Towards Autonomous Driving: Road Surface Signs Recognition using Neural Networks

Stephen Karungaru, Jumpei Yamamoto, Kenji Terada
University of Tokushima, 2-1 Minami Josanjima, Tokushima City, Japan

Abstract: In recent years, the number of traffic accidents has rapidly increased for many reasons. However, the most common cause is the driver carelessness and inattentiveness to road signs. Therefore, the aim of this paper is to automatically recognize road surface markings and by availing the information to the driver, hope to reduce road accidents. In the proposed method, the captured image is transformed and edges information used to extract the target road area. The road marking candidate is then extracted and recognized using a neural network.

Keywords: Autonomous driving, Road surface signs, Neural Network.

I. INTRODUCTION

Recently, although the number of serious car accidents is reducing, the total number of accidents is still high. There are various factors contributing to occurrence of accidents; drunk driving, fatigue, use of mobile devices, carelessness, etc. Careless driving examples include the driver not paying attention to his surroundings (especially checking for cars, people, etc.), not confirming the road markings or signs present, speeding, etc. In addition, because there are dozens of road markings, the driver may not always remember the meaning of all of them. Nevertheless, it is paramount that the driver be careful while driving. These issues have raised the need for autonomous driving systems. As a part of such a system, we propose to automatically recognizing the road markings. We define road marking as the signs and information painted on the road surface. We believe that traffic accidents can be reduced by providing he driver with this information.

Road marking recognition has been studied by many researchers [1-6, 9, 10-12]. A generative learning method is used in [1,9] to solve shape deformation, resolution, blurring, etc. while [2] proposes a feature driven approach using a trained classifier combined with additional rule-based post-processing then facilitates the real-time delivery of road marking information as required. The conventional method of recognition method uses GPS information that can be acquired by a navigation system installed in the vehicle. However, when the road environment has changed due to construction, etc. the recognition in a real-time environment is not possible. Therefore, in this work, using an inexpensive USB camera, the environment can be captured in real time allowing prompt respond to changes in the road environment. Figure 1 shows the view in front of the vehicle as captured by a camera fixed on the dashboard inside the vehicle. From the figure, it can be seen that the road markings to be recognized are distorted in shape. Therefore, to improve the recognition accuracy, we must transform the distorted shapes to an overhead camera view. After, feature extraction, the road markings are recognized using a neural network.

II. PROPOSED METHOD

In the proposed method, the road area is extracted using the line edge information. The extracted area is then transformed to a rectangular view and extract a road marking candidate region. Thereafter, we extract the road markings, and recognize the road markings by using a neural network. The flow of the proposed method is shown in Fig. 2.

A. Road Area Extraction

To identify the road area using the line information, the edge detection on the acquired image is performed. Since we are interested in the road surface and the marking on it, the upper half of the image is not required. The edge detection is performed only the lower half of the image. Edge detection uses the Sobel operator.
Next, we extract lines using the RANSAC algorithm performed on the edge image. This process is described below.

1. Data sampling at random on the edges.
2. Select two random data points.
3. Draw a line between the selected data points.
4. Count the number of data that are within a constant width from lines drawn.
5. Repeat 1-4.

We repeat these steps a specified number of times and select as the best line i.e. the one with the most number of points.

In order to define the road area, we extract the five main lines. The road area candidate is the most inner region bounded by the lines, fig. 3.

![Fig.3 Example of road area extraction](image)

**B. Projective Transformation**

Road markings to be recognized are in a road surface region extracted. However, the image acquired from the vehicle camera on the dashboard has the form of the road markings is distorted. Therefore, transformation is performed for the road region using the projective transformation [6,7], eq. 1.

\[
\begin{bmatrix}
X_1 \\
X_2 \\
X_3 \\
X_4 \\
Y_1 \\
Y_2 \\
Y_3 \\
Y_4
\end{bmatrix} =
\begin{bmatrix}
x_1 & y_1 & 1 & 0 & 0 & -x_2 x_1 & -x_3 x_1 & -x_4 x_1 & a_1 \\
x_2 & y_2 & 1 & 0 & 0 & -x_2 x_2 & -x_3 x_2 & -x_4 x_2 & a_2 \\
x_3 & y_3 & 1 & 0 & 0 & -x_2 x_3 & -x_3 x_3 & -x_4 x_3 & a_3 \\
x_4 & y_4 & 1 & 0 & 0 & -x_2 x_4 & -x_3 x_4 & -x_4 x_4 & a_4 \\
0 & 0 & 0 & x_1 & y_1 & 1 & -y_1 x_1 & -y_1 y_1 & a_5 \\
0 & 0 & 0 & x_2 & y_2 & 1 & -y_2 x_2 & -y_2 y_2 & a_6 \\
0 & 0 & 0 & x_3 & y_3 & 1 & -y_3 x_3 & -y_3 y_3 & a_7 \\
0 & 0 & 0 & x_4 & y_4 & 1 & -y_4 x_4 & -y_4 y_4 & a_8
\end{bmatrix}
\]

\((X_{1-4} \quad Y_{1-4}) : \text{Coordinates before conversion} \quad (X_{1-4} \quad Y_{1-4}) : \text{Coordinates after conversion} \quad a_{1-8} : \text{Projective transformation parameters}\)

When performing projective transformation, it is necessary to also perform pixel interpolation. The linear interpolation method is famous, but to allow more precise pixel interpolation, we use the Bicubic method. Interpolation is performed for each pixel using the surrounding 16 pixels, Fig. 4. Approximation is performed using the eq. (2).

\[
h(t) = \begin{cases} 
(a + 2)|t|^3 - (a + 3)|t|^2 + 1 & \text{When } |t| \leq 1 \\
(8t - 5)|t|^3 + 6a|t| - 4a & \text{When } 1 < |t| \leq 2 \\
0 & \text{When } 2 < |t| 
\end{cases}
\]

\(h(t) : \text{bicubic weight}\)

Figure 5 shows the results of the projective transformation method after pixel interpolation.

**C. Region Segmentation**

On the transformed image, we have to distinguish between the road region and road markings. We perform region segmentation using pyramid images. Image pyramid is a method that changes the resolution of the images to be processed from high to low-resolution, fig. 6.

![Fig.5 Example of Projective transformation](image)

![Fig.6 Image pyramid](image)

Figure 7 shows an example of the image created by the image pyramid from the original image, and the area segmentation.
D. Extraction of marker candidates

The image is divided into regions to extract the first road area. In the method, by assuming that the road surface is the largest cluster, we remove the other clusters whose brightness value is lower than the road surface cluster.

Boundary-tracking process is performed on the remaining region to grasp the size and shape of the region. Moreover, boundary tracking can be used to remove noise or very small areas, when the standard deviation of the contents of the circumscribed square is very small.

E. Feature Extraction

To identify the road markings, the features are extracted using horizontal and vertical histograms, Fig.8. The extracted feature can handle distorted shapes. We then normalize the feature value using the number of features per image, which is 96.

A second set of features are extracted from the image using edge gradient direction. The image is subdivided into 64 blocks. The average edge direction per block is used as the feature. Therefore, an additional 64 features are generated.

III. NEURAL NETWORK

The back propagation method was used as the learning method in a hierarchical neural network. A hierarchical neural network is shown in Fig.9.

The back propagation method is described below.

1. Assign to the input layer the input signal to derive an output signal to the output layer through the intermediate layers (This is referred to as the forward operation).

2. By comparing the teacher signal and the output signal, an error is derived.

3. Back-calculates based on the derived error, then update the weights and threshold.

4. Repeat 1-3.

Finding the ideal weight threshold is the purpose of the back propagation method. Eq. 3 shows the formula used for calculating the intermediate layer nodes from the input layer to the operation forward.

$$\text{net}_j = \sum_i x_i w_{ji}$$

$$y_j = f(\text{net}_j + \theta_{hj})$$

$$f(x) = \frac{1}{1 + \exp(-x)}$$

Where: \( \text{net}_j \) is the sum of the weighted input signal. \( x_i \) is input layer, and \( w_{ji} \) are weights.

By evaluating the function \( f(x) \) on the result of \( \text{net}_j \) and threshold \( \theta_{hj} \), we can determine the output of the intermediate layer \( y_j \). \( f(x) \) is Sigmoid function. The function is used to produce an output of between 0 and 1 using a step function. The back propagation method uses the Sigmoid function to return a value between 0 and 1. Fig.10 shows the Sigmoid function.

Eq. 4 is used to calculate the output layer from the middle layer.

$$\text{net}_k = \sum_j y_j w_{kj}$$

$$\text{out}_k = f(\text{net}_k + \theta_{ok})$$

The calculation method does not change for the middle layers. In the backward direction, the output signal is calculated by comparing the neural network output to the teacher signal, expressed in what is called the square error \( E \). The expression for \( E \) is shown in eq. 5.

$$E = \frac{1}{2} \sum_k (\text{net}_k - \text{out}_k)$$

Eq. 6 shows the method of calculating the error, eq. 7 the formula for weight update eq. 8 and for updating the threshold.

$$\text{error}_{ok} = (\text{net}_k - \text{out}_k) * \text{out}_k * (1 - \text{out}_k)$$

$$\text{error}_{hj} = y_j * (1 - y_j) * \sum_k \text{error}_{ok} * w_{hjok}$$

$$w_{hjok} += \text{error}_{ok} * y_j$$

$$w_{hjij} += \text{error}_{hj} * x_i$$

$$\theta_{ok} += \text{error}_{ok}$$

$$\theta_{hj} += \text{error}_{hj}$$
Error \( \epsilon_{o_k} \) and error \( \epsilon_{h_j} \) are the output layer and the intermediate layer errors. \( t_k \) is teacher signal, which is set in advance to either 0 or 1. Weights and threshold values are updated based on these errors.

A. Learning Data

In this section, we explain the learning data used in the neural network. For road markings, to fix the form of a distorted shape, the projective transformation is used. However, some irregular road markings may still remain. Therefore, the road markings to be used as training data includes some irregular shaped examples to address this.

First, we prepare one original image of an undistorted road marking. On the image, shearing in the transverse direction is done at ±30 degrees and used as learning data, Fig. 11. We performed the same operation on all of the road markings to generate the final training data.

IV EXPERIMENTS AND RESULTS

The experiment used a USB camera and actual driving on real roads. USB camera is placed on the dashboard as shown in Fig. 12.

A notebook PC placed on the passenger sit is used to process the data in real time experiments.

Six types of white road markings (Turn right, Turn right straight, Straight, Turn left straight, Turn left, Crosswalk), and five orange road markings (30km, 40km, 50km, 60km, U-turn ban), a total of 11 different marking are used, Fig. 13.

Recognition is performed by 2 neural networks each for the white and orange signs as shown Fig.14.

![Fig.14 White and orange markings neural networks](image)

In the neural networks for the white markings, the input layer has 160 nodes, the hidden layer 15. There are 6 and 5 output nodes for the white and orange signs respectively. It was trained for 5000 times. Training data was 100 samples each, and the negative data was 2000 samples.

In the neural networks for the orange markings, the input layer has 96, the hidden layer 20, and output layer 5 nodes. It was trained for 500 times. Training data was 100 samples each, and negative data was 1000 samples.

The test data was captured in 2 real driving scenes each about 10km long.

Table 1 and 2 shows the recognition accuracy.

<table>
<thead>
<tr>
<th>TABLE 1: RECOGNITION ACCURACY: PER FRAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scene 1</td>
</tr>
<tr>
<td>Scene 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE 2: RECOGNITION ACCURACY: ROAD MARKINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scene 1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Scene 2</td>
</tr>
</tbody>
</table>

The average accuracy was about 87%. The cause being is that the recognition rate of the some road markings is very low. The reason is that the features are similar especially between the road marking and noise, Fig. 15. Looking at the features, it can be seen that the histogram in the horizontal direction is very similar. Thus, the possibility that it caused the erroneous recognition is high.

Another reason for the low recognition rate was that the recognition result of the orange markings was very low. For the 30km ~ 60km markings, all have the number “0” on the right. Therefore, since the feature amount of the right-hand side will be exactly the same, the features that are different are reduced, causing the low recognition rate.
As another cause, it also may be mentioned that the road markings are not fully visible due to wearing out as shown Fig.16. The edges and the features could not be accurately extracted. There is a possibility that this could be solved using time-series information.

Figures 17 and 18 shows example scenes captured in the experiment.

CONCLUSION

In his work, road marking recognition using Neural Network is proposed. Recognition rate of white markings was high, but the recognition rate of orange markings was not satisfactory. To improve the system, it is necessary to concentrate on recognizing the number to the left of the orange marking. In addition, there is no information of the diagonal features using the histogram. In the future, the number of features used will be increased especially the edge and diagonal information.

In case of nighttime experimentation, the image was blurred and noisy. There is need to change the settings of the camera during this time. In addition, it is very susceptible to the effects of oncoming headlights and streetlights.

In addition, placing the camera on the dashboard produced blurring or distortion in many images. This must be addressed in the future.

REFERENCES