A surface defect identification method based on improved threshold segmentation algorithm

Xinglong Feng,1a* Xianwen Gao,1 Ling Luo1
1Northeastern University, College of Information Science and Engineering, Shenyang, China, 110819
a1610238@stu.neu.edu.cn *Corresponding author’s Email: fengxinglong@vip.163.com

ABSTRACT: For enterprises, defect detection is very important because it is related to the quality of products produced by enterprises. With the development of machine vision, accurate analysis of image data benefits defect detection. In an enterprise that produces electronic cigarettes, professional and technical personnel used to detect a defect in a workpiece by using manual testing. The defect detection rate of this method is only 95%, and the efficiency is low. We use an improved threshold segmentation method to solve this problem in this paper and we have achieved success. Compared with typical methods, the accuracy of our proposed algorithm reaches over 99% and at the same time, the detection efficiency has been improved by more than 50%. Our method also has the advantage of simplicity, practicality and low cost.

1. INTRODUCTION

1.1. The research status of machine vision in the field of surface defects
Machine vision detection method has the characteristics of wide application range, high accuracy, free from the influence of the contour of detection parts, high detection efficiency and automatic detection[1]. Many researchers have studied the surface defect detection based on machine vision. For example, Lin Jian-Chun et used machine vision method to identify the surface defects of photovoltaic modules[2]. Zhu Yan-Yan et used image processing technology to detect the surface defects of rubber tubes[3]. Wu Ting et used laser and machine vision to detect defects in pipelines[4]. Yuan Xiao-Cui et used machine vision detection method to detect defects in the rail[5]. Li Jianyou et adopted binocular vision inspection to conduct binocular vision inspection on welded parts[6]. With the development of artificial intelligence technology, deep learning-based methods have been widely used in product defect detection due to their ability to fit arbitrary complex functions and better feature extraction[7]. Many researchers used deep learning to detect defects in products to improve their quality[8-10].

1.2. the background of this paper
Although the predecessors have done a lot of in-depth research on defect detection and achieved fruitful results. However, in the actual industry, there are still some problems need to be further solved. For example, an enterprise produces electronic cigarettes’ workpieces and assembles them. It is necessary to stick glue on each workpiece in order to put them together. In the area where the workpieces need to be glued, it is qualified if the glued part reaches more than 50%. It is allowed to detect some normal parts as defects, for example: it is acceptable if the area of the adhesive part of a part is 55%, but it is judged to be unqualified. However, it is never allowed to test the defective parts as normal. At present, due to the enterprise has not found a more appropriate method after studying...
this problem, they use manual methods to identify defects. The disadvantage of this method is that it is inefficient and less accurate. According to reliable sources, the manual method is less than 95% accurate. Figure 1 shows some samples of the parts.

![Figure 1 Sample images of the parts](image)

As it is shown in Figure 1, the workpieces are first fixed by workers at specific locations on the testing table, so the workpiece that need to be glued are basically fixed in the image, while the glued parts look very small. Therefore, the image need to be locally enlarged to better show the glued part. In practice, workers also use software to zoom in on images. After regional enlargement of the part surface, the obtained part surface is shown in figure 2. Since it is necessary to proceed to the next step before the glue dries after the test, the test speed should not be too slow, otherwise the glue on the surface of the part will cause adverse effects on the next step.

![Figure 2. Area to be detected of the workpieces](image)

As is shown in Figure 2, in the actual working condition, the glue on the surface of the workpieces is colorless and transparent. However, when illuminated by an auxiliary light source, the glue-covered parts of the workpieces appear black and white in the camera. In the actual test, the principle of the upper and lower parts is the same. Mark the gluing part on the surface of the part red, as shown in Figure 3.

![Figure 3. The area of the glue area](image)

2. Material and method

This paper will propose a reasonable solution to the problem of surface defect detection that has been put forward in front. Since the experts of the enterprise hope that the proposed scheme can effectively improve the accuracy of detection, so engineering effectiveness is the primary factor to be considered in this paper. For the existing detection equipment of the enterprise does not contain high-performance computers, in order to avoid additional economic burden on the enterprise, we first consider the algorithm that can be effective on ordinary hardware. The enterprise provided us with samples of no more than 400 pictures, and for some reasons, it is difficult for us to obtain new samples. So, we need to design an algorithm that works for small samples. In addition, because of the enterprise will produce many different products; we hope that the proposed algorithm can have a certain migration ability. In this way, once similar products need to be tested, the enterprise can safely use the algorithm we proposed after simple modification.
3. Theory and calculation

3.1. Threshold selection based on human experience
Select an image of the part that has been pasted with glue and enlarge it as shown in figure 4. As shown in figure 4, we can divide the ROI area into the following three parts: the part that appears white after gluing, the part that appears black after gluing, and the part that has not been glued.

![Figure 4](image)

First, we find some of the pixel values in the white area after gluing, represented by A. Next, we find some of the pixel values in the black area after gluing, represented by B. Then, we find the part that has not been glued, represented by C. We analyze the above pixels and we can draw the following conclusions: (1) All pixel values in A are above 180, and most are above 200. (2) All pixel values in B are below 35 and most are below 20. (3) Most pixel values are above 35 and below 180, and individual pixel values are below 35. Therefore, the parameter values of 35 and 180 can identify almost all the adhesive parts on the surface of the workpiece while ensuring a relatively small noise ratio. So we chose 35 and 180 as the thresholds.

3.2. Defect identification theory based on improved threshold segmentation
After selecting the ROI area, set the pixel value below 35 to 0 and the pixel value above 180 to 255. Calculate the number of pixel 0 and the number of pixel 255 in the ROI area, that is, preliminarily calculate the area of the gluing part. After this step, we divide the pixel value of the gluing part by the pixel value of the entire ROI area to get the proportion of the gluing part, expressed as k. Of course, the calculated proportion of gluing part has some errors at present. An algorithm is then needed to eliminate this error.

The principles on which the method in this paper is based are as follows: Suppose there are two images A and B. A represents the image of the part after glue was applied, and B represents the image without glue and they both have an area of 1. We use Y to represent the sum of the areas in B with pixel values below 35 and above 180. As a result of the actual situation, the pixel values below 35 and above 180 in B are related to the characteristics of the part and the reflection, so we can assume that the area of A below 35 and above 180 is also Y without glue. We use X to represent the sum of the areas with A pixel value below 35 and the areas with A pixel value above 180. We use U to represent the area of the glued area. Notice that U is the unknown, and X and Y are known. We consider the case where the pixel value in B is below 35 and the pixel value is above 180 as noise. Suppose the noise in A and in B obeys random distribution, so we can get formula (1)

$$U + (1-U) \times Y = X$$  \hspace{2cm} (1)

As it is shown in formula (1), U is the glued area. We transform this formula to get the expression for U, as shown in formula (2).

$$U = (X - Y) / (1 - Y)$$  \hspace{2cm} (2)

The expression of U is valid if it satisfies our hypothesis that the noise (points with pixel values below 35 and points with pixel values above 180) is randomly distributed. In practice, however, noise is not completely random. Instead, there will be a degree of aggregation in the center of the ROI region. Since the actual ROI area is often glued, this can cause the calculated U value to be slightly larger than the actual value. Fortunately, due to the relatively small ratio of noise, the value of U is not much larger than the actual adhesive area. We can use D for this bias. As a rule of thumb, D is usually no more than 0.05. And then the expression for U becomes as formula (3).
U = (X - Y) / (1 - Y) - D                                (3)

Next, we will use the above theory to test the actual images. Note that in practice, because workers need to put more than one workpiece at a time into the designated position of the testing table, and then test the workpiece based on it. Therefore, there may be cases where workers fail to put the workpiece in due to negligence. In this case, we should also judge it as unqualified. We made a similar analysis to the above theory on the image of the unfilled workpiece on the testing platform, and found that in this case, the U value obtained could never be greater than 0.5. Therefore, we can use the theory proposed above to detect both the workpiece with less than 50% glue content and the workpiece not placed on the testing table.

4. Validation and Results

4.1. The validation results of the actual dataset

Select an image of the part to be tested without gluing at all. Find the ROI area of the image and calculate the proportion of pixel values from 0 to 35 in the ROI area and the proportion of pixel values above 180. In the image of workpieces without glue, the proportion with the pixel value below 35 is denoted as x, and the proportion with the pixel value above 180 is denoted as y. Therefore, it can be known that the error of u = x + y will be generated by direct calculation of parts without gluing by using the aforementioned method. We calculated that the value of x is 0.165 and the value of y is 0.012. So u is 0.177. If we assumed that the number of pixels below 35 and above 180 on the surface of parts without glue is random. So there is a conclusion that can be reached in formula (4)

\[ k = M + (1 - M) \cdot u \] (4)

In formula (4), M represents the proportion of the real gluing part. According to the formula, we can get formula (5)

\[ M = \frac{k - u}{1 - u} \] (5)

Since in practice the number of pixels below 35 and above 180 are not very random. So we need to minus one bias which is represented by d in order to avoid that he proportion of sticking glue is too larger. For the enterprise wants to check the quality strictly, that is to say, it does not want the defective products to be tested as qualified products. So let's set a slightly larger value of d, which is 0.03. In this way, we can get the formula of the proportion of the glue area in formula (6)

\[ M = \frac{k - u}{1 - u} - d = (k - 0.177)/(1 - 0.177) - 0.03 = 1.215 \cdot k - 0.245 \] (6)

In this way, the proportion of the area of sticking glue is obtained. According to formula (6), we can calculate the value of k for each image that needs to be detected. Then substitute the k value of each picture into the formula to find the value of M. Finally, the images with M value above 0.5 (including 0.5) are judged as qualified, while the pictures with M value below 0.5 are judged as unqualified.

We test the validity of the proposed algorithm in practice. A total of 335 pictures of parts to be tested were collected from the site. In addition, an image of an unglued part is also required. The method described above is used to build the model of detecting parts. According to the established model, the proportion of glue in each image to be tested can be obtained. Of all the images, there are 42 images that were unqualified.

According to the standards of the enterprise, the 42 images that were judged to be substandard were acceptable. Meanwhile, the remaining 293 pictures that were tested as qualified were all qualified. We haven't missed any unqualified images. Of course, there is a price to be paid for maintaining such high accuracy in identifying defective parts. In other words, we will judge the individual qualified parts as unqualified. There are three images that were judged as unqualified.

4.2. The method in this paper compared with manual method

We compared the results of this paper with the manual testing methods currently used by the enterprise. The comparison results are shown in table 1.
Tabel 1 The method in this paper that compared with the manual method

|                          | the method of this paper | manual detection |
|--------------------------|--------------------------|------------------|
| Rate of detect detection | Almost 100%              | Nearly 95%       |
| False detection of normal workpieces | 1.01%                  | About 2%         |
| Number of tests per hour | About 1800               | 1000             |

According to table 1, we can see that the method proposed in this paper has a higher defect detection rate than the manual method. At the same time, the error detection rate of normal workpieces is lower than that of manual methods. In addition, the method in this paper is also much faster than the manual detection method. At present, due to the expansion of production, the enterprise will add testing equipment. If the enterprise adopts this method, the existing testing equipment is enough. In other words, the method in this paper can also save some equipment costs for the enterprise.

4.3. Conclusion

From the comparison in the previous section, we can see that: the method proposed in this paper is superior to the manual method in detection rate, recall rate and detection speed. In other words, we theoretically prove that the method in this paper has a quota advantage.

At the same time, the method we proposed was also recognized by the technical personnel of the enterprise. In the actual production of the enterprise, the enterprise can combine the method we proposed with the manual scheme at the same time. This application scheme can reduce the identification of unqualified parts as qualified due to human error. Or companies can let algorithms and humans simultaneously mark the proportion of surface area glued to parts. This can lead to more convincing results. Anyway, the proposed scheme can improve the accuracy of surface defect detection. It also provides a new way of thinking for related academic research.

References

[1] Yuan Wei-Qi, Xue Dan. Review of tunnel lining crack detection algorithm based on machine vision. Chinese Journal of Scientific Instrument, 2017,38(12):3100−3111
[2] Lin Jian-Chun, Yang Ai-Jun, Shen Yao-Hui. Evaluation of imaging performance for electroluminescence defect detector. Optics and precision engineering,2017,25(6):1418−1428
[3] Zhu Yan-Yan, Zuo Jian-Hua, Lu Jian-Ping, Xu Dong-Xiao. A on-line detection system development based on image processing for rubber hose defects. Transactions of Beijing Institute of Technology,2017,37(9):937−941
[4] Wu Ting, Lu Shao-Hui, Han Wang-Ming, Hu Ke-Gang, Tang Yi-Ping. In-pipe internal defect inspection method based on active stereo omni-directional vision sensor. Chinese Journal of Scientific Instrument,2015,36(10):2258−2264
[5] Yuan Xiao-Cui, Wu Lu-Chen, Chen Hua-Wei. Improved Image Preprocessing Algorithm for Rail Surface Defects Detection. Journal of Computer-Aided Design & Computer Graphics,2014,26(5): 800−805
[6] Galan U, Orta P, Kurfess T, et al. Surface defect identification and measurement for metal castings by vision system. Manufacturing Letters,2017,15:5−8
[7] Li Shao-Bo, Yang Jing, Wang Zeng, Zhu Shu-De, Yang Guan-Chi. Review of development and application of defect detection technology. ACTA AUTOMATICA SINICA, ,2018, XX(X):X−X.
[8] Tao X, Wang Z, Zhang Z, et al. Wire Defect Recognition of Spring-Wire Socket Using Multitask Convolutional Neural Networks. Ieee Transactions on Components Packaging and Manufacturing Technology, 2018, 8(4): 689−698.
[9] Jiang J, Chen Z, He K. A feature-based method of rapidly detecting global exact symmetries in CAD models. Computer-Aided Design, 2013, 45(8): 155−171
[10] Cheng J C P, Wang M. Automated detection of sewer pipe defects in closed-circuit television images using deep learning techniques. Automation in Construction, 2018, 95: 1081−1094.DOI : https://doi.org/10.1016/j.autcon.2018.08.006