

Big-video Mining of Road Appearances in Full Spectrums of Weather and Illuminations

Guo Cheng, Zheyuan Wang

Department of Computer Science
Indiana University Purdue University Indianapolis
Indianapolis, USA
guocheng@iupui.edu, zheywang@iupui.edu

Jiang Yu Zheng

Department of Computer Science, and
Transportation Active Safety Institute, IUPUI
Indianapolis, USA
jzheng@iupui.edu

Abstract—Autonomous and safety driving require the control of vehicles within roads. Compared to lane mark tracking, road edge detection is more difficult because of the large variation in road and off-road materials and the influence from weather and illuminations. This work investigates visual appearances of roads under a spectrum of weather conditions. We use big-data mining on large scale naturalistic driving videos taken over a year through four seasons. Large video volumes are condensed to compact *road profile* images for analysis. Clusters are extracted from all samples with unsupervised learning. Typical views of a spectrum of weather/illuminations are generated from the clusters. Further, by changing the number of clusters we find a stable number for clustering. The learned data are used to classify driving videos into typical illumination types briefly. The surveyed data can also be used in the development of road edge detection algorithm and system as well as their testing.

Keywords — *Driving video; Weather and Illumination; Clustering; Visual Appearance; Road Departure; Autonomous Vehicle; ADAS;*

I. INTRODUCTION

Current lane mark following has been matured and functional for autonomous vehicles on roads [7]. However, there are still many roads unpaved perfectly and unpainted with lane marks in the world. Road edge detection is a fundamental capability for drivers before modern road infrastructure is built. How to understand road environments during driving? To implement this task, vision sensing is the default tool [3]. By using cameras in road detection, weather and illumination are critical factors that influence the road appearances. We study all appearances of road and off-road under various illuminations for the algorithm/system design and real road construction with apparent features. The original contributions in this paper are as follows: (1) Observing driving videos, we analyze the physics properties of road qualitatively and categorize them in visual appearances related to typical weather and illuminations. The combined weather and illumination factors allow us to handle a reasonable number of situations. (2) Quantitatively, we use a large scale Naturalistic Driving Datasets [10] to model roads, instead of doing simulation using optical equitation [2]. The videos were obtained from a hundred cars running on various roads through different seasons, weathers, and illuminations. (3) We use data mining method to survey the data and find distributions of visual properties. Unsupervised learning by K-means algorithm generates clusters in feature space and a hierarchy of clusters.

The related works on weather identification using in-

vehicle cameras are classification using color HIS and gradient features [11]. Raindrops were detected as a clue of raining [12]. K-NN is used in classifying raining weather [6]. Our work is an attempt to deal all possible weather and illumination conditions in a general framework of data mining. Therefore, clustering method is applied to a huge driving video data set to understand various types of weathers in road environments.

In order to tackle the data mining of large video data size, we first convert a video to a compact image format called *road profile* [1], and performs visual analysis, data mining, and visualization at road and weather sensitive regions in the video. Although the data size is reduced tremendously, the information is extracted and stored for later analysis. This allows us to extend our data analysis to 2TB video dataset, roughly 7500 5-minute video with 9k frames each [10].

The results of this big-data mining answer the critical questions such as how many illumination types frequently appear. Under which illuminations, the road can be detected robustly, barely visible but can be figured out with other clues, or even invisible such that driving has to rely on other navigation tools. For visible road edges, what types of features or their combinations are most stable. For invisible road edges, lane mark painting can be recommended to ensure safe driving. This analysis based on whole spectrum of weather and illumination can reduce the ambiguity in heuristic algorithms of road edge detection, and find missing cases in the learning algorithms trained and tested with partial data. This can also provide real testing scenarios for road edge detection systems.

In the following, we introduce weather and illumination categorization related to visual perception in Sec. II. Sec. III describes the effort to reduce videos to road profile images and illumination sensitive data. Sec. IV applies K-mean algorithm to find major clusters related to weather and illuminations, and the classification of video frames to validate the clustering.

II. WEATHER AND ILLUMINATION CATEGORIZATION

A. Physics based Vision and Weather Categories

Although various road detection algorithms have been developed, they are lacking of physical functions dealing changes in weather and illuminations. In principle, the road appearance is determined from surface reflectance on both road and off-road materials, environment illuminations, and camera sensitivity as shown in Fig. 1. In addition, wet ground creates specular reflection component that may be much stronger than normal diffused surface reflection. Trees and buildings cast shadows in which ambient light is the illumination different

Research partly supported by TOYOTA CSRC, Ann Arbor, MI, USA.

from the sunlight. Sky may full of clouds far away but close road may still be in sunshine. All these variations make the road edge detection unstable because they depend on heuristic clues while scenes keep changing during driving. The simplest reflection model for qualitative analysis of appearances is

$$I(x) = c(d M(X) \bullet L(X) + s S(X) L(X)) \quad (1)$$

where $I(x)$ is the image color at x mapped from road position X , $M(X)$ and $S(X)$ are the diffused and specular surface reflectance, and $L(X)$ is the illumination component. The value c is the camera sensitivity and exposure parameter.

Table I Road, off-road, season, weather, and illumination conditions

Road Materials	Off-road Materials	Seasons	Weather	Illumination	Camera sensitivity
Asphalt new, old, repaired	Grass, Soil/dirt, Gravel, Vegetation/field,	Spring	Rainy/wet	Specular reflection	Dirty windshield
		Summer	Heavy rain		
Concrete	Concrete/curb, Cliff/ditch,	Fall	Snowing/snow-covered,	Dark lit	Normal
		Winter		Direct lighting	auto-exposure
Gravel	Tree/forest, Construction cone,		Sunny,	Shadow	
Soil/dirt	Guardrail/barrier		Cloudy,	Night	
	Other vehicles		Fog		



Figure 1 Various road appearance in driving videos.

The camera captured color is the integration of a variety of factors. However, the combination of all these factors (columns in Table I) creates a large number of cases, even if not every case is concurrent. To handle a small number of cases in real world, we integrate all the columns into three factors as in Table II, where each one has about 10 categories. Among three factors, the weather/illumination is the most significant one that influences the road appearance in our observation. Based on the reflection principle, detailed visual appearances of roads are described accordingly in Table III. These categories are human defined according to the optical principle and have ambiguity when videos are tagged by human subjectively.

Table II Integrated categories for road visual appearances

On-road materials	Off-road materials	Weather/illumination
Asphalt new	Grass green	Sunny – facing sun
Asphalt old	Grass yellow/gray	Sunny – back to sun
Asphalt repaired	Vegetation green	Shadow
Concrete	Vege yellow/brown, soil/dirt	Cloudy
Gravel	Tree/forest green	Dark lit
Wet	Tree/forest brown	Night
Snow covered	Gravel	Direct light
	Concrete/curb	Snowing
	Ditch/cliff	Raining
	Snow-covered	Fog
	Guardrail, barrier, cones, vehicle	Dirty windshield

Table III Description of weather and illumination categories. () num of videos

W & I	Qualitative description of appearance
Sunny back to sun (93)	Including the sun at side and back direction, less shadow, rich color
Cloudy (83)	Including overcast and partly cloudy
Sunny face sun (21)	Sun in the forward direction, may have specular reflection on asphalt road surface, large shadow areas on roadside objects
Shadow (26)	Shadow on road surface casted from trees, etc.
Dark lit (43)	At dusk or sun set, sky is still bright but road and scenes are dark without being sufficiently illuminated. No color visible on road. Entire view becomes dark with camera auto-exposure.
Rainy (20)	Dark illumination, low sensitivity of camera. Some cannot be separated from overcast. We look at wiper on/off to decide rainy state. Wet surface has specular reflection of scenes and sky.
Direct light (13)	Sun at front with glare in image. Even hard for human driver to see road. May have highlight on road surface.
Snowy (1)	Snow intervene the scene and camera, makes image noisy. Road sides are more snow-covered than road.
Fog (5)	A half-transparent layer in front of camera blurs scenes. Camera exposure can enhance contrast.
Night (55)	Night is tagged when vehicle headlights are on. It includes rural night without street light, city with street light, and wet ground with strong reflection of other car lights and city lights



Figure 2 Four sampling regions in video frame for weather clustering.

B. Visual Features in Video That Reflect Weather

For road edge detection, several types of visual features are particularly important, which include the linearity describing road shape extending forward, surface homogeneity from road surface pavement, and color determined by road and off-road materials, i.e., surface reflectance. Four regions are selected from video to examine the influence from weather and illumination. They are indicated in Fig. 2 as a sky area, a road surface area, and two roadside regions. After the horizon is calibrated in video, we locate a horizontal line, l , with a fixed distance from horizon to capture the road surface 10~15m ahead the vehicle. This is a distance suitable for road departure alert on road. The lane width is calibrated from this setting.

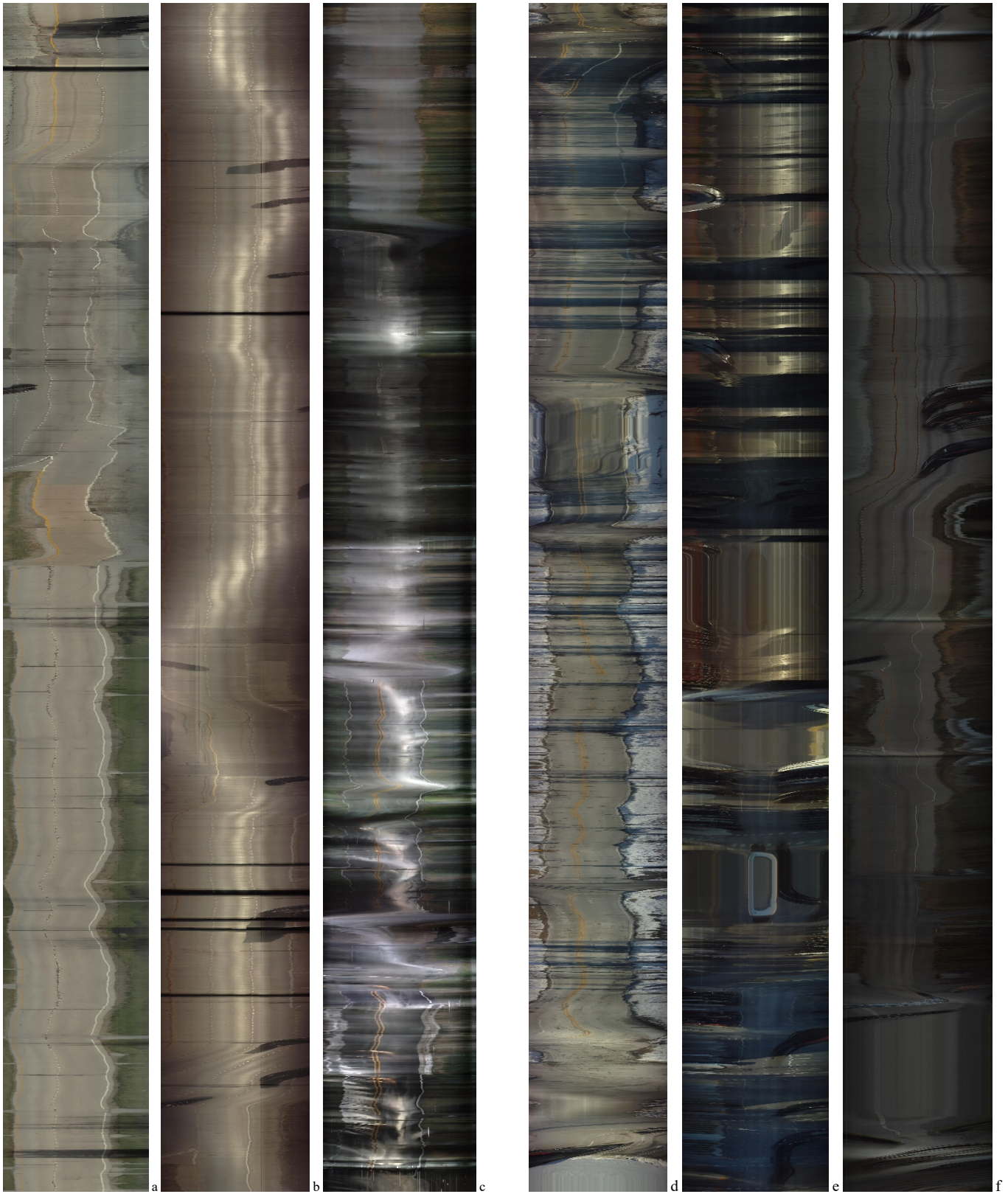


Figure 3 Road profiles captured different weather and illuminations. The time axis is upward. (a) Sunny back to the sun, (b) Sunny facing the sun with highlight on road, (c) Rainy on wet road with specular reflection of sky. Specular reflection is more dominant than road edges. (d) snow-covered road with partial shadow, (e) heavy shadow on urban road with road partly visible, and (f) raining day. Road edge is barely visible.

Side lanes or road sides are covered by roadside regions. The road surface sampled with I is a little wider than a lane width.

Rather than taking the whole image full of dynamic objects into account [6], we sample the image regions that are less

influenced from the color of other vehicles on road. A sky area centered at high position is sampled to avoid high buildings and trees. Two roadside areas between the horizon and sampling line are set to cover large depth, but has a gap from the driving lane so as to avoid inaccurate edge positions that change frequently in driving video, which reduces the color influenced by a special building or sign. The left region has more possibility to cover opposite road. We compute the average color intensity, chroma, and variation in these regions. Such a setting provides more reliable information for understanding weather than a single region in video, and avoid front vehicles and objects being captured in the evaluation.

III. DATA REDUCTION OF VIDEO FOR VIDEO MINING

This work takes a data mining approach to explore visual appearance of roads rather than estimating exact physical parameters of roads using reflection theorem because of the large variations of roads affected by changeable weather. To facilitate the data mining of videos, we first convert long driving videos to congregate images for visualization. From each frame (Fig. 2), we sample horizontal line l and copy its pixels into another image. Sampled pixel lines from consecutive frames are piled along the time axis such that a *road profile* image is created [1] as in Figures 3. A five-minute video yields the road profile of 1280×9000 pixels, which reduces video to 1/720th in size but keeps all the road and roadside colors in about 4 lane width.

In the road profiles, lanes are waved due to vehicle waving in roll during the driving on uneven road. Lane marks are the most distinct sign to identify roads. In road edge investigation, we intentionally ignore lane marks by median filtering pixels of that width and looking at global regions. It is obvious that the largely changed weather and illumination make the road more different in appearance than road materials. Poor illuminations can loss the visibility of road (e.g., direct light) and edges (Fig. 3c, f). For the video captured in rainy days, either wiper movement is captured as horizontal lines in the motion profile, or raindrop forms unique texture in the road profile. We treat these phenomena from data clustering point of view for all kinds of weathers.

IV. CLUSTERING DIFFERENT WEATHER AND ILLUMINATION

A. Features Reflecting Weather and Illumination

This work uses the unsupervised learning to cluster visual features in different weather and illuminations. 360 video clips with 324k locations are collected evenly according to Table III. We find clusters by K-mean algorithm including their centroids, variations, and inter-cluster distances. The features used are average intensity, chroma, and variation from four regions of sky, road, left and right roadsides. We do not use *Hue* as feature since it is more related to the roadside materials other than weather. We use *Chroma* instead of *Saturation* because saturation has a singular value when intensity is zero.

The sampled videos are manually tagged in 10 categories C_i at clip level as indicated in Table III for verifying the consistency to the K-means output S_j , $j=1\sim 10$. Certainly, human tagged weather and illumination categories are

ambiguous in many cases and have large overlap between categories in the feature space. We visualize average image intensity and chroma in three regions (sky, left, right roadsides) as triangles for all categories as in Fig. 4. We can notice the differences between categories in parameters. *Night* is the easiest category to be separated from others.

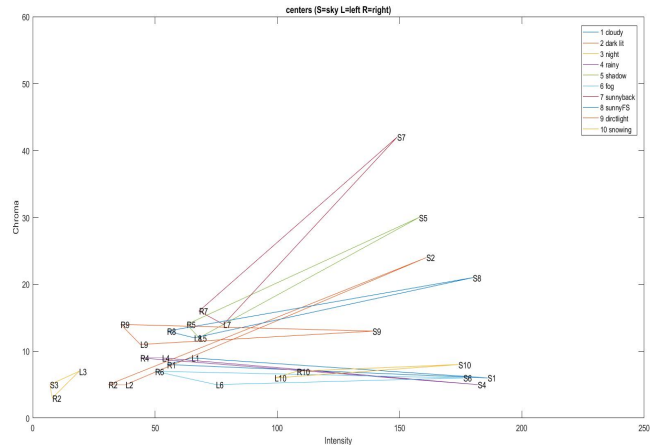


Fig. 4 Visualization of weather/illumination categories tagged by humans in intensity (horizontal) and chroma (vertical) space. Triangles indicate those parameters from sky, left and right roadside regions. The intensity and chroma of sky is more significant than left and right roadsides. We can also notice the similarity among fog, cloudy, rainy, and snowy.

By looking at category C_i as a representative image with averaged features in four regions, we can generate typical views of all categories in Fig. 5. Because manual tagging of *shadow* is at clip level other than detailed to frames, it contains many *sunny* frames. The *shadow* average is like that of *sunny back to the sun*. Raining and cloudy are much similar in nature in human tagging, because the wiper movement monitored above is ignored in averaging colors in the road profiles. This tagging may also miss some small sets of weathers like urban night or even with wet ground.

B. K-means Clustering of Weather and Illumination

Because human tagging is intuitive and qualitative at clip level, we further apply K-means algorithm on all 324k frames to find quantitative results of weather and illumination clusters. K-means is selected because the weather properties are of continuous nature rather than discrete samples. We also calculate likelihood probability $P(S_j|C_i)$ to find out which human tagged categories contribute to each cluster the most. The dominant category is used to name the cluster as a *related name*. For a cluster without any major category contributing to it, it is considered as an outlier. K-mean clusters are clear and non-overlapped partition of the feature space, but are not necessarily consistent to the human tagged categories exactly. Moreover, we compare the clusters (Fig. 6) to the human tagged categories in Fig. 4 with triangles from average image intensity and chroma in three regions.

To examine the cluster structure in various levels of details, we change the number of clusters $K=10, 9, \dots, 2$ to obtain different clusters S_j , $j=1, \dots, K$. By comparing $P(S_j)$, a K-mean tree is generated as in Fig. 7. This tree tells a way to summarize weather and illumination clusters from coarse to

fine levels using their *related names* at each K . Night is always the most distinct cluster to separate from daytime. Figure 8 is the result when cluster number is set to 7. On the other hand, we have also increased K in clustering. The raining and cloudy are unable to be separated up to $K=30$. This indicates that rainy and cloudy are unable to be separated clearly if observation is not down to the level of raindrop [11].

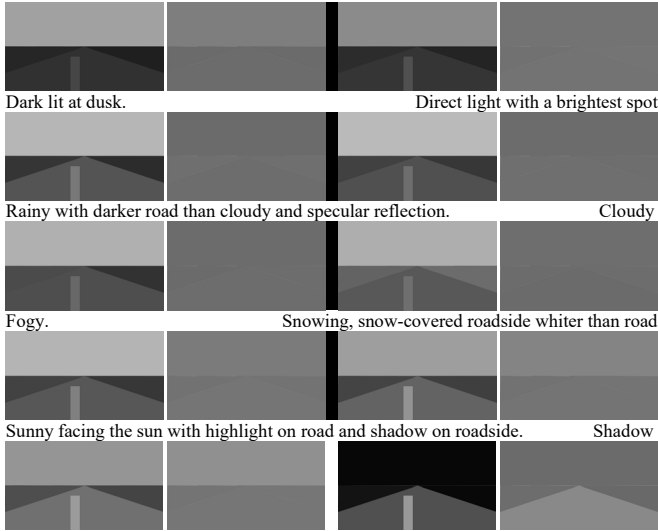


Fig. 5 Intensity and chroma average pairs in human tagged categories. The central vertical bars on road indicate the color variation against surface due to specular reflection, shadow, surface repairing, etc. Compare to the intensity, chroma averages are much smaller (displayed with 10% scale up).

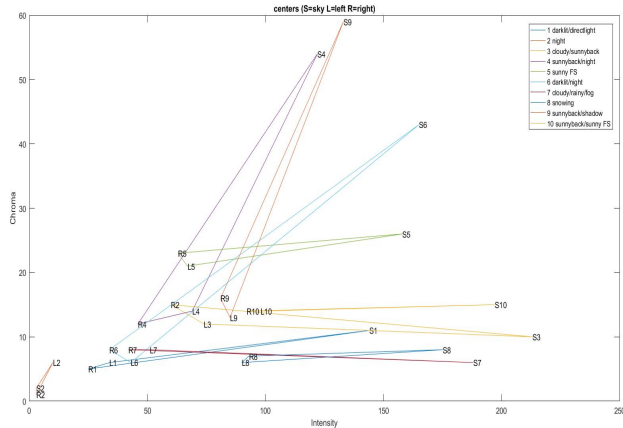


Fig. 6 Visualization of weather/illumination clusters ($K=10$) from K-means algorithm with intensity (horizontal) and chroma (vertical) space. It has a sparser inter-distance between weather clusters.

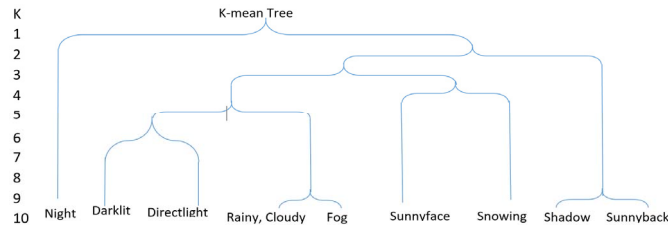


Fig. 7 K-means tree containing a hierarchy of clusters for different K , obtained from analysis of $P(S|C)$ for each K .

If a robust K is to be selected, $K=7$ is a proper choice, since other $K > 7$ generate the same number of clusters as human

tagged except some difference in outliers, when we compute K from 1 to 30. The typical views of the clusters for $K=10, 7, 3$ are also visualized in Fig. 9 by using centroid values of K-mean clusters; their values are consistent across different K . For $K=3$, views are separated to *night*, *bright*, and *dark* daytime, in which *dark road* is researched in road edge detection recently [5,9]. With a coarse clustering, the intra-class distances become larger, meaning such classes have larger variations.

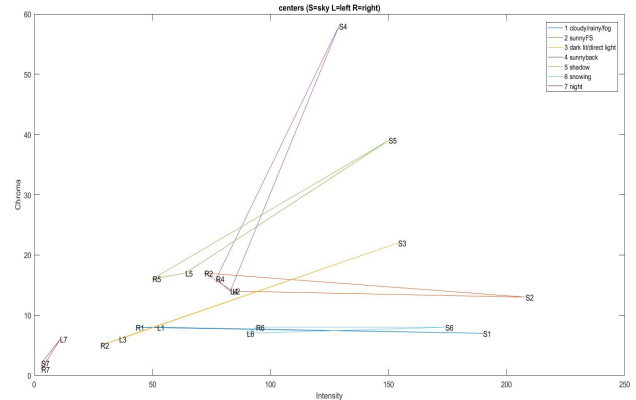


Figure 8 Visualization of weather/illumination clusters ($K=7$) from K-means with intensity (horizontal) and chroma (vertical) space.

C. Classification of frames using learned clusters

The ultimate goal of weather/illumination clustering is the recognition of weather in video frames for the guidance of road edge detection. Based on learned weather clusters, we classify another larger set of video containing 300 clips. The same regions and parameters are extracted from each frame. The distances of a frame to previously obtained cluster centroids are computed in the feature space to find the closest cluster. Note that K-means is a different division of feature space based on equal distance, while human tagging is more qualitative on visual properties even if the measure is imprecise.

V. EXPERIMENTS AND DISCUSSION

The experiments are carried out using Naturalist Driving Data [10] taken from 110 vehicles on various types of roads (highway, urban, rural, residential roads, and so on) for over one year through different seasons, weathers and times. All cameras on vehicles are the same type so that they have the same response to the external brightness. The generated video data reaches 50TB; a subset of 2TB video is selected evenly per date and time in this study. Naturalistic driving videos with the resolution of 1280×720 pixels are taken by the same type of HD cameras at 30 frames per second with auto-exposure function. The horizon in all clips are manually determined and road profiles are generated from them. Each clip has 5-minute length (9k frames or positions).

K-mean algorithm on Matlab starts from the averages of human tagged categories as seeds. The color distributions of road and roadsides obtained from the clusters are then used for better road edge detection method [8]. We also feed all 5800 samples into the K-mean algorithm and compare them with the results from 360 training video clips. Figure 10 shows the chroma views for $K=10$ clusters along with intensity views.

The left roadside is brighter than right roadside, because left regions in the frames may include opposite road surface, barrier, central divider, etc. The right-side difference is thus more reliable. In night, the road surface lit by headlight or street light has high chroma from light source.

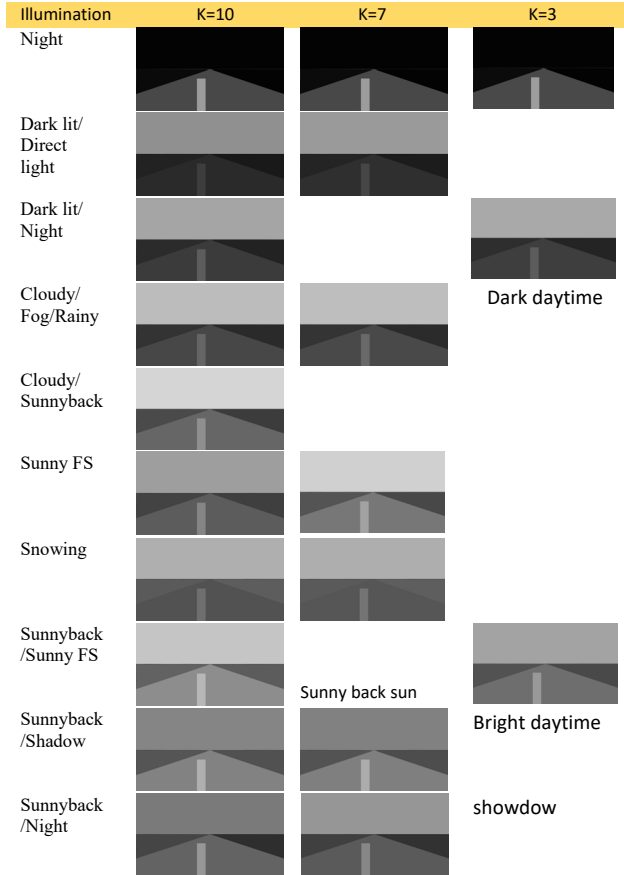


Fig. 9 Typical views of clustered weathers in intensity for $K = 10, 7,$ and 3 clusters. Central bars are standard deviation from the road average color.

In the weather and illumination classification, the results for $K=7$ is obtained at frame level for verification using the road profiles. Table IV shows the consistency measure of the classification as compared to the human tagged labels. For each cluster, we assign it with a *related category name* mentioned above. If human tagged category name is the same in a frame, the frame is consistent; otherwise, it is counted inconsistent. The cluster-wised consistency rates are thus obtained after evaluating another 60 video clips \times 9k frames in a testing set different from the training set. We found some inconstancy from the passing or front vehicles in the road profiles. Again, this is the evaluation via human tagging, which is not an absolute measure because of its ambiguity.

From the video mining, we have observations: (1) White lane marks are the most reliable features to follow. Painting lane mark is an inexpensive way to improve infrastructure for autonomous driving. (2) Road edges may be invisible in poor W/I conditions. Providing wrong sensed edges to the vehicle control will cause even worse consequences. Human drivers memorize road, follow front car, or refer to scenes on roadside in such maneuver. (3) The data mining discriminates different weather, rather than recover the absolute illuminance values as

auto-exposure cameras are used.

TABLE IV Matching rates of weather and illumination labels between human perception and data clustering

	C1	C2	C3	C4	C5	C6	C7	%
	cloudy/ rainy/fog	sunny facing	dark lit/ direct light	sunny back	shadow	snowing	night	
K7	93%	80%	91%	87%	80%	96%	98%	90%

	Night	Dark lit/ Direct light	Dark lit/ Night	Cloudy/ Fog/Rainy	Cloudy/ Sunny back

Fig. 10 Average intensity (upper) and chroma (lower) views for 10 clusters.

VI. CONCLUSION

This paper is a study of how roads look like under full spectrum of weather and illuminations based on mining large driving videos. Video clips are compressed to compact road profile images. The clustering of weather sensitive data results in stable clusters without overlap in the feature space, as compared to ambiguous human weather tagging. The classification based on learned clusters can help weather identification in video frames, and be applied to road edge detection or invisibility declaration in locating road edges.

REFERENCES

- [1] M. Kilicarslan, J. Y. Zheng, "Temporal video profile from driving video," IEEE IV, pp. 529-551, 2014.
- [2] N. Gimonet, A. Cord, G. Saint Pierre, "How to predict real road state from vehicle embedded camera?" IEEE IV 2015, 593-598
- [3] Tobias Kuhn, Jannik Fritsch, "Visio-spatial road boundary detection for unmarked urban and rural roads", IEEE IV 2014, 1251-1256.
- [4] B. Hillel, R. Lerner, D. Levi, and G. Raz, (2014). "Recent progress in road and lane detection: a survey". MVA, 25(3), 727-745
- [5] Guo and S. Mita, "Semantic-based road environment recognition in mixed traffic for intelligent vehicles and advanced driver assistance systems", 15th Int. IEEE ITSC, 444-450, 2012
- [6] H. Song, Y. Chen, Y. Gao, "Weather condition recognition based on feature extraction and K-NN", Advances in Intelligent Systems and Computing, 2014.
- [7] A. Gem, R. Moebus, U. Franke, Vision-based lane recognition under adverse weather conditions using optical flow, IEEE IV 2000.
- [8] Z. Wang, G. Cheng, J. Y. Zheng, "All weather road edge identification based on driving video mining", IEEE ITSC 2017.
- [9] J. Zhang, S. Xia, K. Lu, H. Pan, A. K. Qin, "Robust road detection from a single image", 23th ICPR 2016, 854-859.
- [10] R. Tian, L. Li, K. Yang, S. Chien, Y. Chen, R. Sherony, "Estimation of the vehicle-pedestrian encounter/conflict risk on the road based on TASI 110-car naturalistic driving data collection". IEEE IV 2014, 623-629.
- [11] H. Kurihata, T. Takahashi, I. Ide, Y. Mekata, H. Murase, Y. Tamatsu, T. Miyahara, "Rainy weather recognition from in-vehicle camera image for driver assistance", IEEE IV 2005.
- [12] X. Yan, et. al, "Weather recognition based on images captured by vision system in vehicle", Advances in Neural Networks, 2009.