REAL TIME VEHICLE DETECTION AND TRACKING USING FUSION APPROACH: A REVIEW

[1]THATHUPARA SUBRAMANYAN KAVYA, [2]YOUNG-MIN JANG, [3]ERDENETUYA TSOGTBAATAR, [4]TAE-HUN NAM, [5]SANG-BOCK CHO

[1,2]Department of Electronics Engineering, University of Ulsan, 93, Daehak-ro, Nam-gu, Ulsan, 44610, South Korea
E-mail:kavyats86@gmail.com

Abstract- Vehicle detection and tracking play an important role in road traffic. In this paper, we propose an efficient fusion based technique for vehicle detection and tracking. The design of fusion approach includes two CCD cameras with two different pre-processing techniques and image fusion with DWT is used to avoid the drawback of the other detection methods. The main aim of this paper is more accurate real time detection and tracking system based on the various combination of the techniques .This review paper explains how to improve the efficiency of vehicle detection and tracking using Fusion Approach.

Keywords- Fusion Approach, Image Processing, Real Time vehicle Detection.

I. INTRODUCTION

Thousands of drivers and passengers ,lose their lives each year due to automotive accidents in the last few decades[1], various research approaches have been applied in this particular area of detecting vehicle data but still the gap in these as it needs improvements in detection and tracking for accurate prediction .

Vehicle detection and tracking play an effective and significant role in the field of intelligent transportation, video surveillance , driving assistance etc. It will improve the safety and efficiency of the transportation system.

The main idea of this paper is to make a real time tracking and detection system with 100% accuracy.

- Low quality images from camera
- Most of the vehicle detection system is not sufficient because of shading and shadows of images.
- A Different view of the vehicle is not detecting properly.

To overcome these challenges, in this paper we proposed a fusion based approach.

- High quality CCD cameras are used to capture the videos
- Parallel preprocessing is proposed
- High quality Image Fusion technique implementation
- Efficient post-processing and Features Extraction .

The main focus of this paper is to improve the possibilities for vehicle detection and tracking in real time.

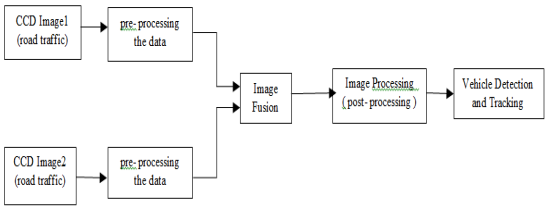


Fig1: Proposed fusion based approach

II. RELATED WORKS

Issam, Elati, Mohamed and Noureddine have proposed a new method for real time approach is established based on the particle filter and background subtraction [2] ,This approach is able to detect and track automatically multiple moving objects without any learning phase or prior knowledge about the size, nature or the initial position. Although the approach has some limitations as background subtraction is generally based on a static background hypothesis which is often not applicable in real environments. e.g due to wind, rain or illumination changes brought by weather[3,4].

Niluthpol,Tamanna presented a novel vehicle tracking methods for a traffic management system that introduces the multiple time spatial images (MTSI) based detection in the stochastic filter based tracking .The MTSI-based tracking employs the concept of a multiple numbers of key vehicles frames(KVFs) for each of the vehicular objects in the traffic. These KVFs provide highly accurate positional information of the vehicle due to the fact that the shape and texture of the vehicles are comparable on the same scale and do not depend on the speed of the traffic, The MTSI based approach can also be augmented to be used in problem like recognition of traffic signals, and localization and re-identification of vehicles [5].

Xuerui,Xue,Jing and Liping Proposed some algorithms of image defogging and color enhancement in order to improve the performance of vehicle detection systems in different weather conditions. However the performance of this system is significantly affected by the training subject is used to train the vehicle detector [6].

Tao and Shijian used a soft discriminative mixture of viewpoint (SDMoV) models for joint vehicle detection and vehicle viewpoint estimation .A two – step mixture model learning approach is proposed, where a set of viewpoint specific component models

are first trained to optimize vehicle viewpoint classification ,and a mixture model is then learned to optimize vehicle detection .[1]

Fernando, David, Arturo and Jose proposed a sensor fusion methodology for intelligent vehicles ,able to overcome the limitations of each sensor, the presented approach achieves safer roads by data fusion technique, especially in single-lane carriageways where casualties are higher than in other road classes, and focuses on the interplay between vehicle drivers and intelligent vehicles. The system is based on the reliability of laser scanner for obstacle detection [7].

Wei, Qian and Jun proposed a real time object detection method based on ORBP features and soft cascade SVM. This approach allows for the complex scenario and illumination change, a new feature descriptor method is put forward to overcome the difficulty. Furthermore the soft cascade SVM multi-level classification threshold is constructed by adaptive feature selection, which improves speech detection robustness and the performance of the system. but this paper has some limitations further studies can be made on improving the accuracy of feature selection and developing more appropriate cascade architectures[8].

III. METHODOLOGY

The proposed real time vehicle detection system makes use of video data acquired from stationary traffic cameras, performing mathematical operations over a set of frames obtained from the video to detect the vehicles present in a scene based on fusion approach.

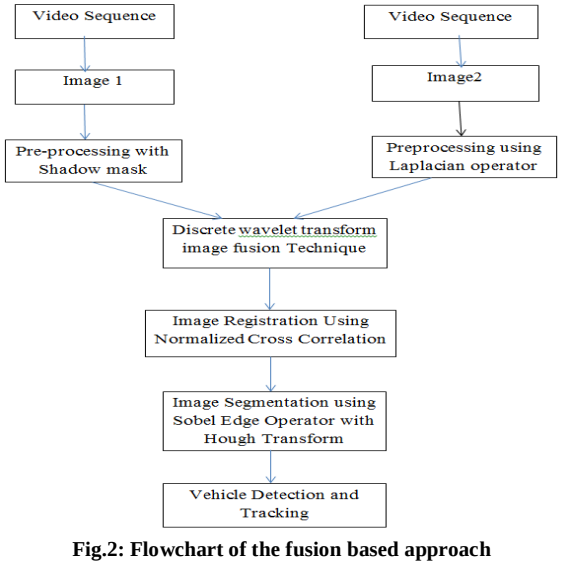


Fig.2: Flowchart of the fusion based approach

3.1 Video Surveillance system

Video surveillance applications are interested in the real time observation of vehicle in the road traffic. There are different approaches suggested in the literature for video surveillance .In this paper we are using Two CCD (Charge Coupled Device) Cameras. CCD is the heart of the digital camera. The primary advantage of CCD imaging is sensitivity. The more sensitive the camera the shorter the exposure required to detect faint detail. CCD cameras have a greater dynamic range than digital SLRs, meaning they can more easily capture both faint and bright detail in a single exposure and less hardware may be required for CCD imaging [9].

3.2 Pre-processing

In this fusion based approach, one of the important parts is the multiple pre- processing techniques, the captured videos are converted into images and these images applied to the preprocessing unit. Image pre-processing is an important phase of the segmentation procedure. Pre-processing of an image includes both data size reduction and image enhancement. In the first preprocessing we proposed a shadow mask processing technique .it will remove the shades and shadow of the image, and second pre-processing technique we are using contrast stretching[10] and Laplacian operator will improve the quality of the image[11].

The Laplacian operator is good at finding the fine details in an image. Any feature with a sharp discontinuity will be enhanced by a Laplacian operator.

The Laplacian of an image f(x,y) ,denoted $\nabla^2 f(x,y)$, is defined as

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2} \quad (1)$$

Commonly used digital approximations of the second derivatives are

$$\frac{\partial^2 f(x,y)}{\partial x^2} = f(x+1,y) + f(x-1,y) - 2f(x,y) \quad (2)$$

And

$$\frac{\partial^2 f(x,y)}{\partial y^2} = f(x,y+1) + f(x,y+1) - 2f(x,y) \quad (3)$$

so that

$$\nabla^2 f(x,y) = [f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1)] - 4f(x,y) \quad (4)$$

This expression can be implemented at all points in an image by convolving the image with the following spatial mask

$$\begin{matrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{matrix}$$

An alternate definition of the digital second derivatives takes into account diagonal elements, and can be implemented using the mask

$$\begin{matrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{matrix}$$

Both derivatives sometimes are defined with the signs opposite to those shown here, resulting in mask that is the negatives of the preceding to masks.

Enhancement using the Laplacian is based on the equation

$$g(x,y) = f(x,y) + c[\nabla^2 f(x,y)] \quad (5)$$

Where f(x,y)is the input image, g(x,y) is the enhanced image, and c is 1 if the center coefficient of the mask is positive or -1 if it is negative

3.3 Image Fusion

Image fusion means combining two or more images into a single image. The resulting image will be more informative than any of the inputs to get a single image [12]. For the mixing operation, we are using a DWT (discrete wavelet transform) approach.

Image fusion methods can be broadly classified into two groups, spatial domain fusion and transform domain fusion [13].

The fusion method such as averaging, brovey method, principal component analysis (PCA) and HS based method fall under spatial domain approaches. Another important spatial domain fusion method is the high pass filtering based technique.

The disadvantages of spatial domain approaches are that they produce spatial distortion in the fused image. The spatial distortion can be very well handled by frequency domain approaches on image fusion .the discrete wavelet transform has become a very useful for image fusion. The DWT fusion method has least spectral distortion; it also provides better SNR than pixel based approaches.

DWT is $, f(n) = \frac{1}{\sqrt{M}}\sum_k W_\emptyset[j_0,k]\emptyset_{j0,k}[n] + \frac{1}{\sqrt{M}}\sum_{j=j_0}^{\infty}\sum_k W_\varphi[j,k]\varphi_{j,k}[n]$ (6)

Here $f(n), \emptyset_{j0,k}[n] and \varphi_{j,k}[n]$ are discrete functions defined in [0,M-1],totally M points. Because the sets $\{\emptyset_{j0,k}[n]\}k \in Z$ and $\{\varphi_{j,k}[n]\}_{(j,k)\in Z^2, j\geq j_0}$ are orthogonal to each other. We can simply take the inner product to obtain the wavelet coefficients.

$W_\emptyset[j_0,k] = \frac{1}{\sqrt{M}}\sum_n f(n)\emptyset_{j0,k}[n]$ (7)

$W_\varphi[j,k] = \frac{1}{\sqrt{M}}\sum_n f(n)\varphi_{j,k}[n] \quad j \geq j_0$ (8)

3.4 Image post processing

Image post processing enhances the quality of the image; it includes techniques to clean up images to make them visually clearer as well as the application of filters and other treatments to change the look and feel of an image

3.4.1. Image Registration

Image registration is the process of overlaying two or more images of the same scene taken at different times, from the different view point, and by different sensors. It is widely used in remote sensing, computer vision etc [14].Jignesh ,Suprava and Salman described a template matching algorithm, for image registration based on NCC(Normalized cross correlation)using Cauchy-Schwartz inequality relying on similarity measures. It is clear that NCC is the ideal approach to image registration by template matching.[15]

Cross Correlation is the basic statistical approach to image registration.

The cross correlation of two continuous functions ,

$\emptyset(u,v) = \int_{-\infty}^{\infty} u^*(\tau - t)v(\tau)d\tau$ (9)

If we define $x(t) = u^*(-t)$then

$x(t) * v(t) = \int_{-\infty}^{\infty} x(t-\tau)v(\tau)d\tau = \int_{-\infty}^{\infty} u^*(\tau - t)v(\tau)d\tau = \emptyset(u,v)$ (10)

Fourier Transform of x(t):

$X(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}dt = \int_{-\infty}^{\infty} u^*(-t)e^{-i2\pi ft}dt$ (11)

$= \int_{-\infty}^{\infty} u^*(t)e^{i2\pi ft}dt = \left(\int_{-\infty}^{\infty} u(t)e^{-i2\pi ft}dt\right)^*$ (12)

$= U^*(f)$

So $\emptyset(u,v) = x(t) * v(t) => \emptyset(f) = X(f)V(f)$

$= U^*(f)V(f)$ (13)

Relation between convolution and cross –correlation

$\emptyset(u,v) = u^*(-t) * v(t)$ (14)

Note that ,unlike convolution, correlation is not associative or commutative

$\emptyset(v,u) = v^*(-t) * u(t) = u(t) * v^*(-t)$ (15)

Normalized Cross –Correlation

Correlation:

$w(t) = \emptyset(u,v) = \int_{-\infty}^{\infty} u^*(\tau - t)v(\tau)d\tau$ (16)

If we define $y(t) = u(t - t_0)$for some fixed t_0,then $E_y = E_u$

$E_y = \int_{-\infty}^{\infty}|y(t)|^2dt = \int_{-\infty}^{\infty}|u(t-t_0)|^2dt$ (17)

$= \int_{-\infty}^{\infty}|u(\tau)|^2d\tau = E_u$ (18)

Cauchy-Schwarz inequality

$:\left|\int_{-\infty}^{\infty} y^*(\tau)v(\tau)d\tau\right|^2 \leq E_yE_v$ (19)

$=\gg |w(t_0)|^2 = \left|\int_{-\infty}^{\infty} u^*(\tau - t_0)v(\tau)d\tau\right|^2 \leq E_yE_v = \leq E_uE_v$

But t_0 was arbitrary ,so we must have $|w(t)| \leq \sqrt{E_uE_v}$ for all t

We can defined the normalized cross correlation

$z(t) = \frac{\emptyset(u,v)}{\sqrt{E_uE_v}}$ (20)

3.4.2 .Image segmentation

Image segmentation is the process of portioning a digital image into multiple segments. The goal of segmentation is to simplify or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries in the image.

The Sobel operator performs a 2-D Spatial Gradient measurement on an image and so emphasis regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient at each point in an input grayscale image.

The operator consists of a pair of 3 x 3 convolution kernels, one kernel is simply the other rotated by 90^0.

Gx

+1	0	-1
+2	0	-2
+1	0	-1

Gy

-1	-2	-2
0	0	0
+1	+2	+1

These kernels are designed to respond maximally to edges running vertically and horizontally relative to the pixel grid, one kernel for each of the two perpendicular orientations. The kernels can be applied to the input image, to produce separate measurements of the gradient component in each orientation(call these Gx and Gy) .these can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient.

The gradient magnitude is given by

$$|G| = \sqrt{Gx^2 + Gy^2} \quad (21)$$

Typically, an approximate magnitude is computed using

$$|G| = |Gx| + |Gy| \quad (22)$$

This is much faster to compute.

The angle of orientation of the edges (relative to the pixel grid) giving rise to the spatial gradient is given by

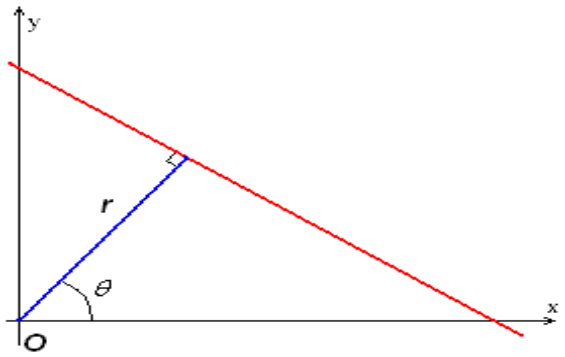
$$\theta = \arctan(Gy/Gx) \quad (23)$$

In this case, orientation 0 is taken to mean that the direction of maximum contrast from black to white runs from left to right on the image and other angles are anti clockwise from this [16].

In practice there may be breaks in boundary and spurious intensity discontinuities due to nonuniform illumination and the presentence of noise. To overcome this limitation in this paper we propose a global processing technique which is Hough Transformation [17].The purpose of the Hough transform is to address this problem by making it possible to perform groupings of edge points into object candidates by performing an explicit voting procedure over a set of parameterized image objects.

Hough transformation is detecting straight lines.In general, the straight line y=mx+b can be represented as a point (b,m) in the parameter space. However vertical lines pose a problem .they would give rise to unbounded values of the slope parameters m.

$$r = x\cos\theta + y\sin\theta \quad (24)$$



Where 'r' is the distance from the origin to the closest point on the straight lines, and θ is the angle between the x-axis and the line connecting the origin with that closest point.it is, therefore possible to associate with each line of the image a pair(r,θ).the (r,θ) plane is sometimes referred to as Hough space for the set of a straight line in two dimensions. This representation makes the Hough Transform.

Given a single point in the plane ,then the set of all straight lines going through that point corresponds to a sinusoidal curve in the (r,θ) plane, which is unique to that point .A set of two or more points that form a straight line will produce sinusoids which cross at the (r,θ) for that line. Thus, the problem of detecting collinear points can be converted to the problem of finding concurrent curves.

CONCLUSION

In this paper explained a new fusion based approach for real time vehicle detection and tracking. To overcome the limitations of the conventional methods we are proposed an image fusion method. It consists in using discrete wavelet transform method for image fusion and Sobel operator with Hough Transformation for image post processing. The advantage of using DWT for image fusion is highlighted in this paper. The illumination changes, shadows and environmental noise etc.. It will create big challenges for researches to make a decision. In this approach with the two preprocessing technique it will remove the complete shadows and also it will improve the quality of the image, and we can easily detect any vehicles without any delay .The Combination of the various technique used in this fusion method to improve the quality of the accurate detection and tracking of the vehicle.

REFERENCES

[1] Tao Chen, Shijian Lu., "Robust Vehicle Detection and Viewpoint Estimation With Soft Discriminative Mixture Model", IEEE Transactions on circuits and systems for Video technology,Vol.27,No.2,February 2017

[2] Issam Elafi, Mohamed Jedra, Noureddine Zahid, "Unsupervised detection and tracking of moving objects for video surveillance applications," Science Direct- Pattern Recognition Letters 84 (2016) 70–77

[3] Swati N.Divatankar,Umesh N.Hivarkar,"A Survey and Comparative Study of Real time Vehicle Detection Methods for Road Traffic applications", International Journal of Computer Applications(0975-8887) National Conference on Digital Image and Signal Processing 2016

[4] Prem Kumar Bhaskar,Suet-Peng Yong,"Image Processing Based Vehicle Detection and Tracking Method", IEEE International conference 2014

[5] Niluthpol Chowdhury Mithun,Tamanna Howlader,S.M. Mahbubur Rahman, "Video-based tracking of vehicles using multiple time-spatial images", Science Direct- Expert Systems With Applications 62 (2016) 17–31

[6] Sebastiano Battiato, Giovanni Maria Farinella, Antonino Furnari, Giovanni Puglisi,"An integrated system for vehicle tracking and classification", Science Direct- Expert Systems with Applications 42 (2015) 7263–7275

[7] Fernando Garcia, David Martin, Arturo de la Escalera, and Jose Maria Armingol,"Sensor Fusion Methodology for vehicle Detection", IEEE Intelligent transportation systems magazine, Spring 2017

[8] Wei Zhu, Qian-Liang Fu, lun-Qi Bai,"The real-time object detection algorithm based on ORBP and cascade SVM",IEEE conference 2016.

[9] Tu-Thach Quach,"Vehicle Track Segmentation Using Higher Order Random Fields", IEEE Geoscience and Remote Sensing Letters,Vol.14,No.3,March 2017

[10] Xuerui dai,Xue Yuan, Jing Zhang ,Liping Zhang,"Improving the performance of Vehicle Detection system in Bad weathers",IEEE International Conference 2016.

[11] Naoya Kosaka , Gosuke Ohashi,"Vision-Based Nighttime Vehicle Detection Using CenSurE and SVM",IEEE Transactions on intelligent transportation syatems,Vol.16,No.5,October 2015.

[12] Yuh-Rau Wang, Wei-Hung Lin, Shi-Jinn Horng, "A sliding window technique for efficient license plate localization based on discrete wavelet transform", Science Direct- Expert Systems with Applications 38 (2011) 3142–3146

[13] K Bavithra ,X Asha,R S Rampriya, "Survey on Object Detection and Tracking Using Fusion Approach", International Journal of Innovative Research in Computer and Communication Engineering Vol.4,Issue 3,March 2016.

[14] Barbara Zitova, Jan Flusser, "Image registration methods:a survey", Elsevier- Image and Vision Computing 21 (2003) 977–1000.

[15] Jignesh N Sarvaiya, Dr. Suprava Patnaik, Salman Bombaywala,"Image Registration Matching Using Normalized Cross- Correlation",International Conference on advances in Computing ,Control, and Telecommunication Technologies 2009

[16] Weihua Wang, "Reach on Sobel operator For Vehicle Recognition", International Joint Conference on Artificial Intelligence 2009

[17] Vmha Kamat , Subramaniam Ganesan, "An Efficient Implementation of the Hough Transform for Detecting Vehicle License Plates Using DSP'S",IEEE International Conference 1995

★★★