EVALUATING FEATURES AND CLASSIFIERS FOR ROAD WEATHER CONDITION ANALYSIS

Yiming Qian, Emilio J. Almazan, James H. Elder

York University
Centre for Vision Research
4700 Keele St, Toronto, ON M3J 1P3

ABSTRACT

Weather-dependent road conditions are a major factor in many automobile incidents; computer vision algorithms for automatic classification of road conditions can thus be of great benefit. This paper presents a system for classification of road conditions using still-frames taken from an uncalibrated dashboard camera. The problem is challenging due to variability in camera placement, road layout, weather and illumination conditions. The system uses a prior distribution of road pixel locations learned from training data then fuses normalized luminance and texture features probabilistically to categorize the segmented road surface. We attain an accuracy of 80% for binary classification (bare vs. snow/ice-covered) and 68% for 3 classes (dry vs. wet vs. snow/ice-covered) on a challenging dataset, suggesting that a useful system may be viable.

***Index Terms*—** Image texture analysis, Classification algorithms, Probability, Feature extraction, Bayes methods

1. INTRODUCTION

The proportion of automobiles equipped with cameras continues to increase, providing opportunities for computer vision algorithms that can automatically extract information from the driving environment to inform the driver and adjust auto-assist functionality, thus reducing the risk of an accident. One factor that greatly affects driving safety is road weather conditions. While these conditions can potentially be inferred from tire slippage, this information can in some cases be slow and unreliable. Dash cams have the potential to provide this information ahead of time so that the driver and automatic control systems can react appropriately. This technology could also be deployed on road maintenance vehicles to provide a real-time log of roadway conditions, guiding the deployment of snowploughs and salt/sand spreaders. With some adaptation, the same technology could be deployed on highway and roadway camera networks that could then broadcast this information to drivers.

In this work we propose a system that categorizes road conditions using static images from a dashboard camera. We employ a dataset of 100 images which are provided and manually classified by our industry partner, who specialize in technologies for road maintenance. (This challenging dataset will be made available publicly for training and evaluation upon acceptance of this paper. We also intend to enlarge this dataset in the near future.) The Figure 1 shows some example images that illustrate several challenges: 1) Camera placement varies widely: the horizon may appear low or high in the image and may have significant roll. 2) The field of view is usually partially blocked by the hood of the vehicle or the bucket of the snowplough. 3) Road geometries are diverse, ranging over highway, urban intersections and parking lots. 4) Lighting conditions vary widely depending upon time of day, cloud cover and street lighting.

In this paper we show that despite these challenges, reasonably good results can be obtained through careful optimization of the region of interest (ROI), features and classifiers. We also use manual segmentation ground truth to quantify potential improvements that will result from improved segmentation algorithms.

This paper is organized as follows: Section 2 reviews prior work, Section 3 details our method for selecting the region of interest, Section 4 describes our classification algorithm, Section 5 reports results and finally Section 6 presents our conclusions and plans for future work.

2. PRIOR WORK

Much of the prior work on road condition classification has focused on the use of polarization and infrared cameras. While these cameras provide valuable information to distinguish different road conditions (e.g. dry, wet, snow, iced) [1, 2, 3, 4, 5, 6], they can be expensive and installation can be complex. In contrast, an algorithm that works with standard RGB imagery can potentially be deployed on existing automobile camera platforms shared by multiple applications.

A number of labs have recently been exploring the use of standard RGB cameras for road condition classification. Omer & Fu [7] used an SVM with RGB and gradient his-

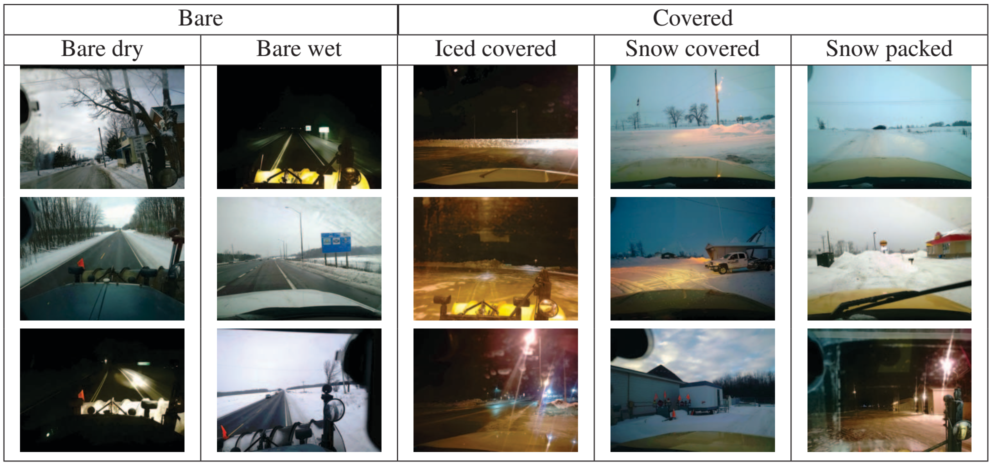


Fig. 1. Example images from training dataset

togram features to classify conditions as bare, covered or covered with bare tire tracks. However their approach required manual cropping of each image to extract the image region projecting from the ground surface, which is impractical for a real system. It is also unclear how well the unnormalized features employed would generalize over large changes in illumination. Kauai et al [8] used colour cues to add some degree of illumination invariance, however their approach depends on detecting white line markings to identify the road area, which will fail under snowy conditions or for roads that are poorly marked. In a very recent paper, Amthor et al. [9] proposed a spatiotemporal approach that integrates over many frames to detect specular reflections indicative of wet conditions. While their method improves over prior approaches, it requires integration over many frames, increasing computational load and delay. Also their system was evaluated on a very stable and carefully calibrated camera system, allowing them to use a hand-selected spatial integration region. This will not generalize well to cameras over changes in camera pose, different installations, or varying roadway structure.

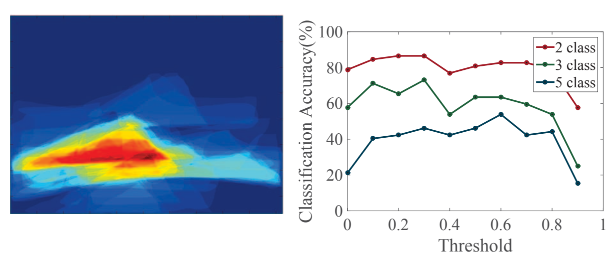
To work toward a more general system with wide application, we begin with a realistic dataset over which camera pose varies considerably and parts of the imagery are sometimes occluded by the vehicle itself (Fig. 1). The data was collected over a diversity of road structures including highways, urban roadways, intersections and parking lots, under diverse weather and illumination, including day and nighttime conditions. To operate reliably under these conditions, the region of interest must be carefully optimized and the system must be adaptive to varying illumination conditions. We detail such a system below.

2.1. Dataset

A challenging dataset of 100 images (Fig. 1) collected by our industrial partner was used to train and evaluate the performance of the algorithm. The dataset contains roads under different weather conditions, from bare dry to snow packed. The pictures were taken at different times of the day, thus covering a wide range of illumination conditions and the camera pose varied considerably. The road condition class was identified manually by our industrial partner. To further enrich the dataset, we manually identified a region of interest for each image, consisting of the pixels projecting from the road (or parking lot) surface. We randomly split the dataset into 50 training images and 50 test images.

3. REGION OF INTEREST

We used the ground truth region of interest labelling for the training data to learn a spatial prior for road pixel locations (Fig. 2(a)). Ultimately, this prior should be combined with geometric cues in the image to adapt the ROI to each camera pose and roadway. For this paper we employ a fixed ROI over all images, derived automatically by optimizing a threshold on the prior.



(a) (b)

Fig. 2. (a) Prior spatial distribution of road pixels. (c-d) Threshold optimization for 2, 3 and 5 Classes.

If we first renormalize the ROI prior to assume a maximum value of 1, varying the threshold from 0 to 1 sweeps out a precision-recall curve for pixel classification as a road/non-road. Threshold could be selected by maximizing some figure of merit on this curve (e.g., F-score), however it is unclear how this will translate to road condition classification. We therefore identified the optimal threshold by maximizing the accuracy of our classification algorithm, detailed in the next section, on the training data. Based on this analysis (Fig. 2(b)) we selected a threshold of 0.3 for the 2- and 3-class conditions and 0.6 for the 5-class condition. Fig. 3 shows some examples of the resulting ROIs using a threshold of 0.3.

(a) (b) (c)

Fig. 3. Example road segmentation output

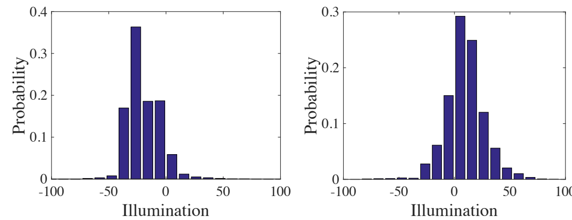
4. ROAD CONDITION CLASSIFICATION

Once an ROI is identified, the problem reduces to 1) computing meaningful features from the ROI pixels and 2) using these features to classify the road weather conditions. Since viewpoint and illumination conditions vary widely, a key to success is to make these features relatively invariant to these parameters.

4.1. Feature Extraction

Varma & Zisserman [10] sought to achieve exactly this objective with their introduction of their MR8 filter. So-named because it derives an orientation-invariant description of the local structure of a texture by selecting the Maximum over the Responses of many oriented filters to generate a concise 8-dimensional feature vector representing structure at multiple scales and phases.It allows proposed algorithm to process image without requiring any priori knowledge about the viewing angle of the dashboard camera. Invariance to illumination is achieved by pre-normalizing imagery and by normalizing local filter responses which could exact the reflectance of the pavement, specularities caused by water or snow. Also due to the small feature sizes, it allows the classification algorithm to perform at real time. These filtered images are segmented into 8 by 8 non-overlap patches and those patches are clustered using K-means to build a codebook, and each image is then represented as a edge histogram. We measured 3-way classification performance on the training dataset as a function of the number k of histograms to determine an optimal codebook size of $k = 74$.

While the wide variation in illumination conditions renders absolute intensity unreliable for road condition classification, intensity of the road pixels *relative* to the rest of the image can be informative, particularly for identifying snow, which is relatively bright. To take advantage of this information, we augment the luminance-invariant MR8 edge histogram with a 20-bin histogram of pixel deviation from the mean image luminance. Fig. 4(a-b) gives examples - notice that dry roads tend to be darker relative to the rest of the image, and snow-covered roads tend to be lighter.



(a) (b)

Fig. 4. Example of luminance of road pixels relative to mean image luminance. (a) Dry road. (b) Snow/ice-covered road.

The result of feature extraction is thus two feature vectors for each image ROI: a 74-dimensional texton vector $\boldsymbol{t}$ and a 20-dimensional relative luminance vector $\boldsymbol{l}$. These features then form the input to the classifier described below.

4.2. Classification

To construct the classifier we first computed the mean class-conditional edge histogram and luminance histograms $\bar{\boldsymbol{t}}_{\boldsymbol{m}}$ and $\bar{\boldsymbol{l}}_{\boldsymbol{m}}$ for each class m. We then computed the exponential distance [11] χ^2 (equation 1) between target image edge histogram h_t and $\bar{\boldsymbol{t}}_{\boldsymbol{m}}$ as well as luminance histograms h_l and $\bar{\boldsymbol{l}}_{\boldsymbol{m}}$. Compare to regular χ^2 calculation, the exponential term effectively projects the input vectors into an infinite dimensional space, where they become linearly separable.

$$\chi^2(x,y) = e^{-\frac{1}{2}\sum_{k=1}^{d}\frac{[x_k-y_k]^2}{x_k+y_k}} \tag{1}$$

The edge histogram and luminance χ^2 distances were concatenated to form a feature vector of the image (equation 2).

$$\chi^2 = \{\chi^2_{t,1}, \dots\chi^2_{t,m}, \chi^2_{l,1}, \dots\chi^2_{l,m}\} \tag{2}$$

The mean μ and standard deviations σ of the χ^2 distance were calculated to compute class-conditional univariate normal models P for each observation Y.

$$P(X = \chi^2|Y) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(\chi^2-\mu)^2}{2\sigma^2}} \tag{3}$$

Finally, assuming class-conditional independence between texton and luminance features, we formed the likelihood of each class $\hat{P}(Y_k|X_1 \dots X_n)$ as the product of the texton and luminance likelihoods $P(X_i|Y_k)$, and multiplied by the prior $P(Y_k)$ over classes, learned from the training data. The posterior probabilities was calculated in following equation:

$$\hat{P}(Y_k|X_1 \dots X_n) = \frac{P(Y_k)\prod_{i=1}^{P}P(X_i|Y_k)}{\sum_{k=1}^{K}P(Y_k)\prod_{i=1}^{P}P(X_i|Y_k)} \tag{4}$$

Where:

Y_k is the random variable corresponding to the k^{th} class of an observation

$X_1 ... X_n$ are random predictor of an observation

$P(Y_k)$ is the prior probability of k^{th} class

The classification of an observation was defined by estimating the posterior for each class, and then assigns the observation to the class with the maximum posterior probability.

5. PERFORMANCE EVALUATION

5.1. Results and Discussion

Classification results for the two classes (bare vs snow/ice covered) , three classes (dry, wet, snow/ice covered) and five classes (dry, wet, snow, ice, packed) are shown in Fig. 5. We evaluate results for road segmentations computed using our prior segmentation algorithm as well as manual road segmentations, the latter providing an estimate of the potential for our classification algorithm as segmentation algorithms improve. We obtain mean accuracies of 80%/98% for prior and manual segmentation in two-class segmentation, 68%/88% for prior and manual segmentation in three-class segmentation, and 46%/62% for prior and manual segmentation in five-class segmentation. Given the challenging nature of the dataset, these are promising results. From these numbers we can also see that better performance will depend upon improvements to both segmentation and classification stages.

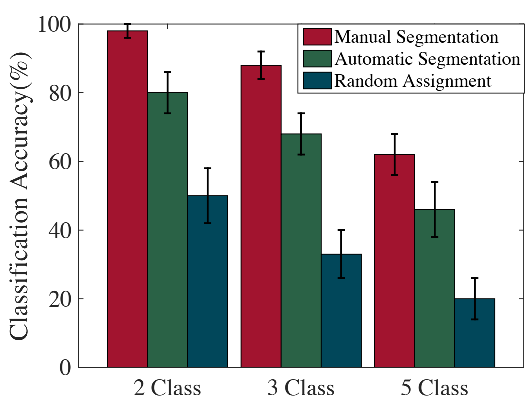


Fig. 5. Results comparison for 2, 3 and 5 classification using mannual, prior segmentation against random guesses

Our algorithm was compared with dense SURF feature [12] and different pooling methods (Improved Fisher Vector[13] and Bag of Visual Words[14], Textons[10]) to measure the effectiveness of algorithm. The SURF feature was extracted in 4 different scales (1.6,3.2,4.8,6.4). The Improved Fisher Vector algorithm was implemented using VLfeat Library [15] to construct a 4096 dimensional feature vector from MR8 and SURF features. The Bag of Visual Words algorithm was implemented using Matlab built-in functions to construct a 74-bin histogram. An 18-Texton dictionary was constructed from the training data and it was used to construct Texton histogram for classification. The optimal size of the Fisher Vector, Bag of Visual Words and Texton dictionary were found by evaluating the classification accuracy on training sets. A few combinations of feature extraction and classification algorithms were tested. (Table 1).

Classification Methods	Features		
	MR8		
	FV	BoW	Texton
Nearest Neighbor	88%(76%)	86%(48%)	78%(50%)
Naive Bayes Boost	26%(26%)	**88%(76%)**	82%(56%)
Linear SVM	84%(76%)	84%(68%)	82%(54%)
RBF SVM	86%(74%)	82%(72%)	82%(54%)
AdaBoost	80%(82%)	84%(76%)	82%(60%)
Decision Tree	74%(69%)	82%(62%)	78%(48%)
	SURF		
	FV	BoW	
Nearest Neighbor	82%(62%)	80%(60%)	
Naive Bayes Boost	28%(26%)	82%(62%)	
Linear SVM	82%(72%)	80%(66%)	
RBF SVM	84%(72%)	80%(66%)	
AdaBoost	78%(80%)	84%(56%)	
Decision Tree	70%(60%)	80%(62%)	

Table 1. Comparison of different classifiers, features and pooling methods in three-class classification on manual segmentation, the numbers shown are the classification accuracies with luminance and without luminance(in the bracket)

6. SUMMARY AND FUTURE WORK

In this paper we have proposed a novel algorithm to classify roads according to their conditions. The algorithm was designed to operate over diverse camera poses, road geometries and conditions. The approach consists of road ROI segmentation and surface classification. To handle the high variability of road layouts and camera pose, we learned the prior information from the training dataset to determine the approximate location of the projection of the road surface in the image. To mitigate problems induced by variable weather and lighting conditions, our classifier was based on intensity-normalized luminance and texture features. Classification performance on a challenging dataset was 80%, 68% and 46% correct for two/three/five-classes respectively. Given the challenging nature of the dataset, these are quite promising results. Our analysis reveals that improvements in performance can be achieve by improving both the segmentation and classification stages. For the former, we are currently developing methods to estimate horizon, road vanish point, and left/right road boundaries, allowing less reliance on the spatial prior. For the latter, we intend to study more powerful classifier frameworks (e.g., convolutional nets), operating on intensity-normalized input.

7. REFERENCES

[1] Sung-Han Lim, Seung-Ki Ryu, and Yeo-Hwan Yoon, "Image recognition of road surface conditions using polarization and wavelet transform," *Journal of The Korean Society of Civil Engineers*, vol. 27, no. 4D, pp. 471–477, 2007.

[2] Hun-Jun Yang, Hyeok Jang, Jong-Wook Kang, and Dong-Seok Jeong, "Classification algorithm for road surface condition," *IJCSNS*, vol. 14, no. 1, pp. 1, 2014.

[3] Maria Jokela, Matti Kutila, and Long Le, "Road condition monitoring system based on a stereo camera," in *Intelligent Computer Communication and Processing, 2009. ICCP 2009. IEEE 5th International Conference on*. IEEE, 2009, pp. 423–428.

[4] Johan Casselgren, *Road surface classification using near infrared spectroscopy*, Ph.D. thesis.

[5] Patrik Jonsson, "Remote sensor for winter road surface status detection," in *Sensors, 2011 IEEE*. IEEE, 2011, pp. 1285–1288.

[6] Patrik Jonsson, Johan Casselgren, and Benny Thornberg, "Road surface status classification using spectral analysis of nir camera images," *Sensors Journal, IEEE*, vol. 15, no. 3, pp. 1641–1656, 2015.

[7] Raqib Omer and Liping Fu, "An automatic image recognition system for winter road surface condition classification," in *Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*. IEEE, 2010, pp. 1375–1379.

[8] Shigeaki Kawai, Ken Takeuchi, Kenji Shibata, and Yuukou Horita, "A method to distinguish road surface conditions for car-mounted camera images at nighttime," in *ITS Telecommunications (ITST), 2012 12th International Conference on*. IEEE, 2012, pp. 668–672.

[9] Manuel Amthor, Bernd Hartmann, and Joachim Denzler, "Road condition estimation based on spatio-temporal reflection models," in *German Conference on Pattern Recognition (GCPR), 2015 37th German Conference on*, pp. 3–15. Springer, 2015.

[10] M. Varma and A. Zisserman, "Classifying images of materials: Achieving viewpoint and illumination independence," in *Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark*. May 2002, vol. 3, pp. 255–271, Springer-Verlag.

[11] Sreekanth Vempati, Andrea Vedaldi, Andrew Zisserman, and CV Jawahar, "Generalized rbf feature maps for efficient detection.," in *BMVC*, 2010, pp. 1–11.

[12] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, "Surf: Speeded up robust features," in *Computer vision–ECCV 2006*, pp. 404–417. Springer, 2006.

[13] Florent Perronnin, Jorge Sánchez, and Thomas Mensink, "Improving the fisher kernel for large-scale image classification," in *Computer Vision–ECCV 2010*, pp. 143–156. Springer, 2010.

[14] Gabriella Csurka, Christopher Dance, Lixin Fan, Jutta Willamowski, and Cédric Bray, "Visual categorization with bags of keypoints," in *Workshop on statistical learning in computer vision, ECCV*. Prague, 2004, vol. 1, pp. 1–2.

[15] A. Vedaldi and B. Fulkerson, "VLFeat: An open and portable library of computer vision algorithms," `http://www.vlfeat.org/`, 2008.