

# Full-Time Monocular Road Detection Using Zero-Distribution Prior of Angle of Polarization

Ning Li<sup>1</sup>[0000-0001-6009-6831], Yongqiang Zhao<sup>1</sup>[0000-0002-6974-7327], Quan Pan<sup>1</sup>[0000-0001-8162-2896], Seong G. Kong<sup>2</sup>[0000-0002-0335-6526], and Jonathan Cheung-Wai Chan<sup>3</sup>[0000-0002-3741-1124]

<sup>1</sup> Northwestern Polytechnical University, Xi'an, 710072, China

ln\_neo@mail.nwpu.edu.cn, {zhaoyq, quanpan}@nwpu.edu.cn

<sup>2</sup> Sejong University, Seoul, 05006, South Korea

skong@sejong.edu

<sup>3</sup> Vrije Universiteit Brussel, Brussel, 1050, Belgium

jcheungw@etrovub.be

**Abstract.** This paper presents a road detection technique based on long-wave infrared (LWIR) polarization imaging for autonomous navigation regardless of illumination conditions, day and night. Division of Focal Plane (DoFP) imaging technology enables acquisition of infrared polarization images in real time using a monocular camera. Zero-distribution prior embodies the zero-distribution of Angle of Polarization (AoP) of a road scene image, which provides a significant contrast between the road and the background. This paper combines zero-distribution of AoP, the difference of Degree of linear Polarization (DoP), and the edge information to segment the road region in the scene. We developed a LWIR DoFP Dataset of Road Scene (LDDRS) consisting of 2,113 annotated images. Experiment results on the LDDRS dataset demonstrate the merits of the proposed road detection method based on the zero-distribution prior. The LDDRS dataset is available at <https://github.com/polwork/LDDRS>.

**Keywords:** Road detection, Polarization prior, Angle of polarization, LWIR DoFP sensor

## 1 Introduction

Road detection is a crucial task for traffic safety and intelligent transportation systems such as Advanced Driver Assistant System (ADAS) [17]. Various sensing modalities such as vision [21], [32], [42], [49] and LiDAR [26], [29] have been used for this purpose. Vision-based methods include passive monocular [21], [32] and the stereo imaging [42], [49], which provides high resolution color, texture and lane marking information for road detection. Stereo imaging can obtain depth information, but still less information than 3D LiDAR. Vision-based methods, however, are unable to perform properly in low illumination conditions, darkness, and strong illumination variations such as headlight of incoming cars at

night. LiDAR is an active modality for lane and road detection, measuring 3D structure of a vehicle nearby using an active light source enabling LiDAR to be employed all the time, day and night. A major drawback of LiDAR is relatively high cost and bulkiness. While thermal infrared cameras are often used for pedestrian/animal detection [5], [23], but not for road detection since radiation difference between road and the background is too small especially at night when thermal equilibrium is reached.

Polarization has a physical property of light that provides characteristic information of an object such as three-dimensional (3D) normal [38], surface smoothness [43], and material composition [18]. Polarization imaging has been widely applied in 3D reconstruction [9], anti-interference object detection [24], visual navigation [40], image dehazing [41], and biomedical imaging [13]. To our best knowledge, no published literatures have investigated the use of polarization for road detection. Division of Focal Plane (DoFP) imaging technology [12], [44] enables acquisition of infrared polarization images in real time with a monocular camera. A DoFP infrared polarization camera consists of traditional infrared focal plane and a micro polarizer array which captures polarization information in real time. DoFP cameras have similar dimensions, weight, and power consumption as traditional uncooled thermal cameras, suitable for installing on vehicles in real situations. We use a long-wave infrared (LWIR) DoFP polarimeter as sensing modality to obtain polarization characteristics of the road.

This paper presents a road detection method based on zero-distribution prior of the road. Zero-distribution prior embodies the zero-distribution of Angle of Polarization (AoP) of the road region, which provides a significant contrast between the road and the background. Using this prior, a coarse mask for the road region is obtained. Then we propose a statistical method with road vanishing assumption to locate the horizon. The road area is located usually in the lower part of the horizon to significantly reduce computational burden. This paper combines zero-distribution of AoP, the difference of Degree of linear Polarization (DoP), and the edge information to segment the road region in the scene. Then we refine the detection result by removing fragmented parts using a confidence map. The definitions of AoP and DoP can be found in the Supplementary Material. To evaluate the proposed method, we captured thousands of LWIR DoFP images of various road scenes, including urban roads and highways during day and night times. A total of 2,113 annotated LWIR DoFP images forms a benchmark database named LWIR DoFP Dataset of Road Scenes (LDDRS).

As contributions, this paper proposes: (1) A zero-distribution prior to represent distinctive characteristics of the road; (2) A statistical method to detect the horizon; (3) A joint road detection approach to segment the road; (4) A database of LWIR DoFP images for benchmarking road detection algorithms.

## 2 Related Works

**Thermal sensor:** As the most related work, LWIR imagers are mainly used in ADAS [5] to detect pedestrians or animals at night since emitted radiation of

human is greater than the environment. Several methods have been developed to detect the pedestrian, including enhanced thresholding segmentation technique [36], adaptive fuzzy C-means clustering, and CNN based approach [19]. LWIR imagers are also used to improve the road detection in low light conditions by thermal-RGB fusion [48], stereo thermal cameras [34], or multi-frames information [47]. Yet thermal characteristics between road and the background may reach equilibrium at night.

**Monocular vision:** Many road detection methods have been developed for conventional RGB cameras based on high-resolution intensity, color, and texture information. Some constrain the road region by detecting lane markings. Wang *et al.* [46] use a straight-curve model to detect curvy road in highways. Some road models are based on the statistics or shape information to segment the road. Lu [30] proposed a self-learned statistical model to re-label each pixel in the input image based on a likelihood ratio classifier. Other road segmentation methods use vanishing point to constrain road boundary. For example, a Gabor filter [21] and a gLoG filter [22] are used to locate the vanishing point of road with a voting-based scheme. Inspired by the Fully Convolutional Network (FCN) [28] designed for semantic segmentation task, several road detection methods are proposed based on deep learning including Up-Conv-Poly [32], FCN-LC [31] and DLT-Net [37]. Utilizing the pixel-wise classification of FCN, these methods have enhanced road detection performances. However, the inherent drawback of visible camera in low light condition limits the performance of a monocular vision system at night or in low illumination conditions.

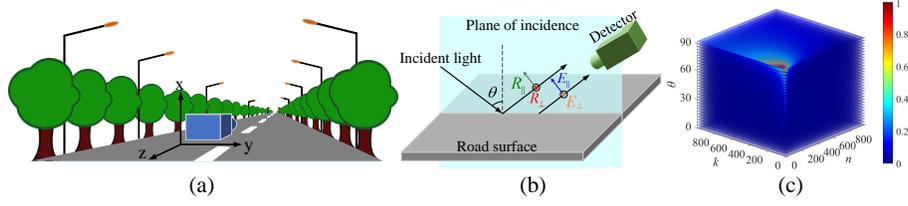
**Stereo imaging & LiDAR:** A disparity map [14] generated by stereo matching of two captured visible images can provide depth information to play a key role in road detection. The traditional vanishing-point-constraint methods [33], [49] and deep learning method [39] based on a disparity map are developed in stereo vision system. Unfortunately, this approach also suffers from the same limitation with the monocular vision system in low light environments. LiDAR, on the other hand, uses an active light source to reconstruct 3D representation of an object. As an active perception modality, LiDAR provides a strong cue to detect road as used in several works [16], [26], [29], [50]. However, the detection accuracy is easily affected by dust, haze or rain [35].

### 3 Zero-distribution Prior

The zero-distribution prior is based on the theory that thermal emissions are partially linearly polarized parallel to the plane of incidence [9]. For a road region, most values of AoP are near zero,

$$A(x) \approx 0, x \in \Omega \quad (1)$$

where  $A$  denotes an AoP image and  $\Omega$  denotes a set of pixels in the road region. This characteristic of AoP in the road region is called *zero-distribution prior of AoP*. In Fig. 1(a), we assume that the road plane is parallel with the  $y$  axis in the image coordinate system, so zero-distribution of AoP means that the thermal

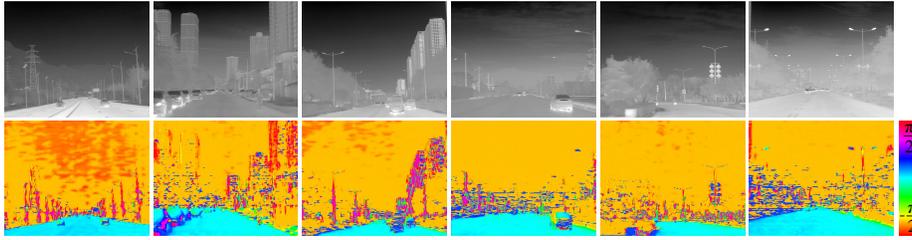


**Fig. 1.** (a) The image coordinate system of LWIR DoFP camera. (b) Mode of how the thermal radiation captured by detector is formed in LWIR. (c)  $\varepsilon_{\parallel}(\theta, n')^2 - \varepsilon_{\perp}(\theta, n')^2$  vary with respect to  $\theta$ ,  $n$  and  $k$

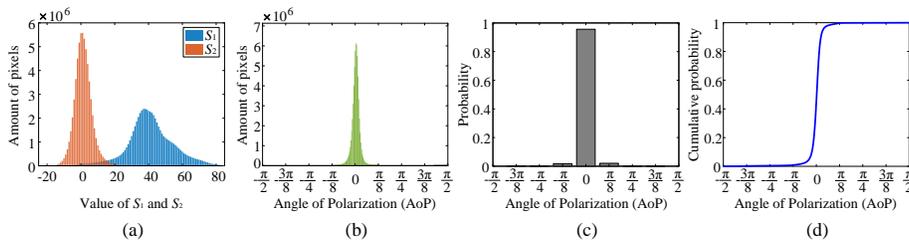
radiations of road are partially linearly polarized parallel to  $x$ - $z$  plane of an incident light in Fig. 1(b). Therefore, the Stokes parameter  $S_1$  of road should be positive and  $S_2$  should near zero (see Supplementary Materials for definition of  $S_1$  and  $S_2$ ). Fig. 1(b) shows that the thermal radiation reaches the DoFP LWIR detector contains reflected radiation  $R$  and emitted radiation  $E$ , and both are expressed as the sum of two orthogonal polarized components, i.e.  $R = R_{\parallel} + R_{\perp}$  and  $E = E_{\parallel} + E_{\perp}$ . For the road scene, the emitted thermal radiation dominates the energies that reach the camera. In [4],  $S_2$  of thermal emissions have been demonstrated to be zero for objects that are large compared to the emitted wavelength [20], so here we only focus on the sign of  $S_1$ . Based on Kirchoff's law [45] and Fresnel's equations [15], we have (See Supplementary Material for derivation)

$$S_1 = n \cos \theta \cdot P(T_1)^2 \cdot (\varepsilon_{\parallel}(\theta, n')^2 - \varepsilon_{\perp}(\theta, n')^2) \quad (2)$$

where  $P(T_1)$  denotes the Planck Blackbody radiance curve at temperature  $T_1$  of road,  $\varepsilon$  is the emissivity and the special subscripts  $\parallel$  and  $\perp$  are added to  $\varepsilon$  that correspond to respectively the polarized components parallel and perpendicular to the plane of incidence, and  $\varepsilon$  is decided by incident angle  $\theta$  and complex index of refraction  $n' = n + ki$  of road. So the sign of  $S_1$  is decided by the sign of  $\varepsilon_{\parallel}(\theta, n')^2 - \varepsilon_{\perp}(\theta, n')^2$  because  $n > 0$  and  $\cos \theta > 0$  ( $\theta \in (0, 90^\circ)$ ). Since the exact refraction of the road is unknown, the intervals are set to  $n \in (0, 1000]$  and  $k \in (0, 1000]$ . We calculate the value of  $\varepsilon_{\parallel}(\theta, n')^2 - \varepsilon_{\perp}(\theta, n')^2$  and plot the result with  $\theta$ ,  $n$  and  $k$  in Fig. 1(c). The quantity  $\varepsilon_{\parallel}(\theta, n')^2 - \varepsilon_{\perp}(\theta, n')^2$  is positive in a sufficiently large interval, so  $S_1$  is also positive. We further make a statistical analysis of the distribution of  $S_1$  and  $S_2$  in road region of 2,113 images in Fig. 3(a) where basically all the values of  $S_1$  in road region are positive and most values of  $S_2$  are around zero. Since the  $S_1$  in road region is positive and  $S_2$  is close to zero, the AoP of road is zero distributed. Fig. 2 shows several road scene images and the corresponding AoP images in false color. A statistical analysis of the distribution of AoP  $A$  in road region of the 2,113 images produces the results in Fig. 3(b)-(d). Approximately 96% of the pixels of road in the AoP images have values in  $(-\frac{\pi}{16}, \frac{\pi}{16})$ , so most AoP values in road region are near zero. Using this property, more generally, the AoP of each position in a plane is the same. Intuitively, the proposed prior can also be effective for other applications such



**Fig. 2.** Top: example images in our DoFP road scene database. Bottom: the corresponding AoP images



**Fig. 3.** (a) The distribution of  $S_1$  and  $S_2$  in road region of the 2,113 road images. (b) Histogram of the AoP of the pixels in road region of all the 2,113 road images, and (c) Corresponding probability distribution and (d) Cumulative distribution

as plane detection or defect detection in roads (e.g. potholes or obstacles). In Section 4, this prior is used to detect the horizon and segment the road.

## 4 Road Detection with Zero-distribution Prior

The proposed zero-distribution prior poses a strong constraint on road detection. Fig. 4 shows the proposed road detection method based on the zero-distribution prior. An input DoFP image is denoised [1] and demosaicked [25] to generate four high resolution polarization images and then the Stokes parameters, AoP and DoP are computed (See Eqs. (1) and (2) in Supplementary Material). We use the zero-distribution prior for horizon detection, and the road region is restricted under the horizon in the image. Then we propose a joint road detection method that combines AoP, DoP, and the edge information to segment the road. The final road detection result is obtained by refining a road confidence map.

### 4.1 Horizon Detection

A coarse road map is obtained based on the strong constraint of zero-distribution prior. Since the value of AoP in road region is near zero, we find a coarse road estimation using

$$R_c(x) = \exp(-\gamma|A(x) - \sigma|) \quad (3)$$

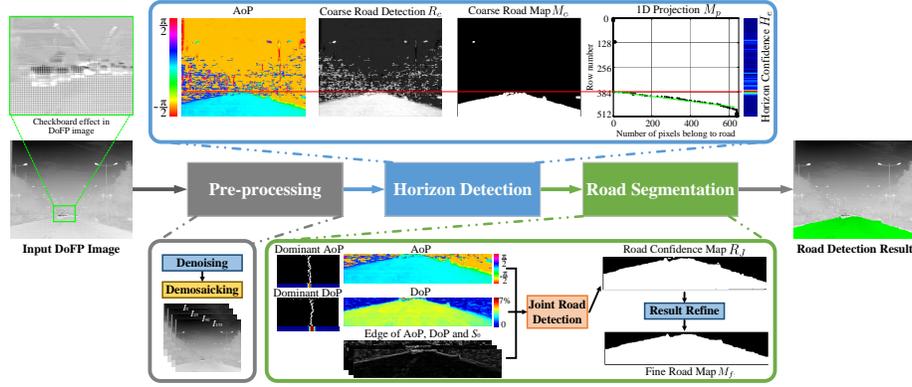


Fig. 4. The proposed road detection method based on the zero-distribution prior

where  $x$  is pixel position,  $\gamma$  was set to 0.01 and  $\sigma$  was set to 0 in the experiment based on the zero-distribution prior. Then, we threshold  $t = 0.75$  as the road, followed by a morphological open operation to remove the small noise regions:

$$M_c = \mathfrak{S}(R_c, t) \odot X, \quad \mathfrak{S}(R_c, t) = \begin{cases} 1, & R_c \geq t \\ 0, & R_c < t \end{cases} \quad (4)$$

where  $\odot$  is the morphological open operation and  $X$  is the structural element.  $M_c$  provides a coarse road detection result, and a horizon locates where the road vanishes. As the road vanishes in Fig. 4, the number of pixels belonging to road in each row decreases from bottom to top. We make a one-dimensional (1D) projection of the coarse road map  $M_c$  by

$$M_p(r) = \sum_{j=1}^w M_c(r, j) \quad (5)$$

where  $w$  is the width of the image,  $r$  is the row number and  $M_p$  represents the number of pixels that belong to road in each row in  $M_c$ . In Fig. 4, the horizon corresponds to the intersection of the green line and the  $x$ -axis in  $M_p$ . Horizon can be obtained by locating where they intersect. Rather than fitting the straight line in  $M_p$  directly, we propose to find the position in  $x$ -axis where the lines (determined by every two points that have fixed rows apart in  $M_p$ ) intersect with  $x$ -axis most. To realize this, a horizon confidence map is generated by

$$H(r) = \sum_{l=1}^{h-s} [\varphi((l, M_p(l)), (l+s, M_p(l+s))) = r] \quad (6)$$

where  $h$  is the height of the image,  $l$  is the row number and  $\varphi(\mathbf{x}, \mathbf{y})$  indicates the row where the straight line (determined by  $\mathbf{x}$  and  $\mathbf{y}$ ) intersects with the  $x$ -axis in  $M_p$ , and here we choose every two points that have two rows apart, that is

$s = 3$ .  $[\cdot]$  equals 1 if the value inside the bracket is true and 0 otherwise. One may notice that the biggest value or the dominant values in  $H$  may lie several pixels away from the ground-truth horizon. To overcome this problem, we propose a simple post-processing of  $H$  by

$$H_e(r) = \sum_{i=r-\rho}^{r+\rho} H(i) \quad (7)$$

where  $H_e$  is the piecewise cumulative energy of  $H$ , and we set  $\rho = 3$ . With this simple step, we improve the robustness of horizon detection and we can locate the horizon by finding the biggest energy in  $H_e$ , as shown by the red line in Fig. 4. The proposed horizon detection method is efficient and robust, and the road is restricted in the lower part of the image.

## 4.2 Road Segmentation

In Section 4.1, we generate a coarse road detection result with the strong constraint of zero-distribution prior, and the horizon restrict the road in a small part of the image which reduces the computation load. The AoP for car surface are similar to that of the road as in Fig. 2 and Fig. 6(c). Therefore, cars can be misclassified as the road if we only use the zero-distribution prior. There exist some differences in DoP between the road and cars. The DoP of the hood and windshield of a car is usually higher than the DoP of the road, while the other surfaces of a car are always lower than that of the road as shown in Fig. 6(b). This observation agrees with previous research for vehicle detection [10], [11]. We selected 200 images that contain various vehicles and computed the DoP distribution of the road and cars. Fig. 5 shows that clear differences of DoP exist between road and vehicle (See Supplementary Material for details). This characteristic helps separate the vehicle and road. This paper combines the AoP, DoP and the edge information for road detection. Inspired by [27], the road confidence map is obtained by the proposed joint road detection as:

$$R_J(x) = \frac{2}{1 + \exp[\eta(1 + \eta_1 C_E(x))(C_A(x) + C_D(x))]} \quad (8)$$

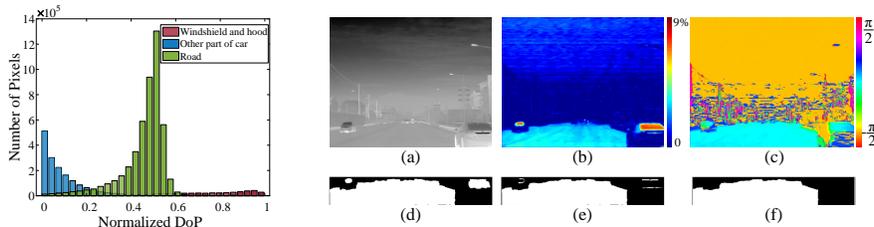
where

$$C_A(x) = \begin{cases} \exp[\eta_2(|A(x) - A_d| - \alpha_1)], & A(x) \geq A_d \\ \exp[\eta_2(|A(x) - A_d| - \alpha_2)], & A(x) < A_d \end{cases} \quad (9)$$

$$C_D(x) = \begin{cases} \exp[\eta_3(|D(x) - D_d| - \beta_1)], & D(x) \geq D_d \\ \exp[\eta_3(|D(x) - D_d| - \beta_2)], & D(x) < D_d \end{cases} \quad (10)$$

$$C_E(x) = \omega_1 E_D(x) + \omega_2 E_A(x) + \omega_3 E_I(x) \quad (11)$$

where  $C_A$  is the constraint of AoP image  $A$ ,  $C_D$  is the constraint of DoP image  $D$ , and  $C_E$  is the edge constraint. In Eq. (11),  $E_D$ ,  $E_A$  and  $E_I$  are edge information of DoP image, AoP image and the intensity image ( $S_0$ ), respectively, and  $\omega_1$ ,  $\omega_2$  and  $\omega_3$  are corresponding weights ( $\omega_1 = 0.3$ ,  $\omega_2 = 0.5$  and  $\omega_3 = 0.2$ ).  $\eta$ ,  $\eta_1$ ,  $\eta_2$

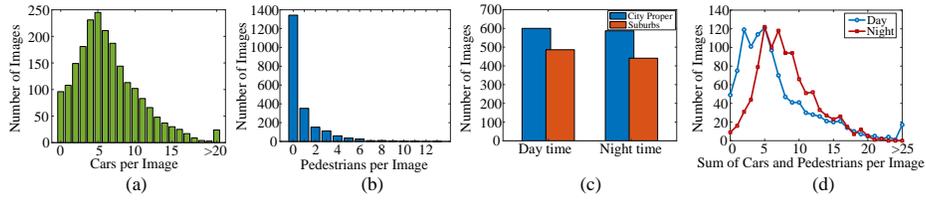


**Fig. 5.** Histogram of range **Fig. 6.** (a) Intensity image. (b) DoP image in false color. normalized DoP in road (c) AoP image in false color. (d) Coarse road map. (e) Road and different part of vehicle confidence map. (f) Refined road detection result. (d)-(f) regions of 200 road scenes only shows the region below the horizon

and  $\eta_3$  are weights of above three constraints ( $\eta = 10^{-7}$ ,  $\eta_1 = 1.9$ ,  $\eta_2 = 0.2$  and  $\eta_3 = 0.2$  in our experiment).  $A_d$  and  $D_d$  are dominate values of AoP and DoP in road region respectively obtained by the statistics of  $A$  and  $D$  of the road region in coarse road map  $M_c$ . The bias terms  $\alpha_1 = \frac{\pi}{50}$ ,  $\alpha_2 = \frac{\pi}{16}$ ,  $\beta_1 = \beta_0 + 0.02$ ,  $\beta_2 = \beta_0 + 0.12$  and  $\beta_0$  is set as the half of the distribution range of the most DoP in road region. With Eq. (8), a pixel belongs to road will get a higher value near 1, and near 0 if otherwise. One can notice that for the constraint of AoP and DoP images in Eqs. (9) and (10), we punish the values greater than  $A_d$  and  $D_d$  more, because the shadow or wet areas of road usually have lower AoP and DoP values. The edges of DoP image, AoP image and the intensity image provide the road boundary information, so  $C_E$  is used to separate the road with other region by giving more punishment on strong edge.

With the road confidence map  $R_J$ , we can get a refined road map. First, the pixels in  $R_J$  are binarized by  $\mathfrak{S}(R_J, \tau)$  and  $\tau = 0.95$ , and we assume that  $m$  separated regions  $K(i)$ ,  $i = 1 \dots m$  are obtained. Then remove all the small pieces whose areas  $N_{K(i)}$  (defined as the number of pixels belong to  $K(i)$ ) are smaller than 2% of the total detected area  $N_t = \sum_{i=1}^m N_{K(i)}$ . We take the biggest one in  $K(i)$  as road and we can compute the mean DoP  $D_m$  and mean intensity  $I_m$  of road region. The remaining small pieces whose mean DoP or intensity value are different from  $D_m$  or  $I_m$  more than the preset threshold will be simply removed ( $\tau_I = 40$  in 8 bit image and  $\tau_D = \beta_0$ ). Finally, the refined road map  $M_f$  is obtained by applying the background region growing to fill the holes in road region.

For the road scene in Fig. 6, the front and back windshields are detected as road in the coarse road map (Fig. 6(d)). But these false detections are mostly removed with the proposed joint road detection method (Fig. 6(e)). And the refined detection result is obtained after removing small speckles in the road confidence map (Fig. 6(f)). The proposed joint road detection is applied only in the part under the horizon detected in the last section. The proposed road detection method is simple yet efficient and robust with the constraint of the zero-distribution prior. The performance of the proposed method will be demonstrated in Section 5.

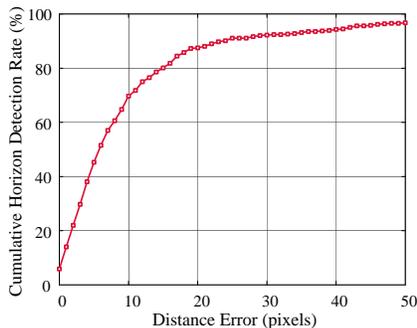


**Fig. 7.** Object geometry statistics of LDDRS. Histograms for the two most predominant categories (a) cars and (b) pedestrians. (c) Data distribution from day to night in city proper and suburbs. (d) Histograms for sum of cars and pedestrians at day and night

## 5 Experiment Results

**Dataset.** The proposed method is based on LWIR polarization imaging. To our best knowledge, there is no existing dataset with this specific information/feature available for the work. The only related LWIR dataset is described in KAIST [8]; its images of road are with only intensity information, no polarization. To test the proposed method, we build a LWIR DoFP dataset of road scene (LDDRS) with 2,113 images which provides both IR intensity ( $S_0$ ) and polarization information. The DoFP images are captured with a self-developed uncooled infrared DoFP camera with  $512 \times 640$  resolution in 14 bits. The dataset includes urban road and highway both day and night. The road regions of all 2,113 images are manually annotated. Statistics of our LDDRS are shown in Fig. 7. LDDRS contains different traffic situations that there are different number of cars and pedestrians in road scene. For the input DoFP images, a BM3D method [1] is used to reduce the noise and then a polarization demosaicking method [25] with a polarization difference model is applied to generate four high resolution images in four orientations. Based on these four images, the AoP, DoP and intensity images can be obtained by using the Stokes equations [15].

**Evaluation on horizon detection.** First, we evaluate the performance of our proposed horizon detection method. The error between the estimation and the ground-truth horizon is defined as the absolute difference in their vertical coordinates. Fig. 8 shows quantitative assessment of the proposed method. The horizontal axis of Fig. 8 represents the distance errors, where about 80% of horizon detection results have distance error smaller than 15 pixels. The horizon detection is used to improve the road detection robustness and effectiveness. To further test the necessity of horizon detection, we evaluate the road detection performance with and without horizon constraint. And three evaluation criteria are used including precision (PRE), recall (REC), and Intersection over Union (IoU) to assess the performance of road detection. Table. 1 shows that the road detection performance is improved by over 20% and execution efficiency increased by about three times when with horizon constraint, which demonstrates the necessity of horizon detection.



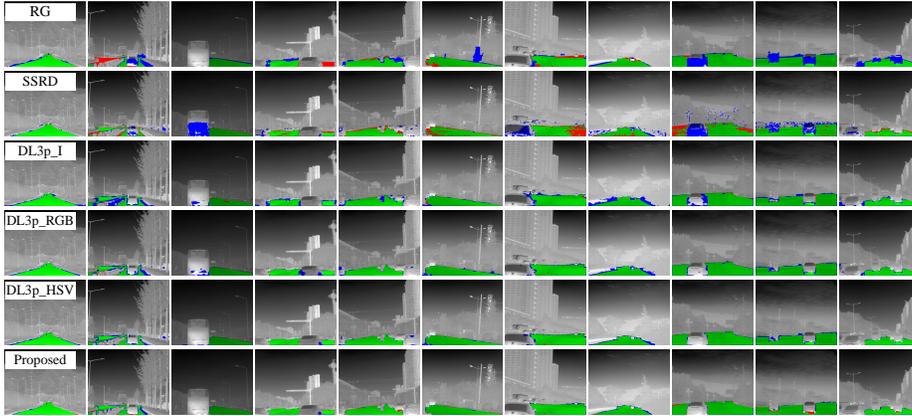
**Fig. 8.** Statistics of the horizon estimation accuracy of the proposed method

**Table 1.** Results of road detection with and without horizon constraint

Methods	PRE (%)	REC (%)	IoU (%)	Running Time (ms)
w/o horizon constraint	72.65	71.87	67.52	194
with horizon constraint	93.48	94.50	88.69	67

**Evaluation on road segmentation.** To evaluate the performance of different road detection methods, PRE, REC and IoU are used as assessment criteria. For comparison, three kinds of single image road detection methods are chosen. First kind of road detection method is designed especially for thermal infrared sensors such as the region-growing based method [47] which applies temporal information of sequential frames to refine the drivable region detection. The second kind of road detection method uses a predefined road mask to build a statistical model for road and background such as SSRD [30] which applies a self-learned statistical model to re-label each pixel in the input image base on a likelihood ratio classifier. The third kind is the deep-learning based semantic segmentation methods such as DeepLabv3 series [6], [7], encoder-decoder network [3], FCN [28] and so on, are most powerful methods for pixelwise classification. Here we use the DeepLabv3+ [7] in our dataset to test the performance of road segmentation.

The region-growing based method [47] is applied to the intensity image, as denoted as RG. The SSRD method is applied on the HSV fusion [2] image (Fig. 10(b)) of intensity image, AoP image and DoP image. The DeepLabv3+ is test on the intensity image, RGB fusion image (fusion of intensity image, AoP image and DoP image in R, G, B channels respectively) and the HSV fusion image and are denoted as DL3p\_I (DeepLabv3+ with the intensity image), DL3p\_RGB (DeepLabv3+ with the RGB fusion image) and DL3p\_HSV (DeepLabv3+ with the HSV fusion image) respectively. For DeepLabv3+, ResNet-101 is used as network backbone. And for dataset split, 1690 images are randomly selected for training, 211 images are used for validation and the rest 212 images are used for test. Considering the limited size of our LDDRS dataset, the DeepLabv3+ is initially trained with KAIST thermal infrared road scene dataset [47], and



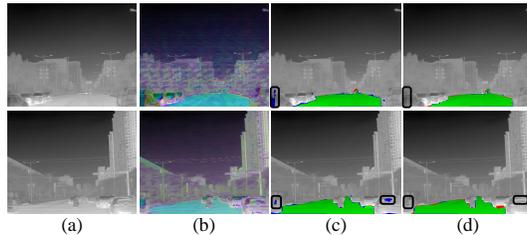
**Fig. 9.** Comparison of the road detection accuracy. Green region corresponds to true positive. Red region represents false negative. Blue region denotes false positive

**Table 2.** Results comparison of different road detection methods, the best results are shown in bold

Methods	PRE (%)	REC (%)	IoU (%)
RG [47]	74.68	87.40	67.69
SSRD [30]	82.45	83.76	71.60
DeepLabv3+ [7] with intensity image	79.65	<b>99.35</b>	79.28
DeepLabv3+ [7] with RGB fusion image	87.88	98.65	86.90
DeepLabv3+ [7] with HSV fusion image	89.34	99.19	88.73
Proposed method on test set	93.19	94.84	<b>88.74</b>
Proposed method on whole set	<b>93.48</b>	94.50	88.69

then the final model is fine-tuned using our training images. The quantitative assessments for these methods are shown in Table 2. Fig. 9 shows some visual comparison results of road region detection by the above methods, and for all the visual results, green region corresponds to true positive, red region represents false negative and blue region denotes false positive.

Table 2 and Fig. 9 show that the proposed method outperforms conventional techniques, and achieve comparable results with the state-of-the-art deep learning method DeepLabv3+. The recall of our method is not as high as other methods because our scheme tends to detect the road as accurate as possible and the low confidence pixels and small pieces that belong to road are discarded in the refinement process, which leads to a lower recall. Besides, the proposed method achieves highest precision especially for the separation between car and road. For example, as shown in the regions delineated in black rectangles in Fig. 10, the windshield of car and road are similar in the HSV fusion image and low occurrence of vehicle hoods compared with the road surface, making this particularity difficult to learn, so the DL3p\_HSV method may wrongly detect



**Fig. 10.** (a) Intensity image. (b) HSV fusion image. (c) Results based on DeepLabv3+ with the HSV image. (d) Ours results

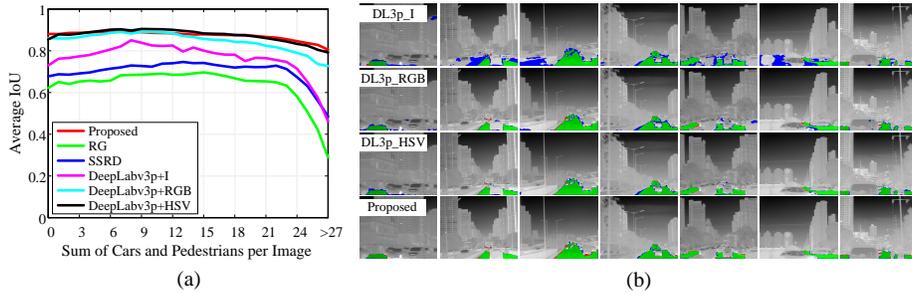
**Table 3.** Comparison of AoP-DoP-edge constraint and AoP constraint only

Constraints	PRE (%)	REC (%)	IoU (%)
AoP only	90.21	94.12	85.41
AoP+DoP+Edge	93.48	94.50	88.69

this car part as road while our method performs well. Methods using intensity image, including RG and DeepLabv3+, have good recall performance but poor precision, because it is difficult to separate road and cars or the background when they have similar thermal radiation. In contrast, methods apply polarization information such as SSRD, DeepLabv3+ and our method perform well both on precision and recall and have higher IoU, which demonstrates the power of the infrared polarization imaging in augmenting the full-time road detection. On the other hand, DeepLabv3+ road detection methods trained with the HSV fusion image outperform that trained with RGB fusion image, because HSV fusion of polarization images can reveal polarization characteristics better than the RGB fusion.

To evaluate how different traffic situations affect the road detection performance of different methods, we test the above road detection methods when the scenes contain different number of cars and pedestrians. Fig. 11(a) shows that most methods cannot detect road well under a complex road condition except for the DL3p\_HSV and our method. The performance of the methods using intensity image and SSRD decline after there are more than 22 cars and pedestrians in scene while our method and the DeepLabv3+ methods trained with HSV fusion image still work well. Fig. 11(b) shows the comparison of DeepLabv3+ methods and our method under several difficult traffic situations. When there are many cars in image, DL3p\_I may wrongly detect parts of car as road because these parts may have similar thermal radiation with road. DL3p\_HSV and DL3p\_RGB benefit from the polarization information and obtain more compelling performance under these difficult situations. The proposed method separates most of cars from road with the strong constraint of the proposed zero-distribution prior and the difference of DoP.

To demonstrate the power of the proposed zero-distribution prior, we test the performance of our road detection method with the AoP constraint only and



**Fig. 11.** (a) Average IoU performance when there are different number of cars and pedestrians in scene. (b) Comparison of DeepLabv3+ methods and our method under several difficult traffic situations

with the AoP-DoP-edge constraint, the results are shown in Table 3. Note that the result of AoP constraint only is obtained by simply removing all the small pieces whose areas are smaller than 2% of the total detected area in  $M_c$  under the detected horizon. The road detection performance of our proposed method with only AoP constraint is still better than most of other methods except the DeepLabv3+ with the HSV and RGB fusion images. The AoP constraint have low precision performance due to the similar AoP between road and vehicle, where the DoP-and-edge constraint is used to solve this problem. More results of road detection videos, horizon detection and different traffic situations can be found in the Supplementary Material.

**Evaluation on execution efficiency.** The computational time is also an important factor for ADAS to evaluate road detection method. Our method is implemented in Matlab and tested on a general PC with 64GB memory and Intel Xeon E3-1225 CPU. The DeepLabv3+ method is tested on a GTX 1080Ti GPU, and other methods are tested in a same PC which runs our method. The average running time of a  $512 \times 640$  input image is shown in Table 4, and the time of denoising and demosaicking is discarded here. On the average, it takes 0.067s for the proposed method to process an image and it is 5 times faster than the DeepLabv3+ method, it is benefited from the powerful zero-distribution prior and the horizon detection.

The resolution of LWIR camera is usually low. To test the robustness and the effectiveness of the proposed method, we down-sample the input image by 2 times ( $2 \times d_s$ ) and 4 times ( $4 \times d_s$ ), and apply our road detection method on these two low resolution inputs. As shown in Table 5, the results under 2 times down-sampled input still outperform the DeepLabv3+ method trained with RGB fusion image and achieve a real-time performance with 33fps cause the 2 times down-sampling barely damage the overall polarization information of scene. With 4 times down-sampled input, the proposed method achieves compelling results even with a higher speed.

It should be noted that the proposed method may fail when the road is wet or the polarization characteristic of car is quite similar with road. When

the road is covered with water, the dominant energy reaches the camera is reflected radiation rather than the emitted radiation as assumed in Section 3. And the DoP of car is affected by the temperature and the viewing angle, so the precision of the proposed method may be influenced when the DoP of car is similar with the road. Our future work aims to solve these limitations.

**Table 4.** Running time comparison of different methods **Table 5.** Performance of the proposed method under 2 times and 4 times down-sampled inputs

Methods	Time (s/frame)	Criteria	Ours-2×d_s	Ours-4×d_s
RG	2.761	PRE (%)	93.03	92.95
SSRD	17.32	REC (%)	93.53	90.98
DeepLabv3+	0.351	IoU (%)	87.43	85.76
Proposed	<b>0.067</b>	Running Time (ms)	30	16
		Frame Rate (fps)	33	62

## 6 Conclusion

This paper presents a road detection technique based on LWIR polarization imaging for autonomous navigation regardless of illumination conditions. We use DoFP infrared imaging technology to acquire infrared polarization information in real time with a monocular camera. The proposed zero-distribution prior of AoP provides a powerful constraint for road detection. Based on this prior, we propose a statistical method with a road vanishing assumption to locate the horizon to reduce computation. We combine zero-distribution of AoP, the difference of DoP, and the edge information to segment the road region in the scene. We developed a LWIR DoFP dataset consisting of 2,113 annotated images. Experiment results on the dataset demonstrate that the proposed method successfully detects the road regardless of illumination conditions, day and night.

## Acknowledgment

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61771391, in part by the Science, Technology and Innovation Commission of Shenzhen Municipality under Grants JCYJ20170815162956949 and JCYJ20180306171146740, Key R & D plan of Shaanxi Province 2020ZDLGY07-11, and by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by Korean government (MSIT) (No. 2019-0-00231, Development of artificial intelligence based video security technology and systems for public infrastructure safety). Yongqiang Zhao is the corresponding author.

## References

1. Abubakar, A., Zhao, X., Li, S., Takruri, M., Bastaki, E., Bermak, A.: A Block-Matching and 3-D filtering algorithm for Gaussian noise in DoFP polarization images. *IEEE Sensors Journal* **18**(18), 7429–7435 (2018)
2. Ainouz, S., Zallat, J., de Martino, A., Collet, C.: Physical interpretation of polarization-encoded images by color preview. *Optics Express (OE)* **14**(13), 5916–5927 (2006)
3. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* **39**(12), 2481–2495 (2017)
4. Bertilone, D.: Stokes parameters and partial polarization of far-field radiation emitted by hot bodies. *JOSA A* **11**(8), 2298–2304 (1994)
5. Bertozzi, M., Broggi, A., Carletti, M., Fascioli, A., Graf, T., Grisleri, P., Meinecke, M.: IR pedestrian detection for advanced driver assistance systems. In: *Joint Pattern Recognition Symposium*. pp. 582–590. Springer (2003)
6. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* **40**(4), 834–848 (2017)
7. Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H.: Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European conference on computer vision (ECCV)*. pp. 801–818 (2018)
8. Choi, Y., Kim, N., Hwang, S., Park, K., Yoon, J.S., An, K., Kweon, I.S.: KAIST multi-spectral day/night data set for autonomous and assisted driving. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **19**(3), 934–948 (2018)
9. Coniglio, N., Mathieu, A., Aubreton, O., Stolz, C.: Characterizing weld pool surfaces from polarization state of thermal emissions. *Optics Letters (OL)* **38**(12), 2086–2088 (2013)
10. Dickson, C.N., Wallace, A.M., Kitchin, M., Connor, B.: Long-wave infrared polarimetric cluster-based vehicle detection. *JOSA A* **32**(12), 2307–2315 (2015)
11. Dickson, C., Wallace, A.M., Kitchin, M., Connor, B.: Improving infrared vehicle detection with polarisation (2013)
12. Garcia, M., Edmiston, C., Marinov, R., Vail, A., Gruev, V.: Bio-inspired color-polarization imager for real-time in situ imaging. *Optica* **4**(10), 1263–1271 (2017)
13. Garcia, M., Edmiston, C., York, T., Marinov, R., Mondal, S., Zhu, N., Sudlow, G.P., Akers, W.J., Margenthaler, J., Achilefu, S., et al.: Bio-inspired imager improves sensitivity in near-infrared fluorescence image-guided surgery. *Optica* **5**(4), 413–422 (2018)
14. Geiger, A., Roser, M., Urtasun, R.: Efficient large-scale stereo matching. In: *Asian Conference on Computer Vision (ACCV)*. pp. 25–38. Springer (2010)
15. Goldstein, D.H.: Polarized light. CRC press (2016)
16. Gu, S., Zhang, Y., Tang, J., Yang, J., Kong, H.: Road detection through CRF based LiDAR-camera fusion. In: *2019 International Conference on Robotics and Automation (ICRA)*. pp. 3832–3838. IEEE (2019)
17. Hillel, A.B., Lerner, R., Levi, D., Raz, G.: Recent progress in road and lane detection: a survey. *Machine Vision and Applications* **25**(3), 727–745 (2014)
18. Hyde, M.W., Cain, S.C., Schmidt, J.D., Havrilla, M.J.: Material classification of an unknown object using turbulence-degraded polarimetric imagery. *IEEE Transactions on Geoscience and Remote Sensing (TGRS)* **49**(1), 264–276 (2010)

19. John, V., Mita, S., Liu, Z., Qi, B.: Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks. In: 2015 14th IAPR International Conference on Machine Vision Applications (MVA). pp. 246–249. IEEE (2015)
20. Klein, L.J., Ingvarsson, S., Hamann, H.F.: Changing the emission of polarized thermal radiation from metallic nanoheaters. *Optics Express (OE)* **17**(20), 17963–17969 (2009)
21. Kong, H., Audibert, J.Y., Ponce, J.: Vanishing point detection for road detection. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 96–103. IEEE (2009)
22. Kong, H., Sarma, S.E., Tang, F.: Generalizing Laplacian of Gaussian filters for vanishing-point detection. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **14**(1), 408–418 (2012)
23. Kwak, J.Y., Ko, B.C., Nam, J.Y.: Pedestrian tracking using online boosted random ferns learning in far-infrared imagery for safe driving at night. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **18**(1), 69–81 (2016)
24. Li, N., Zhao, Y., Pan, Q., Kong, S.G.: Removal of reflections in LWIR image with polarization characteristics. *Optics Express (OE)* **26**(13), 16488–16504 (2018)
25. Li, N., Zhao, Y., Pan, Q., Kong, S.G.: Demosaicking DoFP images using Newton’s polynomial interpolation and polarization difference model. *Optics Express (OE)* **27**(2), 1376–1391 (2019)
26. Li, Q., Chen, S., Wang, C., Li, X., Wen, C., Cheng, M., Li, J.: LO-Net: Deep real-time LiDAR odometry. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 8473–8482 (2019)
27. Li, S., Seybold, B., Vorobyov, A., Fathi, A., Huang, Q., Jay Kuo, C.C.: Instance embedding transfer to unsupervised video object segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6526–6535 (2018)
28. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 3431–3440 (2015)
29. Lu, W., Zhou, Y., Wan, G., Hou, S., Song, S.: L3-Net: Towards learning based LiDAR localization for autonomous driving. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 6389–6398 (2019)
30. Lu, X.: Self-supervised road detection from a single image. In: 2015 IEEE International Conference on Image Processing (ICIP). pp. 2989–2993. IEEE (2015)
31. Mendes, C.C.T., Frémont, V., Wolf, D.F.: Exploiting fully convolutional neural networks for fast road detection. In: 2016 IEEE International Conference on Robotics and Automation (ICRA). pp. 3174–3179. IEEE (2016)
32. Oliveira, G.L., Burgard, W., Brox, T.: Efficient deep models for monocular road segmentation. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 4885–4891. IEEE (2016)
33. Ozgunalp, U., Fan, R., Ai, X., Dahnoun, N.: Multiple lane detection algorithm based on novel dense vanishing point estimation. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **18**(3), 621–632 (2016)
34. Peláez, G., Bacara, D., de la Escalera, A., García, F., Olaverri-Monreal, C.: Road detection with thermal cameras through 3D information. In: 2015 IEEE Intelligent Vehicles Symposium (IV). pp. 255–260. IEEE (2015)
35. Peynot, T., Underwood, J., Scheduling, S.: Towards reliable perception for unmanned ground vehicles in challenging conditions. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1170–1176. IEEE (2009)

36. Piniarski, K., Pawlowski, P.: Efficient pedestrian detection with enhanced object segmentation in far IR night vision. In: 2017 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA). pp. 160–165. IEEE (2017)
37. Qian, Y., Dolan, J.M., Yang, M.: DLT-Net: Joint detection of drivable areas, lane lines, and traffic objects. *IEEE Transactions on Intelligent Transportation Systems (TITS)* (2019)
38. Rahmann, S., Canterakis, N.: Reconstruction of specular surfaces using polarization imaging. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). vol. 1, pp. I–I. IEEE (2001)
39. Rashed, H., El Sallab, A., Yogamani, S., ElHelw, M.: Motion and depth augmented semantic segmentation for autonomous navigation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). pp. 0–0 (2019)
40. Reda, M., Zhao, Y., Chan, J.C.W.: Polarization guided autoregressive model for depth recovery. *IEEE Photonics Journal* **9**(3), 1–16 (2017)
41. Shen, L., Zhao, Y., Peng, Q., Chan, J.C.W., Kong, S.G.: An iterative image dehazing method with polarization. *IEEE Transactions on Multimedia (TMM)* **21**(5), 1093–1107 (2018)
42. Su, Y., Zhang, Y., Lu, T., Yang, J., Kong, H.: Vanishing point constrained lane detection with a stereo camera. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **19**(8), 2739–2744 (2017)
43. Terrier, P., Devlaminck, V., Charbois, J.M.: Segmentation of rough surfaces using a polarization imaging system. *JOSA A* **25**(2), 423–430 (2008)
44. Tyo, J.S., Goldstein, D.L., Chenault, D.B., Shaw, J.A.: Review of passive imaging polarimetry for remote sensing applications. *Applied Optics (AO)* **45**(22), 5453–5469 (2006)
45. WA, R.: Principles of optics (1999)
46. Wang, H., Wang, Y., Zhao, X., Wang, G., Huang, H., Zhang, J.: Lane detection of curving road for structural highway with straight-curve model on vision. *IEEE Transactions on Vehicular Technology (TVT)* **68**(6), 5321–5330 (2019)
47. Yoon, J.S., Park, K., Hwang, S., Kim, N., Choi, Y., Rameau, F., so Kweon, I.: Thermal-infrared based drivable region detection. In: 2016 IEEE Intelligent Vehicles Symposium (IV). pp. 978–985. IEEE (2016)
48. Zhang, W., Cai, X., Huang, K., Zhang, Z.: A novel RGB-T based real-time road detection on low cost embedded devices. In: National Conference on Embedded System Technology. pp. 17–26. Springer (2017)
49. Zhang, Y., Su, Y., Yang, J., Ponce, J., Kong, H.: When Dijkstra meets vanishing point: a stereo vision approach for road detection. *IEEE Transactions on Image Processing (TIP)* **27**(5), 2176–2188 (2018)
50. Zhang, Y., Wang, J., Wang, X., Dolan, J.M.: Road-segmentation-based curb detection method for self-driving via a 3D-LiDAR sensor. *IEEE Transactions on Intelligent Transportation Systems (TITS)* **19**(12), 3981–3991 (2018)