

Vehicle Classification using SIFT

Narhe M. C.

Dept.of Electronics and Telecommunication
MIT College of Engineering, University of Pune
Pune, India.

Prof. Dr. Nagmode M. S.

Dept. of Electronics and Telecommunication
MIT College of Engineering, University of Pune
Pune, India.

Abstract - Vehicle classification is challenging task due to problems like motion blurs, varying image resolution etc. So far numerous algorithms have been developed for vehicle classification and counting. Vehicle classification technique find major application at Toll plaza, Traffic signal etc. This paper proposes an effective Scale Invariant Feature transform (SIFT) algorithm used for moving vehicle classification. This will help to improve efficiency and reliability of vehicle classification technique.

Index Terms - Vehicle classification, SIFT.

I. INTRODUCTION

As one of the important research topic in computer vision based intelligent transportation system (ITS) is vehicle classification. Increase in number of vehicle lead to major problems such as vehicle congestion, accidents vehicle robberies etc. So to deal with such problems intelligent transportation system gives attention on vehicle classification. Many techniques have been proposed using different approach and sensors. These techniques can be categorized by the types of sensors used in the classification. Two common techniques of vehicle classification are as given below:

1. Hardware based vehicle classification
2. Software based vehicle classification

The hardware based vehicle classification technique includes various sensors such as infrared sensors, radar, magnetic sensors etc. The traditional hardware based vehicle classification technique are simple and reliable, but with some shortcomings. The drawbacks of most of the traditional hardware based vehicle classification technique are that they are intrusive and have to be cut into the pavement. This means that the roadways in which they are installed must undergo construction. This also implies that there will be lane closures and added vehicle delay, which is costly for everyone. Another drawback is that they have a fixed location; once put in place they cannot be adjusted to accommodate changing traffic conditions and more accurately measure traffic flow. Additionally, installation of hardware based vehicle classification technique may decrease pavement life. While software based vehicle classification technique uses vision sensors. In this paper, we use Scale Invariant Feature transform for the vehicle classification.

In vehicle classification technique the performance of relies very much on the algorithm instead of the sensor. Vehicle classification technique can either be image-based [1] or video based. Video-based classification [2] has

advantage over image-based because moving object can be separated from the static background reasonably well by background modelling and subtraction. The background subtraction is followed by the segmentation whereby the vehicle is segmented from the background for further processing. In many conventional vehicle classification systems, the features of the vehicle used for the classification task are the height, length and width of the vehicle. More sophisticated algorithm [3] adopted the keypoints detection on the vehicles as the features for the classification task. Keypoint as the features for classification has advantage over the conventional dimensional profile due to the high reliability of the scale invariant keypoint detection algorithms.

The Scale Invariant Feature Transform (SIFT) algorithm introduced by David Lowe in 1999. This algorithm is a widely used for keypoint detection. The method is notable for the reason that the features used are invariant to image scaling, translation, rotation, affine or 3D projection and partially invariant to illumination changes. In the keypoint detection algorithm is adopted and modified to fit in vehicle classification task and return a good result. As per Jian Wu et al [4], SIFT keeps the invariance on scale and rotation change, and illumination change. It maintains a certain degree of stability for image blur and affine transformation. The goal of this research paper is to use SIFT algorithm and classify the vehicles according to class vehicle counting should be done.

II. SCALE INVARIANT FEATURE TRANSFORM OVERVIEW

The SIFT algorithm [5] has four major steps as illustrated in Fig 1.(a) Scale-Space Extrema Detection, (b) Keypoint Localization, (c) Orientation Assignment, (d) Keypoint Descriptor Generation.

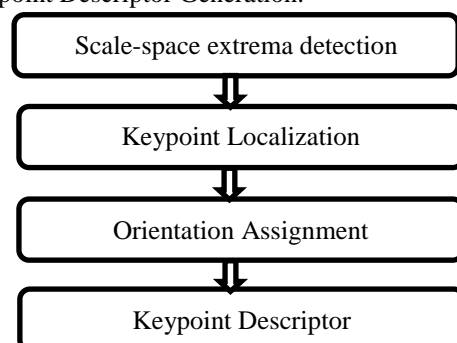


Fig 1. Scale Invariant Feature Transform algorithm

Scale Space Extrema Detection

In this step is extraction of the keypoints takes place. These are the keypoints are invariant to scale change. Hence, it is needed to search for stable features in all possible changes. A function, $L(x,y,\sigma)$, is the scale-space of an image which is obtained from the convolution of an input image, $I(x,y)$, with a variable scale-space Gaussian function, $G(x,y,\sigma)$ as shown in equation (1).

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (1)$$

For the detection of stable keypoint locations, David Lowe has proposed the scale-space extrema in difference-of-Gaussian (DOG) function, $D(x,y,\sigma)$ as shown in equation (2).

$$D(x,y,\sigma) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (2)$$

For the computation of extrema, compute the difference of two nearby scales separated by a constant factor K . The above process for several octaves. The process is shown Figure 2.

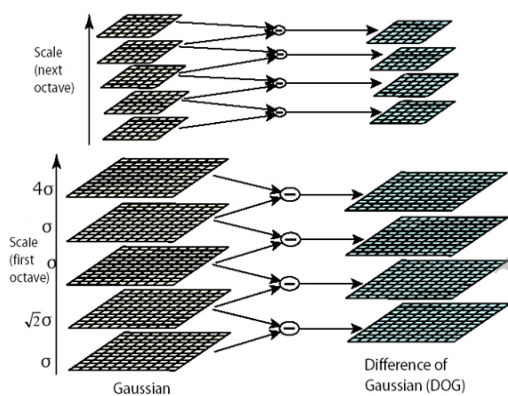


Fig 2. Visual representation of DOG

Then comparison of each sample point (pixel) with its neighbours, is carried out, according to their intensities for finding out whether is smaller or larger than neighbours. Each sample point (pixel) is checked with the eight closest neighbours in image location and nine neighbours in the scale above and below as shown in figure 3. If the sample point is an extrema against all 26 neighbours, is selected as candidate keypoint. The cost of this comparison is reasonably low because most sample point will be eliminated with first few checks.

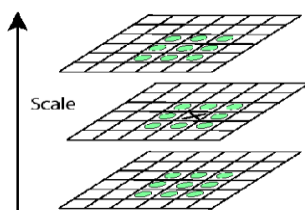


Fig 3. Local extrema detection in DOG

Accurate Keypoint Localization

For each candidate keypoint, interpolation of nearby data is used to accurately determine its position and then the many keypoints which are unstable and sensitive to noise such as the points with low contrast and the points that are on the edge, will be eliminated. Taylor expansion (up to the quadratic terms) of the scale-space function, $D(x,y,\sigma)$, shifted so that the origin is at the sample point:

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (3)$$

Assigning an Orientation

In this step, the keypoint descriptor can be represented relative to this orientation, to obtain invariance to image rotation. For each Gaussian smoothed image sample, the points in regions around keypoint are selected and magnitude m and orientations θ of gradient are calculated. Then computation of created weighted histogram of local gradient directions at selected scale is carried out. Histogram is obtained by quantizing the orientations into 36 bins to covering 360 degree orientation range. The highest peak in the histogram is detected where peaks in the orientation histogram gives dominant directions of local gradient.

Keypoint Descriptor

The above operation gives location, scale, and orientation to each keypoint, which provides invariance to these parameters. For an image sample, $L(x,y)$ at this scale, the gradient magnitude $m(x,y)$, and orientation $\Theta(x,y)$, is calculated using equation (4) and (5):

$$m(x,y) = \sqrt{(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) - L(x,y-1))^2} \quad (4)$$

$$\theta(x,y) = \tan^{-1}((L(x,y+1) - L(x,y-1)) / (L(x+1,y) - L(x-1,y))) \quad (5)$$

The descriptor is based on 16×16 samples and that the keypoint is in the centre of. These samples are divided into 4×4 sub regions in 8 directions around the keypoint. Each point has weighted magnitude and gives less weight to gradients far from keypoint as shown in figure 4. Hence feature vector dimensionality is 128 ($4 \times 4 \times 8$).

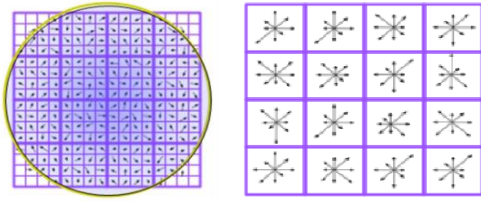


Fig 4. Key point descriptor

At last, vector normalization is applied. The vector is normalized to unit length. A change in image contrast in which each pixel value is multiplied by a constant will multiply gradients by the same constant. Contrast change will be cancelled by vector normalization and brightness change in which a constant is added to each image pixel will not affect the gradient values, as they are computed from pixel differences. Now, we can find the keypoints of an image in the other image and match them together. One image is the training sample and the other image is the world picture that might contain instances of the training sample. These two images have features associated with them across different octaves. Keypoints in all octaves of a one image match with all keypoints in all octaves in other image, independently. A feature matching takes place with the help of nearest neighbour algorithm. The nearest neighbour is defined as, the keypoint with minimum Euclidean distance for the invariant descriptor vector as described upside.

III. VEHICLE CLASSIFICATION AND COUNTING IMPLEMENTATION SCHEME

The implementation flow of proposed technique of vehicle classification and counting is shown in figure 5. The first function is to read the video clip, which is stored in database and convert that video into number of frames. Second function is to find frame differences and identifying the background with background registration and background subtraction. Next post-processing is segmentation is performed, and the vehicles are classified with the help of SIFT algorithm. The goal of the algorithm is to design an efficient classification based on SIFT algorithm on highway.

Background Subtraction

As a part of preprocessing in the proposed method, moving regions are extracted using background subtraction. Better elimination of background region results into the improvement in performance of the next step i.e. segmentation.

Segmentation

The next step is segmentation. The segmentation operation effectively separates the homogeneous regions from the rest of image.

Scale Invariant Feature Transform

SIFT is the state of the art in the field of image recognition and is used in a wide range of content-based image-retrieval applications. The vehicle regions extracted using above procedure is given as input to the SIFT based classifier for classification. SIFT based features are local and these SIFT features are based on the appearance of the vehicle at particular interest regions. The invariant features are detected and extracted, exploring the scale-space structure of an image. Features are localized and filtered, keeping only those that are likely to remain stable over affine transformations, have adequate contrast, and are not along edges.. The latter parameter gives the ratio between the largest and smallest magnitude eigenvalues of a matrix containing curvature information of a DoG function.

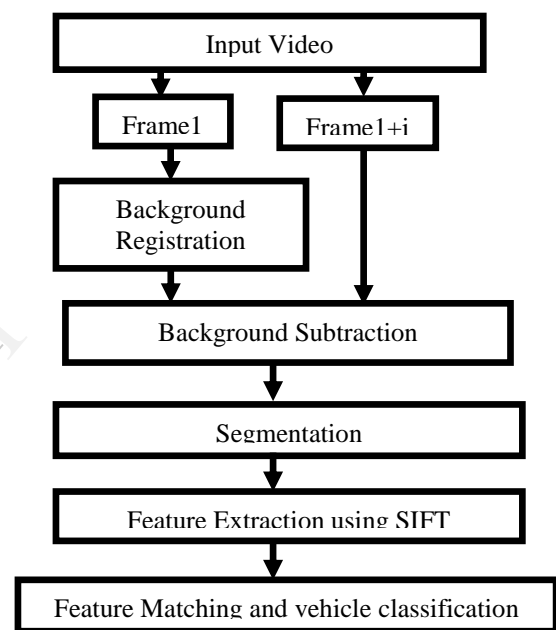


Fig 5. Schematic flow of proposed vehicle classification using SIFT

Feature Matching

With above process, scale invariant key points of the given objects are determined and for each keypoint local descriptors are determined. Hence for each vehicle a set of keypoint, local descriptors are stored in database. when an object is tested for its identity, the SIFT features are extracted and these obtained features are matched to the SIFT feature database which is obtained from the training images. In feature matching[7], for each feature i in the query image, the descriptor is used to search for its nearest neighbour matches among all the features from all the images j contained in a feature database. The nearest neighbours are selected by satisfying a minimum L2 Euclidean distance threshold γ criterion for the descriptor vectors Q_i and DB_{ij} for the query and database image j , respectively.

IV. EXPERIMENTAL RESULTS

As in figure 6. (a), (b), (c), (d) and (e) Original Image, Background subtracted Image, Edge detected Image, Dilated Image and Segmented Image are shown respectively. As a part of preprocessing in the proposed method, moving regions are extracted using background subtraction. Then in next step segmentation is performed. The segmentation operation effectively separates the homogeneous regions from the rest of image. The result of segmentation is shown in figure 6(e). Segmentation step is followed by Feature extraction using SIFT. During this step segmented frame is given as input and keypoint detection of that segmented object takes place. After plotting of keypoints over that object, that obtained keypoints are matched with the database image. Result of the feature matching is shown in figure 6(g).

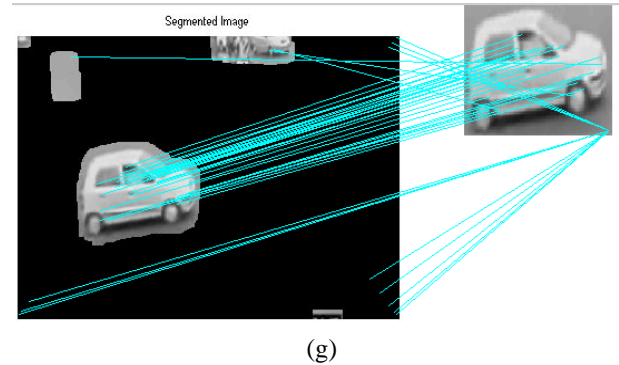
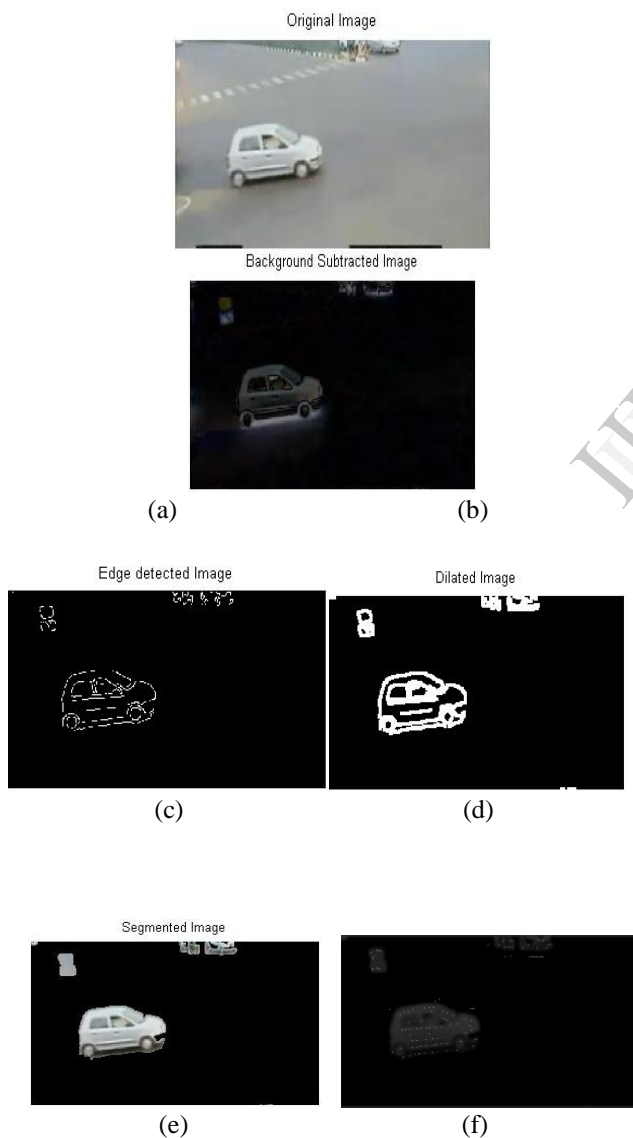


Fig 6.(a) Original Image (b) Background subtracted Image (c)Edge detected Image (d)Dilated Image (e) Segmented Image (f)Keypoints detection of object (g) Feature Matching

V. CONCLUSION

In this paper Scale Invariant Feature Transform (SIFT) is used for vehicle classification technique is described. In this paper keypoint detection and feature matching is done using Matlab R2013a. With the help of SIFT algorithm, extraction invariant image features, that are stable over image translation, rotation, scaling, camera viewpoint and somewhat invariant to changes in the illumination will be possible.

REFERENCES

- [1] Jun Yee Ng, Yong Haur Tay "Image-based Vehicle Classification System" 11th Asia-Pacific ITS Forum & Exhibition, June 2011
- [2] Saroj K.Meher, M.N.Murty"Efficient Method Of Moving Shadow Detection andVehicle Classification "International Journal of Electronics and Communication(AEU),665-670,2013
- [3] David G. Lowe "Distinctive Image Features from Scale-Invariant Keypoints" January 5, 2004
- [4] Jian Wu, Zhiming Cui, Victor S. Sheng, Pengpeng Zhao, Dongliang Su, Shengrong Gong"A Comparative Study of SIFT and its Variants" Measurement Science Review, Volume 13, No. 3, 2013
- [5] Phaneendra Vinukonda, "A Study of the Scale-invariant Feature Transform on a Parallel Pipeline" Theses Submitted to the Faculty of the Louisiana State University and Agricultural and Mechanical College in partial fulfilment of the requirements for the degree of Master of Science in Electrical Engineering, May 2011
- [6] Apostolos P. Psyllos, Christos-Nikolaos E. Anagnostopoulos, Eleftherios Kayafas,"Vehicle Logo Recognition Using a SIFT-Based Enhanced Matching Scheme" IEEE Transactions on Intelligent Transportation Systems, vol. 11, no. 2,June2010
- [7] Dan Middleton, Deepak Gopalakrishna, and MalaRaman "Advances in traffic data collection and management" Texas Transportation Institute Cambridge Systematics, Inc. January, 2003
- [8] Mrs. P.M.Daigavane and Dr.P.R.Bajaj "Real Time Vehicle Detection and Counting Method for Unsupervised Traffic Video on Highways Unsupervised Traffic Video on Highways" IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.8, August 2010