Real-time Arrow Traffic Light Recognition System for Intelligent Vehicle

Zixing Cai¹, Mingqin Gu¹, Yi Li¹
¹ School of Information Science and Engineering, Central South University, Changsha, Hunan Province, China

Abstract - Arrow traffic lights are common at intersections of current urban scenes. However, existing algorithms mainly focus on recognition methods of circle traffic light. A novel system used to detect and recognize arrow traffic lights in urban scenes is proposed in the paper. Firstly, the boards of traffic light are exactly localized by image segmentation and morphology processing, then candidates of traffic light are obtained according to threshold segmentation in YCbCr color space and judged by relative position between candidate and its board. Secondly, for confirming traffic light type, Gabor wavelet transform and 2D independent component analysis (2DICA) are used to extract traffic light candidate’s features. A nearest neighbor classifier identifies arrow direction and removes noise. Experimental results show that overall recognition rate of proposed method exceeds 91%, and system will provide real-time, robust, effective, and accurate traffic light outputs for intelligent vehicle.

Keywords: Traffic light recognition; Intelligent vehicle; Gabor wavelet; 2DICA

1 Introduction

A key component of intelligent vehicles is the perception system, which allows the vehicle to perceive and interpret its surroundings. Humans have engineered the driving problem to make it easier. For example, traffic lights indicate precedence at intersections, and turn signals indicate the driver’s intentions. Traffic lights are a special perception problem. Non-passive traffic light systems that broadcast their current state and other information have been demonstrated in industry and academic settings, but this requires a significant investment in infrastructure. Reliable detection and recognition of the state of traffic lights including color and direction is essential to intelligent vehicle in urban situations. Common failures of traffic light detection are either visual obstructions or false positives, such as those induced by the taillights of other vehicles, false greens which arisen from particular patterns of light on a tree, or from brightly lit billboards and so on.

Recently years, several approaches arose in the detection and recognition of traffic lights. Yung et al. detected the traffic lights in the video sequence of red, amber and green, and identified the traffic lights with the camera fixed on a shelf at the cross. However, the algorithm needed a clear view of the 3-color traffic lights; the distance between the camera and the traffic lights was not far. Chung et al. first estimated background images and mean time illumination. Fuzzy methods together with morphological technique were applied to acquire the traffic light candidate regions. It cannot be applied to moving vehicles due to estimate background images in advance. In [4], Hwang used 6 color thresholds to get the candidate regions. Unfortunately, it could not discriminate red light from amber light and lacked flexibility. A method was proposed a method to detect a traffic light based on the Hough transform in [5]. It considered the structure of traffic light. But this approach was lack of effective recognition of traffic lights in complex environment. The hue and saturation modeling based on Gaussian distributions and their parameters were learned by training images in [6], then candidate regions of the traffic lights in the test images can be extracted and judged by these parameters and shape information. Gong et al. also used the color threshold obtained by the statistics information in HSV color space to image segmentation. The machine learning and CAMSHIFT tracking algorithm classified and tracked traffic light respectively. Such systems were vulnerable to varying illumination conditions. Park et al. introduced the detector which biased toward circular regions of high intensity surrounded by regions of low intensity to attempt to report the state of light. However, their method did not reliably differentiate between multiple lights, or determine the state of more than one light in a single camera frame. Charette et al. applied the spotlight detection and template matching method to identify the traffic lights. Although the detection and recognition of this method was accurate, the algorithm was very time-consuming. Li et al. used morphology filtering and statistical classification to detect circular traffic light industry urban scenes. However, their method did not detect and recognize states of arrow traffic light. Other researchers pre-map traffic light locations to perform online traffic light detection.

Arrow traffic lights are common type in current urban scenes and provide more useful information than circular traffic lights to moving vehicles in the intersections, but above algorithms can merely detect and recognize circle traffic lights. A novel approach is proposed to detect and recognize arrow traffic lights in this paper. Arrow traffic
lights are localized on input image by color and shape information in urban environments. Features used to classify traffic light are extracted by Gabor wavelet transform and 2D independent component analysis, and a nearest neighbor classifier identifies arrow type.

2 Traffic Light Detection

Arrow traffic lights are common in urban scenes, but their backgrounds are very complex. Besides, due to varying illumination with time, luminance of traffic lights and small size, it is very difficult to detect traffic lights.

Figure 1. common traffic lights in China urban scene

To solve above problem, a framework of arrow traffic light recognition is provided in figure 2.

2.1 Board of Traffic Light Localization

Let each pixel value \( v(x, y) = [v_r, v_g, v_b] \) in RGB color space, \((x, y)\) is pixel coordinate. Black pixels are separated from input image by 2 different techniques as following:

\[
\text{Bin}_1(x, y) = \begin{cases} 
1, & \max(v_r, v_g, v_b) < T_1 \\
0, & \text{otherwise}
\end{cases}
\]  
(1)

\[
\text{Bin}_2(x, y) = \begin{cases} 
1, & \max(|v_r - v_g|, |v_g - v_b|, |v_b - v_r|) < T_2 \\
0, & \text{otherwise}
\end{cases}
\]  
(2)

\(T_1, T_2\) are thresholds for image segmentation. Binary image of black regions is the union of 2 obtained binary images:

\[
\text{Bin}(x, y) = \text{Bin}_1(x, y) \cup \text{Bin}_2(x, y)
\]  
(3)

Two morphology technologies, Erosion and Dilation with same structural element, are used to remove noise and broken regions on \(\text{Bin}(x, y)\).

Region labeling is performed on \(\text{Bin}(x, y)\) to extract connected components with sizes over 300 pixels and below 3000 pixels as candidate regions of traffic light boards. Assume that there are \(N_r\) candidate regions, denoted by \(R_i\), and \(R_i\) is height and width of \(R_i\) respectively. \(R_{wh}\) is aspect ratio of \(R_i\) and defined as following.

\[
\frac{R_i\text{height}}{R_i\text{width}}
\]  
(4)

\(\text{Saturation}_i\) represents convexity of region \(R_i\), and has

\[
\text{Saturation}_i = \frac{\text{Area}_i}{\text{ConvexArea}_i}
\]  
(5)

Where, \(\text{Area}_i\) and \(\text{ConvexArea}_i\) are the pixels number of \(R_i\), and convex hull of \(R_i\) respectively.

For an obtained candidate region of traffic light board, the morphology filter of \(R_i\) is designed as following steps.

\[
\text{Bool}(R_{wh}) = \begin{cases} 
1, & \text{T}_{wh} < R_{wh} < T_{wh} \\
0, & \text{otherwise}
\end{cases}
\]  
(6)

\[
\text{Bool}(\text{Saturation}) = \begin{cases} 
1, & \text{Saturation}_i > T_{sat} \\
0, & \text{otherwise}
\end{cases}
\]  
(7)

\[
\text{Filter}(R_i) = \begin{cases} 
R_i, & \text{Bool}(R_{wh}) \cap \text{Bool}(\text{Saturation}) = 1 \\
\text{null}, & \text{otherwise}
\end{cases}
\]  
(8)

\(T_{wh}, T_{sat}\) are the corresponding thresholds. The \text{Filter} is used to remove regions that are not satisfied board morphology. After filtering, the board candidate regions are remained and shown as Figure 3(c).

2.2 Arrow Traffic Light Detection
The color image of candidate board region $R_s$ is cropped from input original image, and converted into YCbCr color space from RGB to segment specific colors of traffic lights.

Assume $p(x, y) = [p_r, p_y, p_c]$ represents pixel value in YCbCr color space, and $(x, y)$ denotes coordinate in cropped image. Since it is difficult to distinguish red and amber traffic lights by threshold segmentation in YCbCr color space, red and amber candidates are gotten together by threshold segmentation in Cb channel:

$$B_{red}(x, y) = \begin{cases} 1, & p_{cb} \leq T_{red} \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where, $T_{red}$ is a threshold for image segmentation. Since positions of red and amber traffic lights in board are obviously different (red lights in the top of board, and amber lights in the middle of board), relative position between candidate region and its board can be used to distinguish red and amber traffic lights candidate in $B_{red}(x, y)$.

Regions are labeled on $B_{red}(x, y)$ to extract connected components with size over 50 pixels and below 1000 pixels, radio of height and width over 0.5 and below 1.5 at the same time. Supposed that there are $Ns$ candidate regions, which are denoted by $SR_j, j = 1, \cdots, Ns$. Assume outside rectangular of $SR_j$ is $Rect_j = \{SR_{x, y_{SR}}, SR_{width}, SR_{height}\}$. If the relative position between $SR_j$ and light board $iR_s$ satisfies:

$$\begin{align*}
&SR_{width} > R_i(width)/2 \\
&R_i(height)/5 < SR_{height} < R_i(height)/2
\end{align*} \tag{10}$$

Then $SR_j$ is a traffic light candidate.

If upper left corner coordinate $SR_{x, y}$ of $SR_j$ satisfies $\max(SR_{x}, SR_{y}) < R_i(width)/3$, then $SR_j$ is a red traffic light candidate. Else if coordinate of $SR_j$ meets:

$$SR_x < R_i(width)/3, \text{ and } R_i(width)/3 < SR_y < 2R_i(width)/3$$

Then $SR_j$ is an amber traffic light candidate, else it is not a traffic light and removed from candidate list.

To detect green traffic light regions in board image, Cr channel image is segmented by threshold $T_{green}$, binary image is obtained by:

$$B_{green}(x, y) = \begin{cases} 1, & p_{c} \leq T_{green} \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

Region labeling method which is implemented for red
light is performed on $B_{green}(x,y)$. Denoted candidate regions as $SR_i, i = 1, \ldots, N_s$. Assume outside rectangular of $SR_i$, Rect$_i = \{SR_i, SR_{i,a}, SR_{i,b}, SR_{i,mp} \}$ satisfies eq.(10). If upper left corner coordinate $SR_i$, $\{x, y \}$ width height $SR_i$, satisfies $SR_i < R_{width}/3$, $SR_i > 2R_{width}/3$, then $SR_i$ is a green traffic light candidate.

An example of candidate region detection after region labeling and size verification is shown in Figure 3(d). Color image of region $SR_i$ is cropped from board image, transformed gray image and normalized to size $30 \times 30$ pixels. The processing image is denoted as $(x,y)$ and sent to step of traffic light recognition.

### 3 Traffic Light Recognition

For classifying traffic light candidate image, 2D Gabor wavelet transform and 2D Independent Component Analysis are used to represent image and reduce features respectively.

#### 3.1 2D Gabor Wavelet Transform

Gabor wavelet transform is a traditional choice for obtaining localized frequency information, and offers the best simultaneous localization of spatial and frequency information[13]. For obtained image $(x,y)$ in detection stage, the Gabor filtered images $O_{uv}(x,y)$ are calculated by convoluting the Gabor functions with image $(x,y)$.

$$G_{uv}(x,y) = \sqrt{\left[\text{real}(O_{uv}(x,y))\right]^2 + \left[\text{imag}(O_{uv}(x,y))\right]^2}$$  \hspace{1cm} (12)

As the outputs $G_{uv}(x,y), u = \{0,1,\cdots,5\}, v = \{0,1,\cdots,5\}$ are orientation and scale of Gabor filter, consisting of different image features. All these features are concatenated into a single column vector in order to derive a feature vector $\chi = \{G_{0,0}(x,y)^T, G_{0,1}(x,y)^T, \cdots, G_{5,5}(x,y)^T\}$.

#### 3.2 Traffic Light Feature Extraction by 2D Independent Component Analysis

2D independent component analysis is an improving method of ICA which can eliminate redundancy and reduce dimension of the samples effectively.

Assume that traffic light samples’ Gabor image feature are $\chi = \{X_i, i \in \{1,2,\cdots, Q\}, X_i \in R^m \}$, each component may be combined by $P(P \leq Q)$ unknown independent component by different coefficient[14]. To reduce dimension of samples, the main task is to obtain optimization projection matrix $S = (s_1, s_2, \cdots, s_m)^T$, according to the method introduced in[14], $m$ is the number of independent components. Independent component $s_i$ must have non-Gaussian distributions with zero mean, unit variance, and has

$$S = W \times \Lambda_{m}^{-1/2} \times U_w^T$$  \hspace{1cm} (13)

Where $\Lambda_w, U_w$ are the largest eigenvalues diagonal matrix and their eigenvectors corresponding to covariance matrix $\Sigma = (1/Q) \sum_{i=1}^{Q} (X_i - \bar{X})^T (X_i - \bar{X})$. $W$ was named separable matrix. To obtain matrix $W$, let $W = (w_1, \cdots, w_m)^T$, the steps of weight vector $w_i$ updates by a learning rule as following:

1. Choose a random initial weight vector $w_i(L)$.
2. Let $w_i(N) = E\{\chi^T (w_i^T(L)\chi)\} - E\{\chi^T (w_i(L)\chi)\} w_i(L)$
3. Let $w_i(N) = w_i(N) - \sum_{j=1}^{i-1} w_j^T(N)w_j$
4. Let $w_i(N) = w_i(N) \sqrt{\sum_{j=1}^{i-1} w_j^T(N)w_j}$
5. If $\|w_i^T(N)w_i(L) - 1\| > 0.001$, go back to (2), else go to (6).
6. Update over, $w_i = w_i(N)$.

Where $w_i(L), w_i(N)$ are values of last and current update respectively, $g(u), g'(u)$ are selected by:

$$g = \tanh(a_0u), g'(u) = 1 - (\tanh(a_0u))^2$$
where $g'(u)$ is the first order derivative of function $g(u)$, $a_i = 1$ is taken.

For a given sample $\chi = (\chi_1, \chi_2, \cdots, \chi_n)$, let

$$Y_k = (\chi - \bar{\chi}) s_k, i = 1, 2, \cdots, n, k = 1, 2, \cdots, m$$ (14)

Then, projected feature vectors $Y_1, \cdots, Y_m$ are called the independent principal component of the sample $\chi$. Feature matrix of the image sample $\chi$ can be reduced to $n \times m$ matrix $B = (Y_1, Y_2, \cdots, Y_m)$.

### 3.3 Traffic Light Classification

After feature extracting by 2DICA, a nearest neighbor classifier is adopted for classification. Supposed that traffic light category $c_i$ ($i = 1, 2, 3$) has $N_j$ training samples $B_j^{(i)} = [Y_1^{(i)}, Y_2^{(i)}, \cdots, Y_n^{(i)}]$, ($j = 1, 2, \cdots, N_j$), $N = \sum_{i=1}^{3} N_j$ is the total number of training samples, and that these samples are assigned $c_i$ class.

Supposed that sample $B$ would be recognized, distance decision function of category $c_i$ is defined as

$$D_i(B, B_j^{(i)}) = \sqrt{(B - B_j^{(i)})^T (B - B_j^{(i)})} = \sum_{k=1}^{m} \|Y_k - Y_k^{(j)}\|$$ (15)

where $\|\|_2$ denotes the Euclidean distance between two vectors.

If $D_i(B) = \min_{0 \leq j \leq 3} \{D_i(B, B_j^{(i)})\}$ and $D_i(B) \leq T$, then $B \in c_i$, otherwise sample $B$ is not a traffic light region. $T$ is called similarity threshold.

### 4 Experiments and Analysis

#### 4.1 Experiments Data

A JAI BB-141 camera which equipped with a 25mm fixed mega-pixel lens with 20.4 degree field of view was used to face straight ahead and mounted to front of car roof. Its resolution and frame rate were 1392×1040, 25fps respectively. Since detection algorithm depended primarily on color because no structure was visible at night, the gain and shutter speeds were fixed to avoid saturation of the traffic lights, particularly bright LED-based green lights.

To demonstrate robustness of traffic light recognition system in various traffic light scenarios and under different illumination conditions such as morning, noon, sunset, we drove an intelligent vehicle with the introduced camera on the road of Changsha, China at each of three different times: morning, noon, sunset. The output recognition results included experiment time, date, number of traffic lights and traffic light type (color, arrow direction) were recorded in a text file.
To test traffic light recognition system, experiments were performed on the 25 typical videos collected on roads in urban scenes. The length of each sequence was 200 frames. Traffic light categories included red, green and amber arrow traffic lights.

4.2 Traffic Light Detection and Recognition Results

Figure 5 shown 3 separate examples in different intersections. A red rectangular box surrounded the detected traffic light plate. Traffic light final state (color, arrow direction) was represented by a small picture which lied upper left of the red rectangular box. Several parallel traffic lights (red, amber, green) were correctly detected and recognized by our proposed algorithm in these images. Output of our algorithm was correct since 3 arrow traffic lights were boxed by red rectangle and its arrow direction were shown in top left of traffic light boards. False positive cases were valid reduced because board of traffic light was considered before locating the traffic light, however false negative rate will slightly improve since board was difficult to fix position from complex background like sombre constructions, tree.

Figure 6 illustrated four different scenes that were difficult to detect, since traffic lights were far from camera, too dim to be visible and occluded by some objects such as tree, signboard, buildings etc., images were overexposed or underexposed, the viewing direction between the camera and the traffic lights was quite slant.

Table 1 listed the image frame number under direct sunlight, backlighting, cloudy, sunny conditions as well as their detection and recognition rate by using the proposed algorithm. The detection rate was up to 97.1% in the peak, the recognition rate was slightly lower since traffic lights often only occupy below 10x10 pixels at remote distance from intersections or were obscure due to motion blur which was caused by rapid moving of intelligent car. Assumed that DR, RR were detection rate and recognition rate of the proposed algorithm. The overall recognition rate was defined as following:

\[ OAR = DR \times RR \]  \hspace{1cm} (16)

Overall recognition rate OAR exceeded 91% in our test. In addition to these quantitative results, a significant implication of this work was that our system was able to provide accurate traffic light information for intelligent vehicle driving through intersections.

4.3 Computational Time Analysis

The proposed algorithm was realized under VC++ net. and running environment was a 2.5GHz Pentium(R) Dual-Core CPU with 3GB RAM. The average computational time for the main steps was listed in Table 2. The average frame rate was 6.57fps and achieved near real-time performance in experiments. The response time of our algorithm almost satisfied making a proper decision of intelligent vehicle in the intersection.

Table 2. Computation Time Analysis

<table>
<thead>
<tr>
<th>Step</th>
<th>Computation time(ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image Acquisition</td>
<td>40</td>
</tr>
<tr>
<td>Board of Traffic Light Location</td>
<td>56</td>
</tr>
<tr>
<td>Arrow Traffic Light Detection</td>
<td>21</td>
</tr>
<tr>
<td>Arrow Traffic Light Recognition</td>
<td>35</td>
</tr>
<tr>
<td>Total Computational Time</td>
<td>152</td>
</tr>
</tbody>
</table>

5 Conclusion

Arrow traffic lights are common type in urban scenes and provide more useful information to moving vehicles in the intersections, but existing algorithms mainly recognize circle traffic lights in limited scenes and lacked robustness. A novel pipeline for arrow traffic light state (color, direction) detection and recognition in urban environments is proposed in this paper. Firstly, the boards of arrow traffic lights are accurately localized by black background segmentation and morphology processing. Then candidates of traffic lights are obtained by threshold segmentation and judging relative position between candidate and their boards. Secondly, Gabor wavelet transform is used to extract traffic light candidate’s features due to its best simultaneous localization of spatial and frequency information. 2D independent component analysis is utilized to effectively eliminate redundancy and reduce dimension of the samples. Nearest neighbor classifier is a simple and time saving method to classify object. Experimental results show that the detection and recognition of multiple arrow lights for each intersection significantly are robustness to noise and have higher accuracy.
There are improvement room to be made on this challenging topic. Chrominance of traffic light captured image will vary widely due to relative position and distance among traffic lights camera, and changing illumination condition. Meanwhile, traffic lights sometimes occupy little pixels at long detection ranges, the actual diameter of traffic light lies in 20–30cm generally. Then it is very difficult to steadily detect traffic light from complex urban scene by state-of-the-art image segmentation and object detection. Future work is improvement of detection performance in varied conditions. On the other side of the time cost spectrum, traffic lights are detected in the most probable regions which traffic lights appear in obtained image, if tracker or other previous knowledge is combined with existing image processing algorithms, as false positives and high time cost may be more easily suppressed. Finally, to enhance safety of intelligent vehicle, subsequent important task is to improve robustness and accuracy of existing system.

6 Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grant 61175064, 90820302 and 60805027, Research Fund for the Doctoral Program of Higher Education under Grant 200805330005, Academician Foundation of Hunan Province under Grant 2011FJ4043.

7 References


