The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Lane Detection and Tracking in Challenging Environments based on a Weighted Graph and Integrated Cues

Chunzhao Guo, *Member, IEEE,* Seiichi Mita, *Member, IEEE*, and David McAllester, *Member, IEEE*

***Abstract*—Real-time lane detection and localization is one of the key issues for many intelligent transportation systems. In this paper, we present a lane detection and tracking approach designed to work in challenging environments where lane boundaries may be low-contrast and changeful with noise due to a number of factors such as wear, type, lighting and weather conditions, etc. In the method, a sophisticated cascade lane feature detector is applied to cope with challenging environments at the very beginning of the detection and a weighted graph is subsequently constructed by integrating intensity as well as geometry cues, reflecting the confidence of each pixel as a lane feature. In order to deal with complex road geometry, we employ Catmull-Rom splines to represent lane boundaries and the left and right lane boundaries are estimated separately in a tracking process using particle filter based on the weighted graph. In the proposed framework, unlike most of previous methods we lay a strong emphasis on accurate and effective lane feature detection since the challenges happen in the very first step of lane detection, and accurately detected lane features can be expected to reduce the complexity and difficulty, as well as improve the accuracy of lane detection in the following steps.**

I. Introduction

ROADWAYS are typically marked with painted and physical boundaries to assist the safe and efficient transportations. Real-time lane detection and localization from a moving vehicle using on-board sensor data is one of the key issues for many intelligent transportation systems. Within the last few years, research on this topic has expanded into a wide variety of applications from navigating a fully autonomous driving system to providing road information to a driver assistance system. The problem of finding lanes can be divided into three sub-problems: lane feature detection, lane boundary estimation, and lane tracking [1]. Lane feature detection refers to the use of on-board sensors to extract the road markings, including painted markings and other environmental markings such as color or texture discontinuities that define the safe and legal regions of driving. Lane boundary estimation is that of using the detected lane features to estimate the shape and positions of the lane boundaries based on a certain representation of the road. Lane tracking is to infer the shape and positions of the lanes using the observations. Among them, lane feature detection is a key component to lane detection. On the one hand, the appearance of lane boundaries can be affected by a number of factors that are not easily measured and change over time, such as wear, type, lighting and weather conditions, etc. Therefore, accurate lane feature detection can be difficult, particularly in challenging environments. Example lane images in Fig. 1 exemplify this kind of challenge, in which, lane boundaries are low-contrast and changeful in the midst of noise. On the other hand, lane boundary estimation and tracking is closely correlated with lane feature detection result. Accurate and effective feature detection can reduce the complexity and difficulty of boundary estimation as well as improve the accuracy and robustness of lane detection while poor lane feature detection result will do everything opposite. Most of previous methods use relatively simple and loose lane feature detection for fast detection and make significant efforts to deal with the noisy observed evidences in lane boundary estimation. In challenging environments, such detectors will either detect too many false positives that are difficult to be discriminated from true positives or produce false negatives that is crucial for lane detection especially in challenging conditions such that the lane boundary can not be found no matter how good their lane boundary estimation system is. In this paper, unlike such methods we lay a strong emphasis on accurate and effective lane feature detection and present a lane detection and tracking approach designed to work in challenging environments.

Figure 1. Examle lane images in challenging environments

Figure 2. Our experimental intelligent vehicle and the stereo camera

Manuscript received March 9, 2010. This work was supported by the Research Center for Smart Vehicles of Toyota Technological Institute, Japan. Partner institutions are the Toyota Technological Institute at Chicago, USA, and Toyota Central R&D Labs, Inc., Japan.

Chunzhao Guo and Seiichi Mita are with the Toyota Technological Institute, Nagoya, Aichi 468-8511 Japan. (e-mail: {guo, smita}@toyota-ti.ac.jp).

David McAllester is with the Toyota Technological Institute at Chicago, Chicago, IL 60637 USA. (e-mail: mcallester@tti-c.org).

The proposed approach is developed in the scope of the stereovision-based navigation system integrated in our experimental intelligent vehicle shown in Fig. 2, which is a hybrid vehicle equipped with six computers, one stereo camera, two four-beam laser scanners, GPS and other sensors. Fig. 3 shows the flow diagram of the proposed approach.

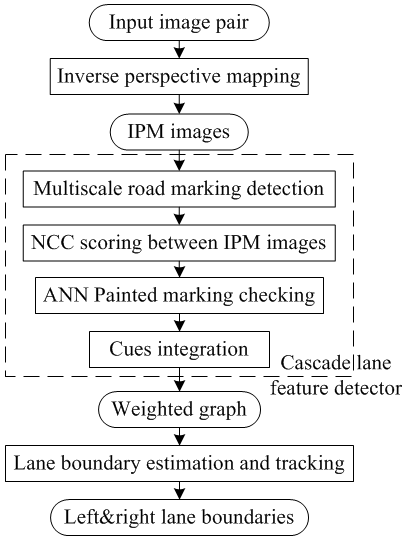

Figure 3. Flow diagram of the proposed approach

First, the Inverse Perspective Mapping (IPM) [2] is performed to transform the input image pair under the flat-road assumption. This procedure removes the perspective effect from the images by remapping the road points in the left and right images into points in the same world coordinates, in which the lane features are homogeneously distributed. Fig. 4 shows an example of IPM. (a)-(b) and (c)-(d) are the image pair and their inverse perspective mapped images (IPM images), respectively. (e) is the superimposed image with the edge images of (c)-(d) obtained by Canny edge detector [3], which shows the geometry cues underlying in the IPM images. Second, a sophisticated cascade lane feature detector is applied to the IPM images to cope with challenging environments. It starts with a multiscale road marking detector based on [4], which is able to find noisy, low-contrast road marking pixels. It employs multiscale elongated filters to measure the responses of difference of oriented means of multiple lengths and orientations in a hierarchical manner, along with a scale-adaptive threshold to eliminate responses due to noise. Pixels whose responses exceed the threshold will be preserved as potential road markings and their responses are used as the intensity cue. Subsequently, for each preserved pixel in the left IPM image, the Normalized Cross Correlation (NCC) is computed with the pixel at the same location in the right IMP image to measure their similarity. Pixels whose NCC values exceed a threshold will be further preserved as the road markings and their NCC values are used as the geometry cue. An Artificial Neural Network (ANN) classifier trained with painted road markings is then applied to check the types of preserved markings. Subsequently, a weighted graph is constructed by integrating the intensity and geometry cues, reflecting the confidence of each pixel as a lane feature. Note that physical road markings near the painted ones will be constrained to have a low confidence since they are not supposed to be used for lane boundary estimation due to traffic rules and regulations. In order to deal with complex road geometry, we employ one of the weakest models, Catmull-Rom splines, to represent the lane boundaries and the left and right lane boundaries are estimated separately in a tracking process using particle filter based on the weighted graph. The objective of the proposed framework is to deal with lane finding problems in challenging environments and overcome noise in the first step of the detection, since the challenges happen at the very beginning, and accurately detected lane features is expected to reduce the complexity and difficulty as well as improve the accuracy of lane detection in the following steps.

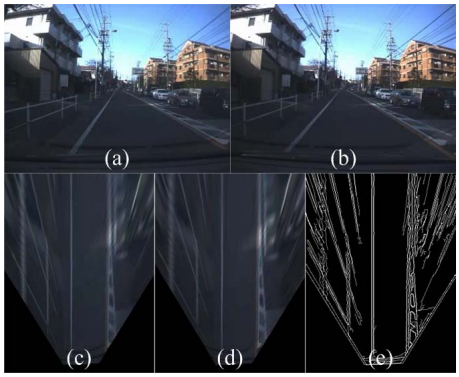

Figure 4. An example of IPM

II. Previous Related Work

Previous related work for lane detection can be categorized into two main types of methods: LIDAR-based methods [5]-[6] and vision-based methods [7]-[8], [10]-[11]. LIDAR sensors are useful in rural areas for helping to detect physical road boundaries, such as roadside fences and curbs. Some systems in DARPA Urban Challenge [9] also use LIDAR to detect painted lane boundaries since LIDAR is insensitive to changing environments so that the reliable detection of the lane boundaries can be expected in practice. However, only changes in the intensity of the returned laser data are indicative of painted road markings while the number of LIDAR beams is very limited. Therefore, such systems usually rely on GPS with other a priori information such as RNDF (Route Network Definition File) or GIS (Geographic Information System). Moreover, the GPS has limitations on the spatial and temporal resolution and the map data may be outdated and inaccurate. While such systems can perform extremely well in certain situations, vision can be utilized to perform well in a wide variety of situations since it can deliver a great amount of information. Vision-based lane detection problems have been extensively studied. McCall and Trivedi provide an excellent survey [10]. Recently, there has been much wok on modeling uncertainty problem in lane detection and tracking. Sehestedt et al. described the use of a particle filter for boundary tracking [11]. ZuWhan Kim presented a system that uses an ANN classifier, RANdom SAmple Consensus (RANSAC) spline

fitting, particle filtering and a combination of dynamic Bayesian network and maximum-likelihood estimation [8]. A. S. Huang et al proposed a probabilistic model of lane curvature called lateral uncertainty model [1]. These systems are concerned with the sub-problem of lane boundary estimation given a set of noisy observed lane features.

The proposed approach differs from previous related work in three key aspects. First, our approach focuses on lane detection in challenging environments, particularly when the low-contrast road markings appear in the midst of noise. Most of previous approaches overcome noise by applying a preprocessing step of isotropic smoothing such as Gaussian smoothing. Such smoothing, however, often reduces the contrast of weak lane boundaries, may blend adjacent boundaries, and may result in missing or poor localization of lane boundaries. We reduce the effect of noise without relying on a prior smoothing by a theoretical estimation of the effect of noise on the response of elongated filters, and derive a scale adaptive threshold which can distinguish between significant real responses and responses due to noise. These enable us to reveal and accurately localize very noisy, low-contrast road markings and consequently reduce the false negatives considerably in the lane feature detection. Second, unlike previous lane feature detectors which are designed for a certain type of systems or environments, our cascade lane feature detection can grasp the information about what it detects and the elongated filters are adjustable in length as well as orientations, which means our approach is potentially to be adaptive to different types of systems and environments. Finally, and most importantly, most of previous algorithms detect the lane features in the "hard" way by classifying the points in the image into two groups of "road markings" and "non-road markings" on the basis of whether they have some property of the lane boundary, such as color, shape, gradient, etc. In this case, the false positives will be used and treated equally for lane boundary estimation just like the true positives, which will increase the complexity and difficulty of boundary estimation and decrease the accuracy. In the same manner, the abandoned false negatives may result in misdetection of the lane boundary. It is a serious problem especially in challenging conditions since road markings are usually low-contrast and changeful with noise which tend to be false negatives in such situations. On the other hand, we utilize geometry cues of the road scene, which are hardly affected by the challenging conditions, together with intensity cues to detect and represent the lane features in a "soft" way by constructing a weighted graph with the integrated cues, reflecting the confidence of each pixel as a lane feature.

III. Lane Feature Detection

Road markings can vary greatly not only between regions but also over nearby stretches of road, making the generation of a single feature extraction technique difficult. Many of the previous algorithms only detect painted or solid and segmented lines by applying machine learning [8] or intensity-bump checking [12]. However, they will fail in situations where lanes are defined by physical boundaries, such as those in Fig. 1 (a), the left painted lane boundary disappears and the following lane boundary is defined by different materials with different colors/intensities, and in (b), the left lane boundary is defined by physical curbs. Besides, as shown in (c) and (d), the appearance of lane boundaries could be affected by a number of factors such as wear, lighting and weather conditions, etc. Therefore, we need to find the common as well as invariant features of both painted and physical lane boundaries as the road markings to be detected. In this paper, we define the road markings as "consistent elongated edge points on the road plane". Despite the noise as shown in Fig. 1 (c) and (d), such low contrast lane boundaries are evident due to their consistent appearance over lengthy curves, while the non-lane edges are not consistent elongated or reside outside the road plane. Occlusions by the nearby leading vehicles may result in a short lane boundary in some frames; however, it can be overcome by tracking.

A. Inverse Perspective Mapping (IPM)

In the proposed approach, we firstly perform the IPM to the input image pair and then apply a sophisticated cascade lane feature detection in the IPM images due to the following three reasons:

a) IPM normalizes the size of the road markings and reduces the range of lane boundary orientations, allowing simplifications and computational load benefit for the multiscale road marking detection. Besides, the normalized painted road markings are more suitable for the use of ANN classifier.

b) IPM remaps the road points in the left and right images into points in the same world coordinates, revealing the geometry relationships of the lane boundaries between the stereo pair. As shown in Fig. 4, the points on the road plane coincide very well (higher intensity) while other points that elevate out from the road appear distorted and not coincided (lower intensity).

c) The IPM image is accordant with the world coordinate so that the detection results of lane boundaries in the IPM image is easy to be used to update the grid map for navigation.

IPM is a linear transformation on homogeneous 3-vectors represented by a 3×3 matrix H, which can correspond each point x on the image plane to a point x' on the ground plane by the following equation,

$$x' = Hx \tag{1}$$

H can be derived from the knowledge of the camera as described in [7], or a simple external camera calibration with four reference points [13]. As shown in Fig. 4, the IPM image has a decreased resolution to further reduce the computational load, and only contain the information in the lower half of the original image since the entire lane boundaries lie in such regions under the current camera configuration.

B. Multiscale Road Marking Detection

Once the IPM images are obtained, we apply the multiscale road marking detector all over the left and right IPM images in parallel to detect road markings. Our goal is to extract consistent elongated road markings, particularly

for noisy, low-contrast ones in challenging conditions. Fig. 5 gives an example of such situation. Row (a) show an example image of a noisy, low-contrast and scattered road markings due to wear, collected in a dark rainy day. (b) shows the histogram of the image intensities in the rectangle. The red bar and the number indicate the mean of histogram. As we can see that, the means of orthogonal square or rectangular patches hardly change while the mean of oriented elongated rectangular patches varies remarkably across the road markings.

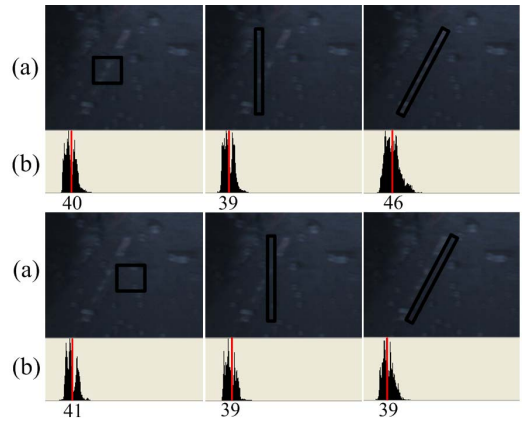

Figure 5. Example noisy, low contrast road markings.

Let $I(x,y)$ be the intensities of the image, and we denote oriented means by $F(x,L,w,\theta)$ which are the averages of $I(x,y)$ along rectangular patches of a given center position (x,y), length L, width w, and orientation θ,

$$F(\mathbf{x},L,w,\theta)=\frac{1}{wL}\sum_{u\in(-w/2,w/2)}\sum_{v\in(-L/2,L/2)} I(x+u\cos\theta-v\sin\theta,\, y+u\sin\theta+v\cos\theta) \tag{2}$$

Furthermore, we define the response of differences of oriented means as

$$D(\mathbf{x},L,w,s,\theta)=\frac{1}{2}F(\mathbf{x}+\frac{s}{2}(-\sin\theta,\cos\theta),L,\frac{w}{2},\theta) -\frac{1}{2}F(\mathbf{x}-\frac{s}{2}(-\sin\theta,\cos\theta),L,\frac{w}{2},\theta) \tag{3}$$

with $s\geq w/2$ to avoid overlap. The response D describes the average difference of neighboring oriented means of the same orientation. In this paper, we utilize such responses to find the noisy, low-contrast road markings. In our implementation, w is set to a small constant, yielding elongated oriented means. Considering the real situations, such setting is reasonable and effective to detect the road markings and reduce false alarms as well.

Given the responses over the image, we apply a scale adaptive threshold, determined by measuring the magnitude of the noise in the image as well as considering the length and width of the measured response, to identify difference of oriented means that elicit significant responses [4]. We assume that the noise does not vary significantly across the image and the pixel noise is normally distributed as a Gaussian: $N(0,\sigma^2)$. The difference of oriented means $D(\mathbf{x},L,w,s,\theta)$ is actually the average difference of two averages of $wL/2$ normal distributions. Consequently, D is distributed as a Gaussian: $N(0,\sigma_L^2)$, where $\sigma_L^2=\sigma^2/(wL)$. In order to find the threshold that can differentiate between responses due to road markings and those due to noise, we estimate the possibility that given a real number $t>>\sigma_L$, a value d drawn from the distribution $D\sim N(0,\sigma_L^2)$ satisfies $|d|<t$, i.e.,

$$\begin{aligned} p(|d|<t)&=\frac{1}{\sqrt{2\pi}\sigma}\sum_{x\in(-t,t)}\exp(\frac{x^2}{2\sigma_L^2}) \\ &\approx 1-\sqrt{\frac{2}{\pi}}\frac{\sigma_L}{t}\exp(-\frac{t^2}{2\sigma_L^2})\doteq 1-\varepsilon \end{aligned} \tag{4}$$

Suppose we produce $O(N)$ responses of length L, where N is the number of pixels in the image. To differentiate between true road markings responses and noise responses, we determine an adaptive threshold $t(w,L,N)$ such that the probability of correct detection in the N responses, i.e. $[p(|d|<t)]^N$, will be high. It means $(1-\varepsilon)^N=O(1)$, according to (4), which implies $1-\varepsilon N=O(1)$. We require $1-\varepsilon N\geq 1/2$ to assure high-probability, and the following approximate relation can be obtained,

$$\frac{t^2(w,L,N)}{2\sigma_L^2}=\ln N+\ln(\frac{\sigma_L}{t})\approx\ln N \tag{5}$$

where the approximation is due to the relation $\ln(\sigma_L/t)<<t^2/(2\sigma_L^2)$, under the assumption $t>>\sigma_L$. This approximation yields the scale adaptive threshold that takes into account the properties of the noise as well as the length and width of the oriented means,

$$t(w,L,N)=\sigma\sqrt{\frac{2\ln N}{wL}} \tag{6}$$

To fast estimate σ, for each pixel we calculate the minimal standard deviation obtained from the patches of 3×3 windows containing the pixel. We then construct a histogram summarizing the empirical distribution of this minimal standard deviation obtained for all the pixels. σ is determined as the 90th percentile value of this histogram, as shown in Fig. 6.

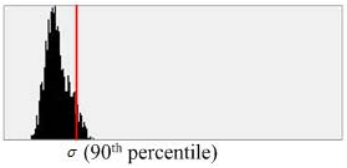

Figure 6. Estimating pixel noise.

A response $D(\mathbf{x},L,w,s,\theta)$ of length L (and fixed w) will be considered significant, potentially indicating the existence of road markings, if it exceeds the length adaptive threshold $t(w,L,N)$. However, long responses may still exceed the threshold due to scattered, short edges of high contrast. If the gaps between the scattered sub-portions are long relative to the length L, the responses may not be contributed by scattered worn road markings or dash

markings but by noise. In order to distinguish between responses due to long edges from responses due to scattered sub-edges, we define two parameters, α and β ($\alpha, \beta \in [0,1]$). We set αT_L to be a low threshold and βT_L to be a threshold on the total gap size for length-L responses. A length-L response cannot be considered significant if the total length of sub-portions that are smaller than αT_L is longer than βT_L.

As mentioned previously, we measure the responses of multiple length and orientations, which is implemented in a hierarchical manner [4], as shown in Fig. 7.

Algorithm 1: Response measurement in hierarchical manner

- INITIALIZATION
  - $L = 1$
  - Compute $D(x, L, w, s, \theta)$ with (3)
  - IF $D > T_L$
    Classify response D as significant
    Gap size = 0
  - ELSE IF $\alpha T_L < D \le T_L$
    Classify response D as nearly significant
    Gap size = 0
  - ELSE
    Classify response D as non-significant
    Gap size = L
  - END IF
- FOR $i = 1$: Number of scales $- 1$
  - $L = 2^i$
  - Compute $D(x, L, w, s, \theta) \leftarrow D(x \pm (L/4)(\cos\theta, \sin\theta), L/2, w, s, \theta)$
  - Gap size = sum of gap sizes associated with the above responses.
  - IF $D > T_L$ & total gap size $< \beta L$
    Classify response D as significant
    Gap size = 0
  - ELSE IF $\alpha T_L < D \le T_L$ & total gap size $< \beta L$
    Classify response D as nearly significant
  - ELSE
    Classify response D as non-significant
    Gap size = L
  - END IF
- END FOR

Figure 7. Hierachical responses computation

Finally, the pixels whose responses are classified as significant will be preserved. Non-maximum suppression is then applied at each length L to ensure well-localized positions, expressed by a single response. Fig. 8 gives an example of challenging road marking detection, in which we use the original road image in order to show the real challenging environment clearly. (a) shows the worn lane boundaries in an image collected in a dark rainy day explicitly (also in Fig.1 (c)), in which the left lane boundary is low-contrast as well as scattered due to wear, and there are many rain drops on the front window. (b) is the detected road marking pixels by the multiscale detection. As demonstrated in the figure, such multiscale detection along with the scale-adaptive threshold performs very well in road marking detection, which assures not only coherent long road markings but also consistent dash or worn road markings can yield significant responses while randomly distributed noise can not. The tiny false positives are due to some consistent rain drops, which is easy to be overcome in the following step of the cascade lane feature detection by using geometry cues. For comparison, (c)-(f) show the results obtained by Canny edge detector with the thresholds of (10,25), (20,50), (30, 75) and (40, 100). (g)-(h) show the result obtained by an ANN classifier. As we can see that it is difficult for Canny edge detector and ANN classifier to differentiate between such low-contrast road markings and noise. Besides, it's difficult to employ ANN classifiers to deal with physical lane boundaries since they are too many to be trained in the real world.

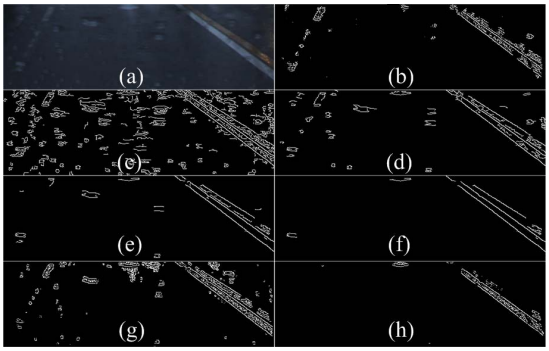

Figure 8. Example road marking detection results.

C. NCC Scoring Between IPM Images

In the second step of the cascade lane feature detection, we compute the NCC value between the left and right processed IPM images to measure the similarity of corresponding pixel locations. The aim of this step is to reveal and utilize the underlying geometry cue. For each detected road marking pixel in the left IPM image, the NCC in (7) is computed with the pixel at the same location in the right IMP image.

$$NCC = \frac{\sum_{(i,j)\in W}[f_1(i,j)-\overline{f_1}][f_2(i,j)-\overline{f_2}]}{\sqrt{\sum_{(i,j)\in W}[f_1(i,j)-\overline{f_1}]^2 \sum_{(i,j)\in W}[f_2(i,j)-\overline{f_2}]^2}} \quad (7)$$

where, W is the computational window, $f_1(i,j)$ and $f_2(i,j)$ are the image blocks in the left and right IPM images, respectively. $\overline{f_1}, \overline{f_2}$ are the average values of the blocks. Pixels whose NCC values exceed a threshold will be further preserved as the road markings and their NCC values are used as confidence of the pixel as a lane feature based on the geometry cue. The threshold is set to a relative small constant to only eliminate obvious non-road markings such as roadside fence.

In real applications, the geometry relationships between cameras and the road plane are actually not fixed since the road may be rough and the cameras may tilt and vibrate. However, the changes are usually small. In this case, even if the road marking pixels are not at the same position in the two IPM images, their scores are still higher than the others so that the road markings can still have more credits from the use of geometry cues. In this way, the proposed NCC scoring for geometry cue is tolerant of camera vibration and vehicle tilting considerably. Fig. 9 shows an example results after the first two steps of the cascade lane feature detection. (a) and (d) are input left images. (b) and (e) are their IPM images, respectively. (c) and (f) are the preserved lane features locations.

Furthermore, we also developed a dynamic stereo calibration algorithm that can estimate the homography matrices of IPM dynamically for each and every frame, since the precise geometry cues are crucial to achieve high

accuracy as well as robustness of lane detection in challenging traffic scenes. Details can be found in [14].

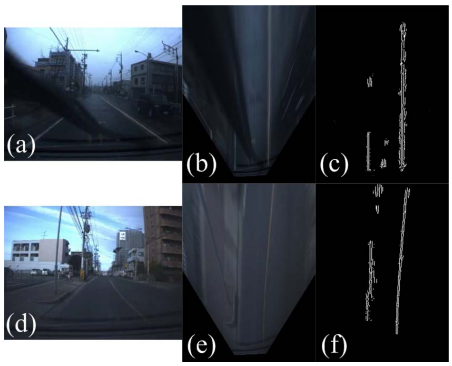

Figure 9. Weighted graph with integrated cues.

D. ANN Painted Marking Checking

Since road markings can vary greatly, as shown in Fig. 1, we employ multiscale road marking detection which can detect both painted as well as physical road markings. However, vehicles must drive not only safely but also legally, especially in urban streets. We should use the painted road markings, if exist, to estimate the lane boundary no matter how significant the responses of nearby physical boundaries are, due to the traffic rules and regulations. Therefore, the identification of the types of detected road marking is necessary to avoid the painted road markings get lost in competing with physical road markings. More specifically, we need to check the detected road markings if they are painted ones. To deal with such a problem, we apply machine learning. A comparative study of both classification performance and computation time on various painted markings classification methods is presented in [8]. Based on that, we choose to use an ANN classifier with two layers and seven hidden nodes since it is fastest, whereas the performances are still good. For training, we have gathered image patches of 100 painted road markings and 100 non-painted markings. For detection, the ANN classifier is applied on a small image patch of 9×3 windows around each and every preserved road marking pixels.

E. Cues Integration

Sensing the real world is an inherently uncertain process. Many previous approaches model uncertainty for lane estimation based on noisy observations of binary classified lane features, in which false positives are treated equally as the true positives. In the proposed approach, we intuitively model the uncertainty in the lane feature detection since the uncertainty, just like the challenges of lane detection in challenging environments, happens at the very beginning. Besides, the modeling of uncertainty in lane feature detection will decrease the uncertainty of boundary estimation. In particular, we construct a weighted graph by integrating the intensity and geometry cues, reflecting the confidence of each pixel as a lane feature, which assures that each pixel has a weight so as to play different role when estimating lane boundaries using particle filter. In this way, the uncertainty of lane feature detection and the one of lane boundary estimation are integrated for the probabilistic reasoning and decision making. Therefore, more accurate as well as robust detection can be expected.

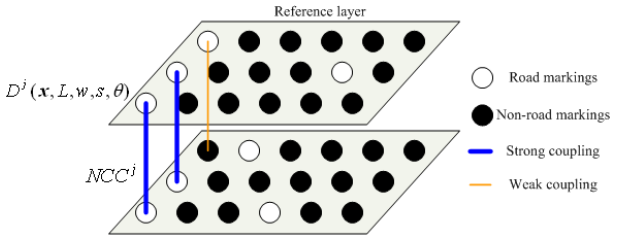

Figure 10. Weighted graph with integrated cues.

The weighted graph $G = (V, E, B)$ is shown in Fig. 10, where each node $v^j \in V$ represents a pixel, associated with the response $D^j(\mathbf{x}, L, w, s, \theta)$ as the node weight. Each edge $e^j \in E$ connects pixels at the same location between the left and right IPM images, associated with the NCC value NCC^j. The integrated weight $b_\theta^j \in B$ of a node in the reference layer, which is constructed from the left IPM image, is computed as follows,

$$b_\theta^j = \lambda_1 \exp(k(1 - NCC^j)) + \lambda_2 L \cdot D^j(\boldsymbol{x}, L, w, s, \theta) / D_{\max} \quad (8)$$

where λ_1, λ_2, k are parameters for the tuning of cues integration. $D_{\max}$ is the maximum value of $D^j(\mathbf{x}, L, w, s, \theta)$. We give more weight to the long response since they are more likely to be part of a lane boundary. Note that the integrated weight reflects not only the confidence of the pixel as a road marking but also the orientation of the road marking segment containing the pixel.

Finally, the integrated weight b_θ^j will be constrained to be a small value if there are painted road marking segments with the same orientation in a certain short distance.

IV. Lane Boundary Estimation using Particle Filter

The following task is to estimate lane boundaries by well utilizing the multiple evidences on the weighted graph. Challenges imposed in this step are maintly due to the complex road geometry, since we have already dealt with challenges due to varying appearances of lane boundaries. In order to deal with this challenge, we employ one of the weakest models, Catmull-Rom splines, to represent the road boundary and estimate the left and right boundaries separately, allowing more flexibility and less restrictions so as to give more weight to the observed evidences. This representation has been proved to be able to deal with challenging scenarios such as a lane curvature, lane changes, and emerging, ending, merging and splitting lanes [8]. As a result of this, we integrate the estimation of lane boundaries in a tracking process using particle filter in the proposed approach since particle filters allow to approximate arbitrary multimodal probability distribution recursively such that it is capable of handling poor process models.

A. Lane Representation

Splines are smooth piecewise polynomial functions, and they are widely used in representing curves. Various spline representations have been proposed, and we use the

Catmull-Rom spline among them because its control points are actually on the curve and it has local control, which means that modifying one control point only affects the part of the curve near that control point. These characteristics enable fast fitting of the curve. In the implementation, we employ two independent Catmull-Rom splines, and each boundary is determined by three control points.

B. Boundary Estimation using Particle Filter

Particle filter is a high versatility and robust stochastic filter. To estimate the state of a system a set of samples X at time t is used. This set $X_t = \langle x_t^i \mid i = 1,...,M \rangle$ and its associated weights w_t^i represent the confidence at time t. The weights are computed according to a sensor model, which contains the information of how likely it is that a sample x represents the true state. The computation of the posterior is then done by the iteration of state prediction and observation update. In the case of the proposed approach, we sample the set of control points L_t and R_t for the left and right lane boundaries hypotheses, respectively. The weights are updated with the weighted graph based on three measurement scores: the likelihood score, curve penalty, and correlation penalty.

1) Likelihood Score: The likelihood score is composed of the confidence and distance support of the weighted graph, which can be calculated as follows,

$$S_{LF} = \sum_p \exp(\mu \cdot b_\theta^p \cos(\theta - \varphi)) DistanceScore(p) \quad (9)$$

where p is a pixel on the lane boundary hypothesis, φ is the tangental angle of the hypothesis at p, and μ is a parameter to adjust the relative weight of the two supports.

The distance support is defined as

$$DistanceScore = \begin{cases} 1 - Distance(p)/K & Distance(p) < K \\ 0 & otherwise \end{cases} \quad (10)$$

where K is a threshold, and $Distance(p)$ is defined as

$$Distance(p) = \| p - NearestRoadMarking \| \quad |\theta' - \varphi| < \tau \quad (11)$$

where, $\| \ \|$ indicates the calculation of Euclidean distance, θ' is the orientation of the nearest road marking, τ is a threshold to ensure the distance is calculated between the pixels with similar orientations. On the weighted graph, if $b_\theta^p = 0$, the first item in (9) equals to 1 and the likelihood score is calculated by the distance score. If $b_\theta^p \neq 0$, the second item in (9) equals to 1 and the likelihood score is calculated by the confidence at p.

2) Curve Penalty: A penalty is imposed when the direction of the curve is changed while the likelihood score is very small, and the penalty is calculated as

$$S_{PC} = \sum_p |(y(p_2) - y(p_1))(\frac{dx}{dy}(p_2) - \frac{dx}{dy}(p_1))| \quad (12)$$

for all points pair p_1 and p_2 where the likelihood score between is smaller than a threshold.

3) Correlation Penalty: The above two scores are just for a single boundary. Here, we employ the correlation penalty to handle the lane width. We assume that the lane width can be linearly increasing or decreasing at small ratio. This assumption is reasonable since dramatic changes are dangerous and very few in the real world. Given a pair of left and right lane boundaries, the lane width is sampled at various distance and fit into a linear equation to obtain the average lane width and the change of the lane width. The lane boundaries penalty S_{PW} is then calculated as the maximum residual distance.

Based on these measurement scores, the boundaries are estimated using particle filter. In the implementation, we generate 100 samples for the left and right, respectively. To choose the best pair of left and right hypotheses as the detected lane boundaries, we define the total likelihood function $S_{l,r}$ as follows,

$$S_{l,r} = (S_{LF}^l - \gamma S_{PC}^l)(S_{LF}^r - \gamma S_{PC}^r) S_{PW} \quad (13)$$

where l, r indicate the scores of the left and right hypotheses, respectively, and γ is a weight parameter.

V. Experimental Results

In the experiments, the proposed approach has been implemented to test a wide variety of typical but challenging road environments without code optimization, which are from five video clips with a total of 6121 images of raining, shinning, cloudy, day and night roads. The experimental specifications are shown in Table I. We employ differences of oriented means $D(\mathbf{x}, L, w, s, \theta)$ at lengths $L = 1,2,4,8,16$ with $s = 5, w = 4, \pi/4 \le \theta \le 3\pi/4$ for the multiscale road marking detection. The NCC is calculated in a 9×9 window. The average computational time is under 200 ms/frame, which can satisfy the autonomous driving at 180km/h given the local navigation map with the road information of 10m. Much of the CPU time is used in the multiscale road marking detection. The computational time will be further reduced by utilizing special image processing hardware, adjusting the image resolution as well as optimizing the codes of the proposed method.

TABLE I. Experimental specifications

CPU	Xeon(R) X5550 @2.67 GHz
Memory (RAM)	12.0GB
Operating system	Window Xp Professional Sp2
Programming language	C++ with OpenCV library
Experimental data	Grayscale images (maximum visible range of the road > 100m)
Resolution	320×240
Computational time	<200 ms/frame

Fig. 11 show a number of example detection results in several typical but challenging environments. In each image block (a)-(f), the first column shows original left images, the second column shows the detected lane boundaries superimposed on the original images, and the third column shows the results in the IPM images. (a)-(c) illustrates our method's ability to overcome noise and localize low-contrast road marking inspire of rain drops, wipers, strong backlight, night scene. The proposed framework of the weighted graph and integrated cues enable the proposed method to keep the accurate and robust detection even in such challenging

Figure 11. Example detection results.

lighting and weather conditions. The effects of noise are considerably eliminated in the cascade lane feature detection, thanks to the multiscale road marking detector and the use of geometry cue. (d) - (f) shows the challenges in the typical road conditions, such as low contrast lane boundaries, occlusions by shadows and nearby vehicles, and irregular boundaries due to temporary road constructions as well. The results illustrate our approach can also work well in such environments.

VI. Conclusion

In this paper we presented a vision-based approach for lane detection and tracking in challenging environments. Our first contribution is a sophisticated cascade lane feature detector, which can reveal and localize the noisy, low-contrast road markings, and grasp the information of the markings it detected, such as type, length, orientations, etc. Our second contribution is the weighted graph constructed with integrated intensity and geometry cues based on a "soft" detection of the lane features. It not only reflects the confidence of detected pixels as a lane feature, but also contains the information of preserved pixels. Such framework can be expected to deal with more complex problems since it has the ability to differentiate nodes thereby decompose the complexity of the problem. Both of the contributions improve the accuracy as well as robustness of lane detection. Experimental results prove the effectiveness of the proposed approach on a wide variety of challenging road environments.

References

[1] A. Huang, S. Teller, "Lane Boundary and Curb Estimation with Lateral Uncertainties," in *Proc.of the 2009 International Conference on Intelligent Robots and Systems*, St. Loius, USA, Oct. 2009, pp. 1729-1734.

[2] H. A. Mallot, H. H. Bülthoff, J. J. Little, and S. Bohrer, "Inverse perspective mapping simplifies optical flow computation and obstacle detection," *Biological Cybernetics* 64, pp. 177-185, 1991.

[3] J. Canny, "A Computational Approach to Edge Detection," in *IEEE Trans. Pattern Analysis Machine Intelligence*, 1986, pp. 679-714

[4] M. Galun, R. Basri, and A. Brandt, "Multiscale edge detection and fiber enhancement using differences of oriented means," *11th IEEE Int. Conf. Computer Vision (ICCV-07)*, Rio de Janeiro, Brazil, 2007.

[5] B. Ma, S. Lakshmanan, and A. O. Hero, "Simultaneous detection of lane and pavement boundaries using model-based multisensor fusion," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 5, pp. 135–147, Sep. 2000.

[6] A.V. Reyher, A. Joos, H. Winner, "A lidar-based approach for near range lane detection," *Proc. Conf. on Intelligent Vehicles Symposium*, 2005, pp. 147 - 152.

[7] M. Bertozzi and A. Broggi, "GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection," *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 62–81, Jan. 1998.

[8] Z. Kim, "Robust lane detection and tracking in challenging scenarios," *IEEE Trans. Intelligent Transportation Systems*, vol. 9, no. 1, pp. 16–26, Mar. 2008.

[9] http://www.darpa.mil/grandchallenge/index.asp

[10] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Transactions on Intelligent Transport Systems*, vol. 7, no. 1, pp. 20–37, Mar. 2006.

[11] S. Sehestedt, S. Kodagoda, A. Alempijevic, and G. Dissanayake, "Robust lane detection in urban environments," in *Proc. IEEE Int. Workshop on Intelligent Robots and Systems*, San Diego, CA, USA, Oct 2007.

[12] S. Ieng, J. Tarel, and R. Labayrade, "On the design of a single lane-markings detectors regardless the on-board camera's position," in *Proc. IEEE Intell. Veh. Symp.*, 2003, pp. 564–569.

[13] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[14] C. Guo, S. Mita, and D. McAllester, "A vision system for autonomous vehicle navigation in challenging traffic scenesusing integrated cues," in *IEEE Proc. Intell. Transp. Syst.*, Madeira Island, Portugal, 2010.