State measurement of isolating switch using cost fusion and smoothness prior based stereo matching

Jinxin Xu¹, Qingwu Li¹,², Ying Luo¹, Yan Zhou¹,² and Jiayu Wang¹

Abstract
To better monitor the state of isolating switches, an efficient binocular vision-based state measurement system is proposed in this article. Two optimal cameras are selected as the vision of our inspection system. Firstly, stereo calibration and distortion rectification are performed on acquired image pair. Secondly, to recover the three-dimensional information of switch, we propose a semi-global stereo matching method by using data- and structure-driven cost volume fusion and then optimizing raw disparity map with weighted- and edge discriminated-smoothness prior. Gradient content is enforced on the weight for suppressing small-weight-accumulation problem in weak-textured regions. Besides, Hough transform with feature constraints is implemented for removing the chaotic lines and extracting center line of the switch arm. Finally, based on the center line and corresponding disparity map of the switch arm, triangulation principle is used for calculating the true angle between the switch arm and insulator such that whether or not the isolating switch is fully closed can be detected. The experimental results demonstrate that the proposed stereo matching method can achieve good performance in Middlebury v.3 data set and switch images, and the system can precisely measure the state of switches.

Keywords
Isolating switch measurement, binocular vision, semi-global stereo matching, Hough transform, triangulation principle

Date received: 11 February 2020; accepted: 14 April 2020

Introductions
In a power system, monitoring the state of the substation equipment is a significant task.¹ Isolating switch is the most frequently used device in high-voltage switchgear.²,³ Although the operating principle and structure of isolating switch are relatively simple, it is relatively prone to failure because of large amount of use. Most of switches operate outdoors, so the equipment components can be inevitably oxidized. The aging of the transmission parts leads to the deviated rotation angle and the incomplete closure of the isolating switch. Whether the switch is completely closed has a great influence on the safety of a substation. Therefore, to ensure the safety and reliability of the substation
work, it is especially considerable to correctly identify the state of the isolating switch.

Traditional inspections mainly rely on the manual observation by telescope and subjective judgment. These methods are time-consuming and susceptible to subjective influence. In recent years, real-time inspection technologies for substation have been rapidly developed. To improve the efficiency of state monitoring for isolating switch, vision-based intelligent inspection technologies have become a new trend, which greatly improves the accuracy and reduces the labor costs. In general, for effectively detecting the switch state, insulators are required to be first located followed by the state determination of the switch. Zhang et al. proposed a simple but efficient isolating switch state recognition method, which determines the state of the switch by directly performing Hough transform on the switch boundary. Lu et al. employed a method based on image geometry to monitor the substation switch. The angle of the arm is calculated by cosine theorem after the contour of the arm image is obtained, whether the switch state is closed or not is determined eventually. Bin et al. introduced a switch detection method based on image recognition technology. The approach performs well on the conditions of image rotation and uneven illumination mainly by extracting and matching scale-invariant feature transform (SIFT) features of the isolating switch. In addition, infrared thermography has also gained much attention in electrical preventive maintenance due to its high precision and sensitivity imaging characteristics. Ullah et al. used multilayered perceptron to classify the thermal conditions of components to address initial prevention mechanism for power substations by infrared thermography. Further work based on deep learning approach and infrared images has also been carried out to defect analysis in high-voltage equipment.

However, the above approaches are all based on two-dimensional (2D) image content, in which the angle between the switch arm and the insulator cannot uniquely reflect the true angle in 3D space especially under the changes of the camera angle. Moreover, by using an infrared device to monitor the closing of the switch, it is impossible to accurately determine whether the closure is completely successful at the moment the switch is closed. Instead, we employ binocular vision technology to significantly detect the state of the isolating switch according to the 3D depth information in this article.

Binocular stereo vision technology has an extensive range of applications in computer vision, such as robotic autonomous navigation, aerospace and remote sensing, and industrial automation systems. The binocular stereo vision system mainly includes the following parts: binocular image acquisition, camera calibration, image correction, stereo matching, and 3D reconstruction process. Dense stereo matching is the key issue of stereo vision, and its purpose is to find the corresponding points in image pairs and obtain the disparity map for recovering 3D information. Four steps are primarily included in the stereo matching procedure, which are matching cost calculation, cost aggregation, disparity calculation, and disparity refinement. There are still many challenging issues existing in stereo matching algorithm, such as radiometric distortions, poor performance in weak-textured and disparity discontinuous regions.

In this article, aiming to precisely measure the state of isolating switch in 3D space, we mainly focus on proposing an effective semi-global stereo matching method by two strategies: the first one is constructing the matching cost by fusing data- and structure-driven cost volumes and the second is optimizing the raw disparity map with weighted-edge discriminated-smoothness constraint. In particular, the weight is combined with gradient content such that the small-weight-accumulation problem in weak-textured regions can be suppressed to a certain extent. Based on our proposed semi-global stereo matching method, we also present a complete binocular vision system that can be employed to calculate the true angle between the switch arm and insulator, which consists of stereo calibration and rectification, switch detection, center line extraction of switch arm, stereo matching, and depth calculation. Firstly, stereo calibration is performed to derive internal and external parameters of cameras. After image rectification, the complete switch inspection is conducted followed by proposed feature constrained-Hough transform for removing the chaotic lines and extracting center line of the switch arm. Moreover, the proposed stereo matching method is implemented from which only the stable disparities are utilized in later 3D coordinate mapping for guaranteeing the accuracy of angle measurement. Finally, the true angle between the switch arm and insulator is obtained by the principle of triangulation and whether or not the isolating switch is fully closed can be identified. The experimental results illustrated that the proposed method not only obtains higher measurement accuracy but also contributes to the substation safety.

The following section outlines the related work of stereo matching. Following that, the methodologies of the proposed methods are presented in the third section. In the fourth section, the performance of the proposed method is tested on Middlebury v.3 data set as well as the switch images. The fifth section concludes the article.

Related work

A detailed classification of stereo matching algorithms is provided by Scharstein and Szeliski, in which stereo matching algorithms can be divided into local and global methods. Generally, global methods can be formulated as an energy minimization problem. The energy function composes of a data term and smoothness term, which is to explicitly penalize disparity discontinuity between neighbor pixels. Dynamic programming (DP), graph cut, and belief propagation are famous global
optimizers, which can obtain comprehensive disparity characterization at the expense of substantial computational complexity.

Different from global methods, local ones are performed by calculating the matching cost between patches in image pairs, followed by aggregation of computed cost volumes over local support windows. Patches are compared by dissimilarity measures, which can be divided into two groups. The first group is traditional dissimilarity measures including sum of absolute differences, normalized cross-correlation, rank and census transform. The second group is learning-based measures, such as discriminative dictionary learning (DDL)\textsuperscript{36} and convolutional neural networks.\textsuperscript{37} Due to better generalization ability for stereo, in this work, cost computation is based on DDL.\textsuperscript{32} However, the DDL tends not to preserve fine details due to the loss of constraint by structural information such as gradients. Hence, in this article, we present a data- and structure-driven cost volume fusion strategy to further improve the robustness of cost volume such that the shortcoming of the single data-driven cost computation can be neutralized.

Besides, cost aggregation is another vital step in local methods. In despite of variable support window (VSM)-based or adaptive support weight (ASW)-based aggregation method, the primary aim is to ensure the implicit smoothness assumption that pixels within the same support window share the similar depth. In VSM-based methods, multiple, fixed, and adaptive windows are employed to prevent them from crossing depth discontinuous regions.\textsuperscript{38–40} However, a rectangular window is not always rigorous enough at discontinuity regions and the computation within adaptive windows can be time-consuming. ASW-based methods assign preferential weights for each pixel within predefined window. Normally, color dissimilarity and distance proximity are chosen for determining the weight. The disadvantage of ASW-based methods is that the matching ambiguity will be unavoidable in low-textured regions, where small-weight-accumulation problem exists.

In the work by Hirschmuller,\textsuperscript{41} semi-global matching (SGM) stereo method is proposed, which works in three steps: first, pixel-wise cost computation is computed; then, the cost volume is aggregated along 8/16 1D search passes with local smoothness constraint between neighbor pixels. Finally, post-processing is employed to remove outliers and obtain only consistent estimation. However, SGM is sensitive to penalty parameter as the smoothness terms need to be big enough to process smooth disparity map and small enough to maintain fine structures at discontinuities. In addition, SGM may result in ambiguous discontinuities around foreground objects in textureless background. Numerous variants of SGM have also been developed to obtain significant performance gains. Mei et al.\textsuperscript{42} performed cost aggregation in dynamic cross-based window and optimized the disparity map in a scanline optimization framework. Liu et al.\textsuperscript{43} proposed a stereo framework and formulated a joint second-order smoothness prior, avoiding the blend of the foreground and background and retaining the disparity discontinuities of strong texture regions. Liu et al.\textsuperscript{44} presented a 3D aggregation strategy based on adaptive support window and good performance in slanted regions can be reached. Chuang et al.\textsuperscript{45} incorporated penalty tuning method and edge content in gradual SGM cost aggregation, balancing the smoothness and fine structures of disparity result. Yao et al.\textsuperscript{46} extended the SGM to a more global matching for overcoming the limitation of local scanline optimization.

However, in these SGM-based methods mentioned above, the selection of optimal disparities from full candidates introduces uncertainty in aggregation passes and the final depth recovery ability. For better handling slanted plans, we propose weighted- and edge discriminated-smoothness constraint to efficiently optimize the raw disparity result. The penalties at edge regions are automatically determined based on confidence measure for penalizing larger disparity changes. The weight combining appearance and gradient content can suppress the small-weight-accumulation problem to a certain extent. In this article, we dedicate to apply binocular stereo vision system to state measurement and improve the matching performance in weak-textured and disparity discontinuous regions.

The contributions of this article are as follows: (1) feature constraints are enforced on Hough transform for obtaining the center line of switch arm; (2) data- and structure-driven cost volumes are fused to construct the matching cost, improving the matching accuracy; and (3) in cost aggregation process, weighted- and edge discriminated-smoothness constraint is proposed to optimize the raw disparity map. The smoothness constraint is determined automatically according to the image confidence.

**Methodology**

**Overview of procedure**

An overview of the proposed method is illustrated in Figure 1. Our system can be divided into three parts: camera calibration and rectification, stereo matching, and center line extraction of switch arm. To start, the first part mainly includes stereo calibration and distortion rectification for camera parameters and subsequently horizontal dissimilarity searching. The second part includes semi-global stereo matching based on data- and structure-driven cost volume fusion and weighted- and edge discriminated-smoothness prior. The third part is for center line extraction of switch arm and we mainly focus on exploiting feature constraints-based Hough transform to remove chaotic lines and extract center line of the switch arm. Ultimately, the vector of the center line in 3D space can be obtained by combining the disparity map and the positions of center line. Consequently, we can get the angle between the arm and the
insulator so as to determine whether or not the isolating switch is fully closed.

**Matching cost computation**

Matching cost measurement is used to compute initial cost volume, which is the basis of the whole stereo matching algorithm. Cost volume is a 3D matrix with each element corresponding to the cost of a relative pixel at a certain disparity level. Considering better generalization ability for stereo, we adopt the DDL method, which belongs to data-driven approaches and utilizes sparse representation to represent each patch over learned discriminative dictionary, to measure the similarity between image patches. The matching cost value \( C_{DDL}(p, d) \) obtained from DDL method for pixel \( p \) at disparity value \( d \) is denoted as

\[
C_{DDL}(p, d) = -s(P_l(x, y), P_r(x - d, y))
\]

where \( s(P_l(x, y), P_r(x - d, y)) \) indicates the similarity score between input patches \( P_l(x, y) \) and \( P_r(x - d, y) \), the negative sign means the similarity score is converted to the matching cost.

Nonetheless, the DDL method tends not to preserve fine details due to loss of constraint by structural information such as gradients. Therefore, in this article, we present a data- and structure-driven cost volume fusion strategy to further improve the robustness of cost volume such that the shortcoming of the single data-driven cost can be neutralized. The fused matching cost measure is defined as follows

\[
C(p, d) = \lambda \min(C_{GRAD}(p, d), \tau_{GRAD}) + (1 - \lambda)C_{DDL}(p, d)
\]

where \( \tau_{GRAD} \) is the truncating value, coefficient \( \lambda \) balances the impact of each measure. \( C_{GRAD}(p, d) \) is the gradient-based cost measurement, which is exploited as follows

\[
C_{GRAD}(p, d) = \sqrt{(\nabla_x G_L(p))^2 + (\nabla_y G_L(p))^2} - \sqrt{(\nabla_x G_R(p_d))^2 + (\nabla_y G_R(p_d))^2}
\]

where \( \nabla_x \) and \( \nabla_y \) are the derivative operators in \( x \) and \( y \) directions, respectively. Grayscale image, \( G(\cdot) \), is obtained by averaging the RGB channels of input color image. \( p_d \) is the corresponding pixel in the right image of pixel \( p(x, y) \) in the left image and satisfy \( p_d = p(x - d, y) \). In our implementation, we set the empirical values \( \tau_{GRAD} = 2, \lambda = 0.3 \).

**Weighted and edge discriminated cost aggregation**

\[ E(D) = \sum_{p \in D} m(p, d_p) + \sum_{(p, q) \in \mathbb{N}} s(d_p, d_q) \]

where \( D \) represents disparity map, \( p \) belongs to image \( I \) and is assigned to disparity \( d_p \). The first term is the data term, while the latter one is regularization function enforced on neighboring pixels \( p \) and \( q \) in accordance with the predefined set \( \mathbb{N} \). \( P_1 \) is used to penalize small jumps in disparity.

**Figure 1.** Flowchart of the proposed measurement system. Our system can be divided into three parts in detail: camera calibration and rectification, stereo matching, and center line extraction of switch arm.
For simplicity and computational efficiency, edge weighted horizontal tree structure is adopted for enforcing the connectivity of adjacent pixels such that the energy optimization can be efficiently implemented via DP. The horizontal tree is constructed as shown in Figure 2. The tree is rooted on pixel $p$, and the horizontal tree is established by extending the node $p$ in the vertical direction, then a batch of horizontal expanding is performed. A complete cost aggregation process combines forward and backward passes in the horizontal phase.

In the edge weighted tree structure, the edge connecting adjacent nodes are generally assigned weight determined by color difference. However, in weak-textured regions, the color difference would be quite small. Many unreliable high weights accumulate along a long path, resulting in small-weight-accumulation drawback. In this article, we propose an enhanced weight function based on color and gradient content. The edge image for weight calculation is obtained by random forest method. The proposed edge weight is defined as follows

$$W_{pq}(I) = \begin{cases} 
-\left(\frac{|I_p - I_q|}{\sigma} + \max\left(\frac{|E_p - E_q|}{E_m}, T_w\right)\right), & ||I_p - I_q|| \leq T_w \\
\exp\left(-\frac{|I_p - I_q|}{\sigma}\right), & \text{else}
\end{cases}$$

where $||I_p - I_q||$ is the maximum color difference computed separately under the RGB three channels between adjacent pixels $p$ and $q$. $I$ and $E$ denote the reference image and edge image, respectively. $E_m$ is the maximum pixel value in the edge image. $\sigma$ is a user-defined parameter that adjusts the smoothness degree and $\beta$ is used to normalize the interval of pixel values. $T_w$ is the truncating value and the empirical parameters are set in our implementation as follows: $\sigma = 0.1$, $\beta = 20$, and $T_w = 2$. When color difference between adjacent pixels is less than $T_w$, the weight only considering the color content is close to 1 and should be reduced to alleviate the problem of small-weight-accumulation. The degree of reduction of the weight is determined according to the gradient difference, and the larger the difference, the lower the weight. By combining the two kinds of information, the weight can be adaptively adjusted such that the matching accuracy will be increased in weak-textured and disparity discontinuous regions.

Confidence-based penalty tuning. How to automatically assign appropriate penalty variables ($P_1, P_2$) is vital to the cost aggregation performance. A small penalty causes sharp boundaries, but many outliers would appear. On the other hand, a large penalty would reduce mismatching in smooth regions but results in over-smooth object boundaries, losing details of the object. To enhance the robustness of the parameters, an adaptive penalty tuning strategy is employed in the proposed method.

Normally, the initial disparity maps are obtained by WTA strategy; subsequently, the left–right consistency check is performed for picking confidence points from the initial disparity maps. A confident ratio $C_{rp}$ of pixel $p$ is defined by

$$C_{rp} = m_{sp}/m_{fp}, p \in n_{cp}$$

where $m_{fp}$ and $m_{sp}$ represent the smallest and secondary small cost values, respectively. $n_{cp}$ is the number of confident points. The greater the confident ratio, the more likely the pixel is matched correctly. Afterward, the penalty $P_1$ is obtained by taking the average of the summation between cost values $m_{fp}$ and $m_{sp}$

$$P_1 = \frac{1}{n_{cp}} \sum_{p \in n_{cp}} (m_{sp} + m_{fp})$$

$$P_2 = (P_1 \cdot ||I_p - I_q|| + P_1)/2$$
where $||I_p - I_q||$ indicates the absolute intensity difference. $P_2$ is a larger constant accounting for penalizing all larger disparity changes that occur at disparity boundaries.

**Edge discriminated cost aggregation.** After constructing the horizontal tree, cost volume is first aggregated from leaves to the root node which locates at same column in horizontal passes. Then, the cost volume is aggregated from the subsidiary nodes to the root node along vertical passes. Therefore, the energy function in equation (4) can be optimized as follows

$$E(D) = \sum_{p\in T} m(p, d_p) + W_{pq}(T) \left( \sum_{q\in\overline{q}(p)} P_1 \left(d(p, d_q), 0 < |d_p - d_q| < d_{step}\right) + \sum_{q\in\overline{q}(p)} P_2 \left(d(p, d_q), d_p - d_q > d_{step}\right) \right)$$

where $P_1$ is a lower penalty for adapting slanted or curved planes and only penalizing the small step change $d_{step} = 1$. For depth discontinuities located by image edge, a more critical penalty $P_2$ is used to penalize all other larger disparity jumps. Specifically, when the disparity jumps smaller than $d_{step}$, small penalty is assigned. And the aggregated cost value comes from the minimum of the adjacent three disparity layers, which include the upper, current, and lower disparity layers. When pixel $p$ locates at image edges where depth discontinuities usually exist, the optimum disparity value is more likely to be generated from layers except from the three disparity layers. Thus, we choose to enforce different penalties on different layers to build appropriate regularization term such that both the slanted planes and depth discontinuities can be handled well.

The cost aggregation process is shown in Figure 2, which is divided into two phases: the horizontal phase and the vertical phase. In each phase, the forward and backward propagation are employed. For the forward pass, the cost volume is renewed recursively from the leftmost pixel of a scanline. The accumulated cost values are preserved in array $C_F$. Similarly, the backward pass calculates the costs successively from the rightmost pixel and the accumulated cost values are stored in $C_B$. Finally, the aggregated cost value $C_{Fin}(p, d_p)$ for pixel $p$ at disparity $d_p$ is calculated by summing up the two reverse passes followed by subtracting the current cost

$$C_{Fin}(p, d_p) = C_F(p, d_p) + C_B(p, d_p) - m(p, d_p) \quad (11)$$

**Disparity computation and refinement**

After the cost volumes are regularized, the raw disparity map is determined by selecting the minimum aggregated disparity value for each pixel utilizing WTA strategy

$$D(p) = \arg \min_{d \in \text{disp}} C_{Fin}(p, d_p) \quad (12)$$

where $\text{disp}$ is the maximum disparity candidate. The disparity map obtained by this stage still has outliers in occluded regions. Therefore, weighted median filter is adopted to correct unstable pixels to a great extent.

**Center line extraction of switch arm**

The rough location of the isolating switch employs the SVM algorithm combined with the Hog eigenvector of the switch, and this rough location is not the focus of this article. Instead, suppose that the rough location is finished and then we focus on exploiting feature constraint-based Hough transform for obtaining the center line of switch arm. Hough transform is an approach for detecting lines, circles, and other shapes. The main principle is that the image is mapped from the original polar coordinate space into parametric space and then to fit the shapes such as lines and circles. A line in the polar coordinate space corresponds to a point in the Hough space, and a line in the Hough space corresponds to a point in the polar coordinate space. Firstly, the contours of switch image from rough location are detected by random forest method. After that, Hough transform algorithm is adopted to search the straight lines of arm edges on the contour image. However, there are chaotic lines in the background of image and small segments also exist, resulting in failed center line extraction. Therefore, a novel method leveraging feature constraints-based Hough transform is proposed for removing the chaotic lines and extracting center line of the switch arm. The specific steps are as follows:

Step 1: As isolating switch has been identified by SVM method, the switch arm takes up the close-up view of the identified isolating switch. Assume the end-to-end coordinates of lines detected by Hough Transform is stored in array $A_N$

$$A_N = \{(x_1, y_1), (x_1', y_1')\}, \{(x_2, y_2), (x_2', y_2')\}, \ldots, \{(x_m, y_m), (x_m', y_m')\}, \ldots, \{(x_N, y_N), (x_N', y_N')\}, \quad 1 \leq m \leq N$$

where $N$ is the total number of lines, $(x, y)$ and $(x', y')$ denote the starting and ending coordinate of each line, respectively.

$$A_N = \{(x_1, y_1), (x_1', y_1')\}, \{(x_2, y_2), (x_2', y_2')\}, \ldots, \{(x_m, y_m), (x_m', y_m')\}, \ldots, \{(x_N, y_N), (x_N', y_N')\}, \quad 1 \leq m \leq N$$

The slope $k_N = (y'_N - y_N)/(x'_N - x_N)$, angle $\alpha_N = \arctan(k_N)$, and the length $l_N = \sqrt{(x'_N - x_N)^2 + (y'_N - y_N)^2}$
of each candidate line are obtained according to the end-to-end coordinates.

Step 2: A line detected by the Hough transform could be divided into several segments, it is necessary to merge the segments by means of integration strategy. If the minimum pixel distance between any two parallel lines belonging to one long line is smaller than integration threshold $T_n$, these two parallel lines are merged as one line. To eliminate inverse influence of vertical lines in the background, we delete the lines of which the angle is larger than $T_c$ degree since the purpose of this article is to detect whether or not the isolating switch is fully closed. $T_b = \text{wid}/2$, $T_v = 80$, wid is the width of the image.

Step 3: To eliminate bad influence of small and cluttered lines, the threshold of the candidate lines is set to $T_o$, that is, the lines are composed of at least $T_o$ pixels. We select the first $m'$ lines in terms of high probabilities such that the two edges of switch arms can be screened out accurately. Then, the corresponding information of the selected lines is stored in array $P_{am}$. Values of $T_o = 50, m' = 5$ are assumed here

$$P_{am} = \{(x_1, y_1), (x_1', y_1'), \alpha_1, l_1\}, \{(x_2, y_2), (x_2', y_2'), \alpha_2, l_2\}, \ldots, \{(x_m, y_m'), (x_m', y_m'), \alpha_m, l_m\}\}$$

where $(x_c, y_c)$ and $(x_c', y_c')$ denote the end-to-end coordinates of middle line, respectively. $\{(x_1, y_1), (x_1', y_1')\}$ and $\{(x_2, y_2), (x_2', y_2')\}$ represent two detected edges of switch arm. Note that if only the longest line is detected, the center line extraction is not necessary.

### Angle measurement of switch arm in 3D space

Stereo imaging system includes binocular image acquisition, camera calibration, image rectification, and 3D information recovery. With reference to Figure 3, the imaging planes of the binocular cameras are coplanar and the distance from cameras to optical centers $(O_1$ and $O_2$) is $f$. The objective point $P(u', v', z')$ in the 3D space is mapped to the points $P_l(u_l, v_l)$ and $P_r(u_r, v_r)$ in the left and right view, respectively. The baseline distance between the two cameras is $B$.

After acquiring the images, the intrinsic, extrinsic, and distortion parameters of the cameras are obtained based on the calibration method proposed by Zhang. To reduce the complexity of matching problem, the epipolar constraint provides an efficient strategy that corresponding points only have horizontal offsets. Once the image pairs have been rectified, the points $P_l$ and $P_r$ satisfy that $u_r = u_l - d, v_r = v_l$, where $d$ is the disparity value between the two points. Assuming that the left and right images are
on the same plane, the following formula can be obtained from the triangular geometry

\[
\begin{aligned}
& u_1 = f \frac{u'c}{z'}, v_1 = f \frac{v'c}{z'} \\
& u_r = f \frac{(u'c - B)}{z'}, v_r = f \frac{v'c}{z'}
\end{aligned}
\]

where \((u', v', z')\) is the 3D coordinate of \(P\), \(B\) is the baseline distance, and \(f\) denotes the focal length of two cameras. The disparity \(d = u_1 - u_2 = fB/\epsilon^c\) is defined as the horizontal position difference of corresponding points in two images.

Therefore, the 3D coordinate of point \(P\) is recovered as follows:

\[
\begin{aligned}
& u' = \frac{B \cdot u_1}{d} \\
& v' = \frac{B \cdot v_1}{d} \\
& z' = \frac{B \cdot f}{d}
\end{aligned}
\]

For isolating switches, when they are in condition “closed,” the measured angle between the switch arm and insulator will be in a certain range. If not, it may indicate that the switch is not fully closed well. The position of the middle line and the corresponding depth information are combined, then the end-to-end coordinates of the middle line can be mapped into 3D space, that is, \((X_{3Dc}, Y_{3Dc}, Z_{3Dc})\), \((X'_{3Dc}, Y'_{3Dc}, Z'_{3Dc})\). Take the horizontal vector \(b\) as the reference and the vector satisfies \(|b| = 1\). The angle is calculated between the center line and the horizontal vector

\[
\theta = \arccos \frac{a \cdot b}{|a||b|}
\]

where \(a = (X_{3Dc} - X'_{3Dc}, Y_{3Dc} - Y'_{3Dc}, Z_{3Dc} - Z'_{3Dc})\) denotes the vector of the center line of switch arm. \(\arccos(\cdot)\) is the anti-cosine function. \(\theta\) is the angle between the horizontal direction and switch arm.

**Experimental results**

In this section, the performance of the proposed stereo matching method is evaluated on the images of the Middlebury stereo benchmark and isolating switches, respectively. To further compare with the proposed method, seven state-of-the-art methods are chosen. In addition, the center line extraction of switch arm leverages the Hough transform and prior constraints, and the evaluation of this content is also illustrated.

**Evaluation on Middlebury v.3 data set**

For evaluation of the proposed stereo matching algorithm, other seven algorithms are selected to compare on the Middlebury v.3 data set which is challenging and representative. The seven methods are weighted cost propagation with smoothness prior (WCPSP), iterative color-depth (MSTCD2), DDL, iterative guided filter (IGF), improvement of stereo matching (ISM), pyramid stereo matching network (PSMNet), and second-order semi-global matching (SGM-SO), respectively. In particular, WCPSP method enforces local smoothness during the weighted cost aggregation on horizontal tree structure, and the smoothness constraints contribute to smooth disparity map, especially on curved surfaces. SGM-SO method utilizes angle direction priors-based second-order smoothness constraint to improve the matching accuracy in weak-textured regions. Then, slanted plane iterative optimization is performed to optimize disparity maps.

The detailed evaluation results of 15 Middlebury v.3 image pairs in non-occluded regions are shown in Table 1. The average error rate of each method is calculated. Our method holds the best performance in accuracy. The average error rate of our method is the lowest among evaluated methods. PSMNet does not work well in the data set, and the average mismatching rate reaches 19.16%, which is twice of the proposed method. Compared to WCPSP, DDL is significantly better due to dictionary learning-based matching cost. However, by implementing data- and structure-driven cost volume fusion and weighted- and edge discriminated-smoothness prior, our method achieves more accurate disparity maps and reliable depth boundaries in comparison with DDL.

For qualitative analysis, Figure 4 presents the disparity maps of Adirondack, ArtL, MotorcycleE, Piano, Pipes, and Playroom. It can be clearly seen that the disparity maps of the proposed method are close to the ground truth. As for MotorcycleE, the disparity map calculated by our method is the best especially on the edge of the motorcycle. Even if the performance of IGF and ISM method is comparable, these methods are not very effective in weak-textured area due to the limitation of local methods. For Piano, fine and smooth results are obtained on piano and stool by WCPSP and the proposed method because of the smoothness constraint. However, WCPSP has a large number of mismatched pixels at top right corner. It may be because the matching cost function is not very competitive for illumination. It can be concluded that the proposed method keeps the distinguished results among these methods especially in the weak-textured and the boundary regions.

**Center line extraction of switch arm**

One pivotal operation in calculating switch angle is the extraction of the center line of switch arm. Since the disparity values of arm edges which locate at depth discontinuities are unstable, it’s indispensable to acquire disparity values that are more stable in the center line of the arm.

In the proposed method, random forest-based edge detection and Hough transform combined with feature
constraints are adopted. Figure 5 presents the results of line detection with different configurations. Images with low light, blurred details, or complex backgrounds are selected for line detection. To prove the effectiveness of the proposed method, four different experimental settings for detecting the edge of the switch arm are adopted. It can be seen from the first row of Figure 5(a) and (c) that when constraints are not enforced, the random forest method could obtain more accurate edges than Canny operator, which is indicated at the lower edge of the switch arm. Then, the center line in Figure 5(d) is obviously more stable than Figure 5(b). For the second and third rows, there are more complex scenarios in the background where chaotic lines occur. In this case, when the constraints are added, the edges of switch arm can be successfully detected, which reveals the positive effect of the proposed feature constraints in combing random forest-based edge detection.

### Switch images

**Stereo calibration.** To recover the 3D information of the object, we adopt Zhang’s calibration method for obtaining the camera intrinsic and extrinsic parameters. The calibration accuracy in binocular stereo vision is a crucial factor and will affect the accuracy of subsequent 3D reconstruction. The intrinsic and extrinsic parameters of the binocular camera calibration are presented in Tables 2 and 3, respectively. $f_x$ and $f_y$ are horizontal and vertical focal lengths, $(u_0, v_0)$ is the center coordinates in the image coordinate system, $k_1$ and $k_2$ are radial distortion parameters, and $p_1$, $p_2$ are tangential distortion parameters.

In the extrinsic parameters, $R$ and $T$ are the pose relationship between binocular cameras, $R$ is the rotation matrix and $T$ is the translation vector. In general, when the cameras are placed horizontally, $R$ is an identity matrix, which means no rotation and scaling. The first parameter of the $T$ matrix is the base distance $B = 9.466$ cm in this article.

**Measurement analysis.** As not all the codes of the compared methods such as IGF and ISM methods are available, the quantitative results in “Evaluation on Middlebury v.3 data set” section of these compared methods are obtained from Middlebury homepage directly. In this section, we utilize four stereo matching methods (i.e., WCPSP, MSTCD2, DDL, and SGM-SO) to analyze the performance of measurement in comparison with our method. These techniques can be used to detect the angle of isolating switch in the power system. Generally, when the angle is 0° or within one preset threshold, the switch arm is regarded as fully closed and operating smoothly. As the accuracy of the angle detection is chiefly affected by angle and distance, the stereo images are captured under different distances and angles using binocular industrial cameras.

Figure 6 presents the captured image pairs under three angles (0° (fully closed), 5°, and 11°) and three distances (1 m, 1.5 m, and 2 m). WCPSP, MSTCD2, DDL, SGM-SO, and the proposed method are employed to generate the disparity map for each stereo pair. In Figure 7, we show the disparity maps of these five stereo matching methods under 1.5 m distance with three different angles (0°, 5°, and 11°). The matching cost function of WCPSP only relies on the linear accumulation of traditional truncated absolute difference (AD) and gradient content, which is sensitive to pattern regions and image noises and is inferior to learning-based dissimilarity measure. Visually, it can be seen from Figure 7(a) that there are obvious error matching points in the leftmost and background regions. In nonlocal MSTCD2 method, minimum spanning tree structure is built for each

### Table 1. Mismatching rate (%) on 15 Middlebury v.3 stereo pairs by eight algorithms.

| Stereo pairs | WCPSP | MSTCD2 | DDL | IGF | ISM | PSMNet | SGM-SO | Proposed |
|--------------|-------|--------|-----|-----|-----|--------|--------|----------|
| Adirondack   | 8.06  | 11.42  | 2.44| 6.62| 6.88| 11.03  | 3.80   | 2.54     |
| ArtL         | 10.24 | 9.99   | 9.22| 9.56| 10.05| 23.39  | 12.18  | 8.14     |
| Jadeplant    | 25.12 | 16.43  | 12.48| 16.21| 16.91| 37.25  | 16.91  | 11.54    |
| Motorcycle   | 4.84  | 3.87   | 4.88| 4.82| 4.96| 8.00   | 6.06   | 4.83     |
| MotorcycleE  | 6.92  | 7.58   | 4.24| 4.51| 4.74| 10.14  | 5.80   | 4.19     |
| Piano        | 17.86 | 19.29  | 9.92| 14.68| 14.47| 17.54  | 11.18  | 8.69     |
| PianoL       | 22.18 | 27.16  | 18.60| 21.09| 21.35| 29.46  | 19.93  | 15.59    |
| Pipes        | 12.33 | 9.61   | 8.50| 10.25| 10.53| 25.56  | 11.58  | 9.17     |
| Playroom     | 13.54 | 16.12  | 9.18| 11.42| 11.24| 18.29  | 16.44  | 9.73     |
| Playable     | 31.13 | 33.54  | 11.57| 28.63| 28.68| 10.50  | 9.06   | 11.24    |
| PlayableP    | 5.33  | 10.64  | 6.37| 11.70| 11.71| 10.16  | 8.10   | 5.85     |
| Recycle      | 3.42  | 6.44   | 4.17| 5.85| 6.20| 6.95   | 4.10   | 3.68     |
| Shelves      | 26.43 | 28.83  | 25.38| 27.28| 25.42| 40.38  | 33.00  | 24.80    |
| Teddy        | 2.53  | 5.06   | 3.82| 3.86| 3.91| 8.24   | 5.68   | 3.25     |
| Vintage      | 24.95 | 36.61  | 17.68| 21.61| 21.02| 30.42  | 20.46  | 16.25    |
| Average error| 14.33 | 16.17  | 9.90| 13.21| 13.20| 19.16  | 12.28  | 9.30     |

*aError threshold is 2.0 pixel and the non-occluded regions are evaluated.

WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; IGF: iterative guided filter; SGM-SO: second-order semi-global matching; PSMNet: pyramid stereo matching network; MSTCD2: iterative color-depth.
Figure 4. Disparity maps of the Middlebury data sets: (a) left image, (b) ground truth, (c) WCPSP, (d) MSTCD2, (e) DDL, (f) IGF, (g) ISM, (h) PSMNet, (i) SGM-SO, and (j) the proposed method. The data below the image are the mismatching rate (%). From the first row to end, the image pairs are ArtL, Jadeplant, MotorcycleE, Piano, and Shelves, respectively. WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; IGF: iterative guided filter; SGM-SO: second-order semi-global matching; PSMNet: pyramid stereo matching network; MSTCD2: iterative color-depth.
pixel and competitive performance against guided image filter on both efficiency and accuracy has been illustrated in this method. However, as the minimum spanning tree structure would become no longer unique in low-textured regions, error disparity values will be obtained. This is consistent with the appearance of black holes in Figure 7(b). In addition, it is hard to guarantee excellent performance in slanted surfaces because cost aggregation is performed on independent cost slices. Instead, our method performs cost aggregation based on edge discrimination strategy and cost slices are incorporated adaptively such that the slanted or curved surfaces can be handled properly. SGM-SO method has a "false obesity" phenomenon between the insulators, which is mainly because of excessive slanted plane iterative optimization. Moreover, compared with DDL, our method has less scanline effects and better distinction on objects edges, especially the edges of the switch arm. It can be concluded that even for these most challenging data, our method remains keeping the best performance in disparity accuracy.

In addition, the quantitative evaluation is successively presented according to two aspects: (i) length comparison of the switch arm under different distances and angles and (ii) accuracy comparison of the angle under different distances and angles. The calculation of the arm’s length and angle is based on 3D coordinate transformation, which is shown in formula (17). Table 4 presents the length comparison of the switch arm calculated in 3D space by five stereo matching methods and nine different situations with regard to distance and angle. In actual measurement, the outermost two screws on the arm are regarded as the end points of arm and the true length is 20.8 cm. As can be seen from the table, our method can estimate accurate and smooth disparity map and the data are closer to the true length in comparison to other methods. In WCPSP, the data are abnormal in condition 1.5/(5,11) and 2/11 because there are outliers in disparity, resulting in huge disparity jump among adjacent pixels. Table 5 illustrates the angle calculation of switch arm by five stereo matching methods. At the angle of 0° and distance of 1 m, 0.95° and 0.60° are calculated by WCPSP and SGM-SO methods, respectively. Compared to the two methods, our method has great advantages and the accuracy is increased nearly by 10 times.

For more intuitively demonstrating the accuracy of each method, absolute errors are calculated. The absolute error $\Delta e$ is defined as follows

$$\Delta e = |e_m - e_c|$$  (19)

Table 2. Intrinsic parameters of camera.

| Parameters | $f_x$ | $f_y$ | $u_0$ | $v_0$ | $k_1$ | $k_2$ | $p_1$ | $p_2$ |
|------------|-------|-------|-------|-------|--------|--------|--------|--------|
| Left camera | 1034.38 | 1035.31 | 393.30 | 294.97 | $-5.69 \times 10^{-1}$ | 3.73 | $-5.52 \times 10^{-3}$ | $-2.83 \times 10^{-2}$ |
| Right camera | 1036.16 | 1036.16 | 427.45 | 276.88 | $-5.06 \times 10^{-1}$ | 0.71 | $-3.18 \times 10^{-3}$ | 5.21 $\times 10^{-5}$ |

Table 3. Extrinsic parameters of camera.

| Parameters | R X | R Y | T |
|------------|-----|-----|-----|
| [0.9990, -0.0160, -0.0457] | [-94.66] |
| [0.0165, 1.0000, 0.0109] | [-0.262] |
| [0.0455, -0.0117, 0.9990] | [-0.661] |

In addition, the results of center line detection with four kinds of experimental settings: (a) Canny, no feature constraint; (b) Canny, feature constraint; (c) random forest, no feature constraint; and (d) random forest, feature constraint. The red lines are the detected lines without constraints, while green lines are detected with constraints. If parallel lines do not exist, the longest line is chosen. The yellow line is the center line extracted by the proposed algorithm.
where $e_m$ is true value and $e_c$ is the calculated value.

The absolute error line chart of Tables 4 and 5 are shown in Figure 8(a) and (b), respectively. As can be seen from Figure 8, WCPSP has the worst performance. With distances increase, the absolute errors become unstable and are more likely to be worse. This is apparent according to the upward trend curves by the five methods. However, our method can still control the errors of length and angle within 0.35 cm and 1.16° and keep the most stable performance.

Figure 6. Binocular images of isolating switch under different angles and distances. Columns from left to right: low (0°), medium (5°), and high (11°) angle. Rows from top to bottom: short (1 m), medium (1.5 m), and long (2 m) distance. (a)-(c) are images captured under 1 m distance with three different angles. (d)-(f) are images captured under 1.5 m distance with three different angles. (g)-(i) are images captured under 2 m distance with three different angles.

Figure 7. Disparity maps of five methods at 1.5 m distance: (a) WCPSP, (b) MSTCD2, (c) DDL, (d) SGM-SO, and (e) the proposed method. Rows from top to bottom represent 3°, respectively. WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; IGF: iterative guided filter; SGM-SO: second-order semi-global matching; MSTCD2: iterative color-depth.
Conclusion

To deal with the safety hazard caused by switch disconnection in power system, an efficient binocular stereo vision-based state measurement system is proposed in this article. The system is carried out in the aspects of camera calibration and rectification, improved SGM algorithm, and the center line extraction of switch arm. In particular, data- and structure-driven cost volume fusion and weighted- and edge discriminated-smoothness prior are proposed in SGM-based stereo matching. The cost volume fusion strategy can neutralize the shortcoming of single data-driven cost and improve the robustness of dissimilarity measures. Meanwhile, the small-weight-accumulation problem in weak-textured regions can be suppressed effectively by the proposed edge weighted tree structure, and the discriminated-smoothness prior construct appropriate regularization terms by enforcing different penalties on different cost layers such that both the slanted planes and disparity discontinuities can be handled well. In terms of line detection, after the rough positioning of the isolating switch by implementing SVM algorithm and Hog eigenvector of switch, random forest method is employed to detect the contours of switch image due to its stability in combination with Hough transform. Then, feature-

| Table 4. Length calculation of the switch arm under different distances (cm) and angles (°). |
|-----------------------------------------------|
| Distance/angle | 1/0 | 1/5 | 1/11 | 1.5/0 | 1.5/5 | 1.5/11 | 2/0 | 2/5 | 2/11 |
| WCPSP          | 20.73 | 20.64 | 20.62 | 20.53 | 139.69 | 131.15 | 20.05 | 20.66 | 54.01 |
| MSTCD2         | 20.35 | 20.22 | 20.18 | 20.27 | 20.19 | 21.34 | 19.92 | 19.48 | 19.34 |
| DDL            | 20.47 | 20.64 | 20.42 | 20.53 | 20.67 | 20.16 | 19.59 | 19.52 | 20.3  |
| SGM-SO         | 20.38 | 20.57 | 20.29 | 20.63 | 20.40 | 19.98 | 19.78 | 19.56 | 20.18 |
| Proposed       | 20.92 | 20.83 | 20.81 | 20.83 | 20.70 | 20.45 | 20.95 | 20.66 | 20.74 |

WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; SGM-SO: second-order semi-global matching; MSTCD2: iterative color-depth.

| Table 5. Angle calculation of isolating switch under different distances and angles. |
|-----------------------------------------------|
| Distance/angle | 1/0 | 1/5 | 1/11 | 1.5/0 | 1.5/5 | 1.5/11 | 2/0 | 2/5 | 2/11 |
| WCPSP          | 0.95 | 5.07 | 11.30 | 0.37 | 14.76 | 13.23 | 1.10 | 6.16 | 8.78 |
| MSTCD2         | 1.28 | 7.23 | 12.08 | 0.37 | 7.14 | 11.68 | 0.32 | 7.86 | 9.88 |
| DDL            | 0.97 | 5.07 | 12.30 | 0.43 | 6.01 | 11.83 | 2.36 | 7.35 | 11.20 |
| SGM-SO         | 0.60 | 6.26 | 12.08 | 0.68 | 7.14 | 11.68 | 1.38 | 7.86 | 12.02 |
| Proposed       | 0.09 | 5.07 | 11.29 | 0.37 | 4.90 | 11.83 | 1.10 | 6.16 | 11.20 |

WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; SGM-SO: second-order semi-global matching; MSTCD2: iterative color-depth.

Figure 8. Absolute error comparison among five methods: (a) absolute error comparison of switch arm’s length and (b) absolute error comparison of switch angle. WCPSP: weighted cost propagation with smoothness prior; DDL: discriminative dictionary learning; SGM-SO: second-order semi-global matching; MSTCD2: iterative color-depth.
constrained Hough transform is adopted to remove chaotic lines from background and extract the center line of the switch arm. By constructing binocular stereo vision framework, the disparity map obtained by semi-global stereo matching and the center line of switch arm are combined that whether the switch is fully closed can be determined based on triangulation principle. Obviously, accurate measurement is also related to the performance of stereo calibration. We adopt Zhang’s calibration method for calculating the intrinsic and extrinsic parameters of cameras. Experimental results on Middlebury v.3 data set prove that the proposed matching method has great advantages and achieves more accurate disparity maps compared with other advanced methods. Moreover, from the quantitative measurement analysis on switch images, we find that our system can control the measurement errors of length and angle within 0.35 cm and 1.16° and keep the most stable performance.

For the convenience of the experiments and the validation, the current work is carried out only in a reasonably laboratory environment, its robustness in real-world scenarios needs to be verified in future research. Moreover, the dynamic processing of visible sequence images captured from closing process and their further cooperatively processing with infrared images can also be exploited for more effective state continuous monitoring of isolating switches.

Declarations of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key R&D Program of China (no. 2018YFC0406900), the Key Research and Development Program of Jiangsu Province (no. BE2018092), and the State Grid Jiangsu Electric Power Co., Ltd Science and Technology Project (no. J2019104).

ORCID iD

Jinxin Xu https://orcid.org/0000-0002-6363-0424

References

1. Pylvanainen JK, Nousiainen K, and Verho P. Studies to utilize loading guides and ANN for oil-immersed distribution transformer condition monitoring. *IEEE Trans Power Delivery* 2007; 22(1): 201–207.
2. Maraaba L, Al-Hamouz Z, and Al-Duwaish H. Estimation of high voltage insulator contamination using a combined image processing and artificial neural networks. In: *2014 power engineering & optimization conference, PEOCO 2014*, Langkawi, Malaysia, 24–25 March 2014, pp. 214–219. Piscataway, NJ, USA: IEEE.
3. Pernebayeva D, Bagheri M, and James AP. High voltage insulator surface evaluation using image processing. In: *2017 international symposium on electrical insulating material, ISEIM 2017*, Toyohashi, Japan, 11–15 September 2017, pp. 520–523. Piscataway, NJ, USA: IEEE.
4. Xie J, Liu CC, Sforna M, et al. On-line physical security monitoring of power substations. *Int Trans Electr Energy Syst* 2016; 26(6): 1148–1170.
5. Liao S and An J. A robust insulator detection algorithm based on local features and spatial orders for aerial images. *IEEE Geosci Remote Sens Lett* 2015; 12(5): 963–967.
6. Bin Y, Hao C, and Wenguang H. Study on the method of switch state detection based on image recognition in substation sequence control. In: *2014 international conference on power system technology, POWERCON 2014*, Chengdu, China, 20–22 October 2014, pp. 2504–2510. Piscataway, NJ, USA: IEEE.
7. Jin L and Zhang D. Contamination grades recognition of ceramic insulators using fused features of infrared and ultraviolet images. *Energies* 2015; 8(2): 837–858.
8. Junzhou L and Xueyi M. Research of isolator’s status image recognition based on Gabor wavelet and SVM. In: *2014 3rd international conference on advanced computer theory and engineering, ICACTE 2010*, Chengdu, China, 20–22 August 2010, pp. 65–68. Piscataway, NJ, USA: IEEE.
9. Liu Y, Yong J, Liu L, et al. The method of insulator recognition based on deep learning. In: *2016 4th international conference on applied robotics for the power industry, CARPI 2016*, Jinan, China, 11–13 October 2016, pp. 1–5. Piscataway, NJ, USA: IEEE.
10. Xia F, Tie HF, Shi EW, et al. Insulator recognition based on mathematical morphology and Bayesian segmentation. In: *2017 10th international congress on image and signal processing, biomedical engineering and informatics, CISP-BMEI 2017*, Shanghai, China, 14–16 October 2017, pp. 1–5. Piscataway, NJ, USA: IEEE.
11. Zuo D, Hu H, Qian R, et al. An insulator defect detection algorithm based on computer vision. In: *IEEE international conference on information and automation, ICIA 2017*, Macau, China, 18–20 July 2017, pp. 361–365. Piscataway, NJ, USA: IEEE.
12. Potnuru SP and Bhima PR. Curvelet transform based statistical pattern recognition system for condition monitoring of power distribution line insulators. *Innov Electron Commun Eng* 2018; 7: 311–317.
13. Zhang G, Zhang D, Li D, et al. The automatic identification method of switch state. *IJSST* 2016; 17(25): 21.1–21.4.
14. Lu J, Hui L, Zhang W, et al. A condition monitoring algorithm based on image geometric analysis for substation switch. In: *International conference on intelligent computing and Internet of Things*, Harbin, China, 17–18 January 2015, pp. 72–76. Piscataway, NJ, USA: IEEE.
15. Jadin M and Taib S. Recent progress in diagnosing the reliability of electrical equipment by using infrared thermography. *Infrared Phys Technol* 2012; 55: 236–245.
16. Ullah I, Yang F, Khan R, et al. Predictive maintenance of power substation equipment by infrared thermography using a machine-learning approach. *Energies* 2017; 10(12): 1987.

17. Ullah I, Khan RU, Yang F, et al. Deep learning image-based defct detection in high voltage electrical equipment. *Energies* 2020; 13(2): 392.

18. Yang L, Wang BQ, Zhang RH, et al. Analysis on location accuracy for binocular stereo vision system. *IEEE Photonics J* 2018; 10(1): 7800316.

19. Zhifeng L, Ke Z, Zhigang W, et al. 3D pose estimation of large and complicated workpieces based on binocular stereo vision. *Appl Optics* 2017; 56(24): 6822.

20. Wang T, Liu B, Wang Y, et al. Research situation and development trend of the binocular stereo vision system. In: 2017 1st international conference on materials science, energy technology, power engineering, *MEP* 2017, Hangzhou, China, 15–16 April 2017, pp. 020220–1–020220–5. AIP Publishing.

21. Weimin L, Siyu S, and Hui L. High-precision method of binocular camera calibration with a distortion model. *Appl Optics* 2017; 56(8): 2368–2377.

22. Sun J, Ma Y, Yang H, et al. Camera calibration and its application of binocular stereo vision based on artificial neural network. In: 2016 9th international congress on image and signal processing, biomedical engineering and informatics, *CISP-BMEI* 2016, Datong, China, 15–17 October 2016, pp. 761–765. Piscataway, NJ, USA: IEEE.

23. Liu W, Huo J, Zhou X, et al. A novel camera calibration method for binocular vision based on improved RBF neural network. In: 2017 2nd CCF Chinese conference on computer vision, *CCCV* 2017, Tianjin, China, 11–14 October 2017, pp. 439–448. Berlin, German: Springer.

24. Jing X, Li C, Yin X, et al. Calibration and stereo image correction for binocular vision system with divergent optical axes arrangement. In: 2013 6th international congress on image and signal processing, *CISP* 2013, Hangzhou, China, 16–18 December 2013, pp. 729–734. Piscataway, NJ, USA: IEEE.

25. Jie L, Jianxun Z, and Yu D. Dense stereo matching based on region growing. *Robot* 2017; 39(2): 182–188.

26. Taniai T, Matsushita Y, Sato Y, et al. Continuous 3D label stereo matching using local expansion moves. *IEEE Trans Pattern Anal* 2018; 40(11): 2725–2739.

27. Ma H, Zheng S, Li Y, et al. Confidence-based iterative efficient large-scale stereo matching. *Cogent Eng* 2018; 5(1): 1427676.

28. Nie E, Li J, Zhang R, et al. Three dimensional point cloud hole repairing strategy for binocular stereo reconstruction. In: 2017 IEEE international conference on robotics and biometrics, *ROBIO* 2017, Macau, China, 5–8 December 2017, pp. 2456–2461. Piscataway, NJ, USA: IEEE.

29. Johnson-Roberson M, Bryson M, Friedman A, et al. High-resolution underwater robotic vision-based mapping and three-dimensional reconstruction for archaeology. *J Field Robot* 2017; 34(4): 625–643.

30. Fokkinga MM. The Hough transform. *J Funct Program* 2011; 21(2): 129–133.

31. Scharstein D and Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int J Comput Vis* 2002; 47(1–3): 7–42.

32. Yao P, Zhang H, Xue Y, et al. AGO: accelerating global optimization for accurate stereo matching. *Multimedia Model* 2018; 10701: 67–80.

33. Zhu S, Gao R, and Li Z. Stereo matching algorithm with guided filter and modified dynamic programming. *Multimedia Tools Appl* 2015; 76(1): 1–18.

34. Guo ZC and Han P. Graph cut stereo matching based on region consistency. *J Northwest Polytech Univ* 2017; 35(1): 160–163.

35. Wang X, Su Y, Tang L, et al. A combined back and foreground-based stereo matching algorithm using belief propagation and self-adapting dissimilarity measure. *Int J Pattern Recognit Artif Intell* 2018; 32(6): 18.

36. Yin J, Zhu H, Yuan D, et al. Sparse representation over discriminative dictionary for stereo matching. *Pattern Recognit* 2017; 71: 278–289.

37. Zbontar J and LeCun Y. Computing the stereo matching cost with a convolutional neural network. In: 2015 IEEE international conference on computer vision and pattern recognition, *CVPR* 2015, Boston, MA, USA, 8–10 June 2015, pp. 1592–1599. Piscataway, NJ, USA: IEEE.

38. Zhang K, Lu J, and Lafruit G. Cross-based local stereo matching using orthogonal integral images. *IEEE Trans Circ Syst Video* 2009; 19(7): 1073–1079.

39. Hong GS and Kim BG. A local stereo matching algorithm based on weighted guided image filtering for improving the generation of depth range images. *Displays* 2017; 49: 80–87.

40. Fu H, Chen G, Li X, et al. Efficient methods using slanted support windows for slanted surfaces. *IET Comput Vis* 2016; 10(5): 384–391.

41. Hirschmuller H. Stereo processing by semiglobal matching and mutual information. *IEEE Trans Pattern Anal* 2007; 30(2): 328–341.

42. Mei X, Sun X, Zhou M, et al. On building an accurate stereo matching system on graphics hardware. In: 2011 IEEE international conference on computer vision, *ICCV* 2011, Barcelona, Spain, 6–13 November 2011, pp. 467–474. Piscataway, NJ, USA: IEEE.

43. Liu J, Li C, and Wang Z. 3D entity-based stereo matching with ground control points and joint second-order smoothness prior. *Vis Comput* 2015; 31(9): 1253–1269.

44. Liu T, Peng X, and Qiao L. Window-based three-dimensional aggregation for stereo matching. *IEEE Signal Process Lett* 2016; 23(7): 1008–1012.

45. Chuang TY, Ting HW, and Jaw JJ. Dense stereo matching with edge-constrained penalty tuning. *IEEE Geosci Remote Sens Lett* 2018; 15(5): 664–668.

46. Yao P, Zhang H, Xue Y, et al. As-global-as-possible stereo matching with adaptive smoothness prior. *IET Image Process* 2019; 13(1): 98–107.

47. Yang Q. Stereo matching using tree filtering. *IEEE Trans Pattern Anal* 2015; 37(4): 834–846.
48. Xu JX, Li QW, Zhou Y, et al. Texture discrimination-enforced matching cost computation and smoothness-weighted cost regularization for stereo matching. *J Electron Imaging* 2019; 28(5): 053025.

49. Dollár P and Zitnick CL. Structured forests for fast edge detection. In: *2013 IEEE international conference on computer vision, ICCV 2013*, Sydney, NSW, Australia, 1–8 December 2013, pp. 1841–1848. Piscataway, NJ, USA: IEEE.

50. Tribak H, Moughyt S, Zaz Y, et al. Remote QR code recognition based on HOG and SVM classifiers. In: *International conference on informatics and computing, ICIC 2016*, Mataram, Indonesia, 28–29 October 2016, pp. 137–141. Piscataway, NJ, USA: IEEE.

51. Zhang Z. A flexible new technique for camera calibration. *IEEE Trans Pattern Anal* 2000; 22(11): 1330–1334.

52. Qi ZS, Wang Z, Huang JH, et al. Invalid-point removal based on epipolar constraint in the structured-light method. *Opt Laser Eng* 2018; 105: 173–181.

53. Scharstein D and Szeliski R. Middlebury stereo evaluation - ver. 2, http://vision.middlebury.edu/stereo/eval (2002, accessed 22 December 2019)

54. Yang QQ. Local smoothness enforced cost volume regularization for fast stereo correspondence. *IEEE Signal Process Lett* 2015; 22(9): 1429–1433.

55. Yao P, Zhang H, Xue YB, et al. Iterative color-depth MST cost aggregation for stereo matching. In: *2016 IEEE international conference on multimedia and expo, ICME 2016*, Seattle, WA, 11–15 July 2016, pp. 1–5. Piscataway, NJ, USA: IEEE.

56. Hamzah RA, Ibrahim H, and Abu Hassan AH. Stereo matching algorithm based on per pixel difference adjustment, iterative guided filter and graph segmentation. *J Vis Commun Image R* 2017; 42: 145–160.

57. Hamzah RA, Kadmin AF, Hamid MS, et al. Improvement of stereo matching algorithm for 3D surface reconstruction. *Signal Process Image* 2018; 65: 165–172.

58. Chang JR and Chen YS. Pyramid stereo matching network. In: *2018 31st IEEE/CVF conference on computer vision and pattern recognition, CVPR 2018*, Salt Lake City, UT, 18–23 June 2018, pp. 5410–5418. Piscataway, NJ, USA: IEEE.

59. Ni JY, Li QW, Liu Y, et al. Second-order semi-global stereo matching algorithm based on slanted plane iterative optimization. *IEEE Access* 2018; 6: 61735–61747.