

RESEARCH

Open Access



Unpaved road detection based on spatial fuzzy clustering algorithm

Jining Bao¹, Yunzhou Zhang^{2*}, Xiaolin Su¹ and Rui Zheng¹

Abstract

Vision-based unpaved road detection is a challenging task due to the complex nature scene. In this paper, a novel algorithm is proposed to improve the accuracy and robustness of unpaved road detection and boundary extraction with low computational costs. The novelties of this paper are as follows: (1) We use a normal distribution with infrared images to detect the vanishing line, and a trapezoid prediction model is proposed according to the road shape features. (2) Road recognition based on connected regions is implemented by an improved support vector machine (SVM) classifier with a normalized class feature vector. According to the recognition results, the road probability confidence map is obtained. (3) With the help of fusing continuous information with the trapezoidal forecasting model and the probability from the confidence map, we present a road probability recognition method based on the trapezoidal forecasting model and spatial fuzzy clustering. Furthermore, the histogram backprojection model is used to solve interference problems caused by shadows on the road. It takes approximately 0.012~0.014 s to process one frame of an image for the road recognition, and the accuracy rate can reach 93.2%. The experimental results show that the algorithm can achieve better performance than some state-of-the-art methods in terms of detection accuracy and speed.

Keywords: Unpaved road detection, Confidence map, Trapezoid prediction, Vanishing line detection, Clustering analysis

1 Introduction

The intelligent vehicle remains a core problem in computer vision technology and has numerous potential applications, such as driver assistance, transportation system scheduling, and searching for the optimum route. There is no doubt that road detection has become one of the most popular topics in computer vision [1–3]. Computer vision technology is very suitable for road detection since it includes a large amount of detection information and accurate sensing [4, 5]. However, road detection is still challenging due to different road types and various background, weather, and illumination conditions.

Over the past few decades, numerous approaches have been developed for road detection. According to the structuralization degree, existing road detections can be classified into two categories: paved road detection and

unpaved road detection [6]. For the well-paved roads with remarkable road borders and lane markings, desirable road detection accuracy can be achieved by many existing methods. However, unpaved roads are most commonly seen in suburban, rural, and battlefield environments. Since there are hardly any variant features to characterize unpaved roads [7, 8], it is challenging to develop a valid algorithm with computer vision. Obtaining accurate road information under unfavorable interferences plays a critical role in unpaved road detection [9]. The main difficulty in unpaved road detection based on computer vision is that the detection algorithm must be able to handle complex and unknown scenes, such as road types with variable illumination, different weather conditions, and varieties of textural characteristics. Furthermore, intelligent transportation systems generally require fast processing since the vehicle speed is bounded by the processing rate [10]. All of these factors impact the unpaved road detection results.

Some algorithms have been proposed for road detection by researchers. The algorithms can be mainly

* Correspondence: zhangyunzhou@mail.neu.edu.cn

²Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110819, China

Full list of author information is available at the end of the article

divided into three categories: feature-based, model-based, and ML (machine learning)-based [11–13]. Algorithms based on features usually extract significant and stable features. These algorithms are very simple and rapid. For example, color features [14] are often applied to extract the road region from an image. However, they do not work well for general road images, especially when the roads have little color differences between their surface and the environment. The traditional feature-based algorithm is not suitable for complicated scenes since it is insensitive to road shapes and is easily impacted by watermarks, shadows, vehicles, and pedestrians. A series of features are analyzed to show their ability to detect road surface from background. Therefore, it is often combined with other algorithms for segmentation processing to obtain proper features. The top-down road recognition algorithm [15] combines both the region and boundary cues of the images. With an off-line classifier, it can detect road regions. Shang et al. [16] proposed an approach to find a way to choose the feature descriptors. Meanwhile, support vector machine (SVM) technology has been added to analyze the importance of the common feature descriptors during the road detection process. Compared with the feature-based algorithm, the model-based algorithm matches the mathematical model according to the prior information, such as the road position, distribution, and shape. Kluge et al. [17] proposed a deformable template model of lane structures to locate lane boundaries without thresholding the intensity gradient information. The Metropolis algorithm is used to maximize a function that evaluates how well the image gradient data support a given set of template deformation parameters. The multi-modal road detection and segmentation algorithm [18] is proposed based on monocular images and HD (high definition) multi-layer lidar data (3D point cloud). The detection algorithm for road boundaries is proposed by using the hyperbolic road model for moving images captured by in-vehicle cameras [19]. This algorithm uses the detected center lines based on edges for parameter estimation, and only the region of validity of the detection result of the center line is applied in the hyperbolic road model. These algorithms have strict demands on the shape of the road region and require precise mathematical models. At present, this type of algorithm is mainly used for simple road detection. The ML-based algorithm needs to collect a large number of samples and training parameters. Typically, SVM-based algorithm is applied to road detection through semi-supervised or supervised online learning [20, 21]. For example, Zhou et al. [22] proposed an effective approach to use SVM for road detection with self-supervised online learning. Yun et al. [23] adopted the boosting, SVM, and random forest classifiers to evaluate the correlation feature set and raw

feature set. To fully utilize potential region feature correlations and improve the classification accuracy, this algorithm also introduces the feature combination method into road detection. *K*-means algorithm and density-based spatial clustering of applications with noise (DBSCAN) algorithms are also significant means of clustering road regions [24, 25]. Road detection is also implemented using a neural network [26]. Although it can achieve high accuracy, its performance often depends on the large amount of training data and complex computations.

The vanishing line is an important component in road detection. With the aid of a set of prior information, road detection is improved. The vanishing line is located horizontally between the road and sky. The upper part is the sky and the lower part is the road in general. The distribution characteristics of the road region assist in segmenting the road region to reduce the error rate of vanishing line detection. Most common methods, such as Prewitt, Sobel, and Gabor filters, detect the vanishing line [27]. To date, many algorithms have attempted to handle off-road conditions. These algorithms have their own advantages and disadvantages. Compared with the significant advances in paved road detection, little progress has been made for unpaved roads [10, 15].

When a vehicle is driven on an unpaved road, illumination variation easily leads to a poor visual condition, which makes it difficult for the driver to distinguish the unpaved road. Compared with the paved road, there is still a large space for improvement in unpaved road detection. First, with respect to road textures, unpaved roads mostly have poor conditions, such as shadows and ruts. Second, the road boundary is fuzzy and difficult to identify. Sometimes the ruts are more obvious than the road boundary. Last, there are some variable factors, such as varying illumination and changing weather conditions. It is significant to propose a good performing algorithm using computer vision for practical applications of unpaved road [10]. Therefore, we propose a new, more efficient method for unpaved road detection based on infrared images in this paper.

In this paper, we propose a novel algorithm for the unpaved road detection based on infrared images that can achieve good performance in terms of detection accuracy and speed. It has a positive theoretical significance in detection of the vanishing line, segmentation, and fuzzy clustering. Recognition and boundary extraction of unpaved roads based on infrared images can overcome illumination and weather changes in the unpaved road. Specifically, we use the detection of the vanishing line as assisting information to segment the road to reduce the error rate. As a result, our paper addresses the above-mentioned problems of unpaved road detection by decomposing the road detection process into several steps: (1) The region of

interest (ROI) is obtained, and the trapezoid mode is established. (2) An improved SVM classifier is constructed based on the unidentified connected region, and a road probability confidence map is also generated according to the recognition results. (3) The road region is determined by the spatial fuzzy clustering algorithm. The fuzzy C-means (FCM) algorithm that combined with the trapezoidal forecasting model and the probability confidence map is adopted to improve the accuracy of road boundary detection. Meanwhile, gray-based histogram backprojection compensates for the missing parts after road clustering due to segmentation and other reasons. Therefore, we can obtain more accurate boundary information.

The remainder of this paper is organized as follows. The proposed algorithm is detailed in Section 2. Section 3 presents the experimental results and a discussion of the proposed algorithm. Finally, Section 4 draws some conclusions for the paper.

2 Methods

In this section, we detail the proposed FCM-based unpaved road detection framework. It is designed with the following steps: (1) vanishing line detection based on the normal distribution with infrared images; (2) an image segmentation method that uses the double-Otsu algorithm and the trapezoid prediction is obtained; (3) image classification with the improved SVM and construction of the probability confidence map according to the classification results; (4) spatial fuzzy clustering algorithm based on FCM combined with the trapezoidal forecasting model and the probability confidence map to complete road recognition; and (5) grayscale-based histogram backprojection to weaken the interference caused by road shadows.

2.1 Vanishing line detection

Normally, an unpaved road boundary is blurry and difficult to recognize. Real-time image information should be utilized as much as possible to improve the detection accuracy. Real-time video collected by an infrared imager contains substantial prior information: (1) the road presents fixed geometrical features, such as triangles or trapezoid; (2) the upper region of the infrared image is the sky, the middle area is the road, and both sides are mostly non-road areas; and (3) consecutive frames of infrared images present spatio-temporal continuity. The road changes are relatively continuous and stable without abrupt changes. The position of the road region in the next frame can be predicted based on the image information in the previous frame.

2.1.1 Image pre-processing

Images captured by a vehicle-mounted infrared imager have 702×576 pixels, as shown in Fig. 1. By removing



Fig. 1 Infrared image

the worthless border, the image is resized to 696×450 pixels to obtain the ROI. Based on image pre-processing, the original infrared image can remove noise to enhance the contrast and weaken the interference in the image. After histogram equalization, the infrared image is shown in Fig. 2. The detailed boundaries of the road are enhanced without affecting the overall contrast of the entire image.

2.1.2 Vanishing line detection

Vanishing line detection can remove irrelevant areas over the horizon to reduce computations and improve accuracy. As the juncture area of the sky and the road area, the vanishing line always appears as a longer line. Furthermore, it has an obvious vertical gradient.

In this paper, the vanishing line detection and tracking method based on the normal distribution is proposed according to the estimate of the vanishing line. Bayesian posterior is used to detect the vanishing line with a priori probability that obeys the normal distribution. When the current frame detects the vertical gradient feature, the probability at position h_m is:

$$p(h_m^{(t)} | l_i^{(t)}) = \frac{p(l_i^{(t)} | h_m^{(t)}) \cdot p(h_m | h_m^{(t-1)})}{p(l_i^{(t)})} \quad (1)$$

where $p(h_m^{(t)} | l_i^{(t)})$ is the probability of the vanishing line at the position h_m when the current frame detects vertical gradient feature. The vertical gradient feature is the probability of vanishing line denoted by $p(l_i^{(t)} | h_m^{(t)})$. $p(h_m^{(t)} | h_m^{(t-1)})$ is a priori probability, that is, the current frame probability is predicted from the last frame normal distribution. $p(l_i^{(t)})$ denotes the i th linear l in vanishing line probability.



Fig. 2 Image after histogram equalization

$$\begin{aligned}
 p(l_i^{(t)}|h) &= p(l_i^{(t)}|h_1^{(t-1)}) \times p(h_1^{(t)}|h_m^{(t-1)}) \\
 &+ p(l_i^{(t)}|h_2^{(t-1)}) \times p(h_2^{(t)}|h_m^{(t-1)}) \\
 &+ \dots + p(l_i^{(t)}|h_n^{(t-1)}) \times p(h_1^{(t)}|h_n^{(t-1)})
 \end{aligned} \tag{2}$$

where $p(l_i^{(t)}|h)$ is a constant to all of the image frames. Bayesian posterior probability can be calculated with the aim of accurately localizing the vanishing line.

2.2 Segmentation method

2.2.1 Double-threshold segmentation based on the Otsu method

As one of the most effective and widely used methods [28, 29], Otsu threshold method is adopted for road segmentation. It selects a threshold using the histogram of a grayscale image to find the image’s optimal threshold, which maximizes the inter-class variance to obtain a larger separation between the foreground and background. Moreover, it has low computation complexity, which is suitable for a real-time system. Given an image I that contains N pixels and gray levels ranging from 0 to $m - 1$, there are n_i pixels for gray level i and its probability is p_i :

$$p_i = n_i/N \tag{3}$$

$$N = \sum_{i=0}^{m-1} n_i \tag{4}$$

We set the threshold as t , the ratio of the number of pixels in the ROI to the entire image area is denoted as $\omega_0(t)$, and the mean grayscale value of the target region is $\mu_0(t)$. The ratio of the pixels in the non-target region to the entire image is $\omega_1(t)$, and the mean of non-target region is $\mu_1(t)$. We obtain the following equations:

$$\begin{cases} \omega_0(t) = \sum_{0 \leq i \leq t} p(i) \\ \mu_0(t) = \sum_{0 \leq i \leq t} ip(i)/\omega_0(t) \end{cases} \tag{5}$$

$$\begin{cases} \omega_1(t) = \sum_{0 \leq i \leq m-1} p(i) \\ \mu_1(t) = \sum_{0 \leq i \leq m-1} ip(i)/\omega_1(t) \end{cases} \tag{6}$$

Average grayscale value of the entire image can be expressed as:

$$\mu = \omega_0(t)\mu_0(t) + \omega_1(t)\mu_1(t) \tag{7}$$

The inter-class variance value ϑ :

$$\vartheta = \omega_0(t)(\mu_0(t)-\mu)^2 + \omega_1(t)(\mu_1(t)-\mu)^2 \tag{8}$$

We choose T as the optimal segmentation threshold when ϑ achieves the maximum value using the traversal method.

Otsu threshold segmentation can effectively extract the road region. However, the intensity between the road and surrounding scenery is always small, which causes poor segmentation results. Furthermore, the shadows are closer to the non-road region. Therefore, threshold segmentation can only be used for the initial classification. To solve this problem, we proposed a double-Otsu threshold segmentation method. For the initial segmentation, Otsu method was used to obtain the threshold T_1 . Then, it is used again to segment the image over the $T_1 \sim 255$ range to achieve the second threshold T_2 . After that, two segmentation images can be obtained by thresholds T_1 and T_2 .

Accordingly, the thresholds ranging from 0 to T_1 are chosen for the same operation. We obtain the third segmentation threshold T_3 , which ranges from 0 to T_1 . Then, T_3 is utilized to continue with the threshold of the segmented image. Meanwhile, the image is divided into blocks using the boundary auxiliary information. Figure 3

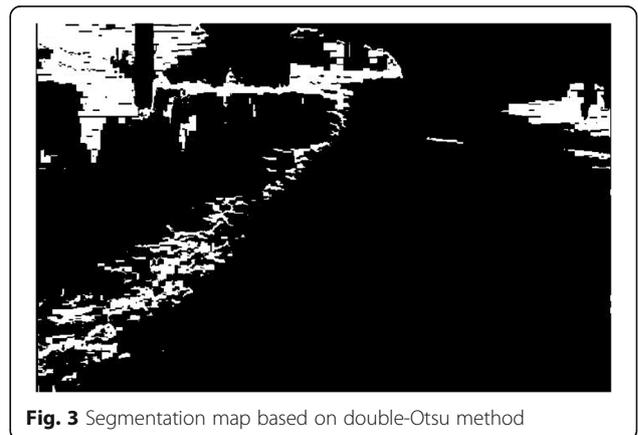


Fig. 3 Segmentation map based on double-Otsu method

shows the results of the double-Otsu method. In Fig. 3, the white region represents the unknown region, the middle black region denotes the road, and the left black region represents the non-road. The obtained result is better than that of one-time Otsu, which can be used to recognize the road region from the non-road region. The obtained segmentation results provide a foundation for further research.

2.2.2 Trapezoidal forecasting model

Since the road region appears to have a trapezoidal geometric shape, there are some correlations between two adjacent frames that will not change in the video [30]. Therefore, the current frame can be used to predict the next frame. The trapezoidal forecasting model is proposed based on the characterization of the positional distribution.

Trapezoidal forecasting model can be established with the following steps: (1) To form a trapezoidal image, as shown in Fig. 4a, b, for road prediction in the next frame, four vertices are extracted from the previous frame of the road region and give the value range. (2) Based on the distance parameter of the trapezoidal forecasting model, the obtained trapezoidal region can be used to determine whether the initial segmentation result is a road region. If the computed distance is far from a certain value, the connected region is assumed to be a non-road region. Otherwise, it is a road region. (3) To extract the road area map as shown in Fig. 4c, we first use the morphological erosion operation with a 3 × 3 disk-shaped structural element to estimate the background of the trapezoid vertex image (see Fig. 4b). Then,

the road boundary result (see Fig. 4d) can be obtained by subtracting the background image from the original trapezoidal vertex image. As shown in Fig. 4c, d, the results of boundary recognition are relatively accurate. The advantage of this method is that it can cut off the non-road region that is similar to the road region, such as a trunk. However, trapezoidal forecasting model cannot independently achieve excellent results.

2.3 Classification method

ML is often used for road recognition. As a supervised learning technique, SVM selects an optimal hyper-plane as a decision function. Taking into account the empirical error and the complexity of the classifier, SVM is widely used because it can optimize the road boundary [31–33]. Considering the complex recognition tasks, SVM is used to classify road regions in this paper.

The training results are shown in Fig. 5. The white regions are considered to be road regions, the black regions are non-road regions, and the other regions are undefined parts of undetected connected areas. The error rates of the trained classifier and positive sample are 7.52 and 1.08%, respectively. In Fig. 5, the road region has large shadow areas and some of it is cut off. Some parts of the road region in the image are unidentified.

2.3.1 Improved SVM classifier

For the SVM classifier, this paper presents two improved methods: feature classification normalization and SVM combination classifier.

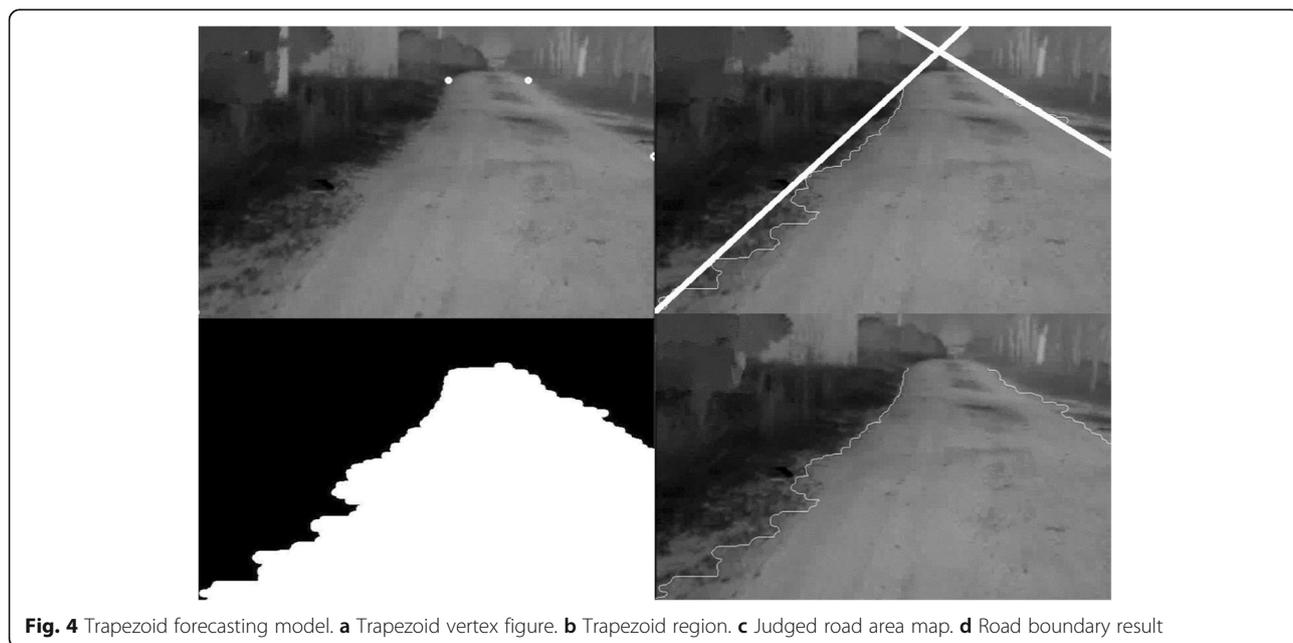


Fig. 4 Trapezoid forecasting model. **a** Trapezoid vertex figure. **b** Trapezoid region. **c** Judged road area map. **d** Road boundary result

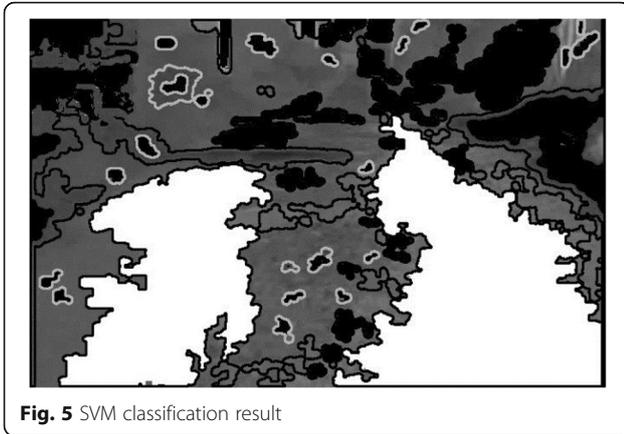


Fig. 5 SVM classification result

Normalization method is adopted to normalize three types of features (grayscale, position, and shape) in the fusion process. This method can retain the dynamic range of the features and affect the training results, but it does not take into account the relationship among several features and each feature within one type of feature. Therefore, we propose three features for classification and normalization. There are a total of 1314 samples, including 467 positive samples (with 413 samples of road regions and 54 road shadows) and 467 negative samples (with 0, 5, and 462 samples of the sky regions, the surrounding trees, and the non-road region, respectively).

Figure 6 shows the classification result of classification-normalized SVM. The obtained result is more accurate and the road region is completely recognized compared with Fig. 5. The error rate of the trained classifier is 6.79% and the positive sample is 6.79%. The testing time is between 0.003 and 0.006 s. Although the classification normalization loses some of the details of the image, it can retain the independence of each feature and weaken the interference among the three features caused by the fusion.

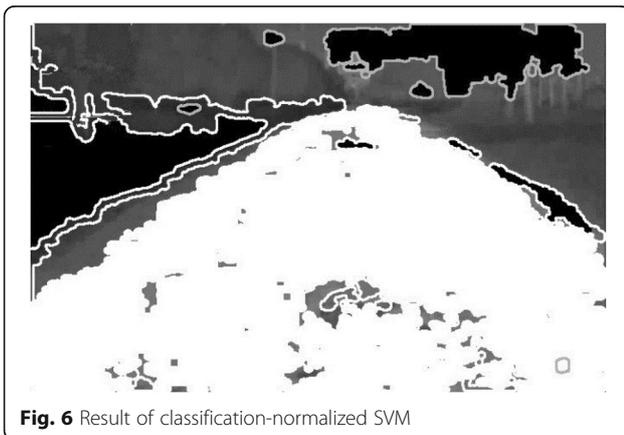


Fig. 6 Result of classification-normalized SVM

The SVM combination classifier separately trains three types of feature information, and each piece of information forms the non-classification normalization and classification normalization features. Therefore, six SVM classifiers are formed depending on the two classifiers and three features' information. Then, we obtain a better SVM classifier through the given weight parameters after the training.

First, the feature information of grayscale, position, and shape are collected. Then, the three features' information and one classifier regarded as the base classifier are constructed. We choose one group with the lowest training error rate in each base classifier as the training set, which can be used to obtain SVM classifier. The initial weight of the base classifier is given with the corresponding parameter calculated as follows:

$$\begin{cases} c_c w_{new} = \alpha \cdot c_c w_{old} + (1-\alpha) \cdot c_c w \\ p_c w_{new} = \alpha \cdot p_c w_{old} + (1-\alpha) \cdot p_c w \\ a_c w_{new} = \alpha \cdot a_c w_{old} + (1-\alpha) \cdot a_c w \\ c_n w_{new} = \alpha \cdot c_n w_{old} + (1-\alpha) \cdot c_n w \\ p_n w_{new} = \alpha \cdot p_n w_{old} + (1-\alpha) \cdot p_n w \\ a_n w_{new} = \alpha \cdot a_n w_{old} + (1-\alpha) \cdot a_n w \end{cases} \quad (9)$$

where $c_c w_{new}$, $p_c w_{new}$, $a_c w_{new}$, $c_n w_{new}$, $p_n w_{new}$, and $a_n w_{new}$, respectively, represent as the training weights of the base classifier in grayscale information non-classification, location information non-classification, shape non-classification, grayscale information classification, location information classification, and shape information classification. $c_c w_{old}$, $p_c w_{old}$, $a_c w_{old}$, $c_n w_{old}$, $p_n w_{old}$, and $a_n w_{old}$, respectively, represent the last corresponding training weight of the base classifier. The last confidence measure is denoted as $\alpha = 0.5$. $c_c w$, $p_c w$, $a_c w$, $c_n w$, $p_n w$, and $a_n w$ denote the calculated weights in this group. The formula is given as:

$$\begin{cases} c_c w = \frac{nc_{cr} - nc_{ce}}{\text{total}} \\ p_c w = \frac{np_{cr} - np_{ce}}{\text{total}} \\ a_c w = \frac{na_{cr} - na_{ce}}{\text{total}} \\ c_n w = \frac{nc_{nr} - nc_{ne}}{\text{total}} \\ p_n w = \frac{np_{nr} - np_{ne}}{\text{total}} \\ a_n w = \frac{na_{nr} - na_{ne}}{\text{total}} \end{cases} \quad (10)$$

where nc_{cr} , nc_{ce} , nc_{nr} , nc_{ne} , np_{cr} , np_{ce} , np_{nr} , np_{ne} , na_{cr} , na_{ce} , na_{nr} , and na_{ne} , respectively, represent the base classifier that predicts the correct or wrong numbers of the normalization of non-classification and classification grayscale information, location information, and shape information in this group. Their summation is calculated as follows:

Table 1 Training results of six basic classifiers on penalty factor c and kernel function parameter σ^2

Adopted normalized feature	c	σ^2	Training time (s)	Number of positive samples	Accuracy (%)
Gray feature (non-classification)	4	256	587	1027	82.9
Location feature (non-classification)	8	8	309	518	41.4
Shape feature (non-classification)	8	512	291	538	43.0
Gray feature (classification)	8	1	291	1314	100.0
Location feature (classification)	8	1	271	1314	100.0
Shape feature (classification)	4	1	339	1144	87.1

$$\begin{aligned} \text{total} = & nc_{cr} - nc_{ce} + np_{cr} - np_{ce} + na_{cr} - na_{ce} \\ & + nc_{nr} - nc_{ne} + np_{nr} - np_{ne} + na_{nr} - na_{ne} \end{aligned} \quad (11)$$

The parameters and the training results of six base classifiers are shown in Table 1. The added contents are the parameter c denotes penalty factor and σ^2 represents the kernel function parameter. Each base classifier uses its own penalty factor and kernel function parameters. The training time is 587 s, and the highest support vector tree is 100%. The final discriminant is as follows:

$$\begin{aligned} SVM = & c_c w \cdot SVM_{cc} + p_c w \cdot SVM_{cp} + a_c w \cdot SVM_{ca} \\ & + c_n w \cdot SVM_{nc} + p_n w \cdot SVM_{np} + a_n w \cdot SVM_{na} \end{aligned} \quad (12)$$

After training, we get the weight parameters of each classifier: $c_c w = 0.1347$, $p_c w = 0.2486$, $a_c w = 0.347$, $c_n w = 0.1100$, $p_n w = 0.2544$, and $a_n w = 0.1176$. The error rate of the combined SVM classifier is 5.04%; the positive sample error rate is 2.05%, and the testing time is between 0.011 and 0.025 s. The result is shown in Fig. 7.

2.3.2 Construction of a probability confidence map

A confidence map is established according to the results of SVM discriminant connected area. We set the confidence values as 0 and 1 to represent the non-road region and road region, respectively. The formula of the probability confidence is as follows:

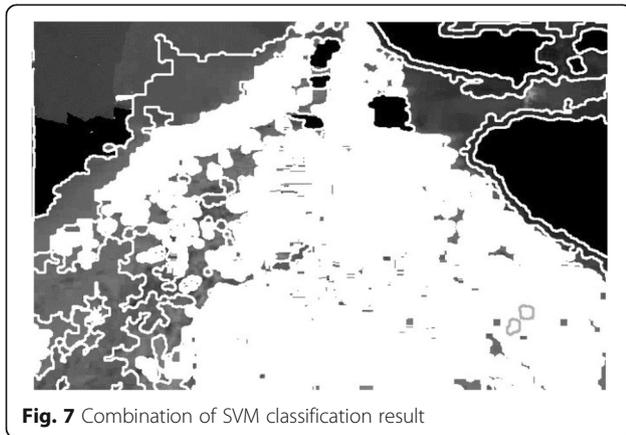


Fig. 7 Combination of SVM classification result

$$p = \begin{cases} 1 & f(x) \geq 1 \\ f(x)/2 + 0.5 & -1 < f(x) < 1 \\ 0 & f(x) \leq -1 \end{cases} \quad (13)$$

The probability confidence map is converted into a visual image. The conversion formula is calculated as:

$$I_{(i,j)} = 256 * p \quad (14)$$

There is an unrecognized background region in the image that has not been subjected to SVM discriminant, except for the connected region. Given that the result of SVM discriminant is 0, the probability confidence is 0.5 which cannot identify road region or not. The grayscale value is 127 in the probability confidence map.

The converted probability confidence map is shown in Fig. 8. (1) In the first group (Fig. 8a, b), the road region and road shadow region are marked as white and near white, respectively. This represents the shadow region belonging to the road region based on the SVM method. (2) In the second group (Fig. 8c, d), the road region and road boundary region are marked as white and near white, respectively. The non-road region is on the right side of the image and is marked as black. The grayscale value of the gray region on the top and left side of the image is smaller than 127. With the substantial color of probability confidence map, SVM method can correctly discriminate the road region from the non-road region. (3) In the last group (Fig. 8e, f), the white region represents the road region, while the other connected regions are marked as black except for the non-discriminated regions. The probability confidence map shows that the connected region is related to the road region, with a more robust stability and fault tolerance.

2.4 Clustering method

We propose to use Otsu method to obtain connected regions and regard each of them as a data point, which can greatly reduce the size of the data set and computational complexity. However, the data points will be sparser and the number of outliers will be increased due to the reduction of data sets. It is crucial to select proper initial parameters and features. The road region is generally in the lower middle image with a relatively stable grayscale value.

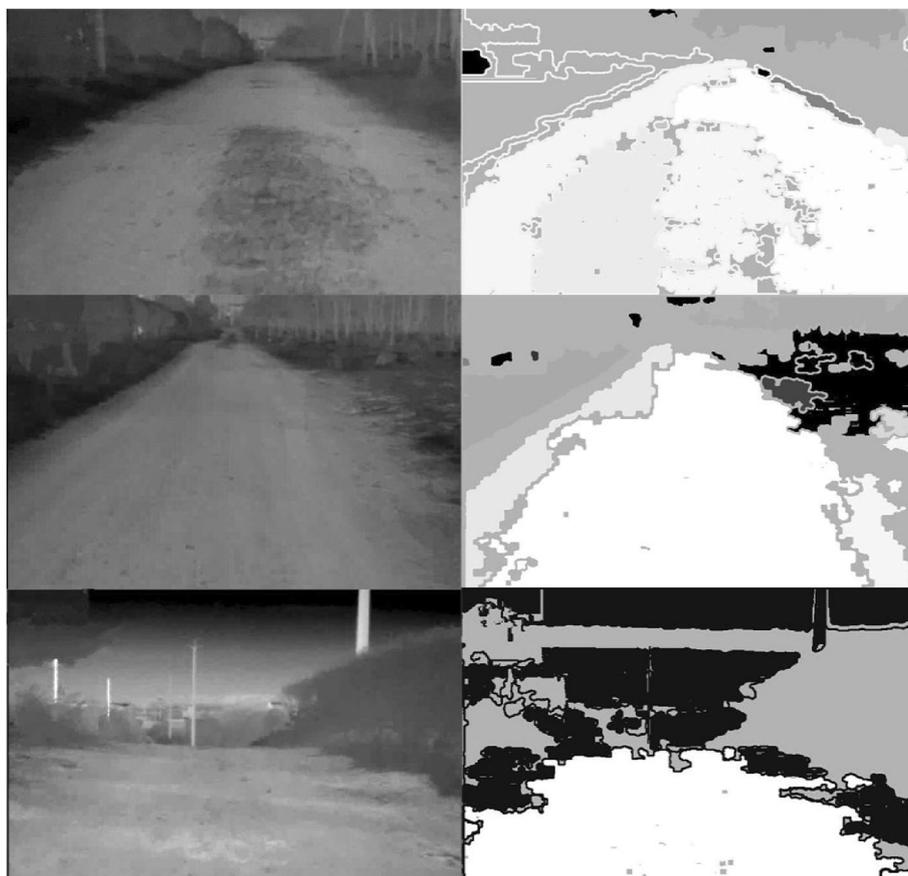


Fig. 8 Probability confidence maps. **a** Original image 1#. **b** Probability confidence map1#. **c** Original image 2#. **d** Probability confidence map 2#. **e** Original image 3#. **f** Probability confidence map3#

It is reasonable to divide the road into three categories, as shown in Fig. 9, which are the ideally segmented road regions. The selection of the initialization has a very significant influence on the performance and results of the clustering analysis. The initialed cluster centers are usually several points that are randomly selected from the given data set or the position of a peak of the histogram. In this paper, the size of the processed data set is relatively small, and therefore, it easily generates empty clusters. The road region clustering problem can be solved by the prototype clustering model. FCM clustering is one of the prototype clustering methods.

2.4.1 Road identification based on FCM algorithm and trapezoidal forecasting model

There are many methods for fuzzy clustering [34, 35], and FCM clustering is adopted in this paper. $X = \{x_1, x_2, \dots, x_m\}$ is defined as a dataset that contains k clusters C_1, C_2, \dots, C_k . w_{ij} is a membership value that indicates that the ability of x_i belongs to C_j , and the sum of the membership values w_{ij} for a given point x_i is 1. $X = \{x_1, x_2, \dots, x_m\}$ represents the input data point set in FCM

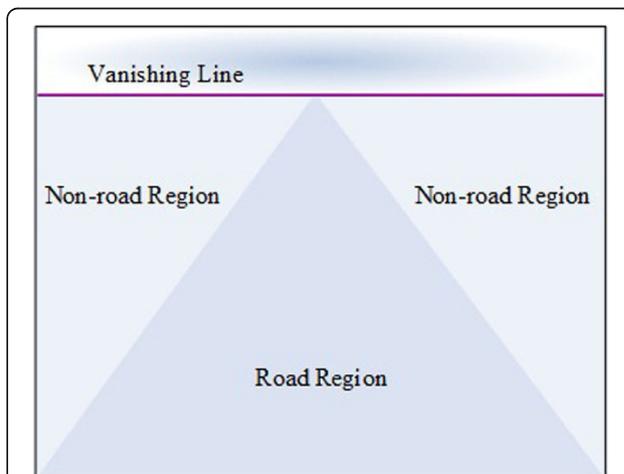


Fig. 9 Prior schematic diagram

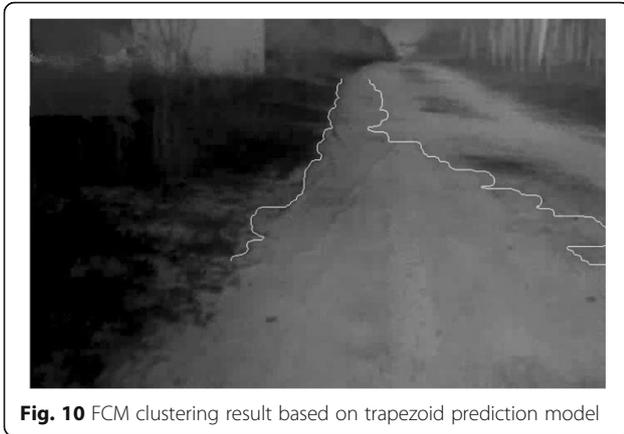


Fig. 10 FCM clustering result based on trapezoid prediction model



Fig. 12 Backprojection probability map of infrared image

algorithm, the number of clusters is C , and the output membership value is w .

In this paper, the points are $[x, y, k * m, d * t]$, where x and y are the central row and column coordinates of the connected regions, respectively. k denotes the proportional coefficient of the grayscale and the positional information is $k = \frac{\sqrt{\text{height}^2 + \text{width}^2}}{256}$. m is the gray scale mean value of the connected regions. d denotes the proportional coefficient of the predictive value and the spatial location information. t is the prediction features. Second, the clusters are separated into three categories and the membership matrix is initialized. When the membership matrix is initialized, we select the virtual center in advance. The center point is defined as: $[110, 100, k * \text{thres}, d * t]$, $[110, 596, k * \text{thres}, d * t]$, $[255, 348, k * \text{thres}, d * t]$, where the predictive parameter value d is set as 100 and $k = 3.24$ by height = 450 and width = 696. Then, the initialized membership matrix is calculated. Finally, the parameter value of p is set to 2. The above clustering procedure can be found in [36]. After the calculation, the connected region is sorted into the highest membership degree.

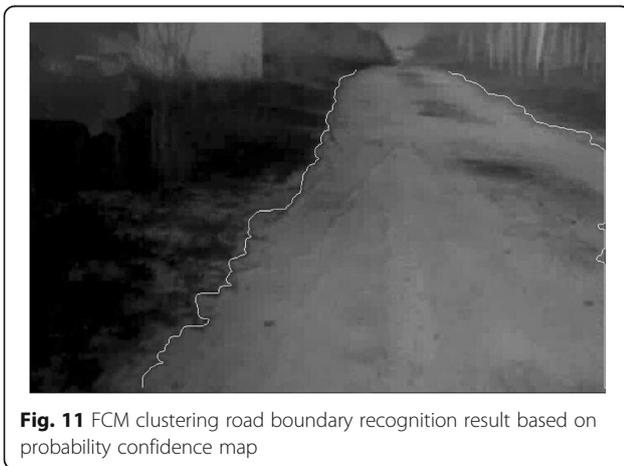


Fig. 11 FCM clustering road boundary recognition result based on probability confidence map

2.4.2 Road probability determination based on FCM algorithm and trapezoidal forecasting model

A single continuity prediction or a single discriminant is prone to generate errors. When the probability confidence map is used as the prediction model, the accuracy of SVM classifier cannot achieve 100%. Although FCM clustering algorithm can reduce the error rate of SVM algorithm, the efficiency improvement is limited. The probability confidence map only addresses each frame image and usually neglects the important characteristic of video continuity in the video.

There are still some problems that occur by only using the trapezoidal forecasting model. For example, if the vehicles' speed increases too fast, the road continuity will be difficult to maintain in some situations, especially in excessively fast turns. In this case, the continuity of the road is greatly influenced by the concentration of traffic. As shown in Fig. 10, some parts of the road region are not recognized by the trapezoidal forecasting model when the road changes rapidly. Moreover, there are some errors in the identification of the road boundary. The combination of the two types of forecasting information is proposed in this paper.

The results from FCM clustering algorithm belong to class membership. Based on the characteristics of FCM results, we propose a discriminant correction model, which is based on FCM algorithm and trapezoidal forecasting model of road probability decision model. After the FCM iteration is finished, the connected region of the largest membership class can be regarded as a preliminary result. Furthermore, the membership degree is recalculated by using the membership probability-based confidence map. Finally, a class of connected regions can be selected by means of the maximum degree of the membership class. The probability confidence map represents the probability of the road region. In this paper, FCM is divided into three categories. The first category is the road region, and the other two categories belong to the non-road region. The recalculated membership formula is as follows:



Fig. 13 Backprojection probability map based on road region model. **a** Original image. **b** Backprojection probability map

$$w_{new} = \begin{cases} w^*p, & \max(w) \in c_1 \\ w^*(1-p), & \max(w) \notin c_1 \end{cases} \quad (15)$$

where c_1 denotes the first class, w is the membership after the stopping iteration, and w_{new} is the final membership of the connectivity region. p is the probability confidence for this connected region, and $\max(w)$ is the category of the maximum membership. With the improved method, road region identification and road boundary recognition obtain higher accuracy.

Compared with Fig. 10, for the same video frame, Fig. 11 shows a more accurate road recognition result that agrees with the actual environment.

2.4.3 Road recognition based on histogram backprojection method

We consider that the above-mentioned algorithms are all based on connected regions. However, they may lead to incomplete recognition of the road boundary in the following three cases:

1. There are undetected connected regions in the video.
2. Morphological erosion is used to segment the connected region, which reduces the undetected regions as well as reduces the road regions.

3. The non-road connected region is too large or too small, which results in inaccurate road regions.

The histogram backprojection method of the road region model can solve the above-mentioned problems. This method, which is based on pixels, effectively compensates for the drawbacks and disadvantages of the connected region as a basic unit.

The histogram backprojection method converts the color histogram of the original image into a color probability distribution. The pixel grayscale value of the original infrared image represents the image intensity. The pixels in the probability map are used to recognize the connected regions to determine the real road region. Figure 12 shows the effect of histogram backprojection.

Although the result of FCM clustering is good, it is not precise enough to use the connected region as the basic unit to recognize the road region. In this paper, the pixel grayscale value of the histogram is used to compensate for the missed road parts using histogram backprojection method. First, the histogram is formed by the statistical distribution of the pixel values in the recognized road region. Then, the brightness of the histogram peak at the corresponding point is calculated, which reserves the grayscale 0.8 to 1.2 times. Histogram is cut off as the probability density curve of the road. The entire

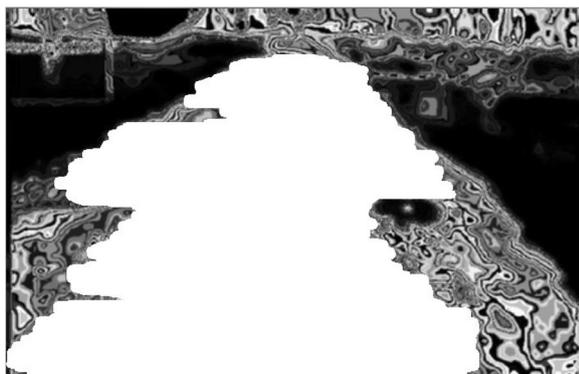


Fig. 14 Road region marking diagram



Fig. 15 Road region based on backprojection probability map

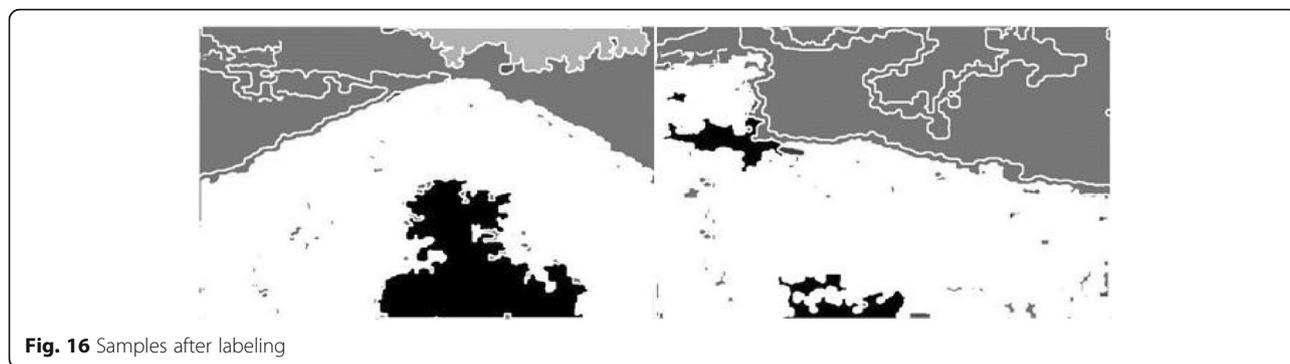


Fig. 16 Samples after labeling

image is transformed into a histogram backprojection probability map. The histogram backprojection probability of the image transformed by FCM cluster for the road's gray-based model is shown in Fig. 13.

On the basis of the histogram backprojection, the road regions and road edges are obtained by combining the results of the road region recognition based on the probability confidence of FCM with the trapezoid prediction model. Road region recognition is complemented by the histogram backprojection. This process involves the following steps: (1) The boundary of the auxiliary backprojection probability map is used to segment the image, while the top region of the vanishing line is removed. (2) Road region recognition based on FCM and the probability confidence map of the trapezoid prediction model is marked by white pixels (255). From Fig. 14, it can be seen that the white area (compared with the real road region) is missing. (3) Whether to keep or delete the connected region depends on the presence of the white mark in the residual connected region of the statistical backprojection probability map.

The new road region is used to recognize the road boundary, and the corresponding recognition result is shown in Fig. 15. It observed that the road boundary is more accurate and accordant with the characteristics of human vision. The histogram backprojection method based on grayscale values has a good effect on road region recognition, which compensates for the deficiency of the connected region.

2.5 Implementation

The system is implemented on a PC (Intel Core i5 at 3.30 GHz with 16 G RAM). The average running time of every image frame is evaluated with Matlab R2015a on Windows 7. The system is tested under complex field conditions for vanishing line detection, image segmentation, and clustering. Grayscale histogram with infrared images is used for feature correspondence. We deliberately selected video clips that were recorded under difficult conditions. All of the images used in our experiment are downloaded from a vehicle-mounted infrared imager.

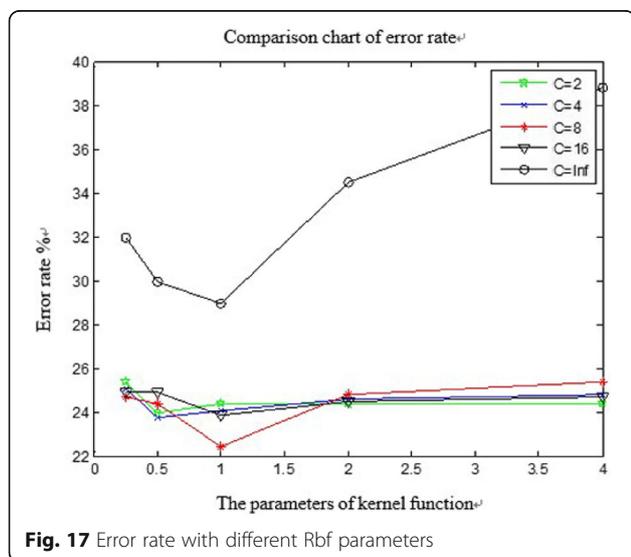


Fig. 17 Error rate with different Rbf parameters

3 Results and discussion

In this section, first, we describe the experimental setting, including the sample collection, the parameters, and classifier choice. Then, the comparison of the vanishing line detection results is depicted. Finally, we compare the proposed approach with the state-of-the-art approaches in terms of accuracy and time consumption.

Table 2 Performance comparison of three SVM classifiers

SVM classifier	Determination time (s)	Accuracy rate (%)
Non-classified normalization	0.0023~0.0043	92.48
Classified normalization	0.003~0.006	93.21
Combination of base classifiers	0.011~0.025	94.96

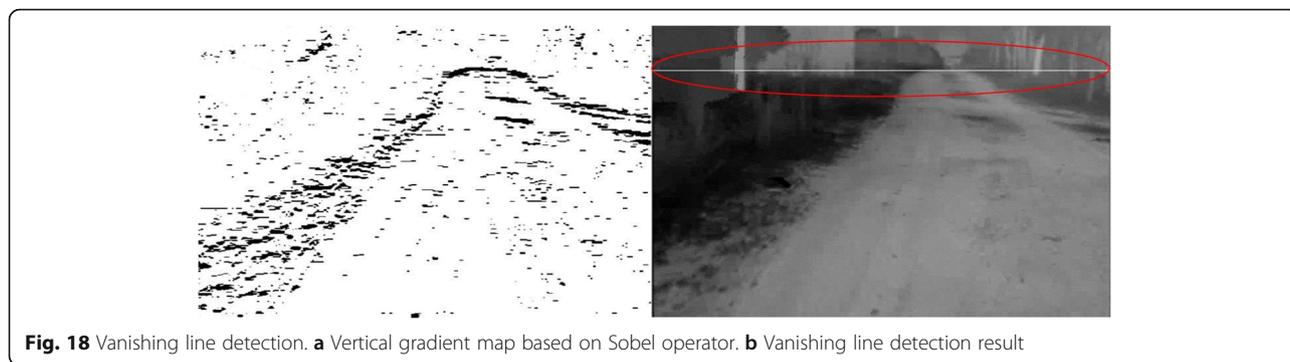


Fig. 18 Vanishing line detection. **a** Vertical gradient map based on Sobel operator. **b** Vanishing line detection result

3.1 The experimental setting

3.1.1 Sample collection

Three features (grayscale, position, and shape) are used as the extracted features. Although there is no RGB (red, green and blue) color information, the infrared image can provide plenty of information in grayscale. Grayscale frames in the road region are different according to changes in the road conditions, but the grayscale distribution in a continuous video is stable. Furthermore, the grayscale value of the road shadow is lower than that of the road region, where the background is on both sides of the road in the image. In a relatively simple scene, location information can also provide a priori information which is very important in road recognition. For example, the road is at the bottom and the middle of the image, while the sky is at the top. In addition, we adopt the shape features due to the meaningful connected regions extracted in our paper.

In this section, extensive experiments are discussed that were used to validate the effectiveness of the proposed approach on a private image database from a Chinese military project. The infrared images captured by the vehicle-mounted infrared imager have a fixed 702×576 resolution with the D1 format. We marked 93 images that were selected from the vehicle-mounted infrared video in natural fields. These selected images contain road bends, forks, and shadows, which are used to validate the effectiveness of the proposed approach by

comparing the results with the ground truth. The recurring objects in the images roughly contain roads, shadows, sky, surrounding trees, and other non-road regions of unknown materials. The marked results are shown in Fig. 16: white (255) represents the road, and black (0) represents the shadowed parts of the road. The sky is marked as 180, the non-road and unknown regions are marked as 120, and the trees are marked as 60. Then, we extracted features from the marked images to generate a total of 1854 samples. Among them, there are 826 road samples and 94 road shadow samples, totaling 920 positive samples. Furthermore, there are 934 negative samples with 0, 10, and 924 samples of the sky regions, surrounding trees and non-road regions, respectively. Ten-fold cross-validation is adopted to select the smallest error as the training set.

3.1.2 Parameter setting

Radial basis function (Rbf) kernel is chosen and the kernel function parameter σ^2 and penalty factor c are determined by the grid method. There are 25 different groups. σ^2 takes $[2^{-2}, 2^{-1}, 2^0, 2^1, 2^2]$, and c takes $[2^1, 2^2, 2^3, 2^4, \text{Inf}]$. We chose the minimum error rate group as the selected parameter in the groups. The number of training samples is reduced to 102 to save time. The classification performance of different parameters is shown in Fig. 17. It can be seen that the error rate is the

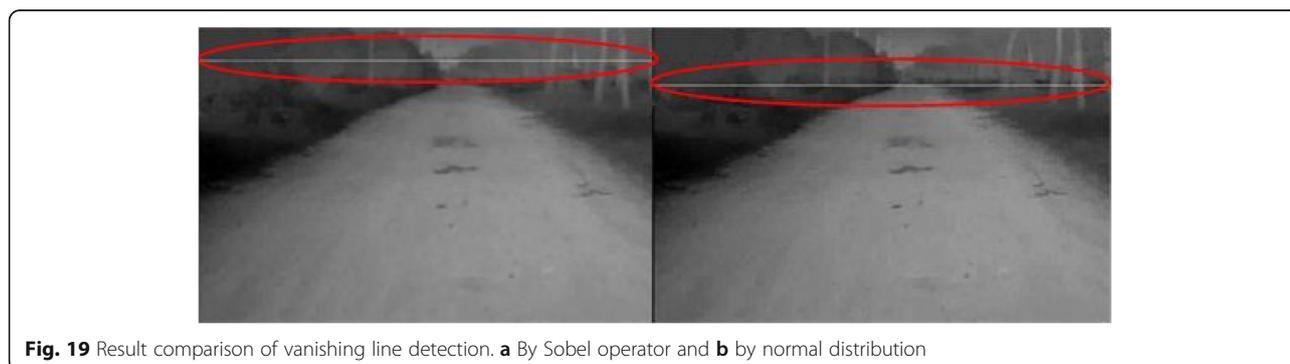


Fig. 19 Result comparison of vanishing line detection. **a** By Sobel operator and **b** by normal distribution

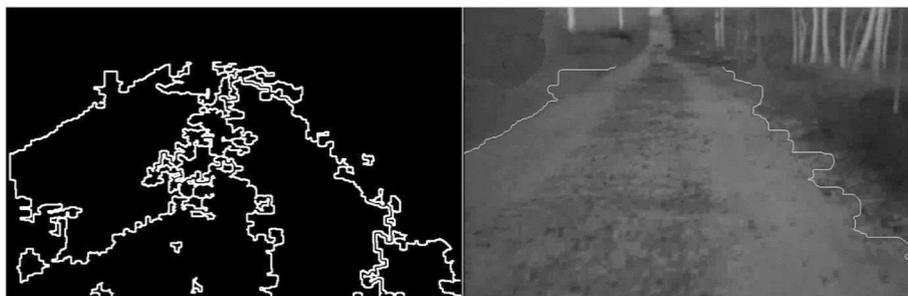


Fig. 20 DBSCAN clustering and road boundary recognition. **a** Clustering result. **b** Road boundary recognition

smallest when setting $\sigma^2 = 1$ and $C = 8$ are used as the parameters for the training set.

3.1.3 Classifier selection

In this paper, we proposed three SVM classifiers: the non-classified normalization classifier, classified normalization classifier, and combination classifier with the above two classifiers, as shown in Table 2. Although the accuracy of the combined classifier is high compared with the other two methods, the discriminant procedure is time-consuming. The accuracy of the classification normalization classifier is relatively higher compared with that of the non-classification one, and the additional time is in an acceptable range. Therefore, we make a trade-off among the three methods to choose the classification normalization as SVM classifier for road recognition in this paper.

3.2 Comparison of vanishing line detection

For vanishing line detection, Sobel operator is usually used because it has important characteristics, such as fast learning and simple operations that can be used to quickly obtain continuous and smooth boundaries. Sobel operator extracts vertical gradient features, as shown in Fig. 18a. It is easily seen that the vanishing line is very obvious in the vertical gradient. In fact, it appears as the longest straight line located in the upper part of the image. Therefore, the vanishing line can be acquired by

scanning the image to search for the longest straight line as shown in Fig. 18b.

The unpaved road suffers from the following issues such as interference caused by shadows and ruts. Sobel operator ignores the complex structural relationships in unpaved roads and therefore can hardly achieve an accurate road detection result. As seen from Fig. 18a, there are some interferences and deviations of the vanishing line. In order to solve this problem, we present a normal distribution method as shown in Fig. 19b. By comparing the results in Fig. 19a with those in Fig. 19b, we observe that the normal distribution method that adds a priori information is more effective than Sobel operator.

3.3 Analysis of image clustering methods

In this work, several methods are considered for detecting road regions. To make a comparison, we also report results for some state-of-the-art detection methods, including those based on (1) DBSCAN method [24] and the result is shown in Fig. 20, (2) K -means method [25] and the result is shown in Fig. 21, (3) least squares method [17], and (4) histogram [15]. To achieve a more valid comparison, we adopt the evaluation mechanism proposed in [37] to analyze the results from the above road detection algorithms. Accuracy is defined as follows:

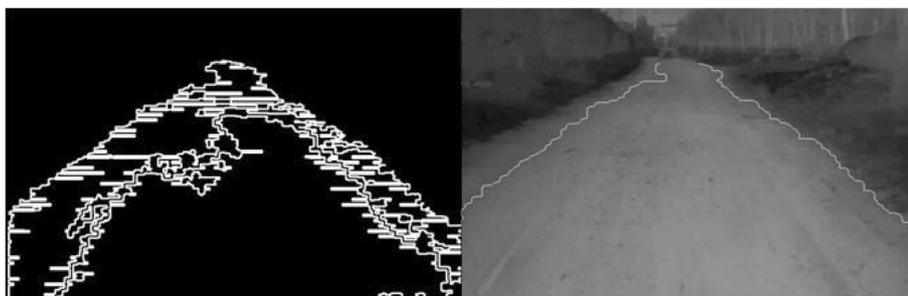


Fig. 21 K -means clustering and road boundary recognition. **a** Clustering result. **b** Road boundary recognition

Table 3 Summary of experimental results on time consumption and assessment

Method	Time consumption (ms)	Accuracy (%)
DBSCAN-based [24]	29	66.80
K-mean-based [25]	12	88.60
Least squares-based [17]	74	77.20
Histogram-based [15]	37	91.10
Ours	14	93.20

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (16)$$

where true positive (TP) is the number of correctly labeled road pixels, true negative (TN) is the number of non-road pixels detected, false positive (FP) is the number of non-road pixels classified as road pixels, and false negative (FN) is the number of road pixels erroneously marked as non-road. The better detection algorithm's performance can obtain the greater accuracy value.

Before calculating the accuracy, we first transform the gray images in Fig. 16 (the ground truth) into binary images. In our experiment, we regard white (255) and black (0) as the road and the other conditions are the non-road. Here, the road pixels are 1 and the non-road pixels are 0.

Moreover, we record the total time consumption and accuracy of all of the methods, as shown in Table 3.

As seen in the aforementioned analysis, it is easy to note that our algorithm achieved obvious improvements and better performance and is more adaptive to real-time road detection.

4 Conclusions

In this paper, we present a method of unpaved road recognition and boundary extraction using infrared images. In the proposed approach, the accuracy and robustness of vanishing line detection can be achieved by means of a normal distribution. The improved SVM classifier is trained with the optimal radial kernel parameters and penalty factor parameters for road detection. Meanwhile, the improved SVM classifier achieves recognition of the unpredictable road from connected regions. Based on the obtained connected regions, the road probability confidence map and trapezoidal forecasting model are formed. Furthermore, we also propose a more suitable FCM cluster to further improve the accuracy of road region recognition. After that, the membership of each connected region can be obtained. The corresponding probability confidence map is able to update the membership degree of the connected region. The accuracy of the results is 93.20%. Finally, histogram backprojection method based on the pixel model is used to complement a road region. The experimental results show the advantages and effectiveness of the proposed algorithm.

Abbreviations

DBSCAN: Density-based spatial clustering of applications with noise; FCM: Fuzzy C-means; FN: False negative; FP: False positive; HD: High definition; ML: Machine learning; Rbf: Radial basis function; RGB: Red, green and blue; ROI: Region of interest; SVM: Support vector machine; TN: True negative; TP: True positive

Acknowledgements

The authors would like to thank the editor.

Availability data and materials

Not applicable.

Funding

This study was supported by the National Natural Science Foundation of China (No. 61471110), Foundation of Liaoning Provincial Department of Education (L2014090), and Chinese Universities Scientific Foundation (N160413002, N16261004-2/3/5).

Authors' contributions

In this section, the contributions of each author are exposed. JB is the main author. She obtained the results, made the research and comparison with other authors, and wrote the article. XS helped in optimizing the code with Matlab and fine-tuning the parameters of the algorithms to obtain the best results. YZ and RZ made the bibliography review, analyzed the results, and proposed the optimization of the parameters. They supervised all the work, reviewed the article, and helped in all the process. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹College of Information Science and Engineering, Northeastern University, Shenyang 110819, China. ²Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110819, China.

Received: 29 August 2017 Accepted: 13 March 2018

Published online: 20 April 2018

References

- H Kong, JY Audibert, J Ponce, General road detection from a single image. *IEEE Trans. Image Process.* **19**(8), 2211–2220 (2010)
- A Kūçükmanisa, G Tarm, O Urhan, Real-time illumination and shadow invariant lane detection on mobile platform. *J. Real-Time Image Proc.*, 1–14 (2017)
- M Bendjaballah, S Graovac, MA Boulahlib, A classification of on-road obstacles according to their relative velocities. *EURASIP journal on image and video processing* **2016**(1), 41 (2016)
- C Yan, H Xie, S Liu, et al, Effective Uyghur language text detection in complex background images for traffic prompt identification. *IEEE Trans. Intell. Transp. Syst.* **99**, 1–10 (2017)
- C Yan, H Xie, D Yang, et al, Supervised hash coding with deep neural network for environment perception of intelligent vehicles. *IEEE Trans. Intell. Transp. Syst.*, 1–12 (2017)
- S Zhou, K Iagnemma, *Self-supervised learning method for unstructured road detection using fuzzy support vector machines, RSJ International Conference on Intelligent Robots and Systems* (IEEE, IROS, 2010), pp. 1183–1189
- M Bertozzi, A Broggi, M Cellario, et al, Artificial vision in road vehicles. *Proc. IEEE* **90**(7), 1258–1271 (2002)
- E Shang, X An, J Li, et al, Robust unstructured road detection: the importance of contextual information. *Int. J. Adv. Robot. Syst.* **10**(3), 179 (2013)
- C Yan, Y Zhang, J Xu, et al, A highly parallel framework for HEVC coding unit partitioning tree decision on many-core processors. *IEEE Signal Processing Letters* **21**(5), 573–576 (2014)

10. J Shi, J Wang, F Fu, Fast and robust vanishing point detection for unstructured road following *IEEE Transactions on Intelligent Transportation Systems*. **17**(4), 970–979 (2016)
11. X Yang, G Wen, *Road extraction from high-resolution remote sensing images using wavelet transform and Hough transform*, 2012 5th International Congress on Image and Signal Processing (IEEE, CISP, 2012), Chongqing Univ Posts & Telecommunications, Chongqing, Peoples Republic of China, pp. 1095–1099
12. Y Wang, C Wen, Vanishing point detection of unstructured road based on Haar texture. *Journal of Image and Graphics* **18**(4), 382–391 (2013)
13. Y Chen, Y Yu, T Li, *A vision based traffic accident detection method using extreme learning machine*, International Conference on Advanced Robotics and Mechatronics (IEEE, ICARM, 2016), Macau, China, pp. 567–572
14. K Lu, S Xia, D Chen, et al., *Unstructured road detection from a single image*, Control Conference (IEEE, 2016), Kota Kinabalu, Malaysia, pp. 1–6
15. W Zuo, T Yao, Road model prediction based unstructured road detection. *Journal of Zhejiang University SCIENCE C* **14**(11), 822–834 (2013)
16. E Shang, X An, L Ye, et al., *Unstructured road detection based on hybrid features*, Proceedings of the 2012 2nd International Conference on Computer and Information Application (ICCIA, 2012), Taiyuan, Peoples Republic of China, pp. 926–929
17. K Kluge, S Lakshmanan, *A deformable-template approach to lane detection*, In Proc Intelligent Vehicle '95 Symposium (IEEE, 1995), Detroit, MI, USA, pp. 54–59
18. X Hu, R FSA, A Geppert, *A multi-modal system for road detection and segmentation*, Intelligent Vehicles Symposium Proceedings (IEEE, 2014), pp. 1365–1370
19. M Nishida, M Muneyasu, *Detection of road boundaries using hyperbolic road model*, 2012 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS, IEEE, 2012), pp. 521–526
20. DY Huang, CH Chen, TY Chen, et al. Vehicle Detection and Inter-vehicle Distance Estimation Using Single-lens Video Camera on Urban/Suburb Roads. *J Vis Commun & Image Repres.* **46**, 250–259 (2017)
21. H Zhang, D Hou, Z Zhou, A novel lane detection algorithm based on support vector machine. *Piera online* **1**(4), 390–394 (2005)
22. S Zhou, J Gong, G Xiong, et al., *Road detection using support vector machine based on online learning and evaluation*, Intelligent Vehicles Symposium (IEEE, IV, 2010), San Diego, CA, pp. 256–261
23. S Yun, Z Guo-Ying, Y Yong, *A road detection algorithm by boosting using feature combination*, 2007 IEEE Intelligent Vehicles Symposium (IEEE, 2007), Istanbul, Turkey, pp. 364–368
24. S Chakraborty, NK Nagwani, Analysis and study of incremental DBSCAN clustering algorithm. Preprint ArXiv **1**(2), 1406.4754 (2014)
25. ME Celebi, HA Kingravi, PA Vela, A comparative study of efficient initialization methods for the k-means clustering algorithm. *Expert Syst. Appl.* **40**(1), 200–210 (2013)
26. P Conrad, M Foedisch, *Performance evaluation of color based road detection using neural nets and support vector machines*, Proceedings. 32nd Applied Imagery Pattern Recognition Workshop (IEEE, 2003), Washington, DC, pp. 157–160
27. ZQ Li, HM Ma, ZY Liu, *Road lane detection with Gabor filters*, International Conference on Information System and Artificial Intelligence (IEEE, 2017), Hong Kong, Peoples Republic of China, pp. 436–440
28. W Wu, G ShuFeng, *Research on unstructured road detection algorithm based on the machine vision*, 2009 Asia-Pacific Conference on Information Processing (APCIP, IEEE, 2009) pp.112-115
29. N Otsu, A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics* **9**(1), 62–66 (1979)
30. RD Gupta, M Jamil, M Mohsin, Discharge prediction in smooth trapezoidal free overfall—(positive, zero and negative slopes). *J. Irrig. Drain. Eng.* **119**(2), 215–224 (1993)
31. C Yan, Y Zhang, J Xu, et al., Efficient parallel framework for HEVC motion estimation on many-Core processors. *IEEE Transactions on Circuits & Systems for Video Technology* **24**(12), 2077–2089 (2014)
32. T Joachims, *Making large-scale SVM learning practical. technical report on Komplexitätsreduktion in Multivariaten Datenstrukturen* (SFB, Universität Dortmund, 1998), Cambridge, USA, p. 475
33. U Aich, S Banerjee, Application of teaching learning based optimization procedure for the development of SVM learned EDM process and its pseudo Pareto optimization. *Appl. Soft Comput.* **39**, 64–83 (2016)
34. X Liu, KK Shang, J Liu, et al., *Unstructured road detection based on fuzzy clustering arithmetic*, 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (IEEE, FSKD, 2014), Xiamen, Peoples Republic of China, pp. 114–118
35. YH Lai, PW Huang, PL Lin, *An integrity-based fuzzy c-means method resolving cluster size sensitivity problem*, 2010 International Conference on Machine Learning and Cybernetics (IEEE, ICMLC, 2010), Qingdao, Peoples Republic of China, pp. 2712–2717
36. J Ji, KL Wang, A robust nonlocal fuzzy clustering algorithm with between-cluster separation measure for SAR image segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**(12), 4929–4936 (2014)
37. JM Ivarez, A Lopez, *Novel index for objective evaluation of road detection algorithms*, 11th International IEEE Conference on Intelligent Transportation Systems (IEEE, ITSC, 2008), Beijing, Peoples Republic of China, pp. 815–820

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com