Jacques G. Verly (ed.), *Enhanced and Synthetic Vision*
Proceedings SPIE, 2736 (1996), pp. 228–239.

Integrated motion and geometry based obstacle detection in image sequences of traffic scenes

Heinrich NIEMANN Włodzimierz KASPRZAK [1] Peter WEIERICH

Bavarian Research Center FORWISS
Knowledge Processing Research Group
Am Weichselgarten 7, D–91058 Erlangen, Germany
E–mail: *name*@forwiss.uni-erlangen.de

Abstract

In this paper a robust method for visual motion estimation under ego–motion is developed. The possible application of this method is image sequence analysis of road traffic or airport runway/taxiway scenes, where the camera is located in a moving vehicle. The method combines an application independent estimation of visual motion with specific methods for instantaneous detection of the vanishing point in the image plane and of the over–road location of the camera. The stationary background is separated form the obstacles while detecting the ego–motion corrected visual motion of on–road objects.

Keywords:

background detection, focus of expansion point, ego-motion, image sequence analysis, obstacle detection, vanishing point, visual motion

1 Introduction

Specific image sequence analysis systems for many object scenes are moving robot guidance ([1]) and vision systems for driver support on roads ([2]). Most of the developed methods for automatic road following and collision avoidance can also be used in airport's runway or taxiway scenes. An *obstacle* in this context can be defined as something rising above the plane. Up to now the classification whether a (relatively) moving object is a specific case of an obstacle or not depends from concrete application. In this paper we define an obstacle as a specific subclass of a moving object with a zero motion vector.

Generally we can classify the methods for object (obstacle) detection into application independent *visual*

[1] The work was partly carried out while this author was staying with: Frontier Research Program RIKEN, Lab. for Artificial Brain Systems, 2–1 Hirosawa, Wako–shi, Saitama 351–01, JAPAN.

motion estimation methods, working in the 2–D image or a constrained plane in space (e.g. ground plane, road plane), and into model–based 3–D object recognition methods ([3]), ([4]). The later methods usually require a dynamic 3–D model of the scene and are outside of our interest here.

Usually the approaches to visual motion (or *optical flow*) estimation ([5]), ([6]) are divided into two classes: the *gradient–based* and *token matching* classes. In the methods of first type the local visual motion estimates are derived from local changes in the image intensity ([7]), ([8]). These methods typically yield *dense* optical flow. In the token matching (or correspondence) techniques discrete features such as corners and line segments are independently detected in each image and features in one image are matched with features in the second image ([9]), ([10]). Complete motion vectors are provided at a few points in the image only, leading to a *sparse* motion field.

The detected visual motion allows to separate the entire image into *stationary background* and *moving* objects. This kind of image segmentation is very often applied in traffic scene analysis, as the objects are usually projected to small image regions what prohibits their detailed recognition, e.g. ([11]), ([12]), ([13]). These methods are especially useful for single object scene analysis.

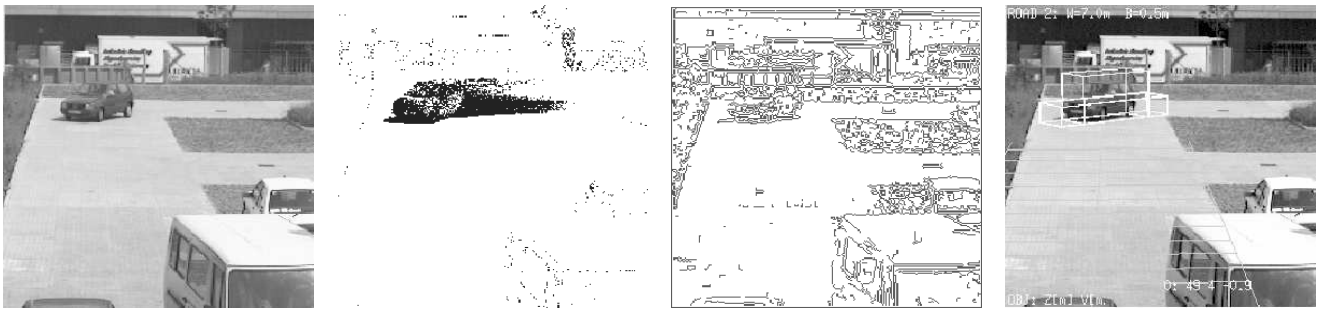
In applications where the observer is moving with a longitudinal motion (the *ego–velocity* case) the object motion can no longer be assumed to be parallel to the image plane. For robot navigation purpose it was proposed to compute the optical flow in a virtual horizontal scene plane instead of the image plane ([14]). Assuming that a stationary obstacle has to be detected only, in the back–projected plane such an object will correspond to high velocity values.

If the observer is moving on roads with large speed the vehicle is turning left or right and the camera may frequently perform an unknown nodding movement (the true *ego–motion* case). Additionally the pitch angle between the optical axis of the camera and the ground plane is relatively small ($< 10^o$) and one has to cope with perspective distortions of scene objects. Despite these difficulties a lot of work on obstacle detection for visual navigation purposes has been done while trying to compute and evaluate the optical flow field in the image plane ([15]). Other class of method in image plane estimates the *dynamic* focus of expansion point (FOE). Image segments whose motions match the FOE point are classified into 'stationary' ones. The remaining segments are "moving" segments. But this method was tested to be not robust enough for real outdoor sequences ([16]).

The main objective of current paper is to propose a robust computational scheme for obstacle detection on the basis of ego–motion corrected visual motion. In order to accommodate above disturbances of conventional visual motion an estimation of instantaneous 3–D ego–motion is performed. This estimation data influences the visual motion scheme leading to necessary corrections of image–to–image correspondence. We proceed as follows. At first, in section 2 we classify the motion estimation schemes according to their working space. In section 3 the new method for pixel–based motion estimation under ego–motion is described. First the geometrical elements of the system leading to a virtual road plane estimation are given. Next the adaptation of camera movements and ego–velocities in the visual motion estimation procedure follows. Some test results on real data are provided in section 4. A summary section concludes the paper.

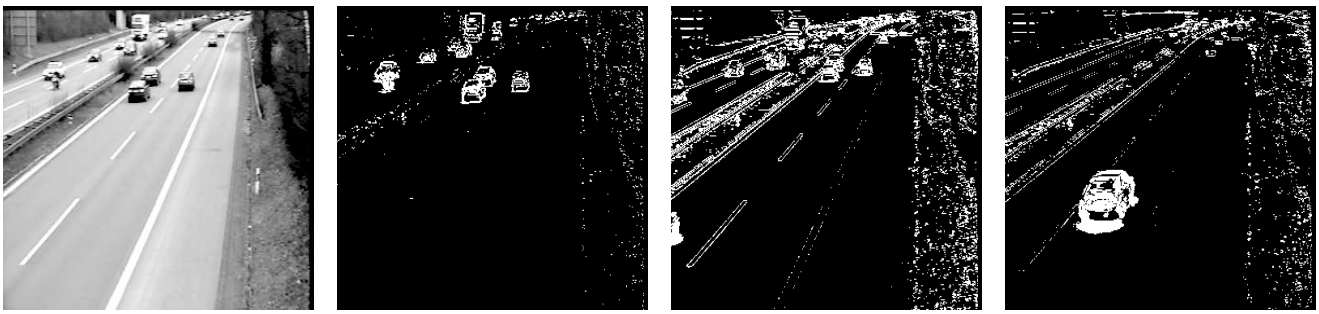
2 Pixel motion based obstacle detection

In this paper we consider *pixel* based methods for visual motion detection (or estimation) only. Other classes of methods like *feature* motion detection (e.g. for corners, edges, line segments, contours) or *frequency* based methods of optical flow estimation are beyond the scope of this paper. We distinguish here three classes of pixel based visual motion: detecting a motion in the image plane, in the ground plane (the ego–velocity case) or in the road plane (the ego–motion case).



(a) (b) (c) (d)

Figure 1: Pixel motion based moving object detection in case of a stationary camera: (a) single image frame, (b) the dynamic mask extracts 'moving" segments from the set (c) of all image segments, (d) the moving segments are grouped to objects by segment growing methods or model–based matching.



(a) (b) (c) (d)

Figure 2: Even a small unknown movement of the camera may lead to wrong results of moving object detection: (a) one original frame from the sequence, (b) detected pixel motion corresponds (nearly) properly to moving objects, (c,d) due to camera movement a wrong pixel motion of the background is detected.

2.1 Visual motion in the image plane

The straightforward solution is to apply the conventional visual motion in the image plane. A pixel based motion (a *dynamic mask* image) is generated by one of the following methods: a simple difference image method, adaptive difference image (**Figure** 1) or one of optical flow methods. The image segments are classified into "moving" or "stationary" depending on the amount of covered "moving" pixel. The neighbor moving pixel can be grouped to moving segments. In case of a stable camera position this method works well. But even a small unknown movement of the camera introduces errors of visual motion estimation as shown in **Figure** 2.

2.2 Visual motion in the ground plane

Let us assume that the camera is moving with constant and known velocity and that no nodding movement occurs. This is approximately true for an indoor environment and for relatively slow moving car or robot. The camera is turned toward the ground, i.e. the vanishing point is located in the image plane over the visible image area (compare the cross–section of a scene given in **Figure** 3).

The vehicle is moving straight ahead along the depth axis $O^g\vec{Z}^g$. The camera depth axis $O^c\vec{Z}^c$ is vertical to the image plane and is projected onto the ground depth axis $O^g\vec{Z}^g$. In the *focus point* of the camera the

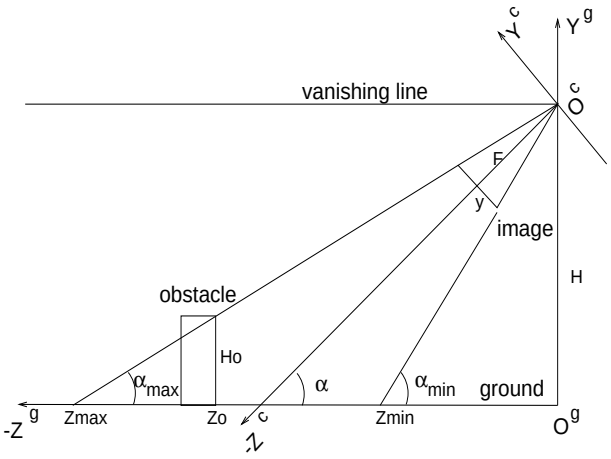


Figure 3: The camera pitch angle α versus the ground plane and the height H determine the projection conditions.

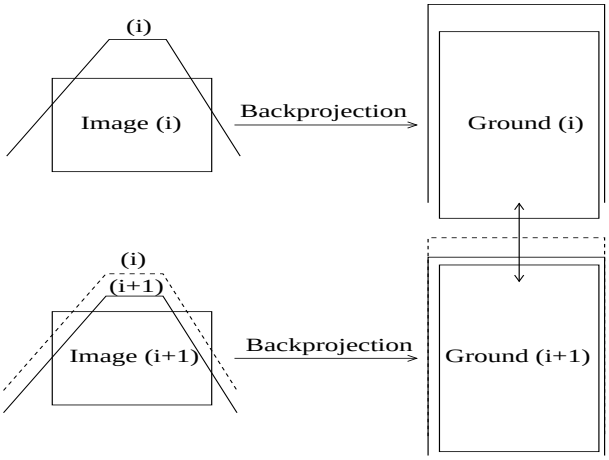


Figure 4: The principle of visual motion in the ground plane.

origin of the camera coordinates $O^cX^cY^c$ is located. The origin of the image coordinates is given as the point $o = (0, 0, -f)^T$ of the camera coordinates (where f is the *focal length*). The image coordinate axes $\vec{ox}$ and $\vec{oy}$ are parallel to appropriate camera coordinate axes $O^c\vec{X}^c$ and $O^c\vec{Y}^c$. The image coordinates $(x, y)^T$ of an image point are equivalent to following camera coordinates $(X^c, Y^c, Z^c)^T = (x * \delta x, y * \delta y, -f)^T$, where $\delta x, \delta y$ are the width or height of one image pixel.

We set $F_x = f/\delta x$ and $F_y = f/\delta y$ and express the focal length in pixel size along both axis directions. Now the mapping between the camera coordinates $P = (X^c, Y^c, Z^c)^T$ of a point in space and its projection $p = (x, y)^T$ onto the image plane is as follows:

$$x = -F_x \frac{X^c}{Z^c}; \qquad y = -F_y \frac{Y^c}{Z^c} \tag{1}$$

As the vanishing point is located over the image window every image can be fully back–projected onto the assumed ground plane $O^gZ^gX^g$ (**Figure** 4). The transformation from camera coordinates into ground coordinates is given as:

$$X^g = X^c \tag{2}$$

$$Y^g = Y^c \cos(\alpha) - Z^c \sin(\alpha) + H \tag{3}$$

$$Z^g = Y^c \sin(\alpha) + Z^c \cos(\alpha) \tag{4}$$

As all points are assumed to be located on the ground, i.e. $Y^g = 0$, the above equations (1) and (2) can be solved for unknown X^g, Z^g (let us notice that the measured depths have negative values):

$$X^g = -\frac{xHF_y}{F_x[y\cos(\alpha) + F_y \sin(\alpha)]} \tag{5}$$

$$Z^g = \frac{H[F_y \cos(\alpha) - y\sin(\alpha)]}{y\cos(\alpha) + F_y \sin(\alpha)} \tag{6}$$

Now a standard visual motion estimation can be applied for two consecutive ground images. This leads to motion vectors in the ground plane. The known ego-velocity of the vehicle provides a threshold for obstacle detection. As an obstacle is violating the planarity assumption it corresponds to high displacement values in the estimated motion field.

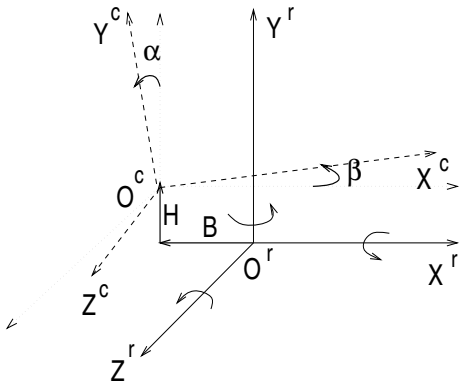


Figure 5: Two coordinate systems in space: the camera– and road systems.

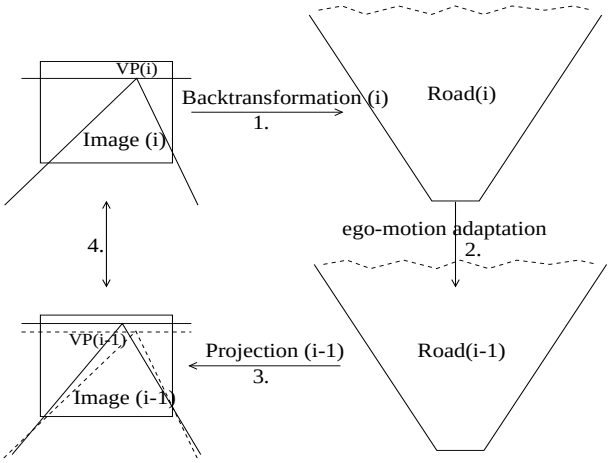


Figure 6: The principle of visual motion under ego–motion.

2.3 Visual motion under ego–motion

If the observer is moving on roads with large speed the vehicle is turning left or right and the camera may frequently perform an unknown nodding movement. Due to large speed a large scene depth range should be covered by the vision system. Hence the pitch angle between the optical axis of the camera and the ground plane should to be relatively small (i.e. below 10^o) and the vanishing point is located inside the image window. Thus this true ego–motion case differs from the previous ego–velocity by two aspects: firstly the back–projection is dynamically changing from image to image and secondly the back-projected ground image has unlimited height.

In case of an unknown camera nodding movement it is required either to have an outer orientation point or to measure the relative position against the ground and/or horizon. In both cases it is useful to consider two coordinate systems in 3–dimensional space, that are dynamically changing in time: the *camera coordinate system* $O^cX^cY^cZ^c$ with origin point O^c and the *road coordinate system* $O^rX^rY^rZ^r$ (**Figure** 5). The origin O^r of road coordinates is defined by the middle road axis and the road cross section at the actual position of the camera. Its origin is located in the center of the road on the road plane and on the same road's vertical cross section as the origin of the camera coordinates. The direction of the depth axis $O^r\vec{Z}^r$ is a tangential line of the road curvature on the road plane $0X^rY^r$.

The principle of proposed approach is to establish a correspondence between current image and a corrected previous image, from which the influence of estimated current ego–motion is eliminated (**Figure** 6). In next section our approach is described in more details.

3 Obstacle detection from visual motion under ego–motion

First a general scheme of our approach is given. Then the dynamic transformations between image and road are explained. Next the estimation of the vanishing point and the (optional) road recognition step are described. At end the applied visual motion estimators are mentioned and the obstacle detection rule is sketched.

The first step is required for estimation of unknown transformation parameters: $\alpha(t)$, $\beta(t)$. Road recognition is an optional step, i.e. it is not required if the rotational velocity of the ego–vehicle ω is already given, otherwise

INPUT: {Image(0), Image(1), ... }
DATA: Motion, VP, ROAD
OUTPUT: ObsMask - binary image of obstacle object pixel
i=0; Motion = **o**; set VP, ROAD to default values
WHILE Input is not empty
i= i+1;
VP = vanishing_point(VP, Image(i)
ROAD = road_recognition(ROAD, Image(i))
PrevImage = transformation(VP, ROAD, Image(i))
Motion = motion_estimation(Motion, PrevImage, Image(i))
ObsMask = obstacle_detection(Motion)

Figure 7: General scheme of obstacle detection from visual motion under ego–motion.

this parameter can be estimated relative to road curvature knowing the dynamics of the on–road position of the camera $B(t)$ and the rotation $\beta(t)$. In any case the proposed scheme requires additional information about translational ego–velocity V and of camera height $H(t)$ over the road plane.

3.1 The procedure

The general algorithm scheme for pixel–based motion estimation in the virtual road plane is shown in (**Figure** 7). After the steps of *vanishing_point* and *road_recognition* the actual space transformation parameters are determined. With this current 3–D transformation a virtual correspondence for every pixel in the image with its predecessor in the previous image, assuming the on–road constrain, can be established. The first step is required for estimation of unknown transformation parameters: $\alpha(t)$, $\beta(t)$. Road recognition is an optional step, i.e. it is not necessary if the rotational velocity of the ego–vehicle ω is already given, otherwise this parameter can be estimated relative to road curvature knowing the dynamics of the on–road position of the camera $B(t)$ and the rotation $\beta(t)$. In any case the proposed scheme requires additional information about translational ego–velocity V and of camera height $H(t)$ over the road plane.

In the step *motion_estimation* different motion detectors/estimators can be applied, for example a difference image method or a gradient based optical flow method. Due to the *obstacle_detection* step an image mask is generated, corresponding to projections of expected obstacles.

3.2 Transformations

A full transformation at time t between road coordinates $X^rY^rZ^r$ and camera coordinates $X^cY^cZ^c$ consists of following sequential steps (**Figure** 5):

- translation of the origin along the Y axis from $Y^r = 0$ to $Y^r = H(t)$ ($H(t)$ is the height location of the origin over the road plane).
- translation of the origin along the X axis from $X^r = 0$ to $X^r = B(t)$ ($B(t)$ is the horizontal distance between road center line and image origin point).
- rotation around the Y axis by angle $\beta(t)$
- rotation around the X axis by angle $\alpha(t)$

- rotation around the Z axis by angle $\gamma(t)$

The transformation from road into camera coordinates of a specific point in space is given by the same sequence of single steps but each step transform parameter has an opposite sign. Such a transformation consists then of appropriate translations by $-H$ and $-B$, and rotations by $-\beta, -\alpha$ and $-\gamma$.

For establishing the image–to–previous image correspondence, as shown in **Figure** 6 some simplifications may be done. As a relative on–road plane position need to be estimated only, it can be set $B = 0$. The angle γ describes the side movement of the ego-car. There is no robust method available for the detection of this movement. The angle γ could be detected, for example, from the analysis of the horizon orientation. But a large view on a nearly linear horizon should be available in nearly every image of the sequence. This can not be guaranteed. As the side movement of the car is small compared to the nodding movement, the angle γ is assumed to be constant all the time.

Let us now explain the transformation steps in **Figure** 6. For every pixel $(x, y) \in Image(i+1)$ a corresponding pixel $(x_p, y_p) \in Image(i)$ is to be find, that represents the same on–road position. This is achieved by following consecutive space transformations:

1. backprojection, i.e. determining the on–road position corresponding to (x, y) (for $y < VP_y$):

$$Y^c = \frac{-Hy}{y\cos(\alpha) + F_y \sin(\alpha)} \tag{7}$$

$$Z^c = -\frac{F_y Y^c}{y}; \quad X^c = -\frac{Z^c x}{F_x} \tag{8}$$

$$X^r = X^c; \quad Z^r = Y^c \sin(\alpha) + Z^c \cos(\alpha) \tag{9}$$

2. ego–motion adaptation, i.e. determining the on–road position of this stationary point at previous time $t_i = t_{i+1} - \tau$:

$$X^r(i) = X^r \cos(\omega\tau) + Z^r \sin(\omega\tau) - \frac{V}{\omega}[\cos(\omega\tau) - 1] \tag{10}$$

$$Z^r(i) = -X^r \sin(\omega\tau) + Z^r \cos(\omega\tau) + \frac{V}{\omega}\sin(\omega\tau) \tag{11}$$

   where V, ω are translational and rotational velocities of the ego-car, τ is the time interval between two consecutive images

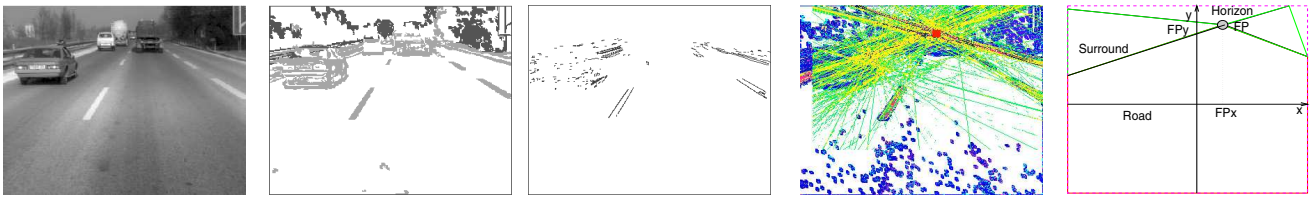
3. projection of previous position onto the image plane ($X^c = X^r$):

$$Y^c = Y^r \cos(\alpha) + Z^r \sin(\alpha) - H\cos(\alpha) \tag{12}$$

$$Z^c = -Y^r \sin(\alpha) + Z^r \cos(\alpha) + H\sin(\alpha) \tag{13}$$

$$x = -\frac{F_x X^c}{Z^c}; \quad y = -\frac{F_y Y^c}{Z^c} \tag{14}$$

For the backprojection selected image points are considered to be located on the road plane, i.e. $Y^r = 0$. The parameter $H(t)$ is assumed to be known and constant, but the remaining three transformation parameters: $B(t)$, $\alpha(t)$ and $\beta(t)$ have to be estimated from the image data. This process is described in next section.



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

Figure 8: Vanishing point measurement steps: (a) image normalization, (b) detection of image segments and their initial classification into road area segments (light color) and surrounding area (dark color), (c) selection of hypothetic vanishing point segments, (d) the search area (large rectangle) is filled with hypothetic VP–lines induced by the longitudinal sizes of selected stripe segments and the VP point is measured as a small rectangle with highest density of VP–lines, (e) the estimated VP point and image area classification.

3.3 Vanishing point recognition

The *vanishing point* VP is defined as the point in the image plane, where the projections of the road border lines vanish (**Figure** 8). If the road curvature is significant the straight lines change to tangential lines of the road stripes and borders in front of the ego–car. The purpose of vanishing point recognition is to estimate the current orientation of the camera in relation to the road plane and the road trajectory. The angle α determines the vertical rotation of the camera and β the horizontal rotation. These angles are estimated from image data as follows:

$$\alpha(k) = -\arctan\left(\frac{FP_y(k)}{F_y}\right); \quad \beta(k) = \arctan\left(\frac{FP_x(k)}{F_x}\right) \tag{15}$$

For dynamic estimation a linear Kalman filter ([17]) is applied. The state vector of VP consists of the position and translation in the image plane: $\boldsymbol{s}_{VP}(k) = [x(k), y(k), \delta x(k), \delta y(k)]^T$. The measurement vector consists of position parameters only: $\boldsymbol{m}_{VP}(k) = [xm(k), ym(k)]^T \leftrightarrow \boldsymbol{s}_{FP\_R}(k)$.

The computation steps, required for vanishing point measurement in each image frame, are exemplary given in **Figure** 8. From the set of all segments the assumed road stripe segments and road border segments are selected first. Such segments are mainly vertically elongated in the image (i.e. the relation of their width to height is larger than a given threshold value), the neighbor region intensities are "white" and the segments are positioned under current lowest vanishing point hypothesis (but at least in the lower image half) (**Figure** 8(c)). Only these segments are finally selected that are located on lines, which are at a small distance to at least one vanishing point hypothesis. Weights of these lines are calculated and their are added in a accumulator image. Next a search is performed for an image area with highest density of these lines, giving the current measurement $\mu_{VP}(k)$ (**Figure** 8(d)). The individual measurements are supplied to the adaptive estimator in order to give an estimated value $\boldsymbol{s}_{VP}(k)$ of the vanishing point ([18]).

3.4 The on–road position

On the basis of initial image region classification (**Figure** 8(e)) the image segments are classified into *road*, *surrounding area* and *heaven* classes. For this a rule–based procedure is used: segments containing a sufficient rate of "road" pixels are classified into "road" segments, the ones without such pixels but containing enough "surrounding" area pixels are classified into surrounding segments. The remaining segments are assumed to be placed above the horizon.

A relatively complex procedure is applied for road recognition ([18]). One of results of this procedure is the

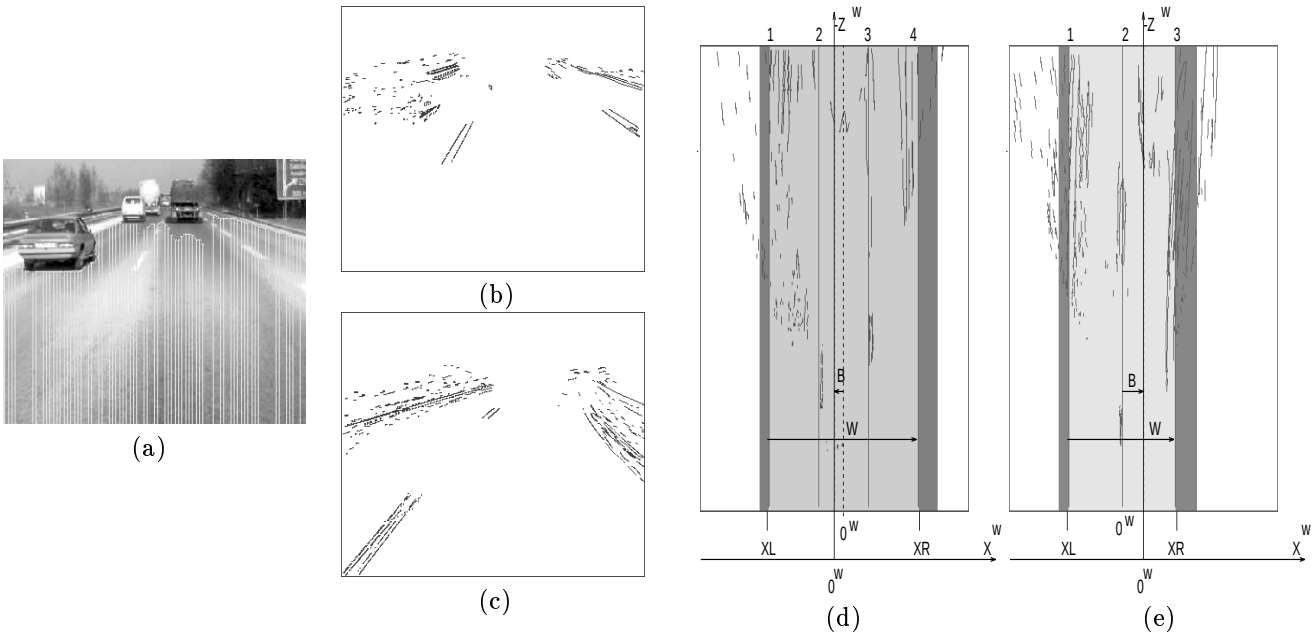


Figure 9: Main road recognition steps: (a) road region search (it gives the minimum road width $X_L\vec{X}_R$, (b,c) selection of line segments that are expected to correspond to road stripes, (d,e) backprojection of these segments onto the expected road plane, their weighted projection onto the road's cross–section and evaluation of densities. The images (b) and (d) are for linear road case, whereas (c) and (e) are for a curved road case.

relative position of the camera against the center line of road B(t). Due to the dynamics of $B(t)$ and $\beta(t)$ a rotational velocity of the ego–car relative to current road curvature can be estimated. The main steps of road recognition are: road region estimation for road width interval determination, (**Figure** 9(a)), selection of final VP–line segments from the set of road segments, that are expected to correspond to road stripes (**Figure** 9(b,c)), their backprojection onto the expected road plane and search for pike densities across the road (**Figure** 9(d,e)).

3.5 Virtual pixel motion and obstacle detection

We applied following methods for visual motion detection/estimation: adaptive difference image ([19]), gradient based optical flow ([20]), p.116, and block–matching ([21]). The first method is adaptive – it processes each image by recursive means. The second method is adaptive to some kind, as the previous optical flow constitutes the starting condition for the next time computation. The third method is non–adaptive – every two consecutive images are processed independently from previous results. Due to ego–motion adaptation the detected visual motion corresponds either to moving objects or to stationary obstacles rising above the ground.

4 Results

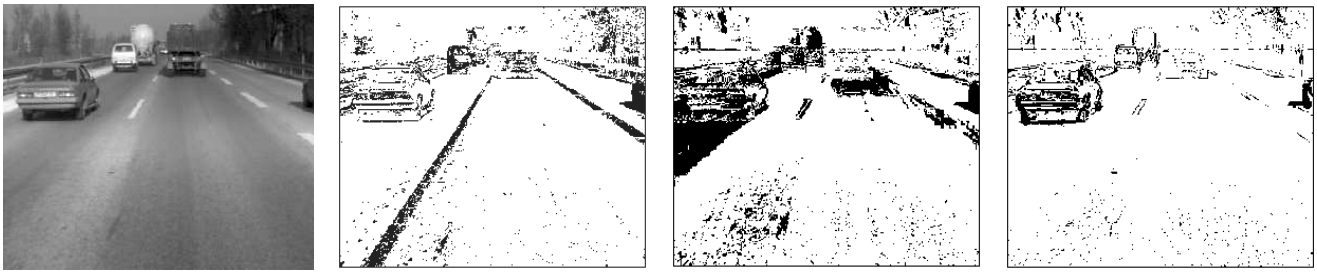
The developed method has been tested on several real image sequences from a moving camera. The other mentioned methods of image plane and ground plane motion, known from the literature, have also been implemented, allowing the comparison of results of different approaches. The frames of monochrome image sequences have usually a resolution of 384x282x8bit each (**Figure** 10).

Figure 10: Single frames of several image sequences, applied in computer simulation experiments.

Figure 11: Image masks used for quantitaive evaluation of visual motion based object detection.

Some preliminary simulation results are already available. At first the above methods for motion detection in the image plane, i.e. adaptive difference image, gradient based optical flow and block–matching have been tested with no modifications (**Figure** 12 (a,b)). In the result images the gray value of a pixel corresponds either to moving pixel (if motion detection is made only) or to the absolute motion value detected for this pixel (if visual motion estimation is provided). Then the influence of known ego–motion on each of the above methods and the obstacle detection in ground plane, according to the scheme proposed in the ([14]), has been tested (**Figure** 12(c)). At the end the adaptation of ego–motion for each of above methods as proposed in current paper has been tested (**Figure** 12(d)). An optical comparison of results from above three approaches already shows the advantage of our method of ego–motion adaptation in visual motion detection.

In the original frame these regions which correspond to moving objects have been marked manually. Then a pixel–wise comparison between these regions and the result images has been made. First quantitative evaluations have shown an increase of obstacle detection quality. The pixel covering rates are summarized in **Table** 1. As the



(a) (b) (c) (d)

Figure 12: Example results of three obstacle detection approaches: (a) single frame, (b) from standard visual motion in the image plane, (c) the ground–plane approach with ego–velocity adaptation, (d) our approach with unknown ego–motion adaptation.

Approach	Properly covered			Wrongly covered		
Visual motion from:	image	ground	ego–motion	image	ground	ego–motion
Adaptive difference image	61%	82%	84%	55%	33%	26 %
Block–matching	70%	75%	84%	56%	32%	26 %
Gradient–based method	66%	75%	85%	52%	25%	22%

Table 1: The rates of properly covered moving regions and non-properly covered background in three approaches.

detection error is still in the range of 20 percent an region based evaluation should be considered in the future.

Due to ego–velocity accommodation the average processing times for one motion frame (image pair) have been increased if compared to the motion estimation in image plane. For given image size the times have been measured on a DEC Alpha Station 250 (4/266) (see **Table** 2). The average time for VP estimation was 0.06 s and for road recognition and transformation calculation - 0.08 s per image. As a parallel processing of all algorithms is realizable these additional processing times are not crucial for real–time requirements.

Approach	Average times per image		
Visual motion from:	image	ground	ego–motion
Adaptive difference image	0.28 s	0.29 s	2.20 s
Block-matching	21.20 s	21.30 s	23.50 s
Gradient-based method	1.09 s	1.10 s	3.10 s

Table 2: Processing times of three approaches for obstacle detection with three visual motion detectors.

5 Conclusions

The specific problem of vision based application–independent obstacle detection scheme was addressed. A moving camera case with unknown nodding movement has been considered. A combination of geometry based analysis for visual motion estimation under ego–motion in both image plane and road plane has been proposed. An automatic estimation of current projection conditions and of the on–road position gives the instantaneous transformation parameters. Hence the unknown camera movement and the transformation between image and road coordinates can be estimated. Three methods of visual motion detection/estimation have been modified by adapting the ego–motion. Tests have shown that in comparison to original methods (without ego–motion) a large improvement in the quality of visual motion obstacle detection may be achieved.

Acknowledgments

The support from the 'Deutsche Forschungsgemeinschaft', Bonn, F.R.G., by Grant Ni-191/8-2 is gratefully acknowledged. The real images with a moving observer are by courtesy of the BMW AG., Munich, F.R.G.

6 REFERENCES

[1] I.J. Cox (ed.). *Autonomous robot vehicles.* Springer, New York, Berlin, Heidelberg etc., 1990.

[2] I. Masaki. *Vision-based Vehicle Guidance.* Springer, New York, Berlin, Heidelberg etc., 1992.

[3] E.D. Dickmanns and V. Graefe. Applications of dynamic monocular machine vision. *Machine Vision and Applications*, 1:241–261, 1988.

[4] D.-B. Gennery. Visual tracking of known three–dimensional objects. *International Journal of Computer Vision*, 7:243–270, 1992.

[5] J.K. Aggarwal and N. Nandhakumar. On the computation of motion from sequences of images - a review. *Proceedings of the IEEE*, 76(8):917–935, 1988.

[6] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. systems and experiment. *International Journal of Computer Vision*, 12(1):43–77, 1994.

[7] B.G. Schunck. The image flow constraint equation. *Computer Vision Graphics and Image Processing*, 35:20–46, 1986.

[8] H.-H. Nagel. On a constraint equation for the estimation of displacement rates in image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:13–30, 1989.

[9] V. Salari and I.K. Sethi. Feature point correspondence in the presence of occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(1):87–91, 1990.

[10] R. Deriche and O. Faugeras. Tracking line segments. *Image and Vision Computing*, 8(4):261–270, 1990.

[11] M. Bober and J. Kittler. Estimation of complex multimodal motion: an approach based on robust statistics and Hough transform. *Image and Vision Computing*, 12(10):661–668, 1994.

[12] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motion. *International Journal of Computer Vision*, 12(1):5–16, 1994.

[13] P. Bothemy and E. Francois. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *International Journal of Computer Vision*, 10(2):157–182, 1993.

[14] H.A. Mallot, H.H. Bülthoff, J.J. Little, and S. Bohrer. Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics*, 64:177–185, 1991.

[15] W. Enkelmann. Obstacle detection by evaluation of optical flow fields from image sequences. *Image and Vision Computing*, 9(3):160–168, 1991.

[16] W. Kasprzak and H. Niemann. Moving segment detection in monocular image sequences under egomotion. In *Signal Processing VII : Theories and Applications*, volume 2, 708–711. EUSIP Association, Lausanne, CH, 1994.

[17] A.H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, New York and London, 1970.

[18] W. Kasprzak, H. Niemann, and D. Wetzel. Adaptive road parameter estimation in monocular image sequences. In *British Machine Vision Conference 1994.*, volume 2, 691–700. BMVA Press, Sheffield, 1994.

[19] K.-P. Karmann and A. von Brandt. Moving object recognition using an adaptive background memory. In V. Cappellini, editor, *Time-Varying Image Processing and Moving Object Recognition*, volume 2, pages 289–296. Elsevier Science Publ., 1990.

[20] Heinrich Niemann. *Pattern Analysis and Understanding.* Springer–Verlag, Berlin Heidelberg New York Tokyo, 1990.

[21] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2:283–310, 1989.