Survey: Vision based Road Detection Techniques

Vipul H. Mistry\(^{\#1}\), Dr. Ramji Makwana\(^{\ast2}\)

\(^{\#1}\)Ph.D. Student, Electronics & Communication Engineering Department, Uka Tarsadia University Maliba Campus, Gopal Vidyanagar, Dist: Surat, Tarsadi- 394 350, Gujarat, India

\(^{\ast2}\)Head of Department, Computer Engineering Department A.D.I.T., Vallabh Vidyanagar, Anand, Gujarat, India

Abstract—Computer Vision based autonomous robot navigation has caught the attention of many researchers because of its huge potential in various industrial and social applications. This paper focuses on various approaches used to detect road region which is fundamental requirement in applications like intelligent vehicles, lane detection and tracking and driver assistance systems. Broadly speaking techniques proposed for road detection are classified in three categories: activity driven, feature driven and model driven. This paper surveys development of last decade in the field of vision based road region detection. Two major components discussed in this paper are methods for structured and unstructured road detection.

Keywords—Computer vision, autonomous navigation, structured roads, unstructured roads.

I. INTRODUCTION

Ability to navigate is very important feature for a fully autonomous mobile robotic system. To ensure this it becomes very critical to recognize road region and stay on the road region while navigating from source to destination. With proximity sensors one can identify the ground plane but it becomes near to impossible to recognize shape and various properties of the ground. Considering these facts, today’s research focuses more on computer vision based navigation approaches. With development in automobile technology and road transportation system accidents are also increasing. It has become essential to develop systems that can assist driver while navigating on the road. Road identification system can immensely help the purpose.

Looking at other perspective, in today’s world the role of robots in our everyday activities is increasing. But the successful completion of a task by robot is highly dependent on how effectively and accurately it can navigate between source and destination. Today the trend is to make the robot understand the environment along with the navigation. That’s why vision based navigation is must. Autonomous road detection is challenging due to range of environmental conditions under which these systems operate: rain, shadow, sunshine, day, night, fog etc. Broadly these techniques can be divided into three categories: activity driven [1, 2], feature driven [3, 4] and model driven [5]. Activity driven approaches benefit from relatively high vehicle motion activity along the roads. These approaches mainly work by extracting an activity map which divides the image into active and non-active regions. Feature driven approach relies upon extraction of image features to detect lane and boundaries. Model driven approaches aim to match a road model to the image. In this paper, we have restricted ourselves to feature driven and road model based approaches only. Broad classification of various feature and model driven approaches is shown in Fig. 1.

Remaining of the paper is organized as follows. Sections 2-7 highlight various vision based road detection methods. Section 8 discusses performance parameters to be used for proper evaluation of road detection approaches. The last section gives outline for the further research work in the field.

II. TEXTURE BASED ROAD DETECTION METHOD

Texture based road detection methods focus on identification of textural differences between road and non-road regions to segregate between them. Mainly texture based methods uses boundary extraction approach using various texture properties of an image and identifies road region between two boundaries. Various articles [6-9] discuss the task of road border extraction using different methods like optical flow measurements, edge detection followed by Hough Transform, segmentation based on color content, color feature combined with the line and road following, texture description alone and texture combined with color features. Authors in [10-11] have discussed various real time constraints like how to reduce the time recognizing the texture as one out of predefined set, how to make acquisition over multiple windows of different size and how to find the subset of relevant features. In 2012, Stevica Graovac Et. Al. [12] took this work further and presented the use of texture descriptors for detection of road regions. There are mainly two types of descriptors used in
IV. ROAD DETECTION BASED ON ILLUMINATION INVARIANCE

The key for vision based road detection is the ability to classify image pixels as belonging or not to the road surface. Identifying road pixels is a major challenge due to the variability caused by lighting conditions. A practically difficult scenario appears when the road surface consists of both shadowed and non-shadowed areas. In 2011, Jose M. Alvarez Et. Al. [18] suggested a novel approach for road detection combining shadow invariant feature space and model based classifier. This method uses both features as well as model based approach for road detection. Here the model is built online to improve the adaptability of the algorithm to the current lighting and the presence of other vehicles in the scene. Work proposed by authors in papers [19-21] reveals that texture and color are potential features to characterize roads. To increase the robustness of the algorithm with variations in imaged road textures use of depth as an additional cue is also proposed, where depth comes from active sensors. Considering this work Ten Et. Al. [19] proposed a method using texture and color features along with histograms in the RG color space to build a model for road and background. Road variability is modeled using various histograms in this method, which are updated frame by frame. Jose M. Alvarez Et. Al. [18] used illumination invariant image introduced by Finlayson Et. A. [22] as a feature space and presented a robust road detection method that can work in the presence of shadows. This method does not depend on either road shape or temporal restrictions. Some of the key steps of this algorithm [18] are as follows:

• Compute I from IRGB.
• Build road model using normalized histogram of the surrounding area of a set N seeds placed at bottom of the image.
• Obtain IROAD by thresholding I according to road model and fixed threshold.
• Perform connected component procedures using same N seed points.
• Fill small holes in the resultant road image.

Considering the Lane Departure and Forward Collision Warming system application Dezhi Gao Et. Al [23] proposed a practical method for road detection which is adaptable to the change in illumination and road conditions. In this approach, sobel operator convolution kernel is introduced in the image preprocessing to enhance the robustness of the algorithm. In order to be feasible against change in illumination and road types adaptive threshold method is used for segmentation and finally Hough Transform and SUSAN algorithm is used to extract lane marks and front vehicles. This method uses lane information to identify road region and tracks lane features using kalman filter for road detection.

V. UNSTRUCTURED ROAD DETECTION PROBLEM

Road detection is a key requirement for unmanned guided vehicles. A lot of approaches for paved road following have been proposed in the last two decades [24-26]. Although this topic has already been documented in technical literature by different research groups,
unstructured road detection poses several interesting and new challenges due to its unstructured nature. [23]. For unstructured roads road boundaries may be unclear and it may have a low intensity contrast. Additionally, the overall road shape may be arbitrary, which leads to a road surface with a degraded appearance. Varying illumination conditions, different viewpoints and changing weather conditions make the problem more complex.

In last sections, we have discussed various features that can be used to extract road region from a given image. Recently, feature combination methods have been widely researched. Three features of a lane boundary starting position, direction and gray level intensity feature comprise a lane vector via simple image processing [27]. In order to improve the effectiveness of the algorithm contextual information is also used in addition to low level feature cues in road detection [28]. How to generate this contextual information and how much information is required was still a matter of concern. In 2012, Erke Shang Et. Al. [29] proposed a novel approach for unstructured road detection on contextual information. A typical low level feature named histogram of RGB value is used. This method combines SVM, KNN and a Bayesian framework using Confidence map generated with the help of GPS data. This method outperforms state of art methods using the same low level feature space.

Given a single image of an arbitrary road, that may not be well paved or not have clear edges or even no prior known color or texture distribution, it is very difficult task for computer to identify road region from it. All the methods discussed so far either solely depends on road features extracted from road boundary or lane markings or combined road region properties with road models. Most of the methods work well for well structured roads but fail in unstructured environments. Apart from region properties some researchers have also worked on global features like vanishing point. Vanishing point based road detection method is discussed in next section.

VI. VANISHING POINT BASED ROAD DETECTION

In previous sections we discussed road detection methods which can be grouped in three categories: edge, region and texture based methods. Edge detection based methods reduce the road detection task to boundary detection or lane marking extraction. These approaches are more appropriate for structured roads, where well painted lane marking or strong edge boundaries are the distinct features. Color cue [30, 31], Hough Transform [32] have been used to find road boundaries. Region based methods assume that road surfaces belong to relatively homogeneous regions such as in well paved roads. Instead of looking for locally distinctive road cues, texture based methods search for global constraints to distinguish road direction [33, 34]. Texture based methods search for local oriented textures and then make them vote for the locations of the road’s vanishing points. Estimating vanishing point plays a pivotal role in the determination of the direction of the road. In 2010, Hui Kong Et. Al. [35] came up with a general road detection method from a single image using vanishing point detection approach. The basic steps involved in vanishing point detection based methods are as follows:

- A texture orientation estimation at each pixel for which confidence level is provided.
- Voting scheme takes into account this confidence level and distance from the voting pixel to the vanishing point candidate.
- Finding road boundaries from vanishing point.

In most of the vanishing point detection methods texture orientation estimation relies on Gabor Filters since they are known to be accurate. The orientation $\phi$ and a scale $\omega$, the gabor wavelets are defined by [36]

$$\psi_{\omega,\phi}(x,y) = \frac{\omega}{\sqrt{2\pi}\omega} e^{-\frac{a^2}{2\omega^2}} e^{-(y\sin\phi + x\cos\phi)^2/c^2}$$

where, $a = x\cos\phi + y\sin\phi, b = -x\sin\phi + y\cos\phi, c = 2.2$. Using Gabor kernel the texture orientation at each pixel is computed. One can make these pixels vote to obtain the vanishing point. Precisely, a pixel $P$ for which the texture orientation is the vector $\vec{O}_P$ can vote for all pixels $V$ above $P$. To overcome this problem authors in paper [35] proposed a soft voting scheme, where the voting score received by a vanishing point candidate from a vector is a value taking into account the distance between the vanishing point and the voter. The correctly detected vanishing point provides a strong cue to the localization of the road region. Based upon the two dominant edges, one can roughly segment the road area and update the vanishing point estimated by above mentioned procedure. In paper [37] a similar straight road segmentation method is given to detect both road borders simultaneously. It is achieved by optimizing a criterion, which is the difference between the average values of some characteristics within the image road region and that characteristics in the region outside the road region. But this method fails when the difference between road and off-road region is very less. Hui Kong Et. Al. [35] proposed road segmentation strategy to find the two most dominant edges by initially locating the first one and the other based upon the first one. Because they utilize both texture and color cues, the proposed method by them exhibits good merits in handling very general road detection tasks.

In order to achieve precise orientation estimation, one needs to apply a large number of orientation filters in all possible directions from $0^\circ$ to $180^\circ$. Designing and applying a bank of differently rotated filters is computationally expensive. To address this problem Freeman and Adelson [38] proposed a steerable filter in which each arbitrary oriented filter can be formed by linear combination of fixed set of basis oriented filters. But still it requires steering the basis oriented filters in all directions with a precise angle step of $1^\circ$. In 2012, Peyman Moghadam Et. Al. [39] proposed a novel methodology based on image texture analysis for the fast estimation of the vanishing point detection in the challenging environments. The key attributes of the methodology consists of the optimal local dominant orientation method that uses joint activity of four
gabor filters, estimate the local dominant orientation at each pixel in the image plane, weighting of each pixel based on its dominant orientation and an adaptive distance based voting scheme for the estimation of the vanishing point. They proposed a novel solution to directly estimate the local dominant orientation based on the joint activity of four gabor filters with only 4 orientations (0°, 45°, 90°, 135°) followed by an efficient and robust voting scheme suitable for real time application.

VII. SUMMARY OF VARIOUS ROAD DETECTION METHODS

In previous sections various types of road detection methods were discussed. Broadly these methods are feature driven or combined road model and feature driven methods. Table 1 shows the comparison between various popular road detection approaches proposed so far and discussed in last sections. All the methods are separated based on types of road structures, feature extraction process, post processing and tracking, evaluation method and remarks.

<table>
<thead>
<tr>
<th>Ref. Paper</th>
<th>Road Model Used</th>
<th>Feature Extraction Method</th>
<th>Post Processing/ Tracking</th>
<th>Evaluation</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>[12]</td>
<td>Structured Roads</td>
<td>Texture Descriptors based features</td>
<td>No Tracking</td>
<td>Single image Qualitative analysis with ground Truth Images</td>
<td>Semi Automatic, Not good with shadows, more descriptors selection is a scope of research</td>
</tr>
<tr>
<td>[13]</td>
<td>Structured Roads</td>
<td>Similarity Search (Region based features)</td>
<td>Knowledge based database is generated, no tracking</td>
<td>Frame by frame comparison with ground truth images</td>
<td>Semi automatic, efficient training is required</td>
</tr>
<tr>
<td>[17]</td>
<td>Structured roads with similar off-road regions</td>
<td>Edge detection based approach</td>
<td>No tracking</td>
<td>Still Image</td>
<td>Automatic</td>
</tr>
<tr>
<td>[18]</td>
<td>Structured roads with shadows</td>
<td>Illumination invariant image space &amp; road model based approach</td>
<td>No tracking</td>
<td>Sequence of images , Qualitative and Quantitative Analysis</td>
<td>Better performance compared to HSI based color space methods, suitable I in Shadowed environment</td>
</tr>
<tr>
<td>[29]</td>
<td>Unstructured Roads</td>
<td>Contextual Information , RGB descriptor based Bayesian framework for confidence map generation</td>
<td>No tracking</td>
<td>SVM, KNN based offline training and testing on images</td>
<td>Use of confidence map to increase the performance of SVM, KNN</td>
</tr>
<tr>
<td>[23]</td>
<td>Structured roads with lane marking</td>
<td>SUSAN Algorithm, Hough Transform (edge detection) adaptive Threshold</td>
<td>Inverse Projection , Tracking using Kalman Filter</td>
<td>Qualitative Evaluation</td>
<td>Lane tracking and vehicle detection application</td>
</tr>
<tr>
<td>[35]</td>
<td>Structured and Unstructured roads</td>
<td>Vanishing point based boundary extraction and region extraction</td>
<td>Region based segmentation , no tracking</td>
<td>Still images</td>
<td>36 Orientation with 5 scale at each pixels using Gabor Filters</td>
</tr>
<tr>
<td>[30]</td>
<td>Structured roads with lane marking</td>
<td>HSI color model based feature extraction, Fuzzy C means algorithm based segmentation</td>
<td>Filtering after segmentation, no tracking</td>
<td>Still images</td>
<td>Provides better results compared to RGB model based method for lane extraction</td>
</tr>
<tr>
<td>[31]</td>
<td>Structured road with lane markings</td>
<td>Color based segmentation for lane extraction</td>
<td>Least square method for extension of lanes, no tracking</td>
<td>Still images</td>
<td>Scope of research work in presence of light reflections, road signs</td>
</tr>
<tr>
<td>[40]</td>
<td>Structured roads with shadows</td>
<td>HSV color space based method</td>
<td>Morphological post processing , no tracking, offline</td>
<td>SVM based road region classification, Qualitative Analysis</td>
<td>Shadow Removal application</td>
</tr>
</tbody>
</table>
VIII. PERFORMANCE EVALUATION OF VISION BASED ROAD DETECTION METHODS

In this section, various qualitative and quantitative measures are discussed which are used to evaluate the performance of various road detection methods. Without proper evaluation scheme it is not possible to justify the effectiveness of proper method for road detection. Considering the application of autonomous navigation on road it is necessary that a method of road region detection is fully autonomous, i.e. there should not be any human interaction for extraction of features. As mentioned in the previous table region based approach proposed in 2012 by Stevica Graivac Et. Al. [12] outperforms previous similar type of texture descriptor based methods, but it is not fully automatic. It works on the basic assumption that the whole image generally consists of three distinguishable regions: road, local (near) background and far background. Human operator is required to select a window of 64x64 from all the three regions. Various qualitative experiments were conducted to evaluate the performance of this algorithm, where results were compared visually after extraction of road regions. The algorithm was tested for structured roads with lane markings and urban roads under various constraints like shadows produced by trees, typical sign marking the road tracks, poor separability of the furthest part of the road etc. Using the positives from this approach a more robust and fully autonomous method can be suggested. No of texture descriptors used in this approach can be still a subject of research. Adding color based features can improve the performance of the algorithm in segmentation stage.

In 2011, Roman Stoklasa Et. Al. [13] suggested a fully autonomous road detection method based on similarity search. It is a content based classification approach where various visual descriptors like edge histogram, color layout descriptors are used to describe image characteristics in a form of vectors. Many quantitative experiments were conducted to measure the performance of the algorithm used. The extracted regions were compared with ground truth regions in each frame of a video sequence. For qualitative analysis, two measures were used: absolute amount of intensity under the mask $S_a$ and relative amount of intensity under the mask $S_r$ [13].

$$S_a = \sum_{p \in \Omega} \min (GT(p), I(p))$$

$$S_r = \frac{\sum_{p \in \Omega} \min (GT(p), I(p))}{\sum_{p \in \Omega} GT(p)}$$

Where, $GT(p)$ : ground truth point, $I(p)$ : Image Point

Several error metrics were also defined this paper which are used as part of quantitative analysis described as follows.

- **FP (False Positive)**: quantifies the proportion of pixels classified as road within non road regions defined as $FP(I) = S_r(I_r, GT)$
- **FN (False Negative)**: quantifies the proportion of pixels classified as non-road within road regions defined as $FN(I) = S_r(I_o, GT)$
- **NP (Non Positive)**: quantifies the proportion of pixels not classified as road within road regions.
- **NN (Non Negative)**: quantifies the proportion of pixels not classified as non-road within non road regions.
- **PA (Positive Accuracy)**: quantifies the proportion of pixels that were correctly classified as road regions.
- **NA (Negative Accuracy)**: quantifies the proportion of pixels that were correctly classified as non road regions.

It was seen that the method [13] was able to detect more than 85% of road area in the input images. Although results show that this method works well but it is not preferable for general autonomous navigation on road as it requires an Adhoc Map to navigate. The quantitative results show that this approach may face problem in tracking road when at edge of the road.

In 2010, silent pixel minimization algorithm was proposed by Falola Et. Al. [17] to increase the performance of region based road detection approaches. One important quantitative measure confusion matrix was used to evaluate the performance of the algorithm. The confusion matrix as shown in Table 2 is a validating tool which contains information about actual and predicted classifications done by a classification system.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Parameter X</th>
<th>Parameter Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Parameter X</td>
<td>Parameter Y</td>
</tr>
<tr>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

The performance accuracy of the minimized salient pixels in a confusion matrix is the proportion of the total no of predictions that were correct. This is expressed as follows.

$$Accuracy(AC) = \frac{\sum (A+C)}{\sum (A+B+C+D)}$$

(4) The quantitative results showed that the proposed scheme work with average 88% of accuracy. But special cases such as influence of environmental factors such as rain on road frames were not investigated here.

Illumination variations affect the performance of various road detection algorithms. In 2011, Jose M. Alvarez Et. Al. [18] proposed an illumination invariant method for road detection. It outperforms the performance of HSI based algorithms. It works by combining the illuminant invariant feature space and the likelihood based classifier. The qualitative evaluation was done on test images by manually segmenting all the images in test database to generate ground truths and later on comparing the results with the ground truths. In this paper quantitative evaluations were done against ground truth using three pixelwise measures: precision, recall and effectiveness as described below [18].
precision($P$) = $\frac{\sum G \cdot I}{\sum I}$

(5)

call($R$) = $\frac{\sum G \cdot I}{\sum G}$

(6)

effectiveness($F$) = $\frac{2PR}{P + R}$

(7)

All three measures range from 0 to 1. Precision and recall provide different insights in the performance of the method: low precision means that many background pixels are classified as road, whereas low recall indicates failure to detect the road surface. Finally effectiveness is the tradeoff using weighted harmonic mean between precision and recall. The results presented in this paper clearly indicates that the proposed method outperforms the HSI based algorithms for images containing strong shadows and other vehicles in the scene. This algorithm is able to run in real time in a per frame basis, it is not constrained to specific road shapes and modeling of the background is also not required. This method provides good results in the presence of shadowing effect. Further this work can be extended for unstructured roads as well.

Unstructured road detection is a challenging problem due to its unstructured nature. In 2012, Erke Shang et. al. [29] proposed a novel method of combining SVM, KNN and confidence map under the Bayesian framework to improve the overall performance of the unstructured road detection. This is a novel approach combining contextual information with confidence map obtained using GPS and combined SVM, KNN classifier to detect the road region in front of the vehicle. It uses histogram of RGB as descriptor for the road region detection feature. Performance of the proposed method was evaluated with various quantitative measures as described in Table 3. The results mentioned in the paper [29] clearly indicates the effectiveness in the detection rate after incorporating confidence map along with SVM and KNN.

### TABLE IIIII

<table>
<thead>
<tr>
<th>Contingency Table</th>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non Road</td>
<td>TN</td>
</tr>
<tr>
<td>Road</td>
<td>FP</td>
</tr>
</tbody>
</table>

**TABLE IIIII RULES OF ROAD DETECTION [29]**

<table>
<thead>
<tr>
<th>Pixelwise Measure</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quality</td>
<td>$G = \frac{TP}{TP + FP + FN}$</td>
</tr>
<tr>
<td>Detection Rate</td>
<td>$DR = \frac{TP}{TP + FP}$</td>
</tr>
<tr>
<td>Detection Accuracy</td>
<td>$DA = \frac{TP}{TP + FN}$</td>
</tr>
</tbody>
</table>

Apart from traditional region based road detection methods, researchers have also proposed vanishing point based boundary extraction and then estimation of road region based on these road boundaries. These approaches are greatly effective when dealing with unstructured road roads with similar road and non road regions. In 2010, Hui Kong Et. Al. [35] proposed a general road detection method based on vanishing point detection approach. It is a texture orientation based approach to find out vanishing point in a given image that is further used to extract boundaries and extract road regions in a given image. Vanishing point detection itself is a computationally costly task. To select minimum no of orientations at each pixel and optimization of the voting scheme is a matter of research in identification of vanishing point. Researchers [33, 34] have suggested novel approaches towards real time identification of vanishing point from a given image. Finally after detection of road boundaries and detection of road regions authors [35] are using same quantitative measures R “recall” and P “precision” as discussed in this section for performance evaluation. Overall a novel locally adaptive soft voting scheme which can reduce the computational cost of vanishing point detection should be an area of research for further work in this approach.

**IX. CONCLUSION AND SCOPE OF FURTHER RESEARCH WORK**

After studying various methods that have been proposed so far for road detection, we can conclude that mainly there are two type of approaches used so far: region features and texture features. As mentioned in section VII most of the region based methods work well for structured road types. Road color properties, lane markings etc are used to differentiate road regions from off road regions. The effectiveness of such algorithms has been improved by various segmentation algorithms and improvements in edge detection algorithms. But these approaches are not promising same results if roads are unstructured. To overcome these problems texture properties based approaches have been proposed. Use of global information like vanishing point can be very useful to extract the boundaries of the roads even in unstructured environments. The main problem to deal with is computational cost of various orientation filters used to extract the texture properties.

Researchers have also highlighted the problem of robust road detection under various environmental and climate changes. As mentioned in section VII illumination invariance based approaches and HSV color model based approaches works well in various illumination conditions. Road region recognition is a main feature that is gaining increasing attention from intellectuals because it helps autonomous vehicle to achieve successful navigation without accident. Different techniques based on computer vision have been used by various researchers and outstanding results have been achieved. Despite their success, there is a huge scope of research work which can further enhance the effectiveness of road detection algorithm and make it more general for all the type of roads as well as more robust to withstand various environmental variations and irregularities in road shapes.

Based on the literature reviewed so far, there is a scope of road detection method that will have all the characteristics mentioned below:

- Road detection algorithm should be completely autonomous without any sort of human interface.
- It should work using global features.
- Detection method should be fast enough to meet real time criteria.
• Algorithm should be general for all the structured and unstructured type of roads.
• Algorithm should be robust.

To meet the requirements of general and robust autonomous road detection we should combine region based features and global orientation based feature like vanishing point. Fast vanishing point detection and selection of unique road region features are two major challenges for the further research work in the field of computer vision based autonomous road detection.

REFERENCES


