*Research Article*

# Robust Locally Weighted Regression for Profile Measurement of Magnesium Alloy Tube in Hot Bending Process

**Zhibin Yao** ,[1,2] **Rongjun Wang,[1,2] Jinning Zhi,[1,2] and Qinglu Shi** [1,2]

[1]*Heavy Machinery Engineering Research Center of the Ministry Education, Taiyuan 030024, Shanxi, China*
[2]*College of Mechanical Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, Shanxi, China*

Correspondence should be addressed to Qinglu Shi; 327606120@qq.com

Section flattening often occurs in the hot bending process of magnesium alloy tube with large curvature. In order to control the forming quality of the tube, it is necessary to measure the section profile of the magnesium alloy pipe online. In this paper, the laser vision system is used to measure the profile of magnesium alloy tube. Due to the influence of the environment and the surface quality of the pipe, there are obviously isolated outliers in the profile data, which seriously affects the accuracy and precision of the tube measurement. An outlier identification algorithm based on robust locally weighted regression and РайТа criterion is proposed. This algorithm is used to identify the typically isolated outliers in the measurement process and discuss its identification ability. Meanwhile, it is compared with the moving mean identifier and the Hampel identifier. Subsequently, the ellipse fitting of profile data was carried out, and the fitting ellipse parameters and fitting precision of the curved section were obtained. At the same time, the fitting results were compared before and after the outliers are eliminated. The experiment proves that the outlier identification method based on robust locally weighted regression and РайТа criterion can effectively identify outliers in profile data, especially for spot outliers. This algorithm is a robust, accurate, and efficient outlier identification method, which can effectively improve the laser profile measurement accuracy of the pipe section and has great significance for the quality control of magnesium alloy tube.

## 1. Introduction

Magnesium alloy has the advantages of low density and high specific strength and specific stiffness, so it is gradually receiving attention from various industries [1–4]. In particular, various types of magnesium alloy tubes are widely used in aviation, aerospace, transportation, and other fields [5–7]. However, various defects often occur in the hot bending forming process of magnesium alloy tubes with large curvatures, such as section flattening, excessive thinning, and wrinkle fracture [8–11]. Section flattening is an unavoidable problem in magnesium alloy tubes with large curvature. As shown in Figure 1, the section flattening of the magnesium alloy tube is mainly due to the action of bending moments $M_1$ and $M_2$, the compression stress $F_c$ on the inner side, and the tensile stress $F_t$ on the outer side, and the pipe is bent due to the uneven force on both sides. As a

result, the section is deformed to become approximately elliptical. For the same specification, the smaller the bending radius, the larger $F_t$ and $F_c$ and the more obvious the trend of section flattening. If it is a bending without a mandrel, the flattening is more serious, which affects the forming quality of the pipe. In order to control the forming quality of the tube, the section profile must be measured online, and the flattening of the magnesium alloy section must be detected in real time.

Pipe profile measurement generally uses contact measurement and noncontact measurement [12–15]. Machine vision is an important method in noncontact measurement. Compared with traditional manual contact measurement methods, machine vision measurement has high accuracy and fast speed and causes no damage [16–19]. In this study, a laser vision system based on line structure light was used to measure the flattening section of magnesium alloy tube
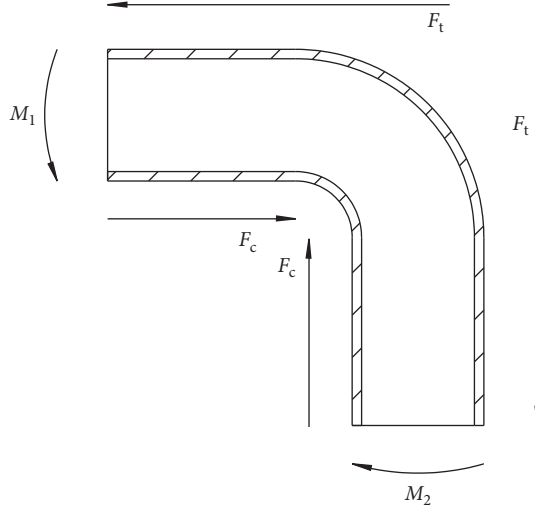
Figure 1: Force condition of bending process.

caused by the large curvature hot bending to obtain the accurate profile data of the magnesium alloy tube section.

The online vision measurement of tube bending is strongly demanded to improve the dimensional accuracy. However, the quality of the profile data deteriorates sharply due to the environment and surface quality of the pipe. In particular, many isolated outliers appear in the measured data. Outliers generally mean that the observed values are significantly different from most observed values, that is, the measured data do not obey the statistical distribution law of the data [20]. In this experiment, the outliers appear in isolation as the form of dot or point, which are not necessarily related to the data quality before and after. Therefore, they are also called isolated outliers. Outliers can be generated for a variety of reasons, such as sensor noise, channel interference, and human factors [12, 21]. These outliers will cause data distortion, leading to the miscalculation of model parameters and wrong analysis results. Therefore, it is necessary to identify and eliminate the outliers in the laser measurement process to improve the precision and accuracy of the measurement.

Commonly used identification methods for outliers include Nair, Grubbs, Dixon, etc [22, 23]. These methods are difficult to apply in laser profile measurement due to the limitation of data distribution and data amount. This paper presents an outlier identification method based on robust locally weighted regression (RLWR), which is applied to laser profile detection. RLWR is developed from locally weighted regression (LWR) and belongs to nonparametric estimation. Nonparametric estimation is an important research direction of modern statistical analysis. It can adapt to more complex nonlinear changes without assuming the specific form of population distribution and error distribution and without directly obtaining data models, which is more flexible, robust, and widely applied than parameter estimation [24, 25].

RLWR is a robust fitting process that integrates local polynomial estimation and locally weighted regression with excellent smoothing performance. This method was first proposed by Cleveland [26] and further elaborated [27, 28], subsequently improved by Jacoby [29] and Loader [30]. It has been gradually applied to different fields of scientific research and engineering applications. Ma et al. [31] used the RLWR to reduce the impact of high-frequency noise on superresolution enhancement of multiangle remote sensing imagery. Leonor et al. [32] minimized the influence of the inhomogeneity effect on tree reradiation pattern by using RLWR. Chen et al. [33] proposed an algorithm based on RLWR and robust z-scores for the construction of a pit-free canopy height models. Nurunnabi et al. [34, 35] and Liu et al. [36] used RLWR techniques to study the filtering of 3D ground cloud point data. Yu et al. [37] applied the method to the smoothing of combustion kinetics data of pine sawdust biochar.

This study is based on the robust smoothing performance of RLWR, which is used to smooth the laser measurement data and realize the identification of isolated outliers. At the same time, profile fitting was carried out to verify the effectiveness of this algorithm in the profile inspection.

This paper organized as follows. Section 2 introduces the laser vision measurement system and its basic principles and analyzes the profile data of magnesium alloy tube, pointing out the typical outliers in the data. In Section 3, the moving average algorithm, Hampel algorithm, and RLWR algorithm are, respectively, adopted to identify the outliers. It focuses on the outlier recognition effect for the RLWR combined with РайТа criterion, median criterion, and quartile criterion, respectively. In Section 4, elliptic fitting of profile data is performed by using the RLWR identification algorithm, and relevant elliptic parameters and fitting error are obtained, at the same time, compared with original data. Lastly, a conclusion is given in Section 5.

## 2. Laser Vision Measurement System and Profile Outliers

*2.1. Laser Vision Measurement System.* The laser vision measurement system adopted in this experiment mainly consists of a line-structured light sensor, image processing unit, and control unit, as shown in Figure 2. The system works as follows: the laser source in the sensor projects line structure light to a measured object. The reflected light is captured by the camera in the sensor. The profile data are recorded and displayed by industrial PC after image processing. The measuring platform can move along the axial direction of the pipe, which is controlled by the motion control system, realizing the profile measurement of each section for the pipe.

Laser triangulation is the basic principle of line-structured light vision measurement. The line-structured light projected by laser source forms a scattered light band on the pipe surface, and the scattered light is imaged in CCD array. The distance and coordinate data of the measured point can be obtained through triangular geometry [38, 39].

As shown in Figure 3, the point $p(x, y, z)$ is a measurement point on pipe surface, point $p\prime(u, v)$ is the imaging point of $p$ in CCD image plane, $f$ is the focal length of
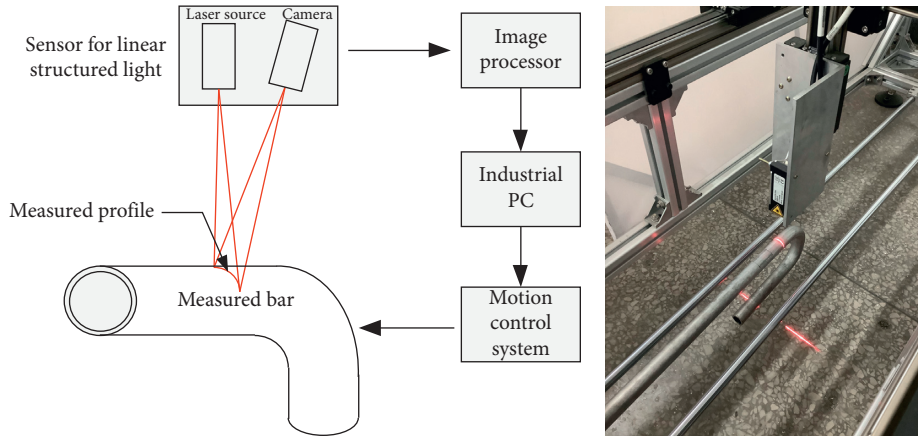
FIGURE 2: The process of laser vision measurement.
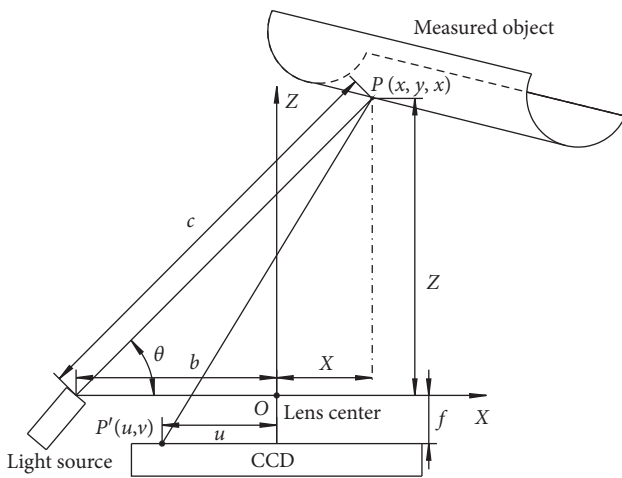


FIGURE 3: Principle of triangulation measurement.

camera, $b$ is the distance between the center of light source and the center of camera lens, and $\theta$ is the angle between the $X$ axis and the construction line, which is formed by the measured point and the center of light source. The precise coordinates of $p(x, y, z)$ on the measured profile can be obtained from the spatial geometry in the figure.

*2.2. Typical Profile Outliers.* Observe the profile data of each section obtained by the measuring system. There are two typical isolated outliers in the measured profile, that is, the spot outliers and the point outliers. The spot outliers or speckle outliers are composed of multiple point outliers, as shown on the right side of Figure 4. The double-point isolated outliers are shown in the middle of Figure 5.

In order to facilitate analysis, as shown in Figure 6, the double-point outliers in Figure 5 are superimposed in Figure 4. Subsequently, the recognition ability of different identification algorithm is investigated for two types of outliers.

## 3. Identification of Profile Outliers

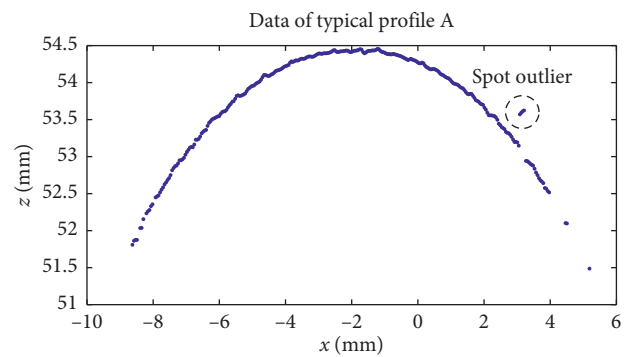*3.1. Moving Mean Identifier Recognizes Isolated Outliers.*
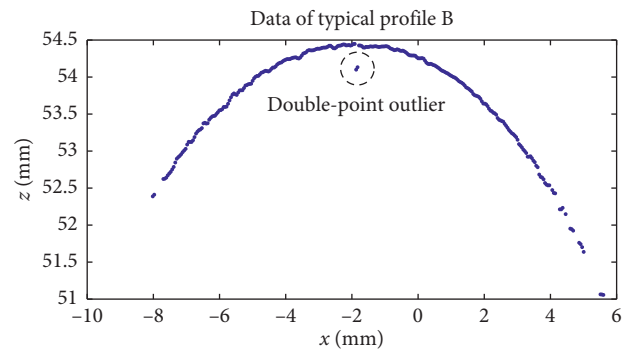


FIGURE 4: Spot outlier in profile A.



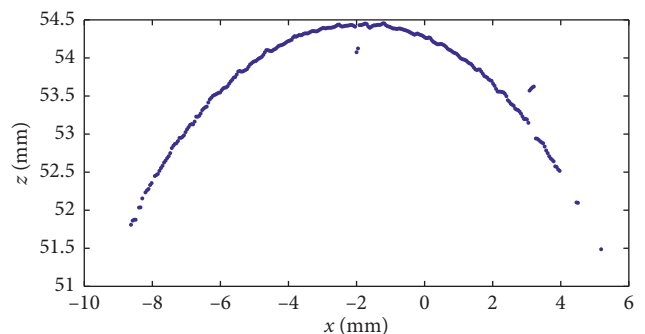FIGURE 5: Double-point outlier in profile B.



FIGURE 6: Typical profile after superposition of outliers.

Moving mean identifier or movmean identifier is the easiest way to identify isolated outliers [40, 41]. The basic definition is as follows. For the data sequence $x_1, x_2, x_3, \ldots, x_n$, moving window length is $wl$, and the outliers are judged by the following equation:

$$y_i = \begin{cases} x_i, & |x_i - \mu_i| \leq 3\sigma_i, \\ \text{outlier}, & |x_i - \mu_i| > 3\sigma_i, \end{cases} \tag{1}$$

where $\mu_i$ is the local mean and $\sigma_i$ represents local standard deviation (LSD) within the moving window. The moving mean identifier is to use the $3\sigma$ criterion to judge the outliers in the data window when the length is $wl$. When the difference between the measured value and the local mean is greater than three times the local standard deviation, it is considered as an outlier.

The moving mean identifier is used to identify outliers in Figure 6, and the recognition results are shown in Figures 7 and 8.

The moving window length is gradually increased. When the window length increases to 19, only one point outlier is identified, as shown in Figure 7. When the window length is increased to 25, the threshold range is reduced in the middle of the profile so that the double-point outliers in the middle could be completely identified (see Figure 8). However, the spot outliers on the right are never identified regardless of the length of the moving window.

It can be seen from Figures 7 and 8 that the distinguishing threshold at the location of outliers is greatly increased due to the existence of outliers in the middle of the profile. With the increase of window length, the threshold curve in the middle of the profile is gradually smooth and the threshold range is gradually reduced. However, the upper and lower threshold ranges on both sides of the profile increase significantly. It indicates that the identification ability of outliers on both sides of the profile decreases with the increase of window width.

### 3.2. Hampel Identifier Recognizes Isolated Outliers.
Hampel identifier is a median identification method, which uses the median and absolute median deviation as a robust estimation of the location and distribution of outliers, with good robustness [42–45].

Hampel identifier is defined as follows: for data sequences $x_1, x_2, x_3, \ldots, x_n$, the number of neighbors on either side of is $k$; then, the moving window length is $2k + 1$, and the local median is $m_i$.

$$m_i = \text{median}(x_{i-k}, x_{i-k+1}, x_{i-k+2}, \ldots, x_i, \ldots, x_{i+k-2}, \\ x_{i+k-1}, x_{i+k}). \tag{2}$$

The local scaled median estimated deviation ($\text{MAD}_l$) is expressed as follows:

$$\text{MAD}_l = k \cdot \text{median}(|x_{i-k} - m_i|, \ldots, |x_{i+k} - m_i|), \tag{3}$$

where $k = 1/(\sqrt{2}\,erfc^{-1}(1/2)) \approx 1.4826$, which is the unbiased estimation of the Gaussian distribution.
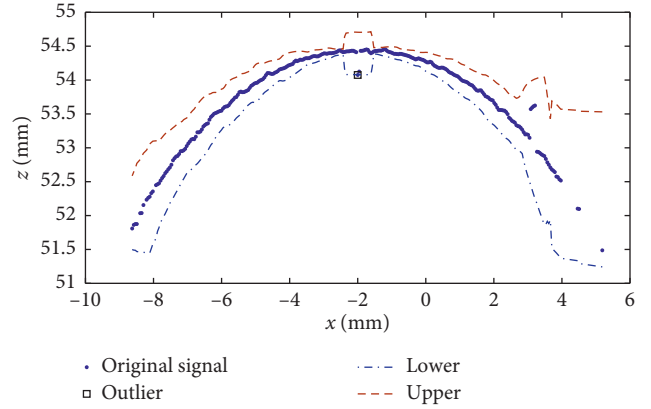


FIGURE 7: Outlier recognition using movmean identifier, window length = 19.
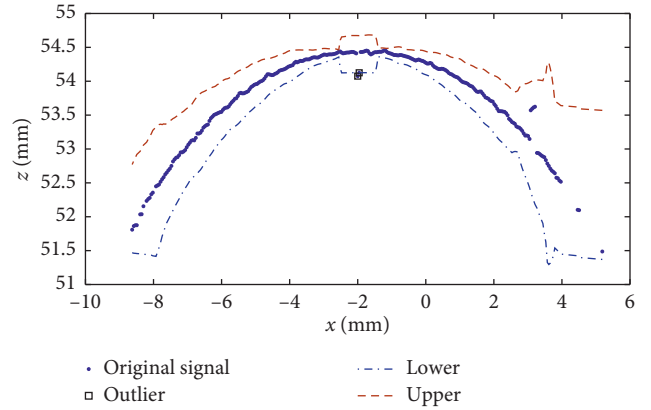


FIGURE 8: Outlier recognition using movmean identifier, window length = 25.

$$y_i = \begin{cases} x_i, & |x_i - m_i| \leq t\text{MAD}_l, \\ \text{outlier}, & |x_i - m_i| > t\text{MAD}_l. \end{cases} \tag{4}$$

When the difference between the measured data and the local median is greater than $t$ times $\text{MAD}_l$, the measured value is considered to be an outlier, as shown in equation (4).

Hampel identifier is used to recognize the outliers in Figure 6. The identification effect is observed by changing the length of moving window when $t = 3$. As shown in Figure 9, when the window length is 5, the threshold range of identification is relatively narrow. Hampel identifier was able to identify the double-point outlier in the middle of the profile, but it is unable to identify the spot outliers on the right side. At the same time, the identification threshold fluctuates significantly with the spot outliers. Furthermore, the misidentification of outliers is observed. Besides, influenced by the discontinuous data on both sides of the profile, the identification threshold fluctuated greatly and the endpoint data are identified as outliers.

When the window length is increased to 25, as shown in Figure 10, the upper and lower identification threshold on both sides of the profile increased significantly and the central double point outliers can be effectively identified, but
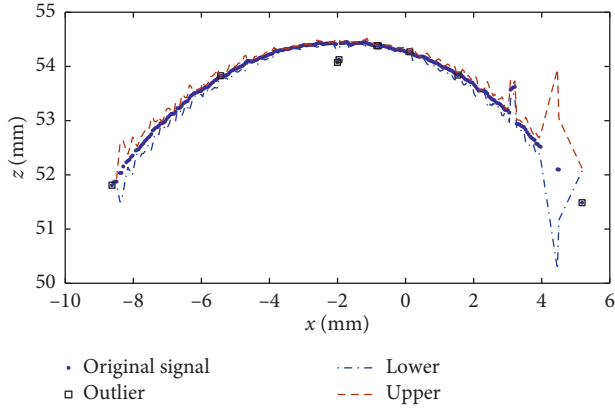
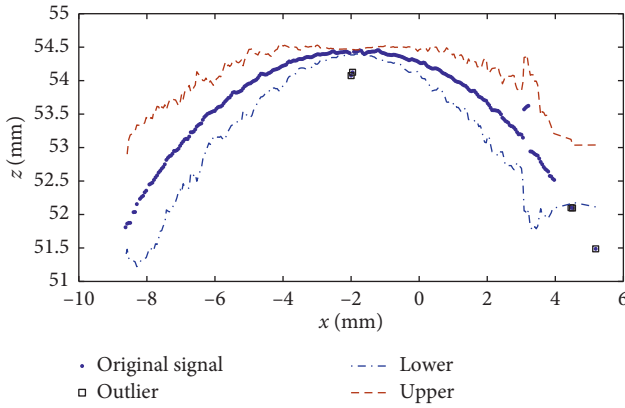FIGURE 9: Outlier recognition using Hampel identifier, window length = 5.



FIGURE 10: Outlier recognition using Hampel identifier, window length = 25.

the spot outliers on the right side of the profile are unrecognized. It is worth noting that the misidentification in Figure 9 no longer appeared. Meanwhile, several discontinuous data at the right end of the profile are identified as outliers.

If the moving window length continues to increase, the identification threshold on both sides of the profile will increase accordingly. Although the double-point outliers can be identified, the spot outliers are still unrecognizable.

Compared with the moving mean identifier, the Hampel identifier significantly reduces the threshold range and the fluctuation phenomenon, which can effectively identify point outliers, but it is still unable to identify spot outliers. In addition, the large interval of profile data will increase the difference between the data and the median, resulting in a large fluctuation for threshold, which affects its ability to identify outliers.

### 3.3. Recognition of Isolated Outliers Based on RLWR. In this section, the RLWR algorithm is combined with РайТа criterion, median criterion, and quartile criterion to identify the isolated outliers and choose the appropriate smoothing window length and observe its identification effect.

### 3.3.1. Identification of Outliers Based on RLWR and РайТа Criterion. The RLWR smoothing algorithm is combined with the РайТа criterion, i.e., $3\sigma$ criterion, to identify the laser profile outliers; this algorithm can be referred to as the RLWRP identifier. The basic approach of this identifier is to smooth the profile data by using the RLWR algorithm firstly, and then the residual between smoothing data and original data is calculated; finally, we use the РайТа criterion to identify outliers.

The algorithm of RLWRP identifier is as follows.

The measured data sequence is $\{x_i, y_i\}, i = 1, 2, \ldots, n.$ The data model assumes the following:

$$y_i = g(x_i) + \varepsilon_i, \tag{5}$$

where $g(x_i)$ is a smooth function of $x_i$ and $\varepsilon_i$ is independent and normally distributed with zero mean and variance.

Set the smoothing coefficient as $f$, where $0 < f \leq 1$, and round $f \cdot n$ to get the data width $r$, $r = [f \cdot n]$. Taking each observation point $x_i$ as the center, select the appropriate $f$ to determine the smoothing window length $wl$, $wl = x_i \pm r$.

Subsequently, a weight function is selected for the locally weighted regression (LWR). LWR typically uses tricube weight function $W(x)$ for weighted least squares fit, defined as

$$W(x_i) = \begin{cases} \left[1 - \left(\dfrac{d(x_i, x_j)}{\max_{j \in N(x_i)} d(x_i, x_j)}\right)^3\right]^3, & j \in N(x_i), \\ \\ 0, & j \in N(x_i), \end{cases} \tag{6}$$

where $N(x_i)$ is local neighborhood in smoothing window which is closest to $x_i$ and $d(x_i, x_j)$ is the distance between $x_i$ and $x_j$ in smoothing window. The value of $W(x_i)$ is a maximum for the point closest to $x_i$ and reduces to zero for the point farthest to $x_i$ in smoothing window.

Use weighted least squares method get estimates of parameters. The parameters estimates of equation (5) are the values of the parameters that minimize

$$\sum_{i=1}^{n} W_i(x)(y_i - g(x_i))^2. \tag{7}$$

The coefficients from each local neighborhood are used to estimate the fitted values $\widehat{g}(x_i)$ at $x_i$.

Generally, the RLWR selects the bisquare weight function $Q(z)$ as follows:

$$Q(z) = \begin{cases} \left(1 - |z|^2\right)^2, & |z| < 1, \\ 0, & |z| \geq 1, \end{cases} \tag{8}$$

where $z_i = e_i/(6 \cdot s)$, in which $e_i$ is the fitting residual, i.e., $e_i = y_i - \widehat{y}_i$. $s = \text{Median}(|e_i|)$, in which $s$ is the median of $|e_i|$.

Replace $W(x_i)$ with $Q(z_i) \cdot W(x_i)$ as new bisquare weight, which is used to estimate the new set of RLWR coefficients by minimizing the error sum of squares:

$$\sum_{i=1}^{n} Q(z_i) W(x_i)(y_i - g(x_i))^2. \tag{9}$$

The new RLWR fitting value $\widehat{y}_i$ is calculated using weighted least squares method. Repeat the above steps of

robust enhancement, and the final robust locally weighted fitting value is obtained.

Next, the outliers are identified according to the $3\sigma$ criterion. The residual $g_i$ is calculated by using the smoothing value obtained by the RLWR algorithm and the original measurement data, and then we get the mean of residual, i.e., $\mu = 1/n \sum_{i=1}^{n} g_i$. Finally, the standard deviation $\sigma$ is obtained:

$$\sigma = \left[ \frac{\sum (g_i - \mu)^2}{n - 1} \right]^{1/2}. \qquad (10)$$

When the difference between the residual and the mean is greater than three times the local standard deviation, it is considered to be an outlier, as shown in the following equation:

$$y_i = \begin{cases} \text{normal data}, & |g_i - \mu| \leq 3\sigma, \\ \text{outlier}, & |g_i - \mu| > 3\sigma. \end{cases} \qquad (11)$$

RLWRP identifier is used to identify the outliers in Figure 6, and its identification ability is observed under different smoothing windows. The length of smoothing window increases gradually from 5. This method shows good recognition effect when the length of smoothing window increases to 11.

As shown in Figure 11, the blue dot is the original data and the red dash dot line is the RLWR smoothing curve. The RLWR smoothing curve retains the characteristics of original profile without the risk of excessive smoothing. In Figure 12, the distribution trend of the original data is removed, which is obtained by using the residual between the smoothing data and the original data, and the outliers in the residual are identified by the $3\sigma$ criterion, as shown in the box. The identification result also is plotted in Figure 11. It can be seen from the figure that the method successfully identifies the double-point outlier and the spot outlier. Meanwhile, a few discontinuous data at the right end of the profile are recognized as outliers.

### 3.3.2. Identification of Outliers Based on RLWR and Median Criterion.
This section uses RLWR smoothing algorithm combined with median criterion to identify outliers; it can be called the RLWRM identifier. The algorithm is as follows.

According to the residual $g_i$ obtained in above section, the median of the residual sequence is calculated, i.e., $m_g = \text{median}(g_i)$.

The scaled median absolute deviation (MAD) is defined as follows:

$$\text{MAD} = k \cdot \text{median}\left(\left|x_i - \text{median}(g_i)\right|\right), \quad i = 1, 2, \ldots, n, \qquad (12)$$

where $k = 1/(\sqrt{2}\, erfc^{-1}(1/2)) \approx 1.4826$.

When the data element is greater than three times MAD, it is considered to be outlier, as shown below.

$$y_i = \begin{cases} x_i, & |x_i - m_g| \leq 3\text{MAD}, \\ \text{outlier}, & |x_i - m_g| > 3\text{MAD}. \end{cases} \qquad (13)$$
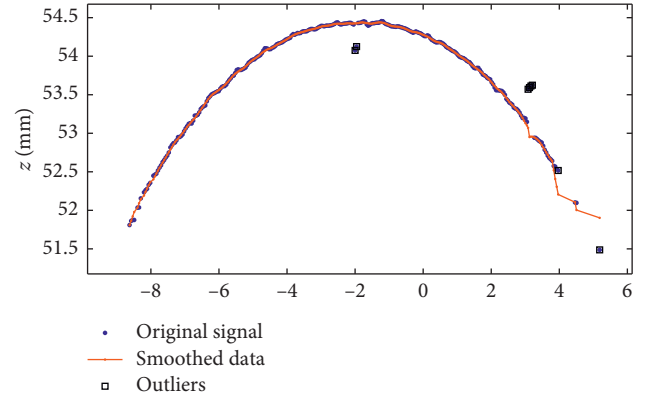


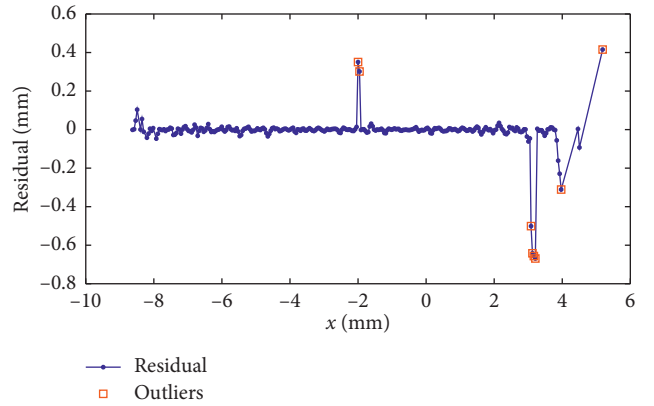FIGURE 11: Outlier recognition using RLWRP identifier, window length = 11.



FIGURE 12: Residual of original and smoothed data using RLWRP identifier.

The RLWRM identifier is used to identify outliers in Figure 6. The RLWR smoothing data are used when the smoothing window length is 11. The identification effect is shown in Figures 13 and 14. In Figure 13, the blue point is the residual between RLWR smoothing data and original data and the red boxes are the outliers identified by the median criterion. The identification result and the smoothed value are plotted in Figure 14 for observation. It can be seen from the figure that although the RLWRM identifier can recognize the double-point outliers and the spot outliers, there are many misidentifications about the isolated outliers.

### 3.3.3. Identification of Outliers Based on RLWR and Quartile Criterion.
This section uses RLWR smoothing algorithm combined with quartile criterion to identify outliers; it can be called the RLWRQ identifier. Quartile criterion is a relatively robust identification method; the algorithm divides sorted data into quarters; $Q_1$, $Q_2$, and $Q_3$ are their break points. $Q_1$ is lower quartile (25 percentile), $Q_2$ is median (50 percentile), and $Q_3$ is upper quartile (75 percentile). The interquartile range (IQR) is introduced here as a statistic for checking outliers, i.e., $\text{IQR} = Q_3 - Q_1$.

Outliers are defined as elements more than 1.5 IQR above $Q_3$ or below $Q_1$.
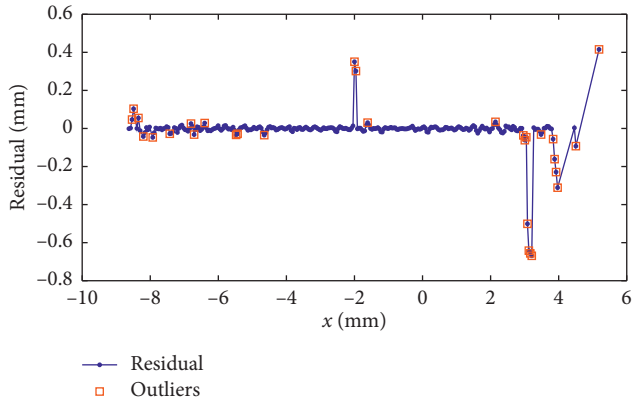
FIGURE 13: Residual of original and smoothed data using RLWRM identifier.
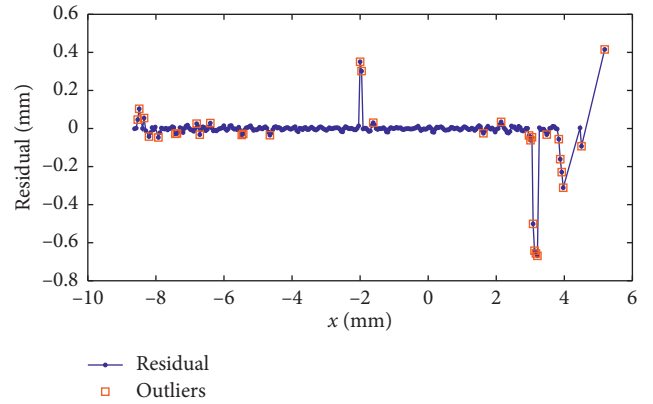


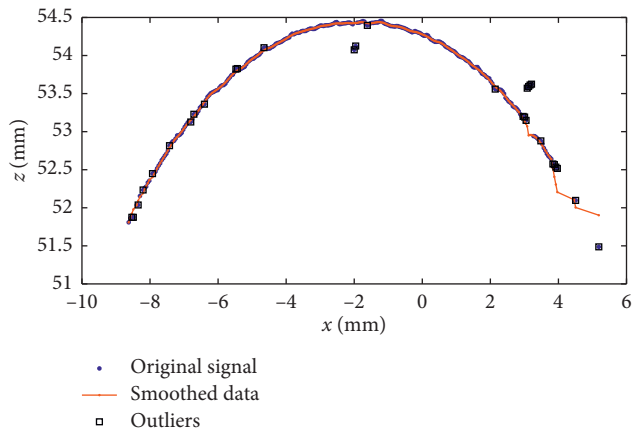FIGURE 15: Residual of original and smoothed data using RLWRQ identifier.



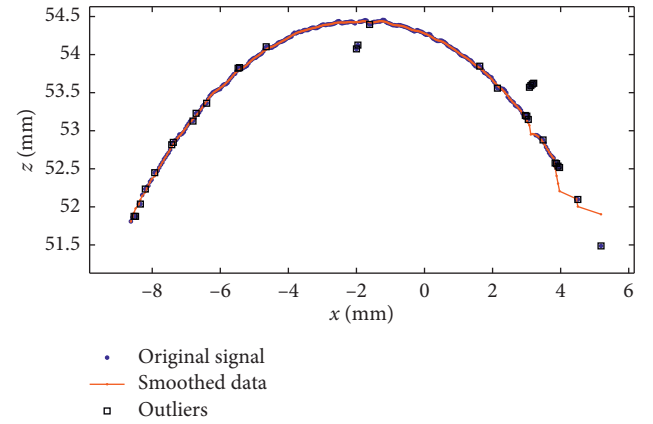FIGURE 14: Outlier recognition using RLWRM identifier, window length = 11.



FIGURE 16: Outlier recognition using RLWRQ identifier, window length = 11.

$$y_i = \begin{cases} x_i, & Q_1 - 1.5 \, (\text{IQR}) \le g_i \le Q_3 + 1.5 \, (\text{IQR}), \\ \text{outlier}, & \text{otherwise}. \end{cases}$$

$$(14)$$

The RLWRQ identifier is used to identify outliers in Figure 6. Similarly, the smoothing window length of RLWR is 11. The identification results are shown in Figures 15 and 16. It can be seen from the figures that the identification effect of the RLWRQ identifier is similar to the RLWRM identifier. This algorithm also has misidentification of outliers and identifies more normal data as outliers. If these values are removed, the profile will not be truly reflected, forming new errors and affecting the measurement accuracy.

## 4. Profile Fitting and Error Analysis

According to the previous section, the RLWRP identifier can obtain better identification results for the isolated outlier. The identification results in Figure 13 are used to remove outliers, and the ellipse fitting experiment is performed by the least squares method.

The least squares method is one of the most important methods of data fitting. The least squares method has the characteristics of simple, effective, and strong applicability. Therefore, this method is selected to conduct ellipse fitting research. This paper chooses the algebraic least squares method to carry out ellipse fitting research, which is to determine the ellipse parameters by measuring the smallest algebraic distance squared from the fitting ellipse to the ellipse.

The elliptic algebraic equation is expressed as follows:

$$x^2 + Axy + By^2 + Cx + Dy + E = 0. \quad (15)$$

According to the principle of least squares method, its objective function is minimized as follows:

$$F(\alpha) = \sum_{i=1}^{n} [f(A, B, C, D, E)]^2$$

$$= \sum_{i=1}^{n} \left( x_i^2 + Ax_i y_i + By_i^2 + Cx_i + Dy_i + E \right)^2. \quad (16)$$

To minimize $F(\alpha)$ on the basis of the extreme value principle, the following equation exists:

$$\frac{\partial f}{\partial A} = \frac{\partial f}{\partial B} = \frac{\partial f}{\partial C} = \frac{\partial f}{\partial D} = \frac{\partial f}{\partial E} = 0. \tag{17}$$

Thus, a linear equation is obtained. Then, by solving the linear equations and combining the constraints, the values of the equation coefficients can be obtained. Get the elliptic equation and draw the fitted ellipse. Finally, get the ellipse equation and draw the fitted ellipse.

The fitting results are compared before and after outlier eliminating, as shown in Figure 17. Meanwhile, the fitting ellipse parameters are shown in Table 1.

In Figure 17, the blue dots are the original data, the boxes are the outliers that are identified, the dotted line is the fitted ellipse of the original data, and the dash-dotted line is the fitting ellipse after removing the outliers.

Before removing outliers, "∗" is the center of the ellipse fitted with the original data, the diameter of the ellipse's major axis is 11.7197 mm, the eccentricity is 0.9263, and the ellipticity is 0.6233. The ellipticity is defined as follows:

$$\text{ellipticity} = \frac{D_{\max} - D_{\min}}{D_{\max}}. \tag{18}$$

After eliminating the outliers, the center of the fitted ellipse is "+," the major axis diameter is 15.8586 mm, the eccentricity is 0.6739, and the ellipticity is 0.2611.

At the same time, the profile was measured by the co-ordinate measuring machine (CMM). The measurement data are shown in Table 1. By comparison, it is found that the fitting ellipse with outliers removed is similar to the measurement results of CMM.

From the above figure and table, it can be found that the fitting result of the original profile has deviated from the actual situation. After removing the outliers by using the RLWRP identifier, the fitted ellipse conforms to the measurement reality.

In addition, the fitting error before and after removing outliers is analyzed, and the results are shown in Table 2. For the original profile, the sum of squares due to error (SSE) is 7.1938$e$ − 06 and the root mean square error (RMSE) is 1.5750$e$ − 04. After eliminating the outliers, the SSE and the RMSE of the fitted ellipse are reduced to 2.1157$e$ − 06 and 8.6617$e$ − 05, respectively. It shows that the elimination of outliers greatly reduces the fitting error and improves the fitting accuracy, and the measurement results are more accurate.

## 5. Conclusions

In this paper, the profile of magnesium alloy tubes with large curvature is measured by a laser vision system based on line structure laser. There are two typical outliers in measurement, that is, point outliers and spot outliers. For these outliers, the moving average method, the Hampel method, and the RWLR method were, respectively, adopted to identify the outliers of profile data. And their ability to identify outliers is discussed for the above methods with different window lengths.

The experiment found that all the above methods could identify the isolated point outliers, but neither the moving mean method nor the Hampel method could identify the isolated spot outliers. In this article, the RWLR method is studied emphatically for the isolated outliers, which was
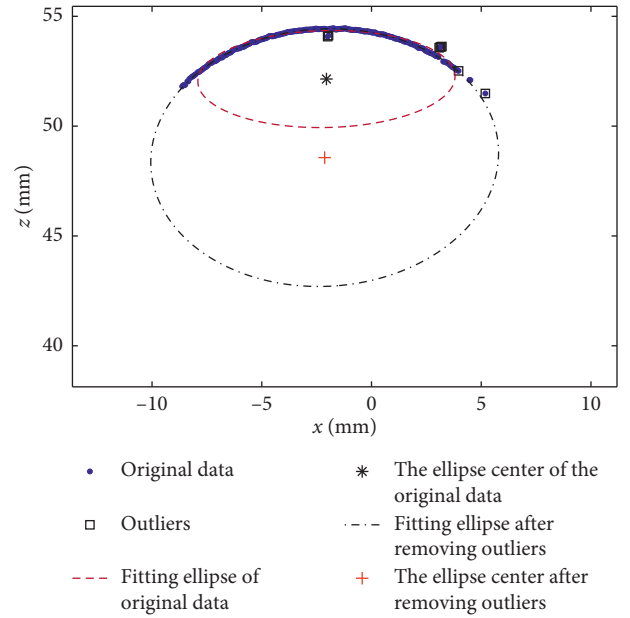


FIGURE 17: Comparison of ellipse fitting result for original profile, window length = 11.

TABLE 1: Elliptical parameters obtained by profile fitting.

| Elliptical parameters | Original data | After removing outliers | The CMM data |
|---|---|---|---|
| Center coordinates | (−2.0536, 52.1455) | (−2.1344, 48.5674) | — |
| Major axis (mm) | 11.7197 | 15.8586 | 15.683 |
| Minor axis mm) | 4.4145 | 11.7173 | 11.378 |
| Eccentricity | 0.9263 | 0.6739 | 0.688 |
| Ellipticity | 0.6233 | 0.2611 | 0.283 |

TABLE 2: Ellipse fitting error before and after removing outliers.

| Fitting error | Original data | After removing outliers |
|---|---|---|
| SSE | 7.1938$e$ − 06 | 2.1157$e$ − 06 |
| RMSE | 1.5750$e$ − 04 | 8.6617$e$ − 05 |
| MSE | 2.4806$e$ − 08 | 7.5024$e$ − 09 |

combined with the РайТа criterion, the median criterion, and the quartile criterion. The research shows that the RWLR smoothing algorithm combined with РайТа criterion, i.e., RLWRP identifier, can more accurately identify different types of outliers with a lower misidentification rate.

At the same time, according to the outlier identification result of the RWLRP identifier, the profile fitting was carried out by the algebraic least squares method. Then, the main parameters of the fitted ellipse are obtained, and the fitting errors are calculated. After the comparison and analysis of the fitting results before and after the outlier processing, it is found that the data contaminated by outliers will lead to a great deviation of profile fitting and wrong profile shape

parameters. The RWLRP identifier is a robust, accurate, and efficient outlier identification method, which can effectively deal with outliers in profile data, especially for spot outliers. This algorithm is suitable for data cleaning in the line structure light measurement, which can effectively improve the precision and accuracy of online profile measurement in the process of hot bending of magnesium alloy tube.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## Acknowledgments

## References

[1] T. M. Pollock, "Weight loss with magnesium alloys," *Science*, vol. 328, no. 5981, pp. 986-987, 2010.

[2] H. E. Friedrich and B. L. Mordike, *Magnesium Technology: Metallurgy, Design Data, Applications*, Springer, Berlin, Germany, 2006.

[3] A. A. Luo, "Magnesium casting technology for structural applications," *Journal of Magnesium and Alloys*, vol. 1, no. 1, pp. 2–22, 2013.

[4] S. You, Y. Huang, K. U. Kainer, and N. Hort, "Recent research and developments on wrought magnesium alloys," *Journal of Magnesium and Alloys*, vol. 5, no. 3, pp. 239–253, 2017.

[5] B. Landkof, "Magnesium applications in aerospace and electronic industries," *Magnesium Alloys and Their Applications*, Wiley-VCH, Weinheim, Germany, pp. 168–172, 2000.

[6] W. J. Joost and P. E. Krajewski, "Towards magnesium alloys for high-volume automotive applications," *Scripta Materialia*, vol. 128, pp. 107–112, 2017.

[7] L. Wang, G. Fang, L. Qian, S. Leeflang, J. Duszczyk, and J. Zhou, "Forming of magnesium alloy microtubes in the fabrication of biodegradable stents," *Progress in Natural Science: Materials International*, vol. 24, no. 5, pp. 500–506, 2014.

[8] T. F. Kong, X. Z. Lu, and L. C. Chan, "Analysis and reduction of wrinkling defects for tube-hydroforming magnesium alloy components at elevated temperatures," *Materials & Design*, vol. 173, Article ID 107761, 2019.

[9] Z.-J. Tao, X.-G. Fan, H. Yang, J. Ma, and H. Li, "A modified Johnson–Cook model for NC warm bending of large diameter thin-walled Ti–6Al–4V tube in wide ranges of strain rates and temperatures," *Transactions of Nonferrous Metals Society of China*, vol. 28, no. 2, pp. 298–308, 2018.

[10] Y. J. Gou, Y. H. Shuang, Y. Zhou et al., "Damage and wall thickness variation of magnesium alloy tube in large curvature and no-mandrel bending process," *Rare Metal Materials and Engineering*, vol. 47, no. 8, pp. 2422–2428, 2018.

[11] T. Hilditch, D. Atwell, M. Easton, and M. Barnett, "Performance of wrought aluminium and magnesium alloy tubes in three-point bending," *Materials & Design*, vol. 30, no. 7, pp. 2316–2322, 2009.

[12] J. P. Bentley, *Principles of Measurement Systems*, Pearson Education, London, UK, 2005.

[13] W. Liu, K. Fan, P. Hu, and Y. Hu, "A parallel error separation method for the on-line measurement and reconstruction of cylindrical profiles," *Precision Engineering*, vol. 51, pp. 1–9, 2018.

[14] M. Rahayem, N. Werghi, and J. Kjellander, "Best ellipse and cylinder parameters estimation from laser profile scan sections," *Optics and Lasers in Engineering*, vol. 50, no. 9, pp. 1242–1259, 2012.

[15] C. Munzinger, G. Lanza, D. Ruch, and J. Elser, "Component-specific scale for inline quality assurance of spatially curved extrusion profiles," *Production Engineering*, vol. 4, no. 2-3, pp. 193–201, 2010.

[16] Z. Gan and Q. Tang, *Visual Sensing and its Applications: Integration of Laser Sensors to Industrial Robots*, Springer, Berlin, Germany, 2011.

[17] S. Egorov, A. Kapitanov, and D. Loktev, "Turbine blades profile and surface roughness measurement," *Procedia Engineering*, vol. 206, pp. 1476–1481, 2017.

[18] S. Martínez, E. Cuesta, J. Barreiro, and B. Álvarez, "Analysis of laser scanning and strategies for dimensional and geometrical control," *The International Journal of Advanced Manufacturing Technology*, vol. 46, no. 5–8, pp. 621–629, 2010.

[19] Y.-C. Zhang, J.-X. Han, X.-B. Fu, and H.-B. Lin, "An online measurement method based on line laser scanning for large forgings," *The International Journal of Advanced Manufacturing Technology*, vol. 70, no. 1–4, pp. 439–448, 2014.

[20] D. M. Hawkins, *Identification of Outliers*, Springer, Berlin, Germany, 1980.

[21] Y. Wang and H.-Y. Feng, "Outlier detection for scanned point clouds using majority voting," *Computer-Aided Design*, vol. 62, pp. 31–43, 2015.

[22] N. M. R. Suri, N. M. Murty, and G. Athithan, *Outlier Detection: Techniques and Applications*, Springer Nature, Berlin, Germany, 2019.

[23] C. C. Aggarwal, "Outlier analysis," *Data Mining*, Springer, Berlin, Germany, pp. 237–263, 2015.

[24] P. M. Robinson, "Nonparametric estimation of time-varying parameters," *Statistical Analysis and Forecasting of Economic Structural Change*, Springer, Berlin, Germany, pp. 253–264, 1989.

[25] S. Efromovich, *Nonparametric Curve Estimation: Methods, Theory, and Applications*, Springer, Berlin, Germany, 2008.

[26] W. S. Cleveland, "Robust locally weighted regression and smoothing scatterplots," *Journal of the American Statistical Association*, vol. 74, no. 368, pp. 829–836, 1979.

[27] W. S. Cleveland and S. J. Devlin, "Locally weighted regression: an approach to regression analysis by local fitting," *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 596–610, 1988.

[28] W. S. Cleveland, S. J. Devlin, and E. Grosse, "Regression by local fitting: methods, properties, and computational algorithms," *Journal of Econometrics*, vol. 37, no. 1, pp. 87–114, 1988.

[29] W. G. Jacoby, "Loess: a nonparametric, graphical tool for depicting relationships between variables," *Electoral Studies*, vol. 19, no. 4, pp. 577–613, 2000.

[30] C. Loader, "Smoothing: local regression techniques," *Handbook of Computational Statistics*, Springer, Berlin, Germany, pp. 571–596, 2012.

[31] J. Ma, J. C.-W. Chan, and F. Canters, "Robust locally weighted regression for superresolution enhancement of multi-angle remote sensing imagery," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 4, pp. 1357–1371, 2014.

[32] N. R. Leonor, R. F. S. Caldeirinha, T. R. Fernandes, and M. G. Sanchez, "A simple model for average reradiation patterns of single trees based on weighted regression at 60 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 11, pp. 5113–5118, 2015.

[33] C. Chen, Y. Wang, Y. Li, T. Yue, and X. Wang, "Robust and parameter-free algorithm for constructing pit-free canopy height models," *ISPRS International Journal of Geo-Information*, vol. 6, no. 7, p. 219, 2017.

[34] A. Nurunnabi, G. West, and D. Belton, "Robust locally weighted regression for ground surface extraction in mobile laser scanning 3D data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. II-5/ W2, pp. 217–222, 2013.

[35] A. Nurunnabi, G. West, and D. Belton, "Robust locally weighted regression techniques for ground surface points filtering in mobile laser scanning three dimensional point cloud data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 4, pp. 2181–2193, 2016.

[36] K. Liu, W. Wang, R. Tharmarasa, J. Wang, and Y. Zuo, "Ground surface filtering of 3D point clouds based on hybrid regression technique," *IEEE Access*, vol. 7, pp. 23270–23284, 2019.

[37] Y. Yu, X. Fu, L. Yu, R. Liu, and J. Cai, "Combustion kinetics of pine sawdust biochar," *Journal of Thermal Analysis and Calorimetry*, vol. 124, no. 3, pp. 1641–1649, 2016.

[38] G. Sansoni, M. Trebeschi, and F. Docchio, "State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation," *Sensors*, vol. 9, no. 1, pp. 568–601, 2009.

[39] J. Santolaria, J. J. Pastor, F. J. Brosed, and J. J. Aguilar, "A one-step intrinsic and extrinsic calibration method for laser line scanner operation in coordinate measuring machines," *Measurement Science and Technology*, vol. 20, no. 4, Article ID 045107, 2009.

[40] T. F. Townley-Smith and E. A. Hurd, "Use of moving means in wheat yield trials," *Canadian Journal of Plant Science*, vol. 53, no. 3, pp. 447–450, 1973.

[41] B. Bucciarelli, L. G. Taff, and M. G. Lattanzi, "A generalization of the moving mean*." *Journal of Statistical Computation and Simulation*, vol. 48, no. 1-2, pp. 29–46, 1993.

[42] F. R. Hampel, "A general qualitative definition of robustness," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1887–1896, 1971.

[43] F. R. Hampel, "The influence curve and its role in robust estimation," *Journal of the American Statistical Association*, vol. 69, no. 346, pp. 383–393, 1974.

[44] R. K. Pearson, Y. Neuvo, J. Astola, and M. Gabbouj, "Generalized hampel filters," *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 1, p. 87, 2016.

[45] Z. B. Yao, J. Q. Xie, Y. Q. Tian, and Q. X. Huang, "Using hampel identifier to eliminate profile-isolated outliers in laser vision measurement," *Journal of Sensors*, vol. 2019, Article ID 3823691, 12 pages, 2019.