Shared-Use Motion Vector Algorithm for Moving Objects Detection for Automobiles

Chup-Chung Wong[1,2], *Member, IEEE*, W.C. Siu[3], *Fellow, IEEE*, S. Barnes[1], P. Jennings[1] and B. Fong[2]
[1]University of Warwick, Coventry, UK. [2]Hong Kong Productivity Council.
[3]Hong Kong Polytechnic University.

***Abstract*--This paper presents a new motion vector based moving object detection system for use in Advanced Driver Assistance Systems. By the shared-use of motion vectors with a video encoder, the computational cost for motion estimation using optical flow method can be saved, making the system more cost effective. Moving objects are detected by evaluating the planar parallax residuals of macroblocks in the video. The proposed algorithm involves newly proposed "APD" constraints on hypothesis generation and template matching in hypothesis verification. Test results show that moving objects can be detected effectively in scenarios that pose danger to the ego vehicle.**

I. Introduction

H.264/AVC encoders [1] are widely used in many consumer products. The successful re-use of the MVs from the encoder for moving object detection enables the recording function by default, saves the computational cost for motion estimation (ME) using optical flow and reduces the hardware cost by 30% by eliminating the Digital Signal Processor for optical flow estimation [2]. There have been many proposed methods to detect moving objects on the road by detecting the inherent features of these objects or using optical flow [3]. But there have been few studies on the use of block based motion vectors (MVs) from video encoders for moving object detection on a moving platform. This algorithm can work in companion with previously proposed methods for relatively slow speed moving object detection [4].

The major contributions of this paper is on the techniques developed to re-use the erroneous and limited precision MVs from a typical H.264/AVC encoder for moving object detection on a moving platform.

II. The Algorithm

MVs from the encoder are output for each 8x8 block which is the same as the technique proposed in [4].

A. Planar Parallax Residual (PPR) Estimation

In this paper, the subscripts "*t-1*" and "*t*" denote previous and current frame respectively, superscripts "*c*" and "*s*" denote camera and screen coordinates respectively.

Fig. 1 shows the conceptual flow chart of the proposed algorithm. For a 3D scene point P_w being seen by a moving camera at p^s_{t-1} and p^s_t respectively, there is a screen point $p_{2G}=M^{-1}p^s_t$ in the previous frame which is the point correspondence of the ground plane projection of p^s_t calculated by the homography matrix M^{-1}. The PPR μ is defined as the difference between the induced 2D image motion of a 3D scene point between two successive images [5], as expressed in (1). The technique mentioned in [6] was used in this paper for the estimation of matrix M^{-1} using (2) and (3), where R_c and R_w are the rotational matrix of the camera coordinates, and the World to the camera coordinates respectively, T_c is the translational matrix, n^T and h are the normal vector and mounting height of the camera to the ground plane respectively, I is a 3x3 identity matrix..

$$\mu = p_{2G} - p^s_{t-1} \tag{1}$$

$$p^c_t = Ap^c_{t-1} = R_c\left(I + R_w T_c n^T / h\right) p^c_{t-1} \tag{2}$$

$$p^s_t = Mp^s_{t-1} = KAK^{-1} p^s_{t-1} \tag{3}$$

The FOE P_{FOE} at (x_v,y_v) can be expressed by (4) [7], where f is the focal length of the camera, V_x, V_y and V_z are the velocity along *X*-, *Y*- and *Z*-axis of the World coordinates respectively. (x_0,y_0) are the vanishing point of the camera. It can be expressed as $[x_0\ y_0\ 1]^T = KR[0\ 0\ 1]^T$, where R is the camera's rotational matrix.

$$\left(x_v, y_v\right) = \left(x_0 + fV_x / V_z, y_0 + fV_y / V_z\right) \tag{4}$$

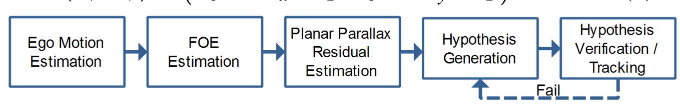

Fig. 1. Conceptual block diagram of the proposed algorithm

B. HG for Detection

Each MV represents an image block of size 8x8. Each PPR needs to pass the "APD" constraints for hypothesis generation.

1) The APD Constraints

"A" for Amplitude: Those PPRs with strong planar parallax with amplitudes larger than (*w*/2) and the threshold $\mu_e(x,y)$ expressed in (5) are retained. t_d is the frame interval time, *w* is the block size which is equals to 8 in this case.

"P" for Position: PPRs represent the displacement of the screen point in the frame interval time. It can be extrapolated to estimate its position on the screen after a time period *T*. An area in the image in front of the ego vehicle is defined as the alert zone. Those extrapolated PPRs entering to the zone after time *T* are retained as they will pose danger to the ego vehicle.

"D" for Direction: Knowing that PPRs pointing to the FOE are either moving in parallel to the ego vehicle or belonging to static objects [7], a PPR $\mu(x,y)$ can be excluded if the angle between the PPR and the line from *(x,y)* to the FOE is smaller

This research is part of the project funded by the Innovation and Technology Commission of HKSAR Government and the Hong Kong Automotive Parts and Accessories R&D Centre under the R&D Project ITP/015/13AI.

978-1-4673-8364-6/16/$31.00 ©2016 IEEE

than a threshold m_{thres}.

$$\mu_e(x,y) = (t_d / 2)\sqrt{(x-x_0)^2 (y-y_0)^2} \tag{5}$$

The PPRs obtained by the APD constraints are further clustered to represent different objects. PPRs on the same rigid object have similar amplitude and direction, the clustering of PPRs is done by comparing the amplitude, direction and distance of PPRs to some thresholds. The cluster with the largest area is selected for HV.

C. HV and Tracking

For the selected cluster in HG, its corresponding Canny edge image is created as the reference template. The sum of absolute difference (SAD) is used as the cost function for block matching to the Canny edge image created from the newly captured image. A simple spiral search of window size ±16 pixels is performed around the average displacement of the MVs of the selected cluster. The HV is successful and a moving object is detected if a local minimum is found inside the search window. The next offset for template matching is updated to the displacement found in the current frame. The template is updated from the Canny edge image of the current frame. Tracking is then performed by the same template matching technique.

III. Test Results

The JM18.4 H.264/AVC encoder was used for encoding and MV output. Six sequences shown in Fig. 2 were prepared with the image and dynamic data captured simultaneously. Since the algorithm is targeted to detect relatively fast speed moving objects, there is no known test result that can be compared directly. The scenarios for testing include a moving object near road junctions (A-C), travelling in opposite direction crossing the driving lane (D) and fast cut-in (E-F). The evaluation parameters include the distance between the detected object and the ego vehicle (denoted by Z_D), time-to-collision (TTC) upon the first successful detection, and the number of false positive frame (FP). TABLE I shows the evaluation result. The last row of the table showing "A to F" is the result by combining all the test sequences.

The test results indicate that the algorithm is able to detect moving objects successfully, leaving enough time for the driver to respond. All sequences have TTC larger than 1.9 second, leaving enough time for the driver to react. Since the detection is non-parametric, the cluster size varies according to the number of detected PPRs, rather than the size of the moving object. With the use of those realistic image sequences, the false positive rate per frame was 0.0072. It can be reduced in the future by checking the moving direction of the detected object.

IV. Conclusion

A novel algorithm to detect moving objects is proposed in this paper. It makes use of the MVs from H.264/AVC encoder for the evaluation of PPRs. The proposed "APD" constraints can effectively filter out unwanted PPRs so that only useful PPRs are clustered for subsequent HV and tracking. Test results show that the algorithm is able to detect moving objects under different driving scenarios without prior knowledge of the shape of the object to be detected. This method solves the problem on the shared use MVs from video encoder for moving object detection. It is also suitable to work in companion with other vehicle detection algorithms.

TABLE I
TEST RESULTS USING THE PROPOSED ALGORITHM

Seq.	No. of Frame	FP	Z_D	TTC	Seq.	No. of Frame	FP	Z_D	TTC
A	24	0	18.5	2.0	D	89	1	27.5	3.1
B	66	0	21.2	1.9	E	179	0	10.4	>4
C	59	2	26.9	3.0	F	139	1	12.1	>4
D	89	1	27.5	3.1	A to F	556	4	n.a.	n.a.

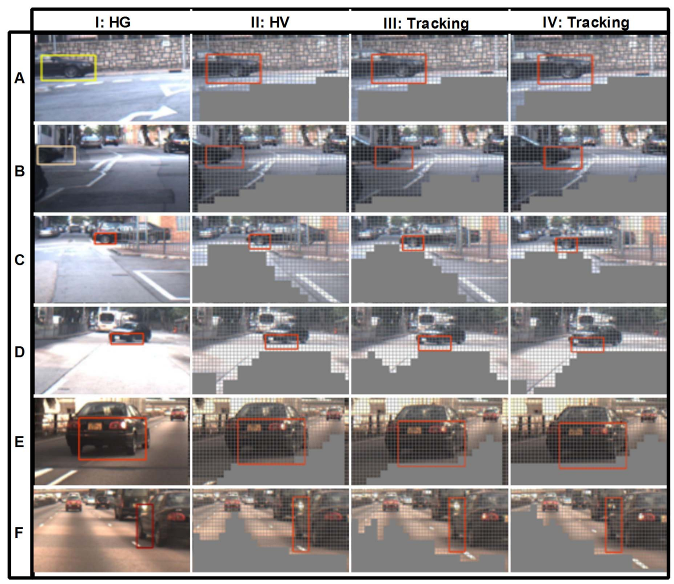

Fig. 2: Test results on different scenarios showing successful HG(I) to HV(II), followed by successful tracking (III) to (IV)

V. Reference

[1] M. Y. Chiu and W. C. Siu, "Computationally-scalable motion estimation algorithm for H.264/AVC video coding," *IEEE Trans. on Consumer Electron.,* vol. 56, pp. 895-903, 2010.

[2] F. Zhang, Y. Gao, and J. D. Bakos, "Lucas-Kanade Optical Flow estimation on the TI C66x digital signal processor," in *IEEE High Perf. Extreme Comput. Conf. (HPEC)*, pp. 1-6, 2014.

[3] S. Sivaraman and M. M. Trivedi, "Looking at Vehicles on the Road: A Survey of Vision-Based Vehicle Detection, Tracking, and Behavior Analysis," *IEEE Trans. on Intell. Transp. Syst.,* vol. 14, pp. 1773-1795, 2013.

[4] C. C. Wong, W. C. Siu, P. Jennings, S. Barnes, and B. Fong, "A smart moving vehicle detection system using motion vectors and generic line features," *IEEE Trans. on Consumer Electron.,* vol. 61, pp. 384-392, 2015.

[5] M. Irani, P. Anandan, and M. Cohen, "Direct recovery of planar-parallax from multiple frames," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* vol. 24, pp. 1528-1534, 2002.

[6] C.-C. Wong, W.-C. Siu, S. Barnes, and P. Jennings, "Low relative speed moving vehicle detection using motion vectors and generic line features," in *IEEE Intl. Conf. on Consumer Electron.*, Las Vegas, pp. 208-209, 2015.

[7] M. I. Lourakis and S. C. Orphanoudakis, "Using planar parallax to estimate the time-to-contact," in *IEEE Comput. Soc. Conf. on Comput. Vision and Pattern Recognition*, 1999.