{yx}(x, y, \lambda) & L_{yy}(x, y, \lambda) \end{bmatrix}, \quad (6)$$

where  $L_{xx}$  is the convolution of the Gaussian second order partial derivative (Gaussian filter), and the input image  $I(x, y)$  at point  $(x, y)$ , the meaning of  $L_{xy}$ ,  $L_{yy}$ , and  $L_{yx}$  are the same as  $L_{xx}$ . The Gaussians is defined as

$$g(\lambda) = 0.5 \exp\left(-\frac{x^2 + y^2}{2\lambda^2}\right) / (2\pi\lambda^2). \quad (7)$$

To facilitate the calculation, the convolution  $D_{xx}$ ,  $D_{xy}$ , and  $D_{yy}$  of the box filters and the input image was used to replace  $L_{xy}$ ,  $L_{yy}$ , and  $L_{yx}$  to construct the fast Hessian matrix. Box filters in size  $9 \times 9$  were substituted by the Gaussian filter when parameter  $\lambda$  was 1.2. The relationship between  $D_{xx}$ ,  $D_{yy}$  and  $L_{xx}$ ,  $L_{yy}$  is as follows:

$$\omega = \frac{\|L_{xy}(1.2)\|_F \|D_{xx}(9)\|_F}{\|L_{xx}(1.2)\|_F \|D_{xy}(9)\|_F} \approx 0.9, \quad (8)$$

where  $\|\cdot\|_F$  is the Frobenius norm, and  $\omega$  is used to balance the relative weights in the expression for the Hessian's determinant. In practical application, a value of 0.9 was adopted to obtain an approximate Hessian matrix determinant:

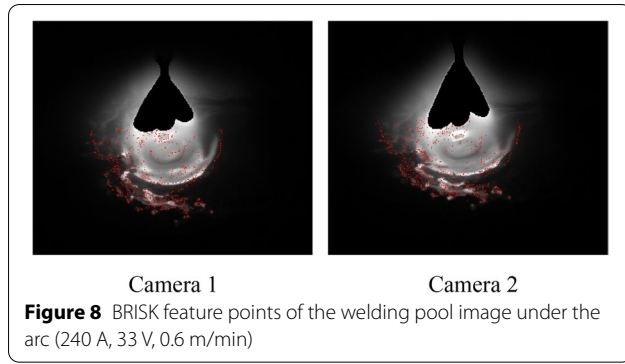
$$\det(H_{essian}) = D_{xx}D_{yy} - (0.9D_{xy})^2. \quad (9)$$

A threshold was needed to estimate the feature points detected by the Hessian matrix. When the value of Hessian's determinant for the test point was greater than this threshold, a non-maximum suppression in a  $3 \times 3 \times 3$  neighborhood was applied. When the value of a test point was greater than that of 26 neighbor points, it was selected as an optimal feature point. When the maxima detected by Hessian matrix was less than the threshold, this maximum was excluded, so the speed of maxima detection was accelerated. Figure 7 shows the feature points extracted by the speeded up robust features (SURF) algorithm. It can be seen that the SURF feature points (red dots) were evenly distributed in the welding pool, having the advantage of affine invariance; however, it cannot fully reflect the edge profile.

### 5.2 BRISK Feature Point Detection

In order to obtain the characteristic of scale invariance on the edges of the welding pool, the scale space was composed of four inner layers  $c_i$  and four middle layers  $d_i (i=0,1,2,3)$  in the frame structure of binary robust invariant scalable keypoint (BRISK) feature detection [20]. Each inner layer image was obtained by 0.5-times down sampling of the previous inner layer image, where the original image corresponded to the  $c_0$  layer. Each intermediate layer  $d_i$  was located between two adjacent inner layers  $c_i$  and  $c_{i+1}$ . The first intermediate layer  $d_0$  was obtained by 1.5-times down sampling of the original image  $c_0$ , and the remaining intermediate layers were obtained by 0.5-times down sampling of the previous intermediate layer. It was assumed that  $\delta$  represented the scale, where  $\delta(c_i) = 2^i$  and  $\delta(d_i) = 1.5 \times 2^i$ .

After the scale space of BRISK was constructed, the feature points were extracted on each scale. The extreme points detected in the spatial domain and the scale



domain were regarded as feature points, so the BRISK corner points had scale invariance for the welding pool edges. BRISK feature point detection was determined by the following equation:

$$N = \sum |I(x) - I(R)| < \varepsilon, \quad (10)$$

where  $I(R)$  was the gray value of the central pixel  $R$ ,  $I(x)$  was the gray value of the pixel points surrounding  $P$ , and  $\varepsilon$  was the threshold, which was 0.00001 in this study. If  $N$  was greater than  $\varepsilon$ , the candidate point was the feature point. BRISK feature points on the welding pool images are shown in Figure 8. It can be seen that the BRISK feature points reflected the edge profile well. However, the resolution on each layer of scale space calculated by the SURF and BRISK algorithms was blurry, which affected the accuracy of feature points. It also affected the number of feature points detected in the blurry area, created by welding pool oscillation.

### 5.3 KAZE Feature Point Detection

The additive operator splitting algorithm and the variable conductance diffusion method were used to construct nonlinear scale space in this study. The logarithmic steps were used in nonlinear scale space to generate the  $O$  group and  $S$  group [21]. Different groups and layers were identified by group index  $o$  and layer index  $s$ , respectively. The corresponding relationship between group, layer and scale parameter  $\sigma$  was expressed as follows:

$$\begin{aligned} \sigma_i(o, s) &= \sigma_0 2^{o+s/S}, o \in [0, \dots, O-1], \\ s &\in [0, \dots, S-1], i \in [0, \dots, O \times S-1], \end{aligned} \quad (11)$$

where  $\sigma_0$  was the initial scale parameter, which flattens the original image to reduce the noise induced by the arc's magnetic field. The relationship between time parameter  $t$  and scale parameter  $\sigma_i$  is:

$$t_i = \frac{1}{2} \sigma_i^2, i \in [0, \dots, O \times S-1]. \quad (12)$$

The nonlinear scale space is expressed as:

$$L^{i+1} = (I - (t_{i+1} - t_i) \cdot \sum_{i=1}^m A_i L^i)^{-1} L^i, \quad (13)$$

where  $L$  is the luminance of the image and  $A_i$  is a matrix that encodes the image conductivities for each dimension.

Combining Eqs. (11) and (13), the KAZE feature points on two images captured by cameras 1 and 2 are shown in Figure 9. The results show that this method was effective to extract feature points from smooth surface with small brightness differences, especially for the white-dotted-line area.

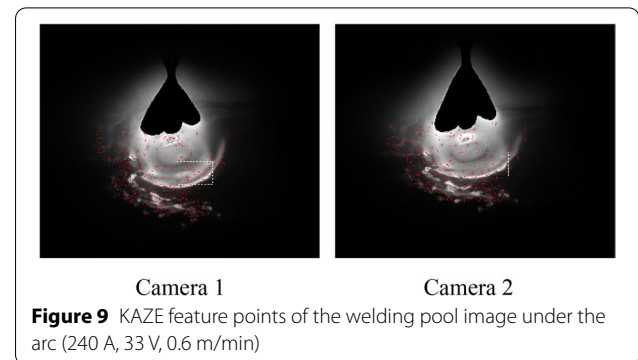
### 5.4 SURF-BRISK-KAZE Feature Description and Matching

In order to match the feature points in the SURF, BRISK, and KAZE methods, a descriptor was determined for the extracted points. The details of the descriptor, which is a 64-dimensional vector, has been discussed in reference [19].

The 64-dimensional vector obtained by the above method only included grayscale information of the welding pool images; the color information of images were not considered, which might decrease the matching accuracy of welding pool images.

The information of the original welding pool images (color) was much richer than the pre-processed welding pool images (gray) in Section 4. The signal interference was also included in the original images. To utilize the color information of the original images effectively, only the information of feature points in Sections 5.1–5.3 and their eight neighborhoods were used to improve the above 64-dimensional vector. The steps were as follows:

(1) It was assumed that  $r, g, b$  were the values of color information for the feature point  $(x, y)$ . The value  $n_c$  was the total number of the feature points  $(x, y)$  and their



neighborhood pixels. Here,  $r, g, b \in [0, 255]$ , and  $n_c$  was set at 9 in this study.

(2) Calculating the mean values  $\mu_r, \mu_g$ , and  $\mu_b$  of the feature point  $(x, y)$  and its eight neighborhoods. The value  $E$  was the 3D vector composed by  $\mu_r, \mu_g$  and  $\mu_b$ . The equations were detailed as follows:

$$\begin{cases} \mu_r = \frac{1}{n} \sum_{i=1}^n r_i, \\ \mu_g = \frac{1}{n} \sum_{i=1}^n g_i, \\ \mu_b = \frac{1}{n} \sum_{i=1}^n b_i, \\ E = (\mu_r, \mu_g, \mu_b). \end{cases} \quad (14)$$

(3) Calculating the variances  $\delta_r, \delta_g$  and  $\delta_b$  of the feature point  $(x, y)$  and its eight neighborhoods as follows.  $F$  was the 3D vector composed by  $\delta_r, \delta_g$  and  $\delta_b$ .

$$\begin{cases} \delta_r = \frac{1}{n} \sum_{i=1}^n (r_i - \mu_r)^2, \\ \delta_g = \frac{1}{n} \sum_{i=1}^n (g_i - \mu_g)^2, \\ \delta_b = \frac{1}{n} \sum_{i=1}^n (b_i - \mu_b)^2, \\ F = (\delta_r, \delta_g, \delta_b). \end{cases} \quad (15)$$

(4) Normalizing vector  $E$  and  $F$ , respectively.

$$\begin{cases} E_n = \left( \frac{\mu_r}{255}, \frac{\mu_g}{255}, \frac{\mu_b}{255} \right), \\ F_n = \left( \frac{\delta_r}{255}, \frac{\delta_g}{255}, \frac{\delta_b}{255} \right). \end{cases} \quad (16)$$

(5) The normalized mean value vector  $E_n$  and variance vector  $S_n$  were combined to form a six-dimensional RGB color classification descriptor vector  $V_c$ .

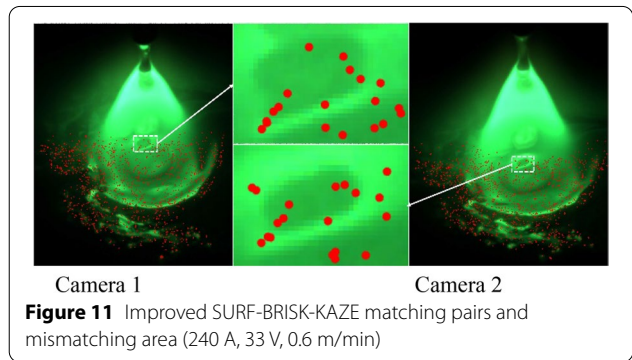
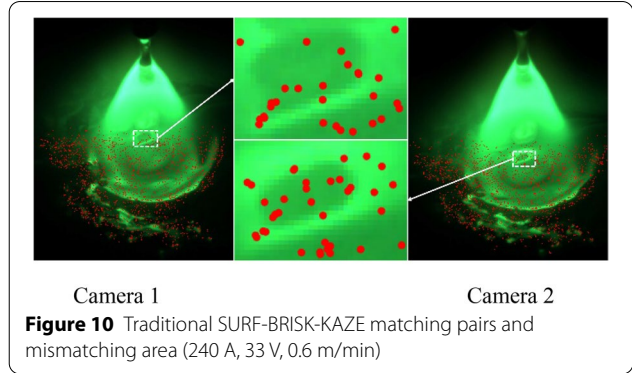
$$V_c = \left( \frac{\mu_r}{255}, \frac{\mu_g}{255}, \frac{\mu_b}{255}, \frac{\delta_r}{255}, \frac{\delta_g}{255}, \frac{\delta_b}{255} \right). \quad (17)$$

(6) By substituting Eq. (17) into  $V_s = (i_1, i_2, \dots, i_{64})$ , the descriptor  $V$  was expressed as follows:

$$V = \left( i_1, i_2, \dots, i_{64}, \frac{\mu_r}{255}, \frac{\mu_g}{255}, \frac{\mu_b}{255}, \frac{\delta_r}{255}, \frac{\delta_g}{255}, \frac{\delta_b}{255} \right). \quad (18)$$

The matching process was carried out as follows. A feature point  $P_1$  (image of camera 1) was compared with a feature point  $P_2$  (image of camera 2) by calculating the Euclidean distance between their descriptor  $V$ . The Euclidean distance is:

$$d_E = \sqrt{(i'_1 - i''_1)^2 + (i'_2 - i''_2)^2 + \dots + \left( \frac{\delta'_g}{255} - \frac{\delta''_g}{255} \right)^2 + \left( \frac{\delta'_b}{255} - \frac{\delta''_b}{255} \right)^2}, \quad (19)$$



where  $V' = (i'_1, i'_2, \dots, \delta'_g/255, \delta'_b/255)$  is the descriptor of  $P_1$  and  $V'' = (i''_1, i''_2, \dots, \delta''_g/255, \delta''_b/255)$  is the descriptor of  $P_2$ . A matching pair was detected if the ratio of their distance to the distance of the second nearest neighbor was less than 0.6.

Figure 10 shows the matching pairs obtained by the SURF-BRISK-KAZE algorithms whose descriptor was 64-dimensional. Figure 11 shows the matching pairs obtained by the improved 70-dimensional descriptor  $V$ . Though the number of mismatching pairs was decreased because of the use of  $V$ , some mismatching pairs still remained, and mismatch elimination was needed.

## 6 Improved RANSAC Algorithm

There were some mismatching pairs (gross errors) in the above matching pairs. A random sample consensus algorithm was used to eliminate the mismatches in this work. The traditional RANSAC algorithm can complete the estimation for the model parameters through data iteration and result validation. It can also reduce the number of gross errors to increase the matching accuracy

[22]. However, the traditional RANSAC algorithm also has shortcomings. For example, the iteration time of the traditional RANSAC algorithm depends on the experimental data and the valid data rate. When the number of valid data is small, the number of iterations of the algorithm will increase exponentially because of a large number of mismatching pairs, which greatly increases the running time of the algorithm. Meanwhile, the initial model parameters are calculated based on a subset of the experimental data. When the valid data rate of the subset is not high, the initial model will be extremely unreasonable, and the verification of this unreasonable model will consume a lot of time, which seriously affects the overall efficiency of the algorithm. The traditional RANSAC algorithm was improved as follows.

### 6.1 Data Preprocessing

The data pre-processing model was defined as  $M(P, C)$ .  $P$  was the matching pairs ( $P_1, P_2$ ) obtained by Section 5.4, which was written as:

$$P = \{P_1[i], P_2[j] | i = j = 0, 1, \dots, n-1\}. \quad (20)$$

$C$  was the criterion for eliminating mismatching point pairs of  $P$ , and it was expressed as:

$$C = (|k_i - k| \leq \Delta k) \& \& (|d_i - d| \leq \Delta d), \quad (21)$$

$$i = 0, 1, \dots, n-1,$$

where  $k_i$  is the ratio calculated by pixel coordinates, and  $k$  is the median of  $k_i$ .  $\Delta k$  is the threshold value, which was 0.9 in this study. If  $d_i^*$  is the square root calculated by pixel coordinates,  $d_i$  is the top 80% of  $d_i^*$  value from small to large.  $d$  is the median of  $d_i$ , and  $\Delta d$  is the threshold value, which was 13 in this study. The  $k_i$  and  $d_i$  were expressed as:

$$k_i = \frac{y_j - y_i}{x_j - x_i}, \quad (22)$$

$$d_i = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \quad (23)$$

where  $x_i$  and  $y_i$  are the pixel coordinates of  $P_1[i]$  captured by camera 1, and  $x_j$  and  $y_j$  are the pixel coordinates of  $P_2[i]$  captured by camera 2.

To increase the proportion of valid data in the data set of  $P$ , all the matching pairs were first calculated based on Eqs. (22) and (23). Any matching pair that did not satisfy Eq. (21) was deleted from the data set of  $P$ , and others were selected as a new data set of  $Q$ . The  $Q$  contained enough matching pairs, which can be used to estimate the truest homography matrix parameter model as much as possible. The proportion of valid data in the  $Q$  became larger, which greatly reduces the number of iterations to

calculate the maximum valid data. As a result, the efficiency of the algorithm was improved. The above work completed the pre-purification of the raw data for the RANSAC algorithm. Figure 12 shows the matching pairs pre-processed by the data pre-processing model. Compared to Figure 11, the number of mismatching pairs was decreased.

### 6.2 Pre-test Model

In the traditional RANSAC algorithm, the subset ( $S$ ) is randomly extracted from experimental data ( $P$ ) of Section 5.4, and the corresponding initial model is estimated by the subset. The initialization model is tested by all the remaining matching points ( $C_P S$ ), which are not belong to the above subset. This circulative progress needs much verification time because many initialization models are not reasonable, especially when the proportion of mismatching pairs is high.

To increase efficiency of the RANSAC algorithm, a pre-testing model was proposed to estimate the initial models before the traditional testing of RANSAC algorithm. The steps of the pre-testing model were as follows.

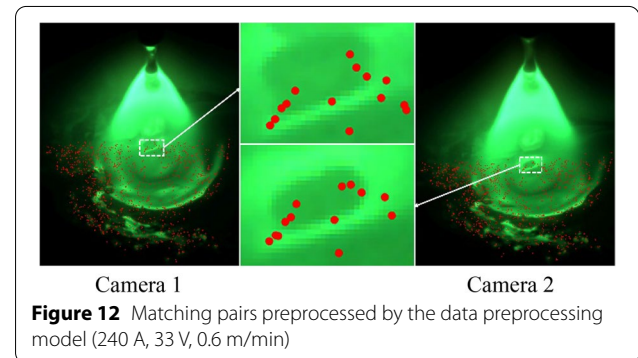
(1)  $n$  pairs of matching points were randomly selected from the new data set ( $Q$ ) of Section 6.1.

(2) The improved initial model  $T$  was estimated by  $m$  pairs of matching points, which was written as follows, where  $m \subseteq n$ :

$$T = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix}, \quad (24)$$

$$\beta \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}, \quad (25)$$

where,  $t_{11}, t_{12}, \dots, t_{33}$  are the parameters of  $T$ ,  $(x_1, y_1)$  is the pixel coordinate of  $P_1[i]$ , and  $(x_2, y_2)$  is the pixel



coordinate of  $P_2[j]$ .  $m$  is the number of matching pairs needed to solve the Eq. (25), which was 4 in this study.

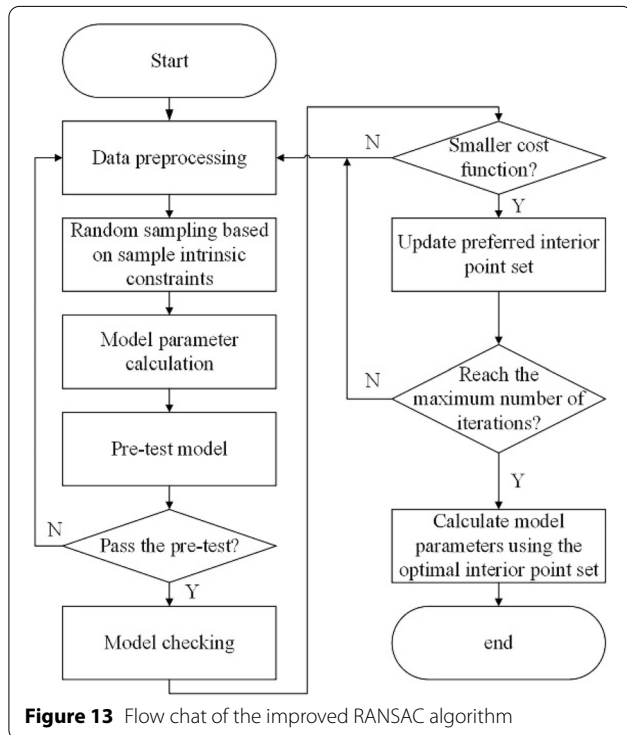
(3) The remaining  $n-m$  pairs were used to verify the reasonability of  $T$  by the following equation:

$$E_s = \sum_{i=1}^{n-m} \left[ \left( x_{1i} \frac{t_{11}x_{2i} + t_{12}y_{2i} + t_{13}}{t_{31}x_{2i} + t_{32}y_{2i} + t_{33}} \right)^2 + \left( y_{1i} \frac{t_{21}x_{2i} + t_{22}y_{2i} + t_{23}}{t_{31}x_{2i} + t_{32}y_{2i} + t_{33}} \right)^2 \right], \quad (26)$$

where  $E_s$  is test error calculated by the matching pairs of  $C_n m$ . The threshold  $t_p$  was 0.6 in this study.

(4) If the  $E_s$  values of all  $n-m$  pairs were less than  $t_p$ , the initial model  $T$  was started to be tested by the remaining matching points ( $C_Q n$ ) of Section 6.1 as the traditional RANSAC algorithm; otherwise, the initialization model  $T$  was discarded directly, and  $n$  pairs of matching points were re-selected for the next cycle of iterative estimation.

The unreasonable initial models were discarded quickly by the above calculation, which reduced the detection time to verify the initial model. A flow chart of the improved RANSAC algorithm is shown in Figure 13. The real detection time of traditional and improved RANSAC algorithms are discussed in Section 6.3. Figure 14 is the matching point pairs of the welding pool under the arc processed by the improved RANSAC algorithm.



### 6.3 Estimation of Commutating Time

To obtain at least one subset that could pass through the verification of the traditional RANSAC algorithm under a certain confidence  $p$  (0.95–0.99), the minimum sampling

times  $k$  must meet the following requirement:

$$1 - (1 - (1 - \omega)^{m_t})^k = P, \quad (27)$$

where  $\omega$  is the proportion of the mismatching pairs in  $P$ , and  $m_t$  is the minimum sample number for all subsets to calculate the initial model parameters of verification which was 4 in this study.

Supposing that the number of  $P$  is  $n_p$ , it takes  $t_s$  seconds to conduct a random sampling from the data set, and it takes  $t_c$  seconds to calculate the initial model.  $t_j$  is the time to test the initial model by one couple of remaining matching pairs. The total time to test the initial model is  $(n_p - 4)t_j$ . Therefore, the calculation time  $T_t$  of traditional RANSAC algorithm with  $k$  random sampling times is:

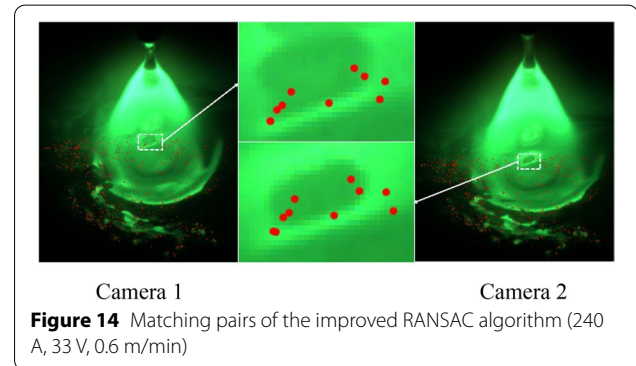
$$T_t = k(t_s + t_c) + k(n_p - m_t)t_j. \quad (28)$$

In this study, the average  $k$  was 500,  $t_s$  was  $2.21 \times 10^{-4}$  s,  $t_c$  was  $5.45 \times 10^{-5}$  s, and  $t_j$  was  $3.67 \times 10^{-5}$  s. The computer processor used was an Intel(R) Core(TM) i7-9750H CPU @ 2.60 GHz with a memory of 8 GB.

For the improved RANSAC algorithm, the minimum sampling times  $k'$  must meet Eq. (27) as well, where  $m_t$  is  $n$  mentioned in Section 6.2. Eq. (28) is modified as:

$$T' = k'(t_s + t_c) + k'(n - m)t_j + k'(n_Q - n)(1 - \omega_Q)^n t_j, \quad (29)$$

where  $T'$  is the calculation time of the improved RANSAC algorithm,  $n_Q$  is the number of  $Q$ , and  $\omega_Q$  is the proportion of the mismatching pairs in  $Q$ .



**Table 3** Comparison between the improved RANSAC algorithm and the traditional RANSAC algorithm

Image group	Total feature points	Traditional RANSAC			Improved RANSAC			Time difference $\Delta T$ (s)	Increasing efficiency (%)
		Interior points	Rate (%)	Algorithm time (s)	Interior points	Rate (%)	Algorithm time (s)		
1	5900	638	10.81	43.77	858	14.54	12.54	31.23	249.04
2	6501	640	9.84	55.55	790	12.15	18.78	36.77	195.79
3	5817	616	10.59	54.05	764	13.13	19.87	34.18	172.02
4	6434	650	10.10	53.69	771	11.98	19.32	34.37	177.90
5	4698	549	11.69	52.13	701	14.92	19.82	32.31	163.02
6	5454	613	11.24	52.71	744	13.64	19.37	33.34	172.12
7	5472	648	11.84	54.64	973	17.78	20.40	34.24	167.84
8	5650	600	10.62	54.30	943	16.69	19.88	34.42	173.14

The time reduced by the improved RANSAC algorithm is as follows:

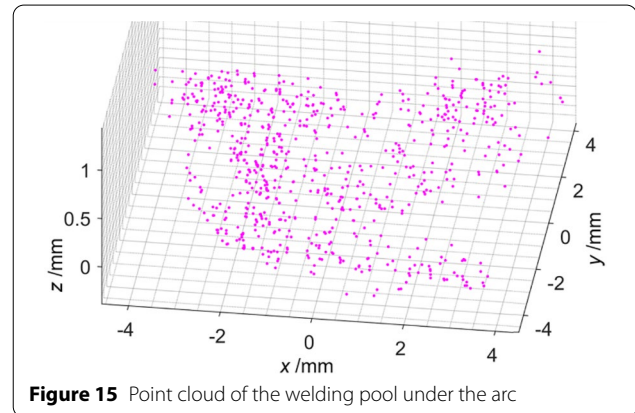
$$\Delta T = T_t - T' = (k - k')(t_s + t_c) + [k(n_p - m_t) - k'(n - m) - k'(n_Q - n)(1 - \omega_Q)^n] t_j. \quad (30)$$

In this study,  $k'$  was 1500. Though the total time of  $t_s$  and  $t_c$  was increased if  $k'$  was larger than  $k$ , it was an order of magnitude smaller than the time for testing the initial model. Therefore, the improved RANSAC algorithm reduced a series of unnecessary operations and improved the efficiency of the algorithm. While its calculation accuracy was not smaller than that of the traditional RANSAC algorithm, the values of  $k'$ ,  $n$ , and  $m$  were not optimal yet and can be optimized in the future.

Compared to the traditional RANSAC algorithm, the calculating efficiency in the improved RANSAC algorithm was increased by at least 160% in this study, as shown in Table 3, which showed a large potential to reduce the calculating time.

## 7 Reconstruction Results and Validation

The GMAW experiments were carried out to capture pairs of the welding pool images under the arc. The parameters of the welding process are shown in Table 2. The feature-matching points on each pair of welding pool images were extracted according to Sections 5 and 6. The world coordinates of the welding pool surface were reconstructed according to Section 3. The results for 240 A and 33 V are shown in Figure 15.

**Figure 15** Point cloud of the welding pool under the arc

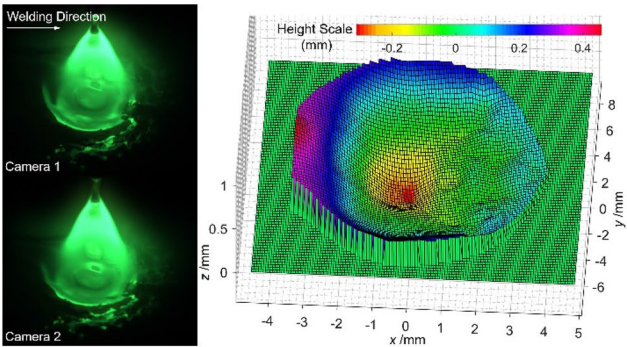
### 7.1 Surface Reconstruction

The way to transfer the point cloud into the 3D reconstructed surface was through the LOWESS regression smoothing filtering algorithm, which was a local scatter weighted estimation algorithm depending on the adjacent points [23]. The algorithm added a sliding window on the basis of the least square method. If the sliding window width was large, there would be more scattered points in the covered window, resulting in a smooth surface of welding pool and loss of large original data information. On the contrary, if the sliding window width was small, the welding pool surface was rough, which also increased the accuracy of reconstructed surface.

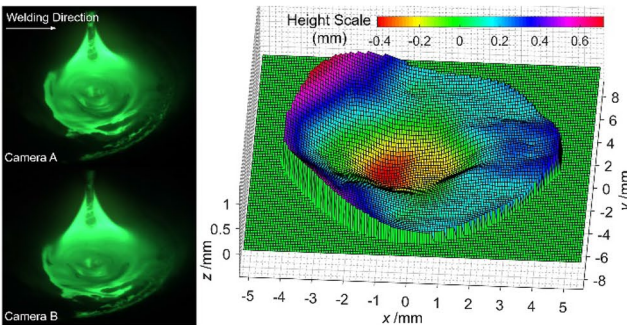
In this experiment, the window width was selected as 6 pixels based on the comparative tests, which provided high accuracy for the 3D reconstructed data and

(See figure on next page.)

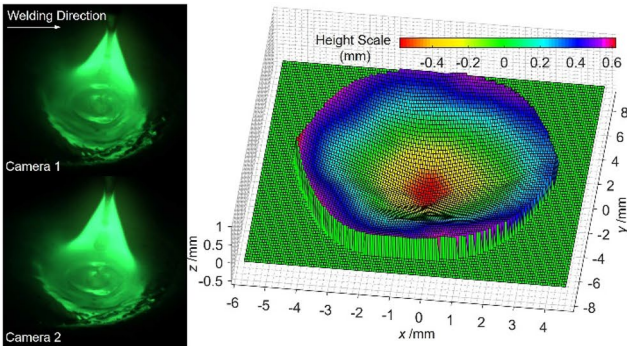
**Figure 16** 3D reconstruction of the welding pool under the arc



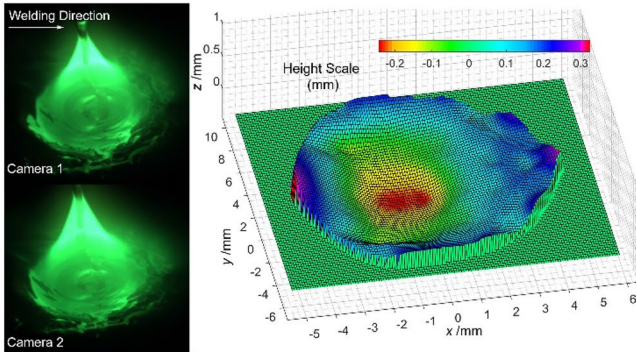
a 240 A, 33 V, 0.6 m/min



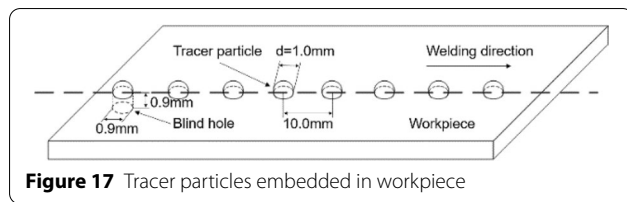
b 240 A, 33 V, 0.6 m/min



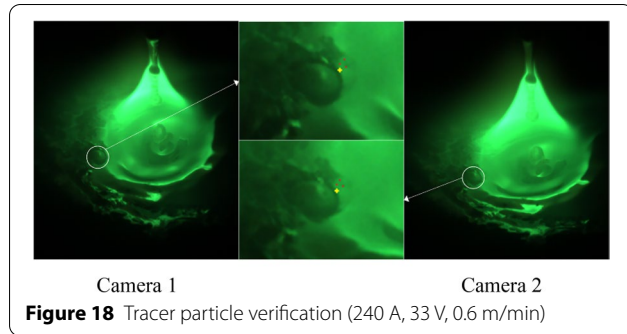
c 280 A, 33 V, 0.6 m/min



d 260 A, 33 V, 0.8 m/min



**Figure 17** Tracer particles embedded in workpiece



**Figure 18** Tracer particle verification (240 A, 33 V, 0.6 m/min)

produced a smoothly reconstructed surface without trimming. The point with the smallest  $z$  value in the point cloud was selected as the center of the welding pool. It is assumed that the world coordinate of the point was  $(x_0, y_0, z_0)$ , and the final surface reconstruction was the symmetric treatment of the smoothly reconstructed surface in the plane where  $y=y_0$ . The final surface reconstruction of Figure 15 is shown in Figure 16. It can be seen that the reconstructed welding pool surface was consistent with the actual welding pool surface in GMAW.

## 7.2 Validation

In order to verify the validity of the above feature point extracting and matching algorithms for welding pool images pairs during GMAW, a verification experiment was carried out. The diameter of silicon nitride tracer particle was a little smaller than that of blind hole, so the tracer particle could be embedded tightly in the

workpiece as shown in the Figure 17. These tracer particles would float on the surface of the welding pool when the solid metal near the blind hole was melted. This phenomenon was captured by the above binocular vision system, as shown in Figure 18. The feature points at the junction of the tracer particle and the welding pool can be selected manually, as shown in Figure 18 (yellow dots). It was clear that the yellow dots were the correct feature matching points. Their world coordinate values, calculated directly by Section 3, can be used to verify the accuracy of their neighbor feature matching points (red dots in Figure 18) extracted by the above improved algorithms. Here, the heights of the yellow dots and red dots were assumed to be similar as they were close to each other on the smooth surface of the welding pool, though their world coordinates in the  $x$  and  $y$  directions were different.

Table 4 provides a comparison between the heights of the yellow dots and red dots. The maximum absolute error in the  $z$  direction was less than 0.07 mm, and the maximum relative error was smaller than 6.0%. All the results indicated that the proposed algorithms mentioned above can reconstruct the 3D surface of welding pool for GMAW with high efficiency.

## 8 Conclusions

- (1) A new method was proposed to determine the feature matching of welding pool captured by a binocular stereo vision system. It included improved SURF-BRISK-KAZE algorithms, improved RANSAC algorithms, and surface reconstruction algorithms, which realized the 3D reconstruction of the welding pool surface under the GMAW arc.
- (2) The feature points descriptor was improved by considering the color information of images, which increased the matching accuracy of welding pool

**Table 4** Comparison between the tracer particle and the matching point pairs

Image group	Tracer particles (mm)	Matching point pair 1 (mm)	Matching point pair 2 (mm)	Matching point pair 3 (mm)	Maximum abs (error) (mm)	Maximum relative error (%)
1	-1.31727	-1.36893	-1.33676	-1.33862	0.05136	3.92
2	-1.32219	-1.29371	-1.33483	-1.37148	0.05199	3.78
3	-1.19575	-1.23198	-1.23604	-1.12156	0.06419	5.37
4	-1.26499	-1.28316	-1.32741	-1.25829	0.06243	4.93
5	-1.17624	-1.16843	-1.21358	-1.16923	0.03736	3.17
6	-1.42836	-1.47268	-1.43864	-1.43269	0.04432	3.10
7	-1.27823	-1.29906	-1.24084	-1.24821	0.03748	2.93
8	-1.36529	-1.32842	-1.31868	-1.35937	0.04863	3.59

images. A data preprocessing and pre-test model were added to the RANSAC algorithm to improve the calculating efficiency of the algorithm.

- (3) The experimental results exhibited high accuracy and efficiency of the new method in reconstruction of welding pool surface. For the reconstructed data, the maximum relative error was smaller than 6.0%.
- (4) The quantitative relationship between welding pool surface and weld quality should be established based on the experience of a skilled welder, and the parallel computing method should be improved by the CUDA to realize the real-time 3D reconstruction of welding pool surface in the future.

#### Acknowledgements

Not applicable.

#### Authors' contributions

JC was in charge of the whole research; ZG and JC wrote the manuscript; CW assisted with sampling and laboratory analyses. All authors read and approved the final manuscript.

#### Authors' information

Zunan Gu, born in 1995, is currently a master candidate at Key Laboratory for Liquid-Solid Structural Evolution & Processing of Materials, Ministry of Education, Institute of Materials Joining, Shandong University, China.

Ji Chen, born in 1982, is currently a professor at Key Laboratory for Liquid-Solid Structural Evolution & Processing of Materials, Ministry of Education, Institute of Materials Joining, Shandong University, China. He received his PhD degree from Shandong University, China, in 2009. His research interests include intelligent control in welding processes and numerical simulation.

Chuansong Wu, born in 1959, is currently a professor at Key Laboratory for Liquid-Solid Structural Evolution & Processing of Materials, Ministry of Education, Institute of Materials Joining, Shandong University, China. He received his PhD degree from Harbin Institute of Technology, China, in 1988. His main research interests include intelligent control in welding processes, solid-state welding, and numerical simulation.

#### Funding

Supported by National Natural Science Foundation of China (Grant No. 51775313), Major Program of Shandong Province Natural Science Foundation (Grant No. ZR2018ZC1760), and Young Scholars Program of Shandong University (Grant No. 2017WLJH24).

#### Competing interests

The authors declare no competing financial interests.

Received: 28 March 2020 Revised: 29 March 2021 Accepted: 8 May 2021  
Published online: 27 May 2021

#### References

- [1] Y Liu, Y Zhang. Control of 3D welding pool surface. *Control Engineering Practice*, 2013, 21(11): 1469-1480.

- [2] B C Wang, S J Hu, L Sun, et al. Intelligent welding system technologies: State-of-the-art review and perspectives. *Journal of Manufacturing Systems*, 2020, 56: 373-391.
- [3] L Caprio, A G Demir, B Previtali. Observing molten pool surface oscillations during keyhole processing in laser powder bed fusion as a novel method to estimate the penetration depth. *Additive Manufacturing*, 2020, 36: 101470.
- [4] K R Thomas, S Unnikrishnakurup, P V Nithin, et al. Online monitoring of cold metal transfer (CMT) process using infrared thermography. *Quantitative Infrared Thermography Journal*, 2017, 14(1): 68-78.
- [5] M Vasilev, C MacLeod, Y Javadi, et al. Feed forward control of welding process parameters through on-line ultrasonic thickness measurement. *Journal of Manufacturing Processes*, 2021, 64: 576-584.
- [6] R Zong, J Chen, C S Wu, et al. Undercutting formation mechanism in gas metal arc welding. *Welding Journal*, 2016, 95: 174s-184s.
- [7] L Yang, E Li, T Long, et al. A welding quality detection method for arc welding robot based on 3D reconstruction with SFS algorithm. *International Journal of Advanced Manufacturing Technology*, 2018, 94: 1209-1220.
- [8] J Chen, Z L Feng. Strain and distortion monitoring during arc welding by 3D digital image correlation. *Science and Technology of Welding and Joining*, 2018, 23(6): 536-542.
- [9] W J Zhang, J Xiao, Y M Zhang. A mobile sensing system for real-time 3D welding pool surface measurement in manual GTAW. *Measurement Science and Technology*, 2016, 27: 045102.
- [10] J K Huang, W Pan, J S Chen, et al. The transient behaviours of free surface in a fully penetrated welding pool in gas tungsten arc welding. *Journal of Manufacturing Process*, 2018, 36: 405-416.
- [11] Z Chen, J Chen, Z Feng. 3D welding pool surface geometry measurement with adaptive passive vision images. *Welding Journal*, 2019, 98: 379s-386s.
- [12] J Zhong, C Yang, Y Xu, et al. Research on reconstruction of welding pool surface based on shape from shading during robot aluminum alloy pulse GTAW. *Advances in Intelligent Systems and Computing*, 2015, 363: 525-538.
- [13] C Mnich, F Al Bayat, C Debrunner, et al. In situ welding pool measurement using stereovision. *Japan-USA Symposium on Flexible Automation*, Denver, USA, 2004.
- [14] Z Liang, H Chang, Q Wang, et al. 3D reconstruction of welding pool surface in pulsed GMAW by passive biprism stereo vision. *IEEE Robotics and Automation Letters*, 2019, 4(3): 3091-3097.
- [15] J Xiong, Y Liu, Z Yin. Passive vision measurement for robust reconstruction of molten pool in wire and arc additive manufacturing. *Measurement*, 2020, 153:107407.
- [16] L Wang, J Chen, X H Fan, et al. Influence of fluid flow on humping bead during high-speed GMAW. *Welding Journal*, 2019: 315-327.
- [17] R Mehrotra, S M Zhan. Computational approach to zero-crossing-based two-dimensional edge detection. *Graphical Models and Image Processing*, 1996, 58(1): 1-17.
- [18] T Janumala, K B Ramesh. Development of an algorithm for vertebrae identification using speeded up robust features (SURF) technique in scoliosis x-ray images. *Image Processing and Capsule Networks(ICIPCN 2020)*, Bangkok, Thailand, May 6-7, 2020: 54-62.
- [19] A Khare, B R Mounika, M Khare. Keyframe extraction using binary robust invariant scalable keypoint features. *Twelfth International Conference on Machine Vision (ICMV 2019)*, Amsterdam, Netherlands, Nov 16-18, 2019: UNSP 1143308.
- [20] B Ramkumar, R Laber, H Bojinov, et al. GPU acceleration of the KAZE image feature extraction algorithm. *Journal of Real-Time Image Processing*, 2020, 17(5): 1169-1182.
- [21] H C Shih, C H Ma, C L Lu. An efficient fragment reconstruction using RANSAC algorithm. *2019 IEEE 8th Global Conference on Consumer Electronics*, Osaka, Japan, Oct. 15-18, 2019: 529-530.
- [22] W S Cleveland. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 1979, 74(368): 829-836.