Landmark Detection and Recognition based on Adaboost and SVM

Yucui Ju¹ ², Hua Zhang¹ ² and Yanbing Xue¹ ², Jin Wang³

¹ Key Laboratory of Computer Vision and System, Tianjin University of Technology, Tianjin 300191, China
² Tianjin Key Laboratory of Intelligence Computing and Novel Software Technology, Tianjin University of Technology, Tianjin 300191, China
³ Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information Science & Technology, Nanjing, 210044, China

Abstract. This paper proposes a robust real-time artificial landmarks detection and recognition system for indoor mobile robot. First, histograms of oriented gradient (HOG) features are extracted to resolve the illumination changes in indoor environment. Second, AdaBoost based algorithm is used in detection phase to increase the processing speed. Finally, RBF-SVM classifier is used for recognition. Experimental results show a high detection and recognition accuracy of the proposed system.

Keywords: landmarks detection and recognition; AdaBoost; HOG; SVM

1 Introduction

Landmark detection and recognition is an important task in mobile robot environment perception [1] which is also the key step to ensure the robot can be positioned correctly.

In detection stage, color threshold segmentation and Hough transform [2-3] are usually used to detect the landmarks. However, the method of using color threshold segmentation will be influenced by the illumination changes and using the shape analysis (Hough transform) will bring an error positioning due to the similarity. In recognition stage, the system evaluates regions found in the detection stage and identifies the landmarks. Template matching is widely used for its simplicity [4], which has the disadvantage of low accuracy. These algorithms can obtain better results in a certain extent, but have high false alarm and slow detection speed.

This paper presents a fusion of AdaBoost [5] and SVM [6] algorithm for landmark recognition. First of all, filtering out the most likely candidate collection of images through AdaBoost, and then, using RBF-SVM to recognize on the collection of candidates. Both of the detection system and the recognition system use HOG features [7] which can describe contour feature of landmarks and is not sensitive to illumination. The proposed method greatly increases the recognition speed, and ensures the accuracy of recognition to some extent.
2 Feature Extraction

In order to make the dimension of feature vector suitable for machine learning, all the landmarks are gray images and normalized to the size of 32 by 32 pixels. The extraction steps of landmark HOG feature are as follows:

1) Gradient Computation. The horizontal and vertical gradient obtained by convolving the gradient operator [-1, 0, 1] and [1, 0, -1] with the landmark image. The gradient direction and gradient magnitude of each pixel is calculated.

2) The division of sub-blocks and histogram acquisition. The landmark is divided into nine image blocks, each block is divided into four cells and the size of each cell is 8 by 8 pixels. The gradient direction of each pixel is 0° ~ 180° and if the gradient magnitude of a pixel is greater than 180°, its final direction value is 180°. Then the final gradient orientation is divided into nine bins and gets the corresponding 9-dimensional feature vector of each cell.

3) Normalization. In order to make the feature vector space having robustness to local illumination changes, shadows, and the edge changes, we will need to normalize the gradient intensity. In this paper, L1-norm is used for normalizing as followed:

$$v \leftarrow v / (\|v\| + \epsilon)$$  \hspace{1cm} (1)

4) Generation of HOG feature vectors. HOG Feature extraction process as shown in Figure1, high dimensional vectors of 9× 9 × 4 = 324 data are generated by above steps. The first “9” means that there are nine bins in each cell, the second “9” means that there are nine blocks, “4” means that there are four cells in each block and then the whole HOG feature vectors are generated.

![Fig. 1. The extraction process of HOG features](image)

3 Landmark Detection Based On AdaBoost

Inspired by HAAR feature, we exploit a fast way of calculating the HOG feature. First, we discretize each pixel’s orientation (including its magnitude) into 9 histogram bins. We compute and store an integral image for each bin of the HOG and use them to compute efficiently the HOG for any rectangular image region. This requires 4 × 9 image access operations.

The weak classifier is the focus of the AdaBoost algorithm. Combination of weak classifiers which are selected according to certain rules, you can construct different strong classifiers and cascade classifier. Not all the features can adapt to the AdaBoost algorithm, so the following describes the structure of weak classifier of HOG-AdaBoost proposed by this paper.
HOG feature is a 36D feature vector and we cannot consider a HOG feature as a weak classifier. The HOG feature of each cell contains important information on how to separate landmarks from other objects, therefore, the set of weak classifiers are created for each cell in this paper. As the HOG is a histogram with bins indicating local gradient distribution, we set a threshold $T$ and then compare the value of one bin with $T$. If the value is larger, we consider it as landmarks otherwise not. The histogram has nine bins in this paper, and then we have nine weak classifiers corresponding to each bin. The weak classifier is defined as follows:

$$h_{k,i}(x) = \begin{cases} 
1 & \text{if } p_i \text{Hist}_{k,i}(x) < p_i \theta_i \\
0 & \text{otherwise}
\end{cases}$$

(2)

In the above equation, $x$ indicates the input image. For the cell $k$, the feature $p_i \text{Hist}_{k,i}(x)$ presents the value on $i$th bin of the histogram in that cell. $\theta_i$ is the decision threshold corresponds to the $i$th feature. $p_i = \pm 1$ determines the direction of the inequality.

## 4 The Landmark Recognition Based On RBFSVM

The one-against-one classification method is used for multi-class SVM classifier of landmark recognition system. First, scaling the initial feature vector to the range [0, 1]. Second, training the best classification parameters. The paper uses the RBF kernel to build classifier. There are two parameters while using RBF kernels: $C$ and $r$, different $C$ and $r$ will produce different accuracy. The paper uses “web searching” to find the optimal $C$ and $r$. 75% of the total samples are used as the training samples for cross validation which are divided into 5 groups. Each group takes turns as test samples, the rest as the training samples.

![Fig. 2. The optimization of landmark classifier](image)

Figure 2 shows the parameter values obtained from the optimization and the classification accuracy of training set. The best parameters can be seen from the
figure: $C=8$, $\Gamma=0.0078125$, the correct rate is 100%. After getting the best parameters, we can obtain the optimal classifier model by training on the training samples. Finally, the optimum classification model is used to recognize the remaining 25% test samples.

5 Experimental Results and Analysis

In the detection phase, the landmark samples are acquired by a camera mounted on a vehicle at the height of 0.50 m (above the ground). We extract the HOG features of the prepared samples and use the AdaBoost and SVM algorithm to train respectively. In the recognition phrase, the RBFSVM, which are obtained in the training process, are used to provide the most effective representation of the landmarks. Finally, we test the performance of our overall system.

Experimental Results of Landmark Detection

Samples are divided into positives and negatives. 1) positive samples: as shown in Figure 3(a), we performed the cut operation on preserved landmark pictures by taking four points roughly and making the bounding rectangle of a circle landmark through the fine adjustment. Finally, the pictures which have been taken out of images are normalized to the size of 32 by 32 pixels. 2) negative samples: Negative samples consist of two parts, one part is the arbitrary regions excluding landmarks which are cropped from the captured photos, and another part is some background images which are downloaded from the Internet. We try to insure that the negative samples have diversity and negative samples are normalized to 64 by 128 pixels. We make 1762 positive samples and 5000 negative samples. Figure 3(b) shows the part of the detection results.

![Fig. 3. The cropping ways of positive samples and the part of detection results](image)

Experimental Results Based On SVM Landmark Recognition

We divide the samples into five groups in the experiment. As shown in Figure 4, the five groups of samples are normalized to the size of 32 by 32 pixels and the HOG features are extracted from the five groups of samples which are used for training SVM. We received a total number of 1794 samples, in which the 1446 samples are used for training samples and the rest of 648 samples are used for test samples.
The Performance of Our Overall System

The performance of overall system is shown in table 1. The “false recognition” in the overall system is the sum of the number of the landmarks which are not detected in the detection phrase and the number of the landmarks which are detected in the detection phrase but not recognize in the recognition phrase.

Table 1. The performance of overall system

<table>
<thead>
<tr>
<th>type</th>
<th>detection</th>
<th>Recognition</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>True positive</td>
<td>False positive</td>
<td>Correct recognition</td>
<td>False recognition</td>
</tr>
<tr>
<td>forward</td>
<td>99.2%</td>
<td>0.8%</td>
<td>95.1%</td>
<td>4.9%</td>
</tr>
<tr>
<td>left</td>
<td>98.4%</td>
<td>2.4%</td>
<td>96.0%</td>
<td>4.0%</td>
</tr>
<tr>
<td>right</td>
<td>99.2%</td>
<td>2.3%</td>
<td>97.0%</td>
<td>3.0%</td>
</tr>
<tr>
<td>rotation</td>
<td>96.9%</td>
<td>4.8%</td>
<td>95.3%</td>
<td>4.7%</td>
</tr>
<tr>
<td>stop</td>
<td>99.3%</td>
<td>2.2%</td>
<td>96.4%</td>
<td>3.6%</td>
</tr>
<tr>
<td>forward</td>
<td>99.2%</td>
<td>0.8%</td>
<td>95.1%</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

Table 1 show that the AdaBoost classifier has higher true positives and low false positives in the detection phrase, and the overall system has higher correct recognitions and lower false recognitions. The average recognition time of the AdaBoost and RBF SVM is 99.84 ms.

6 Conclusions

The paper presents a real-time artificial landmark detection and recognition system for mobile robot. The detection method is based on HOG feature and...
AdaBoost learning algorithm. In the recognition phase, we use HOG feature combined with RBF SVM classifier. An experimental comparative study on the two algorithm verified that the proposed algorithm is not only able to meet real-time requirements, but also to obtain a very high recognition rate by applying the whole system to the mobile robot.

Acknowledgements

This research has been supported by Key project in Science and technology pillar program of Tianjin (10ZCKFGX00400), and Tianjin Education Committee science and technology development Foundation (No. 20120802).

References