Journal of Southeast University (English Edition) Vol. 24, No. 4, pp. 517 - 521 Dec. 2008 ISSN 1003—7985

Traffic light detection and recognition in intersections based on intelligent vehicle

Zhang Ning He Tiejun Gao Zhaohui Huang Wei

(ITS Research Center, Southeast University, Nanjing 210096, China)

Abstract: To ensure revulsive driving of intelligent vehicles at intersections, a method is presented to detect and recognize the traffic lights. First, the stabling siding at intersections is detected by applying Hough transformation. Then, the colors of traffic lights are detected with color space transformation. Finally, self-associative memory is used to recognize the countdown characters of the traffic lights. Test results at 20 real intersections show that the ratio of correct stabling siding recognition reaches up to 90%; and the ratios of recognition of traffic lights and divided characters are 85% and 97%, respectively. The research proves that the method is efficient for the detection of stabling siding and is robust enough to recognize the characters from images with noise and broken edges.

Key words: intelligent vehicle; stabling siding detection; traffic lights detection; self-associative memory; light-emitting diode (LED) characters recognition

Intelligent vehicles perceive the circumference using video systems with cameras, sensors and computers for finally realizing video revulsive driving[1]. For such purposes, it is necessary to detect the scene of the roadway, including the road signs and obstacles[2-3]. Among them, the detection of stabling siding and traffic lights at intersections is of significant importance in determining proper driving behavior according to the conditions of traffic lights and stabling siding.

Currently, the most famous intelligent vehicle products are the NavLab series[4] developed by CMU (Carnegie Mellon University) Robot Research Center. The NavLab series used video image processing technologies for computing the distance of a vehicle deviating from the lane centerline, and detecting road obstructions. CMU also developed a Ranger intelligent vehicle that could be applied to wild environments. In addition, the U. S. military developed DEMO Ⅲ combining various intelligent vehicle technologies, including CCD, laser, radar, ultrasound, infrared, microwave and so on[5], which could be applied to the daytime, nighttime and other complex weather and roadway environments (e. g., rain, a dirty road). Moreover, the DEMO Ⅲ can detect and avoid irregular obstacles. The University of Wisconsin at Madison used the regional road color information to separate images, and used Hough transformation to detect the edges of the roadway. The University of Parma, Italy developed the ARGO (a Lancia Thema passenger car)[6], as a critical part of the PROMETHEUS (Program for Europe Traffic with Highest Efficiency and Unprecedented Safety) project using the GOLD (generic obstacle and lane detection) system for detecting the obstacles in front of a vehicle and locating the vehicle using the symmetric characteristics of edge gray values around vehicles and the roadway edge structure.

Compared with the recent studies, researches on intelligent vehicles in China are far behind in applying self-revulsive driving techniques. Although Tsinghua University and Jilin University have made rapid progress in this field[7], current researches mainly focus on extracting the lane line, detecting the obstacles, and detecting road signs. The detection of traffic lights is still limited to red light running detection at a fixed scene. As a result, the detection is merely designed for conventional tri-color circular traffic lights[8]. According to the authors' best knowledge, little investigation has been put into detecting the currently and frequently applied LED tri-color digital traffic lights. Therefore, it is meaningful and important to study the detection and recognition techniques of video-based digital traffic lights at intersections.

For simplicity, some hypotheses are made in this paper for actual intersections. These hypotheses include: 1) There are no obstacles, such as other vehicles and pedestrians; 2) There is only one group of traffic lights; and 3) The traffic revulsive driving is straight. Therefore, for an intersection with traffic lights where the stabling sidings are perpendicular to lanes, whether the vehicles approach the intersection or not is judged through the detection of the stabling sidings in this paper. If there is intersection stabling siding, the detection of traffic lights is carried out, and the LED character is recognized with self-associative memory[9]. According to the recognition results, proper driving actions are triggered, such as accelerating the vehicle to pass the intersections or decelerating the vehicle to stop at a proper location behind the stabling sidings.

1 Detection of Stabling Siding with Hough Transformation

Hough transformation is usually used for straight detection, by transforming the spatial domain into a parametric domain. The curve of the images is first depicted with some proper parametric points. Then, the necessary information is statistically obtained. Finally, a satisfactory result can be obtained even under the background of noise and deformation. The usually applied parametric equation of Hough transformation is given as

$$\rho = x\cos(\theta) + y\sin(\theta) \tag{1}$$

where ρ is the distance from the present coordinate to the or-

Received 2008-03-21.

Biography: Zhang Ning (1972—), male, doctor, associate professor, ningzhang1972@yahoo.com.cn.

Foundation item: The Cultivation Fund of the Key Scientific and Technical Innovation Project of Higher Education of Ministry of Education (No. 705020).

Citation: Zhang Ning, He Tiejun, Gao Zhaohui, et al. Traffic light detection and recognition in intersections based on intelligent vehicle[J]. Journal of Southeast University (English Edition), 2008, 24(4): 517 - 521.

igin, and θ is the angle between the straight line and the X-axis.

Based on this, the Hough transformation of a point in a spatial domain corresponds to the curve in the ρ-θ domain. Consequently, the collinear points of the spatial domain transforming to the ρ-θ domain will be a cluster of concurrent curves. Moreover, the parameter of the point with a maximum deposited value is the parameter of the corresponding straight line.

During the experiment, a camera with maximum pixels of 640×480 is employed. Fig. 1(a) shows an image before the intersection, where the stabling detection domain is defined as a vertical distance of 80 pixels in front of the vehicle head in the image, which is about 4 m in the actual scene. In order to extract the stabling siding from the images, a binary process for current frame image is carried out for extracting the white stabling siding in an RGB space in this paper. For this, a threshold T is set in the three sub-spaces of R, G and B as 190. If the values of the pixel point on the three subspaces are larger than T, the point will be set as 1; otherwise, it will be set as 0, as shown in Fig. 1(b). Because traditional methods for edge detection result in many edge lines[10] and bring more problems for the transformation and detection, the skeleton extraction in the morphology is used for thinning the stabling siding in the binary image, as shown in Fig. 1(c). The Hough transformation result of the line in Fig. 1(c) is shown in Fig. 1(d). If the maximum value of the deposited detection values (the brightest point in Fig. 1(d)) has a θ range of 80 to 100, it is believed to be stabling siding. Since the stabling siding is vertical to the driveway, the θ of the stabling siding is nearly 90°. The distance from the stabling siding to the vehicle head can be estimated as the distance of the pixel distance.

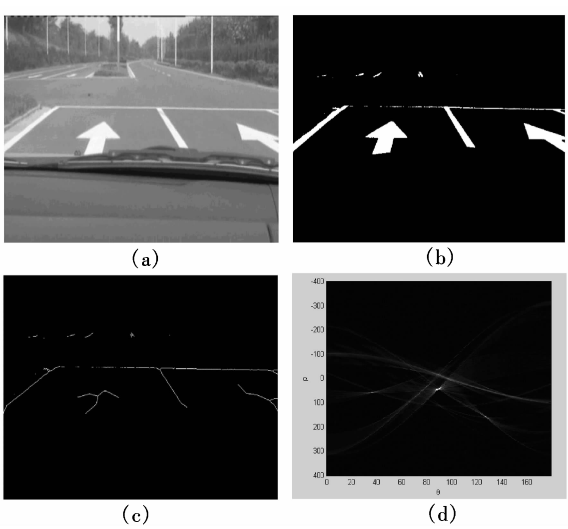

Fig. 1 Stabling siding detection process. (a) Scene image; (b) Binary result; (c) Thinning result; (d) Hough detection result

2 Detection of Traffic Lights

2.1 Transformation for RGB space to HSI space

The image color space obtained by a camera from the scene is defined as an RGB model, i. e., the integrants of R (red), G (green) and B (blue) values. For this, a color spatial model HSI (hue, saturation, and intensity, respectively) is used for image processing. Since the HSI space consists of independent H, S and I, the special traffic light color is searched for. The transformation of the RGB images to the HSI images[11] is given as

$$H = \begin{cases} \theta & B \leqslant G \\ 360° - \theta & B > G \end{cases} \tag{2}$$

$$\theta = \arccos \frac{\frac{1}{2}[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}} \tag{3}$$

$$S = 1 - \frac{\min(R, G, B)}{I} \tag{4}$$

$$I = \frac{R+G+B}{3} \tag{5}$$

where θ is a middle variable; H is the hue to describe the pure color property; S is the saturation to describe the degree of hue diluted by the white light; and I is the intensity showing the brightness of color, which is independent of H. Thus, the color of the images is described with the integrants of H and S.

2.2 Position acquirement of RYG traffic lights

The most obvious characteristics of the traffic lights from an environmental scene are their intensity and their colors of red, yellow and green. The position of traffic lights was acquired through detecting these two characteristics in an HSI space[12].

Through the color distribution in an HSI space, it is known that the defined values of hues at red, yellow and green are 0, $\pi/3$ and $2\pi/3$, respectively. Consequently, the positions of red, yellow and green traffic lights are detected through the following three equations:

Red zone

$$R = \{f_t(x, y): |f_{t,H}(x, y)| < t_h \text{ and } f_{t,H}(x, y) > t_v\} \tag{6a}$$

Yellow zone

$$Y = \left\{f_t(x, y): \left|f_{t,H}(x, y) - \frac{\pi}{3}\right| < t_h \text{ and } f_{t,H}(x, y) > t_v\right\} \tag{6b}$$

Green zone

$$G = \left\{f_t(x, y): \left|f_{t,H}(x, y) - \frac{2\pi}{3}\right| < t_h \text{ and } f_{t,H}(x, y) > t_v\right\} \tag{6c}$$

Here, the given detection color domain is determined by the threshold t_h, which is set as $\pi/6$; t_v is the threshold of intensity (0.5 in this paper) that is dependent on the LED traffic lights.

Fig. 2 shows the detection process of the red traffic light, where Fig. 2(a) is the scene image of the intersection. After transforming the image into an HSI space, Fig. 2(b) is obtained through binary image processing by Eq. (6a) where two characteristic pixels are detected. Since the position of a traffic light is relatively high in Fig. 2(b), the coordinate searching can be determined by one third of the image from the upside for effectively decreasing the error and improving calculation efficiency. After acquiring the coordinates, the image of the traffic light can be separated from Fig. 2(a), as shown in Fig. 2(c). For the yellow and green traffic lights, the detection is similar to that for the red lights.

Fig. 2 Detection of red traffic lights. (a) Intersection traffic light; (b) Traffic light detection; (c) Character extraction

3 Recognition of Time Characters of Traffic Light

3.1 Characters segmentation and normalization

According to the displayed traffic lights in Fig. 2(c), the time of the traffic lights is found having two digital characters. After extracting the traffic light position, projection segmentation is applied for ensuring that each recognition is only for one character. Because separated characters have different sizes, normalization is carried out on the separate characters. In this paper, each separate character is normalized into 90×50 pixels by bilinear interpolation.

3.2 Character feature extraction

In order to facilitate the feature extraction and obtain the weights for the neural network for later self-associative memory of the character recognition, the normalization of characters are divided into 9×5 lattice blocks[13] and each lattice block is set as 10×10 pixels. If the pixel value is 1, each lattice block's pixels are accumulated. If the cumulative value is more than 40, the eigenvalue of this lattice block is set as 1, and otherwise it is set as 0. Fig. 3 shows that the character "2" is the image after the character feature extraction, in which black represents that the eigenvalue is 1 and white 0. Therefore, the corresponding vector is given as $\boldsymbol{P}^{\mathrm{T}} = [0\ 1\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ \ldots\ 1\ 1\ 1\ 1\ 0]$.

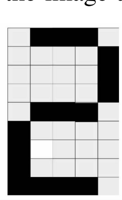

Fig. 3 The eigenvalue of lattice block

3.3 Self-associative memory of character recognition

For the self-associative memory of character recognition, almost all learning of the neural network could be considered as a transformation of the Hebb learning rules, where the expected outputs are equal to the inputs of the neural network. Even when the characters are broken and noised, they still can be recognized correctly by storing a group of standard characters[9].

The rules of the supervised Hebb is described as

$$\boldsymbol{W}_q^{\mathrm{new}} = \boldsymbol{W}_q^{\mathrm{old}} + \boldsymbol{t}_q \boldsymbol{P}_q^{\mathrm{T}} \tag{7}$$

where $\boldsymbol{W}_q^{\mathrm{new}}$ and $\boldsymbol{W}_q^{\mathrm{old}}$ are the weight matrices, $\boldsymbol{t}_q$ is the q-th target vector, and $\boldsymbol{P}_q$ is the q-th input vector. If the weight matrices are initialized to 0 and the Q input/output pairs are applied once to Eq. (7), Eq. (7) can be described as

$$\boldsymbol{W} = \boldsymbol{t}_1 \boldsymbol{P}_1^{\mathrm{T}} + \boldsymbol{t}_2 \boldsymbol{P}_2^{\mathrm{T}} + \ldots + \boldsymbol{t}_Q \boldsymbol{P}_Q^{\mathrm{T}} = \sum_{q=1}^{Q} \boldsymbol{t}_q \boldsymbol{P}_q^{\mathrm{T}} = \boldsymbol{T}\boldsymbol{P}^{\mathrm{T}} \tag{8}$$

When the input vector is non-orthogonal, the Hebb rules will have errors. We use a pseudoinverse rule to reduce such errors. The pseudoinverse rule is given as

$$\boldsymbol{W} = \boldsymbol{T}\boldsymbol{P}^{+} \tag{9}$$

where $\boldsymbol{P}^{+}$ is the Moore-Penrose pseudoinverse. When the rows number of $\boldsymbol{P}$ is greater than the columns number, and the column vectors of $\boldsymbol{P}$ are linearly independent, then the pseudoinverse can be computed by

$$\boldsymbol{P}^{+} = (\boldsymbol{P}^{\mathrm{T}}\boldsymbol{P})^{-1}\boldsymbol{P}^{\mathrm{T}} \tag{10}$$

In training the network weights matrix as a self-associative memory, the input and output of feature vectors are a standard character template, that is, $\boldsymbol{T} = \boldsymbol{P}$. So the weight matrix can be expressed as

$$\boldsymbol{W} = \boldsymbol{P}(\boldsymbol{P}^{\mathrm{T}}P)^{-1}\boldsymbol{P}^{\mathrm{T}} \tag{11}$$

4 Experimental Analyses

The experiment consists of three parts: 1) Detection of stabling siding; 2) Detection of traffic lights; 3) LED character recognition where the detections of the stabling siding and traffic lights have been addressed in the above sections. Therefore, only LED character recognition is detailedly described in this section as shown below.

Whether a vehicle is approaching the intersection is determined through detecting the stabling siding within a given zone. If the vehicle is in the intersection, the traffic lights are detected, and the character image is separated. In the LED traffic light, the characters of 0 1 2 3 4 5 6 7 8 9 A b and C exist. Among them, A, b, C represent 10, 11, 12, respectively. As a result, the number of input and output characters $Q = 13$. A standard board as shown in Fig. 4 is used for training the weight of $\boldsymbol{W}$.

Considering the characteristic vector from a standard board as an input, each characteristic vector, $\boldsymbol{P}_q$, consists of

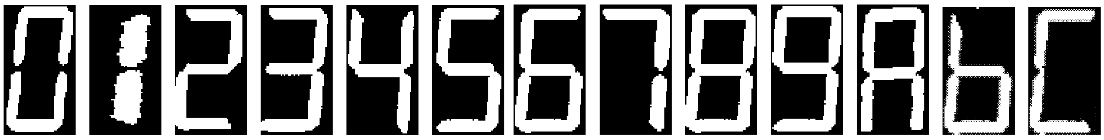

Fig. 4 Standard template characters

9×5 characters of 0 and 1, and thus the 45 characteristic points make up a 45×1 characteristic vector. Since altogether there are 13 input vectors, $\boldsymbol{P} = [\boldsymbol{P}_1\ \boldsymbol{P}_2\ \boldsymbol{P}_3 \ldots\ \boldsymbol{P}_{13}]$, the final input is a 45×13 matrix. According to Eq. (11), the network weight $\boldsymbol{W}$ is calculated. Then the characteristic $\boldsymbol{P}_q$ of detected traffic lights is used as inputs, as shown in Eq. (12) where the output of $\boldsymbol{t}_q$ is a standard characteristic vector of the character. Then the corresponding character is recognized.

$$\boldsymbol{t}_q = \boldsymbol{W}\boldsymbol{P}_q \tag{12}$$

Fig. 5 shows some poor character images of traffic lights from actual roads with some noise and broken edges. Using the extracted characteristic of the characters as self-associative memory inputs, the standard characteristics of the characters are recovered according to Eq. (12). The outputs are shown in Fig. 6, which proves that this method is robust enough to recognize the characters from images with noise and broken edges.

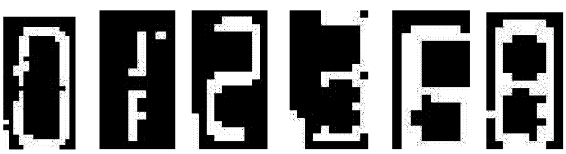

Fig. 5 Input character images

Fig. 6 Output character images

In order to test the efficiency of the method proposed in this paper, twenty scenes are selected. Selected typical urban intersections include sidewalk intersection, cross-type intersection, ring and T-shaped intersection, as shown in Fig. 7, in which the black line is the result of stabling siding detection and the white box is the result of signal light detection. For the intersections with clear stabling siding and traffic lights, this method can successfully detect the stabling siding and the location of red traffic lights, and recognize the time characters. However, when the intersection stabling siding is fuzzy and the LED is aging, some negative impacts are found. Based on the 20 actual scenes, the test results are shown in Tab. 1, where the types of intersections are corresponding to those shown in Fig. 7. It is shown that the detection for sidewalk intersections, ring and T-shaped intersections is much more accurate than that for cross-type intersections. The reasons are identified as the comparatively wider range of cross-type intersections. Consequently, the signal light is imaged in a relatively small image due to the erection on the opposite side of intersections, and the location accuracy of the lights is much lower. Fig. 7(a) and Fig. 7(b) are two typical major intersections in urban cities, which are the focuses of this paper for identifying objects. The sidewalks and cross-type intersections are classified into 18 groups, of which experiments show that two stabling sidings detections fail, while the two groups of ring and T-shaped intersections can be correctly detected and recognized. In the whole scene of experimental tests, the results show that the detection ratios of stabling siding and traffic lights are 90% and 85%, respectively. Because the green of traffic lights is similar to the color of leaves in the actual scene, the detection ratios of yellow and red traffic lights are relatively higher than that of green traffic lights. The character recognition ratio reaches 97% on the basis that the signal characters separate correctly in which most of the characters can be correctly identified except for the noisy and particularly "damaged" individual characters.

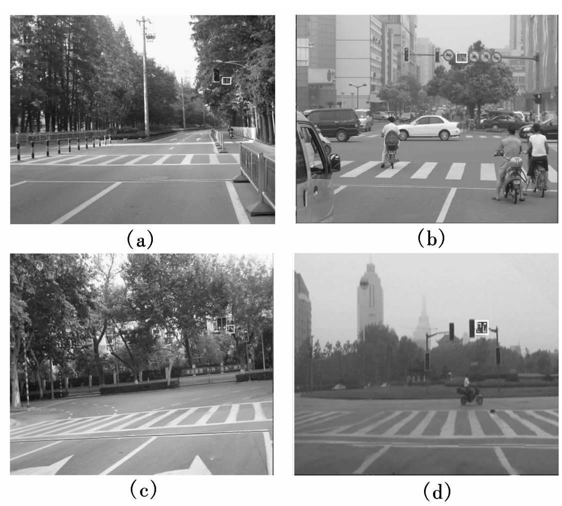

Fig. 7 Typical intersection scene. (a) Sidewalk intersection; (b) Cross-type intersection; (c) T-shaped intersection; (d) Ring-shaped intersection

Tab. 1 Intersection signal light detection and recognition results

Intersection type and quantity	Correct number of stabling siding detection	Correct number of signal light location	Correct number of character recognition
Sidewalk-type(10)	9	9	9
Cross-type(8)	7	6	5
Ring and T-shaped(2)	2	2	2

5 Conclusion

Detection and recognition of traffic lights at intersections is studied in this paper with a detection and recognition method for LED tri-color character traffic lights. Whether the vehicles approach the intersection or not is determined through the detection of stabling siding, and then the traffic lights are detected. Finally, the time characters of the traffic lights are identified with a self-associative memory. The experimental results show that the method is efficient to detect the stabling siding and traffic lights even under a background of noise and broken edges.

This paper proposes a new kind of approach to detect and identify traffic lights for revulsive driving under a simple traffic scene. Future researches will be carried out for more complicated traffic scenes, such as the scenes with the front stabling siding occupied by other vehicles, passengers or other obstacles. How to determine the left turn or right turn for intelligent vehicles will also be studied.

References

[1] Broggi A, Cattani S. An agent based evolutionary approach to path detection for off-road vehicle guidance[J]. *Pattern Recognition Letters*, 2006, **27**(11): 1164 − 1173.
[2] Zuo Xiaoqing, Li Qingquan, Xie Zhiying. Lane-based road data model[J]. *Journal of Chang'an University*: *Natural Science Edition*, 2004, **24**(2): 73 − 76. (in Chinese)
[3] Wang Rongben, Zhao Yibing, Li Linhui, et al. Approach review of obstacle detection for intelligent vehicle[J]. *Journal of Highway and Transportation Research and Development*, 2007, **24**(11): 109 − 113. (in Chinese)
[4] Bayouth M, Nourbakhsh I, Thorpe C. A hybrid human-computer autonomous vehicle architecture[C]//*Third ECPD International Conference on Advanced Robotics*, *Intelligent Automation and Active Systems*. Bremen: International Program Committee, 1997: 1 − 7.
[5] Bellutta P, Manduchi R, Matthies L, et al. Terrain perception for DEMO Ⅲ [C]//*Proceedings of the IEEE Intelligent Vehicles Symposium*. Bergh: IEEE, 2000: 326 − 331.
[6] Broggi A, Bertozzi M, Fascioli A, et al. Automatic vehicle guidance: the experience of the ARGO autonomous vehicle [R]. Singapore: World Scientific, 1999.
[7] Xu Youchun, Wang Rongben, Li Keqiang, et al. A linear model based road identification algorithm[J]. *Journal of Image and Graphics*, 2004, **9**(7): 858 − 864. (in Chinese)
[8] Li Liang. Study on monitor techniques of vehicles peccancy based on moving image recognition[D]. Dalian: School of Electronic and Information Engineering of Dalian University of Technology, 2005.
[9] Hagan M T, Demuth H B, Beale M. *Neural network design* [M]. Beijing: China Machine Press, 2005: 99 − 111.
[10] Sarangi S, Rath N P. Performance analysis of fuzzy-based canny edge detector[C]//*Conference on Computational Intelligence and Multimedia Applications*. Sivakasi: IEEE, 2007, **3**: 272 − 276.
[11] Dandawate Y H, Joshi M A, Chitre A V. Quality analysis of color images compressed with enhanced vector quantizer designed using HSI color space[C]//Conference on Computational Intelligence and Multimedia Applications. Sivakasi: IEEE, 2007, **3**: 138 − 142.
[12] Yung N H C, Lai A H S. An effective video analysis method for detecting red light runners[J]. *IEEE Transactions on Vehicular Technology*, 2001, **50**(4): 1074 − 1084.
[13] Shen H, Roysam B, Stewart C V, et al. Optimal scheduling of tracing computations for real-time vascular landmark extraction from retinal fundus images[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2001, **5**(1): 77 − 91.

智能车辆的交叉口数字信号灯检测与识别

张　宁　何铁军　高朝晖　黄　卫

(东南大学 ITS 研究中心,南京 210096)

摘要:为了保证智能车辆在交叉口内的诱导行驶,提出了一种交叉口信号灯检测与识别方法. 首先利用 Hough 变换检测交叉口内的停车线,然后采用颜色空间变换检测红、黄、绿三色信号灯,最后建立自联想存储器以识别切分出来的信号灯时间字符. 通过 20 个实际交叉口场景测试数据验证,采用所提出的方法,停车线检测正确率达 90%,信号灯检测正确率为 85%,在信号灯字符正确分割出来的基础上,字符的识别率达 97%. 结果表明提出的方法能够十分有效地进行数字信号灯的检测,并具有足够的鲁棒性识别"破损"及带噪声的字符.

关键词:智能车辆;停车线检测;信号灯检测;自联想存储器;LED 字符识别

中图分类号:U491