The Railway Detection via Adaptive Multi-scale Fusion Processing

Qian Peng\(^1\), Shiwei Ren\(^1\), Weijiang Wang\(^1,2\), and Yueting Shi\(^1,3\)
\(^1\)School of Information and Electronics, Beijing Institute of Technology, Beijing, People's Republic of China
\(^2\)3120180691@bit.edu.cn, \(^3\)renshiwei@bit.edu.cn, \(^4\)wangweijiang@bit.edu.cn
\(^5\)shiyuetings@bit.edu.cn

Abstract. One of the main problems for safe autonomous driving vehicles that have not been solved completely is the high-precision and timely lane detection. In this work, we present a novel operator for railway detection to settle these tasks based on lane detection for the first time, called adaptive multi-scale fusion Sobel operators. The new operators can eliminate the noises generated by the environment in the railway image and derive more integrated edge feature information from the 0\(^\circ\), 45\(^\circ\), 90\(^\circ\), and 135\(^\circ\) detection via 4 matrixes of 3 * 3 operators for permutation and summation. The image processing for railway detection includes the preprocess for images, railway edge detection, and track line polynomial fitting. Our experiment has validated that this improved detection method has realized the high accuracy and efficiency for rail detection. The dynamic rail detection and identification in the video of the railway track prove that this method has a significant effect on the left and right curved railway detection. It has good robustness and applicability.

1. Introduction

In recent years, a good deal of researchers has been contributed to developing automatic driving systems or assisted driving systems [1-2] to reduce traffic accidents caused by the lane deviating from the correct transition lane. Self-driving vehicles collect information through various sensors and then make decisions on the surrounding environment and lane lines by scientific analysis. The lane detection technology is generated to ensure that the vehicle is on a safe route and runs correctly by providing the most basic lane and environment information for automatic driving vehicles to maintain or change the right lanes. Therefore, the lane line detection technology plays an important role in the automatic driving system, and it directly determines the instantaneity and accuracy of the system.

The lane detection technology is very challenging. Firstly, estimating lanes needs to require a semantic understanding of the scene. Moreover, the environment is diverse and subject to several conditions of weather and illumination, which might affect the driving change. Up to now, there are mainly two developed directions for lane estimation. One is the traditional approach which consists of the extraction of hand-crafted features followed by a curve-fitting process, including the camera correction, filtering [3], ROI (region of interest) selection [4], operators (Canny [5,8], Sobel [6], etc.) for feature extraction, the polynomial fitting of lanes and other processes. Many researchers also use the Hough transform [7] to quickly identify lane lines. [9] selected the alternate lane through object segmentation, utilized the Sobel operator to extract the excess edges, obtained lane labels by selecting as many edge thresholds as possible, and finally adopted segmented fitting. [10] learned real-time fast high-speed lane detection based on OpenCV. image processing algorithm. With the boom of artificial intelligence, the lane detection technology based on deep learning is developing rapidly [11]. [12] proposed the task of 3D lane line detection for the first time, and realized end-to-end lane detection through the anti-perspective transformation inside the innovative network and the newly created anchor points. [13] is proportional to the past and can be trained end-to-end, with superior parameters and faster speed, which established a parametric model of the lane line detection problem, and used Transformer model to capture the lane line features and symbol features on the road pictures.

In this context, the method is based on the traditional image process, focusing on improving the accuracy of edge detection benefitting self-driving train systems and advanced driver assistance systems (ADAS). The general Sobel operator always extracts limited feature information. This paper proposes an adaptive multi-scale fusion Sobel operator(mul_sobel) based on [14], collecting more
complete features information by convolution from different angles. The approach needs to preprocess the rail images while observing the characteristic information firstly. Then the adaptive multi-scale Sobel operator that could smooth the noise is employed for feature extraction. Finally, the appropriate ROI is selected to serve histograms and sliding windows are utilized to fit each railway. We have obtained better inspection results in the straight, left, and right railway.

2. Proposed Solution

The railway detection approach contains the main three steps: preprocess for original images, edge detection of the railway, and polynomial fitting for each railway line.

It is vital to preprocess the original image captured by the camera firstly. The influence of factors such as illumination, obstructions, and road defects increase the difficulty of trajectory feature extraction. The intrinsic of the camera also affects the quality of the original image taken too. Moreover, the color is easily affected by factors such as light, which makes it difficult to provide critical information. The grey-scale processing of the color image can reduce the size and greatly increase the calculation speed in unnecessary color-based detection. In the detection of rail lines, our experiments have proved that important image features can be obtained only by greying the original image and then filtering and detecting feature information.

The most important step of railway detection is to detect the edges of the railway and the background. Since different incident points have different grey values, the boundaries between different objects usually have obvious edges including gradient information. Therefore, different operators can use this feature to segment the image for edge detection. Currently, railway edge detection is mostly based on the Canny operator and the Sobel operator. However, because the Canny operator is easy to misjudge the noise point as the boundary, the extracted lane edge is too thin to lose important edge information. Compared with the railway track perpendicular to the highway lane, the former will generate more noise. The rail line is different from the highway lane line feature extraction, so we choose the Sobel operator that can smooth the noise for feature detection.

We propose adaptive multi-scale fusion operators. The basal Sobel operator consists of duplicate matrixes of variables 3*3, which are the horizontal and vertical convolution kernels convolved with the image respectively to convert the horizontal and vertical discrete difference approximations. If A is the original image, Gx and Gy are the image grey values of the horizontal and vertical edge detection, the values of Gx and Gy can be calculated by formulas (1-2). The gradient of every pixel in the image is calculated by the formula (3) combining the horizontal and vertical gradient approximations.

\[
G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A 
\]

\[
G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * A 
\]

\[
G = \sqrt{G_x^2 + G_y^2} 
\]

Although the Sobel operator can smooth noise, it only considers the gradient value of the level and the correct direction. The edge information acquisition degree of the image may be insufficient. In edge detection, there is usually no obvious difference between the size of the image object, causing surrounding objects to interfere with the edge of the target and the poor positioning accuracy of the target edge. [19] provided a multi-scale Sobel operator shown in Figure 3, which is used to plus the gradient of the different directions including 0°, 45°, 90° and 135°.
The gradient of the general multi-scale Sobel operator is calculated by the formula (4).

\[ G = \sqrt{\sum a_i G_i^2} \]

(4)

Figure 4 is a comparison diagram of the edge detection results from the four direction convolution kernels based on Figure 3. We found that there are more complete rail edge information and less interference information extracted from 0° and 45° directions. Therefore, we propose an adaptive multi-scale fusion algorithm, calculated by formulas (5-6). We add weight values multiplied with the gradient of different directions and summed to calculate the pixel total gradient, which will get a large weight value if it performs well in extracting edge features, for retaining more complete feature information and enhancing the edge detection effect.

\[ G = \sqrt{\sum a_i G_i^2} \]

(5)

\[ \sum a_i = 1, i \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\} \]

(6)

After the edge detection of the railway image, we can polish the rail lines edge information from the result obviously, and we select ROI (Region of Interest) as a method of intelligent video coding, its function is to improve the efficiency of the processing process without damaging the image quality and continue to fit the railway lines according to the feature information.
3. Experiments

3.1. Image Pre-processing

3.1.1. Camera Correction. We use the checkerboard method provided by the OpenCV API to detect the image points of the black and white checkerboard calibration images from different angles, and then combine the known object points to get the camera calibration matrix and distortion coefficient, etc., the final corrected image is shown in Figure 1(b) according to the original trajectory image Figure 1(a). The subsequent image processing operations are based on the corrected image.

3.1.2. Image Grayed and Median Filtering. The image captured by the camera is severely affected by discrete pulses, pepper-salt noise, and zero-mean Gaussian noise. Median filter separation sorts all pixels inside the window function to obtain the median value to represent the value of the center of the window. The suppression of noise is particularly good while retaining edge details. Formula (7) is a two-dimensional median calculation formula.

\[
g(x, y) = \text{median}[f(x - k, y - l), (k, l \in W)]
\]

where \(f(x, y)\) and \(g(x, y)\) are the original image and the processed image, \(W\) is a two-dimensional template, and \(K\) is usually 3*3, 5*5 convolution kernels. Figure 2 shows the results of grayscale and median filtering, we found that some noise in the image is blurred after filtering.

3.2. Adaptive Multi-scale Fusion Edge Detection

The image edge information is extracted according to the ordinary Sobel operator and ordinary multi-scale Sobel detection, as shown in Figure 6 (a) (b). The two methods increase noise or edge gradient interference of other objects, and cannot accurately detect the rail edge.

\[
G = \sqrt{0.8G_{0^\circ}^2 + 0.1G_{45^\circ}^2 + 0.05G_{90^\circ}^2 + 0.05G_{135^\circ}^2}
\]

Experiments proved that when the weight values of 0°、45°、90° and 135° as 0.8, 0.1, 0.05, and 0.05 to extract the feature effect as formula (7), the image not only retains more useful edge information but also enhances the edge detection results of the track. Figure 5 (d) shows the effect of using the [-1,0,1] and [1, 0, -1] operators to enhance edge features after adaptive multi-scale Sobel operator edge detection, which is beneficial to Histogram feature statistics and accurate positioning of the left and right track line coordinates.

![Figure 5. Comparison charts of edge detection](image-url)
3.3. Railway Lines Fitting

3.3.1 ROI Selection and Perspective Transformation. ROI (Region of Interest) as a method of intelligent video coding, its function is to improve the efficiency of the processing process without damaging the image quality, there is useless information when detecting lane lines, and the redundancy is reasonably eliminated Information, as shown in Figure 6 (a), and then remove some noise and abnormal values in the ROI, and finally get Figure 6 (c), then eliminate noise interference and improve the efficiency of feature detection.

The installation position of the camera is at a certain angle with the ground, which causes the rail image to appear near and far, so that the parallel left and right rails first produce an inclination angle that affects the railway detection. Therefore, the perspective transformation is used to convert the actual rail image into a bird’s eye view of the rail firstly, shown in Figure 6 (b), and restore the parallel state of the track for fitting the rail.

![ROI masked](image1.png) ![BirdViews](image2.png) ![cleaned](image3.png) ![histogram](image4.png)

**Figure 6.** Results of ROI selection, bird’s views, and histogram

3.3.2 Histogram and Polynomial Fitting of Railway Lines. According to the bird’s eye view, the histogram is calculated step by step at the vertical y-axis, and the result is shown in Figure 6 (d). The peak points at the left and right from the histogram statistics are the starting point coordinates of the left and right rail lines. We design ten small windows, which move slightly to the horizontal direction, find the effective steering of the left and right rail line, and effectively remove the surrounding interference factors at the same time.

Without the interference of other noises, the rail line inside the collection area of the sliding window will not have a sudden change in direction. Therefore, the displacement point inside the window is greater than a certain threshold, and the point is used as the center of the next sliding window. The sliding window result is shown in Figure 7.

![sliding left](image5.png) ![sliding right](image6.png)

**Figure 7.** Sliding windows with left and right lines
Finally, we use the quadratic multivariate formula and formula (8) to fit these center points to obtain the orbit curve. The fitting result is shown in Figure 8. Figure 8 (a) is the fitted curve, and (b) is the result of curve mapping to the bird’s eye view.

\[ y = a_2 x^2 + a_1 x + a_0 \] (9)

Figure 8. Quadratic polynomial fitting and predicted line

The orbital curvature R can be calculated by the formula (9) based on the fitted quadratic multiplex orbital curve. Then through the anti-perspective transformation, the rail fitting result is compared with the rail image. The final result is shown in Figure 9.

Figure 9. Final result

4. Result and Analysis

The experiment also tests the performance of adaptive multi-scale fusion edge detection on other linear rail images. Figure 10 is divided into ordinary multi-scale fusion detection. Adaptive multi-scale detection can detect the rail line more completely and is also applicable. The detection results of the left and right curved rails show in Figure 11.
In order to verify the real-time performance of the whole system, we intercept the video of the straight rail and the left and right curved rail of the same train when the train was running. This experiment is the first time to conduct dynamic rail line detection on track railway video, we define the difference $S$ between the midpoint $R_c$ of the left and right curve of the track fitting, and the midpoint coordinate $P_c$ of the image itself is used as the evaluation criteria, formula (10). And we intercept part of the straight track video to align the training set to train the difference range of the straight rail $S (-0.34~-0.28)$. When the difference is less than the range, it is the left track, and when the difference is greater than the range, it is the right track, the judgment is shown in equation (11).

$$S = R_c - P_c$$

$$Label = \begin{cases} 
  Left & S \geq -0.28 \\
  Straight & -0.34 < S < -0.28 \\
  Right & S < -0.34 
\end{cases}$$

Table 1 and Table 2 are the detection of the straight track, the left and right rails in the railway track videos which respectively take 10 seconds, and every 200 frames of images total. the Labels in the first column are the testing rail videos, and the first row is the predicted labels, the numbers represent prediction results. Table 1 shows the ordinary multi-scale Sobel operator for orbit detection, and Table 2 uses the adaptive multi-scale Sobel operator for rail detection.

Experimental results show that the detection accuracy of the adaptive multi-scale algorithm is improved to 85% in the linear orbit compared with the 82% accuracy of the ordinary multi-scale algorithm. Meanwhile, the accuracy in the left and right orbits is also as high as 91% and 94%. The
tables proved the usability of criteria S, the robustness and applicability of the adaptive multi-scale fusion detection algorithm based on the Sobel operator.

### Table 1. The ordinary multi-scale Sobel operator

| Labels | Straight | Left | Right | Success Rate |
|--------|----------|------|-------|--------------|
| Straight | 164 | 33 | 3 | 83% |
| Left | 40 | 158 | 2 | 79% |
| Right | 47 | 2 | 151 | 75.5% |

### Table 2. The adaptive multi-scale Sobel operator

| Labels | Straight | Left | Right | Success Rate |
|--------|----------|------|-------|--------------|
| Straight | 171 | 29 | 0 | 85% |
| Left | 18 | 182 | 0 | 91% |
| Right | 12 | 0 | 188 | 94% |

5. **Conclusion and Future Scope**

The lane line detection requires image preprocess, edge detection, feature extraction, fitting and other processes. In addition, the lane features are affected easily by factors such as illumination and obstructions to increase the detection difficulty. An adaptive multi-scale fusion algorithm is proposed here for railway track detection. Sobel operators from different scales are used to extract the proportions of important feature offsets from different feature information, which ensures the integrity of feature information, reduces the impact of noise jamming, and gets better detection accuracy on a straight rail and left and right curved track.

The detection technology of railway tracks is still weakening to the lane detection of ordinary highways. On the one hand, there are no complete data set, mature detection methods, and evaluation indicators. On the other hand, researchers have investigated its use to tackle lane detection with the success of deep learning. Whether it’s suitable for rail detection via deep learning, machine learning, and other AI technologies is still needed in-depth discussion and research. We will continue to explore these aspects further.

6. **References**

[1] FAN Chao, DI Shuai, HOU Li-long. A lane recognition algorithm based on line model[J]. Application Research of Computers, 2012, 29(1): 326-328, 332

[2] Kumar A M, Simon P. Review of lane detection and tracking algorithms in advanced driver assistance system[J]. Int. J. Comput. Sci. Inf. Technol, 2015, 7(4): 65-78

[3] X. Wang, Y. Wang, C. Wen, Robust lane detection based on gradient-pairs constraint, in: Proceedings of the 30th Chinese Control Conference, IEEE, 2011, pp. 3181–3185

[4] P.-C. Wu, C.-Y. Chang, C.H. Lin, Lane-mark extraction for automobiles under complex conditions, Pattern Recognition 47 (8) (2014) 2756–2767 29.

[5] C. Mu, X. Ma, Lane detection based on object segmentation and piecewise fitting, TELKOMNIKA Indones, J. Electr. Eng, TELKOMNIKA 12 (5) (2014) 3491–3500.

[6] P.-C. Wu, C.-Y. Chang, C.H. Lin, Lane-mark extraction for automobiles under complex conditions, Pattern Recognit 47 (8) (2014) 2756–2767 29.

[7] J. Niu, J. Lu, M. Xu, P. Lv, X. Zhao, Robust lane detection using two-stage feature extraction with curve fitting, Pattern Recognit 59 (2016) 225–233.

[8] Li H, Nashashibi F. Robust real-time lane detection based on lane mark segment features and general a priori knowledge [C]// 2011 IEEE International Conference on Robotics and Biomimetics. IEEE, 2011: 812-817.
[9] Mu C, Ma X. Lane detection based on object segmentation and piecewise fitting[J]. TELKOMNIKA Indones. J. Electr. Eng. TELKOMNIKA, 2014, 12(5): 3491-3500.

[10] Z. Wang, Y. Fan and H. Zhang, "Lane-line Detection Algorithm for Complex Road Based on OpenCV," 2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), Chongqing, China, 2019, pp. 1404-1407, doi: 10.1109/IMCEC46724.2019.8983919.

[11] B. Jig ang Tang A, S. L. A. B, and P. L. A. B. "A Review of Lane Detection Methods based on Deep Learning." Pattern Recognition (2020)

[12] N. Garnett, R. Cohen, T. Pe'er, R. Lahav and D. Levi, "3D-LaneNet: End-to-End 3D Multiple Lane Detection," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 2921-2930, doi: 10.1109/ICCV.2019.00301.

[13] Zou, Qin, et al. "Robust Lane Detection From Continuous Driving Scenes Using Deep Neural Networks." IEEE Transactions on Vehicular Technology 69.1(2020):41-54.

[14] Luo J H, Lan F C and Chen J Q. "A lane detection method that combines image processing and target constraints." Chinese Journal of Automotive Engineering. 9.01(2019):30-38.