

# Automatic Fastener Classification and Defect Detection in Vision-Based Railway Inspection Systems

Hao Feng, Zhiguo Jiang, Fengying Xie, Ping Yang, Jun Shi, and Long Chen

**Abstract**—The detection of fastener defects is an important task in railway inspection systems, and it is frequently performed to ensure the safety of train traffic. Traditional inspection is usually operated by trained workers who walk along railway lines to search for potential risks. However, the manual inspection is very slow, costly, and dangerous. This paper proposes an automatic visual inspection system for detecting partially worn and completely missing fasteners using probabilistic topic model. Specifically, our method is able to simultaneously model diverse types of fasteners with different orientations and illumination conditions using unlabeled data. To assess the damages, the test fasteners are compared with the trained models and automatically ranked into three levels based on the likelihood probability. The experimental results demonstrate the effectiveness of this method.

**Index Terms**—Fastener, latent Dirichlet allocation (LDA), railway, structure modeling, visual inspection.

## I. INTRODUCTION

RAILWAY inspection is a very critical task for ensuring the safety of railway traffic. Traditionally, this task is operated by trained human inspectors who periodically walk along railway lines to search for any damages of railway components. However, the manual inspection is slow, costly, and even dangerous. With the extension of high-speed railway network, the inspection and maintenance face more challenges than ever before. Recently, the railway companies of all over the world are interested in developing automatic inspection systems, which are specialized trains and are able to detect railway defects very efficiently [1].

An automatic railway inspection system is composed of a number of functions such as gauge measurement [2], track-profile measurement [1], [3], track-surface defects detection [4], and fastener defects detection [5], [6]. Our research focuses on automatically finding and assessing the partially worn and missing fasteners based on computer vision technologies.

Manuscript received December 17, 2012; revised April 1, 2013; accepted April 2, 2013. Date of publication October 11, 2013; date of current version March 6, 2014. This work was supported by the National Natural Science Foundation of China under Grant 61071137, Grant 61071138, Grant 61027004, and the 973 Program of China under Project 2010CB327900. The Associate Editor coordinating the review process was Dr. Shervin Shirmohammadi.

The authors are with the Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing, 100191, China, and also with the Beijing Key Laboratory of Digital Media, Beijing, 100191, China (e-mail: fenghao@sa.buaa.edu.cn).

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.2013.2283741

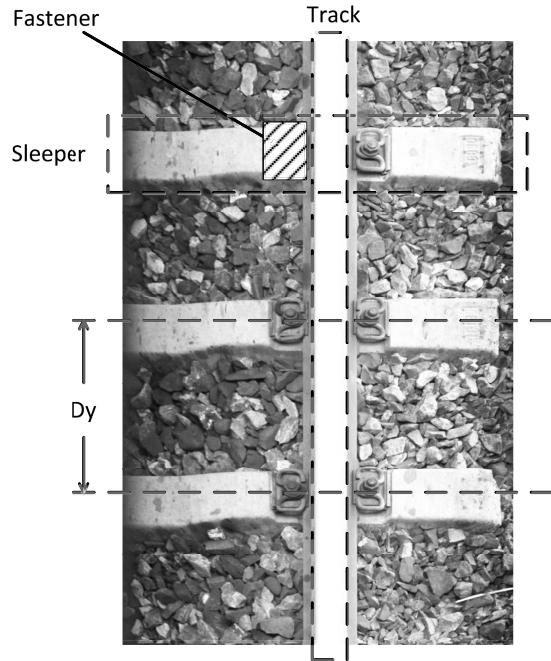


Fig. 1. Image captured by onboard camera.

The fasteners are used to hold the track on sleepers, as shown in Fig. 1, while the worn and missing fasteners shown in Fig. 2 are hazardous defects, which would cause the displacement of track and even threaten the safety of train operation. Generally speaking, there are two kinds of fasteners: 1) hexagonal-headed bolts and 2) hook-shaped fasteners. The hook-shaped fasteners are widely used in current railway lines. However, their diverse shapes give rise to significant difficulties in both modeling and inspecting. In this paper, multiple types of hook-shaped fasteners and hexagonal-headed bolts are jointly concerned.

In the past decade, some researchers have devoted into developing fastener inspection methods. For hexagonal-headed bolts, Marino *et al.* [7] used a multilayer perceptron neural classifier to detect missing bolts. For hook-shaped fasteners, Stella *et al.* [8] employed wavelet transformation and principal component analysis to preprocess railway images and searched for the missing fasteners using the neural classifier. Similarly, Yang *et al.* [9] took advantage of direction field as the template of fastener. For matching, they use linear discriminant analysis to obtain the weight coefficient matrix. To achieve real-time performance, Ruvo *et al.* [10] applied the



Fig. 2. Worn and missing fasteners. (a) Left fastener is partially worn. (b) Right fastener is completely missing.

error backpropagation algorithm to model two types of fastener. They implemented the detection algorithm on graphical processing units. Ruvo *et al.* [11] also introduced a FPGA-based architecture for automatic hexagonal bolts detection using the same algorithm. However, the methods mentioned above cannot identify the partially worn fasteners. Recently, Xia *et al.* [5] and Rubinsztein [12] have successfully applied the AdaBoost algorithm to the fastener detection work. Specifically, Xia *et al.* [6] departed the hook-shaped fastener into four parts and each part was independently trained by AdaBoost. Thus, this method has the ability of detecting partially worn fasteners. Similarly, Li *et al.* [13] used image processing methods to detect the components of fastener. However, these two methods can only handle specific fastener types and the robustness on illumination variation was not discussed. Other technologies that have been used for modeling and detecting fasteners include support vector machine (SVM) [14], Gabor filters [15], and edge detection [16]. In summary, most of the earlier methods take advantage of discriminative models (classifiers) to classify the fastener and nonfastener samples, but it is difficult for them to identify the partially worn ones, because there is no uniform representation of the worn cases. Although, the part-based methods [5], [13] can solve this problem to some extent, they require multiple classifiers and can only handle specific type of fastener. On the other hand, to train classifiers, numerous labeled fastener samples including worn and intact fasteners must be collected. However, the number of partially worn fasteners is very limited. Our method is designed for detecting the defects of various types of fastener and is able to find both the partially worn and the completely missing instances. Different from earlier methods, which rely on discriminative model, we solve this problem using generative model. Hence, the training stage requires only intact fastener samples. Meanwhile, our training set is unlabeled and composed of various fastener types, orientations and illumination conditions.

We conclude three major problems in current fastener inspection systems.

- 1) There are lots of partially worn fasteners that have already lost effectiveness. However, most of the researches aim at searching for missing fasteners, and they usually fail to identify the partially worn ones.
- 2) A railway line is always installed with various types of fasteners for various reasons (some examples are shown in Fig. 3). Therefore, it is necessary for an automatic inspection system to recognize all these types without manual intervention.

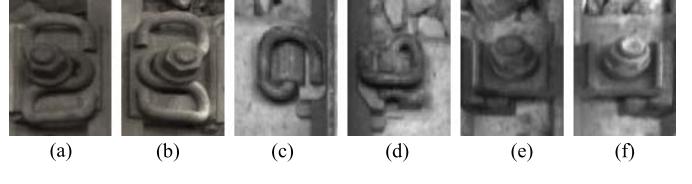


Fig. 3. Three types of fasteners. (a) Type-I left. (b) Type-I right. (c) Type-II left. (d) Type-II right. (e) Type-III left. (f) Type-III right.

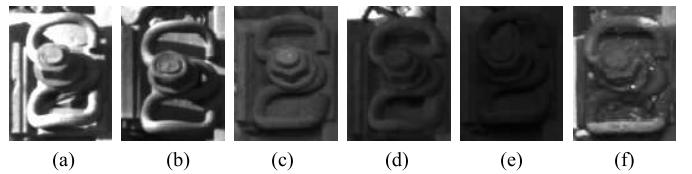


Fig. 4. (a)–(e) Fasteners with different illumination conditions. (f) Fastener polluted by leaked grease.

- 3) The qualities of the captured images are not uniform in illumination. Generally speaking, the cameras are installed in the open area under a train coach. The sun light will significantly affect the brightness of the acquired images, as shown by Fig. 4. Although active light sources and sun shields are used, this problem is still unsolved. On the other hand, fasteners are always shielded by stone, leaked grease, or litter. These issues affect the performance of some fastener detection algorithms.

To handle these problems, we propose a new probabilistic structure topic model (STM) to model fasteners. This model is generative, data driven, and it can simultaneously learn the probabilistic representations of different objects using unlabeled samples. We train the fastener models using a collection of intact fastener samples. Compare with classifier-based approaches, STM only interests in the intrinsic features of fastener. The likelihood probability can be used to measure the similarity between a test fastener and a model. Generally speaking, the worn fastener has lower likelihood probability than intact ones. We rank fasteners into three levels based on their likelihood probabilities in descending order. The intact fasteners are ranked into high level; the fasteners in middle level may be partly worn or polluted and the fasteners ranked into low level are severely worn or completely missing.

STM is an extension of latent Dirichlet allocation (LDA) [17], which is a probabilistic topic model for unsupervised extracting and modeling semantic topics from language words or other discrete data. Many modeling problems in computer vision field have been solved using the LDA model [18]–[20]. To the best of our knowledge, this is the first research that offers the LDA model with the ability of structure analysis. STM has the same advantages as LDA, it is able to cluster the samples with the similar structure into a class and simultaneously generate the probabilistic model for this class.

In summary, our approach has the following three advantages.

- 1) Multiple types of fasteners can be recognized. The models of all types of fasteners in each orientation are simultaneously learned from unlabeled samples.

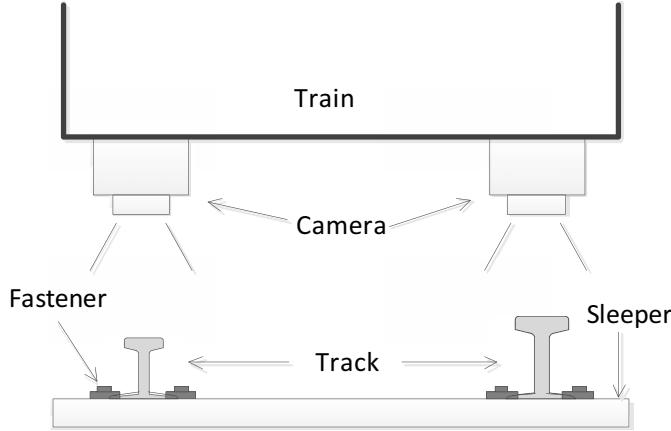


Fig. 5. Configuration of our fastener inspection system.

- 2) The fasteners of different illumination conditions are independently modeled. Thus, the illumination variations will not affect the final performance.
- 3) Our system outputs the states of fasteners automatically. This design offers the convenience for human inspectors to assess the damages. Specifically, the method can effectively detect the partially worn fasteners.

This paper is organized as follows. In Section II, the overview of our system is presented. In Section III, we briefly introduce the fastener localization algorithm. The STM model is introduced in Section IV. The experimental results are illustrated in Section V. Finally, we conclude this paper in Section VI.

## II. SYSTEM OVERVIEW

The configuration of our system is shown in Fig. 5. There are two cameras hanged below a train coach, each of which monitors a side of track. The size of an acquired image is  $560 \times 900$  pixels. An example of the acquired image is shown in Fig. 1. The image is sent to the onboard high-performance computers as the input of data processing module. The data processing module is composed of three major components: 1) fastener localization; 2) fastener classification; and 3) score ranking. The final results are summarized on an interactive user interface for further manual inspection and analysis. We show the flow chart of the data processing module in Fig. 6.

Overviews of these steps are given as follows.

1) *Fastener localization*: Fasteners are permanently installed on each sleeper at both sides of a track. Therefore, the localization of fasteners can be replaced with the detection of sleepers and tracks. This detection is achieved by searching some specific geometric relations between line segments. To alleviate the computational burden, we predict the positions of fasteners based on the constant distance between two adjacent sleepers.

2) *Fastener classification*: The image regions that contain fasteners are then transmitted to the classification module. In this module, the type of each test fastener is determined according to the learned models.

3) *Wear ranking*: We rank the fasteners into three levels based on the likelihood probability, which measures the consistency between a fastener and the corresponding model. We show that the fasteners ranked into the low level are severely worn or missing, which need to be immediately replaced. Most of the fasteners ranked into middle level are partly worn. These fasteners require further assessment by human inspectors.

It should be noted that the scope of this paper is limited to the development of software algorithms of data processing module. The detailed description of the image acquisition subsystem will not be provided.

## III. FASTER LOCALIZATION

The positions of fasteners can be indirectly determined by the positions of sleepers and tracks. In this paper, we take advantage of the robust line segment detection algorithm and the geometric relationships to localize sleepers and tracks. We first introduce the track detection and sleeper detection algorithms and then describe a sleeper prediction approach, which considerably accelerates the detection speed and improves the robustness.

### A. Track Detection

In an acquired image, a track is viewed as a long rectangle vertically located nearby the middle of image. Mostly, it is overexposure due to the high reflection rate of the smooth track head. The detection of a track can be simplified to the detection of two longest vertical lines. First of all, line segment detector (LSD) [21] is used to extract lines. The vertical lines that close to the middle of the horizontal axis are preserved. Then, the pixel values are projected onto the  $x$ -axis to generate accumulated intensity histogram. Finally, the overlapped positions of the vertical lines and the sharp increasing or decreasing in the histogram are identified as the edges of the track.

### B. Sleeper Detection

For most of the railway infrastructures, sleepers are symmetry with respect to the track and periodically arranged along the railway line. Therefore, the sleepers can be detected by simply searching the symmetrical line pairs at the same  $y$ -coordinate. This algorithm is composed of the following three steps.

- 1) The LSD algorithm is performed on railway images to extract line segments. Only the horizontal lines are preserved.
- 2) Among these lines, the parallel lines are identified. The distance between two parallel lines is not longer than the width of a sleeper. For our image, the maximum width of a sleeper is 180 pixels.
- 3) For each pair of the parallel lines, if one or more corresponding parallel lines with the same  $y$ -coordinate can be found at the other side of the track, the region enclosed by these parallel lines is recognized as a sleeper.

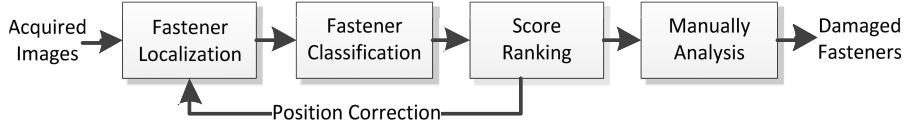


Fig. 6. Flow chart of the data processing module.

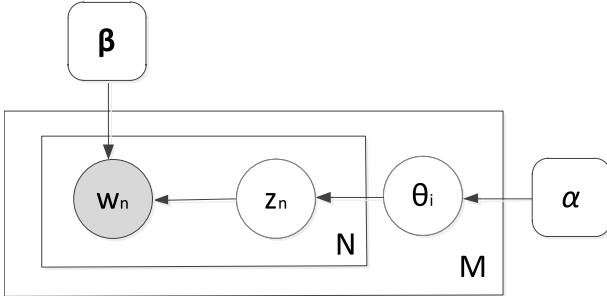


Fig. 7. Graphic model representation for LDA.

### C. Sleeper Prediction

The position of a fastener can be predicted by simply adding  $Dy$  in terms of the position of the previously detected sleeper. As shown in Fig. 1,  $Dy$  is the constant distance between adjacent sleepers. In our implementation,  $Dy$  is initialized by performing sleeper localization algorithm for the first 100 frames, and then the positions of following sleepers are predicted in coming frames. The sleeper detection algorithm is also performed in every 500 frames to rectify  $Dy$  and prevent the accumulated error. Furthermore,  $Dy$  is recomputed when all the adjacent fasteners are given very low likelihood probabilities in the classification stage.

## IV. FASTENER CLASSIFICATION

To effectively model fasteners, we propose a structure analysis approach, which employs the advantages of the LDA model. We named our model as STM. In the first two parts of this section, we first give a brief introduction to LDA and then detail our STM model. The feature computation and other implementation details are given in Sections IV-C and D, respectively.

### A. Latent Dirichlet Allocation

There are two understandings of LDA [17]: 1) LDA is a probabilistic clustering method, which can be used to cluster words into semantic topics based on the co-occurrence property and 2) LDA is a data-driven model, and it can automatically explore the latent topics from unlabeled discrete data. On the other hand, LDA suffers from some weaknesses. The most obvious one is that the spatial relationship between words is ignored. To clearly understand the LDA topic model, we first show the probabilistic graphic model representation in Fig. 7.

Given a collection of  $M$  documents denoted by  $I_m = \{I_1, I_2, \dots, I_M\}$ , each document has  $N$  words. LDA groups words  $\omega_n, n = \{1, 2, \dots, N\}$  into  $K$  topics, which is equivalent to assign a latent topic to each word. In Fig. 7,  $z_n$  is an

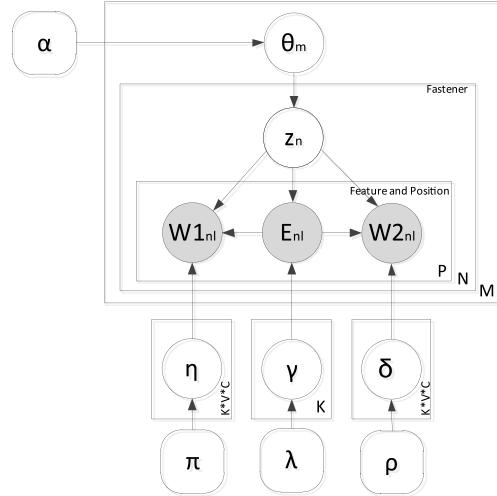


Fig. 8. Graphic model representation of STM.

index, which shows the topic label of word  $\omega_n$ ,  $\theta_i$  represents the distribution of topics for a document (document-topic distribution) and itself has a Dirichlet prior with parameter  $\alpha$ , and  $\beta$  is a matrix for the word distributions of each latent topic (topic-word distribution).

To apply this model for solving computer vision problems, the concepts of images must be translated to the corresponding concepts of languages. There are two aspects: 1) image features are translated to words (these words are usually called visual words) using bag of features [20]. This issue will be presented in Section IV-D; and 2) in STM, the topics are translated to fastener classes. The description of the model is given in Section IV-B.

### B. Fastener Modeling With STM

STM that considers the spatial information of visual words is an extension of generative topic model. We model the structures of fasteners in topic level. The STM model has the following two advantages when handling our fastener modeling problem: 1) it can simultaneously learn multiple types of fasteners from unlabeled samples and produce the models for each fastener class (topic) and 2) the learned model can be used to classify fasteners and offer the consistency scores for assessing the damages.

The graphic model representation of STM is shown in Fig. 8. Suppose that the data set contains  $M$  railway images denoted by  $\theta_m = \{\theta_1, \dots, \theta_M\}$ . Each image contains  $N$  fasteners.  $z_n = \{1, \dots, K\}$  is the class label of a fastener. In fact,  $\theta_m$  represents the distribution of fastener classes in the  $m$ th image. A fastener class is represented as the composition of  $P$  triples ( $W1_{nl}, W2_{nl}, E_{nl}$ ). Specifically,  $E_{nl}$  is an index

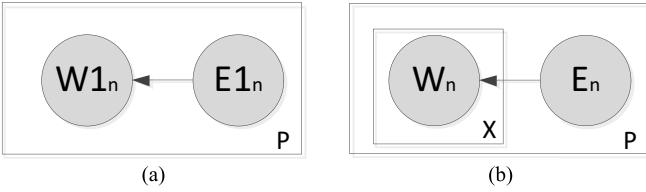


Fig. 9. Graphic model representations of other STM setups (only the modified parts are shown). STM model with (a) single feature and (b) multiple features.

points to two different coordinates denoted by  $C_{E_{nl}}^{(1)}$  and  $C_{E_{nl}}^{(2)}$ , from which the visual words  $W_{1nl}$  and  $W_{2nl}$  are sampled, respectively. In other words, this triple expresses a truth that for a type of fastener, there must be two specific visual words simultaneously occurred at two given coordinates. It should be noted that  $E_{nl}$  can suggest only one or multiple coordinates by simply changing the model, as shown in Fig. 9. In Fig. 9(a),  $E_{nl}$  contains only one coordinate. This simplification results in underfitting when modeling the objects with similar structure configurations. On the contrary, the model in Fig. 9(b) places very strict constraints on object structures and leads to overfitted representations. In addition, this model also consumes huge memory and computational resources in the inference and parameter estimation procedures.

From the viewpoint of generative process,  $\theta_m$  is sampled from a Dirichlet distribution with parameter  $\alpha$ . For each image, fastener class  $z_n$  is drawn from a multinomial distribution with parameter  $\theta_m$ . For a fastener class  $z_n$ , the index of a coordinate pair is first drawn from the multinomial distribution parameterized by  $\gamma_{z_n}$  (class-coordinate distribution). Assuming that  $C_{E_{nl}}^{(1)}$  and  $C_{E_{nl}}^{(2)}$  are the two coordinates shown by  $E_{nl}$ , the visual words  $W_{1nl}$  and  $W_{2nl}$  are then simultaneously sampled from multinomial distributions with parameters  $\eta_{C_{E_{nl}}^{(1)}, z_n}$  and  $\delta_{C_{E_{nl}}^{(2)}, z_n}$ , respectively. Both  $\eta_{C_{E_{nl}}^{(1)}, z_n}$  and  $\delta_{C_{E_{nl}}^{(2)}, z_n}$  are the distributions over visual words with respect to the fastener class and coordinate. In summary, the generative process of STM is given as follows.

- 1) For each image, draw a fastener class distribution  $\theta_m$  according to Dirichlet ( $\alpha$ ).
- 2) Draw a fastener class  $z_n$  from multinomial ( $\theta_m$ ).
- 3) For each fastener sample, do the following steps  $P$  times.
  - a) Draw a class-coordinate distribution  $\gamma$  according to Dirichlet ( $\lambda$ ).
  - b) Draw an index  $E_{nl}$  according to multinomial ( $\gamma_{z_n}$ ). This is equivalent to sample two coordinates  $C_{E_{nl}}^{(1)}$  and  $C_{E_{nl}}^{(2)}$ .
  - c) Draw class-word distributions  $\eta$  and  $\delta$  according to Dirichlet ( $\pi$ ) and Dirichlet ( $\rho$ ), respectively.
  - d) Draw visual words  $W_{1nl}$  and  $W_{2nl}$  according to multinomial ( $C_{E_{nl}}^{(1)}, z_n, \eta$ ) and multinomial ( $C_{E_{nl}}^{(2)}, z_n, \delta$ ), respectively.

Given hyperparameters  $\alpha$ ,  $\pi$ ,  $\rho$ , and  $\lambda$ , the joint distribution of hidden variables  $\theta_m$ ,  $z_n$ ,  $\eta$ ,  $\gamma_{z_n}$ , and  $\rho$  and observed triples

$(W_{1n}, W_{2n}, E_n)$  can be written as follows:

$$P(\theta, z, E, W_1, W_2 | \alpha, \pi, \lambda, \rho) = P(\theta_m | \alpha) \prod_{n=1}^N \left[ P(z_n | \theta_m) \sum_{l=1}^P P(E_{nl} | z_n, \lambda) P(W_{1nl} | E_{nl}, z_n, \pi) P(W_{2nl} | E_{nl}, z_n, \rho) \right]. \quad (1)$$

Exact inference in STM is intractable, but it is possible to use a collapsed Gibbs sampler to approximate the distribution [22], [23]. Gibbs sampler alternatively samples a hidden variable following its distribution conditioned on all the data and other variables sampled at the previous iteration [22]. In our application, we only interested in sampling class label (topic)  $z$ . Other hidden variables, such as  $\theta$ ,  $\eta$ ,  $\gamma$ , and  $\delta$ , can be analytically marginalized out without explicit sampling. The conditional posterior distribution given  $z$  and all the observations ( $E$ ,  $W_1$ ,  $W_2$ ) is given by

$$p(z_i^d = j | z_{-i}^d, \mathbf{w}, \mathbf{e}) \propto \frac{n_{-i,j}^{(d)} + \alpha}{n_{-i,-}^{(d)} + K\alpha} \sum_{l=1}^P \left( \frac{n_{-l,j}^{(E_{il})} + \lambda}{n_{-l,-}^{(E_{il})} + C\lambda} \times \frac{n_{-r,j}^{(w_{E_{il},c1})} + \pi}{n_{-r,-}^{(w_{E_{il},c1})} + V\pi} \times \frac{n_{-q,j}^{(w_{E_{il},c2})} + \delta}{n_{-q,-}^{(w_{E_{il},c2})} + V\delta} \right) \quad (2)$$

where  $i$  is the index of a fastener sample and  $j$  is the label of fastener class (topic).  $z_i^d$  is the class label of the  $i$ th patch in image  $d$  and  $z_{-i}^d$  denotes the class label of all image patches, not including the  $i$ th patch in image  $d$ .  $n_{-i,j}^{(d)}$  is the number of triples assigned to class  $j$  in image  $d$ , not including the current one,  $n_{-i,-}^{(d)}$  is the total number of triples in image  $d$ , not including the current one,  $n_{-l,j}^{(E_{il})}$  is the number of instances of coordinate pair  $E_{il}$  assigned to class  $j$ , not including the current one,  $n_{-r,j}^{(w_{E_{il},c*})}$  is the number of instances of word  $w_{C_{E_{nl}}^{(*)}}$  appeared at coordinate  $C_{E_{nl}}^{(*)}$  in class  $j$ , not including the current word (\* represents one or two), and  $n_{-r,-}^{(w_{E_{kl},c*})}$  is the total number of words occurred at coordinate  $C_{E_{nl}}^{(*)}$  in class  $j$ , not including the current word (\* represents one or two).  $\mathbf{w}$  represents the observed words and  $\mathbf{e}$  represents the coordinate list. The constants are defined as follows:  $C$  is the total number of coordinate pairs,  $V$  is the total number of visual words,  $P$  is the number of coordinate pairs considered in a fastener class, and  $K$  is the total number of classes (topics).

The sampling algorithm runs (2) for a number of iterations. The class (topic) of each sample is updated in each iteration. Finally, the class labels and some statistical variables are recorded. Here, the class-word and class-coordinate distributions can be computed using the following equations:

$$\eta_{j,E,w} = \frac{n_j^{(w,E,c1)} + \pi}{n_j^{(w,E,c1)} + V\pi} \quad (3)$$

$$\delta_{j,E,w} = \frac{n_j^{(w,E,c2)} + \rho}{n_j^{(w,E,c2)} + V\rho} \quad (4)$$

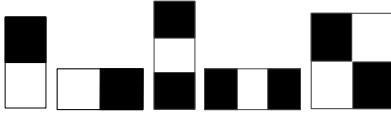


Fig. 10. Haar-like features.

$$\gamma_{j,E} = \frac{n_j^{(E)} + \lambda}{n_j^{(\cdot)} + C\lambda} \quad (5)$$

where  $n_j^{(w, E, c*)}$  is the number of instances of word  $w$  occurred at coordinate  $C_{E_nl}^{(*)}$  in class  $j$  (\* represents one or two),  $n_j^{(\cdot, E, c*)}$  is the total number of words occurred at coordinate  $C_{E_nl}^{(*)}$  in class  $j$  (\* represents one or two),  $n_j^{(E)}$  is the number of instances of coordinate pair  $E$  assigned to class  $j$ , and  $n_j^{(\cdot)}$  is the total number of coordinate pairs assigned to class  $j$ .

Given an unknown fastener image and its word map  $\mathbf{w}$ , the likelihood for a class  $j$  can be computed using

$$p(\mathbf{w}, \mathbf{e}|z = j) \propto \sum_{E=1}^C \gamma_{j,E} \cdot \eta_{j,E,w} \cdot \delta_{j,E,w}. \quad (6)$$

The class of a fastener is defined as the one with the maximum likelihood

$$z = \max_{j=1 \dots K} (p(\mathbf{w}, \mathbf{e}|z = j)). \quad (7)$$

### C. Normalized Haar-Like Feature

Haar-like feature [24] is a simple but effective representation of local structures. In our implementation, we use five layouts, as shown in Fig. 10. To represent structures of multiple scales, we link the Haar-like features of five different types with the size of 2–10 pixels into 45-dimension descriptor.

The multisize Haar-like feature is normalized according to the maximum values of each dimension. For example, the value of the second layout in Fig. 10 can be calculated using the following equation:

$$\left| \sum_{j=0}^{H-1} \sum_{i=W}^{2W-1} I(i, j) - \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} I(i, j) \right| / \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} 255 \quad (8)$$

where  $W$  and  $H$  are the width and height of a subrectangle and  $I(i, j)$  is the intensity value at the coordinate  $(i, j)$ . Other dimensions are normalized using the similar manner.

### D. Implementation Details

1) *Training Samples*: The training samples can be easily selected using the fastener localization method, which has been introduced in Section III. The left and right regions at the intersections of the detected rails and sleepers are captured as the training sample. The size of samples is  $84 \times 110$  pixels. All types of fasteners at both sides can be trained together without labeling. In our experiments, the training set is composed of three types of fasteners, and each type contains 200 samples.

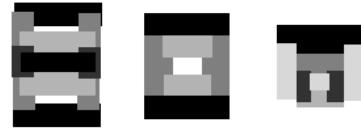


Fig. 11. Sketches of the considered fasteners.

---

**Algorithm 1** Training Process of STM

---

**Initialize:**

Randomly assign topic to each sample  $z_i^d = j_0$ .

**In the mth iteration:**

For the  $i$ th sample in the  $d$ th image:

- 1) Remove the topic  $z_i^d = j_{m-1}$ ,
- 2) For each topic  $j (j = 1, \dots, K)$ , compute the predictive likelihood  $p(z_i^d = j | z_{-i}^d, \mathbf{w}, \mathbf{e})$  using (2);
- 3) Sample new topic assignment from the predictive likelihoods;
- 4) Assign the new topic  $j_m$  to the current patch, let  $z_i^d = j_m$ .

Compute Eq. (3), (4), and (5) for fastener models.

---

2) *Parameter Selection*: In the proposed model, the STM hyperparameters and the topic number need to be provided. The STM hyperparameters control the smoothness of the multinomial distributions [25]. These parameters are chosen experimentally. We first train the model using recommended parameter values [23], [25] and tune the parameters based on the topic assignment result. We chose the parameters, which result in the lowest classification error. In our implementation,  $\alpha$  is set to 0.1 and other hyperparameters are set to one. The topic number  $K$  should be larger than the actual number of the considered fastener types to reserve enough topics to represent the appearance variances. In our application, we modeled three types of fasteners using 30 topics.

3) *Coordinate-List Generation*: To alleviate the computational burden, the coordinates, which are indicated by  $E_{nl}$ , are selected from a list (coordinate list). In fact,  $E_{nl}$  is the index of an element in this list. The coordinate list is generated as follows. For each type of fastener, we draw a sketch using different gray scales, as shown in Fig. 11. Then, the algorithm randomly selects a number of pixel pairs with different gray scales, and the two coordinates of pairs are recorded as an element of the coordinate list. It should be noted that all the pairs from different sketches are mixed together. In our configuration, the coordinate list contains 11451 coordinate pairs.

4) *Training Process*: A visual dictionary is first built by performing K-means clustering on all the Haar features extracted from the training samples. Each cluster center is recognized as a visual word. In our implementation, there are total 200 visual words in the dictionary. In addition, some visual words correspond to flat features (the value of feature is very low) are meaningless, these words are manually identified and removed in advance. Then, each feature descriptor in an image is mapped to a certain visual word through the

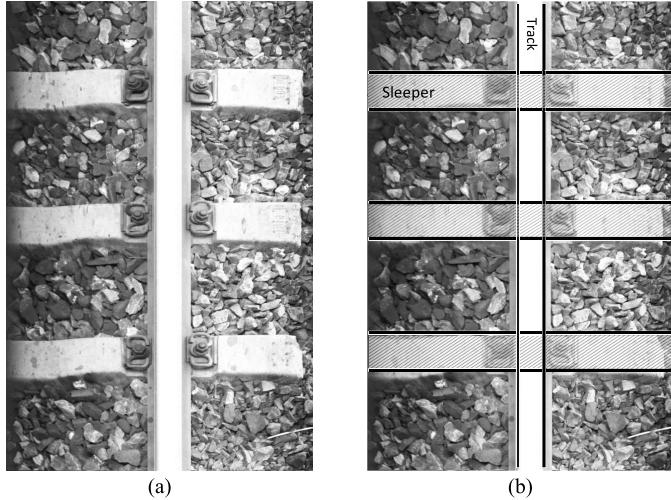


Fig. 12. Detection result for track and sleepers. (a) Original image. (b) Detection result.

TABLE I  
DETECTION RATE OF THE PROPOSED FASTENER DETECTION ALGORITHM

	Use sleeper detection	Use sleeper prediction
Detection Rate	95.2%	100%

nearest neighborhood criterion, by which each image can be represented by a sequence of visual words.

The training process can be summarized in Algorithm 1. It should be noted that to compute  $n_{-l,j}^{(E_{il})}$  and  $n_{-l,j}^{(.)}$ , the occurrence of a coordinate pair is defined as the condition that both the visual words at coordinates  $C_{E_{nl}}^{(1)}$  and  $C_{E_{nl}}^{(2)}$  are not meaningless.

5) *Fastener Classification*: For each test sample, Haar-like features are first extracted and mapped to the visual dictionary. Then, we scan its word map in pixel-by-pixel manner and compute the likelihood probabilities for each topic using (6). After scanning the entire image, a set of probability maps can be achieved. The type of a fastener is determined by the maximum value in these probability maps using (7).

## V. EXPERIMENTAL RESULTS

In this section, both the fastener detection method and the STM fastener model are evaluated. The test set contains 57 124 images covers 110-km railway line. Each image has six or eight fasteners. The total number of fasteners is 399 078.

### A. Fastener Localization

In this experiment, we use two schemes to localize fasteners. First, the sleeper and track detection algorithms are performed on all test images. The second scheme uses sleeper prediction approach and performs periodical rectification for  $D_y$ . An example of detection result of sleeper and track is shown in Fig. 12, where the regions enclosed by horizontal lines show

the detected sleepers and the vertical lines show the edges of detected track.

The evaluation results are shown in Table I. From this table, we can observe that the fastener detection algorithm achieves the detection rate of 95.2% and the sleeper prediction method helps to improve the detection rate to 100%. It is apparent that with the prior knowledge about the railway configuration, the fastener localization algorithm achieves very reliable performance.

### B. Fastener Classification and Score Ranking

In this experiment, we model three types of widely used fasteners shown in Fig. 1. The training set contains 600 images with the size of  $84 \times 110$  pixels. The implementation details are introduced in Section IV-D. The model is configured as follows: 1) the hyperparameters of Dirichlet distributions are  $\alpha = 0.1$ ,  $\pi = 1$ ,  $\lambda = 1$ , and  $\rho = 1$ ; 2) the topic number  $K$  is set to 30; and 3) the number of iterations is 100. The parameter settings affect the training speed. Under the current configuration, the training phase consumes 2 h on our computer, and the maximum memory consumption is 7.2 GB.

In the following parts, the topic assignments of the training samples are first displayed. Furthermore, the effectiveness of the proposed method will be verified by classifying the fasteners. Finally, we will illustrate the ranking result of the test fasteners.

1) *Topic Assignment Result*: The training process of STM clusters the samples into topics according to their structure configurations, and the models of each topic can be simultaneously acquired. We display the assignment results of 11 topics in Fig. 13, where only four samples are given for each topic due to the lack of space. It is obvious that each topic represents a type of fastener with the same orientation and illumination condition. The labels of each topic are manually provided.

2) *Fastener Classification*: We display the classification results for the test samples in this section. The test set is a mixture of different types of fasteners as well as several worn and missing ones. The test images are achieved by the fastener detection algorithm, and the size is  $100 \times 160$  pixels. As we introduced in Section IV-D, a subwindow of the size  $84 \times 110$  pixels is scanned on the test images to compute the likelihoods for each fastener model (topic). These probabilities can be displayed after normalization, as shown in Fig. 14. Then, the class (topic) of the test sample is determined by the maximum likelihood among all the maps. In this case, the fastener is classified to Topic 6.

The performance of fastener classification is evaluated by recall and precision [26], [27]. Recall and precision are widely used criteria for evaluating the performance of classification or detection algorithms and they are defined as

$$\text{Recall} = \frac{tp}{N} \quad (9)$$

$$\text{Precision} = \frac{tp}{tp + fp} \quad (10)$$

where  $tp$  is to the number of correctly classified samples for a class,  $fp$  is the number of unexpected samples classified

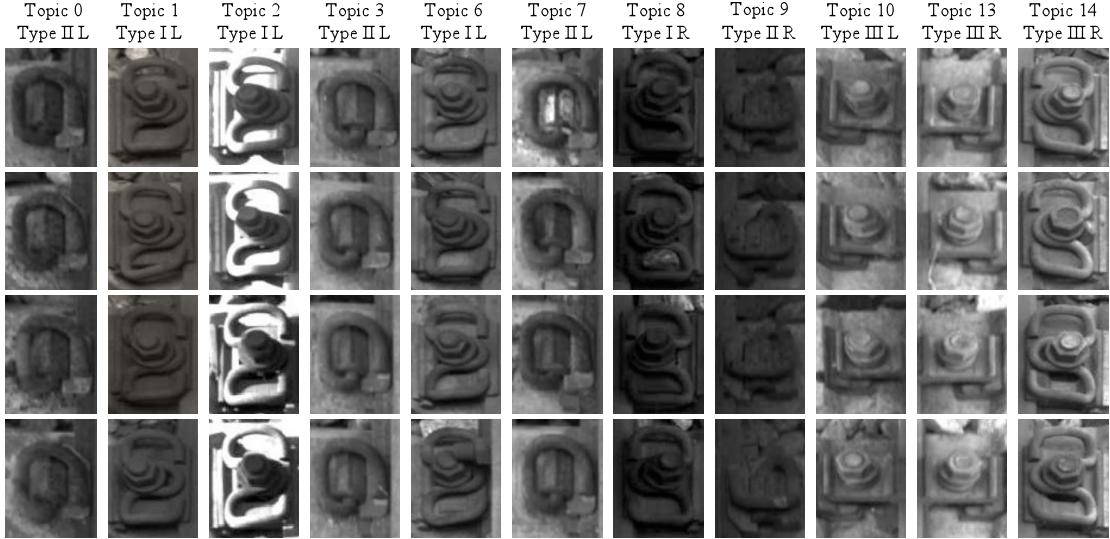


Fig. 13. Topic assignment results for some training samples. Here, we select 11 topics. Four samples are displayed for each topic. The labels of each topic are provided. L: the left fastener. R: the right fastener.

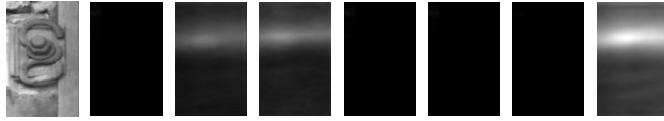


Fig. 14. Normalized probability maps for a test image. The probability maps are normalized to 0–255. Lighter pixels show higher probabilities. From left to right, we display the original test image and the probability maps for Topic 0–6. Due to the lack of space, only first six maps are displayed.

to this class, and  $N$  is the total number of the test samples for this class. In other words, recall refers to the percentage of correctly classified instances and precision reflects the accuracy of the classification for a class. The evaluation result is shown in Table II. It should be noted that the fasteners of the same type are counted together. From this table, we can observe that all the types have high precisions and recalls.

We evaluate the classification results for partially worn and missing fasteners. The partially worn fastener has a few missing or obscured parts, but they are visually recognizable. On the other hand, the missing fastener loses its major component such as hook. They are unrecognizable and most of them are represented as flat image region. Examples of these two cases are shown in Fig. 2. All the worn fastener samples are selected for test. There are 3312 partially worn fasteners and 746 missing fasteners. Our method achieves 99.5% classification precision for partially worn fasteners. The examples of misclassified samples are shown in Fig. 15(a). It can be observed that they are obscured by stones or wires. On the other hand, the classes of missing fasteners are unable to determine. The average classification precision is only 21.3%. Some examples of missing fastener are shown in Fig. 15(b). It should be noted that the misclassification of missing fasteners will not affect the final wear ranking (see Section IV-D, Fig. 16), since these fastener cannot match to any model. Meanwhile, the misclassification can be rectified if necessary. Generally speaking, the fasteners, which installed on the same

TABLE II  
EVALUATION RESULT FOR FASTENER CLASSIFICATION USING STM

Type	Recall	Precision
Type-I Left	99.6%	99.9%
Type-1 Right	99.7%	99.9%
Type-II Left	99.4%	99.9%
Type-II Right	99.3%	99.3%
Type-III Left	97.3%	97.9%
Type-III Right	97.4%	96.5%

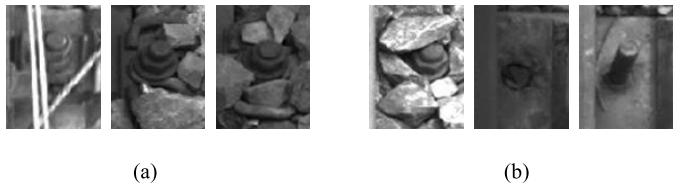


Fig. 15. Examples of misclassified samples. (a) Partially worn fasteners. (b) Missing fasteners.

sleeper, are of the same type. Thus, the misclassified samples can be easily identified and corrected.

For comparison, the classifiers such as boosted tree [6], latent SVM [28], and neural network [7] are trained and tested

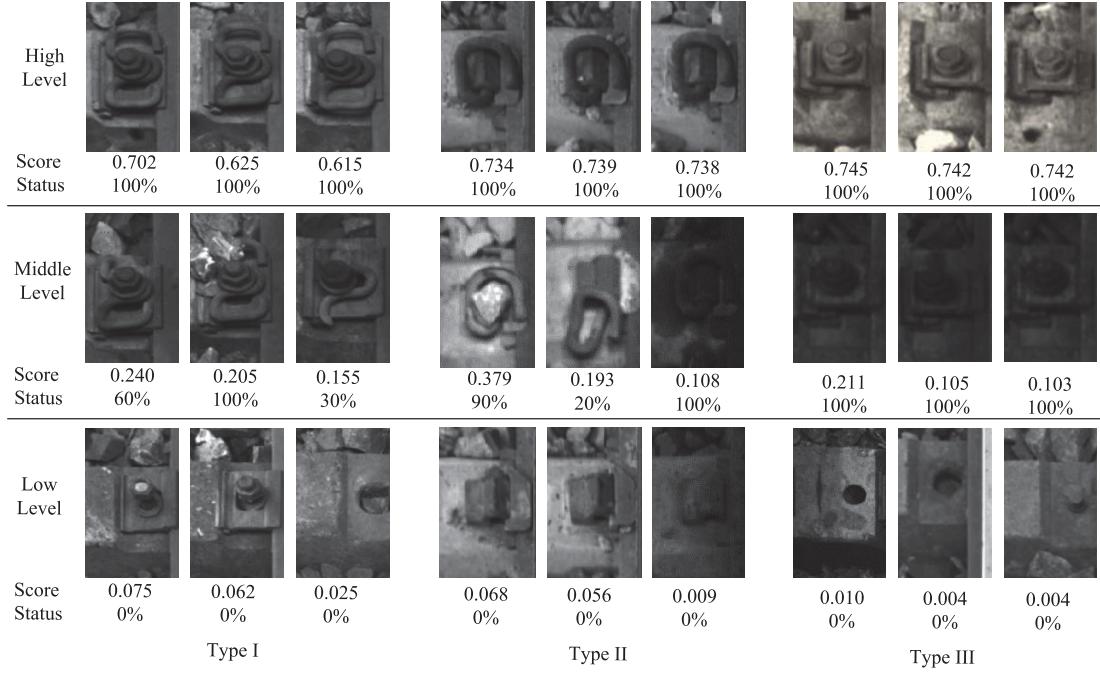


Fig. 16. Examples of ranking result.

TABLE III  
TOTAL CLASSIFICATION PRECISION OF STM, BOOSTED TREE, LATENT SVM, AND NEURAL CLASSIFIER

Method (Feature)	STM (Haar-like)	Boosted Tree (HOG)	Latent SVM (HOG)	Neural Network (Gabor wavelet)
Precision	98.9%	93.4%	94.4%	92.5%

on the same data set. In this experiment, the classifiers consider seven classes corresponding to three types of fasteners with two orientations and a nonfastener class. There are 100 training samples for each fastener class and 1000 samples for nonfastener class. Other samples are used for test. We use the HOG descriptor [28] and Gabor wavelet [15] to extract structural information of fasteners. The total precisions are shown in Table III. Compare with STM, other methods achieve lower precision rates. There are two reasons for this result: 1) the illumination variations of the training samples significantly affect the performance of classifiers and 2) the shapes of some fasteners are similar, and the classifiers and features cannot distinguish them effectively. For STM, the number of classes is automatically determined by the training samples. As a result, the fasteners of different illumination conditions are modeled separately. On the other hand, STM places strict spatial constraints on fastener shapes. Thus, the model is able to identify the differences between similar shapes. We can conclude that the proposed method is able to model various types of fasteners in different visual conditions.

3) *Wear Ranking*: To identify the worn fasteners, the likelihoods of samples are normalized and ranked. The likelihood

can be understood as a score, which measures the consistency between a fastener and a model. The lower the score is, the worse the fastener condition. As we introduced above, the likelihood are ranked and grouped into three levels. The middle and low levels need to be concerned. In the low level, the fasteners are significantly different from their models, while most of them correspond to the missing ones. The fasteners in the middle level are partially worn, polluted, or covered with dirty. We suggest that these fasteners require further examination by human inspector to accurately determine their states. The score ranges of each level are defined as follows. The score of high level is in the range 0.5–1.0, which means that more than half of a fastener can be matched to the model. The score range of middle level is between 0.1 and 0.5. This range denotes that only a small part of the fastener can be recognized. Finally, the score range of low level is from zero to 0.1.

Some examples of the ranking result are shown in Fig. 16. Due to the lack of space, we only show the results of left fasteners. The normalized scores are given below the image. We also offer the manual inspection results (ground truth) shown as the percentage of completeness. Among the

TABLE IV  
DEFECT DETECTION PRECISIONS OF DIFFERENT METHODS

Method	Our method (Likelihood ranking)	Our method (SVM)	Adaboost [6]	BP [7] [8]	Point Matching [13]
Precision	99.4%	98.7%	98.1%	93.2%	91.1%

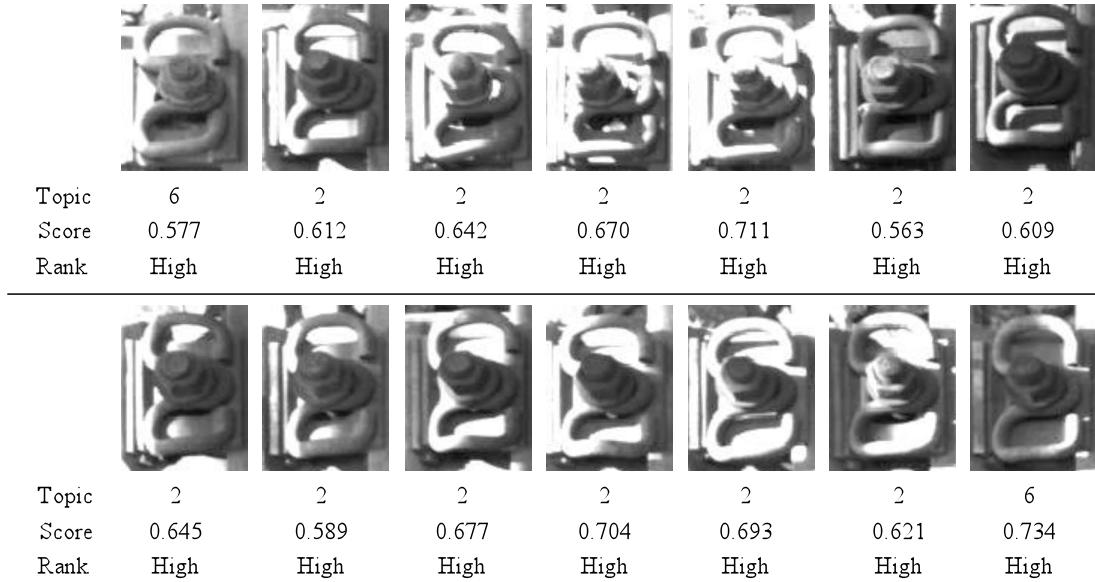


Fig. 17. Classification and ranking results for the fasteners with illumination variation.

399 078 samples, there are 392 518 (98.4%) fasteners ranked into high level. We have verified that all these fasteners are in good condition. Second, there are 5726 (1.4%) samples ranked into middle level. These fasteners belong to four major situations: 1) partly worn fasteners; 2) partly shielded fasteners; 3) polluted fasteners; and 4) good fasteners. It should be noted that the shielded fasteners are considered as worn ones since their true conditions are unknown. In addition, there is also a large amount of good fasteners (2502 fasteners) ranked into middle, and this may be due to the severe variations in shape and illumination condition. Finally, 834 (0.2%) fasteners are ranked as low, they are all seriously damaged. As mentioned above, the missing fasteners are always misclassified because no information is available. This problem did not affect the ranking result, because they cannot match to any model and their likelihood probabilities are very low. As a result, they are recognized as low level.

Compare with discriminative model, our approach offers a fuzzy class (middle level). We believe that this fuzzy class is useful in this application. The fasteners in middle level are in border condition and they are possibly misclassified by traditional discriminative models. As a result, some railway companies still require periodical reexamination of all the fastener images by human inspector. In our approach, the conditions of fasteners can be evaluated quantitatively.

The fasteners in border condition are also collected. Thus, the inspectors can focus on these indecisive cases. On the other hand, the thresholds between each two levels are flexible. They can be adjusted based on other statistical criteria or learned by classifiers.

To prove the effectiveness of our approach, we consider the wear ranking as classification problem and compare the total precision with some state-of-the-art methods. To calculate the precision, the worn fasteners in the middle and low levels are considered as correct classification. Furthermore, we also trained SVM linear classifier on the normalized likelihood values for each class (topic). We select 1000 samples to train the classifier-based approaches (500 intact fasteners and 500 worn fasteners), others are used for evaluation. Because other methods can only handle single type of fastener, this test is performed on Type-I left samples. The total precisions are listed in Table IV. It can be seen from this table that the performance of our method is better than others. There are two reasons: 1) other methods cannot identify the partly worn fasteners and 2) it is difficult for them to handle the severe illumination variation.

It should be noted that the test samples cover the fasteners with continuous brightness variation. We select 677 samples, which are affected by sunlight to evaluate the performance of our method. In Fig. 17, we show the topic assignment results

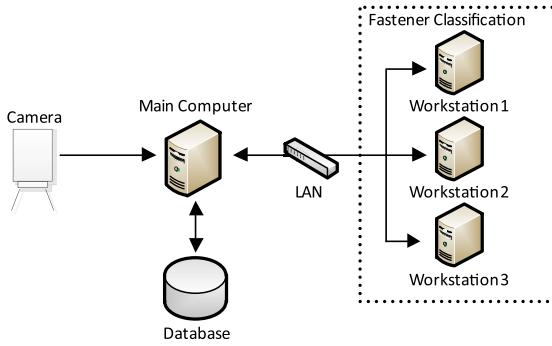


Fig. 18. System configuration for parallel computation.

and ranking results of a serial of fasteners shot under different sun azimuth. It can be observed that most of the fasteners are assigned Topic 2, which corresponds to the fasteners of very high brightness. For these samples, the ranking precision is 100%. On the other hand, the dark fasteners can be matched to other topics, which correspond to their brightness level. We believe that our method is effective to identify the fasteners with various illumination conditions. There are two main reasons for this success. First, STM uses multiple models to represent the fasteners in different appearances. On the other hand, the bag-of-feature allows a word to represent a range of feature values. Then, the images with different brightness may correspond to the same word assignment.

### C. Speed

The speed of our method can be easily accelerated using parallel computing technologies such as multicore processor and distributed system. In our implementation, the images are distributed to different computers (processor cores) to run the fastener classification algorithm (inference of STM) in parallel. There are four computers in our system. The main computer, which is installed with a single Intel Core-i7-3770K processor (four cores and eight threads), is responsible for image acquisition, task distribution, and ranking. Others are 2U blade workstations, and each of which is configured with eight Intel Xeon L5520 processors (32 cores and 64 threads). They are used to perform fastener classification.

First, the main computer acquires images and saves them to database. Once 1200 images are arrived, they are allocated to a workstation through LAN. For each workstation, the fastener detection algorithm is first performed, and then the fastener images are sequentially allocated to different cores for calculating the likelihood probabilities based on the trained STM. The class labels and likelihood probabilities are finally returned to the main computer. The current system configuration is shown in Fig. 18. The fastener localization speed for a single computer is 0.007 s per fastener, the fastener classification speed for a single computer is 0.065 s per fastener, and the average defect detection speed of our system is 262.3 km/h. It should be noted that this speed is computed by the distance between the first and last fasteners divided by the total processing time. The speed can be further accelerated when more workstations are available.

## VI. CONCLUSION

The detection of worn and missing fasteners is an important task in railway inspection. However, the manual inspection is of poor efficiency. On the other hand, the earlier automatic inspection systems based on classifiers are of low reliability.

In this paper, a novel railway inspection system is proposed, which is able to simultaneously assess the damage of multiple types of fasteners. Relying on the topic model, the proposed inspection system has the following three major advantages: 1) different types of fasteners can be simultaneously modeled using unlabeled data; 2) the system is robust to illumination changes; and 3) the statuses of fasteners are ranked. Technically, we introduce a new topic model named STM to model the structures of fasteners. Possibly, STM is the first probabilistic topic model aiming at representing object structure. By which, the modeling of diverse types of fasteners becomes much easier.

The detailed evaluation on railway lines is provided. The proposed method has very high performance on recognizing good fasteners as well as detecting worn ones.

## ACKNOWLEDGMENT

The authors would like to thank the China Academy of Railway Sciences for offering the data.

## REFERENCES

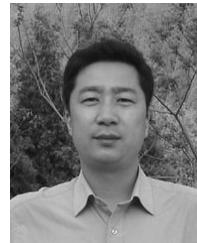
- [1] C. Alippi, E. Casagrande, F. Scotti, and V. Piuri, "Composite real-time image processing for railways track profile measurement," *IEEE Trans. Instrum. Meas.*, vol. 49, no. 3, pp. 599–564, Jun. 2000.
- [2] Q. Yang and J. Lin, "Track gauge dynamic measurement based on 2D laser displacement sensor," in *Proc. Int. Conf. Mech. Autom. Control Eng.*, Chengdu, China, Jul. 2011, pp. 5473–5476.
- [3] C. Alippi, E. Casagrande, M. Fumagalli, F. Scotti, V. Piuri, and L. Valsecchi, "An embedded system methodology for real-time analysis of railways track profile," in *Proc. 19th IEEE Instrum. Meas. Technol. Conf.*, vol. 1. Anchorage, AK, USA, May 2002, pp. 747–751.
- [4] R. Edwards, S. Dixon, and X. Jian, "Characterisation of defects in the railhead using ultrasonic surface waves," *NDT & E Int.*, vol. 39, no. 6, pp. 468–475, Sep. 2006.
- [5] 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, Apr. 2004.
- [6] Y. Xia, F. Xie, and Z. Jiang, "Broken railway fastener detection based on adaboost algorithm," in *Proc. Int. Conf. Optoelectron. Image Process.*, vol. 1. Haiko, China, Nov. 2010, pp. 313–316.
- [7] 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.
- [8] E. Stella, P. Mazzeo, M. Nitti, G. Cicirelli, A. Distante, and T. D'Orazio, "Visual recognition of missing fastening elements for railroad maintenance," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, Singapore, Sep. 2002, pp. 94–99.
- [9] J. Yang, W. Tao, M. Liu, Y. Zhang, H. Zhang, and H. Zhao, "An efficient direction field-based method for the detection of fasteners on high-speed railways," *Sensors*, vol. 11, no. 8, pp. 7364–7381, Jul. 2011.
- [10] P. De Ruvo, A. Distante, E. Stella, and F. Marino, "A GPU-based vision system for real time detection of fastening elements in railway inspection," in *Proc. 16th IEEE Int. Conf. Image Process.*, Cairo, Egypt, Nov. 2009, pp. 2333–2336.
- [11] G. De Ruvo, P. De Ruvo, F. Marino, G. Mastronardi, P. L. Mazzeo, and E. Stella, "A FPGA-based architecture for automatic hexagonal bolts detection in railway maintenance," in *Proc. 7th Int. Workshop Comput. Archit. Mach. Percept.*, 2005, pp. 219–224.

- [12] Y. Rubinsztejn, "Automatic detection of objects of interest from rail track images," M.S. thesis, School Comput. Sci., Univ. Manchester, Manchester, U.K., 2011.
- [13] Y. Li, C. Otto, N. Haas, Y. Fujiki, and S. Pankanti, "Component-based track inspection using machine-vision technology," in *Proc. 1st ACM Int. Conf. Multimedia Retr.*, Trento, Italy, 2011, no. 60.
- [14] P.L. Mazzeo, N. Ancona, E. Stella, and A. Distante, "Visual recognition of hexagonal headed bolts by comparing ICA to wavelets," in *Proc. IEEE Int. Symp. Intell. Control*, Houston, TX, USA, Oct. 2003, pp. 636–641.
- [15] 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, Oct. 2004.
- [16] M. Singh, S. Singh, J. Jaiswal, and J. Hempshall, "Autonomous rail track inspection using vision based system," in *Proc. IEEE Int. Conf. Comput. Intell. Homeland Security Personal Safety*, Alexandria, VA, USA, Oct. 2006, pp. 56–59.
- [17] D. Blei, A. Ng, and M. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Jan. 2003.
- [18] B. Zhao, L. Fei-Fei, and E. P. Xing, "Image segmentation with topic random field," in *Proc. 5th Eur. Conf. Comput. Vis.*, Heraklion, Greece, Sep. 2010, pp. 785–798.
- [19] X. Wang, X. Ma, and W. E. L. Grimson, "Unsupervised activity perception in crowded and complicated scenes using hierarchical Bayesian models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 3, pp. 539–555, Mar. 2009.
- [20] L. Fei-Fei and P. Perona, "A Bayesian hierarchical model for learning natural scene categories," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2005, pp. 524–531.
- [21] R. G. von Gioi, J. Jakubowicz, J. Morel, and G. Randal, "LDA: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, Apr. 2010.
- [22] P. Orbanz and J. M. Buhmann, "Nonparametric Bayesian image segmentation," *Int. J. Comput. Vis.*, vol. 77, nos. 1–3, pp. 25–45, May 2008.
- [23] T. Griffiths and M. Steyvers, "Finding scientific topics," *Nat. Acad. Sci. USA*, vol. 101, pp. 5228–5235, Apr. 2004.
- [24] E. B. Sudderth, A. Torralba, W. T. Freeman, and A. S. Willsky, "Describing visual scenes using transformed objects and parts," *Int. J. Comput. Vis.*, vol. 77, nos. 1–3, pp. 291–330, May 2008.
- [25] G. Herinrich, "Parameter estimation for text analysis," Faculty Math. Comput. Sci., Univ. Leipzig, Leipzig, Germany, Tech. Rep. 2.4, 2008.
- [26] D. Olson and D. Delen, *Advanced Data Mining Techniques*, 1st ed. Berlin-Heidelberg, Germany: Springer-Verlag, 2008.
- [27] D. M. W. Powers, "Evaluation: From precision, recall and F-factor to ROC, informedness, markedness & correlation," *J. Mach. Learn. Technol.*, vol. 2, no. 1, pp. 37–63, 2011.
- [28] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1627–1645, Sep. 2010.



**Hao Feng** received the B.Eng. degree from the Beijing University of Technology, Beijing, China, in 2007. He is currently pursuing the Ph.D. degree with the Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing.

His current research interests include computer vision, pattern recognition, and machine learning.



**Zhiguo Jiang** received the B.Eng., M.S., and Ph.D. degrees from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 1987, 1990, and 2005, respectively.

He is currently a Professor with the Image Processing Center, Beijing University of Aeronautics and Astronautics. He has been a Vice Dean with the School of Astronautics since 2006. His current research interests include medical image processing and remotely sensed image processing.



**Fengying Xie** received the B.Eng. degree from the Department of Computer Science, Daqing Petroleum Institute, Daqing, China, in 1997, and the M.S. and Ph.D. degrees in Pattern Recognition and Intelligent System from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2002 and 2009, respectively.

She was a Visiting Scholar with the Laboratory for Image and Video Engineering, University of Texas at Austin, Austin, TX, USA, from 2010 to 2011. She is currently an Associate Professor with the School of Astronautics, Beijing University of Aeronautics and Astronautics. Her current research interests include biomedical image processing, image analysis and segmentation, object detection, and recognition.



**Ping Yang** received the B.Eng. degree from the Beijing University of Chemical Technology, Beijing, China, in 2010. She is currently pursuing the M.S. degree with the Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing.

Her current research interests include computer vision pattern recognition and machine learning.



**Jun Shi** received the B.S. degree from Huainan Normal University, Huainan, China, in 2007, and the M.S. degree from Yangzhou University, Yangzhou, China, in 2011. He is currently pursuing the Ph.D. degree with the Image Processing Center, School of Astronautics, Beijing University of Aeronautics and Astronautics, Beijing, China.

His current research interests include computer vision, pattern recognition, and machine learning.



**Long Chen** received the B.Eng. degree from the Beijing University of Aeronautics and Astronautics, Beijing, China, in 2005, where he is currently pursuing the Ph.D. degree with the Image Processing Center, School of Astronautics.

His current research interests include object detection and recognition and probabilistic graphical models.