MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

Improved Wavelet Feature Extraction Methods Based on HSV Space for Vehicle Detection

Xuezhi Wen, Wei Liu, Nan Wang, Huai Yuan, Hong Zhao
Software Center, Northeastern University
AAC, A2 Building, Neusoft Park, Hun Nan Industrial Area, Shen yang 110179, China
(wenxz, lwei, wangnan, yuanh, zhaoh)@neusoft.com

Abstract

The focus of this work is on the problem of feature extraction for vehicle detection. Feature extraction is a key point of pattern recognition. In particular, we propose using improved wavelet feature extraction approaches based on HSV space for rear-vehicle detection. Wavelet features are attractive for vehicle detection because they form a compact representation, encode edges, capture information from multi-resolution, and can be computed efficiently. Currently, the wavelet features based on coefficients and grayscale space are easily affected by the varied surroundings and illumination conditions and cause high intra-class variability. In order to deal with this problem, three improved wavelet feature extraction approaches based on HSV space are proposed. The experimental results indicate that the improved approaches based on HSV show super performance compared with the current methods based on both HSV space and Grayscale space. Furthermore, they also show better results than themselves based on Grayscale space.

1 Introduction

Vehicle detection is the basis for driver assistance system (DAS), and it is a challenging task because it is not only affected by the shape, size, color and pose of vehicles, but affected by illumination, weather, surroundings and the surface of different roads. So it needs good adaptive pattern recognition algorithms which lead to higher request for feature extraction. Matthews et al. [1] gave an introduction on how to use Principal Components Analysis (PCA) to extract features of Region of Interest (ROI) for vehicle detection. It scaled the vehicle training images and non-vehicle training images to 20×20 grayscale images, each 20×20 image was then divided into 25 4×4 subwindows, each 4×4 subwindow was extracted by PCA. The PCA features were then fed to NN for classification. Goerick et al. [2] proposed using Local Orientation Code (LOC) for extracting edge information of ROI. The histogram of LOC was then fed to NN for classification. Features extraction using Gabor filters was investigated in [3] [4], Gabor filters provide a mechanism for obtaining edge and line information of ROI by tuning orientation and scale. An over-completed dictionary of Haar wavelet features was utilized in [5] for vehicle detection. They argued that the over-completed representation provided a richer model and spatial resolution and was more suitable for capturing complex patterns. Sun et al. [6] [7] performed the comparison of three methods of wavelet feature extraction which are all based on coefficients and grayscale space.

Wavelet features based on coefficients and grayscale space not only indicate the difference in average intensity between local regions along different orientations, but include the information on the sign of the intensity gradient [6] [7] [9]. This feature pattern means that two attributes which are belonging to the same class with same absolute values but different signs denote two different features. So this feature pattern will lead to high intra-class variability and increase the complexity of classification. Furthermore, Wavelet features based on grayscale space are sensitive to illumination changes and noise in a sense.

The rest of the paper is organized as follows: In Section 2, we provide the description of the improved methods of wavelet feature extraction in detail. Experimental results are listed in Section 3. Finally, Section 4 contains our conclusions and directions for future research.

2 Improved wavelet feature extraction

In this section, we first introduce the wavelet pyramid decomposition process on 2-Dimensional signals, and then introduce the improved approaches based on wavelet coefficient magnitudes and HSV space in detail.

2.1 Wavelet pyramid decomposition

Wavelets have broken through the constraint of traditional analysis methods of signal – Fourier transform. They can analyze the different regions of a signal in different resolutions and have been applied successfully in many fields including objects detection [5], face recognition [8] [9] and image retrieval [10].

The basic steps of 2-Dimensional discrete wavelet pyramid decomposition are as follows:

First, each row of N×N image is filtered by lowpass filter and highpass filter respectively and then half down sampled. The half-down-sample removes one column every other column of N×N image to obtain an N × (N/2) image.

Second, each column of N × (N/2) filtered image is filtered by low-pass filter and high-pass filter respectively and then half down sampled. Down sampling removes one row every other row after two filters. Figure1 shows the process of decomposition, each branch denotes an (N/2) × (N/2) image and makes different contribution during analysis process for N×N image. The lowpass filter for rows of image makes every row blurry. Then the lowpass

filter for columns of image produces the low pass approximation of the entire image. The fourth branch denotes the approximation of the entire image. If columns are filtered by highpass filter after rows are filtered by lowpass filter, the changes between rows are captured, and that is, horizontal detail is disclosed. The third branch denotes the horizontal detail. The top branch denotes filtering the image firstly with the highpass filter. Rows filtered firstly by highpass filter means highlighting change between two neighbor pixels of given rows, then columns filtered by lowpass filter blurs the changes between rows, which produces vertical detail. The second branch shows the vertical detail. When columns are filtered by highpass filter after rows are filtered by highpass filter, the filtered results represent neither the horizontal detail nor the vertical detail. In fact, they produce diagonal detail of the original images, as is shown in the first branch. If we process the same steps on the (N/2×N/2) lowpass detail as on the original N×N image, we can get four （N/4）×（N/4） subimages: lowpass detail (LL), horizontal detail (HL), vertical detail (LH) and diagonal detail (HH). The analysis may go on until the subimage just has one pixel.

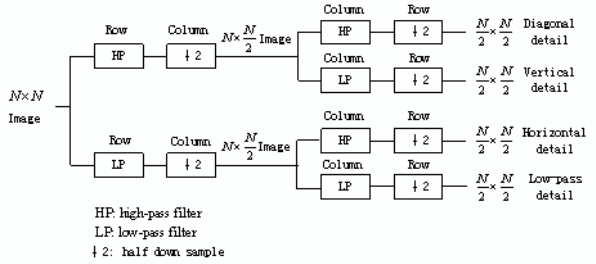

Figure 1. Wavelet pyramid decomposition process.

Haar wavelet is the simplest to implement and demands the least computational load. Moreover, since Haar basis forms an orthogonal basis, the transform provides a non-redundant representation of an input image. So we adopt Haar wavelet for pyramid decomposition.

2.2 HSV color model

Color is one of the most important information for vision. Human vision can be described in various color model forms, e.g. RGB color model, YIQ color model and HSV color model. The HSV color model is closer to the description of human sight sensation.

HSV stands for Hue, Saturation, and Value. Formula (1) gives the expression of RGB color model transformed to HSV color model.

$$\begin{cases} S = \dfrac{MAX - MIN}{MAX} \\ H = \begin{cases} 60 * \dfrac{G - B}{MAX - MIN} & R = MAX \\ 120 + 60 * \dfrac{B - R}{MAX - MIN} & G = MAX \\ 240 + 60 * \dfrac{R - G}{MAX - MIN} & B = MAX \end{cases} \\ V = MAX \end{cases} \quad (1)$$

2.3 Improved methods

For an input RGB image, firstly, it is scaled to a 32×32 RGB image. The scaled RGB image is then transformed into HSV image according to above (1). Finally 5 levels wavelet decomposition is performed based on the "V (value)"-channel of HSV space, and then the absolute values of each of the coefficients are obtained. For the 5 levels wavelet coefficient magnitudes, we propose three improved approaches to produce the wavelet feature vector.

Removing HH Sub-band of The First Level: Firstly, keep all the coefficient magnitudes except the ones in the HH sub-band of the first level, because the removed ones are likely to encode noise. Then, normalize the data to [0, 1]. Finally form the normalized data to a feature vector of size 768.

Hard-threshold Filtering: Firstly, normalize all coefficient magnitudes to [0, 1], Then, set the coefficient magnitudes which are smaller than a threshold value to zeros while keep the others invariable. Finally form the processed data to a feature vector of size 1024. Because small data indicate mostly noise or fine details that are not essential for vehicle detection.

Level Filtering: Firstly, normalize all coefficient magnitudes to [0, 1]. Then filter the normalized data of different levels by using different thresholds according to above Hard-threshold Filtering. Finally form the processed data to a feature vector of size 1024. Because the coefficient magnitudes from different levels represent different resolution feature information, and the coefficient magnitudes in the same level have better contrast than that of all levels.

Above three improved wavelet feature extraction methods, the main advantage of normalizing the coefficient magnitudes to [0, 1] is to avoid the problem that attributes in greater numeric ranges dominate those in smaller numeric ranges. Another advantage is to avoid numerical difficulties during the classification calculation.

3 Experimental results

The proposed features extraction approaches were used to detect vehicles in our rear-vehicle detection system based on monocular vision. The system consists of two stages: ROIs generation and vehicles validation. At first, used the shadow underneath a vehicle and aspect ratio to segment ROIs, and then verified the vehicles' existence in ROIs. In nature, the validation of vehicle's existence is a two pattern recognition problem: vehicle vs. non-vehicle. In the application, the improved approaches were adopted to extract features of the selected vehicles and non-vehicles. Then the obtained features set were used to training SVM classifier. Finally, the SVM classifier was used to verify the vehicles' existence in ROIs segmented from testing frames.

Different real-world scene videos taken from a camera mounted on a moving vehicle were used for evaluating the presented algorithm, and they were taken on different daytime scenes, including highway, urban common road, sparse road, etc. Sometimes roads are covered with japanning, smear, snow, etc. 7,446 training frames and 3,500 testing frames are extracted respectively from different videos. We selected a total of 3,058 samples from the segmented ROIs for training which include 1,171 vehicle

samples (positive samples) and 1,887 non-vehicle samples (negative samples) , the testing frames contain 6,457 ROIs which include 4,157 vehicles and 2,300 non-vehicles. Figure 2 shows the vehicle and non-vehicle training examples. The vehicles in the training and testing frames includes various kinds of vehicles such as cars, 1boxs, trucks and buses as well as different colors such as red, blue, black, gray, etc. Furthermore, it not only includes vehicles near the ego vehicle but the ones far from the ego vehicle. The non-vehicles in the training and testing images includes roads, buildings, green plant, advertisements boards, bridges, traffic signs, guardrails, and so on.

SVM classifier with RBF kernel was performed to evaluate the proposed approaches and current methods in the rear-vehicle detection system. The optimal parameters were selected via 5-fold cross-validation.

Figure 2. Examples of vehicle and non-vehicle training samples.

To evaluate the performance of the approaches, the true positive rate (or vehicle detection) t_p and false positive rate f_p were recorded. They are defined as follows.

$$t_p = \frac{N_{TP}}{N_{TP} + N_{FN}}, f_p = \frac{N_{FP}}{N_{FP} + N_{TN}} \quad (2)$$

Where N_{TP} , N_{FP} , N_{TN} and N_{FN} are the number of objects identified as true positives, false positives, true negatives and false negatives respectively. The following six methods were evaluated based on the same training and testing set.

1. Keep all signed coefficients except the ones in the HH sub-band of the first level [9].
2. Select 100 largest signed coefficients [6] [7].
3. Keep 100 largest signed coefficients and quantize them to -1, 0, and 1[6] [7].
4. Proposed removing HH Sub-band of the first level based on unsigned coefficients.
5. Proposed hard-threshold filtering based on unsigned coefficients.
6. Proposed level filtering based on unsigned coefficients.

Here we perform two group experiments based on HSV space and Grayscale space respectively.

The vehicle detection results of the first group experiments are shown in Table 1, and the results of the second group experiments are shown in Table 2. The ROC curves of the first and the second group experiments are shown in Figure 3 and Figure 4 respectively.

From Table 1 and Table 2, especially from figure 3 and figure 4, we can find that the proposed feature extraction methods based on HSV space have super performance compared with the current methods based on both HSV space and grayscale space. Furthermore, they also show better results than themselves based on grayscale space. On the other hand, from figure 4, we can find that the proposed approach 1) has the best performance among the three proposed approaches, because it has more classification feature information than the others.

4 Conclusion and future work

We have proposed three improved approaches of Haar wavelet feature extraction based on wavelet coefficient magnitudes and HSV space for rear-vehicle detection. Our experimental results and comparisons have shown that the proposed approach 1) based on HSV space has better performance than the others based on both HSV space and grayscale space. Moreover, it can also be adapted to front-vehicle detection. However, the best feature extraction approach produces 768-dimensional feature vector, for future work, we will investigate the methods to reduce the dimension of feature vector in order to reduce the computational load and enhance the generality ability for classification.

Acknowledgments

The authors would like to thank their colleagues for their help with the data collection.

Table 1. The vehicle detection results based on grayscale space.

Metrics / Methods	t_p	f_p
1	91.80%	17.35%
2	96.78%	22.39%
3	79.07%	4.04%
4	97.14%	7.04%
5	96.61%	6.39%
6	97.04%	6.70%

Table 2. The vehicle detection results based on HSV space.

Metrics / Methods	t_p	f_p
1	93.82%	16.30%
2	96.44%	19.04%
3	79.46%	5.83%
4	96.61%	4.74%
5	96.80%	6.35%
6	96.42%	5.00%

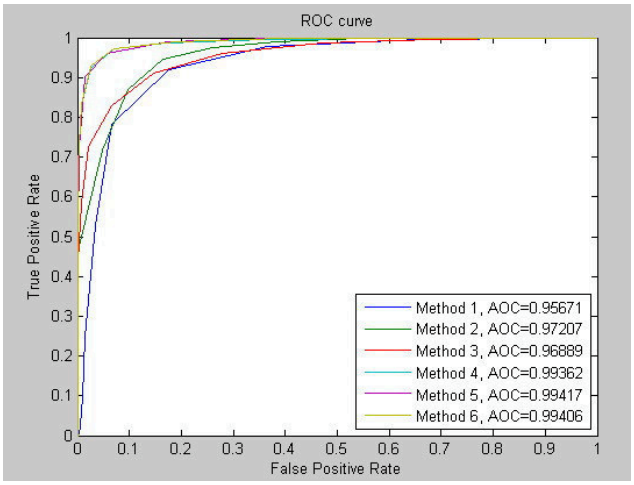

Figure 3. ROC curve of the six methods based on gray-scale space.

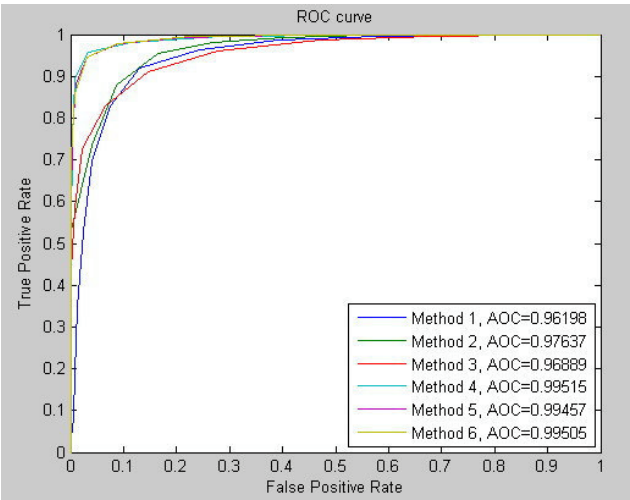

Figure 4. ROC curve of the six methods based on HSV space.

References

[1] N.D.Matthews, P.E.An, D.Charnley, and C.J.Harris: "Vehicle Detection and Recognition in Greyscale Imagery," Control Eng.Practice, vol.4, no.4, pp.473-479, 1996.

[2] C. Goerick, N. Detlev and M. Werner: "Artificial Neural Networks in Real-time Car Detection and Tracking Application," Pattern Recognition Letters, vol. 17, pp.335-343, 1996.

[3] Thiang, R. Lim, and A. T. Guntoro: "Car Recognition Using Gabor Filter Feature Extraction," Circuits and Systems, APCCAS'02. (2), pp.451-455, 2002.

[4] Zehang Sun, George Bebis and Ronald Miller: "On-road Vehicle Detection Using Gabor Filters and Support Vector Machines," IEEE 14th International Conference on Digital Signal Processing, pp.1019-1022, 2002.

[5] C.Papageorgiou and T.poggio: "A Trainable System for Object Detection," International Journal of Computer Vision, vol.4, no.4, pp.15-33, 2000.

[6] Z. Sun, G. Bebis and R. Miller: "Quantized Wavelet Features and Support Vector Machines for On-road Vehicle Detection," 7th International Conference on Control, Automation, Robotics and Vision, vol.3, pp.1641 – 1646, 2002.

[7] Z. Sun, G. Bebis and R. Miller: "Monocular Precrash Vehicle Detection: Features and Classifiers," IEEE Transactions on Image Processing, vol.15, no.7, pp.2019-2034, 2006.

[8] G.Garcia, G. Zikos, and G. Tziritas: "Wavelet Packet Analysis for Face Recognition," Image and Vision Computing. 18, pp.289-297, 2000.

[9] H. Schneiderman: "A Statistical Approach to 3D Object Detection Applied to Faces and Cars," CMU-RI-TR-00-06, 2000.

[10] C. Jacobs, A. Finkelstein and D. Salesin: "Fast Mutiresolution Image Querying," Proceedings of SIG-GRAPH, pp. 277-286, 1995.