# Nighttime Vehicle Light Detection on a Moving Vehicle using Image Segmentation and Analysis Techniques 

YEN-LIN CHEN<br>Department of Computer Science and Information Engineering<br>Asia University<br>500 Liufeng Rd., Wufeng, Taichung 41354<br>TAIWAN<br>ylchen@asia.edu.tw


#### Abstract

This study proposes a vehicle detection system for identifying the vehicles by locating their headlights and rear-lights in the nighttime road environment. The proposed system comprises of two stages for detecting the vehicles in front of the camera-assisted car. The first stage is a fast automatic multilevel thresholding, which separates the bright objects from the grabbed nighttime road scene images. This proposed automatic multilevel thresholding approach provide the robustness and adaptability for the system to operate on various illuminated conditions at night. Then the extracted bright objects are processed by the second stage the proposed knowledge-based connected-component analysis procedure, to identify the vehicles by locating their vehicle lights, and estimate the distance between the camera-assisted car and the detected vehicles. Experimental results demonstrate the feasibility and effectiveness of the proposed approach on vehicle detection at night.


Key-Words: - Computer vision, vehicle detection, image segmentation, image analysis, multilevel thresholding, autonomous vehicles

## 1 Introduction

A vision-based for detecting the road environment for autonomous vehicle guidance is an emerging research area. Accordingly, many researchers have developed valuable techniques for recognizing interesting vehicles and obstacles from the images of the road environments [1]-[6], to facilitate applications on the camera-assisted system that helps the driver understanding the possible dangers on the road, and automatically controlling the facilities of the vehicles, such as low-beam and high-beam switching of the headlights. A visionbased vehicle and obstacle detection system is aiming at identification of vehicles, obstacles, traffic signs and other patterns on the road from grabbed image sequences by means of image processing and pattern recognition techniques. Until recently, researchers in this field still open new questions and concepts [7]-[9]. By adopting different concepts and definitions on interesting objects on the road, different techniques are applied on the grabbed image sequences to detect them as vehicles or obstacles. For locating the vehicles in the images, the task can be carried out by searching for specific patterns on the images based on the typical features of the vehicles, such as shape, symmetry, bounding boxes, and shadows [10]-[13].

However, until recently, most of these works focused on detecting the vehicles under the daytime road environment. Under bad-illuminated condition in the nighttime road environment, the obvious features of vehicles which are effective for detecting in daytime become invalid in nighttime road environment. Hence a detection approach for detecting vehicles at nighttime environment is practically a necessary demand for the autonomous camera-assisted car on night driving.

At night, as well as under dark illuminated condition in general, the only visual features of the vehicles are their headlights and rear-lights. Headlights and rear-lights are visible if a vehicle lies in the visible range of the camera. Besides, there are also many other illuminant sources coexisted with the vehicle lights in the nighttime road environment, such as street lamps, traffic lights, and road reflector plates on ground. Hence an efficient approach for identification and extraction of the actual vehicle lights from the grabbed night scene images is necessary for night driving of the camera-assisted car.

In this study, we propose an efficient vehicle detection system for identifying the vehicles by locating their headlights and rear-lights. This system comprises of two stages. The first stage applies a
fast automatic multilevel thresholding to separate the bright objects from the grabbed image sequences of nighttime road scene. The advantage of this proposed automatic multilevel thresholding approach is its robustness and adaptability on various illuminated conditions at night. Then the second stage applies a knowledge-based connectedcomponent analysis on the bright objects obtained by the first stage, for identifying the vehicles by locating their vehicle lights, and estimating the distance between the camera-assisted car and the detected vehicles. Figure 1 sketches the flow diagram of the proposed nighttime vehicle detection method. Experimental results demonstrate that the proposed system is feasible and effective on vehicle detection in the nighttime road environment.


Fig. 1. Block diagram of the proposed method

## 2 Extracting Bright Objects Using Automatic Multilevel Thresholding

The input image sequences grabbed from the vision system, which is mounted behind the windshield inside the car, consist of the nighttime environment shown in front of the car. The image sequences are grabbed with the $720 \times 480$ resolution with 24 -bit true colours. Figure 2 shows one sample nighttime road scene taken from the vision system. In this sample scene, there are two vehicles appeared on the road, where the left one is approaching in the opposite direction on the neighbouring lane, and the right one is moving in the same direction with the camera-assisted car. The left approaching car shows its bright headlights, while the front moving one shows its smaller and slightly gloomier rear-lights. In addition to the head and rear lights of the vehicles, some lamps, traffic lights and signs are also the visible illuminant appeared in the image sequences of the nighttime environment.


Fig. 2. An example of nighttime road environment

Hence, the first task is to extract these bright objects from the road scene image to facilitate further knowledge-based analysis. To save the computation cost on extracting bright objects, we firstly extracted the gray-intensity image, i.e. the $Y$ channel, of the grabbed image by performing a $R G B$ to $Y$ transformation. For extracting these bright objects from a given transformed gray-intensity image, bright objects must be separated from other objects with different illuminations. For this purpose,
the discriminant criterion, for measuring separability among the segmented images with different objects, is introduced in this section. The discriminant criterion, used on bi-level image thresholding was firstly presented by Otsu [14]. In Otsu's paper, the optimal threshold is determined by maximizing the between-class variance between dark and bright regions of the image. In this study, we extend and adopt the properties of discriminant analysis on multilevel thresholding. By evaluating the separability using the discriminant criterion, the number of objects, into which the image should be segmented, can be automatically determined. As a result, the bright objects will be appropriately extracted from other illuminated objects.

Let $f_{i}$ denote the observed occurrence frequencies (histogram) of pixels in a given image $\boldsymbol{I}$, with a given gray level $i$, and $N$ denotes the total amount of pixels in the image $\boldsymbol{I}$, and can be given by $N=$ $f_{0}+f_{1}+\ldots+f_{L-1}$, where $L$ is the number of gray values in the histogram. Hence, the normalized probability $P_{i}$ of one pixel having a given gray level $i$ can be denoted as,

$$
\begin{equation*}
P_{i}=f_{i} / N \text {, where } P_{i} \geq 0, \sum_{i=0}^{L-1} P_{i}=1 \tag{1}
\end{equation*}
$$

To segment homogeneous foreground objects and background components from the image $\boldsymbol{I}$, the pixels of the image $I$ must be partitioned into a suitable number of classes. For multilevel thresholding, with $k$ thresholds to partition the image into $k+1$ classes, pixels of the image $\boldsymbol{I}$ are segmented by applying a threshold set $T$, which is composed of $k$ thresholds, where $\boldsymbol{T}=\left\{t_{l}, \ldots, t_{n}, \ldots, t_{k}\right\}$. These classes are represented by $C_{0}=\left\{0,1, \ldots, t_{1}\right\}, \ldots$, $C_{n}=\left\{t_{n}+1, t_{n}+2, \ldots, t_{n+1}\right\} \quad, \ldots, \quad C_{k}=$ $\left\{t_{k}+1, t_{k}+2, \ldots, L-1\right\}$. The between-class variance, denoted by $v_{B C}$, an effective criterion for evaluating the results of segmentation, is utilized to measure the separability among all classes, and is expressed as,

$$
\begin{equation*}
v_{B C}(\boldsymbol{T})=\sum_{n=0}^{k} w_{n}\left(\mu_{n}-\mu_{T}\right)^{2} \tag{2}
\end{equation*}
$$

The within-class variance, denoted by $v_{W C}$, of all segmented classes of pixels is computed as,

$$
\begin{equation*}
v_{W C}(\boldsymbol{T})=\sum_{n=0}^{k} w_{n} \sigma_{n}^{2} \tag{3}
\end{equation*}
$$

The total variance $v_{T}$ and the overall mean $\mu_{T}$ of pixels in the image $I$ are computed as,

$$
\begin{equation*}
v_{T}=\sum_{i=0}^{L-1}\left(i-\mu_{T}\right)^{2} P_{i}, \text { and } \mu_{T}=\sum_{i=0}^{L-1} i P_{i} \tag{4}
\end{equation*}
$$

where $k$ is the number of selected thresholds to segment pixels into $k+1$ classes; $w_{n}$ is the cumulative probability mass function of class $C_{n} ; \mu_{n}$ and $\sigma_{n}$ represent the mean and the standard deviation of pixels in class $C_{n}$, respectively. They are defined as,

$$
\begin{gather*}
w_{n}=\sum_{i=t_{n}+1}^{t_{n+1}} P_{i}, \mu_{n}=\frac{\sum_{i=t_{n}+1}^{t_{n+1}} i P_{i}}{w_{n}}, \text { and } \\
\sigma_{n}^{2}=\frac{\sum_{i=t_{n}+1}^{t_{n+1}} P_{i}\left(i-\mu_{n}\right)^{2}}{w_{n}} \tag{5}
\end{gather*}
$$

where a dummy threshold $t_{0}=0$ is utilized for simplifying the expression of equation terms.

The aforementioned criterion functions can be considered a measure of separability, among all existing classes, decomposed from the original image $\boldsymbol{I}$. We introduce this concept as a criterion of automatic image segmentation, denoted by the "separability factor" - $S F$ in this study, which is defined as,

$$
\begin{equation*}
S F=\frac{v_{B C}(\mathbf{T})}{v_{T}}=1-\frac{v_{W C}(\mathbf{T})}{v_{T}} \tag{6}
\end{equation*}
$$

where $v_{T}$ is the total variance of the gray-level values of the image $\boldsymbol{I}$ and serves as the normalization factor in this equation. The $S F$ value measures the separability among all existing classes, and the $S F$ value lies within the range $0 \leq S F \leq 1$. Maximizing the $S F$ value is the objective to optimize the segmentation result. Through observation of the terms comprising $v_{W C}(\boldsymbol{T})$, if the pixels in each class are broadly spread, i.e. the contribution of the class variance $\sigma_{n}{ }^{2}$ is large, then the corresponding $S F$ measure becomes small. Hence, when $S F$ approaches 1.0, all classes of gray levels decomposed from the original image $\boldsymbol{I}$ are ideally and completely separated. This property also satisfies the concept of uniformity of the segmented regions, as presented by Levein and Nazif [15].

Accordingly, based on this efficient discriminant criterion on measuring the separability of the object regions of homogenous gray levels, an automatic multilevel thresholding method can be developed to recursively segment homogeneous objects from the image $\boldsymbol{I}$, regardless of the number of objects and the complexity of the image. The multilevel thresholding process can be recursively performed on the gray levels of the image $\boldsymbol{I}$ until the $S F$ measure is large enough, i.e. $S F$ approaches 1.0 , to reflect that the appropriate discrepancy among the resultant classes of gray levels is achieved so that the homogeneous objects are completely segmented into separate thresholded images.

Through the aforementioned properties, this objective can be reached by minimizing the total within-class variance $v_{W C}(\boldsymbol{T})$. This can be achieved by the scheme that selects the class with the maximal contribution ( $w_{n} \sigma_{n}{ }^{2}$ ) of the total withinclass variance for performing the bi-level thresholding procedure to partition it into two more classes in each recursion. Thus, the SF measure will most rapidly achieve the maximal increment to satisfy sufficient separability among the resultant classes of gray levels. Furthermore, objects with homogeneous gray levels, will be well-separated.

Based on the above definitions, a new automatic multilevel thresholding method is developed. The details of the proposed method are presented below.

Step 1: In the beginning, compute the histogram of gray values of the image $\boldsymbol{I}$, and all the gray values in $\boldsymbol{I}$ are assigned to one initial class $C_{0}$. Let $q$ denote the number of currently determined thresholds in the threshold set $\boldsymbol{T}$, which classify the gray values into $q+1$ classes; Initially, $\boldsymbol{T}$ comprise of no thresholds and $q=0$.

Step 2: In current recursion, $q$ thresholds have been determined, i.e. $\boldsymbol{T}=\left\{t_{l}, \ldots, t_{n}, \ldots, t_{q}\right\}$, which partition the gray values of the image $\boldsymbol{I}$ into $q+1$ classes $\left(C_{0}, \ldots, C_{n}, \ldots, C_{q}\right)$. Compute the class-mean $\mu_{n}$, the cumulative probability mass function $w_{n}$, and the standard deviation $\sigma_{n}$ of each existing class $C_{n}$ using Eq. (5), respectively, where $n$ denotes the index of the present classes and $n=0 \sim q$.

Step 3: From all classes $C_{n}$, determine the class $C_{p}$ with the maximal contribution ( $w_{n} \sigma_{n}{ }^{2}$ ) of the total within-class variance $v_{W C}$, which is to be partitioned in the following step to achieve the maximal
increment of $S F$.
Step 4: Determine the optimal threshold $t_{s}^{*}$ to partition $C_{p}:\left\{t_{p}+1, t_{p}+2, \ldots, t_{p+1}\right\}$ into two classes $C_{p 0}$ and $C_{p l}$ which comprise the subsets of gray values decomposed from $C_{p}$. The $t_{s}^{*}$ is obtained by maximizing the between-class variance $v_{B C}^{\prime}$ of partitioned $C_{p 0}$ and $C_{p 1}$ with respect to $t_{s}$, and is computed as,

$$
\begin{gather*}
v_{B C}^{\prime}\left(t_{S}^{*}\right)=\operatorname{Max}_{t_{p}<t_{s} \leq t_{p+1}} v_{B C}^{\prime}\left(t_{S}\right)  \tag{7}\\
v_{B C}^{\prime}\left(t_{S}\right)=w_{p 0}\left(\mu_{p 0}-\mu_{p}\right)^{2}+w_{p 1}\left(\mu_{p 1}-\mu_{p}\right)^{2},  \tag{8}\\
w_{p 0}=\sum_{i=t_{p}+1}^{t_{s}} P_{i}, w_{p 1}=\sum_{i=t_{s}+1}^{t_{p+1}} P_{i}  \tag{9}\\
\mu_{p 0}=\sum_{i=t_{p}+1}^{t_{s}} i P_{i} / w_{p 0}, \mu_{p 1}=\sum_{i=t_{s}+1}^{t_{p+1}} i P_{i} / w_{p 1} \tag{10}
\end{gather*}
$$

where $w_{p}$ and $\mu_{p}$ are the class-probability and class-mean of $C_{p}$, respectively.

Then $t_{S}^{*}$ is put into the threshold set $\boldsymbol{T}$, and is applied to partition $C_{p}$ into $C_{p 0}$ and $C_{p l}$, i.e. $C_{p 0}:\left\{t_{p}+1, t_{p}+2, \ldots, t_{s}^{*}\right\}$, and $C_{p l}:\left\{t_{s}^{*}+1, t_{S}^{*}+2, \ldots, t_{p+1}\right\}$. Hence the gray values of the image $\boldsymbol{I}$ are then divided into $q+1$ classes. Then re-label all the thresholds in $\boldsymbol{T}$ and let $q=q+1$.

Step 5: Computed the $S F$ measure of all currently obtained classes using Eq. (6). If the following "Objective Condition" is satisfied,

$$
\begin{equation*}
S F \geq T h_{S F} \tag{11}
\end{equation*}
$$

then go to Step 6; otherwise, go back to Step 2 to perform further partition process on the obtained classes.

Step 6: Apply the largest threshold value $t_{q}$ to extract the brightest class of pixels $C_{q}$, which contains the bright objects of interest, into a separate bright object plane; and terminate the thresholding procedure.

This work employs $T H_{S F}=0.9$, determined from the training, using numerous images, such that all existing classes are almost completely separated. As a result, the bright objects of interest are extracted into an individual thresholded bright object plane, as shown in Fig. 3.


Fig. 3. Extracted bright objects by automatic multilevel thresholding

## 3 Projection-based Component Analysis on Bright Objects

To obtain the vehicle-light-like components from the bright object plane, a connected-component extraction technique [16] is then performed on the bright object plane to locate the connectedcomponents of the bright objects. By extracting the connected-components, the location and dimension of each connected-component are obtained as well.

The location and dimension of a connectedcomponent are represented by the bounding box which encloses it. We are interested in looking for the horizontal-aligned vehicle lights; hence a spatial clustering process is applied on the connectedcomponents to cluster them into several groups. A resultant group is comprised of a set of connectedcomponents, and may be vehicle-lights, traffic lights, road signs, and some other illuminant objects in the nighttime environment. These connected-component groups are then processed by the vehicle light identification process to obtain the real moving vehicles.

The definitions used in the text detection and extraction process are:
(a). $C_{i}$ denotes a connected-component of the current, and the bounding box which encloses $C_{i}$ is denoted as $B_{i}$.
(b). $G_{j}$ denotes a group of connected-components, $G_{j}=\left\{C_{i}, i=0,1,2, \ldots, p\right\}$, the number of connectedcomponents contained in $G_{j}$ is denoted as $N_{c c}\left(G_{j}\right)$.
(c). The location of the bounding boxes of $C s$ employed in the spatial clustering process are their top, bottom, left and right coordinates, and they are denoted as $t\left(B_{i}\right), b\left(B_{i}\right), l\left(B_{i}\right)$, and $r\left(B_{i}\right)$, respectively.
(d). The width and height of the bounding boxes are denoted as $W\left(B_{i}\right)$ and $H\left(B_{i}\right)$, respectively.
(e). The horizontal distance $D_{h}$ and vertical distance $D_{v}$ between two bounding boxes are defined as,

$$
\begin{align*}
& D_{h}\left(B_{i}, B_{j}\right)= \\
& \quad \max \left[l\left(B_{i}\right), l\left(B_{j}\right)\right]-\min \left[r\left(B_{i}\right), r\left(B_{j}\right)\right] \tag{12}
\end{align*}
$$

$$
\begin{align*}
& D_{v}\left(B_{i}, B_{j}\right)= \\
& \quad \max \left[t\left(B_{i}\right), t\left(B_{j}\right)\right]-\min \left[b\left(B_{i}\right), b\left(B_{j}\right)\right] \tag{13}
\end{align*}
$$

If the two bounding boxes are overlapping in the horizontal or vertical direction, then the value of the $D_{h}\left(B_{i}, B_{j}\right)$ or $D_{v}\left(B_{i}, B_{j}\right)$ will be a negative value.
(f). Hence the measure of overlapping between the vertical projections of the two bounding boxes can be defined as,

$$
\begin{align*}
& P_{v}\left(B_{i}, B_{j}\right)= \\
& -D_{v}\left(B_{i}, B_{j}\right) / \min \left[H\left(B_{i}\right), H\left(B_{j}\right)\right] \tag{14}
\end{align*}
$$



Fig. 4. The processing area determined by the virtual horizon and the labelled connectedcomponents

To preliminarily screen out the non-vehicle-light objects such as street lamps and traffic lights, we firstly filter out the bright connected-components which are located above the one-third of the vertical y -axis, i.e. only the bright connected-components located under the constraint line as shown in Fig. 4 will be taken into account. This is because the vehicles which are located at the distant place on the road become very small light "points", and will "converge" into a virtual horizon. Hence we utilize the constraint line in Fig. 4 as this virtual horizon.


Fig. 5. The spatial clustering of bright components

Based on the functions and notation defined above, the connected-components of bright objects are recursively merged and clustered into connected-component group Gs if they are horizontally close to each other, vertically overlapped and well-aligned. In other words, if the two neighboring connected-components satisfy the following conditions, they are merged with each other and clustered as the same group $G$ :

1) They are horizontally close to each other, i.e.:

$$
\begin{equation*}
D_{h}\left(B_{1}, B_{2}\right)<2.0 \times \max \left(H\left(B_{1}\right), H\left(B_{2}\right)\right) \tag{15}
\end{equation*}
$$

2) They are highly overlapped in vertical projection profiles, i.e.:

$$
\begin{equation*}
P_{v}\left(B_{1}, B_{2}\right)>0.8 \tag{16}
\end{equation*}
$$

3) They have similar heights, i.e.:

$$
\begin{equation*}
H\left(B_{s}\right) / H\left(B_{l}\right)>0.8 \tag{17}
\end{equation*}
$$

where $B_{s}$ is the shorter among the two bounding boxes $B_{1}$ and $B_{2}$, and $B_{l}$ is the larger one.

Figure 5 illustrates the spatial clustering process of bright connected-components. After performing the spatial clustering process, several groups of bright components are obtained, and they are called candidate vehicle light groups. In Fig. 5, the meaningful bright components are grouped into two candidate vehicle light groups, the left one contains the headlights of the approaching vehicle, and the right one contains the rear-lights of the front vehicle. Then a rule-based vehicle-light identification process stated in the following section is conducted on these candidate groups to extract actual vehicle lights.

## 4 Rule-Based Vehicle Light Identification

Then the rule-based identification process is utilized to determine whether the candidate vehicle light groups contain actual vehicle lights or other illuminated objects based on the statistical features of their contained bright components.

If one candidate group is an actual vehicle light set, then,

1) The enclosing bounding box of the candidate group must form a horizontal oblong shape, i.e. the
ratio of the width $W$ and the height $H$ of the enclosing box of the candidate group must satisfy the condition,

$$
\begin{equation*}
W / H \geq 2.0 \tag{18}
\end{equation*}
$$

2) Its contained bright components should be wellaligned, i.e. the ratio of the total area of contained bright components of the candidate group to the area of its enclosing box must satisfy the condition,

$$
\begin{equation*}
0.4 \leq \sum_{C_{i} \in G} A\left(C_{i}\right) /(W \times H) \leq 0.95 \tag{19}
\end{equation*}
$$

where $A\left(C_{i}\right)$ is the area of the bounding box of the $i$-th contained bright component of the candidate group.
3) The number of these bright components should also be within a reasonable range, because the vehicle lights are mostly appeared in pairs, and some types of compound light set may comprise of at most four lights, so that the light number condition is defined as,

$$
\begin{equation*}
2 \leq N_{c c} \leq 4 \tag{20}
\end{equation*}
$$

The above-mentioned decision rules are obtained by analyzing many experimental results of processing image sequences of the real road environment having vehicle lights with various types, directions and distances. The constant values utilized under the above decision rules are determined experimentally and yield good performance in most general cases.

Besides, to determine the moving directions of the detected vehicles, we need to distinguish the detected vehicle lights into headlights and rearlights. Since the headlights have much more variation in colours and sizes then those of the rearlights, so that we can utilize some typical characteristics of rear-lights to distinguish them from head-lights. The typical characteristic of the rear-lights is that they mostly are red illuminated lights. Hence we utilize a criterion to check if the candidate vehicle light group is a rear-light set, i.e.

$$
\begin{equation*}
R_{a}-8>\text { both } G_{a} \text { and } B_{a}, \tag{21}
\end{equation*}
$$

where $R_{a}, G_{a}$, and $B_{a}$ denote the average intensities of the $R, G$, and $B$ channels of the pixels comprised
of the bright components contained in the candidate vehicle light group.

## 5 Vehicle Distance Estimation

To estimate the distance between the cameraassisted car and detected vehicles, we apply the perspective projection model of the CCD camera as introduced in Betke et. al's formulation model [11]. The origin of the virtual vehicle coordinate system is placed at the central point of the camera lens. The $X$ and $Y$-coordinate axes of the virtual vehicle coordinate are parallel to the $x$ and $y$-coordinates of the grabbed image, and the $Z$-axis is placed along the optical axis and perpendicular to the plane formed by the $X$ and $Y$ axes. In the nighttime environment, only the vehicle lights can be detected, so that only the width of the vehicles can be properly estimated by measuring the width covered by their vehicle light pairs. We utilize the perspective projection equation in [11], and hence the $Z$-distance in meters between the cameraassisted car and one detected vehicle can be obtained by,

$$
\begin{equation*}
Z=k_{\text {hor }} \cdot f \cdot(W / w) \tag{22}
\end{equation*}
$$

where $k_{h o r}$ is a given factor for converting from pixels to millimetres for the CCD camera; $f$ is the focal length in millimetres; $w$ is the width in pixels on the image; and $W$ is the estimated width of a typical car in Taiwan, i.e. 1.8 meters.

## 6 Experimental Results

In this section, we describe the implementation of the proposed method on our experimental cameraassisted car, and conduct various representative realtime experiments to make the performance evaluation of the proposed method.
The proposed system is implemented on a Pentium-4 2.4 GHz platform which is set up on our experimental camera-assisted car. The vision system for acquiring input image sequences of road environments, as shown in Fig. 6, is mounted behind the windshield inside the experimental camera-assisted car. The frame rate of the vision system is 30 frames per second and the size of each frame of grabbed image sequences is 720 pixels by 480 pixels per frame.


Fig. 6. The vision system mounted in the experimental car

The proposed system has been tested on several videos of real nighttime road scenes in various conditions. Figures 7 - 9 exhibit the most representative ones of the experimental samples on performance evaluation. As shown in Fig. 7, the oncoming vehicle is correctly detected by locating its headlight pair, although some other non-vehicle illuminated objects also coexist with the vehicle in this scene. The distance between this oncoming vehicle and the camera-assisted car is estimated as about 21 meters by the proposed system, which is close to the actual distance obtained by manual measurement.


Fig. 7. Result of vehicle detection on the nighttime road scene with one oncoming vehicle


Fig. 8. Result of vehicle detection on the nighttime road scene with oncoming and preceding vehicles

Figure 8 exhibits a sample of the condition when two vehicles appeared in the scene. Here the headlight set comprised of four headlights at the left is determined as an oncoming vehicle, and its distance to the experimental car is estimated as 9.4 meters. Although some other illuminant objects are coexisted with the taillights of the right vehicle on its body, the taillight pair is still correctly located and identified as a preceding vehicle, and its distance is estimated as 10.7 meters.


Fig. 9. Result of vehicle detection on the nighttime road scene comprised of vehicles and many other non-vehicle lights

As shown in Fig. 9, a more complicated scene is illustrated. The vehicle lights of the two vehicles are very close to each other in this scene. As well as a series of lamps appears above the left oncoming vehicle and many small illuminated light objects occur above and near to the right preceding vehicle. Although interfered by many non-vehicle illuminant objects in this scene, the proposed method still successfully detect these two vehicles by locating their vehicle light pairs. The distances of the left oncoming vehicle and the right preceding vehicle to the experimental car are estimated as about 23 meters and 10 meters, respectively.
The computation time spent on processing one input frame depends on the road scene complexity of the frame. Most of the computation time is spent on the connected-component analysis and the projection-based spatial clustering process on bright objects. For an input video sequence with $720 \times 480$ pixels per frame, the proposed system takes an average of 16 milliseconds processing time per frame. Thus, this frugal computation cost ensures that the proposed system can effectively satisfy the demand of real-time processing with 30 frames per second. From the results of our numerous outdoor tests on many different road environments at night, the system shows that it can provide fast, real-time, and correct vehicle detection results to facilitate the autonomous camera-assisted car on night driving.

## 7 Conclusions

This study has presented an efficient vehicle detection system for identifying the vehicles by locating their headlights and rear-lights in the nighttime road environment. By extracting the bright objects using the automatic multilevel thresholding approach, the interesting illuminated objects including vehicle lights are extracted. Then these bright objects are grouped according to their location and size features using a connectedcomponent spatial clustering process. Hence the vehicle lights are identified using their typical characteristics, such as pairing and alignment, from the grouped bright objects. Experimental results show that the proposed system can effectively detect the vehicles by locating their vehicle lights in the nighttime road environment.

## 8 Acknowledgements

This paper was supported by the National Science Council of R.O.C. under Contract No. NSC-97-2218-E-468-007, and NSC-96-2622-E-468-003-

CC3, and Asia University under Contract No. 97-I01, and 97-I-06.

## References:

[1] I. Masaki (Ed.), "Vision-based Vehicle Guidance", Sppringer-Verlag, New York, 1992.
[2] M. Maurer, R. Behringer, S. Furst, F. Thomarek, and E.D. Dickmanns, A compact vision system for road vehicle guidance", in Proc. $13^{\text {th }}$ Int'l Conf. Patt. Recognit., Vol. 3, 1996, pp. 313-317.
[3] M. Bertozzi and A. Broggi, "Vision-based vehicle guidance", IEEE Comput., Vol. 30, 1997, pp. 49-55.
[4] A. Broggi, M. Bertozzi, A. Fascioli, G. Conte, "Automatic Vehicle Guidance: The Experience of the ARGO Autonomous Vehicle", World Scientific, Singapore, 1999.
[5] H. Zheng, "Morphological neural network vehicle detection from high resolution satellite imagery", WSEAS Trans. Computers, Vol. 5, 2006, pp. 2225-2231.
[6] J. Janta, P. Kumsawat, K. Attakitmongkol, A. Srikaew, "A pedestrian detection system using applied log-Gabor", Proc. 7th WSEAS Int'l Conf. Signal, Speech and Image Processing, 2007, pp. 55-60.
[7] A. Broggi, M. Bertozzi, A. Fascioli, C.G.L. Bianco, A. Piazzi, "Visual perception of obstacles and vehicles for platooning", IEEE Trans. Intell. Trans. Syst., Vol. 1, 2000, pp. 164-176.
[8] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, C. Pocol, R. Schmidt, T. Graf, "High accuracy stereo vision for far distance obstacle detection", in Proc. IEEE Intell. Vehicle Symp., 2004, pp. 292-297.
[9] Z. Sun, G. Bebis, N. Bourbakis, "Decisionlevel fusion for vehicle detection", Proc. 11th WSEAS Int'l Conf. Computers, 2007, pp. 624629.
[10] U. Franke, and S. Heinrich, "A study on recognition of road lane and movement of vehicles using vision system", SICE, Nagoya, 2001, pp. 38-41.
[11] M. Betke, E. Haritaoglu, and L. S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle", Mach. Vis. Appl., Vol. 12, 2000, pp. 69-83.
[12] M.Y. Chern and B.Y. Shyr, "Detecting vehicles on highway from the driver's front view", Proc. $15^{\text {th }}$ Conf. Comput. Vis., Graph., Image Process., 2002, pp. 779-786.
[13] M. Krips, A. Teuner, J. Velten, A. Kummert, "Camera based vehicle detection and tracking using shadows and adaptive template matching", Proc. 2nd WSEAS Int'l Conf. Electronics, Control and Signal Processing, 2003.
[14] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Trans. Sys., Man, Cybern., Vol. SMC-9, 1979, pp. 62-66.
[15] M.D. Levine, and A.M. Nazif, "Dynamic measurement of computer generated image segmentation", IEEE Trans. Patt. Anal. Mach. Intell., Vol. 7, 1985, pp. 155-164.
[16] R. C. Gonzalez and R. E. Woods, "Digital Image Processing, 2/e", New Jersey: Prentice Hall, 2002.

