Proceedings of the 1991 IEEE
International Conference on Robotics and Automation
Sacramento, California - April 1991

UNSCARF, A Color Vision System for the Detection of Unstructured Roads

Jill D. Crisman
Dept. of Electrical and Computer Engr.
Northeastern University
Boston, MA 02115

Charles E. Thorpe
Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

This research addresses the problem of navigating a robot vehicle on unstructured roads, which have no lane markings, may have degraded surfaces and edges, and may be partially obscured by strong shadows. These conditions cause many road following systems to fail. We have built a system, UNSCARF, that is based on pattern recognition techniques to successfully navigate on a variety of unstructured roads. UNSCARF does not need a road location prediction to find the location of the road; therefore, UNSCARF can be used as a bootstrapping system. UNSCARF uses a clustering technique to group pixels with similar color and location. UNSCARF then matches models of road shape to locate the roads in the image. These methods are more robust in noisy conditions than other road interpretation techniques. UNSCARF has been integrated into a navigation system that has successfully driven a test vehicle in many types of weather conditions.

1. Introduction

This paper discusses a color vision system that detects roads for an intelligent mobile robot. To navigate a mobile robot, a navigation system consists of perception systems to first sense the environment, path planning systems that decide on a pathway, and vehicle control systems that actuate the motion of the robot. For road navigation, at least one of the perception systems must sense the location the roads, and the path planning system must generate paths that keep the robot in the proper lane while avoiding any obstacles. This work focuses on the detection of roads in color images for robot navigation systems.

Current road detection systems perceive certain types of roads under a limited number of conditions. Many systems can detect highway lane markings, but are unable to perceive rural roads that do not have these features. Some systems rely on detecting the road edges, but will often fail if these edges are degraded and broken. Some systems rely on having a map to describe the shape of the road and to predict when and where new roads will appear. This specialization is due to the real-time nature of the task where often more general capabilities are traded for vehicle speed. In general, road detection systems are specialized for specific scenarios.

This research focuses on detecting the most difficult road scenario without giving up the capability of perceiving less complex situations. In particular, we are able to navigate a real robot vehicle on unstructured roads which may have:

- no lane or edge lines painted on the road surface,
- degraded road edges,
- road surface scars,
- strong shadow conditions,
- no map information.

These roads lack a clearly defined characteristic which can be used for their detection such as road edge lines. This makes these roads the most difficult to detect.

Specifically, this paper describes an algorithm called UNSCARF (UNSupervised Clustering Applied to Road Following) for detecting unstructured roads. We first highlight other road detection systems and compare their approaches in Section 2. We then present the details of the UNSCARF algorithm in Section 3. In Section 4, we show results of this algorithm running on several unstructured road sequences.

2. Overview of Road Detection Systems

Road detection systems localize roads in camera images using a variety of methods. In this section, we will discuss several systems that have been designed to navigate unstructured roads. We begin with systems designed for structured roads and progress to the detection of unstructured roads.

Many methods rely on the presence of specific structured features of roads such as lane markings and road edge lines. These include the General Motors Lanelok System [1], the VaMoRs System [2], and the YARF (Yet Another Road Follower) System [3]. These systems usually have very fast processing speeds and are very well suited for navigation on structured roads. However, these systems are not directly applicable to unstructured roads since they rely on road edge lines.

Some systems extended the detection of road markings to the detection of road edges [4,5], typically using gradient operators [6]. These systems detect road edges quickly, but only under good conditions. Gradient operators are high pass filters, and as such, tend to be noisy. These operators also have a stronger response to shadow edges than to road edges. Therefore, these systems have difficulties in the presence of degraded road edges, high texture, and shadows.

To avoid the difficulties with noisy or cluttered road edges, other systems detect road surfaces. Several systems use a histogram and threshold approach [7,8]. These systems label pixels in the image as "road" or "off-road" by first histogramming a one-dimensional color feature of the image, and then thresholding to separate groups of pixels. Pixels within certain ranges of values are labeled as "road" while the others are labeled as "off-road". When shadows are present, shaded road and shaded off-road often have very similar features, more so than sunlit road and shaded road. This approach works well for many unstructured roads, but has difficulties with degraded surfaces, leaves, and shadows.

Other systems extend the idea of histogram and threshold approaches by using a multi-dimensional classification approach [8,9]. This approach is based on modeling the colors of the road, and comparing the color of each pixel with each color model. Road colors are sampled in an input image region that lies well within the predicted road region. This way the systems are fairly sure that a pure road sample is attained. The pixels are labeled as "road" or "off-road" depending on how well they match the road colors. These systems work well in many difficult unstructured road scenarios, but they rely on tracking results from image to image.

UNSCARF is a classification-based system that can detect difficult roads without tracking results from image to image. In each new image, it clusters pixels having similar colors into regions. Then it selects the a set of regions whose combined shape forms the shape expected for a road surface. In this way, it no longer relies on tracking color information from image to image.

3. UNSCARF

UNSCARF collects similar pixels in the input color image using a modified clustering technique. The edges between groups of pixels are matched to road edges to determine the location of the road in the image. As outlined in Figure 1, the pre-processing first reduces the size of the input color images and appends the images with the spatial location of each pixel forming the feature images. A clustering module collects the pixels in the feature images into groups and then extracts the edges between groups of pixels forming a class edge image. The interpretation generation selects a road interpretation whose road edges best match the edges in the class edge image.

3.1 Feature Identification

UNSCARF feature extraction first reduces the size of the input color images to reduce the amount of data to be processed, thereby reducing the time required to process the input images. The reduction first averages the input image to reduce the noise and then subsamples to the desired size. This results in a reduced color image. It then appends the (row, column) location to each color pixel in the reduced image, forming a five-dimensional feature image where each pixel $\mathbf{x}$ is:

$$\mathbf{x} = [\text{red, green, blue, row, column}]^T$$

An alternative color representation could use intensity, hue, and saturation images rather than the red, green, and blue color images. However computing these images is both costly in computation time and also susceptible to singularities in the transformations. Therefore, we decided to use the input red, green, and blue bands of the color image.

3.2 Unsupervised Clustering

This module groups pixels in the feature images into homogeneous regions using a modified ISODATA clustering algorithm [10]. The clustering algorithm first assigns all of the pixels in the image to arbitrary classes. The color model consisting of a mean feature vector $\mathbf{m}_i$ and a covariance matrix $\mathbf{C}_i$ is computed for each class, ω_i. All the covariance matrices are normalized so that they have a unity determinant. This is accomplished by dividing each element of the covariance matrix by the fifth root of the determinant (since the matrices are 5x5):

$$\mathbf{C'}_i = \mathbf{C}_i / \| \mathbf{C}_i \|$$

Next the pixels are re-classified by nearest Mahalanobis distance using the most recent mean and normalized covariance matrices:

$$\mathbf{x} \in \omega_i \quad \text{if } \alpha_i \leq \alpha_j \;\; i \neq j \text{ where}$$

$$\alpha_i = (\mathbf{x}-\mathbf{m}_i)^T \, \mathbf{C'}_i^{-1} \, (\mathbf{x}-\mathbf{m}_i)$$

These steps are repeated until only a few pixels change class labels in the re-classification step. An example of the clustering algorithm is shown in Figure 2. The input feature image is shown first, at the top left. The middle pair of images in Figure 2 depicts the initial random classification. The right image of this pair shows arbitrary intensities for each class and the left shows the current mean value for the classes. The next lower pair of images in the figure shows the classification after the color model is computed using the initial arbitrary classification. Figure 2 also shows subsequent steps of the clustering process and the bottom pair of images of this figure show the resulting class labeling of the original image. At the end of the clustering, each pixel in the image is labeled with a class, forming a class image.

3.3 Class Edge Detection

This module takes the class image from the unsupervised classification module and cleans up the clustering result. Since the clustering can result in noisy class images (i. e. class images with many small class label groups) the small regions are removed from the class image using a shrink and expand` algorithm [6]. The edges between class labels are then detected in the class image, and collected in a class edge image. An example of a class edge image is shown in Figure 1.

3.4 Road Model Matching

The resulting road interpretation is the one whose road edges best match the class edges between pixels having different class labels. Each candidate interpretation is compared with the class edge image by stepping along each edge of the candidate and determining the distance to the nearest class edge. This distance is accumulated along the length of the candidate road edges, and the average distance between the interpretation and the class edges is determined by dividing the accumulated distance with the length of the edges. The best interpretation is the one whose average distance between its road edges and the class edges is the smallest.

To avoid searching for the nearest class edge for each location along a candidate road edge, a new image is generated whose pixels record the distance to the closest class edge. This is done by applying a Grassfire transformation [11] to the class edges, forming a distance image. The total distance is computed by summing values in the distance image lying on the road edge lines. The average distance is again determined by dividing the accumulated distance by the length of the edge lines.

UNSCARF picks the road interpretation which has the minimum total distance between the candidate road edges and the edges in the class image. Figure 3 shows the results and the edges between the class labels on a sequence of unstructured road images. In these cases, the road edges are visible in the class image. In all four of these examples, UNSCARF finds a good straight road approximation to the road location.

5. Results and Conclusions

We have run UNSCARF on a variety of example images using six color classes. UNSCARF used approximately 8 steps for clustering in the examples presented here, requiring about 1 minute on a single Sun 4 computer. We have implementations of UNSCARF that divide the image over multiple Sun computers for the clustering. We have seen proportional decreases in computation time with the number of computers used.

Figure 4 shows UNSCARF results on a sequence of darkly shadowed images. UNSCARF correctly identifies the road in all but the second image. In the this image, the road is still located well enough to correctly steer the vehicle on the road.

An example of UNSCARF processing a sequence of unstructured roads surrounded by dirt and snow is shown in Figure 5. Again, UNSCARF finds a good straight road approximation to the roads in the image in most of these cases. In the examples that are not exact, the results are good enough to steer the vehicle to stay on the road.

UNSCARF is insensitive to changes in illumination since the colors are clustered independently from image to image. It allows multiple colors to represent the road surface colors, and does not rely on having a previous or predicted road description input from the navigation system.

UNSCARF has been integrated into a navigation system that has driven an actual robot vehicle, the Carnegie Mellon Navlab [3], on an unstructured road in various conditions. UNSCARF has detected unstructured roads, including degraded, dirt, and paved roads.

The first strength of UNSCARF comes from representing multiple color classes as Gaussian distributions in full RGB color. Having multiple classes allows UNSCARF to represent different colors of the road (e.g. asphalt, wet patches, shadowed pavement, and leaves) and off-road objects (trees, sunlit grass, shaded grass, dirt, and leaves). Using full color instead of monochrome images (or some combination of colors) keeps all of the image information that may be useful in discrimination. The Gaussian representation of each color represents whether a particular variation in color is significant. Sunlit asphalt tends to be homogeneous in color and is represented by a color model with a small variance; grass has more variety, and is represented with correspondingly larger variances.

Acknowledgements

We wish to thank Dr. Takeo Kanade at Carnegie Mellon University who supervised this research. I would also like to thank the numerous members of the Navlab and related projects at Carnegie Mellon, without whom this work would not have been possible.

This research was supported by the DARPA Road Following contracts DACA 76-89-C-0014 and DACA 76-85-C-0003.

References

[1] S. Kenue, "Lanelok: Detection of Lane Boundaries and Vehicle Tracking using Image-Processing Techniques.", *Mobile Robots IV*, Proc. SPIE, Vol. 1195 , Nov. 1989.

[2] B. Mysliwetz and E. Dickmanns, "Distributed Scene Analysis for Autonomous Road Vehicle Guidance", *Mobile Robots II*, Proc. SPIE, Nov.1987, Cambridge MA, pp. 72-79.

[3] C. Thorpe, (editor), *Vision and Navigation: The Carnegie Mellon Navlab*, Kluwer Academic Publishers, 1990.

[4] S. Liou and R. Jain, "Road Following Using Vanishing Points", *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Miami Beach, FL, June 1986, pp. 41-46.

[5] A. Waxman, J. LeMoigne, L. Davis, B. Srinivasan, T. Kushner, E. Liang, and T. Siddalingaiah, "A Visual Navigation System for Autonomous Land Vehicles", *IEEE J. Robotics and Automation*, vol. RA-3, no.2 April 1987, pp.124-141.

[6] D. Ballard and C. Brown, *Computer Vision*, Prentice-Hall Inc. 1982.

[7] D. Kuan, G. Phipps, and A. Hsueh, "Autonomous Land Vehicle Road Following", *Proc. First Inter. Conf. Computer Vision*, June 1987, London England, pp. 557-566.

[8] M. Turk, D. Morgenthaler, K. Gremban, and M. Marra, "VITS: A Vision System for Autonomous Land Vehicle Navigation", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-10, no 3, May 1988, pp. 342-361.

[9] J. Crisman, "Color Vision for the Detection of Unstructured Road and Intersections", Ph.D. Thesis, Dept. of Electrical and Computer Engineering, Carnegie Mellon University, May, 1990.

[10] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, Inc., 1973.

[11] K. Castleman, *Digital Image Processing*, A. Oppenheim, editor, Prentice-Hall, Inc., 1979.

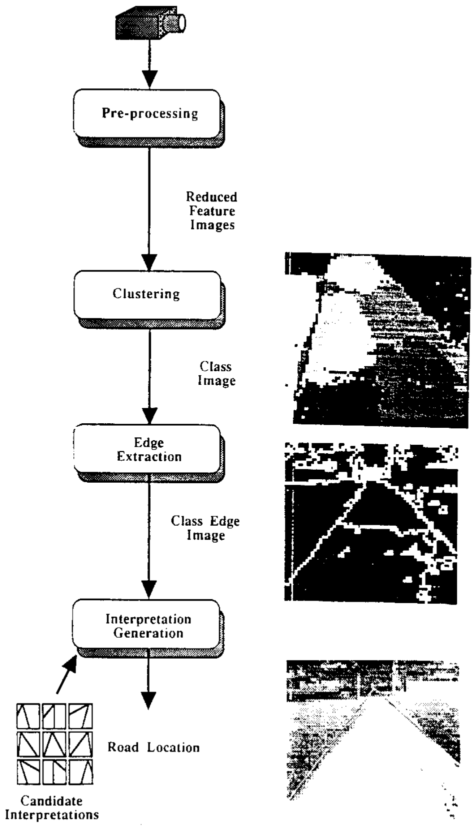

Figure 1: UNSCARF Block Diagram

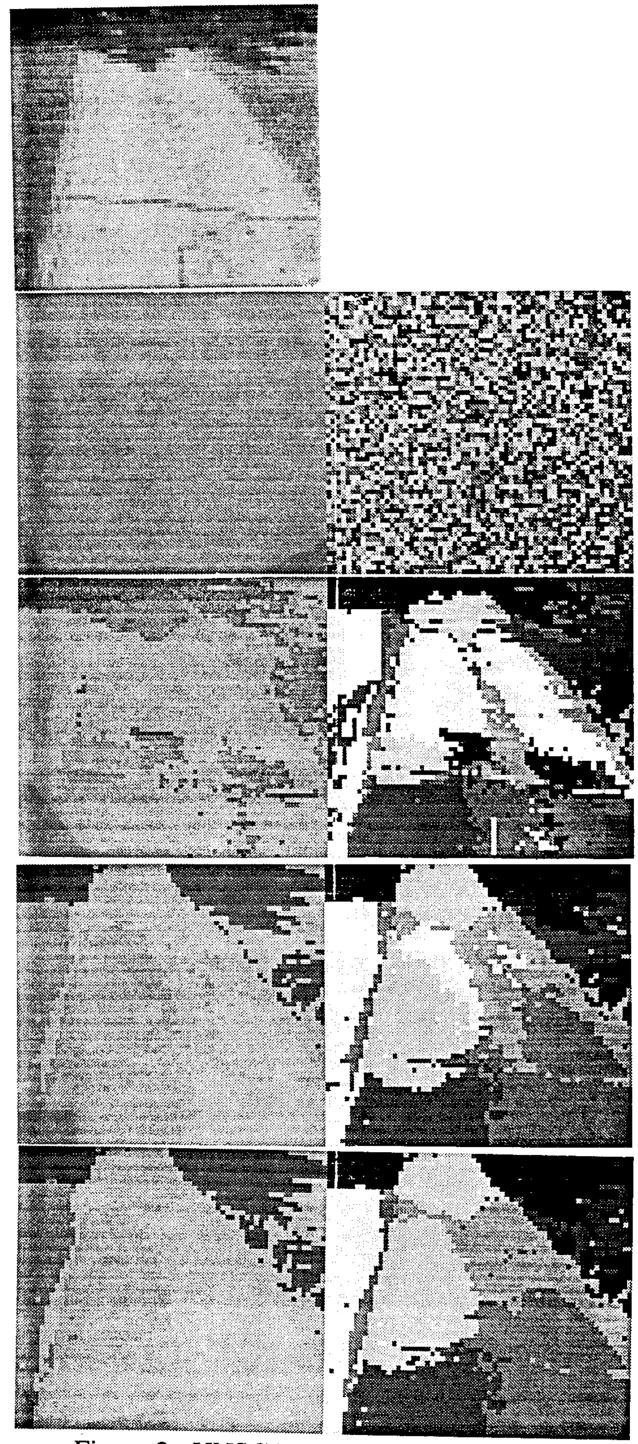

Figure 2: UNSCARF Clustering Example

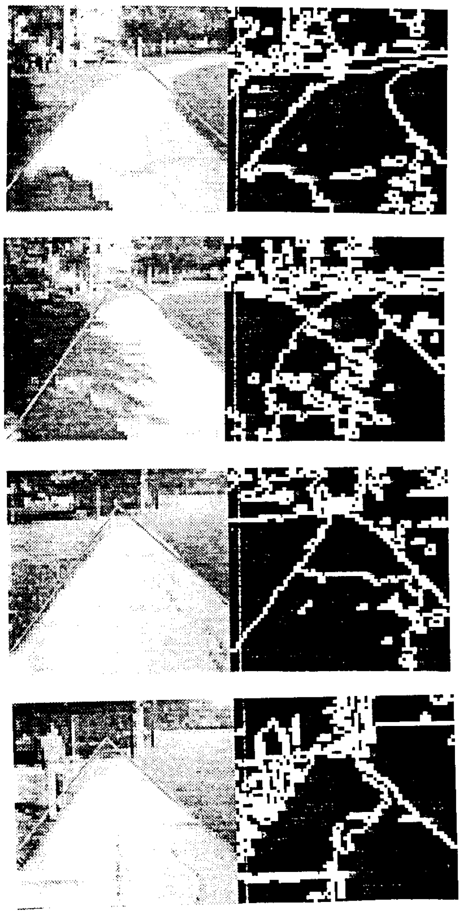

Figure 3: UNSCARF Results on an Unstructured Road: UNSCARF finds a good straight road approximation to the roads in the image. In these cases, the road edges are visible in the class edge image.

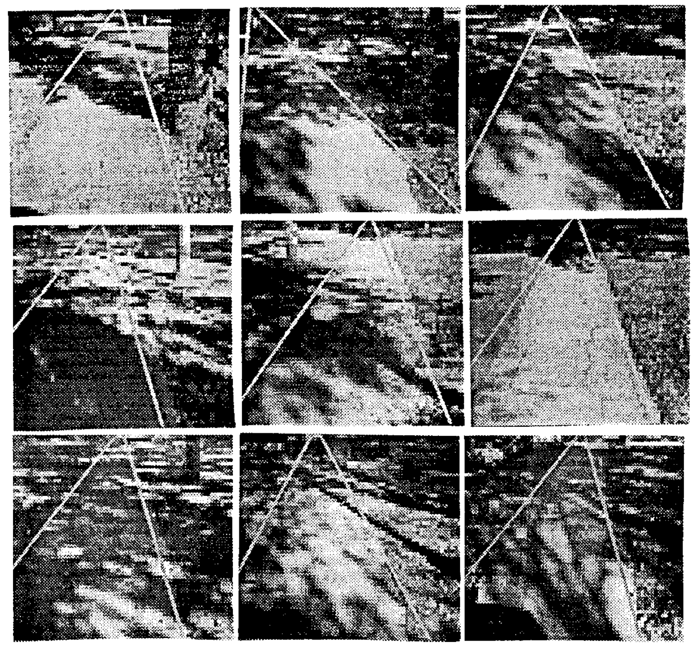

Figure 4: UNSCARF Results in Dark Shadows: UNSCARF correctly locates the road in all by the second of these images. The result, however, is adequate to steer the vehicle on the road until the next image is processed.

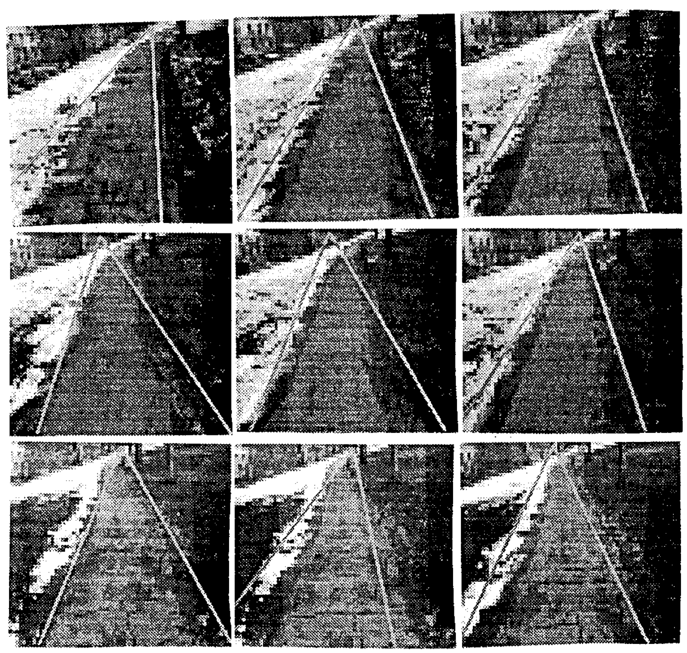

Figure 5: UNSCARF Results with Dirt and Snow: UNSCARF Adequately detects the road when the roads are surrounded by dirt and snow.