2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India

Vehicle Trajectory Prediction using a Catadioptric Omnidirectional Camera

Vigneshram Krishnamoorthy
Dept. of ECE
National Institute of Technology, Trichy
Tiruchirappalli, India
ram7sachin@gmail.com

Saksham Agarwal
Dept. of Electrical Engineering
Indian Institute of Technology Kanpur
sakshamagarwals@gmail.com

K.S. Venkatesh
Dept. of Electrical Engineering
Indian Institute of Technology Kanpur
venkats@iitk.ac.in

***Abstract*—A practical method is presented to predict the future spatial-temporal trajectories of multiple vehicles at road intersections in real time using a catadioptric omnidirectional camera equipped with an Equiangular mirror. Tracking is done using CamShift algorithm running alongside a Kalman Filter to handle occlusions. Domain transformation of the tracked objects location and velocity from image space to real world is done using a geometrical model. A computationally effective model for trajectory prediction has been presented along with the experimental results obtained using it. Applications such as collision prediction and vehicle tracking or any other event of interest using a dual-camera system are also discussed briefly.**

***Keywords*-Trajectory prediction, Omnidirectional camera,Kalman and CAMShift tracking.**

I. Introduction

With the increase of vision based traffic surveillance, substantial developments have been made in the sensors which are used to capture images. Omnidirectional cameras enable us to monitor a wide area without having to stitch images from multiple cameras to get the same information. Catadioptric systems made of standard camera combined with a parabolic, hyperbolic, elliptical or a custom shaped mirror provide a 360-degree FOV in horizontal and about a 100-degree in elevation. Object tracking implemented using the Camshift algorithm, such as in [1], work well except in the case of occlusion which is common in a traffic scene. Motion based tracking algorithms, such as a Kalman filter provides a way to handle occlusion of vehicles for a short duration of time due to its robust motion prediction model.Dual-camera systems, such as an omnidirectional camera paired with a PTZ camera, are often used to track vehicles [2]. The former can provide a wide field of view while the latter has a good resolution. Trajectory prediction for vehicles can be crucial in video surveillance of such systems, since the future trajectory of the vehicle is essential for a PTZ camera to track objects in real time.

Omnidirectional camera was used first in 1990 for robot localisation [3]. Since 2000, with improved manufacturing techniques, their reduced size has allowed their widespread use in surveillance, robotics, panoramic art, etc. [2], [4], [5], [6] and [7] discuss the image formation and projection models of catadioptric omnidirectional cameras in detail. Daniillidis and Geyer [8] developed a unified model for all conventional kinds of central catadioptric systems - parabolic, hyperbolic and elliptical. [9] discusses a calibration technique for omnidirectional cameras independent of mirror shapes. Equiangular and other non-central systems have been discussed in-depth in [7] and [5]. Much work has been done in object tracking [10]. [11] and [12] are widely used algorithms for Background Subtraction. CamShift is described in [13] and [14]. Not much work has been done in prediction of future trajectories of vehicles realtime using an omnidirectional cameras.

In our work, we provide a real-time traffic surveillance tool,which uses background subtraction followed by an object tracking algorithm integrating CamShift and Kalman Filter, to make the model immune to occlusion.The use of omni-directional cameras is justified as they provide a wide field of view gathering more scene information in one shot.A projection model for the omni-directional camera helps map the image space coordinates to the real world. The experiment results show the accuracy of our tool, which has a error rate of less than 5 percent. The applications of our work such as collision prediction based on these predicted trajectories, tracking a vehicle using our system coupled with a PTZ camera are also discussed below.

II. Tracking Multiple Objects

A. Background Subtraction and Foreground Extraction

Background subtraction algorithm using improved adaptive Gaussian Mixture Models [11] has been used to detect moving objects. It is invariant to the variable background conditions and allows detection of shadows. Morphological opening, closing, filling and thresholding operations are applied to the resulting binary mask for de-noising. The connected blobs are joined and are considered as a single object. Hence, the background and the foreground are separated and the moving blobs are identified and labeled.

B. CAMShift and Kalman Filter integration

The image now consists of a set of labeled blobs which correspond to moving objects. Individual tracks are maintained for each blob centroid. The tracking is accomplished using CamShift algorithm running alongside a Kalman Filter

978-1-5090-2029-4/16/$31.00 ©2016 IEEE

with a constant acceleration model. In case of occlusion, the color information of the object is not available and hence CamShift fails. Thus, as illustrated in [15], we calculate the Bhattacharya Coefficient for each frame. If it exceeds the threshold, we assume the object is occluded and therefore use Kalman Filters prediction for the next frame. Otherwise the CamShift prediction is used.

C. Track Updation

The tracks are maintained using a suitable cost function [16]. New tracks are created for unassigned detections and are labeled. The tracks which go undetected for a number of frames greater than a specific threshold are assumed to have gone out of the scene and are deleted.

III. Catadioptric Omnidirectional Camera

Catadiopric cameras combine a standard camera with a shaped mirror, like parabolic, hyperbolic and elliptical mirrors. These provide 360-degree field of view in the horizontal plane and more than 100-degree in the vertical plane [9]. Many cameras also have mirrors having specialised shapes depending on the applications. We use a camera with an Equiangular mirror as it provides better resolution than conventional mirrors. Equiangular mirrors are designed so that each pixel spans an equal angle irrespective of its distance from the center of the image. A unified projection model is provided in [8] for all central conventional catadioptric systems. The equiangular mirror is a non-central imaging system where the viewpoint varies with the angle of incoming rays and lies along a short locus within the mirror known as caustic [5]. But in practice, the variation in viewpoint is very small and the following central model very well approximates such mirrors. Point **P**(x,y,z) in the real world forms an image at **Q**(u,v) in the image plane. The mirror center is at the origin. **P** makes angle θ from the principal axis of the mirror and an angle ϕ with the x-axis. The reflected ray makes an angle ψ with the principal axis. The parameters α and k depend on the mirror and the focal length of the camera respectively.

$$\theta = \alpha\psi \tag{1}$$

$$r = \sqrt{u^2 + v^2} \tag{2}$$

$$\theta = \tan^{-1}(\sqrt{x^2 + y^2}/z) \tag{3}$$

$$r = k\tan(\psi) \tag{4}$$

$$u = r\cos(\phi) \quad v = r\sin(\phi) \tag{5}$$

$$\phi = \tan^{-1}(y/x) \tag{6}$$

$$x = \beta u \qquad y = \beta v \tag{7}$$

$$\beta = z\tan(\alpha\tan^{-1}(r/k)) \tag{8}$$

IV. Spatial Trajectory Prediction

Each tracked object is assumed to be a point object located at its centroid.The history of the locations of the centroid of the track are used in trajectory prediction. The tracked coordinates (u,v) obtained are filtered and the image space velocity (u,v) is computed using a simple differentiator.

The mathematical model mentioned in the previous section is used to map the location and velocity of the tracked objects from the image space to real world. The spatial temporal trajectories for the objects are found in the real world using the laws of motion. The predictions are then back-projected to the image space.

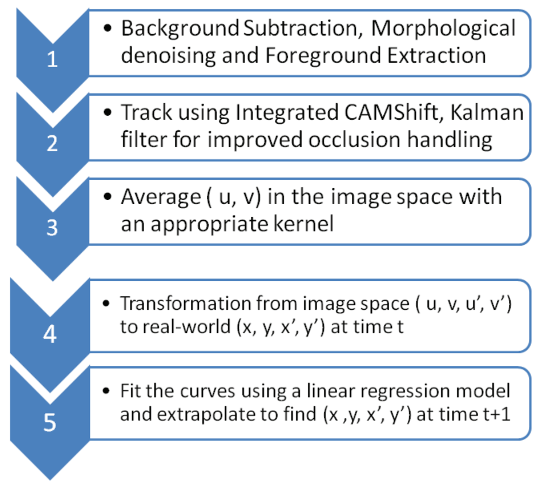

Figure 1. Implementation steps

A. Calibration to find the parameters

The unknown parameters are found using calibration making use of the visual cues available in the scene. A set of curves were found in the image space which on transformation should give straight lines. This is established with the help of markers on the road which are straight lines in the real world. The approximate slope of this line was estimated by finding the slope of the curve, in the region of the image space that is least distorted.

B. Spatial Trajectory Prediction

The real world locations of the tracked objects are found using the transformation (u,v) to (x,y) given above (z being a known constant). We have assumed that the motion of vehicles on the road for a very short duration of time to be a straight line. The observed real world trajectory for the past few frames is best fit using a basic linear regression model which extrapolates the curve and gives us the prediction (x,y) for the future frames. We can safely assume that the car will trace this predicted straight line for the next few frames. However, the probability of the predicted trajectory to be in line with the real world trajectory drops over time due to uncertainty in human actions. Hence, a correction mechanism is required in order that the trajectory will correct itself over time. This is achieved by looping the trajectory prediction algorithm over every few frames, and hence comparing the latter predictions with the former and correcting the error.

C. Spatial-Temporal Matching

Finding the time coordinate is extremely crucial for surveillance of the object or for any event of interest. Now that we know the future trajectory of the object, we should be able to know the time coordinate for each point of the predicted trajectory.

This time coordinate will depend on the velocity of the object. The real world velocity vectors x and y are obtained by taking the derivative of x and y, and the found as a function of time using linear regression.The velocity vectors are taken for a past few frames and extrapolated, assuming constant acceleration (if any) of the vehicle.

The extrapolation gives the expected velocity vectors for the upcoming frames. The accuracy of the extrapolated curve decreases with time and hence needs to be updated over every few frames to make the necessary corrections.

Integrating this curve over the current frame to i frames ahead gives us the predicted location (x, y) of the object in the real world. Back-projecting (x,y) using the projection model gives us the coordinates (u, v) in the video for the ith frame ahead. Hence (u,v,t) coordinates(t denoting time) can be tabulated.

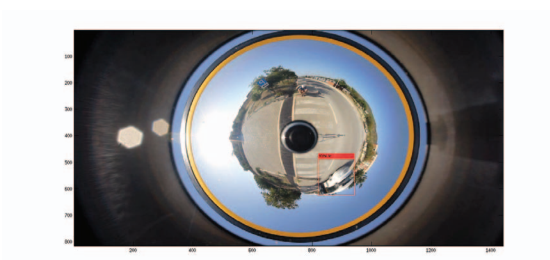

Figure 2. Tracked object

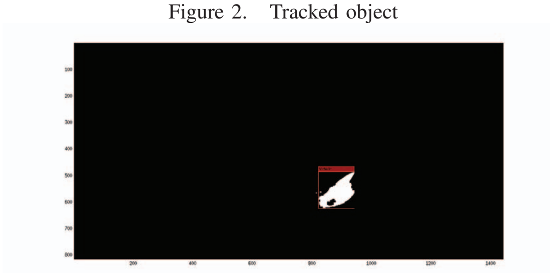

Figure 3. Resulting Binary Mask

V. Experimental Results

We tested our model on the omnidirectional video dataset available on [17]. The results are provided for the tracking and trajectory prediction of vans taking different paths. The frame rate of the videos in the dataset is 15 fps. The resolution of the video is 1440 x 816 pixels. The binary mask of the vehicle is obtained as shown in figure 3 based on background subtraction, followed by preprocessing. For each frame n, the predicted trajectory for the next 10 frames was found using the data of the trajectory traversed by the object in the previous 10 frames as can be seen from figure 4. The predicted points (u,v) are compared with the points from the real trajectory , obtained from frames n+1 to n+10 of the video as shown in figure 5.The mean squared error was calculated taking into account that the probability of prediction decreases over time as given in table I. The error was found to be minimal in cases where the vehicles do not change course suddenly. Under ideal conditions, this error can be safely assumed to be constant and time-invariant.

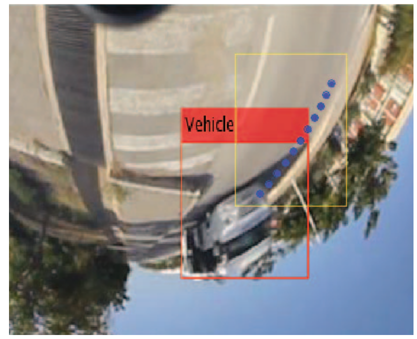

Figure 4. Frame 70-Object tracking result is shown with the bounding rectangle on the vehicle with the predicted locations of the centroid of the tracked object for the frames 75 to 85 shown in circles.

Figure 5. Frame 80-Object tracking result is shown with the bounding rectangle on the vehicle with the object traversing the previously predicted locations of the centroid of the tracked object (frames 75 to 85)

Table I
The Mean squared error in the predicted trajectory for various paths.Maximum error is less than 5 pixels

Average error rate (in pixels)	Path 1	Path 2
Vehicle I	4.69	3.83
Vehicle II	3.97	2.15
Vehicle III	2.89	3.65

VI. Applications and Future Work

Given that the system provides the spatial-temporal coordinates, a system can be made to predict collisions between moving objects in the scene (vehicles, pedestrians etc.). The probability that the prediction is true increases when more data points are available.The intersection of the spatial-temporal trajectories provides us the space and time of collision. Not only that, we can also find cues as to which vehicle was at fault as we have the critical data such as the speed of the vehicle and the path in which it travelled on the road.

By calculating the space and time coordinates using our model, we can record any event of interest with a PTZ camera. Projection models are available [18] to get pan and tilt angles needed to focus the camera towards the collision area. Thus, we can confirm if the collision has indeed occurred, enabling us to send messages in case of accidents to emergency services more quickly.

VII. Conclusion

The system works in real-time predicting the trajectory of a fast moving vehicle.The projection model for the catadioptric camera maps the image coordinates to real world coordinates. The model presented can also be extended for other kinds of omni-directional cameras using the projection models given in [2],[4],[5],[6] and [7]. The model can be improved taking into account various previously known scene parameters like road width, intersections and pedestrian area, etc.

References

[1] Hongxia Chu, Shujiang Ye, Qingchang Guo, and Xia Liu. Object tracking algorithm based on camshift algorithm combinating with difference in frame. In *2007 IEEE International Conference on Automation and Logistics*, pages 51–55. IEEE, 2007.

[2] Simon Baker and Shree K Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):175–196, 1999.

[3] Yasushi Yagi and Shinjiro Kawato. Panorama scene analysis with conic projection. In *Intelligent Robots and Systems' 90.'Towards a New Frontier of Applications', Proceedings. IROS'90. IEEE International Workshop on*, pages 181–187. IEEE, 1990.

[4] O Faugeras, Ryad Benosman, and Sing Bing Kang. *Panoramic vision: sensors, theory, and applications*. Springer Science & Business Media, 2013.

[5] Peter Corke. *Robotics, vision and control: fundamental algorithms in MATLAB*, volume 73. Springer, 2011.

[6] Davide Scaramuzza. Omnidirectional camera. In *Computer Vision*, pages 552–560. Springer, 2014.

[7] Peter Sturm, Srikumar Ramalingam, Jean-Philippe Tardif, Simone Gasparini, and Joao Barreto. Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends® in Computer Graphics and Vision*, 6(1–2):1–183, 2011.

[8] Christopher Geyer and Kostas Daniilidis. A unifying theory for central panoramic systems and practical implications. In *European conference on computer vision*, pages 445–461. Springer, 2000.

[9] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *Fourth IEEE International Conference on Computer Vision Systems (ICVS'06)*, pages 45–45. IEEE, 2006.

[10] Alper Yilmaz, Omar Javed, and Mubarak Shah. Object tracking: A survey. *Acm computing surveys (CSUR)*, 38(4):13, 2006.

[11] Pakorn KaewTraKulPong and Richard Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. In *Video-based surveillance systems*, pages 135–144. Springer, 2002.

[12] Chris Stauffer and W Eric L Grimson. Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 2. IEEE, 1999.

[13] Gary R Bradski. Real time face and object tracking as a component of a perceptual user interface. In *Applications of Computer Vision, 1998. WACV'98. Proceedings., Fourth IEEE Workshop on*, pages 214–219. IEEE, 1998.

[14] John G Allen, Richard YD Xu, and Jesse S Jin. Object tracking using camshift algorithm and multiple quantized feature spaces. In *Proceedings of the Pan-Sydney area workshop on Visual information processing*, pages 3–7. Australian Computer Society, Inc., 2004.

[15] Shengluan Huang and Jingxin Hong. Moving object tracking system based on camshift and kalman filter. In *Consumer Electronics, Communications and Networks (CECNet), 2011 International Conference on*, pages 1423–1426. IEEE, 2011.

[16] MathWorks. *Motion-based multiple object tracking*. 2015.

[17] Y. Bastanlar. Omnidirectional video dataset for vehicle classification. 2015.

[18] Chung-Hao Chen, Yi Yao, David Page, Besma Abidi, Andreas Koschan, and Mongi Abidi. Heterogeneous fusion of omnidirectional and ptz cameras for multiple object tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 18(8):1052–1063, 2008.