MOTORCYCLE HELMET DETECTION IN VIDEO

ABSTRACT

In order to ensure the safety measures, the detection of traffic rule violators is a highly desirable but challenging task due to various difficulties such as occlusion, illumination, poor quality of surveillance video, varying whether conditions, etc. The helmet is the main safety equipment of motorcyclists, however many drivers do not use it. The main goal of helmet is to protect the drivers head in case of accident. In case of accident, if the motorcyclist does not use can be fatal.
In general, the processing framework of this kind of visual surveillance in dynamic scenes includes thefollowing stages: modeling of environments, detection of motion,classification of moving objects(mtorcycle), tracking, extraction of region of interest, encoding into feature, classification of head and helmet. We review recent developments andgeneral strategies of all these stages. Finally, we analyze possible research directions, e.g. applying new techniques of object segmentation like yolo and MaskRCNN, dealing with occlusion etc.

Index Terms: Helmet; Feature extraction; Classification; Traffic Surveillance; Deep Learning; Convolutional Neural Network.

INTRODUCTION

In the last couple of years alone most of the deaths in accidents are due to damage in the head. Because of this wearing helmet is mandatory as per traffic rules, violation of which attract hefty fines. Despite this, a large number of motorcyclists do not obey the rules. Presently, all major cities already deployed large video surveillance network to keep a vigil on a wide variety of threats. Thus using such already existing system will be a cost efficient solution, however these systems involve a large number of humans whose performance is not sustainable for long periods of time. Recent studies have shown that human surveillance proves ineffective, as the duration of monitoring of videos increases, the errors made by humans also increases.

Over the past years many works were carried out in traffic analysis, including vehicle detection and classification, and helmet detection. Intelligent traffic systems was usually implemented using vision computer algorithms, such as: background and foreground image detection to segment moving objects in scene and image descriptors to extract features. Computational intelligence algorithms are used too, like machine learning algorithms to classifier the objects. Next, some related works to helmet detection are presented.

Wen et al. suggested a circle arc detection method based upon the circular Hough transform. They applied it to detect helmet on the surveillance system of the Automatic Teller Machine. The weakness of this work is that they only use geometric features to verify if any safety helmet exists in the scene. Geometric features are not enough to find helmet. The head can be mistaken with a helmet because both had a circular-like form.

In another paper Chiu et al. proposed a computer vision system aims to detect and segment motorcycles partly occluded by another vehicle. A helmet detection system is used, and the helmet presence determines that there is a motorcycle in a scene. In order to detect the helmet presence, the image edges are computed in the possible helmet region. The Canny edge detector is used. The quantities of edge points which are similar to a circle define a helmet region. The method needs so much information (helmet radius, camera angle and height) that must be provided by user.

In another paper Chiverton described and tested a system for tracking and automatic classification of riders with and without helmet. The system uses the SVM classifier, which is trained with the attribute vectors of the Histogram of Oriented Gradients (HOG) descriptor. The HOG descriptor is executed with the Sobel operator for calculating the edges with a neighborhood of 3 × 3 pixels. The algorithm proposed in is used to calculate the background. Still photographs and individual image frames from the video data are used on the histograms extraction. The method achieved a total accuracy rate of 83% for classification of motorcycles and 85% for helmet detection. It is noticed that the amount of images in the testing phase is small.

Silva et al. proposed a system in which he tracks the vehicles using Kalman filter. An important advantage of this Kalman tracking system is the ability to continue to track objects even if they are lightly occluded but when there were more than two or three motorcyclists appear in a same frame, Kalman filter fails because Kalman filter mostly works well for linear state transitions (i.e tracking single objects/one object at a time). But to track multiple objects, we need non-linear functions to track them. Recently, Dahiya et al. proposed a system which first uses Gaussian mixture model to detect moving objects. This model is robust to slight variations in the background. It uses two classifier in serial, one for separating motorcyclist from moving objects and another for separating without helmet from the upper one fourth part of the
motorcyclists.

SYSTEM OVERVIEV

The system is usually divided in three steps: 1) Moving objects segmentation; 2) Moving objects classification; 3) Helmet detection.

A. Moving Objects Segmentation

The moving objects segmentation in scene allows that only objects of interest in the image are evaluated. This causes a reduction in processing time due to the fact that the algorithm works in a small area of the image. It also reduces the probability of false-positives in classifiers. A background (BG) image is necessary to segment the moving objects. From BG is possible know the moving pixels. We used a video camera to capture the traffic images. In environments where the static objects change during the time (parked vehicles along the roads, changes position of the shadow, etc), the algorithms to calculate adaptive BGs are necessary. This way, we update the image background using the Adaptive Mixture of Gaussians (AMG).

B. Recognition of Motorcyclists from Moving Objects

The next step consists of classify the moving objects resulting from the segmentation procedure. In pattern recognition works it is necessary to extract image features at first. This way, we use descriptors to extract image features.

The descriptor returns a set of values as result. This set is called “feature vector”. After this, classifiers are usedto recognize the image. The feature vector is used by the classifier to separate the objects in two classes: motorcycle and non-motorcycle. The objects are classified only into two groups just because it is enough to know that each object corresponds to a motorcycle or not. In the moving objects detection and classification steps we use the same strategy of . Thus, we do not show the results of this classification.

C. Helmet Detection

After the motorcycle detection, we made the helmet detection step is perfomed. This step is divided in four phases:

1) Region of interest (RoI) determination;
2) Sub-window calculation;
3) Features extraction;
4) Image classification.

The RoI aims to compute a small region to reduce the computational cost and improve the helmet search. After, it is obtained a sub-window that contains the motorcyclist head region. Next, the features were extracted from the sub-window the classifiers are used.

REGION OF INTEREST DETERMINATION

The RoI determination is an important step in our problem. It is used to search the helmet. The motorcyclist head must be inside the RoI. Using RoI decreases the computational cost the search area. In our database, all images present the motorcyclist head region inside the RoI. The RoI was specified at the image top and corresponds to 1/5 of the image height. This value was chosen empirically. For improving the coverage some other researchers have also used the top 1/4 th portion but 1/5 proved to be enough for many applications.

SUB - WINDOW CALCULATION

In order to improve further the helmet search we computed a sub-window, that contains the exact motorcyclist head region. This way, the image classification is improved. The reason is the descriptor will extract features of the head region, only. Before computing the descriptor a set of preprocessing is required, all this preprocessing was made in the RoI. At first, it is computed the grayscale image.
After the grayscale image be calculated, a average filter 5 × 5 is applied to reduce the noise. At next, the Otsu threshold is computed. It is applied in grayscale image to get a binary image (black and white). In the binary image is applied the Sobel algorithm, that returns the image edges. The next step is the CHT (Circular Hough Transform) calculation to find circle-like regions. It was used because the motorcyclist head region has a circle-like shape. The CHT computation with edges is computationally faster and has a better precision. The CHT was set with the computed edges and get the possible circles. One of the parameters required by the CHT is the circle size fetched. The radius size was half of the RoI height. That value was chosen because a helmet has not a size larger than the RoI. From the found circles a search is made to return the best circle (circle with more points in the CHT computation).

A strategy that use only geometric information, like CHT, do not returns good results because the head and the helmet have similar shape. Therefore, other features are necessary to separate the problem class. For this, a sub-window from RoI is computed, that will correspond to the square circumscribed in the circle found.

FEATURE EXTRACTION

A simple **marsk rcnn** or **yolo** approach can be used to directly get out the segment containg the subjects head or helmet other than that previously tried approaches use image transformations to achieve this. In order to extract image features, Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP) and Wavelet Transform (WT), were used as descriptor.

A. Histogram of Oriented Gradients

The Histogram of Oriented Gradients (HOG) algorithm is a feature descriptor that calculates an image histogram of oriented gradients. The final descriptor is an one-dimensional array of histograms extracted from the image. The algorithm is based on the local object shape and appearance, which in an image can be represented by intensity gradients or edge directions. The feature extraction can be done without an edge position foreknowledge.

B. Local Binary Pattern

The Local Binary Pattern (LBP) has a good performance in many applications, including classification and texture, image recovery and surface inspection. The original LBP labels the image pixels for a threshold neighborhood of 3 × 3. Each pixel is compared with the central pixel and the result is the binary number.

C. Wavelet Transform

The Wavelet Transform (WT) is an alternative approach in relation to Fourier transform for image processing. The WT allows that both frequency and time information are presented.

CLASSIFICATION OF HELMETED AND HEAD

In this step we use six classifiers to compare the results: Support Vector Machines (SVM), MultiLayer Perceptron (MLP), Radial Basis Function Network (RBFN), Naive Bayes, Random Forest and K-Nearest Neighbors (KNN). These classifiers were used because they are the most present in the literature. A feature vector is generated for each sub-window. The input of the classifiers is the feature vector computed in the feature extraction step. The images were divided into two groups: with-helmet and without-helmet. We can also use CNN's instead of the usual classifiers and HOG Features. A convolutional neural network (CNN) is a variant of feed forward neural networks using back propagation algorithm. It learns high-level features from the spatial data like image. The recent widespread success of convolutional neural networks is in it's ability to extract inter-dependant information from the images i.e localization of the pixels which are highly sensitive to other pixels. The convolutional neural network training consist of convolution layers, relu layers maxpooling layers, fully connected layers and a loss function (e.g. SVM/Softmax) on the last (fully-connected) layer. In the primary layers we get the edge information of the images similar to some of the handcrafted algorithms but, In the final layers, we start getting texture and ridge information which helps us in getting sensitive information usefull for classification . To recognize motorcyclists without helmet, from the images of motorcyclists, we cropped only the top one fourth part of the image as that was the region where the motorcyclist's head is located most of the time. From this, we locate the portion of the head by subtracting the binary image of the foreground of same region. Then we build a CNN model in order to separate the without-helmet from the with-helmet images. This model is trained for the binary classification of helmet and head. Fig. 3 shows the feature maps of the sample helmets. These feature maps illustrate that the CNN learns the common hidden structures among the helmets in the training set and thus able to distinguished between a helmet and a head.

DIRECTION OF RESEARCH

Most of the researchers haven't used the new techniques of MaskRCNN and Yolo for object segementation due to the limitation of the type of dataset and satisfactory results with other techniques however more exploration shall be done in the field of CNN which could lead to better performance on occluded images and dark images.

REFERENCES

1. K. Dahiya, D. Singh, and C. K. Mohan, "Automatic detection of bike-riders without helmet using surveillance videos in real-time," in Proc. Int. Joint Conf. Neural Networks (IJCNN), Vancouver, Canada, 24–29 July 2016, pp. 3046–3051.

2. D. Singh, C. Vishnu, and C. K. Mohan, "Visual big data analytics for traffic monitoring in smart city," in Proc. IEEE Conf. Machine Learning and Application (ICMLA), Anaheim, California, 18–20 December 2016.

3. C. Behera, R. Ravi, L. Sanjeev, and D. T, "A comprehensive study of motorcycle fatalities in south delhi," Journal of Indian Academy of Forensic Medicine, vol. 31, no. 1, pp. 6–10, 2009.

4. W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews, vol. 34, no. 3, pp. 334–352, 2004.

5. C.-C. Chiu, M.-Y. Ku, and H.-T. Chen, "Motorcycle detection and tracking system with occlusion segmentation," in Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services, Santorini, Greece, 6–8 June 2007

6. J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," IET Intelligent Transport Systems (ITS), vol. 6, no. 3, pp. 259–269, 2012.

7. R. Silva, K. Aires, T. Santos, K. Abdala, R. Veras, and A. Soares, “Automatic detection of motorcyclists without helmet,” in Proc. Latin American Computing Conf. (CLEI), Puerto Azul, Venezuela, 4–6 October 2013

8. W. Rattapoom, B. Nannaphat, T. Vasan, T. Chainarong, and P. Pattanawadee, “Machine vision techniques for motorcycle safety helmet detection,” in Proc. Int. Conf. Image and Vision Computing New Zealand (IVCNZ), Wellington, New Zealand, 27–29 November 2013

9. R. V. Silva, T. Aires, and V. Rodrigo, “Helmet detection on motorcyclists using image descriptors and classifiers,” in Proc. Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, 27–30 August 2014