Robust and Real-Time Traffic Lights Recognition in Complex Urban Environments

Chunxiang WANG, Tao JIN
Research Institute of Robotics, Shanghai Jiao Tong University
Shanghai, 200240, China

Ming YANG[*], Bing WANG
Department of Automation, Shanghai Jiao Tong University
Key Laboratory of System Control and Information Processing, Ministry of Education of China
Shanghai, 200240, China

Abstract

The traffic lights play an indispensable role in urban road safety and researches on intelligent vehicles become more popular recently. In this paper an automatic system for robust and real-time detection and recognition of traffic lights for intelligent vehicles based on vehicle-mounted camera is proposed. The method applying image processing and pattern recognition theory mainly works in three stages. First, the candidate regions of traffic lights are extracted using the color threshold segmentation method. Secondly, noise removal and two types of filtering which take account of shape information are applied to the candidate regions. Thirdly, template matching using normalized cross correlation techniques is adopted to validate the traffic lights candidate. Experimental results show that the proposed algorithm works effectively and robustly for traffic lights recognition in complex urban environments.

Keywords: Traffic lights recognition, Image processing, Pattern recognition, Color threshold segmentation, Candidate regions filtering, Template matching.

1. Introduction

Road traffic crashes are one of the world's largest public health and injury prevention problems. The problem is all the more acute because the victims are overwhelmingly healthy prior to their crashes. According to the World Health Organization (WHO), more than a million people are killed on the world's roads each year. Traffic lights play an indispensable role in urban road safety, and a lot of researches on traffic safety[1-3] focused on other aspects have been done. Also in recent years, more and more attention has been paid to intelligent vehicles driving in urban environments. The intelligent vehicle must comply with traffic rules in the complex urban environment, and traffic lights are important signals for a vehicle driving in the complex urban environment. Traffic lights usually locate at intersections of the roads and provide rich information about the location of intelligent vehicle. They also play an important role in car navigation systems.[4] So the detection of the existence and the confirmation the status of the traffic lights is an important part for intelligent vehicles driving in complex urban environments.

There are quite a lot of researches on the detection and recognition of traffic signs[5-7], and some researches on the recognition of traffic lights also have been done only for the suspended traffic lights in semi-urban environment[8-12]. However, few researches have been conducted on other types of traffic lights in complex urban environments. But in fact, there are significant differences from place to place in how traffic lights are mounted or positioned in complex urban environment. Traffic lights may be mounted on poles situated on street corners, hung from horizontal poles or wire strung over the road way, or installed within large horizontal

gantries that extend out from the corner and over the right-of-way. Also, there are numerous interference for traffic lights recognition in complex urban environment, including vehicle tail lights, store billboard lights, and traffic lights partly obscured by other objects.

Y. Chung et al.[13] proposed a method that background images were generated by the system at first and in the mean time illumination parameters were estimated. Then they employed the HSI color model, fuzzy methods together with morphological technique to acquire the candidate traffic light areas. With the relative spatial and temporal information, the scales, positions and timing sequences of traffic lights were obtained. Since their method needed to estimate background images, it could not apply to moving vehicles. Similar application condition can also be found in the work of N. Yung and A. Lai[14], in which the cameras were installed in a static place and these algorithms were not applicable to the traffic lights recognition for intelligent vehicles.

T. Hwang et al. proposed an approach to detect traffic lights which consisted of a color-based thresholding and center finding by Gaussian mask, and verification the candidates of traffic lights using the suggested existence-weight map. Y.K Kim et al.15 performed image adjustment first and then applied two types of color filtering and a thresholding algorithm to detect traffic lights. Although these color-based recognition methods can effectively detect and identify traffic lights in a relatively simple environment that the background is almost static and generally include a sky area, a large number of false detections will occur in the complex urban environment.

Y. Shen et al.16 proposed an algorithm modeled the hue and saturation according to Gaussian distributions and learned their parameter with training images. But their candidate regions extraction method was too computational expensive to satisfy the real-time recognition requirements. R. d. Charette and F. Nashashibi17-18 introduced a real-time traffic lights recognition system mainly relied on their generic "Adaptive Templates" to recognize different kinds of traffic lights. However, usual template matching based on simple covariance or the sum of the absolute differences (SAD) is not robust enough to reach good matching results.

Color and shape are dominant visual features of traffic lights with distinguish characteristics and are key information for drivers to process when driving along the road. So this information should be utilized effectively and effectively even in the knowledge that color and shape vary with the change of lighting conditions and viewing angles. Generally, the traffic lights recognition methods mainly include the color-based method, the shape-based method and the method based on template matching. In the following sections, new algorithms will be presented for color and shape representation, robust template match description, leading to traffic lights recognition. Section 2 gives detailed description of color thresholding segmentation method. Noise removal and two kinds of filtering based on shape property will be given in Section 3. Section 4 addresses the application of the normalized cross correlation for a robust template matching. The experimental results and conclusions are given in Section 5 and Section 6, respectively.

Fig. 1.is the source image, presented for future demonstration. Four traffic lights in status 'stop' are all emphasized by red ellipses only to be more easily visual by readers.

Fig. 1. Source image will be used as reference for further demonstrations. Traffic lights are emphasized only to be more easily visible by the readers.

2. Extraction of color information

The candidate regions of the image in which the traffic lights are more likely to be found are selected by using the information on color. As we know, RGB (Red, Green and Blue) color model is one of the most normal color space models, but it is not most suitable for the traffic lights detection in color image. The reason for this is that the color value in RGB model is decided by

the RGB three primary color's value plus result, and they are correlated strongly. HSI (Hue, Saturation, Intensity) color space is a color model based on human vision, and many researchers have used color thresholding with the HSI color model. It can be used as a measure for distinguishing predefined colors of traffic lights. In this paper the authors use HSI color model to analyze colors of traffic lights. At first the components of HSI model can be transformed from the general camera RGB model [19].After transforming RGB color model into HSI model, the authors execute the thresholding with regard to red, green and yellow pixels to get traffic lights candidates, as

$$\begin{aligned}
\mathrm{Re}\,d &= \{p(x,y)\,|\,(T_{HR1} < p_H(x,y) < T_{HR2}) \\
&\quad \wedge (p_I(x,y) > T_I) \wedge (p_S(x,y) > T_S)\} \\
Green &= \{p(x,y)\,|\,(T_{HG1} < p_H(x,y) < T_{HG2}) \\
&\quad \wedge (p_I(x,y) > T_I) \wedge (p_S(x,y) > T_S)\} \\
Yellow &= \{p(x,y)\,|\,(T_{HY1} < p_H(x,y) < T_{HY2}) \\
&\quad \wedge (p_I(x,y) > T_I) \wedge (p_S(x,y) > T_S)\}
\end{aligned} \tag{1}$$

where $p_H(x,y)$, $p_S(x,y)$, $p_I(x,y)$ is the hue, saturation and intensity value of the pixel $p(x,y)$.

However, the colors of the traffic lights are not commonly considered pure red, green, and yellow. Because the red light contains some orange in its hue, and the green light contains some blue, to provide some support for people with red-green color blindness. Duo to this and the fact that the hue value satisfy with Gaussian distribution, the authors collect training images with red, yellow, green traffic lights respectively and label the traffic lights regions manually. With all the pixels in manually labeled traffic lights regions, the mean value and the standard deviation of the hue value of each color can be estimated. And according to the formula $t_h = \mu \pm 3\sigma$ the authors can determine the threshold for the hue value, which will contain about 99.7% of the detected pixels. The thresholds for the saturation and intensity value are also manually assigned based on the results of the training image.

As illustrated in Fig. 2. , all pixels with the same color (red, green, yellow) of the traffic lights are reserved for future procession.

3. Filtering of Candidate Regions

The filtering of candidate regions is carried out in two stages. First, the noise removal is carried out on the image using the morphology operation, and scatter

Fig. 2. Result image of color thresholding on the source image.

small blocks are removed and connected regions are preserved accurately. Secondly, each candidate region is tested according to their shape feature and validated or discarded depending on whether they satisfied the dimension property and the border property of the real traffic lights.

3.1. *Noise removal*

The color thresholding step generates the candidate regions for the red, green and yellow traffic lights considering only the color information. But some noise s may be produced from lens distortion, vibration of camera, diffraction of sunlight and background objects with similar color. In order to eliminate scattered small blocks of candidate regions and preserve the area of connected regions more accurately, the authors apply the opening morphological operation. The opening morphology is the dilation of the erosion of a set A by a structuring element B as shown in:

$$A \bullet B = (A - B) + B \tag{2}$$

where " – " and "+" denote erosion and dilation morphological operation, respectively.

The kernel size used in the opening operation and the iteration times of the opening operation are determined on the basis of the resolution of the original image and the focal length of the lens. With high resolution image acquisition and larger focal length, the size of the traffic lights in image is relatively larger, a larger kernel can be chosen and the opening operation can be conducted twice. Otherwise, a smaller kernel should be chosen and the opening operation should be

conducted once. Fig. 3. is the result image applying noise removal on the color thresholding image, and the great improvement can be noticed compared to Fig. 2.

3.2. *Filtering based on shape feature*

Although noise removal can eliminate small scattered candidate regions, it cannot deal with large components which may be mistakenly treated as candidate regions of traffic lights. So the authors perform connected component extraction of the 8-connected image graph using the two-path algorithm first and then filtering these candidate regions according to their shape feature.

3.2.1. *Dimension based thresholding*

Though traffic lights are generally round, we cannot expect the detected regions to be exact circles for image sampling reasons and due to noises and lens distortions. So the authors extract the connected component and enclose it with a bounding box. And according to the dimension properties of the connected component extraction, the candidate regions can be filtered.

First, the dimension ratio DR of the connected component is used to reject some candidates:

$$DR = \frac{Max(width, height)}{Min(width, height)} \quad (3)$$

Dimension ration DR defines the ratio between the width and the height of the bounding box. In the case the connected component is an exact circle region, the dimension ratio is 1. We should choose the thresholding value t_{DR} larger than 1 and if the dimension ration DR
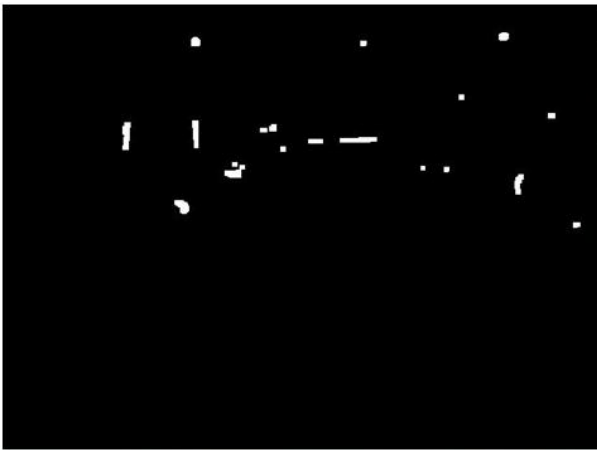
is greater than the thresholding value t_{DR}, the connected component should be eliminated.

Secondly, the area of the bounding box A of the connected component is used to reject some candidates:

$$A = width * height \quad (4)$$

If the area of the bounding box A is greater than the thresholding value t_A^G or less than the thresholding value t_A^L, the connected component should be eliminated.

Thirdly, the filling ratio FR of the connected component is used to reject some candidates:

$$FR = \frac{Area(C)}{width * height} \quad (5)$$

The filling ratio FR defines the ratio between the area of the connected component and that of its bounding box. In the case the connected component is an exact circle region, the filling ratio FR is 0.785. We should choose the thresholding value t_{FR} smaller than 0.785 and if the filling ratio FR is less than the thresholding value t_{FR}, the connected component should be eliminated.

As with the noise removal section, all the thresholding values ($t_{DR}, t_A^G, t_A^L, t_{FR}$) are determined by the resolution of the original image and the focal length of the lens and can be manually assigned based on the training image. And the result after filtering based on dimension properties is shown in Fig. 4.

Fig. 3. Result image of noise removal using opening morphological operation with a 5x5 rectangle structural element (kernel).

Fig. 4. Result image with bounding box of the candidate regions after filtering based on dimension properties. Notice that there are many false alarms but all the four candidate regions of traffic lights are detected.

3.2.2. *Border Property based thresholding*

Typical for any traffic light is a brighter blob surrounded by a darker box, so the authors can apply this border property to check whether the connected component candidate is traffic light or not. The authors transform the initial image into the gray-scale image and then apply modified region growing algorithm[20] using seed points lied in every connected component candidates to achieve this.

The crucial part of our modified region growing algorithm is the determination of the seed points and the range of the threshold. The purpose of applying region growing on this image is that the authors want to find a new connected component which corresponding to the originated connected component and then compare them. Therefore, the authors choose the point having the highest gray scale value in the originated connected component candidates as seed point. And the range of the threshold is determined based on the fact that the whole traffic light blob should be marked.

Due to the border property discussed ahead, the originated connected component should resemble the extracted component. Consequently, if the originated connected component and the connected component extracted from our modified region growing algorithm are similar, the originated connected component is accepted as a traffic light candidate. Otherwise, the connected component extracted from our region growing algorithm is much larger and the originated connected component is rejected. So some candidate regions can be rejected according to the degree of their similarity. Fig. 5 shows the result after filtering based on the border properties.

4. Template Matching Using Normalized Cross Correlation

Template matching was used previously in different detection and recognition process. Template matching is computationally expensive and slow when it is applied on the whole image to find the best match. However the authors just need to estimate the matching degree for the position where the candidate regions have been extracted. Based on the matching degree, the authors decided whether the candidate regions are traffic lights or not and the status of traffic lights. And in our case, the information about the position of the candidate regions can be used to reduce the computational effort significantly.

Fig. 5. Result image with bounding box of the candidate regions after filtering based on border properties. Notice that there is only one false alarm which is the taillight of the car. Still, no candidate regions of traffic lights are missed.

The template matching degree is evaluated by calculating the normalized cross-correlation[21] coefficient between the selected templates and the candidate region. If the normalized cross-correlation value is higher than a fixed threshold, the region is considered as corresponding to a traffic light. Otherwise it is rejected.

In order to get a good and robust matching result, the authors perform template matching using the normalized cross correlation method, and it is motivated by the distance measure (squared Euclidean distance):

$$d_{f,t}^2(u,v) = \sum_{x,y} [f(x,y) - t(x-u, y-v)]^2 \qquad (6)$$

where f is the image and the sum is over x, y under the window containing the feature t positioned at u, v. In the expansion of d^2,

$$d_{f,t}^2(u,v) = \sum_{x,y} [f^2(x,y) - 2f(x,y)t(x-u,y-v) + t^2(x-u,y-v)] \qquad (7)$$

the term $\sum t^2(x-u, y-v)$ is constant. If the term $\sum f^2(x,y)$ is approximately constant then the remaining cross-correlation term

$$c(u,v) = \sum_{x,y} f(x,y)t(x-u,y-v) \qquad (8)$$

is a measure of the similarity between the image and the feature.

There are also several disadvantages to using "Eq. (8)" for template matching:

- If the image energy $\sum f^2(x, y)$ varied with position, matching using "Eq. (8)" can fail. For example, the correlation between the feature and an exactly matching region in the image may be less than the correlation between the feature and a bright spot.
- The range of $c(u,v)$ is dependent on the size of the feature.
- "Eq. (8)" is not invariant to changes in image amplitude such as those caused by changing lighting conditions across the image sequence.

The correlation coefficient overcomes these difficulties by normalizing the image and feature vectors to unit length, yielding a cosine-like correlation coefficient

$$\gamma = \frac{\sum_{x,y}(f(x, y) - \overline{f_{u,v}})(t(x-u, y-v) - \overline{t})}{\sqrt{\sum_{x,y}(f(x, y) - \overline{f_{u,v}})^2 \sum_{x,y}(t(x-u, y-v) - \overline{t})^2}} \quad (9)$$

In "Eq. (9)" $\overline{f_{u,v}}$ denotes the mean value of $f(x, y)$ within the area of the template t shifted to (u, v) which is calculated by

$$\overline{f_{u,v}} = \frac{1}{N_x N_y} \sum_{x=u}^{u+Nx-1} \sum_{y=v}^{v+Ny-1} f(x, y) \quad (10)$$

With similar notation $\overline{t}$ is the mean value of the template t.

Fig. 6. Result image of the whole traffic lights recognition algorithm. All four traffic lights in status "stop" are recognized correctly without false alarms and labeled with the red rectangle.

Due to the normalization, the use of "Eq. (9)" for the calculation of the matching degree is more robust than the other similarity measures, like simple covariance or the sum of the absolute differences (SAD). The result image after template matching based on normalized cross correlation is shown in Fig. 6.

5. Experiments Results

5.1. *Traffic Lights Recognition Results*

In order to analyze the recognition performance of the proposed method, the experiments have been performed. The video was provided by the Robotics Centre of Mines ParisTech and collected in a dense urban environment around Paris. The Marling F-046C video camera was mounted behind the interior rear-view mirror of the C3 vehicle. The frame rate of the video sequences is 25 fps and the resolution is 640 x 480.

The video sequences were downloaded to a standard PC and tested on our algorithm offline. Processing was performed on a 2.2GHz Core processor and the size of the main memory is 2GB. The ground truth file lists the real traffic lights on each frame and also provided by the Robotics Centre of Mines ParisTech. And finally the authors compared the detected traffic lights with the ground truth. Detailed results of the algorithm are shown in Fig. 7.

5.2. *Traffic Lights Recognition Accuracy*

In this Section, the authors will provide a quantitative analysis about the recognition accuracy of the proposed algorithm compared with human traffic lights labeling results in the ground truth file in 2 sequences. During the labeling process our human operator noticed several ambiguous regions for which they had issue to decide whether it was a real traffic light (with circle light) or not. The authors thus decided to simply ignored these ambiguous regions during the evaluation. Therefore, any traffic light detected in these regions won't be taken into account neither as a "false positive" nor as a "true positive".

Indeed, there were 9168 instances of traffic lights in our experiments, and among them 449 were "ambiguous" regions and they were strictly labeled "ambiguous" only if they validate one of the following conditions:

Fig. 7. Result of The traffic lights recognition experiment.

- Reflection distortion. The region is a reflection of an object which seems to be a traffic light
- Light shape not valid. The light of the traffic light appears circle were it is in fact a rectangle
- Too blurry. The traffic light is 'too' blurry during its whole timeline
- Too small. The traffic light is too small during its whole timeline.
- Not facing the vehicle. The traffic light is not facing the vehicle but the light is still visible.
- Lower traffic light. The small and lower traffic lights under the big one are ignored.

Table 1. Recognition Statistics of the Proposed Algorithm

Index	Number of Total Frames	Precision	Recall
1	1581	97.2%	96.5%
2	3972	96.7%	92.3%

The Precision and Recall used in Table 1 are computed as follows:

$$\Pr ecision = \frac{truePositives}{truePositives + falsePositives} \quad (1)$$

$$\operatorname{Re} call = \frac{truePositives}{truePositives + falseNegatives} \quad (2)$$

5.3. *Computational Time Analysis*

Table 2. lists the computation time step by step of the proposed algorithm. Note that the traffic lights candidate regions extraction is the most time consuming step. Conversely, the template matching using normalized cross correlation is almost insignificant because it is only applied on the candidate regions found in the traffic lights candidates' detection step.

Table 2. Computational Time Analysis

Step	Average Computational Time (ms)
Extraction of color information	24
Filtering based on shape feature	13
Robust template matching using NCC	3
Total	40

6. Conclusions

In this article, an algorithm for the detection and recognition of traffic lights in complex urban environment has been proposed. The color segmentation is completed in the HSI color space and the segmentation threshold is obtained through the statistical analysis of the training image. Using the filtering based on the dimension property and the border property the authors are able to eliminate the number of candidate regions effectively. And then the authors use the robust template matching algorithm which applied the normalized cross correlation to verification whether the candidate region is a traffic light. Finally, the whole system presented has been proved on urban sequences.

References

1. W. Wang, F. Hou, H. Tan, A Framework for Function Allocation in Intelligent Driver Interface Design for Comfort and Safety, International Journal of Computational Intelligence Systems, 3(5) (2010),531-541.
2. W. Wang, H. Guo, K. Ikeuchi, Numerical simulation and analysis procedure for digital driving dependability in intelligent transport system, KSCE Journal of Civil Engineering, 15(5) (2011),891-898.
3. W. Wang, Y. Mao, J. Jing, Driver's various information process and multi-ruled decision-making mechanism: a fundamental of intelligent driving shaping model, International Journal of Computational Intelligence Systems, 4(3) (2011),297-305.
4. T. Hwang, I. Joo, and S. Cho, Detection of traffic lights for vision-based car navigation systems, in Pacific-Rim Symposium on Image and Video Technology (2006) 682-691.
5. A. Escalera, J. Armingol, M. Mata, Traffic sign recognition and analysis for intelligent vehicles, Image and Vision Computing, 21(2003), 247-258.
6. A. Soetedjo and K. Yamada, "Fast and robust traffic sign detection," in IEEE International Conference on Systems, Man and Cybernetics (2005) 1341-1346
7. X. W. Gao, L. Podladchikova, D. Shaposhnikov, Recognition of traffic signs based on their colour and shape features extracted using human vision models, Journal of Visual Communication and Image Representation (2005) 675-685.
8. K. Lu, C. Wang, S. Chen. Traffic light recognition [J]. Journal of the Chinese Institute of Engineers, 31(6) (2008) 1069-1075.
9. J. Gong, Y. Jiang, G. Xiong, The recognition and tracking of traffic lights based on color segmentation and CAMSHIFT for intelligent vehicles, IEEE Intelligent Vehicles Symposium (University of California, San Diego, CA, USA, 2010), pp. 431-435.
10. M. Omachi, S. Omachi, Traffic light detection with color and edge information, Computer Science and Information Technology (2009) 284-287
11. M. Omachi, S. Omachi, Detection of traffic light using structural information, The IEEE International Conference on Signal Processing (ICSP) (2010) 809-812
12. Lindner F, Kressel U, Kaelberer S. Robust recognition of traffic signals. IEEE Intelligent Vehicles Symposium [C]. (Piscataway NJ USA, 2004), pp. 49-53.
13. Y. Chung, 1. Wang, and S. Chen, A vision-based traffic light detection system at intersections, Journal of National Taiwan Normal University: Mathematics, Science & Technology, vol. 47 (2002) pp. 67-86.
14. N. Yung and A. Lai, An effective video analysis method for detecting red light runners, IEEE Trans. Vehicular Technology, vol. 50 (2001) 1074-1084.
15. Y. K. Kim, K. W. Kim, X. Yang, Real time traffic light recognition system for color vision deficiencies, IEEE International Conference on Mechatronics and Automation (Washington, DC, 2007), pp. 76-81.
16. Y. Shen, U. Ozguner, K. Redmill, A robust video based traffic light detection algorithm for intelligent vehicles. IEEE Intelligent Vehicles Symposium [C]. (Piscataway NJ USA, 2009), pp. 521-526.
17. R. d. Charette, F. Nashashibi, Traffic light recognition using image processing compared to learning processes, The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (St. Louis USA, 2009), pp. 333-338.
18. R. d. Charette, F. Nashashibi, Real time visual traffic lights recognition based on spot light detection and adaptive traffic lights templates, IEEE Intelligent Vehicles Symposium [C]. (Piscataway NJ USA, 2009), pp. 358-363.
19. R. Gonzalez and R. Woods, Digital Image Processing, (Addison-Wesley, 1992).
20. Y. Chang, X. Li, Adaptive image region-growing, IEEE Transactions on Image Processing (1994) 868-872.
21. J. P. Lewis Fast normalized cross-correlation, Industrial Light and Magic.