Obstacle Detection by Evaluation of Optical Flow Fields from Image Sequences

Wilfried Enkelmann
Fraunhofer-Institut für Informations- und Datenverarbeitung (IITB)
Fraunhoferstraße 1, D-7500 Karlsruhe 1

INTRODUCTION

Image sequences contain information about the dynamic aspects of the recorded scene. Optical flow fields describe the temporal shift of observable gray value structures in image sequences. Various approaches have been suggested to estimate optical flow fields from image sequences - see, for example, Nagel 87. In most cases, this optical flow field is a good approximation to the temporal displacement of the image of a depicted surface element. Optical flow fields contain not only information about the relative displacement of image points but also about the spatial structure of the recorded scene. Several investigations in the literature show how these 2D-fields can be interpreted to infer information about the 3D-environment (Aggarwal & Nandhakumar 88).

To detect obstacles, some authors try to segment single images (Olin et al. 87), evaluate range data (Daily et al. 87) or use divergence fields (Koenderink & van Doorn 77, Nelson 88). Storjohann et al. 88 evaluate spatial disparities to detect obstacles whereas Dickmanns & Christians 89 extract only a few edges and decide wether the extracted edges are projections of an obstacle or not by tracking them. In this contribution, we investigate an approach for the detection of obstacles by evaluation of optical flow fields.

OBSTACLE DETECTION

The basic procedure we used to detect stationary as well as non-stationary obstacles in front of a moving vehicle consists of three steps (Enkelmann 89):

a) Calculate optical flow vectors $\mathbf{u}(\mathbf{x},t)=(u(\mathbf{x}), v(\mathbf{x}))^T$ which link the pixel at image location $\mathbf{x}=(x,y)^T$ in one frame to the corresponding pixel position in a consecutive frame from the recorded image sequence.

b) Estimate model vectors $\mathbf{u}_M(\mathbf{x},t)$ that describe the expected optical flow field without any obstacle in the scene. The current implementation assumes that the camera is translating on a planar surface.

c) Evaluate the differences $\mathbf{u}_D$ between calculated optical flow vectors $\mathbf{u}$ and estimated model vectors $\mathbf{u}_M$.

Optical Flow Calculation

To test the obstacle algorithm we used the first five frames of an image sequence to calculate optical flow vectors. During this time interval the vehicle carrying the camera passed a distance of about 2.5 meters. The optical flow vectors calculated from the monotonicity operator (Kories & Zimmermann 86), correspondences of countour elements (Rehfeld 88), and with an analytical approach (Enkelmann 88) are shown in Figs. 1-3, respectively.

Estimation of a model vector field

The two-dimensional model vector field $\mathbf{u}_M(x,t)$ assigns to each image location $\mathbf{x}=(x,y)^T$ the maximal possible shift in the image plane that will be acceptable if the projected 3D-point does not belong to an obstacle. If we assume only a translational camera motion parallel to the road plane, the components of the model vector $\mathbf{u}_M=(u,v)^T$ are given by:

$$u = x'-x = \left(\frac{X'_C f}{Z'_C} + p_x\right) - x \qquad v = y'-y = \left(\frac{Y'_C f}{Z'_C} + p_y\right) - y \tag{1}$$

Variables in lower case denote image plane coordinates. Quoted variables correspond to the end of the time interval used to calculate the optical flow vectors. In general, the direction of sensor motion is not parallel to the optical axis, so that the focus of expansion (FOE) is different from the projection point (p_x, p_y) of the optical axis onto the image plane. Subscripts C correspond to the camera coordinate system.

Insertion of the coordinate transformation of 3D-scene points on the road plane and the camera translation into Eq. (1) results in the following equations (where $Z'_C = Z_C - v_{CZ}\Delta t$):

$$u = \frac{(x-p_x) - f v_{CX}/v_{CZ}}{Z_C/(v_{CZ}\Delta t) - 1} = \frac{(x-p_x) - (FOE_x - p_x)}{Z_C/(v_{CZ}\Delta t) - 1} \qquad v = \frac{(y-p_y) - f v_{CY}/v_{CZ}}{Z_C/(v_{CZ}\Delta t) - 1} = \frac{(y-p_y) - (FOE_y - p_y)}{Z_C/(v_{CZ}\Delta t) - 1} \tag{2}$$

Equations (2) relate the components of the model vector $\mathbf{u}_M(\mathbf{x})=(u,v)^T$ to the image location $\mathbf{x}=(x,y)^T$, the FOE, the distance Z_C of the depicted 3D-scene point from the camera, and the component $v_{CZ}\Delta t$ of the sensor translation vector. The FOE was determined from the calculated optical flow field using the approach of Bruss & Horn 83.

The distance Z_C of a 3D-scene point on the road plane can be calculated from the intersection of a 3D-line of sight with the road plane if we know the transformation matrix between camera and vehicle coordinate system from a calibration procedure. At those image locations where the line of sight has crossed the horizon, the Z_C component of an intersection point with a virtual plane parallel to the road plane is used instead. This virtual plane is given a priori by the height of the moving vehicle.

The last unknown term in Eq. (2) is the component $v_{CZ}\Delta t$ of the camera translation vector. In the current implementation a vehicle state model is not yet available which can directly be used to estimate the model vector field. Therefore, we used in our experiments described in this contribution a 3D-translation vector which had been determined by backprojecting manually established point correspondences.

Evaluation of Differences

To detect obstacles, all differences between calculated optical flow and estimated model vectors have to be evaluated. Image locations where the length of model vector $|\mathbf{u}_M|$ is larger than the length of the corresponding optical flow vector $|\mathbf{u}|$ are not considered as an image of an obstacle, since the distance of the projected scene point to the moving vehicle is larger than those points of the volume where obstacles have to be detected. If the length of the model vector is less than the length of the optical flow vector, the ratio of the absolute values of difference vector $|\mathbf{u}_D|$ and model vector $|\mathbf{u}_M|$ will be compared to a threshold $\theta=0.3$ for all experiments. If ratio $|\mathbf{u}_D| / |\mathbf{u}_M|$ is larger than threshold θ, the image location will be considered a projection of a 3D-obstacle point. In areas around the FOE this ratio would become larger than θ even if the absolute differences were small. Therefore, another test is

performed to make sure that the denominator of the ratio is significantly different from zero. The detection results for optical flow vectors in Figs. 1-3 are shown in Figs. 4-6.

It can be seen clearly that the monotonicity operator (Fig. 4) extracted only few features in the image of obstacles. It seems very difficult to reliable detect obstacles using these features alone. The contour element correspondence approach uses much more information. Contours belonging to the box in the center of the road (Fig. 5) are clearly detected to be an image of an obstacle as well as the points in the upper right image. However, some false alarms were raised due to problems with optical flow components along straight line contours varying in length. With the analytic approach (Fig. 6) image areas which cover the parking cars, trees, and the box in the center are correctly detected to be an image of an obstacle. However, image locations in the vicinity of the box in the center are incorrectly marked as obstacles because the hierarchical multigrid approach spreads some information across contours into a small neighborhood around the image of an obstacle. We expect to overcome this problem with a theoretically well-founded optical flow estimation approach (Schnörr 89) which responds much more strong to discontinuities in the optical flow field.

In Figs. 4-6 we also see that obstacles beside the road are marked due to the fact that the corresponding scene points are not in the road plane. To concentrate the detection of obstacles to the volume that will be swept by the moving vehicle, additional limitations of the so-called 'motion tunnel' (Zimmermann et al. 86) have to be inserted. Two additional virtual planes perpendicular to the road plane are introduced to reduce the volume (Fig 7). The 3D-coordinates of these additional virtual planes are determined by the width of the moving vehicle. This approach assumes that the velocity vector of the vehicle remains constant. Using this bounded motion tunnel, a model vector field can be estimated with the procedure described above. Due to shorter distances to the camera this results in larger model vectors to the side of the road and, therefore, scene points outside this volume are not considered to be a projection of an obstacle (Fig. 8).

CONCLUSION

The approach discussed in this contribution is an example for the interpretation of temporal variations in image sequences recorded by a translating camera. The encouraging results show how obstacles can be detected in image sequences taken from a translating camera by evaluation of optical flow vectors estimated with independently developed approaches. Further developments are neccessary to extend the approach to more general motion and more complex environments.

ACKNOWLEDGEMENTS

This work has been partially supported by the "Bundesministerium für Forschung und Technologie" of the Federal Republic of Germany and the Daimler-Benz AG during the "Verbundprojekt Autonom mobile Systeme" as well as by the Eureka project PROMETHEUS. I thank Daimler-Benz AG for supplying the facility to record image sequences from a moving bus. I thank N. Rehfeld for providing the contour element correspondences shown in Fig. 3, and G. Hager, H.-H. Nagel and G. Zimmermann for their discussions and comments on a draft version of this paper.

REFERENCES

Aggarwal, J.K, N. Nandhakumar (1988).
On the Computation of Motion from Sequences of Images - A Review. Proceedings of the IEEE, Vol. 76, No. 8 (1988) 917-935

Bruss, A.R. , B.K.P. Horn (1983).
Passive Navigation. Computer Vision, Graphics, and Image Processing 21 (1983) 3-20

Daily, M.J., J. G. Harris, K. Reiser (1987).
Detecting Obstacles in Range Imagery, Image Understanding Workshop, Los Angeles/California February 23-25, 1987, 87-97

Dickmanns, E.D., Th. Christians (1989).
Relative 3D-State Estimation for Autonomous Guidance of Road Vehicles, Intelligent Autonomous Systems 2, T. Kanade, F.C.A. Groen, L.O. Hertzberger (eds.), 11-14 December, 1989, Amsterdam, 683-693

Enkelmann, W. (1988).
Investigations of Multigrid Algorithms for the Estimation of Optical Flow Fields in Image Sequences. Computer Vision, Graphics, and Image Processing (1988) 150-177

Enkelmann, W. (1989).
Interpretation of Traffic Scenes by Evaluation of Optical Flow Fields from Image Sequences, Control, Computers, Communications in Transportation, CCCT 89, Sept. 19-21, 1989, Paris, 87-94

Koenderink, J.J., A.J. van Doorn (1977).
How an ambulant observer can construct a model of the environment from the geometrical structure of the visual inflow, Kybernetik 1977, G. Hauske, E. Butenandt (eds.), Oldenbourg Verlag München Wien 1978, 224-247

Kories, R., G. Zimmermann (1986).
A Versatile Method for the Estimation of Displacement Vector Fields from Image Sequences. Workshop on Motion: Representation and Analysis, Kiawah Island Resort, Charleston/SC, May 7-9, 1986, IEEE Computer Society Press, 1986, 101-106

Nagel, H.-H. (1987).
On the estimation of optical flow: relations between different approaches and some new results. Artificial Intelligence 33 (1987) 299-324

Nelson, R.C. (1988)
Visual Navigation, PhD Thesis, Center for Automation Research, University of Maryland, August 1988, CAR-TR-380

Olin, K.E., F.M. Vilnrotter, M.J. Daily, K. Reiser (1987)
Development in Knowledge-Based Vision for Obstacle Detection and Avoidance, Image Understanding Workshop, Los Angeles/California February 23-25, 1987, 78-86

Rehfeld, N. (1988).
Dichte Verschiebungsvektorfelder entlang von Kantenzügen für zeitliche und stereoskopische Bildpaare. 10. DAGM Symposium, Zürich, 27.-29. September 1988, H. Bunke, O. Kübler, P. Stucki (Hrsg.), Mustererkennung 1988 Informatik-Fachberichte 180, Springer-Verlag Berlin Heidelberg New York, 1988, 97-103

Schnörr, C. (1989).
Zur Schätzung von Geschwindigkeitsvektorfeldern in Bildfolgen mit einer richtungsabhängigen Glattheitsforderung, 11. DAGM Symposium, Hamburg, 2.-4.10. 1989, H. Burkhardt, K.H. Höhne, B. Neumann (Hrsg.), Mustererkennung 1989, Informatik-Fachberichte 219, Springer-Verlag Berlin Heidelberg New York 1989, 294-301

Storjohann, K., E. Schulze, W. v. Seelen (1988).
Segmentierung dreidimensionaler Szenen mittels perspektivischer Kartierung. 10. DAGM Symposium, Zürich, 27.-29. September 1988, H. Bunke, O. Kübler, P. Stucki (Hrsg.), Mustererkennung 1988 Informatik-Fachberichte 180, Springer-Verlag Berlin Heidelberg New York, 1988, 248-254

Zimmermann, G., W. Enkelmann, G. Struck, R. Niepold, R. Kories (1986).
Image Sequence Processing for the Derivation of Parameters for the Guidance of Mobile Robots. Proc. Int. Conf. on Intelligent Autonomous Systems, Amsterdam, Dec. 8-10, 1986, Elsevier Science Publisher B.V., Amsterdam, 1986, 654-658

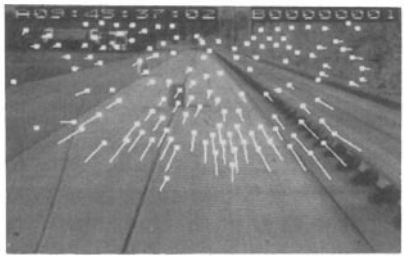

Fig. 1: Optical flow vectors calculated from features determined by the monotonicity operator (Kories & Zimmermann 86).

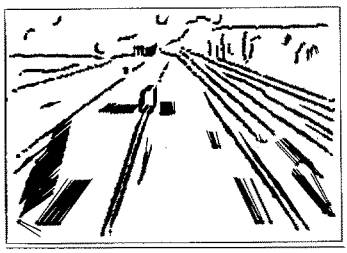

Fig. 2: Optical flow vectors at contour elements which could be matched by the approach of Rehfeld 88.

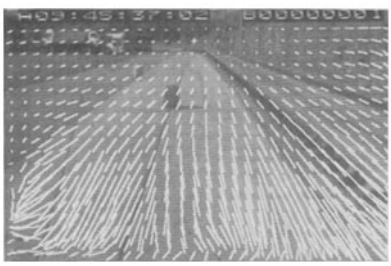

Fig. 3: Optical flow field calculated using a multigrid approach (Enkelmann 88). The FOE is marked by a small cross in the upper center of the image.

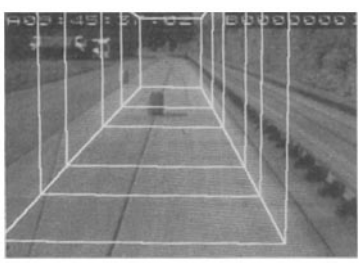

Fig. 7: The 'motion tunnel'. The lines perpendicular to the road axis have a distance of approximately 6m, 8m, 10m, 15m, 20m, and 30m, respectively.

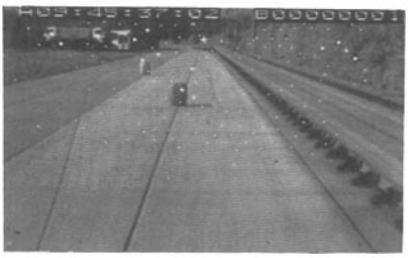

Fig. 4: Tracked feature points extracted by the monotonicity operator (Fig.1) which are considered to be projections of 3D-obstacles are marked with a square. Otherwise, the image location is marked with a little dot.

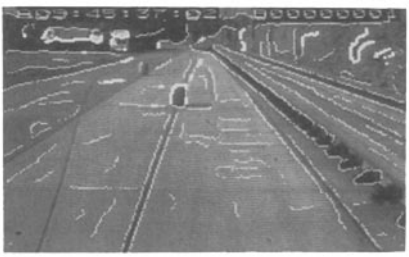

Fig. 5: as Fig. 4, but optical flow vectors were calculated from contour correspondences in Fig. 2.

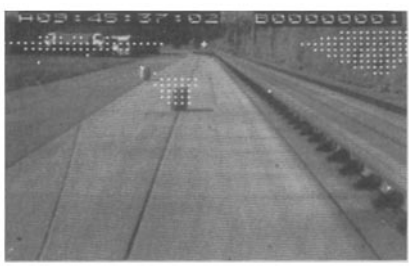

Fig. 6: Image locations which are considered to be projections of obstacles. The obstacles were detected by evaluating the optical flow field in Fig. 3.

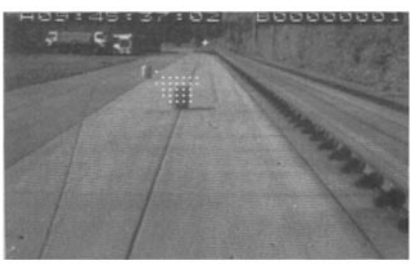

Fig.8: as Fig. 6, but only those image locations are marked which are considered to be projections of 3D-objects located within the 'motion tunnel'.