Proceedings of the 2005 CIKM Conference: Workshop on Research in Knowledge Representation for Autonomous Systems, October 31 - November 5, 2005, Bremen, Germany.

Towards an Approach for Knowledge-based Road Detection

Mike Foedisch
National Institute of
Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
+1-(301) 975-5516
mike.foedisch@nist.gov

Craig Schlenoff
National Institute of
Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
+1-(301) 975-3456
craig.schlenoff@nist.gov

Michael Shneier
National Institute of
Standards and Technology
100 Bureau Drive
Gaithersburg, MD 20899
+1-(301) 975-3421
michael.shneier@nist.gov

ABSTRACT

Our previous work on road detection suggests the usage of prior knowledge in order to improve performance. In this paper we will explain our motivation for a novel approach, define requirements and point out issues, particularly concerning the representation of road depending on the use, which need to be addressed. The proposed system will provide symbolic data for high-level processes and guidance for low-level processes. Furthermore, we will outline the recognition approach based on previously discussed requirements and issues. This paper has a visionary character based on our experience with road detection for autonomous road vehicles.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics - a*utonomous vehicles, sensors.*

General Terms

Algorithms, Design

Keywords

Road detection, road recognition, autonomous driving, road representation, model-based recognition, tree search, constraints

1. INTRODUCTION

In the last two decades a lot of effort has been spent in developing autonomous road vehicles, including significant research into methods to extract information about the road. Hereby, we can differentiate between road following, road detection, road recognition and road reconstruction depending on how the information about the road is used (see Bertozzi et al. [2], and Dickmanns [6]). While road following is the task of tracking the road (edges) over a sequence of images, e.g. using prediction techniques like Kalman filters, the road detection process tries to find areas in sensor data that represent road. Road recognition and reconstruction can be characterized as aiming to understand the type and structure of the road as well as its location in the world. However, several of these terms can be found synonymously in the literature.

Our previous approaches [8, 9] for road detection (i.e. detecting the road area in color images), which were merely bottom-up approaches, focused on the system's adaptability to new or changing environments.

In order to adapt the system continuously to the environment, we defined feature extraction windows. Three windows were placed over the road region in the image (green frames in Figure 1). Features taken from these windows were labeled as road. Similarly, three windows were placed over non-road regions and result in similar regions being labeled as non-road (red frames in Figure 1). The extracted features plus the labels defined by windows make samples, which are used for periodic update of the road detection system.

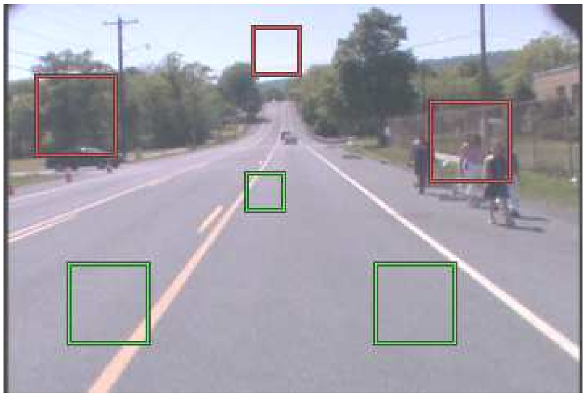

Figure 1: Pre-defined feature extraction windows

However, this static window structure is based on the assumption that the windows that represent road regions will continue to represent such regions in subsequent frames, and similarly non-road windows. But this assumption is frequently violated, for example, when there is a curve in the road (see Figure 2).

KRAS'05, November 5, 2005, Bremen, Germany.
ACM 1-59593-202-X/05/0011.

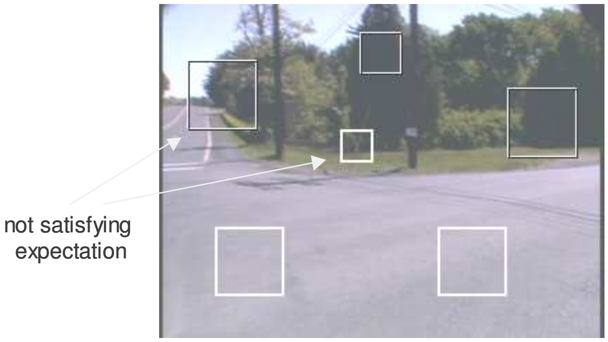

Figure 2: Example of windows violating assumptions due to a change of curvature of the road

To overcome this problem, we developed a variant in which the positions of the road windows are not fixed. Initially, all windows are placed in the image as depicted in Figure 3(left). The systems' continuous adaptation to the appearance of road and non-road areas causes the road windows to smoothly cover the entire road area in the image (see Figure 3(right)).

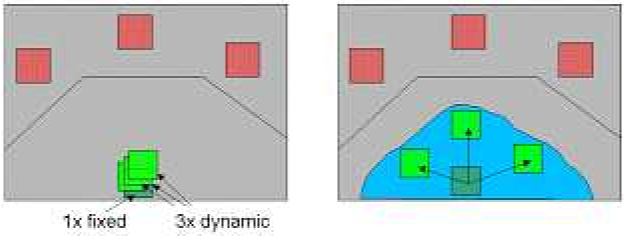

Figure 3: Dynamic windows approach

The evaluation of both variants showed that the static structure approach allows better for sudden changes (like shadows) if the current road type complies with the static windows' position. The dynamic approach, however, covers better smooth changes of the road's appearance as well as arbitrary road shapes. We concluded that the placement of feature extraction windows is the key for performance improvement. Therefore, we suggest a tight coupling of road detection and road recognition through a knowledge-based approach to road detection. Knowledge about the type of road in geometrical and topological terms would help us to extract sample data for the adaptation purpose in a more informed and certain way. The knowledge of the exact location of the road in the image would prevent us from extracting erroneous sample data.

Another argument for a knowledge-based approach can be derived from the typical design of autonomous systems (e.g. RCS in Albus et al. [1]). The path from sensor data up to modules making decisions and down again to those that control the system is typically hierarchically organized. The higher the level where decisions are made the more symbolic the representation of data becomes. The approach discussed in this paper will provide this higher level of symbolic information about road. Figure 4 shows the placement of the new approach between low-level sensor processing (which it supports by guidance) and high-level control (for which it provides symbolic data).

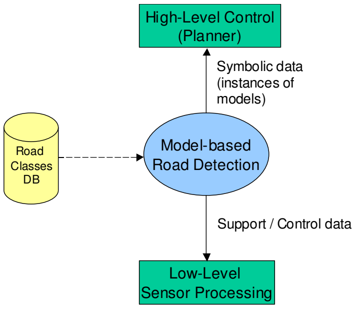

Figure 4: Placement of our knowledge-based road detection approach

Several previous approaches (e.g. [3, 10, 13]) incorporate model information in order to improve the performance of their algorithms. But the term "model" in the field of road detection is not unambiguous, e.g. it could stand for a statistical or geometrical representation rather than a symbolic description. Our models will incorporate symbolical (topological) as well as geometrical information.

In Gengenbach et al. [10] a geometrical model is extracted from a-priori digital maps from commercial automatic navigation systems. The model contains information about the lane structure, lane widths and distances between road junctions and intersections. The focus of the approach is on road tracking and steering control, especially reducing the longitudinal uncertainty in the position of the vehicle. In contrast, our approach is based on extracting topological information and only rough geometrical data, while road following and accurate geometrical processing of data is left to other modules. However, we believe such modules benefit from guidance by the symbolic data our system provides. Instead of using a-priori maps, our approach is based on an abstract knowledge base describing basic concepts, limitations, and constraints of the road domain.

In Chapuis et al. [3] a statistical model containing (among others) the coordinates of the left and right road edge is created. The model is designed to use Kalman filters for prediction in order to track road edges. The algorithm employs tree search in order to find coherent features. We follow a similar approach but employ tree search to find consistent interpretations of the road structure.

In Luetzeler et al. [13] a road recognition approach is presented, which employs near and far range models in order to support road detection on a set of multi-focal cameras. A straight skeleton line centered on the road describes the near range model. On multi-lane roads, lanes are described by an offset from the skeleton line and an individual lane width. The approach uses structural constraints (in order to reconstruct the lane structure) and real-time constraints for efficiency reasons. We follow a more explicit road model approach, which also covers transitional segments (e.g. from two lanes to one lane) and intersections. However, the use of constraints will be an essential part of our system as well.

Another approach employing constraints for road reconstruction was introduced in DeMenthon et al. [5]. The approach was based on finding opposite points on the road (points which face each other on opposite sides of the road). We will use a similar type of features. But, instead of assuming a known and constant road width, we allow the width of the road to be close to that of any of the models in our model base and we allow the width of road to change according to the legal connectivity (i.e. through transitional segments) in our model base.

Our goal is to develop a stable and reliable road detection, which is also adaptable (to changes in environments) and extendable (by new road types). In the following, we define some of the requirements the new road detection system for an autonomous road vehicle should meet.

The system:

a) needs to understand the current topology of the road in terms of legal traversability, e.g. through the lane structure;

b) needs to recognize the geometry of the road in world coordinates in order to control the vehicle's steering. (partly a road following problem);

c) should be able to adapt to changes of the environment, e.g. changes in the appearance of the road depending on the road surface and lighting conditions;

d) should be easily extendable to new types of road, e.g. through simple models;

e) has to run in real-time.

We will first discuss issues related to the representation of roads for different purposes in section 2. Then, in section 3, we will describe our choice of recognition approach as well as feature data extraction and model representation.

2. REPRESENTATIONAL ISSUES

Typical knowledge-based recognition systems employ a database of recognizable objects. The representation of such objects depends heavily on their character and appearance on sensor data and is therefore critical to the system's performance.

There are several ways to represent a road, depending on how the information is to be used. In the following, we will discuss issues and situations that directly or indirectly affect the representation of roads in order to support *road recognition*, *high-level planning* and *performance evaluation*.

2.1 Recognition - related Issues

The representation of road models as well as feature data derived from sensor data need to be suitable to handle the challenges of a real world environment.

Noise

Noise is a typical problem, which has to be addressed in a sensor processing system. In the case of road detection, noise could confuse the system and cause it to incorrectly interpret the structure of road. Therefore, the recognition system should allow for alternative interpretations.

Image Edges

A special situation occurs in road images (taken from the driver's point of view) when the vehicle is passing an intersection. The areas representing road in the image will eventually touch the edge of the image. The same holds for road areas in the image showing road that is directly in front of the vehicle (see Figure 5). Therefore, the topological interpretation of road areas touching the image edge depends on the context. In order to handle this uncertainty the recognition system needs a conceptual representation of *unknown* (or not-yet-known) road segments.

Figure 5: Example of road areas touching the edge of the image (left) and projection in world coordinates (right)

Relationships between Concepts

More complex structures, e.g. intersections, can be expressed by relationships between simpler concepts. For example, the simple concepts of *regular road, road widening* and *road narrowing* could be used to describe an intersection. Higher-level concepts can be enriched by additional information, for example in order to express the order in which simpler concepts appear, or geometrical constraints like minimal and maximal lengths.

Normal Orientation

Another issue concerns the orientation of the vehicle on the road. Figure 6 shows examples of a vehicle's orientation on the road. A *normal orientation* (Figure 6(left)), where the vehicle is limited to sit only in lanes for which the legal driving direction agrees with the vehicle's direction, provides a canonical form of the appearance of road on images and may therefore simplify the representation process. All other orientations of the vehicle do not comply with the normal orientation (Figure 6(center/right)). We can allow the limitation of a normal orientation if we assume that the autonomous system will be aware of when it leaves the normal orientation (e.g. due to avoidance of obstacles on the road).

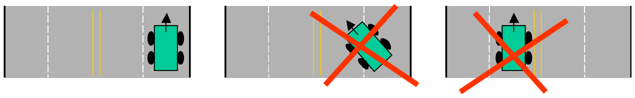

Figure 6: Normal orientation, legal and illegal orientations

Hierarchical Representation

Details of the road (e.g. lane markings) might generally not be as reliably extractable as coarse structures, like the width of the road. Figure 7 shows an example of two road types, which appear similar in width but different in topology. A hierarchical representation of road types containing more general concepts (i.e. class of roads with certain widths) would allow the system to act on at least the level of information available so far, e.g. to continually adjust the system to the environment (see section 3).

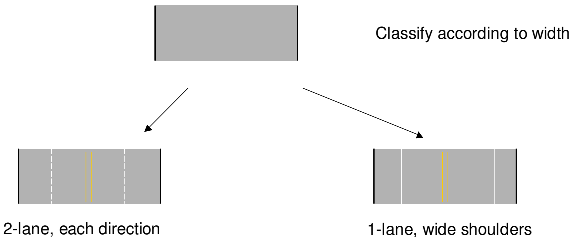

Figure 7: Hierarchical Representation

2.2 High-level Planning – related Issues

High-level control and planning for an autonomous vehicle must be aware of and must respond appropriately to any object it encounters. This includes other vehicles, pedestrians, debris, construction, accidents, emergency vehicles, and it also includes the roadway itself. The road network must be described in such a way that an autonomous vehicle knows, with great precision and accuracy, where the road lies, rules dictating the traversal of intersections, lane markings, road barriers, road surface characteristics, and other relevant information.

A Road Network Database is being developed at the National Institute of Standards and Technology (see Schlenoff et al. [14]). Its purpose is to provide the data structures necessary to capture all of the information necessary about road networks so that a planner or control system on an autonomous vehicle can plan routes along the roadway at any level of abstraction. At one extreme, the database should provide structures to represent information so that a low-level planner can develop detailed trajectories to navigate a vehicle over the span of a few meters. At the other extreme, the database should provide structures to represent information so that a high-level planner can plan a course across a country. Each level of planning requires data at different levels of abstraction, and as such, the Road Network Database must accommodate these requirements.

The fundamental components of the Road Network Database are:

- Junctions, Intersections
- Lane Junctions
- Road, Road Segment, Road Element
- Lane Cluster, Lane, Lane Segment, Junction Lane Segments
- Time Varying Attribute Tables, Lookup Tables

2.3 Performance Evaluation Issues

For the purpose of performance evaluation the representation of road depends on the level and structure of the ground truth (i.e. reference data) and vice versa. For example, in order to compare our new system's results with previous approaches we need to provide a similar representation.

Figure 8: Example of previous results

Figure 8 shows a sample result from a previous road detection system (see Conrad et al. [4]). White dots depict road areas recognized by the system. While small white dots describe image areas, which were correctly (according to ground truth) recognized as road, large white blocks show areas erroneously recognized as road. The same holds for small and large black dots, which describe non-road areas. The ratio of wrongly recognized areas to the overall number of areas gives a measure of the correctness of the classification result (see Hong et al. [12]).

While the road representation for the ground truth consisted of a polygon (which described the boundary between road and non-road areas in the image), the recognition result was represented by a binary road / non-road label image. However, the proposed approach to road detection will provide additional information beyond mere road areas on images, in the form of the topology and basic geometry of the road in world coordinates. This holds the potential for new approaches to performance evaluation on a qualitative and quantitative level. On the one hand the correctness of the symbolic level of the road representation can be measured by comparing the symbolic structure (e.g. a chain of road types) with the corresponding ground truth. On the other hand geometrical data describing the occurrence of road structures can be compared with geometrical earth data collected by satellite or overflight.

We defined requirements for a knowledge-based approach to road detection in section 1 and discussed issues with the representation of roads in section 2. Both requirements and representational issues contributed to our choice of a recognition approach, which we discuss in the next section.

3. RECOGNITION APPROACH

In this section we will describe our proposed knowledge-based road detection system in terms of the requirements and issues we pointed out above. We will explain the general approach as well as components of the recognition system.

Figure 9 gives an overview of the proposed approach. The left side of the graphic depicts the steps (low-level image processing, segmentation, classification) of a classical bottom-up approach for a recognition system. The segmentation step of the proposed system uses a trainable classifier to segment the image. The trainability allows the continuous update of the separation function between road and non-road areas in the image according to the current environment (see requirement "c" in section 1). After segmentation, another classifier is applied, which will recognize the type and structure of the road. In our case of a model-based approach, this classifier tries to fit models from a database of known road types to the segmentation results (in Figure 9 the road type "Straight Road" was recognized).

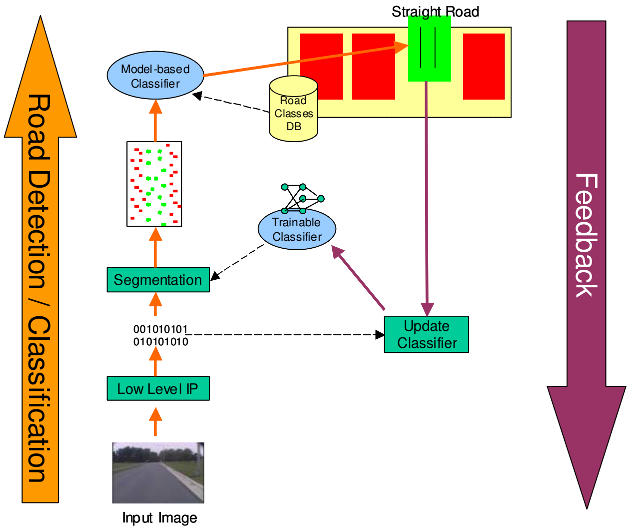

Figure 9: System overview

The resulting structural and topological description of the current road will now be used to control the update of the trainable classifier used for segmentation. This feedback in terms of symbolic knowledge about the road structure allows an informed extraction of training samples, which describe road or non-road areas.

In the following we will discuss the features extracted from data, the simple approach to represent road models and the matching principle in more detail. Finally, we will describe example results.

3.1 Feature Extraction

Assuming the proposed normal orientation of the vehicle on the road (see section 2.1) a simple set of features, which are easily extracted and well-understood can be derived (DeMenthon et al. [5] showed the applicability for the purpose of road detection). The features are based on "slices" of the road perpendicular to the direction of the vehicle. They can be extracted by applying one of several approaches for detecting the road area in images or road edge detecting algorithms (e.g. [3, 5, 7-9,13]).

Figure 10 gives an example of the features. Starting at the bottom image row, the left and right road edge points in each row are determined. A pair of road edge points described in both image and world coordinates (through camera calibration) describes one feature item. The process continues bottom-up row-by -row until the world coordinates of the road edges reach a given maximum distance in front of the vehicle (e.g. more than 55 meters).

Furthermore, additional data will be associated with a feature item, e.g. information about lane markings and indicators that describe if the road edge touches the image edge (see section 2.1).

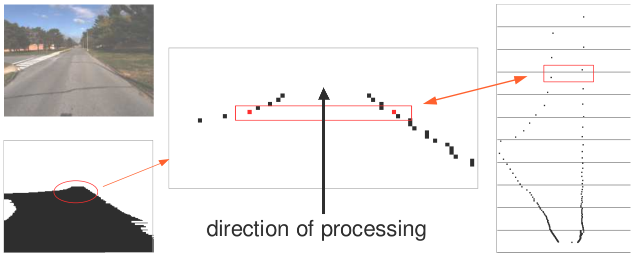

Figure 10: Feature Extraction - original image and road areas (left), road edge points (center), and projection into world coordinates (right)

3.2 Model Representation

Figure 11 depicts our approach for representing road model primitives. A "slice" of road is described by its width (geometrical component) and lane structure in terms of number of lanes and their legal direction (topological component). This representation of road primitives is compatible with the type of feature data (see section 3.1). While the parameter values (width and lane structure) will appear relatively constant in the case of a regular road, a range of legal values is needed in the cases of road widening and narrowing (e.g. as part of intersections or change of lane numbers).

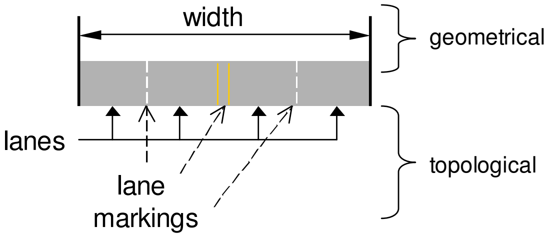

Figure 11: Geometrical and topological representation of a "slice" of road

A road type consists of an ordered group of primitive road model items. For such groups additional constraints apply. A road type might require a minimal and/or maximal lateral length or, in the case of road widening and narrowing, a certain monotonic behaviour. Other constraints limit the connectivity between (primitive) road types, e.g. a two-lane road segment can connect to a three-lane road segment only through a transitional segment.

Primitive road items and road types are organized hierarchically. Starting from abstract concepts like regular road, road widening and narrowing, concrete models, e.g. a four-lane road, can be derived. The requirement of simple extendibility (see section 1) is met by simply adding new primitive model items that represent new types of road.

Additionally, primitive model items are grouped by the type of driving environment, e.g. highway driving, rural road or urban road driving. Appropriate connectors describe transitions from one environment to another one (e.g. a highway exit transfers the vehicle from highway driving to rural road driving).

3.3 Matching

The purpose of the matching process is to find associations between feature data and models, and to create an interpretation of the road structure seen in sensor data.

In order to handle noise (see section 2.1) we follow an approach that provides alternative interpretations and structural information (see requirements “a” and “b” in section 1). Our choice is a constraint tree search, similar to the interpretation tree approach in Grimson [11].

The tree structure is built by associating each feature item with potentially all of the primitive model items in our model base. The potentially exponential search space is limited by applying constraints. Unary constraints limit the association of a feature item to model items, e.g. by selecting only models of a similar width. Additionally, context-based constraints limit legal associations by forcing compliance to group-based requirements, e.g. in terms of minimal and/or maximal length of the road type, as well as connectivity consistency.

Leafs of the tree will represent global interpretations of the road structure. After applying appropriate global consistency measures, a single global interpretation is selected.

3.4 Example Results

In this section we describe exemplary results of our proposed system for the typical road image shown in Figure 12 (upper left). After applying a road segmentation algorithm (lower left in Figure 12), road edge features are extracted as described in section 3.1. Each road edge pair is now potentially associated with all primitive models (see section 3.2) in out model base. Matches that do not comply with unary constraints (e.g. similar road width) are rejected. Each single association is connected with a group of neighboring associations of the same type. If the new association is of the same type as the association of the previous road edge pair, the new association has to survive context related constraints too. For example, assuming road edge pair *e1* was associated to road type *road narrowing* (see center of Figure 12), associating road edge pair *e2* also to road type *road narrowing* requires compliance to a context-related constraint which demands the road to shrink in width monotonously. If a new association creates a new group, then the previous group must be closable according to that group’s constraints, e.g. a widening segment must be of a minimum length.

Certain sequences of groups allow a high-level interpretation of the road. In the example in Figure 12 the occurrence of *a road widening segment* and a *road narrowing segment* plus an undefined *image edge touching segment* allows to conclude (in the context of a two-lanes road) the existence of a T-intersection.

4. SUMMARY

In this paper we motivated the need of a knowledge-based road detection based on our previous work. We defined requirements for such an approach and discussed issues with the representation of road. Both lead us to the conclusion to follow a model-based recognition approach using constraint tree search.

5. ACKNOWLEDGMENTS

We are grateful for the support of the Army Research Laboratory and the Defense Advanced Research Projects Agency.

6. REFERENCES

[1] Albus, J., Huang, H.M., Messina, E., Murphy, K., Jubets, M., Lacaze, A., etc., “4D/RCS Version 2.0: A Reference Model Architecture for Unmanned Vehicle Systems*”, NISTIR 6910, National Institute of Standards and Technology*, Gaithersburg, MD, August 2002

[2] Bertozzi, M., Broggi, A., and Fascioli, A., “Vision-based intelligent vehicles: State of the art and perspectives”, in Robotics and Autonomous Systems, Vol. 32, 2000

[3] Chapuis, R., Aufrere, R., and F. Chausse, “Accurate Road Following and Reconstruction by Computer Vision”, in

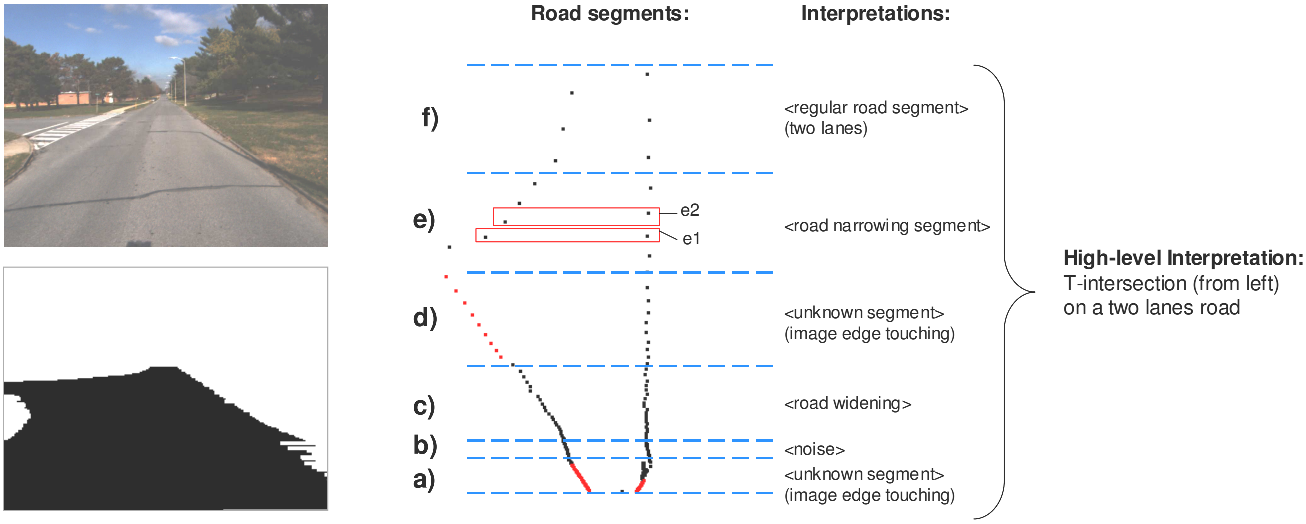

Figure 12: Exemplary Results

IEEE Transactions on Intelligent Transportation Systems, Vol. 3, No. 4, December 2002

[4] Conrad, P., and Foedisch, M. ,"Performance Evaluation of Color Based Road Detection Using Neural Nets and Support Vector Machines", in Proc. *Applied Imagery Pattern Recognition Workshop*, Washington D.C., 2003.

[5] DeMenthon, D., and Davis, L.S., "Reconstruction of a road by local image matches and global 3d optimization", In Proc. International Conf. on Robotics and Automation, 1990

[6] Dickmanns, E., "Vehicles capable of dynamic vision", *Proc. Int. Joint Conf. On Artificial Intelligence*, pp. 1577-1592, 1997.

[7] Dickmanns, E., and Mysliwetz, E., "Recursive 3-D Road and Relative Ego-State Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, February 1992

[8] Foedisch, M., and Takeuchi, A., "Adaptive Real-Time Road Detection Using Neural Networks", in *Proc. 7th Int. Conf. on Intelligent Transportation Systems*, Washington D.C., 2004

[9] Foedisch, M., and Takeuchi, A., "Adaptive road detection through continuous environment learning", in *Applied Imagery Pattern Recognition Workshop*, 2004. Proceedings. 33rd, Vol., Iss., 13-15 Oct. 2004

[10] Gengenbach, V., Nagel, H.-H., Heimes, F., Struck, G., and Kollnig, H., "Model-based recognition of intersections and lane structures*", Intelligent Vehicles '95 Symposium., Proceedings of the*, Vol., Iss., 25-26 Sep 1995

[11] Grimson, W. E. L., *Object Recognition by Computer: The Role of Geometric Constraints.* The MIT Press, Cambridge, Mass., 1990

[12] Hong, T., Takeuchi, A., Foedisch, M., and Shneier, M., "Performance evaluation of road detection and following systems*", Proc. SPIE Vol. 5609, Mobile Robots XVII*; Douglas W. Gage; Ed., Dec 2004

[13] Luetzeler, M., and Dickmanns, E., "Road recognition with MarVEye", in Proc. IEEE International Conference on Intelligent Vehicles, Stuttgart, Germany, October, 1998

[14] Schlenoff, C., Balakirsky, S., Barbera, A., Scrapper, C., Ajot, J., Hui, E., etc., "The NIST Road Network Database: Version 1.0", *NISTIR 7136, National Institute of Standards and Technology*, Gaithersburg, MD, July, 2004