



HAL
open science

Color-Based Road Detection and Its Evaluation on the KITTI Road Benchmark

Bihao Wang, Vincent Fremont, Sergio Alberto Rodriguez Florez

► **To cite this version:**

Bihao Wang, Vincent Fremont, Sergio Alberto Rodriguez Florez. Color-Based Road Detection and Its Evaluation on the KITTI Road Benchmark. IEEE Intelligent Vehicles Symposium (IV 2014), Jun 2014, Dearborn, United States. pp.31-36. hal-01010998

HAL Id: hal-01010998

<https://hal.science/hal-01010998v1>

Submitted on 21 Jun 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Color-based Road Detection and its Evaluation on the KITTI Road Benchmark

B. Wang^{1,2} V. Frémont^{1,2} and S. A. Rodríguez^{3,4}

Abstract—Road detection is one of the key issues of scene understanding for Advanced Driving Assistance Systems (ADAS). Recent approaches has addressed this issue through the use of different kinds of sensors, features and algorithms. KITTI-ROAD benchmark has provided an open-access dataset and standard evaluation mean for road area detection. In this paper, we propose an improved road detection algorithm that provides a pixel-level confidence map. The proposed approach is inspired from our former work based on road feature extraction using illuminant intrinsic image and plane extraction from v-disparity map segmentation. In the former research, detection results of road area are represented by binary map. The novelty of this improved algorithm is to introduce likelihood theory to build a confidence map of road detection. Such a strategy copes better with ambiguous environments, compared to a simple binary map. Evaluations and comparisons of both, binary map and confidence map, have been done using the KITTI-ROAD benchmark.

Index Terms—Road detection, Color images, binary map, confidence map, KITTI-ROAD benchmark

I. INTRODUCTION

Nowadays, Advanced Driving Assistance Systems have achieved great developments, and several new algorithms have been proposed to obtain a better understanding of the environment, in order to improve traffic safety and efficiency. Road detection is one of the key issues of environment understanding for intelligent vehicles. Many approaches that have been developed, involve many kinds of sensors like lidar, radar, camera; many different features like texture features [1], [2], and spatial feature[3]; and different algorithms like optical flow[4], or neural networks[5]; even the results presentations are various: perspective mapping[6], occupancy grid[7], and bird eye view mapping[3], [8]. Most of these approaches are evaluated on different datasets with different measurements. Fortunately, [8] has introduced an open-access dataset and benchmark which is called KITTI-ROAD for road area detection. They also provide a web interface to evaluate road detection approaches in the 2D Bird's Eye View (BEV) space.

In recent years, many vision-based road detection approaches have been proposed such as [1], [2], [9]. The versatility of vision systems provides rich information like colors, shapes, and depth at low cost with reduced power consumption. However, vision-based systems are sensible to illumination conditions such as shadows, back-lighting and low rising sun conditions. In [6], a stereo vision based

fast road detection approach has been proposed to handle this problem. It uses the illuminant invariance theory on color images to extract road surface feature, and classify the pixels using confidence interval, finally a stereo vision based road profile extraction helps to confirm the detection result. However, in real traffic scenario, especially with unstructured road, a simple binary classifier presents some limits. Ambiguities often happen in the real driving scene. To handle this issue, we proposed a likelihood-based algorithm which provides a confidence map of the detection result inspired by [6]. In [6], there are two main parts in the algorithm: pre-detection from illumination intrinsic image and plane extraction from the v-disparity map segmentation. The idea is to build a likelihood for each pixel after these two main procedures, and to calculate a confidence map by fusing the two likelihood results. The objective is to show that the confidence map should be more flexible than a simple binary map in complex environments.

Both the results of binary map and confidence map are evaluated on the KITTI-ROAD benchmark [8]. The evaluation results support our hypotheses in the sense that the confidence map is more adaptive to ambiguous scenery, while binary map outperforms in regular road scenes configuration. Comparisons are also made with the other algorithms published on the KITTI-ROAD website. Our binary map take the second place in F-measure, and the confidence map algorithm still needs to be improved.

The paper is organized as follows: Section II gives a general description of the fast road binary map generation algorithm; in Section III, a likelihood-based road confidence map generation algorithm is introduced in details with a discussion of its application. Section IV shows the experiment results and their evaluation on the KITTI-ROAD benchmark. Finally, a conclusion is given in Section V.

II. FAST ROAD DETECTION WITH BINARY MAP

As described in [6], the method pipeline starts with a first processing for sky and shadows removal applied to the dataset. An image conversion into a log-chromaticity space provides an efficient and stable identification (i.e. illuminant invariant) of road surface intrinsic features. Then, sample pixels are randomly selected from an assumed “road” area. Next, a confidence interval is defined using the samples to classify pixels of the intrinsic image into the road and non-road surfaces. Finally, stereo-vision based extension grants access to the 3D road profile estimation and enhances the detection precision. The whole system is show in Fig. 1

The authors are with ¹Université de Technologie de Compiègne (UTC), ²CNRS Heudiasyc UMR 7253, ³Université Paris-Sud, ⁴CNRS Institut d'Électronique Fondamentale UMR 8622.

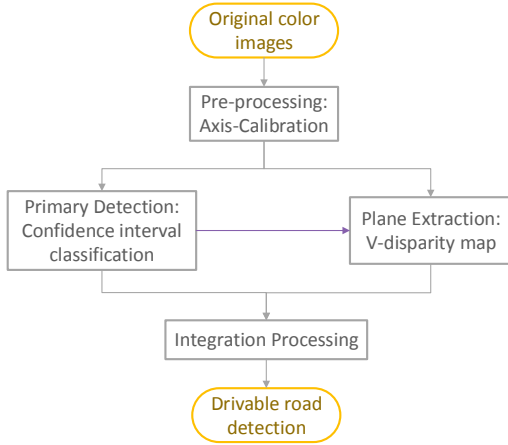


Figure 1. System of the fast road detection algorithm from color image

A. Illumination intrinsic image generation

In feature-based detection, shadows usually lead to wrong detections. Therefore, invariance properties to illumination conditions must be extracted.

Firstly, a geometric mean based log-chromaticity space is built from the RGB information of the color images [10]. In order to avoid favoring one particular channel, the R, G, B factors are divided by their geometric mean.

$$c_i = R_i/C_{ref} \text{ where, } C_{ref} = \sqrt[3]{R * G * B} \quad (1)$$

where, c_i defined as the chromaticity, is then transformed into log space as $\rho_i = \log(c_i)$. The color space ρ is orthogonal to $u = 1/\sqrt{3}(1, 1, 1)^T$.

Therefore, the transformation from the original coordinates to the geometric mean divided space is given by:

$$\chi \equiv U\rho, \quad \chi \text{ is } 2 \times 1 \quad (2)$$

where $U = [v_1, v_2]^T$, and $v_1 = (1/\sqrt{2}; -1/\sqrt{2}; 0)$, $v_2 = (-1/\sqrt{6}; -1/\sqrt{6}; 2/\sqrt{6})$.

Thus, the two dimensions of χ form a log-chromaticity space. In this space, the pixels on the same surfaces under different illuminations build a straight line. The lines which represent different surfaces are almost parallel, Fig.2. Their directions only correspond to the spectral power distribution. Their displacements are only related to the surface spectral reflectance function [6]. Hence, an image I_θ with suppressed shadows, can be obtained by projecting the lines on their common orthogonal axis which make an angle θ with the horizontal axis.

$$I_\theta = \chi_1 \cos \theta + \chi_2 \sin \theta \quad (3)$$

Entropy minimization[11] is used to find the correct axis θ (see Fig.2). After the geometric log-space transformation and projection onto this axis, as a result, the image I_θ is lighting independent and is also shadow-free.

In this research, the calibrated angle is defined off-line to

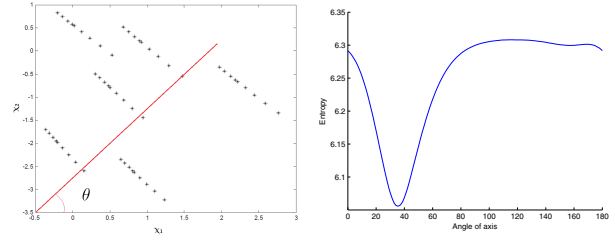


Figure 2. Axis calibration in log-chromaticity space. Left image: example of chromaticities for 6 surfaces under 9 lights in log-chromaticity space; Right image: Entropy plot along angle axis.

reduce time consumption. The result can be directly used for various driving scenes [9]. Thus, as a preparation work, calibration could be carried out, and would not affect the real-time detection performance. The algorithm is summarized in [6].

B. Confidence interval-based pre-detection

Practical fast road detection should be adaptive to dynamic environment. For this reason, it is necessary to reduce the dependency on the prior knowledge about road's ground truth. From this consideration, confidence interval theory is introduced and helps to determine an interval $[\lambda_1, \lambda_2]$ to separate the pixels into road or non-road class.

Since I_θ has eliminated the influence of shadows, the histogram composed by pixels on road surface is expected to be uni-modal with low dispersion and skewness. Therefore, the normalized histogram follows the empirical form of a normal distribution for a random variable, i.e. $I_\theta(\text{road}) \sim \mathcal{N}(\mu, \sigma^2)$. In the KITTI-ROAD dataset, there exist several surfaces similar to road surface, and in order to correctly separate them, a lower confidence level $1 - \alpha = 0.75$ is applied to calculate the confidence interval $[\lambda_1 = \mu - 0.6745 \frac{\sigma}{\sqrt{n}}, \lambda_2 = \mu + 0.6745 \frac{\sigma}{\sqrt{n}}]$ of $I_\theta(\text{road})$.

On image I_θ , pixels whose grayscale values fall outside this interval, are considered as background or obstacles. Therefore the classifier is redefined as:

$$\begin{cases} I_R = 1 \text{ Road,} & \text{if } \lambda_1 \leq I(p) \leq \lambda_2 \\ I_R = 0 \text{ non Road,} & \text{otherwise} \end{cases} \quad (4)$$

Eq. 4 is a classifier, and provides binary images of road detection results. Because the thresholds are based on a confidence level, some pixels can be mis-classified. Holes filling and 'majority' morphological operations can cope with false negative errors. For the false positives, we need fusion information to refine them. This is how stereo vision works in the proposed algorithm, see Section II-C.

Notably, when the vehicle stops right behind the front vehicle, the assumption of bottom road may not stand. For videos or continuous image sequences, a tracking processing is recommended to detect this situation. However, for the dataset composed of discrete frame from different sequence like KITTI-ROAD, it is still a tough issue need to be discussed. A possible way to solve this problem is using

disparity region regroup to decide if the bottom area is on a quasi vertical plane or quasi-horizontal plane which helps instructing sample selecting region.

C. Road extraction by stereo vision

In this paper, the former method presented in Section II-B is employed with stereo vision to assist the plane extraction and conversely the v -disparity map helps to correct the false-detected pixels. When driving in a complex environment, especially in urban scenery, artificial constructions are all along the road, where the materials are possible to show a similar intrinsic feature with the road surface. In this case, only confidence interval based classification is not enough to exclude these artificial structures, while in disparity map, these buildings represent obvious difference with road structure. Thus, stereo vision helps decreasing the false positive detection.

With two images of the same scene captured from slightly different viewpoints, a disparity map $I\Delta$ can be computed. $I_v\Delta$ is called the v -disparity map and is built by accumulating the pixels of same disparity in $I\Delta$ along the v -axis [12]. A projected plane in the v -disparity image follows the equation:

$$\Delta = av + b \quad (5)$$

Thus, for the road plane profile, the disparity value is linearly related to the v -axis.

In the difference of classical methods[13], [14], only the pixels classified as road area in the binary image I_R will be accumulated to the disparity map in our algorithm. The limitation of Region-of-Interest (ROI) (where, $I_R = 1$) will greatly reduce the run-time consumption. Reminding that I_R gives an pre-detection of drivable road area, most of the obstacles (e.g. buildings and pedestrians) can be avoided in the v -disparity computation. Thus, a regular sloping line as a representation of the drivable area is achieved as shown in Fig.4. According to Eq.5 which is extracted by the Hough Transform, the ground plane of image I_G could be constructed by:

$$\text{if } \Delta_p \in [\Delta_v \pm \varepsilon_v], I_G = 1, \text{ else, } I_G = 0 \quad (6)$$

Where, $\varepsilon_v = c \cdot v$ is a dynamic variance, which depends on v with a constant scale factor c . The variable c is a positive parameter which indicates that the closer the layout is, the greater the variance becomes. It is determined by preserving the pixels in a 0.95 leveled confidence interval. The variance of disparity on each line is calculated during the accumulation which is used to calculate a proper factor c . Once the parameter is fixed, it is adaptable to most of the driving scenes. Finally, the intersection of I_G and I_R represent a double verified road detection result, i.e. $I_{final} = I_R \cap I_G$.

However, the v -disparity maps are not always as ideal as expected. Walk sides, or the deformation of the road edge (usually depression) represent a bunch of lines slightly different from road profile in v -disparity space. Fortunately, they only take a small portion of the ROI. A dominate factor preservation step is added to refine the road profile in v -disparity image. The three upper images in Fig.4 shows



Figure 3. Binary map detection results on the KITTI-ROAD dataset. Each line presents two images from UM, UMM and UU dataset separately.

comparison of v -disparity images before and after preserving pixels with high intensity values. As a result, the line indicating road profile becomes more clear and precise.

Fig.3 shows the final result of the algorithm integrated with illumination invariant image, confidence interval and v -disparity map. This result is represented in form of a binary map, which means there is no need for further training or threshold determination on this detection result.

III. LIKELIHOOD METHOD WITH CONFIDENCE MAP

Binary map detection result requires a strict precision of each parameter in the algorithm. However, in some other cases, even the same road might be composed of different materials with different surface textures. Thus, we not only need to be able to separate the surfaces different from road but also need to be tolerant to different textures on the same road. This would be a cruel request for binary map detection. To solve this problem a confidence map is built to provide a more flexible and still reliable road detection result. A confidence map might be much more practical in unstructured roads and high variability in scene layout and in strong changes of illumination conditions.

For generating a confidence map of the road surface, we introduced the concept of likelihood in both the pre-detection processing (Section II-B) and plane extraction processing (Section II-C) in Fig.1.

A. Likelihood distribution of pre-detection

As in Section II-B, a pre-detection binary image I_R is obtained from the intrinsic image I_θ . As mentioned above, there exist surfaces with similar intrinsic features to the road surface, and also the road surface itself might show different textures caused by materials, extreme illumination conditions like over-saturation. The former problem will lead to false positive detection, and the latter one will cause false negative detection. To deal with these two conflicting situations in a robust algorithm, a likelihood value is assigned to each pixel of I_R with a 3-by-3 filter matrix composed of 1. For every pixel, the likelihood is calculated by the sum of its neighbors in the 3-by-3 operator, and then normalized:



Figure 4. Examples of v -disparity map in different situations (selected from UU dataset in KITTI-ROAD). Up left : Image of which stereo baseline is not parallel to the ground plane. Bottom left : Image taken from common situation. Middle: v -disparity maps generated from pre-detected ROI. Right: refined v -disparity with high intensity pixels preserved.

$$L_R(v, u) = \sum_{i=v-1}^{v+1} \sum_{j=u-1}^{u+1} I_R(i, j) / 9 \quad (7)$$

Since I_R is a binary image, only the valid pixels (where, $L_R = 1$) after pre-detection will be accumulate to the likelihood computation. The more valid pixels around, the more likely they are on the road surface. As to the false positive pre-detected pixels, they are commonly spread sparsely. After filtered by the operator, they are distributed only with a small likelihood value. On the other hand, for the false negative pre-detected pixels, they can gain some likelihood values thanks to their correctly detected neighbors. Thus, Eq.7 successfully transforms a binary result into likelihood distribution map.

B. Likelihood distribution of plane extraction

During the driving scene, the stereo baseline is not always parallel to the ground. Hence, the ground plane represented in the picture might be non-horizontal. For example, the left side is lower/higher than the right side. In this case, the disparity plane values along the image u -axis are not centered around a specific value Δ_v , but differ in a broader range, as shown in Fig.4. The disparity value of plane pixels stands on the same row v of the image, is more likely following a uniform distribution, thus, a dominant disparity value with high intensity does not exist, Fig.5. Also, it is hard to define a range of disparity values along the road profile in the v -disparity map.

When there is an obstacle along the road standing on the lower side of ground plane, the disparity value of the obstacle might give ambiguity in the v -disparity accumulation. As in Fig.4, the upper line shows an image taken by non-horizontal cameras. Right next to the image, the v -disparity map extracted from pre-detection result is illustrated. A bunch of lines spread almost uniformly in a broad range. After the refinement step for preserving dominant disparity values along the v -axis, two lines appear causing an ambiguous situation.

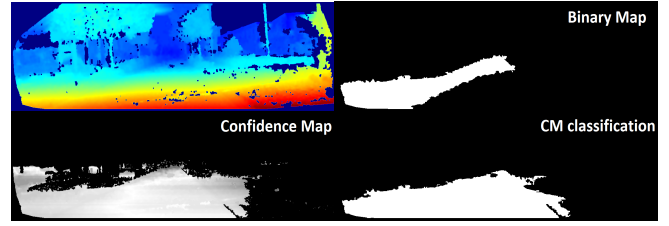


Figure 5. Detection results of non-horizontal image. Up left: Disparity map of the non-horizontal image presented in Fig.4. Up right: Binary map generate by original algorithm (Section II), which, directly represents the road detection result. Bottom left: Confidence map generated by improved algorithm (Section III). Bottom right: Road detection result by applying a proper threshold on confidence map (CM).

Hence, a simple binary classifier is not enough to handle this problem. Firstly, the range of disparity values on the plane $[\Delta_v \pm \varepsilon_v]$ is widen, so it is possible to have mis-detected obstacles as road surface, as long as their disparity value falls into this wide range. Secondly, the disparity value corresponding to the v -axis Δ_v might even deviate by a wrong Hough line extraction. However, building a complex binary classifier is time consuming, and probably need to verify multiple cues like road's topological and morphological characters. For example, the continuity of the pixels disparity value on the same v -axis might need to be considered.

Another way is to build a likelihood distribution for plane extraction, which measures the deviation of the disparity values to its expectation Δ_v . Since the dominant value extraction is not reliable, another criterion need to be proposed. The median factor is a proper candidate for the new criterion Δ_v . It is because road is a sloping plane, then disparity on it on the same v -axis should follow a uniform disparity. Notice that the v -disparity map is built on pre-detected result of Section II-B, even if there exist bias and noise, the median value will fall on the ground plane. For reducing the influence of false positive pre-detection, and speed up the algorithm, only the biggest connected component in I_R is preserved as a new Region-of-Interest I_{new_ROI} .

In binary map detection algorithm, a range of disparity values along v -axis is calculated as $[\Delta_v \pm \varepsilon_v]$, but here: $\Delta_v = \text{median}(\Delta(p_v))$, $\Delta(p_v)$ is the disparity value of the pixels on the v^{th} row of the image and $p_v \in I_{new_ROI}$. Pixels are detected as road surface in I_R , if their disparity value fall out of this range. Then, they will be distributed with a likelihood value depending on their disparity difference to Δ_v .

$$L_G(v, u) = (1 - |\Delta_{I_R}(v, u) - \Delta_v| / \Delta_v) \quad (8)$$

With Eq.8, every positive pixels in I_R has a value up to 1; then we add a unit step function to eliminate those negative likelihood. Thus, we get a likelihood map of ground plane L_G within the range of $[0, 1]$;

$$L_G = \frac{1}{2} L_G \cdot (1 + \text{sgn}(L_G)) \quad (9)$$

C. Confidence map generation

With the two likelihood maps L_R and L_G , a confidence map of road detection could be generated based on the following fusion function:

$$L_C(v, u) = L_R(v, u) \cdot L_G(v, u)$$

Thus, every potential road surface pixel has been distributed with a confidence value. After an evaluation on the training set of KITTI-ROAD benchmark, a best threshold for the confidence value will be found. Fig.5 compares the detection performance of binary map and confidence map on the non-horizontal image presented in Fig.4. In Fig.5, the binary map detection result is deviated due to the ambiguity of the road profile line in v-disparity map. On contrary, in confidence map, every pixel that is likely to be on the road is presented with a confidence value. After applying a proper threshold on the confidence map, the road area is very well classified. Therefore, confidence map is more reliable in complex situation.

Furthermore, the introduction of a confidence map not only provides a good development based on the original algorithm, but also provides a compatible result with the other algorithms have been evaluated on the KITTI-ROAD benchmark.

IV. EVALUATION RESULTS AND DISCUSSIONS

The two algorithms in this paper are evaluated on the KITTI-ROAD benchmark[8]. It contains three different categories of road scenes: UU - urban unmarked road, UM - urban marked two direction road and UMM - urban multiple marked lanes' road. The whole dataset has 600 annotated training and test images captured from broad spectrum of urban road scenes. Binary map detection (BM) is evaluated both on training dataset in perspective space and on testing dataset in Bird Eye View (BEV) space, see Fig.6. Considering the confidence map (CM) generation function is an ongoing research which still needs to be improved, here, we only evaluate it on training dataset as a reference to binary map and baseline algorithm. Perspective evaluation is carried out on a local machine. Bird Eye View results are evaluated on KITTI web server.

Our local processing platform is a standard PC with Windows 7 Enterprise OS, Intel CPU of 2.66 GHz. The computation environment is MATLAB R2013b. The run-time for binary map method is about 2s per frame. A complete confidence map generation algorithm takes about 4s per frame. To speed up the algorithm, we add a processing of maximum connected area preservation, which is firstly used in the ROI calculation to reduce the calculation area for plane extraction. Secondly it is used in the final result definition to refine the result. We also improved the code in details, that is how the computing time is reduced by a factor two from [6].

Table I
RESULTS [%] OF PIXEL-BASED ROAD AREA EVALUATION ON TESTING DATASET.

	URBAN - BEV space					
	F _{max}	AP	Prec.	Rec.	FPR	FNR
SPRAY	86.33	90.88	86.75	85.91	7.55	14.09
BM	82.32	68.95	76.15	89.56	16.15	10.44
CNN	78.92	79.14	76.25	81.79	14.67	18.21
BL	75.61	79.72	68.93	83.73	21.73	16.27

A. Binary map evaluation

In the KITTI-ROAD benchmark, the images are selected from different sequences, which demands a more reliable detection result for single images. However, it leads to a problem for usual axis-calibration. The intrinsic angles are different from sequences according to changes in the camera parameters. In the test with UM dataset, axis-calibration varies from images to images, while for the UMM dataset it does not change a lot. This is why our algorithm performs better on UMM dataset in Tab I. In order to provide a reliable result on KITTI-ROAD benchmark, the ground truth of the training dataset is used to calculate the axis angle θ , which is approximately equal to 33° . Detection results are evaluated by F-measure, average precision, accuracy, and other standard measures like: precision, recall and false positive/negative rate.

Methods listed in Tab I are Spatial ray classification (SPRAY[3]); Convolutional neural network (CNN[5]) and Base line (BL), they all have been introduced in [8]. In the comparison, binary map detection (BM) performs the best in the measurement of recall and false negative rate. However, since the intrinsic features with a high confidence level detect every possible pixels on the road, in some special situations, binary map will simply lead to false positive detection, for example, when sidewalk shows similar intrinsic features with the road surface, it has a strong probability of being detected as road area. In general, binary map detection methods provides a relative high value on F-measure among the compared algorithms in BEV space. According to the F-measure, fast road detection from color image in binary map won the second place comparing with the other algorithms mentioned in [8]. The strength of binary map is its independency from prior knowledge of ground truth: it doesn't need PR curve analysis for a satisfying threshold of confidence. This makes it a portable algorithm which can be applied directly into dynamic environment.

B. Confidence map evaluation

Considering that the KITTI-ROAD dataset is composed of discrete frames from different sequences, (in different weather condition, with different camera parameters), a higher confidence level is assigned during the pre-detection step for likelihood distribution. Thus, more potential road surface pixels will be taken into consideration.

It is interesting that the average precision of the binary map generation algorithm is quite low, which can be hardly

Table II
RESULTS [%] OF PIXEL-BASED ROAD AREA EVALUATION ON TRAINING SET.

UM perspective space							
	F_{max}	AP	Acc	Prec.	Rec.	FPR	FNR
BL	89.27	92.18	96.53	88.93	90.17	2.26	9.83
BM	85.67	72.21	94.89	77.83	95.26	5.18	4.74
CM	81.69	80.46	94.09	81.06	82.33	6.67	17.67
UMM perspective space							
	F_{max}	AP	Acc	Prec.	Rec.	FPR	FNR
BM	88.76	81.29	94.55	87.04	90.55	4.20	9.45
CM	85.28	82.08	92.99	85.09	85.46	4.67	14.54
BL	82.81	89.21	91.23	77.54	88.86	8.02	11.14
UU Perspective space							
	F_{max}	AP	Acc	Prec.	Rec.	FPR	FNR
BL	80.79	86.13	94.70	79.00	82.67	3.42	17.33
BM	80.50	62.53	94.19	73.44	89.07	5.02	10.93
CM	75.88	71.48	93.18	72.53	79.55	4.69	20.45

compared to the other algorithms. Actually, this measurement, average precision, is a description of Precision-Recall curve with different thresholds to classify the confidence maps. While the binary map directly provides a definite Precision and Recall value, this measurement might not be suitable for it. Under this consideration, we develop the confidence map based on original algorithm so to evaluate our approach on this measure. According to the evaluation result in perspective view (Tab II), the confidence map greatly improved the average accuracy of the detection compare to binary map. Even though, the performance of confidence map general compared to binary image and baseline, it still shows great potential developing space to provide reliable results in complex environment, as been proved in Fig. 5. Besides, in the subset of UMM, confidence map even overpass the baseline. If the confidence map algorithm is improved further more, a good performance is promising, and the evaluation in BEV space will be proceeded later.

V. CONCLUSION

In this paper, we proposed a novel method which combined likelihood theory with our former work presented in [6]. The results of the original algorithm using a binary map and the improved one based on a confidence map are evaluated on the KITTI-ROAD benchmark. The experimental results show that the binary map provides a high value on the F-measure compare to the other algorithms (second place, only behind the SPRAY algorithm). Another strong point of this algorithm is it can provide a straightforward information of free road area without any training. Nevertheless, when drive in complex environments, the detection performance using the binary map falls sharply. As an improved approach, the likelihood-based confidence map performs better in these situations, such as non flat road surface and over-saturation images. However, according to the general evaluation result, there are still improvements need to be done for confidence

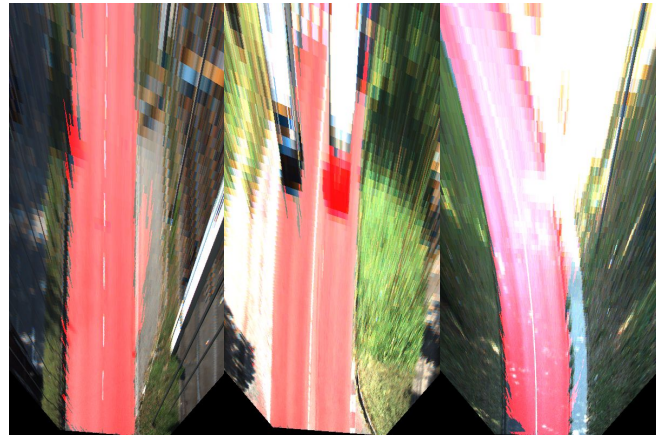


Figure 6. Detection results transformed in Bird Eye View (BEV) space.

map estimation. From this consideration, our future works will focus on the likelihood model construction.

REFERENCES

- [1] Hui K., J.-Y. Audibert, and J. Ponce. General road detection from a single image. *Image Processing, IEEE Transactions on*, 19(8):2211–2220, 2010.
- [2] S. Graovac and A. Goma. Detection of Road Image Borders Based on Texture Classification. *International Journal of Advanced Robotic Systems*, 9:1, 2012.
- [3] T. Kuehnl, F. Kummert, and J. Fritsch. Spatial ray features for real-time ego-lane extraction. In *Proc. IEEE Intelligent Transportation Systems*, 2012.
- [4] W. Yoshizaki, Y. Mochizuki, N. Ohnishi, and A. Imiya. Free Space Detection from Catadioptric Omnidirectional Images for Visual Navigation using Optical Flow. In *The 8th Workshop OMNIVIS*, 2008.
- [5] Jose M. Alvarez, Theo Gevers, Yann LeCun, and Antonio M. Lopez. Road scene segmentation from a single image. In *ECCV 2012*, volume 7578 of *Lecture Notes in Computer Science*, pages 376–389. Springer Berlin Heidelberg, 2012.
- [6] Bihao Wang and Vincent Frémont. Fast road detection from color images. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 1209–1214. IEEE, 2013.
- [7] Mathias Perrollaz, J-D Yoder, Anne Spalanzani, and Christian Laugier. Using the disparity space to compute occupancy grids from stereo-vision. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 2721–2726. IEEE, 2010.
- [8] Jannik Fritsch, Tobias Kuehnl, and Andreas Geiger. A new performance measure and evaluation benchmark for road detection algorithms. In *International Conference on Intelligent Transportation Systems (ITSC)*, 2013.
- [9] J.M.A. Alvarez and A.M. Lopez. Road detection based on illuminant invariance. *Intelligent Transportation Systems, IEEE Transactions on*, 12(1):184–193, 2011.
- [10] G.D. Finlayson, M.S. Drew, and L. Cheng. Intrinsic images by entropy minimization. In *European Conference on Computer Vision*, 2004.
- [11] Cheng Lu G.D.Finlayson, M.S.Drew. Entropy minimization for shadow removal. *International Journal of Computer Vision*, 2009.
- [12] David Pritchard. Cloth parameters and motion capture. Technical report, 2003.
- [13] Nicolas Soquet, Didier Aubert, and Nicolas Hautiere. Road segmentation supervised by an extended v-disparity algorithm for autonomous navigation. In *Intelligent Vehicles Symposium, 2007 IEEE*, pages 160–165. IEEE, 2007.
- [14] Basam Musleh, Arturo de la Escalera, and José María Armingol. Uv disparity analysis in urban environments. In *Computer Aided Systems Theory–EUROCAST 2011*, pages 426–432. Springer, 2012.