

Received January 9, 2020, accepted January 28, 2020, date of publication February 3, 2020, date of current version February 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971242

# Flexible Printed Circuit Fracture Detection Based on Hypothesis Testing Strategy

DEJIAN LI<sup>1,2</sup>, SHAOLI LI<sup>1,2</sup>, AND WEIQI YUAN<sup>1</sup>

<sup>1</sup>Computer Vision Group, Shenyang University of Technology, Shenyang 110870, China

<sup>2</sup>School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110870, China

Corresponding author: Shaoli Li (lishaoli@sut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61271365.

**ABSTRACT** A detection algorithm based on hypothesis testing strategy is presented for the detection of the fracture of FPC track. First, the sandwiched region between any two break points of the track skeleton is hypothesized as the alternative fracture to wait for “testing”. Then, the language semantic judgment (LSJ) algorithm is proposed to verify (“testing”) the hypothetical connecting track which is the union of the two incomplete tracks located by the ends of the alternative fracture. And the structure reasonable degree of the hypothetical connecting track compared with the FPC layout is tested by the longest-relative sequence (LRS) algorithm we proposed. Then the error control technology is used to conduct the posterior verification, to solve the problem that circuits with the same topology structure interfere with the evaluation process of LSJ algorithm. In addition, we have established an automatic optical detection system for FPC of the laptop keyboard. We also conduct experiments with the images captured by this system. The results demonstrate that, compared with the state-of-the-art methods, our proposed algorithm can provide a better detection effect for the track fracture—the Recall, Precision and FNR are 0.9185, 1, and 0.0815 respectively under the condition of without missing detection. And the time consumption is 1.6915 s, which is a relatively short time-consuming compared with the other methods.

**INDEX TERMS** FPC, fracture detection, hypothesis testing, language semantic judgment, longest-relative sequence.

## I. INTRODUCTION

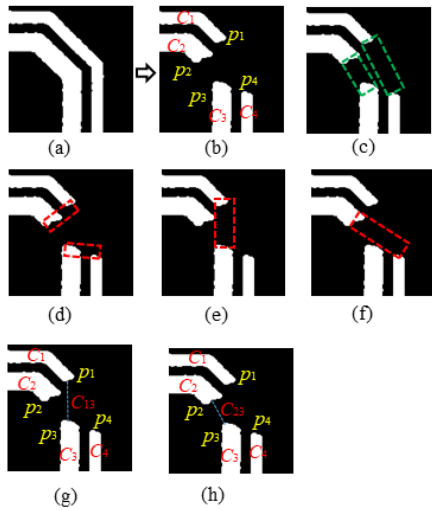
The circuit break is a common defect of printed circuit board (PCB) [1], which damages the electrical function of the circuit directly and affects the quality of electronic products seriously. The traditional quality assurance method of electronic products is detecting the finished PCB and removing the ones with break. But the method of weeding out defective products also causes economic losses. A more ideal method is to locate the broken region and repair the fracture (In this paper, “fracture” indicates the gap of the tracks, and the specific definition is shown in Fig. 1 (c). In addition, to make the description easier and clearer, many concepts which will be used frequently are clarified or defined in Fig. 1 also), so as to avoid resource waste.

At present, the repairing of PCB track fracture is completed by means of “manual + automatic”. First, the position of the fracture is detected manually and the coordinate information

The associate editor coordinating the review of this manuscript and approving it for publication was Mehul S. Raval.

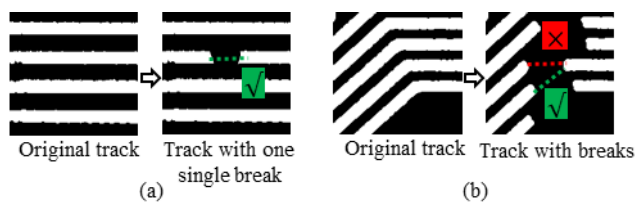
of the fracture is input into the manipulator control system. Then, the manipulator drives a device called silver-paste-pen to complete the reprinting. However, this approach is extremely inefficient, and with the increase of the wiring density, the speed, accuracy and stability of human visual observation can hardly meet the requirements of the automatic production. Therefore, an effective automatic fracture location method is urgently needed. In recent years, based on the rapid development of image processing technology, it becomes a mainstream to detect PCB defects using machine vision. The idea of replacing human vision with machine vision has been widely applied in the detection of various defects on PCB, including the break defect [2].

At present, the detection methods for PCB break defect based on machine vision are mainly divided into three categories: the reference comparison method, the rule-based method and the hybrid method. (1) The basic idea of the reference comparison method is to compare the tested image with the reference image after matching, and the defects are detected through the differences [3]–[7]. But the accuracy



**FIGURE 1.** Clarification or definition of the concepts frequently mentioned hereinafter: (a) Normal circuit track, (b) Circuit track with break defect.  $p_1 - p_4$  are the "broken ends" caused by the break defect, and  $C_1 - C_4$  are the "incomplete track sections". (c) Fracture: the region sandwiched between two broken ends generated from one break, i.e. the region enclosed by the dotted green line. (d)-(f) Pseudo fracture: the region sandwiched between the broken ends generated by different breaks, i.e. the region enclosed by the dotted red line. (g)-(h) Hypothetical connecting track: a complete new circuit track formed by any two incomplete track sections.

of image registration restricts the detection effect [8]–[15], especially for FPC images with flexible deformation. FPC has a good flexible characteristic, and it usually exists in a warped state, especially which is based on polyester film. And this results in the flexible deformation of the acquired images. Thus the accuracy of the image registration is impaired, which affects the performance of the current methods. (2) The rule-based method designs a series of digital image processing rules according to the basic criterion of circuit printing to detect defects. However, a part of these researches can only achieve qualitative inspection [16], [17]. The other parts can realize the location of the fracture, but they only consider the "simple" case that the fracture occurs on a single track. And the "complicated" case that two or more adjacent tracks break at the same time and the fractures interference with each other (as shown in Fig. 2) are not considered [18]–[20]. (3) The definition of the hybrid method is not strict +. It combines the theory of the reference comparison method and the rule-based method [21]–[23]. However, it can only realize the qualitative detection [24]–[26].



**FIGURE 2.** Track fracture with different complexity: (a) "Simple" fracture, and (b) "Complex" fracture.

In summary, the current methods still cannot solve the problem of the detection (or location) of the "complex" fracture in FPC effectively. Therefore, to address the problem, a method based on hypothesis testing strategy and language semantic judgment (LSJ) is proposed in this research. The proposed method is as follows: First, OTSU, geometric parameter constraints, iterative thinning, break point detection and other image processing operations are used to detect the broken end of the track. Then, all the broken ends of the track are matched in pairs to locate the alternative fracture, that is, fracture hypothesis. Then, based on the LSJ algorithm we proposed, the alternative fracture is tested to recognize the true fracture. Finally, posterior verification is carried out to improve the accuracy of the fracture recognition. The main contributions of this research are as follow:

- (1) A novel idea of detecting FPC track break based on hypothesis testing strategy is proposed.
- (2) This research proposes an innovative feature description method which realizes the function of transforming FPC track into a language.
- (3) A FPC track fracture detection algorithm based on language semantic judgment (LSJ) and error control technique is proposed.

This paper is organized as follows: Section II introduces the relevant work and the basic theory. Section III presents the detailed methods of the fracture detection. Section IV presents the experimental equipment and the image data set. The experiments and results are detailed in section V. We conclude this paper in section VI.

## II. RELATED WORK AND BASIC THEORY

### A. RELATED WORK

The related work is presented from three perspectives depending on the three PCB break defect detection methods mentioned above.

#### 1) THE REFERENCE COMPARISON METHOD

The researches of this method have three emphases. (1) Image registration before the comparison. Accurate registration of the images is the prerequisite of the effectiveness of this method. Therefore, many scholars have made in-depth researches on it. However, current methods are still sensitive to external factors such as illumination intensity, mechanical errors and image flexible deformation particularly [8]–[10]. When the substrate of printed circuit is flexible, it is called FPC generally, which is a category of PCB of the most widely used. The image flexible deformation generated by the warping deformation of the FPC affects the registration accuracy strongly, which is not considered. (2) The indirect comparison method of the images. Although current research on image registration has achieved good results, there are still some limitations. Therefore, some researchers try to use an indirect comparison approach instead of the direct comparison to abandon the adverse effects of poor registration. Reference [11] proposes to divide the defects of FPC into global defects and local defects. The global defects

are calculated by histogram matching, and the local defects are calculated by projection matching. However, it can only realize the qualitative detection instead of exact location. A method of matching and comparing an established template with the reference template is proposed by [12], which uses the feature of minimum line width. Reference [13] uses the Hausdorff distance value of the image edges for comparison. But the effect also depends on the accuracy of image matching. (3) Comparison of the image features. This category of method recognizes defects by feature comparison, and the fundamental purpose is still to overcome or avoid the adverse effects of poor image registration. In [14], the author first acquires the gray distribution curve of a parallel line in the image to be tested, and the curve is reconstructed by Fourier transform to remove noise. And then it is compared with the gray distribution curve of the corresponding reference image to detect defects. In [15], the adjoint matrix of the two images to be compared is used to calculate the symmetric matrix, and the rank of the symmetric matrix is used as the similarity measurement index of defect detection. Finally, the defect position is determined by mapping back to the corresponding position through the rank. However, the registration accuracy of the target image and the reference image are still highly required.

## 2) THE RULE-BASED METHOD

According to whether it is only used for qualitative detection, these methods can be divided into two categories. (1) The first category is for the purpose of qualitative detection. Reference [16] calculates the number of connected domains in the image to analyze the number of circuits statistically, so as to determine whether there is any break defect. In [17], the number of the track borders in each small local area is counted, and then the defect is detected qualitatively according to the number of borders perpendicular to the track extension direction. (2) Another category is for the precise location of the defect. In [18], the author uses the distribution state of information entropy in the neighborhood gradient direction of the track area as a rule to detect the fracture, but the “complex” fractures (as shown in Fig. 2) would affect the value of the information entropy. In [19], according to the characteristic that the break will result in broken ends of the track, the fracture is recognized by detecting the break point of the track skeleton line. However, it can only recognize one of the two broken ends of the fracture. And when there are more than one fracture in a local area (as shown in Fig. 2 i(b)), it is difficult to affirm which two broken ends belong to a certain fracture. In [20], the track border is fitted to get the fitting line firstly, and then the break is recognized through the difference between the actual track border and the fitting line, which uses the track characteristic of continuity. But the situation in Fig. 2 (b) is still not resolved.

## 3) THE HYBRID METHOD

Reference [21] extracts the features of LBP and the gradient direction information entropy. And they are combined to train

SVM classifier for the identification of the defects. In [27], features such as gray scale, track width and track area are extracted from the local region of the FPC image and they are input into the neural network for training to realize qualitative detection of the defects. In [22] and [23], the features of LBP and HOG are extracted within the local area of the image to train a classifier for recognizing defects. However, it can only realize the qualitative detection also [24]–[26]. In addition, due to the random state of “complex” fracture, it is difficult to collect a large number of training samples with the same category labels. And this limits the effect of deep learning methods [28]–[30].

In summary, there are some problems in the current research, which is as follows (1) The reference comparison method is sensitive to the flexible deformation of the image. (2) The researches on the rule-based method do not involve the “complex” fracture of two or more adjacent track break at the same time. As is shown in Fig. 2 (b), there are more than two broken ends and fractures, and the requirement of recognizing the pseudo fracture (the red dotted line) and true fracture (the green dotted line) can still not be realized by the current methods. (3) The researches based on the hybrid method take a small local region of the image as the subject, which can only realize the qualitative detection of the defects instead of locating the fracture accurately, especially when multiple fractures are gathered in the local area (as shown in Fig. 2 (b)).

## B. BASIC THEORY

In this section, several key techniques or theories used in this research are briefly introduced. The kernel of the method in this paper, LSJ algorithm, is designed based on the idea of statistical language model (SLM). Therefore, we first introduce the basic theory of SLM in section II-B-1). In the module of probability calculation of the semantic rationality in LSJ algorithm, we refer to the evaluation theory of DNA sequence similarity, so it is introduced briefly in section II-B-2).

### 1) THEORY OF STATISTICAL LANGUAGE MODEL

Statistical language model (SLM) [31] is a mathematical model established to deal with the problem of natural language with contextual characteristics. SLM is originally designed to solve the problem of speech recognition. Different from the traditional mainstream method of determining whether a word sequence conforms to the rules of a certain grammar, it judges the semantic meaning of a sentence from the perspective of statistics. The basic idea of SLM is to determine the rationality of a sentence  $S$  by calculating the probability of its occurrence. Suppose that  $S$  stands for a meaningful sentence consisting of a sequence of words  $w_1, w_2, \dots, w_n$ , where  $n$  is the length of the sentence. Then the probability  $P(S)$  of  $S$  can be described as

$$P(S) = P(w_1, w_2, \dots, w_n) \quad (1)$$

According to conditional probability formula [32] and the Markov hypothesis [33], (1) can be rewritten as

$$P(S) = P(w_1) \cdot P(w_2|w_1) \cdot P(w_3|w_2) \cdots P(w_i|w_{i-1}) \cdots P(w_n|w_{n-1}) \quad (2)$$

where  $i$  represents the sequence number of the words in the sentence. According to (2),  $P(S)$  can be easily calculated when the conditional probability  $P(w_i|w_{i-1})$  is obtained. And according to conditional probability formula,  $P(w_i|w_{i-1})$  is as follows:

$$P(w_i|w_{i-1}) = \frac{P(w_{i-1}, w_i)}{P(w_{i-1})} \quad (3)$$

According to law of large numbers, the relative frequency of  $(w_{i-1}, w_i)$  and  $(w_{i-1})$  appearing in the corpus (which are called  $f(w_{i-1}, w_i)$  and  $f(w_{i-1})$  respectively) are approximately equal to the joint probability  $P(w_{i-1}, w_i)$  and  $P(w_{i-1})$  respectively. And the corpus above is a sentence database constituted by all “reasonable” sentences generated by the corresponding language. And the basic unit of the sentence is called morpheme. Therefore, when the corpus is determined, the relative frequency  $f(w_{i-1}, w_i)$  and  $f(w_{i-1})$  can be obtained through the statistics, as follows.

$$f(w_{i-1}, w_i) = \frac{\#(w_{i-1}, w_i)}{\#} \quad (4)$$

$$f(w_{i-1}) = \frac{\#(w_{i-1})}{\#} \quad (5)$$

where  $\#(w_{i-1}, w_i)$  is the frequency of the two morphemes  $(w_{i-1}, w_i)$  appearing adjacently in the corpus,  $\#(w_{i-1})$  represents the frequency of morphemes  $(w_{i-1})$  appearing in the corpus, and  $\#$  represents the size of the corpus.

Equation (2) describes the simplest binary model of SLM, the more general expressions is as (6), which is called the  $N$ -element model.  $N$  is a natural number and satisfies  $2 \leq N \leq n$ . But in actual application,  $N$  is usually set to 2 or 3.

$$P(S) = P(w_1) \cdot P(w_2|w_1) \cdots P(w_i|w_{i-N+1}, w_{i-N+2}, \dots, w_{i-1}) \cdots P(w_n|w_{n-N+1}, w_{n-N+2}, \dots, w_{n-1}) \quad (6)$$

where when  $i - N + \alpha \leq 0$  ( $\alpha = 1, 2, 3, \dots$ ), the parameter item  $w_{i-N+\alpha}$  does not participate in the operation.

## 2) THEORY OF THE SIMILARITY EVALUATION OF THE DNA SEQUENCE

For a sequence  $S$ , if it is a subsequence of two or more known sequences respectively and it satisfies the condition that it is the longest one in all the subsequences, we call it the longest common subsequence (LCS) of the known sequences [34], [35]. Based on LCS, an effective method to evaluate the similarity of two DNA sequences is defined as follows [36], [37]:

$$H(x, y) = \frac{B(x, y)}{D(x, y)} \quad (7)$$

where  $x$  and  $y$  represent two DNA sequences respectively,  $B(x, y)$  is the length of the LCS of  $x$  and  $y$ , and the definition of  $D(x, y)$  is

$$D(x, y) = \max\{d(x), d(y)\} \quad (8)$$

where  $d(x)$  and  $d(y)$  represent the length of  $x$  and  $y$  respectively. According (7), the closer  $H(x, y)$  is to 1, the more similar these two DNA sequences are, and the more possible that  $x$  and  $y$  belong to the same species. On the contrary, the closer  $H(x, y)$  is to 0, the less similar the sequence  $x$  and  $y$  are.

Finally, we will expound the role of the above two theories in this research here. Since the track fracture is inevitable to produce the broken end, the fracture is the region sandwiched between two broken ends of the track. Therefore, the key for locating the fracture accurately is to recognize the correct broken end pair from one track, that is, to match the two incomplete track sections which belong to the same original track. Inspired by SLM, we creatively transform the problem of matching two incomplete track sections to the problem of calculating the semantic rationality of the sentence. And the sentence is represented by the new track constituted by the two incomplete track sections. And the newly constituted track is represented by a one-dimensional symbol sequence creatively, that is, generating a language. Then, on the basis of DNA sequence similarity evaluation theory, a longest-relative sequence (LRS) model is defined to evaluate the rationality of the language semantics of the new track, and the conclusion of whether the fracture on the new track is true or not is determined.

## III. FPC TRACK FRACTURE DETECTION BASED ON THE STRATEGY OF HYPOTHESIS TESTING AND LSJ ALGORITHM

In this section, we present the detection method based on the strategy of hypothesis testing and LSJ for the track fracture of FPC. The proposed detection method mainly consists of four parts: detection of track broken end, hypothesis of the alternative fracture, hypothesis testing based on LSJ, and posterior verification. The first two parts are used to complete the content of “hypothesis” of the hypothesis testing strategy, and the last two parts are used to complete the content of “testing”. The overall flow of this method is shown in Fig. 3, and the details are presented in the following sections.

### A. DETECTION OF FPC TRACK BROKEN END

Since the occurrence of the fracture will inevitably result in the broken ends of the track, we first detect the broken ends to realize the basic location of the fractures.

The detection of the broken end is implemented by extracting the “break point” of the track skeleton. First, some image preprocessing methods are applied to extract the skeleton of the track. Since the track shows as an obvious bright intensity region in the image, it can be easily extracted by OTSU [38]. Thus, the binary image  $I_b$  is obtained. Then the area threshold  $T_a$  is used to remove the small non-electrical function

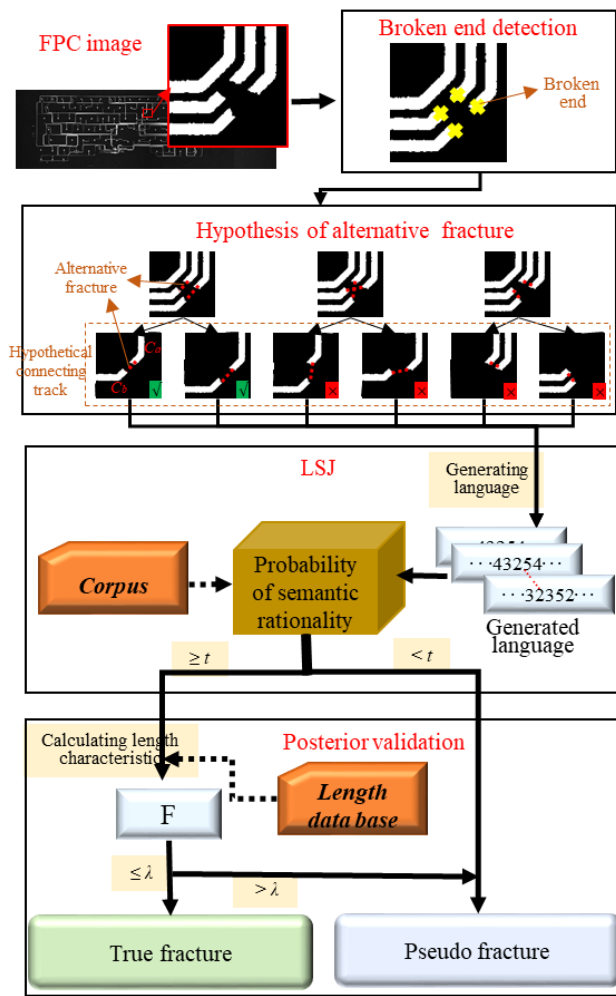


FIGURE 3. Overview of the proposed FPC track fracture detection method.

identifiers and other noise regions for obtaining  $I_c$ . A circular structure element with radius  $r$  is then used to perform the morphological operation on  $I_c$ , so as to detect the circular welding pad  $R_p$ . Subsequently, the circular connection end  $R_h$  is detected through the template matching [39]. And the region of  $I_c$  excepting  $R_p$  and  $R_h$  is the track, which is marked as  $R_{line}$ . Next, the  $R_{line}$  is thinned according to the structural element “L” in the Goly character set [40] to obtain the skeleton  $S_c$ . Finally, burr noise of  $S_c$  is removed by the burr removal algorithm in [19], and the obtained result is marked as  $S_l$ . If  $S_l$  is stored as an image  $I_{S_l}$  with a foreground of 1 and background of 0, the endpoint  $p_{ij}$  of  $S_l$  can be extracted according to

$$\sum_{u=-1}^1 \sum_{v=-1}^1 G(p_{i+u,j+v}) - G(p_{i,j}) = 1 \quad (9)$$

where  $i$  and  $j$  represent the horizontal and vertical coordinates of point  $p_{ij}$  respectively, and  $G(p_{ij})$  is the gray value of  $p_{ij}$  in the image  $I_{S_l}$ .

There are two reasons for the generation of the endpoints of the track skeleton line: the normal end of the track such as the welding pad and the connecting end, and the end of the line caused by the break in the track. Therefore, the break point of the track skeleton line can be extracted as follows:

$$P_{break} = P_{end} \odot (P_{end} \cap (R_p \cup R_h)) \quad (10)$$

where  $a \odot b$  represents the remainder of set  $a$  and set  $b$ , and  $P_{end}$  is the set of points formed by skeleton line endpoints.

### B. HYPOTHESIS OF THE ALTERNATIVE FRACTURE

After the processing of section III-A, all the break points of the track skeleton line are obtained, and the sandwiched regions between these break points all could be the true fracture theoretically, as shown in the red dotted line in the module “Hypothesis of alternative fracture” in Fig. 3. However, only the fractures marked with “√” are true, while the ones marked with “×” are pseudo. Therefore, the method we proposed recognizing the true fracture from the alternative fractures is presented in section III-C.

### C. ALTERNATIVE FRACTURE TESTING BASED ON LSJ

The purpose of testing the hypothetical fracture is to recognize the true ones. And for a true fracture  $F_{ab}$ , it is the region sandwiched between the two incomplete track sections ( $C_a$  and  $C_b$ ) of a certain track after being segmented by  $F_{ab}$ , as shown in the module “Hypothesis of alternative fracture” in Fig. 3. Therefore, the true fracture  $F_{ab}$  can be recognized by judging whether the two incomplete track sections  $C_a$  and  $C_b$  are originally derived from a complete circuit.

In order to verify whether  $C_a$  and  $C_b$  belong to a complete circuit originally, we first generate the hypothetical connecting track  $C_{ab}$  by connecting  $C_a$  and  $C_b$  at the corresponding fracture. And then  $C_{ab}$  is compared with track  $C'$  which comes from the normal FPC image. If  $C_{ab}$  and  $C'$  are identically, or  $C_{ab}$  is part of  $C'$ , then  $C_a$  and  $C_b$  are derived from a complete circuit originally. And the corresponding fracture  $F_{ab}$  is a true fracture. Otherwise,  $F_{ab}$  is a pseudo fracture.

A method called language semantic judgment (LSJ) algorithm is proposed to implement the idea above. The following part is organized as follows: Section III-C-1) presents our innovative feature description method of transforming the FPC track into a language. Section III-C-2) presents the calculation method of linguistic semantic rationality we proposed. Section III-C-3) describes the specific process of the fracture detection based on the theories in sections III-C-1) and III-C-2).







#### 1) VERBALIZATION OF THE FPC TRACK

Language is a symbolic system composed of words according to a certain grammar. Therefore, we illustrate the established verbalization method from the aspects of vocabulary and grammar.

Printed circuit can be regarded as a diagram composed of some basic line structure units connected in accordance with

certain rules. Therefore, we transform the basic line structure unit into vocabulary, that is, morpheme. And the connection rules are transformed into grammar. Then the printed circuits can be expressed in language. The transformation we defined from basic line structure unit to morpheme is shown in Table 1. Through observation and analysis of FPC circuit diagram structure, we split the FPC diagram into six basic structure units. In other words, any common FPC diagram can be represented by these six basic track structure units defined in Table 1. We realize the transformation from basic track structure unit to vocabulary by assigning symbols 1-6 in Arabic numerals to different structure units uniquely, which is the first step of expressing FPC with a language.

TABLE 1. FPC morpheme definition.

NO.	Name of the basic track structure unit	Pattern of the basic track structure unit	Morpheme symbol
1	Weld/Connecting end		“1”
2	Vertical moulding		“2”
3	Right slash		“3”
4	Ledgement		“4”
5	Left slash		“5”
6	“T”/Cross structure		“6”

The second step of FPC verbalization is formulating the grammar rules which are the rules of sentence structure. The grammar rules we proposed are described according to a formalized system which is called the rewriting system and used in the structural analysis of natural language commonly [41], [42]. And the language of track  $C$  (marked with  $L_{ANG}(C)$ ) is represented by a one-dimensional array. The specific rules are as follows:

1.  $L_{ANG}(C) \rightarrow [L(B_1), L(B_2), \dots, L(B_i), \dots, L(B_\zeta)]$
2.  $L(B_i) \rightarrow [L'(B_{i,left}), L'(B_{i,up}), L'(B_{i,right}), L'(B_{i,down})]$
3.  $L'(B_{i,j}) \rightarrow [M_1, M_2, \dots, M_k, \dots, M_6]$
4.  $M_k \rightarrow 1|2|3|4|5|6$

In rule 1,  $L(B_i)$  is the morpheme symbol sequence formed by the track branch segment which is directly connected by the  $i$ 'th bifurcation point  $B_i$  in the circuit. Taking  $B_1$  in Fig. 4 as an example, the directly connected track branch segments include  $a, b$  and  $c$ . In addition, in the rule 1,  $B_i$  and  $B_{i+1}$  satisfies  $\rho(B_i) < \rho(B_{i+1})$ , or  $[\rho(B_i) = \rho(B_{i+1}) \text{ AND } x_i > X_{i+1}]$ . And  $\rho(B_i)$  is defined in (11), and  $\zeta$  represents the number of branch points.

$$\rho(B_i) = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \quad (11)$$

where  $(x_i, y_i)$  represents the coordinate of bifurcation point  $B_i$ . And  $(x_0, y_0)$  represents the coordinate of the upper left corner of the smallest surrounding rectangle with any orientation

of the effective track area in FPC [43], which is called the reference position.

In rule 2,  $B_{i,j}$  ( $j$ =left, up, right, down) represents the track branch directly connected by the bifurcation point  $B_i$  in the direction of  $j$ .  $L'(B_{i,j})$  is the sequence of morpheme symbols formed by the track branch  $B_{i,j}$ . In particular, when  $B_{ij} = B_{i-1,\nu}$  ( $\nu$  = left, up, right, down), rule 3 no longer needs to be performed.

In rule 3,  $M_k$  represents the morpheme value of the  $k$ -th basic track structure unit of  $B_{i,j}$  (denoted as  $B_{LSU}(M_k)$ ) according to Table 1. In addition,  $M_k$  and  $M_{k+1}$  satisfy that  $B_{LSU}(M_k)$  and  $B_{LSU}(M_{k+1})$  are adjacent. And  $M_1=6$ , that is,  $B_{LSU}(M_1)$  is the “T”/crossing structure. 6 represents the number of the basic structure unit.

There is a special case that the track to be verbalized has no bifurcations. Then it can be verbalized from rule 3 directly, that is, the above rules can be simplified as follows:

1.  $L_{ANG}(C) \rightarrow [M_1, M_2, \dots, M_k, \dots, M_6]$
2.  $M_k \rightarrow 1|2|3|4|5|6$

where  $M_1$  is no longer equal to 6, and it only needs  $M_1$  and  $M_\beta$  satisfying the condition of the minimum distance between  $B_{LSU}(M_1)$  and  $(x_0, y_0)$  being greater than that of  $B_{LSU}(M_6)$  and  $(x_0, y_0)$ .

The primary process of FPC verbalization according to the grammar rules defined above is shown in Fig. 4.

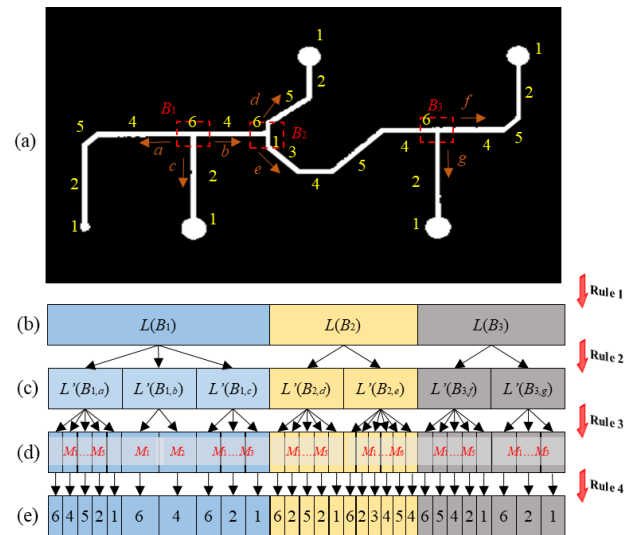


FIGURE 4. Diagram of the process of FPC verbalization: (a) A FPC track, marked C, (b) Intermediate results of rule 1, (c) Intermediate results of rule 2, (d) Intermediate results of rule 3, and (e) Final result  $L_{ANG}(C)$ .

In Fig. 4 (c), we replace symbols “left”, “up”, “right”, and “down” with serial numbers of the track branch segments such as  $a, b, c, \dots$ , so as to represent the whole process more clearly.

## 2) PROBABILITY CALCULATION OF SEMANTIC RATIONALITY BASED ON LRS

In the classical statistical language model, the judgment of sentence semantic rationality is evaluated by the occurrence

probability of the  $N$  adjacent words in the corpus, which is described in (6). When the value of  $N$  in the  $N$ -element model described in (6) is equal to the length of the sentence to be tested, the probability value is the most accurate. But at present, in the field where the application of SLM is relatively mature, such as speech recognition,  $N$  is set to 2 or 3 merely. One reason is that the space complexity and time complexity of  $N$ -element model are almost exponential functions of  $N$ , and the resource consumption increases sharply with the increase of  $N$ . Another reason is that when the number of morpheme is order of magnitude of thousands and the corpus is order of magnitude of millions, a small  $N$  can obtain a very approximate effect compared with a large  $N$ . However, in this paper, the number of morpheme is few (6 different morphemes) and the corpus is small (the size is 1180). This results an obvious difference between the semantic rationality probability calculated by (6) at a small  $N$  and a large  $N$  [44], [45]. And if a large  $N$  is adopted to improve the probability calculation effect of (6), statistical unreliability and unsmooth situation will occur due to the small corpus [46]. Therefore, inspired by the method of DNA sequence similarity evaluation (presented in section II-B) [47], we propose a probabilistic calculation model of Longest-relative sequence (LRS). And it is as follows:

$$P(u) = \max_{v \in \Omega} \frac{B(u, v)}{d(u)} \quad (12)$$

where  $u$  represents the sentence to calculate the probability of its semantic rationality,  $d(u)$  represents the length of the  $u$ ,  $v$  is the sentence in the corpus  $\Omega$ ,  $B(u, v)$  is the length of the longest common subsequence of  $u$  and  $v$ , and  $0 \leq P(u) \leq 1$ . The closer  $P(u)$  is to 1, the more reasonable the semantics of  $u$  is. Oppositely, the closer  $P(u)$  is to 0, the more unreasonable the track structure expressed by sentence  $u$  is.

We explain the corpus  $\Omega$  established in this research here. The circuit in the FPC image is composed of several independent subcircuits, and there is no crossover or connection between subcircuits. Therefore, according to the grammar rules defined in section III-C-1), each individual subcircuit can be expressed in a sentence, and then the set composed of these sentences is the corpus  $\Omega$ .

### 3) FPC FRACTURE INSPECTION BASED ON LSJ

Inspired by the working mechanism of SLM, we transform the problem of judging whether the structure of hypothetical connecting track  $C_{ab}$  is correct to the problem of judging whether the language semantic of  $C_{ab}(L_{ANG}(C_{ab}))$  is reasonable relative to the sentences in  $\Omega$ . Then, based on the reasonable probability of  $L_{ANG}(C_{ab})$ , whether  $F_{ab}$  on  $C_{ab}$  is a true fracture is determined. The specific steps are as follows.

Step1: Generation of language.

According to section III-C-1), the obtained hypothetical connecting track  $C_{ab}$  is expressed with the language  $L_{ANG}(C_{ab})$ .

Step2: Probability calculation of semantic rationality.

According to (12) in section III-C-2), the rationality probability  $P[L_{ANG}(C_{ab})]$  of  $L_{ANG}(C_{ab})$  is calculated.

Step3: Authenticity determination of alternative fracture.

The authenticity of the alternative fracture  $F_{ab}$  on the hypothetical connecting track  $C_{ab}$  is determined as follows:

$$\begin{cases} P[L_{ANG}(C_{ab})] \geq t, & \text{True} \\ P[L_{ANG}(C_{ab})] < t, & \text{False} \end{cases} \quad (13)$$

where True means that  $F_{ab}$  is a true fracture, False means that  $F_{ab}$  is a pseudo fracture, and  $t$  is a threshold whose specific value is determined through the experiments.

The specific experimental data and results of this section are detailed in Section V below.

### 4) POSTERIOR VALIDATION BASED ON ERROR CONTROL TECHNIQUE

According to the experimental results of section V-B, the LSJ algorithm we proposed has the Precision of 100%, but the Recall is low. The reason for the low Recall is that the high similarity of some track structures on FPC results in false detection, which is detailed in section V-B. From the perspective of digital signal processing, the process of the false detection event can be described as follows: The track to be detected can be regarded as the signal source, and the morpheme sequence obtained by linguistic transformation from the track is equivalent to the code by coding the signal source. The highly similar signal sources results in the highly similar coding sequences, which leads to errors in the processing results of the signal processing system (that is, LSJ). The process mentioned above is a common problem in digital signal processing, and the corresponding solution is error control.

The core of error control technology is anti-interference coding or re-identification of confusing codes by introducing additional information. According to our observation and analysis of the track wrongly detected in section V-B), although the morpheme sequence is similar to the track of true fracture, the length of the basic track structure unit represented by each morpheme is different. Therefore, the length information is introduced to control the error. And the specific method is as follows:

$$\begin{cases} F(C_{ab}) \leq \lambda, & \text{True} \\ F(C_{ab}) > \lambda, & \text{False} \end{cases} \quad (14)$$

where  $F(C_{ab})$  represents the length characteristic of the hypothetical connecting track  $C_{ab}$ , and the specific definition is shown in (15).  $\lambda$  represents the threshold whose value is obtained through the experimental data. True means the fracture on  $C_{ab}$  is true, and False means fracture on  $C_{ab}$  is pseudo.

$$F(C_{ab}) = \sqrt{\sum_i^n \{\Phi[B_{LSU}(M_i)] - \bar{\Phi}[B_{LSU}(M_i)]\}^2} \quad (15)$$

where  $B_{LSU}(M_i)$  is defined in section III-C-1). The definitions of  $\Phi[B_{LSU}(M_i)]$  and  $\bar{\Phi}[B_{LSU}(M_i)]$  are shown in (16) and (17)

respectively. The smaller the  $F(C_{ab})$  value is, the more likely  $C_{ab}$  is the hypothetical connecting track of a true fracture.

$$\Phi [B_{LUS}(M_i)] = \Psi [B_{LUS}(M_i)] - \Psi' [B_{LUS}(M_i)] \quad (16)$$

$$\bar{\Phi} [B_{LUS}(M_i)] = \frac{1}{n - k_{16}} \sum_{i=1}^n \Phi [B_{LUS}(M_i)] \quad (17)$$

where  $\Psi [B_{LSU}(M_i)]$  represents the length of  $B_{LSU}(M_i)$ , and the calculation method is (18).  $\Psi' [B_{LSU}(M_i)]$  is the length of the basic track structure unit of the morpheme corresponding to  $M_i$  in the corpus  $\Omega$ . In (17),  $k_{16}$  represents the sum of the number of the morphemes whose values are 1 or 6 in  $L_{ANG}(C_{ab})$ .

$$\Psi [B_{LSU}(M_i)] = \begin{cases} \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, & M_i = 2, 3, 4, 5 \\ 0, & M_i = 1, 6 \end{cases} \quad (18)$$

where  $(x_1, y_1)$  and  $(x_2, y_2)$  are the coordinates of the two endpoints of  $B_{LSU}(M_i)$  respectively.

Specially, we explain the length of the basic track structure unit represented by the morpheme in  $\Omega$ , that is  $\Psi' [B_{LSU}(M_i)]$  in (16). When the corpus is generated according to the method in section III-C-1), the length information of each  $B_{LSU}(M_i)$  is obtained according to (18). We store it in the order corresponding to the morphemes in the corpus, and it is used as the length information base to be invoked at any time.

TABLE 2. Steps of FPC track fracture detection algorithm.

Input	All the broken ends $p_{break}$ , the track skeleton lines, and the weld/connecting ends $R_{pi}=R_p \cap R_b$ of one certain FPC image I.
Output	The broken end pairs $(p_u, p_v)$ of the true fractures in image I.
Step1:	<b>for</b> each two broken ends $p_a$ and $p_b$ , <b>do</b>
Step2:	Generate the hypothetical connecting track $C_{ab}$ by connecting $C_a$ and $C_b$ (incomplete track sections with breakpoint of $p_a$ and $p_b$ ) at $p_a$ and $p_b$ respectively.
Step3:	Express $C_{ab}$ with $L_{ANG}(C_{ab})$ according to rule 1- rule 4.
Step4:	Calculate $P[L_{ANG}(C_{ab})]$ of $L_{ANG}(C_{ab})$ by (12).
Step5:	Determine whether the region sandwiched by $p_a$ and $p_b$ is a true fracture according to (13), and output "true" or "false".
Step6:	<b>end</b>
Step7:	<b>for</b> each broken end pair $(p_u, p_v)$ which is determined "true" in Step 5, <b>do</b>
Step8:	Verify $(p_u, p_v)$ by (14).
Step9:	<b>end</b>

## IV. IMAGING EQUIPMENT AND IMAGE DATA SET

### A. IMAGING SYSTEM AND IMAGING METHODS

In this paper, the laptop keyboard FPC based on polyester film is taken as the experimental object. The specific version is GCNO52C1 and the size is 158mm×340mm, which is shown in Fig. 5.

We construct an image acquisition system in the position after the keyboard FPC printing station to acquire the FPC

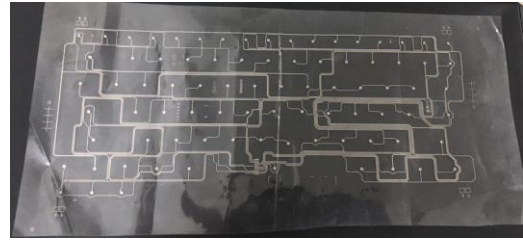


FIGURE 5. Picture of the laptop keyboard FPC.

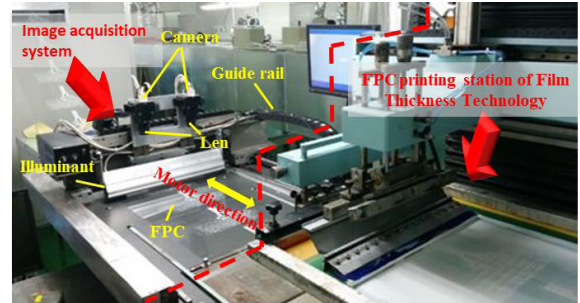


FIGURE 6. Picture of hardware device.

surface image. As shown in Fig. 6, the system is mainly composed of two linear cameras (DALSA p4-cm-08k070), two lenses (Schneider Apo-Componon 4.0/60), one high uniform strip illuminant (OPT-1st562-w), and one homemade mobile platform with guide rails. The camera and the illuminant are fixed on the mobile platform, and they are driven by the guide rail to move back and forth at a uniform speed along the direction indicated by the yellow arrow in Fig. 6. The FPC image is acquired as the camera moving forth, and the restoration of the camera position is performed as the camera moving back. The camera contains 8192 active pixels, and its working distance is set to 270mm. In addition, since the GCNO52C1 FPC is larger-size, we placed two cameras parallelized on the mobile platform to acquire the left and right sides image of the FPC to ensure the sufficient resolution. The acquired images are shown in Fig. 6 (a) and (b), and the resolutions are both 8192×4972.

### B. IMAGE DATA SET

There are 189 pieces of FPC with break defect collected in a year in the production workshop are imaged by the system in Fig. 6. Then we obtain an image data set with size of 189. It is worth noting that the sample of the established data set is the complete image (as shown in Fig. 7) obtained by the image stitching [48] of the left and right image of one FPC. The established data set contains 823 fractures, and the example of the image containing the break in the track is shown in Fig. 8.

## V. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, the proposed FPC fracture detection algorithm is evaluated. The image data set used for the experimental evaluation is established in section IV, which includes 189 FPC images and contains 823 fractures. In addition,



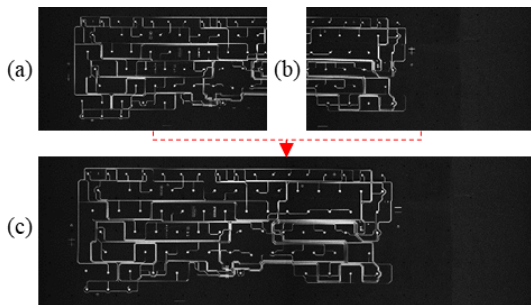


FIGURE 7. Image in the data set: (a) Image acquired by the left camera, (b) Image acquired by the right camera, and (c) Stitched image.

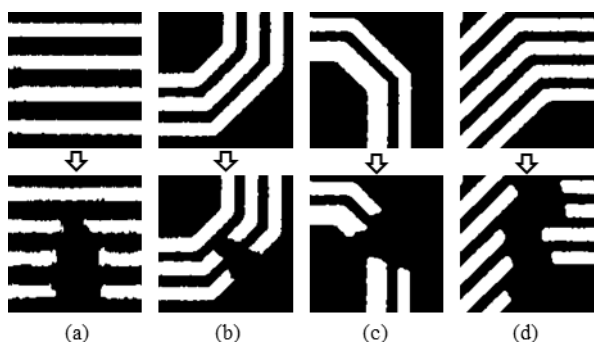


FIGURE 8. Track break: (a) Straight-line break, (b) Oblique track break, (c) Turning track break, and (d) Composite track break.

in order to carry out the experimental evaluation, we established the corpus for the GCNO52C1 keyboard FPC which is our experimental subject.

The Recall, Precision, and False negative rate (FNR) are three measurement indexes commonly used to evaluate the performance of defect detection. And the specific definitions are as follows.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (19)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (20)$$

and

$$\text{FNR} = \frac{FN}{TP + FN} \quad (21)$$

where TP represents the number of correctly recognized fractures, FP is the number of fractures which are recognized as pseudo fractures, and FN is the number of the pseudo fractures recognized as true.

In addition, through our observation and statistics about the samples in the image data set, the area value  $T_a$  and the radius  $r$  of the circular structure elements in section III-A are set to 10000 and 20 respectively. The following experiments are all executed on the basis of the above threshold value.

### A. HYPOTHESIS OF THE ALTERNATIVE FRACTURE

According to the theory in section III, the sandwiched region between all the track broken ends of FPC can be a true fracture

theoretically. Supposing that the number of track broken ends is  $\delta$ , then  $C_\delta^2$  alternative fractures and hypothetical connecting tracks can be obtained. Then, the true fracture will be recognized by judging the correctness of these hypothetical connecting tracks. But in fact, through the observation of the fracture in the image data set, the size of all the fractures is small. Therefore, for any detected broken end  $p_a$ , it only needs to make a hypothetical connecting with the broken ends in a certain range nearby for further verification. In other words, the broken end  $p_b$  which may form a true fracture with  $p_a$  satisfies the following conditions:

$$\sqrt{(x_{pa} - x_{pb})^2 + (y_{pa} - y_{pb})^2} < R \quad (22)$$

where  $(x_{pa}, y_{pa})$  and  $(x_{pb}, y_{pb})$  represent the coordinates of the points  $p_a$  and  $p_b$  respectively,  $R$  is the distance threshold, and the specific values are obtained according to experimental observation and statistics. Under the above limitation, the number of alternative fractures will be much less than  $C_\delta^2$ , which greatly reduces the calculation amount and consuming time of the algorithm. And it is very necessary in the practical application.

In order to determine the specific value of  $R$  in (22), the distribution of the length  $l$  of 823 fractures on 189 images in the data set is statistically analyzed, and the relevant data is shown in Fig. 9 with the blue line. To facilitate the display, the abscissa of Fig. 9 is set as  $\tau = l/W$  (the physical meaning is the multiple of track width), where  $W$  represents the average track width, and the value of  $W$  is 14 (pixels) in the dataset we established corresponding to the actual physical distance of 0.5mm. As shown in Fig. 9 (a), the length of most fractures is 2-3 times of the track width, accounting for 63.54% of the total number of fractures. The distribution trend is approximately a Gaussian distribution with a mean value of 2.5. The length value of the longest fracture is  $13W$ .

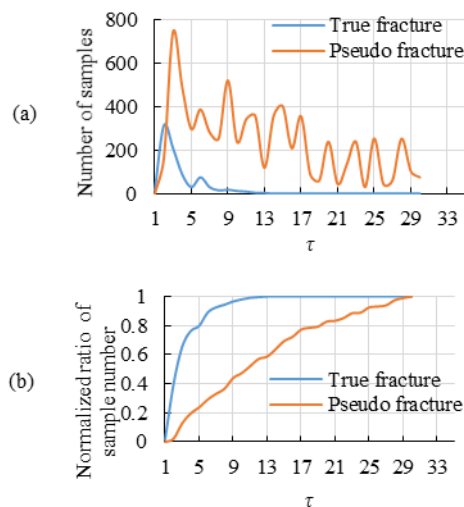


FIGURE 9. Relationship between the number of fractures and  $\tau$ : (a) Distribution of the fracture number with various length values, and (b) Cumulative distribution of the fracture number with various length values.

Fig. 9 (b) is the cumulative distribution curve of Fig. 9 (a). We also count the pseudo fracture number under different  $\tau$ , which is what the orange curve shows. As shown in Fig. 9 (a), the number of pseudo fracture reaches the most in  $\tau = 3$ , then decreases with the increase of  $\tau$ . As shown in Fig. 9 (b), the number of pseudo fractures is linear with the increase of the  $\tau$  approximately.

In conclusion, a greater  $\tau$  will result more pseudo fractures, and there is no true fracture when  $\tau > 13$ . Therefore, when  $R$  is set to  $13W$  (that is, 182 pix), the algorithm we proposed will deal with the least amount of data under the condition of meeting the testing requirements, and the number of pseudo fractures is 4087.

**B. LSJ ALGORITHM EFFECT VERIFICATION**

In this section, we evaluate the performance of the algorithms in section III-C, and the parameters  $R$  of (15) is set to  $13W$  (182pix).

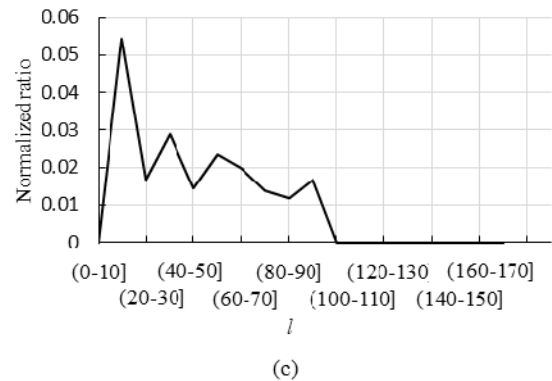
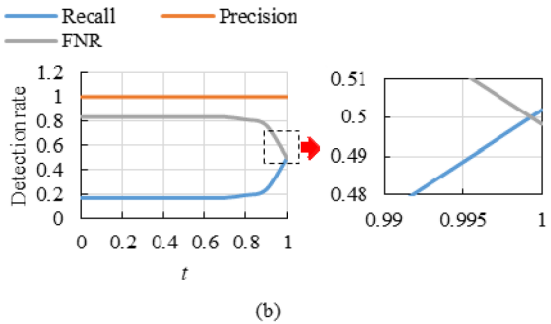
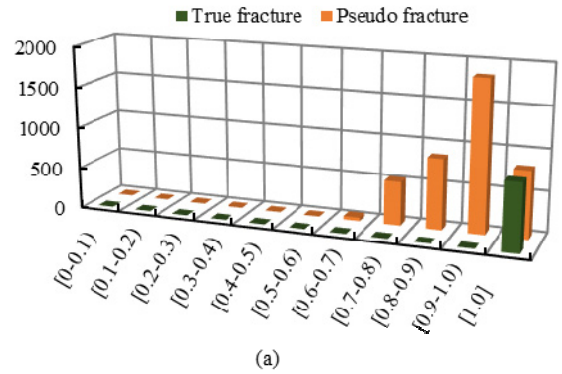
Fig. 10 (a) and (b) show the distribution of the semantic rationality probability  $P$  of the hypothetical connecting track of the 823 true fractures and 4087 pseudo fractures when  $R=182$ . As shown in Fig. 10 (a), for the hypothetical connecting tracks of 823 true fractures, the probability  $P=1$ . However, there are 817 hypothetical connecting tracks of pseudo fractures whose probability  $P$  is 1 also, which means that they cannot be distinguished with the true fractures.

Fig. 10 (b) shows the effect of LSJ algorithm under various  $t$ . The Precision is 100% at any value of  $t$ , which reflects the advantages and characteristics of the LSJ algorithm. When  $t=1$ , Recall reaches the maximum and FNR reaches the minimum, which are 0.5018 and 0.4982 respectively. Obviously, from these two indexes of Recall and FNR, the effect of LSJ is not ideal enough, but it has a perfect advantage in Precision, which means that there will be no missing detection for any true fracture. And this is also the main and core requirement in actual enterprise applications. According to the characteristic of indexes Precision, the cause of poor Recall and FNR is the false detection of 817 pseudo fractures. Therefore, on the basis of LSJ, an algorithm with considerable effect is expected to be obtained if some additional information can be provided for further verification of the detected 817 pseudo fractures. This is the purpose of section III-D exactly.

Deservedly, a further verification is a good approach for the actual requirements of fracture detection. However, in terms of the theory itself, it is necessary to analyze the algorithm characteristics of LSJ and the reasons for the above effects.

1) ANALYSIS OF LSJ ALGORITHM CHARACTERISTICS

Table 3 shows the proportion of the ones that satisfy various  $l$  and  $P$  among 4087 pseudo fractures, where  $l$  represents the length of the morpheme sequence of the hypothetical connecting track sentence. As shown in Table 3, 1) 4087 pseudo fractures all satisfy  $P \geq 0.6$ . 2) The  $P$  of the pseudo fracture with larger  $l$  is lower, whereas the  $P$  of the pseudo fracture with smaller  $l$  is higher. 3) For the hypothetical connecting



**FIGURE 10. Detection performance of the proposed LSJ algorithm: (a) Distribution of  $P$  of the hypothetical connecting track, (b) Relationship of the fracture detection effect and  $t$ , and (c) Distribution curve of the pseudo fracture number in Table 3,  $P=1$ .**

tracks of 817 pseudo fractures whose  $P$  value is 1, they are satisfying  $l \leq 0.5R$ , which indicates that the LSJ algorithm we proposed is more likely to recognize falsely when the length of the branches are short. As is shown in the distribution of the fractures number when  $P=1$  in Table 3 that it is basically in the trend of decreasing with the increase of  $l$  (we show it as a curve in Fig. 10 (c)). This is actually the systematic error of LSJ.

2) ANALYSIS OF THE POOR RECALL OF LSJ ALGORITHM

Since the Precision of LSJ algorithm is always 1, that is, all the true fractures are recognized successfully, so the unsatisfactory Recall and FNR are both caused by the false recognition of 817 pseudo fractures. Through our observation, the 817 pseudo fractures are all similar to certain true FPC

TABLE 3. Number of pseudo fracture under various  $l$  and  $P$ .

$l \setminus P$	[0-0.1)	[0.1-0.2)	[0.2-0.3)	[0.3-0.4)	[0.4-0.5)	[0.5-0.6)	[0.6-0.7)	[0.7-0.8)	[0.8-0.9)	[0.9-1.0)	1	Cumulative Sum
(0-10]	0	0	0	0	0	0	0	0	0	0	0	0
(10-20]	0	0	0	0	0	0	0	0	0	0	0.0543	0.0543
(20-30]	0	0	0	0	0	0	0	0	0	0	0.0166	0.0166
(30-40]	0	0	0	0	0	0	0	0	0	0.0282	0.0289	0.0571
(40-50]	0	0	0	0	0	0	0	0	0	0.0489	0.0144	0.0633
(50-60]	0	0	0	0	0	0	0	0	0	0.0506	0.0234	0.0740
(60-70]	0	0	0	0	0	0	0	0	0.0117	0.0272	0.0198	0.0587
(70-80]	0	0	0	0	0	0	0	0	0.0131	0.0352	0.0137	0.0620
(80-90]	0	0	0	0	0	0	0	0	0.0174	0.0471	0.0117	0.0762
(90-100]	0	0	0	0	0	0	0	0	0.0161	0.1106	0.0169	0.1436
(100-110]	0	0	0	0	0	0	0	0	0.0302	0.0387	0	0.0689
(110-120]	0	0	0	0	0	0	0	0	0.0276	0.0418	0	0.0694
(120-130]	0	0	0	0	0	0	0	0	0.0421	0.0218	0	0.0639
(130-140]	0	0	0	0	0	0	0	0	0.0201	0	0	0.0201
(140-150]	0	0	0	0	0	0	0	0.0021	0.0213	0	0	0.0234
(150-160]	0	0	0	0	0	0	0	0.0247	0.0103	0	0	0.0350
(160-170]	0	0	0	0	0	0	0.0017	0.0409	0	0	0	0.0426
(170-180]	0	0	0	0	0	0	0.0083	0.0626	0	0	0	0.0709
Cumulative Sum	0	0	0	0	0	0	0.0100	0.1303	0.2099	0.4501	0.1997	1

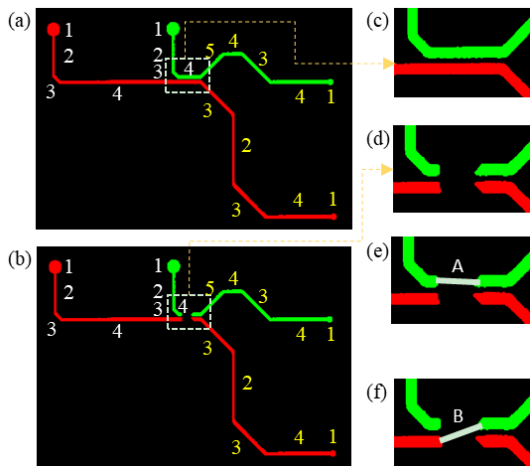


FIGURE 11. False detection caused by tracks with same topology: (a) Normal FPC track, (b) FPC track with fracture, (c) Local area of normal FPC track, (d) Local area of the fracture, (e) True fracture, and (f) Pseudo fracture.

track in topological structure, as shown in Fig. 11 (a). In the figure, the white numbers of morpheme sequences marked on the red track and the green track are both [1]–[4].

Therefore, when the two tracks are broken simultaneously in the white dotted line rectangular box (as shown in Fig. 11 (b) and (d)), the morpheme sequence of the hypothetical connecting modes of the fracture in Fig. (e) and (f) are the same. Compared with the corpus, the rationality of the two morpheme sequence are both 100%. However, the connecting mode in Fig. 11 (f) is incorrect obviously.

Although the pseudo fracture satisfies the LSJ algorithm condition due to the same morpheme sequence, it is obvious that the length of the basic track structural units represented by each morpheme is different. Taking Fig. 11 as an example still, the morpheme in the hypothetical connecting track

of true fracture A is [1]–[4], and the corresponding length is [0,180,32,110]. And the corresponding length of pseudo fracture B is [0,192,40,675]. It shows that these two length sequences are different, and the next verification can be performed based on this information to recognize the pseudo fracture, so as to realize the error control of LSJ algorithm.

C. EFFECT EVALUATION OF ERROR CONTROL ALGORITHM

In this section, we evaluate the performance of the algorithms in section III-D, whose results also represent the final effect of this research. According to the experimental results in section V-B, the algorithm in section III-C achieves the best effect when  $t=1$ . However, excepting the 823 true fractures recognized, there are still 817 pseudo fractures detected by mistake. In this section, the 1640 (817+823=1640) fractures selected in section V-B are used to evaluate the proposed error control algorithm.

Fig. 12 shows the distribution of eigenvalue  $F$  with 823 fractures and 817 pseudo fractures. As is shown that the distribution range of  $F$  values of true fractures and pseudo

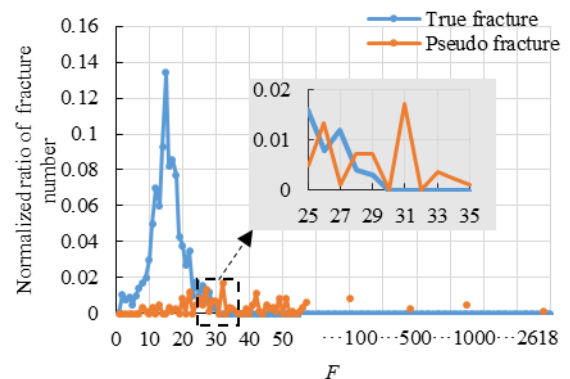


FIGURE 12. Normalized ratio of fracture number at various  $F$ .

fractures is 0-29 and 7-2618 respectively. Since the distribution state of the normalized ratio of pseudo fractures number in the range of  $F=[50-2618]$  is similar to that in the range of  $F=[0-50]$ , we only give the values at  $F$  being 100, 500, 1000 and 2618 schematically for the range of  $F > 50$  in Fig. 12 to ensure the readability of the graph.

Fig. 13 provides the Recall, Precision and FNR of the algorithm at various  $\lambda$  value. When  $\lambda = 24$ , Recall and Precision are equal nearly, and they are 0.9458 and 0.9550 respectively, achieving the best average performance. Meanwhile, the FNR is 0.0542. When  $\lambda$  is 29, the Recall=0.9185 and Precision=1, and the algorithm recognizes all the true fracture completely. And the FNR=0.0815.

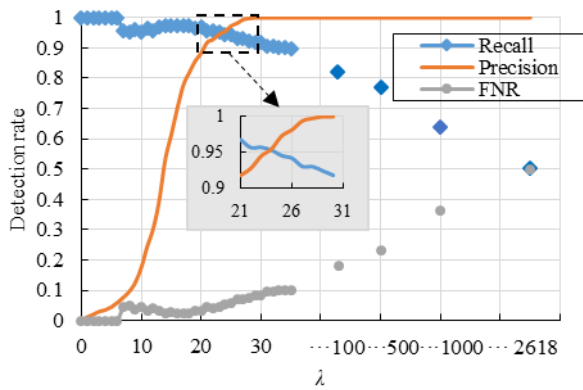


FIGURE 13. Relationship of Recall, Precision, and FNR with  $\lambda$ .

As shown in Fig. 13, when  $\lambda = 29$ , there are some pseudo fractures detected falsely. And when  $\lambda = 24$ , the algorithm has both missing detection and false detection. Therefore, we observe the wrongly detected samples and analyze the reasons below.

(1) Analysis of missing detection. According to the theory in section III-D, the immediate cause of missing detection is that the  $F$  of the hypothetical connecting track of the fracture is larger than the set threshold value  $\lambda$ . However,  $F$  is obtained according to the difference between the length of the hypothetical connecting track  $l_r$  and the corresponding reference length  $l_p$  in the length information base. In theory, for a true fracture, there is  $l_r \equiv l_p$ . Therefore, we further observe the image where the hypothetical connecting track is. It indicates that the reason for  $l_r \neq l_p$  is that the flexible deformation of the image results in the stretching and compression of the track. And this leads to the missing detection finally.

(2) Analysis of false detection. According to the theory in section III-D, the immediate cause of the false detection is that the  $F$  value of the hypothetical connecting track of the fracture is smaller than the set threshold value  $\lambda$ . This means that the length of the hypothetical connecting track of the false detected pseudo fracture is very close to the length of a certain sentence in the length information base. The illustration of this situation is shown in Fig. 14. In the figure, there are two groups of fracture areas, as marked by the yellow dotted line 1 and 2. For the fracture area 1, the true fracture is the

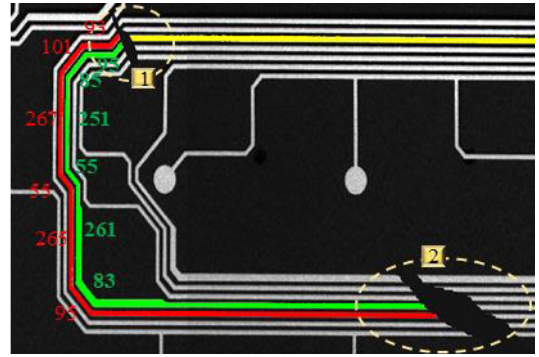


FIGURE 14. Legend of false detection.

one between the yellow and green tracks. But the existence of fracture area 2 results in the occurrence of both the same morpheme sequences and the similar length sequences of the red track and green track. The length sequence of the red track and green track are  $\psi_{red}=[95,85,251,55,261,83]$  and  $\psi_{green}=[93,101,267,55,265,95]$  respectively. Compared with  $\psi_{green}$ , the  $F$  of  $\psi_{red}$  is 22.80, which is less than  $\lambda$ , and it is incorrectly detected.

It should be noted that the 189 images used in our experiment are collected under the natural production state of the production line, which indicates the robustness of the proposed method for environmental variables such as FPC position on the production line, illumination, and deformation of FPC substrate. In addition, Precision could achieve 1 according to the experimental results, which indicates that all the true fractures are recognized. This proves the strong robustness of the proposed algorithm for the identification of the fractures.

#### D. COMPARISON OF THE DETECTION PERFORMANCE

To illustrate the classification performance of the algorithm we proposed, it is compared with the methods described in [9], [13], [15], [17], and [20]. The reason for choosing the methods in the above six references is that: 1) The result of the method or its certain step can obtain the exact location of the fracture. 2) The selected methods for comparison represent each category of current fracture detection methods as comprehensively as possible. 3) Publication date is as new as possible. In the compared methods, [9] is the representative of the direct comparison method (DCM), [13] is the representative of indirect comparison method (ICM), and [15] is the representative of feature comparison method (FCM). Reference [17] and [20] are two different rule-based methods, where [20] takes the difference between the track edge and the fitted line as the detection rule, called straight line fitting method (SLFM). And in [17], the edge shape of the local region of the track is used as the detection rule, called local edge configuration method (LECM).

Table 4 shows the comparison performance between our method proposed in this research and the other state-of-the-art methods mentioned above. As is shown that our method achieves better results in Recall, Precision and FNR com-

**TABLE 4. Effect comparison between the method we proposed and other methods.**

Methods	Recall	Precision	FNR	Time consumption(s)
Our method	0.9185	1	0.0815	1.6915
FCM[15]	0.7284	0.8408	0.2716	5.8539
ICM [13]	0.6073	0.8250	0.3927	7.6698
DCM [9]	0.5529	0.8190	0.4471	2.6469
SLFM[20]	0.7910	0.7497	0.2090	2.0602
LECM[17]	0.5976	0.5395	0.4024	0.5200

pared with the others. In terms of the time consumption, our method is second only to that of [17]. The specific time consumption of each step in the algorithm we proposed is given in Table 5.

**TABLE 5. Time consumption of each step of the algorithm in this research.**

Algorithm process	Time consumption(s)
Broken end detection and alternative fracture hypothesis	0.6195
Language generation	0.4985
Probability calculation of semantic rationality	0.0526
Error control	0.5209

## VI. CONCLUSION

This paper has presented a detection method of FPC track fracture. First, an imaging system has been constructed to capture the images of the surface of FPC, and then the image data set is established. A hypothesis testing strategy is used to detect the fracture in this research. In this strategy, the region sandwiched between any two track broken ends is hypothesized to be the candidate. Then we propose the LSJ algorithm to recognize the true fracture from the alternative fractures. In LSJ algorithm, we innovatively put forward the vocabulary definition method and grammar rules for FPC verbalization, and define the semantic rationality probability calculating model LRS. In addition, we propose an error control method with the length information of track to improve the detection accuracy. We conduct experiments to evaluate our proposed algorithm. The results have demonstrated that the proposed algorithm can provide a better detection effect. The Recall, Precision and FNR are 0.9185, 1 and 0.0815 respectively in the condition of without missing detection. In addition, the average time consumption of the algorithm in the testing data set is 1.6915s, which is relatively less compared with the state-of-the-art methods. In future work, we plan to apply this algorithm to other PCB data sets and the keyboard production workshop.

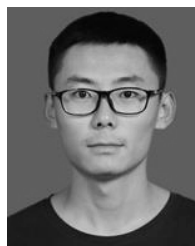
## ACKNOWLEDGMENT

The authors were grateful to Xingxiehe Technology Company Ltd., for their contributions to the establishment of the imaging system of the FPC. They also provided us with sufficient experimental samples.

## REFERENCES

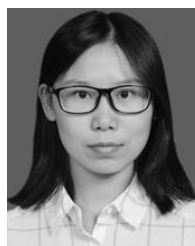
- [1] L. Wang, Y. Zhao, Y. Zhou, and J. Hao, "Calculation of flexible printed circuit boards (FPC) global and local defect detection based on computer vision," *Circuit World*, vol. 42, no. 2, pp. 49–54, May 2016.
- [2] F. Wu, J. Guo, X. Zhang, S. Li, and T. Wu, "3D quality inspection method for PCB solder joint surface," *Chin. J. Sci. Instrum.*, vol. 39, no. 5, pp. 233–240, 2018.
- [3] Y. P. Wang and F. L. Guo, "Key technology research of the machine vision-based PCB defect inspection system," *Bull. Sci. Technol.*, vol. 33, no. 1, pp. 101–105, 2017.
- [4] J. P. R. Nayak, K. Anitha, B. D. Parameshachari, R. Banu, and P. Rashmi, "PCB fault detection using image processing," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 225, Aug. 2017, Art. no. 012244.
- [5] W. B. Wang, D. Y. Liu, and Y. Q. Yao, "Defects detection of printed circuit board based on the machine vision method," *Appl. Mech. Mater.*, vols. 494–495, pp. 785–788, Feb. 2014.
- [6] W. B. Huang and P. Wei, "A PCB dataset for defects detection and classification," *Comput. Vis. Pattern Recognit.*, vol. 14, no. 8, pp. 1–9, 2018.
- [7] J. Zhu, A. Wu, and X. Liu, "Printed circuit board defect visual detection based on wavelet denoising," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 392, Aug. 2018, Art. no. 062055.
- [8] J. Huang, X. Chen, K. Luo, Y. Ji, S. He, L. Lin, and J. Li, "A modified SURF algorithm for PCB image registration," *J. Fujian Univ. Technol.*, vol. 16, no. 3, pp. 275–280, 2018.
- [9] G. Hua, W. Huang, and H. Liu, "Accurate image registration method for PCB defects detection," *J. Eng.*, vol. 2018, no. 16, pp. 1662–1667, Nov. 2018.
- [10] J.-X. Huang, D. Li, F. Ye, and Z.-J. Dong, "Flexible printed circuit defective detection based on image registration," in *Proc. 3rd Int. Congr. Image Signal Process.*, Oct. 2010, pp. 2570–2574.
- [11] L. D. Kong, J. Tan, and N. Chu, "Research of detection algorithm of flexible printed circuit boards based on global and local defects," *J. China Univ. Metall.*, vol. 23, no. 3, pp. 257–263, 2012.
- [12] C. L. Ma, "Automatic inspection system for FPC board based on multi-template reference contrast," *Comput. Meas. Control*, vol. 13, no. 5, pp. 421–422, 2005.
- [13] C.-T. Liao, W.-H. Lee, and S.-H. Lai, "A flexible PCB inspection system based on statistical learning," *J. Signal Process. Syst.*, vol. 67, no. 3, pp. 279–290, Jun. 2012.
- [14] J. Du, F. Chen, and Y.-M. Hu, "Automatic defect inspection of patterned FPC board based on 1-D fourier reconstruction," in *Proc. 36th Chin. Control Conf. (CCC)*, Jul. 2017, pp. 10109–10112.
- [15] V. H. Gaidhane, Y. V. Hote, and V. Singh, "An efficient similarity measure approach for PCB surface defect detection," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 277–289, Feb. 2018.
- [16] C. X. Xia, "Surface detection and analysis of PCB based on image processing," *Video Eng.*, vol. 42, no. 8, pp. 28–32, 2018.
- [17] K. Ren and Y. Hong, "The automatic detection system of flexible printed circuit board's defect in computer equipment based on picture processing and pattern recognition," *Circuit World*, vol. 40, no. 4, pp. 121–126, Oct. 2014.
- [18] Y. F. Li and S. Y. Li, "Defect detection for PCBs based on gradient direction information entropy," *China Mech. Eng.*, vol. 28, no. 6, pp. 695–701, 2017.
- [19] W. Q. Yuan, D. J. Li, and S. L. Li, "The circuit break detection of flexible deformation FPC based on skeleton," *Chin. J. Sci. Instrum.*, vol. 38, no. 4, pp. 996–1004, 2017.
- [20] N. S. Qiao and P. Sun, "Circuitry detection in printed circuit board," *Acta Photonica Sinica*, vol. 42, no. 11, pp. 1355–1359, 2013.
- [21] Y. Li and S. Li, "Defect detection of bare printed circuit boards based on gradient direction information entropy and uniform local binary patterns," *Circuit World*, vol. 43, no. 4, pp. 145–151, Nov. 2017.
- [22] Z. Lu, Q. He, X. Xiang, and H. Liu, "Defect detection of PCB based on Bayes feature fusion," *J. Eng.*, vol. 2018, no. 16, pp. 1741–1745, Nov. 2018.
- [23] L. Zhang, Y. Jin, X. Yang, X. Li, X. Duan, Y. Sun, and H. Liu, "Convolutional neural network-based multi-label classification of PCB defects," *J. Eng.*, vol. 2018, no. 16, pp. 1612–1616, Nov. 2018.
- [24] P. Wei, C. Liu, M. Liu, Y. Gao, and H. Liu, "CNN-based reference comparison method for classifying bare PCB defects," *J. Eng.*, vol. 2018, no. 16, pp. 1528–1533, Nov. 2018.

- [25] C. Zhang, W. Shi, X. Li, H. Zhang, and H. Liu, "Improved bare PCB defect detection approach based on deep feature learning," *J. Eng.*, vol. 2018, no. 16, pp. 1415–1420, Nov. 2018.
- [26] R. Heriansyah, "Neural network paradigm for classification of defects on PCB," *J. Technol.*, vol. 39, no. 1, pp. 87–104, 2003.
- [27] L. J. Guo, Y. X. Xie, and B. J. Luo, "Surface quality inspection of PCB based on subpix," *Electron. Quality*, vol. 2018, no. 6, pp. 8–13, 2018.
- [28] X. Dong, J. Shen, D. Wu, K. Guo, X. Jin, and F. Porikli, "Quadruplet network with one-shot learning for fast visual object tracking," *IEEE Trans. Image Process.*, vol. 28, no. 7, pp. 3516–3527, Jul. 2019.
- [29] W. Wang and J. Shen, "Deep visual attention prediction," *IEEE Trans. Image Process.*, vol. 27, no. 5, pp. 2368–2378, May 2018.
- [30] W. Wang, J. Shen, and H. Ling, "A deep network solution for attention and aesthetics aware photo cropping," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 7, pp. 1531–1544, Jul. 2019.
- [31] L. C. Yuan, "A comparison of several statistical word clustering methods," *J. Central South Univ. (Sci. Technol.)*, vol. 47, no. 9, pp. 3079–3084, 2016.
- [32] T. Wen, "Predicted feature error mapping and its application in multi-static underwater target recognition," *Acta Acustica*, vol. 44, no. 1, pp. 57–67, 2019.
- [33] G.-H. Zhang, C. Han, F. Lian, and L. Zeng, "Cardinality balanced multi-target multi-Bernoulli filter for pairwise Markov model," *Acta Automatica Sinica*, vol. 43, no. 12, pp. 2100–2108, 2017.
- [34] J. H. Li and P. Z. Zhang, "Research and application on similarity word searching method," *Microcomput. Inf.*, vol. 28, no. 9, pp. 417–418, 2012.
- [35] Q. D. Wang, "A matching path constrained longest common subsequence length algorithm," *J. Electron. Inf. Technol.*, vol. 39, no. 11, pp. 2615–2619, 2017.
- [36] C. Chen, "Statistical algorithm for DNA repeats frequency based on finite state automaton," *Comput. Eng.*, vol. 37, no. 11, pp. 184–186, 2011.
- [37] D. G. Wu and X. A. Zhou, "Similarity analysis of DNA sequences based on longest common subsequence," *Intell. Comput. Appl.*, vol. 8, no. 6, pp. 22–26, 2018.
- [38] T. Y. Goh, S. N. Basah, H. Yazid, M. J. A. Safar, and F. S. A. Saad, "Performance analysis of image thresholding: Otsu technique," *Measurement*, vol. 114, pp. 298–307, Jan. 2018.
- [39] E. Taghizadeh, A. Terrier, F. Becce, A. Farron, and P. Büchler, "Automated CT bone segmentation using statistical shape modelling and local template matching," *Comput. Methods Biomech. Biomed. Eng.*, vol. 22, no. 16, pp. 1303–1310, Dec. 2019.
- [40] G. L. Wang, "Extracting method of skeleton for characteristics curve of shoe-last based on mathematical morphology," *J. Qingdao Univ. (Natural Sci. Ed.)*, vol. 25, 2, pp. 43–46, 2012.
- [41] X.-W. Zhang, J. Wang, G.-D. Lu, S.-M. Fei, and D.-L. Zhang, "Extraction and reuse of pattern configuration based on ontology and shape grammar," *J. Zhejiang Univ. (Eng. Sci.)*, vol. 52, no. 3, pp. 461–472, 2018.
- [42] W. Y. Wang, "Extraction and reuse of pattern configuration for handicrafts personalized customization," *J. Graph.*, vol. 40, no. 3, pp. 583–590, 2019.
- [43] Y. W. Cao, "Rapid evaluation method of shape characteristics of aggregate particle based on the minimum outer rectangle," *J. Chongqing Jiaotong Univ. (Natural Sci.)*, vol. 38, no. 6, pp. 61–65, 2019.
- [44] Z. M. Chen, M. X. Li, and M. W. Wang, "Sentence-level machine translation quality estimation based on neural network features," *J. Comput. Res. Develop.*, vol. 54, no. 8, pp. 1804–1812, 2017.
- [45] J. Xiong, Y. Fan, D. Teng, C. Ma, G. Dai, and H. Wang, "An online synchronization input method based on paper forms and electronic ones," *J. Comput.-Aided Des. Comput. Graph.*, vol. 24, no. 9, pp. 1125–1133, 2012.
- [46] X. F. Wang and L. M. Du, "Automatic segmentation of Chinese using overlaying ambiguity examining method and statistics language model," *J. Electron. Inf. Technol.*, vol. 25, no. 9, pp. 1168–1173, 2003.
- [47] D. H. Lv, "Study on the genus *Aspidistra* based on DNA barcoding," *Plant Sci. J.*, vol. 36, no. 6, pp. 784–789, 2018.
- [48] L. Liu, Y. Wu, M. Luo, and B. Ma, "An image mosaic algorithm based on coordinate transformation," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 382, pp. 338–343, Jul. 2018.



**DEJIAN LI** received the B.S. and Ph.D. degrees from the School of Information Science and Engineering, Shenyang University of Technology, Shenyang, China, in 2013 and 2018, respectively.

He is currently a Researcher with the Electrical Engineering Postdoctoral Research Station, Shenyang University of Technology, where he is also a member of the Computer Vision Group. In recent years, he published over ten technical articles. His current research interests include machine vision, image processing, and industrial product quality inspection.



**SHAOLI LI** received the B.S. and Ph.D. degrees from the School of Information Science and Engineering, Shenyang University of Technology, Shenyang, China, in 2013 and 2018, respectively.

She is currently a Researcher with the Electrical Engineering Postdoctoral Research Station, Shenyang University of Technology. Her current research interests include machine vision and biometric identification.



**WEIQI YUAN** received the B.S. degree from Hunan University, Changsha, China, in 1982, and the M.S. and Ph.D. degrees from Northeastern University, Shenyang, China, in 1988 and 1997, respectively.

In 2000, he was selected to the second batch of Hundreds of Thousands of Talents Project in Liaoning, China. In 2001, he was awarded the title of Outstanding Scientific and Technological Workers of Shenyang by the Shenyang Municipal Party Committee and Shenyang Municipal Government. In 2003, he was selected as the Senior Software Talents of Shenyang. He is currently a Professor with the Shenyang University of Technology. He is the Executive Director of the Chinese Instrument Association and the Director of the Biomedical Electronics Branch of the Chinese Electronics Association. In recent years, he published over 100 technical articles and has authored one book. He holds six patents. His current research interests include machine vision, biometric identification, and nondestructive testing.

• • •