

# Visual Defect Inspection for Deep-aperture Components with Coarse-to-fine Contour Extraction

Xinyi Gong, Hu Su, De Xu, *Senior Member, IEEE*, Jiabin Zhang, Lei Zhang and Zhengtao Zhang

**Abstract**— This paper investigates automatic quality inspection for the components with small diameter and deep aperture. An automatic pick-and-place system is constructed, which employs an endoscope to achieve better image quality aiming at the characteristics of the component. A coarse-to-fine contour extraction algorithm with four steps is presented to inspect the component's quality. Firstly, approximate locations of the targets are estimated using faster region-based convolutional neural networks (faster RCNN). Secondly, the corresponding edge image is obtained by using the multiscale probability boundary (mPb) detector. Thirdly, edge enhancement is performed, which is based on Brownian motion model. Fourthly, the corresponding contours are finely extracted by edge grouping. A shape analyzing algorithm is utilized to classify the components based on the extracted contours. Comparison experiments fully demonstrate the superiority of the proposed inspection method over existing methods. Meanwhile, successful inspection results on challenging real-world image data prove that the system is of practical significance to industrial applications.

**Index Terms**— Defect inspection, image processing, edge grouping, coarse-fine positioning, deep-hole component

## I. INTRODUCTION

Quality inspection is a significant concern in many manufacturing processes. Until now, this time-consuming task is still mainly accomplished by human workers and is increasingly becoming the bottleneck of productivity. Recently, the replacement of manual labor using an automatic detection system has attracted a good deal of attention, not only to cut labor costs, but also to increase detection accuracy. Various automatic detection systems are constructed and applied to defect inspections of railway surface [1], Ball Grid Array component [2], texture surface [3], printed circuit boards [4], and flat panel displays [5]. In their special application scenarios, the systems work automatically and efficiently without human intervention. Meanwhile, it should be noted that the use of vision-based measurements (VBM) is almost intuitive in defect inspection, which is of definite advantages of non-contact and thus preventing secondary injury.

Electric connector, as shown in Fig. 1, is a basic component and has found an increasingly wide utilization in multiple areas such as aerospace and military industry [6], whose quality

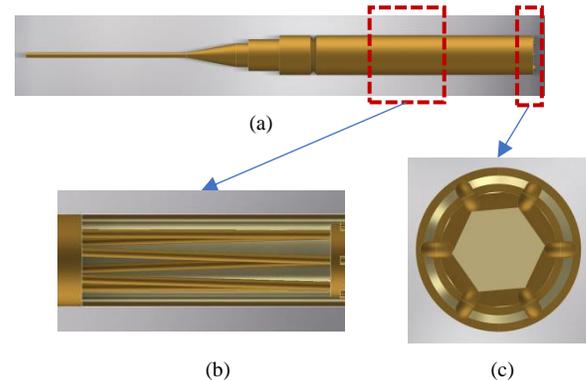


Fig. 1. Mechanical structure of the component, (a) stereo view, (b) partially-sectioned view, (c) cross-section view.

inspection is investigated in this study. The component includes two parts: a small diameter body and an inner sleeve with deep aperture. In the sleeve, several copper wires connecting the top and the bottom of the component are evenly distributed. This kind of component is produced in various forms with the diameters of the apertures ranging from 0.4 mm to 3 mm and the depths of the apertures ranging from 5 mm to 10 mm. The internal structures are also different, i.e., the amount of wires inside the sleeve varies from 5 to 12. Complex internal structure makes the component vulnerable to various damages or defects, leading to serious problems in the conductive performance. Thus, the component's quality needs to be inspected before use. In the inspection, VBM strategy is adopted for the advantages including non-contact and nondestructive. However, the above-mentioned systems do not work in the application due to the natural characteristics of the components with small diameter and deep aperture. Actually, it is even difficult to get clear image of the component with these systems.

In the inspection of electric connector, shape recognition of the internal structure is the major task [7]. According to previous publications, shape recognition methods for VBM can be roughly divided into two categories: those that depend upon recognition of feature points or landmarks [4, 7], and those that depend on the contour and skeleton of the object [1, 2, 3, 5, 8, 9]. For the components considered in the study, only feature points or landmarks cannot reflect the complex structure. In most cases, this would result in wrong judgment and decision. By contrast, richer features are obtained by extracting the contour or skeleton based on which higher inspection accuracy could be achieved. However, due to the small aperture and the complex structure of the component, it is hard to get satisfactory images through the traditional image acquisition system. For example, shadows and noises are almost inevitable in the image and the

Manuscript received, 2018. This work was supported in part by the National Key Research and Development Program of China (2017YFB1302303), National Natural Science Foundation of China under Grants 61503378 and 61733004.

All authors are with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences (CASIA), Beijing 100190, China, and also with the University of Chinese Academy of Sciences, Beijing 101408, China. (e-mail: zhengtao.zhang@ia.ac.cn).

contrast of the image is relatively low. Hence, contours often break into fragments and cannot be completely detected by the existed extraction algorithms [10, 11]. In this sense, it is a primary issue to completely and accurately extract the contour of the object in such a low-quality image.

For this problem, edge grouping and completion algorithms are emerged as solutions. The grouping process collects individual edges together to form continuous contours. In [12], a general grouping approach was proposed which could be used for the grouping of various types of measurements, and to incorporate different grouping cues. In [13], the concepts of link saliency and contour saliency were introduced and used to identify smooth closed contours bounding objects of unknown shape in real image. In [13], the method, named as stochastic completion fields [14, 15], was adopted in the calculation of transition probabilities. Besides, based on the optimization framework pioneered by Jermyn and Ishikawa [16], the method called ratio contour was presented in [17, 18] for extracting salient region boundaries. Especially, the Symmetric Ratio Contour (SRC) method in [18] was for detecting boundaries with symmetry. In [19], the path-based grouping method was proposed, in which connectivity of edge elements via mediating elements was sufficiently explored. Some other publications emphasized the closure principle, which is supposed to play a critical role in human perception. Meaningful attempts addressing this issue were reported in [20, 21], both of which adopted Conditional Random Field (CRF) model to achieve contour closure. However, the differences between them lie in the completion mode of contours across gaps and the potential function employed in the CRF models. With these methods, the accuracy requirement in contour detection could be relaxed. However, in many of the approaches, global optimization is conducted, in which way information from the whole image could be taken into consideration at the same time. The difficulty is, in most cases, the optimization problems are NP-hard, which makes it unlikely to find the global optimum in polynomial time.

An orthogonal line of work regards contours as boundaries of regions, where region segmentation is the primary concern. Regions of constant or slowly varying texture or color are firstly identified, and contours are then straightforwardly determined as closed region boundaries. In the task of region segmentation, deep convolutional neural networks (DCNNs) have recently shown impressive performances. Fully Convolutional Network (FCN) [22] converted the fully-connected layers of VGG into convolutional ones and attempted to harness information from multiple layers to better estimate the object boundaries. Within the same FCN framework, denser score maps were obtained by using atrous (dilated) convolutions [23-25] and the label map was refined with a fully connected CRF in [23, 24]. The networks were shown to produce high-resolution segmentation. The upsampling layer employed firstly in FCN was extended to deep deconvolutional networks in [26, 27]. As demonstrated, the network [26] performed significantly better than FCN although at the cost of complex training and inference. The pressure could be alleviated with the architecture proposed in [27] by discarding the fully con-

nected layers in the VGG net. Recent work in segmentation is the Pyramid Scene Parsing network (PSP net) [28] in which multi-scale pooling features were incorporated to make the final prediction more reliable.

Compared with contour, bounding box is a coarse representation indicating the location and area of an object within an image. Obviously, if bounding box is provided, the efficiency and accuracy of contour detection would be certainly improved by confining the detection to the box. In the definition of bounding box, Girshick *et al.* [29] proposed the method named RCNN. RCNN was the pioneering work to apply DCNN to the detection task, achieving far much better accuracy than existing methods. Despite that, the method takes up 10s to process one image and thus the real-time performance is poor. Several advanced variants [30-31] are proposed to overcome this limitation. [30-31] computed the feature maps from the entire image only once which were 24-102 faster than RCNN while achieving better or comparable accuracy. In [32], a Region Proposal Network (RPN) was introduced to address the bottleneck of region proposal computation in the RCNN. The method in [32] referred to as faster RCNN is the first end-to-end deep learning framework and could serve as a practical system. Besides, additional efforts are devoted to accuracy improvement. For example, [33] suggested position-sensitive score maps to address the dilemma between translation-invariance in classification and translation-variance in object detection. In [34], an architecture, called Feature Pyramid Network (FPN) was proposed which exploits the inherent multi-scale, pyramid hierarchy of DCNNs to construct feature pyramids. As demonstrated therein, state-of-the-art results were achieved by using FPN in faster RCNN.

Along with DCNNs achieving good performance on image related tasks, efforts have been made to apply DCNN to defect inspection. Masci *et al.* [35] creatively proposed a multi-scale pyramidal pooling network for the classification of steel defects. The network could adapt to the images of different sizes to extract multi-scale features and performed far better than traditional methods. A flexible defect inspection framework was proposed in [36], which was general and could be applied to different scenarios. A majority voting mechanism was employed to combine results derived by features from different layers of VGG. Wang *et al.* [37] suggested a fast and robust inspection method that utilized traditional CNN with a sliding window to localize the product damage. While these methods are primarily aimed at the classification task of the defect image, exact sizes and locations of defects are not referred at all. In [38], a method of realizing five types of surface damages was developed based on Faster R-CNN. The method identified the types and the locations of defects simultaneously. In [39], Lin *et al.* built a CNN for light emitting diode chip detection. The defect regions were localized by using a class activation mapping technique without region-level human annotations. Recently, Ren *et al.* [40] proposed a generic approach for surface inspection, in which the features are transferred from a pre-trained deep learning network to overcome the limitation of highly dataset-dependent for traditional CNNs. Besides, Chen *et al.* [9] and Tao *et al.* In [40-41], the authors divided the whole

inspection process into different stages, leading to cascading architectures.

This study investigates defect inspection of electric connector, the component with small diameter and deep aperture. To achieve automatic quality inspection, a pick-and-place system is established. A VBM is designed to capture the image of the component's internal structure which employs endoscope for image quality improvement and is essential to the system. With captured images, the subsequent is to extract the contours, which however, often break into fragments due to noises, shadows and low contrast. To solve the problem, a contour grouping method, the SRC method, is adopted for contour extraction. To improve the performances of the SRC method, the following strategies are designed and applied. Firstly, faster RCNN is utilized to define the bounding boxes of the targets. With the boxes that roughly indicate the location and area, contour extraction is significantly restricted. Secondly, a Brownian motion model is proposed to calculate the probabilities of edges belonging to the contours. The edges having high probabilities are enhanced while others are suppressed most of which arise from noises, shadows and textures. Based on the contours, rules arisen from quantified expertise determine whether the component is qualified to indicate the next action of the system. The contribution of the study can be summarized as follows:

- 1) A VBM including an automatic component pick-and-place system is developed, in which an endoscope is employed to improve the quality of the captured image of the component with small diameter and deep aperture.
- 2) A hierarchical Brownian motion model is developed, based on which an edge enhancement method is developed. The method enhances the edges arisen from the boundaries of the wires and suppresses others.
- 3) A novel contour extraction strategy incorporating coarse locating and fine contour extraction is proposed, which is efficient and universal.
- 4) Arising the rules from quantified expertise to determine component's quality and to indicate the next action, the constructed system achieves automatic inspection in experiments with satisfactory speed and accuracy.

The remainder of this paper is organized as follows. The component and the entire system are introduced in Section II. The Brownian motion model and path-based probability calculation are presented in Section III. In Section IV, the proposed inspection method is detailed. Implementation details are provided in Section V. The effectiveness and superiority of the proposed method is experimentally verified in Section VI. Finally, the study is concluded in Section VII.

## II. SYSTEM OVERVIEW

### A. Introduction of the component

The component to be studied in this paper contains an aperture with diameter of 1.1 mm and depth of 5.3 mm. Inside, it contains five copper wires which are evenly distributed along the aperture and cover from the top to the bottom. The wires' diameter is 0.12mm and the angle of every two adjacent wires

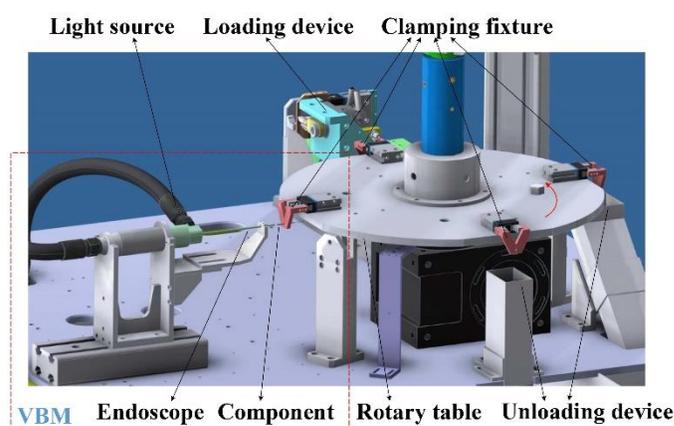


Fig. 2. Automatic component pick-and-place system.

is 72 degree.

### B. Hardware platform

As shown in Fig. 2, a component pick-and-place system is constructed. The system includes four main portions: loading device, the VBM, a rotary table and unloading device. Among them, the VBM is the core of the system that captures the image of the component and then estimates its quality. While conventional lens and lighting patterns cannot satisfy the imaging requirements for the component with small size and deep aperture, the designed VBM employs an endoscope to improve the quality of the captured images. As experimentally demonstrated, the distance between the component and the endoscope should be about 1mm and the angle between the axis of the component and the endoscope should be less than  $\pm 10$  degree. Additionally, other three portions take charge of auxiliary functions such as uploading or offloading the component. With all of these portions together, automatic quality inspection for the component is achieved.

For the components to be inspected, they are delivered into the system by loading device one by one. And for each of them, it is clamped by one fixture placed on the rotary table. Then, the rotary table rotates 90 degrees anticlockwise around its center shaft to align the component with the endoscope. The VBM captures the images of the component and estimates its quality. After that, the table continues to rotate and delivers the component to unloading device. According to the inspection result, unloading device releases the component in different areas. So far, the entire pick-and place process is accomplished. For ease of analysis, typical defects are explained in detail below.

- 1) *Wire inhomogeneous*. The wires' heads are not evenly distributed along the border of the aperture. Their distribution is chaotic.
- 2) *Wire tortuous*. One or more wire(s) has too much bend in the middle.
- 3) *Wire linking*. Two adjacent wires are so close that they are linked together.
- 4) *Wire lacking*. One or more wire(s) disappears, leading to the amount of wires less than five.
- 5) *Wire broken*. Fissure or deletion occurs in the wire(s) while the wire head(s) still exits.

These defects will cause poor connection or even electrical

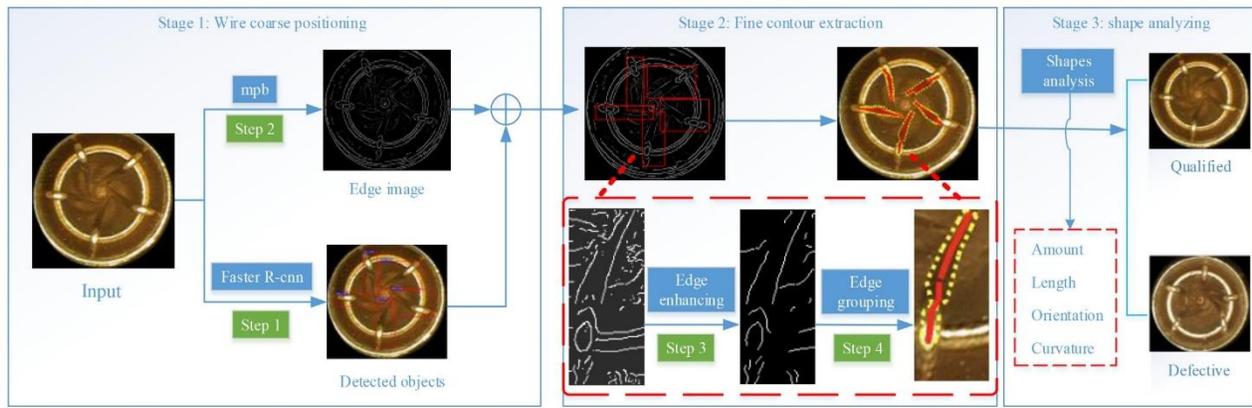


Fig. 3. Flowchart of the inspection algorithm.

equipment failure. However, it is hard to distinguish all these types of defects due to their overall ambiguity. From the images, we conclude that the imaging differences appear obvious and that shadows and noises are almost inevitable in the images. Both of these contribute to increased difficulty in the discrimination of the defects. Existing detection algorithms can hardly deal with the complex problem.

### C. Flowchart of the inspection method

According to the analysis of the typical defects, we can conclude that the morphologies of the wires completely determine the quality of the component. The morphological information could be appropriately represented by the contours. In this sense, extraction of the wires' contours is crucial. However, contours often break into fragments due to shadows, noises and low contrast. To solve the problem, a coarse-to-fine contour extraction algorithm is proposed. With the extracted contours, a shape analyzing algorithm is utilized to classify the components. Based on these concerns, a novel inspection method is proposed and is illustrated in Fig. 3. The method is divided into the three stages.

The wires' positions are coarsely identified in Stage 1 which includes two steps. In Step 1, coarse locations of the wires recognized by the bounding boxes surrounding them are generated by faster RCNN. In Step 2, the corresponding edge image is obtained by using the mPb detector [42]. Combining the edge image and the bounding boxes, several regions of interest (ROIs) are obtained each of which contains the edges for one single wire.

Consequently, the accurate contours of the wires are finely extracted in Stage 2 with another two steps. In Step 3, the edges belonging to the contour are enhanced and others arisen from noises and shadows are suppressed by an edge enhancement strategy. In Step 4, the wires' contours are extracted with a grouping algorithm.

Finally, Stage 3 utilizes a shape analyzing algorithm to inspect the component's quality. Various features of the contours are obtained, such as amount, length, orientation and curvature. Rules are arisen from quantized expertise and utilized to classify the component.

## III. CONNECTIVITY MODELING AND PROBABILITY CALCULATION

### A. Brownian Motion Modeling

The model quantifies the connectivity between pairs of edges by simulating the movement of a particle that moves with constant speed in directions undergoing Brownian motion [13]. Here, we assume that the change in direction obeys Gaussian distribution. And a fixed speed is chosen which performs well in experiments. The process of Brownian motion is represented as

$$\begin{cases} \dot{x} = v \cdot \cos \theta \cdot \tau \\ \dot{y} = v \cdot \sin \theta \cdot \tau \\ \dot{\theta} \sim N(\bar{m}, \delta^2) \end{cases} \quad (1)$$

where  $\dot{x}$  and  $\dot{y}$  specify change in position.  $v$  is the speed of the particle.  $\tau$  is the decay rate.  $\theta$  is the change in direction, which stands for the variation in orientation (*i.e.* curvature) of two edges and we consider it obeying the Gaussian distribution  $N(\bar{m}, \delta^2)$ .

To make the motion model more suitable for discrete pixels in the image, a hierarchical propagation strategy is developed. The initial position of the particle is supposed to be 0 layer, denoted as  $O_0$ . The eight neighbors, or simply neighbors for short in the following, are uniformly called as 1st layer denoted as  $O_1$ . Analogically,  $O_{k+1}$  is expressed as

$$O_{k+1} = C_I \left( \bigcup_{j=0}^k O_j \right) \cap E(O_k) \quad (2)$$

where  $E(A) = \{\cup_i E(a_i) | a_i \in A\}$ ,  $E(a_i)$  is the set of the neighbors of  $a_i$ .  $I$  represents the whole image and  $C_I(A)$  is the complementary set of  $A$  in  $I$ . The moving of the particle from 0 to  $k$ -th layer is completed step by step where, in each step, it moves from  $n$ -th layer to the neighbors on the  $n+1$ -th layer. The strategy is illustrated in Fig. 4.

Given initial and terminal positions, Brownian motion is carried out. Before proceeding, we explain several notations. The initial position and direction are denoted as  $P_0$  and  $\theta_0$ , respectively. The terminal position and direction are  $P_e$  and  $\theta_e$ .

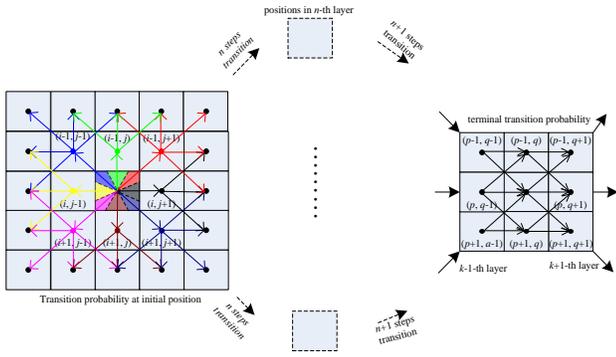


Fig. 4. Transition probability of the Brownian motion.

In the motion,  $P_e$  is supposed to be in  $k$ -th layer. One certain position located in  $n$ -th layer is generally represented as  $P_n$  while its moving direction is  $\theta_n$ . Its neighbors are labeled as  $P_{ni}$ ,  $i=1, 2, \dots, 8$ .  $l(\cdot)$  is the layer in which the position locates. The change of the moving direction is  $\Delta\theta$  which obeys the Gaussian distribution in (2). For the particle moving from  $P_n$ , the next to arrive is

$$P^* = \underset{P_{ni}}{\operatorname{argmin}} \left\| \theta(\overrightarrow{P_n P_{ni}}) - (\theta_n + \Delta\theta) \right\| \quad (3)$$

where  $\theta(\cdot)$  represents the inclined angle of the vector. Subsequently, the particle would proceed next movement step at  $P^*$  in the direction  $\theta(\overrightarrow{P_n P^*})$ . Based on the above definitions, the transition probabilities of one-step,  $N$ -step and terminal are respectively constructed, which form the basis of the calculation of probability that the particle moving from the initial position to the terminal position.

**One-step transition probability.** It refers to the probability of a particle moving from current position to each of its eight neighbors. Based on (3), the transition probability  $Pr(P_{ni}|P_n, \theta_n)$  of the particle moving from  $P_n$  in the direction  $\theta_n$  to  $P_{ni}$  is computed as

$$Pr(P_{ni}|P_n, \theta_n) = \int_{\theta(\overrightarrow{P_n P_{ni}}) - \pi/8 - \theta_n}^{\theta(\overrightarrow{P_n P_{ni}}) + \pi/8 - \theta_n} N(\theta; \overline{m}, \delta^2) d\theta \quad (4)$$

**$N$ -step transition probability.** It refers to the probability that the particle moves from  $P_0$  to  $P_n$  in  $n$ -th layer ( $k-1 \geq n \geq 2$ ). In the moving procedure from  $n-1$ -th to  $n$ -th layer, the particle walks along one or more paths, arriving at the same position, as illustrated in Fig. 4. In the figure, different colored arrows represent different paths and the intersections imply that more than one paths could reach the position. Moreover, it is also found that  $n$ -th layer can only be reached by  $n-1$ -th layer. Therefore, the probability calculation is only concerned with the paths that pass through  $n-1$ -th layer. Based on one-step transition probability,  $N$ -step transition probability  $Pr(P_k)$  can be computed as

$$Pr(P_n) = \sum_{P_c \in E(P_n) \cap O_{n-1}} Pr(P_c) \cdot \int_{\theta(\overrightarrow{P_c P_n}) - \pi/8 - \theta_c}^{\theta(\overrightarrow{P_c P_n}) + \pi/8 - \theta_c} N(\theta; \overline{m}, \delta^2) d\theta \quad (5)$$

**Terminal transition probability.** It refers to the probability that the particle moves from  $P_0$  to  $P_e$  in  $k$ -th layer. Considering the directions of the particle at  $P_0$  and  $P_e$  and using the results in (4) and (5), the probability  $Pr(P_e)$  is represented as

$$Pr(P_e) = \sum_{P_c \in E(P_e) \cap O_{k-1}} Pr(P_c) \int_{\theta(\overrightarrow{P_c P_e}) - \pi/8 - \theta_c}^{\theta(\overrightarrow{P_c P_e}) + \pi/8 - \theta_c} d\theta \cdot \int_{\theta_e - \pi/8 - \theta(\overrightarrow{P_c P_e})}^{\theta_e + \pi/8 - \theta(\overrightarrow{P_c P_e})} d\theta' N(\theta'; \overline{m}, \delta^2) N(\theta; \overline{m}, \delta^2) \quad (6)$$

## B. Probability Calculation

This section considers the probabilities of the detected edges belonging to the contour with a prior marker. For each edge, the probability is modeled as the connectivity to the initial edge. The connectivity mentioned here concerns the path constituted of intermediate edges and could be computed by the connectivity of each two consecutive edges on the path. Inspired by the work of Fischer [19], the path-based connectivity is computed as

$$D_i = \max_{\Gamma_p \in \text{Path}(E_i, E_0)} \max\{Pr_0(E_i), \min_{1 \leq e < |\Gamma_p|} S_{\Gamma_p[e]\Gamma_p[e+1]}\} \quad (7)$$

where  $E_0$  and  $E_i$  are the initial edge and one detected edge in the image, respectively.  $D_i$  is the path-based connectivity of  $E_i$  which represents the probability of  $E_i$  belongs to the contour.  $\text{Path}(E_i, E_0)$  is the set of paths from  $E_0$  to  $E_i$  and  $\Gamma_p$  is one of the paths.  $Pr_0(E_i)$  is the connectivity between  $E_i$  and  $E_0$ , which is calculated with the algorithm described in the previous section.  $S$  is the connectivity matrix. Formula (7) means that the path-based connectivity could be measured by the larger of  $Pr_0(E_i)$  and the lowest connectivity of the path  $\Gamma_{p^*}$ . The  $\Gamma_{p^*}$  is the path with maximum lowest connectivity out of the set of all admissible paths between  $E_i$  and  $E_0$ .

## IV. QUALITY INSPECTION

### A. Wire coarse positioning

Coarse locations of the wires recognized by the bounding boxes surrounding them are identified. As detailed in Fig. 3 (left), faster RCNN is used in this process. The bounding boxes each of which contains a single wire are determined by faster RCNN. The bounding boxes could roughly represent the wires, based on which an initial judgment of whether the component is wire lacking is deduced. However, they would not work in the inspection of other defects such as wire tortuous or wire linking. It is necessary to further extract accurate contours of the wires. For the purpose, edge detection is carried out on the entire image. In this process, the mPb detector [42] is adopted. The detector explores different types of features including brightness, color, texture gradient and multi-scale information to minimize the impact of noise and texture. The method assigns a probability to each pixel, indicating whether it belongs to a boundary. By using the mPb detector, edges arisen from texture and noise could be suppressed. The corresponding edge image is obtained by using the mPb detector. Combining the edge

image with the bounding boxes determined by faster RCNN, several ROIs are identified, each of which contains the detected edges for one single wire. The following contour extraction is conducted independently for each wire in the corresponding ROI. In this way, the disturbances caused by noise and shadows are minimized and the efficiency and accuracy could thus be improved. The determination of coarse locations of the wires is essential and forms the basis for the following steps.

### B. Accurate contour extraction

In the process, the accurate contours and skeletons of the wires are extracted. This process is conducted independently for each ROI and involves the following steps.

**Connectivity Calculation.** For ease of description, the calculation of connectivity  $\varphi$  between edges  $l_1$  and  $l_2$  is taken as an example. The two endpoints of  $l_1$  are respectively denoted as  $p_1$  and  $p_2$  while those of  $l_2$  are  $q_1$  and  $q_2$ . Their orientations are  $\theta_1$  and  $\theta_2$ , respectively. The connectivity between  $p_i$  and  $q_j$  is computed ( $i=1, 2, j=1, 2$ ) which is respectively denoted as  $\varphi_{p_1q_1}$ ,  $\varphi_{p_1q_2}$ ,  $\varphi_{p_2q_1}$  and  $\varphi_{p_2q_2}$ . Then,  $\varphi$  is approximated by

$$\varphi \approx (\varphi_{p_1q_1} + \varphi_{p_1q_2} + \varphi_{p_2q_1} + \varphi_{p_2q_2})/4 \quad (8)$$

In addition, we further explain the calculation of the connectivity between edge endpoints. Without loss of generality, we take  $\varphi_{p_1q_1}$  for instance.  $\varphi_{p_1q_1}$  denotes the probability of a particle moving from  $p_1$  in the direction  $\theta_1$  to  $p_2$  heading the direction  $\theta_2$ . Based on the Brownian motion model, it is computed with (4), (5) and (6). Similarly,  $\varphi_{p_1q_2}$ ,  $\varphi_{p_2q_1}$  and  $\varphi_{p_2q_2}$  can be calculated and  $\varphi$  is then obtained. In this way, the connectivity matrix  $S$  in which the connectivity between every two edges in the image is involved can be constructed.

**Probability Determination.** The probability is regarded as the path-based connectivity between the edge and the initial edge. The initial edge in each ROI is defined in the following way. The starting point of the initial edge is the farthest from the image center among the four vertices of the ROI. Its orientation is parallel to the diagonal that crosses the starting point and points to the image center. Its length is set to be 5. The calculation of the probabilities is based on the connectivity matrix  $S$  and connectivity between the initial edge and edges in the image. We consider this process as a shortest path problem in a directed graph and the Dijkstra's method is adopted to solve this problem. With this algorithm, an edge-in-contour figure in which the gray-values of the edges represent the probabilities of belonging to the contour can then be established.

**Edge Grouping.** We use the criterion SRC algorithm [18] to implement the contour extraction. The SRC method is an edge grouping method for detecting closed boundaries with symmetry when considering the boundary information of proximity and symmetry, and the region information of the enclosed area. The edge-in-contour figure mentioned above is processed by binarization and then is used as the input of the SRC.

### C. Quality inspection based on shape analysis

As illustrated in Fig. 3 (right), the component's quality could truly be judged by analyzing the shape of the extracted contours and skeletons. Different features associated with the shape are involved in the analysis, such as curvature, length and orienta-

tion. Among the features, the curvature is crucial especially in the inspection of the defects of wire linking and wire tortuous. However, it could not be derived directly from the coordinates of the points on the skeletons due to noises and coordinate jumping between discrete points in the image. Curve smoothing is required before accurate curvature estimation.

Inspired by the work of Wang [17], a global optimization strategy is proposed to smooth the curves which are originally represented by a number of points obtained by edge detection. We refer to these points as base points and associate a quadratic spline to each of them. A curve could then be represented with a set of connected quadratic splines. Considering the overlap in the positions of the adjacent splines, a loss function for the whole curve is determined to derive the optimal attributes for the splines. The base points corresponding to the splines of the curve are corrected by gradient descent method. The coordinate of the base point is iterated and updated until it converges or reaches the maximum number of iterations. The proposed smoothing strategy can eliminate the aliasing and noises on the curve while maintaining the primitive shape of the curve.

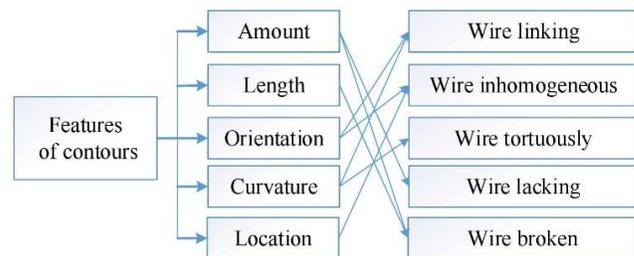


Fig. 5. Classification of defects based on contours' features.

After curve smoothing, the curvature, length and orientation of the skeleton are computed and are combined with amount and location of the contour to judge the component's quality. In the inspection of different defects, different features are employed as illustrated in Fig. 5. Intuitively, the amount determines whether or not the component is wire lacking. In the inspection of wire broken, both the amount and length are involved. The curvatures determine whether or not the component is wire linking or wire tortuous. Based on the orientations, wire linking could be appropriately inspected. In the inspection of wire inhomogeneous, the combination of the orientations and the locations are considered. As discussed above, rules are arisen from quantized expertise and are illustrated in Table I. In the table, several notations need to be explained. Where,  $n$  is the amount of the contours and  $D_{max}$  represents the maximum distance between any two adjacent con-

TABLE I  
REGULATIONS TO DISTINGUISH DEFECTS

Features	Feature Criterion	Defects
Amount	$n \neq 5$	Wire broken or Wire lacking
Length	$l < R/2$	Wire broken
Orientation	$\theta > 90$ or $\theta < 40$	Wire inhomogeneous or Wire linking
Curvature	$C_{max} > 0.15$	Wire tortuous or Wire linking
Location	$D_{max} > R/2$ or $D_{max} < R/4$	Wire inhomogeneous

tours. For one wire, the length of the skeleton is  $l$ .  $C$  represents the curvature of any point on the skeleton and  $C_{max}$  is the maximum one.  $\theta$  is the angle between the orientations of two adjacent skeletons.

## V. IMPLEMENT DETAILS

Our component detection dataset includes 2307 labeled images. 2000 randomly selected images and their flipped ones are used to train faster R-CNN. For anchors, we use three naive scales with box areas of  $64^2$ ,  $128^2$ , and  $256^2$  pixels, and three aspect ratios of 1:1, 1:2 and 2:1. The network was implemented on the PyTorch framework. We chose ResNet101 as our base stone network. We train the model with the Stochastic Gradient Descent (SGD) solver 80k iterations with a fixed learning rate of 0.001 and after 30k iterations reducing the learning rate to 0.0001. We also use a momentum of 0.9 and a weight decay of 0.0001. The training time of Faster R-CNN is about 29 hours. The accuracy is 97.13% while the IoU threshold is 0.5.

We used the mPb detector in edge detection, and SRC algorithm in edge grouping, leaving the parameters at their default values. To determine the parameters in Brownian motion, we begin with a study of the statistical properties with a number of human marked images. The curvatures of approximate 150,000 edge points in 100 images are calculated. The change of angle obeyed Gaussian distribution  $N\sim(0.083, 3.0262)$ . The decay rate  $\tau$  is set to 1.

## VI. EXPERIMENTS AND RESULTS

According to the configuration provided in Section II, an experimental system was set up, as shown in Fig. 6. In this system, the camera consisted of one RM17-10-000-65 endoscope and one CAM-200 CCD. The size of test image captured by the system was  $400 \times 368$  pixels. The whole algorithm was implemented in Python, running on a computer with Ubuntu 16 operating system, Intel Xeon Processor E5-2620, 128-GB RAM and NVIDIA GeForce GTX 1080.

### A. Evaluation metrics and test dataset

The entire inspection algorithm consists of several stages, each of which is experimentally evaluated in the section. Different metrics are employed in the evaluation of different stages, all of which are measured by TP, FP, TN, and FN. As commonly defined, TP is true positive. FP is false positive. TN is true negative and FN is false negative. Utilizing TP, FP, TN, and FN, the precision and recall can then be calculated. The precision is calculated as  $TP/(TP+FP)$  while the recall is calculated as  $TP/(TP+FN)$ . The accuracy is calculated as  $(TP+TN)/(TP+FP+TN+FN)$ . The performance of different stages can be quantified based on the precision and recall.

To validate the performance of the methods, a test data set was constructed including images of both qualified components and typical defective ones. To be fair, in each condition, the amount was set to 50. Totally 300 components were selected. 50 of them were qualified while the rest were the ones with different defects.

### B. Contour extraction

The proposed contour extraction algorithm was verified on

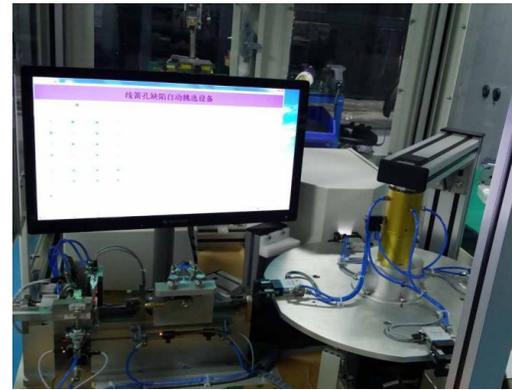


Fig. 6. Experimental system configuration.

the images captured in industrial locale. Hundreds of captured images, in which all possible types of defect introduced in Section II were involved, were employed to quantitatively evaluate the proposed contour extraction algorithm. Several examples are presented in Fig. 7 and from the first row to the sixth row respectively are qualified product, wire inhomogeneous, wire linking, wire tortuous, wire lacking and wire broken. For better illustration, the intermediate results are shown as well. Figs. 7(b) to 7(e) are the results of Steps 1 to 4, respectively. As illustrated in Fig. 3, Steps 1 and 2 constitute Stage 1 while Steps 3 and 4 constitute Stage 2. In the following, each stage is explained and the result is analyzed.

In Fig. 7(b), the coarse positions represented by the red rectangles are accurately detected. We tested Faster R-CNN on the test dataset in which a total of 1387 wires are included. Note that, the performance of Faster R-CNN is the detection result of bounding box. To evaluate this performance, we adopt the evaluation metric in Faster R-CNN [32]. Faster R-CNN utilizes the IoU (Intersection over Union) to describe the divergence between the predicted result and the ground truth. IoU refers to the fraction of the overlap of the prediction and ground truth to the union of the prediction and ground truth. In Faster R-CNN, the author conducted several rules to determine whether there is a target in the anchor. One of the rules is that if the IoU is bigger than the threshold, then it can be determined that there is an object in the anchor. This threshold is called IoU threshold. When the IoU threshold is fixed, the precision and recall is different according to the different confidence threshold. These corresponding precision and recall constitute a PR curve at a certain IoU threshold. The area of the PR curve is the AP (Average Precision) which can be evaluated the detection performance at the current threshold. The AP with different IoU thresholds is listed in Table II. The average AP for 10 IoU thresholds in Table II is 73.41%. As can be seen from Table II, when the IoU threshold is 0.5, the network has the best detection performance and the AP is 95.02%. Therefore, in the proposed method, the IoU threshold in Faster R-CNN is set to be 0.5. The metrics demonstrate that the Faster R-CNN is capable to detect the coarse locations of the wires accurately and robustly. Fig. 7(c) is the detection result of the mPb. As can be concluded, mPb is effective to overcome noise and texture to some extent and even low contrast boundaries of the wires are not neglected. The operator would take about 640 milliseconds to process an image.

Fig. 7(d) is the enhancement result. From the comparison of

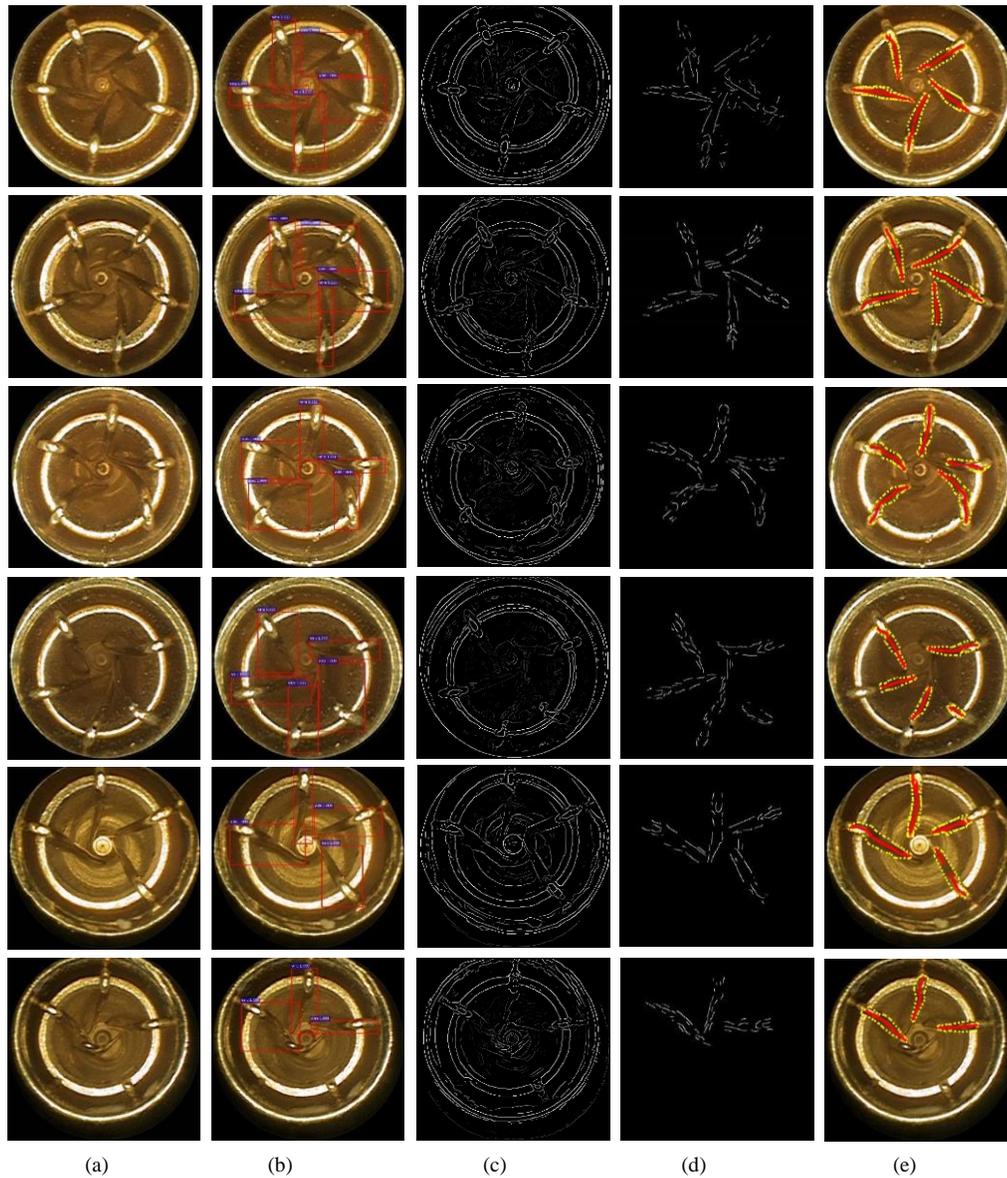


Fig. 7. Detection results. (a) Original image, (b) detection result of step 1, (c) detection result of step 2, (d) enhancement result of step 3, (e) contour extraction result of step 4

Figs. 7(c) and 7(d), the edges belonging to the contours are enhanced while others most of which arise from the noises and shadows are suppressed. After thresholding, the amount of the edges is decreased enormously, leading to a significant improvement in computational efficiency while ensuring accuracy of the contour extraction. In certain areas of the image, the luminance and contrast were so low that the edges were totally neglected. Hence, the contours of the wires broke into fragments. The problem was solved by the grouping algorithm with which the complete and accurate contours were extracted, as shown in Fig. 7(e). In the figure, yellow dashed lines and red solid lines respectively represent the contour and skeleton of

wires. It can be seen that in most cases, the proposed contour extraction algorithm can detect the boundaries and skeletons completely and accurately. In few cases, the algorithm extracts the contour of the wire's body where its head is not included. However, this does not affect the results of the final quality inspection. This is due to the fact that the quality of the component is only influenced by the wires' bodies. As long as the contours of the wires in the aperture are extracted, the quality of the component can then be effectively inspected.

### C. Comparison experiments on contour extraction

Comparison experiments were conducted for further demonstration. First, ablation experiment is conducted for the proposed enhancement method. The proposed contour extraction method is compared with itself but without the enhancement. The comparison results are illustrated in Fig. 8. Figs. 8(b) and 8(c) are the results of the algorithm without and with the enhancement method, respectively. In the figure, the red solid lines represent the skeleton and yellow dashed lines represent

TABLE II

AP OF FASTER R-CNN WITH DIFFERENT IOU THRESHOLD

IoU	0.5	0.55	0.6	0.65	0.7
AP(%)	<b>95.02</b>	95.02	94.60	94.29	93.22
IoU	0.75	0.8	0.85	0.9	0.95
AP(%)	88.81	77.55	54.74	18.76	0.53

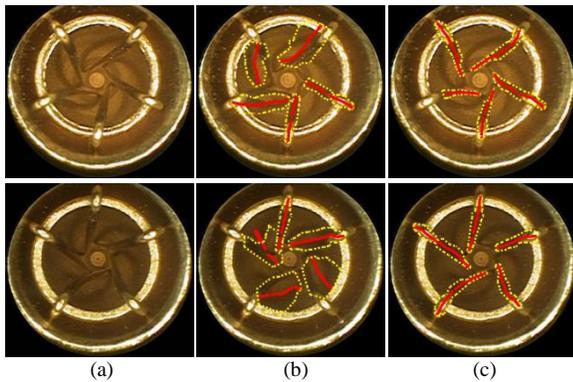


Fig. 8. Comparison results of the proposed algorithm with and without enhancement method. (a) Original image, (b) results without enhancement method, (c) results with enhancement method.

the contours. It can be concluded from the comparison that the enhancement method contributes to the better result shown in Fig. 8(c). The experiment verifies the necessity and effectiveness of the enhancement method.

In addition, the proposed contour extraction algorithm was compared with two deep learning methods, Mask R-CNN and the PSP net. Both the methods achieve state-of-the-art performances but for different tasks. Mask R-CNN is for instance segmentation while the PSP net is for semantic segmentation. In the experiments of Mask R-CNN and the PSP net, the contours were regarded as boundaries of the segmented targets and could be straightforwardly obtained with the results provided

by Mask R-CNN and PSP net. The extracted results were then compared with the result of the proposed algorithm. The instance segmentation dataset and the semantic segmentation dataset were constructed for training Mask R-CNN and the PSP net, respectively. The instance segmentation dataset includes 2429 labeled images. Among them, 1967 randomly selected images were used for training and the remaining 462 images were used for validation. The semantic segmentation dataset includes 6250 labeled images. For training, 5650 images were randomly selected and cropped to the size of  $368 \times 368$  pixels. In the experiments, we trained the two networks with the SGD solver 130k iterations with a fixed learning rate 0.001. We also used a momentum of 0.9 and a weight decay of 0.0001.

The three methods were tested the test dataset. The extracted results of Mask R-CNN, the PSP net and the proposed method are shown in Fig. 9. The yellow solid lines in Figs. 9(b) and 9(c) and the yellow dashed lines in Fig. 9(d) are the contours extracted by Mask R-CNN, the PSP net and the proposed method, respectively. Besides, the red solid line in Fig. 9(d) represents the skeleton of the wire. Actually, in most cases, all the methods achieve satisfactory results. An example is shown in the first row of Fig. 9. But in some cases, Mask R-CNN and the PSP net would lead to poor performance. For the results of Mask R-CNN shown in Fig. 9, one of the extracted contours is incomplete and deviations occur in several of the contours such that they are not in complete accord with the boundaries of the wires. This is due to the low contrast of the target. In addition,

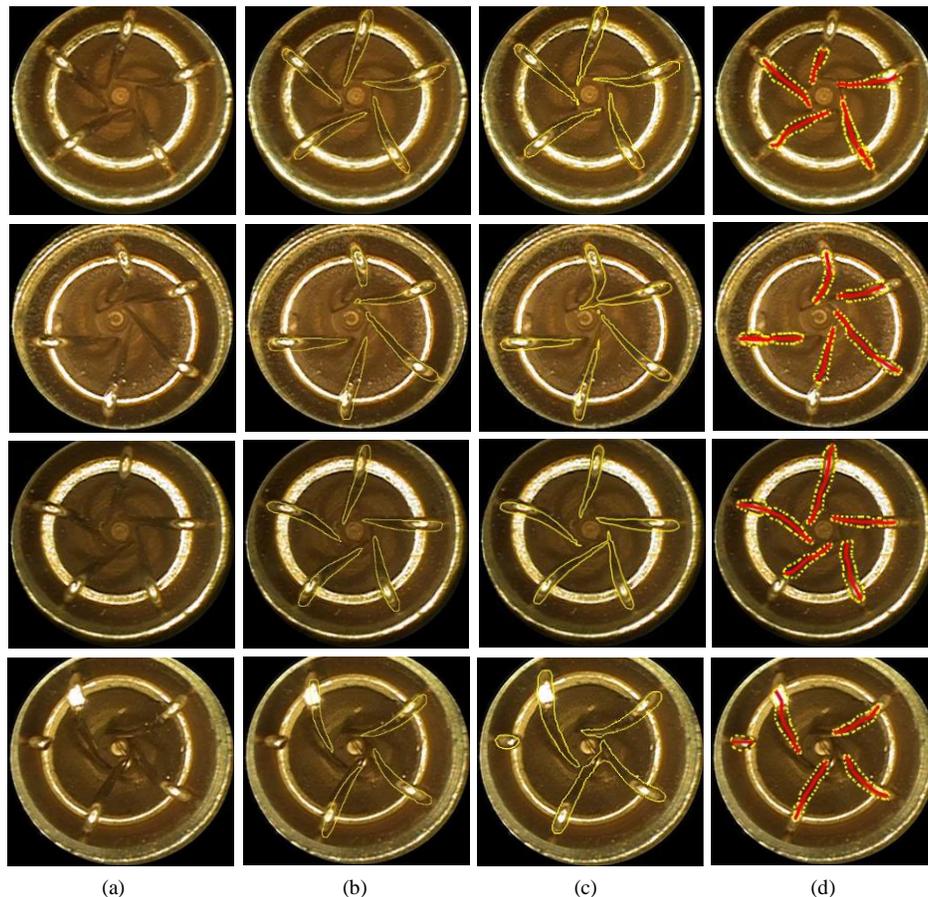


Fig. 9. Comparison results of the proposed algorithm with the Mask R-CNN and PSP net. (a) Original image, (b) results of the Mask R-CNN, (c) results of PSP net, (d) results of the proposed algorithm.

as shown in the last row of the figure, Mask R-CNN neglects the contour of the broken wire. In contrast, the proposed method achieves satisfactory result in all of the cases and thus the superiority is sufficiently demonstrated. For the PSP net, the wires could not be recognized precisely occasionally. One case is particular worthy of note, namely when two wires are intersected, because in such case the PSP net would mistake them for one single target. We would only get one overall contour for the two wires, as shown in the second to fourth rows of Fig. 9 (c). The proposed algorithm could extract each wire’s contour in the emergence of intersecting wires.

Additionally, the performances of these three methods are quantified with the IoU in all test images despite of the true or false segmentation. The IoU of the segmentation results with different methods on the test dataset is illustrated in Table III. As shown in the table, PSP net achieves the best segmentation performance and the performance of the proposed method is better than that of the Mask R-CNN. Since the PSP net cannot separate two connected wires properly, the proposed contour extraction is the most effective aiming at the component investigated in this paper.

Experiment demonstrates that the proposed algorithm occu-

TABLE III  
IoU OF SEGMENTATION RESULTS WITH DIFFERENT METHODS

Method	IoU
Mask R-CNN [43]	0.70
PSP net [28]	0.87
Proposed method	0.78

pies the capability of dealing with low-contrast, noises and shadows in the particular images, while other algorithms do not. Moreover, the proposed algorithm can extract the complete morphological information of the target. Thus, it is more accurate in object representation. It should be emphasized that pixel-level annotated images are required during training Mask R-CNN and the PSP net. Acquiring such data is a time-consuming annotation effort while the bounding boxes needed in training faster RCNN are far easier to collect. Therefore, the proposed contour extraction algorithm is more suitable to solve the problem studied in this paper.

#### D. Defect Inspection

The quality inspection algorithm was tested on the test dataset. The performance of inspection method which is quantized by precision, recall and accuracy refers to the classification result of all test images. The confusion matrix of the inspection result is shown in Fig. 10. The inspection accuracy is 92.67%. As to the proposed method, it is worthy to note that, contour extraction is not necessary for all the inspections. In the inspection of wire lacking or wire broken, the result could be completely determined by the bounding boxes provided by faster RCNN and thus contour extraction is not necessarily performed. For these two types of defects, the mean detection time is 0.15 seconds and the maximum is 0.16 seconds. For other defects, the mean detection time is 1.84 seconds and the maximum is 2.21 seconds. Generally, it is required in industrial

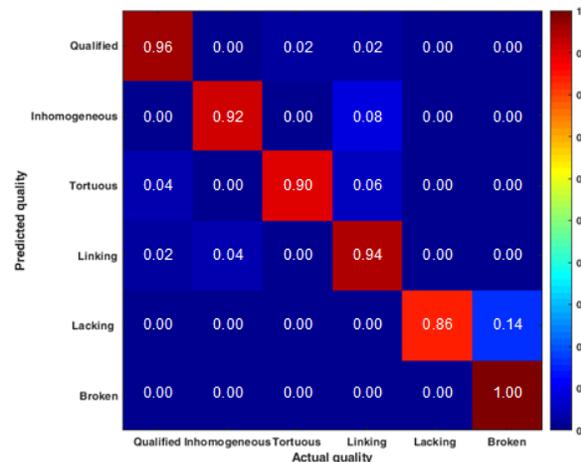


Fig. 10. Confusion matrix of inspection result.

site that the detection time should be no more than 3 seconds. Therefore, both the real-time and accuracy requirements of practical applications are fully satisfied with the proposed method.

The proposed inspection method was compared with several state-of-the-art methods. The method was firstly compared with the contour-based inspection method in which the PSP net was adopted to extract the contours. The comparison was conducted in the following way. The contour extracted by the PSP net was analyzed by the proposed shape analyzing algorithm. The inspection result was obtained according to the regulations laid down in Section IV and was compared with the result of the proposed method. Mask R-CNN was compared with the proposed method in a similar way. Besides, we compared the proposed method with the end-to-end method, the multi-task CNNs [8] which inspects different defects with different network architectures. The precision, recall and accuracy of these four methods are shown in Tables IV, V and VI respectively.

It can be seen from the tables that the proposed method achieves better inspection result than Mask R-CNN mainly due to the higher accuracy of contour extraction. The inspection accuracy of the PSP net is the lowest mainly because the PSP net cannot separate the two connected wires properly leading to inspection errors. The multitask CNN increases the accuracy to over 90% which however, encounters a bottleneck in industrial inspections. Compared with it, the proposed method improved the inspection accuracy by 1.34 percentages. Note that the multi-task CNN determines the quality of the component by training different types of samples in the dataset. It cannot prevent that the inspection result is influenced by other factors, such as brightness while the proposed method can. Compared to multi-task CNN, the proposed method could be conveniently improved to satisfy requirements of the industrial site by parameter adjustment. Moreover, considering the cost of a large number of pixel-level annotations required to the segmentation network such as Mask R-CNN and PSP net, the proposed method is more suitable for industrial applications. With these experiments, the superiority of the proposed method over previous methods is effectively illustrated.

TABLE IV  
PRECISION OF DIFFERENT INSPECTION METHODS

Method	Qualified	Inhomogeneous	Tortuous	Linking	Lacking	Broken	Mean precision
PSP net [28]	95.23%	<b>97.06%</b>	94.44%	0.00%	36.23%	<b>100.00%</b>	70.49%
Mask R-CNN [43]	85.96%	96.89%	<b>95.92%</b>	91.11%	75.76%	<b>100.00%</b>	90.94%
Multitask CNN [8]	<b>100.00%</b>	92.16%	81.48%	91.84%	<b>95.45%</b>	88.68%	91.60%
Proposed method	95.92%	92.31%	90.74%	<b>93.33%</b>	86.21%	<b>100.00%</b>	<b>93.09%</b>

TABLE V  
RECALL OF DIFFERENT INSPECTION METHODS

Method	Qualified	Inhomogeneous	Tortuous	Linking	Lacking	Broken	Mean recall
PSP net [28]	80.00%	66.00%	68.00%	0.00%	<b>100.00%</b>	<b>96.00%</b>	68.33%
Mask R-CNN [43]	<b>98.00%</b>	92.00%	94.00%	82.00%	<b>100.00%</b>	68.00%	89.00%
Multitask CNN [8]	<b>98.00%</b>	94.00%	88.00%	<b>90.00%</b>	84.00%	94.00%	91.33%
Proposed method	94.00%	<b>96.00%</b>	<b>98.00%</b>	84.00%	<b>100.00%</b>	84.00%	<b>92.67%</b>

TABLE VI  
ACCURACY OF DIFFERENT INSPECTION METHODS

Method	Accuracy
PSP net [28]	68.33%
Mask R-CNN [43]	89.00%
Multitask CNN [8]	91.33%
<b>Proposed method</b>	<b>92.67%</b>

## VII. CONCLUSION

The paper addressed the problem of quality inspection for the deep-aperture component, for which an automatic pick-and-place system was established. Aiming at the characteristic of the component, an endoscope was employed to improve the quality of captured images. With the images, a novel quality inspection method was presented. In the method, a coarse-to-fine contour extraction algorithm was proposed to obtain the morphological information of the wires inside the component. The coarse locations of the wires recognized by the bounding boxes surrounding them were firstly generated by faster RCNN algorithm. Then, in each bounding box, the corresponding wire's contour was finely extracted by using a contour grouping algorithm. The real-time and accuracy performances of the contour extraction algorithm were improved with the edge enhancement algorithm. The components were ultimately classified by evaluating the morphological information of the extracted contour. Experiments demonstrated the efficiency and superiority of the proposed inspection method. The entire pick-and-place system was verified in industrial applications to be effective with the average rate on accuracy 92.67% and with the max time-to-detection 2.21 seconds, which fully meets the industrial requirement.

## REFERENCES

- [1] H. Zhang, X. Jin, Q.M.J. Wu, Y. Wang, Z. He, and Y. Yang, "Automatic Visual Detection System of Railway Surface Defects with Curvature Filter and Improved Gaussian Mixture Model," *IEEE Transactions on Instrumentation and Measurement*, (2018, early access).
- [2] L. F. Bai, X.Q. Yang, and H.J. Gao, "A Novel Coarse-Fine Method for Ball Grid Array Component Positioning and Defect Inspection." *IEEE Transactions on Industrial Electronics*, vol.65, no. 6, pp. 5023-5031, 2018.
- [3] S. Mei, H. Yang, Z. Yin, "An Unsupervised-Learning-Based Approach for Automated Defect Inspection on Textured Surfaces," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 6, pp. 1266-1277, 2018.
- [4] D. M. Tsai, Y.C. Hsieh, "Machine Vision-Based Positioning and Inspection Using Expectation-Maximization Technique," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 11, pp. 2858-2868, 2017.
- [5] G. Nam, H. Lee, S. Oh, M.H. Kim, "Measuring color defects in flat panel displays using HDR imaging and appearance modeling," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 2, pp. 297-304, 2015.
- [6] X. Tao, Z. H. Wang, Z. T. Zhang, D. P. Zhang, D. Xu, X. Y. Gong, L. Zhang, "Wire Defect Recognition of Spring-Wire Socket Using Multitask Convolutional Neural Networks," *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 8, no. 4, pp. 689-698, 2018.
- [7] K. T. Maddala, R.H. Moss, W.V. Stoecker, "Adaptable ring for vision-based measurements and shape analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 66, no. 4, pp. 746-756, 2017.
- [8] Y. Park, I.S. Kweon, "Ambiguous surface defect image classification of AMOLED displays in smartphones," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp.597-607, 2016.
- [9] J. Chen, Z. Liu, H. Wang, A. Núñez, Z. Han, "Automatic defect detection of fasteners on the catenary support device using deep convolutional neural network," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 2, pp. 257-269, 2018.
- [10] P. Arbelaez, J. Pont-Tuset, J. T. Barron, F. Marques, J. Malik, "Multiscale combinatorial grouping," *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 328-335.
- [11] Y. Ming, H. Li, X. He, "Winding number constrained contour detection," *IEEE Transactions on Image Processing*, vol. 24, no.1, pp. 68-79, 2015.
- [12] A. Amir and M. Lindenbaum, "A generic grouping algorithm and its quantitative analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 168-185, 1998.
- [13] S. Mahamud, L.R. Williams, K.K Thornber, and K. Xu, "Segmentation of multiple salient closed contours from real images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 4, pp. 433-444, 2003.
- [14] L. R. Williams and D. W. Jacobs, "Stochastic completion fields: A neural model of illusory contour shape and salience," *Neural Computation*, vol. 9, no. 4, pp. 837-858, 1997.
- [15] K. K. Thornber and L.R. Williams, "Analytic solution of stochastic completion fields," *Biological Cybernetics*, vol. 75, no. 2, pp. 141-151, 1996.
- [16] I. H. Jermyn and H. Ishikawa, "Globally optimal regions and boundaries as minimum ratio weight cycles," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1075-1088, 2001.
- [17] S. Wang, T. Kubota, J.M. Siskind, and J, "Wang. Salient closed boundary extraction with ratio contour," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 546-561, 2005.
- [18] J. S. Stahl and S. Wang, "Globally optimal grouping for symmetric closed boundaries by combining boundary and region information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 3, pp. 395-411, 2008
- [19] B. Fischer and J.M. Buhmann, "Path-based clustering for grouping of smooth curves and texture segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 4, pp. 513-518, 2003.

[20] Y. Ming, H. Li and X. He, "Contour Completion Without Region Segmentation," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3597-3611, 2016.

[21] X. Ren, C.C. Fowlkes, and J. Malik, "Learning probabilistic models for contour completion in natural images," *International Journal of Computer Vision*, vol. 77, no. 1-3, pp. 47-63, 2008.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640-651, 2017.

[23] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," *arXiv preprint arXiv: 1412.7062*, 2014.

[24] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *arXiv preprint arXiv: 1606.00915*, 2016.

[25] F. Yu, and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv: 1511.07122*, 2015.

[26] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," *IEEE International Conference on Computer Vision*, Santiago, Chile, pp. 1520-1528, 2015.

[27] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *arXiv preprint arXiv: 1511.00561*, 2015.

[28] H. S. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, "Pyramid scene parsing network." In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, Honolulu, HI, USA, 2017.

[29] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 580-587, 2014.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *European Conference on Computer Vision*, Zurich, Switzerland, pp. 346-361, 2014.

[31] R. Girshick, "Fast r-cnn." In *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, pp. 1440-1448, 2015.

[32] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, Montreal, Canada, pp. 91-99, 2015.

[33] J. Dai, Y. Li, K. He, J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," In *Advances in neural information processing systems*, Barcelona Spain, pp. 379-387, 2016.

[34] T.Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, "Feature pyramid networks for object detection," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, Honolulu, HI, USA, vol. 1, no. 2, pp. 936-944, 2017.

[35] J. Masci, U. Meier, G. Fricout, and J. Schmidhuber, "Multi-scale pyramidal pooling network for generic steel defect classification," In *Proceedings of the International Joint Conference on Neural Networks*, Dalls, TX, USA, pp. 1-8., 2013

[36] V. Natarajan, T.Y. Hung, S. Vaikundam, L.T. Chia, "Convolutional networks for voting-based anomaly classification in metal surface inspection," In *Proceedings of the IEEE International Conference on Industrial Technology*, Toronto, ON, Canada, pp. 986-991, 2017.

[37] T. Wang, Y. Chen, M. Qiao, and H. Snoussi, "A fast and robust convolutional neural network-based defect detection model in product quality control," *International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9-12, pp. 3465-3471, 2018.

[38] Y.J. Cha, W. Choi, G. Suh, S. Mahmoudkhani, and O. Büyüköztürk, "Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 9, pp. 731-747, 2017.

[39] H. Lin, B. Li, X. Wang, Y. Shu, and S. Niu, "Automated defect inspection of LED chip using deep convolutional neural network," *Journal of Intelligent Manufacturing*, online available.

[40] R. Ren, T. Hung, and K. C. Tan, "A generic deep-learning-based approach for automated surface inspection," *IEEE Transactions on Cybernetics*, vol. 48, no. 3, pp. 929-940, 2018.

[41] X. Tao, D. Zhang, W. Ma, X. Liu, and D. Xu, "Automatic metallic surface defect detection and recognition with convolutional neural networks," *Applied Science*, vol. 8, no. 9, pp. 1575, 2018.

[42] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898-916, 2011.

[43] K. He, G. Gkioxari, P. Dollár, R. Girshick, "Mask R-CNN," *IEEE International Conference on Computer Vision*, Venice, Italy, pp. 2980-2988, 2017.



where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interests include intelligent control and optimization.



where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interests include intelligent control and optimization.



where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interest includes robotics and automation, such as visual measurement, visual control, intelligent control, visual positioning, microscopic vision, and micro-assembly.



where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interest includes robotics and automation, such as visual measurement, visual control, intelligent control, visual positioning, microscopic vision, and micro-assembly.

where he is currently a Professor with the Research Center of Precision Sensing and Control. His current research interests include industrial appearance inspection, computer vision, and weakly supervised learning systems for computer vision tasks.



**Lei Zhang** received the B.Sc. degree in mechanical and automotive engineering from the China University of Geosciences, Beijing, China. He is currently a Mechanical Engineer with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing. His current research interests include mechanical design and mechanical

analysis.



**Zhengtao Zhang** received the B. Sc. degree from China University of Petroleum, Dongying, China, in 2004, and M. Sc. degree from Beijing Institute of Technology, Beijing, China, in 2007, and the Ph. D. degree from IACAS, Beijing, China, in 2010, all in control science and engineering.

He is a Professor in the Research Center of Precision Sensing and Control, IACAS. His research interests include visual measurement, micro-assembly, and automation.