International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

Dynamic Urban Road Detection and Vehicle Verification for Vehicle Detection

Okmin Lee, Insu Won, Sangmin Lee, Jangwoo Kwon

***Abstract*—This paper presents a vehicle detection system for urban roads comprising dynamic road detection, moving object detection using background subtraction, and vehicle verification with vertical asymmetry measurement. The proposed system shows better performance than Gaussian mixture model at detecting vehicles on the road.**

***Keywords*—vehicle detection, vehicle verification, dynamic road detection, perspective mapping, Gaussian mixture model**

I. Introduction

In modern society, it is customary to take a car when travelling a long distance. With increases in the use of cars, the car is not just a machine anymore, but complex modern technology that is an amalgamation of machine engineering, telecommunications engineering, electronics, and computer science. A vehicle detection system can be used to avoid crashes, determine the volume of traffic, and so on. There are many methods to detect vehicles—mostly digital image processing and computer vision: horizon and vertical edge analysis, stereo vision by using distance information, background deletion to separate moving objects, and so on. These days, edge-based constraint filters that assist in the segmentation of vehicles from background clutter[4], non-negative matrix factorization (NMF)[5] and shadow detection/object verification[6] are used to detect vehicles. An edge-based constraint filter is just for automobiles. NMF and shadow detection/object verification is slow because it calculates too many things to detect vehicles. To solve these problems, we propose using dynamic road detection with separate region Hough transform (SRHT), perspective mapping, an improved Gaussian mixture model (IGMM), and vehicle verification to detect vehicles on the road. The proposal in this paper is a traffic flow monitoring system for the future.

II. Proposed System

Usually, off-road regions in images cause unnecessary calculations. To avoid that, we propose advanced road area separation with a geometric concept. As shown in Fig.1, the road was detected after canny edge detection and a Hough transform in separated regions. We applied perspective mapping around the vanishing point, then the IGMM is applied to detect vehicles in the road area. Finally, we applied vehicle verification to verify whether there are vehicles or not. SRHT and IGMM are methods proposed in this paper.

Inha University
Incheon, Korea

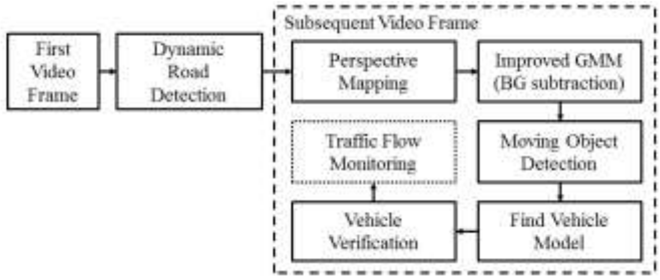

Figure 1. Proposed vehicle detection system.

A. Road Detection

The system should detect the road as the region of interest (ROI). There are two methods for ROI detection: simple ROI detection and dynamic ROI detection. Simple ROI detection specifies a rectangular area without further calculations. This method can calculate the ROI within a constant time. However, a simple rectangular ROI cannot find the exact road area. In contrast, dynamic ROI detection needs additional computation. But it is possible to detect the correct road area. In this paper, an ROI with a hexagonal area is applied to detect the road.

There are several steps to detecting the ROI. First, apply canny edge detection to the input image, as shown in Fig.2 (a) and (b). Second, divide the image with left–right division to apply SRHT to obtain the outline of the road. Third, apply SRHT. After this, the non-linear lines are eliminated. [1]

But unexpected linear properties, like building edges, can still remain. In the fourth step, a slope filter is applied to remove them. In this step, linear lines at 20 ~ 80 degrees to the left and 100 ~ 180 degrees to the right remain. Fig.2 (c) to (e) show results after SRHT is applied. A merged image is shown in Fig.2 (f).

The fifth step is comparison of two linear lines. The lower end point between two linear lines at the center is the starting point of the upper side of the ROI. The upper side is parallel to the y-axis, as shown Fig.2 (g). The detected ROI for the input image is shown in Fig.2 (h).

After the detection of the road, the system proceeds to vehicle detection. Three stages are needed to identify a vehicle: (i) perspective mapping based on the road region, (ii) moving object detection using background subtraction, and (iii) vehicle verification.

B. Perspective Projective Transform

A homographic mapping method is used to illustrate the relationship between two different views of the same real-world scene. Let p and p′ be the corresponding projected image points on the image plane of two different views of the same point located in the 3D real-world coordinate system. Assume the coordinates of this pair of matching

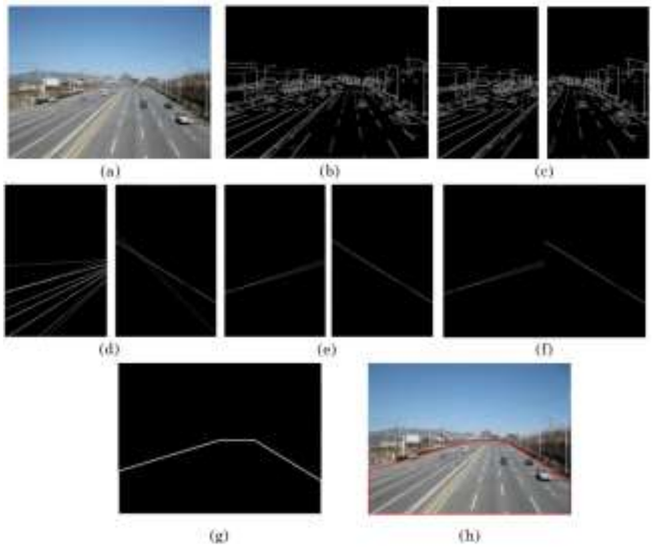

Figure 2. Dynamic road detection. (a) Input image, (b) canny edge detection, (c) image division, (d) Hough transform, (e) slope filter application, (f) SRHT, (g) result of dynamic road detection, and (h) apply to input image

points, p and p′, in inhomogeneous form are denoted as $(x_1, y_1)^T$ and $(x_2, y_2)^T$. Homographic mapping of the two points is then a planar projective transformation.

Under the homograph, we can write the transformation of a point in 3D to 2D coordinates. To estimate H, start from $\mathbf{X}_2 \sim H\mathbf{X}_1$. Written element by element, in homogenous coordinates we get the following (1):

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}. \quad (1)$$

In inhomogeneous coordinates ($x'_2 = x_2 / z_2$, $y'_2 = y_2 / z_2$):

$$(x'_2, y'_2) = \left(\frac{h_{11}x_1 + h_{12}y_1 + h_{13}}{h_{31}x_1 + h_{32}y_1 + h_{33}}, \frac{h_{21}x_1 + h_{22}y_1 + h_{23}}{h_{31}x_1 + h_{32}y_1 + h_{33}}\right). \quad (2)$$

Without loss of generality, set z_1=1 and rearrange:

$$\begin{aligned} x'_2(h_{31}x_1 + h_{32}y_1 + h_{33}) &= h_{11}x_1 + h_{12}y_1 + h_{13} \\ y'_2(h_{31}x_1 + h_{32}y_1 + h_{33}) &= h_{21}x_1 + h_{22}y_1 + h_{23}. \end{aligned} \quad (3)$$

To solve for H, even though these inhomogeneous equations involve the coordinates nonlinearly, the coefficients of H appear linearly, and after rearranging (3), we get (4):

$$\begin{aligned} &\mathbf{a}_{x,y}{}^T\mathbf{h} = 0 \\ &where \\ &\mathbf{h} = (h_{11}, h_{12}, h_{13}, h_{21}, h_{22}, h_{23}, h_{31}, h_{32}, h_{33})^T \\ &\mathbf{a}_x = (-x_1, -y_1, -1, 0, 0, 0, x_2'x_1, x_2'y_1, x'_2)^T \\ &\mathbf{a}_y = (0, 0, 0, -x_1, -y_1, -1, y'_2x_1, y'_2y_1, y'_2)^T. \end{aligned} \quad (4)$$

Given at least four or more corresponding points, it can from the following linear system of (5),

$$\mathbf{Ah} = 0. \quad (5)$$

Where $\mathbf{A}$=[$\mathbf{a}_x{}^T$, $\mathbf{a}_y{}^T$,…, $\mathbf{a}_{xn}{}^T$, $\mathbf{a}_{yn}{}^T$]^T. Equation 5 can be solved using homogeneous linear least squares. That is, H is the eigenvector corresponding to the minimum eigenvalue of **A**. Fig.3 is the result of a perspective projective transform using four points of the ROI that are calculated in dynamic road detection.

C. *Improved GMM for detect moving object*

Background subtraction is mainly used to detect moving objects in video sequences. In an outdoor environment like a road, there are many changeable environmental factors, such as changing illumination, the camera shaking, and suddenly moving objects, which must be considered for robust processing. Normally, GMM is used to subtract the background by adaptively considering the various changes in the scenes. We use GMM with M components [3][4]:

$$p(\vec{x}, \vec{\theta}) = \sum_{m=1}^{M} \hat{\pi}_m N(\vec{x}; \hat{\vec{u}}_m, \hat{\sigma}_m^2 I). \quad (6)$$

where $\vec{\theta}$ is parameter vector of a Gaussian mixture model, $\vec{\theta} = \{\pi_1, \ldots \pi_m, \vec{\mu}_1, \ldots, \vec{\mu}_m, \sigma_1^2 I, \ldots, \sigma_m^2 I\}$, $\hat{\vec{u}}_M$ are the estimates of the means, and $\hat{\sigma}_M$ are the estimates of the variances that describe the m-th Gaussian components. The covariance matrices are assumed to be diagonal, and the identity matrix I has proper dimensions. The mixing weights denoted by $\vec{\pi}_m$ are non-negative and add up to one. Given a new pixel value $\vec{x}^{t+1}$ at time (t+1) the recursive update equations are:

$$\begin{aligned} \hat{\pi}_m^{(t+1)} &= \hat{\pi}_m^{(t)} + \alpha(o_m^{(t)}(\vec{x}^{(t+1)}) - \hat{\pi}_m^{(t)}) \\ \hat{\vec{\mu}}_m^{(t+1)} &= \hat{\vec{\mu}}_m^{(t)} + o_m^{(t)}(\vec{x}^{(t+1)})(\alpha / \hat{\pi}_m^{(t)})\vec{\delta}_m \\ \hat{\sigma}_m^{2(t+1)} &= \hat{\sigma}_m^{2(t)} + o_m^{(t)}(\vec{x}^{(t+1)})(\alpha / \hat{\pi}_m^{(t)})(\vec{\delta}_m^T\vec{\delta} - \hat{\sigma}_m^2). \end{aligned} \quad (7)$$

where $\vec{\delta}_m = \vec{x}^{(t)} - \hat{\vec{\mu}}_m$. Instead of the time interval T that was mentioned above, here, the constant α describes an exponentially decaying envelope that is used to limit the

Figure 3. Perspective mapping results (left: input image, right: perspective mapping image).

influence of the old data. It keeps the same notation, bearing in mind that α = 1/T (approximately). For new data, the ownership $O_m^{(t)}$is set to 1 for the 'close' component with the largest $\hat{\pi}_m$, and the others are set to zero.

Usually, the intruding foreground objects will be represented by some additional distributions with small weights $\vec{\pi}_m$. Therefore, we can approximate the background model with the first B largest distribution. If the components are sorted to have descending weights, $\vec{\pi}_m$, we have:

$$B = \arg\min_b \left(\sum_{m=1}^{b} \hat{\pi}_m > (1 - c_f) \right). \qquad (8)$$

where c_f is a measure of the maximum portion of the data that can belong to foreground objects without influencing the background model. If the object remains static long enough, its weight becomes larger than c_f and it can be considered part of the background. If we look at (7), we can conclude that the object should be static for approximately log(1-c_f) / log(1-α) frames. For example for c_f = 0.1and α = 0.001, we get 105 frames. In Fig.4, we present the moving object detection results using GMM. We get a stable foreground object as shown in part (d).

Using GMM, moving objects are detected. There are too many objects not the vehicle. We should remove the noises. Because only the road area is detected as the ROI, it is not necessary to consider off-road areas, but camera shake, movement of the car, and so on, should be considered for better performance.

In this paper, we add three procedures to solve these problems. First, objects within the average vehicle size are considered a vehicle. This is about 1200 pixels, and noise is a lot smaller than the value. Second, when detected regions overlap, we determine if there is either a vehicle or noise by using (9).[5]

$$o(R_1, R_2) = \begin{cases} true \quad if & pt_{11} < pt_{21} \text{ and } pt_{12} > pt_{22} \\ & pt_{11} > pt_{21} \text{ and } pt_{12} < pt_{22} \\ false \quad else & \end{cases} \qquad (9)$$

where R_1 and R_2 are each regions. R_1 has pt_{11} and pt_{12}. This means that pt_{11} is the top-left point of R_1, and pt_{12} is the bottom-right point of R_2. So does R_2. There are four cases for that, as shown in Fig.4. If two regions overlap as in Fig.4 (a) and (d), $o(R_1,R_2)$ returns true. Otherwise, if two regions overlap, as in Fig.4(b) and (c), $o(R_1,R_2)$ returns false. Third, when ratios of a region are almost square, it is considered a vehicle. We calculated average ratios of each region. The ratio of regions for the vehicle is almost square. But if it is not, the ratios of regions are various rectangle ratios, as seen in Table 1. We determine that by using (10).

(a) (b) (c) (d)

Figure 4. Moving object detection using IGMM: (a) first frame, (b) unstable background subtraction image, (c) 110th frame, and (d) stable results.

$$T = \arg_{i=0,1,\ldots,n-1} R(i) = \begin{cases} true \quad if \quad 0.9 < R(i) < 1.1 \\ false \quad else \end{cases}$$

$$R(i) = \frac{height(i)}{width(i)} \qquad (10)$$

T determines if there is a vehicle or not by making use of $R(i)$. $R(i)$ is ratio of each region. If $R(i)$ is bigger than 0.9 or smaller than 1.1, T is true; otherwise, T is false.

D. Vehicle Verification

The system verifies if detected moving objects are a vehicle or not. The object region is turned into an edge difference map by comparing all columns to the first column of the possible regions, and the difference map equation is defined as (11). $VS(i)$ is the vertical asymmetry measure with the asymmetry axis located at y=i. [6]

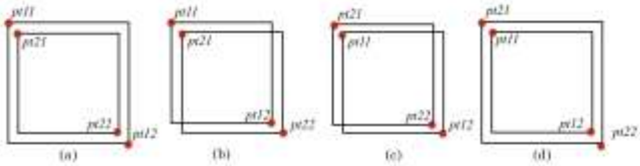

Figure 5. Cases when each object is stacked.

TABLE I. THE RESULTS OF RATIOS OF A RECTANGULAR REGION

	Image 1	Image 2	Image 3	Image 4	Image 5
True Positive	1.06	1.02	1.04	1.02	1.03
False Positive	3.12	3.6	0.64	4.74	0.43

Figure 6. Detected candidate object (cyan) and verified vehicle (green)

$$VS(i) = \sum_{i=y_h}^{y_h+H} \sum_{j=x_l}^{x_l+W} |G(i,j) - G(i,1)|$$
$$i_{sym} = \arg\max VS(i) \quad . \qquad (11)$$

The same idea is applied to horizontal symmetry detection with some modification.

$$HS(j) = \sum_{j=X_l}^{X_l+W} \sum_{i=Y_h}^{Y_h+H} \sum_{\Delta x=1}^{W/2} |G(i, j+\Delta x) - G(i, j-\Delta x)|$$
$$j_{sym} = \arg\min HS(j) \quad . \qquad (12)$$

III. Conclusion And Future Works

Using pattern recognition, the system can find vehicles. It is better than just using IGMM. In the future, we will test the system for application to a variety of images and optimize it to improve performance.

Acknowledgment

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1019) supervised by the IITP(Institute for Information & communications Technology Promotion)

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2010-0020163)

This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the C-ITRC(Convergence Information Technology Research Center) (IITP-2015-H8601-15-1003) supervised by the IITP(Institute for Information & communications Technology Promotion)

References

[1] Byeoung-su Kim and Whoi-Yul Kim, "Robust Lane Detection Method in Varying Road Conditions," *Journal of the Institute of Electronics Engineers of Korea*, vol. 49, no. 1, pp. 88-93, , 2012.

[2] Gangnet Michel, Didier Perny and Philippe Coueignoux, "Perspective mapping of planar textures," *Proceedings of Computers & Graphics*, vol. 8, no. 2, pp. 115-123, 1984.

[3] Zivkovic Zoran, "Improved adaptive Gaussian mixture model for background subtraction,", *Proceedings of the International Conference on Pattern Recognition*, vol. 2, pp. 28-31, 2004.

[4] Srinivasa Narayan, "Vision-based vehicle detection and tracking method for forward collision warning in automobiles," *Proceeding of the IEEE Intelligent Vehicle Symposium*, vol. 2, pp. 626-631, 2002.

[5] Jae Min Ban, Byeong-rae Lee and Hyunchul Kang, "Vehicle recognition using NMF in urban scene," *Journal of Korean Institute of Communications and Information Sciences*, vol. 37, no. 7, pp. 554-564, 2012.

[6] King Hann Lim and et al. "Lane-vehicle detection and tracking," *Proceedings of the International MultiConference of Engineers and Computer Scientists*, vol. 2, pp. 1124-1129, 2009.

[7] Y. Ma, S. Soatto, J. Kosecka, and A. A. Sastry, An Introduction to 3-D Vision From Images to Geometric Models, Springer,2004.