Implementation of Attentive Automotive Vision System

Ch. Divya Bhargavi [1], Maddu Srinivasa Rao [2], Dr. D.N.Rao [3]

[1]M.Tech student, Joginpally BR Engineering College, JNTU Hyderabad, India. [2]Associate Professor and HOD, Department of ECE, Joginpally B.R. Engineering College, Hyderabad, India. [3]Principal, Joginpally B.R. Engineering College, Hyderabad, India.

Abstract

In many intelligent transportation systems the applications must reliably work in real-time in a widespread environments, covering rural to urban areas. In this paper we address particular problems in the area that can have a significant impact on people's lives, namely, the detection of sudden pedestrians crossing and traffic signal lights to assist drivers in accord to avoid accidents and to comply with traffic rules. The traffic light signals are detected by CMOS sensor camera in both day and night conditions using predefined HSV values. The detection of sudden pedestrians crossing using Haar cascade Classifier, as early as possible just as they enter the view of the car-mounted camera within the region of interest and to maintain a false alarm rate as low as possible for practical purposes.

1. Introduction

The human activity and action detection, and its analysis has obtained much attention in computer vision because of its wide-range of applications, including content-based image or video retrieval, robotics, assisted living, surveillance, intelligent vehicles, and also the advanced user interfaces. Here, we address certain problems that can have a major impact on people's lives when a vehicle is moving on road, they are the detection of sudden pedestrians crossing and detection of traffic signal lights to avoid the mishaps. According to the traffic safety data from the National Highway Traffic Safety Administration, many people are killed or injured each year in pedestrian–motor accidents; most of which occur when pedestrians attempt a road crossing at non-intersections. The drivers must be alerted about the crossing pedestrians as early as possible for immediate action to be most effective. For this reason, the crossing pedestrians should be identified even before they come into full view. The need for a combination of high processing speed, detection of partially visible pedestrians as they enter the scene who are within the region of interest, rejection of pedestrians who are out of region of interest, and handling of unrestrained camera motion distinguishes this application from related work from current methods in intelligent vehicle systems.

Traffic signal lights are the special perception problem for humans. Efforts have been made to broadcast traffic signal lights state over radio of the locations, and another method is a prior map indication to know presence of the traffic signal lights by the vehicle in the travelling paths, but both of these requires lot of infrastructure and bit tedious to handle. The vision is the only way to detect the state of the traffic signal lights, it may include detection of the sub-elements of the light which illuminate. Although any vision task may be challenging due to the variety of outdoor conditions, the traffic signal lights have been engineered to be highly visible, emissive light sources that eliminate or greatly reduce illumination-based appearance variations. A CMOS sensor camera with an exposure, and aperture can be directly calibrated to traffic light colour levels. Safety is of vital importance in the automotive field. The most common failure conditions in a traffic light detection system are either visual obstructions or false positives such as those induced by the brake lights of other vehicles. Both of these are fail-safe conditions. By using the region of interest of traffic lights the vehicle can detect traffic lights and take conservative action, such as braking gradually to a stop while alerting the driver, when it is unable to observe any other lights. The false positives may rise, which includes false greens, may arise from particular patterns of light on a tree, or from brightly lit billboards. In order to avoid the false positives, the vehicle should predict it in a specific region where the light should appear in the camera image. Tight priors on the prediction region, strict classifiers, temporal filtering, and interaction with the driver can help to reduce these false positives. When there are multiple alike lights visible at a junction, the chance of multiple simultaneous failures is geometrically reduced.

2. Image based or Video based Detection of the Pedestrian

The pedestrian detection can be characterised into two types according to the data source, i.e., image based and video based. Detecting pedestrians in images is a challenging task that has attracted much attention in computer vision since it is only limited to the stationary appearance. In contrast, video based detection consists of additional information in the forms of motion data and depth data, which can be used efficiently within the region of interest (ROI) and can be used together with appearance features to increase classification accuracy. However, both of these methods share the same general outline, which includes candidates region of interest, features representation, and classification. There are some pre-processing (e.g., image smoothing or enhancement and video stabilization) and post processing (e.g., tracking) methods which are essential in achieving high performance. Relevant to other object detection problems, features (representation, selection, and dimensionality reduction) and classification algorithm highly effect detection precision and play vital roles in the whole system. Candidate selection is usually used to balance speed and accuracy, i.e., by quickly filtering out most of the non-pedestrian areas. The final decision is made to further improve precision by removing terminated detections. We analysis related work from the characteristics in the following sections.

2.1. Detection of Candidate in the Region of Interest

The Image or video is retrieved in terms of frames, and the detection of pedestrian depends on the processing speed of the total number of frames per second. Sliding windows at various scales and locations are first examined to detect ROIs based on certain features, which may be global or local, and single or multiple. Then, single or multiple classifiers (e.g., variants of boosting algorithms) are used to judge whether the sliding windows in ROIs bound a person or not. Since the basic sliding window method performs a thorough search, it may obtain a higher accuracy; however, the speed is comparatively low. To elevate speed, some candidate selection methods are used to roughly locate the ROIs that have a higher possibility of containing pedestrians. The OpenCV toolbox, consists of Canny pruning which is used as the candidate selection module for Haar-like feature-based pedestrian detection which is widely used, and this method is used in our system.

2.2. Pedestrian Detection Features

Different features have been used for pedestrian detection, which may be characterised into appearance features or stationary features and motion features. For image-based methods, only appearance features can be used, whereas motion features usually offer more information for video-based pedestrian detection. The Viola and others used Haar-like features to represent both appearance and motion, and trained a cascaded classifier to detect walking pedestrians. However, motion information is hard to use since it also brings in more noise, particularly in unconstrained videos; therefore, some existing systems use only stationary information to detect pedestrians in videos. We can also classify current pedestrian detection systems according to the types of features. Solitary features for instance shapelets, edges, wavelet coefficients, histograms of image patches (e.g., LBP), local representative fields, and Haar-like features. Haar-like features are widely used for pedestrian detection and general object detection in the literature and in our system. Many methods have also been proposed to combine several types of features. For example, different kinds of histogram features can be directly joined to form new features. Diverse features can be used to train some classifiers individually, and a final decision is made by widely held voting or in a cascaded manner. Different features can be carefully chosen and combined by boosting, and a single classifier or a cascaded classifier can be obtained. Furthermore, the structures can be categorized into global ones and local ones. Global features are obtained directly from each sample image, while local features are obtained by dividing a sample image into different sub regions, where each section can be taken as an element from which to obtain one or more kinds of features.

2.3. The Differences of Our Work

In general if the pedestrians are coming across the moving vehicle, then the human or the driver will detect the pedestrian with the naked eye and controls the vehicle manually. Even the sensors are developed to detect the pedestrians. In our system detection of sudden pedestrians crossing is done through the camera and is processed by the ARM 9 processor and controls the vehicle automatically. There are variations from general image based or video based pedestrian detection in two considerable ways. One is the necessity of fast processing to alert the driver as early as possible. Although composite form and motion features from closely sampled sliding windows have

often been used to assure a great recognition level, and this leads to very low recognition speeds that are not suitable for our real-time application purposes. We notice that, for the purpose of notifying drivers, a pedestrian need not be discovered in every single frame in which he or she is present; therefore, we propose a sparse sliding window to speed up the detection process. Another difference from previous works is that our system aims to attentive the drivers only to sudden pedestrian crossings that may take the driver by surprise. So, in these cases the warning alarm plays a most crucial role. We don't consider pedestrians who cross at a far distance from the view of the camera. Our system considers only pedestrians that come in the region of interest over the course of several frames. By considering only pedestrians crossing near to the camera, we can assume that the sudden pedestrian crossing events initiate at the sides of the camera view, which limits the area of each video frame that needs to be examined. By considerably reducing the search space in this manner, it is possible to integrate various computationally expensive features without a major sacrifice in processing speed.
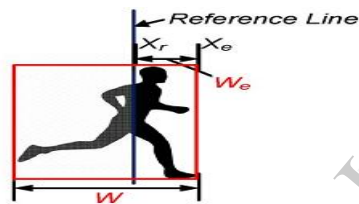

Figure 1. An event of detecting the crossing pedestrian.

3. The Pedestrian Crossing Event

In this paper, a pedestrian crossing action is defined as a spatiotemporal volume that covers a given range of pedestrian perceptibility as she or he enters the camera view. To calculate pedestrian visibility, we first describe the pedestrian entering ratio. As shown in Figure 1, when a pedestrian enters the camera view from the left side, the entering ratio α is defined as

$$\alpha = (X_e - X_r)/W$$

Where X_e is the value of the X-axis right edge of the pedestrian's bounding box, X_r is the value of an X-axis vertical reference line, and W is the horizontal width of the bounding box. In real applications of our system, the reference line is taken as the right or left edge of each video frame. But, here for the testing purpose only the reference line is placed within the frame so that the bounding box is totally observable and the definite entering ratios are known for assessment purposes. We

also define $W_e = X_e - X_r$ as the entering width. We use this arbitrary characterisation of entering ratio to know more easily the variations in entering style. Based on the definition of the entering ratio α, the spatiotemporal volume of a pedestrian crossing event starts from a predefined threshold α_e and ends when the entering ratio reaches a certain threshold α_l.

4. Pedestrians Datasets for Training and Testing

Many pedestrian images have been collected by many research groups, and numerous videos have recently become available. However, these data sets do not cover many sudden pedestrian crossings and are not applicable for our application.

Figure 2. Various cases of samples, with different entering angles/directions, with/without obstructions, and with/without crosswalks.

4.1. Image Data Set Training

The classifiers used for training at frame level, and compiled a set of 2500 grayscale side view images of pedestrian, which differs from standard pedestrians of the front view. To acquire the training image set, we firmly crop the side-view pedestrian images from many pedestrian training or testing sets, Internet images and movies. As shown in Figure 2, the samples of training consists of people with diverse entering angles and poses, obstructions, and crosswalks.

4.2. Video Data Set Testing

The video data set testing was attained by using a vehicle mounted with the high-definition video camera of 1440 × 1080 resolution and 25 fps frame rate. The video clips of people, who were instructed to walk, run from left to right and right to left in front of the moving vehicle are taken. As per our awareness, this might be the first HD pedestrian data set captured from a moving vehicle. Distinct from 640× 480 videos in which adults are less than 80 pixels of height around a distance of 30 m from the camera, adults in our data set are about 120 pixels tall at the same distance. The statistics of this data set are given in the following.

4.2.1. Sudden Pedestrian Crossing Event Statistics.

A sudden crossing event is defined to start at an entering ratio of α_e=0.25 and to end at an entering ratio of α_l =1.5. For an entering ratio of less than 25%, it is problematic even for humans to say whether it is a pedestrian or not. So the pedestrian detection size limitations in terms of W x H or in terms of pixels are predefined.

5. Traffic Light detection

Traffic signal lights detection is based on colour, using either the red–green–blue (RGB) or the hue–saturation–value (HSV) colour space makes to have high detection rates. But, they also tend to return a large number of false positives as it covers the whole viewing angle of the camera. In contrast, methods that include templates are less prone to errors in the day time, and are not good at night (i.e., most of these methods assume three dominant circular edges, and some also model the rectangular backboard, which are hard to observe at night time due to the low dynamic range in cameras), and tend to have lower overall detection rates.

Figure 3. Traffic light detection in the region of interest to control the vehicle.

The processing used herein for traffic light detection and feature association. Special attention is given to the processing time, detection rate, and reliability of detection (i.e., false association between expected and noticed features). Most modern traffic lights use an array of LEDs sitting behind a lens that gives a selected wavelength, depending on the LEDs and the lens. Light output from a red light will also have some orange hue, and the light output from a green light will also have some blue hue, this may cause colour blindness. To identify candidate light regions from a colour camera that represents each pixel on an image as RGB components, a rule based approach is applied to each pixel on the image. Let I (u, v):$I^2 \rightarrow I^3$ be a pixel on the image I, where the output vector$[I_R, I_G, I_B]$ represents the RGB components of the

pixel at location (u, v). Each of the output components is 8 b, with integer values between 0 and 255.

Traffic light detection utilizing both Region of Interest (ROI) and template matching. The intelligent vehicle must also comply with traffic rules in an urban environment, and traffic lights are important signals for a vehicle driving in an urban environment. The automatic red light signal detection on a video or in real time environment in the candidate region of interest of the traffic light. The CMOS camera is used to carry on a recognition testing of an LED light. The traffic light detection methods are based on the intelligent transportation system, with the development of intelligent vehicle technology, few traffic light recognition methods based on the vehicle-mounted camera have been proposed, and made use of the vehicle mounted with camera to carry out the video sequence processing and realize the traffic signal recognition for each image frame. The algorithm used a hue and saturation model according to a Gaussian distribution and acquired the statistical model parameter by analysing the collected samples of data. The process of the recognition was done with the feature of the template in the candidate region. Then the traffic signal recognition using detection and classification was done. This method made use of the colour space to get the candidate region of the traffic light and mark the region. For this approach, the recognition distance should not be far. Mostly, the traffic signal recognition methods mainly include the colour-based method, and the method based on template matching.

Figure 4. Detection of traffic light colors to control the vehicle

For an instance, consider sky as the background, the colour-based recognition method can effectively notice and identify traffic light. For a relatively complex situation, such as an urban road environment, false detections will appear easily using the colour-based recognition method. The shape-based feature recognition method can reduce the false detections of the colour-based feature recognition. But, the different shape characteristic rule has to be created

for the different styles of traffic lights. This limits the flexibility of the algorithm. Even though the styles of traffic lights are different, they are mainly composed of red, yellow and green. In the urban environment, the vehicle has to comply with the instruction of the traffic light. Therefore, this study aims to the recognition and following of the traffic light signals based on vehicle mounted with cameras. The traffic light recognition algorithm combined with colour segmentation for a complex urban environment is presented to achieve recognition. So in this project within the region of interest the traffic signal lights are detected, by using HSV values and the predefined size constraints. If the detected colour of the traffic light signal is red then the ARM9 processor controls the moving vehicle automatically by applying brakes, and if the detected colour is green then the vehicle goes to normal state, by releasing the brakes of the vehicle.

6. Conclusion

In this paper, the development process of pedestrian detection and traffic light signals detection is introduced. The system is developed on ARM 9 processor by using Linux operating system in which Qt creator is used which provides integrated development environment (IDE) within that C++ language and OpenCV libraries are used for writing the code. By utilizing necessary algorithms and library functions, the automatic pedestrian and traffic light signals detection, and controlling of the vehicle is done.

References

[1] Stephan Matzka, Andrew M. Wallace, and Yvan R. Petillot, "Efficient Resource Allocation for Attentive Automotive Vision Systems," IEEE Transactions on Intelligent Transportation Systems, Vol. 13, No. 2, JUNE 2012, pp. 859 – 872.

[2] N. Bellotto and H. S. Hu, "Multisensor-based human detection and tracking for mobile service robots," IEEE Trans. Syst., Man, Cybern, vol. 39, no. 1, Feb. 2009, pp. 167–181.

[3] Cheng-Chin Chiang, Ming-Che Ho, Hong-Sheng Liao, Andi Pratama and Wei-Cheng Syu, "Detecting and Recognizing Traffic Lights by Genetic Approximate Ellipse Detection and Spatial Texture Layouts," IJICIC, Vol. 7, No. 12, Dec. 2011, pp. 6919 – 6934.

[4] Jianfeng Wang, "Video-Based Face Detection Using New Standard Deviation," Advances in Future Computer and Control Systems, Advances in Intelligent and Soft Computing Vol. 159, 2012, pp 187-192.

[5] http://opencv.willowgarage.com

[6] http://www.friendlyarm.net/products/mini2440

[7] http://qt-project.org/wiki/Category:Tools::QtCreator

[8]http://docs.opencv.org/modules/objdetect/doc/cascade_classification.html

[9]http://docs.opencv.org/modules/core/doc/drawing_functions.html#rectangle

[10]http://www.alldatasheet.com/datasheet-pdf/pdf/83786/SAMSUNG/S3C2440A.html

[11]http://www.andahammer.com/assets/Uploads/All2440/LinuxDNW.pdf

Divya Bhargavi. CH: She is Pursuing Master's in Electronics and Communication Engineering specialization in Embedded Systems from Joginpally BR Engineering College affiliated to JNTU Hyderabad. She has obtained B.Tech degree in Electronics and Communication Engineering from Sri Padmavathi Mahila University, Tirupati.

Maddu Srinivasa Rao: He has obtained Master Degree in Electronics and Communication Engineering specialization in Computers and Communications from JNTU Kakinada Andhra Pradesh and BE degree in Electronics and Communication Engineering from Sri Ramanand Theerth Maratwada University , Nanded, Maharastra. He is pursuing Ph.D in the field of Computer Networks in JNTUH. He was ratified as Associate Professor by Jawaharlal Nehru Technological University, Hyderabad, Andhra Pradesh. He is presently working as Associate Professor & HOD in the Department of ECE in Joginpally B.R. Engineering College, Moinabad, Hyderabad and has an experience of 14 years both in Industry & Teaching**.**

Dr. D.N.Rao: His carrier spans nearly three decades in the field of teaching, administration, R&D, and other diversified in-depth experience in academics and administration. He has actively involved in organizing various conferences and workshops. He has published over 11 international journal papers out of his research work. He presented more than 15 research papers at various national and international conferences. He is currently approved reviewer of IASTED International journals and conferences from the year 2006. He is also guiding the projects of P.G. /Ph.D students of various universities.