Research on Product Appearance Detection System Based on Image Sparse Representation

Xiaoning Chen1,* and Jian Zhao2
1Dongguan Polytechnic, Dongguan, China
2Northwest University, Xi’an, China

*Corresponding author

Abstract. The USB camera is used as the front device of the image acquisition. The host computer is a PC for the detection system. Aiming at low resolution of USB camera, the super-resolution image reconstruction method based on sparse representation is used to overcome the low resolution for the low-cost imaging sensor and improve the image resolution. In order to reduce the influence due to shooting light, angle, etc., visual keywords are used for image matching to improve the robustness of the system. The experiment proves that the system is simple in structure and easy to operate, and the detection accuracy of the system can reach 98.67%.

Keywords: USB camera; sparse representation; super-resolution reconstruction; visual keyword.

1. Introduction

With the development of digital equipment and the machine vision in recent years, the traditional manufacturing industry relies on manual detection of product appearance have been difficult to adapt to current production requirements due to misjudgment and low efficiency. In recent years, some researchers have also done some research on product appearance inspection. Takahashi et al. [1] used virtual equipment to develop the design and operation support technology of the equipment for product detection, which provided a reference for the appearance detection of products, but the technology basically stayed in theoretical research. Huang et al. [2] used image processing and neural networks to detect the appearance defects of polarizing films. The research has certain practical value, but the object of detection is single and cannot be popularized and applied. Zhang et al. [3] proposed the technology of appearance detection for pipeline products based on LabVIEW, which has a simple structure, is conducive to ready-made algorithms, and is easy to implement, but the detection is not good. Machine vision-based detection devices have been used in various fields of industrial manufacturing. However, the detection of product appearance includes the inspection of packaging boxes of products, labels, and the product itself. Due to the variety of defects, the detection has increased difficulty. This research proposes a device for product appearance detection based on a USB camera. The USB CCD camera is used to collect images of specific areas of the product. In conjunction with a supporting automatic determination software system, the collected images are compared with images stored in the standard image library for judging of the local appearance of the product. However, the resolution of USB cameras is low, and the obtained images are not sharp enough. The process of super-resolution image reconstruction (SRIR or SR) overcomes the inherent low resolution of low-cost imaging sensors. At the same time, in order to improve the accuracy of
image detection, we take into account the color and texture characteristics of the image, and use the visual keywords that match the local area of the image to implement the image detection. The rest of this article is organized as follows. The second section describes the system's composition, including the system's hardware composition and the user interface. The third section describes the system's basic principles for image processing algorithm and implementation. In section 4, we present the experimental results and analysis. Finally, Section 5 summarizes this article.

2. System Composition

2.1. System Hardware Composition
The hardware structure of the product appearance detection system based on the USB camera is shown in Figure 1. This system uses a USB digital camera as the front-end acquisition for image, which resolution is 640pixel × 1480pixel, and the focal length is 2.8 ~ 12mm. The host computer is a PC, using an i3 processor and windows7 operating system. It constitutes a system of the product appearance detection with simple hardware system and convenient operation. The real device is shown in Figure 2. In this system, the analysis software module is the key for the system and which is also the focus and difficulty of the research.

![Figure 1. System composition.](image1)

![Figure 2. The real device.](image2)

2.2. System User Interface
This system uses a PC as the host computer and performs system development under the Windows operating environment. Because the product is aimed at the operator, they are not proficient in computer operation knowledge, so the humanized design of the human-machine interface is very important. The user interface should be simple and clear, and the operation is simple and convenient. The detection system consists of three types of detection interfaces: login interface, image library entry interface and appearance detection interface.

Login interface: The system enters the login interface when the system is running, and requires the user to log in with an authorized account, as shown in Figure 3. The system login includes administrators and ordinary users. The specific user login process is shown in Figure 4. The system only opens the standard library entry function for administrators, while ordinary users only open the test function.
Figure 3. Login system UI.

Library entry interface: Figure 5 is the system library entry interface. The administrator logs in to the gallery entry interface. By this system, the user can enter the appearance of the product to be detected through the camera. The input image is a large-scale image. The user can select the local area that needs to be detected according to the actual needs. As shown in Figure 5, the left image is the camera shooting area, the green dotted frame is the image of the selected detection area, and the right is the selected image area, which is the standard image of the library.

Figure 4. Login system Flower.

Appearance detection interface: Figure 6 shows the system for appearance detection interface. The appearance detection is mainly to replace the tedious manual operation. Because the appearance detection is relatively intuitive, it is not necessary to record any detection data during the detection, as long as it is judged as qualified or not. As shown in Figure 8, the left side of each picture is the standard image from library, the right side is the detection image, and the lower left side is the detection result.

Figure 5. Library entry interface.

Figure 6. Appearance detection interface and result judgment. (a) Test results of qualified appearance and (b) Test results of unsatisfactory appearance.
3. Principles and Implementation
The overall workflow of the system for product appearance detection based on the USB camera is shown in Figure 7, which is mainly divided into two parts. 1) Image pre-processing. The image is pre-processed by super-resolution reconstruction of the acquired image. 2) Image matching. First, the visual keywords of the image are extracted, and then the visual keywords are used to perform feature matching between the test image and the standard image, and the threshold of the number of keyword matches is set to realize the detection result output.

3.1. Image Pre-processing
The image acquisition system based on the USB camera has a low resolution for the obtained image, and the quality of the obtained image is not high without an auxiliary light source and isolated from ambient light, and the image needs to be pre-processed. In order to obtain high-quality acquired images, super-resolution reconstruction (SR) is performed on the first acquired images [4-6]. Currently, super-resolution image reconstruction mainly uses three types of algorithms based on interpolation, reconstruction, and learning. The interpolation method has a simple algorithm. The brightness value of the pixel to be interpolated is calculated from the brightness values of neighboring pixels, such as the bicubic interpolation method [7]; Based on the reconstruction method, the image degradation model is used to constrain the reconstructed image to achieve reconstruction [8]; based on the learning method, the prior knowledge of high-resolution images is used to construct a training set, and the image is reconstructed by the training model. After experimental comparison, this system uses an image super-resolution reconstruction algorithm based on sparse representation.

3.1.1. Principle of Super-resolution Image Reconstruction Algorithm Based on Sparse Representation.
The super-resolution image reconstruction method based on sparse representation has better robustness and flexibility [9-11], and it has better effect on low-resolution image reconstruction. The core principle and process of the algorithm are shown in Figure 8.

Figure 7. The workflow for the system.

Figure 8. Image sparse representation principle and super-resolution reconstruction process.
The super-resolution image reconstruction algorithm based on sparse representation is mainly divided into a dictionary training phase and a sparse coding phase. The high-resolution and low-resolution images extracted from the product appearance are used as the training set, \( P = \{ X^h, Y^l \} \) in the dictionary training phase, where \( X^h = \{ x_1, x_2, x_3, \ldots x_n \} \) is the image feature set extracted from the high-resolution image. \( Y^l = \{ y_1, y_2, y_3, \ldots y_n \} \) is a collection of image features extracted from a low resolution image. The optimized models for obtaining high-resolution dictionaries \( D_h \) and low-resolution dictionaries \( D_l \) are as follows:

\[
D_h = \arg \min_{\beta, Z} \left( \| X^h - D_h Z \|^2_2 + \lambda \| Z \|_1 \right)
\]

(1)

\[
D_l = \arg \min_{\beta, Z} \left( \| Y^l - D_l Z \|^2_2 + \lambda \| Z \|_1 \right)
\]

(2)

Making the image features of high resolution and the low resolution have the same sparse representation. Therefore, it is realized that the complete dictionary \( D_h \) and \( D_l \) are learned by using the image feature training set.

The sparse coding phase is a process of realizing the reconstruction of the low resolution image \( Y \) into the high resolution image \( X \). According to Donoho's research results, as long as the coefficients \( \alpha \) are sufficiently sparse, for low-resolution image block \( Y \), sparse representation in the dictionary \( D_l \) can be implemented and equivalently converted to the following expression:

\[
\min_{\alpha} \left( \| FD_l \alpha - F y \|^2_2 + \lambda \| \alpha \|_1 \right)
\]

(3)

Where \( F \) is the operator that extracts image features, commonly used to extract high-frequency components in the image, \( \lambda \) is the penalty parameter, balances the sparsity of \( \alpha \) and the precision of approximate solution of \( Y \). The above formula is optimized multiple times and finally transformed into the following optimal solution problem:

\[
\alpha^* = \arg \min_{\alpha} \left( \| D \alpha - \bar{y} \|^2_2 + \lambda \| \alpha \|_1 \right)
\]

(4)

Where \( \bar{D} = \begin{bmatrix} FD_l \\ P \end{bmatrix} \), \( \bar{y} = \begin{bmatrix} FY_l \\ P \end{bmatrix} \), \( \alpha^* \) is the optimal sparse representation of the low resolution image block in the low resolution dictionary. \( P \) is the overlap area between the last reconstructed image block and the image block currently being reconstructed. \( w \) is the value of the overlapped region of the last reconstructed high resolution image block. Finally, a high-resolution image block is obtained according to the following equation, and finally a high-resolution image \( X \) is obtained.

\[
x^* = D_h \alpha^*
\]

(5)

3.1.2. The effect of super-resolution image reconstruction based on sparse representation. According to the super-resolution image reconstruction principle mechanism introduced above, the super-resolution reconstruction of the low-resolution product appearance image acquired by the USB camera is performed, and the image clarity is obviously improved. As shown in Figure 9.
3.2. Image Matching Algorithm

3.2.1. Image matching process. In general, useful information in an image is often expressed in local areas [12]. Therefore, the image matching algorithm in this system uses a feature matching algorithm based on visual keywords [13] to achieve image detection. The visual keyword extraction of the image uses the method I described in [13] for visual keyword extraction for the collected image. Therefore, the matching of the two images eventually becomes the matching of the visual keywords. The specific algorithm flow is shown in Figure 10 below:

\[ P = (1 - \frac{n}{N}) \times 100\% \]  

(6)

Figure 9. Product’s appearance image reconstruction effect. (a) USB camera capture image, (b) Double cubic interpolation and (c) This method.

Figure 10. Image matching principle and process.

3.2.2. The number of visual keywords affects verification. According to the results of research in [14], an image is expressed as 50-80 visual keywords, which can basically express the image completely. We take 50, 60, 70, 80, 90, and 100 visual keywords to detect labels, patterns, text, and QR codes, respectively. We use detection accuracy to measure the effect of detection, as shown in Figure 11. Let the number of samples misjudged in the experiment be \( n \) and the number of test samples be \( N \), then the accuracy is
According to the analysis, it can be seen that when the number of visual keywords increases, the detection accuracy rate also increases. However, when the number of visual keywords is increased to 80, the increase in accuracy is very small. Considering that the larger the number of visual keywords, the larger the amount of calculation data, which affects the detection speed, and finally we take 80 visual keywords for image matching detection.

4. Results and Analysis

In the appearance inspection system, we set the image matching degree in the target image and the standard image library to more than 90%, that is, when $80 \times 90\% = 72$ visual keywords are consistent, the system judges to be "qualified", at $80\% - 90\%$, it is judged as "defect". Less than $80\%$ is "failed" and the matching threshold can be modified according to product requirements.

The system uses a USB camera with a resolution of $640 \times 1480$ pixels and a focal length of 2.8 to 12 mm. The host computer uses i3 processor, windows7 operating system. This paper carries out experimental verification through the collected product appearance image database. The image library contains 100 kinds of product appearance images, and each product captures 10 images. The top 5 images of similar products are used to train the model, and the last 5 are used as test samples. Randomly extract 4 different types of product appearance images, each type includes 5 images, including labels, patterns, texts and QR codes, as shown in Figure 12.

Figure 13 is a random selection of four types of appearance for verification, including labels, patterns, and texts. Each type extracts 10 of product appearance images, and total of 150. Due to the small number of acquisitions, the QR code is extracted from 5 products and total 25 for verification. Firstly, it is manually detected, and the detection result of the algorithm is in good agreement with the manually detection result, which indicates that the detection accuracy of the algorithm is high. In order to verify the accuracy of the algorithm, three different algorithms (including the algorithm) are used for detection, and manual detection is used as a standard. And the number of misjudgments for each method is counted. There are only two cases in this judgment result, that is, when the matching keyword is greater than or equal to 90%, the result is set to pass (non-defective), otherwise it is unqualified (defective). There are two cases of misjudgment. One is to misjudge the non-defective as a defect; the other is to judge the defect as non-defective. And the detection accuracy is calculated according to the formula (6).
5. Conclusion
The product appearance detection device based on USB camera obtains a better quality of the acquisition image by reconstructing the low-resolution image by super-resolution. A matching algorithm based on the visual keywords of the local area of the image is used to match the collected image with the database image, and the final image discrimination result is output according to the matching degree of the keywords. It can be seen from the experimental results that this system has a high accuracy rate of product appearance detection. This paper first applied the key visual theories to the product appearance detection, and provided a reference for further research. It can be seen that the system has a low detection accuracy for QR codes. Therefore, in the later research, we will focus on the detection methods for QR codes.

Acknowledgments
This work was supported by the “Dongguan City Social Science and Technology Development (General) Project”, (Project number: 2019507156720), and also by Guangdong Provincial Department of Education 2018 Guangdong Province University's Characteristic Innovation Project (Project No. 2018GKTSCX062).

References
[1] Takahashi F and Abe H 2015 Development of design and operation supporting techniques for product inspection devices using virtual devices[J].(2015 14th IAPR International Conference on Machine Vision Applications (MVA), May 2015) pp 271-274
[2] Changchiun H, Chungfeng K and Jeffrey 2016 Inspection of appearance defects for polarizing films by image processing and neural networks[J]. Textile Research Journal. Vol 86(15) pp 1565-1573
[3] Yiwei Z and Jianxin Z 2014 Research on Appearance Inspection Technology of Pipeline Products Based on LabVIEW [J]. Journal of Zhongyuan Institute of Technology, 2014. Vol 12 PP 20-23
[4] Xinlei W and Naifeng L 2016 Super-resolution of remote sensing images via sparse structural manifold embedding [J]. Neurocomputing, 2016 Vol 173 pp 1402-1411
[5] Xiaoning C and Changming L 2019 Research on 4G wireless network law enforcement system based on super-resolution image reconstruction [J]. Electronic World, Vol 567(2019(05)) pp 46-48
[6] Ganjun W and Yanfeng S 2017 Sparse representation of depth image super-resolution reconstruction research and simulation [J]. Computer simulation, Vol 34(05) pp 234-237
[7] Huipeng W , Lili Z and Jie Z 2010 A region-based bicubic image interpolation algorithm [J]. Computer Engineering, Vol 36(19) pp 216-218
[8] Linwei Y and Huanfeng S 2016 Image super resolution: the techniques, applications, and future [J]. Signal Processing, Vol 128 pp 389-408
[9] Li S and Yanfang H 2015 Super-resolution image reconstruction based on sparse representation [J]. Electronic technology, Vol 28(9) pp 144—147
[10] YANG J, MEMBER S and WRIGHT J 2010 Image super-resolution via sparse representation [J]. IEEE transactions on image processing, Vol 19(11) pp 2861-2873
[11] Park S C and Park M K 2003 Super-resolution image reconstruction: a technical overview[J]. IEEE Signal Processing Magazine, Vol 20(3) pp 21–36
[12] Yong L 2009 Research on Content-based Image Retrieval Technology [D]. (Jilin, Jilin University) pp 23-32
[13] Xiaoning C and Jian Z 2017 Research on new method of visual keyword extraction based on point of interest detection[J]. Modern Electronic Technology, Vol 2017(21) pp 62-65