Morphologic recovery method of product label noises based on image features and application in defect detection

Yuanhao Chen¹, Jianjun Yi¹*, Yajun Zhang¹ and Sheng Tao²

¹ School of Mechanical Engineering, East China University of Science and Technology, Shanghai, 200237, China
² Shanghai Composite Material Technology Co., Ltd, Shanghai, 201100, China

* Corresponding author’s e-mail: jjyi@ecust.edu.cn

Abstract: To ensure the reliability of carbon fiber material products and eliminate all kinds of possible defects, it is necessary to take effective means to check their quality. X-ray based nondestructive testing technology combined with image recognition algorithm is considered to be a fast and efficient solution. However, the surface of the processed material is often accompanied by labels which contains various information. These labels are confused in the image and make most of the defect detection methods easy to misidentify. This paper mainly studies the recovery method of product label noises based on image features and application in defect detection. What’s more, the proposed method can effectively and efficiently eliminate noises without affecting the rest of the image information which ensures the correct identification of defects in materials.

1. Introduction

Carbon fiber composite material is a cutting-edge new composite material. Because of its excellent performance, it is widely used in the production and processing of some important equipment, such as nuclear industry equipment, aerospace, ships, etc.[1]. The quality of composite material workpieces is affected by factors such as casting process and production conditions. During the production process, there may be manufacturing defects on the surface or inside of the workpiece [2], mainly including: inclusions, cracks and looseness, etc., which greatly affect the service life and performance of the equipment.

Most of the castings use non-destructive testing before they are officially put into use, because they do not need to damage the structure of the workpiece to be tested. At present, the defect identification of carbon fiber composite materials in the industry is mainly done manually. The operator observes the radiographic images obtained by stretching the local contrast to classify and label the defects [3]. However, this method is limited by the operator's experience and subjective factors, and the obtained images often have problems such as low contrast, other noise interference, fuzzy edge information of defects, and different shapes and sizes [4], resulting in low recognition efficiency. In recent years, intelligent image recognition algorithms have provided new directions for defect detection. Experts and scholars have conducted relevant research on the above-mentioned difficult problems in image analysis.

Aiming at the problem of high similarity between foreground and background images, Lan Yeshen et al. built a low-rank matrix decomposition model to decompose bearing surface defect detection into low-rank background images and sparse foreground images, reducing the effect of background factors in the defect detection process [5]; Feng Ming et al. improved an adaptive image enhancement algorithm...
based on unsharp mask technology, which can quickly and efficiently complete the internal defect
detection of cylindrical aluminum alloy castings [6]; Zhang Xiaoguang Et al. chose piecewise sine
function as the degree of fuzzy membership, and proposed a generalized fuzzy enhancement algorithm
suitable for radiographic inspection welding images [7]; Sun Jiting et al. performed a threshold
segmentation process on bamboo-plastic composites, and selected part of the original image for image
fusion processing, in which the weighting average fusion method, pyramid fusion method and wavelet
transform fusion method are used [8].

The above-mentioned processing algorithms and defect detection methods mentioned in most
documents can realize defect detection and segmentation tasks for the X-ray images obtained by
shooting. But they tend to focus on the defect recognition task itself, and the obtained images are
relatively simple and clear. However, tag noise is unavoidable during the shooting and scanning of
industrial scenes. Products are often labeled with relevant information. These tags will be confused in
the image, and it is easy to cause misrecognition through traditional recognition algorithms. In response
to this problem, this paper proposes a method of product surface label topography restoration based on
image features, which can greatly improve the accuracy of image-based carbon fiber composite defect
detection under label noise.

2. Extraction of suspicious label noise area based on gray-scale scanning analysis

The type and position of these labels attached to the surface of the workpiece are not fixed, and they are
imaged together with the product features during the X-ray scanning process, which is difficult to
distinguish. The existence of label noise often leads to the two non-ideal states of missed detection or
misidentification of defects.

Through analysis, when the label noise signal appears, it will often cause large gray value fluctuations
in a small range. When the label noise disappears, the gray curve in the parallel direction will continue
the original curvature; when the defect noise appears, the gray value will have small fluctuations in a
relatively large range. Therefore, we can define the label noise as a strong signal, and the defect feature
as a weak signal. Aiming at the characteristics of the strong signal of label noise, the location of label
noise can be quickly locked through the parallel direction adaptive gray-scale scanning analysis method.

Step 1: Perform horizontal line scan line by line on the X-ray image of the product to obtain h groups
of gray-scale transformation curves, where h is the height of the X-ray image. Figure 1 shows a
schematic diagram of the line scan result of a certain row. In most cases, the gray value changes
smoothly between 0 and 65535.

Step 2: Search and locate in the gray-scale transformation curve of each line according to the strong
signal characteristics of the label noise. First, because the color of the x-ray imaging part of the label
signal is darker, the search for each local minimum gray value point in the gray transformation curve of
the i-th row is marked as \( G(x_i) \), where \( j = 1, 2, \ldots, n \), \( x_i \) represents the The pixel coordinates of
the point.
Step 3: The strong signal characteristics of the tag signal are often accompanied by the characteristics of small single-peak fluctuation area and high peak value. Therefore, we start from the selected minimum point and search for the nearest minimum point and maximum point of the gradient transformation to the left and right sides respectively, denoted as $H_{\text{min}}(x_{ij})$ and $H_{\text{max}}(x_{ij})$. When calculating the gradient value, considering that the gray level will also have slight fluctuations under normal circumstances, in order to ensure the smoothness of the overall gradient, we calculate the gradient transformation through the 10 adjacent pixels on the left and right sides of $x_{ij}$.

$$H(x_{ij}) = \frac{\sum_{k=0}^{1} x_{i(j-k)} - \sum_{k=0}^{1} x_{i(j+k)}}{10}$$  \hspace{1cm} (1)

The minimum and maximum points of the gradient transformation on the left and right sides of the local minimum gray value point are the positions where the signal may undergo abrupt changes. At the same time, the distance between the mutation position and the gray minimum point is recorded as: $\Delta d_{l}$ and $\Delta d_{r}$. Combine the information from each local minimum gray point to the minimum point and maximum point of the left and right gradient changes in the i-th row, and the set is obtained as follows:

$$\{(G_{\text{min}}(x_{ij}), H_{\text{min}}(x_{ij}), H_{\text{max}}(x_{ij}), \Delta d_{l}, \Delta d_{r}, ... )\}$$  \hspace{1cm} (2)

Step 4: This set is considered to be a set of points where label noise may occur, and the suspicious label area is locked by filtering out the location of strong signal features in the set. The selection strategy is as follows:

Gray condition:

$$G_{\text{min}}(x_{ij}) < \frac{G(x_{i(j-\Delta)}) + G(x_{i(j+\Delta)})}{2} - \alpha$$  \hspace{1cm} (3)

Gradient conditions:

$$\left| H_{\text{min}}(x_{ij}) + H_{\text{max}}(x_{ij}) \right| < \beta$$  \hspace{1cm} (4)

Distance condition:

$$\left| \Delta d_{l} - \Delta d_{r} \right| \leq \gamma, \Delta d_{l} \leq \epsilon, \Delta d_{r} \leq \epsilon$$  \hspace{1cm} (5)

When the points $G_{\text{min}}(x_{ij})$ in the set meet the above three conditions, they are considered to be label noise, and the area $(j-\Delta d_{l}, j+\Delta d_{r})$ is considered to be a label noise area. Among them $\alpha, \beta, \lambda, \gamma, \epsilon$ is the corresponding threshold, which is set to 1000, 650, 425, 200, and 50 in our experiment. So far, we have filtered out the area of label noise.

3. Adaptive Shape Recovery Strategy Based on Shape Representation

After analyzing by grayscale scanning, locate the position where there is single or multiple continuous peaks of label noise. In this chapter, we will discuss how to restore the topography of the area affected by these label noises.

The main discussion here is the symmetrical parts that are common in the industry. The horizontal gray curve of symmetrical parts also has corresponding symmetry characteristics. According to this feature, the gray distribution of the area not affected by the label noise is mapped to the selected label area, so as to realize the shape restoration of the specific area according to the shape characterization. We propose an adaptive center-finding strategy to find symmetric centers that may deviate in the image. And according to the matching degree of the curve on both sides of the center point, adaptive compensation is performed to eliminate the label noise. Specific steps are as follows.

Step 1: Since the center position of the product does not necessarily coincide with the imaging position, the symmetry axis is quickly searched based on the gray level curve. For the gray level curve extracted from the i-th row, the peripheral area of the center of the curve is taken as the hypothetical domain of the symmetry axis. Traverse all the values in the area, and calculate the similarity of the
curves on the left and right sides of the central axis, and take the point with the highest similarity as the symmetry point. Assuming \( j \) is a value in the hypothesis domain, the similarity calculation formula is as follows:

\[
S_i = \sum_{\nu=1}^{n} \{ c = 1 \text{ if } |G(x_{i,j}) - G(x_{i,j})| < \nu \text{ else } 0 \}
\]

In our experiment, it is set to 200.

**Step 2:** Through the above steps, the suspicious area of label noise and the center point position of each row in the horizontal direction and its corresponding similarity have been obtained. If there is a suspicious noise area in the \( i \)-th row, when the similarity \( S_i \geq 440 \), the value of the label noise area is mapped to the other half of the center point to restore the shape, and each connected area is expanded to the left and right by 1/20 of the image width. To ensure completeness; When \( 400 \leq S_i \leq 440 \) and the curves on both sides are slightly deviated, the gray value of the symmetrical row and the gray value of the adjacent two rows are weighted and combined as the new gray value of the suspicious area, the weight is 0.8 and 0.2, and finally each connected area is expanded by 1/30 to the left and right to ensure completeness; When \( 380 \leq S_i \leq 400 \), the deviation on both sides becomes larger, and the weighted combination is also performed, but the weights of two adjacent rows are increased. The weights are set to 0.6 and 0.4, and the expansion is 1/40 to ensure integrity. In the experiment, it was found that the parts with lower similarity were concentrated on the parts with changeable shapes and more complex contours at the top and bottom ends of the product. The data set includes a total of 109 x-ray images, all of which contain label defects, and the similarity \( S_i \) at both ends of the symmetry center point is higher than 380.

**Step 3:** After the initial recovery of the defective area, all the lines are recombined into an image. However, even if the similarity between the two ends of the symmetry is very high, there will be a certain difference, which will cause a sudden change in the gray value of the adjacent rows in the restored image, which does not conform to the normal law. Therefore, the suspicious area of the label is filtered by median to eliminate such effects.

4. **Intelligent defect recognition and segmentation algorithm**

At this time, the label noise in the image has been effectively eliminated, and then the traditional digital image processing technology is used to achieve the segmentation and extraction of the industrial radiographic defect of the carbon fiber composite material.

**Step 1:** First, the contrast enhancement of the original image is performed. Define the original pixel value of a certain point in the picture as Orig, the average value of the pixel points in the 3x3 cells around the point is defined as Mean, the contrast factor (enhancement degree) is defined as Factor, and the new pixel value size is defined as \((\text{Orig} - \text{Mean}) \times \text{Factor} + \text{Orig}\). After this process, the texture and defects in the image become clear.

**Step 2:** Then use the Sobel operator [10] to obtain the first derivative of the enhanced image as an edge detector to detect edges. The size of the convolution kernel is 3x3, and the filter masks:

\[
A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & -1 \end{bmatrix}
\]

Use the masks to perform the convolution operation on the enhanced picture and then extract the edge information according to the calculation method \( (|A| + |B|) / 4 \).

**Step 3:** Perform a binarization and inverse operation on the picture to make the edge information different from the background and facilitate the subsequent threshold segmentation, and then perform the mean filtering on the picture to remove some unnecessary noise in the picture. Finally, threshold segmentation and connected domains are performed according to defective pixel values, and unconnected areas are defined as separate instances.

**Step 4:** Using the geometric characteristics of the region instance (such as area, perimeter, roundness,
etc.), the region instances obtained in the above steps are further screened and classified, and finally the defect part is segmented and extracted.

Figure 2. X-ray image of tagged tailpipe parts

5. Experiment and analysis
In this section, the tail pipe parts made of carbon fiber composite materials are selected as examples for detailed experimental verification. Figure 2 shows the X-ray images of different tail tube products. The red circled part is the location of the label. It can be seen that the location of the label is disordered and irregular.

Figure 3. Grayscale and gradient information distribution
First, we locate the suspicious area based on the strong signal characteristics of the tag. The resolution of the collected x-ray image is 1024x1024 pixels. Line scanning is performed line by line to obtain 1024 sets of horizontal gray distribution data with 1024 data, which is drawn as a curve as shown in the green curve in Figure 3. It is found that the gray scale curve in the area without label noise or far from the edge of the contour is very smooth. We have also studied the defects such as cracks and inclusions. The gray level fluctuation of cracks and inclusions is small, so it is still relatively smooth in a large range. The blue curve in the figure is the gradient transformation curve of the gray value. According to the screening strategy, we can accurately identify the area corresponding to the label noise, as shown in the red labeled area.

Determine the symmetry center point of each line according to the content in the second subsection, and record its similarity, as shown in Figure 4. In order to show the similarity more vividly, the symmetrical point is taken as the center, and the gray value of the defect-free side is mapped to the other side. Among them, (a) is the part with more top contours of the tailpipe product. It can be seen that its similarity is at a low value. In (b), (c), (d), despite the presence of label noise, the similarity is still at a relatively high level.

According to the similarity value obtained in each row, different strategies are performed on the defect part in each row to restore the shape. Finally, the image is reorganized for median filtering, and the topographic restoration image of the label area obtained is shown in Fig. 5, where the red frame is the original label area. Comparing with the original picture, we found that the label noise was effectively eliminated, and the remaining important information in the picture, such as edge information and defect information, was not affected in any way.

![Figure 4. Comparison of different areas of workpieces and mapping compensation strategies](image1)

![Figure 5. Restoration image of label noise profile](image2)
A large number of experiments have shown that our test results are considered reliable. Among the 109 tail tube X-ray images, there are 87 with defects and 22 without obvious defects. The algorithm correctly identified 102 images. Figure 6 shows part of the experimental results. Figures (a) and (b) have crack defects, Figure (c) has no defects, and Figure (d) shows the recognition effect of Figure (a) without label topography recovery.

Figure 6. Defect detection results and comparison

6. Concluding remarks
This paper proposes a label shape restoration method based on product shape characterization when the image has low contrast and label noise interference problems. The method of extracting the suspicious label noise region based on gray-scale scanning analysis and the adaptive recovery strategy based on shape symmetry are proposed. The label noise signal in the image is effectively eliminated, which greatly avoids the possibility of directly using the original image for defect recognition.

Acknowledgement
This paper was supported by the Major Program of National Natural Science Foundation of China under Grant No. 61690214, Shanghai Science and Technology Action Plan under Grant No. 18DZ1204000, 18510745500, 18510750100, 18510730600, Shanghai Aerospace Science and Technology Innovation Fund (SAST) under Grant No. 2019-080, 2019-116, Shanghai Sailing Program under Grant No. 20YF1417300, the Natural Science Fund of China (NSFC) under Grant No. 51575186, Shanghai Industrial Collaborative Innovation Special Fund under Grant No. 2021-cyxt1-kj 13, 2021-3, 2023-3, Shanghai International Science and Technology Cooperation Fund under Grant No. 19510730600, 2019-1201, 2022-1130, and Pudong New Area Science and Technology Development Industry-University Project Special Fund (Artificial Intelligence) under Grant No. PKX2020-R10, 2020-8, 2022-7.

References
[1] Tian Yong. (2015) Research on Drilling Defects of Carbon Fiber Composite Material and Tool Optimal Design[D]. Harbin University of Science and Technology.
[2] Zhang Huiling. (2013) Research on non-destructive testing and identification of main defects of carbon fiber composite materials [D]. National University of Defense Technology.
[3] Xiao Hui. (2017) Research on Automatic Detection and Recognition System of Welding Defect Radiographic DR Image[D]. Nanchang Hangkong University.
[4] Yu Yongwei, Yin Guofu, Yin Ying, Du Liuqing. (2013) Weak signal recognition method of X-ray image of weld defects[J]. Computer Integrated Manufacturing System, 19(10): 2557-2561.
[5] Lan Yeshen, Zhou Mingan, Liu Wenjun. (2020) Research on Bearing Surface Defect Detection Based on Low Rank Matrix Decomposition[J], Southern Agricultural Machinery, 51(21):17-19.
[6] Feng Ming. (2020) Research and implementation of X-ray digital imaging system for cylindrical aluminum alloy castings[D]. North University of China.
[7] Zhang Xiaoguang, Gao Ding, Zhang Xinggan. (2005) Radiation inspection welding image enhancement based on generalized fuzzy operator[J]. Journal of Harbin Institute of Technology, 62-65.
[8] Sun Jiting. (2019) Research on X-ray non-destructive testing of bamboo-plastic composite materials based on image fusion technology[D]. Northeast Forestry University.

[9] Zhao Haifa, Zhu Rong, Du Changqing. (2020) Image enhancement algorithm combining global tone mapping and local contrast processing[J/OL]. Journal of Wuhan University, 597-604.

[10] Abdel Karim M. Baareh, Ahmad Al-Jarrah, Ahmad M. Smadi, et al. (2018) Performance Evaluation of Edge Detection Using Sobel, Homogeneity and Prewitt Algorithms. 11(11):537-551.