An Application for Stereo Vision based Vehicle/Obstacle Detection for Driver Assistance

K S Chidanand Kumar
Tata Elxsi
Bangalore, INDIA

ABSTRACT

A stereo vision based vehicle/obstacle detection system has been proposed that generates alarms when vehicles/obstacles are detected in vicinity of near/mid region. Numerous techniques have been applied to detect vehicles/obstacles using a forward facing monocular camera mounted inside a vehicle. This paper presents a methodology for vehicle/obstacle detection using high resolution stereo camera. Stereo cameras are calibrated; stereo images are undistorted and rectified using stereo calibration parameters. Stereo disparity image is then generated using stereo matching algorithm. To localize vehicles/obstacles, search space reduction forms a preliminary step by eliminating sky region retaining only road regions on which vehicles/obstacles are embedded. Periodic peaks in the histogram of stereo disparity image are used as a cue in vehicle/obstacle detection process. Line profiles corresponding to each periodic peak are extracted using vertical strokes. Statistical features are extracted and analyzed for each line profiles to determine the presence/absence of vehicles/obstacles. If vehicles/obstacles are detected in a line profile, then bounding box of blobs are detected using vertical projection technique. Blobs detected from all the gray level profiles are merged together and connected component analysis is applied to count the number of vehicles/obstacles. Color mapping of vehicles/obstacles detected are done to indicate the presence of vehicles/obstacles in near region or mid region thereby generating appropriate alarms.

Keywords

Camera calibration, Stereo rectification, stereo matching, Inverse perspective mapping (IPM), Vertical projection, Vertical stroke, Feature extraction.

1. INTRODUCTION

One of the major challenges of the next generation road transportation vehicles is to increase the safety of the passengers and of pedestrians. Over 10 million people are injured yearly worldwide in road accidents. These include two to three million severely injured and 400,000 fatalities. The financial damage of accidents is estimates as 1-3% of world GDP [1]. Lack of attention by the driver is identified as the cause for 91% of driver related accidents. According to a 1992 study by Daimler-Benz (cited in [1]), if car drivers have a 0.5-second additional warning time, about 60% of collisions can be prevented. An extra second of warning time can prevent about 90% of collisions. This places collision warning system high on the list of solutions that can contribute significantly to reduce the number and the severity of driving accidents.

Vehicle/Obstacle detection methods using Collision warning system can be categorized as either active or passive. Active sensor-based methods include millimeter wave radar-based (Park et al.,) [2], Lidar-based (Wang et al.,) [3], and acoustic-based (Chellappa et al., 2004) [4]. Optical sensors, such as normal cameras, are usually referred to as passive sensors because they acquire data in a non-intrusive way. Generally, passive sensors offer some advantages over active sensors: low cost, wide spatial coverage, high spatial resolution, fast scanning speed, and no interference that might be caused when a large number of vehicles move simultaneously in the same direction using the same type of sensor.

Sun et al [5] made an overview for vision-based on-road vehicle analysis, which is strictly divided into hypothesis generation and hypothesis verification corresponding to detection and recognition respectively. Hypothesis verification step is separated into two methods: the template and appearance based method. Template-based methods [6] utilize predefined vehicle template to verify suspected patterns through correlation. However, their performance may decisively rely on the created templates. Appearance based techniques uses features like Haar-like [7, 8] and Gabor [9, 10] cooperated with classifiers like SVM and Adaboost proved to yield a decent performance in the recent literatures. But the execution time of methods using databases tends too slow to be applied to an embedded system.

The hypothesis generation step uses the knowledge, stereo vision and motion based methods. Knowledge-Based method uses many features (symmetry, color [11], shadow, corner, edge, and texture).Motion-based methods use motion vectors such as optical flow [12] to locate objects with large displacement but such methods suffer from correspondence problems. There are two types of stereo-based methods: disparity map-based and IPM (Inverse Perspective Mapping)-based methods. Although dense stereo matching algorithms are progressing continuously and are able to recognize the shapes and judge the distances of objects precisely and minutely, they are not useful in automotive collision warning/avoidance because the development direction is not aimed at real-time computation and target selection (Brown et al., 2003 [13]; Scharstein and Szeliski, 2002 [14]; Kastrinaki et al., 2003 [15]). This may be a reason why Sun's review mentions only two kinds of stereo vision based methods.

The IPM-based method generates two virtual images respectively for left and right images by assuming that every object is located on a flat ground plane. Then, the difference between the remapped left and right images reveals the location of obstacles, because anything located above the road makes large clusters of nonzero pixels in the difference image (Bertozzi and Broggi, 1998) [16]. Because of the flat ground plane assumption, the performance of the IPM-based method is degraded severely when the tilt angle of the camera changes on uneven roadways.

Another direction is the so called U/V disparity, which consists in two projection images of the disparity map. In these projections, the U respectively the V image axis is the axis on which the projection is applied, the other axis is the disparity value and the pixel intensity represents the number of pixels having a specific coordinate and a specific disparity value. The V disparity is used to detect the road surface and obstacles lying on it since a surface in the disparity map are mapped to a line in V disparity [17, 18]. The object detection however is not always reliable enough, so another approach helps segmenting by using preprocessed radar signals [19] in order to generate master points that validate the result.

In contrast to state of art methods, the proposed segmentation algorithm uses directly the disparity map, and works independently of the image content. The method detects vehicle/obstacles in the near/mid region reliably. Section 2 discusses stereo camera calibration, rectification and disparity generation techniques; Section 3 discusses vehicle/obstacle detection technique. Section 4 discusses simulation results followed by conclusion and references.

2. STEREO CAMERA CALIBRATION, RECTIFICATION AND DISPARITY MAP GENERATION

This involves three main steps. First step is to compute stereo calibration parameters of the camera to relate points in real-world to pixels in image space and relative rotation and translation between left and right camera using [20]. Secondly the images from two cameras are aligned such that both left and right image should be coaxial and parallel to baseline [21]. In the third step, stereo matching algorithm [22] is used to compute Disparity map. The results of stereo matching algorithm are shown in Fig.1(c).

Fig 1: (a) Left Stereo image (b) Right stereo image (c) Stereo Disparity image

3. VEHICLE/OBSTACLE DETECTION

Fig.2 shows the architecture of Vehicle/Obstacle detection using stereo vision. The proposed method reduces the search space by eliminating sky region and retaining only the ground regions on which obstacles are located. A histogram of stereo disparity image is taken in which periodic repetitive peaks are extracted. Line profiles corresponding to periodic repetitive peaks are extracted using vertical strokes. Features of every line profile are estimated the presence/absence of vehicle/obstacle or ground region. Blobs obtained from all the line profiles are merged and connected component analysis was carried to detect the number of vehicles/obstacles.
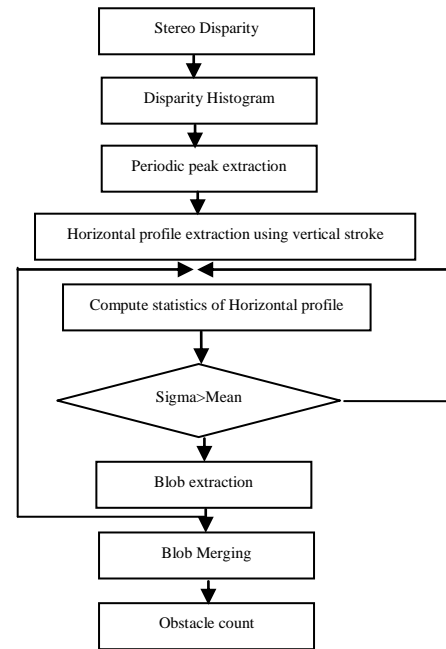

Fig 2: Architecture of stereo based obstacle detection

3.1 Search space reduction

In vehicle/obstacle detection process, sky region is not a region of interest (ROI). At the preliminary stage, a traffic scene image I(x, y) is divided into sky region and road region using vertical mean distribution [23]. Vertical mean distribution is measured by averaging the gray values of each row on road image and the row means are plotted in the graph depicted in Fig. 3(a).

The threshold value of horizon line is obtained through a minimum search along the vertical mean curve, where the first minimum occurs from the upper curve is the regional dividing intensity than road pixels, and it might have a big jump of intensity difference as sky pixels approaches ground. The horizon line threshold is applied to generate a road image (Rroi) as demonstrated in Fig 3(b), where all vertical coordinates below the threshold are discarded.
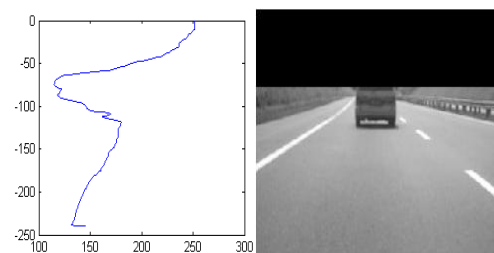

Fig 3: (a) Vertical mean distribution (b) Road region image (Rroi)

3.2 Ground plane estimation

In order to detect vehicles/obstacles, estimation of ground plane needs to be carried out first. In order to estimate ground plane, histogram of the disparity map is used as a cue. A plot of histogram for the disparity stereo image is shown in Fig 4(b) corresponding to Fig 4(a) is shown below.
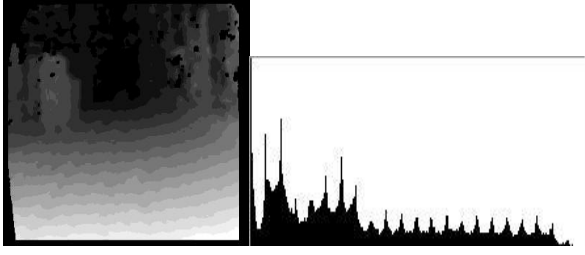
Fig 4: (a) Stereo disparity image (b) Disparity Histogram

3.3 Stroke based profile extraction

Let A be a stereo disparity image as shown in Fig 5(a). We now describe the process of vertical strokes on stereo disparity image A. For each peak gray value ($Peak_{Gray}$) in the list of $Peaks$ obtained from disparity histogram ($Disp\_Hist$), construct a binary image E consisting of object pixels in A that are visible from the north or top neighbor and their gray level equal to $Peak_{Gray}$. Let each of these vertical strokes be represented by $profile\_strokes_{vertical}$. Each vertical stroke is a digital curve. Shape of these vertical strokes are analyzed for extraction of features

3.4 Feature extraction for road extraction

One of the major factors for the success of road extraction using stereo disparity histogram technique is its feature extraction part. The feature should be selected in such a way that it should reduce the intra-class variability and increase the inter-class discriminability in the feature space. From each horizontal profile of vertical stroke as in Fig 5(b), mean (μ) and standard deviation (σ) of Y co-ordinate scalar features are extracted.
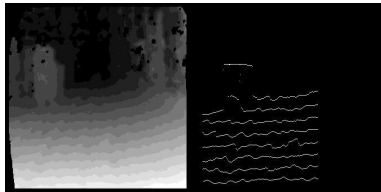

Fig 5: (a) Stereo disparity image (b) Profile of vertical strokes

A horizontal profile is said to contain vehicle/obstacle if $\sigma > (k*\mu)$ otherwise it contains only road region.

$$Obstacle\_profile = \begin{cases} 1 & \sigma > (k*\sigma) \\ 0 & otherwise \end{cases} \quad (1)$$

$$Road\_profile = \begin{cases} 1 & \sigma < (k*\sigma) \\ 0 & otherwise \end{cases}$$

In our experiments, we assumed $k = 2$ to get satisfactory results.

3.5 Blob roi extraction, merging and Blob counting

Once vehicle obstacle has been detected in $Obstacle\_profile$ from a set of line profiles obtained using vertical strokes ($profile\_strokes_{vertical}$) as in Fig.6 (a), the next step is to identify the blob bounding box in the $Obstacle\_profile$. A vertical profile is constructed on the $Obstacle\_profile$ region which is the sum of vertical disparity pixel values perpendicular to the Y axis and is represented by the vector V_p of size N columns and is defined by:

$$P_v[i] = \sum_{j=1}^{M} S(i,j) \quad (2)$$

Also from Fig.6 (b) vertical disparity gray level profile, it is observed that high peak region indicates bounding box of $Obstacle\_profile$.
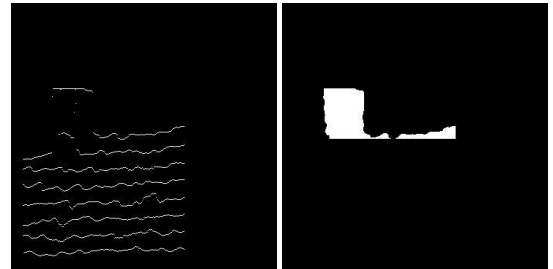

Fig 6: (a) Profile of vertical stroke (b) Vertical projection of obstacle profile

All the blobs obtained from $Obstacle\_profile$ are merged, and then connected component analysis is done to count the number of detected vehicles/obstacles.

4. RESULTS AND DISCUSSIONS

Figure 7, 8,9 (a)–(f) shows the various stages in detecting the vehicles/obstacles using narrow baseline stereo for driver assistance starting from input image. Search space reduction is done first on the stereo input image to eliminate sky region retaining road region in which segmentation of vehicles/obstacles needs to be carried out. Figure 7, 8, 9(b) shows the stereo disparity image. Figure 7, 8, 9(c) result of line profile extraction corresponding to periodic repetitive peaks in the histogram of stereo disparity image. Line profile features are extracted to determine the presence/absence of vehicles/obstacles. Vehicle/Obstacle bounding boxes are extracted using vertical projection technique and the result of which is shown in Figure 7, 8, 9(d). Figure 7, 8, 9(e) depicts vehicles/obstacles detected from all the profiles are merged and connected component analysis algorithm is carried out to detect the number of vehicles /obstacles.

The proposed method has been implemented on Intel Dual Core processor with 1.6GHZ, CPU 256 MB RAM running on windows vista operating system. The program was developed using Visual C++ language and OpenCV2.4.2. Stereo image dataset of resolution 1348*374 has been collected on highway

roads. To evaluate the performance of the above proposed approach, we tested with three videos downloaded from [24].From the Table 1, it is observed for Video 2 and Video 3 vehicle/obstacle detection rate is better than Video 1 because Video 2 & 3 contains more number of vehicles/obstacles in near/mid region and less in far region whereas in Video 1, more number of vehicles/obstacles were also present in far region thus creating more false alarms.

5. CONCLUSION

This paper presents a novel approach for detecting vehicle/ obstacles using stereo vision. Stereo camera is calibrated to get intrinsic, extrinsic and relative rotation and translation between stereo cameras, stereo images are rectified, undistorted and stereo matching algorithm is then carried out to obtain stereo disparity image. Search space reduction is carried on the input image to eliminate sky region retaining only road regions in which vehicles/obstacles are
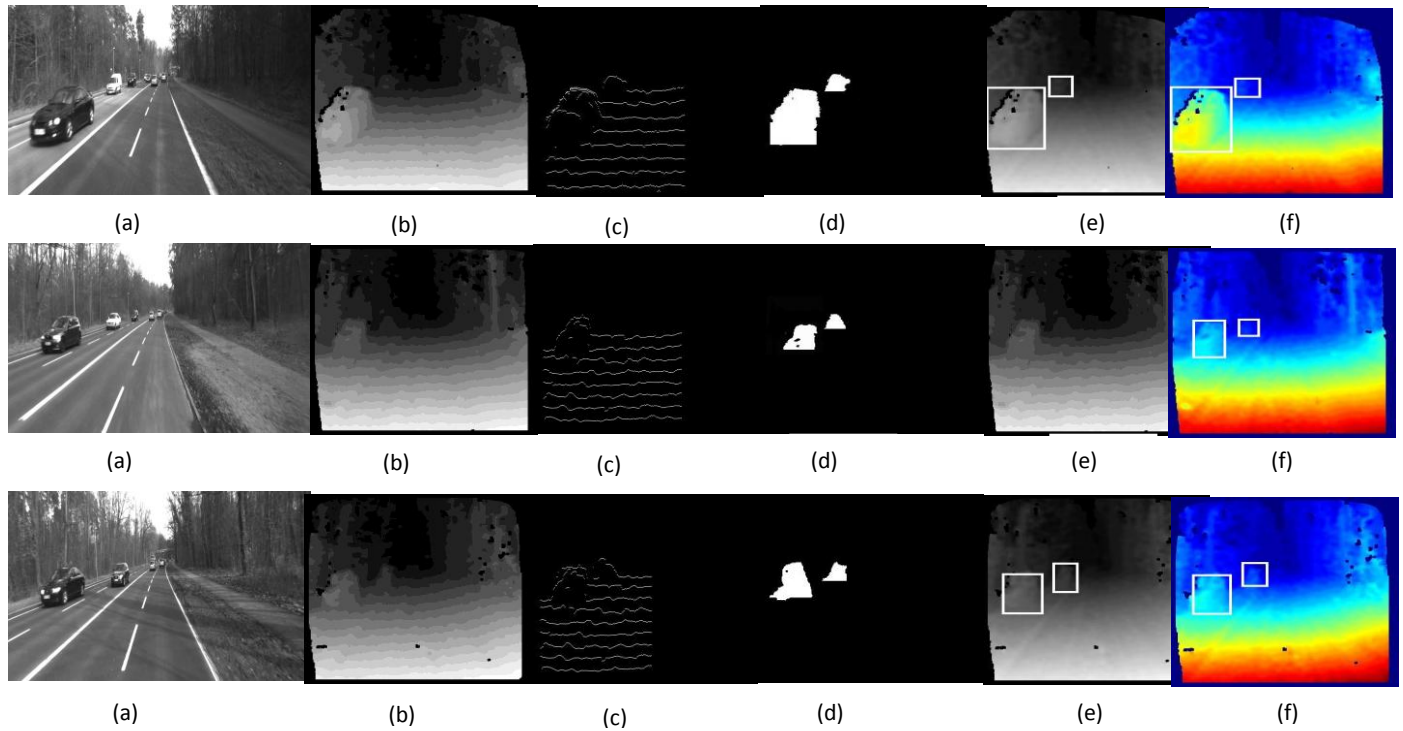

(a)   (b)   (c)   (d)   (e)   (f)

Fig.7, 8, and 9: Stereo based obstacle detection results

Table 1. Stereo based vehicle/obstacle detection results

	Video1	Video2	Video3
Total frames	354	306	289
Detected	560	514	439
Missed	63	32	27
False alarm	89	65	31
% Detection rate	78.65	84.12	88.33
% Average correct rate – 83.7			

embedded. Periodic repetitive peaks in the histogram of stereo disparity image are extracted which forms a cue in identifying vehicles/obstacles. Line profiles corresponding to periodic peaks are extracted using vertical strokes. Features are extracted on each and every profile to detect the presence/absence of a vehicle/obstacle. Vehicle/Obstacle bounding boxes are extracted using vertical projection, blobs from all the line profiles are merged and connected component analysis is applied to count the number of detected vehicles.

6. REFERENCES

[1] National Transportation Safety Board, Special Investigation Report - Highway Vehicle- and Infrastructure-based Technology for the Prevention of Rear-end Collisions. NTSB Number SIR-01/01, May 2001

[2] Park, S., Kim, T., Kang, S. and Heon, K. (2003). A novel signal processing technique for vehicle detection radar, 2003 IEEE MTT-S Int. Microwave Symp. Digest 607-610.

[3] Wang, C., Thorpe, C. and Suppe, A. (2003). Ladar-based detection and tracking of moving objects from a ground

vehicle at high speeds. Proc. IEEE Intelligent Vehicle Symp. 416-421.

[4] Chellappa, R., Qian, G. and Zheng, Q. (2004). Vehicle detection and tracking using acoustic and video sensors. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, 793-796.

[5] Z. Sun, G. Bebis, and R. Miller, On-road vehicle detection: A review, IEEE Trans. on PAMI, pp. 694–711, May 2006.

[6] A. Bensrhair, M. Bertozzi, A. Broggi, P. Miche, S. Mousset, and G. Toulminet, A cooperative approach to vision-based vehicle detection, in Proc. IEEE Intelligent Transportation Systems, Aug. 2001, pp. 207–212.

[7] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, in Proc. IEEE CVPR, 2001, vol. 1, pp. I–511–I– 518.

[8] P. Negri, X. Clady, S. Hanif, and L. Prevost, A cascade of boosted generative and discriminative classifiers for vehicle detection, EURASIP Journal on Advances in Signal Processing, 2008.

[9] H. Cheng, N. Zheng, and C. Sun, Boosted gabor features applied to vehicle detection, in Proc. IEEE ICPR, 2006, vol. 1, pp. 662–666.

[10] Z. Sun, G. Bebis, and R. Miller, On-road vehicle detection using evolutionary Gabor filter optimization, IEEE Trans. on Intelligent Transportation Systems, vol. 6, pp. 125–137, June 2005.

[11] J. Rojas, J. Crisman, Vehicle Detection in Color Images, in Proc. IEEE Conference Intelligent Transportation, pp. 403-408, November 1998

[12] J. Wang, G. Bebis, and R. Miller, Overtaking vehicle detection using dynamic and quasi-static background modeling, in Proc. IEEE CVPR, June 2005, pp. 64−64.

[13] Brown, M. Z., Burschka, D. and Hager, G. D. (2003). Advances in computational stereo. IEEE Trans. Pattern Analysis and Machine Intelligence 25, 8, 993-1008.

[14] Scharstein, D. and Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. Int. J. Computer Vision, 47(1/2/3), 7-42.

[15] Kastrinaki, V., Zervakis, M. and Kalaitzakis, K. (2003). A survey of video processing techniques for traffic applications. Image and Vision Computing, 21, 359-381.

[16] Bertozzi, M. and Broggi, A. (1998). GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection. IEEE Trans. Image Processing 7, 1, 62-81.

[17] Nicolas Soquet and Didier Aubert, "Road Segmentation Supervised by an Extended V-Disparity Algorithm for Autonomous Navigation," in 2007 IEEE Intelligent Vehicles Symposium, pp. 160-165, June 2007.

[18] R. Labayrade, D. Aubert, and J.P. Tarel. Real time obstacle detection in stereo vision on non flat road geometry through v-disparity representation. In Proc. IEEE Intelligent Vehicle Symposium, Versailles, France, June 2002.

[19] Michael Teutsch, Thomas Heger, Thomas Schamm, and Johann Marius Zöllner,"3D-segmentation of traffic environments with u/v-disparity supported by radar-given masterpoints", Intelligent Vehicles Symposium, page 787-792. IEEE, (2010)

[20] Zhang, Zhengyou. "A flexible new technique for camera calibration." Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.11 (2000): 1330-1334.

[21] Andrea Fusiello, Emanuele Trucco, Alessandro Verri: A compact algorithm for rectification of stereo pairs. Mach. Vis. Appl. 12(1): 16-22 (2000)

[22] Geiger, Andreas, Martin Roser, and Raquel Urtasun. "Efficient large-scale stereo matching." Computer Vision–ACCV 2010. Springer Berlin Heidelberg, 2011. 25-38.

[23] W. Lu, H. F. Wang, and Q. Z. Wang, "A Synchronous Detection of the Road Boundary and Lane Marking for Intelligent Vehicles," in Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2007. SNPD 2007. Eighth ACIS International Conference, 2007, pp. 741-745.

[24] http://www.cvlibs.net/datasets/karlsruhe_sequences.html