PCA based Computation of Illumination-Invariant Space for Road Detection

Taeyoung Kim
KAIST
retupmoc@kaist.ac.kr

Yu-Wing Tai
Sensetime group limited
yuwing@gmail.com

Sung-Eui Yoon
KAIST
sungeui@kaist.edu

Abstract

Illumination changes such as shadows significantly affect the accuracy of various road detection methods, especially for vision-based approaches with an on-board monocular camera. To efficiently consider such illumination changes, we propose a PCA based technique, PCA-II, that finds the minimum projection space from an input RGB image, and then use the space as the illumination-invariant space for road detection. Our PCA based method shows 20 times faster performance on average over the prior entropy based method, even with a higher detection accuracy.

To demonstrate its wide applicability to the road detection problem, we test the invariant space with both bottom-up and top-down approaches. For a bottom-up approach, we suggest a simple patch propagation method that utilizes the property of the invariant space, and show its higher accuracy over other state-of-the-art road detection methods running in a bottom-up manner. For a top-down approach, we consider the space as an additional feature to the original RGB to train convolutional neural networks. We were also able to observe robust performance improvement of using the invariant space over the original CNN based methods that do not use the space, only with a minor runtime overhead, e.g., 50 ms per image. These results demonstrate benefits of our PCA-based illumination-invariant space computation.

1. INTRODUCTION

Recent developments of Advanced Driving Assistance System (ADAS) aim to detect driving environments efficiently and provide various useful supports related to safety and convenience to drivers. One of the most important tasks is to identify road areas ahead of a driving vehicle. Identifying road areas is one of the fundamental steps for other tasks such as vehicle [29] and road sign [23] detection.

ADAS can use different types of sensors such as radar, LIDAR, and stereo cameras. In this paper, we focus on vision-based road detection techniques with a monocular camera mounted in front of a vehicle, as this is one of the most cost-effective and practical ways to install ADAS to a wide variety of existing vehicles.

Early work utilizes various boundary and region features [10, 21, 17, 20], to detect road boundaries on well-marked roads [18]. When roads have severe illumination changes such as shadows, these features in a simple RGB color space can easily generate false detection. To address this issue, different color spaces [28, 25], convolutional neural nets [5], and utilizing stereo images [27] have been proposed. A few work [2, 3] suggested illumination-invariant spaces for handling shadows with a particular camera, and utilized a histogram classifier for detecting road area. Nonetheless, investigating the illumination-invariant space for road detection has not been actively studied, and detecting road areas robustly under severe illumination changes remains a main challenge in the field.

Main contributions. To efficiently, yet accurately consider such illumination changes, we propose a Principal Component Analysis(PCA)-based computation for computing the illumination-invariant space for road detection (Section 3.1). To demonstrate its wide applicability to various road detection methods, we have tested the computed space both with bottom-up and top-down road detection methods. We observe that in the illumination-invariant space, patches on the road area share similar texture appearance and are connected well within the road area. Based on this observation, we develop a simple, bottom-up road detection method using patch propagation (Section 3.2). Furthermore, we apply the illumination-invariance space as an additional feature to recent CNN-based road detection methods that pass context information in a top-down manner.

Overall, our PCA-based computation method, PCA-II, works 20 times faster than the prior entropy-based method, while it improves the accuracy (Section 4.1). To demonstrate benefits of the illumination-invariant space for road detection, we test various detection methods with the KITTI dataset (Section 4.2). We first compare our simple patch-propagation method against other state-of-the-art road detection methods running in a bottom-up manner, and found that our simple method with the invariant space shows

(a) FCN LC w/o II [24] (b) FCN LC w/ II (c) CN24 w/o II [5] (d) CN24 w/ II

Figure 1. Visual comparison of recent CNN approaches w/ and w/o considering the proposed illumination-invariant (II) images. Detected road pixels are visualized in the green. By considering II images, accuracy of these tested CNN methods has been improved.

higher accuracy and takes 450 ms on average in a Matlab implementation. For using the space as the additional feature within recent CNN based road detection methods, we were able to observe robust accuracy improvements, up to 1.5% point higher, over the original CNN methods only with a minor runtime overhead, e.g., 50 ms per each image (Fig. 1). These results demonstrate that our illumination-invariant space is a robust feature that can be used in many different road detection methods irrespective of whether they run in a bottom-up or top-down manner.

2. RELATED WORK

We discuss illumination-invariant images followed by road detection.

2.1. Illumination-Invariant Images

In many camera-based vision systems, variations on illumination such as shadows and interreflections can create unwanted artifacts and thus significantly affect results for their applications. As a result, separating effects of lighting on the image has been considered actively in general.

Barrow and Tenenbaum [4] introduced the concept of separations called "intrinsic images" to decompose the lighting effect given an input image. Weiss *et al*. [31] proposed a method to generate an intrinsic reflectance image from a sequence of images based on the maximum-likelihood estimation. Finlayson *et al*. [12, 14, 13] suggested a shadow removal method using an entropy minimization from a 2D chromaticity representation.

Corke *et al*. [7] and Maddern *et al*. [22] showed effects of illumination-invariant images on mobile robotics problems such as outdoor scene localization and scene parsing. They applied invariant images only to execute those tasks. Álvarez *et al*. [2, 3] suggested road detection based on illumination-invariant images.

We suggest a novel method to generate illumination-invariant images using Principle Component Analysis (PCA), which achieves high quality in a faster performance over prior methods. We also demonstrate that our proposed PCA-based invariant images can become a new feature to CNN based structures for road detection.

2.2. Road Detection

In monocular image based road detection, various techniques use boundary and region features. Methods using boundary features (e.g., road edges) mainly considered low-level information of input images [6, 20]. These techniques commonly extract edges to find boundaries or line paths of roads, and determine road area between detected boundaries. These methods are appropriate for structured roads like the highway, which have well-painted lane marks or evident edge boundaries. Unfortunately, they have been identified to be vulnerable to drastic illumination changes such as shadow [19].

Other approaches with region features utilized raw information such as color or texture of images. Texture based road detection methods used the textural difference of road and non-road regions. Most commonly used texture descriptors are statistical and structural descriptors [17]. Simple color spaces like RGB, unfortunately, do not represent road features well, because road textures in the RGB space vary significantly as a function of illuminations [3]. Due to the vulnerability of the RGB color space, some methods used different color spaces as features. For example, Sotelo *et al*. [28] used hue-saturation-intensity (HSI) color space to find luminance-robust features. Álvarez *et al*. [3] utilized shadow removal techniques using the illumination-invariant methods [14] to classify road area.

Some of recent approaches use convolutional neural networks. Álvarez *et al*. [1] learned road-texture patterns by color plane fusion and applied neural networks to label transfer. They combine extracted general information and the Naive Bayes framework to classify images. Brust *et al*. [5] presented convolutional patch networks with spatial information of the patch. They classify image patches at each pixel position. Mendes *et al*. [24] proposed a large contextual window using a network-in-network architecture to label the road area.

We found that most prior road detection techniques are rather sensitive to the variation of illumination. In this paper, we propose to use an novel PCA-based illumination-invariant image, and show its benefits against various road detection methods.

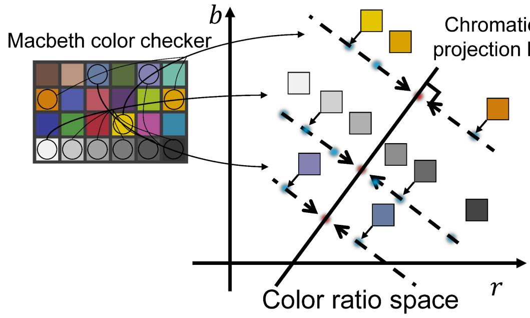
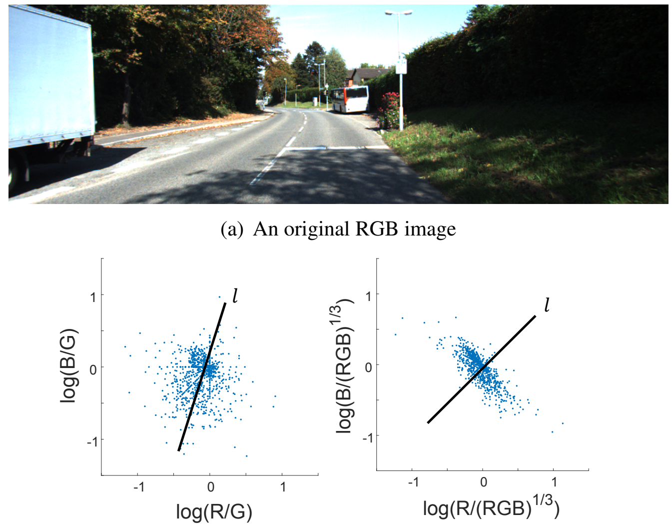


Figure 2. This figure shows an ideal log color ratio plot from the Machbeth color checker. Patches of the same chromaticity are mapped on a dotted line. We compute an illumination-invariant space by identifying the chromaticity projection line, l, and projecting patches onto the solid line l.



(a) An original RGB image

(b) (c)

Figure 3. (a) shows an RGB image from the KITTI dataset. (b) and (c) show plots of log color ratio and geometric log color spaces, respectively. We also show chromaticity projection line l in each space. (c) can lead to the smallest variance on the line, resulting in a more compact set of chromaticity and canceling illumination changes better.

3. OUR APPROACH

In this section, we describe our illumination-invariant image generation by finding a chromaticity projection line with the minimum variance. We also suggest a simple road detection method using patch propagation that utilizes the property of the illumination-invariant space.

3.1. Illumination-invariant space

Color based road detection techniques can be sensitive to the variation of illumination. Especially, illumination changes such as the shadow on roads pose challenging problems. For designing a robust and efficient technique with a monocular camera, we propose to compute an illumination-invariant color space by using the Principal Component Analysis (PCA). By using PCA, we efficiently find a chromaticity projection line, l, that realizes a compact set of chromaticities.

To generate illumination-invariant images, three assumptions commonly known as PLN-assumptions are used [14]. It assumes that an image is captured under *Planckian illumination* and *Lambertian surfaces* obtained by *Narrowband camera sensors*. Under the PLN-assumptions, we can remove or reduce the influence of illumination variations using an appropriate color ratio space.

After converting RGB colors of input images into a 2-D color ratio space, we can now find a line space that is independent to illumination. In the ideal case shown in Figure 2, the same chromaticity with different intensities are mapped on the same dotted line in the color ratio space. The solid line l, which is orthogonal to all the dotted lines, is a one-dimensional line space that we want to find. To obtain an illumination-invariant image, we project all the colors to the solid line l in the direction of the dotted arrow.

Ideally, we aim to find an optimal projection direction that results in the largest variance along the projection direction, since the variance is correlated with illumination changes and canceling them reduces the influence of illumination changes effectively.

We now describe each step of computing the illumination-invariant space. The first step is to transform the RGB space into a color ratio space. We found that two modern color ratio spaces are useful for our goal. A recent approach [3] proposed to compute log color ratios for each pixel as the following:

$$r = \log\left(\frac{R}{G}\right), \quad b = \log\left(\frac{B}{G}\right). \tag{1}$$

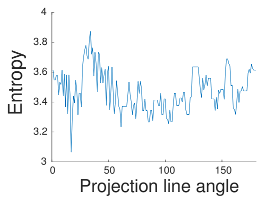
This function computes log color ratio values using the G channel as the normalizing channel. However, it works under the ideal narrowband sensor camera. Certain real cameras have broadband sensors, and in that case, Finlayson *et al*. [11] suggested a geometric mean log color space, as follows:

$$r = \log\left(\frac{R}{(RGB)^{1/3}}\right), \quad b = \log\left(\frac{B}{(RGB)^{1/3}}\right). \tag{2}$$

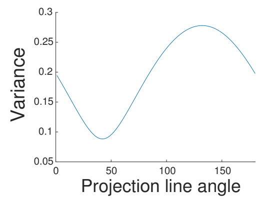
Before we find the projection line, determining a proper mapping function is essential for computing a high-quality illumination-invariant space. We discovered that the log mapping function (Equation (1)) works well for certain types of camera sensors such as Sony ICX084, while the geometric mean (Equation (2)) works better for other kinds of cameras such as ICX204.

Figure 3 shows two different color ratio plots of an input image from the well-known KITTI dataset [16]. This

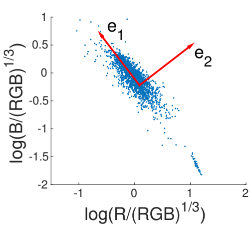
Figure 4. Original RGB images (first row), and the illumination-invariant images computed by our PCA-based variance minimization (second row). Our method removes shadows reasonably well in the visual inspection. Furthermore, our PCA approach improves the road detection accuracy over the prior entropy method (Table 2).



(a) Entropy



(b) Variance



(c) PCA

Figure 5. (a) and (b) show entropy and variance curves as a function of the projection line angle. (c) shows two eigenvectors from PCA. We pick e_2 as the projection line l for creating the shadow-free image.

input image is obtained by Point Gray Flea 2 Video Camera, which has SONY ICX204 sensor. It shows that the log geometric mean (c) is more proper than log (b), since it can reduce chromaticity variance on a computed chromaticity projection plane. Among available options of computing color ratio spaces, we pick a space for each camera sensor that results in the largest variance along the projection direction using PCA; its detail method is given later in this section. Note that this process of choosing a proper color ratio space is one-time process for each camera sensor, since it is invariant across different images captured from the camera sensor.

Once we pick the color ratio space for each sensor type, we now determine a projection direction for each input image. Prior methods [13] relied on sampling process, and tried out many projection directions and picked the one with the minimum entropy. We find that this approach is slow for our purpose and sensitive to outliers in the chosen color ratio space. We find that variance measure is more robust than the entropy measure, especially under noise and outliers. Figure 5 shows entropy and variance curves with different projection lines parameterized by a line angle.

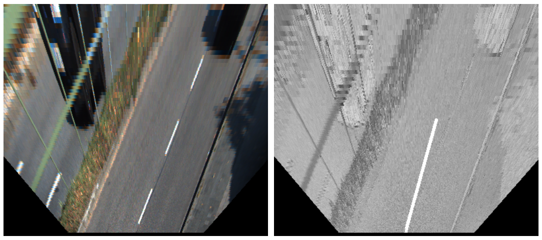
To efficiently identify the projection direction with the maximum variance, we propose to use PCA as a global approach that identifies a semi-major axis with the maximum variance (Figure 5). To apply PCA, we consider the color ratio values as $2 \times n$ matrix X, where n is a number of sampled pixels. The covariance matrix $C = XX^T$ is decomposed using the singular value decomposition as follows:

$$C = XX^T = \begin{pmatrix} e_1 & e_2 \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} {e_1}^T \\ {e_2}^T \end{pmatrix}, \quad (3)$$

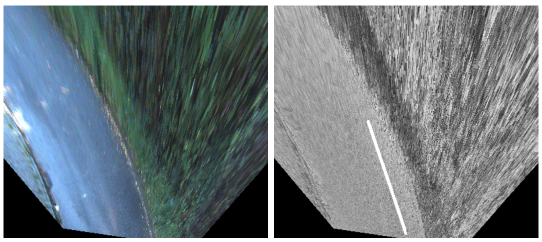
where e_1 and e_2 are the two eigenvectors with their corresponding eigenvalues λ_1 and λ_2, respectively. We choose the largest variance λ_1 and its eigenvector e_1 as the projection direction. In other words, its orthogonal direction e_2 and the center of samples define the computed chromaticity projection line l. Figure 4 shows results of our method. As can be seen, our PCA-based approach successfully removes the effect of shadow casted area.

Our PCA based method shows extremely fast generation of the illumination-invariant space than the prior entropy-based method. The faster running performance is acquired, because our method accesses the data two times during the PCA computation, while the entropy based approach performs projections many times (e.g., 180 times).

We have also applied our method to a simple road detection method, which is described in the next section. We found that our method achieves similar road detection accuracy over the entropy-based method, mainly thanks to the robustness of our method to noise and outliers. Furthermore, when we apply our illumination-invariance space as an additional feature to recent CNN based road detection methods [5, 24], we achieve more accurate detection accuracy than the original CNN approachs that do not use our space. More detailed comparisons and analysis in terms



(a) Straight road



(b) Curved road

Figure 6. RGB images in bird's-eye view (BEV) on the left, and their corresponding BEVs in the illumination-invariant space on the right with the computed seed lines.

of accuracy and runtime performance are available in Section 4.

3.2. Road detection using patch propagation

We observed that in the illumination-invariance space, patches on road areas have similar texture appearance and are connected to each other within the ego-lane. Based on these observations, we suggest a simple patch-based road detection, which runs in a bottom-up manner of detecting the ego lane. Additionally, we apply our method to recent CNN-based road detection methods that take a top-down manner for detection, and show benefits of the illumination-invariance space in Sec. 4.2.

Our simple road detection methods are composed of two parts. This approach identifies road regions by efficiently propagating patches from seed road patches. First, we select appropriate seed patches given an image. For extracting those seed patches, we use the illumination-invariant space, which is a shadow-free grayscale image of an input image. Second, starting from the seed patches extracted from the seed line, we propagate those seed patches by identifying similar, nearby patches.

Many prior methods [21] also utilize the idea of using seeds that are in the road region. These methods commonly identify such seed pixels based on a heuristic that the bottom middle area is usually located in the road. This approach is very simple and reasonable for most roads. The heuristic, however, fails, when some obstacles are located in front of the camera or a road shape is strongly curved.

Instead of relying solely on the heuristic, we propose a more robust seed selection method that utilizes the illumination-invariant image. Our method utilizes the idea that the road in the illumination-invariant image exhibits similar texture appearances, and considers the variance measure to find similar road textures. Specifically, after generating the illumination-invariant image, we convert the image into a Bird's-Eye View (BEV) image using the reverse perspective projection provided by KITTI benchmark [16], to reduce perspective effects of on the road; note that BEV images are still in the illumination-invariant space. We then attempt to find the seed line in the BEV image for extracting seed patches.

We define 2 D positions (x, y) of lines as a function a line angle, θ, as follows:

$$y = \tan\theta(x - cam.x) + cam.y, \tag{4}$$

where $cam.x$ and $cam.y$ are the camera location in the image. To compute the seed line on the road, we test lines with pre-defined line angles that cover various kinds of road shapes, e.g. straight or left-curved road. We then compute the variance of pixel values that are located on each line, and pick the angle θ with the minimum variance and use it as the seed line. Fig. 6 shows computed seed lines based on our method; the computed seed lines are located well within the road area, regardless of road's curvature and obstacles.

Once we identify the seed line, we then extract non-overlapping $k \times k$ road patches along the seed line and use them as seed patches for identifying road patches. We found that k values in the range $[10, 50]$ show high accuracy, but we empirically chose $k = 22$ that strikes a good balance between the accuracy and performance for all the tests.

Starting from the seed patches, we propagate them by identifying similar, nearby patches. This propagation step is performed by considering the illumination-invariant image as well as a color image consisting of only HS (hue and saturation) components. The reason considering HS features additionally is that the illumination-invariant image does not provide the color information, and considering the color can differentiate different patches even though they have a similar illumination level. For example, suppose a road with curb. We found that it is difficult to differentiate the road area and the curb, since they tend to have the similar illumination level. By considering HS components, there is a higher chance to differentiate them during our patch propagation.

Given the input seed patches, we apply a simple and greedy propagation method to achieve a high runtime performance, summarized below:

1. Insert all the seed patches into a patch queue.
2. Dequeue the front patch and compare its similarity with nearby adjacent patches.
3. For similar patches, enqueue them into the patch queue.

Table 1. Measures of road detection performance.

Pixel-wise measure	**Definition**
Recall (R)	$R = \frac{TP}{TP+FN}$
Precision (P)	$P = \frac{TP}{TP+FP}$
F-measure (F)	$F = \frac{2PR}{P+R}$

4. Repeat step 2) and 3) until the patch queue is empty.

Each patch, I, is defined by $k \times k$ size with its center position $[c_x, c_y]$ in the illumination-invariant (II) and HS spaces, respectively. As a result, each patch is represented by a $3k^2$ dimensional feature vector. Given the patch I, its adjacent patches are defined in the same size along four diagonal directions from the center position of I. For the patch similarity test, we use the L2 distance between two patches. Since each patch is represented by three different components (II, H, and S), we normalize each component separately. This technique is known as the joint equal contribution technique [26]. We found that it is simple, yet effective in terms of improving the detection accuracy among many other alternatives (e.g., metric learning) for our problem.

We approach the problem of testing whether two patches are similar or not as statistical hypothesis testing (e.g., computing p-value). In other words, based on those initial seed patches, we know their similarity distribution, and we can compute a probability that a nearby patch belongs to the distribution of the seed patches. To perform the statistical hypothesis in an efficient manner, we first measure the average L2 distance, α, and its standard deviation, σ, between all the initial seed patches. We then treat a nearby patch, p, similar to be those seed patches, when p and its nearby seed patch have a L2 distance less than $\alpha + 1.1\sigma$. Once the patch p is identified to be similar to seed patches, we then incrementally re-compute α and σ by considering p as a newly added seed patch. We use the constant of 1.1 to allow minor patch variations for the similarity test, and found that it works well in the tested benchmark.

4. EVALUATION

We have implemented our illumination-invariant (II) space and patch propagation methods with MATLAB. We use Intel i7 machine with 3.4 GHz CPU for testing our approach and others, unless mentioned otherwise.

To evaluate benefits of the II space, we consider the space at a particular application, road detection, and thus tests all methods with the KITTI road dataset [16]. This road image data is acquired using an onboard Point Gray Flea video camera based on the Sony ICX204 sensor. The resolution of captured images is 1241 $\times$ 376. The data set contains 289 training images and 290 testing images. Only training images have their ground truth binary labeling, i.e., road or non-road per each pixel. The KITTI dataset is composed of three types of road structures: UM (urban marked), UMM (urban multiple marked), and UU (urban unmarked).

Table 2. Detection result (F-measure) of two shadow removal methods on the KITTI training dataset

Method	Overall	UM	UMM	UU	Time(ms)
Entropy based	83.19%	84.49%	86.12%	79.15%	24.896
Ours	83.25%	84.19%	86.73%	79.03%	1.129

We use F-measure to define the road detection accuracy, since it has been widely adopted as the most important measure [15]. This measure uses True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) (Table 1).

We compare our propagation method using II images with recent other bottom-up road detection algorithms that start from pixels or low-level features, and merge similar features:

- ANN [30]: This technique uses image segmentation based on watershed transform in 2D and label classification using v-disparity in 3D.
- Splane [8]: This method is based on a new scheme of block matching stereo that relaxes the fronto-parallel assumption using a disparity set.
- RES3D [27]: It supports fast obstacle estimation and road detection methods using stereo images.

Furthermore, to see benefits of our II space in top-down approaches for road detection methods, we apply the space as an additional feature space to recent CNN approaches considering context information:

- CN24 [5]: This method uses convolution neural net with spatial information, which serves as top-down information on images.
- FCN LN [24]: This approach explicitly utilizes the contextual information of given input patches to determine road area.

These two methods proposed their particular network architectures specialized in road detection for achieving high detection accuracy. At a high level, these network architectures adopted only several convolution and pooling layers (e.g., 4 to 6 layers) for high runtime performance, and chose to use the KITTI road datasets as the training dataset. They use 3-dimensional RGB images as inputs. Since the network architectures and training schemes of these approaches are available with high runtime performance and detection accuracy, we have implemented their methods and tested how the accuracy of these techniques behaves with the additional feature of our space. We have implemented these CNN approaches using TensorFlow r0.10 [9] and two of GTX TITAN X GPUs.

Table 4. Detection accuracy of different spaces used for our patch propagation with the KITTI training dataset

Method	Overall	UM	UMM	UU
HS color space	67.16%	63.28%	72.26%	66.08%
YCbCr color space	65.78%	68.12%	66.72%	62.59%
CIE Lab color space	64.85%	68.96%	63.22%	62.47%
Illumination-Invariant (II)	82.99%	84.17%	86.27%	78.72%
II + YCbCr	80.71%	84.00%	78.87%	79.42%
II + CIE Lab	77.31%	81.61%	73.25%	77.35%
Ours (II + HS)	**83.25%**	**84.19%**	**86.73%**	**79.03%**

Figure 7. Result of our approach. Green area is true positive, red area is false negative, and blue area is false positive.

Table 3. Road area detection accuracy in the online KITTI benchmark

Method	**Overall**	UM	UMM	UU
Ours	**81.23%**	81.04%	84.53%	78.30%
RES3D-Stereo [27]	**80.40%**	78.98%	83.62%	78.75%
Splane [8]	**77.83%**	78.19%	82.28%	73.30%
ANN [30]	**65.68%**	62.83%	80.95%	54.07%

Table 5. Detection accuracy of CNN-based methods considering top-down information

	CN24 [5]			FCN LC [24]		
	UM	UMM	UU	UM	UMM	UU
w/o II	88.94%	89.66%	80.45%	89.61%	92.98%	81.20%
w/ II	**89.00%**	**91.52%**	**80.64%**	**89.65%**	**93.11%**	**82.75%**

4.1. Illumination-invariant space evaluation

We compare our PCA-based variance minimization against the prior entropy-based approach. Unfortunately, there are no ground truth results for shadow removal in the tested road image dataset. As a second-to-best choice, we compare different methods based on road detection accuracy and computation time with the training dataset that has the ground-truth labels.

Overall, our proposed PCA-based method generates II images drastically faster even with a slightly higher accuracy compared to the entropy-based method (Table 2). Especially, our method performs 20 times faster on average than the prior one. This improvement is achieved mainly because our PCA based method finds the globally optimal projection direction without testing many projections. Thanks to its high accuracy and performance, we choose our PCA based method for computing II images and use them for our road detection.

4.2. Road detection evaluation

We report benefits of our II space both with bottom-up and top-down road detection approaches.

We first report the road detection accuracy of our simple propagation approach against other bottom-up methods. Road detection accuracy of our method is measured by submitting our detection results to the KITTI benchmark suite [15]. Detection accuracies of all the other tested methods in the bottom-up approach are adopted from the road detection homepage of the KITTI benchmark.

Accuracy of bottom-up methods. Table 3 shows F-measures of different methods with three types of roads, which are available at the benchmark suite [15]. Our method handles shadowed and non-shadowed road images well and achieves the highest detection accuracy on average compared to other tested methods. For the category of UM and UMM, our method achieves the highest accuracy, while achieving the second-to-best accuracy in category of UU. While our method shows slightly lower accuracy (0.45%) over RES3D-Stereo under the UU category, our method achieve higher accuracy in all the other cases. Figure 7 shows results of our proposed patch propagation.

While our method is implemented with MATLAB, it achieves fast processing time. The average processing time of our method breakdowns into the following components: 1ms to find the II projection line, 70ms to generate the II image, 280ms to bi-directional bird's-eye view conversion and find the seed line, and 100ms to 250ms to propagate patches for finding the entire road area. Overall processing time is 450 ms on average. We expect that we can achieve faster running time, when our method is implemented with C/C++ or in GPU.

Different color spaces. We have tested different color spaces to see their impacts on detection accuracy. Tested color spaces include HSI, YCbCr, and CIE Lab. We also combine our proposed II space with other color spaces to see their behaviors.

Table 4 shows F-measure values of different combinations of tested color spaces. The best performance is achieved by using our II space with HS color space. Note that we do not use the illumination component of the common HSI space, since our II space covers the same illumination after removing shadows. Interestingly, all of color space combinations achieve higher detection accuracy, when combined with our II space. It verifies again that removal of shadow influence by using the illumination-

(a) Light direction

(b) PCA-based illumination invariant image

(c) Entropy-based illumination invariant image

Figure 8. The red and green arrows show the sunlight direction and indirect indirection reflected by the white wall, respectively.

invariant space is very critical for achieving high road detection accuracy.

Accuracy of top-down approaches. We also apply the II space as an additional new feature to two recent CNN top-down road detection methods. Table 5 shows road detection accuracy w/ and w/o using our II feature. For methods w/ using our II feature, we concatenate the original RGB and the II image as the fourth dimension to train CNNs w/o changing any other network architectures. By performing detection with the II image as a new feature, we can see the improvement of detection accuracy across all the tested cases. This consistent accuracy improvement supports benefits of our illumination-invariant space again. Furthermore, considering the additional feature requires only a minor, 10% higher, computational overhead over those techniques that do not the feature.

Figure 1 shows visualization comparison of two CNN methods w/ and wi/o considering the II space. Those two CNN methods considering the additional II space detect road area more robustly over the original ones that do not consider the additional feature.

Note that the overall accuracy reported in this paper is not the highest one reported in the detection homepage of the KITTI benchmark. Unfortunately, techniques showing the highest accuracy are hidden and submitted by anonymous authors. As a result, we took our two tested CNN methods, which were published in the academic field with their available network architectures and training schemes. We were thus able to re-implement their approaches and to test our II space within these methods.

Indirect illumination. Our method shows improved accuracy in both bottom-up and top-down road detection methods. Nonetheless, we also found that the illumination-invariant space is not perfect for all the cases. Especially, we found that our method does not handle the effect of indirect lighting well. Figure 8 shows an example of indirect lighting, which is reflected from the white wall. In this case, our method does not remove all the shadow effects. Nonetheless, we also found that the prior, entropy method does not handle the case properly either, and our method shows higher detection accuracy overall.

5. CONCLUSION

In this paper, we have proposed a fast, yet accurate PCA-based computation for the illumination-invariant space. The computed illumination-invariant space removes shadows and other illumination changes on road area from an input RGB image. We showed its benefits in a simple, patch-based road detection method, running in a bottom-up manner, compared to other recent bottom-up road detection methods. Furthermore, we applied the illumination-invariant space as an additional feature to recent CNN-based road detection methods, which also utilize context information and have benefits of top-down approaches. We were able to observe that our feature improves the overall road detection method over those two CNN methods only with a minor computational overhead.

As future work, we would like to address hard cases for our method. Especially, our current approach did not consider about global and local indirect illuminations. We expect to obtain more accurate illumination-invariant space, when we consider such indirect illumination. Finally, we would like to combine top-down and bottom-up approaches together to fully utilize various information available at the original input RGB space and our illumination-invariant space.

Acknowledgements

We would like to thank anonymous reviewers for constructive comments. S.-E. Yoon is a corresponding author of the paper. This project was supported in part by MSIP/IITP (R0101-16-0176) and MSIP/NRF (No. 2013-067321).

References

[1] J. M. Alvarez, T. Gevers, Y. LeCun, and A. M. Lopez. Road scene segmentation from a single image. In *Computer Vision–ECCV 2012*, pages 376–389. Springer, 2012.

[2] J. M. Alvarez, A. López, and R. Baldrich. Illuminant-invariant model-based road segmentation. In *Intelligent Vehicles Symposium, 2008 IEEE*, pages 1175–1180. IEEE, 2008.

[3] J. M. Álvarez and A. M. Ĺopez. Road detection based on illuminant invariance. *Intelligent Transportation Systems, IEEE Transactions on*, 12(1):184–193, 2011.

[4] H. Barrow and J. Tenenbaum. Recovering intrinsic scene characteristics. *Comput. Vis. Syst., A Hanson & E. Riseman (Eds.)*, pages 3–26, 1978.

[5] C.-A. Brust, S. Sickert, M. Simon, E. Rodner, and J. Denzler. Convolutional patch networks with spatial prior for road detection and urban scene understanding. *arXiv preprint arXiv:1502.06344*, 2015.

[6] K.-H. Chen and W.-H. Tsai. Vision-based autonomous land vehicle guidance in outdoor road environments using combined line and road following techniques. *Journal of Robotic Systems*, 14(10):711–728, 1997.

[7] P. Corke, R. Paul, W. Churchill, and P. Newman. Dealing with shadows: Capturing intrinsic scene appearance for image-based outdoor localisation. In *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, pages 2085–2092. IEEE, 2013.

[8] N. Einecke and J. Eggert. Block-matching stereo with relaxed fronto-parallel assumption. In *Intelligent Vehicles Symposium Proceedings, 2014 IEEE*, pages 700–705. IEEE, 2014.

[9] M. A. et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015.

[10] O. Falola, I. Osunmakinde, and A. Bagula. Supporting drivable region detection by minimising salient pixels generated through robot sensors. *PRASA*, 2010.

[11] G. D. Finlayson and M. S. Drew. 4-sensor camera calibration for image representation invariant to shading, shadows, lighting, and specularities. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 473–480. IEEE, 2001.

[12] G. D. Finlayson, M. S. Drew, and C. Lu. Intrinsic images by entropy minimization. In *Computer Vision-ECCV 2004*, pages 582–595. Springer, 2004.

[13] G. D. Finlayson, M. S. Drew, and C. Lu. Entropy minimization for shadow removal. *International Journal of Computer Vision*, 85(1):35–57, 2009.

[14] G. D. Finlayson, S. D. Hordley, C. Lu, and M. S. Drew. On the removal of shadows from images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(1):59–68, 2006.

[15] J. Fritsch, T. Kuhnl, and A. Geiger. A new performance measure and evaluation benchmark for road detection algorithms. In *Intelligent Transportation Systems-(ITSC), 2013 16th International IEEE Conference on*, pages 1693–1700. IEEE, 2013.

[16] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013.

[17] S. Graovac and A. Goma. Detection of road image borders based on texture classification. *International Journal of Advanced Robotic Systems*, 9(242), 2012.

[18] Y. He, H. Wang, and B. Zhang. Color-based road detection in urban traffic scenes. *Intelligent Transportation Systems, IEEE Transactions on*, 5(4):309–318, 2004.

[19] Z. He, T. Wu, Z. Xiao, and H. He. Robust road detection from a single image using road shape prior. In *Image Processing (ICIP), 2013 20th IEEE International Conference on*, pages 2757–2761. IEEE, 2013.

[20] T. Kühnl, F. Kummert, and J. Fritsch. Monocular road segmentation using slow feature analysis. In *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pages 800–806. IEEE, 2011.

[21] K. Lu, J. Li, X. An, and H. He. A hierarchical approach for road detection. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 517–522. IEEE, 2014.

[22] W. Maddern, A. Stewart, C. McManus, B. Upcroft, W. Churchill, and P. Newman. Illumination invariant imaging: Applications in robust vision-based localisation, mapping and classification for autonomous vehicles. In *Proceedings of the Visual Place Recognition in Changing Environments Workshop, Robotics and Automation (ICRA), 2014 IEEE International Conference*, volume 2, page 3, 2014.

[23] S. Maldonado-Bascón, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gómez-Moreno, and F. López-Ferreras. Road-sign detection and recognition based on support vector machines. *Intelligent Transportation Systems, IEEE Transactions on*, 8(2):264–278, 2007.

[24] C. C. T. Mendes, V. Frémont, and D. F. Wolf. Exploiting fully convolutional neural networks for fast road detection. *Image*, 16(1000x1x1):16x15x15, 2016.

[25] 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.

[26] S. Sarin and W. Kameyama. Joint equal contribution of global and local features for image annotation. In *CLEF (Working Notes)*, 2009.

[27] P. Y. Shinzato, D. Gomes, and D. F. Wolf. Road estimation with sparse 3d points from stereo data. In *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*, pages 1688–1693. IEEE, 2014.

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

[29] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection using evolutionary gabor filter optimization. *Intelligent Transportation Systems, IEEE Transactions on*, 6(2):125–137, 2005.

[30] G. B. Vitor, D. A. Lima, A. C. Victorino, and J. V. Ferreira. A 2d/3d vision based approach applied to road detection in urban environments. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*, pages 952–957. IEEE, 2013.

[31] Y. Weiss. Deriving intrinsic images from image sequences. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 68–75. IEEE, 2001.