

Visual Abnormalities Detecting based on Similarity Matching

Liuchuan Yu ^a, Erjing Zhou ^b, Baomin Xu, Shuangyuan Yu

School of Computer and Information Technology, Beijing Jiaotong University, Beijing 100044, China

^a17127141@bjtu.edu.cn, ^b17127099@bjtu.edu.cn

Abstract. Abnormalities detecting is one important application in the field of image processing and pattern recognition. It can alleviate human workload and improve productivity that employing computer graphic image theory and image processing technology analyzes and matches images in order to detect the abnormal region in image which has broad application prospects. In this paper, we propose a new abnormality detecting method based on similarity matching to address whether either missing or error abnormalities existing in bound books in industrial situation. First of all, we denoise the image by means of digital image processing and transformation, extract the sub rectangular region containing bound books using contour matching and locate the area exactly matching the template image using template matching. After that, we get a binary denoised image to detect the missing abnormality and the error abnormality using shape matching. In addition, we introduce some thresholds to improve the performance. The experiments show that the method we proposed achieve a better or the same performance comparing with the state-of-the-art methods.

Keywords: Digital Image Processing, Pattern Recognition, Computer Vision, Abnormalities Detecting, Template matching.

1. Introduction

Computer image processing has been as a common instrument for dealing many kinds of problems. Nowadays, image processing has been applied in plenty of fields such as aerospace technology, biomedical engineering, communication engineering, industry and engineering, military public security, culture and art, robot vision, video and multimedia systems, scientific visualization and electronic business, etc. Among them, abnormalities detecting is one important purpose of image processing in real applications. For example, this paper addresses the problem of missing or error abnormalities of bound books in printing factory. When books are printed, labeled, bound and clipped, they are going into next pipeline to binding covers. Before that, it takes a lot of manual labours as well as material resources that employee pick out the books with error label. Therefore, detecting abnormalities based on image significantly improve productivity. Theoretically, abnormalities detecting is defined as confirming whether specific object appears in predefined area or not. Commonly, methods of abnormalities detecting consist of four steps: preprocessing, window-sliding, feature extraction and postprocessing. While template matching is a fundamental approach to detect objects in particular image area. This paper proposes a contour matching algorithm combining already existing algorithms such as template matching and shape matching to compose the method to implement detecting abnormalities.

2. Related Work

Object detecting and extracting is one of the fundamental tasks in the field of image processing and pattern recognition. Many researchers have proposed various algorithms to address this problem. Suzuki[1] proposed an extended version of the border following algorithm which can extract the topological structure of a given binary image which includes contours that are a useful tool for shape analysis and object detecting and recognition. Carsten Rother[2] proposed an image segmentation method based on graph cuts, called GrabCut. Starting with a user-specified bounding box around the object to be segmented, the algorithm estimates the color distribution of the target object and that of the background using a Gaussian mixture model[3]. This is used to construct a Markov random field[4] over the pixel labels, with an energy function that prefers connected regions having the same

label, and running a graph cut based optimization to infer their values. Eva-Marie Nosal[5] used flood-fill algorithm for passive acoustic detection and tracking. The flood-fill algorithm takes three parameters: a start node, a target color, and a replacement color. The algorithm looks for all nodes in the array that are connected to the start node by a path of the target color and changes them to the replacement color. Luke Cole[6] used Template Matching to address the scaling problem by reduction of the size of the set against which new images are compared. Template matching is a technique in digital image processing for finding small parts of an image which match a template image.

3. Approach

We develop a software named TemplateEditor to draw the bounding rectangle of bound books and the black rectangular label area of each book side and then output a template image as one input of our method. The template image's size approximately equals to the size of bounding rectangle of bound books in the input image. Besides, the template image is transparent except the area which is corresponding to the black approximate rectangular label in the bound books' side. Each template image matches one kind of bound books. Hence, it's necessary to draw a new template image for a kind of bound books unseen before. An example procedure is shown as follows.

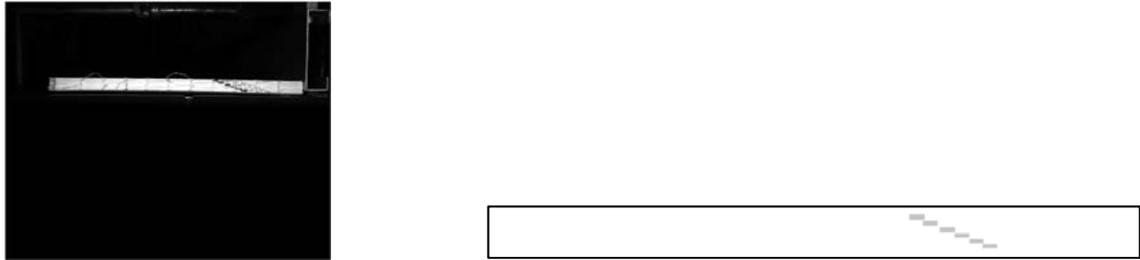


Fig. 1 The input image and the output image (Since the output image is transparent, the black border lines just indicate the image border.)

3.1 Contour Matching Algorithm

In order to extract the rectangular area containing bound books, this paper introduces a difference function which is formularized as:

$$\text{difference} = \frac{\Delta w^2 + \Delta h^2 + \Delta \text{area}^2}{w_t^2 + h_t^2 + \text{area}_t^2} \quad (1)$$

where Δw , Δh , Δarea and area_t are separately computed as:

$$\Delta w = w_i - w_t \quad (2)$$

$$\Delta h = h_i - h_t \quad (3)$$

$$\Delta \text{area} = w_i \times h_i - w_t \times h_t \quad (4)$$

$$\text{area}_t = w_t \times h_t \quad (5)$$

where w_i and h_i stand for the width and height of bounding rectangle of i-th contour, while w_t and h_t represent the width and height of template image which is drawn by hand referring to known correct image.

The difference is smaller, the similarity and the degree of matching is more. We use the contour whose difference between itself and template is the smallest one.

3.2 Shape Matching Algorithm

For distinguishing abnormalities from noise, we use Hu invariants to filter the abnormal area. The formula defines as follows [7].

$$I(A, B) = \max_{i=1...7} \frac{|m_i^A - m_i^B|}{|m_i^A|} \quad (6)$$

where

$$m_i^A = \text{sign}(h_i^A) \cdot \log h_i^A \quad (7)$$

$$m_i^B = \text{sign}(h_i^B) \cdot \log h_i^B \quad (8)$$

and h_i^A, h_i^B are the Hu moments[8] of object A and object B.

3.3 Detect Whether Existing Missing

As beginning, we use Contours Finding algorithm[1] to find contours in the input image, which is followed by selecting the best contour covering the bound books according to Contour Matching. Then, we apply the GrabCut algorithm[2] to exactly segment the bound books regarding the bounding rectangle of the best contour as one input. After that, we transform the last output to binary image where bounding area is black and most central area is white, fill four corners with white employing flood-fill algorithm and inverse the image to make white pixel black and black pixel white. Since the transformation needs a threshold, we introduce a threshold, named binary threshold (BT). Afterwards we apply Template Matching algorithm[6] to discovery the rectangle in the last output which is most similar to template image generated by hand drawing. And Then, we fill the pixels with white along the border of the rectangle on the image before filling four corners. The same to before, we fill four corners with white and inverse the image. Continuously we cut out the rectangular region using the best matched rectangle as the target region. Now, the region equals to the template in size. Hereafter, we traverse every blob in template image, find the corresponding area in the region and count the non-zero pixel in this area. If the number of non-zero pixel is not zero, the area is to be a matched one. Otherwise, the area is to be marked as missing. With the procedure of traversing, we fill all pixels in the matched area with white and employ seed-growing algorithm to fill with black in the inversed and uncut image. This operation is much beneficial to denoising.

3.4 Prepare for Next Detecting

First of all, we initialize a mask image filled with zero and equal to the target region image in width and height. In experiment, we find that scattered salt and pepper noises distribute in the region. Therefore, we apply Median filter algorithm with a 3x3 size window to wipe off such noises. One step closer to denoising, we summarize the maximum width, the maximum height and the minimum area defined as width multiplied by height among the matched areas. Otherwise, in order to decrease the misjudgement, we introduce a coefficient named minimum area coefficient (MAC) used to multiply by the min area. We call the multiplied result minimum non-zero pixels count (MNZPC). And then, we traverse every matched rectangular area, apply Template Matching algorithm to find the best matched area in the target region and count the non-zero pixel in this area. If the number is less than MNZPC, this area is to be noise. Otherwise, we fill all pixels in this area with white and employ seed-growing algorithm. And it will be also treated as noise that either the width or the height of the bounding rectangle of seed-grown area exceeds the maximum width or the maximum height. Therefore, if the area is judged as noise, we fill the area with black to update the target region. If not, the rectangular area in the mask image corresponding to the best matched area is filled with one. At last, we multiply the target region image by the mask image to extract the most probable abnormalities area.

3.5 Detecting Whether Existing Error

In practice, we found that some stubborn noises still lie in the fringe of the output image of the pre processing, which are survived in previous denoising procedures. Thus, for increasing the correctness, we introduce two thresholds, one is called left right ignoring ratio(LRIR), which stands for the percent of unnecessary considering width of the left and right sub area of the output of last section, the another is called top bottom ignoring ratio(TBIR) representing the percent of unnecessary considering height of the top and bottom sub area. As to the output of last section, we loop through every area filled with non-zero pixel, match with matched rectangular areas by means of Shape Matching algorithm[7]. If the area is neither in the LRIR nor in the TBIR. Since the matched result is a format of float value, we introduce another threshold, called max Hu invariants difference ratio (MHIDR). If the float value is no more than the MHIDR, the area is to be an abnormality.

3.6 Make Decision

Until here, we have already detected missing abnormality and error abnormality in source image. Consequently, what we only have to do is combining the outputs of the section 3.1 which indicates whether existing missing and the section 3.3 which indicating whether existing error to make decision. If either missing or error exists, the final judgement is abnormal. Otherwise it is normal.

Table 1. Experiment result

Book name	Image quantity	Correct judgement	Accuracy
1949	17	14	82.35%
Les Misérables	33	33	100%
The View of my Father's Back	9	9	100%
Dawn Blossoms Plucked at Dusk	20	20	100%
Stray Birds	14	14	100%
Воскресение	14	10	71.43%
Vingt mille lieues sous les mers	8	5	62.5%
The Field of Life and Death and Tales of Hulan River	14	14	100%
Silent Spring	20	19	95%
The Shop of Lin Family	23	21	91.30%
Bible	5	5	100%
Once Upon a Time	29	23	79.31%
Summary	206	187	90.78%

4. Experimental Evaluation

We collect experimental images from the pipeline in the factory which are automatically photoed when bound books once reach one predefined position where the shutter will be trigger in a specific angle and illumination condition. We use the accuracy as the performance metric. The ground truth is given by visual inspection. As to the five thresholds, we separately set BT to 80, MHIDR to 0.5, LRIR to 0.15, TBIR to 0.2 and MAC to 0.5. In total, we apply our method in twelve species of bound books, whose quantities vary from five to thirty-three. Accordingly, we draw twelve template images before launching our experiment. The experiment result is shown as Table 1. A sample result is show as Fig.2a, Fig 2b and Fig 2c.

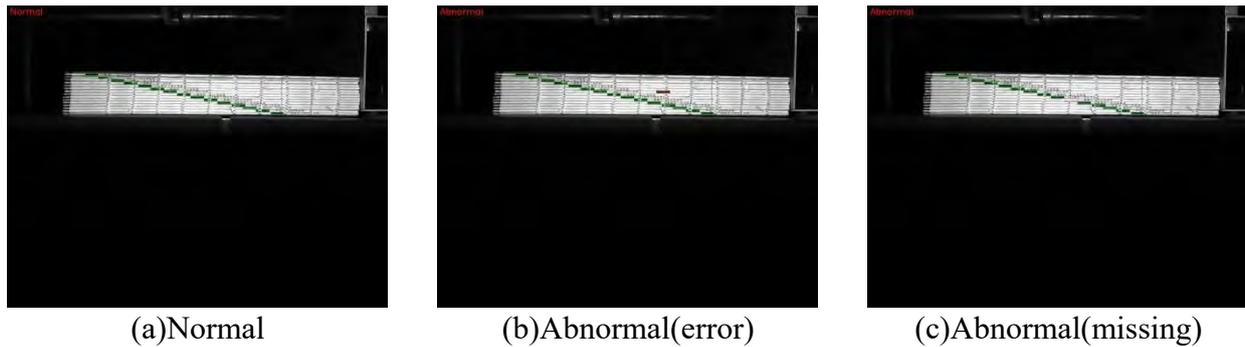


Fig. 2 Sample results. Green rectangles mark the matched area corresponding to template, red rectangles mark the error abnormality, purple rectangles mark the missing abnormality.

5. Conclusion

This paper proposes a practical method based on similar matching including contour matching, template matching and shape matching to address the problem of abnormalities detecting in production environment, which makes full use of the known structural and relative positional information. The experiment shows that this method is better enough to put into practice in the factory.

References

- [1]. Suzuki S. Topological structural analysis of digitized binary images by border following[J]. *Computer vision, graphics, and image processing*, 1985, 30(1): 32-46.
- [2]. Rother C, Kolmogorov V, Blake A. Grabcut: Interactive foreground extraction using iterated graph cuts[C]//*ACM transactions on graphics (TOG)*. ACM, 2004, 23(3): 309-314.
- [3]. Reynolds D. Gaussian mixture models[J]. *Encyclopedia of biometrics*, 2015: 827-832.
- [4]. Chellappa, Rama, and Anil Jain. "Markov random fields. Theory and application." Boston: Academic Press, 1993, edited by Chellappa, Rama; Jain, Anil (1993).
- [5]. Nosal E M. Flood-fill algorithms used for passive acoustic detection and tracking[C]//*New Trends for Environmental Monitoring Using Passive Systems*, 2008. IEEE, 2008: 1-5.
- [6]. Cole L, Austin D, Cole L. Visual object recognition using template matching[C]//*Australian conference on robotics and automation*. 2004.
- [7]. Information on https://docs.opencv.org/2.4/modules/imgproc/doc/structural_analysis_and_shape_descriptors.html#matchshapes.
- [8]. Hu M K. Visual pattern recognition by moment invariants[J]. *IRE transactions on information theory*, 1962, 8(2): 179-187.