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

Chup-Chung Wong1,2, Member, IEEE, W.C. Su1, Fellow, IEEE, S. Barnes1, P. Jennings1 and B. Fong2
1University of Warwick, Coventry, UK. 2Hong Kong Productivity Council. 3Hong 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 $p'_{t-1}$ and $p'_t$, respectively, there is a screen point $p_{2G} = M^{-1} p'_t$ in the previous frame which is the point correspondence of the ground plane projection of $p'_t$, calculated by the homography matrix $M^{-1}$. The PPR $\mu$ 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, $n^2$ and $h$ are the normal vector and mounting height of the camera to the ground plane respectively, $I$ is a 3x3 identity matrix.

$$\mu = p_{2G} - p'_{t-1} \tag{1}$$

$$p'_t = Ap'_{t-1} = R_c \left( I + R_c T_c n^2 / h \right) p'_{t-1} \tag{2}$$

$$p'_t = M p_{ss} = K A K^{-1} p'_{t-1} \tag{3}$$

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 $[x_0, y_0]^T = KR[0 0 1]^T$, where $R$ is the camera’s rotational matrix.

$$\begin{bmatrix} x_v \ y_v \end{bmatrix} = \begin{bmatrix} x_0 + fV_x / V_z \ y_0 + fV_y / V_z \end{bmatrix} \tag{4}$$

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 $\mu(x,y)$ expressed in (5) are retained. $t_i$ 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
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.

### 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).

#### TABLE I. TEST RESULTS USING THE PROPOSED ALGORITHM

<table>
<thead>
<tr>
<th>Seq. No. of Frame</th>
<th>No. of Frame</th>
<th>FP</th>
<th>ZD</th>
<th>TTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>24</td>
<td>0</td>
<td>18.5 2.0</td>
<td>D</td>
</tr>
<tr>
<td>B</td>
<td>66</td>
<td>0</td>
<td>21.2 1.9</td>
<td>E</td>
</tr>
<tr>
<td>C</td>
<td>59</td>
<td>2</td>
<td>26.9 3.0</td>
<td>F</td>
</tr>
<tr>
<td>D</td>
<td>89</td>
<td>1</td>
<td>27.5 3.1</td>
<td>A to F</td>
</tr>
</tbody>
</table>

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

V. REFERENCE


