Research and Design on Vehicle Pedestrian Detection in Natural Scene

Pei-ran ZHAO¹, Xue-wu ZHANG¹,*, Yu-bo XIE¹, Ling-li XU¹, Yan XIANG² and Jin-bao SHENG²

¹College of Internet of Things Engineering, Hohai University Changzhou, China
²Nanjing Hydraulic Research Institute, China
*Corresponding author

Keywords: Intelligent driving, Real-Time pedestrian detection, Feature extraction, Integral channel.

Abstract. We propose a two-layer pedestrian feature extraction algorithm based on multi features fusion in integral channel, which is gained by intelligent driving system in natural environment. The algorithm extracts the gradient direction, gradient magnitude and LUV color channel of the image as upper layer feature, by using the fast operation speed of integral channel. Edgelet features that describe pedestrian local information are used as lower level features for validation. While ensuring real-time detection, we improved the low robustness of single layer features. Experimental results show the algorithm can reduce the false detection rate to a great extent under the premise of ensuring the detection speed in dealing with pedestrians in the background of natural and complicated vehicles.

Introduction

Pedestrian detection system as a part of an intelligent vehicle-assisted driving system, an important technology to reduce road traffic accidents and traffic pressure caused by pedestrian collision events have potential application value and significance [1]. However, the road image information obtained by the vehicle embedded device is affected by factors such as changes in ambient light, pedestrian non-rigid body characteristics, pedestrian size in the vehicle image, and mutual interference between people, vehicles and roads, so that there are many problems in the promotion of pedestrian detection system [2].

In recent years, researchers have proposed a lot of solutions based on machine learning, according to the research of feature extraction in pedestrian detection. Among them, the most representative is that Dalal proposed HOG characteristics to detect pedestrians in 2005, but because of its low accuracy, it is difficult to apply in vehicle detection system; After that, Dollar and other researchers proposed the integral channel feature (ICF) [5] is used to describe pedestrians, and good results are obtained; Benenson and other people improved the method and increased its detection speed to 100fps [6]; Recently, Dollar improved ICF by putting forward the concept of aggregation channel feature (ACF) [7]. The new Edgelet feature [9] is introduced to describe the pedestrian, it shows great performance in some public datasets.

Double Layer Feature Extraction Algorithm

Integral Channel Features

Integral channel features (ICF) was proposed by Dollar et al, which itself is not an independent new features, but the image of nonlinear or linear transformation. They naturally integrate heterogeneous histogram of oriented gradient (HOG) [4], have few parameters and are insensitive to exact parameter. The channel types include color space, such as gray, GRB and Luv, linear filters, nonlinear transformations, pointwise transformations, integral Histogram, gradient histogram and so on.
**Edgelet Feature**

Edgelet features are mainly described local features of pedestrians, its main idea is to divide the human body into four core parts (head, arms, body, legs) for feature extraction, the various parts of the corresponding input as a classifier, a multi classifier to detect image window [3].

The Edgelet feature extracts the local contour information of pedestrians and describes the contour features of each part of the pedestrian. Then, the extracted eigenvalues match the constructed pedestrian feature sets to obtain the response value. The larger the response value, the higher the similarity between the detection window and the pedestrian:

\[
S(x, y) = \frac{1}{s} \sum_{i=1}^{k} I_i(x+u_i, y+v_i) < N_i(x+u_i, y+v_i), n_i >
\]

\(S(x, y)\) is the value of response, \(i\) is The number of interior points in a neighborhood, \(I_i(x+u_i, y+v_i)\) is the edge strength for corresponding point \((u_i, v_i)\), \(N_i(x+u_i, y+v_i)\) is the gradient normal vector of point \((x+u_i, y+v_i)\).

**Double Layer Feature Extraction Algorithm**

Using the method of multi feature fusion integration channel, in the case of complete pedestrian contour, the detection rate can reach 89%, but in practical application, where pedestrian will be blocked, and the overlap between multiple pedestrian phenomena will affect the detection. When the method is used, the detection rate is reduced to 67.6%, which shows that the detection performance of this method is poor in complex environment. Edgelet features perform well in a variety of situations, but due to its large computational complexity, it is difficult to implement real-time monitoring in embedded devices.

![Figure 1. Double layers feature extraction processing.](image)

To further improve the robustness of the detection, the algorithm can be improved on the basis of the above detection methods, and we propose the algorithm of feature extraction based on double difference in integral channel. That is to say, the upper layer takes the advantage of fast calculation in the integral channel, and extracts the features of the image in the LUV, the gradient magnitude and the gradient direction channel. Through fast sliding windows in the image, several candidate regions can be obtained, and preliminary positioning is achieved. Edgelet feature is introduced into the lower layer to detect the human local features of the candidate pedestrian target window, which has better detection performance than the global feature when dealing with pedestrian overlap or occlusion. By combining these two methods with their respective advantages, forming a detection system of two graded independently, and the introduction of Bayesian decision-making algorithm, through the confidence to carry out the discrimination of the candidate target computing candidate window, implementation of pedestrian detection in complex environment, this algorithm can greatly
reduce the detection error rate and robustness of the algorithm. The double layers feature extraction process is shown in figure 1.

Considering the problem of slow speed of Edgelet feature extraction, double feature fusion algorithm is proposed in this paper will be arranged in the ICF Edgelet feature detection, which can reduce the detection area of Edgelet features a large area, to achieve fast positioning requirements, reduce the detection time of second layers.

The Bayesian decision strategy includes the following two steps: first the pedestrian is divided into four parts (head, arms, body and legs), using physiological and structural characteristics of pedestrian detection, so that the head of the window above, the body, the arm detector middle part of the leg detection window under part four parts the detection results, which can improve the detection speed; then calculate the likelihood probability judgment and confidence to the detection results, the window of the confidence formula is as follows:

$$l(x) = \log \frac{P\{y = 1|x\}}{P\{y = 0|x\}}$$  \hspace{1cm} (2)

$$P\{y = 1|x\} \quad \text{and} \quad P\{y = 0|x\} \quad \text{represent the probability of pedestrians and the probability of non-pedestrians.}$$

The final test output formula is as follows:

$$H(x) = \begin{cases} 
1 & l(x) > T \quad \text{pedestrians} \\
0 & \text{otherwise} \quad \text{non-pedestrians} 
\end{cases}$$  \hspace{1cm} (3)

Among them, $T$ represents the threshold value of pedestrian judgment, the normal range of values is $[0.6, 0.8]$, in this paper after a large number of experimental effect analysis, select the threshold $T = 0.75$ to determine the pedestrian.

**Experiments**

Experiments are conducted on two public datasets: The Caltech and the INTIA pedestrian dataset, and one self-built datasets, which data is collected from a tachograph by recording the daily driving record photos and reorganized. These two public datasets are the most two popular datasets for pedestrian detection. Data shows that by adding additional data, the detection rate can be improved [10]. We train the detectors on all the datasets above, and test in our self-built dataset.

**Pedestrian Detection Analysis under Different Scenes**

We first divide the self-built pedestrian image database according to the background of the complexity of classification into two parts: the simple background (small number of pedestrians) and the complex background (pedestrian intensive, bad weather scene), and then chooses the representative image as the test set. Experiments are done by using HOG+SVM, integral channel features, Edgelet features and our algorithm. The experimental results under different scenarios are shown in Table 1 and Table 2:

<table>
<thead>
<tr>
<th>Features</th>
<th>Miss Rate(%)</th>
<th>Process Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOG</td>
<td>33.7</td>
<td>0.532</td>
</tr>
<tr>
<td>Edgelet</td>
<td>16.3</td>
<td>2.360</td>
</tr>
<tr>
<td>ICF</td>
<td>24.3</td>
<td>0.039</td>
</tr>
<tr>
<td>Ours</td>
<td>14.7</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Table 1. Performance of different algorithms in simple context.

<table>
<thead>
<tr>
<th>Features</th>
<th>Miss Rate(%)</th>
<th>Process Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOG</td>
<td>46.3</td>
<td>0.547</td>
</tr>
<tr>
<td>Edgelet</td>
<td>23.7</td>
<td>2.532</td>
</tr>
<tr>
<td>ICF</td>
<td>32.5</td>
<td>0.041</td>
</tr>
<tr>
<td>Ours</td>
<td>22.8</td>
<td>0.049</td>
</tr>
</tbody>
</table>

Table 2. Performance of different algorithms in complex context.
From Table 1 and Table 2, the use of HOG single features, the detection rate is relatively low, and the missing rate compared to the other algorithms by nearly 20%; and although Edgelet features can ensure the missing rate, but due to its complicated algorithm causes the detection speed is only 0.4 frames per second, this is obviously unable to meet the requirements of real-time detection; ICF algorithm has obvious advantages in the detection speed and detection rate in the simple background is relatively good, but in the complicated situation of pedestrian occlusion and a large number of passengers, the detection rate fell by nearly 10%.

The double features of the proposed fusion algorithm, the detection rate and error rate, whether in simple background or complex background, compared to the existing feature extraction algorithm, are improved, and with the help of the integral image technique and AdaBoost cascade classifier detection, classification speed advantage, will the detection rate of up to 20 frames per second, far faster than the traditional feature extraction algorithm, which can meet the requirement of real-time detection.

Figure 2 is the three different algorithms to test the effect of some samples of the graph, it shows that the algorithm in dealing with multiple pedestrians walking abreast of the situation, can be very good for pedestrians to distinguish; rainy day umbrella, raincoat, by increasing the sample type, the detection rate can be to achieve a higher; on the pedestrian lighting the night, can achieve the most pedestrian detection scenarios, the whole system has good adaptability.

![Figure 2. The performance of three different algorithms in the same background.](image)

**Summary**

This paper makes improvement based on integral image technology, presents a double layer pedestrian detection algorithm based on multi feature fusion is a kind of integral channel. The algorithm makes use of the advantage of rapid speed of integral channel, extract the image of the LUV channel, channel gradient amplitude and gradient direction of channel characteristics, to achieve fast positioning of pedestrians, and then for the candidate pedestrian area, is selected to describe the local characteristics of the human body Edgelet feature detection, finally the detection results were finally diagnosed by using Bayesian decision the idea of improving the performance of the algorithm in error detection. The experiment shows that the proposed feature extraction of pedestrian algorithm in the case of occluded pedestrian processing complex background and when the rate of performance in the detection rate and false detection has greatly improved, with the help of integral calculation. At the same time channel and AdaBoost cascade classifier [8], which can meet the real-time requirements in measuring speed.
Acknowledgements
This paper is supported by: the Fundamental Research Funds for the Central Universities (2015B25214, Y716019, Y716015, Y716011, 51679151; National Natural Science Foundation of China (61671202, 6157312, 61273170); National Key Research Program of China (2016YFC0401606, 2016YFC0401605).

Reference


