

# Automatic Visual Detection System of Railway Surface Defects With Curvature Filter and Improved Gaussian Mixture Model

Hui Zhang<sup>1</sup>, Xiating Jin, Q. M. Jonathan Wu, *Senior Member, IEEE*, Yaonan Wang, Zhendong He, and Yimin Yang, *Member, IEEE*

**Abstract**—Rails are among the most important components of railway transportation, and real-time defects detection of the railway is an important and challenging task because of intensity inhomogeneity, low contrast, and noise. This paper presents an automatic railway visual detection system (RVDS) for surface defects and focuses on several key issues of RVDS. First, in view of challenges such as complex condition and orbital reflectance inequality, we put forward a region-of-interest detection region extraction algorithm by vertical projection and gray contrast algorithm. In addition, a curvature filter equipped with implicit computing and surface preserving power is studied to eliminate noise and keep only the details. Then, an improved fast and robust Gaussian mixture model based on Markov random field is established for accurate and rapid surface defect segmentation. Additionally, an expectation–maximization algorithm is applied to optimize the parameters. The experimental results demonstrate that the proposed method performs well with both noisy and railway images, which enables identification and segmentation of the defects from rail surface, achieving detection performance with 92% precision and 88.8% recall rate on average, and is robust compared with the related well-established approaches.

Manuscript received September 19, 2017; revised November 24, 2017; accepted December 21, 2017. Date of publication February 28, 2018; date of current version June 7, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61401046 and Grant 61473049, in part by the National Technology Support Project under Grant 2015BAF11B01, in part by the Scientific Research Fund of Hunan Provincial Education Department under Grant 17C0046, and in part by the Hunan Province Key Laboratory of Videometric and Vision Navigation under Grant TXCL-KF2013-001. The Associate Editor coordinating the review process was Dr. Emanuele Zappa. (*Corresponding authors:* Hui Zhang; Xiating Jin.)

H. Zhang is with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha 410012, China, also with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China, and also with the Computer Vision and Sensing Systems Laboratory, Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: zhanghui@uwindsor.ca).

X. Jin and Y. Wang are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, China (e-mail: xtchin@hnu.edu.cn; yaonan@hnu.edu.cn).

Q. M. J. Wu is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada, and also with the School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China (e-mail: jwu@uwindsor.ca).

Z. He is with the School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450000, China (e-mail: hezhendong\_itl@163.com).

Y. Yang is with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON N9B 3P4, Canada (e-mail: eyyang@uwindsor.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIM.2018.2803830

**Index Terms**—Curvature filter, improved Gaussian mixture model (GMM), Markov random field (MRF), railway surface defect, visual detection.

## I. INTRODUCTION

RAILWAY transportation is considered as one of the safest transportation modes all over the world. The quality of railway track is essential for the safety of train operation, so the inspection of railway defects has become increasingly important and urgent. However, traditional human inspection and contact measurement techniques are still being used to detect failures in the rails, resulting in low sensitivity and low accuracy rate. There are inadequate to meet the needs of today's advanced railway technology [1]. The challenges for existing technologies include early detection, reliability, and the cost of detection systems. A rail inspection system based on ultrasonic guided waves and noncontact ultrasound transduction [2] can predict the expected detection performance of the transverse defect in the presence of various defects of different sizes and positions. Recently, there have been a number of nondestructive testing methods applied in railroad damage detection such as the eddy current testing method [3], an acoustic emission (AE) method [4], an ultrasonic method [5]–[7], and so on. High-speed flaw detection using eddy current [3] can evaluate the location and extent of the rail damage. However, there are difficulties in intuitively displaying the different types of rail defects. Zhang *et al.* [4] present an improved AE rail defect detection method by multilevel adaptive noise cancellation (ANC) with variable step-size least mean square, where the features of noise signals and defect signals are analyzed for effective detection. Meanwhile, structural health monitoring (SHM) of aging railway tracks is important for safety, to prevent catastrophic failure and reduce maintenance costs. In recent years, interest has grown in performing SHM by measuring structural deformation [5], full-field strain, and geometry profiles using 3-D digital image correlation systems [6]. Mariani *et al.* [7] developed a railroad monitoring system that can detect cracks in rails while in motion, using ultrasonic waves and noncontact (laser) probing to detect the flaws.

Visual detection [8], [9] of the most critical components for the safe operation of trains is necessary. Feng *et al.* [10] propose an automatic visual inspection system (VIS) for detecting partially worn and completely missing fasteners using a

probabilistic topic model. Molodova *et al.* [11] present a new measuring system for the automatic detection of squats. The detection algorithm [12], [13] relies on the signature tunes of the squats, which were identified from numerical simulations in a previous work and were validated by field measurements. Principal component analysis (PCA) with a histogram depth peaks peak checking approach [14] is utilized to locate the fastener. To detect significant rail components such as ties, tie plates, and anchors, Li *et al.* [15] propose a novel global optimization framework to combine evidence from multiple cameras, global positioning system, and distance measurement instruments to further improve the detection performance. Similarly, Mazzeo *et al.* [16] and Marino *et al.* [17] use several combinations of wavelet transform and PCA methods to detect the presence as well as the absence of the bolts on the sleepers or some other critical components. However, for component detection, these proposed methods can only handle specific types, whereas the robustness on illumination variation was not discussed.

Different from component detection, Li and Ren [18], [19] develop a real-time VIS for discrete surface defects of rail heads. Salvador *et al.*'s work [20] delves into data acquisition and processing techniques in order to enhance track surveying processes, which successfully run on the Metropolitan Rail Network of Valencia, Spain. Similar efforts were also reported by Resendiz *et al.* [21]; they took advantage of image and 1-D signal and 1-D signal processing methods to complete three tasks: component detection, turnout detection, and tie detection on a railroad track. The specific image-to-signal conversion and the final periodicity detection or classification vary for the three algorithms. The three most related works have been proposed [22]–[24]. Specifically, Alippi *et al.* [25] apply laser scanning and image analysis approach for track profile extraction with real-time image processing. Liang *et al.* [26] also introduce an adaptive noise canceling (ANC) and time-frequency application for railway wheel flat and rail surface defect detection. Other image processing technologies have been used for preprocessing and classification of the defects, including Gabor filtering [27], learned partial differential equation [28], and maximally stable extremal region marking [29]. In our previous work [30], we use an inverse P-M diffusion model to enhance the railway surface defect boundaries and suppress noises. Similarly, Taştimur *et al.* [31] also apply image processing techniques such as image enhancement, noise cancellation, feature extraction, and morphological operations to rail images for rail surface defect detection. In summary, most of the present methods take advantage of image gray features to inspect the track surface. However, when the input railway images are noisy and generally concentrated in high-frequency part of the image, this affects extraction of image details and edge. Hence, railway surface image defects are a challenging problem.

Moreover, some literature collects numerous labeled railway surface defect samples such as small spots, welds, and insulated joints. Then, based on train classifiers, several machine learning approaches, for instance, the semisupervised techniques [32], deep convolutional neural networks [33], and deep multitask learning [34], have recently been applied for



Fig. 1. Some examples of rail surface defects.

automatic detection of rail surface defects. Although these learning methods can achieve high accuracy, the training stage requires huge labeled samples. Different from this recently learning method, our samples set in unlabeled and composed of various types of defects, low signal-to-noise ratio (SNR), and illumination images. Rail defects are almost always caused by fatigue from wheel–rail interactions and the presence of defective materials. Rail surface defects are any imperfection, damage, or deformation at or near the exterior surface of a rail. Any type of surface condition can be a significant obstacle in the detection of the underlying rail defect. The most common surface defects [35] are squats, spalling, flaking, cracks, wheel burns, running band deviation, broken rail, rolling contact fatigue cracks, and so on. For example, one common defect type is the so-called squat, which is rail defects on the running band on the rails. The displacement of parent metal from the rail head is generally referred to as spalling. Flaking is a progressive horizontal separation of the running surface of the rail near the gauge corner, with scaling or chipping of small slivers. Fig. 1 show some examples of rail surface defects. Moreover, the rail defect rate is nonlinear. Rail track defects per mile or defect initiation grows with age but the growth rate is not linear. The crack also grows nonlinearly from initiation to failure. It has been shown that the past defect history appears to be important for the prediction of future rail defects [36]. In [37], statistical rail defect distribution is demonstrated. Based on this situation, the railway detection vehicle needs to run periodically. For our acquired railway images, we conclude four major problems in surface inspection systems.

- 1) *Illumination Inequality:* The image detection system working condition is complicated. Railway images are captured by a camera installed under a train coach. This system must withstand natural light and mechanical vibration, because the interference of hostile environments leads to the captured image being distorted or noise. Natural light will significantly affect the brightness of the acquired images, as shown in Fig. 2(a). Although specific LED light sources and sun shields are used, this problem is still unsolved.
- 2) *Orbital Reflectance Inequality:* Because of the curved geometry of the rail head, the reflection properties of different surfaces are different. For example, the bulging central area is smooth, when the reflected light is heavy, and both sides of the region are diffusely reflected, if the reflected light is less, so that the captured image is a central region of the rail's bright, dark side areas, which makes it hard for defects and background segmentation.

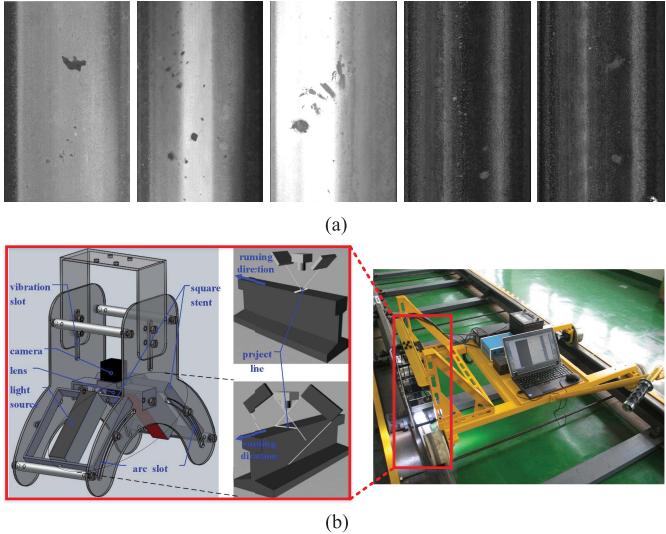


Fig. 2. Designed imaging system and acquired illumination images. (a) Railway surface with different illumination conditions. (b) Prototype of the proposed defects detection system.

- 3) Limited features are available for defect recognition. Adverse factors 1) and 2) give the image processing a bad ill-posed problem, i.e., denoising and segmentation. Several effective methods are widely used in this aspect, such as total variation (Tv) and curvature geometric flow, but generally they are computationally complex, slow, and difficult to realize. In addition, the Gaussian mixture model (GMM) of the segmentation effect is very sensitive to external interference. This paper is based on inspiration from these methods but puts forth a far superior algorithm.
- 4) The requirement is of high-speed detection, high accuracy, and high reliability. Detection systems are expected to be in real time in high-speed railway systems, so a preferable detection system should have a low computational complexity. Consequently, automated inspection offers a consistent and fast process for visual detection.

In ideal conditions, the gray level of surface defects is significantly different from the background [27]. However, this order is often broken by illumination inequality, low SNR, positional deviation, and so on. This disorder of gray images brings out several challenges for a railway visual detection system (RVDS). To handle these problems, we propose a new defects detection structure: region extraction and gray equalization-fast filter and denoising-robust segmentation.

This paper is organized as follows. In Section II, the railway detection imaging system and its principle are explained in detail. In Section III, a brief analysis of the algorithms implemented for the main issues is described. Experimental results and implementation details obtained by applying the proposed method are given in Section IV. Finally, Section V provides the conclusions.

## II. SYSTEM OVERVIEW

### A. System Configuration

As we know, many factors influence the imaging quality, such as light intensity, exposure time, incident angle, camera

motion velocity, line rate and vibration, and so on. Hence, a specially designed imaging mechanism is used to acquire the desired images. A 3-D illustration of the imaging system, which is installed in a specially designed car, is shown in Fig. 2(b). The key component is a DalsaSpyder3 line-scan camera with a resolution of 1024 pixels and a maximum line rate of 68 000 lines/s, which can get a set of frames and finally generates a panorama. It is to be noted that the mismatch between the camera velocity and the line rate will lead to image distortion. To synchronize data acquisition, the line-scan camera is triggered by a wheel encoder. The arc slots are used to adjust the incident angle. The light intensity and the exposure time are controlled by the light controller and the software, respectively. A panorama is obtained based on a set of frames. Vibration slots are used to reduce vibration through the damper. Moreover, the quality of the rail surface image is inevitably affected by natural light, stain, and rust; this results in challenges in detecting defects from these images [38]. Then, the captured images are processed by the processor for real-time detection.

### B. Framework of the Proposed Detection Approach

This section presents an overview of the proposed detection approach with several parts. To be specific, in image pre-processing, extract region-of-interest (ROI) detection region for high-resolution image utilizing gray vertical projection and enhance the image by gray image equalization due to inequality in illumination and variation of reflection properties. In addition, the image regularization term of the variational model is calculated as a curvature filter to rapidly denoise. Moreover, in the proposed segmentation method, the improved GMM based on Markov random field (MRF) accurately extracts the contour of the surface defects. The framework of our railway surface defect detection approach procedure is shown in Fig. 3.

## III. DEFECTS DETECTION METHOD

What we are concerned with here is how to process railway images quickly and effectively and obtain the defects where the edge details of the input images are preserved with reduced noise. The process of the proposed approach is shown in Fig. 3. In this section, we first describe the curvature filter model for image denoising of RVDS and then present the improved GMM based on MRF for fast segmentation of defects.

### A. Curvature Filter

Different from the traditional method [39], [40], the curvature filter first studies the developable or minimal surface instead of curvature and then selects the optimal surface to approximate the original data, and in this way, the curvature energy could be minimized. As a unique optimization algorithm, the curvature filter not only inherits the superiority of the traditional method, that is to use curvature regularization, but also provides a grand new way to minimize the regularization energy, which is to use developable or minimal surface. The success is that it makes full use of the continuity of differential geometry and the discreteness of data, bringing

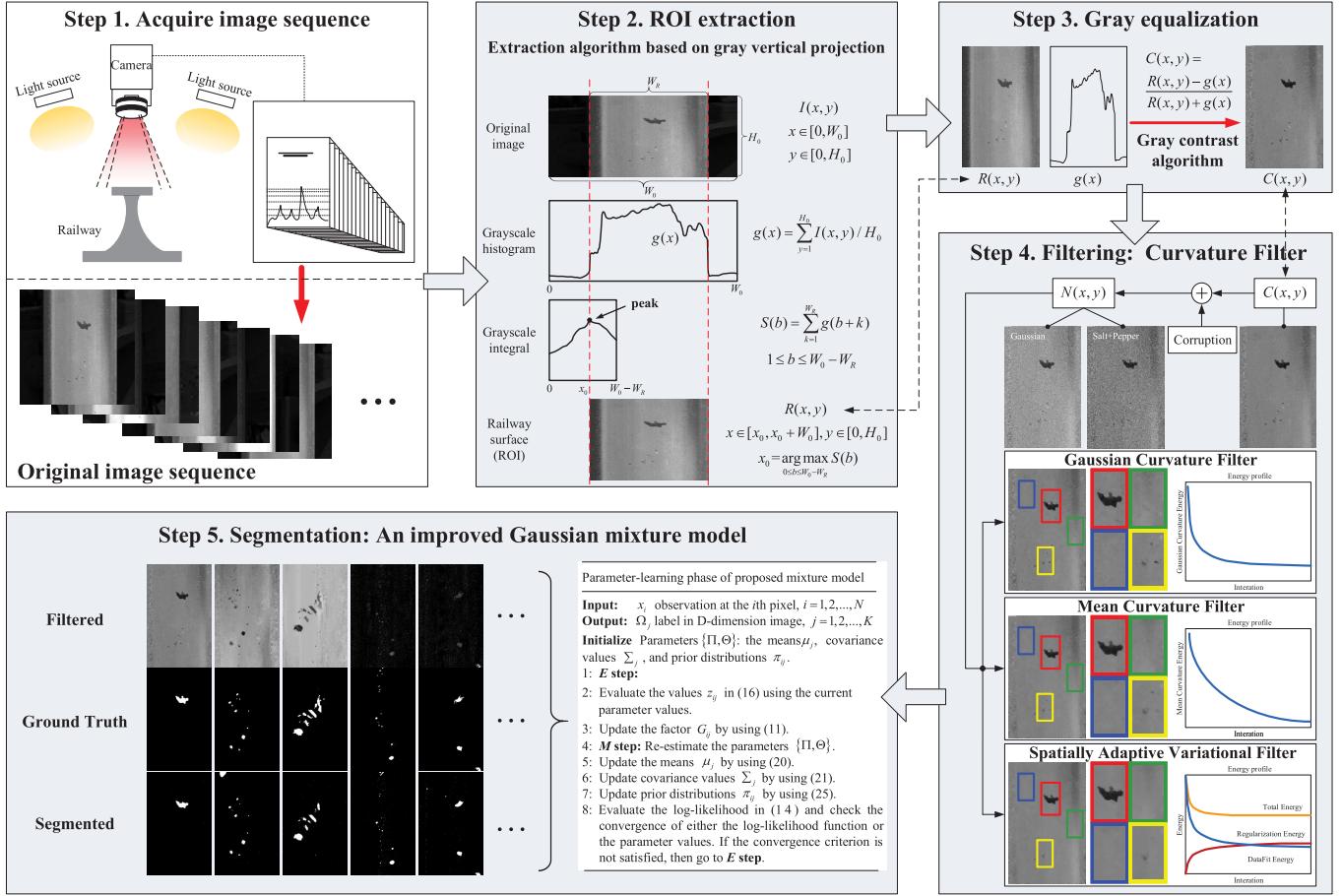


Fig. 3. Framework of the railway surface detection approach procedure.

out that it could implicitly optimize energy instead of directly calculating curvature, as well as maintain the image gradient and edge details [41].

Compared with the curvature-based geometric flow, curvature filter contributes to avoiding the direct calculation for curvature and other complex problems, like numerical stability, step length, time discretization, and so on. Thus, it would compute less, converge faster, and realize more easily [42], [43].

*1) Optimize Energy Functional:* From variational regularization methods, we learn that an energy functional should be defined for an image processing problem. Moreover, when this energy functional is smaller, the variable is going to be closer to what we expect. Thus, in every variational formulation, one aims to find a minimizing function  $\bar{U} = \arg \min_U \varepsilon(U)$  to an energy functional

$$\varepsilon(U) = \varepsilon_d(U, I) + \lambda \varepsilon_r(U). \quad (1)$$

Fig. 4 shows the evolution of energy functional when solving a variational model. In this case, data fit energy has been increasing while regularization energy and overall energy have been decreasing. It is proved that the regularization energy takes a central role in the optimization process. From the earlier deduction, what we have to do first is to minimize the regularization energy and then decide whether to converge or continue to iterate by data fit. In this paper, Gaussian

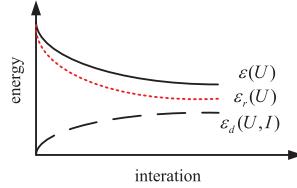


Fig. 4. Energy functional evolution.

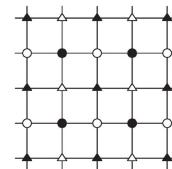


Fig. 5. Illustration of domain decomposition.

curvature filter will be introduced as follows, which is proved to be a filtering algorithm with computational efficiency and data locality.

*2) Domain Decomposition:* What the Gaussian curvature filter do first is to decompose the discrete domain. There is dependence among the adjacent pixels, blocking the principal curvature to be minimized locally. To avoid this problem, we propose a domain decomposition algorithm shown in Fig. 5. There are four subsets decomposed from the discrete domain of original image  $U$  such as white triangle, white circle, black triangle, and black circle. In this way, our domain

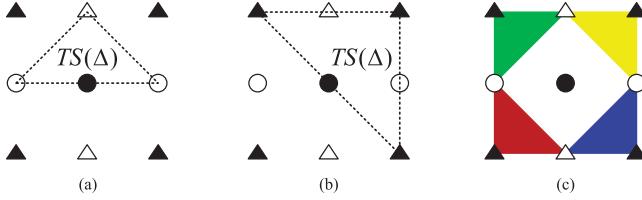
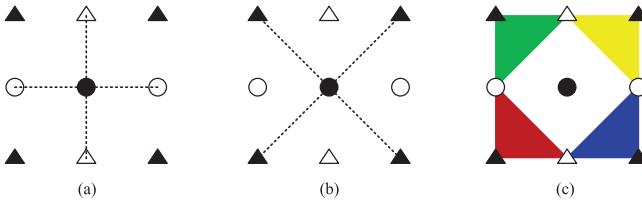


Fig. 6. 12 possible tangent planes for  $N(x)$ . (a) Example TS connected with white  $N(x)$ . (b) Example TS connected with black  $N(x)$ . (c) Example TS connected with the mixture  $N(x)$ .



decomposition could: 1) eliminate dependence among adjacent pixels and accelerate filtering; 2) coverage easily due to pixels independence; and 3) enumerate all tangent planes in a  $3 \times 3$  local window.

3) *Tangent Plane Projection*: If one pixel is defined as  $x$ , the pixel value is defined as  $U(x)$ . A new surface could be acquired through projecting from the current pixel value  $U(x)$  to a new pixel value  $\bar{U}(x)$ . Note that  $\bar{U}(x)$  is on the optimal tangent plane TS of adjacent pixels  $N(x)$ . In a word,  $U(x)$  and  $\bar{U}(x)$  would meet the relationship as follows:

$$\bar{U}(x) = U(x) + d. \quad (2)$$

As shown in Fig. 6, we enumerate all possible triangular tangent planes for  $N(x)$  in a  $3 \times 3$  window. Note that these planes would not use the current pixel  $x$  (center point) as a vertex. Among them, four planes connect with white  $N(x)$ , as shown in Fig. 6(a), four planes connect with black  $N(x)$ , as shown in Fig. 6(b), and four planes connect with the mixture  $N(x)$ , as shown in Fig. 6(c).

For the 12 triangular tangent planes listed in Fig. 6, we know that there are only eight  $d$ , namely, two  $d_i$  for  $TS(W)$ , as shown in Fig. 7(a), two  $d_i$  for  $TS(B)$ , as shown in Fig. 7(b), and four  $d_i$  for mixture, as shown in Fig. 7(c).

4) *Minimum Projection Operator*: The next step is to consider which  $d_i$  to choose. Learning from Euler's theorem, we could see

$$d_i \approx k_1 \cos 2\theta_i + k_2 \sin 2\theta_i. \quad (3)$$

Note that  $k_{1,2}$  is the principal curvature and  $\theta_i$  is the angle toward the principal plane. If  $k_1 k_2 \geq 0$  then  $|d_m| \approx \min\{|k_i|\}$  under the condition  $\theta_i \in (-\pi, \pi)$ . Thus, the minimum value  $d_m = \min\{|d_i|, i = 1, \dots, 8\}$  is the best projection from  $U(x)$  to  $\bar{U}(x)$ . In the end, a new surface would be acquired by  $\bar{U}(x) = U(x) + d_m$ . To conclude, we define a minimum projection operator  $P_g$  and Algorithm 1 to normalize the statement above. Moreover, the pseudocode is in the following.

5) *Gaussian Curvature Filter*: The final step is to gain  $\bar{U}(x)$  for each pixel  $W_T$ ,  $W_C$ ,  $B_T$ , and  $B_C$ . Similarly, we define a

---

**Algorithm 1** Minimal Projection Operator  $P_g$ 


---

```

Require:  $U(i, j)$ 
1:  $d_1 = (U(i - 1, j) + U(i + 1, j))/2 - U(i, j)$ 
2:  $d_2 = (U(i, j - 1) + U(i, j + 1))/2 - U(i, j)$ 
3:  $d_3 = (U(i - 1, j - 1) + U(i + 1, j + 1))/2 - U(i, j)$ 
4:  $d_4 = (U(i - 1, j + 1) + U(i + 1, j - 1))/2 - U(i, j)$ 
5:  $d_5 = U(i - 1, j) + U(i, j - 1)U(i - 1, j - 1)U(i, j)$ 
6:  $d_6 = U(i - 1, j) + U(i, j + 1)U(i - 1, j + 1)U(i, j)$ 
7:  $d_7 = U(i, j - 1) + U(i + 1, j)U(i + 1, j - 1)U(i, j)$ 
8:  $d_8 = U(i, j + 1) + U(i + 1, j)U(i + 1, j + 1)U(i, j)$ 
9: find  $d_m$ , such that  $d_m \min\{|d_i|, i = 1, \dots, 8\}$ 
Output:  $\bar{U}(i, j)U(i, j) + d_m$ 

```

---

**Algorithm 2** Gaussian Curvature Filter  $G_c$ 


---

```

Require:  $U(i, j)$ 
1:  $\forall x \in B_T, P_g(U(x))$ 
2:  $\forall x \in B_C, P_g(U(x))$ 
3:  $\forall x \in W_T, P_g(U(x))$ 
4:  $\forall x \in W_C, P_g(U(x))$ 
Output:  $\bar{U}(i, j)$ 

```

---

Gaussian curvature operator  $G_C$  and Algorithm 2 to normalize statement. Moreover, the pseudocode is in the following.

### B. Improved Gaussian Mixture Model Based on MRF

There are many uncertain descriptions in a rail image such as reflection, shadow, and noise, which cause a great difficulty in image segmentation. As a classical approach, the GMM is widely used due to its simplicity and ease of implementation. However, since all the pixels are assumed to be distributed independently, the prior distribution of the GMM does not depend on the spatial relationship between pixels and neighboring pixels, making it much more sensitive for noise or illumination. MRF is also a successful graph model, which can effectively reduce interference from noise. Compared with the GMM, the MRF mixture model is much hard to calculate mainly because the M step of expectation–maximization (EM) algorithm cannot directly estimate the prior distribution for maximization of log-likelihood function.

This paper proposes a new mixture model, which is improved by the GMM and MRF. We incorporate the spatial relationship between neighboring pixels in a much simpler standard, making it faster and easier to implement. In addition, our method could directly apply the EM algorithm for parameter optimization and is robust toward noise, as well as accurate and efficient.

1) *Theoretical Basis*: Before the new mixture model is introduced, a description of the traditional MRF mixture model for image segmentation is indispensable, which is also the theoretical basis for our new method.

Let  $x_i, i = (1, 2, \dots, N)$  denote an observation at the  $i$ th pixel in a D-dimension image. Then,  $\hat{o}_i$  denotes the neighborhood of the  $i$ th pixel. Moreover,  $(\Omega_1, \Omega_2, \dots, \Omega_K)$  would denote the labels. To divide the image composed of  $N$  pixels into  $K$  labels, the GMM in [44] and [45] assumes

that each observation  $x_i$  is independent of the label  $\Omega_j$ . So the density function  $f(x_i|\Pi, \Theta)$  at an observation  $x_i$  would be defined as

$$f(x_i|\Pi, \Theta) = \sum_{j=1}^K \pi_{ij} \phi(x_i|\Theta_j) \quad (4)$$

where  $\Pi = \{\pi_{ij}\}, i = (1, 2, \dots, N), j = (1, 2, \dots, K)$  is the data set of prior distribution modeling the probability that pixel  $x_i$  belongs to label  $\Omega_j$ , which satisfies the following constraints:

$$0 \leq \pi_{ij} \leq 1 \text{ and } \sum_{j=1}^K \pi_{ij} = 1. \quad (5)$$

Moreover, Gaussian distribution is given by  $\phi(x_i|\Theta_j)$ , also called a component of the mixture. For each Gaussian distribution, there is a uniform form of expression

$$\begin{aligned} & \phi(x_i|\Theta_j) \\ &= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma_j|^{1/2}} \exp \left\{ -\frac{1}{2} (x_i - \mu_j)^T \Sigma_j^{-1} (x_i - \mu_j) \right\} \end{aligned} \quad (6)$$

where  $\Theta_j = \{\mu_j, \Sigma_j\}, j = (1, 2, \dots, K)$ . The D-dimension vector  $\mu_j$  is the mean, the  $D \times D$  matrix  $\Sigma_j$  is the covariance, and  $|\Sigma_j|$  is the determinant of  $\Sigma_j$ . Note that the observation  $x_i$  in formula (4) is modeled statistically independent. Thus, the joint conditional density of the data set  $X = (x_1, x_2, \dots, x_N)$  could be modeled as

$$p(X|\Pi, \Theta) = \prod_{i=1}^N f(x_i|\Pi, \Theta) = \prod_{i=1}^N \sum_{j=1}^K \pi_{ij} \phi(x_i|\Theta_j). \quad (7)$$

Since the observation  $x$  is considered to be independent of pixel label, the spatial correlation between adjacent pixels would not be considered, making the segmented image much sensitive for noise and illumination [45]. To overcome this problem, the MRF distribution [46] is used to incorporate the spatial correlation among label values

$$p(\Pi) = Z^{-1} \exp \left\{ -\frac{1}{T} U(\Pi) \right\} \quad (8)$$

where  $Z$  is a normal constant,  $T$  is a temperature constant, and  $U(\Pi)$  is a smoothing prior. Then, the posterior probability density function based on Bayes' rules is given by

$$p(\Pi, \Theta|X) \propto p(X|\Pi, \Theta)p(\Pi). \quad (9)$$

We incorporate the formula (6), and log-likelihood function can be derived as

$$\begin{aligned} L(\Pi, \Theta|X) &= \log(p(\Pi, \Theta|X)) \\ &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} \phi(x_i|\Theta_j) \right\} + \log p(\Pi) \\ &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij} \phi(x_i|\Theta_j) \right\} - \log Z - \frac{1}{T} U(\Pi). \end{aligned} \quad (10)$$

Depending on the type of energy  $U(\Pi)$  in formula (10), there are different kinds of models with different performances and effects. In the Bayesian autologistic model [46], [47], the smoothing prior  $U(\Pi)$  has a simple form, and thus it is easy to optimize the parameter set  $\{\Pi, \Theta\}$  to maximize the log-likelihood function. However, one of its main drawbacks is that the segmentation result is not robust to noise. Other mixture models based on MRF have been successfully applied to image segmentation and different ways are adopted to select the energy  $U(\Pi)$ . For more details regarding these smoothing priors  $U(\Pi)$  and mixture models, we refer interested readers to [45] and [48]–[50].

As is shown above, the MRF distribution and various selected  $U(\Pi)$  are efficient to incorporate local information as well as to add complexity. An iterative EM algorithm is applied to maximize the likelihood of parameters  $\Pi$  and  $\Theta$  in (10). However, due to the complexity of the log-likelihood function, the M-step of EM algorithm cannot be directly applied to prior distribution  $\pi_{ij}$ . Note that the prior distribution  $\pi_{ij}$  should satisfy the constraint as formula (5). As a result, the final algorithm is computationally complex and tremendous computing cost is expended to solve the constrained optimization problem of prior distribution  $\pi_{ij}$ . For details regarding the maximization of the log-likelihood function in (10), we refer the readers to [45], [48], and [49].

**2) Proposed Algorithm:** The ways to derive smoothing prior  $U(\Pi)$  differ considerably for various mixture models. The parameter set  $\{\Pi, \Theta\}$  could be optimized to maximize the log-likelihood function in [46]. However, one of its main drawbacks is that the segmentation result is not robust to noise. A complex smoothing prior has been applied to the model in [45] and [48]–[50]. What the primary disadvantage lies in is its additional training complexity. Specifically, the M-step of EM algorithm cannot be directly applied to the prior distribution, increasing the complexity of the algorithm as a result. To overcome this drawback, a novel factor  $G_{ij}$  is introduced as

$$G_{ij}^{(t)} = \exp \left[ \frac{\beta}{2N_i} \sum_{m \in \partial_i} (z_{mj}^{(t)} + \pi_{mj}^{(t)}) \right] \quad (11)$$

where  $\beta$  is the temperature parameter for smoothing prior control,  $\partial_i$  is the neighborhood of the  $i$ th pixel including itself,  $N_i$  is the total number in neighborhood  $\partial_i$ ,  $z_{mj}$  is the posterior probability,  $\pi_{mj}$  is the prior distribution, and  $t$  is the iteration step. In this paper,  $\beta$  has been set to 12 ( $\beta = 12$ ). A square window of size  $5 \times 5$  is used and thus  $N_i$  is equal to 25 ( $N_i = 25$ ). By taking a closer look at (11), it can be visualized that the factor  $G_{ij}$  is defined as a multiplication of both posterior probabilities and prior distributions. Based on a fact that neighboring pixels in an image are similar in some sense, we can use this kind of relationship by replacing each posterior probability  $z_{ij}$  and posterior probability  $\pi_{ij}$  in an image with the average value of their neighbors, including themselves. Note that the factor  $G_{ij}$  only depends on the value of priors and posteriors at the previous step (at the  $t$  step). It plays a role as a linear filter for smoothing and restoring images corrupted by noise. For this reason, one of the main

advantages is the ease of implementation and incorporation of the spatial relationships among neighboring pixels in a simpler metric.

Next, we propose a novel approach to incorporate spatial information into a smoothing prior. The new smoothing prior  $U(\Pi)$  is given by

$$U(\Pi) = - \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \quad (12)$$

where  $N$  is the total number of pixels in D-dimension image,  $K$  is the total number of labels, and  $t$  is the iteration step. In formula (12), the derivative of smoothing prior  $U(\Pi)$  concerning prior distribution  $\pi_{ij}^{(t+1)}$  at current step (at  $t + 1$  step) only depends on the term  $\pi_{ij}^{(t+1)}$ . Thus, for our approach, the M-step of EM algorithm is simple and computationally efficient. The MRF distribution  $p(\Pi)$  in formula (8) is given by

$$p(\Pi) = Z^{-1} \exp \left\{ \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \right\} \quad (13)$$

where  $Z$  is a normalizing constant and  $T$  is a temperature constant.

Having the MRF distribution, the log-likelihood function in formula (10) can be written as

$$\begin{aligned} L(\Pi, \Theta|X) &= \sum_{i=1}^N \log \left\{ \sum_{j=1}^K \pi_{ij}^{(t+1)} \phi(x_i|\Theta_j^{(t+1)}) \right\} \\ &\quad - \log Z + \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \end{aligned} \quad (14)$$

where the log-likelihood function  $L(\Pi, \Theta|X)$  is based on Bayes rules in (9),  $\phi(x_i|\Theta_j^{(t+1)})$  is the Gaussian distribution, also called a component of the mixture, observation  $x_i$  is the  $i$ th pixel in image, and  $\Theta_j^{(t+1)}$  is a parameter set.

According to the appliance for complete data condition in [45], maximizing the log-likelihood function  $L(\Pi, \Theta|X)$  in formula (14) would increase the value of objective function  $J(\Pi, \Theta|X)$

$$\begin{aligned} J(\Pi, \Theta|X) &= \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \{ \log \pi_{ij}^{(t+1)} + \log \phi(x_i|\Theta_j^{(t+1)}) \} \\ &\quad - \log Z + \frac{1}{T} \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (15)$$

Then, conditional expectation values  $z_{ij}$  of the hidden variables can be given by

$$z_{ij}^{(t)} = \frac{\pi_{ij}^{(t)} \phi(x_i|\Theta_j^{(t)})}{\sum_{k=1}^K \pi_{ik}^{(t)} \phi(x_i|\Theta_k^{(t)})}. \quad (16)$$

The next objective is to optimize the parameters' set  $\{\Pi, \Theta\}$  for the maximum of objective function  $J(\Pi, \Theta|X)$  in formula (15). Similar to the MRF-based method,  $Z$  and  $T$  in

formula (15) are set as 1 ( $Z = 1$  and  $T = 1$ ). As a result, a novel objective function is given as

$$\begin{aligned} J(\Pi, \Theta|X) &= \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \{ \log \pi_{ij}^{(t+1)} + \log \phi(x_i|\Theta_j^{(t+1)}) \} \\ &\quad + \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)}. \end{aligned} \quad (17)$$

According to formula (6), (17) could be rewritten as

$$\begin{aligned} J(\Pi, \Theta|X) &= \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ \log \pi_{ij}^{(t+1)} - \frac{D}{2} \log(2\pi) - \frac{1}{2} \log |\Sigma_j^{(t+1)}| \right\} \\ &\quad + \sum_{i=1}^N \sum_{j=1}^K z_{ij}^{(t)} \left\{ -\frac{1}{2} (x_i - \mu_j^{(t+1)})^T \Sigma_j^{(t+1)} (x_i - \mu_j^{(t+1)}) \right\} \\ &\quad + \sum_{i=1}^N \sum_{j=1}^K G_{ij}^{(t)} \log \pi_{ij}^{(t+1)} \end{aligned} \quad (18)$$

where the D-dimension vector  $\mu_j$  is the mean, the  $D \times D$  matrix  $\Sigma_j$  is the covariance,  $|\Sigma_j|$  is the determinant of  $\Sigma_j$ , and the parameter set  $\Theta_j = \{\mu_j, \Sigma_j\}$ ,  $j = (1, 2, \dots, K)$ .

To maximize this function, the EM algorithm is applied. Then, we consider the derivation of function  $J(\Pi, \Theta|X)$  with respect to  $\mu_j$  at the  $(t + 1)$  interation step, as shown by

$$\frac{\partial J}{\partial \mu_j^{(t+1)}} = \sum_{i=1}^N z_{ij}^{(t)} \left[ -\frac{1}{2} (2\Sigma_j^{-1(t+1)} - 2\Sigma_j^{-1(t+1)} x_i) \right] \quad (19)$$

where  $J$  represents the objective function  $J(\Pi, \Theta|X)$ .

When  $(\partial J / \partial \mu_j) = 0$ , the minimum of  $\mu_j$  at  $(t + 1)$  interation step is given by

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} x_i}{\sum_{i=1}^N z_{ij}^{(t)}}. \quad (20)$$

Thus, setting the derivative of the function in (17) with respect to  $\Sigma_j^{-1}$  at the  $(t + 1)$  iteration step, we have  $(\partial J / \partial \Sigma_j^{-1(t+1)}) = \sum_{i=1}^N z_{ij}^{(t)} [(1/2)\Sigma_j^{(t+1)} - (1/2)(x_i - \mu_j^{(t+1)})(x_i - \mu_j^{(t+1)})^T]$ . Moreover, equating it to zero yields

$$\Sigma_j^{(t+1)} = \frac{\sum_{i=1}^N z_{ij}^{(t)} (x_i - \mu_j^{(t+1)})(x_i - \mu_j^{(t+1)})^T}{\sum_{i=1}^N z_{ij}^{(t)}}. \quad (21)$$

An important consideration is that the prior distribution should satisfy the constraints in (5). To enforce these constraints, we use the Lagrange multiplier  $\eta_i$  for each data point

$$\frac{\partial}{\partial \pi_{ij}^{(t+1)}} \left[ J - \sum_{i=1}^N \eta_i \left( \sum_{j=1}^K \pi_{ij}^{(t+1)} - 1 \right) \right] = 0 \quad (22)$$

where  $J$  represents the objective function  $J(\Pi, \Theta|X)$  and  $\eta_i$  is a Lagrange multiplier.

The formula (22) can be rewritten in the following form:

$$\frac{z_{ij}^{(t)}}{\pi_{ij}^{(t+1)}} + \frac{G_{ij}^{(t)}}{\pi_{ij}^{(t+1)}} - \eta_i = 0. \quad (23)$$

The constraint  $\sum_{j=1}^K \pi_{ij} = 1$  enables the Lagrange multiplier  $\eta_i$  to satisfy the following condition:

$$\eta_i = 1 + \sum_{j=1}^K G_{ij}^{(t)}. \quad (24)$$

The necessary condition for determining the prior distribution  $\pi_{ij}$  at the  $(t+1)$  iteration step becomes

$$\pi_{ij}^{(t+1)} = \frac{z_{ij}^{(t)} + G_{ij}^{(t)}}{\sum_{k=1}^K (z_{ik}^{(t)} + G_{ik}^{(t)})}. \quad (25)$$

So far, the discussion has focused on estimating  $\{\Pi, \Theta\}$  of the model to assign a label  $\Omega_j$  to the pixel  $x_i$ . The various steps of the proposed mixture model incorporating spatial information based on MRF can be summarized as Algorithm 3.

---

**Algorithm 3** Parameter-Learning Phase of the Proposed Mixture Model

---

**Input:**  $x_i$  observation at the  $i$ th pixel,  $i = 1, 2, \dots, N$

**Output:**  $\Omega_j$  label in D-dimension image,  $j = 1, 2, \dots, K$

**Initialize** Parameters  $\{\Pi, \Theta\}$ : the means  $\mu_j$ , covariance values  $\sigma_j$ , and prior distributions  $\pi_{ij}$

1: **E step:**

2: Evaluate the values  $z_{ij}$  in (16) using the current parameter values.

3: Update the factor  $G_{ij}$  by using (11).

4: **M step:** Re-estimate the parameters  $\{\Pi, \Theta\}$ .

5: Update the means  $\mu_j$  by using (20).

6: Update covariance values  $\sigma_j$  by using (21).

7: Update prior distributions  $\pi_{ij}$  by using (25).

8: Evaluate the log-likelihood in (14) and check the convergence of either the log-likelihood function or the parameter values. If the convergence criterion is not satisfied, then go to **E step**.

---

Once the parameter-learning phase is complete, every pixel  $x_i$  is assigned to the label with the largest posterior probability  $z_{ij}$

$$x_i \in \Omega_j : \text{IF } z_{ij} \geq z_{ik}; \quad j, k = (1, 2, \dots, K) \quad (26)$$

where  $\Omega_j$  is the labels in the D-dimension image and  $K$  is the total number of labels.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

To evaluate the performance of the proposed algorithm, it was implemented in the MATLAB R2017a programming environment on a PC with Intel(R) Core(TM) i5-7200 CPU at 2.50-GHz running Windows 10. All images used in this experiment are obtained by the RVDS described in Section II-A in actual Hunan province railways. The system adopts a DalsaSpyder3 line-scan camera with a resolution of 1024 pixels for image capture. Its maximum line rate is 68 000 lines/s, which can capture a set of frames and finally generates a panorama. High-scanning frequencies of the line-scan camera allow resolutions of currently  $1 \text{ mm} \times 0.15 \text{ mm}$  at a traveling speed of 240 km/h. A data set with images containing at least one defect is constructed to evaluate performance [30], [51].

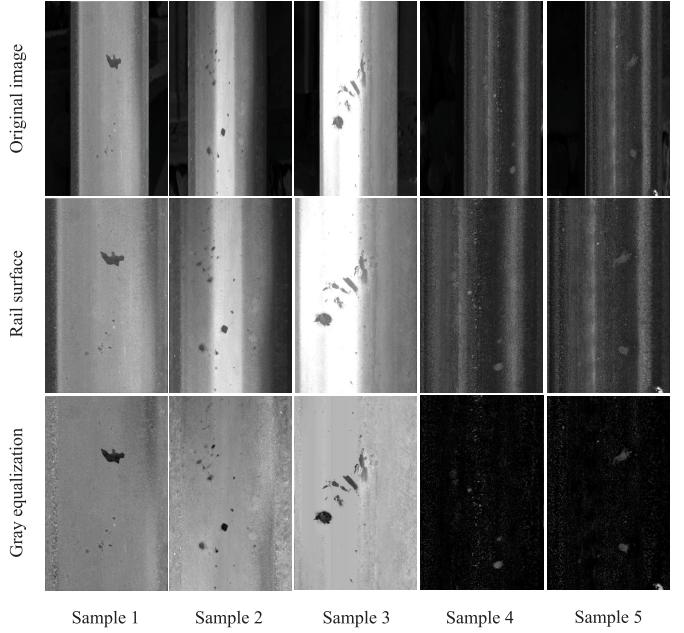


Fig. 8. Preprocessing results of different samples.

#### A. Results of Image Preprocessing

1) **ROI Extraction:** The original image acquired by the camera (for a rail scene image) also includes other backgrounds in addition to the rail. To reduce the detection range and avoid the redundant operation of the following processing, we conduct an extraction algorithm based on the gray vertical projection, extract the valuable rail area  $R(x, y)$  from the scene image  $I(x, y)$ , and form the detection ROI area.

The average gray value of the input image  $I(x, y)$  in the  $x$ th column is calculated as

$$g(x) = \sum_{y=1}^{H_0} I(x, y) / H_0 \quad (27)$$

where  $H_0$  is the height of the image and  $I(x, y)$  is the grayscale value of input image  $I$  at the location  $(x, y)$ .

Then,  $G = [g(1), g(2), \dots, g(W_0)]$  is the projection of the input image  $I$  in the vertical direction.

Assuming the rail width is  $W_R$ , the grayscale integral of each width in the projection curve along the  $x$ -axis is the integral of the interval  $S(b)$

$$S(b) = \sum_{k=1}^{W_R} g(b+k) \quad (28)$$

where  $1 \leq b \leq W_0 - W_R$ . The coordinates  $b$  of the maximum value of  $S(b)$  on the  $x$ -axis are taken as the starting position of the rail on the  $x$ -axis. Moreover, the end position  $b + W_R$  is taken as the end position so that the subimage between the sections  $[b, b + W_R]$  on the  $x$ -axis is cut out from the input image  $I$ , that is the detect rail region  $R$ .

2) **Gray Equalization:** Observing the rail area  $R$ , we found that surface defects are submerged in a large volume of noise. Note that the noise is generated by the compound conditions such as natural and nonnatural light, train mechanical

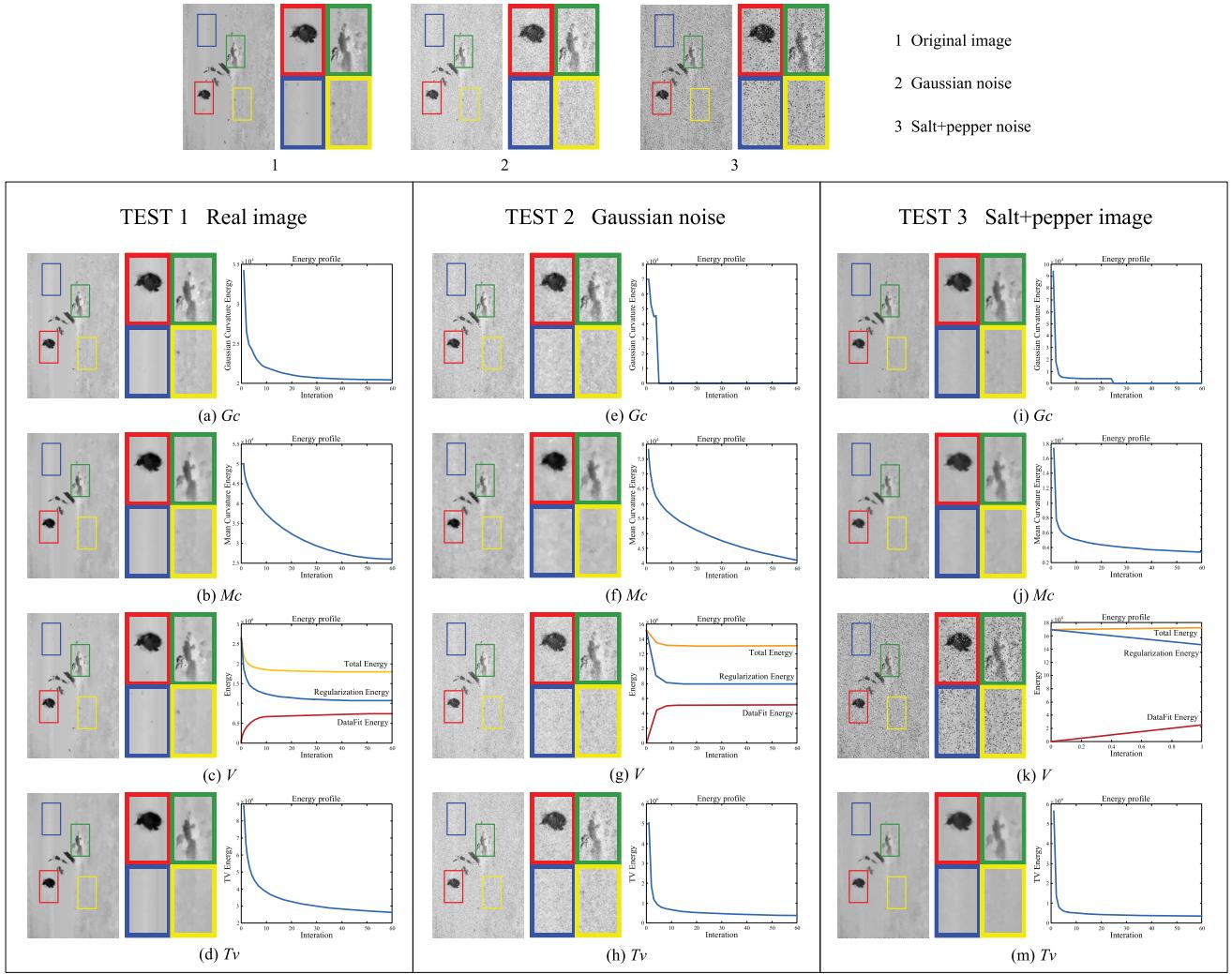


Fig. 9. Illustration of curvature filters for different noise types. (a)–(d) Filtering results for original image with Gc, Mc, V, and Tv, respectively. (e)–(h) Filtering results for Gaussian noise image with Gc, Mc, V, and Tv, respectively. (i)–(m) Filtering results for Salt + pepper noise image with Gc, Mc, V, and Tv, respectively.

vibration, rail geometry, and orbital reflector material. These corruptions are mainly reflected on the rail image, where the steel surface appears alternately as light and dark, resulting in defects difficult to distinguish from the railroad background. Therefore, in order to eliminate the interference of inherent noise, the second step of image preprocess is gray equalization for rail region R, making the defect more obvious on the surface of railway. The rail area image R is converted to the rail gray contrast image C, and the grayscale equalization is achieved

$$C(x, y) = \frac{R(x, y) - g(x)}{R(x, y) + g(x)} \quad (29)$$

where  $R(x, y)$  is the detection rail area,  $C(x, y)$  is the contrast image after the equalization of grayscale, and  $g(x)$  is the average grayscale value of input image I in column  $x$ . In Fig. 8, the experimental results are shown for preprocessing the five sample images of different lighting conditions and obtaining the experimental data.

#### B. Image Denoising Results by Curvature Filter

Considering the versatility of the curvature filter, on the one hand, we filter the preprocessed image to remove the noise and keep the details [52], [53], while on the other hand, we maintain the visible surface of the overall image. The following test object is the real images without any noise and the synthesized images with artificial noise. The synthesized images are Gaussian noise images and salt and pepper noise images. We use three kinds of curvature filters, namely, Gaussian curvature filter (Gc), Mean curvature filter (Mc), and spatially adaptive variational filter (V), for processing and analysis of experimental results in detail. It should be noted that total variation (Tv) as a classic variational regularization method has, in essence, the same idea with the curvature filter. So in Sections IV-B1 and IV-B2, we use the Tv filter and three kinds of curvature filter in contrast to conducting a preliminary qualitative evaluation. Moreover, the more contrast experiment and more quantitative assessment are shown in Section IV-B3.

1) *Filter of the Natural Real Images:* We use the sample 3 for testing. Fig. 9(a)–(d) is the real image of the filtering results

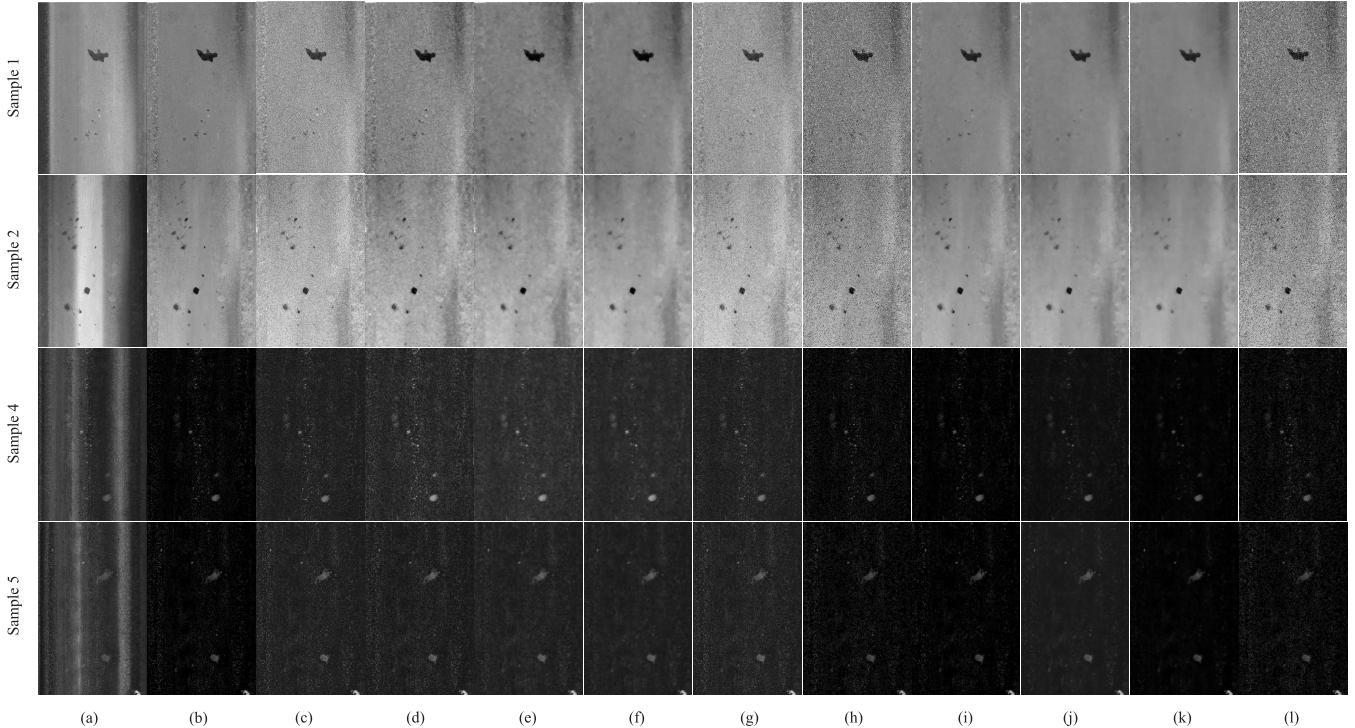


Fig. 10. Results of the curvature filter of different samples. (a) Original image. (b) Gray equalization. (c) Gaussian noise image. (d)–(g) Filtering results for Gaussian image with Gc, Mc, Tv, and V filters, respectively. (h) Salt + pepper noise image. (i)–(l) Filtering results for Salt + pepper image with Gc, Mc, Tv, and V filters, respectively.

and energy evolution. To facilitate observation, we magnify the image details in four images, namely, four typical representations: regular defects (red box), irregular defects (green box), rough rail surface (blue box), and smooth rail surface (yellow box). We find that: 1) although Gc, Mc, and V are not very different, corresponding Gaussian curvature energies are 246.98, 102.4, and 97.15, and thus the effectiveness of curvature filter can be seen and 2) comparing Gc with V, curvature filter can work alone or work with a data fitting (imaging model). It does not require a gradient to calculate the imaging model, which makes it very versatile.

*a) Gaussian noise image results:* Fig. 9(e)–(h) shows that all the four filters have a good capacity to identify and remove out Gaussian noise. However, the differences among them reflect in the following aspects. Gc as well as V enables to maintain more edges and details because they are designed to retain the developable surface of image. Mc, corresponding to retain the minimum surface of image, filters out noise and smooths image much aggressively, causing more edges of defect to blur and more details missing. In addition, Tv has the same effect as Mc.

The energy evolution in Fig. 9(e)–(h) shows that as the number of iterations increases, the four kinds of filters can effectively reduce the energy of the regular term and keep the convergence. The difference is that for Gc, Tv, and V after iteration of about 5 times, 45 times, and 3 times, the energy tends to be stable, indicating the convergence of Gc, Tv, and V is good, with convergence of V being the fastest. Mc, however, has not been fully converged in the entire experiment process.

In summary, different filters have different performances on Gaussian noise image. Gc and V maintain more edges and

details for defects when working because they are designed to retain the developable surface of image. Compared with other traditional methods, they are suitable for our real-time railway detection system, with the faster speed by 10 to 1000 times and a good convergence. Mc retains the minimum surface of image, thus focusing on smoothing image and removing noise as much as it can, but it cannot converge well and meet the demand for small defect detection.

*b) Salt and pepper noise image results:* Fig. 9(i)–(m) shows that: Gc, Mc, and Tv have a good salt and pepper noise recognition and filtering function. The difference is that Gc completely retains the visible surface and can restore more details at the same time. Because of the smallest curved surface, Mc smoothing effect is the strongest, and Tv is easy to make image blur. It is worth noting that the V result still has much noise, indicating that V cannot remove the salt and pepper noise.

The energy evolution in Fig. 9(i)–(m) shows that the four kinds of filters can effectively reduce the energy of the regular term and keep the convergence as the number of iterations increases. The difference is that for Gc and Tv, the energy tends to be stable after iterations of about 15 times and 50 times, indicating that convergence of Gc and Tv is good, while convergence of Gc is the fastest. However, during the whole experiment process, Mc has not been fully converged, whereas V cannot make the regular items and data achieve a balance, resulting in the overall energy rise and curve divergence, and thus convergence conditions are not established.

*2) Other Samples' Results:* In Fig. 10, we use the curvature filter to filter the other samples. The experimental results verify

TABLE I  
COMPARISON OF IMAGE FILTER RESULTS ON VARIOUS SAMPLES

Image	Parameter	WLS	Guided Filter	Bilateral Filter	BF	RTV	Gc	MC	V
Sample1	PSNR(dB)	19.59	19.62	17.57	<b>28.82</b>	25.67	<b>29.51</b>	<b>28.82</b>	<b>29.08</b>
	MSE	715.10	710.14	1.14E+03	85.23	176.41	72.75	85.23	80.34
	Time(s)	1.13	<b>0.22</b>	4.20	5.96	2.30	<b>0.01</b>	<b>0.41</b>	<b>0.26</b>
Sample2	PSNR(dB)	19.42	19.58	17.58	<b>27.46</b>	24.86	<b>27.30</b>	<b>27.46</b>	<b>26.86</b>
	MSE	742.39	716.13	1.14E+03	116.83	212.38	121.08	116.83	133.96
	Time(s)	1.05	<b>0.20</b>	2.73	2.37	2.30	<b>0.24</b>	<b>0.42</b>	<b>0.20</b>
Sample3	PSNR(dB)	19.47	19.81	17.86	<b>26.40</b>	20.97	<b>26.70</b>	<b>26.40</b>	<b>25.16</b>
	MSE	734.59	679.73	1.07E+03	149.00	520.65	139.16	149.00	198.29
	Time(s)	1.04	<b>0.12</b>	2.69	2.50	2.31	<b>0.01</b>	<b>0.41</b>	<b>0.30</b>
Sample4	PSNR(dB)	18.19	18.87	18.02	16.75	16.34	16.78	16.75	15.58
	MSE	986.02	843.94	1.03E+03	1.38E+03	1.51E+03	1.37E+03	1.38E+03	1.80E+03
	Time(s)	5.22	<b>0.28</b>	7.38	8.88	6.75	<b>0.04</b>	2.18	1.75
Sample5	PSNR(dB)	18.19	18.85	18.02	19.27	19.12	19.37	19.27	18.77
	MSE	986.94	846.72	1.03E+03	768.82	796.59	752.18	768.82	863.86
	Time(s)	5.54	<b>0.29</b>	7.62	8.65	6.63	<b>0.04</b>	2.15	1.76

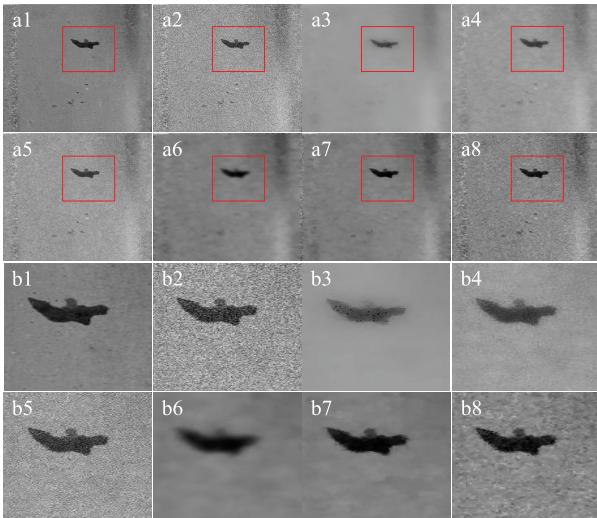


Fig. 11. Results of other filters in sample 1. (a1) Original image. (a2) Gaussian noise image. (a3)–(a8) Filtering results with WLS, guided filter, bilateral filter, BF, total variance (Tv), and our Gc filter, respectively. (b1)–(b8) Amplifications of the block portions in (a1)–(a8), respectively.

the correctness of the qualitative evaluation of each of the curvature filters.

3) *Performance Comparison and Statistics:* The test experiments in Section IV-B parts 1)-2) only verify the feasibility of the curvature filter in noise removal and edge preserving and cannot reflect its superiority. Now, we use several classic edge preserving filters in the sample 1 on the comparative experiment, which is shown in Fig. 11.

By contrast, we find that our Gc and other filters have good denoising capabilities, but other filters are more likely to lose the details of the image, i.e., poor edge retention. For example, weighted least squares (WLS) in Fig. 11(b3) appears smoothed, resulting in incomplete defects. Fig. 11(b4) and (b5) shows the guided filter and bilateral filter caused by the gray value of the defect changes, reducing the contrast between the defect and the background of the grayscale. In Fig. 11(b6), Bernstein filter (BF) produces a blurring phenomenon so that

the edge of the defect is not clear. The RTV in Fig. 11(b7) produces various ladder and glitch interferences at the edge of the defect. In contrast, in Fig. 11(b8), due to the implicit calculation of the curvature of the regular term and the retention of the surface, Gc preserves the edge of the defect, suppressing the ladder effect of interference. The contrast between the defect and the background reflects the robustness and accuracy of the curvature filter algorithm.

Next, samples 1–5 were tested using various filters. The experimental results are presented in Table I. In the table, we mark some of the better performance data, where the red marker data belong to our curvature filter and the green marker data belong to the other filters. We find the following.

- 1) Our filter in the various types of samples on the test to occupy most of the advantages of data.
- 2) In samples 1–3, due to retaining the image of the surface or the smallest surface, the SNR and peak SNR (PSNR) of our filter are higher than other filters. Because of the implicit calculation of the curvature of the regular term, the time cost of our filter is 10 to 1000 times less than most filters.
- 3) In samples 4 and 5, our filter loses the advantage over other filters and only Gc runs fast. In summary, due to the implicit calculation and the retention of the unique surface, our filter makes the denoising performance and image recovery capability better, the time cost and the convergence step are small, and the calculation speed is faster than other filters 1–3 orders of magnitude; reflecting the curvature filter algorithm is fast, accurate, and easy to implement. However, these advantages are not found in similar samples 4 and 5, whose overall grayscale values of images are minimal.

### C. Image Segmentation Results by Improved Gaussian Mixture Model Based on MRF

Defect detection difficult in Section I describes the VIS for the requirement of quickness, accuracy, and robustness, so we use a fast and robust space-constrained GMM for the purpose of image segmentation [54], [55].

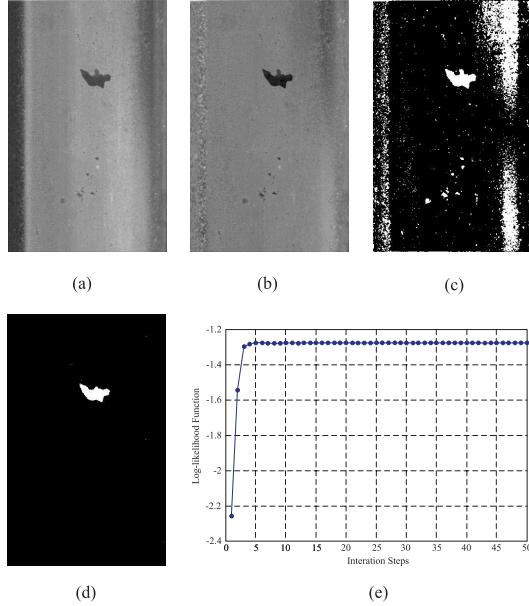


Fig. 12. Defect segmentation result of sample 1. (a) Original image. (b) Gray equalization. (c) Segmentation result with standard GMM. (d) Segmentation result with our proposed method. (e) Maximization progress of log-likelihood function in (d).

In the following experiment, we first use this method to segment the filtered image, and then test the temperature coefficient  $\beta$  of the proposed method, and finally compared with other segmentation methods. Our method uses the standard GMM to initialize, so Sections IV-C1–IV-C3 in the standard GMM with our method to form a qualitative evaluation comparison. However, the other contrast experiment and more intuitive quantitative assessment are shown in Sections IV-C4.

1) *Sample Segmentation Test*: Fig. 12 shows the segmentation result for sample 1. As shown in Fig. 12(a), the filtered image includes defects, rails, noise, and other information, and the purpose is to divide the image into two categories: defective area and nondefective areas, and the difficulty lies in the boundaries between the two classes. As shown in Fig. 12(b), although the standard GMM effectively classifies the defect, at the same time to separate a number of noise and part of the rail surface is difficult, standard GMM will cause serious misclassification, because it is sensitive to noise. In Fig. 12(c), our method has a good segmentation effect, accurately classifying the image into defective and nondefective classes and suppressing a significant amount of noise, indicating that the robustness to noise is excellent. Fig. 12(d) is the logarithmic probability maximization process; the method can converge quickly after five iterations. It shows that our method optimizes the parameters directly by using the EM algorithm, which is faster, more accurate, and easier to implement than other methods.

2) *Temperature Coefficient  $\beta$  Test*: Fig. 13 shows the segmentation result of our method under different temperature values  $\beta = 2, 4, 6, 8, 12, 16, 20$ . By taking a closer look, it can be visualized that when  $\beta$  is smaller, the sharpness and details of defects lose less, but a large volume of noise remains, which

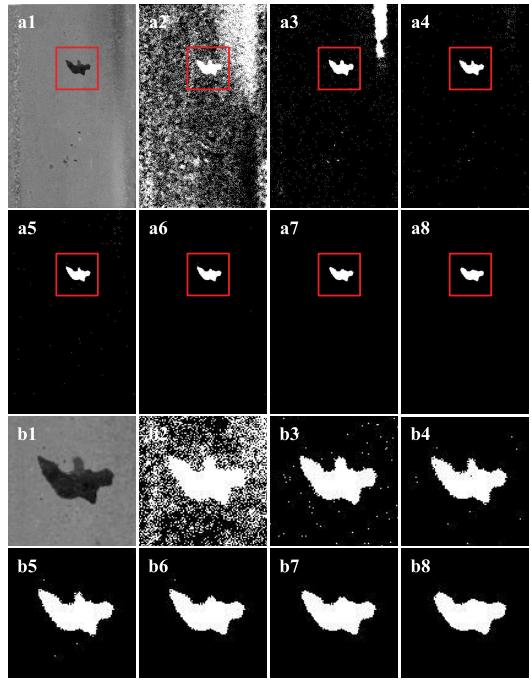


Fig. 13. Proposed method with different temperature values  $\beta$ . (a1) Original image. (a2)–(a8)  $\beta = 2$  (MCR = 67.43%),  $\beta = 4$  (MCR = 17.88%),  $\beta = 6$  (MCR = 5.39%),  $\beta = 8$  (MCR = 2.65%),  $\beta = 12$  (MCR = 1.86%),  $\beta = 16$  (MCR = 1.55%), and  $\beta = 20$  (MCR = 1.54%), respectively. (b1)–(b8) Magnifications of block portions in (a1)–(a8).

is shown in Fig. 13(a2), (a3), (b2), and (b3). Additional, our method is robust but easier to lose details in the case of a larger  $\beta$ , which is shown in Fig. 13(a8) and (b8). In a word,  $\beta$  is chosen small enough to prevent from losing much sharpness and details but misclassify, and  $\beta$  is chosen large enough to tolerate noise but result in a low accuracy.

3) *Test Results and Performance Comparison for Other Samples*: The test of Sections IV-C1 and IV-C2 verifies the feasibility of using the MRF GMM for defect segmentation but does not reflect the superiority of the algorithm. Therefore, we then use *K*-mean, GMM, FCM, FFCM, and Student-T MM classical clustering method. A comparative experiment was conducted on five samples, and we use the misclassification rate (MCR), Rand index (RI), PSNR, mean square error, variation of information (VI), time cost (Time), and global consistency error (GCE) to analyze of segmentation performance quantitatively.

Fig. 14 shows the results of the comparative experiment. Since our sample is an unmarked rail image, the ground truth is replaced by manual segmentation. Table II records the performance data for the various methods with samples 1–5. In Table II, we mark some of the better performance information, where the red marker data belong to our method and the green marker data belong to other methods. It is worth noting that when the MCR, RMSE, VI, and GCE are smaller, PSNR, NPR, and PRI are larger, indicating that the segmentation quality is better. In contrast, we find the following.

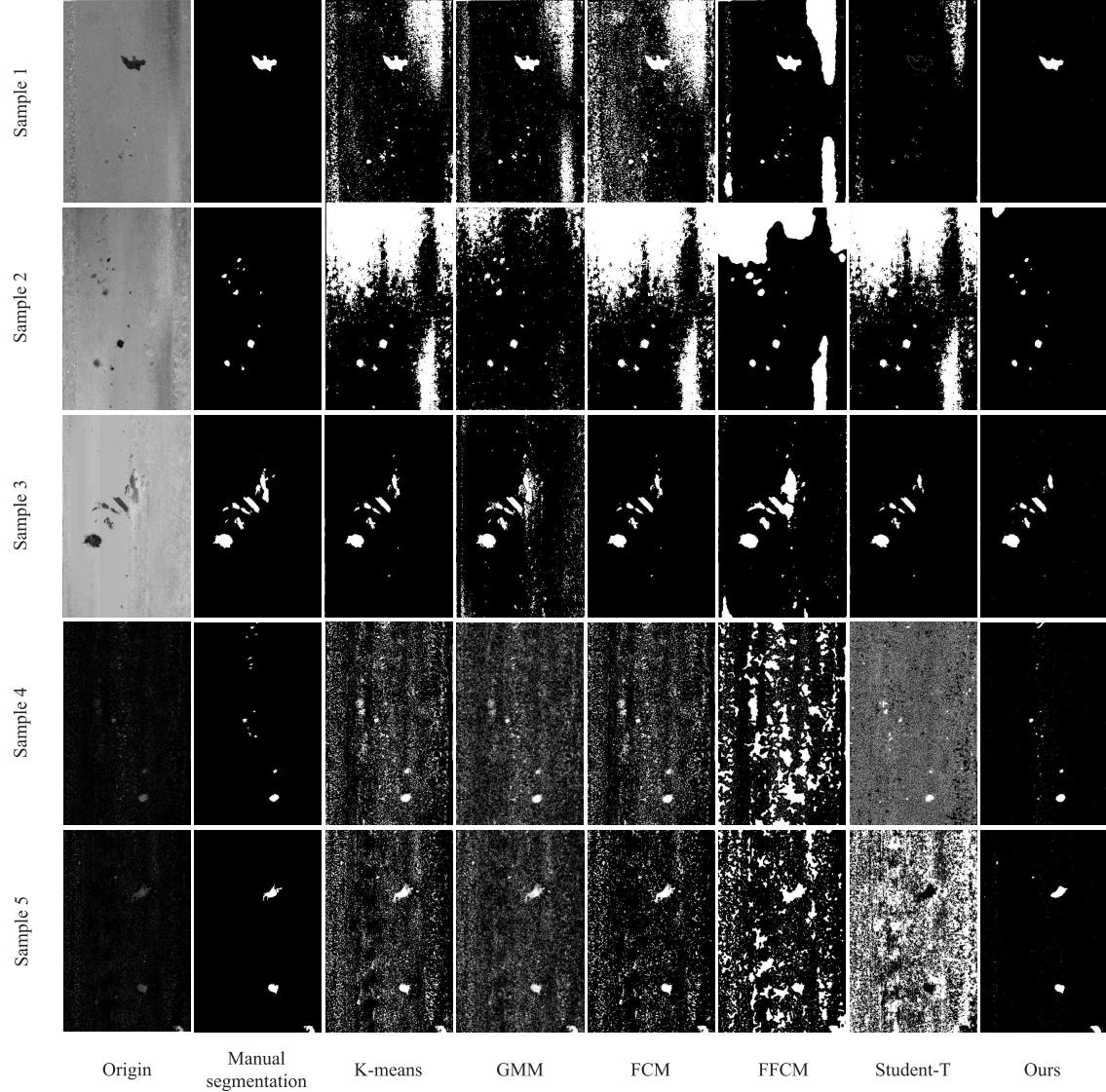


Fig. 14. Comparison results of defects detection using different methods.

- 1) Our method works well in all segmentation tests of various samples.
- 2) Regardless of the existence of special cases, segmentation results of *K*-means, GMM, FCM, FFCM, and Student-T show very serious misclassification and much residual noise, resulting in their MCR, VI, and GCE being very high and RI and PSNR being very low; in contrast, our method is effective.
- 3) Since our method can directly use the EM algorithm to optimize the logarithmic probability, the other methods, such as MCR, VI, and GCE are deficient, and RI and PSNR are very high.
- 4) *Quantitative Analysis of Defects Detection:* To more quantitatively compare the detection accuracy of the manual method [56], four batches of different railway image sequences were tested. In those images, defects contained line type and small-block defects, and so on with illumination variation. Among the detection results, true positive (TP), true negative (TN), false positive (FP), and false negative (FN) correspond to the count of defective unqualified, defect-free

qualified, defective qualified, and defect-free unqualified products, respectively. The precision (Pr), recall (Rc), and accuracy (Acc) are used to investigate the defect detection performance

$$\begin{aligned} \text{Pr} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{Rc} = \frac{\text{TP}}{\text{TP} + \text{FN}} \\ \text{Acc} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}. \end{aligned} \quad (30)$$

The defect detection performance is verified using the obtained experimental data in Sections IV-A and IV-B, and the quantitative indexes are listed in Table III with each row corresponding to different defects types and illumination. We have achieved a very good performance with 92% recall and 88.8% precision rates. As evident, most oblong defects similar to sample 1 are easily detected with the proposed algorithm, even though some exceptions might happen to tiny and ambiguous defects or the very noisy interferences are occasionally discriminated as defects. Moreover, the algorithm has achieved the best performance for block-type defects with 92.4% precision and 87.6% recall rates. The errors are mainly due to the largely varying gray-level distribution of

TABLE II  
QUANTITATIVE COMPARISON OF IMAGE SEGMENTATION RESULTS ON VARIOUS SAMPLES

Image	Algorithm	MCR(%)	RI(%)	PSNR(dB)	MSE	VI(pixel)	Time(s)	GCE(%)
Sample 1	K-means	35.48	45.04	8.66	8.85E+03	2.63	7.68	2.96
	GMM	28.54	54.03	9.53	7.24E+03	2.24	<b>0.12</b>	2.94
	FCM	52.36	28.13	6.03	1.62E+04	3.41	15.29	2.99
	FFCM	16.75	72.29	9.50	7.30E+03	1.19	10.61	2.94
	Student-T	<b>7.71</b>	<b>86.17</b>	<b>14.70</b>	2.21E+03	0.81	5.26	2.98
Sample 2	Ours	<b>1.64</b>	<b>98.43</b>	<b>29.80</b>	68.15	<b>0.15</b>	<b>0.13</b>	<b>1.04</b>
	K-means	49.17	37.91	4.19	2.48E+04	2.49	6.61	3.09
	GMM	19.13	67.81	<b>11.47</b>	4.63E+03	1.68	<b>0.13</b>	3.05
	FCM	51.07	37.13	3.98	2.60E+04	2.52	13.96	3.09
	FFCM	24.25	63.35	6.73	1.38E+04	1.26	17.50	3.01
Sample 3	Student-T	46.93	39.46	4.40	2.36E+04	2.42	9.78	3.09
	Ours	<b>2.12</b>	<b>97.50</b>	<b>23.71</b>	276.49	<b>0.24</b>	<b>0.22</b>	<b>2.28</b>
	K-means	<b>5.55</b>	<b>94.02</b>	<b>22.34</b>	379.09	<b>0.56</b>	6.60	<b>4.18</b>
	GMM	20.73	70.26	14.75	2.18E+03	1.87	<b>0.12</b>	8.02
	FCM	<b>5.56</b>	<b>94.01</b>	<b>22.25</b>	387.01	<b>0.56</b>	16.49	<b>4.19</b>
Sample 4	FFCM	<b>9.90</b>	<b>88.01</b>	15.93	1.66E+03	0.97	12.17	7.44
	Student-T	<b>5.32</b>	<b>93.84</b>	<b>21.04</b>	511.97	<b>0.54</b>	6.84	<b>3.58</b>
	Ours	<b>5.45</b>	<b>93.77</b>	<b>21.29</b>	482.98	<b>0.56</b>	<b>0.25</b>	<b>3.96</b>
	K-means	46.51	32.10	8.48	9.23E+03	3.29	9.06	2.68
	GMM	54.69	24.52	7.92	1.05E+04	3.69	<b>0.34</b>	2.69
Sample 5	FCM	45.01	33.60	8.80	8.56E+03	3.22	31.28	2.67
	FFCM	39.97	40.27	7.02	1.29E+04	2.75	27.54	2.66
	Student-T	66.14	18.58	4.65	2.23E+04	3.99	131.48	2.70
	Ours	<b>3.91</b>	<b>93.50</b>	<b>25.03</b>	204.10	<b>0.44</b>	<b>0.51</b>	2.35
	K-means	43.96	34.69	8.86	8.45E+03	3.12	8.15	2.41
Sample 5	GMM	52.93	25.89	8.32	9.57E+03	3.59	<b>0.24</b>	2.45
	FCM	28.10	<b>54.26</b>	<b>11.04</b>	5.11E+03	2.22	28.84	2.34
	FFCM	34.68	<b>46.16</b>	7.73	1.10E+04	2.48	24.16	2.41
	Student-T	66.66	19.98	2.93	3.31E+04	3.79	142.51	2.47
	Ours	<b>3.47</b>	<b>95.00</b>	<b>25.20</b>	196.33	<b>0.36</b>	<b>0.46</b>	<b>1.48</b>

TABLE III  
TEST RESULTS OF THE PRECISION, RECALL, AND RECOVERY ACCURACY

Detection sample	Rc (%)	Pr (%)	Acc (%)	Time (s)				
				ROI extraction	Gray equalization	Filtering	Segmentation	Total
Oblong(sample1)	96.6	90.1	83.3	0.15	0.18	0.01	0.13	0.46
Block(sample2/4/5)	92.4	87.6	75.1	0.62	0.37	0.11	0.40	1.37
Square(sample3)	87.0	88.7	79.3	0.16	0.21	0.01	0.25	0.62
Average	92.0	88.8	79.2	0.43	0.30	0.07	0.31	1.04

the background. In the block-type defects, there are quite a few FP error defects in Fig. 14 samples 2 and 5. The primary reason is that segmentation is wrongly taken as a block area, which causes abrupt gray-level changes in projection profiles. Moreover, the small defects near the region boundaries are occasionally considered as defects. With the proposed method, all the critical and major defects, including oblong, block, square, and so on, are easily detected. The main errors are the small, ambiguous defects or the noisy background.

## V. CONCLUSION

In this paper, a visual detection method for rail surface defect using curvature filter and improved GMM is proposed. The first step is image preprocessing, including rail surface

extraction and gray equalization. An extraction algorithm based on vertical projection is used to capture the rail surface image from original scene image. For gray equalization, the contrast image is acquired from gray contrast algorithm to handle problems such as illumination inequality and orbital reflectance inequality. We take advantage of the implicit calculation and differential geometry of three curvature filters, so that Gc, Mc, and V enable removal of noise while retaining the details and maintaining the developable surface or minimal surface of an image. In order to segment the surface defect rapidly, accurately and robustly, we propose a new way to incorporate spatial information between neighboring pixels into the GMM based on MRF. Distinct from other mixture models, the EM algorithm in our model could be directly applied to the prior distribution for maximization of the log-

likelihood function. In the experiment, the applied curvature filters achieve higher PSNRs and are 10 to 1000 times faster than other filters; the proposed mixture model performs well with the average rate on precision (88.8%), recall (92.0%), and computational time (1.04 frame/s), which is also 10 to 1000 times faster than other approaches. The results demonstrate that the proposed combined filter and MRF-based detection system can not only enhance the capability of RVDS in faint defects but also reduce the processing time cost. The qualitative and quantitative results indicate that the method is effective and efficient in railway track defects detection.

## REFERENCES

- [1] H. A. Toliyat, K. Abbaszadeh, M. M. Rahimian, and L. E. Olson, "Rail defect diagnosis using wavelet packet decomposition," *IEEE Trans. Ind. Appl.*, vol. 39, no. 5, pp. 1454–1461, Sep. 2003.
- [2] R. Clark, "Rail flaw detection: Overview and needs for future developments," *NDT &E Int.*, vol. 37, no. 2, pp. 111–118, 2004.
- [3] X. Li, B. Gao, W. L. Woo, G. Y. Tian, X. Qiu, and L. Gu, "Quantitative surface crack evaluation based on eddy current pulsed thermography," *IEEE Sensors J.*, vol. 17, no. 2, pp. 412–421, Jan. 2017.
- [4] X. Zhang, Y. Cui, Y. Wang, M. Sun, and H. Hu, "An improved AE detection method of rail defect based on multi-level ANC with VSS-LMS," *Mech. Syst. Signal Process.*, vol. 99, pp. 420–433, Jan. 2018.
- [5] I. Bartoli, F. L. di Scalea, M. Fateh, and E. Viola, "Modeling guided wave propagation with application to the long-range defect detection in railroad tracks," *NDT &E Int.*, vol. 38, no. 5, pp. 325–334, 2005.
- [6] A. Sabato and C. Niziołek, "Feasibility of digital image correlation for railroad tie inspection and ballast support assessment," *Measurement*, vol. 103, pp. 93–105, Jun. 2017.
- [7] S. Mariani, T. Nguyen, X. Zhu, and F. L. di Scalea, "Field test performance of noncontact ultrasonic rail inspection system," *J. Transp. Eng., A, Syst.*, vol. 143, no. 5, p. 04017007, 2017.
- [8] J. Wang, T. Li, Y.-Q. Shi, S. Lian, and J. Ye, "Forensics feature analysis in quaternion wavelet domain for distinguishing photographic images and computer graphics," *Multimedia Tools Appl.*, vol. 76, no. 22, pp. 23721–23737, 2017.
- [9] B. Gu, V. S. Sheng, K. Y. Tay, W. Romano, and S. Li, "Incremental support vector learning for ordinal regression," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 7, pp. 1403–1416, Jul. 2015.
- [10] H. Feng, Z. Jiang, F. Xie, P. Yang, J. Shi, and L. Chen, "Automatic fastener classification and defect detection in vision-based railway inspection systems," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 4, pp. 877–888, Apr. 2014.
- [11] M. Molodova, Z. Li, A. Núñez, and R. Dollevoet, "Automatic detection of squats in railway infrastructure," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 1980–1990, Oct. 2014.
- [12] Z. Zhou, Y. Wang, Q. M. J. Wu, C.-N. Yang, and X. Sun, "Effective and efficient global context verification for image copy detection," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 48–63, Jan. 2017.
- [13] H. Liu, J. Qin, F. Sun, and D. Guo, "Extreme kernel sparse learning for tactile object recognition," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4509–4520, Dec. 2017.
- [14] Ç. Aytekin, Y. Rezaeitabar, S. Dogru, and İ. Ulusoy, "Railway fastener inspection by real-time machine vision," *IEEE Trans. Man., Cybern., Syst.*, vol. 45, no. 7, pp. 1101–1107, Jul. 2015.
- [15] Y. Li, H. Trinh, N. Haas, C. Otto, and S. Pankanti, "Rail Component detection, optimization, and assessment for automatic rail track inspection," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 2, pp. 760–770, Apr. 2014.
- [16] P. L. Mazzeo, M. Nitti, E. Stella, and A. Distante, "Visual recognition of fastening bolts for railroad maintenance," *Pattern Recognit. Lett.*, vol. 25, no. 6, pp. 669–677, 2004.
- [17] F. Marino, A. Distante, P. L. 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, May 2007.
- [18] Q. Li and S. Ren, "A real-time visual inspection system for discrete surface defects of rail heads," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 8, pp. 2189–2199, Aug. 2012.
- [19] Q. Li and S. Ren, "A visual detection system for rail surface defects," *IEEE Trans. Syst., Man., Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 1531–1542, Nov. 2012.
- [20] P. Salvador, V. Naranjo, R. Insa, and P. Teixeira, "Axlebox accelerations: Their acquisition and time-frequency characterisation for railway track monitoring purposes," *Measurement*, vol. 82, pp. 301–312, Mar. 2016.
- [21] E. Resendiz, J. M. Hart, and N. Ahuja, "Automated visual inspection of railroad tracks," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 2, pp. 751–760, Jun. 2013.
- [22] M. Sun, X. Lin, Z. Wu, Y. Liu, Y. Shen, and N. Feng, "Non-destructive photoacoustic detecting method for high-speed rail surface defects," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2014, pp. 896–900.
- [23] S. Grassie, "Squats and squat-type defects in rails: The understanding to date," *Proc. Inst. Mech. Eng., F, J. Rail Rapid Transit*, vol. 226, no. 3, pp. 235–242, 2012.
- [24] H.-M. Thomas, T. Heckel, and G. Hanspach, "Advantage of a combined ultrasonic and eddy current examination for railway inspection trains," *Insight-Non-Destructive Test. Condition Monitor.*, vol. 49, no. 6, pp. 341–344, 2007.
- [25] C. Alippi, E. Casagrande, F. Scotti, and V. Piuri, "Composite real-time image processing for railways track profile measurement," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 3, pp. 559–564, Jun. 2000.
- [26] B. Liang, S. Iwnicki, A. Ball, and A. E. Young, "Adaptive noise cancelling and time-frequency techniques for rail surface defect detection," *Mech. Syst. Signal Process.*, vols. 54–55, pp. 41–51, Mar. 2015.
- [27] C. Mandriota, M. Nitti, N. Ancona, E. Stella, and A. Distante, "Filter-based feature selection for rail defect detection," *Mach. Vis. Appl.*, vol. 15, no. 4, pp. 179–185, 2004.
- [28] L. Wang, Y. Hang, S. Luo, X. Luo, and X. Jiang, "Deblurring Gaussian-blur images: A preprocessing for rail head surface defect detection," in *Proc. IEEE Int. Conf. Service Oper., Logistics, Informat. (SOLI)*, Jul. 2011, pp. 451–456.
- [29] A. K. Dubey and Z. A. Jaffery, "Maximally stable extremal region marking-based railway track surface defect sensing," *IEEE Sensors J.*, vol. 16, no. 24, pp. 9047–9052, Dec. 2016.
- [30] Z. He, Y. Wang, F. Yin, and J. Liu, "Surface defect detection for high-speed rails using an inverse P-M diffusion model," *Sensor Rev.*, vol. 36, no. 1, pp. 86–97, 2016.
- [31] C. Taştimur, M. Karakoç, E. Akin, and İ. Aydin, "Rail defect detection with real time image processing technique," in *Proc. IEEE 14th Int. Conf. Ind. Inform. (INDIN)*, Jul. 2016, pp. 411–415.
- [32] S. Hajizadeh, A. Núñez, and D. M. J. Tax, "Semi-supervised rail defect detection from imbalanced image data," *IFAC-PapersOnLine*, vol. 49, no. 3, pp. 78–83, 2016.
- [33] S. Faghih-Roohi, S. Hajizadeh, A. Núñez, R. Babuska, and B. De Schutter, "Deep convolutional neural networks for detection of rail surface defects," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2584–2589.
- [34] X. Gibert, V. M. Patel, and R. Chellappa, "Deep multitask learning for railway track inspection," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 1, pp. 153–164, Jan. 2017.
- [35] N. O. Attoh-Okine, *Big Data and Differential Privacy: Analysis Strategies for Railway Track Engineering*. Hoboken, NJ, USA: Wiley, 2017.
- [36] J. W. Palese and T. W. Wright, "Risk-based ultrasonic rail test scheduling on burlington northern santa fe," in *Proc. AREMA Annu. Conf.*, 2000, pp. 1–35.
- [37] D. F. Cannon, K.-O. Edel, S. L. Grassie, and K. Sawley, "Rail defects: An overview," *Fatigue Fracture Eng. Mater. Struct.*, vol. 26, no. 10, pp. 865–886, 2003.
- [38] S. Shirmohammadi and A. Ferrero, "Camera as the instrument: The rising trend of vision based measurement," *IEEE Instrum. Meas. Mag.*, vol. 17, no. 3, pp. 41–47, Jun. 2014.
- [39] S. Tania and R. Rowaida, "A comparative study of various image filtering techniques for removing various noisy pixels in aerial image," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 9, no. 3, pp. 113–124, 2016.
- [40] N. Remenyi, O. Nicolis, G. Nason, and B. Vidakovic, "Image denoising with 2D scale-mixing complex wavelet transforms," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5165–5174, Dec. 2014.
- [41] Y. Gong, "Spectrally regularized surfaces," Ph.D. dissertation, MOSAIC Group, ETH Zurich, Zürich, Switzerland, 2015.
- [42] Y. Gong and I. F. Sbalzarini, "Curvature filters efficiently reduce certain variational energies," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1786–1798, Apr. 2017.

- [43] Y. Gong, "Bernstein filter: A new solver for mean curvature regularized models," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2016, pp. 1701–1705.
- [44] T. Akilan, Q. M. J. Wu, and Y. Yang, "Fusion-based foreground enhancement for background subtraction using multivariate multi-model Gaussian distribution," *Inf. Sci.*, vols. 430–431, pp. 414–431, Mar. 2018.
- [45] S. Sanjay-Gopal and T. J. Hebert, "Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm," *IEEE Trans. Image Process.*, vol. 7, no. 7, pp. 1014–1028, Jul. 1998.
- [46] S. Z. Li, *Markov Random Field Modeling in Image Analysis*. London, U.K.: Springer, 2009.
- [47] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Roy. Stat. Soc., B (Methodological)*, vol. 36, no. 2, pp. 192–236, 1974.
- [48] C. Nikou, N. P. Galatsanos, and C. L. Likas, "A class-adaptive spatially variant mixture model for image segmentation," *IEEE Trans. Image Process.*, vol. 16, no. 4, pp. 1121–1130, Apr. 2007.
- [49] K. Blekas, A. Likas, N. P. Galatsanos, and I. E. Lagaris, "A spatially constrained mixture model for image segmentation," *IEEE Trans. Neural Netw.*, vol. 16, no. 2, pp. 494–498, Mar. 2005.
- [50] T. M. Nguyen and Q. M. J. Wu, "Fast and robust spatially constrained Gaussian mixture model for image segmentation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 621–635, Apr. 2013.
- [51] X. Gibert, V. M. Patel, and R. Chellappa, "Robust fastener detection for autonomous visual railway track inspection," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2015, pp. 694–701.
- [52] X. Tao, Z. Zhang, F. Zhang, and D. Xu, "A novel and effective surface flaw inspection instrument for large-aperture optical elements," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 9, pp. 2530–2540, Sep. 2015.
- [53] W. Zhao and H. Lu, "Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 9, pp. 2283–2294, Sep. 2017.
- [54] A. R. J. Fredo, R. S. Abilash, and C. S. Kumar, "Segmentation and analysis of damages in composite images using multi-level threshold methods and geometrical features," *Measurement*, vol. 100, pp. 270–278, Mar. 2017.
- [55] Z. Ji, Y. Huang, Q. Sun, and G. Cao, "A spatially constrained generative asymmetric Gaussian mixture model for image segmentation," *J. Vis. Communun. Image Represent.*, vol. 40, pp. 611–626, Oct. 2016.
- [56] D. Zhang, Q. Li, Y. Chen, M. Cao, L. He, and B. Zhang, "An efficient and reliable coarse-to-fine approach for asphalt pavement crack detection," *Image Vis. Comput.*, vol. 57, pp. 130–146, Jan. 2017.



**Hui Zhang** received the B.S., M.S., and Ph.D. degrees in pattern recognition and intelligent system from Hunan University, Changsha, China, in 2004, 2007, and 2012, respectively.

He is currently an Assistant Professor with the College of Electrical and Information Engineering, Changsha University of Science and Technology, Changsha. He is also a Post-Doctoral Fellow with the College of Electrical and Information Engineering, Hunan University. He is also a Visiting Scholar with the Computer Vision and Sensing Systems Laboratory, Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. His current research interests include machine vision, sparse representation, and machine learning.



**Xiating Jin** was born in Wenzhou, China, in 1995. He received the B.S. degree in automation from the Changsha University of Science and Technology, Changsha, China. He is currently pursuing the master's degree with the College of Electrical and Information Engineering, Hunan University, Changsha.

His current research interests include machine vision, intelligent image processing, and industry inspection.



**Q. M. Jonathan Wu** (M'92–SM'09) received the Ph.D. degree in electrical engineering from the University of Wales, Swansea, U.K., in 1990.

He was with the National Research Council of Canada for ten years beginning in 1995, where he became a Senior Research Officer and a Group Leader. He is currently a Professor with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He is also a Visiting Professor with the School of Automation, Hangzhou Dianzi University, Hangzhou, China. He has authored over 300 peer-reviewed papers in computer vision, image processing, intelligent systems, robotics, and integrated microsystems. His current research interests include 3-D computer vision, active video object tracking and extraction, interactive multimedia, sensor analysis and fusion, and visual sensor networks.

Dr. Wu holds the Tier 1 Canada Research Chair in Automotive Sensors and Information Systems. He is an Associate Editor of the *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*, and the *Cognitive Computation*. He has served on technical program committees and international advisory committees for many prestigious conferences.



**Yaonan Wang** received the Ph.D. degree in electrical engineering from Hunan University, Changsha, China, in 1994.

He was a Post-Doctoral Research Fellow with the Normal University of Defense Technology, Changsha, from 1994 to 1995. From 2001 to 2004, he was a Visiting Professor with the University of Bremen, Bremen, Germany. Since 1995, he has been a Professor with the College of Electrical and Information Engineering, Hunan University. His current research interests include robotics and image processing.

Dr. Wang was a Senior Humboldt Fellow in Germany from 1998 to 2000.



**Zhendong He** received the B.S. and M.S. degrees in robot intelligence and system from Northeast Dianli University, Jilin, China, in 2003 and 2006, respectively, and the Ph.D. degree in pattern recognition and intelligent system from Hunan University, Changsha, China, in 2016.

He is currently an Assistant Professor with the School of Electrical and Information Engineering, Zhengzhou University of Light Industry, Zhengzhou, China. His current research interests include machine vision, robotics, and machine learning.



**Yimin Yang** (S'10–M'13) received the M.Sc. and Ph.D. degrees in electrical engineering from the College of Electrical and Information Engineering, Hunan University, Changsha, China, in 2009 and 2013, respectively.

He is currently an Assistant Professor with the College of Electric Engineering, Guangxi University, Nanning, China. He is also a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Windsor, Windsor, ON, Canada. He has authored or co-authored over 20 refereed papers. His current research interests include computer vision, robotics, and machine learning.