Fusion of Different Sensors and Algorithms for Segmentation

Uwe Handmann, Gesa Lorenz, Thomas Schnitger, and Werner v. Seelen
Institut für Neuroinformatik, Lehrstuhl für Theoretische Biologie, Ruhr-Universität Bochum, Germany

ABSTRACT

On this poster we present a system for coupling different basic algorithms and sensors for segmentation. Three different solutions for image segmentation by fusion are described, compared and results are shown. The fusion of basic algorithms with color-information and a sensor fusion process of an optical and a radar sensor including a feedback over time is realized. A feature-in decision-out fusion process is solved. For the fusion process a multilayer perceptron (MLP) with one hidden layer is used as a coupling net. The activity of the output neuron represents the membership of each pixel to an initial segment.

INTRODUCTION

- Traffic Scene Analysis
- System Architecture
- Different Algorithms

Fully or partly autonomously guided vehicles, particularly for road-based traffic, impose high demands on the development of suitable algorithms. This is due to the conditions imposed by natural environments. At the Institut für Neuroinformatik in Bochum, Germany, present projects are concerned with the analysis of traffic scenes. The great variety of different scenarios as well as the high degree of reliability necessary for the given task require an encompassing and flexible system architecture (figure 1).

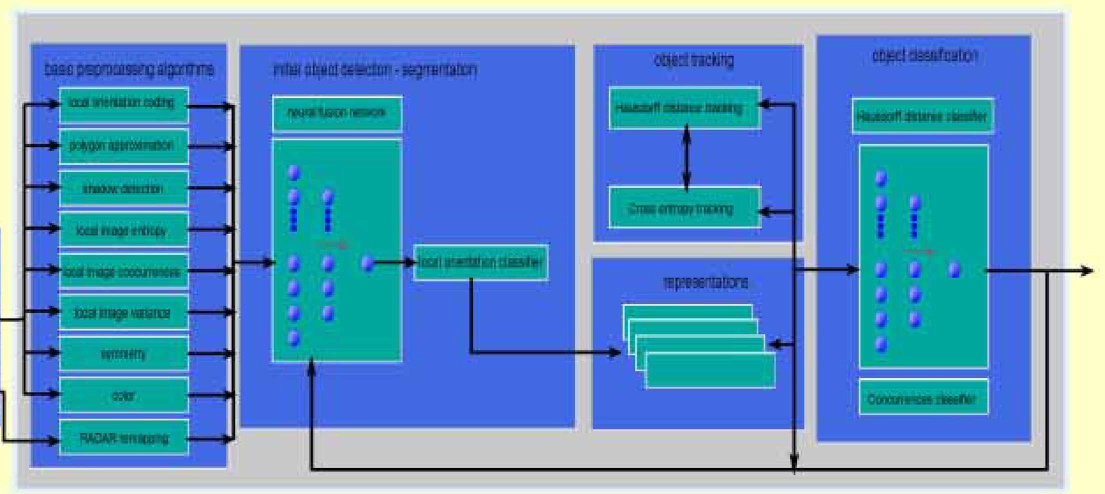

Figure 1

- Segmentation
- Classification
- Tracking

In principle the analysis of these scenes is a hierarchical process with a segmentation, a classification, and a tracking task (figure 2).

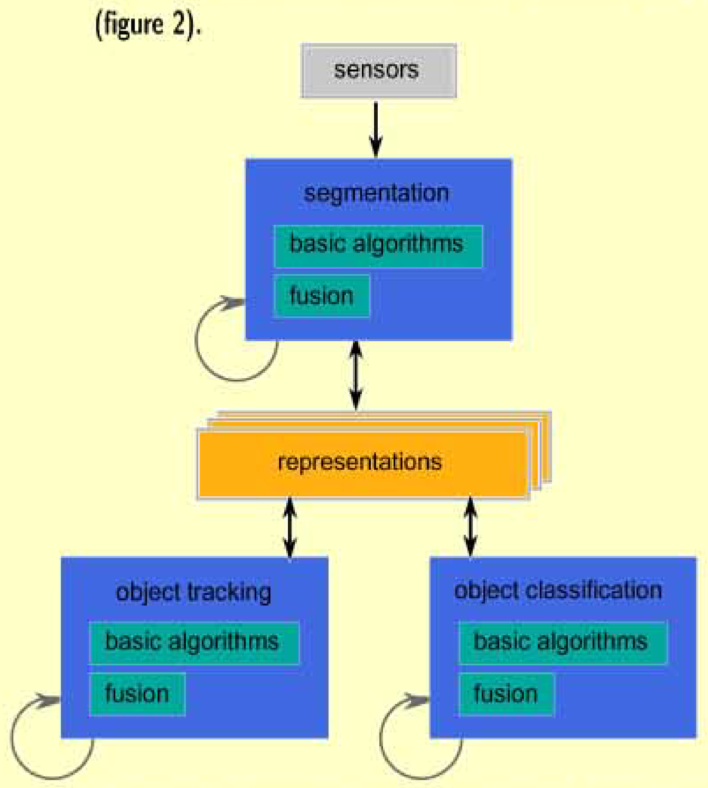

Figure 2

- Basic Algorithms
- Representations
- Adaptation

The variety of geometric appearances of involved objects and the variety of environmental constraints of both deterministic as well as statistical nature necessitate a multitude of partial solutions based on different representations of the environment. Consequently the structure of the system has to be adaptable to allow accommodation of additional modules without degeneration of already accomplished partial solutions. For this reason, even 'simple' applications are encumbered by considerations concerning the system architecture.

- Object Hypotheses
- Background

The segmented picture represents a substantial aspect of the automatic scene analysis. By segmenting a partitioning of the picture in object hypotheses and background is understood (figure 3). The generated object hypotheses are classified and tracked within the total system in further processing steps.

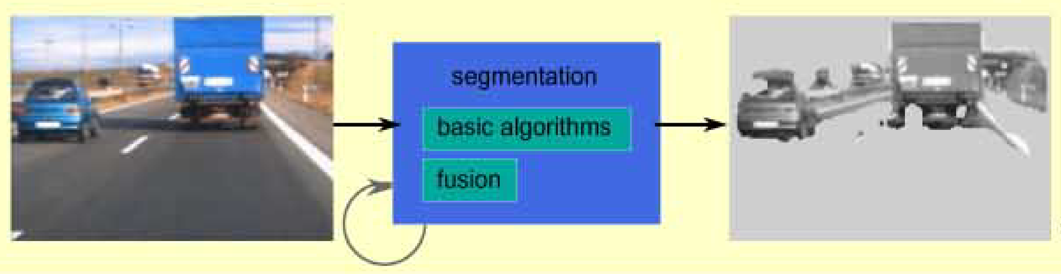

Figure 3

FUSION PROCESS

- Fusion Process
- Feature-In Decision-Out
- Multilayer Perceptron

On this poster three solutions of the fusion process for segmentation are presented. A feature-in decision-out fusion process is solved. For the fusion process a multilayer perceptron (MLP) with one hidden layer is used as a coupling net. The activity of the output neuron represents the membership of each pixel to an initial segment.

- Basic Algorithms
- Feature Vector
- Fusion

Essentially the total system can be divided into basic algorithms and algorithms for the fusion process. The basic procedures supply special partial solutions with given boundary conditions. The results of the individual algorithms are not independent, so the fusion of the results entails an increase of redundancy making the total system safe and reliable. The algorithms of the fusion process provide for a flexible interaction and for an integrative result of the basic components (figure 4).

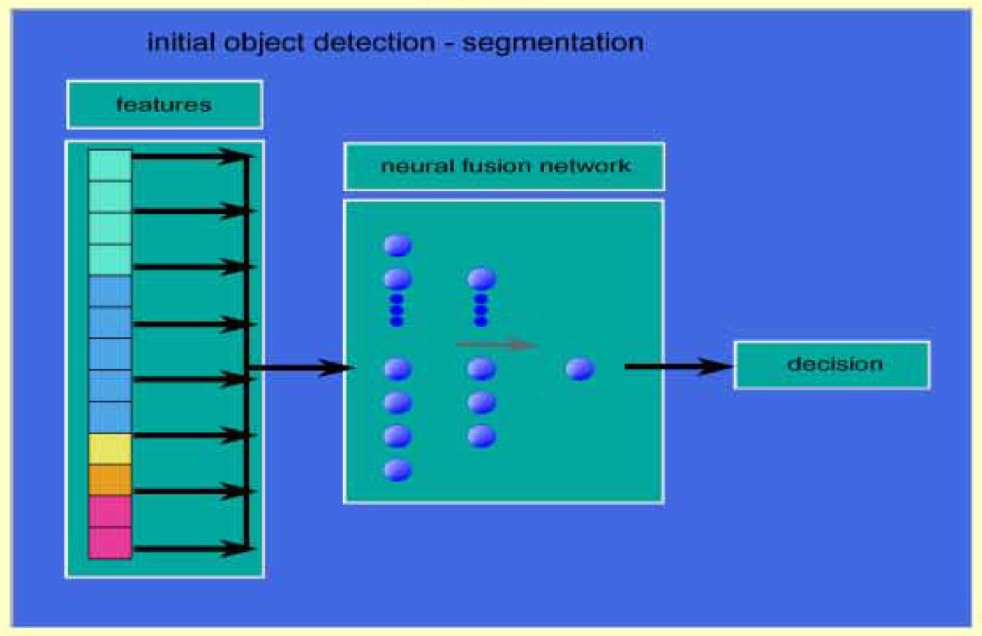

Figure 4

- On-Line Learning
- Feedback
- Sensorfusion

With a multilayer perceptron as coupling structure the coupling is learnable and flexible. An on-line learning is just as possible as a feedback over time. Additional sensor information (radar, lidar, position) or other basic algorithms can easily be integrated at this level. The flexibility of the coupling structure is clarified with special solutions. Three examples of the segmentation task using the fusion process are implemented. First, a simple application is described. Second an integration of color information is shown and third a sensor fusion process is solved.

- Fusion Process
- Feature Vector
- Relevant Segments

A fusion process can take place on different hierarchical levels. In general, a fusion process can be established on the data -, feature or decision level. In the present case a fusion process on the feature level is selected, in order to use the advantage of the data reduction compared with a fusion on the data level. Relevant result features of the basic algorithms are assembled to build a feature vector. The features are fused with the help of a neural net. A threshold at the output activity (decision) assigns the individual pixels of a frame to the background or a relevant segment.

SEGMENTATION

- Realizations of the Fusion Process

In this chapter three realizations of the fusion process to solve the segmentation task are described. The results are shown on different traffic scenes with specific requirements. First a simple fusion of differentiating and integrating basic algorithms is shown to clarify the principle of the fusion process on the feature level. In the second part the integration of color information into the fusion process is realized. A third implementation shows a complex feature fusion process. A sensor fusion, using optical and radar information, is realized. Additional basic algorithms and a feedback over time are integrated.

Three Realizations:

- Segmentation based on intensity image sequences
- Integration of color saturation for segmentation
- Radar and optical sensor fusion

Segmentation based on intensity image sequences

- Differentiating Features
- Integrating Features
- 12-5-1 Structure

In the first example differentiating and integrating basic algorithms (figure 5, figure 6) couple into a MLP (with a 12-5-1 structure), in order to solve the segmentation problem. The polygonal approximation of the contour and the LOC describe the differentiating features. The local image entropy is used as integrating feature. For each pixel a twelve-dimensional input vector

$$\mathbf{f}(x,y) = (\mathbf{f}_1(x,y)^T, x, y)^T \qquad (1)$$

for the coupling net is generated. The 10-dimensional vector $\mathbf{f}_1(x,y)$ includes the contour and texture information, the variables x and y represent the pixel coordinates. The vector is defined as

$$\mathbf{f}_1(x,y) = \sum_{(i,j)\in R} \mathbf{u}(i,j). \qquad (2)$$

whereby R describes the local neighborhood (9×9) of a pixel (x,y) and $\mathbf{u}(i,j)$ is a binary vector. The vector items $u_1(x,y),\ldots,u_4(x,y)$ encode reduced LOC features, $u_5(x,y),\ldots,u_9(x,y)$ encode the different entropy values and $u_{10}(x,y)$ is set, if the pixel is part of a polygon.

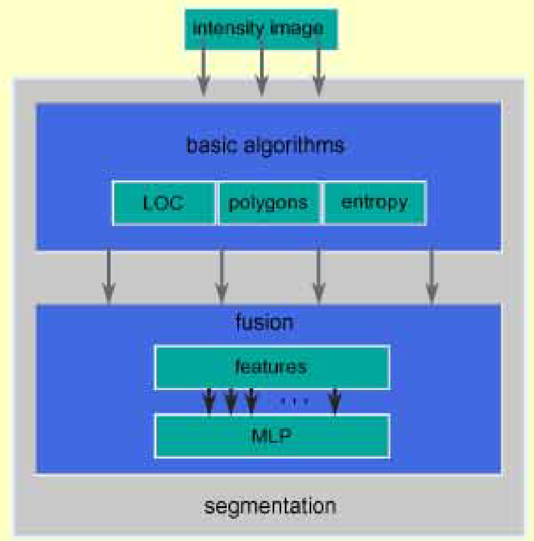

Figure 5

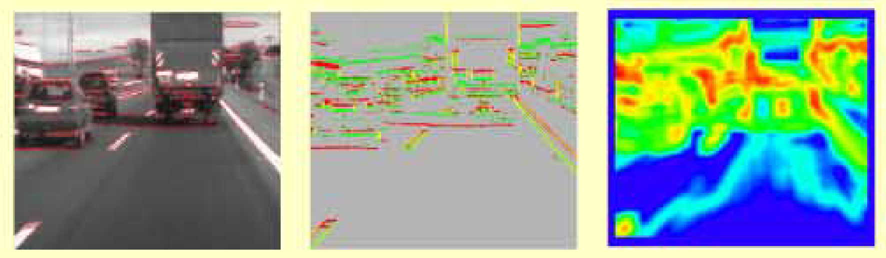

Figure 6

- Results

Figure 7 clarifies the segmentation result of vehicles in a sequence of 200 frames of a traffic scene, whereby each 50th frame is represented. All relevant objects are segmented stably. However partial false segmentation of objects with small contrast or missing structure is possible.

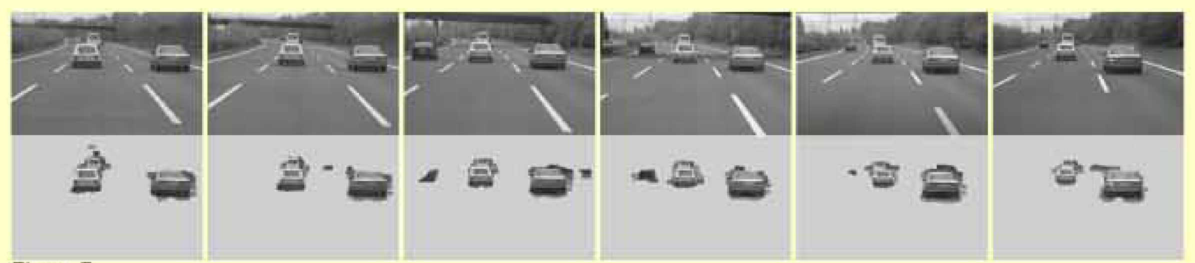

Figure 7

- Discussion

For analyzing more complex scenarios a solely standard camera (CCD-Camera) based approach is not sufficient. On the one hand the dynamic range of these cameras is small and on the other hand only the intensity of the scenarios is analyzed. These disadvantages can be avoided by the use of other cameras (e.g. HDRC, color camera) and the inclusion of further sensors.

Integration of color saturation for segmentation

- Color Features
- Improvement
- Homogeneous Areas

Color features have been used rarely so far to solve general detection tasks since color is generally not a specific property of objects. Vehicles are characterized mainly by a distinct form. In image analysis this form is often represented by contours coded here as lines and LOC features. Furthermore, they possess a distinct texture in comparison to the surrounding road surface. This property is expressed by the local image entropy. Numerous vehicles are characterized by a salient color. The use of color in addition to form and texture information improves the segmentation in situations where the exclusive use of the other features often fails, e.g. the segmentation of objects with large homogeneous areas.

The fusion concept proposed here offers the possibility to integrate color information into the coupling net apart from the already available differentiating and integrating features (figure 8).

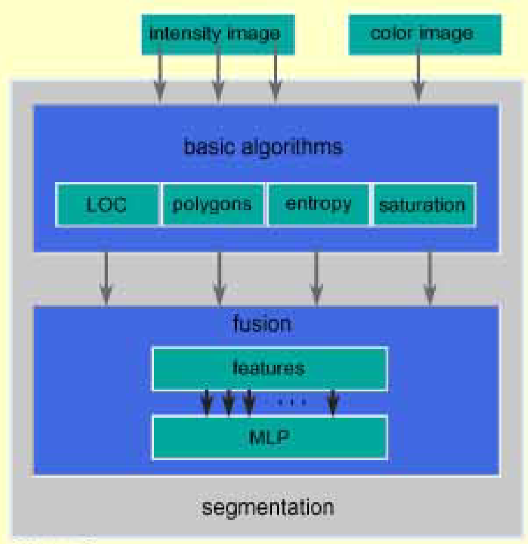

Figure 8

Color is described adequately by three quantities. Therefore, to allow a convenient specification of colors a 3-D coordinate system has to be established. In order to avoid intensive computation often the hardware-oriented RGB color space is chosen. We are interested in extracting highly saturated image areas independent of their actual hue or intensity value. Therefore, a color description based on the HSV color space describing color by the attributes hue, saturation and brightness is used.

- HSV
- Color Saturation
- 13-5-1 Structure

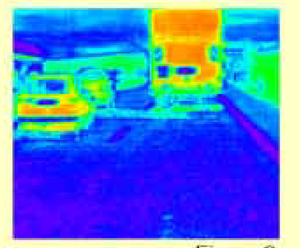

Figure 9

For each pixel only the saturation coordinate (figure 9) is evaluated

$$S = \frac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} \quad . \qquad (3)$$

Thresholding the saturation image $s(x,y)$ with s_{min} results in a binary image

$$v(x,y) = \begin{cases} 1 & \text{if } s(x,y) > s_{min} \\ 0 & \text{otherwise} \end{cases} \quad , \qquad (4)$$

serving as a basis for the new color feature

$$\mathbf{f}_2(x,y) = \sum_{(i,j)\in R} v(i,j). \qquad (5)$$

For each pixel a 13-dimensional vector is created

$$\mathbf{f}(x,y) = (\mathbf{f}_1(x,y)^T, \mathbf{f}_2(x,y), x, y)^T \qquad (6)$$

and fed into the coupling net (MLP with a 13-5-1 structure).

- Results
- Discussion

Figure 10 shows a comparison of the segmentation results with and without using the color feature in a sequence of 200 images of a traffic scenario presenting every 50th image.

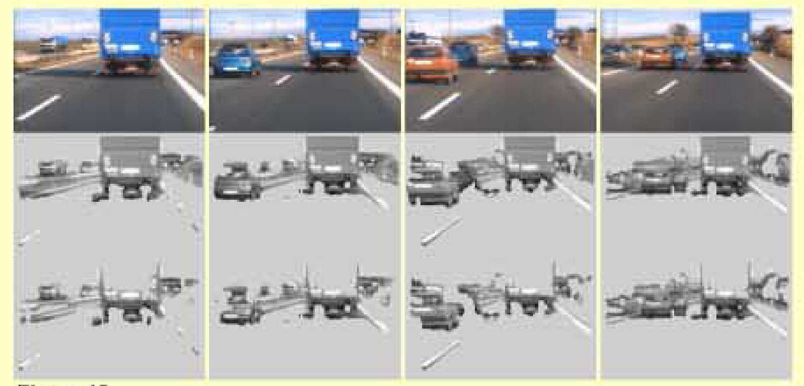

Figure 10

Whereas both methods succeed in segmenting the cars with good contrast, the additional color feature improves the segmentation results in the case of badly illuminated areas (figure 10, column 3) as well as the segmentation of trucks with large homogeneous areas (figure 10, row 3).

Radar and optical sensor fusion

- Complex Fusion
- Sensorfusion
- Feedback
- 16-5-1 Structure

A third complex implementation of a robust image segmentation uses besides the already discussed base algorithms a feedback over time, a local variance analysis, a shadow detection algorithm, and additional radar information to realize a sensor fusion process (figure 11, figure 12). The fusion combines data from independent sensors (optical and radar) to derive information that would be unavailable from the individual sensors.

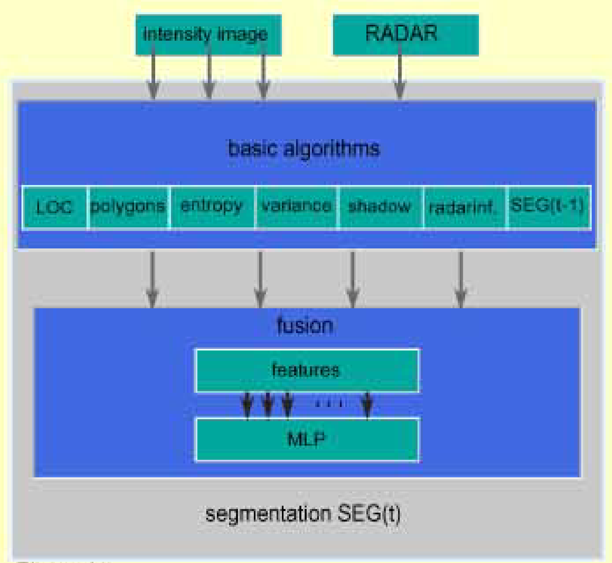

Figure 11

The fusion of the features $\mathbf{w}$ (feedback over time $w_1(x,y)$, local variance $w_2(x,y)$, shadow detection $w_3(x,y)$, and radar information $w_4(x,y)$) can be extended in this way to a multi sensor fusion. The vector $\mathbf{f}_3(x,y)$ is defined as

$$\mathbf{f}_3(x,y) = \sum_{(i,j)\in R} \mathbf{w}(i,j). \qquad (7)$$

whereby R describes the local neighborhood (9×9) of a pixel (x,y). A 16-dimensional input vector

$$\mathbf{f}(x,y) = (\mathbf{f}_1(x,y)^T, \mathbf{f}_3(x,y)^T, x, y)^T \qquad (8)$$

for the coupling net (MLP with a 16-5-1 structure) results for each pixel.

Figure 12

- Results
- Discussion

In figure 13 a sequence (every 250th frame) of 1500 frames with the result of the fusion process is shown. All vehicles are segmented, even if more than three vehicles (three radar beams) have to be detected. The system is robust with respect to wrong input of one of the coupled sensors or algorithms.

Figure 13