

Detection algorithm of rail surface defects based on multifeature saliency fusion method

Hua Zhai

Anhui Province Key Lab of Aerospace Structural Parts Forming Technology and Equipment, Hefei University of Technology, Hefei, China and Institute of Industry and Equipment Technology, Hefei University of Technology, Hefei, China, and

Zheng Ma

Anhui Province Key Lab of Aerospace Structural Parts Forming Technology and Equipment, Hefei University of Technology, Hefei, China

Abstract

Purpose – Effective rail surface defects detection method is the basic guarantee to manufacture high-quality rail. However, the existed visual inspection methods have disadvantages such as poor ability to locate the rail surface region and high sensitivity to uneven reflection. This study aims to propose a bionic rail surface defect detection method to obtain the high detection accuracy of rail surface defects under uneven reflection environments.

Design/methodology/approach – Through this bionic rail surface defect detection algorithm, the positioning and correction of the rail surface region can be computed from maximum run-length smearing (MRLS) and background difference. A saliency image can be generated to simulate the human visual system through some features including local grayscale, local contrast and edge corner effect. Finally, the meanshift algorithm and adaptive threshold are developed to cluster and segment the saliency image.

Findings – On the constructed rail defect data set, the bionic rail surface defect detection algorithm shows good recognition ability on the surface defects of the rail. Pixel- and defect-level index in the experimental results demonstrate that the detection algorithm is better than three advanced rail defect detection algorithms and five saliency models.

Originality/value – The bionic rail surface defect detection algorithm in the production process is proposed. Particularly, a method based on MRLS is introduced to extract the rail surface region and a multifeature saliency fusion model is presented to identify rail surface defects.

Keywords Machine vision, Defect detection, Maximum run-length smoothing, Multifeature saliency fusion, Rail inspection

Paper type Research paper

1. Introduction

Defects on rail surface are the major factors that affect the manufacturing quality. Complex local grinding system is developed to remove defects after scanning and distinguishing products surface in the rail factory. Therefore, surface inspection system plays a pivotal role in the rail manufacturing process. Manual detection is inefficient, less sensitive and not suitable for harsh production environments. Different from the existing online manual rail surface defect inspection system (Yu *et al.*, 2019; He *et al.*, 2016), the manufacturing process automatic nondestructive inspection system of rail products faces the following problems: the harsh working environment, the unsteady position of the rolling rail on the production line, the random generation and distribution of defects and the vibration rail transmission.

Nondestructive testing techniques such as ultrasonic (Cruz *et al.*, 2017), eddy current (Zhu *et al.*, 2018) and computer vision (Zhang *et al.*, 2021; He *et al.*, 2016) have been

developed. Ultrasonic testing cannot detect surface defects and eddy current testing is sensitive to the vibration detection environment (Zhang *et al.*, 2021). In contrast, computer vision inspection is more suitable for surface defect inspection because of merits such as fast detection speed, high detection precision and easy installation.

© Hua Zhai and Zheng Ma. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial & non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Anhui Province Key Project.

201903c08020008.

National Natural Science Foundation of China.

61701162.

Science and Technology guidance fund of Bengbu.

Development of grinding robot for special steel ba.

Received 15 October 2021

Revised 1 February 2022

15 March 2022

Accepted 18 April 2022

The current issue and full text archive of this journal is available on Emerald Insight at: <https://www.emerald.com/insight/0260-2288.htm>



Sensor Review
42/4 (2022) 402–411
Emerald Publishing Limited [ISSN 0260-2288]
[DOI 10.1108/SR-10-2021-0363]

More attention to the automatic nondestructive inspection algorithms for rail manufacturing quality is paid based on computer vision inspection in recent years. Current defect detection algorithms can be divided into two categories: prior knowledge and deep learning.

Some detection algorithms of specific image information can be guided to locate defects based on prior knowledge. Yu *et al.* (2019) presented row consistency, phase-only Fourier transforms and pixel consistency to extract features of different scales to identify defects. Gan *et al.* (2020) proposed a background-oriented defect inspection to improve defect detection by considering specified characteristics of the rail during the inspection. However, some special constraints on the number and uniformity of images in the data set must be satisfied. Further concern for image enhancement strategies was obliged to reinforce the distinction between defects and backgrounds (Li and Ren, 2012; He *et al.*, 2016). Li and Ren (2012) introduced the local normalization method for image enhancement and proposed a projection profile algorithm to identify possible defects, although the prediction result is only a bounding box. In He *et al.*'s (2016) work, the Perona–Malik diffusion model was presented for defect boundary enhancement and noise suppression and a nearest-neighbor difference scheme was designed to select proper defect boundaries. Morphological operations are often used in rail inspection systems (Nieniewski, 2020; Min *et al.*, 2018). Nieniewski (2020) presented a rail defect detection system and shape extraction method using morphological pyramid. By considering the geometric features of defects, morphological processing was applied to remove the interference of redundant information, and the direction chain code was used to identify the defect shape feature (Min *et al.*, 2018). The method in Niu *et al.*'s (2021) work is proposed based on the bi-level super pixel-based framework and bag-of-words feature extractor. Ni *et al.* (2021) discussed a novel defect detection algorithm based on a partitioned edge feature. In Zhang *et al.*'s (2018) work, a curvature filter was embedded to retain relevant details and eliminate noise. Furthermore, an improved fast and robust Gaussian mixture model based on Markov random field was established for surface defect segmentation. Although it has high accuracy and strong robustness, the computation complexity is high and not conducive for implementation.

Although prior knowledge-based detection algorithms are easy to implement without the training process, most of them are proposed for rail surface defect online detection and verified using the publicly available rail surface discrete defect (RSDD) data set. The difference between railway images will limit the application scenarios of the algorithm, which makes the prior knowledge-based detection algorithms not suitable for rail manufacturing process.

Recently, some deep learning methods have been proposed for rail surface defects detection. Deep learning detection algorithm has an advantage to automatically extract features based on training samples. A deep neural network (Faghih-Roohi *et al.*, 2016) is applied for defect detection and classification using convolution layers to extract suitable features. In Jin *et al.*'s (2020) work, a multimodel rail inspection system is established for surface defect where fast and robust spatially constrained Gaussian mixture model (FRGMM) is presented for segmentation proposal and Faster RCNN is used

for objective location in a parallel structure. Zhang *et al.* (2021) proposed a limited sample rail surface defects detection scheme using line-level label. This scheme regarded defect images as sequence data and classified pixel lines with neural networks to solve the problem of a small number of defect images, but only the y -direction position of the defect can be computed. Liang *et al.* (2018) proposed a deep convolution neural network of the SegNet architecture to detect the surface defects of rails. Yuan *et al.* (2019) designed a novel network to classify and locate defects based on MobileNetV2 and ensures real time by multiscale defect detection.

The deep learning-based algorithms have excellent fitting capabilities with large number of training samples. There are two difficulties in sample collection; first, it is difficult to collect samples due to the small number of defects and the harsh environment, which takes a lot of manpower and time, and second, pixel-level labeling requires specialized knowledge of the rail manufacturing process and certain computer skills.

The human visual nervous system can quickly find interesting objects in complex scenes. This selective visual ability is called the visual attention mechanism. The saliency algorithm is an important solution path to simulate the human visual mechanism to find prominent objects. Itti *et al.* (1998) proposed a visual saliency model based on the cognitive psychology theory of Koch and Ullman (1987). Liu *et al.* (2011) formally aimed at the saliency object detection task and introduced the visual attention mechanism into the object segmentation task.

Subsequently, a large number of saliency algorithms such as AC (Achanta *et al.*, 2008), frequency-tuned (FT) (Achanta *et al.*, 2009), LC (Zhai and Shah, 2006) and HC (Cheng *et al.*, 2014) were proposed. Because of good anti-interference ability in the field of defect identification, the saliency algorithm has been widely used in various surface defect detection systems. For example, FT (Achanta *et al.*, 2009) and ITTI (Itti *et al.*, 1998) models can be used for steel strip surface inspection (Song and Yan, 2013; Guan, 2015; Song *et al.*, 2014) and AC (Achanta *et al.*, 2008) models for welding inspection (Ben Gharsallah and Ben Braiek, 2015). In the field of rail defect detection, Hu *et al.* (2018) created the block local contrast measure (BLCM) saliency model to detect the peeling of the rail surface. To improve the accuracy and efficiency of defect identification in the rail manufacturing process, a bionic rail surface defect detection algorithm based on the multifeature saliency fusion is investigated in this paper.

The remaining of this paper is organized as follows. Section 2 gives details of the proposed model. Section 3 evaluates the performance of the method through experiment and comparison. Section 4 provides the conclusions.

2. Proposed method

2.1 Overview

The rail blank will be rough ground after roll forming. This process will be performed to achieve high precision surface to remove some surface discrete defects such as rolling scar, scratches and oxide skin.

During the vibrating movement of the rail on the production line, the position of the camera relative to the rail cannot be guaranteed to be constant; three categories can be divided in

those sampled rail images: the rail is not in the center of the image, the rail is inclined and the camera is not perpendicular to the rail, as shown in Figure 1. Because the curvature shape of the rail surface does not reflect light uniformly, the rail surface region has alternating light and dark stripes. Various defects are randomly formed on the rail surface during the production process. The rail surface image has the following characteristics:

- The rail position is not fixed.
- The rail surface partially has oversaturated regions.
- The defect-free area of the rail surface has a complex texture.

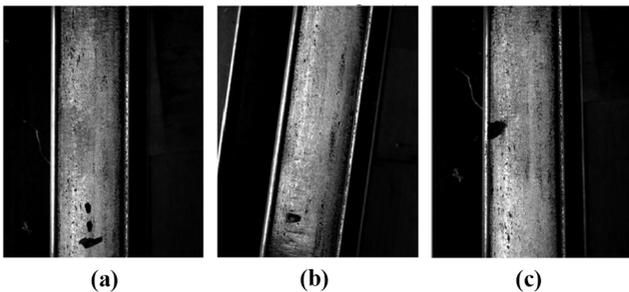
Traditional machine image processing methods based on global brightness or edge detection cannot accurately extract surface defects. When the human vision system is dealing with complex information, it will quickly focus on the important areas with larger saliency value and allocate limited neural computing resources to the key parts of the scene. Saliency detection is to simulate the human visual attention mechanism through the establishment of a suitable calculation model to get the results consistent with human visual cognition. The defect area is only a small part of the rail surface and the saliency value of the discrete defects of the rail is higher. The human eye can quickly identify the defect area, so the saliency detection method can be implemented to detect the surface defect of the rail. For the special situation of the rail surface, three image saliency features are proposed to identify rail surface defects:

- Local grayscale feature, the pixel value of the defect is lower than surroundings.
- Local contrast feature, the defect color is distinguishable from the surroundings.
- Edge corner effect, the color changes of defect edges and corners are very dramatic.

Huang *et al.* (2020) introduced the above three features and background texture feature into a neural network to detect the surface defects of the magnetic tile.

In this paper, the maximum run-length smearing (MRLS) algorithm is improved to locate the rail surface region, and a preprocessing algorithm is proposed to correct uneven reflection of rail images based on background difference. The rationality of applying above features to rail surface defects is discussed, and the corresponding saliency images are generated with three saliency features. Defects are identified by features fusion, meanshift clustering and adaptive threshold segmentation. The pipeline image of the algorithm is shown in Figure 2.

Figure 1 Examples of rail surface images



Notes: (a) Rail is not in the center of image; (b) rail is inclined; (c) camera is not perpendicular to rail

2.2 Target region location based on maximum run-length smearing

Rail images obtained by cameras usually contain irrelevant surrounding region which should be eliminated to avoid interference. Therefore, the rail surface region in the image should be located before defect detection. The MRLS algorithm is improved to achieve rail surface region precise position in the image from the production line.

The run-length is a continuous area with the same pixel value in the same row (or column) in the image. As shown in Figure 3, run-length smearing is defined as the black run-length between two white pixels is replaced with a white run-length when its length is less than a threshold len .

The flowchart of rail surface region location algorithm based on MRLS is shown in Figure 4. As shown in Figure 4(a), the ideal value of each row of rail surface region is a continuous white run-length when the original image is binarized. But white run-length is often interrupted by noise, defects and uneven reflection, forming a black run-length as shown in Figure 4(b). Each row in the image is scanned and replaced, completing the run-length smearing.

The white run-length of the rail surface region can be connected as a whole by run-length smearing while not including irrelevant areas. But the black run-length between the rail surface and the rail bottom is easily smeared to the rail surface region, which will cause misidentification. This black run-length is much longer than the black run-length in rail

Figure 2 Pipeline image of the algorithm

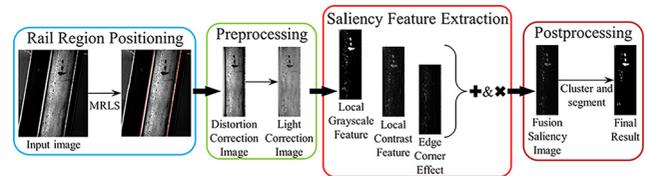
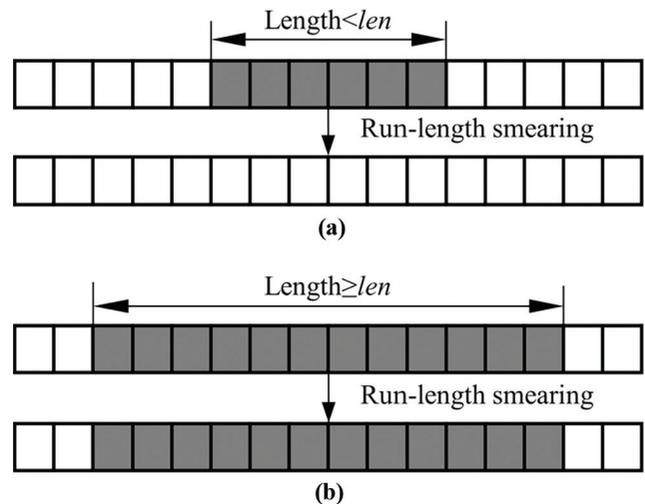
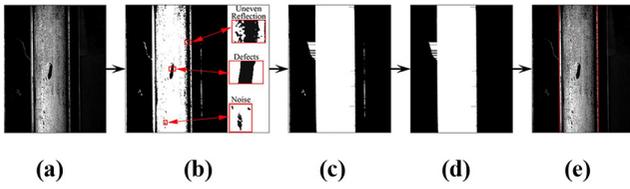


Figure 3 Principle diagram of maximum run-length smearing



Notes: (a) Short run-length is removed; (b) long run-length is reserved

Figure 4 Flowchart of rail surface region location algorithm based on MRLS



Notes: (a) Original image; (b) binary image; (c) run-length smearing image; (d) NMS image; (e) Hough transform detection image

surface region, so the value of the threshold *len* is adjustable with the change of environment. In the [Figure 4\(c\)](#), *len* is set to 0.8 times the distance from the rail surface to the bottom of the rail. It can be seen that the rail bottom in the irrelevant area is identified, but its area is smaller than the rail surface region. So nonmaximum suppression (NMS) is performed on the image, and only the white run-length with the longest length of each row is retained. The results obtained after NMS are shown in [Figure 4\(d\)](#). The boundaries of rail surface are clearly visible, which could be recognized via Sobel edge detection and Hough transformation. Note that to remove lateral edge interference, we only use a single direction Sobel detection operator.

2.3 Preprocessing method of rail surface image

2.3.1 Distortion correction of rail surface image

According to the characteristics of approximately flat rail surface, the image is corrected by perspective transformation. The perspective transformation maps the points in the coordinate system to a new coordinate system. The transformation function is defined by:

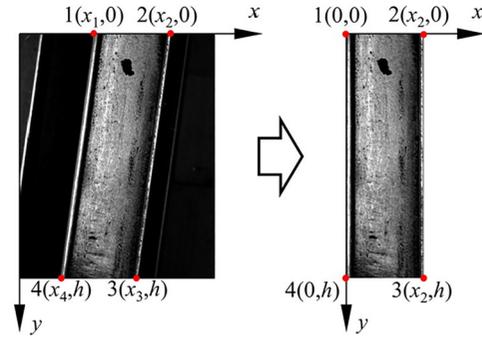
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \quad (1)$$

where *x* and *y* are the pixel coordinates before transformation and *u*, *v* and *w* are the corresponding coordinates after transformation. Perspective transformation matrix has just eight degrees of freedom, so let $a_{33} = 1$. The remaining eight parameters in the transformation matrix, $a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, a_{31}$ and a_{32} , can be solved according to the four vertices corresponding to the image before and after the transformation. The vertices of the quadrilateral detected by MRLS are taken as the initial coordinates. The vertices of the target rectangle are taken as target coordinates. The correction result is shown in [Figure 5](#).

2.3.2 Uneven reflection correction of rail surface image

The recognition result of the algorithm will be affected by uneven reflection, so the illumination correction is performed on the rail surface image. Background difference method is a widely used approach for detecting moving objects from static cameras. The rationale in the approach is that of detecting the moving objects from the difference between the current frame and a reference frame, often called the “background image” ([Piccardi, 2004](#)). Background image is obtained by learning

Figure 5 Distortion correction of rail surface image



temporal sequence of the frames. To highlight defects and reduce the influence of uneven light reflection, the background image is subtracted from the rail surface image to obtain a corrected image. Only one image can be taken for each area of the rail surface during the movement of the rail, so the background difference method cannot be directly used for rail background modeling. The rail surface image has the feature of small pixel value changes along the rail direction, which can be used to construct a background image. The direction perpendicular to the rail is defined as the *x*-axis, and the direction along the rail is defined as the *y*-axis. The background image is defined as:

$$I_m(x, y) = \sum_{y=1}^h \frac{I(x, y)}{h}, \quad x \in [1, w], \quad (2)$$

where $I_m(x, y)$ is the background image, $I(x, y)$ the input image, *h* is the height of the image and *w* is the width. The difference correction image is generated by subtracting the background image from the rail surface image. The difference correction image is defined as:

$$\Delta I(x, y) = I(x, y) - I_m(x, y), \quad (3)$$

where $\Delta I(x, y)$ is the differential correction image. The results in [Figure 6\(c\)](#) show that the influence of uneven reflection is corrected to a certain extent while the defects are retained.

2.4 Multifeature saliency fusion defect detection method

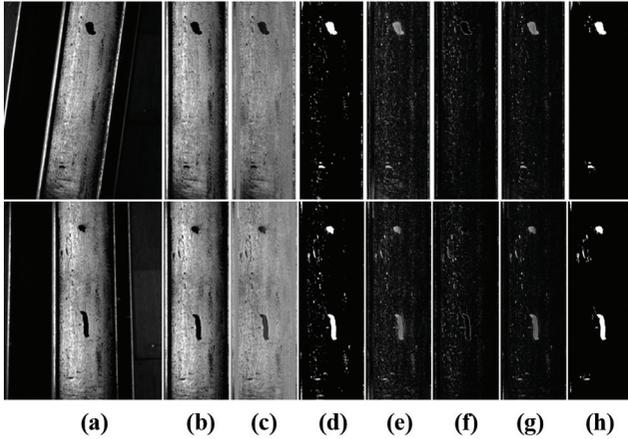
2.4.1 Local grayscale feature

The human visual system is very sensitive to pixels with outliers. The occurrence of defects will change the geometry of the rail surface, resulting in changes in the diffused reflection of light. Therefore, the pixel value of defect pixels is usually less than that of nondefect pixels. According to some qualitative differences between the defect characteristics and its background, the adaptive binary method can be used to identify the defect area, which is defined as:

$$S_D(x, y) = \begin{cases} 1, & I_R(x, y) - \Delta I(x, y) > t \\ 0, & I_R(x, y) - \Delta I(x, y) \leq t \end{cases} \quad (4)$$

where $I_R(x, y)$ is a mean filter blurred image of the difference correction image ΔI in a $R \times R$ blur window. The pixel value range of the corrected image is 0–255, so we define *t* as a

Figure 6 Examples of preprocessed images and saliency images



Notes: (a) Original images; (b) perspective transformation images; (c) light correction images; (d) local grayscale feature saliency images; (e) local contrast feature saliency images; (f) edge corner effect saliency images; (g) feature fusion saliency image; (h) final result images

constant value in the range of 0–255. It can be deduced from equation (4) that when the pixel value is significantly different from its surrounding pixels, effective defect identification can be achieved. The local grayscale feature saliency image is shown in Figure 6(d).

2.4.2 Local contrast feature

The color of defect is distinguishable from the surroundings in a local area, so the AC (Achanta *et al.*, 2008) saliency model is selected to locate the defect location. This algorithm defines saliency as the local contrast of a pixel relative to its surrounding regions on different scales. The saliency value is obtained by calculating the Euclidean distance between the central pixel and the mean of the surrounding regions. The calculation function is defined as:

$$S_{AC}(x, y) = \frac{1}{3} \left[\frac{F_W(x, y)}{2} + \frac{F_W(x, y)}{4} + \frac{F_W(x, y)}{8} \right],$$

$$F_t(x, y) = d[\mathbf{c}(x, y), \mathbf{s}_t(x, y)] \quad (5)$$

where $F_t(x, y)$ is defined as a Euclidean distance d between the *Lab* pixel vector $\mathbf{c}(x, y)$ (center) and the average *Lab* pixel vector $\mathbf{s}_t(x, y)$ in window t (surround). The square surround region is varied as $t = \{W/2, W/4, W/8\}$. W is the smaller pixel size of the two dimensions of the image. The local contrast feature saliency image is shown in Figure 6(e).

2.4.3 Edge corner effect

The above two features can effectively identify the defect area, but the accuracy of the defect edge recognition is not exact because the defects edge area is close to the background. Rail surface defects and the nondefect region form a strong edge corner response due to the drastic change in reflected light. The human visual system is easily attracted by the more differentiated edges. Therefore, a Strukturtensor (ST) can be

introduced to improve the accuracy of defect edge recognition. The ST of the image is defined as:

$$M = \begin{bmatrix} A & B \\ B & C \end{bmatrix} = \begin{bmatrix} I_X I_X & I_X I_Y \\ I_X I_Y & I_Y I_Y \end{bmatrix}, \quad (6)$$

where I_X and I_Y are gradient image of the two dimensions of original image, respectively. Let λ_1 and λ_2 be two eigenvalues of M , respectively. Harris (Harris and Stephens, 1988) proved that, edge responses will occur when one eigenvalue is large while the other one is small, and corner responses happen if and only if both eigenvalues are large. Using λ_1 and λ_2 directly as the edge corner effect will greatly increase the computational difficulty, so let $D = (\lambda_1 - \lambda_2)^2$ be used to represent the edge information and let $E = |\lambda_1 + \lambda_2|$ be used to represent the corner information. Then with equation (6), we get:

$$D = (A - C)^2 + B^2$$

$$E = A + C \quad (7)$$

The edge corner effect S_{ST} is defined as:

$$S_{ST} = D + E \quad (8)$$

When a pixel is in the edge region with a large pixel value change, it will be given a higher saliency value. The edge corner effect image is shown in Figure 6(f).

2.4.4 Feature fusion algorithm for rail surface defects images

Image pixel-wise addition can complement the shortcomings between different feature images. Image multiplication can only enhance the areas with higher saliency values in both images and suppress the areas with lower saliency values. Combining the feature fusion method (Huang *et al.*, 2020), the feature fusion function is defined by:

$$S = (S_D + 1) \times (S_{AC} + S_{ST}), \quad (9)$$

where S_D , S_{AC} and S_{ST} , respectively, represent saliency images of local grayscale feature, local contrast feature and edge corner effect. For consistency with different features, all saliency values are scaled to the range of 0–1 by min-max normalization. S is the final saliency map.

The defect area can be correctly identified with the local grayscale feature and the local contrast feature, but some nondefective areas also have higher saliency values. Edge corner effect can effectively identify the edge part in the image. The saliency value of the nondefective area in the local grayscale feature is zero, which will invalidate the other features in the pixel multiplication. Therefore, the value of S_D is incremented by 1. Figure 6(g) is the final fusion saliency image. The multiplication of pixels makes the region with high saliency value of multiple features be highlighted. The high saliency region in the final saliency image is closer to the real defect area.

2.4.5 Image segmentation

Meanshift (Comaniciu and Meer, 2002) clustering is a well-established algorithm that has been applied successfully in image processing and computer vision. Cluster centers are derived by local mode seeking identifying maxima in the normalized density of the image. Through the meanshift clustering, the original pixel

value is replaced by the pixel value of the clustering center. Thus, the local similar texture in the image is removed, and the features with large differences such as edges are retained.

To segment the defect area in the cluster image, define the Niblack (He *et al.*, 2016) threshold TH :

$$TH = \mu_S + k \cdot \delta_S \quad (10)$$

where μ_S and δ_S are the mean and variance of the saliency image S , respectively, and k is a control coefficient. The defect recognition performance can be controlled by adjusting the parameter k value, the final images are demonstrated in Figure 6(h) after clustering with meanshift algorithm.

3. Evaluation

3.1 Data set

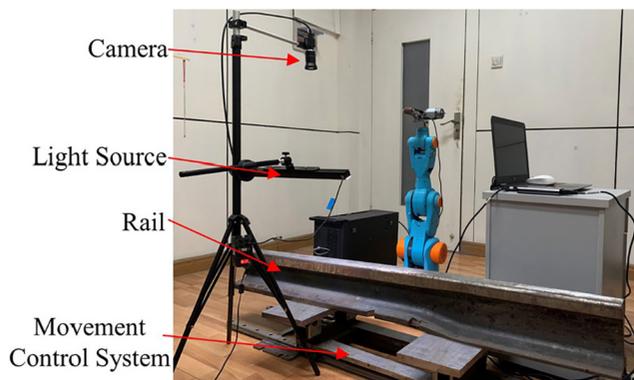
3.1.1 RSDDS-126 data set

Under laboratory conditions, an automatic grinding platform is built to simulate the production environment. The data set samples are taken from an actual industrial production line of one section-steel factory. The experimental equipment is shown in Figure 7. The essential equipment is a Yuanqi charge coupled device (CCD) camera with a resolution of $1,280 \times 1,024$ pixels. The lens is mounted vertically downwards, and an light-emitting diode (LED) light source controlled by the software reduces the effect of natural light variation on imaging quality. The resolution of one pixel is $0.19 \times 0.19 \text{ mm}^2$. In total, 126 images with resolution of $1,280 \times 1,024$ were captured by CCD camera to form RSDDS-126 data set. All defects are visually labeled. Note that defects with area less than 15 mm^2 can be ignored according to the requirements of rail production process.

3.1.2 Rail surface discrete defect data set

The RSDD is a public railway image data set, which is mainly composed of grayscale images captured from express rails and heavy haul rails, including two subdata sets: Type I and Type II. the Type I RSDD data set contained 67 challenging images acquired from express rails and the Type II RSDD data set contained 128 challenging images acquired from ordinary/heavy haul rails. Each surface image contained one or more defects that were difficult to identify owing to the noisy backgrounds. These defects included squats at different levels and other related damages produced by rolling contact fatigue.

Figure 7 Automatic rail grinding platform



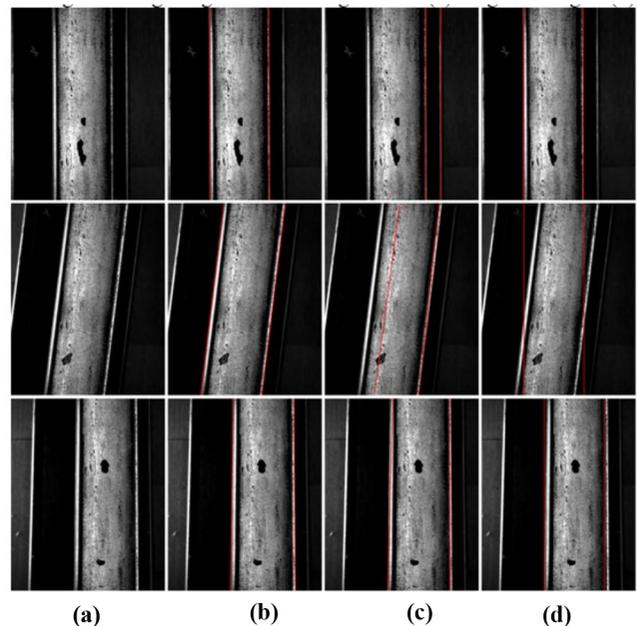
3.2 Rail surface region location experiment

To verify the accuracy of the proposed rail surface location algorithm, intersection over union (IoU) is applied to evaluate the algorithm. MRLS method is compared with two classic algorithms of Hough detection method (Hu *et al.*, 2018) and the track extraction based on projection profile (TEBP) (Li and Ren, 2012) method. Hough detection method detects the line of rail edge to locate the rail surface region, and the TEBP (Li and Ren, 2012) method used the characteristics of the rail surface and the irrelevant surrounding area to have different pixel average values along the rail direction to segment the target region. The results are illustrated in the Figure 8 and Table 1. The location results formed by MRLS method are most similar to the real boundary of the rail. The Hough detection method will identify irrelevant areas due to the influence of uneven reflection and the edges of the rail bottom. Due to the limitation of the principle, TEBP method is only effective when the rail is vertical. The edge of the rail surface cannot be positioned when the rail is inclined or the camera is not vertical to the rail surface. MRLS method has the highest IoU value and the strongest robustness.

3.3 Testing experiment of rail surface defect

For a comprehensive assessment, pixel-level index (precision, recall and F -measure) and defect-level index (precision', recall', F' -measure) are the important factors to evaluate the experimental results of RSDDS-126 and RSDD data set (Yu *et al.*, 2019). Precision reflects information on how many generated masks are true defects among all the detections. Recall reveals information on how many defects were detected among all the defects. The F -measure is a criterion, which tries to capture both precision and recall in a single number.

Figure 8 Examples of recognition results using three target region location algorithms



Notes: (a) Original images; (b) MRLS; (c) Hough detection; (d) TEBP

Table 1 Experimental results of three target region location algorithms

Approach	Rail is not in the center (%)	Rail is inclined (%)	Camera is not perpendicular (%)	Total (%)
MRSL	99.14	99.41	99.20	99.24
Hough detection	46.44	86.80	65.22	65.51
TEBP	98.82	80.05	93.77	91.18

3.3.1 RSDDS-126 data set

In defect recognition, both t in equation (4) and k in equation (10) directly determine the precision of defect recognition. If t and k are too high, more defect information will be filtered out and the number of missing defects will increase. On the contrary, if t and k are too small, noise will be generated and the number of false detection of defects will increase. A balance between precision and recall should be established. The algorithm has the best comprehensive performance when F -measure is the largest. Figure 9 shows the F -measure of defect identification under different t and k values. Reported values are average values for the whole data set. The best F -measure can be achieved when $t = 53$ and $k = 3.5$.

The determined parameters were obtained to identify the RSDDS-126 data set, and the results were shown in Figure 10. In the saliency image as show in Figure 10(c), all defect areas are visible and have high saliency value. Large-scale defects that pose a potential threat to the rails are identified in the binary image as shown in Figure 10(d).

The experiments have been performed to compare the effects among BLCM (Hu et al., 2018), P-M (He et al., 2016), local normalization (LN) + defect localization based on projection profile (DLBP) (Li and Ren, 2012) and the presented detection algorithm in this paper. The results in Figure 10(e)–(g) depict some examples of the detection results of three methods. The defect area smaller than 15 mm^2 has been removed. The detection results of BLCM are poor. A large number of overextracted regions and underextracted regions are produced. The edge of defects detected by P-M will break, so the defect region cannot be explicitly found. And the edge area of the rail is erroneously extracted due to the uneven reflection. LN + DLBP cannot identify the contours of defects and will give false positives when the illumination is uneven along the rail. The evaluation indexes are compared in Table 2. Because P-M and LN + DLBP cannot accurately detect the contours of the defects, only the defect-level index is evaluating indicator. The results shows that six metrics of the bionic algorithm are better than other algorithms.

Figure 9 Effect of different t and k value on algorithm performance index

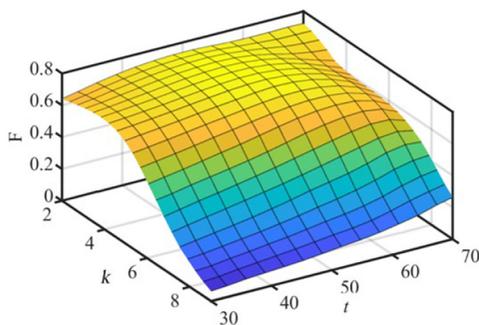
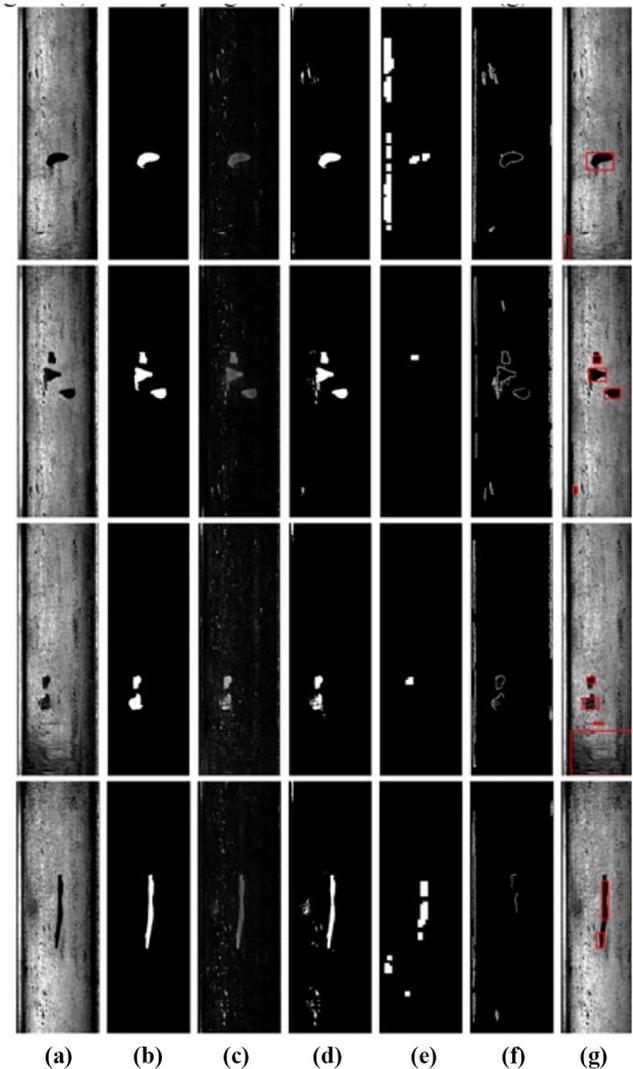


Figure 10 The sample defect images and inspection results of different methods from RSDDS-126 data set



Notes: (a) Original images; (b) ground truth; (c) saliency images; (d) binary images; (e) BLCM; (f) P-M; (g) LN + DLBP

Five classical saliency models were compared on the RSDD-126 data set. It includes three models which use global color rarity, such as FT (Achanta et al., 2009), LC (Zhai and Shah, 2006) and HC (Cheng et al., 2014). Besides, other two models can also be compared, such as ITTI (Itti et al., 1998) and BMS (Zhang and Sclaroff, 2016), which use visual attention mechanism. The evaluation metrics is shown in Table 3. It can be ascertained that our method is better than other saliency models.

Table 2 Experimental results of different methods on the RSDD-126 data set

Approach	Pre (%)	Rec (%)	F (%)	Pre' (%)	Rec' (%)	F' (%)
Our method	69.83	86.42	76.53	85.71	90.57	88.07
BLCM	40.68	26.68	20.48	34.12	19.81	25.07
P-M	–	–	–	12.83	63.68	21.35
LN + DLBP	–	–	–	55.73	70.80	62.36

Table 3 Experimental results of different saliency models on the RSDD-126 data set

Approach	Pre (%)	Rec (%)	F (%)	Pre' (%)	Rec' (%)	F' (%)
ITTI	73.62	73.92	72.15	78.66	85.38	81.88
FT	60.64	88.16	69.52	66.78	91.51	77.21
LC	50.66	84.68	62.32	66.90	91.98	77.45
HC	50.15	86.52	62.14	61.86	89.15	73.04
BMS	39.69	53.31	43.72	60.08	46.70	52.55

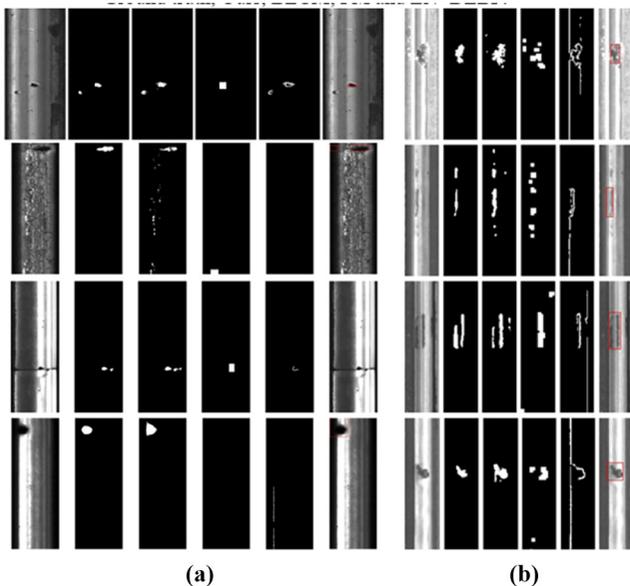
3.3.2 Rail surface discrete defect data set

We again evaluate the dependency on the parameter t and k . The appropriate parameter is selected using the same way as that in RSDDS-126 data set. The results of the RSDD data set are shown in [Figure 11](#) and the evaluation metrics of comparison are shown in [Tables 4](#) and [5](#). It can be ascertained that our method is better than other algorithms in Type-II. In Type-I, LN + DLBP had the best detection result, but the F' of the four tested methods were all low.

3.3.3 Failure case analysis

The bionic algorithm in this paper cannot detect some special defects in the data set. Several reasons leading to poor performances are as follows:

Figure 11 The results of different methods for the Type I (a) and Type II (b) of RSDD data set. From left to right, each column in (a) and (b) is: original image, ground truth, ours, BLCM, PM and LN + DLBP



Notes: (a) and (b) is: original image, ground truth, ours, BLCM, PM and LN + DLBP

Table 4 Experimental results of different methods on the Type-I RSDD data set

Approach	Pre (%)	Rec (%)	F (%)	Pre' (%)	Rec' (%)	F' (%)
Our method	37.37	43.30	36.34	76.00	25.55	38.24
BLCM	16.97	26.47	18.59	5.41	25.55	8.92
P-M	–	–	–	36.96	12.41	18.58
LN + DLBP	–	–	–	56.18	39.42	46.33

Table 5 Experimental results of different methods on the Type-II RSDD data set

Approach	Pre (%)	Rec (%)	F (%)	Pre' (%)	Rec' (%)	F' (%)
Our method	61.27	64.36	59.28	86.55	71.82	78.50
BLCM	15.17	16.90	11.25	3.35	10.50	5.08
P-M	–	–	–	11.55	40.33	17.96
LN + DLBP	–	–	–	86.92	61.14	71.79

- Defects with low saliency values. The saliency values of some special defects in dark background are lower than that in bright background; when the brightness of the background changes greatly, only defects in bright background can be identified.
- Background interference. Some surfaces have features similar to defects, resulting in invalid saliency image.
- Minor defects. Black spots in the rail background are identified as defects, but the areas of these defects are small. They can be removed with area filtering.

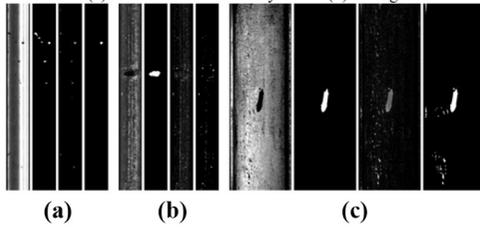
Some examples of failure cases are shown in [Figure 12](#). Each column in [Figure 12\(a\)–\(c\)](#) from left to right is: original images, ground truth, saliency images and binary images. Reason (a) and (b) are the main reasons for the poor performance of the algorithm on Type-I RSDD data set. But the above phenomena will rarely happen in the railway manufacturing process.

3.4 Ablation study

The core features of the saliency algorithm are the local grayscale feature and the local contrast feature because they can all obtain the whole saliency image of the defects image. To compare the effect of single feature, the saliency images of the above two features are segmented and evaluated by the same postprocessing method on the RSDDS-126 data set. The algorithm verification image and evaluation metrics are shown in [Figure 13](#) and [Table 6](#). Blue areas represent overextracted regions and green areas represent underextracted regions.

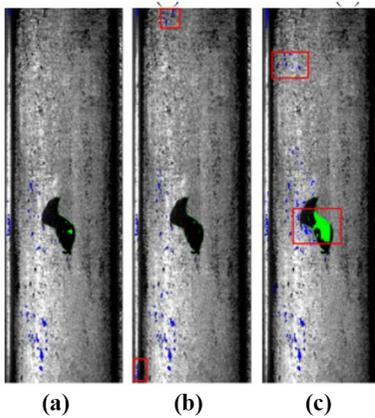
When recognizing defects with only single feature, more false positives are generated, so Pre and Pre' are lower. As shown in the red rectangle in [Figure 13\(b\)](#) and [\(c\)](#), part of the overextracted regions and underextracted regions are eliminated after feature fusion, which verifies the rationality of the feature fusion method. The evaluation metrics after removing the edge corner effect are also shown in [Table 6](#). It can be seen that the edge corner effect can slightly improve the detection effect. The fusion feature has the highest F and F' , both of F and F' reach 76.53% and 88.07%, respectively, which shows that all of three features in the bionic algorithm are necessary.

Figure 12 Examples of failure cases



Notes: (a) Defects of low saliency value; (b) background interference; (c) minor defects

Figure 13 Example of verification image with different features



Notes: (a) Fusion feature; (b) local grayscale feature; (c) local contrast feature

Table 6 Evaluation metrics of defect detection for different features

Approach	Pre (%)	Rec (%)	F (%)	Pre' (%)	Rec' (%)	F' (%)
Fusion feature	69.83	86.42	76.53	85.71	90.57	88.07
S _D	62.76	86.57	71.93	81.30	90.57	85.68
S _{AC}	50.66	84.68	62.32	78.91	89.62	83.92
Remove S _{ST}	68.96	86.32	75.98	83.61	90.57	86.95

4. Conclusion

A bionic rail surface defect detection algorithm applied in the production process is proposed in this paper. According to the characteristics of rail surface defect images, a rail surface region positioning algorithm based on MRLS and a light correction algorithm based on background difference are proposed. The fusion saliency image can be generated with local grayscale feature, local contrast feature and edge corner effect, and then be clustered and segmented with meanshift algorithm to obtain accurate defect location and contour information. On the RSDDS-126 data set, pixel-level index and defect-level index are the important factors to evaluate the effectiveness of several models. Two indexes of F and F' reach 76.53% and 88.07%, respectively, which exceed the existed three advanced methods

and five saliency models. Experimental results indicate that the algorithm can effectively identify surface defects in railway manufacturing process. Furthermore, the bionic algorithm is tested on the public railway image data set (RSDD data set) and several failure reasons are analyzed. Finally, the ablation experiment verifies that the detection ability of a single feature is lower than that of the fusion feature, which proves the necessity of three saliency features.

It is worth noting that the method in this paper is only suitable for locating rail surface defects, more research is needed to determine the defect attributes. Therefore, the defect classification will be studied in the future to provide more detailed defect information for the grinding process.

References

Achanta, R., Estrada, F., Wils, P. and Sússtrunk, S. (2008), "Salient region detection and segmentation", in Gasteratos, A., Vincze, M. and Tsotsos, J.K. (Eds), *International Conference on Computer Vision Systems*, Springer, Berlin, Heidelberg, pp. 66-75, doi: [10.1007/978-3-540-79547-6_7](https://doi.org/10.1007/978-3-540-79547-6_7).

Achanta, R., Hemami, S., Estrada, F. and Susstrunk, S. (2009), "Frequency-tuned salient region detection", *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1597-1604, doi: [10.1109/CVPR.2009.5206596](https://doi.org/10.1109/CVPR.2009.5206596).

Ben Gharsallah, M.B. and Ben Braiek, E.B. (2015), "Weld inspection based on radiography image segmentation with level set active contour guided off-center saliency map", *Advances in Materials Science and Engineering*, Vol. 2015, pp. 1-10, doi: [10.1155/2015/871602](https://doi.org/10.1155/2015/871602).

Cheng, M., Mitra, N.J., Huang, X., Torr, P.H. and Hu, S. (2014), "Global contrast based salient region detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37 No. 3, pp. 569-582, doi: [10.1109/TPAMI.2014.2345401](https://doi.org/10.1109/TPAMI.2014.2345401).

Comaniciu, D. and Meer, P. (2002), "Mean shift: a robust approach toward feature space analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24 No. 5, pp. 603-619, doi: [10.1109/34.1000236](https://doi.org/10.1109/34.1000236).

Cruz, F.C., Simas Filho, F.E., Albuquerque, M.C., Silva, I.C., Farias, C.T. and Gouvêa, L.L. (2017), "Efficient feature selection for neural network based detection of flaws in steel welded joints using ultrasound testing", *Ultrasonics*, Vol. 73, pp. 1-8, doi: [10.1016/j.ultras.2016.08.017](https://doi.org/10.1016/j.ultras.2016.08.017).

Faghih-Roohi, S., Hajizadeh, S., Núñez, A., Babuska, R. and De Schutter, B. (2016), "Deep convolutional neural networks for detection of rail surface defects", *2016 International joint conference on neural networks (IJCNN)*, pp. 2584-2589, doi: [10.1109/IJCNN.2016.7727522](https://doi.org/10.1109/IJCNN.2016.7727522).

Gan, J., Wang, J., Yu, H., Li, Q. and Shi, Z. (2020), "Online rail surface inspection utilizing spatial consistency and continuity", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Vol. 50 No. 7, pp. 2741-2751, doi: [10.1109/TSMC.2018.2827937](https://doi.org/10.1109/TSMC.2018.2827937).

Guan, S. (2015), "Strip steel defect detection based on saliency map construction using gaussian pyramid decomposition", *ISI³ International*, Vol. 55 No. 9, pp. 1950-1955, doi: [10.2355/isijinternational.ISIJINT-2015-041](https://doi.org/10.2355/isijinternational.ISIJINT-2015-041).

Harris, C. and Stephens, M. (1988), "A combined corner and edge detector", *Proc Alvey vision conference*, Vol. 1988 No. 3, pp. 147-151.

- He, Z., Wang, Y., Yin, F. and Liu, J. (2016), "Surface defect detection for high-speed rails using an inverse P-M diffusion model", *Sensor Review*, Vol. 36 No. 1, pp. 86-97, doi: [10.1108/SR-03-2015-0039](https://doi.org/10.1108/SR-03-2015-0039).
- Hu, Z., Zhu, H., Hu, M. and Ma, Y. (2018), "Rail surface spalling detection based on visual saliency", *IEEE Transactions on Electrical and Electronic Engineering*, Vol. 13 No. 3, pp. 505-509, doi: [10.1002/tee.22594](https://doi.org/10.1002/tee.22594).
- Huang, Y., Qiu, C. and Yuan, K. (2020), "Surface defect saliency of magnetic tile", *The Visual Computer*, Vol. 36 No. 1, pp. 85-96, doi: [10.1007/s00371-018-1588-5](https://doi.org/10.1007/s00371-018-1588-5).
- Itti, L., Koch, C. and Niebur, E. (1998), "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20 No. 11, pp. 1254-1259, doi: [10.1109/34.730558](https://doi.org/10.1109/34.730558).
- Jin, X., Wang, Y., Zhang, H., Zhong, H., Liu, L., Wu, Q.M.J. and Yang, Y. (2020), "DM-RIS: deep multimodal rail inspection system with improved MRF-GMM and CNN", *IEEE Transactions on Instrumentation and Measurement*, Vol. 69 No. 4, pp. 1051-1065, doi: [10.1109/TIM.2019.2909940](https://doi.org/10.1109/TIM.2019.2909940).
- Koch, C. and Ullman, S. (1987), "Shifts in selective visual attention: towards the underlying neural circuitry", in Vaina, L.M. (Ed.), *Matters of Intelligence*, Springer, Dordrecht, pp. 115-141, doi: [10.1007/978-94-009-3833-5_5](https://doi.org/10.1007/978-94-009-3833-5_5).
- Li, Q. and Ren, S. (2012), "A real-time visual inspection system for discrete surface defects of rail heads", *IEEE Transactions on Instrumentation and Measurement*, Vol. 61 No. 8, pp. 2189-2199, doi: [10.1109/TIM.2012.2184959](https://doi.org/10.1109/TIM.2012.2184959).
- Liang, Z., Zhang, H., Liu, L., He, Z. and Zheng, K. (2018), "Defect detection of rail surface with deep convolutional neural networks", *2018 13th World Congress on Intelligent Control and Automation (WCICA)*, pp. 1317-1322, doi: [10.1109/WCICA.2018.8630525](https://doi.org/10.1109/WCICA.2018.8630525).
- Liu, T., Yuan, Z., Sun, J., Wang, J., Zheng, N., Tang, X. and Shum, H. (2011), "Learning to detect a salient object", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 33 No. 2, pp. 353-367, doi: [10.1109/TPAMI.2010.70](https://doi.org/10.1109/TPAMI.2010.70).
- Min, Y., Xiao, B., Dang, J., Yue, B. and Cheng, T. (2018), "Real time detection system for rail surface defects based on machine vision", *EURASIP Journal on Image and Video Processing*, Vol. 2018 No. 1, pp. 1-11, doi: [10.1186/s13640-017-0241-y](https://doi.org/10.1186/s13640-017-0241-y).
- Ni, X., Liu, H., Ma, Z., Wang, C. and Liu, J. (2021), "Detection for rail surface defects via partitioned edge feature", *IEEE Transactions on Intelligent Transportation Systems*, pp. 1-17, doi: [10.1109/TITS.2021.3058635](https://doi.org/10.1109/TITS.2021.3058635).
- Nieniewski, M. (2020), "Morphological detection and extraction of rail surface defects", *IEEE Transactions on Instrumentation and Measurement*, Vol. 69 No. 9, pp. 6870-6879, doi: [10.1109/TIM.2020.2975454](https://doi.org/10.1109/TIM.2020.2975454).
- Niu, M., Song, K., Huang, L., Wang, Q., Yan, Y. and Meng, Q. (2021), "Unsupervised saliency detection of rail surface

- defects using stereoscopic images", *IEEE Transactions on Industrial Informatics*, Vol. 17 No. 3, pp. 2271-2281, doi: [10.1109/TII.2020.3004397](https://doi.org/10.1109/TII.2020.3004397).
- Piccardi, M. (2004), "Background subtraction techniques: a review", *2004 IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3099-3104, doi: [10.1109/ICSMC.2004.1400815](https://doi.org/10.1109/ICSMC.2004.1400815).
- Song, K. and Yan, Y. (2013), "Micro surface defect detection method for silicon steel strip based on saliency convex active contour model", *Mathematical Problems in Engineering*, Vol. 2013, pp. 1-13, doi: [10.1155/2013/429094](https://doi.org/10.1155/2013/429094).
- Song, K., Hu, S., Yan, Y. and Li, J. (2014), "Surface defect detection method using saliency linear scanning morphology for silicon steel strip under oil pollution interference", *ISIJ International*, Vol. 54 No. 11, pp. 2598-2607, doi: [10.2355/isijinternational.54.2598](https://doi.org/10.2355/isijinternational.54.2598).
- Yu, H., Li, Q., Tan, Y., Gan, J., Wang, J., Geng, Y. and Jin, L. (2019), "A coarse-to-Fine model for rail surface defect detection", *IEEE Transactions on Instrumentation and Measurement*, Vol. 68 No. 3, pp. 656-666, doi: [10.1109/TIM.2018.2853958](https://doi.org/10.1109/TIM.2018.2853958).
- Yuan, H., Chen, H., Liu, S., Lin, J. and Luo, X. (2019), "A deep convolutional neural network for detection of rail surface defect", *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*, pp. 1-4, doi: [10.1109/VPPC46532.2019.8952236](https://doi.org/10.1109/VPPC46532.2019.8952236).
- Zhai, Y. and Shah, M. (2006), "Visual attention detection in video sequences using spatiotemporal cues", *Proceedings of the 14th ACM international conference on Multimedia*, pp. 815-824, doi: [10.1145/1180639.1180824](https://doi.org/10.1145/1180639.1180824).
- Zhang, D., Song, K., Wang, Q., He, Y., Wen, X. and Yan, Y. (2021), "Two deep learning networks for rail surface defect inspection of limited samples with line-level label", *IEEE Transactions on Industrial Informatics*, Vol. 17 No. 10, pp. 6731-6741, doi: [10.1109/TII.2020.3045196](https://doi.org/10.1109/TII.2020.3045196).
- Zhang, H., Jin, X., Wu, Q.M.J., Wang, Y., He, Z. and Yang, Y. (2018), "Automatic visual detection system of railway surface defects with curvature filter and improved gaussian mixture model", *IEEE Transactions on Instrumentation and Measurement*, Vol. 67 No. 7, pp. 1593-1608, doi: [10.1109/TIM.2018.2803830](https://doi.org/10.1109/TIM.2018.2803830).
- Zhang, J. and Sclaroff, S. (2016), "Exploiting surroundedness for saliency detection: a Boolean map approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 38 No. 5, pp. 889-902, doi: [10.1109/34.1000236](https://doi.org/10.1109/34.1000236).
- Zhu, J., Min, Q., Wu, J. and Tian, G.Y. (2018), "Probability of detection for eddy current pulsed thermography of angular defect quantification", *IEEE Transactions on Industrial Informatics*, Vol. 14 No. 12, pp. 5658-5666, doi: [10.1109/TII.2018.2866443](https://doi.org/10.1109/TII.2018.2866443).

Corresponding author

Hua Zhai can be contacted at: jxzhaihuajx@hfut.edu.cn