3-2016

Detection of bird nests in overhead catenary system images for high-speed rail

Xiao WU

Ping YUAN

Qiang PENG

Chong-wah NGO
Singapore Management University, cwngo@smu.edu.sg

Jun-Yan HE

Pattern Recognition 51 (2016) 242–254

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Detection of bird nests in overhead catenary system images for high-speed rail

Xiao Wu [a,*], Ping Yuan [a], Qiang Peng [a], Chong-Wah Ngo [b], Jun-Yan He [a]

[a] School of Information Science and Technology, Southwest Jiaotong University, No. 111, North Section 1, 2nd Ring Road, Chengdu, China
[b] Department of Computer Science, City University of Hong Kong, #83, Tat Chee Avenue, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 5 September 2014
Received in revised form
13 June 2015
Accepted 13 September 2015
Available online 28 September 2015

Keywords:
Bird nest detection
Image classification
Overhead catenary system
High-speed rail
Intelligent transportation system

ABSTRACT

The high-speed rail system provides a fast, reliable and comfortable means to transport large number of travelers over long distances. The existence of bird nests in overhead catenary system (OCS) can hazard to the safety of the high-speed rails, which will potentially result in long time delays and expensive damages. A vision-based intelligent inspection system capable of automatic detection of bird nests built on overhead catenary would avoid the damages and increase the reliability and punctuality, and therefore is attractive for a high-speed railway system. However, OCS images exhibit great variations with lighting changes, illumination conditions and complex backgrounds, which pose great difficulty for automatic recognition. This paper addresses the problem of automatic recognition of bird nests for OCS images. Based on the unique properties of bird nests, we propose a novel framework, which is composed of five steps: adaptive binarization, trunk/branch detection, hovering point detection, streak extraction and pattern learning, for bird nest detection. Two histograms, Histogram of Orientation of Streaks (HOS) and Histogram of Length of Streaks (HLS), are novelly proposed to capture the distributions of orientations and lengths of detected twig streaks, respectively. They are modeled with Support Vector Machine to learn the patterns of bird nests. Experiments on different high-speed train lines demonstrate the effectiveness and efficiency of the proposed work.

1. Introduction

High-speed trains are well developed around the world due to their numerous advantages. Indeed, they are safe, reliable, sustainable, convenient and comfortable for passengers. Over the past few years, China's high-speed rail (HSR) network has progressively expanded and now become the world's longest high-speed rail network with around 9300 km of routes [1]. With the fast development of high-speed rail network, railway officials and managers are facing arduous tasks to ensure that the high-speed rail system is operating in an orderly and reliable way. Among them, safety has the highest priority for high-speed rail system, especially after a fatal high-speed railway accident happened near Wenzhou, China on July 23, 2011, which has caused great concerns on the safety of high-speed rail network.

The Pantograph–Catenary (PAC) system is the dominant form for supplying the vital power to railway electrical trains. A pantograph is an apparatus mounted on the roof of an electric train to collect power through contact with an overhead catenary equipment called the Overhead Catenary System (OCS) [2]. The steel rails on the tracks act as the electrical return. The OCS is a high voltage system consisting of contact wire and catenary wire suspended via supports primarily on poles placed along the railway. The OCS includes messenger wire, contact wire, droppers, and supporting structure, which consists of metallic poles, cross-arms, and running rails. The structure of OCS is illustrated in Fig. 1. In order to achieve good current collection, the contact wire has to be placed geometrically within defined limits, which is usually achieved by supporting the contact wire from above by a second wire known as the messenger wire or catenary. This wire is attached to the contact wire at regular intervals by vertical wires known as droppers or drop wires. The messenger wire is supported regularly at structures, by a pulley, link, or clamp. Due to space limitation, we will not elaborate the details of each component.

Despite offering a balance in cost effectiveness and system reliability, the important railway power supply chain is a major cause of train failure faults. The defects in OCS, for example, down hanging, ripped off droppers, bondings, broken insulators and bird nests, can result in long time delays, expensive damages, and even disasters. The existence of bird nests greatly threats the safety of the high-speed rail network. It will lead to a defective uptake of energy by the locomotive, resulting in significant energy loss and damage to the overhead contact wire and pantograph. While this

* Corresponding author.
E-mail address: wuxiaohk@home.swjtu.edu.cn (X. Wu).

http://dx.doi.org/10.1016/j.patcog.2015.09.010
0031-3203/

Fig. 1. Structure of overhead catenary system of high-speed rail.

is a problem common to all trains of electric traction, it is especially critical for high speed tracks, since the reliability of power collection decreases as train speed increases. It is the permanent aim of all railway line operators to detect faults at an early stage, and to correct any damage, fault or wear on time in order to prevent serious disturbances of railway traffic. Therefore, it is greatly desired that intelligent approaches can be designed and implemented to provide support for technicians. In this paper, we will focus on the automatic detection of bird nests for high-speed rails.

The demand for safety and high reliability grows with the development of high-speed railway lines, leading to the invention of new monitoring devices. The OCS inspection system is one of the key items for high-speed rail system. Traditionally, it is implemented by laser scanner on top of the train. Compared to the extremely expensive laser scanner, a breakthrough video monitoring system with on-board high-resolution cameras equipped in the high-speed trains becomes a promising solution due to its high functionality, flexibility and interoperability. It can directly capture the OCS images/videos to ensure the reliability of the OCS. Conventionally, the images/videos have to be visually checked and evaluated offline by trained technicians, often frame by frame, to inspect the status of the OCS. It is time-consuming, laborious, and impractical to manually monitor millions of overhead catenary supporters along thousand miles of railways. In this work, we deploy an OCS image inspection system featured with non-contact cameras for high-speed trains.

Unfortunately, we notice that OCS images exhibit great variations in lighting changes, illumination conditions, occlusion, complex structures, and mixture of foreground and background, which make automatic detection of bird nests a challenging task. First, the images from different train lines are captured with varied weather, lighting changes, viewpoints and illumination conditions. Second, bird nests exist at different locations of the OCS. And the regions having bird nests are relatively small and not easy to be noticed. Third, the overhead catenary structures are complex with messy crossing lines, especially when the foreground and background are mixed together, so that many lines are crossed over and overlapped. Moreover, discerning catenary system from complex background such as mountains, trees and buildings, is generally difficult. Therefore, it is extremely challenging for an automatic OCS image inspection system to function as expected under various practical concerns. Several representative OCS images with bird nests are demonstrated in Fig. 2, while some difficult examples are also shown in Fig. 10.

In this paper, we explore an automatic detection of bird nests for high-speed rails, which analyzes the captured OCS images and assists technicians to make decisions. Experiments demonstrate that the proposed approach achieves a promising performance for bird nest recognition. The contributions of this paper are as follows:

- To the best of our knowledge, this is the first work to systematically analyze the properties of bird nests in OCS images, and the first to automatically detect bird nests by image processing technology for OCS inspection in high-speed rail system.
- Based on the features of bird nests in OCS system, a five-phase framework is novelly proposed to detect bird nests in image sequences, which includes binarization, trunk and branch identification, hovering point detection, streak extraction and pattern learning.
- To model the properties of unordered, non-parallel, distributed and diverse twigs of bird nests, HOS and HLS histograms are proposed to represent the distributions of orientations and lengths of detected streaks, which are exploited to detect the presence of bird nests.
- Experiments on multiple sequences of images from real high-speed rail lines demonstrate the effectiveness and efficiency of the proposed approach.

This paper is organized as follows. Section 2 gives a brief overview of related work. Section 3 introduces the system architecture of the OCS inspection system, and the proposed framework of automatic bird nest detection system. Section 4 elaborates the detailed process of bird nest detection. Section 5 describes the experimental setup and empirical results. Finally, Section 6 concludes this paper.

2. Related work

2.1. Pantograph-Catenary inspection system

Conventional OCS inspection systems use physical instruments mounted on an inspection vehicle to measure the status of pantograph and catenary. The Fiber Bragg Grating (FBG) sensors on a pantograph are used in [3] to monitor the underground pantograph-catenary system, which measure the contact force and the vertical acceleration of the pantograph head. An automatic diagnostic system is installed in a special trolley running up to 100 km/h to check the health conditions of catenary [4]. Particular considerations are given on the values and the trends of voltage, height, stagger, wear, forces and so on. These instruments are mounted on the pantograph itself and signal wires are attached to each instrument [4,5], which affect the dynamic characteristics of the pantograph. An optical radar system [6] is equipped on the inspection vehicle to record catenary related parameters like contact wire position, wire wear, pole position as well as distance between pole and track. However, these physical inspection systems mainly focus on checking the status and properties of catenary systems, which are not applicable for bird nest detection.

With the recent development of computerized image recognition, current works begin to focus on video based detection to discover defects in OCS [7–15], taking advantage of the video

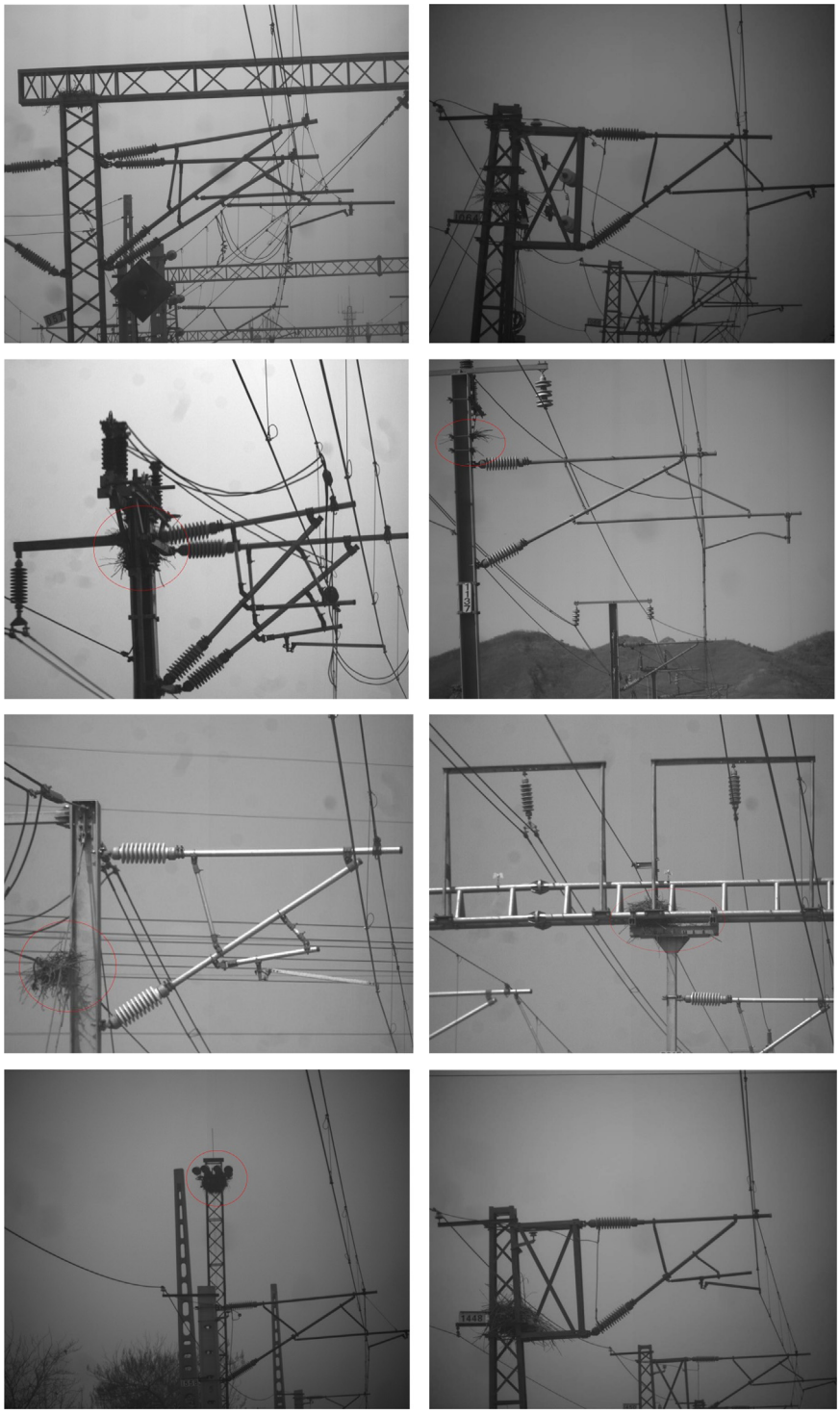

Fig. 2. Examples of bird nests appearing at different locations of the overhead catenary system of high-speed rail. The image qualities are suffered from different variations in lighting changes, complex structures, mixture of foreground and background, which make automatic detection of bird nests a challenging task.

camera that offers higher resolution and lower cost. The most related works to our research are [8,9]. With IRIS 320 inspection vehicle in France [8], image analysis strategy is adopted for catenary inspection for preventive maintenance. It locates the catenary sections between two supporting arms, and extracts all surrounding elements such as contact wire, carrying wire, supporting arms and simple droppers. The work [9] deals with the automatic recognition of the catenary elements, such as supporting arms, droppers, and droppers with electrical connection. It consists of two steps: a classical scene analysis is conducted to identify elements by segmentation in vertical and horizontal components, followed by feature extraction and classification. Next, the results are checked and analyzed by analyzing the catenary element sequence using a Markov model. In [11], a video-based obstacle detection system is developed in Germany, which automatically detects obstacles in the pantograph gauge and retracts the pantograph before collision with an obstacle. With the recognition of steady arms, the reliability of the system is improved considerably. A system for the automatic detection of droppers in catenary staves is proposed in [13]. Based on a

top-down approach, the system exploits priori knowledge to perform reliable extraction of droppers. In [10], the problematic of contact wire wear in railways is first presented and a computer vision system is studied to improve the precision of the systems for wear measurement of contact wire. A sensor system with a line-scan camera is used to monitor the interaction between the catenary and the pantograph [14], which can detect the occurrence of defects in the catenary-pantograph interaction. In [15], the pantograph-overhead contact wire system is investigated by using an infrared camera. In order to detect the temperature along the strip from a sequence of infrared images, a segment-tracking algorithm based on the Hough transform is employed. It helps maintenance operations in the case of overheating of the pantograph strip, bursts of arcing, or an irregular positioning of the contact line. Although the aforementioned works focus on identification of the components and defects for OCS systems, direct comparison with them is not possible because these works are not tailored for bird nest detection. To the best of our knowledge, few research works and systems are dedicated to the bird nest detection for the topic of automatic video based diagnostics at catenaries.

2.2. Image classification and abnormal pattern detection

Bird nest detection for OCS images belongs to the research issues of image classification and object detection, which are key research areas in computer vision and image processing. Image classification [16] and object detection [17,18] have been extensively studied for decades and been applied to different areas, such as adverse weather detection [19], human detection [20–22], and intelligent transportation [23,24]. An image is classified according to its visual content, e.g., the existence of bird nests or not. Element-independent features such as color, shape, texture, gradient or contour [16] are extracted from the image in the first place. Then in the recognition process, these features are gathered and knowledge is incorporated in order to identify the objects.

Recently, an object detection system is described in [17], which represents highly variable objects using mixtures of multiscale deformable part models. These models are trained using a discriminative procedure that only requires bounding boxes for the objects in a set of images. The system attains state-of-the-art results in terms of efficiency and accuracy. In order to localize and segment objects, R-CNN is proposed in [18], which applies high-capacity convolutional neural networks (CNNs) for bottom-up region proposals. In addition, supervised pre-training on a large auxiliary dataset, followed by domain-specific fine-tuning on a small dataset, is an effective paradigm for learning high-capacity CNNs when data are scarce. This method yields a significant performance boost compared to state-of-the-art technologies. A compositional model for human detection is built in [20] by exploiting the analogy between human body and text. A discriminative alphabet is automatically learnt to represent body parts. Based on this alphabet, the flexible structure of human body is expressed by means of symbolic sequences, which correspond to various human poses and allow for robust and efficient matching. Experiments on standard benchmarks demonstrate that the proposed algorithm achieves state-of-the-art or competitive performance. A simple yet effective detector for pedestrian detection is proposed in [21], which incorporates common sense and everyday knowledge into the design of simple and computationally efficient features. A statistical model of the up-right human body is deployed where the head, the upper body, and the lower body are treated as three distinct components, from which a pool of rectangular templates is tailored to this shape model. Since different kinds of low-level measurements are incorporated, the resulting multi-modal and multi-channel Haar-like features represent characteristic differences between parts of the human body, which are robust against variations in clothing or environmental settings.

To identify bird nests in an image of OCS scene, previously explored works on image classification and object detection could be beneficial for our study. To represent the distinct patterns of various applications, different kinds of histograms are extracted and trained for image classification. A system based on computer vision is presented in [19] to detect the presence of rain or snow. A histogram of orientations of rain or snow streaks is computed with the method of geometric moments, which is assumed to follow a model of Gaussian uniform mixture. The orientation of the rain or the snow is represented with Gaussian distribution whereas the orientation of the noises is represented with uniform distribution. Expectation maximization (EM) is used to separate these two distributions. An image-based vehicle-type recognition is proposed in [24], which uses Gabor wavelet transform and the Pyramid Histogram of Oriented Gradients (PHOG) features. A reliable classification scheme is proposed by cascade classifier ensembles. A novel approach for near-duplicate keyframe identification is proposed in [25] by matching, filtering and learning of local interest points. Owing to the robustness consideration, the matching of local points across keyframes forms vivid patterns. Pattern entropy is proposed to capture the matching patterns with the histogram of matching orientation, and then learn the patterns with SVM for discriminative classification. Although the existing approaches on image classification may not be directly applicable for bird nest detection of OCS images, the idea of capturing the patterns in the form of histograms enlightens this research. In [26], a robust abnormal event detection framework based on sparse reconstruction over the normal bases is proposed. Given a collection of normal training examples, the sparse reconstruction cost (SRC) is proposed to measure the normalness of the testing sample. By introducing the prior weight of each basis during sparse reconstruction, the proposed measurement is more robust compared to other outlier detection criteria. An approach is proposed in [27] to detect aberrations in video streams using Entropy. It is estimated on the statistical treatments of the spatiotemporal information of a set of interest points within a region of interest, by measuring their degree of randomness of both directions and displacements. A framework is proposed in [28] to robustly identify local motions of interest in an unsupervised manner by taking advantage of group sparsity. In order to robustly classify action types, local motion is emphasized by combining local motion descriptors and full motion descriptors, and then group sparsity is applied to emphasize motion features using the multiple kernel method. Overall, there is no prior works on bird nest detection system presently available. The feasibility and effectiveness of the commonly used features and classification solutions are worth exploration.

3. System architecture and framework

In this section, we will first present the system architecture of the adopted OCS inspection system, followed by an introduction of the properties of bird nest. Finally, we will describe the proposed framework for the automatic detection of bird nests.

3.1. Architecture of OCS inspection system

The architecture of OCS inspection system is shown in Fig. 3. The inspection system is equipped with two on-board CCD cameras with resolutions of 2456×2058 (5 million pixels) and 1392×1040 (1 million pixels), respectively. The high resolution camera points toward the catenary to capture the images of OCS, while the low resolution one focuses on the supporter to capture

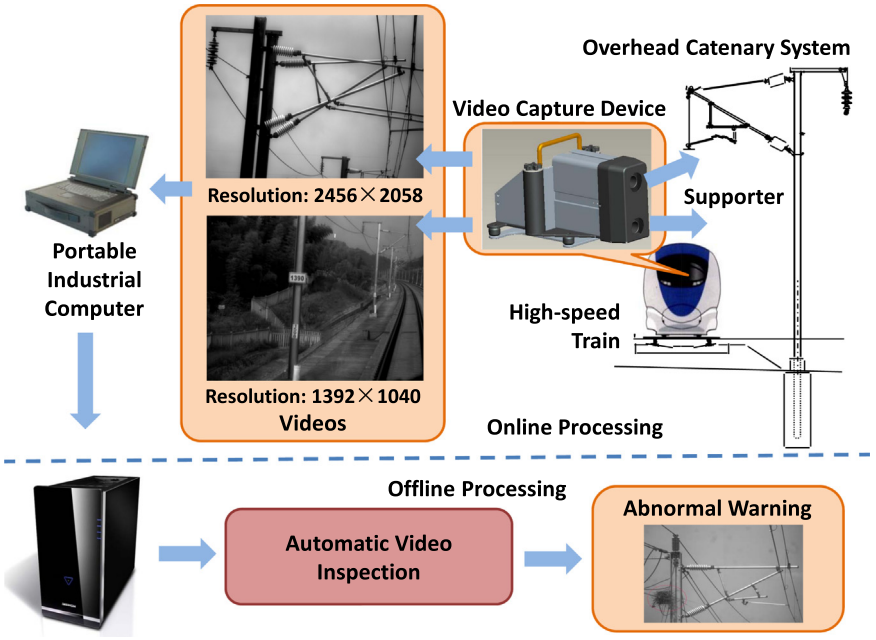

Fig. 3. System architecture of OCS inspection system.

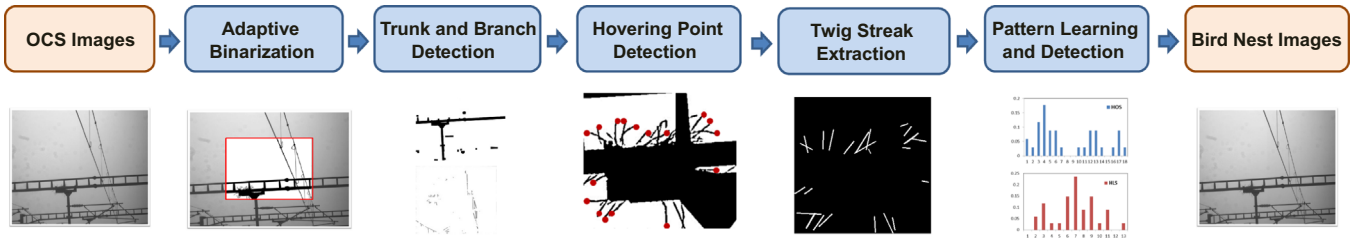

Fig. 4. System framework of the intelligent bird nest detection.

the pole ID and milestones, so that the technicians can locate the position of the supporters having bird nests. In this work, we only consider the OCS images captured with the high resolution camera. The images usually include the content of contact wires, messenger wires, insulators, and supporters. In order to best capture the OCS components in the middle of an image, the cameras are required to point directly toward the OCS structure. This device is placed in high-speed trains along the direction of the movement with a certain angle to capture the still images of the OCS, while the train is running with a maximum speed of 350 km/h. The two cameras simultaneously take photos with a rate of 17 frames per second. It produces a large number of images in each examination. The on-board inspection device is connected with a portable hard disk to store the captured images. The portable hard disk can be conveyed and connected to the offline PC for data analysis and processing, including automatic detection of bird nests, which is the focus of our work.

3.2. Properties of bird nest

We will first study the properties of bird nest. A bird nest is the place where a bird incubates its eggs and raises its youngs. It popularly refers to a specific structure made by the bird itself. For the bird nests in high-speed rail, the grassy cup nest is the most common type. The cup nest is smoothly hemispherical inside, with a deep depression to house the eggs. It is made of pliable materials, mainly grasses and plant fibers. In flight, the bird breaks a small twig from a tree and presses it into the saliva, angling the twig downwards, so that the central part of the nest is the lowest. It continues adding globs of saliva and twigs until it has made a crescent-shaped cup [29]. By our observation, bird nest demonstrates three prominent aspects:

- Bird nest is usually constructed on a foundation with strong and solid basis, such as the supporter or the portal structure in the OCS.
- A bird nest is composed of a bunch of short or long twigs with unordered, non-parallel organization. In other words, the groups of streaks belonging to a bird nest should be with distributed and diverse orientations.
- At the same time, the groups of streaks of a bird nest should be with inconsistent lengths.

Motivated by the properties of bird nest, we propose a novel framework for bird nest detection.

3.3. Framework of the proposed approach

Based on these features of a bird nest, the proposed approach mainly consists of five phases: pre-processing, trunk and branch identification, hovering point detection, twig streak extraction, and pattern learning of streaks. The purpose of these phases is to verify the presence of bird nest. The system framework of the proposed approach is illustrated in Fig. 4. To smooth the background and highlight the discrimination between background and catenary areas, an adaptive binarization step is first conducted. Trunk and streak identification is then performed to separate the trunks that the bird nest will be hosted and the detailed branches

that are the potential twigs of a bird nest. The hovering points are detected so that the twigs can be located. After that, the streaks are extracted using Hough transform to convert the twigs to straight lines. Two histograms, so-called HOS and HLS, are built by accumulating the orientations and lengths of different components obtained by the hovering point detection. The distributions of HOS and HLS are then modeled using SVM to discriminate the potential regions with or without a bird nest. A decision criterion on the histograms allows detecting the presence or absence of a bird nest. In this paper, branches, twigs and streaks have the same meaning, and are interchangeably mentioned depending on context.

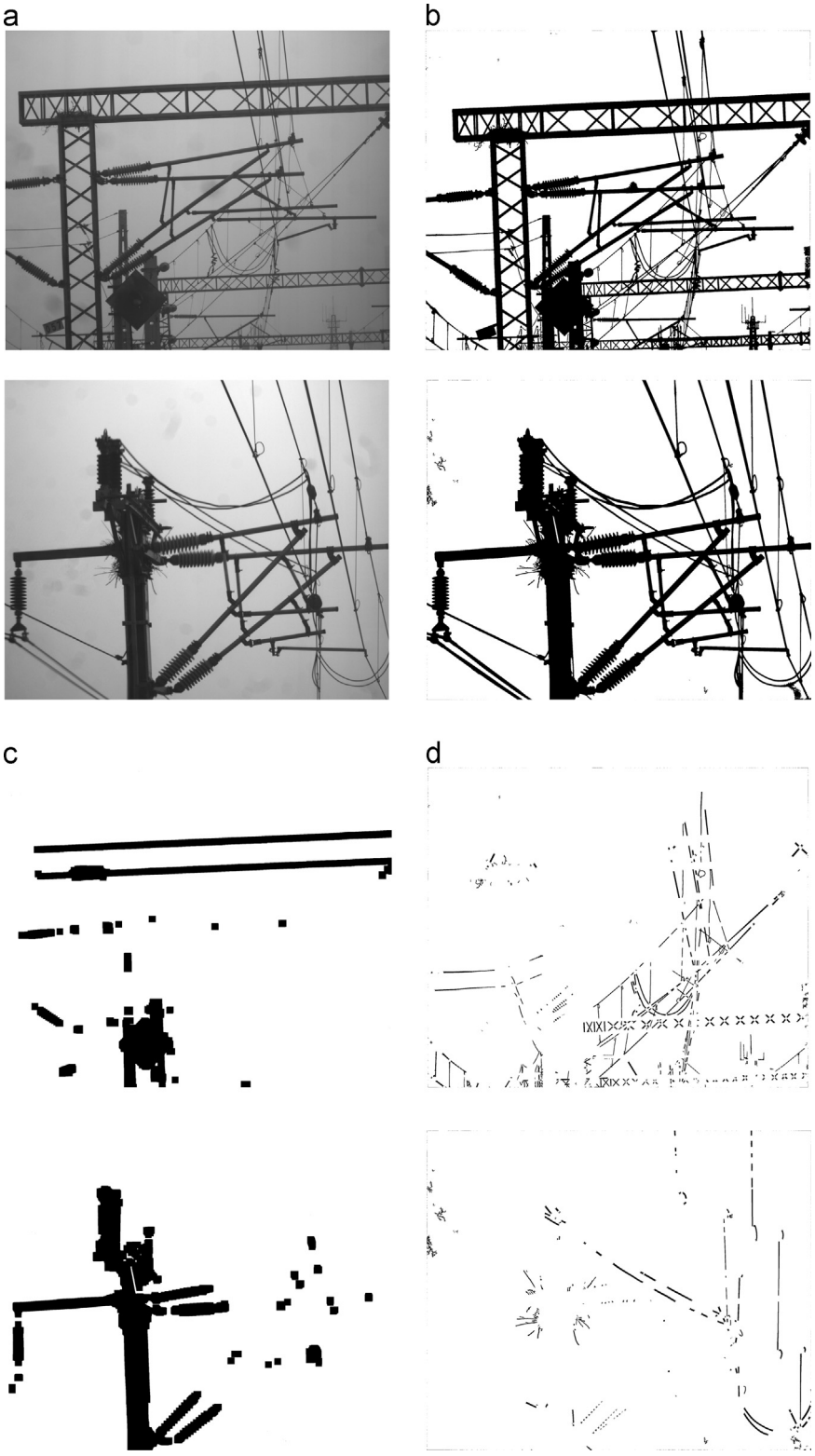

Fig. 5. Trunk and branch identification after image binarization. (a) Original images (b) Images after binarization. (c) Detected trunks (d) Detected branches.

4. Automatic bird nest detection for OCS images

4.1. Pre-Processing: Image binarization

Due to the convenience of non-contact on-board OCS inspection system, the portable device with high resolution cameras is mounted inside the driver's cabin instead of mounted on the roof of the train. Therefore, the dusts on the train windows are easily captured by the high-resolution cameras. In addition, the OCS images are always suffering from noises, uneven lighting and compression artifacts. The brightness distribution of various positions in an OCS image may vary because of the condition of catenary and the effect of lighting environment. These reasons make noise a crucial factor affecting the performance.

Binarization is the initial and a key step for automatic detection of bird nests, since it is the base for successful localization and recognition of bird nests. Its purpose is to distinguish the overhead catenary areas from the background areas and remove noises. In order to avoid the disadvantage of global threshold methods, which use a single threshold value to classify image pixels into objects and background classes, we use local adaptive binarization. It deploys multiple values selected according to the local area information. An OCS image is divided into multiple blocks according to an $n \times n$ sliding window, and for each block, the average value of all pixels in the sliding window is selected as the threshold.

In addition, we adopt two granularities of block size for the sliding window to detect the main structures and small details, respectively. The one with a larger size (e.g., 400×400) is to retain the shape of relatively large objects, such as the supporter and the portal structure, but ignore small details. On the contrary, the other one with a smaller size (e.g., 10×10) is to keep details and thin edge information, such as the streaks of bird nests. These two binarized images are then combined by a union operation to form the final image. Their combination perfectly contains the complete content of the overhead catenary and the details such as lines and strokes of bird nests, which improves the quality of catenary regions and preserves stroke connectivity by removing isolated pixels. A post-processing technique is used to eliminate noise pixels. Small connected regions with connected component labeling [30] are treated as noises and then removed. The original images and the effects after adaptive binarization are shown in Fig. 5(a and b), respectively. After this step, each image is converted from a gray scale to a binary image I_B with reduced noises and sharpened objects.

4.2. Trunk and branch identification

After image binarization, the next step is to identify the trunk and branch regions, which are the potential foundation of a bird nest built and the twigs of a bird nest, respectively. The trunk in the OCS images refers to the large regions such as the supporter and the portal structure of the overhead catenary system. Meanwhile, the branch denotes thin lines and small details.

To identify the trunk and branch, Canny edge detector [31] is performed on the binarized image to detect the edges, forming the edge image I_E, which makes the identification of the overhead contact lines, the poles, and the streaks of bird nests easier. The main part of the catenary and the portal structure is detected by morphological opening operation with rectangle kernel of $n \times n$, which tends to enlarge small holes, remove small objects, and separate objects. The pixels between two edges whose distance is less than $n/2$ will be filled. Otherwise, it will be kept unchanged. Finally, the trunk I_T can be achieved by combining the image I_E after the aforementioned operations with the binarized image I_B by a logical union operation. That is, the intersection of the detected main part and the binarized image will form the trunk regions, which is shown in Fig. 5(c). From this figure, we can see that the main part of the supporter and the portal structure of OCS can be correctly detected, which are the base where bird nests will be built up. Once the trunks are extracted, the left part is the thin lines and the twigs. The binarized image I_B is subtracted with the trunk image I_T, which forms the branches and the thin lines. The detected branch lines are shown in Fig. 5(d). We can see that the thin structures, including the streaks of bird nests, are well identified.

4.3. Hovering point detection

Once the catenary structure is identified, the following step is devoted to locate possible regions of a bird nest. The regions near the main truck are potential areas that the bird nests will be constructed. In addition, since the twigs of a bird nest are hovering in the air, the hovering points for each twig are first identified to check the existence of twigs, that is, the end of a twig. Ossification is undergone for the image after the aforementioned steps. Therefore, the width of each twig corresponds to one pixel. The next step is to find connected components. The points satisfying the following two conditions, *connectivity* and *end point*, are treated as hovering points. First, the point itself is on the twig. It is connected with other points and it is not an isolated point. Second, this point is at the end of a twig or a line. In other words, the number of connected points among its 8 neighbors is just one. The condition of a hovering point is defined as follows:

$$\text{(a). } Connectivity: \sum_{m=-1}^{1}\sum_{n=-1}^{1} P_C(x_0+m, y_0+n) = 2 \tag{1}$$

$$\text{(b). } End\ po\ int: P_C(x_0, y_0) = 1 \tag{2}$$

where (x_0, y_0) is the location of a point and $P_C()$ is the number of connected points in its 8 neighbors. For a hovering point, it has only one connected point, such as point t in Fig. 6. On the contrary, a non-hovering point has at least 2 connected points, such as points p and q in Fig. 6. An example of the detected hovering points is shown in Fig. 7, which are labeled in red color. If there are several hovering points in a sliding window, it is a potential region having a bird nest. However, it does not mean that it must be a bird nest in this region, such as the insulator regions.

4.4. Twig streak extraction using Hough transform

Once the hovering points have been identified, we will extract the twigs of a bird nest, which can be roughly represented as a set of straight lines with differentiated lengths. Due to imperfections in either the noises or the edge detector, there may be missing

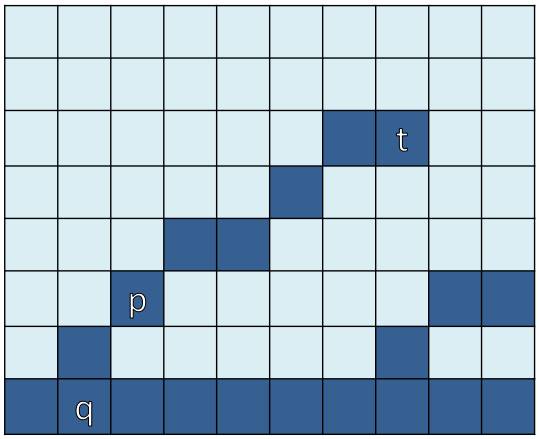

Fig. 6. Hovering point detection. A hovering point (e.g., point t) has only one connected point, while other points (e.g., points p and q) have more than two connected points.

points on the desired curves, as well as spatial deviations between the ideal line and the noisy edge points obtained from the edge detector. The image may contain multiple edge fragments corresponding to a single whole streak. The classic Hough transform [32] is famous for its capability with the identification of regular curves in an image, such as lines, circles, etc. The main advantage of the Hough transform is that it is tolerant to gaps in feature boundary descriptions and is relatively unaffected by image noise. To increase algorithm efficiency and to reduce computational burden, the straight line search is limited inside the potential rectangular regions of the OCS image. Therefore, identifying twigs is converted to the extraction of straight lines using Hough transform, which is illustrated in Fig. 8. We can see that most of the streaks are correctly identified although a small amount of curves and short lines are ignored. When the straight lines are detected, we can learn their patterns to check if there exists a bird nest in the OCS image.

Fig. 7. Detected hovering points of a bird nest candidate, which are labeled in red dots. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

4.5. Pattern learning with HOS and HLS

Based on the properties of a bird nest, a bird nest is composed of a bunch of short or long twigs with unordered, non-parallel organization. In other words, the groups of streaks in images should be with distributed and diverse orientations, at the same time, with inconsistent lengths. On the contrary, for regions such as insulators, the streaks usually demonstrate uniform and parallel orientations and they generally have equal lengths, as shown in Fig. 8. To represent the orientation and length patterns of streaks in a bird nest, we construct Histogram of Orientation of Streaks (HOS) and Histogram of Length of Streaks (HLS), respectively, and then learn the patterns using SVM for discriminative classification.

The histogram of orientations of streaks (HOS) can be constructed by quantizing the angles of the detected streaks and the horizontal axis, and then accumulating the angles for all streaks in the region. The angle is in the range of 0°–180°. We quantize the angles into 18 bins with a step size of 10° from 0° to 180° to form a histogram. The HOS histogram is constructed by counting the number of streaks at a particular range of angles. The diverse distribution of HOS potentially stands for the orientation of a bird nest while a uniform distribution represents other cases such as insulators.

Similarly, the histogram of lengths of streaks (HLS) can be constructed by quantizing the lengths of the detected streaks, and then accumulating the lengths for all streaks in the region. We quantize the lengths equally into 13 bins with a step size of 5 pixels to form a

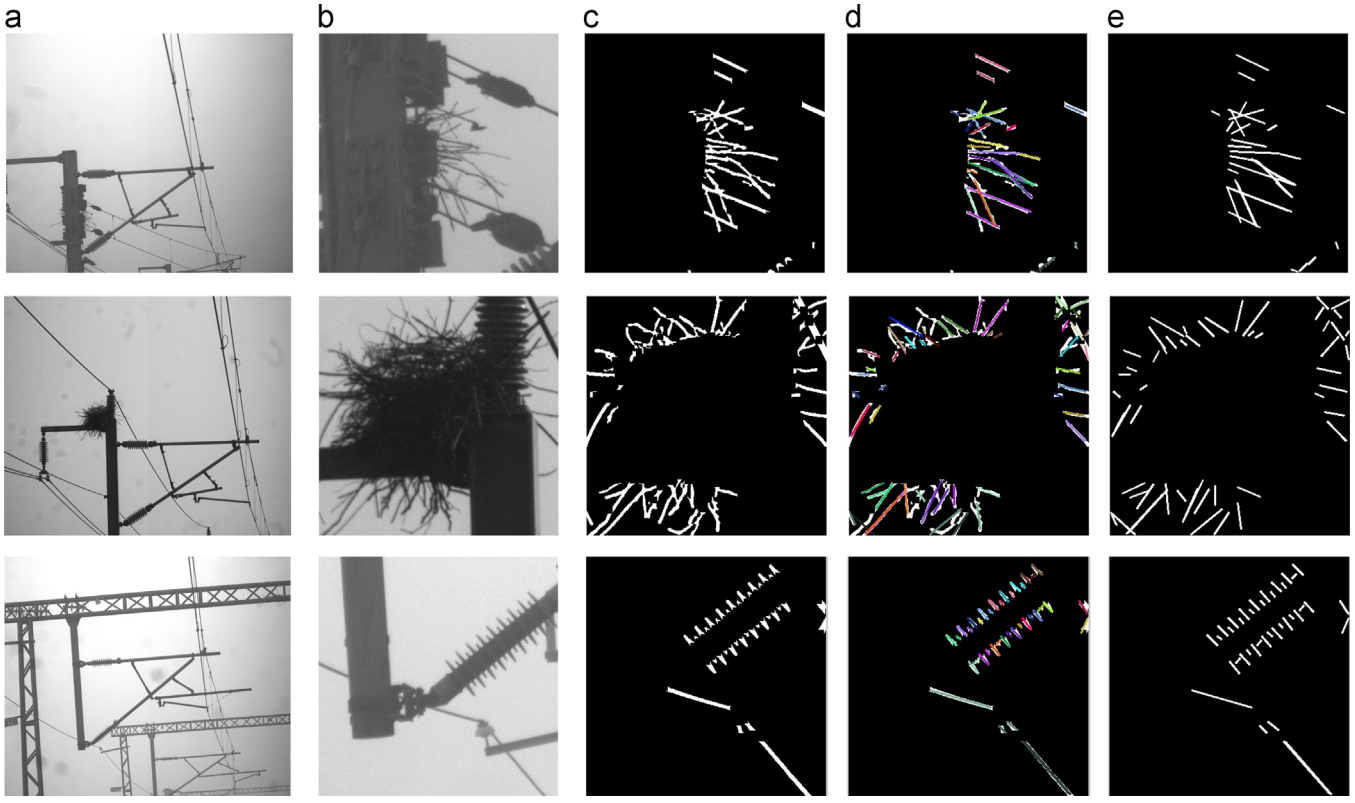

Fig. 8. The detected streaks in each region using Hough transform. (a) the original images, (b) the located regions, (c) the thin branches, (d) the detected lines using Hough transform are labeled with different colors, (e) the final detected streaks. We can see that most of the streaks can be correctly detected.

histogram. The HLS histogram is constructed by counting the number of streaks at a particular range of lengths. The HLS represents the length distribution of the streaks of a bird nest, whereas the diverse distribution potentially indicates a bird nest while a uniform distribution refers to other cases.

Fig. 9 demonstrates the HOS and HLS histograms of three OCS images, in which two of them have bird nests while the other one has the insulator. We can see that the HOS and HLS histograms of images having bird nests (the first two rows in Fig. 9) are different from the one with the insulator (the last row in Fig. 9). The regions of bird nests have diverse HOS and HLS distributions. The orientations and lengths of bird nests are diverse and inconsistent. And the streak orientations are dominated by two major directions, towards left and right, but have less opportunity to be vertical. So we can see an obvious gap at the 9^{th} bin of HOS. On the contrary, the HOS and HLS histograms for images of insulators usually have a couple of prominent peaks, indicating that they have dominant orientations and lengths. That is, the streaks of insulators have consistent orientations and lengths. From this figure, we can see that the distributions of HOS and HLS for images having bird nests are different from other cases.

The HOS and HLS for regions are trained using SVM to learn the patterns of orientation and length, respectively. The regions with and without bird nests are input as positive and negative examples, respectively. We deploy *radial basis function* (*RBF*) kernel to train the models. With SVM, the detection of the presence of a bird nest becomes straightforward in our approach. The HOS and HLS are directly computed and then fed into SVM to perform binary classification. The classifications based on HOS and HLS are finally combined by late fusion to determine the presence of bird nest in OCS images.

5. Experiments

5.1. Dataset

To evaluate the performance of the proposed automatic detection of bird nests for OCS images, we use the non-contact OCS inspection system to capture the OCS images. This inspection device with high resolution camera is placed in high-speed trains to capture still OCS images with 2456×2058 resolution (5 million pixels). In each train line, the same bird nests will appear in multiple images but with different positions, since the inspection system is mounted in the moving high-speed train, with a rate of 17 frames per second. The OCS inspection system has been run on several high-speed train lines across China. Totally, there are 39 videos captured from 9 high-speed railways in China during January 2012 to April 2012. The images having bird nests and their neighbor images before and after them are included in the testing dataset, which consist of 6330 images. The number of bird nests in the dataset is 86. Totally, there are 716 OCS images having bird nests, while the rest of 5614 images do not have. The data information of OCS images is listed in Table 1. These OCS images are challenging since they are captured from different train lines with complex backgrounds such as trees, buildings and towers.

In order to be consistent with the real world condition and to evaluate the performances of different approaches, the experiments

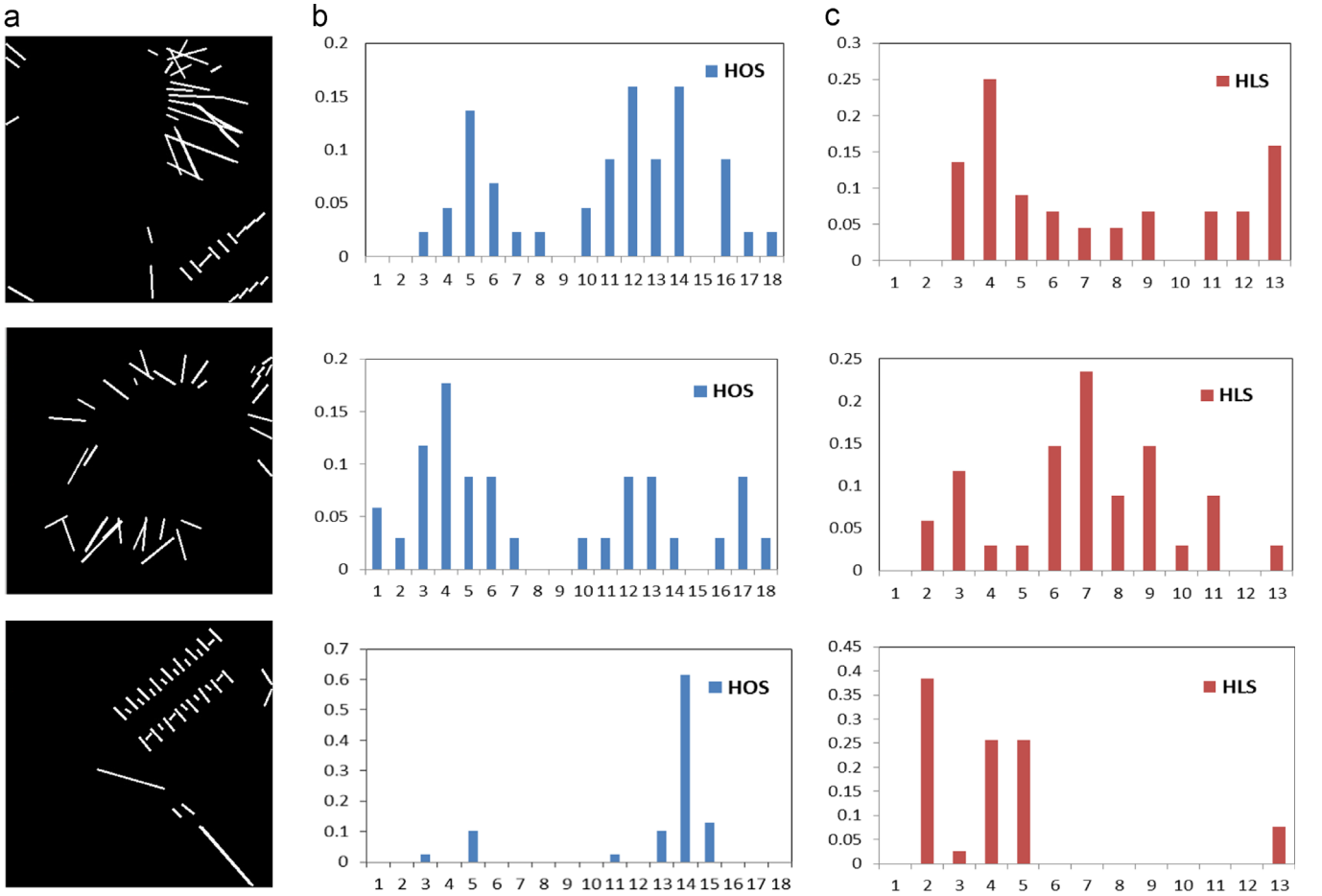

Fig. 9. The histograms of orientation of streaks (HOS) and the histograms of length of streaks (HLS) for images of bird nests and insulators demonstrate different patterns. The images of bird nests have diverse HOS and HLS, while the HOS and HLS for images of insulators usually have few obvious peaks, which mean that they have dominant orientations and lengths. (a) the detected streaks (b) the histograms of orientation of streaks (c) the histograms of length of streaks.

Table 1
Data Information of OCS Images.

Train Lines	# of images	# of bird nests	# of images with bird nests
Beijing–Shanghai	800	10	91
Beijing–Jinan	2900	33	210
Nanjing–Wuhan	350	5	57
Nanjing–Shanghai	100	1	17
Jinan–Qingdao	1156	20	199
Shanghai–Jinan	400	4	23
Shijiazhuang–Taiyuan	183	3	25
Hefei–Nanjing	61	1	16
Hefei–Wuhan	380	9	78
Total	**6330**	**86**	**716**

are carried out in 9-fold leave-one-out cross-validation, since there are 9 train lines in the whole dataset. Each time, the images of an "unseen" line is treated as the testing set and the data of the remaining eight lines are used as the training set. After nine rounds of experiments, the results are averaged as the final performance score.

We use *Precision* and *Recall* as the metric for performance evaluation, which are defined as follows.

$$\text{Precision} = \frac{\#\text{ of correctly detected images having bird nests}}{\text{Total }\#\text{ of detected images having bird nests}} \quad (3)$$

$$\text{Recall} = \frac{\#\text{ of correctly detected images having bird nests}}{\text{Total }\#\text{ of images having bird nests}} \quad (4)$$

5.2. Performance evaluation

This work aims at detecting the presence of bird nests in image sequences for OCS inspection purpose. To the best of our knowledge, this is the first work for automatic detection of bird nests for high-speed rail OCS system. Therefore, there are limited approaches which we could compare in the literature. We resort to other image classification approaches for comparison. Since Histogram of Oriented Gradient (HOG) has been widely used in object recognition and image classification [16,33], we compare our proposed approach with the famous HOG based image classification method [33]. The HOG is applied to the binary enhanced images using the same conditions as the proposed approach. The HOG features are extracted from patches densely located by every 8 pixels on the region, under different scales, 16×16, 24×24 and 32×32 and so on. Each HOG descriptor is 1764 dimension. The extracted HOG histogram is then trained by SVM for detection. To evaluate the performance with the state-of-the-art techniques in object detection, we also compare the proposed method with *Deformable Part Models (DPM)* [17], which has achieved state-of-the-art results on the PASCAL and INRIA person datasets. It is based on mixtures of multiscale deformable part models to represent highly variable object classes. The original implementation of DPM provided by [17] is used for this experiment.

The performance comparison of bird nest detection at the image level is listed in Table 2. From this table, we can see that the performance of HOG based method is poor. It only achieves 26.3% precision and 60% recall. Only 432 bird nest images are correctly detected. Although DPM has achieved state-of-the-art performance in general object detection, the performance for bird nest detection is not satisfactory. The precision is 30.72% and the recall is only 44.41%. In order to guarantee good precision, a large number of negative samples (around 60,000 regions from 300 images) are randomly extracted for the learning of DMP. Among them, there should be certain similarity between the backgrounds of positive and negative samples. Since the backgrounds of OCS images are complex and the regions with bird nests are not so conspicuous, DPM cannot correctly detect the images having bird nests, causing the low recall. In addition, DPM treats the location of a part as a latent variable. However, the shapes of bird nest vary and are not regular, causing the location of parts having bird nests irregular too. Therefore, the latent variable of DPM cannot improve the performance of bird nest detection. Although DPM adopts mixtures of multiscale deformable parts, it does not consider the property of bird nests, which is incompetent for the bird nest detection.

Table 2
Performance of the bird nest detection at image level.

Methods	Correctly detected	Falsely detected	Precision (%)	Recall (%)
HOG [33]	432	1211	26.29	60.36
DPM [17]	318	717	30.72	44.41
HOS+HLS	652	1168	35.82	91.06

Table 3
Performance of HOS and HLS.

Methods	Correctly detected	Falsely detected	Precision (%)	Recall (%)
HOS	628	1251	33.42	87.71
HLS	614	1093	35.96	85.75
HOS+HLS	652	1168	35.82	91.06

For the proposed HOS+HLS method, 652 out of 716 bird nests are correctly detected. The recall is good which reaches 91%, since it is the most critical factor that inspectors care about. They prefer higher recall so that more potential security risks can be notified. With this intelligent bird nest detection system, the inspectors can be released from the time-consuming and arduous task, as they only need to check these abnormal situations visually. Of course, we can see that the precision is not ideal, which is only 35.8%.

In addition, we also compare the performance using only HOS or HLS individually and their combination, which is listed in Table 3. We can see that although HOS can detect more images having bird nests, it also brings more false positives. On the contrary, HLS generates less falsely detected images. It can be inferred that histograms of orientations and lengths have different merits for the classification of bird nests. The combination of HOS+HLS has the best performance.

Furthermore, we apply a statistical *t*-test to verify the significance of the difference between HOG and HOS+HLS. The results of the *t*-test experiment are listed in Table 4. Alpha is the significance level. The *confidence interval (CI)* for the average difference is computed using a *Student-t* model. The pair sample *t*-test experiment is conducted to distinguish the difference between them. For the *t*-test experiment on the performance of precision, the images that are correctly detected as having bird nests or correctly detected as without bird nests are labeled with 1, otherwise they are labeled with 0. The average value of labels is the precision. *t*-Test on precision result demonstrates that 95% confident interval for the average difference. Since the images with bird nests only take up a small proportion of the whole dataset, the average difference confident interval is within the range of 3.57–5.40%. Similar to the precision *t*-test experiment, *t*-test experiment on recall only considers the images having bird nests. The average difference confident interval is within the range of 25.44–32.87%. The difference is considered as significant since the interval does not include zero. From Table 4, we can see that the performance is significantly improved with the HOS+HLS method.

Since the OCS images are pretty complex, automatic bird nest detection is a challenging task. Due to the reasons such as complex background (e.g., mountain, buildings, trees, etc.), mixture of

foreground and background, crossing lines, and so on, it is difficult to discriminate the catenary from the complex background. Examples of failure cases are shown in Fig. 10. For example, in Fig. 10(a), the branches of the bird nest are pretty short and not obvious, so there are not enough hovering points detected in the potential region. This region is excluded by this step. For Fig. 10(d), the background is rather complex, in which many components are mixed together. In the process of trunk and detail detection, the wires are intersected with the steel tower, which increases the number of hovering points. The binary image of this region is very similar to the bird nest region. Therefore, it is falsely detected as a bird nest.

5.3. Speed analysis

In addition to the effectiveness of the proposed method, we also verify speed efficiency. Currently, the proposed approach is performed offline and in a non-real-time way. To monitor the status of the OCS, the high resolution image (2456×2058) is adopted, which is much larger than other images/videos, such as the traditional TRECVID news videos with resolution of 320×240. All methods are implemented with VS 2008 coding, and tested on a PC with 2.66 GHz Core 2 CPU and 4 G memory.

The average speed of each stage for bird nest detection is listed in Table 5. Totally, the time for bird nest detection is around 5.5 seconds. Due to the high resolution of OCS images, the bottleneck of this method is adaptive binarization, which takes much time (around 5 s) to convert OCS images to binary format. For other steps, it is pretty efficient. For the HOG based classification method, the first few steps are the same, and thus their time costs are the same. The difference is the extraction of HOG feature and SVM classification steps. HOG is more efficient compared to HOS and HLS. To extract the streaks, Hough transform is conducted, which takes some time. However, they can be finished within 0.15 second. Overall, the proposed method is about the same speed as HOG based method except feature extraction.

Table 4
Paired sample *t*-test between HOG and HOS+HOL.

t-test type	Alpha	H	CI
***t*-Test on precision**	0.05	1	[3.57%, 5.40%]
***t*-Test on recall**	0.05	1	[25.44%, 32.87%]

6. Conclusion

The development of innovative OCS inspection system based on intelligent image analysis is a meaningful way to improve efficiency, reliability, and safety of the high-speed train. In this paper, we propose a novel framework and solution for bird nest detection of OCS images for high-speed rail. The characteristics of bird nest are first studied. Based on the intuition, the patterns of the streak orientations and lengths are modeled and fed into SVM to classify the OCS images. Experiments on different high-speed train lines have demonstrated promising performance.

This is the first exploration on bird nest detection of OCS images for high-speed rail, and good performance has been

Table 5
Speed efficiency for bird nest detection.

Methods	Binarization (ms)	Hovering point detection (ms)	Feature extraction (ms)	SVM classification (ms)
HOG [33]	5016	328	5	12
HOS+HLS	5016	328	156	10

Fig. 10. Examples of failure cases. Due to the effects of extremely complex background (e.g., mountain, buildings, trees, etc.), mixture of foreground and background, crossing lines, and so on, some bird nest images are missed.

achieved. Unfortunately, due to the complex structure of OCS and the messy background, automatic bird nest detection remains a challenging task. The performance is still not perfect, especially the precision. There are still many images being falsely detected. We will explore other techniques to eliminate the false positives and further improve the precision and recall. Current work does not consider inter-frame relationship, which will be beneficial for elimination of false alarms. The temporal feature will be taken into account in our future work. Except bird nests in OCS, there are other defects including down hanging, ripped off droppers and broken insulators, which potentially affect the safety of high-speed rail network. In our future work, we will explore automatic detection of these defects by intelligent image analysis using the OCS inspection system.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (Nos. 61373121, 61036008 and 61272290), Program for Sichuan Provincial Science Fund for Distinguished Young Scholars (Nos. 2012JQ0029 and 13QNJJ0149) and the Fundamental Research Funds for the Central Universities of China.

References

[1] High-speed rail in China. Available at: ⟨http://en.wikipedia.org/wiki/Chinese_high_speed_rail⟩.
[2] Pantograph. Available at: ⟨http://en.wikipedia.org/wiki/Pantograph_(rail)⟩.
[3] M. Bocciolone, G. Bucca, A. Collina, L. Comolli, Pantograph-catenary monitoring by means of fibre Bragg grating sensors: results from tests in an underground line, Mech. Syst. Signal Process. 41 (1–2) (2013) 226–238.
[4] A. Fumi, A. Forgione, A new complete system for catenary's checking, in: Proceedings of the World Congress on Railway Research (WCRR), 2001.
[5] X. Zhu, H. Wang, L. Fang, et al, Dual arms running control method of inspection robot based on obliquitous sensor, in: Proceedings of IEEE International Conference on Intelligent Robots and Systems, 2006.
[6] H. Hofler, M. Dambacher, N. Dimopoulos, V. Jetter, Monitoring and inspecting overhead wires and supporting structures, in: Proceedings of IEEE Intelligent Vehicles Symposium, 2004, pp. 512–517.
[7] B. Hulin, M. Pfarrdrescher, W. Krötz, B. Sarnes, H. Möller, Video-based onboard surveillance of catenaries of railways, (2007).
[8] R. Kouadio, V. Delcourt, L. Heutte, C. Petitjean, Video based catenary inspection for preventive maintenance on Iris 320, in: Proceedings of the World Congress on Railway Research (WCRR), 2008.
[9] F. Montreuil, R. Kouadio, C. Petitjean, et al, Automatic extraction of information for catenary scene analysis, in: Proceedings of the EUSIPCO, 2008.
[10] J.R. Price, F. Meriaudeau, Precision of computer vision systems for real-time inspection of contact wire wear in railways, Proc. SPIE 5679 (2005) 18–26.
[11] I. Puhlmann, S. Schussler, B. Hulin, Improvements on obstacle detection in the pantograph gauge due to the recognition of steady arms, in: Proceedings of IEEE Intelligent Vehicles Symposium, 2004, pp. 518–521.
[12] C. Caraffi, S. Cattani, P. Grisleri, Off-road path and obstacle detection using decision networks and stereo vision, IEEE Trans. Intell. Transp.Syst. 8 (4) (2007) 607–618.
[13] C. Petitjean, L. Heutte, R. Kouadio, V. Delcourt, A top-down approach for automatic dropper extraction in catenary scenes, Pattern Recognit. Image Anal. 5524 (2009) 225–232.
[14] C.A.L. Vázquez, M.M. Quintas, M.M. Romera, Non-contact sensor for monitoring catenary-pantograph interaction, in: Proceedings of IEEE International Symposium on Industrial Electronics (ISIE'10), July 2010, pp. 482–487.
[15] A. Landi, L. Menconi, L. Sani, Hough transform and thermo-vision for monitoring pantograph-catenary system, Proc. IMechE 220 (2006) 435–447 , part F.
[16] Y. Lin, L. Cao, F. Lv, S. Zhu, et al., Large-scale image classification: fast feature extraction and SVM training, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 2011, pp. 1689–1696.
[17] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, Object detection with discriminatively trained part based models, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1627–1645.
[18] R. Girshick, J. Donahue, T. Darrell, J. Malik, Rich feature hierarchies for accurate object detection and semantic segmentation in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'14), 2014, pp. 580–587.
[19] J. Bossu, N. Hautière, J.-P. Tarel, Rain or snow detection in image sequences through use of a histogram of orientation of streaks, Int Journal of Comput Vision 93 (3) (2011) 348–367.
[20] C.Yao, X. Bai, W. Liu, L.J. Latecki, Human Detection using Learned Part Alphabet and Pose Dictionary in: Proceedings of IEEE Conference on European Conference on Computer Vision (ECCV'14), 2014, pp. 251–266.
[21] S. Zhang, C. Bauckhage, A.B. Cremers, Informed haar-like features improve pedestrian detection, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'14), 2014, pp. 947–954.
[22] Q. Ye, J. Liang, J. Jiao, Pedestrian detection in video images via error correcting output code classification of manifold subclasses, IEEE Trans. Intell. Transp. Syst. 13 (1) (2013) 193–202.
[23] N.C. Mithun, N.U. Rashid, S.M.M Rahman, Detection and classification of vehicles from video using multiple time-spatial images, IEEE Trans. Intell. Transp. Syst. 13 (3) (2013) 1215–1225.
[24] B. Zhang, Reliable classification of vehicle types based on cascade classifier ensembles, IEEE Trans. Intell. Transp. Syst. 14 (1) (2013) 322–332.
[25] W.-L. Zhao, C.-W. Ngo, H.-K. Tan, X. Wu, Near-duplicate keyframe identification with interest point matching and pattern learning, IEEE Trans. Multimed. 9 (5) (2007) 1037–1048.
[26] Y. Cong, J. Yuan, J. Liu, Abnormal event detection in crowded scenes using sparse representation, Pattern Recognit. 46 (2013) 1851–1864.
[27] M.H. Sharif, C. Djeraba, An entropy approach for abnormal activities detection in video streams, Pattern Recognit. 45 (2012) 2543–2561.
[28] J. Cho, M. Lee, H.J. Chang, S. Oh, Robust action recognition using local motion and group sparsity, Pattern Recognition 47 (2014) 1813–1825.
[29] Bird Nest: Available at: ⟨http://en.wikipedia.org/wiki/Bird_nest⟩.
[30] L. He, Y. Chao, K. Suzuki, A run-based two-scan labeling algorithm,", IEEE Trans. Image Process 17 (5) (2008) 749–756.
[31] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal Mach Intell 8 (6) (1986) 679–698.
[32] R.O. Duda, P.E. Hart, Use of the Hough transformation to detect lines and curves in pictures, Commun. ACM, 15, (1972) 11–15.
[33] N. Dalal, B.Triggs, Histograms of oriented gradients for human detection," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 886–893.

Xiao Wu (S'05-M'08) received the B.Eng. and M.S. degrees in computer science from Yunnan University, Yunnan, China, in 1999 and 2002, respectively, and Ph.D. degree in the Department of Computer Science from City University of Hong Kong, Hong Kong, in 2008. Currently, he is an associate professor, the Assistant Dean of School of Information Science and Technology, and the Department Head of the Department of Computer Science and Technology, Southwest Jiaotong University, Chengdu. Now he is a Visiting Associate Professor at University of California, Irvine, CA, USA from Sep. 2015. He was a research assistant and a senior research associate at the City University of Hong Kong, Hong Kong from 2003 to 2004, and 2007 to 2009, respectively. From 2006 to 2007, he was with the School of Computer Science, Carnegie Mellon University, Pittsburgh, USA, as a visiting scholar. He was with the Institute of Software, Chinese Academy of Sciences, Beijing, China from 2001 to 2002. His research interests include multimedia information retrieval, image/video processing, and data mining.

Ping Yuan received her B.Eng and M.Eng. degrees in computer science from School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China, in 2011 and 2014, respectively. Her research interests include multimedia information retrieval, intelligent transportation system, and image processing.

Qiang Peng received the B.E. degree in automation control from Xi'an Jiaotong University, the M.Eng degree in computer application and technology, and the Ph.D. degree in traffic information and control engineering from Southwest Jiaotong University, Chengdu, in 1984, 1987, and 2004, respectively. He is currently a professor at the School of Information Science and Technology, Southwest Jiaotong University. He has been in charge of more than 10 national scientific projects, published over 100 papers and holds 10+ Chinese patents. His research interests include digital video compression and transmission, image/graphics processing, traffic information detection and simulation, virtual reality technology, multimedia system and application.

Chong-Wah Ngo (M'02) received the B.Sc. and M.Sc. degrees in computer engineering from Nanyang Technological University, Singapore, and the Ph.D. degree in computer science from the Hong Kong University of Science and Technology, Hong Kong. He was a Post-Doctoral Scholar with the Beckman Institute, University of Illinois at Urbana-

Champaign, Champaign. He was a Visiting Researcher with Microsoft Research Asia. He is currently a Professor with the Department of Computer Science, City University of Hong Kong, Kowloon, Hong Kong. His current research interests include large-scale multimedia information retrieval, video computing, and multimedia mining. Prof. Ngo is currently an Associate Editor of the IEEE Transactions on Multimedia. He is the General Co-Chair of ACM International Conference on Multimedia Retrieval 2015, the Program Co-Chair of the ACM Multimedia Modeling Conference 2012 and the ACM International Conference on Multimedia Retrieval 2012, and the Area Chair of the ACM Multimedia 2012. He was the Chairman of the ACM (Hong Kong Chapter) from 2008 to 2009.

Jun-Yan He received his B.Eng in Software Engineering in 2013, and is currently pursuing the Ph.D. degree in School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China. His research interests include multimedia information retrieval, image/video processing, and intelligent transportation system.