

Sparse Coding of Weather and Illuminations for ADAS and Autonomous Driving

Guo Cheng, Jiang Yu Zheng, *Senior Member, IEEE* and Hiroshi Murase, *Fellow, IEEE*

Abstract— Weather and illumination are critical factors in vision tasks such as road detection, vehicle recognition, and active lighting for autonomous vehicles and ADAS. Understanding the weather and illumination type in a vehicle driving view can guide visual sensing, control vehicle headlight and speed, etc. This paper uses sparse coding technique to identify weather types in driving video, given a set of bases from video samples covering a full spectrum of weather and illumination conditions. We sample traffic and architecture insensitive regions in each video frame for features and obtain clusters of weather and illuminations via unsupervised learning. Then, a set of keys are selected carefully according to the visual appearance of road and sky. For video input, sparse coding of each frame is calculated for representing the vehicle view robustly under a specific illumination. The linear combination of the basis from keys results in weather types for road recognition, active lighting, intelligent vehicle control, etc.

I. INTRODUCTION

In current autonomous driving, many tasks cannot perform well under poor illuminations and the algorithms are not even aware of what type of weather a vehicle is driving in (Fig. 1). Some weather changes the appearance of road even more significantly than scene color (surface materials), and thus it is critical to vision tasks such as road and traffic detection. This paper deals with weather modeling and recognition in driving views for automatic driving.



Figure 1 Various road appearances in driving videos under different weather and illumination conditions.

Weather and illuminations have a great diversity and they lack discriminative features in images. Previous methods on weather recognition have assumed that weather category falls into separable classes, mostly two-classes, among huge differences in illumination. For example, recognizable individual weather types are fog [10], rain [11], snow [12], shadow [13], and night. Raindrops were detected on windshield as a clue of raining [2]. Though to some extent, those methods work well, they are unable to satisfy real

demand of multiclass weather classification in autonomous driving. To achieve a robust recognition of road environment and surrounding traffic in vision, remind driver safety actions, and even control the vehicle in a hazard weather, identifying weather and illumination is indispensable. Based on that, we can adapt thresholds in road segmentation, report a poor illumination and alert drivers a proper speed, and apply active lighting in night. Another effort on weather study is to estimate physical parameters, which is hard to connect to sensing and decision systems of the vehicle directly. Some assumptions and devices are not realistic to acquire on vehicles, which is difficult to fit into the real scenario for ADAS.

We adopt the data mining approach to describe multi-class weather and illuminations using a large set of naturalistic driving videos [5] captured by many cars. The related works so far on weather identification through in-vehicle cameras use color HIS and gradient features [1]. K-NN is further employed in the classification of raining weather [3]. The weathers are more daily used category measured in the entire view, rather than our road related category including illuminations.

Different from above works, this paper uses sparse representation that has dictionary learning and sparse coding for multiple-class weather classification. We classify different weather classes based on multiple features such as intensity and chroma from traffic-insensitive regions in driving videos. For a video frame, we extract its principal weights as a linear combination of typical views from weather and illumination classes. With these sparse coefficients obtained using a carefully learned dictionary, we classify the weather and illumination in every frame during the vehicle motion. Our input data consists of weather data profile [5] and road profile [6, 8] obtained from driving video in real time. As reported in other research [14], sparse coding is less influenced from partial corruption in image recognition. Our experiment also shows that the sparse coding approach is robust in the weather and illumination classification.

In the following Sec. II, we introduce vehicle views under various weather and illumination as preparation. Section III describes sparse coding applied to the view representation. Section IV discusses the weather and illumination classification, followed with experiments in Section V.

II. WEATHER AND ILLUMINATION IN VEHICLE VIEW

The weather and illumination we considered are not identical to the daily used categories in weather forecasting. They are related to the road view from a vehicle captured by a forward camera and recorded in driving video. The

G. Cheng is with Indiana University-Purdue University Indianapolis, Indianapolis (IUPUI), IN46202, USA (e-mail: guocheng@iu.edu).

J. Y. Zheng is with Department of Computer Science, and TASI, Indiana University-Purdue University Indianapolis, Indianapolis (IUPUI), IN46202, USA (e-mail: jzheng@iupui.edu).

H. Murase is with Nagoya University, Nagoya, 464-8603, Japan. (e-mail: murase@is.nagoya-u.ac.jp).

weather and illumination affect driver perception of road and traffic ahead, and influence vision tasks for autonomous driving. According to computer vision theorems, the light intensity casted on objects is additive for multiple light sources. The camera sensed irradiance is a linear combination of different light components ranging from ambient light, diffused reflection on object surfaces, specular reflection, and direct light from illuminants. For a vehicle on road, these light sources do not appear at the same moment. The specular reflection is from wet and asphalt road surface. Direct light comes into the camera at sunrise and sunset. The diffused reflection component is also determined from the surface reflectance as coefficients, and the specular reflection on wet road is further related to the light source direction. Night is another scenario where active lighting can be casted from vehicle headlight. Moreover, surface reflectance is determined from materials even related to seasons, e.g., grass turns to green and yellow in spring and fall respectively. In addition, all camera views are affected by intervened water particles in rain, snow, and fog, which blurs, whitens, and darken scenes.

To reduce the number of cases for consideration, we combine weather and illumination in our categorization, leaving material changeable on road and off-road as in Table I [5]. More sunny cases are included here as in Fig. 2, where the sun has directions from back, side, and front with respect to a vehicle. Figure 3 intuitively shows the spectrum of weather and illuminations in vehicle views with possible highlight and specular reflection. Most vision tasks on vehicles work well for the cases from sunny to cloudy so far.

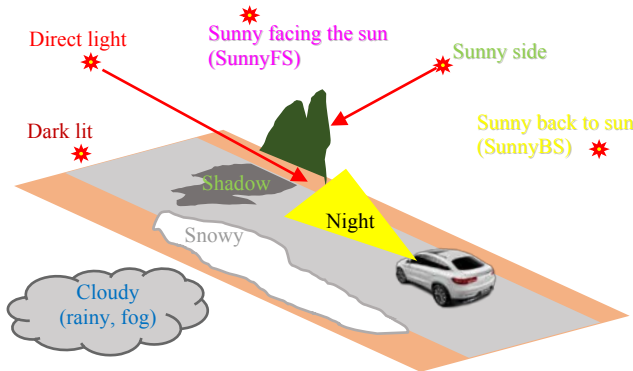


Figure 2 Direction of light source (the sun) with respect to the forward vehicle camera.



Figure 3 Iconic vehicle views of weather and illumination spectrum [5] summarized from a large naturalistic driving dataset.

To quantify the weather and illumination in vehicle views, we sample several traffic insensitive regions in video frames for color properties. The road and off-road colors are correlated with sky light, while the sky may vary drastically according to viewing distance and direction. We thus

consider the sky and ground together [5].

Table I Human categorized weather and illuminations

<i>W & I</i>	Qualitative description of road appearance	Appearance of vehicles on road
<i>Sunny back to the sun (BS)</i>	Including the sun in back and side directions without shadow, rich color, diffused reflection on road.	Vehicles and pedestrians have saturated color.
<i>Sunny facing the sun (FS)</i>	Sun in the camera forward direction, may have specular reflection on asphalt road surface. The scene is not colorful due to a narrow dynamic range of cameras. Large shadow areas are on objects.	Frontal vehicles have less color and dark shadows against bright road regions. Strong highlight spots on vehicles.
<i>Cloudy</i>	Including overcast and partly cloudy. Mostly ambient light from sky illuminate scenes.	Vehicle colors are well visible except the overcast.
<i>Shadow</i>	Shadow on road surface casted from trees, buildings, etc. Other vehicles may also cast shadow on road. Soft shadow from sparse trees and small objects has less influences on appearance, while hard shadow from high buildings and upper bridge coverage make the road area like dark lit.	Vehicles have small shadow regions casted by their bodies. The shadow regions below bumpers become an important clue for vehicle detection.
<i>Dark lit</i>	At dusk or after sun set, sky is still bright as ambient light, but road and background are dark without sufficient illumination. No color visible on road.	Vehicle lower boundary is not clear, and a vehicle is identified from tail or headlights.
<i>Direct light</i>	Sun at front with glare in image. Even hard for human driver to see road. May have highlight on road surface. Camera may close exposure due to strong sun light such that the ground becomes very dark.	A vehicle can be seen on top partially but less details in its region. Some direct light from opposite vehicle headlight happens in night.
<i>Rainy</i>	Dark illumination, low sensitivity of camera. Some cannot be separated from overcast. We look at wiper on/off to decide rainy state. Wet road surface has specular reflection of scenes (dark gray) and sky (bright).	Frontal vehicles may have unclear shape at lower part because of water spray.
<i>Snowy</i>	Snow intervene the scene and camera, makes image noisy. Road sides are more snow-covered and melt later than road surface. Road edges may not be straight.	Vehicle boundary may not clear when snow-covered.
<i>Foggy</i>	A half-transparent layer in front of camera blurs scenes. Camera exposure can enhance contrast to a certain extent.	Vehicles at front may have low contrast.
<i>Night</i>	Night is tagged by human when vehicle headlights are on. It includes rural night in dark, city night under street light, and wet ground with strong reflection from other cars and street lights.	The complete vehicle shape is hard to obtain; vehicle lights become main clue for vehicle detection.

In all video frames where the horizon is fixed by camera calibration (Fig. 4), we sample the road surface at a *line* below the horizon, which reaches approximately 15m ahead for road detection. The linear data are stored temporally in the road profile image [8] as shown in Fig. 5. The road region sampled in the road profile covers the width of one lane. Additional three regions in the frame such as sky, left and right roadsides between the horizon and the sampling line are also sampled. The sky region is set at the very top of

frames to avoid buildings in forward direction. Roadside regions cover large areas from 15m to infinity. Although they may still capture opposite vehicles on the left, and passing vehicles on both sides, the time duration in the video is relatively short.



Figure 4 Sampling video frames to obtain the road profile and feature data related to weather and illuminations. Video frame where horizon and sampling line are indicated.

From these four regions, *intensity* and *chroma* are calculated and recorded along the time axis. We are not using *saturation* to describing color because of its singular value with zero-intensity in night scenario. The color variation on the road surface is computed as well from the road profile. All these nine features, f_i , $i = 1, \dots, 9$, form a feature vector and the whole video sequence generates a feature matrix, $f_i(j)$, proportional to the video length, where j is the frame number of the video clip. We use these features to measure the weather and illumination conditions, which avoids far road beyond 15m (bothered by other vehicles frequently) and roadside buildings, trees, and mountains as much as possible.

Rather than relying on the semantic weather and illumination categories C annotated by humans, our previous work has investigated the discriminable clusters among the feature space to describe the weather and illuminations. K-mean clustering method was applied to large naturalistic driving videos [9] to obtain seven stable clusters K as shown in Fig. 6. These data mining clusters in K are not the exact overlaps with classes C but have an 80% consistency by classifying human tagged video clips [5]. As a result, cloudy and rainy are indistinguishable in the feature space [5], though they can be separated from wiper movement or rain drop on windshield in the videos. Direct light and dark lit are in the same category. More clusters will increase the difficulty of classification, while fewer clusters may not provide sufficient information to road and object recognition. The chroma are all low, except for *sunny back to the sun* in daytime and on *night* road surface (lit by yellow headlight in the dataset). The further detailed changes in materials and illumination can be considered as their sub-clusters further.

III. SPARSE CODING OF WEATHERS IN VEHICLE VIEWS

A. Sparse Coding by Sample Basis

With the clusters learned from large naturalistic driving videos in the feature space, new views from driving video can be classified into one of these clusters. The obtained cluster can be further used to retrieve the visual properties for guiding the road segmentation [6], active lighting in

night, control the vehicle in poor weather, and report the invisibility of road edges for switching to other modes such as vehicle following.

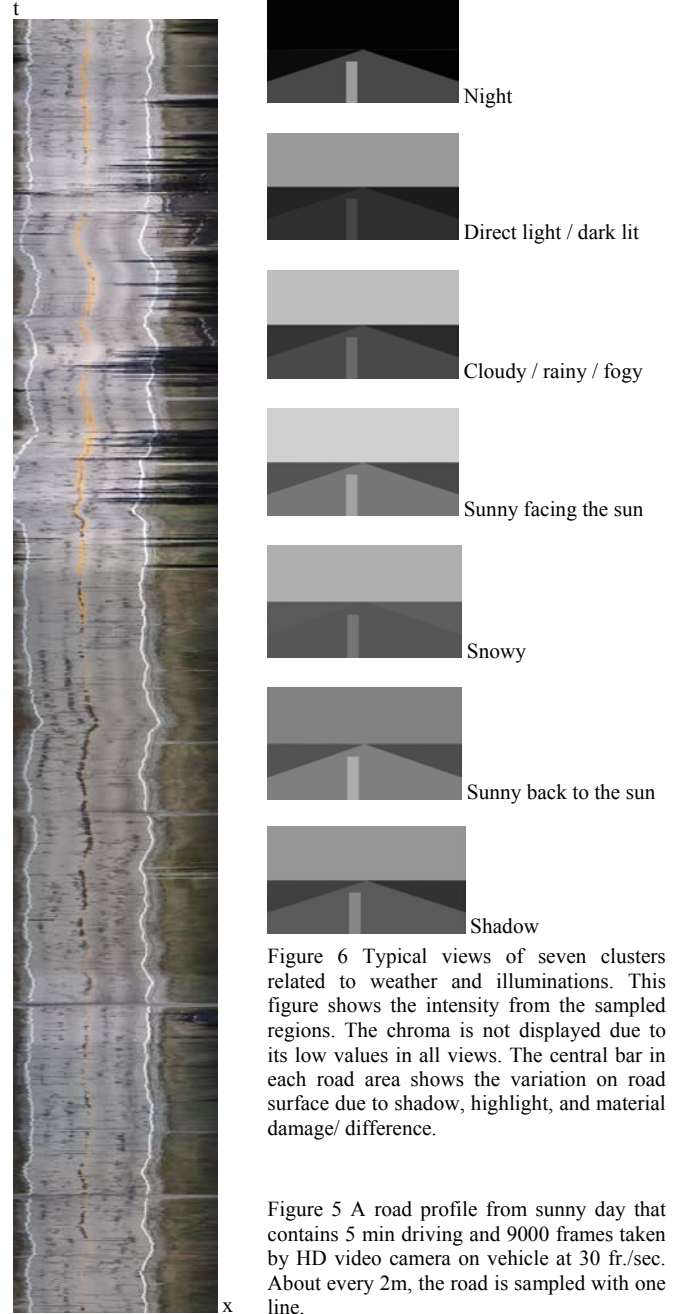


Figure 6 Typical views of seven clusters related to weather and illuminations. This figure shows the intensity from the sampled regions. The chroma is not displayed due to its low values in all views. The central bar in each road area shows the variation on road surface due to shadow, highlight, and material damage/ difference.

Figure 5 A road profile from sunny day that contains 5 min driving and 9000 frames taken by HD video camera on vehicle at 30 ft./sec. About every 2m, the road is sampled with one line.

Here we explore the *sparse coding* representation of weather and illuminations in the vehicle views. A new vehicle view (features) is formed as a linear combination of typical views (features) among various weather clusters, according to the light additive principle under multiple light sources. A few non-zero coefficients in this linear combination indicate the belonging cluster of this new view, and the properties of the clusters can also be inherited to the view directly.

In image recognition, *sparse coding* [7] has been widely applied in finding group presentations of data sets, either labeled or unlabeled. The process is divided into two phases:

dictionary learning and *sparse coding*. The former plays a role in extracting keys from input data (samples) to form base vectors of a *code book*. While the latter results sparse coefficients through optimization. This captures a sparse expression of a data sample based on the dictionary. There are two points distinguishable from PCA: (1) All the keys in the dictionary have no requirement to be orthogonal. (2) The dictionary can be an over-complete spanning set. As an unsupervised method, the sparse coding aims at representing input data by sample basis in the dictionary space. More precisely, given a set of bases, $\Phi_i, i = 1, 2, \dots, k$ in the dictionary learning process, sparse coding aims at representing an input set X as a linear combination of Φ_i with corresponding coefficients a_i as:

$$X = \sum_{i=1}^k a_i \Phi_i \quad (1)$$

Note that the number of basis Φ_i usually is larger than the dimension of feature space. This over-completeness ensures a full possibility to present various patterns in input data. In addition, we need a restriction to fulfill the “sparseness” on coefficients, i.e. the coefficients far from zero should be as few as possible. Therefore, a cost function is added to Eq. (1) to penalize a_i for being non-zero. For all input samples $X^j, j=1, \dots, m$, we obtain coefficients $a_i^{(j)}$ of the linear combination by

$$\underset{a_i^{(j)}, \Phi_i}{\text{minimize}} \sum_{j=1}^m \|X^j - \sum_{i=1}^k a_i^{(j)} \Phi_i\|^2 + \lambda \sum_{i=1}^k S(a_i^{(j)}) \quad (2)$$

where m is the number of input samples and $S(a_i^{(j)})$ is the penalty for sparse coefficients. In sparse coding, $S(a_i^{(j)})$ is measured with L_1 norm, since L_1 overrides L_0 in the property of differentiable and is easier to be optimized in general. For example, we can adopt $S(a_i^{(j)}) = |a_i^{(j)}|$ and its log penalty is $S(a_i^{(j)}) = \log(1+a_i^2)$.

Moreover, the sparsity penalty can be affected by the scale of Φ_i , such that the basis or the dictionary learned have an impact on L_1 norm. Here we constrain the value of basis. The full sparse coding cost function thus becomes

$$\underset{a_i^{(j)}, \Phi_i}{\text{minimize}} \sum_{j=1}^m \|X^j - \sum_{i=1}^k a_i^{(j)} \Phi_i\|^2 + \lambda \sum_{i=1}^k S(a_i^{(j)}), \quad \|\Phi_i\| \leq C, \forall i = 1, \dots, k \quad (3)$$

B. Dictionary from Weather Views and Sparse Coding

Now we introduce the view representation under different weather and illuminations with sparse coefficients combined from a set of base views. With the whole spectrum of weather and illumination K obtained in K-mean clustering [4], we try to compute a fewer number of coefficients that can adapt typical views to an arbitrary frame in the driving video.

Although various algorithms of sparse coding learn the dictionary from input data, we expect the basis to be as primitive as they can so that any complex input can be represented as a sparse combination of them. On the other hand, the number of basis in the dictionary also plays a role. Here, we select the centroids of clusters from K-means [5] for the basis. In addition, we use K-mean clustering in the obtained clusters to further obtain sub-clusters, which implement a hierarchical and sequential learning. The sub-clusters cover the further variations in the clusters due to the

different roadside materials, illumination styles, and weather differences. The basis we extract from K-means clusters should have as small overlap as possible in the feature distributions. Among the clusters and sub-clusters, we pick up their peak values as the basis for codebook. In the high-dimension feature space, the larger the distance between the basis, the sparser in representing each weather view, and the easier for classification. Based on the volume of each cluster and its distribution, we obtained $k=17$ basis as the atoms in codebook from seven clusters ($K=7$). They are displayed in Fig. 7 in views. We put them into the sparse coding algorithm to produce the sparse coefficients for input samples. Here the term “sparse” mostly refers to the division of zero and non-zero values of coefficients in the expression.

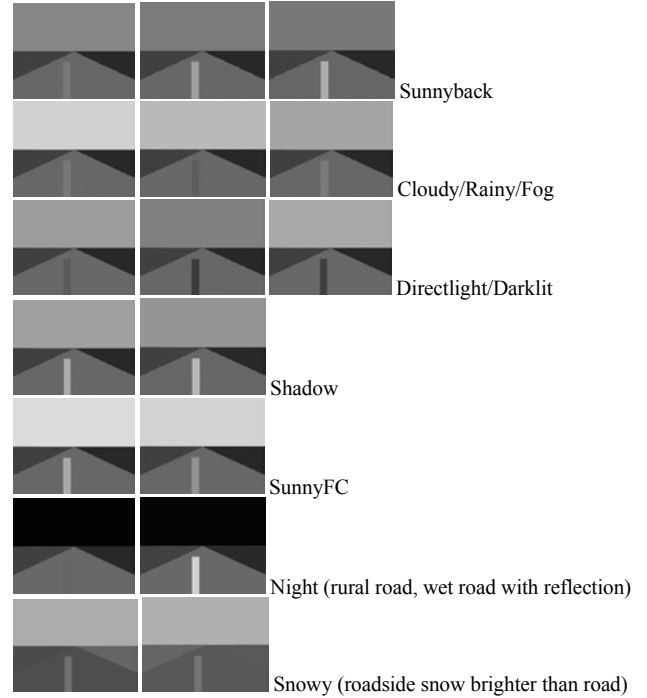


Figure 7 Intensity of 17 sub-cluster keys from K-means clustering based on 300 sample videos of different weathers. The brightness of central bar in each image indicates the variation within road area.

For example, we took 300 profiles in the K-Means ($K=7$) clustering into this sparse coding model. Each frame with 9 features is expressed as a linear combination of 17 basis, among which only a few are non-zero values. For examining the results, we seek for a visualization to compare the original road profile with the view approximated by sparse coding. Figure 8 compares the ground in estimated views with the road profiles. We can notice that they are close, although the sky is not displayed. Ten road profiles from different weather categories are selected for comparison. Non-zeros coefficients are associated with the corresponding basis.

IV. WEATHER AND ILLUMINATION CLASSIFICATION

Given a batch of unlabeled view frames, we have learned their sparse representation in terms of the dictionary. Each input will yield a projection to the basis. Sparse coding uses the dictionary to learn sparse coefficients, among which only a few of them are non-zero. The following supervised learning in a pipeline is done for weather classification of video frames. We take another 300 labeled profiles for

experiments, and adopt sparse coding for the testing of weather recognition:

a) *Learning sparse coefficients.* Input the 9-feature vector of a frame and project it onto the dictionary space. The sparse coding algorithm minimizes the difference of linear combination with the input view (features), and this generates sparse coefficients with only a few non-zero values.

b) *Classification.* Use non-zero coefficients to classify the input into a related weather cluster, which has the basis that non-zero coefficients associated with.

c) *Visualization and Verification.* the classified frame is compared with the road profile briefly and small differences indicate a good approximation by the linear combination.



Figure 8 Comparison of road profile with approximated views on road and off-road for different weather clusters. Left of each figure is road profile and right column is a series of frames generated from sparse coding (linear combination of 17 basis). The sky is brighter than lower ground except night. The frame top is aligned with the frame position in road profile.

Several recognition results covering different weather

clusters are displayed with the road profiles in Fig. 9. Different colors are marked as the weather ID on both sides of a road profile for verification of weather and illumination types. Along with the final classification and road profile on left and middle columns, the right column consists of three small columns that represent the most obvious three sparse coefficients. We can notice that the second coefficient is much smaller than the first one, and it may still be correct if the first coefficient is wrong.

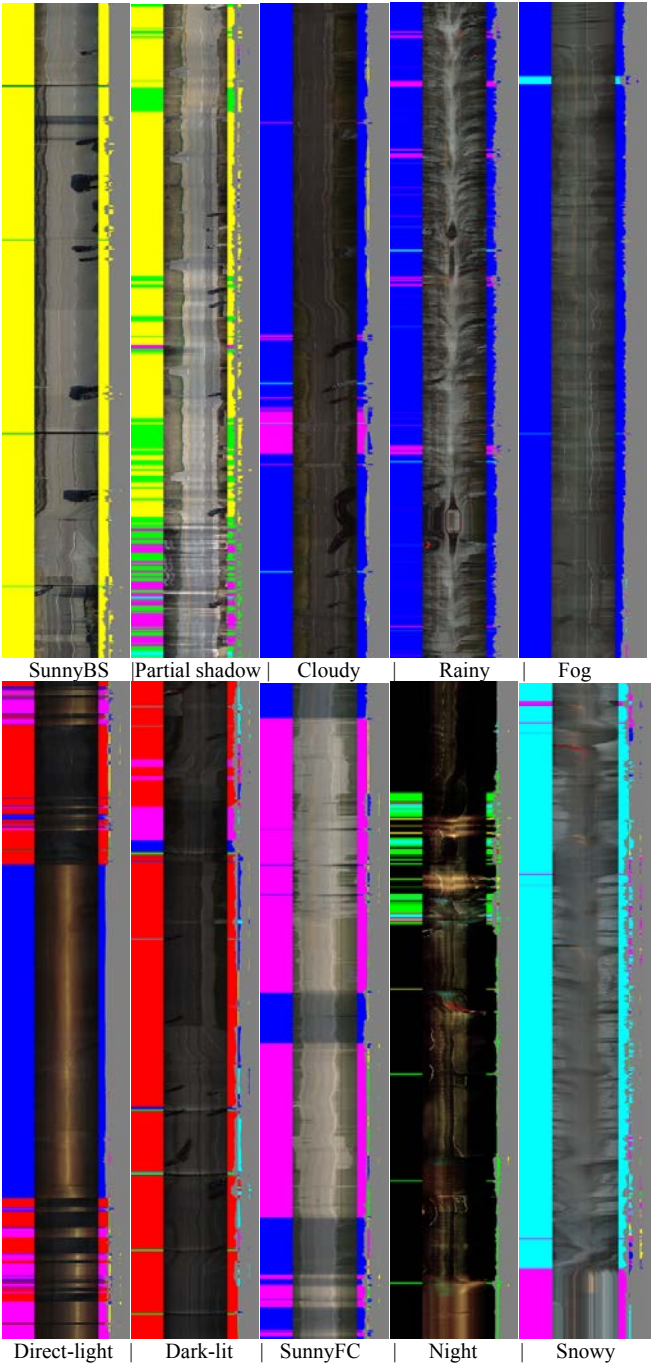


Figure 9 Classification results of all weather and illuminations for all frames in video clips. Yellow-Sunny back to the sun, Green-Shadow, Magenta-Sunny facing the sun, Blue-Cloudy/raining/fog, Cyan-Snow, Red-Dark lit/direct light, Black-Night. The cluster each frame belongs to is colored on left side of road profile. Right of each road profile shows top three sparse coefficients in consecutive narrow columns. Their scales are displayed by horizontal lengths with the remaining part painted in gray.

V. EXPERIMENTS AND DISCUSSION

The naturalistic driving video [9] uses 110 cameras of the same type. Auto-exposure is applied based on the average brightness in the view. Among the dataset, 300 videos covering all spectrums of weather and illumination described in Table I were selected for experiment. Each clip has 5 min driving and the 9 features are extracted from every frame for over 9000 frames along with the road profile [6]. The road types include urban, rural, highway, and local roads in different seasons. For narrow streets with high buildings on sides, they have more chance to be categorized as shadow in sunny day if road is casted with shadow.

Table II Confusion matrix of weather and illumination recognition using {C}. CRF stands for Cloudy/Rain/Fog, and DL/DL stands for direct light and dark lit. The result counts frame numbers. The F1 score is **89.9%**.

	Sunny BS	Shadow	Sunny FS	CRF	Snowy	DL/ DL	Night
SunnyBS	80636	835	22	321	1950	58	21
Shadow	1006	7590	154	97	0	99	42
SunnyFS	103	0	8675	818	604	0	0
CRF	15	252	1584	55797	3216	886	44
Snowy	7	737	978	81	36090	11	1
DL/DL	82	277	552	4130	658	82052	143
Night	293	514	22	85	226	27	62193
Recall	0.96	0.84	0.85	0.90	0.95	0.93	0.98
Precision	0.98	0.74	0.72	0.90	0.84	0.98	0.99

The software used for sparse coding is CVX tool in Matlab and running time in calculating the sparse coefficients for a road profile with 9027 frames is about 929 seconds, i.e., 0.1 sec/fr in average. This means that the weather identification is sufficient for real time video processing before other modules start working. The λ in Eq. (3) is set to 100 in the minimization. Using 60 testing videos including all weather and illuminations, the recognition accuracy by sparse coding is given in Table II resulting F1 score around 90% as compared to the recognition accuracy from 60%-85% using K-mean distance in the feature space [5]. That is, the sparse coding result using K codebook is more consistent to the human annotation C than using K-mean distance from clusters K .

As shown in Table II and Fig. 9, some miss-classification with the largest coefficient can be examined further with the second coefficient, which may be correct. The error can be considered as corruption due to the sampling regions affected by buildings, surrounding vehicles, and unexpected scenes. Direct light and sunnyFS can be mixed as they are adjacent cases with different sun position in the view. The frame-level result also shows a sensitive change to illumination on road. We are not considering a temporal filtering to remove inconsistent frames locally with the results from frames before and after; thus, we can capture real details on road.

The merits using the sparse coding are (1) only a few non-zero coefficients are used in the linear combination to represent a single frame, which has more robust results in some corruption [14]. (2) the direct use of the coefficients in determining the sensing strategy and parameters. For example, it can guide the road edge detection with a proper threshold for different weather and illuminations. For side area with only a weak contrast in the image, a vehicle could

not judge so far whether an edge exists in dark, or it is a wide road of the same material if no weather and illumination above the ground was identified. By referencing the sky and surroundings, the illumination is classified and then thresholds can be set accordingly [6]. If it is a direct light, or night driving on wet road, the system should report the invisibility of road edges, and switch to tracking mode to follow other frontal vehicles.

VI. CONCLUSION

This paper challenges weather recognition in a vehicle view with the sparse coding for a robust result. The weather and illumination have been modeled through clustering in an unsupervised learning. Based on the additive nature of lights, the sparse coding identifies a weather type, as a road image is input from the forward camera on vehicle, using a series of basis from sample views. The coefficients can be used to guide road segmentation in ADAS, active lighting in dark illumination, and driving speed control in a poor weather and illumination.

ACKNOWLEDGEMENT

This work is partially supported by Dept. of Transportation, USA under UTC-CrIS mini-grant, TOYOTA CSRC, USA, and Future Value Creation Research Center, Graduate School of Informatics, Nagoya University.

REFERENCES

- [1] X. Yan, et. Al, "Weather recognition based on images captured by vision system in vehicle", Advances in Neural Networks, 2009.
- [2] 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 Intelligent Vehicle Symposium 2005.
- [3] H. Song, Y. Chen, Y. Gao, "Weather condition recognition based on feature extraction and K-NN", Advances in Intelligent Systems and Computing, 20014.
- [4] Adam Coates, Andrew Y. Ng, "Learning Feature Representations with K-means", G. Motavon, G. B. Orr, K. -R. Muller (Eds.), Nueral Networks: Tricks of the trade, 2nd edn, springer LNCS 7700, 2012.
- [5] G. Cheng, Z. Wang, J. Y. Zheng, "Big-video mining of road appearances under full spectrums of weather and illuminations", IEEE ITSC 2017, 1-6.
- [6] Z. Wang, G. Cheng, J. Y. Zheng, "All weather road edge identification based on driving video mining", IEEE ITSC 2017, 1-6.
- [7] D. Cai, H. Bao, X. He, Sparse concept coding for visual analysis, CVPR, 2905-2910, 2011.
- [8] M. Kilicarslan, J. Y. Zheng, "Temporal video profile from driving video," IEEE Intelligent Vehicle Symposium, pp. 529-551, 2014.
- [9] 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 Intelligent Vehicle Symposium, 623-629, 2014.
- [10] K. Y. Choi, K. M. Jeong, B. C. Song, Fog detection for de-fogging of road driving images, IEEE ITSC 17, 1833-1838, 2017.
- [11] S. Hasirlioglu, A. Riener, Introduction to rain and fog attenuation on automotive surround sensors, IEEE ITSC, 661-667, 2017.
- [12] A. Gem, R. Moebus, U. Franke, Vision-based lane recognition under adverse weather conditions using optical flow, IEEE IV, 2000.
- [13] J. M. Alvarez, A. M. Lopez, Road detectin based on illuminant invariance, IEEE Transaction ITS, 12(1), 184-193, 2011.
- [14] J. Wright, A. Yang, A. Ganesh, S. Sastry, Y. Ma, Robust face recognition via sparse representation. IEEE T-PAMI, 31(2), 2009.