Detection of Vehicles Using Gabor Filters and Affine Moment Invariants from an Image

Tadayoshi Shioyama¹, Haiyuan Wu², and Atsushi Iwai¹

¹ Department of Mechanical and System Engineering
Kyoto Institute of Technology
Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan
shioyama@ipc.kit.ac.jp

² Department of Computer and Communication Sciences
Wakayama University

Abstract. This paper proposes a new algorithm for detecting vehicles from an image. An image is at the first segmented into regions by using not only color information but also Gabor transformation of grayscale image. Second, candidate regions corresponding to a vehicle are extracted using affine moment invariants. Third, a true region for a vehicle is selected from candidate regions using normalized cumulative histogram of grayscale in a window which is set for a candidate region of interest, and from the selected region the area of a vehicle is detected.

1 Introduction

This paper addresses vehicles detection for the purpose of aiming at a travel aid for the blind pedestrian. In general, vehicle detection is important for traffic surveillance of the traffic flow and also useful for intelligent transportation control or collision avoidance in vehicles. For the purpose of realizing a travel aid for the blind, it is also important to detect vehicles because vehicles are dangerous moving obstacles for a pedestrian.

Many vision-based methods have been proposed for detecting vehicles under various viewing situations. Some of them view the road from a moving vehicle [1–3], or from an airborne camera [4], and some look at traffic intersections from above [5]. Gabor filters [6,7], frame differences of image sequences [5], HMM (Hidden Markov Model) [1,4] and optical flows [2,3,8] have been used for detecting objects.

In this paper, we present a new algorithm for detecting vehicles from an image viewed from a pedestrian, in order to aim at realizing a travel aid for the blind pedestrian. In our method, at the first an image is segmented into regions by using not only color information but also outputs of Gabor filters operated on a grayscale image. Second, we find candidate regions corresponding to a vehicle as follows. At each region in the segmented image, we calculate the affine moment invariants for a contour of the region, and compare the invariants with the invariants of reference contours which typically exist in a vehicle region. If the region of interest has the same invariant as one of reference contours, then

the region is treated as a candidate region corresponding to a part of a vehicle. Third, we select a true region corresponding to a vehicle from candidate regions as follows. For each candidate region, we set a rectangular window including the region and compute the normalized cumulative histogram of grayscale for pixels belonging to the window. We select a region corresponding to a vehicle by removing spurious regions other than a vehicle using the normalized cumulative histogram. For the purpose of evaluating the proposed method for detecting vehicles, experimental results are shown for many real images.

2 Segmentation

Segmentation is performed using not only color information but also outputs of Gabor filters.

2.1 HLS Color Space

We use the HLS color model where H denotes Hue angle, L lightness and S saturation. The lightness L takes a value in the range $[-1, +1]$, the S takes a value in the range $[0, 1]$ and the H takes a value in the range $[0, 2\pi)$ as illustrated in Fig. 1.

$$HLS(x, y) = (L, S \cos H, S \sin H), \quad (1)$$

where (H, L, S) are values at an image coordinate (x, y) .

2.2 Gabor Filter

Let $f(x, y)$ be a signal at (x, y) in a input image. Then a Gabor transformation of $f(x, y)$ is given by a convolution:

$$z(x, y) = \int \int_{-\infty}^{\infty} f(x - t, y - s) \exp \left(-\frac{t^2 + s^2}{s\sigma^2} \right) \times \exp \{ -j2\pi(u_0 t + v_0 s) \} dt ds, \quad (2)$$

Fig. 1. HLS color model

where (u_0, v_0) are coordinates in a 2 dimensional frequency space, and a Gabor function is given by

$$g(x, y) \equiv \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \exp\{-j2\pi(u_0x + v_0y)\}. \quad (3)$$

A Fourier transformation of $g(x, y)$ is given by

$$G(u, v) = 2\pi\sigma^2 \exp[\{(u - u_0)^2 + (v - v_0)^2\}/(2\sigma_{uv}^2)], \quad (4)$$

here, $\sigma_{uv} \equiv 1/2\pi\sigma$. Thus, a Gabor filter is a band pass filter with its center of (u_0, v_0) . The radial frequency r_0 is given by

$$r_0^2 = u_0^2 + v_0^2. \quad (5)$$

A Gaussian window of a Gabor function has a width σ . A Gabor transformation is considered to be a frequency analysis (Fourier transformation) within a local area which has a width σ . In order to keep a number of waves within the width σ constant, we introduce a following relation:

$$\sigma r_0 = 1/\kappa. \quad (6)$$

For the purpose of detecting vehicles, we empirically use two radial frequencies r_1 and r_2 , and four directions θ 's where θ is determined by $\theta = \tan^{-1}(v_0/u_0)$. Eight Gabor filters are allocated in a frequency space as follows [9]. Let two adjacent directions be $\theta = q_1$ and q_2 and an angle between those be q . In a frequency space, let $c_i, i = 1, 2, 3, 4$ be centers of Gabor filters with radial frequencies of r_1 and r_2 . and with directions of q_1 and q_2 . Let P be a point in a frequency space where all outputs of the four Gabor filters have an equal value ϵ . Given r_1, q, ϵ and κ , the another radial frequency r_2 is determined by [9]

$$r_2/r_1 = A/(1 - A), \quad A = \frac{1}{2} \left(1 - \sqrt{1 - \frac{1 + 2\kappa^2 \ell m \epsilon}{\cos^2(q/2)}} \right). \quad (7)$$

In this paper, we set as $\epsilon = 0.3, q = 45^\circ$ and

$$\text{Low: } r_0 = 0.14, \quad \sigma = 2.43, \quad \text{High: } r_0 = 0.33, \quad \sigma = 1.10.$$

An allocation of Gabor filters used in this paper is illustrated in Fig. 2.

2.3 Gabor Features

We denote by $GE_{\theta r_0}$ the output of Gabor filter with a radial frequency r_0 and with direction θ for a grayscale image $f(x, y)$, here $r_0 = h(\text{High})$ and $\ell(\text{Low})$, and $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$. We define a feature vector $GE(x, y)$ as a vector with 8 components $GE_{\theta r_0}(x, y)$:

$$GE(x - y) = (GE_{0^\circ \ell}(x, y), (GE_{45^\circ \ell}(x, y), (GE_{90^\circ \ell}(x, y), (GE_{135^\circ \ell}(x, y), \\ GE_{0^\circ h}(x, y), (GE_{45^\circ h}(x, y), (GE_{90^\circ h}(x, y), (GE_{135^\circ h}(x, y)). \quad (8)$$

The examples of $GE(x, y)$ are shown in Fig. 3.

Fig. 2. Gabor filter allocation

Fig. 3. Examples of $GE(x, y)$

2.4 Segmentation

For segmentation, we use both the feature vector $GE(x, y)$ and the color vector $HLS(x, y)$. The segmentation is carried out by merging similar regions neighboring to a considered region.

First, the initial regions are generated in the following steps:

1. Assign a label to an unlabelled pixel found by scanning.
2. Compute the distance D^2 between the last labelled pixel and its 4-neighboring pixel, and assign the same label to the pixel where the distance is less than a threshold T_1 :

$$D^2 \equiv \|GE(x_a, y_a) - GE(x_b, y_b)\|^2 + \eta \|HLS(x_a, y_a) - HLS(x_b, y_b)\|^2, \quad (9)$$

where (x_a, y_a) and (x_b, y_b) are the coordinates of 4-neighboring pixels in an original image, and η denotes a weighting coefficient.

3. Iterate the same process as the step 2 for the pixel which is assigned the label in the step 2. When there is no pixel which is to be assigned the same label, go to the step 4.
4. If all pixels are already labelled, then stop. Otherwise, return to step 1.

For each initial region, we compute the average of the feature vectors, and also compute the center of gravity and the second central moments such as covariances from the coordinates of pixels belonging to the region. Utilizing these information, we efficiently merge regions by the following algorithm:

1. For each region i , find a candidate region for merging which satisfies the following condition:

$$\min_j E_{ij} \equiv D_{ij}^2 \{R_{ij} + \kappa(M_{ij}^2 + M_{ji}^2)\} \quad (10)$$

$$\text{and } E_{ij} < T - 2. \quad (11)$$

where κ and T_2 are constants, and

$$D_{ij}^2 \equiv \|GE_i - GE_j\| + \eta \|HLS_i - HLS_j\|^2, \quad (12)$$

GE_i, HLS_i : the average of GE and HLS in region i ,

$$R_{ij} \equiv N - iN_j / (n_i + n_j), \quad (13)$$

n_i : the number of pixels in region i ,

$$M_{ij}^2 \equiv \frac{1}{(1 - \rho_i^2)V_{x_1}V_{y_i}} \{V_{y_i}(x_i - x_j)^2 - 2\rho_i\sqrt{V_{x_1}V_{y_i}}(x_i - x_j)(y_i - y_j) + V_{x_i}(y_i - y_j)^2\}, \quad (14)$$

(x_i, y_i) : the center of gravity of region i ,

V_{x_1}, V_{y_i} : variances in x-axis and y-axis in region i

ρ_i : correlation coefficient in region i ,

M_{ij} : Maharanobis distance from region i to region j .

2. Merge two regions which are mutually candidates each other. Compute the average of feature vector and the second central moments for the new region after merging.
3. Iterate the procedure of steps 1 and 2 until there is no region which satisfies the condition (11).

The criterion E_{ij} includes the factor of similarity D_{ij} , the size factor R_{ij} and the Maharanobis distance M_{ij} which depends on the shape of the region. In the algorithm, the region with small size (i.e. small R_{ij}) is superiorly merged. The example of segmentation is shown in Fig. 4 where pixels belonging to the same region are shown to have the same grayscale, i.e. the average grayscale of the region.

Fig. 4. Example of segmentation:(a) original image, (b) result

3 Detection of Vehicles

In this section, we find candidate regions including a vehicle by using affine moment invariants, and select a region including truly a vehicle by using normalized cumulative histogram.

3.1 Extraction of Candidate Regions Using Affine Moment Invariants

After the segmentation, a vehicle is in general segmented into multiple regions. Some among such regions have typical contours corresponding to a vehicle, such as side body, side window, front window, bonnet and so on as illustrated by dotted black lines in Fig. 5. The typical contours are used for finding candidate regions corresponding to a vehicle.

3.1.1. Affine Moment Invariants

In general, when contours of planar objects are observed by a camera, the shapes of the observed contours change depending on the viewpoint of a camera. However, the changes of observed shapes are known to be according to affine transformation if the distances from a viewpoint to objects are sufficiently far in comparison with the differences of depths of objects. Therefore, we use affine

Fig. 5. (a)–(d) Some of typical contours shown by dotted black lines and (e) an extracted region for a vehicle

moment invariants, which are invariant under the affine transformation, in order to find candidate regions corresponding to a vehicle. Affine moment invariants are given by [10]

$$I - a = \mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4, \quad (15)$$

$$I - b = \mu_{30}^2\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{30}\mu_{12}^3\mu_{03} - 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10}, \quad (16)$$

$$\mu_{pq} = \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q \rho(x, y) dx dy, \quad p, q = 0, 1, 2, \dots, \quad (17)$$

$(\bar{x}, \bar{y})$: the center of gravity of a counter,

$\rho(x, y) = 1$ at contour,

$= 0$ otherwise. (18)

(19)

3.1.2 Extracting Candidate Regions with Affine Moment Invariants

In learning process, we calculate affine moment invariants I_{am}^* , $m = 1, 2, \dots, M$, $I_{bm'}^*$, $m' = 1, 2, \dots, M'$, for the typical contours, and store those values as reference models. For each region, which satisfies the conditions $n_i > 40$ and $s_y < s_x < 75s_y$, in test segmented image, we calculate affine moment invariants I_a and I_b , and find candidate regions which satisfy the following conditions:

$$|I_a - I_{am}^*| < \epsilon_{am}, \quad m = 1, 2, \dots, M, \quad |I_b - I_{bm'}^*| < \epsilon_{bm'}, \quad m' = 1, 2, \dots, M', \quad (20)$$

where ϵ_{am} and $\epsilon_{bm'}$ are thresholds. Figure 5e shows an example of candidate region which is found from an original image in Fig. 4a.

3.2 Detection of Vehicles Using Normalized Cumulative Histogram

In the candidate regions for a vehicle extracted in the preceding section, there may be few spurious regions other than a vehicle. Hence we remove those spurious regions using normalized cumulative histogram of grayscale in a detection window.

3.2.1 Setting of Detection Window

We set a rectangular frame so as to fit a candidate region and refer to the frame as a “region frame”. We denote by (x_c, y_c) the center of the gravity of a candidate region, and by (x_f, y_f) the center of the gravity of the region frame whose width and height are w and h as in Fig. 6. We denote by y_{min} and y_{max} the coordinates of the upper and lower sides of the region frame. We denote by s_x^2 and s_y^2 the variances of candidate region in x and y directions. We determine the “detection window” from the region frame as follows.

(Case 1). In the case of $y_c < y_f$, the upper side of detection window is set as $y_{min} - h$, and the height is set as $2h$ as illustrated in Fig. 6.

Fig. 6. A region frame and a detection window

(Case 2). In the case of $y_c > y_f$, the lower side of detection window is set as $y_{max} + h$, and the height is set as $2h$.

(Case 3). In the case of $s_x > 6s_y$, the upper and lower sides of detection window are set as $y_{min} - 1.5h$, and $y_{max} + 1.5h$, respectively, and the height is set as $4h$.

In the above all cases, the width of detection window is the same w as a region frame.

3.2.2. Removing Spurious Regions by Geometric Characteristics

In this section, among candidate regions we remove regions which have geometric characteristics of spurious regions, that is, we remove regions which satisfy at least one of the following cases:

(Case 1). The number of pixels in the considered region less than one tenth of the area of the corresponding region frame.

(Case 2). Consider the rectangle whose center coincide with the center of the region frame and whose sizes are a half width and a half height of the region frame of the considered region. Then, in the rectangle the considered region contains pixels whose number is less than two thirds of the rectangle area.

(Case 3). The considered region contains another region whose area is greater than 5 pixels.

(Case 4). The width w and height h of the region frame satisfy the relation of $w < 2.5h$, and the considered region and the region frame have the almost same size and shape.

(Case 5) The considered region has a shape of parallelogram and a boundary of the parallelogram has the same direction as the y -axis of image coordinate.

The region satisfying cases (1–3) corresponds to a region with a complicate shape which for instance consists in many parts each having concave shape. The region satisfying cases (4) and (5) corresponds to a region such as a window of building. The removed candidate regions by geometric characteristics in this section are not checked using normalized cumulative histogram in the next section.

Fig. 7. Normalized cumulative histograms (NCH's) for three cases other than a vehicle

3.2.3. Detection of a Vehicle Using Normalized Cumulative Histogram

In this section, we detect a vehicle by removing spurious regions other than a vehicle using cumulative histograms of grayscale in a detection window.

In normalized cumulative histograms (NCH's) as illustrated in Fig. 7, we define a step S as the interval of grayscale where there are more than ten grayscales with gradient greater than 0.005 and the total increment in the interval is greater than 0.160. We denote by S_n the number of steps S . If the distance between the centers of two steps is shorter than 35 grayscale, these two steps are counted by only one into S_n . We denote by ℓ_z the center of step S with the greatest grayscale and by g_z the increment in the step S with ℓ_z . The spurious regions other than a vehicle have a NCH among three typical NCH's which satisfy the following conditions:

1. $S_n = 0$,
2. $S_n = 1$ and $g_z > 0.810$,
3. $S_n = 2$ and $\ell_z > 135$.

Here, condition (1) implies a uniform grayscale histogram and corresponds to a background region including for instance a sign board and so on, condition (2) implies a grayscale histogram with one dominant peak and corresponds to a background region including for instance shadows and so on, and condition (3) implies a grayscale histogram with two dominant peaks and corresponds to for instance a pedestrian crossing region including white lines and so on. A vehicle region does not correspond to any one of these three conditions. We remove the spurious regions satisfying one of three conditions above

3.2.4. Overlap Rate of Two Detection Windows

Since a vehicle may have multiple detection windows, we merge two detection windows when they have an overlap rate (OR), which is defined in the following, greater than a threshold ($=1.33$): $OR = (X_1 + X_2) / X_1 \cup X_2$, where $X_i, i = 1, 2$ denote widths of two detection windows, and $X_1 \cup X_2$ denotes the width of a union of two detection windows. Thus obtained detection windows are the areas of detected vehicles which correspond to the selected regions among the candidate regions for vehicle.

4 Experimental Results

To evaluate the performance of the proposed method, we use 57 real images of road scenes including vehicles taken by 3-CCD camera (Sony DCR-VX1000 with specification of 1/3 inch CCD and 5.9 mm focal length) at the height of human eye. To obtain images under various illumination conditions, we took images in different weathers. The total real images include 71 vehicles which have areas more than 40 pixels. The parameters are empirically set as follows: $\eta = 80$, $\kappa = 0.01$, $T_1 = 500$, $T_2 = 50000$. We use the real part of Gabor transformation because its performance for vehicle detection is superior to the absolute value of Gabor transformation. We use the reference models of affine moment invariants I_{am}^* , $m = 1, 2, \dots, M$, $I_{bm'}^*$, $m' = 1, 2, \dots, M'$, where $M = 5$ and $M' = 15$. The values of $I_{bm'}^*$, $m' = 1, 2, \dots, 15$, are calculated for the 15 typical contours of regions included in vehicles as illustrated in Fig. 5a–d, for example. Since some of I_{am}^* , $m = 1, 2, \dots, 15$, calculated in the same manner as $I_{bm'}^*$, have similar values, these similar values are averaged and the 5 averages are used as the reference models. This is the reason for the difference between M and M' . The thresholds in equation (19) are empirically set as follows: $\epsilon_{a1} = 14.7$, $\epsilon_{a2} = 21.5$, $\epsilon_{a3} = 20.5$, $\epsilon_{a4} = 4.0$, $\epsilon_{a5} = 75.0$, $\epsilon_{b1} = 7.0$, $\epsilon_{b2} = 1.3$, $\epsilon_{b3} = 2.1$, $\epsilon_{b4a2} = 1.7$, $\epsilon_{b5} = 7.0$, $\epsilon_{b6} = 1.0$, $\epsilon_{b7} = 4.0$, $\epsilon_{b8} = 40.5$, $\epsilon_{b9} = 100$, $\epsilon_{b10} = 50.0$, $\epsilon_{b11} = 100$, $\epsilon_{b12} = 250$, $\epsilon_{b13} = 2.1$, $\epsilon_{b14} = 1.7$, $\epsilon_{b15} = 7.0$. Some of experimental results are shown in Fig. 8. In the figure, detection windows are shown by white rectangles for the areas which are detected as vehicles. All these 71 vehicles were detected as vehicles by the proposed method, and 3 spurious regions other than a vehicle

Fig. 8. Some of the experimental results

were wrong detected as a vehicle as illustrated in Fig. 8l for instance. We decided that the proposed method detects a vehicle when a detection window includes some part of a vehicle, because the objective in this paper is to discriminate whether or not there exists a vehicle.

5 Conclusions

We have proposed a new algorithm for detecting a vehicle using Gabor filter and affine moment invariants. From experimental results, it is found that all vehicles can be detected by the proposed method, that is, there is no dangerous error such that a real existing vehicle is not detected as a vehicle.

References

1. Kato, J., et al., *An HMM-based segmentation method for traffic monitoring movies*, IEEE Trans. on PAMI, Vol. 24, No. 9 (2002) pp. 1291–1296.
2. Duric, Z., et al., *Estimating relative vehicle motions in traffic scenes*, Pattern Recognition, Vol. 35 (2002) pp. 1339–1353.
3. Murray, D.W. and Buxton, B.F., *Scene segmentation from visual motion using global optimization*, IEEE Trans. on PAMI, Vol. 9, No. 2 (1987) pp. 220–228.
4. Tao, H., Sawhney, H.S. and Kumar, R., *Object tracking with Bayesian estimation of dynamic layer representations*, IEEE Trans. on PAMI, Vol. 24, No. 1 (2002) pp. 75–89.
5. Li, X., Liu, Z.Q. and Leung, K.M., *Detection of vehicles from traffic scenes using fuzzy integrals*, Pattern Recognition, Vol. 35 (2002) pp. 967–980.
6. Lau, H.F. and Levine, M.D., *Finding a small number of regions in an image using low-level features*, Pattern Recognition, Vol. 35 (2002) pp. 2323–2339.
7. Jain, A.K., Ratha, N.K. and Lakshmanan, S., *Object detection using Gabor filters*, Pattern Recognition, Vol. 30 (1997) pp. 295–309.
8. Adiv, G., *Determining three-dimensional motion and structure from optical flow generated by several moving objects*, IEEE Trans. on PAMI, Vol. 7, No. 4 (1985) pp. 384–401.
9. Kawada, K. and Arimoto, S., *Hierarchical texture analysis using Gabor expansion*, J. of the Institute of Electronics, Information and Communication Engineers, Vol. J78-DII, No. 3 (1995) pp. 437–444.
10. Flusser, J. and Suk, T., *Pattern recognition by affine moment invariants*, Pattern Recognition, Vol. 26 (1993) pp. 167–174.