Fusing Geometry and Appearance for Road Segmentation

Gong Cheng, Yiming Qian, James H. Elder
Centre for Vision Research
York University, Toronto, Canada
{gongcheng, yimingq}@eecs.yorku.ca, jelder@yorku.ca

Abstract

We propose a novel method for fusing geometric and appearance cues for road surface segmentation. Modeling colour cues using Gaussian mixtures allows the fusion to be performed optimally within a Bayesian framework, avoiding ad hoc weights. Adaptation to different scene conditions is accomplished through nearest-neighbour appearance model selection over a dictionary of mixture models learned from training data, and the thorny problem of selecting the number of components in each mixture is solved through a novel cross-validation approach. Quantitative evaluation reveals that the proposed fusion method significantly improves segmentation accuracy relative to a method that uses geometric cues alone.

1. Introduction

Vehicle-mounted cameras can provide critical visual data to support assisted and autonomous driving as well as road condition assessment. An important initial task is to segment the portion of the image that projects from the road, as opposed to portions of the vehicle on which the camera is mounted, other vehicles, the sidewalk or shoulder, overpasses, etc. This task is challenging particularly for cameras for which pan/tilt parameters are variable and/or unknown, so that a region of interest cannot be pre-defined. Other complications include occlusions caused by other vehicles and variability in road appearance due to weather conditions (rain, snow, etc) - see Fig. 1 for examples.

Recent methods use either appearance cues or geometric cues to estimate the road surface, but not both. Unfortunately appearance cues are fallible since non-road surfaces (e.g., sidewalk, buildings, overpasses) can be made of material that is very similar to the road surface, and with fresh snow cover it can be difficult to distinguish the road surface from neighbouring regions that are also covered with snow. While geometric methods are not subject to these limitations, they can fail when the road geometry is obscured (e.g., due to snow), or less well-defined (e.g., in parking lots). Moreover, geometric methods do not provide a means for excluding non-road objects such as vehicles or pedestrians. To address these limitations we propose here a novel probabilistic method for fusing both geometry and appearance cues and show that this leads to more reliable road segmentation.

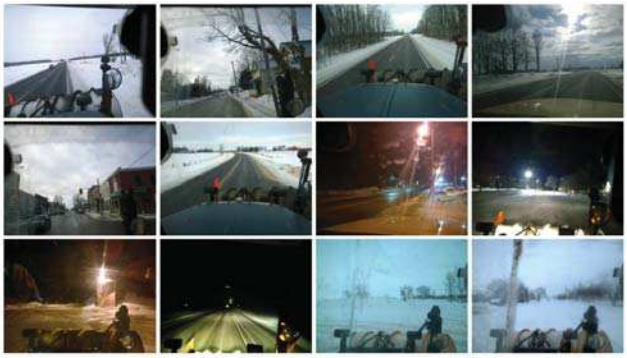


Figure 1. Example images from the road weather conditions dataset of Qian *et al*. [15].

To model geometric cues, we follow the approach of Almazen *et al*. [1], who used estimates of the road vanishing point or horizon to deliver a map of the probability that each image pixel projects from the road surface. To model road surface appearance we use a Gaussian mixture model (GMM) in RGB space. While simpler than many currently popular discriminative models for appearance [6, 14, 7, 8, 19] this probabilistic approach has the advantage of allowing appearance cues to be fused with probabilistic geometric cues in a rigorous way without resorting to ad hoc weighting functions.

The two main challenges in real-world computer vision application of GMMs is 1) adaptation to varying scene conditions and 2) selection of the number of components. A major contribution of our work is a novel supervised cross-validation method that optimizes the number of GMM components for each image in the training set and adaptively selects the optimal model for a new image at inference time.

2. Prior Work

Prior appearance-based road surface segmentation algorithms have modeled road pixel appearance in HSV [21, 18] or RGB [12, 4, 23] space. However, these methods become problematic when the road surface is covered by ice or snow and when other nearby objects (e.g., sidewalks, buildings) have similar appearance.

An alternative is to use geometric cues such as the road vanishing point [9, 16, 17], horizon [2], and/or road boundaries [10, 22, 24]. Here we employ the recent geometric approach of Almazen *et al*. [1], which uses either the detected vanishing point of the road or the horizon to identify the probability that each pixel of the image projects from the road surface. While Almazen *et al*. showed that this geometric segmentation method improved road weather classification performance, the resulting segmentations are approximate and do not exclude objects such as other vehicles that may be occluding the road surface. Indeed, Almazen *et al*. noted that residual error in road segmentation remains one of the factors limiting road weather classification accuracy.

Recently, Deep Neural Network (DNN) methods have been applied to the road segmentation problem. StixelNet [13] uses a 5-layer convolutional architecture inherited from LeNet [11], trained from on the KITTI dataset [5] for obstacle detection and road segmentation. SegNet [3] is a VGG16 [20]-based deep convolutional encoder-decoder network for semantic road scene segmentation that includes the road surface as one of 12 classes. The StixelNet system has only been evaluated on the KITTI dataset which lacks the more realistic and challenging diversity in camera pose, weather, illumination and road environments present in the Qian *et al.* dataset [15], and the StixelNet code is not publicly available. However, the SegNet code is publicly available, and we compare its performance against our proposed method on the Qian *et al.* dataset (Section 5).

Our proposed method uses the posterior probability map generated by the geometry-based method of Almazen *et al*. [1] as a prior and combines with appearance likelihoods based upon our mixture models to generate a fused posterior that combines both geometric and appearance cues. A final segmentation can then be generated by thresholding the posterior.

3. Dataset

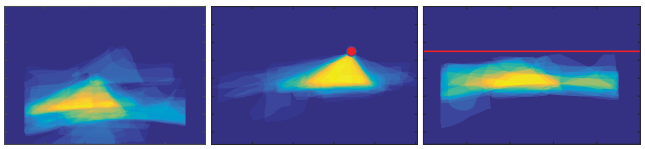
To train and evaluate the proposed method, we use the dataset of road images introduced by Qian *et al*. [15], obtained directly from the authors. The dataset consists of 100 2048×1536 pixel images in a 50/50 training/test split. For each of the images the portion of the image projecting from the road has been segmented by hand. It covers a broad range of road weather conditions, including dry, wet, snow and ice-covered as well as a broad range of illumination conditions, from full sunlight to night. The pose of the camera varies considerably - note that the horizon may appear either toward the top or toward the bottom of the image. Finally, the images were obtained from a diversity of road maintenance vehicles, and large portions of the road surface can be occluded by the hood of the vehicle and mounted snow-clearing equipment.

All parameter tuning was performed on the training dataset: the test dataset was used only for final evaluation.

4. Methods

4.1. Geometric Prior

Let $x_i \in \{F, B\}$ represent the provenance of pixel i, where F indicates foreground (i.e., the road surface) and B indicates background, and let $\mathbf{y}_i$ represent the colour of this pixel, in RGB space. Almazen *et al*. [1] observed that while the unconditional probability map $p(x_i = F)$ estimated from the training subset of the Qian *et al*. dataset [15] is quite broad due to the variability in camera pose and road geometry (Fig. 2(a)), anchoring the prior to a detected vanishing point (Fig. 2(b)) or horizon (Fig. 2(c)) tightens the distribution considerably. With this motivation, their method proceeds by first using dominant lines in the image to estimate the road vanishing point. If the resulting estimate is judged to be reliable, the vanishing point prior (Fig. 2(b) is employed to estimate the road segmentation. Otherwise the horizon line is estimated and the the horizon prior (Fig. 2(c)) is employed. In either case, Almazen *et al*. threshold the prior at 0.5 to deliver a final segmentation, which is then used to estimate road conditions. For our purposes we will retain the real-valued probability map to allow probabilistic fusion with appearance cues.



(a) (b) (c)

Figure 2. (a) Unconditional prior. (b) Prior anchored to road vanishing point. (c) Prior anchored to horizon.

4.2. Learning Appearance Models

The training subset of the Qian *et al*. dataset [15] provides ground truth segmentations of foreground (road) and background from which appearance models can be learned. Here we elect to model the appearance $\mathbf{y}_i$ of a pixel conditioned on its provenance x_i (foreground or background) as Gaussian mixture model (GMM) in RGB space. Unlike

discriminative methods this approach delivers a probability that allows principled fusion with our geometric prior.

In prior approaches [21, 18, 12, 4, 23], a single model for road appearance is learned from training data, and we could certainly follow this approach by learning a single GMM for foreground and a single GMM for background over all training data. However, this naïve approach would fail to account for the non-stationarity of these conditional distributions due to diversity in geography, weather and illumination conditions (Fig. 1). To address this challenge we instead adopt a nearest-neighbour approach. Specifically, we learn separate foreground and background models for each image in the training dataset. At inference time we then apply each model to the test image and select the model with highest likelihood. This allows the model to adapt to changing geography, weather and illumination conditions. Note that foreground and background models can be and often are drawn from different training images.

A second issue that arises with all mixture models is how to choose the number of components. Maximizing likelihood leads to overfitting, while maximizing performance on training data can be slow and lead to models that are too task-specific and do not generalize. Here we adopt a novel cross-validation approach that is simple but effective. Let Y_i^F and Y_i^B represent the set of appearance vectors for all pixels in foreground and background regions of training image i and let $\mathcal{M}_{ik}^F$ and $\mathcal{M}_{ik}^B$ represent k-component maximum likelihood GMMs for these appearance vectors.

We first determine these models for $k = [1 \ldots 20]$ and all training images. Our goal is then to identify, for each training image i, the optimal model complexity k_i^F and k_i^B for foreground and background regions, respectively. Here we define optimality in terms of the ability of the model to generalize, selecting the model that maximizes the log likelihood over the other images in the training dataset, assuming conditional independence of appearance observations over pixels:

$$k_i^F = \underset{k}{\operatorname{argmax}} \sum_{j \neq i} \log p\left(Y_j^F | \mathcal{M}_{ik}^F\right), \quad (1)$$

$$\text{where } p\left(Y_j^F | \mathcal{M}_{ik}^F\right) = \prod_{\mathbf{y}_l \in Y_j^F} p\left(\mathbf{y}_l | \mathcal{M}_{ik}^F\right) \quad (2)$$

$$k_i^B = \underset{k}{\operatorname{argmax}} \sum_{j \neq i} \log p\left(Y_j^B | \mathcal{M}_{ik}^B\right), \quad (3)$$

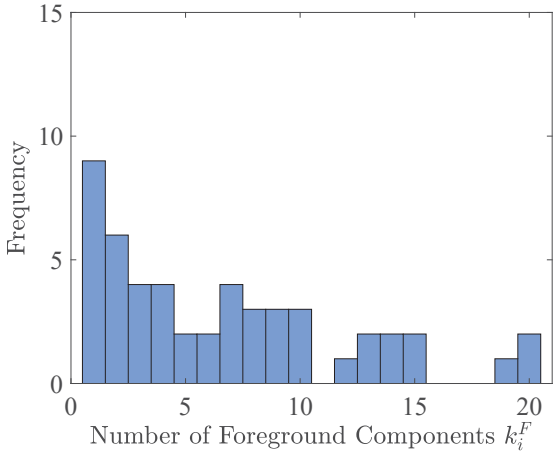
$$\text{where } p\left(Y_j^B | \mathcal{M}_{ik}^B\right) = \prod_{\mathbf{y}_l \in Y_j^B} p\left(\mathbf{y}_l | \mathcal{M}_{ik}^B\right) \quad (4)$$

In this way we identify a dictionary of optimal GMMs, one for each training image:

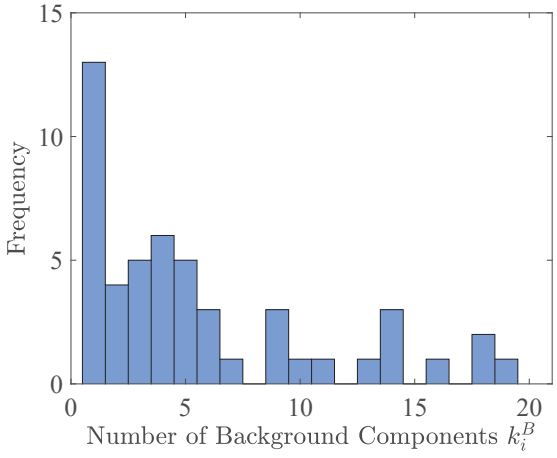
$$\mathcal{M}_i^F = \mathcal{M}_{ik_i^F}^F \quad (5)$$

$$\mathcal{M}_i^B = \mathcal{M}_{ik_i^B}^B \quad (6)$$

In effect this approach uses the other training images as surrogates for test images, optimizing the complexity of the model to maximize its predictive capacity. Fig. 3 shows the distributions over the number of mixture components k_i^F and k_i^B for the optimized foreground and background models derived from the training images.



(a)



(b)

Figure 3. Distributions over the number of mixture components for the optimized foreground k_i^F (a) and background k_i^B (b) models, over the training images.

4.3. Selecting an Appearance Model

Given a novel test image, we wish to select appropriate foreground and background appearance models from

our GMM dictionary learned from training data. This is a chicken-and-egg problem: If we knew which test pixels were foreground and background, we could evaluate each of the models from our foreground dictionary on the set of foreground test pixels and each of the models from our background dictionary on the set of background test pixels, selecting the models with maximal likelihood. However, to discriminate foreground from background pixels we first need to apply our appearance model.

We resolve this chicken-and-egg problem with a simple bootstrap approach. First, we use the methods of Section 4.1 to compute the geometric prior $p(x_i = F)$ over all pixels of the test image. We then use this geometric prior to identify samples $\hat{Y}^F$ and $\hat{Y}^B$ of the colours in the test image that we believe to be drawn from foreground or background regions:

$$\hat{Y}^F = \{\mathbf{y}_i : p(x_i = F) > 0.5\} \quad (7)$$

$$\hat{Y}^B = \{\mathbf{y}_i : p(x_i = F) = 0\} \quad (8)$$

The asymmetry between these two thresholds arises from the asymmetry in the geometric priors (Fig. 2) derived from the training data. While there are many pixels with a foreground probability of zero, there are very few with a foreground probability of one.

Finally, we select from our dictionary foreground and background appearance models that maximize the likelihood over these samples, again assuming conditional independence of appearance observations over pixels:

$$\mathcal{M}^F = \mathcal{M}_{\hat{i}}^F, \quad (9)$$

$$\text{where } \hat{i} = \underset{i}{\operatorname{argmax}}\, p\left(\hat{Y}^F | \mathcal{M}_i^F\right) \quad (10)$$

$$\text{and } p\left(\hat{Y}^F | \mathcal{M}_i^F\right) = \prod_{\mathbf{y}_i \in \hat{Y}^F} p\left(\mathbf{y}_i | \mathcal{M}_i^F\right) \quad (11)$$

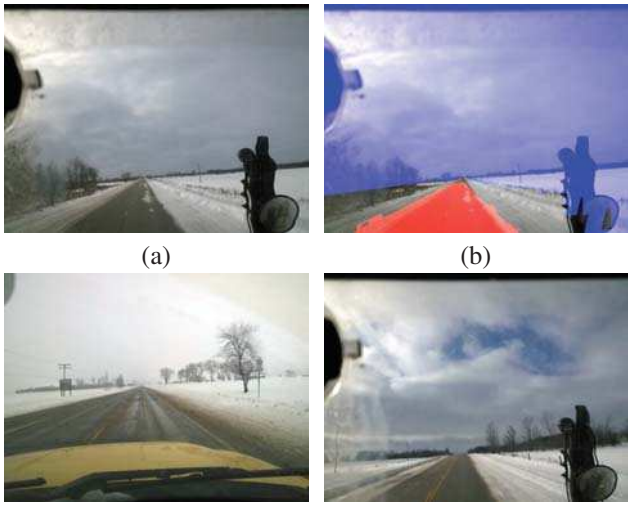
$$\mathcal{M}^B = \mathcal{M}_{\hat{i}}^B, \quad (12)$$

$$\text{where } \hat{i} = \underset{i}{\operatorname{argmax}}\, p\left(\hat{Y}^B | \mathcal{M}_i^B\right) \quad (13)$$

$$\text{and } p\left(\hat{Y}^B | \mathcal{M}_i^B\right) = \prod_{\mathbf{y}_i \in \hat{Y}^B} p\left(\mathbf{y}_i | \mathcal{M}_i^B\right) \quad (14)$$

These models allow us to compute appearance likelihoods $p\left(\mathbf{y}_i | x_i = F, \mathcal{M}^F\right)$ and $p\left(\mathbf{y}_i | x_i = B, \mathcal{M}^B\right)$ for each pixel in the test image.

Fig. 4 shows foreground $\hat{Y}^F$ and background $\hat{Y}^B$ appearance samples for an example test image, and the training images from which optimal foreground $\mathcal{M}^F$ and background appearance models $\mathcal{M}^B$ are drawn for this test image.



(a) (b) (c) (d)

Figure 4. Selecting an appearance model. (a) Target test image. (b) Foreground $\hat{Y}^F$ (red) and background $\hat{Y}^B$ (blue) appearance samples. (c-d) Training images from which optimal foreground $\mathcal{M}^F$ and background appearance models $\mathcal{M}^B$ are drawn for this test image.

4.4. Fusion of Geometry with Appearance

Given a novel image, geometric and appearance cues are fused by computing the posterior odds of the pixel labels:

$$\frac{p\left(x_i = F | \mathbf{y}_i, \mathcal{M}^F\right)}{p\left(x_i = B | \mathbf{y}_i, \mathcal{M}^B\right)} = \frac{p\left(\mathbf{y}_i | x_i = F, \mathcal{M}^F\right) p\left(x_i = F\right)}{p\left(\mathbf{y}_i | x_i = B, \mathcal{M}^B\right) p\left(x_i = B\right)} \quad (15)$$

An example is shown in Fig. 5.

4.5. Segmentation and Refinement

We segment road from non-road regions by thresholding the posterior odds. To optimize the threshold, we apply our appearance model selection method to each of the training images, using a dictionary drawn from the remaining training images, and compute the log odds of the posterior. Sweeping over a broad range of thresholds, we select the threshold that maximizes the average intersection-over-union (IOU) of the segmented road region with ground truth over all training images.

Fig. 6 (a) shows the results of this optimization. Note that the optimal empirical threshold is close to the theoretically optimal value of 0.

Since our appearance model lacks a smoothness term this initial segmentation contains small foreground fragments and holes. We therefore incorporated a refinement stage in which we filled all holes and removed segments below a threshold size, again optimizing the threshold by maximizing IOU of the detected foreground with ground truth over the training dataset. Fig. 6 (b) shows the results

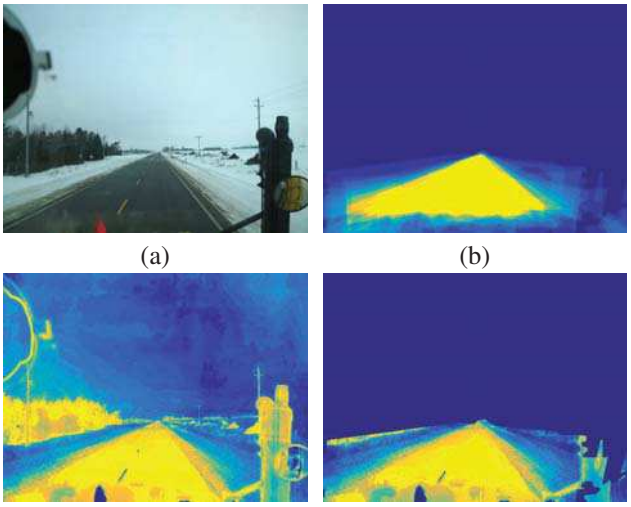


Figure 5. Fusion of geometric and appearance cues for an example test image (a). (b) Geometric prior ratio map $p(x_i = F)/p(x_i = B)$. (c) Appearance likelihood ratio map $p\left(\mathbf{y}_i|x_i = F, \mathcal{M}^F\right)/p\left(\mathbf{y}_i|x_i = B, \mathcal{M}^B\right)$ (d) Posterior map $p\left(x_i = F|\mathbf{y}_i, \mathcal{M}^F\right)/p\left(x_i = B|\mathbf{y}_i, \mathcal{M}^B\right)$

of this optimization.

4.6. Implementation Details

We implemented our method in MATLAB. To fit the GMMs, we sampled 5,000 pixels from both foreground and 5,000 pixels of each training image. Full-covariance GMMs were estimated using the EM algorithm (MATLAB's fitgmdist) with 10 replications per model, initializing with k-means, and employing a small regularization term (10^{-6}) to ensure covariance matrices were positive definite. Morphological close operations (MATLAB imclose) with progressively larger structuring elements were applied until all holes were eliminated. This was assessed by counting the number of foreground regions (MATLAB function bwconncomp) and comparing with the Euler number (MATLAB function bweuler).

All experiments were conducted on a 2.6 GHZ Pentium i7 with 16GB RAM. The geometric method of Almazen *et al.* [1] takes about 18 sec per image. Total run time for appearance model selection, fusion and refinement is roughly 3.7 sec per image.

5. Results

Fig. 7 compares the performance of our fusion method against the geometric method of Almazen *et al.* [1] on the test dataset, in terms of foreground IOU with ground truth. We find that fusing appearance cues with the geometric method of Almazen *et al.* [1] increases IOU by about 15% (from 43.8 to 50.3). Matched-sample t-tests between the

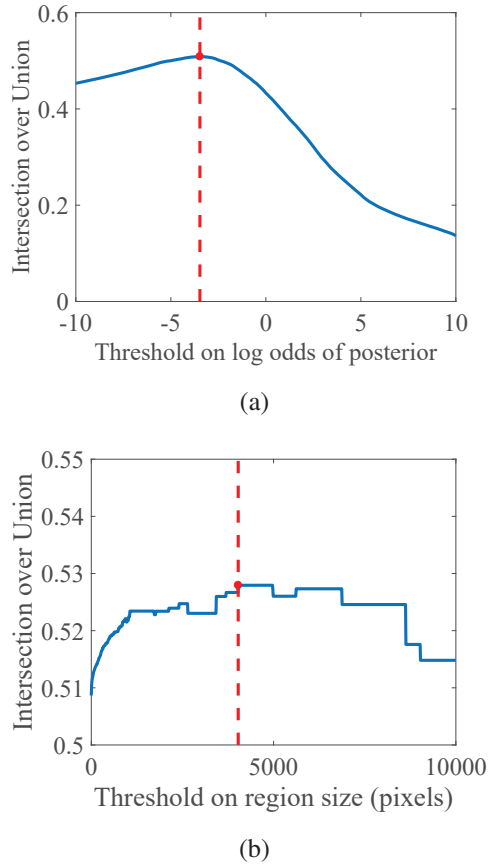


Figure 6. Optimizing segmentation and refinement. (a) Optimal threshold on log posterior odds. (b) Optimal threshold on region size.

mean IOU for our fusion method (refined) and the geometric method of Almazen et al [1] confirm that these improvements are statistically significant ($t(49) = 2.6, p = 0.012$).

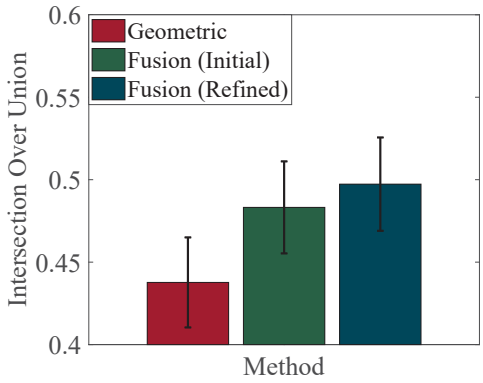
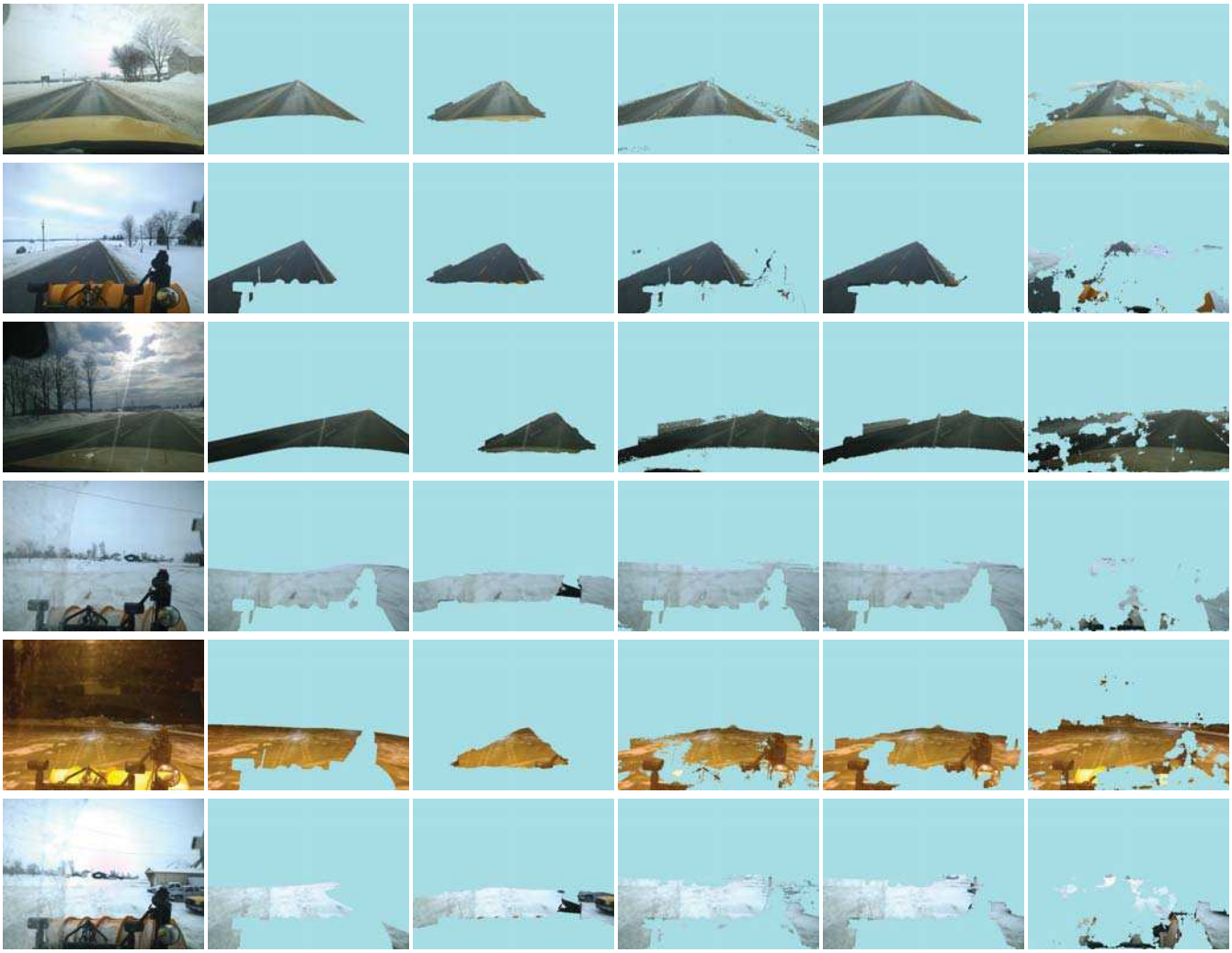


Figure 7. Mean performance on test set. Error bars indicate standard error of the mean.

Table 1 compares these results against SegNet [3]. We



(a) (b) (c) (d) (e)

Figure 8. Example results. (a) Test image (b) Ground truth segmentation. (c) Geometric method [1]. (d) Fusion method (initial). (e) Fusion method (refined). (f) SegNet method (road + pavement)

noticed that SegNet sometimes confuses the road and pavement categories and therefore assess both of these individually as well as their union. SegNet achieves an average IOU of only 29.3, performing best when road and pavement categories are combined.

Table 1. Quantitative comparison (IOU for foreground) with SegNet [3] on test set.

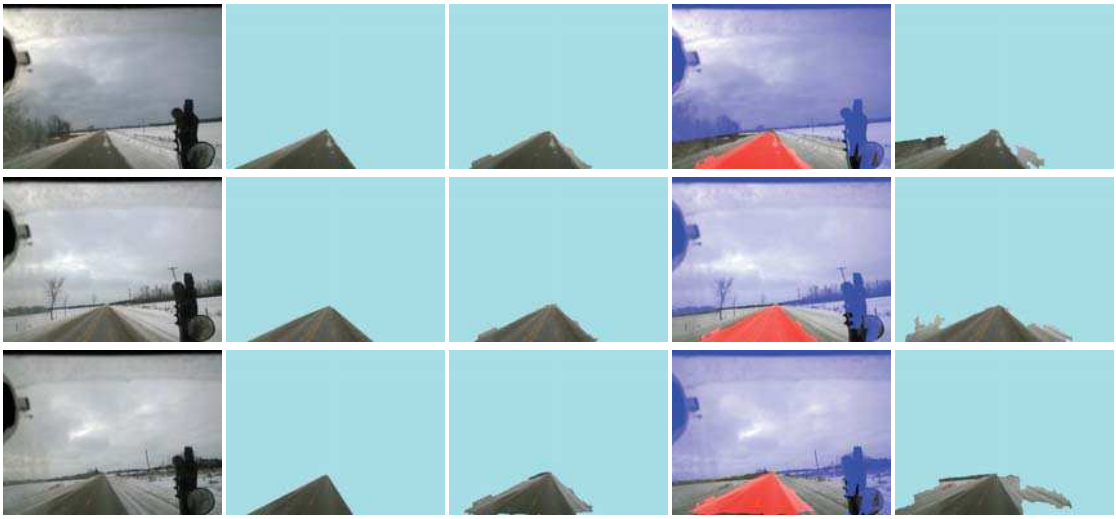
Method	Mean	Std. Err.
Geometric	43.8	2.7
Fusion (initial)	48.6	2.9
Fusion (refined)	**50.3**	**3.1**
SegNet (road)	28.9	3.5
SegNet (pavement)	10.0	2.1
SegNet (road + pavement)	29.3	2.7

Fig. 8 shows sample results from the test set.

6. Limitations & Future Work

SegNet [3] did not perform well on the Qian *et al.* test set [1]. Some of the errors may be due to lack of familiarity with wet and snowy ground conditions and occlusions by the vehicle, although it also commits errors where the road surface seems relatively clear (e.g., Fig. 8, second row). Its performance might be improved by fine-tuning on the Qian *et al.* training dataset; this is future work. However, it is still striking that our relatively simple geometric and fusion methods outperform SegNet by such a large margin.

Our fusion method fails to improve the geometric segmentation on some of the test images. Fig. 9 shows the three test cases where fusion causes the greatest *% decrease*



(a) (b) (c) (d) (e)

Figure 9. Example failure cases. (a) Input image (b) Ground truth segmentation. (c) Geometric segmentation [1]. (d) Foreground $\hat{Y}^F$ (red) and background $\hat{Y}^B$ (blue) appearance samples. (e) Fusion method (refined).

in foreground IOU relative to the geometric method. Generally these failures tend to be cases where the geometric method is already quite accurate, and the fusion method serves to broaden the foreground region to extend beyond the ground truth. In some cases the decline in IOU may be partly due to error in the manual ground truth segmentation in the Qian *et al*. dataset, as it is difficult to identify the exact right and left road boundaries under snowy conditions. However, some of these errors (e.g., the inclusion of some of the tree bottoms in the foreground region for the first failure mode) may be due to the simple colour appearance model we employ. Our framework can easily accommodate an extension of the appearance model to include texture cues that would help to discriminate between trees and the road surface; this is future work.

Since our current training sets consist of only 50 images each it is feasible to include a model from each of the training images in our foreground and background dictionaries. Scaling up to larger and more diverse datasets will require a little more thought in forming the dictionary if the algorithm is to remain efficient. One option is to simply cluster the models, representing each cluster by its centre. Another is to organize the models in a decision tree so that only a logarithmic number of models need be evaluated for a given test image.

References

[1] E. Almazen, Y. Qian, and J. Elder. Road segmentation for classification of road weather conditions. *Workshop on Computer Vision for Road Scene Understanding and Autonomous Driving, ECCV 2016*, pages 96–108, 2016. 1, 2, 5, 6, 7

[2] J. M. Alvarez, T. Gevers, and A. M. Lopez. 3D scene priors for road detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 57–64, 2010. 2

[3] V. Badrinarayanan, A. Kendall, and R. Cipolla. Segnet: A deep convolutional encoder-decoder architecture for scene segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PP(99):1–1, 2017. 2, 5, 6

[4] H. Dahlkamp, A. Kaehler, D. Stavens, S. Thrun, and G. Bradski. Self-supervised monocular road detection in desert terrain. In *Robotics: Science and Systems*, volume 38, Philadelphia, USA, 2006. 2, 3

[5] J. Fritsch, T. Kuehnl, and A. Geiger. A new performance measure and evaluation benchmark for road detection algorithms. In *International Conference on Intelligent Transportation Systems (ITSC)*, 2013. 2

[6] S. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on Gabor filters. *IEEE Transactions on Image Processing*, 11(10):1160–1167, 2002. 1

[7] H. Jegou, M. Douze, C. Schmid, and P. Perez. Aggregating local descriptors into a compact image representation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3304–3311, 2010. 1

[8] H. Jegou, F. Perronnin, M. Douze, J. Sanchez, P. Perez, and C. Schmid. Aggregating local image descriptors into compact codes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9):1704–1716, 2012. 1

[9] H. Kong, J. Y. Audibert, and J. Ponce. Vanishing point detection for road detection. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 96–103, 2009. 2

[10] H. Kong, J. Y. Audibert, and J. Ponce. General road detection from a single image. *IEEE Transactions on Image Processing*, 19(8):2211–2220, 2010. 2

[11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. 2

[12] J. Lee and C. D. Crane. Road following in an unstructured desert environment based on the EM (expectation-maximization) algorithm. In *SICE-ICASE International Joint Conference*, pages 2969–2974, 2006. 2, 3

[13] D. Levi, N. Garnett, and E. Fetaya. Stixelnet: A deep convolutional network for obstacle detection and road segmentation. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 109.1–109.12. BMVA Press, September 2015. 2

[14] F. Perronnin, J. Sanchez, and T. Mensink. Improving the Fisher kernel for large-scale image classification. In Daniilidis, K and Maragos, P and Paragios, N, editor, *European Conference on Computer Vision*, volume 6314 of *Lecture Notes in Computer Science*, pages 143–156, 2010. 1

[15] Y. Qian, E. J. Almazan, and J. H. Elder. Evaluating features and classifiers for road weather condition analysis. In *IEEE International Conference on Image Processing (ICIP)*, pages 4403–4407, 2016. 1, 2

[16] C. Rasmussen. Grouping dominant orientations for ill-structured road following. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 470–477, 2004. 2

[17] C. Rasmussen. Texture-based vanishing point voting for road shape estimation. In *British Machine Vision Conference*, pages 7.1–7.10, 2004. 2

[18] C. Rotaru, T. Graf, and J. Zhang. Color image segmentation in HSI space for automotive applications. *Journal of Real-Time Image Processing*, 3(4):311–322, 2008. 2, 3

[19] J. Sanchez, F. Perronnin, T. Mensink, and J. Verbeek. Image classification with the Fisher vector: Theory and practice. *International Journal of Computer Vision*, 105(3):222–245, 2013. 1

[20] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014. 2

[21] M. A. Sotelo, F. J. Rodriguez, and L. Magdalena. Virtuous: Vision-based road transportation for unmanned operation on urban-like scenarios. *IEEE Transactions on Intelligent Transportation Systems*, 5(2):69–83, 2004. 2, 3

[22] B. Southall and C. J. Taylor. Stochastic road shape estimation. In *IEEE International Conference on Computer Vision.*, volume 1, pages 205–212, 2001. 2

[23] C. Tan, T. Hong, T. Chang, and M. Shneier. Color model-based real-time learning for road following. In *IEEE Intelligent Transportation Systems Conference*, pages 939–944, 2006. 2, 3

[24] C. J. Taylor, J. Malik, and J. Weber. A real-time approach to stereopsis and lane-finding. In *Conference on Intelligent Vehicles*, pages 207–212, 1996. 2