Obstacle Detection Based on Color Blob Flow

B. Heisele & W. Ritter
Daimler-Benz AG
Research and Technology
Image Understanding Systems (F3M/B)
P.O.Box 2360, 89013 Ulm, Germany
Phone (+49) 731-505-2140
Fax (+49) 731-505-4110
{heisele,ritter}@dbag.ulm.DaimlerBenz.COM

Abstract

In this paper we present a new robust approach to extract moving objects in traffic scenes. It allows moving obstacles (cars, motorcycles, ...) to be detected from a moving car without any a priori information about the shape or location of these moving objects. This technique is based on the concept of tracking color blobs in image sequences.

1 Introduction

Obstacle detection is one of the key functions in an autonomous driving vehicle. For vision based systems several approaches have been suggested to perform this task. Following [1], these approaches can be divided into three classes.

One class uses a priori knowledge of the obstacle's appearance in the 2-D image plane. They detect obstacles by localizing certain constellations of straight lines [2] or some symmetry features [2], [3] which describe the rear side of a car. Some of these algorithms work well in interstate ("Autobahn") scenarios where the shape as well as the object classes of potential obstacles are strongly restricted. In highway or country road scenes, however, these algorithms will probably fail due to large variations in obstacle classes and their 2-D shapes.

A second class of approaches uses stereo information [4], [5] to detect obstacles. It generates depth estimations from stereo correspondences of significant image structures such as corners or edges. In many cases the depth map is only sparsely filled with depth cues so that the information is not sufficient to extract the complete obstacle from the image.

The third class of the approaches detects obstacles based on motion information. A simple method which requires a stationary camera is the evaluation of the difference between new frames and the background image [6]. Most approaches, however, are based on optical flow. Here *continuous* and *discrete* methods, which will be explained in the next paragraphs, are distinguished.

Continuous approaches usually generate a displacement vector for each pixel by computing a spatio-temporal gradient of the local intensity distribution. The existing approaches resort to a smoothness constraint [7] to make the underdetermined problem of optical flow computation solvable. This is why errors occur at discontinuities in the velocity field, which hinder motion segmentation.

In contrast to continuous methods, discrete methods use image features such as corners [8] or local intensity minima and maxima [9] which are matched in adjacent images to obtain the displacement vectors. These displacement vectors are very precise. However, the resulting vector field is often not dense enough for object extraction.

At this point the question arises as to how to obtain a precise displacement vector field which allows us to segment all moving objects of interest. Analogous to the discrete optical flow, where displacement vectors are generated by computing the motion of corner or edges of adjacent images, we generate displacement vectors for color blobs. The displacement vector field of the image is what we call color blob flow. Motion segmentation is now reduced to combining adjacent color blobs with similar motions.

To compute the color blob flow, we perform following basic steps:

- Color segmentation.
- Connectivity analysis to get a symbolic description of the color blobs.
- Tracking of the blobs on a symbolic level over a sequence of images to obtain the displacement vectors.

Based on the color blob flow, we finally extract objects by combining color blobs with similar motion into motion segments.

2 Color segmentation

One suitable technique for real-time unsupervised color segmentation is clustering in color space. This technique determines a fixed number of reference vectors in such a way as to optimize the representation of the color distribution in the original image. Then the color vector of each point in the original image is replaced by ist nearest reference color vector in the sense of Euclidian distance. In choosing suitable cluster techniques, a trade-off must always be made between the quality of quantization on the one hand and the amount of computation on the other hand. The cluster technique published in [10] is a good trade-off, where realtime computation is important.

As described above this method of “hard“ quantization can result in undesirable color speckles in the segmented image. This affect becomes increasingly noticeable as the number of clusters is increased. On the other hand if the number of clusters is too small, surfaces of different objects cannot be distinguished. In addition to these considerations, the color distribution in the original image must be considered when determining the number of clusters. As the number of significant colors in the image increases, the number of clusters must be increased in order to maintain fineness of segmentation. A clustering with sixteen reference colors has proven to be a practical value in images so far investigated.

3 Connectivity analysis

The color segmentation described above produces a color labeled image with a fixed number of (currently sixteen) different colors. After that, neighboring pixels must be grouped into areas of common color. To compute the *color-connected components*, we use a fast algorithm proposed in [11].

It is an efficient, sequential, one-pass algorithm for generating the border line chain for each component. Based on the border-line chain, it also computes region attributes such as the area, the centroid, and the bounding box of the color-connected component with little additional effort.

Simultaneously the algorithm produces for each component a list of all adjacent components that are encountered during a walk around the border line, thus providing full topological information, which can be used for matching components in adjacent images.

4 Generation of the color blob flow

The decisive step in this technique is correctly matching corresponding color blobs in adjacent images. From the connectivity analysis described above, we have a number of attributes at our disposal which may be used in surface matching. Color is one similarity feature, which for the most part is independent of the motion of the objects and is thus especially well suited for matching. This one feature, however, is not sufficient. For additional features we used color blob area and aspect ratio of its bounding box. At this point we must note, however, that both area and aspect ratio are subject to change through occlusions and rotations. In these cases they can be used as matching criteria only if the image-to-image differences are relatively small, i.e. if the image frame rate is sufficiently large relative to the rate of motion in the objects in question.

From two color blobs A and B, one in each of two adjacent images, we define the difference measure for each of these three features as follow:

- color difference:

$$D_{C_{AB}} = \|\mathbf{w}_A - \mathbf{w}_B\| \text{ with } \mathbf{w} = (r,g,b)^T \qquad (1)$$

- relative surface difference:

$$D_{A_{AB}} = \left| \frac{A_A - A_B}{A_A} \right| \quad (2)$$

- difference measure for the aspect ratio:

$$D_{S_{AB}} = \left| \frac{S_A - S_B}{S_A + S_B} \right|, \quad \text{with } S = \frac{width}{height} \quad (3)$$

The total deviation is given as

$$D_{AB} = w_C D_{C_{AB}} + w_A D_{A_{AB}} + w_S D_{S_{AB}}, \quad (4)$$

where the w_i are arbitrary weighting constants for the corresponding features.

If

$$D_{AB} = \min\left\{ D_{AX} \,\middle|\, X \in \Phi, \|\mathbf{g}_A - \mathbf{g}_X\| \le d \right\}, \quad (5)$$

then the representation of a surface in two adjacent images P and Q may be associated to each other, where Φ is a set of all color blobs in Q, $\mathbf{g}$ is the centroid vector of the surface, and d the maximum allowable Euclidean distance of the centroid of two color blobs to be associated with each other.

How d is selected depends on the maximum movement to be detected in the image and also on the temporal and spatial resolution of the image in the sequence. Merging and occlusions of color blobs may make an association impossible. Tracking of these areas in cases like this is halted and tracking of newly arising color blobs is initiated as necessary. In order to carry out the motion segmentation described below, the centroid motions of color blobs are computed over several images.

5 Motion segmentation

Two color blobs are considered to form one motion segment if they are adjacent to one another and have similar centroid motions. A necessary condition for this is that the period of observation and the path transversed be sufficiently long. The measure for similar motion is the linear correlation of the centroid coordinates over the period of observation of the two color blobs and is given by:

$$cor_{AB} = \frac{\sum_n (\mathbf{g}_A(n) - \overline{\mathbf{g}}_A)^T (\mathbf{g}_B(n) - \overline{\mathbf{g}}_B)}{\sqrt{\sum_n \|(\mathbf{g}_A(n) - \overline{\mathbf{g}}_A)\|^2 \sum_n \|(\mathbf{g}_B(n) - \overline{\mathbf{g}}_B)\|^2}}, \quad (6)$$

where $\overline{\mathbf{g}}$ is the mean of the centroid over the period of observation.

Two color blobs are combined into one motion segment if

$$\begin{aligned} &cor_{AB} > c_{\min} \text{ and } B \in \theta_A, \\ &\quad \text{and } T_{AB} > T_{\min}, \\ &\quad \text{and } \min(p_A, p_B) > p_{\min}, \end{aligned} \quad (7)$$

where θ is the set of all color blobs whose bounding boxes overlap with that of surface A. T_{AB} is the period of observation of color blobs A and B, and p is the path length of the centroid motion during the period of observation.

Constants with the subscript “min“ designate minimal values of the corresponding variables.

The centroid motion of a motion segment is computed as the weighted mean of the centroid motions of the combined color blobs. Thus according to eqs. (6) and (7) an independent surface may be combined to an already existing motion segment.

6 Results

These experiments were carried out on traffic scenes on highways and interstates. Images were taken inside a test vehicle with a 3-chip color movie camera, which was connected to a digital video recorder (Y:U:V = 4:2:2, 720 x 576 pixels, 25 frames/sec.).

In Fig. 1 we see a typical four-frame sequence which shows two cars. The car in the foreground has just passed our test vehicle. The vehicle behind, which is also on the left side, is traveling at a much higher speed than our test vehicle, i.e. it is moving away from us. The test vehicle was traveling at about 80 km/h (50 mph).

The first frame (a) shows the results of color clustering with overlapping contours. In the second frame (b) those color blobs which could be matched to corresponding color blobs in the first frame are shown with a dark outline. Those which could not be matched are shown with a light outline. In the next frame (c) we see how the individual color blobs have been combined into motion segments: The vehicle in the foreground has been detected. The second vehicle, in the background, is not detected until frame d, because it is further away from our test vehicle and therefore appears to be moving slowly.

Fig. 2 shows a cross-road scene where a white car is passing the crossing. Fig. 2a shows the original image. Fig. 2b shows the contours that belong to the color blobs combined into the motion segment. The dark lines represent the tracks of the color blob centroids after a six-frame period of observation.

As is shown in the last two figures the detection is independent of the particular view from the vehicle.

7 Summary

In our approach we presented a new method for isolating moving objects from a non-stationary background. The main idea is to track color blobs, generated by color space clustering, over an arbitrarily long image sequence. Motion segmentation is now reduced to combining adjacent color blobs with similar motions. The criteria for matching color blobs from two adjacent frames are color, area, position of centroid, and the aspect ratio of the bounding box. Two adjacent color blobs in a given frame are combined to a single motion segment whenever the correlation between the motions of the centroid points over several frames is sufficiently high.

First experiments with traffic scenes taken with a non-stationary camera show that this technique produces relatively good results with both small and large object displacements.

References

[1] Enkelmann, W., Gengenbach, V., Krüger, W., Rössle, S., Tölle, W., "Obstacle Detection by Real-Time Optical Flow Evaluation", *IEEE Symposium on Intelligent Vehicles*, pp. 97 - 102, Paris, 1994

[2] Thomanek, F., Dickmanns, E. D. and Dickmanns, D., "Multiple Object Recognition and Scene Interpretation for Autonomous Road Vehicle Guidance," *IEEE Symposium on Intelligent Vehicles*, pp. 231 - 236, Paris, 1994

[3] Schwarzinger, M., Zielke, T., Noll, D., Brauckmann, M., von Seelen, W., "Vision-Based Car-Following: Detection, Tracking, and Identification," *IEEE Symposium on Intelligent Vehicles*, pp. 24 - 29, Detroit, 1992.

[4] Saneyoshi, K., Hanawa, K., Kise, K., Sogawa, Y., "3-D Image Recognition System for Drive Assist," *IEEE Symposium on Intelligent Vehicles*, pp. 60 - 65, Tokyo, 1993.

[5] Brauckmann, M. E., Goerick, C., Groß, J. and Zielke, T., "Towards All Around Automatic Visual Obstacle Sensing for Cars," *IEEE Symposium on Intelligent Vehicles*, pp. 79 - 84, Paris, 1994.

[6] Kilger, M., "A shadow handler in a video-based real-time traffic monitoring system," *Proc. IEEE Workshop on Applications of Computer Vision*, Palm Springs, CA, pp. 1060 - 1066, 1992

[7] Horn, B. K. P. and Schunck, B., "Determining optical flow," *Artificial Intell.*, vol. 17, pp.185-203, 1981

[8] Weng, J., Ahuja, N., Huang, T. S., "Matching two perspective views," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAMI-14, no. 8, pp. 806 - 825, 1992

[9] Koller, D., Heinze, N., Nagel, H.-H., "Algorithmic characterization of vehicle trajectories from image sequences by motion verbs," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Maui, HI, pp. 90 - 95, 1991

[10] Wu, X. and Witten, I. H., "A fast k-means type clustering algorithm," Dept. Computer Science, Univ. of Calgary, Canada, 1985

[11] Mandler, E. and Oberländer, M., "One-pass Encoding of Connected Components in Multi-Valued Images," *10th International Con ference on Pattern Recognition (ICPR)*, pp. 64 - 69, Atlantic City, 1990

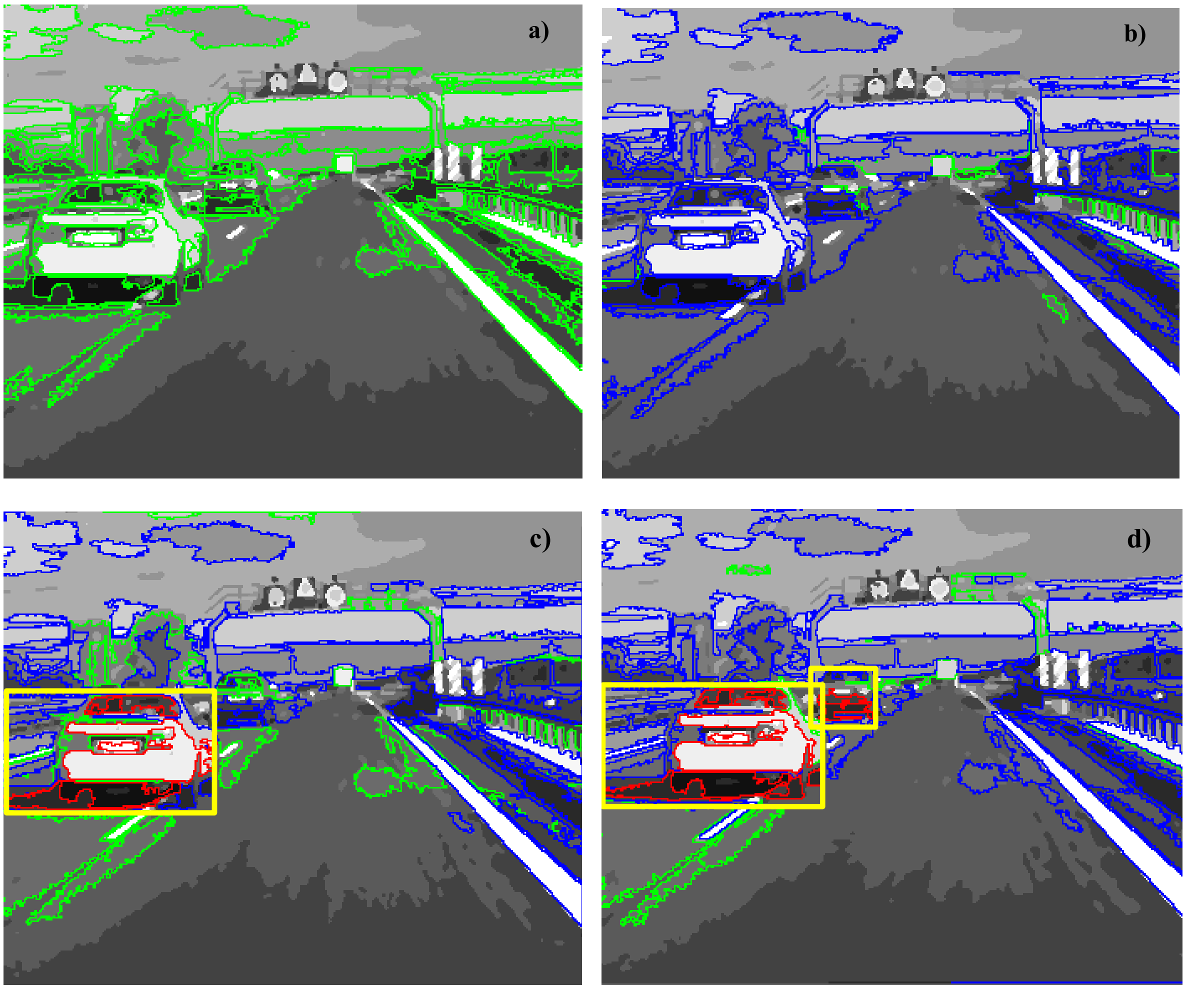

Fig. 1: Obstacle detection in a traffic scene by tracking color blobs. Dark outline denote blobs which correspond to color blobs in the preceding frame. Color blobs which have been combined into a single motion segment are enclosed in yellow rectangles.

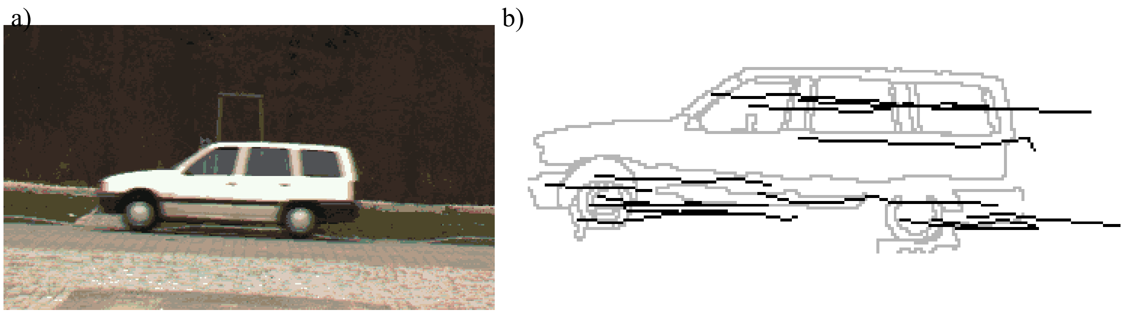

Fig. 2: A cross-road scene where a white car is passing the crossing. a) shows the original image, b) shows the contours of the color blobs belonging to one motion segment. The dark lines represent the tracks of the color blob centroids after a six-frame period of observation**.**