# Vehicle Detection from Aerial Images Using Local Shape Information

Jae-Young Choi and Young-Kyu Yang⋆

College of IT, Kyungwon University,
Seongnam, Gyeonggi, 461-701, Republic of Korea
{jychoi,ykyang}@kyungwon.ac.kr

**Abstract.** Detection and extraction of vehicle objects in high resolution satellite imagery are required in many transportation applications. This paper presents an approach to automatic vehicle detection from aerial images. The initial extraction of candidate vehicle is based on Mean-shift algorithm with symmetric character of blob-like car structure. By fusing the density and the symmetry, the method can remove the ambiguous blobs and reduce the cost of the detected ROI processing in the subsequent stage. To verify the blob as a vehicle, log-polar shape descriptor is used for measuring similarity. The edge strengths are obtained and represented as its spatial histogram by the orientation and distance from the center of blob. The proposed algorithm is able to successfully detect the vehicle and very useful for the traffic surveillance system.

**Keywords:** Vehicle detection, Aerial imagery, Traffic monitoring, Mean shift, Shape description, symmetry.

## 1 Introduction

In recent years, analysis of the aerial or satellite imagery play an important role as improving in the performance of equipment and growing the application fields. However, there is not much review on vehicle detection using aerial image. All of them are categorized into a variety of aspects such types of sensors, target vehicle types or types of measurements etc [1].

This paper deals with automatic detection of vehicle in high resolution aerial imagery of approximately 0.25 meter resolution. Mean-shift clustering algorithm extracts the candidate blob with symmetric property of car shape. By using both geometric and radiometric characteristics, the ambiguous blobs are eliminated and the cost of detected ROI processing is reduced in the subsequent step. To verify the blob in the next step, log-polar shape descriptor is used for measuring similarity. The edge strengths are obtained and represented as its spatial histogram by the orientation and distance from the center of blob. The phase symmetry information from above stage helps to compensate the orientation of the shape context. In our last algorithm, the candidate blobs are merged in case that the blobs are extracted from same vehicle.

---

⋆ Corresponding author.

The remainder of this paper is organized as follows. We first introduce related work on vehicle detection in Section 2. Section 3 describes the proposed vehicle detection approach. In this section, we present the blob clustering method, measuring phase symmetry of blob, and shape description of detected blob as well as overview of proposed algorithm. Section 4 shows the experiment result and Section 5 contains conclusion.

## 2   Related Works

Detection and segmentation are arguably the most important operations in low-level image processing. Especially, Interesting of the vehicle detection has been increased in the computer vision area because this item is such a significant part of our life.

Several authors suggested the use of feature vectors from image region. Lowe's SIFT algorithm is a popular algorithm for extracting the salient features from object [2]. Invariant features are robust to image scale, rotation, and partially invariant to changing viewpoints, and change in illumination [3]. However, if the object or region is too small to detect, it cannot be known a priori if it will be represented by any key point.

A variety of boosting algorithms have been developed for detecting object by machine learning [4]. After remarkable success of the face detector, boosting method has been widely used for solving recognition problems. The main task of this work is to train a classifier. Therefore, the method depends on the selection and learning of exemplar such as car model or features instead of pixel values. In addition to training, it exhaustively scans the whole image to find the exact object. Some efforts use statistical learning to resolve the variance of appearance, but the complex relationship of the different appearance is difficult to learn.

Neural network or Support Vector Machine algorithm is applied to detect the vehicle from low-resolution image [5]. In this work, they only address detection of cars aligned with road direction. This approach may be sufficient for their data. Generalized Hough Transform is used for extract road side. A certain rectangular band width is examined in search position where GHT response is high. The number of edge pixels inside the search band is taken to be the likelihood of the presence of a vehicle in that area. They use the mask which consists of four elongated operators with first-derivative of Gaussian cross sections in recent work [6,7].

Model based approach is introduced in [8]. The method uses an explicit model that consists mainly of geometric features and radiometric properties. The vehicle is modeled as 3D object by a wire frame representation. They also use vehicle queue model to find stands on a road. In this case, detection relies on matching the model. If there is sufficient support of the model in the image, a vehicle is assumed to be detected.

Above approaches have their advantages and disadvantages. Most of the previous work regards a vehicle as a 2D pattern. Vehicle detection in aerial images is relatively constrained by the resolution. Therefore, we need a variety of information of vehicle in order to extract outstanding characteristics.

# 3   Vehicle Detection Approach

This section describes a proposed approach to detect vehicles. With respect to the detection of a vehicle, we can consider following conditions;

- Vehicle is aligned with road which sides run parallel each other
- Car has symmetric structure as geometric aspect
- Size of car is approximately constant length
- Road surface is almost biggest and homogenous pavement (in case of urban)

To detect a vehicle model, there are two techniques: "top-down" and "bottom-up". The surface of a vehicle is made of pixels. But the vehicles form also very frequently lines themselves settled in meta-lines [9].

Proposed method is intended to use "bottom-up" way that we measure the pixels in an input image, and cluster the pixels which consist of car-structure as a geometric aspect and have a similar value of car as a radiometric point of view. After that, we verify the shape of vehicle using the shape description as shown in Fig. 1.
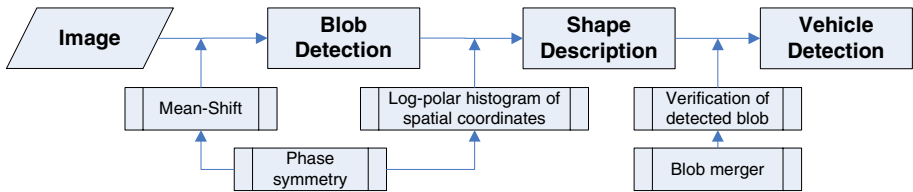


**Fig. 1.** Flow chart of the suggested algorithm

## 3.1   Detection and Clustering of Blob

The mean shift clustering algorithm first applied to image segmentation by Comaniciu and Meer in 1997, whereas the original idea was proposed in 1975 by Fukunaga and Hostetler [10].

This algorithm is designed to find modes, centers of the regions of high concentration, of data represented as arbitrary dimensional vectors. The major steps in the computation of the algorithm as follows [11].

1. Choose the radius r of the search window.
2. Choose the initial location of the window.
3. Compute the mean shift vector and translate the search window by that amount.
4. Repeat till convergence.

The mean shift vector is described in (1).

If $y_j$ is instead of $x$, and $\{y_j\}_{j=1,2,...}$ denotes the sequence of successive locations of the kernel $G(x)$, the equation can be the weighted mean at $y_j$.

$$m(x) = \left[\sum_{i=1}^{n} x_i g\left(\frac{\|x - x_i\|^2}{h}\right)\right] \bigg/ \left[\sum_{i=1}^{n} g\left(\frac{\|x - x_i\|^2}{h}\right)\right] - x \qquad (1)$$

In case of the color image clustering like our application, the RGB color image is mapped into the $L^*u^*v^*$ color space model. The mean shift method clusters this multi-dimensional data set by associating each point to a peak of the data set's probability density. One of the drawbacks in using mean shift algorithm is choice of the fixed bandwidth. That is, once the kernel bandwidth decided by the size of window, it is invariable in the entire tracking process. Recent method modified the bandwidth in $\pm 10\%$ intervals, the tracks object separately using three different kernel bandwidths in the current frame.

In our approach, we use geometric (symmetry of car shape) property as well as radiometric (intensity of pixel level) characteristics. During the probability density function estimates a density in color space, each point which is peak of data set is examined as local maximum of phase symmetry. The measurement of phase symmetry will be introduced in next section in detail.

The advantage of using above fusion method is that it is able to eliminate the ambiguous blobs. As a result of this, the cost of next step is also reduced because it does not need to check whole detected blobs in order to verify the vehicle.

### 3.2   Symmetric Property of Blob

Symmetry is an important property we identify the structure of objects. In this work, the Wavelet Transform is used to obtain local frequency information. Because we are interested in phase information in signals, the method uses wavelets based on complex valued Gabor functions to modulate sine and cosine waves by Gaussian [12]. Let $I$ denote the signal and $M_n^e$ and $M_n^o$ denote the even-symmetric (cosine) and odd-symmetric (sine) wavelets at a scale $n$, then the responses of each quadrature pair of filters as in (2).

$$[e_n(x), o_n(x)] = [I(x) * M_n^e, I(x) * M_n^o] \tag{2}$$



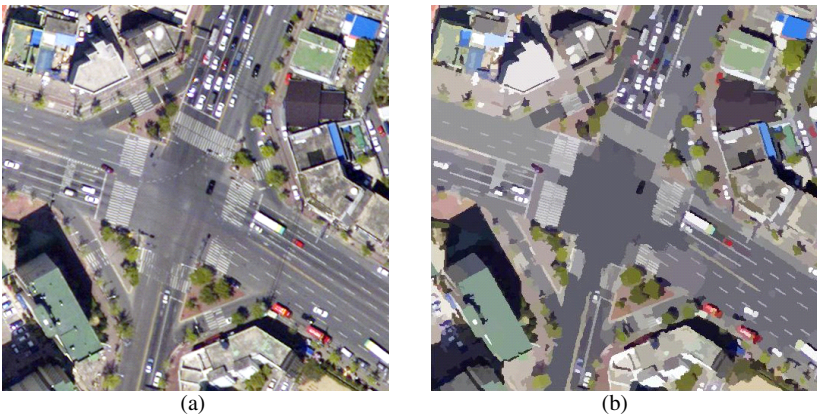**Fig. 2.** Mean-shift clustering. (a) Input aerial image, (b) Clustering of blob using Mean-shift algorithm.

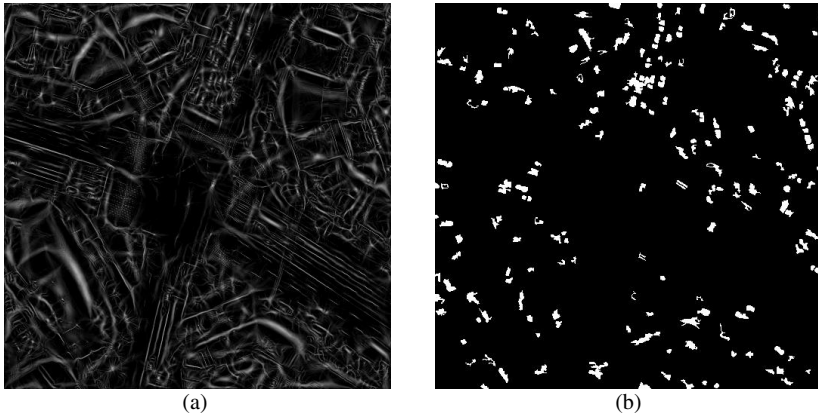(a)                                                              (b)

**Fig. 3.** Results using symmetric property. (a) Phase symmetry output of Fig. 2(a), (b) Blob detection using symmetry information with Mean-shift density.

At each point $x$ in a signal, the response vector for each scale of filter will be obtained. For example, at a point of symmetry, the absolute value of the even-symmetric filter output will be large and another output will be small. This produces a function that varies between $\pm 1$ and varies linearly with phase deviation. In case of multiple scales, the difference of the values of the even and odd filter responses at each scale is weighted by the magnitude of the filter response vector at each scale $n$. This equation is shown in (3). The amplitude of the transform at a given wavelet scale is $A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2}$ and the phase is $\phi_n(x) = atan2(e_n(x), o_n(x))$.

$$Sym(x) = \frac{\sum_n \lfloor A_n(x)[|cos(\phi_n(x))| - |sin(\phi_n(x))|] - T \rfloor}{\sum_n A_n(x) + \varepsilon}$$
$$= \frac{\sum_n \lfloor [|e_n(x)| - |o_n(x)|] - T \rfloor}{\sum_n A_n(x) + \varepsilon} \tag{3}$$

The factor $T$ is a noise compensation term representing the maximum response generated from noise in signal. This 1D equation can extend to 2D by applying the 1D analysis in multiple orientations and forming a weighted sum of the result. This can be used as a line and blob detector. Phase symmetry is an illumination and contrast invariant measure of symmetry in an image.

Fig. 3(a) illustrates phase symmetry measure from the Fig. 2(a). As geometric property of vehicles, there is a bilateral symmetry even though a vehicle has its shadow. Therefore, we calculate the phase symmetry value and apply it to the mean shift process to detect blob which has car-like structure as shown in Fig. 3(b).

### 3.3   Shape Description

In order to verify the detected blob as vehicle, we treat a blob as a point set and assume that the shape of an object is represented by discrete pixel set
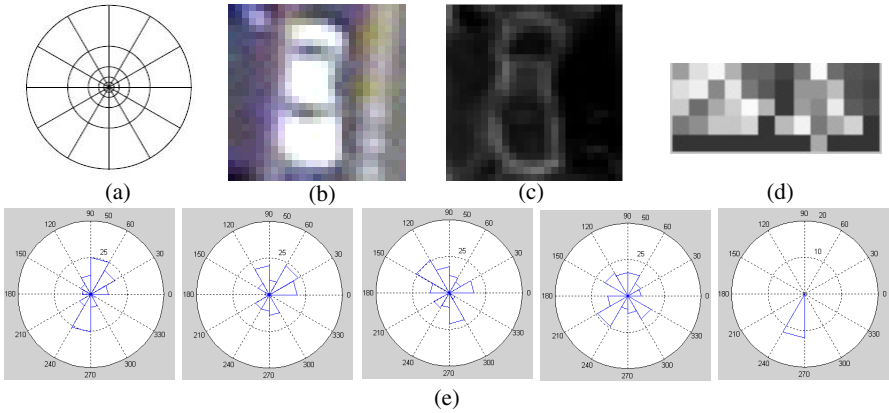
**Fig. 4.** Shape description. (a) Diagram of log-polar histogram, (b) Input region for shape description, (c) Edge image, (d) Shape context represented 5×12 bins, (e) The value of the each shape context bin (from left to right of (e) corresponds to from top row to the bottom row of (d), respectively).

which has the information such as geometric relationship between pixels and the different value among its adjacency pixels. A conventional approach of matching between a given model and target object is an exhaustive search on whole region or histograms for every possible points. In case the search should be done at different orientations, the whole process should be repeated as many times as the number of directions.

In this section, we propose the shape description that is flexible in rotation of object. As mentioned in the beginning of Section 3, the size of a vehicle is approximately constant length and the shape of one is almost uniform. Therefore we can estimate the geometric shape of vehicle by measuring the distance and orientation between the center of blob and its surrounding edges.

Consider an edge of image $I(x, y)$ with a blob and its neighborhood (the size of search mask is 25×25 and the center of mask is a centroid of blob in this experiment), the orientation $\theta$ is defined by $\theta(x, y) = arctan\lfloor d_y(x, y)/d_x(x, y)\rfloor$ where $d_x, d_y$ are the distance from the center of blob. Distance $r$ is also obtained by $r = \sqrt{d_x^2 + d_y^2}$.

We use a log-polar coordinate system which has 5 bins for log $r$ and 12 bins for $\theta$ because the descriptor should be more sensitive to differences in nearby pixels in Cartesian spaces [13]. Fig. 4(a) depicts the diagram of log-polar histogram bins. Each bins of log-polar diagram is represented as a 5×12 array as shown in Fig. 4(b). For normalizing the histogram $H$, each accumulated bin $H(i)$ is divided by the number of detected pixel in each sector.

The direction of blob can be estimated from the phase symmetry information treated in previous section. These directions of interest (DOI) will be used for compensating the orientation of the shape context. In our approach, we rotate the inspect region which contains a blob and its neighborhood by the value of

compensated orientation. That is, after the direction of symmetry axis of the inspect region is obtained, its shape context bin is shifted from direction of symmetry to the vertical axis by the difference of orientation.

Reference bins are made by strength of grey level difference along the boundary of general vehicle, assuming that it is approximately a rectangle and ladder. A compensated shape context bins have same direction with reference one, and are matched histogram similarity by Euclidean distance.

### 3.4   Post Processing

As can be seen from Fig. 4(b), the separate detected blob from same vehicle is due to the front and rear windshield of car. For example, in case of a bright car we expect a strong gray value edge since a windshield in usually very dark, while in case of a dark car the gray value edge may disappear completely.

Therefore, the proposed algorithm merges the blobs which are located around and have color constancy between hood(trunk) color and roof color as a radiometric feature.
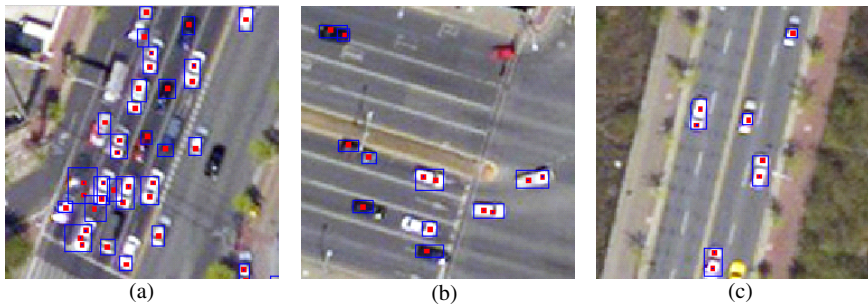


(a)                              (b)                              (c)

**Fig. 5.** Examples of blob merge (the red dot is a centroid of blob and the blue rectangle is a merged blobs)

Fig. 5 depicts the result of blob merge. The red dot indicates the center of initial blob and the blue rectangle expresses the integration of blobs. Some vehicles such as a truck have a different color of cargo. For this reason, same vehicle is counted as two objects as shown in the center image of Fig. 5.

## 4   Experimental Results

The implemented method is intended to detect vehicles on the aerial images of 0.25 meter resolution and performed on Matlab 7.0 platform. The detector is evaluated on 512×512 sub-images from the whole aerial image. In the test image, a typical vehicle (medium sedan) is around 15 to 25 pixels in length and around 10 to 12 pixels in width.

Some of results show that almost vehicles have been detected by a red rectangle(see Fig. 6). The evaluation does not refer to real ground truth. We marked
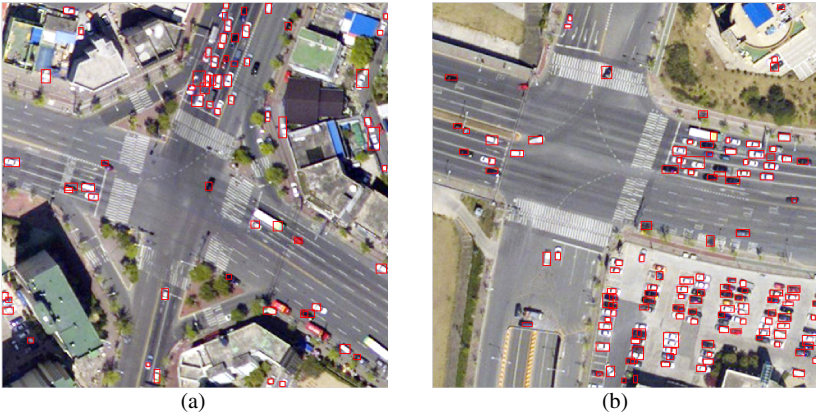
(a)                                    (b)

**Fig. 6.** Results of vehicle detection

**Table 1.** Evaluation of test images

| Evaluation criteria | Left image | Right image |
|---|---|---|
| Number of detection | 63 | 129 |
| False alarms | 5 | 11 |
| False negative alarms | 12 | 16 |
| Detection rate (%) | 84.0 | 88.9 |
| Accuracy rate (%) | 92.1 | 91.5 |

all cars in each sub-images by hand. In both images, the suggested method detected 176 correct vehicles and missed 28 vehicles, hence the detection rate is 87.3% and false alarm rate is 7.8%. There is always a tradeoff between false alarm rate and mis-detection rate. The more detailed descriptions are used, the more number of features and conditions are needed to cover all types of vehicles. Table 1 summarizes the numerical evaluation of the both images in Fig. 6.

In Fig. 6, false negative alarm (missing detection) occurs in regions where the road or parking lot is darkened by building or tree shadows. Large vehicles are also missing because our method focused on the typical vehicles and examined around 25×25 pixels neighborhood.

In addition to above environment in input image, the detection rate is relative to the setting of parameters during mean-shift algorithm processing. Over-segments cause false negative alarms because the different part of vehicle such as hood, roof, trunk, windshield ,and shadow is not merged. For these reasons, a small blob is eliminated as a noise or an isolated blob. For addressing these problems, the suggested method has to be extended by considering the global information in order not to damage a component of a vehicle structure.

However, the proposed method detects vehicle with higher accuracy rate although there exist many objects in urban scenes with a similar appearance.

## 5    Conclusion

Object detection become challenging due to time-consuming manual detection. We are developing a novel approach for detection of vehicle from aerial image.

In order to extract initial candidate vehicle, Mean-shift clustering algorithm is used for detecting the dense and symmetric blob. Fusing geometric and radiometric characteristics helps to reject ambiguous blobs and save the cost of ROI processing in the subsequent step. To verify the detected blob, we apply the log-polar shape descriptor to measure similarity. It is able to avoid the sensitivity of differences in nearby pixels.

From the phase symmetry information, DOI can be used for compensating the orientation of the shape context to match with reference vehicle. Since the vehicles are represented by a few pixels, their detection is very sensitive to the surrounding context. In post processing step, the candidate blobs are merged in case that the blobs are extracted from same vehicle. Our work will be continued by refining the process and making the robust and flexible shape context bin to verify a variety of vehicles.

We believe that similar approaches can also be useful for other object detection as well as transportation application.

## Acknowledgements

## References

1. Punvatavungkour, S., Shibasaki, R.: Three Line Scaner Imagery and On-street Packed Vehicle Detection. Int'l Archives of Photogrammetry Remote Sensing and Spatial Information Sciences 35(3), 355–359 (2004)
2. Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints. Int'l J. Computer Vision 60(2), 91–110 (2004)
3. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
4. Viola, P., Jones, M.: Rapid Object Detection using a Boosted Cascade of Simple Features. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 511–518 (2001)
5. Zhao, T., Nevatia, R.: Car Detection in Low Resolution Aerial Images. Image and Vision Computing 21, 693–703 (2003)
6. Chellappa, R., Burlina, P., Davis, L.S., Rosenfeld, A.: SAR/EO Vehicular Activity Analysis System Guided by Temporal and Contextual Information. In: Proc. 1994 ARPA Image Understanding Workshop, pp. 615–620 (1998)
7. Moon, H., Chellappa, R., Rosenfeld, A.: Performance Analysis of a Simple Vehicle Detection Algorithm. Image and Vision Computing 20(1), 1–13 (2002)
8. Hinz, S.: Integrating Local and Global Features for Vehicle Detection in High Resolution Aerial Imagery. Photogrammetry Remote Sensing Spatial Information 34(3W/8), 119–124 (2003)

9. Ruskone, R., Guigues, L., Airault, S., Jamet, O.: Vehicle Detection on Aerial Images: A Structural Approach. In: Proc. Int'l Conf. Pattern Recognition, pp. 900–904 (1996)
10. Fukunaga, K., Hostetler, L.D.: The Estimation of the Gradient of a Density Function with Applications in Pattern Recognition. IEEE Trans. Information Theory 21(1), 32–40 (1975)
11. Comaniciu, D., Meer, P.: Robust Analysis of Feature Spaces: Color image Segmentation. In: IEEE Conf. Computer Vision and Pattern Recognition, pp. 750–755 (1997)
12. Kovesi, P.: Symmetry and Asymmetry from Local Phase. In: 10th Australian Joint Conf. on Artificial Intelligence, pp. 185–190 (1997)
13. Belongie, S., Malik, J., Puzicha, J.: Shape Matching and Object Recognition using Shape Contexts. IEEE trans. Pattern Analysis and Machine Intelligence 24(24), 509–522 (2002)