

Comprehensive Vehicle Motion Analysis using Optical Flow Optimization Based on Pulse-Coupled Neural Network

Yanpeng Cao*, Alasdair Renfrew**, and Peter Cook***

*Control System Centre, University of Manchester, Manchester,
M60 1QD, UK (y.cao-3@postgrad.manchester.ac.uk)

**Power Conversion Group, University of Manchester, Manchester,
M60 1QD, UK (a.renfrew@manchester.ac.uk)

*** Control System Centre, University of Manchester, Manchester,
M60 1QD, UK (p.cook@manchester.ac.uk)}

Abstract: This paper presents the application of a novel optical flow optimization algorithm for a comprehensive on-road vehicle motion analysis. Optical flow, which contains abundant local motion information, has been extensively studied for vehicle motion estimation in the last decades. How to generate a reliable optical flow at a low computation cost is always a challenging task. The primary aim of this paper is to enhance the accuracy and efficiency of optical flow estimation for a reliable vehicle motion analysis. In the paper, an innovative optical flow optimization algorithm is proposed based on a 3-D Pulse-Coupled Neural Network (PCNN) model. Because of the excellent information clustering ability of PCNN, the proposed algorithm can significantly improve the quality of optical flow. Moreover, a sparse motion flow field is generated to boost the computation efficiency. We employ a preliminary processing to detect the Region of Interest (ROI) in the image, and optical flow is only calculated and optimized in the ROI to save computation resources. Finally, the improved sparse optical flow field is exploited for a systematic on-road vehicle motion analysis. The proposed methodology has been evaluated under various challenging traffic situations to demonstrate its excellent performance.

Keywords: Optical flow, intelligent vehicle, motion estimation, motion detection, Pulse-Coupled Neural Network

1. INTRODUCTION

Vehicle motion analysis is an important component in intelligent vehicle systems (Sun et al., 2006). It can be used to alert drivers about abnormal driving conditions, potential collisions or even trigger an automatic emergency control (Large et al., 2004), (Cheng et al., 2006). Optical flow (Bereziat et al., 2000), which contains valuable local motion information, has been widely investigated for the task. Giachetti, Campani and Torre have already proved the feasibility of reliable vehicle motion estimation on a basis of optical flow interpretation (Giachetti et al., 1998).

Until now, various methods have been proposed for optical flow computation. These methods are mainly classified into two categories: differential techniques (Bauer et al., 2006), (Papenberg et al., 2006) and correlation techniques (Weickert and Schnorr, 2001), (Alvarez et al., 2000). Unfortunately, the popular differential techniques often fail in the traffic scene analysis, since the large displacement between consecutive frames overturns their fundamental temporal/spatial consistency assumption. The correlation techniques can successfully overcome the problems incurred by the large displacement. However, they also encounter some critical problems: (1) texture insufficiency brings in serious instability; (2) noise critically damages the image quality; (3) a large amount of computations are involved in the correlation calculation. Many efforts have been made to

improve the performance of optical flow computation based on different principles (Kruger et al., 2000), (Kenney et al., 2005) and (Chivilo et al., 2004). However, none of them can produce the reliable estimation which is required for vehicle motion analysis.

To solve the above problems, we propose a systematic methodology for on-road vehicle motion analysis based on optical flow optimization. To enhance the accuracy and robustness, an innovative 3-D Pulse-Coupled Neural Network (PCNN) model is constructed to perform optical flow optimization. Furthermore, a Region of Interest (ROI) is defined in the image and only the points within the detected ROI are considered for optical flow optimization to improve the efficiency of our method. Experimental results show that the improved optical flow field is precise enough to perform a quantitative vehicle motion analysis.

The rest of this paper is organized as follows: Section 2 briefly outlines the architecture of the proposed methodology. Details of the proposed optimization methodology, which include ROI generation, preliminary optical flow calculation and optical flow optimization, are presented in Section 3. In Section 4, we further discuss the application of the improved optical flow field for comprehensive vehicle motion analysis. Finally, experimental results are demonstrated in Section 5 and conclusions are given in Section 6.

2. OVERVIEW OF METHODOLOGY

The proposed methodology comprises four major steps: (1) Region of Interest (ROI) generation; (2) preliminary optical flow calculation; (3) optical flow optimization; (4) vehicle motion analysis. Fig. 1 systematically demonstrates the architecture of the method. Details of each step will be presented in the following sections.

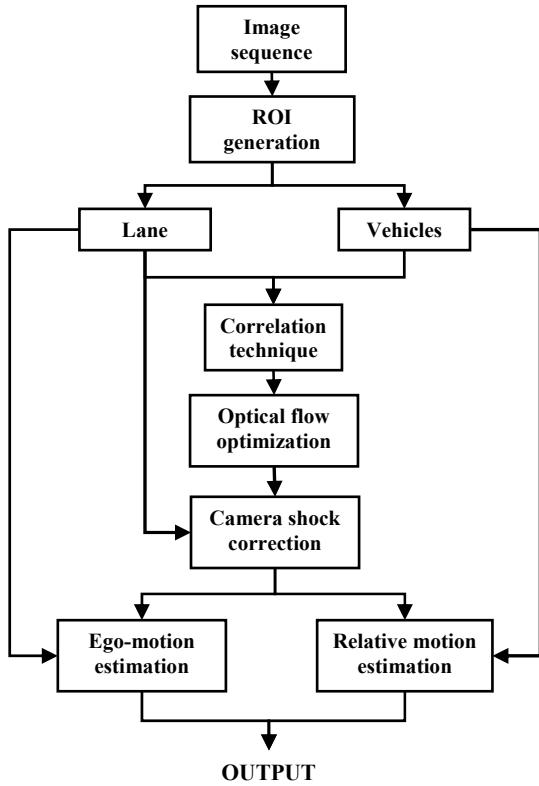


Fig. 1. Architecture of the methodology

3. SPARSE OPTICAL FLOW OPTIMIZATION

Overall, three major improvements are made in the method to enhance the robustness and efficiency: (1) a correlation-based technique is employed to handle the large displacements in a vehicle-recorded image sequence; (2) sparse optical flow is generated in the detected ROI to boost processing speed; (3) a novel optimization algorithm is proposed to improve the quality of optical flow.

3.1 ROI generation

An important aspect of computation efficiency improvement is the removal of irrelevant outlier information (i.e. sky, trees or buildings). In our case, the Region of Interest (ROI) includes lane markers and moving vehicles. An explicit example of ROI generation is shown in Fig. 2.

For most structured roads, in the area as far as 50m ahead, lane markers can be approximately portrayed by straight lines with certain orientation angles. Therefore, a linear method is applied for the lane detection due to its simplicity and robustness. The method consists of gradient-based edge detection and Hough Transform. For a gray-level image

$I(x, y)$, its gradient function $\nabla I(x, y)$ can be approximately estimated as follows:

$$\nabla I(x, y) \approx \sqrt{D_x^2 + D_y^2} \quad (1)$$

$$\nabla I(x, y) = \begin{cases} \nabla I(x, y), & \text{if } \nabla I(x, y) > T_\nabla \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where D_x and D_y are the intensity differences computed in x and y directions using the Sobel operator (Pratt, 2001). The threshold T_∇ is set to filter out edge points with a trivial gradient value. The gradient-based edge detection result is shown in Fig. 2 (b). Then, we employ a Hough Transform technique (Pratt, 2001) to locate the lane markers in the image. The Hough Transform technique is robust to noise disturbance and obstruction by taking into account all edge points. This is particularly useful to detect dashed lanes or partially blocked lane. The gradient accumulation function $C(\rho, \theta)$ is calculated as:

$$C(\rho, \theta) = \sum_i \nabla I(x_i, y_i) \quad (3)$$

(x_i, y_i) are all the edge pixels satisfying $\rho = x_i \cos\theta + y_i \sin\theta$. A look-up table of $\sin\theta$ and $\cos\theta$ ($\theta \in (-90, 90)$) would be generated to speed up the computation. Assuming lane markers are the most significant linear structure in the image, local maxima can be found in the 2-D accumulation function $C(\rho, \theta)$ to determine the position of lane markers. In our case, two prominent boundaries of the current driving lane can always be successfully detected.

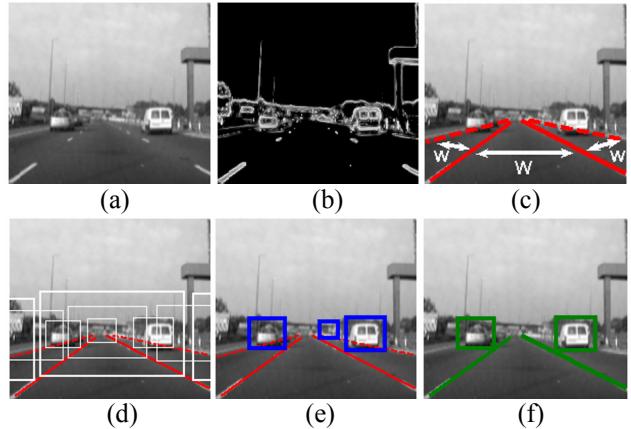


Fig. 2. An example of ROI generation. (a) the input frame, (b) gradient-based edge detection result, (c) lane detection and search area definition, (d) search windows generation, (e) vehicle detection based on the edge analysis (false detection can be eliminated by checking the symmetry), (f) the final defined ROI which includes the vehicles and lane markers.

After the lane detection, a feasible search region is defined for vehicle detection. The search region is defined as shown in Fig. 2 (c), where the road width w is calculated based on lane detection and camera calibration. It's noted that the defined search area adequately covers the current driving lane and two adjacent lanes. Along each lane, a series of search windows are automatically generated, as shown in Fig. 2 (d). Each search window has a rectangular shape, whose width and height are set sufficient to enclose the largest vehicle. It's

noted that the rear part of a vehicle usually contains many distinct horizontal and vertical structures, such as window, bumper, and registration plate. The existence of a vehicle will incur significant edge constellations into an image. Therefore, the edge information is applied as the key factor for vehicle detection. In a search window, the gradient magnitudes of all edge points are accumulated as:

$$Ga = \sum_j \nabla I(x_j, y_j) \quad (4)$$

If the gradient accumulation Ga exceeds a pre-defined threshold, then the window is considered to contain a vehicle. In addition, the symmetry level is inspected to eliminate false vehicle detection results. The symmetry level of a search window is calculated as:

$$Gs = \sum_{x=1}^{H-W/2} \sum_{y=1}^{W/2} |\nabla I(W/2-x, y) - \nabla I(W/2+x, y)| \quad (5)$$

where H and W denote the height and width of the search window. The higher Gs is, the more symmetric is the window. In the near vision field, moving vehicles can be robustly detected on the basis of edge and symmetry investigation.

The detection results of lane markers and vehicles are combined to generate the final ROI, as shown in Fig. 2 (f). To save computation cost, only the pixels within the detected ROI are considered for optical flow computation and optimization in the next step.

3.2 Preliminary optical flow calculation

Instead of the poorly performed differential technique, a correlation technique is employed in our algorithm for optical flow estimation. The correlation technique computes the optical flow by searching the displacement which produces the minimum Sum of Absolute Difference (SAD). In a $M \times N$ image, the SAD of displacement $r(r_x, r_y)$ at pixel (i, j) is defined as:

$$SAD(i, j, r) = \sum_{dx=-b}^{dx=b} \sum_{dy=-b}^{dy=b} |I(i+dx, j+dy) - I(i+dx+r_x, j+dy+r_y)| \quad (6)$$

$(i=1, 2, \dots, M, j=1, 2, \dots, N, r=1, 2, \dots, R)$

where b is the size of the matching mask, M and N are the size of the image, R is the number of all possible displacements. Moreover, the Block Matching (BM) is defined as:

$$BM(i, j, r) = \frac{1}{SAD(i, j, r)} \quad (7)$$

The higher the BM is, the lower is the SAD. However, exclusively searching for the maximum BM (minimum SAD) usually cannot guarantee the correctness of optical flow estimation in a complex traffic scene. Hence a new concept, Block Matching Probability (BMP), is introduced as:

$$BMP(i, j, r) = \frac{BM(i, j, r)}{BM(i, j, r^*)} \times 100\% \quad (8)$$

where r^* is the displacement with the highest BM. $BMP(i, j, r)$ measures the probability of displacement r being selected as the motion description at pixel (i, j) . A $M \times N \times R$ matrix is created to store the probability information of R displacements at $M \times N$ pixels. In addition, the Block Matching Certainty (BMC) is defined as:

$$BMC(i, j) = \frac{|BMP(i, j, r^*) - \overline{BMP}(i, j)|}{|BMP(i, j, r^*) + \overline{BMP}(i, j)|} \quad (9)$$

where \overline{BMP} is the average of all BMP values. The meaning of BMC is quite explicit. It measures the difference between the highest BMP and others. A large difference implies that one displacement has a distinct advantage over others to be selected as the motion estimation. Consequently, the higher the BMC is, the more reliable is the motion estimation. In the next section, the BMP matrix will be re-calculated by referring to the more reliable adjacent motion estimates.

3.3 Optical flow optimization

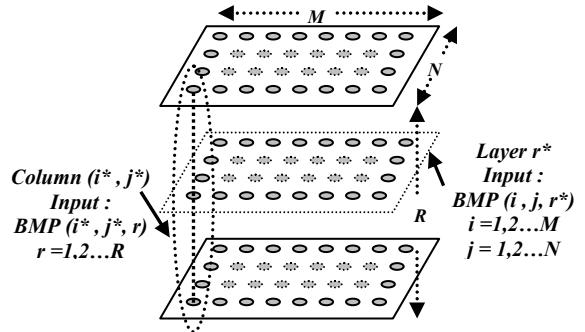


Fig. 3. 3-D PCNN model structure

Pulse-Coupled Neural Network (PCNN) is a neural model based on the biological structure of cat's visual cortex (Eckhorn, 1999). Until now, PCNN has mainly been applied in two-dimensional image processing tasks (Yu and Zhang, 2004) and (Gu et al., 2005). Here, a novel 3-D PCNN model (Cao et al., 2007) is constructed to perform three-dimensional optical flow optimization. The input of the PCNN model is the 3-D ($M \times N \times R$) BMP matrix. Accordingly, the model comprises R layers and each layer has $M \times N$ neurons. Neuron (i, j, r) receives $BMP(i, j, r)$ as its input. In the model, each layer corresponds to one displacement for all pixels, while each column refers to all possible displacements for one pixel. Fig. 3 shows the structure of the 3-D PCNN model.

Principally, the BMP re-calculation attempts to apply the reliable estimates to rectify the uncertain ones. It firstly identifies the column with the highest BMC. In this column, the neuron with the highest BMP will pulse and trigger a firing wave in the iteration T . In the PCNN model, the neuron fired at step t will pass a lift-up to its adjacent neurons. The neurons who received the lift-up will fire at the next step $t+1$ if they have already had high inputs. Consequently, the firing wave will expand in the layer until no neuron has a high enough input to fire any more. Then the model will find the second highest BMC and start another firing wave in the iteration $T+1$. After a neuron fired, its BMP value is replaced

with the higher internal status. As a consequence, the BMP will be re-calculated. The optimization will be terminated when the highest BMC is lower than a pre-defined threshold. The neuron dynamic activity is described as follows:

$$F(i, j, r) = \text{BMP}(i, j, r) \quad (10)$$

$$\begin{aligned} Y(i(T), j(T), r(T)) &= 1 \\ \text{if } \text{BMC}(i(T), j(T)) > \theta(T) \& \end{aligned} \quad (11)$$

$$\text{BMP}(i(T), j(T), r(T)) = \max(\text{BMP}(i(T), j(T), r))$$

Equation (11) determines the first neuron fired in the iteration T .

REPEAT Equation (12) – (16) in the iteration T

$$L(i, j, r(T), t) = \sum_{k, l \in N_8(i, j)} w(k, l) Y(k, l, r(T), t-1) \quad (12)$$

$$U(i, j, r(T), t) = F(i, j, r(T)) [1 + \beta L(i, j, r(T), t)] \quad (13)$$

$$Y(i, j, r(T), t) = \begin{cases} 1 & \text{if } U(i, j, r(T), t) > \delta(i, j, t-1) \\ 0 & \text{else} \end{cases} \quad (14)$$

$$\delta(i, j, t) = \begin{cases} \text{BMP}(i, j, r^*) & \text{initially} \\ \delta(i, j, t-1)[1 - Y(i, j, r(T), t)] \\ + U(i, j, r(T), t) Y(i, j, r(T), t) \end{cases} \quad (15)$$

$$\text{BMP}(i, j, r(T)) = \delta(i, j, t) \quad (16)$$

UNTIL no neuron fires any more, then update $\theta(T)$ to $\theta(T+1)$

$$\theta(T+1) = \begin{cases} e^{-\alpha\theta} \theta(T) & T \geq 1 \\ \max(\text{BMC}) & T = 1 \quad (\theta_{\min} < \theta(T)) \end{cases} \quad (17)$$

move to (11) to operate in the iteration $T+1$

where $F(i, j, r)$ and $L(i, j, r)$ are the feeding and linking channels of the neuron (i, j, r) , $w(k, l)$ is a weighted functional mask, $N_8(i, j)$ denotes the eight neighborhoods of neuron (i, j) , β is the linking strength, $U(i, j, r, t)$ is the internal activity status, $\delta(i, j, t)$ is the activity threshold and $\theta(T)$ is the certainty threshold. $U(i, j, r, t)$ is compared with $\delta(i, j, t)$ to decide the firing status $Y(i, j, r, t)$ of the neuron. $\delta(i, j, t)$ is initially set to the highest BMP value in column (i, j) and kept updated to track the highest activity in the column.

The underlying principle of the BMP re-calculation is very intuitive. Assuming the adjacent points are likely to have the same motion, the algorithm attempts to improve the uncertain optical flow by referring to the more reliable adjacent motion information. When local information is disturbed or insufficient, spatial supports from neighborhoods become essential to ensure the accuracy of the optical flow estimation. It's also noted that a motion estimates will not be affected by its neighbourhoods if its reliability level is higher. This property will preserve the discontinuous but true optical flows in the optimization.

Finally, the re-calculated BMP matrix is investigated to produce the final optical flow field. For each pixel, the displacement with the highest BMP is selected as its motion estimation.

4. VEHICLE MOTION ANALYSIS

In this section, we perform a comprehensive vehicle motion analysis by interpreting the improved optical flow field. To simplify the mathematical computation, we assume the host vehicle is moving on a planar structured road. This

assumption, which is usually referred to as the passive navigation, stands true in many driving situations (Giachetti et al., 1998).

4.1 Vibration correction

To apply the optical flow for the vehicle motion analysis, the camera vibration needs to be compensated. Since the horizontal vibration is much smaller than the vertical one, only the latter one is considered. If there is no vibration disturbance, the vertical position of the vanishing point is expected to be fixed in the image. Based on the observation, the camera shock can be effectively rectified by estimating the vertical position variance s_v of vanishing point and then subtracting it from the optical flow field. Previously, two straight lines have been defined to describe the lane markers in the near vision field. The intersection of these two lines can approximate the vanishing point.

4.2 Ego motion estimation

In the case of passive navigation, the relative motion between the camera and the road is exclusively caused by the host vehicle movement. Therefore, the ego-motion of host vehicle can be fully recovered from the optical flow in the road area. The motion (u, v) of a pixel (x, y) on the road surface can be expressed as (Giachetti et al., 1998):

$$\begin{aligned} u &= t_z xy / Hf + \omega_y (x^2 / f + f) \\ v &= t_z y^2 / Hf + \omega_y xy / f \end{aligned} \quad (18)$$

where t_z is the instantaneous translational velocity and ω_y is the perpendicular angular velocity. The camera height H and the focal length f can be measured through intrinsic and extrinsic camera calibrations. At each pixel, a pair of motion estimates (t_z, ω_y) is retrieved from its corresponding optical flow (u, v) based on (18). In the end, all estimates are averaged to improve the robustness of final vehicle ego-motion estimation.

4.3 Relative motions estimation

The optical flow of a vehicle is investigated to calculate its relative motion to the host vehicle. Based on the relative motion, a moving vehicle on the road can be classified into either a departing vehicle or an approaching vehicle. A departing vehicle produces a converging optical flow, while an approaching vehicle generates a diverging flow in the image. Due to the large variability of vehicle appearance, it is difficult to recover the relative motion accurately. We apply the same method for ego-motion estimation to approximately calculate the relative motion. A minimum 10 km/h relative movement can be successfully detected. This is accurate enough for many intelligent vehicle applications (e.g., overtaking vehicle detection, collision avoidance, and emergency brake control). However, when a vehicle is getting close to the host vehicle and covers a large portion of the frame, its motion estimation tends to become unstable.

5. EXPERIMENT RESULTS

The proposed method has been tested in various challenging motorway driving circumstances. The evaluation results of

optical flow optimization and vehicle motion estimation are reported in this section.

5.1 Data set

The video camera is fixed on the dashboard of the host vehicle. Several video sequences are recorded under different illumination conditions on the M60 motorway in Manchester, UK. The capturing rate of the video is 15 frame/sec. The ground truth of pixel displacement is obtained by manual matching. The ground truth value of the host vehicle speed is obtained from the on-board odometer and checked manually. The speeds of other vehicles on the road are approximately estimated by measuring the length of lane markers.

5.2 Optical flow optimization

We first evaluate the performance of proposed optimization algorithm in several typical driving scenes, as shown in Fig. 4 (a) – (d). The optical flow estimation using the traditional correlation technique are shown in Fig. 5 (a) – (d). In comparison, the results of the proposed optimization algorithm are shown in Fig. 6 (a) – (d). It is evident that the proposed method significantly improves the quality of optical flow in terms of accuracy and robustness. A more than 90 percent accuracy rate can be achieved in most normal driving circumstances.

5.3 On-road vehicle motion analysis

The speeds of the host vehicle and other vehicles on the road are then estimated based on the improved optical flow field. Since the ground truth of angular velocity ω_y is unavailable, we only examine the instantaneous velocity t_z . As shown in Fig. 7, the method produces the ego-motion estimation at more than 95% precision rate. Meanwhile, the maximum error for relative motion estimation is less than 10 km/h. The detection rates of approaching and departing vehicles at different relative speeds are shown in Table I. The result is precise enough for most intelligent vehicle applications.

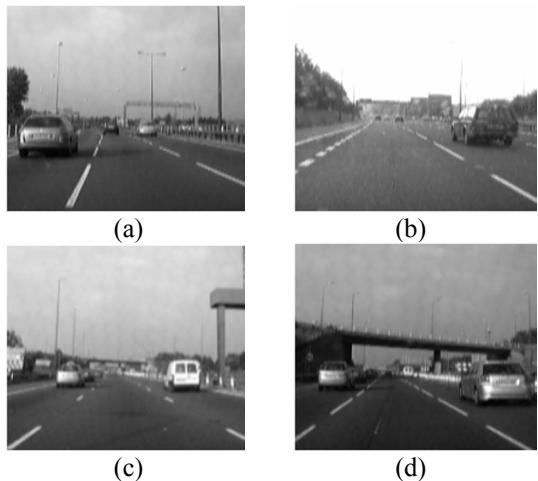


Fig. 4. Typical traffic scenes. (a) a vehicle is approaching from the left, (b) a vehicle is overtaking from the right, (c) two vehicles are both approaching on the two adjacent lanes, (d) a vehicle is approaching while another is overtaking.

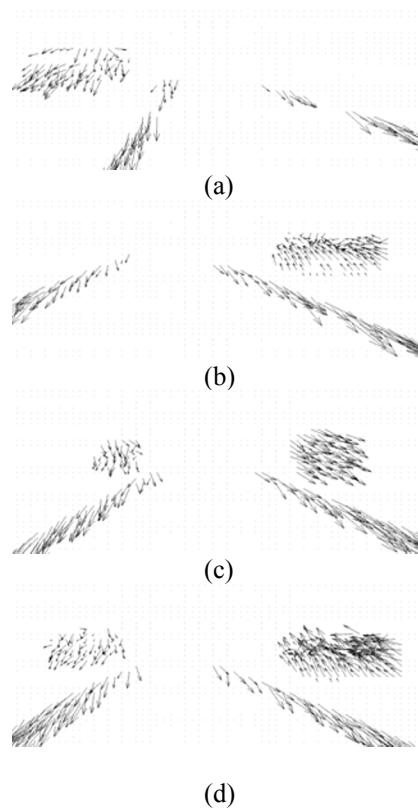


Fig. 5. The optical flow estimation using the correlation technique in the traffic scenes shown in Fig. 4 (a) – (d)

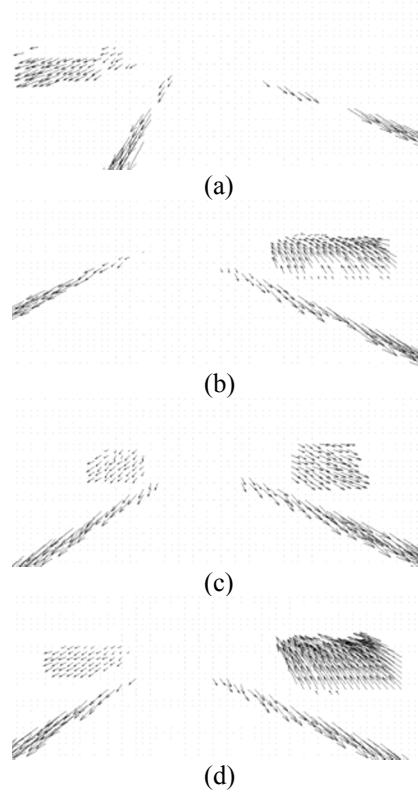


Fig. 6. The optical flow estimation using the proposed optimization algorithm in the traffic scenes shown in Fig. 4 (a) – (d)

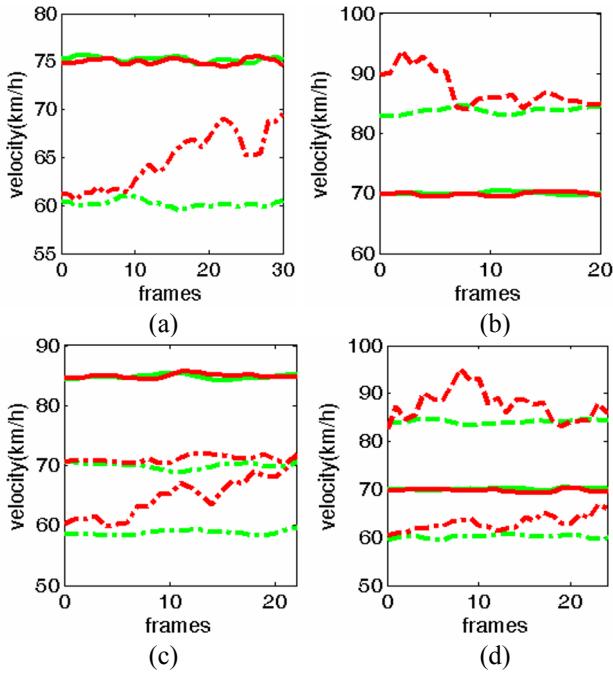


Fig. 7. Vehicle speed estimation based on the improved optical flow field for the traffic scenes shown in Fig. 4 (a) – (d). (green/red solid lines denote the true/estimated speeds of the host vehicle; green/red dashed lines denote the true/estimated speeds of overtaking vehicles; green/red dash-dot lines denote the true/estimated speeds of approaching vehicles)

Relative speed	Overtaking vehicle detection	Approaching vehicle detection
5-10 km/h	67.3%	75.1%
10-20 km/h	91%	94.3%
20 km/h	100%	100%

Table I. Detection rates of the approaching and departing vehicles at different relative speeds

6. CONCLUSIONS

In this paper, a novel optical flow optimization methodology is proposed for a systematic vehicle motion analysis. In the methodology, three significant improvements are carried out to enhance the robustness and efficiency: (1) The correlation technique is applied to handle the large displacement in the image sequence; (2) The region of interest (ROI) is detected to improves the processing speed of the algorithm; (3) a novel 3-D PCNN model is constructed to improve the quality of optical flow. Our experiments demonstrate that the proposed algorithm has excellent performance in many challenging traffic circumstances. The accurate ego-motion estimation is achieved at less than 5% error rate. Meanwhile the maximum relative motion estimation error is less than 10 km/h. The reliable vehicle motion estimation results can be further applied in various intelligent vehicle applications.

REFERENCE

- L. Alvarez, J. Weickert, and J. Sanchez (2000). Reliable estimation of dense optical flow fields with large displacements. *International Journal of Computer Vision*, **39**(1), 41-56.
- N. Bauer, P. Pathirana and P. Hodgson (2006). Robust optical flow with combined Lucas-Kanade/Horn-Schunck and automatic neighborhood selection. *ICIA 2006*, 378-383.
- D. Bereziat, I. Herlin, and L. Younes (2000). A generalized optical flow constraint and its physical interpretation. *CVPR 2000*, **2**, 487-492.
- Y. Cao, A. Renfrew, P. Cook (2007). Novel optical flow optimization using Pulse-Coupled Neural Network and Smallest Univalue Segment Assimilating Nucleus. *IEEE ISPAKS 2007*, 264-267.
- H. Cheng, B. Jeng, P. Tseng and K. Fan (2006). Lane detection with moving vehicles in the Traffic Scenes. *IEEE Transaction on Intelligent Transportation Systems*, **7**(4), 571-582.
- G. Chivilo, F. Mezzaro, A. Sgorbissa and R. Zaccaria (2004). Follow the leader behaviour through optical flow minimization. *IEEE Intelligent Robots and Systems*, **4**, 3182-3187.
- R. Eckhorn (1999). Neural mechanisms of scene segmentation: recordings from the visual cortex suggest basic circuits for linking field models. *IEEE Transaction on Neural Networks*, **10**(3), 464-479.
- A. Giachetti, M. Campani and V. Torre (1998). The use of optical flow for road navigation. *IEEE Transaction on Robotics and Automation*, **14**(1), 34-48.
- X. Gu, D. Yu and L. Zhang (2005). Image shadow removal using pulse coupled neural network. *IEEE Transaction on Neural Networks*, **16**(3), 692-698.
- C. Kenney, M. Zuliani and B. Manjunath (2005). An axiomatic approach to corner detection. *CVPR 2005*, **1**, 191-197.
- W. Kruger, W. Enkelmann, and S. Rossle (2000). Real-time estimation and tracking of optical flow vectors for obstacle detection. *IEEE Intelligent Vehicles Symp*, 304-309.
- F. Large, D. Vasquez, T. Fraichard and C. Laugier (2004). Avoiding cars and pedestrians using velocity obstacles and motion prediction. *IEEE Intelligent Vehicles Symp*, 375-379.
- N. Papenberg, A. Bruhn, T. Brox, S. Didas and F. Weickert (2006). Highly accurate optic flow computation with theoretically justified warping. *International Journal of Computer Vision*, **67**(2), 141-158.
- W. Pratt (2001). *Digital Image Processing*. 3rd edition. John Wiley & Sons, Inc.
- B. Yu and L. Zhang (2004). Pulse-coupled neural networks for contour and motion matchings. *IEEE Transaction on Neural Networks*, **15**(5), 1186-1201.
- Z. Sun, G. Bebis and R. Miller (2006). On-road vehicle detection: a review. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **28**(5), 694-711.
- J. Weickert and C. Schnorr (2001). Variational optic flow computation with a spatio-temporal smoothness constraint. *Journal of Mathematical Imaging and Vision*, **14**(3), 245-255.