Research Article

A Method of Surface Defect Detection of Irregular Industrial Products Based on Machine Vision

Mengkun Li,1 Junying Jia,2 Xin Lu,2 and Yue Zhang1

1School of Management, Capital Normal University, Beijing/100089, China
2Shenyang Fengchi Software Co. LTD, Shenyang/110167, China

Correspondence should be addressed to Mengkun Li; limengkun@cnu.edu.cn and Xin Lu; luxin@fchsoft.com

Received 23 November 2020; Revised 8 December 2020; Accepted 9 April 2021; Published 10 May 2021

Academic Editor: Amr Tolba

Copyright © 2021 Mengkun Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, the surface defect detection technology of irregular industrial products based on machine vision has been widely used in various industrial scenarios. This paper takes Bluetooth headsets as an example, proposes a Bluetooth headset surface defect detection algorithm based on machine vision to quickly and accurately detect defects on the headset surface. After analyzing the surface characteristics and defect types of Bluetooth headsets, we proposed a surface scratch detection algorithm and a surface glue-overflowed detection algorithm. The result of the experiment shows that the detection algorithm can detect the surface defect of Bluetooth headsets fast as well as effectively, and the accuracy of defect recognition reaches 98%. The experiment verifies the correctness of the theory analysis and detection algorithm; therefore, the detection algorithm can be used in the recognition and detection of surface defect of Bluetooth headsets.

1. Introduction

In the mass production process of Bluetooth headsets, it is necessary to glue the connection of the Bluetooth headset. During the bonding process, the problem of glue overflow may occur. In addition, the surface of the headset is also prone to scratches during the processing process. Therefore, after the production of the Bluetooth headset, it is necessary to check whether the appearance of the headset is scratched and the glue-overflowed. The traditional manual detection efficiency is low, labor and time costs are high, the work is boring and tedious, and it is difficult for the human eye to recognize the relatively small scratches and glue-overflowed, and it is easy to make mistakes when working for a long time.

In view of this, this article proposes a Bluetooth headset surface defect detection algorithm based on machine vision.


2. Design of the Detection Algorithm

The flow chart of the Bluetooth headset surface defect detection algorithm is shown in Figure 1. Firstly, we use an industrial camera to collect the headset image and then preprocess the input image. The preprocessing includes threshold segmentation and image enhancement. The purpose of threshold segmentation is to find the area that needs to be detected, and the image enhancement is to make the defects on the headsets surface more obvious. Next, perform surface scratch detection and surface glue-overflew detection on the preprocessed image to confirm whether there are scratches and glue-overflew on the headset surface, and finally, output the detection results. If no defects are found in the two detections, it is determined that there is no defect on the surface of the headsets.

2.1. Preprocess Image

2.1.1. Threshold Segmentation. After we use the industrial camera to collect the headset picture, the image needs to be preprocessed to find out the area to be detected, that is, the area that needs to detect scratches and glue overflow. The method of threshold segmentation is used to extract the region to be detected.

The basic principle of threshold segmentation is:

\[
G(i, j) = \begin{cases} 
0, & f(i, j) < T \\
255, & f(i, j) > T
\end{cases}
\]

\( (1) \)

\( G(i, j) \) is the generated image after threshold segmentation, \( f(i, j) \) is the input image, and \( T \) is the threshold used for segmentation value. The selection of threshold is very important for the detection of the image. Currently, the commonly used threshold selection methods mainly include fixed threshold method and automatic threshold method. The threshold in the fixed threshold segmentation is generally set manually, and the method is simple, but the threshold set the manual threshold set needs to be based on a large amount of experimental data, which is low in efficiency and cannot adapt to environmental changes. Automatic threshold segmentation automatically selects the segmentation threshold through image data statistics. There are mainly the maximum between-class variance segmentation method, the maximum entropy threshold segmentation method, and so on.

(1) Maximum Between-Class Variance Segmentation Method. The maximum between-class variance method was first proposed by a Japanese scholar Otsu Zhanzhi in 1979, so it is also called Otsu method. The basic idea is to divide the image grayscale histogram into two groups of background area and target area at a certain threshold according to the grayscale characteristics of the image. The variance between the two groups increases with the gray value difference between the background area and the target area. When the variance between the two groups is the largest, the threshold is determined and segmented [22].

(2) Maximum Entropy Threshold Segmentation Method. The maximum entropy threshold segmentation method is mainly based on the information entropy in the image, designing a reasonable entropy criterion and optimizing it, so that the threshold when the image entropy is maximum can accurately segment the target area and the background area in the image.

Use Eq. (1) to segment the image, and the result is shown in Figure 2. (a) is the input image; (b) is the threshold segmentation result.

2.1.2. Image Enhancement. There is a difference in the gray value between the defective area on the headset surface and the nondefective area, but the difference is not obvious enough. Increasing this difference through image enhancement algorithms can help to improve the accuracy of subsequent specific defect detection. The specific method is as follows: Firstly, we transform the input image from spatial domain image to frequency domain image by Fourier transform. The high frequency components of frequency domain image are usually defect edge and noise. The low-pass filter is used to filter the frequency domain image to remove the high-frequency components in the frequency domain, and then, the filtered frequency domain image is transformed into the spatial domain image through the inverse Fourier transform. Compared with the original image, the area with larger difference may be the glue-overflew area, and the surface defect of the difference image is more obvious than the original image, and the image enhancement is realized.

The Fourier transform formula of picture \( f(x, y) \) with image size of \( M \times N \).

\[
F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)e^{-j2\pi((ux/M)+(vy/N))}.
\]

\( u = 0, 1, 2, \cdots, M - 1, v = 0, 1, 2, \cdots, N - 1, u \) and \( v \) are the frequency domain variables and \( x \) and \( y \) are the spatial domain variables.
Corresponding inverse Fourier transform formula.

\[
f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi((ux/M) + (vy/N))}.
\] (3)

We get the frequency domain image after Fourier transform, then filter the frequency domain image. Commonly used methods of filtering include mean filtering, Gaussian filtering, median filtering, and so on.

Mean filtering is a typical linear filtering algorithm. It refers to giving a template to the target pixel on the image. The template includes neighboring pixels around it (8 pixels around the target pixel as the center to form a filter template). The average value of all pixels in the template replaces the original pixel value. Mean filtering uses a linear method to average the pixel values in the entire window range. Mean filtering has inherent defects. It cannot protect image details well. It also destroys the details of the image while denoising. The image becomes blurred, and the noise points cannot be removed well. Mean filtering is better for Gaussian noise.

Gaussian filtering is to scan each pixel in the image with a template, and use the weighted average gray value of the pixels in the neighbor-hood determined by the template to replace the value of the center pixel of the template. Gaussian filtering is the process of weighted averaging the input image. The value of each pixel is obtained by weighted averaging of itself and other pixel values in the neighbor-hood.

Median filtering is to take the points of adjacent pixels, and then sort the points of adjacent pixels, and take the gray value of the midpoint as the gray value of the pixel. The median filter uses a nonlinear method, which is very effective in smoothing impulse noise, and it can protect the sharp edges of the image, but it performs poorly against Gaussian noise.

Compared with mean filtering and median filtering, Gaussian filtering can keep more features of the overall gray distribution of the image when smoothing the image, so we choose Gaussian filtering in this process. Through Fourier transform, Gaussian filtering, and inverse Fourier transform, get the filtered image, after subtracting from the original image, the image enhanced image is obtained as shown in Figure 3(b). Compared to Figure 3(a) without any image processing, the defects on the headset surface in Figure 3(b) are more obvious.

2.2. Scratch Detection. During the processing of the Bluetooth headset, scratches may occur on the surface, which affects the appearance of the headset. The scratches are linear, so the usual method in scratch detection is to detect the lines on the surface of the object. For example, some edge detection operators are used to detect the lines on the surface of the object. The more commonly used methods are Sobel operator, Canny operator, etc. The Sobel operator detection method has a better effect on image processing with gray gradient and more noise. When the accuracy requirements are not very high, it is a more commonly used edge detection method. Canny operator [23] is a detection operator proposed by computer scientist John F. Canny in 1986. It is currently the most comprehensive detection algorithm theoretically. Canny operator is not easy to be interfered by noise. Canny operator is a multistage optimization operator with

![Flow chart of the surface defect detection algorithm.](image1)

![The result of threshold segmentation.](image2)
filtering, enhancement, and detection. Before processing, the Canny operator first uses Gaussian smoothing filter to smooth the image to remove noise. The operator uses the finite difference of the first-order partial derivative to calculate the gradient amplitude and direction. In the process, the Canny operator also goes through a process of nonmaximum suppression. Finally, the Canny operator also uses two thresholds to connect the edges. We use the canny operator to detect the scratches on the surface of the headset, and the results are shown in Figure 4(b).

Figure 3: The result of image enhancement.

From a morphological point of view, a scratch is composed of a series of adjacent pixels with a large difference in gray value from the background, usually a connected area. However, in the process of image preprocessing, the breakpoints with a small area divided by the shallow scratch may be filtered out, causing the scratch to break, which affects the result of line detection. At this time, performing a morphological expansion operation on the scratch to fill the scratch defect cavity can connect some adjacent fracture scratches, thereby solving the problem of scratch truncation. Then, through the skeleton extraction algorithm, a connected area is refined into the width of one pixel, and all the obtained skeleton subsets are combined to form the final scratch skeleton. The skeleton is constructed in a way that each point on it can be seen as the center point of a circle with the largest radius possible while still being completely contained in the region. The result of skeleton extraction is as shown in Figure 4(c).

From the detection results in Figure 4, the performance of the skeleton extraction algorithm is better than the line detection algorithm.

2.3. Glue-Overflow Detection

2.3.1. Watershed Algorithm. The watershed segmentation method is a mathematical morphology segmentation method based on topological theory. The basic idea is to regard the image as a topological toposgraphy in geodesy. The gray value of each pixel in the image indicates the altitude of the point. Each local minimum and its affected area are called collection basins, while the boundary of the collection basin forms a watershed.

Firstly, we compute the watersheds without applying a threshold $T$, resulting in the same basins that would be obtained when calling watersheds. Secondly, the basins are successively merged if they are separated by a watershed that is smaller than threshold $T$. Let $B_1$ and $B_2$ be the minimum gray values of two neighboring basins and $W$ the minimum gray value of the watershed that separates the two basins. The watershed is eliminated, and the two basins are merged if

$$\max\{W - B_1, W - B_2\} < T.$$ (4)

Through testing, we can get better image segmentation results by setting $T$ to 5. Figure 5 shows the detection results of watershed detection algorithm.

Watershed algorithm is an image region segmentation algorithm. In the process of segmentation, it will take the similarity with adjacent pixels as an important reference basis. Pixels with similar spatial positions and similar gray values are connected to each other to form a closed contour. Through the watershed algorithm, the headset is divided into several small areas according to the similarity of the gray values, which are used as the input of the subsequent gray level cooccurrence matrix.

2.3.2. Gray Level Cooccurrence Matrix. The gray level cooccurrence matrix (GLCM) was first proposed by Haralick et al. in 1973. GLCM describes the spatial relationship of gray levels in texture images and has been widely used in texture statistics and analysis. The texture is formed by alternating gray levels in spatial positions, so there is a certain spatial relationship between two pixels separated by a certain distance in the image. GLCM is a commonly used method to express the spatial correlation of pixel gray levels, mainly describing the image from the distance, direction, and degree of change between adjacent pixels. Its essence is to calculate the appearance frequency of two gray pixels under a certain spatial relationship, which can indicate the regional consistency and relativity of the image. GLCM changes quickly in fine textures with distance, while coarse textures change slowly. GLCM defines a square matrix whose size represents the probability of the gray value $i$ from a fixed spatial position relationship (size and direction) to another gray value $j$. 
Suppose \( f(x, y) \) is a 2D grayscale image, where \( S \) is a set of pixels with a certain spatial relationship in the area and \( P \) is GLCM, which can be expressed as

\[
P(x, y) = \# \{ (x_1, y_1), (x_2, y_2) \in S \mid f(x_1, y_1) = i, f(x_2, y_2) = j \}.
\]

\(#(X)\) represents the number of elements in the set \( X \).

Applying GLCM to describe texture features is based on two-order statistical parameters as texture metrics. Haralick [26] proposed 14 feature statistics to describe image texture features. However, we usually use the following 4 statistical types: energy (ASM), contrast (CON), correlation (COR), and entropy (ENT). The arc second moment (ASM) reflects the regularity and uniformity of image distribution. Contrast (CON) reflects the depth and smoothness of the image texture structure. Correlation (COR) reflects the similarity of image texture in the horizontal or vertical direction. Entropy (ENT) is a measure of image information, reflecting the complexity of texture distribution. The 4 types of statistics are as follows:

\[
ASM = \sum_i \sum_j P(i, j)^2,
\]

\( i \) \& \( j \) range from 1 to \( N \).

\[
CON = \sum_i \sum_j (i - j)^2 P(i, j),
\]

\( i \) \& \( j \) range from 1 to \( M \).

\[
COR = \frac{\sum_i \sum_j (i - \bar{i})(j - \bar{j})P(i, j)}{\sigma_x \sigma_y},
\]

where \( \bar{i} = \frac{1}{N} \sum_i i \), \( \bar{j} = \frac{1}{M} \sum_j j \), \( \sigma_x = \sqrt{\frac{1}{N} \sum_i (i - \bar{i})^2} \), and \( \sigma_y = \sqrt{\frac{1}{M} \sum_j (j - \bar{j})^2} \).

\[
ENT = -\sum_i \sum_j P(i, j) \log P(i, j).
\]

We pick out 100 glue overflow areas and 100 normal areas from the small areas extracted by the watershed algorithm and count their energy (ASM), contrast (CON), and correlation (COR), respectively. The results are as follows:

In the gray level cooccurrence matrix, the energy value reflects the uniformity of the image gray level distribution and the texture thickness. If the element values of the gray level cooccurrence matrix are similar, the energy value is smaller, which means the texture is detailed, and the energy value is large, which indicates a more uniform and regular texture pattern. Therefore, the area without glue-overflowed is more regular and usually has a larger energy value, while the area with glue-overflowed has more texture and smaller energy value.

As shown in Figures 6 and 7, the ASM score of the glue-overflowed area is almost less than 0.2, while the ASM score of the area without glue-overflowed is almost greater than 0.3.

Therefore, set the ASM threshold \( T_{asm} \) to 0.2; ASM(\( i \)) represents the ASM value of the area \( i \).

The area \( i \) that satisfies the following formula is judged to be a possible glue-overflowed area.

\[
ASM(i) < T_{asm},
\]
In the gray level cooccurrence matrix, the contrast (CON) reflects the sharpness of the image and the depth of texture. The deeper the texture, the greater the contrast.

As shown in Figures 8 and 9, the CON score of the area without glue-overflowed is almost less than 0.15, while the ASM score of the glue-overflowed area is almost greater than 0.2.

Therefore, set the CON threshold $T_{\text{con}}$ to 0.2; $\text{CON}(i)$ represents the CON value of the area $i$.

The area $i$ that satisfies the following formula is judged to be a possible glue-overflowed area.

$$\text{CON}(i) > T_{\text{con}}.$$  \hfill (11)

In the gray level cooccurrence matrix, the correlation (COR) reflects the local gray-scale correlation in the image. If there is texture distribution along a certain direction, the correlation value of GLCM is larger.

As shown in Figures 10 and 11, the COR score of the glue-overflowed area is almost greater than 0.9, while the COR score of the area without glue-overflowed is almost less than 0.9.

Therefore, set the COR threshold $T_{\text{cor}}$ to 0.9. $\text{COR}(i)$ represents the COR value of the area $i$.

The area $i$ that satisfies the following formula is judged to be a possible glue-overflowed area.

$$\text{COR}(i) > T_{\text{cor}}.$$  \hfill (12)

If the area $i$ satisfies Eq. (10), Eq. (11), and Eq. (12) at the same time, we judge the area $i$ as the glue-overflowed area. Finally, the adjacent glue-overflowed areas are merged, and the areas that cannot be merged and are particularly small are removed. The final test results are shown in Figure 12. (a–d) are the detection results of glue-overflowed at different positions of the headsets. From the figure below, the glue-
overflowed detection effect based on the watershed algorithm, and the gray level cooccurrence matrix algorithm is good.

3. Experimental Results and Analysis

In order to verify whether the image enhancement algorithm improves the defect detection ability, we compare the image of the headset without image enhancement algorithm and image enhancement algorithm. As shown in Figure 13, (a) is the original picture of the headset, (b) is the picture after image enhancement, and (c) and (d) are the results of glue-overflowed detection on (a) and (b), respectively. It is proved from the figure that through the image enhancement algorithm, the glue-overflowed area in the image is more obvious, and the misdetection caused by the illumination is reduced. Therefore, the image enhancement algorithm proposed in this paper helps to improve the accuracy of the headset surface defect detection.

To further verify the proposed surface defect detection algorithm performance, we have defined the precision ($Pr$) and recall ($Re$) which are calculated as Eq. (13) and Eq. (14). The corresponding results are reported in Tables 1 and 2.

$$Pr = \frac{TP}{TP + FP}, \quad (13)$$
$$Re = \frac{TP}{TP + FN}, \quad (14)$$

where $TP$, $FP$, and $FN$ denote True Positive, False Positive, and False Negative which correspond to the count of actual defective products detected with defects, actual
Figure 12: The result of glue-overflowed detection.

Figure 13: The comparison of the image enhancement algorithm.
We draw the precision and recall of the two detections into a line graph, as shown in Figures 14 and 15.

Through actual tests, the scratch detection precision is approximately 98%, and the recall is approximately 100%. The glue-overflow detection precision is approximately 98%, and the recall is approximately 100%. In addition, the precision and recall of the detection are not affected by the number of images. Basically meet the testing requirements of industrial testing.

4. Conclusions

Taking the surface defect detection of Bluetooth headset as an example, this paper studies a surface defect detection algorithm for irregular industrial products based on machine vision, which can detect both surface scratches and glue-overflowed of irregular industrial products. At present, the Bluetooth headset surface detection algorithm has been applied in the actual industrial production; the practice proves that the detection algorithm proposed in this paper is effective. Meanwhile, it is prone to be applied to the surface defect detection of other products with similar surface features.

Data Availability

Data available on request. The data are available by contacting Mengkun Li (limengkun@cnu.edu.cn).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This research was funded by the Major Technology Project of China National Machinery Industry Corporation (SINOMACH): “research and application of key technologies for industrial environment monitoring, early warning and intelligent vibration control (SINOMAST-ZDZX-2017-05),” and partially supported by the Scientific Research Foundation of Beijing Municipal Education Commission (KM201810028021).

References


