LETTER *Special Section on Machine Vision and its Applications*

Thresholding Based on Maximum Weighted Object Correlation for Rail Defect Detection

Qingyong LI[†a)], **Yaping HUANG**[†], **Zhengping LIANG**[††], ***Nonmembers***, ***and*** **Siwei LUO**[†], ***Member***

SUMMARY Automatic thresholding is an important technique for rail defect detection, but traditional methods are not competent enough to fit the characteristics of this application. This paper proposes the Maximum Weighted Object Correlation (MWOC) thresholding method, fitting the features that rail images are unimodal and defect proportion is small. MWOC selects a threshold by optimizing the product of object correlation and the weight term that expresses the proportion of thresholded defects. Our experimental results demonstrate that MWOC achieves misclassification error of 0.85%, and outperforms the other well-established thresholding methods, including Otsu, maximum correlation thresholding, maximum entropy thresholding and valley-emphasis method, for the application of rail defect detection.

key words: *automatic thresholding, defect detection, correlation*

1. Introduction

Defect inspection for rail head surface is one of the most important tasks for railway maintenance. To efficiently detect such defects, inspection systems based on computer vision have been attracting more and more attention [1], [2]. These systems capture rail surface images with a high-speed line-scan digital camera and automatically inspect the images using a customized image processing software. For the captured images, intensity is the most distinguishing feature that can be relied on to detect defects, and thresholding techniques are potential tools for segmenting defects from background.

Thresholding algorithms automatically select an optimal gray-level value according to certain criterion to separate objects of concerned from background in an image. It is a fundamental technique for image segmentation, and many methods have been proposed in literature [3]. Many universal thresholding methods can be applied in defect detection. Nacereddine et al. [4] compared four traditional thresholding methods for weld defect detection, including Otsu method [5], Maximum Entropy thresholding (ME) approach [6], minimum error thresholding [7] and moment preserving method [8], and found that ME approach was the best. Yen et al. [9] put forward the Maximum Correlation (MC) criterion to substitute for maximum entropy. MC thresholding can achieve comparable performance to ME method, but it is faster than the latter.

Most universal thresholding approaches, however, are not perfect for defect detection, and some revisions were made to fit the characteristics of this application. Ng [10] put forward the Valley-Emphasis Method (VEM). VEM selects a threshold that has a small probability of occurrence, and meanwhile maximizes the between-class variance as in Otsu [5] method. VEM is verified to be an effective revision of Otsu method for applications of defect detection.

Though there are many thresholding methodologies can be adopted for rail defect detection, they face challenges for this application because of the following factors.

1. The rail images are unimodal distribution as shown in Fig. 1 (a). As we known, Otsu and many other methods are weak for unimodal images [3].
2. The proportion of defects is always small for a railway in service, even zero for most rail images. VEM [10] restricts a threshold to locate at a valley with a low occurrence frequency, but it cannot guarantee the defect proportion to be small.

Taking these factors into account, this paper puts forward the Maximum Weighted Object Correlation (MWOC) thresholding method. The basic idea of MWOC is to maximize the product of object correlation and exponen-

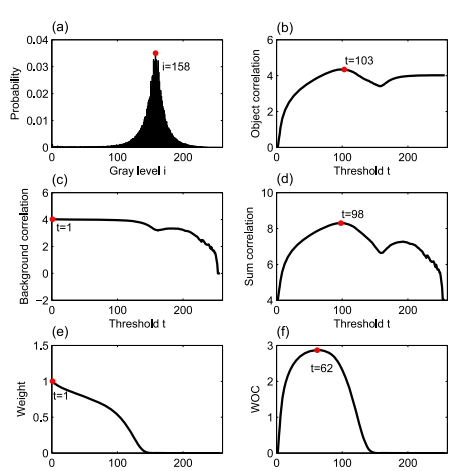

Fig. 1 (a) The PMF of the normalized image in the middle of Fig. 2. (b) The object correlation $C_O(t)$. (c) The background correlation $C_B(t)$. (d) The sum of object and background correlation. (e) The weight curve $W(t)$. (f) The weighted object correlation $WOC(t)$. The red points denote the gray level of the curve peak.

Manuscript received October 27, 2011.
Manuscript revised January 9, 2012.

[†]The authors are with the School of Computer and Information Technology, Beijing Jiaotong University, Beijing, 100044, China.
[††]The author is with the College of Computer Science and Software Engineering, Shenzhen University, Guangdong, 518060, China.
a) E-mail: liqy@bjtu.edu.cn

DOI: 10.1587/transinf.E95.D.1819

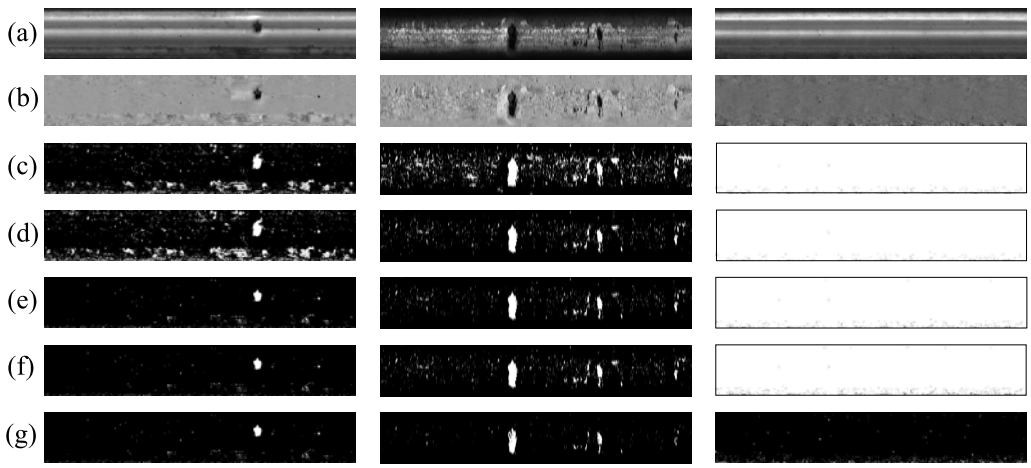

Fig. 2 Examples of thresholding result with the five methods. (a) Original rail images. (b) Normalized images. (c) Otsu. (d) VEM. (e) MC. (f) ME. (g) MWOC. For the left image, the thresholds of Otsu, VEM, MC, ME and MWOC are 134, 140, 105, 99 and 98 respectively; they are 128, 104, 97, 102 and 61 for the middle image; and they are 155, 158, 151, 146 and 61 for the right image.

tially tuned proportion of thresholded background. MWOC can achieve a threshold that maximizes the correlation and meanwhile keeps the proportion of defects in a low level. Our experimental results validate the advantages of MWOC for rail defect detection, compared with other well-established thresholding methods.

2. The MWOC Thresholding Method

Consider a 256 gray-level image and its histogram or Probability Mass Function (PMF) $P_G \equiv \{p_i | i \in [0, 255]\}$, where p_i denotes the normalized frequency of gray level i. For a given gray level t, its cumulative probability is given by

$$P(t) = \sum_{i=0}^{t-1} p_i. \tag{1}$$

If $P(t)$ is greater than zero and less than one, t refers to a threshold dividing pixels into two classes $O = \{0, 1, \cdots, t-1\}$ and $B = \{t, t+1, \cdots, 255\}$. O and B often denote objects of interested and background, respectively. The PMFs of O and B can be derived from P_G as follows:

$$P_O(t) \equiv \left\{ \frac{p_0}{P(t)}, \frac{p_1}{P(t)}, \cdots, \frac{p_{t-1}}{P(t)} \right\},$$
$$P_B(t) \equiv \left\{ \frac{p_t}{1-P(t)}, \frac{p_{t+1}}{1-P(t)}, \cdots, \frac{p_{255}}{1-P(t)} \right\}.$$

The correlations of P_O and P_B are defined as

$$C_O(t) = -\ln \sum_{i=0}^{t-1} \left(\frac{p_i}{P(t)} \right)^2, \tag{2}$$

$$C_B(t) = -\ln \sum_{i=t}^{255} \left(\frac{p_i}{1-P(t)} \right)^2. \tag{3}$$

According to MC criterion [9], an optimal threshold t^* should maximize the sum correlation of C_O and C_B, and is computed by

$$t^* = \arg \max_{t \in [0,255]} (C_O(t) + C_B(t)). \tag{4}$$

As discussed in Sect. 1, rail images have two characteristics: 1) unimodal histogram; 2) small defect proportion. Fitting these characteristics, MWOC obtains a threshold that maximizes the object correlation $C_O(t)$ and keeps the defect proportion in a low level meanwhile.

On one hand, the unimodal P_G greatly impacts the feature of $P_O(t)$ and $C_O(t)$. We obtain the following observation from our dataset.

- $P_O(t)$ is likely to approximate a uniform distribution for small t, since most pixels with gray level $i < t$ are noise or defects. It derives small $C_O(t)$.
- $P_O(t)$ becomes a non-Gaussian distribution with a high tail for t near the peak of P_G, and generates large $C_O(t)$.
- $P_O(t)$ approximates a normal distribution for large t and gets a little changed $C_O(t)$.

Figure 1 (b) shows the curve of $C_O(t)$ for the image in the middle of Fig. 2. We can observe that the value of $C_O(t)$ increases with the rise of t for $t < 103$, whereas it become somewhat flat for $t > 103$. On the contrary, $C_B(t)$ is flat for low t and decreases with large t, as illustrated in Fig. 1 (c). As a result, the change of the sum correlation in (4) is dominated by that of $C_O(t)$ when t falls in a low range. Figure 1 (d) illustrates the curve of the sum correlation. So we take $C_O(t)$ as the criterion instead of the sum. This substitution has two advantages: 1) $C_O(t)$ restricts the threshold in a low gray-level range implicitly, because its maximum often locates at the left side of a histogram peak; 2) it is simpler and need less computation than the sum correlation.

On the other hand, a weight is assigned to $C_O(t)$ to keep the thresholded defect proportion be small. The *defect proportion* equals to the cumulative probability $P(t)$ in (1) since

pixels with gray level $i < t$ are divided into defects. Generally, a preferred threshold t derives a low $P(t)$, or equivalently a large $(1 - P(t))$ that denotes the *background proportion*. In addition, we add an exponential parameter α to the background proportion in order to tune its significance. So the weight function $W(t)$ is defined as

$$W(t) = (1 - P(t))^{\alpha}. \tag{5}$$

Figure 1 (e) illustrates the curve of $W(t)$ with $\alpha = 15$. $W(t)$ is a monotonically decreasing function. Especially, the decreasing slope is high when t takes a low value, though it is very flat for great t. Note that this slope can be tuned by α and more discussion about α is seen in Sect. 3.

Multiplying these two terms, the Weighted Object Correlation (WOC) is defined as

$$WOC(t) = C_O(t)(1 - P(t))^{\alpha}. \tag{6}$$

Further, the formulation for the MWOC thresholding is given by

$$t^* = \arg \max_{t \in [0,255]} (WOC(t)). \tag{7}$$

Figure 1 (f) presents the curve of WOC, which shows notably different shape compared with Fig. 1 (b) and Fig. 1 (d). The obtained threshold is lower than those of C_O and MC thresholding, and achieves better segmentation results as shown in the middle of Fig. 2.

3. Experiments and Analysis

In the experiments, we first compare the proposed MWOC with Otsu [5], ME [6], MC [9] and VEM [10], and then analyze the influence of the parameter α.

The dataset comprises 100 rail images that are captured by a Dalsa Spyder-2 line-scan camera in actual railways. There are 80 images with defects and 20 images without defect. The images with defects were manually labeled and converted to binary images as ground-truth. Note that all images are cropped with size 140×1120 and preprocessed by a local normalization procedure [11] as following:

$$L_{(x,y)} = \frac{F_{(x,y)} - \mu_y}{\sigma_y}, \tag{8}$$

where $L_{(x,y)}$ and $F_{(x,y)}$ are the normalized value and gray-level value at position (x, y) respectively, and μ_y and σ_y denote the mean and standard variance of gray values in the line $Y = y$.

We evaluate the thresholding performance with the widely used misclassification error (Err) [10], [12], which is defined as

$$Err = 1 - \frac{|B_G \cap B_T| + |F_G \cap F_T|}{|B_G| + |F_G|},$$

where B_G and F_G denote the background and foreground of the ground-truth image, B_T and F_T denote the background and foreground in the thresholded image, and $|\cdot|$ is the cardinality of the set.

3.1 Comparison of Performance

Figure 2 shows three examples of thresholding result for the five methods. MWOC achieves the smallest thresholds, which bring out the best segmentation results. Especially, all methods except MWOC get too high thresholds and wrongly separate most of background into defect for the right image in Fig. 2. We have to note that most rail images are defect-free for a rail system in service and WMOC can avoid wrong alarm compared with MC and ME from a viewpoint of practical use. VEM gets larger thresholds than Otsu for the left and right images. This is because the weight term $(1 - p(t))$ of VEM is based on occurrence probability $p(t)$. This weight just makes threshold t prefer to a valley value in a histogram, but the defect proportion may be large. On the contrary, the weight $(1 - P(t))$ of MWOC is a cumulative probability and it is able to restrict the defect proprotion in a low level.

We further compare the thresholding performance of these methods by misclassification error. Table 1 presents the average misclassification error for the five methods on the dataset. From this table, we can observe that MWOC overwhelms the other methods and gets a very small misclassification error. VEM is better than ME, MC and Otsu, because it benefits from the weight $(1 - p(t))$ for most images whereas it may get worser thresholds occasionally. ME and MC are better than Otsu, since they are more suitable to unimodal images than Otsu as discussed in [3].

3.2 Influence of Parameter α

The parameter α of the weight function in (5) is an important factor impacting the performance of MWOC, because it tunes the significance of the background proportion in (6). Figure 3 (a) illustrates the curve of average Err on the im-

Table 1 Average misclassification error for the five methods.

	Otsu	VEM	MC	ME	MWOC
Err	0.1361	0.0502	0.0538	0.0513	0.0085

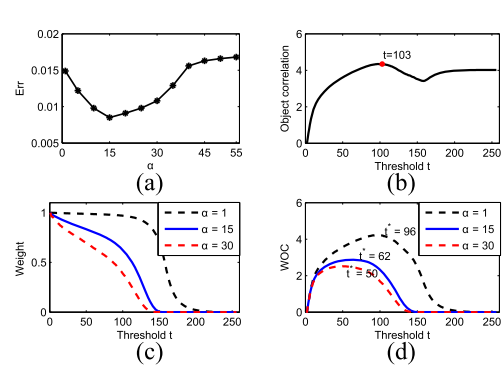

Fig. 3 (a) The curve of misclassification error for parameter α. (b) Object correlation function $C_O(t)$. (c) Weight function $W(t)$. (d) Weighted object correlation function $WOC(t)$.

age set for different α. From this figure, we can find that Err decreases with the rise of α for $\alpha < 15$, and then it increases for somewhat greater α. Lastly, it gets close to the constant that is the average defect proportion of the dataset. This result can be explained from two aspects.

First, the object correlation is regarded as a criterion, and background proportion is a dynamic weight, which is tuned by α. A large α generates a $W(t)$ that decreases very fast in the low range of t. Such deep $W(t)$ pushes the peak of $WOC(t)$ towards the low range of gray level, and a small t^* is obtained. A small α derives a $W(t)$ that changes little in the low range of t, and the $WOC(t)$ is dominated by C_O. So a large t^* is gotten. An appropriate α can bring out a good trade-off between background proportion and object correlation, and achieve the best threshold. Figure 3 (c) and Fig. 3 (d) show the curves of $W(t)$ and $WOC(t)$ with different parameter α for the middle image in Fig. 2.

Second, taking the logarithm of (6), we can get

$$LWOC(t) = \ln C_O + \alpha \ln(1 - P(t)). \tag{9}$$

There are two terms: $\ln C_O$ and $\ln(1 - P(t))$. The former expresses the object correlation, and the latter represents the background (or defect) proportion. The weight α tunes the significance of the latter. There should be a trade-off between these two terms, for example, $\alpha = 15$ in our experiments.

4. Conclusion

This paper has put forward the maximum weighted object correlation thresholding method, fitting the characteristics of rail defect detection, such as unimodal histogram and small defect proportion. MWOC selects a threshold that maximizes the thresholded object correlation and also keeps the defect proportion in a low level. Our experimental results validate that WMOC is an attractive thresholding method for rail defect detection with misclassification error of 0.85%, and it outperforms Otsu [5], maximum entropy [6], maximum correlation [9] and valley-emphasis method [10].

Acknowledgements

The authors acknowledge the support from the R&D foundation of Shenzhen (JC201005280432A) and Fundamental Research Funds for the Central Universities (2011JBZ005). They also thank Yongliang Wang, Jie Lin, Guohao Lv, Shengwei Ren and Qiang Han for their work on collecting and labeling rail images.

References

[1] F. Marino, A. Distante, P. Mazzeo, and E. Stella, "A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection," IEEE Trans. Syst. Man. Cybern. C, Appl. Rev., vol.37, no.3, pp.418–428, 2007.

[2] M. Papaelias, C. Roberts, and C. Davis, "A review on non-destructive evaluation of rails: state-of-the-art and future development," Proc. Institution of Mechanical Engineers, Part F: J. Rail and Rapid Transit, vol.222, no.4, pp.367–384, 2008.

[3] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," J. Electronic Imaging, vol.13, no.1, pp.146–156, 2004.

[4] N. Nacereddine, M. Hamami, and N. Oucief, "Non-parametric histogram-based thresholding methods for weld defect detection in radiography," World Academy of Science, Engineering and Technology, vol.9, pp.213–217, 2005.

[5] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst. Man. Cybern., vol.9, no.1, pp.62–66, 1979.

[6] J. Kapur, P. Sahoo, and A. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," Comput. Vis. Graph. Image Process., vol.29, no.3, pp.273–285, 1985.

[7] J. Kittler and J. Illingworth, "Minimum error thresholding," Pattern Recognit., vol.19, no.1, pp.41–47, 1986.

[8] W. Tsai, "Moment-preserving thresholding: A new approach," Comput. Vis. Graph. Image Process., vol.29, no.3, pp.377–393, 1985.

[9] J. Yen, F. Chang, and S. Chang, "A new criterion for automatic multilevel thresholding," IEEE Trans. Image Process., vol.4, no.3, pp.370–378, 1995.

[10] H. Ng, "Automatic thresholding for defect detection," Pattern Recognit. Lett., vol.27, no.14, pp.1644–1649, 2006.

[11] X. Xie and K. Lam, "An efficient illumination normalization method for face recognition," Pattern Recognit. Lett., vol.27, no.6, pp.609–617, 2006.

[12] W. Yasnoff, J. Mui, and J. Bacus, "Error measures for scene segmentation," Pattern Recognit., vol.9, no.4, pp.217–231, 1977.