Design of traffic light identification scheme based on TensorFlow and HSV color space

To cite this article: Xiaoqiang Wang et al 2018 J. Phys.: Conf. Ser. 1074 012081

View the article online for updates and enhancements.
Design of traffic light identification scheme based on TensorFlow and HSV color space

Xiaoqiang Wang, Xianghui Cheng¹, Xue Wu, Huili Zhou, Xiah Chen and Lin Wang

College of Mechanical and Electrical Engineering, China Jiliang University, Hangzhou 310052, China

¹E-mail: xachen@cjlu.edu.cn

Abstract. This paper proposes an effective traffic light identification scheme based on TensorFlow API and HSV color space. The program uses a distortion-corrected wide-angle camera to collect the road images in real time and utilizes the RCNN target detection interface, which is provided by the Google TensorFlow framework, to track the precise bounding box of the traffic light in the image. The obtained corner information is recorded for cropping ROI image from the original. Then using OpenCV to smooth the ROI image and enhance the contrast. After converting the ROI image from BGR to HSV color space, the result of the traffic light can be consulted on the H channel according to the area of the connected domain.

1. Introduction

The vigorous development of information technology, sensor technology, deep learning and artificial intelligence has given rise to the continuous emergence of technological innovations in autonomous driving and unmanned vehicles [1]. The practical applications are also deepening. In recent years, numerous smart devices and systems have been applied to the automotive industry for either assisted driving or some degree of automatic driving. One of the most important research topics is how to effectively identify traffic signals in urban areas, and timely convey the data to drivers or autonomous driving systems. This will help to reduce traffic accidents and improve driving safety.

Traffic signs in real life are complex and diverse. Except for the common circle of motor vehicle lights, there are pedestrian traffic lights, directional lights and so on. In addition, the image color performance captured by wide-angle cameras varies widely under different weather conditions like day, night or rainy. Ordinary image processing methods often use different image pre-processing methods or filtering pre-treatments, and then identify the result based on the BGR color space [2, 3]. Practicality of such methods is always weak, and the test results are not satisfactory. To solve these problems, this paper divides the recognition of traffic signs into two steps. Firstly, locate the traffic light accurately and cut out the area of interest (ROI) around the location to reduce the calculation, and then through the image processing to achieve the final identification of traffic light.

In this paper, a wide-angle camera is utilized to acquire a wider field of view. The bounding box of the traffic light is available from the image by using the Region Convolutional Neural Network algorithm interface provided by the TensorFlow [4]. Then the ROI region is cut out by using OpenCV and image processing is performed in the HSV color space. After experimental verification, the algorithm can effectively and accurately identify the common types of traffic signals under various weather conditions.
2. Proposed system

The program designed in this paper divides the tasks into two parts: positioning and identification. The positioning section calls the TensorFlow API and then frames out all possible objects in the image, if the target belongs to the traffic light category, only the location information of the traffic light is retained. Then depending on the location information, the cutting, image processing and other work are performed by OpenCV. So, this part belongs to the image pre-processing. The overall design block diagram is given in figure 1.

![Overall system design flow chart.](image)

2.1. Wide-angle distortion correction

The camera in the visual system, due to the lens itself, the captured images tend to have some distortion, especially for the wide-angle camera. To truly restore the objects in the real world, we must correct the distortion. Figure 2 is the most commonly used coordinate system in image processing. The image appears on the two-dimensional plane as the horizontal and vertical axes u and v. If we know the parameters of the camera's internal optical axis, we can make the corresponding conversion [5].

![Conversion of image processing related coordinate system.](image)

Through the above transformation of the relationship, we can get the camera's internal parameter matrix K:
The purpose of calibration is to get these five unknown parameters \((fx, fy, s, u, v)\). Zhang Zhengyou checkerboard grid calibration method can not only avoid equipment demanding, cumbersome operation and other shortcomings, which are the most traditional methods have, but also provide a higher accuracy [6]. Figure 3 is the result of calibration using MATLAB “Camera Calibration” toolbox.

\[
K = \begin{bmatrix}
    fx & s & u \\
    0 & fy & v \\
    0 & 0 & 1
\end{bmatrix}
\]  \hspace{1cm} (2)

2.2. Positioning

TensorFlow (TF) is an artificial intelligence learning system from Google. The latest version of TF provides Faster RCNN-based object detection APIs. Since there is no dedicated public traffic signal dataset currently, TF uses Microsoft's MSCOCO dataset, which contains 90 categories, including traffic light [7]. Call this TF API, we can get the specific four corner coordinates (Bounding Box) of traffic signals on the image [8, 9].

As shown in the figure 4, the RCNN contains a regional generation network that maps the complex features produced by convolution networks back to the original image, generates possible regions, and then classifies each area into a network, which results in a classification. If the category belongs to the traffic light, the network returns the minimum external moment of four corners of the block area.

There are several ways to avoid negative samples from RCNN networks. First of all, for the same resolution and shooting distance, we can set the minimum external moment area threshold to exclude part of the negative sample. Secondly, the aspect ratio of the external moment and the position in the picture can be used to discriminate. Traffic lights are generally horizontal rectangular or vertical
rectangular, in addition, for the front of the traffic lights, the two sides of the position of the external moment can be ruled out.

2.3. Identification
As shown in the figure 5, the ROI image is transferred from the BGR space to the HSV color space, and then the H (hue) component is separated therefrom to make a traffic light determination [10, 11].

![Figure 5. HSV color model.](image)

HSV (Hue / Saturation / Value) is a color space model, according to the visual characteristics of color. HSV is a hexagonal pyramid model. The six vertices and borders of the hexagon represent the color, the horizontal axis represents purity, and the vertical axis represents brightness. Compared with the BGR model, HSV mainly focuses on hue and saturation professional direction. Therefore, HSV color space is more conducive to identify the three colors red, green and blue. The model conversion formula is as follows:

\[
\begin{align*}
    h &= \begin{cases} 
        \text{undefined}, \max = \min \\
        60^\circ \frac{g-b}{\max - \min} + 0^\circ, \max = r, \ g \geq b \\
        60^\circ \frac{b-r}{\max - \min} + 360^\circ, \max = r, \ g < b \\
        60^\circ \frac{r-g}{\max - \min} + 120^\circ, \max = g \\
        60^\circ \frac{r-g}{\max - \min} + 240^\circ, \max = b
    \end{cases} \\
    s &= \begin{cases} 
        0, \max = 0 \\
        \max - \min \quad \text{other} \\
        \frac{\min}{\max} = 1 - \frac{\min}{\max}
    \end{cases} \\
    v &= \max
\end{align*}
\]

Figure 6 shows the original picture of the traffic light, the HSV color space map, and the H/S/V three single channel graphs. It can be seen from the single channel contrast diagram of the red light and the green light that it is difficult to get the correct result through the S and V channels. For H channel, different angle ranges correspond to different colors, for example, \(0^\circ, 120^\circ\) and \(240^\circ\) correspond to red, green and blue respectively. According to the empirical data, the H channel data is red when the value ranges from 160 to 179 and green when it’s between 40 and 70. Since \(360^\circ\) exceeds the range that 8-bit binary can represent, it is common to halve the H value to indicate the color. From the H single channel diagram can also be seen that the green light area is darker, and the red light area is much brighter.
3. Experimental results
To test the final performance of the system, a total of 150 pictures were selected for daytime, rainy and night. The processing time of the picture and the final recognition rate are the main test aspects, the results are shown in table 1:

| Conditions | Resolution  | Total Samples | Correct Numbers | Correct Rate | Time Consuming(s) | Average Time(s) |
|------------|-------------|---------------|-----------------|--------------|-------------------|-----------------|
| Sunny      | 1920*1080   | 50            | 49              | 98%          | 87.72753          | 1.75455         |
| Rainy      | 1920*1080   | 50            | 49              | 98%          | 90.872164         | 1.81744         |
| Night (22:00) | 1920*1080 | 50            | 46              | 92%          | 100.75020         | 2.01500         |

It can be seen from the table that the system has a high recognition accuracy for daytime and rainy pictures. The night's recognition rate declined due to the dark light, which caused the shutter time to be prolonged during filming and a large number of halos appear near the traffic lights. In addition,
because the system model is relatively complex, the processing time of each picture is very long, this is the place where the system needs to focus on optimization in the future.

4. Conclusions
This paper presents a traffic light identification scheme which combines traditional image processing method and TensorFlow API. The RCNN region convolution neural network is used to locate the traffic light position in the image, then the region is cropped out by OpenCV, and the result is identified by analysing the H channel of HSV. Through the identification of daytime, rainy and night traffic lights, the system has high recognition accuracy under high resolution picture and can run any hardware platform with TensorFlow. This system can be used as auxiliary driving system and unmanned technical support, thus greatly improve the driving safety.

Acknowledgements
This research work is supported by National Natural Science Foundation of China (Grant No. 51504228), 2016 postdoctoral research projects of Zhejiang Province of China and the research work described in this paper was supported by the Public Welfare Technology Application Research Project of Zhejiang Province Science and Technology Department (LGG18F030010).

References
[1] Mu K L 2017 Development status and future prospects of driverless cars J. Software Engineering 17 112-114
[2] Yang G and He W X 2005 Study of system on image recognition and tracking of moving object J. Computer Automated Measurement & Control 13 267-269
[3] Sun H, Zhou H X and Li Z H 2002 Study on edge detection technique in image processing J. Computer Development & Applications 15 7-9
[4] Huang, J and Guadarrama, S 2016 Speed/accuracy trade-offs for modern convolutional object detectors J. IEEE Computer Society 3296-3297
[5] Lee, T. Y and Chang, T. S 2011 C. Visual Communications and Image Processing. Wide-angle distortion correction by Hough transform and gradient estimation.
[6] Zhang, Zhengyou 2002 Flexible Camera Calibration by Viewing a Plane from Unknown Orientations C. The Proceedings of the Seventh IEEE International Conference on Computer Vision 11 666-673
[7] Lin, T. Y and Maire, M 2014 J. Microsoft coco: common objects in context 8693 740-755
[8] Felzenszwalb and Pedro F 2010 Object detection with discriminatively trained part-based models J. IEEE Transactions on Pattern Analysis & Machine Intelligence 32 1627-2010
[9] Han F F, Duan F J and Wang K 2010 Study and model construction for camera depth-of-field in machine vision inspection system J. Chinese Journal of Sensors and Actuators 23 1744-1747
[10] Su X, Chen X D and Xu H Y 2017 HSV color space and adaptive shape support window based local stereo matching algorithm J. Laser & Optoelectronics Progress 17 1-13
[11] Rashmi H and Prashanth K G 2014 Automatic Tracking of Traffic Signs Based on HSV C. International Journal of Engineering Research and Technology