FEATURE++: CROSS DIMENSION FEATURE FUSION FOR ROAD DETECTION

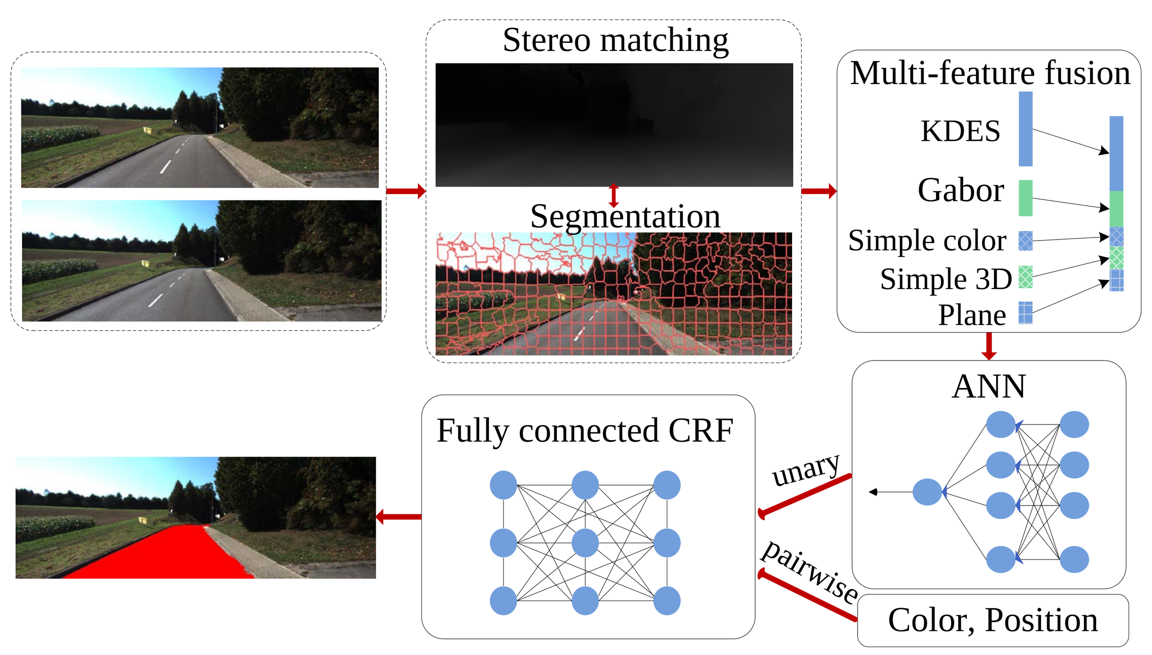
Wenli He[1], Guorong Cai[2], Zhun Zhong[1], and Songzhi Su[1]*

[1] Cognitive Science Department, Xiamen University [2] Computer Engineering College, Jimei University, Xiamen, Fujian, China

ABSTRACT

Road detection is a key component of Advanced Driving Assistance Systems, which provides valid space and candidate regions of objects for vehicles. Mainstream road detection methods have focused on how to extract discriminative features. In this paper, we propose a robust feature fusion framework, called "Feature++", which is combined with superpixel feature and 3D feature extracted from stereo images. Then a neural network classifier is trained to decide whether a superpixel is road region or not. Finally, the classified results are further refined by conditional random field. Experiments conducted on the KITTI ROAD benchmark show that the proposed "Feature++" method outperforms most manually designed features, and are comparable with state-of-the-art methods that based on deep learning architecture.

METHOD



Features:

- The 2D features include RGB, Gabor, Gradient kernel descriptors (GKDES) and RGB kernel descriptors (RGBKDES).
- The 3D features consist of Plane(angles and inliers percentage), Pos3D , Depth gradient kernel descriptors (DGKDES) and SPIN kernel descriptors(SPINKDES).

Fully connected CRF:

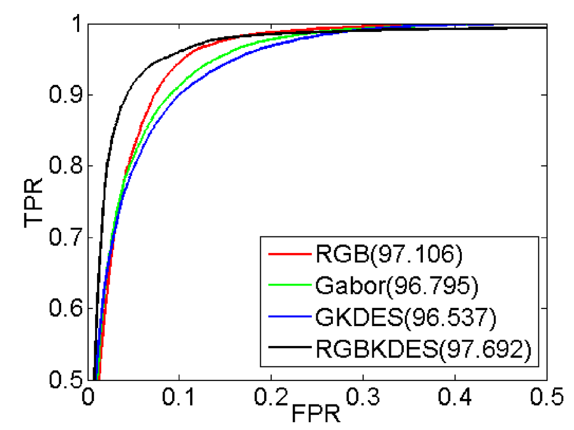
- Unary potential takes the negative log-likelihood of the road confidence.
- Pairwise potential= $\omega^{(1)} exp\left(-\frac{|p_i-p_j|^2}{2\theta_\alpha^2}-\frac{|I_i-I_j|^2}{2\theta_\beta^2}\right) + \omega^{(2)} exp\left(-\frac{|p_i-p_j|^2}{2\theta_\gamma^2}\right)$

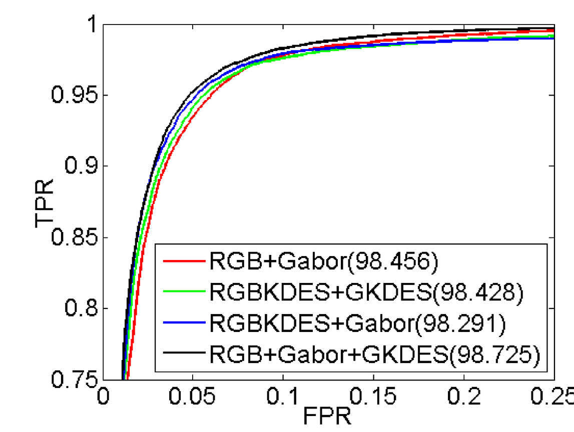
I is the color vector and *p* represents position.

RESULTS

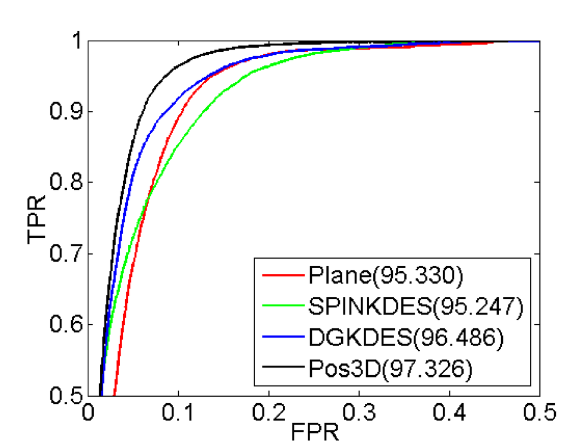
Feature Evaluation

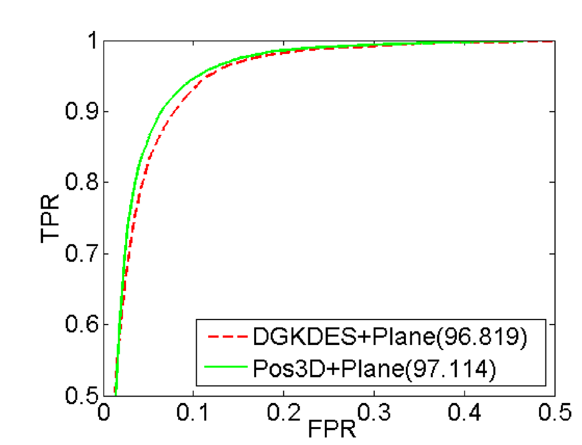
We first evaluate the 2D and 3D features respectively, then selectively fuse them to get the best feature fusion.



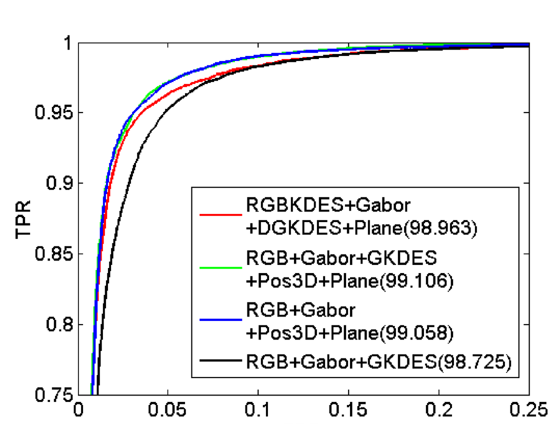


- Among 2D features, RGBKDES is the most discriminative for road.
- RGBKDES and RGB represent color information, Gabor provides texture and GKDES describes gradient. According to the appearance characteristics of the road, we process different combination, and the "RGB + Gabor + GKDES" performs the best.



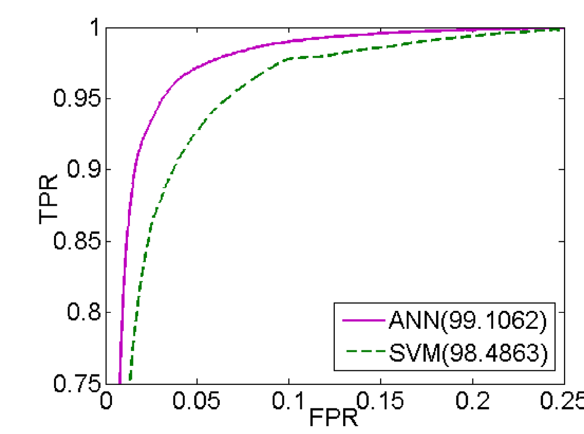


- For 3D features, Pos3D is the most discriminative.
- Both SPINKDES and Plane represent the angle information, and SPINKDES computes more details. Due to that stereo matching would have some error more or less, the SPINKDES, more affected by stereo noise, performs not as well as Plane.
- Both Pos3D and DGKDES represent the 3D position changes. Combined with angle feature, we process two fusions for road. The "Pos3D + Plane" performs better.



- Fuse the 2D and 3D features. "RGB + Gabor + GKDES + Pos3D + Plane" is shown the best fusion.
- Comparison of monocular and stereo performance shows the importance for the use of stereo information.

Classifier Evaluation



We evaluate how the classifier affects the performance. Experiment shows that ANN outperforms SVM in our method. ANN is an efficient tool for feature fusion and classification.

KITTI RAOD Benchmark Submission

With CRF[%]						
Benchmark	**MaxF**	**AP**	**PRE**	**REC**	**FPR**	**FNR**
UMM_ROAD	93.55	92.34	92.77	94.34	8.08	5.66
UM_ROAD	90.43	88.83	89.63	91.26	4.81	8.74
UU_ROAD	87.40	85.48	86.19	88.64	4.63	11.36
URBAN_ROAD	**91.12**	**89.51**	**90.16**	**92.10**	**5.54**	**7.90**
Without CRF[%]						
URBAN_ROAD	**89.48**	**90.45**	**87.74**	**91.28**	**7.02**	**8.72**

- Verify the necessity of the CRF post-process. The CRF can help the performance improve 2-3%.
- the performance in unmarked roads(UU_ROAD) is not as good as in marked roads. The most unmarked roads are in the rural areas and the surrounded environments are relatively complex.

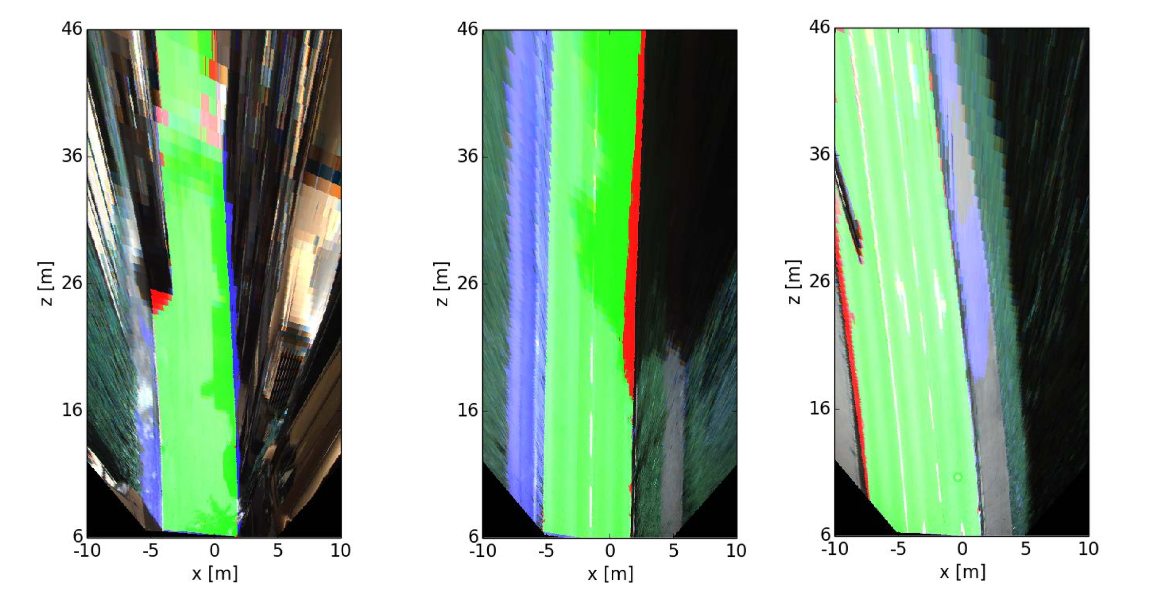
URBAN – BEV space[%]						
Method	**MaxF**	**AP**	**PRE**	**REC**	**FPR**	**FNR**
Feature++	**91.12**	**89.51**	**90.16**	**92.10**	**5.54**	7.90
NNP	89.68	86.50	89.67	89.68	5.69	10.32
CB	88.97	79.69	89.50	88.44	5.71	11.56
FusedCRF	88.25	79.24	83.62	93.44	10.08	**6.56**
SRF	82.44	87.37	80.60	84.36	11.18	15.64
BL	75.89	79.28	71.56	80.77	5.65	19.23

- Compare our "Feature++" with other methods that based on manually designed features .
- BL, CB and SRF are monocular based methods; FusedCRF fuses the LIDAR and monocular information; NNP is based on stereo.

Sample results in the perspective image

Red denotes false negatives, blue areas correspond to false positives and green represents true positives(top to bottom: UU, UM, UMM).

Corresponding results in the BEV image



CONCLUSION

In this paper, we have proposed a robust feature for road detection, which selectively fuses the Gabor, kernel descriptors, simple color and 3D spatial information. Meanwhile, we have evaluated each component of our system. Experiments show that our method outperforms the most of methods which use the manually designed features on the KITTI ROAD benchmark.

However, Illumination conditions usually affect the performance, especially in unmarked road scenes. In our future works, illumination invariant features would be included in the road feature fusion and used in the pairwise potentials of CRF.

ACKNOWLEDGMENTS

This work is supported by the Nature Science Foundation of China (No. 61572409, No.61402386 & No. 61571188), Fujian Province 2011 Collaborative Innovation Center of TCM Health Management and Collaborative Innovation Center of Chinese Oolong Tea Industry-Collaborative Innovation Center (2011) of Fujian Province.