Camera Image Dehazing and Target Detection for Autonomous Vehicles

Keizo MIYAHARA[1] Yucheng XU[2] and Natsuki OHARA[3]

***Abstract*— This paper describes a simple target detection system for autonomous vehicles to perform an emergency stop. We especially focus on the system only with a digital camera as the perception unit for the situation. The detecting performance of the imaging sensors, including the digital cameras, often suffer from "haze" due to atmospheric obscuration. In this paper, we review the image "dehazing" algorithms aiming for the in-vehicle usage, and we propose a system configuration for the safety measure with applying a selected algorithm. A series of experimental results depicted the feasibility of the system configuration and its applicability to real-time processing.**

I. INTRODUCTION

Governments of European countries, United states, Japan, and other countries are leading development in "Automated Driving Systems (ADS)" technologies [1]–[4]. Under these assistance, varieties of ADS researches have been conducted [5]. An emergency deceleration and stop strategy, or "Minimum Risk Maneuver (MRM)," is one of the key technologies for ADS safety. It minimizes the risk in case that the driver does not respond to a taking-over request from ADS [6]. Major sensors for ADS are shown in Table I. Generally 20-30 of these sensors are implemented on ADS and expected to function consistently [7], particularly for level 4 and 5 operations [8].

TABLE I: Sensors utilized for "Automated Driving Systems (ADS)" [7].

Sensor	Characteristic	Detection target
Camera	Object recognition with position / orientation Low performance at hazy scene	Car Pedestrian Sign Marking
Light detection and ranging (Lidar)	Dark scene acceptable Low performance at hazy scene	Car Pedestrian
Millimeter wave radar	Hazy scene acceptable Low performance with near / small objects	Car Pedestrian Obstacles
Ultrasonic range sensor	Almost all weather acceptable Low performance with far objects	Near obstacles

This paper aims at operating the MRM with digital cameras. These image sensors are often installed interior of vehicles, and therefore even in traffic accident cases, we can expect that the cameras are operational being kept from harmful damages comparing to other sensors attached to vehicle exterior. The cameras, however, often face difficulties in acquiring images with foggy, rainy and/or misty conditions (hereinafter referred to as "haze"). Regardless the haze, the object recognition system should be reliable to perform the MRM strategy. To achieve this final goal, this paper studies an object recognition system only with a camera as the perception unit. Especially we propose a system configuration that performs affirmatively in the hazy scenes.

The structure of this paper is as follows: Section II shows former works on "dehaze" algorithms and another related area. Section III describes the proposed recognition system composed of image dehazing and target detection subsystems. A series of experimental results listed in Section IV shows the feasibility of the system for real-time applications. Note that a part of basic design of the proposed recognition system was presented by the authors at a conference [9]. This paper summarizes the former contents, and describes the newly performed experimental results, together with a review/study of dehaze algorithms.

II. RELATED RESEARCH

Dehaze algorithms can be categorized into two groups: image enhancement and image restoration. In short, the algorithms in the former group focus on the image quality itself. They adjust parameters of the image to reduce noises and to enhance the edges. In contrast, the latter adopts certain physical models to estimate the real images. Representing algorithm from each group will be examined in the following subsections to apply for the object recognition system.

A. Image enhancement

1) "Retinex" algorithm: One of the active research field on dehazing with image enhancement is the "Retinex" algorithm [10]. Retinex is a coined word combining "retina" and "cortex." It is modeled after the human visual function, and has been proposed as an effective processing technique for improving quality of the input image.

Several dehaze processes with the Retinex algorithm have been studied, including FMC [11], SSR [12], MSR [13] and MSRCR [14]. The processes SSR and MSR were examined here with sample images. Figure 1 shows the results of the experiment. With this image example, MSR (c) provides better detection results than SSR (b).

It is also known that Retinex algorithm sometimes encounter with noise reduction difficulties since it re-

[1]Keizo MIYAHARA and [3]Natsuki OHARA are with Graduate School of Science and Engineering, Kwansei Gakuin University, Hyogo, Japan [1]miyahara.keizo@kwansei.ac.jp, [3]gin02492@kwansei.ac.jp

[2]Yucheng XU is with Japan Global Gateway, Fujitsu Limited, Kanagawa, Japan xu.yucheng@jp.fujitsu.com

(a) Original image

(b) After SSR

(c) After MSR

Fig. 1: Experiments with Single-scale Retinex (SSR) and Multi-scale Retinex (MSR).

quires computational resources and ample parameter adjustment. [15].

2) Histogram equalization (HE): The "Histogram equalization" is another widely applied method to image quality improvement. It can be re-categorized into two sub-group according to the size of divided region: GHE (Global histogram equalization) and LHE (Local histogram equalization).

Although the GHE is effective for uniform noise over the image, it frequently results in so-called "halo effect" [16]. The LHE has subtypes: POSHE [17], BPDFHE [18], and CLAHE [19], those utilize different region dividing algorithms. In particular, CLAHE is able to suppress the "block effect" [18]. Figure 2 shows the results of the experiment with CLAHE for the same original image in Figure 1. It provides better detection results than the Retinex cases with the sample image. The advantage of CLAHE is originated from the intrinsic adoptability of the method to suppress image noise.

(a) After CLAHE

Fig. 2: Experiments with Contrast-Limited Adaptive Histogram Equalization (CLAHE).

B. Image restoration

1) Physical model for image restoration: The model of "atmosphere scattering" or "air light scattering" was proposed in [20]. According to this model, haze can be formulated with following equation as depicted in Fig. 3:

$$I(x) = J(x)t(x) + A\left(1 - t(x)\right), \tag{1}$$

where I, J, t, A and x are: an observed intensity (obtained camera image), scene radiance (direct light from the object), medium transmission (non-scattered light ratio), atmospheric light (globally/uniformly over-laid air light) and pixel of an input image, respectively.

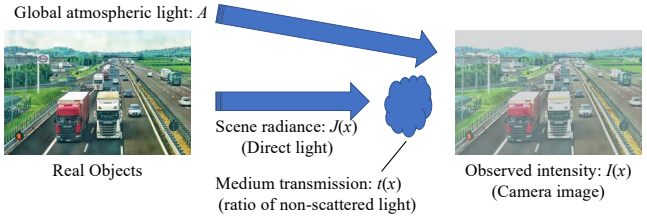


Fig. 3: The atmosphere scattering model.

We note that the camera image $I(x)$ is under the influence of haze. That is, the model describes $I(x)$ as a linear sum of the direct light $J(x)$ (ideal image of the object without haze) and the atmospheric light A. We see that every $I(x)$, $J(x)$, and A holds intensity of RGB values. $t(x)$ is a parameter, called "transmission" that directly describes the influence of haze condition. According to the equation, we understand that our target image $J(x)$ can be restored from the obtained image $I(x)$ with information of $t(x)$ and A.

2) Dark channel prior (DCP): "Dark channel prior (DCP)" [21] is an image restoration concept that is based on the statistical study of massive number of hazy scenes in order to compute $t(x)$ and A. It was found in many cases that the hazy scenes include a few pixels with which luminance values were significantly low. This fact was named "Dark channel prior."

We applied DCP for the proposed system to make use of its straightforward formulation and memory-friendliness as shown in the next section. The mathematical stability of DCP process was proved in [22]. According to the experimental results of DCP presented in Section IV show that the algorithm is able to provide affirmative dehazing process in real-time.

III. PROPOSED SYSTEM

A. Image "dehaze" processing sub-system

We formulate the dehaze process with the DCP concept in this subsection. The DCP over the input image $(I^{DC}(x))$ can be defined as follows:

$$I^{DC}(x) = \min_{y \in \Omega(x)} \left(\min_{C \in \{R,G,B\}} I^{C}(y) \right), \tag{2}$$

where $I^{DC}(x)$ and $\Omega(x)$ are: a dark channel with respect to a pixel x and a neighborhood area of x, respectively. Setting a neighborhood area Ω of every x, pixel y sweeps over Ω. Computing minimum intensity among any channel (R,G,B) at every y, we pick out the minimum value among Ω, and the picked value is set as $I^{DC}(x)$ at the pixel x. We note that this value describes the influence of haze over the image (small value = less influence).

Applying this computation on $J(x)$, we obtain the following equation:

$$J^{DC}(x) = \min_{y \in \Omega(x)} \left(\min_{C \in \{R,G,B\}} J^{C}(y) \right) \simeq 0. \tag{3}$$

Substituting this into eq. (1), we obtain the following equation:

$$I^{DC}(x) \simeq A\left(1 - t(x)\right), \tag{4}$$

and this directly results in the transmission $t(x)$ as follows:

$$t(x) \simeq 1 - \frac{I^{DC}(x)}{A}. \tag{5}$$

As shown in Fig. 3, the atmospheric light A uniformly laid over the image $J(x)$. According to this physical model, we consider that the following holds:

$$A_0 \simeq \max_{C \in \{R,G,B\}} I^{C}(x), \tag{6}$$

where A_0 is the estimate of A. Together with eq. (5), we can obtain the dehazed target image $J(x)$ as below:

$$J(x) \simeq \frac{I(x) - A_0}{t(x)} + A_0. \tag{7}$$

B. Target detection sub-system

Among the target detecting algorithm in public domain, YOLO.v5 [23] was implemented for the proposed system. YOLO is based on the deep learning framework "PyTorch," with a neural network "Darknet." YOLO supports "Computer Unified Device Architecture (CUDA)" framework, a parallel computing environment, to perform multithreaded processing using GPUs. Our paper aims at contributing to ADS research field through a proposal of a system configuration that implements DCP dehaze processing into the CUDA environment seamlessly with YOLO algorithm as shown in the following subsection.

For the learning scheme, we applied YOLOv5s (Table II) in terms of the balance of the recognition precision and the learning speed. The learning curve of the applied dataset with 20 leaning classes is shown in Fig. 4. Horizontal and vertical axes show the learning iterations (times) and model evaluation index (mean average precision: mAP), respectively. From the figure we see that mAP increases uniformly according to the learning times, and that the learning effect saturates around 50 iterations.

TABLE II: Learning dataset for "YOLO" [24].

DataSet	mean Average Precision mAP	Learning speed example (Skylake)	Learning speed example (NVIDIA V100)
YOLOv5n	45.7	45	6.3
YOLOv5s	56.8	98	6.4
YOLOv5m	64.1	224	8.2
YOLOv5l	67.3	430	10.1
YOLOv5x	68.9	766	12.1

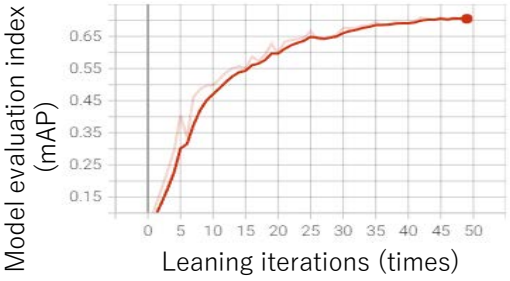


Fig. 4: Learning curve of the applied dataset.

C. System configuration

The configuration of the proposed object recognition system is shown in Fig. 5. Input data can be both still images and moving images (movies). The input image will be processed in the CUDA environment shown in the middle box of the figure. From the view point of hardware, the main part of the process will be performed on GPU (Fig. 6). The image data will be transmitted to GPU memory, and then distributed to CUDA cores by the thread-scheduler (Warp) as CUDA tasks. At this moment, the DCP dehaze process should possess

priority to the YOLO target detection tasks in order for image consistency. This programing is the key for the seamless implementation of the DCP into the YOLO algorithm.

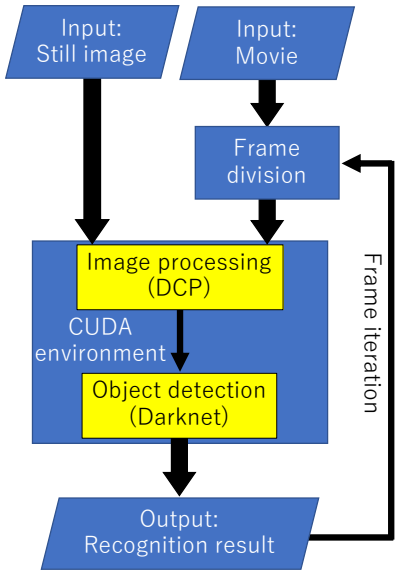


Fig. 5: Proposed system (software configuration).

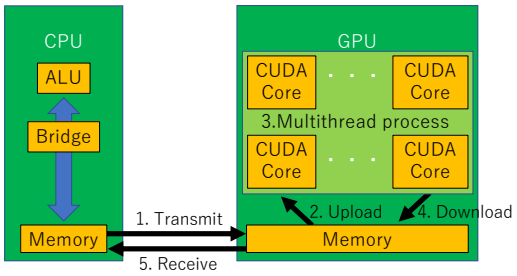


Fig. 6: "CUDA" concurrent process environment (device-based configuration).

IV. EXPERIMENTS

A series of experiments were performed to verify the feasibility and the real-time processing capacity of the proposed system.

A. System parameter Ω

The system parameter Ω introduced in Section III was examined for foundational setting. Applying a set of images recorded in hazy condition on real roads, object recognition accuracy was examined with respect to the size of Ω to optimize its value. Some example images are shown in Fig. 8. The experimental results showed similar trend among the sample set. An example result of the experiments is shown in Fig. 7. According to these results, the size of Ω was set to 51×51 for the proposed system.

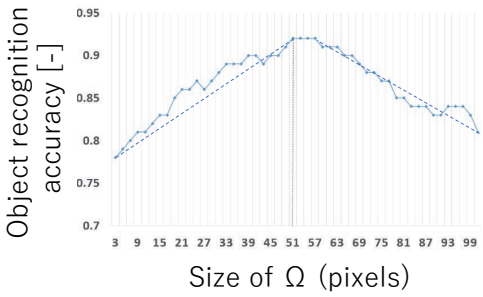


Fig. 7: Ω adjustment with sample images (example result).

B. Basic feature

The basic feature of the system, the object recognition performance with the dehaze process, was examined. The DCP dehaze image processing sub-system and the target detection sub-system with YOLO was implemented in a laptop PC, after the learning process stated in Section III and Ω setting mentioned above. The specification of the laptop is shown in table III. The applied images for this experiment are the same set for the Ω setting.

Fig. 8 shows four example cases of the experimental results. The bounding box with class identifier and recognition accuracy is shown in the figure as the output of the system. We see that a target object (car) was detected on every sample image, and that the recognition accuracy is improved after the DCP process ((b),(d),(f),(h)), comparing to the original images ((a),(c),(e),(g)).

TABLE III: Specification of the applied PC for the experiments.

CPU	AMD Ryzen 7, 3.20GHz
GPU	NVIDIA GeForce RTX 3070
Memory	DDR3 SDRAM 16GB

C. Contrast evaluation

Throughout the basic feature experiment shown above, the target recognition accuracy and image contrast were collated.The results showed similar trend again among the sample set. Figure 9 shows the collation of improvement results among the four example images. We confirmed the correlation between recognition accuracy and contrast values, which can be considered as the dehazing effect.

D. Processing time

The processing time of the proposed system was examined with digital camera input. Table IV shows the specification of the applied USB camera.

Instead of camera traversing real haze road, recorded movie in hazy environment was applied for the experiments as an alternative. The movie was displayed on a monitor, and the USB camera shoot the moving images to input them for

(a) Sample image 1 (before DCP)

(b) Sample image 1 (after DCP)

(c) Sample image 2 (before DCP)

(d) Sample image 2 (after DCP)

(e) Sample image 3 (before DCP)

(f) Sample image 3 (after DCP)

(g) Sample image 4 (before DCP)

(h) Sample image 4 (after DCP)

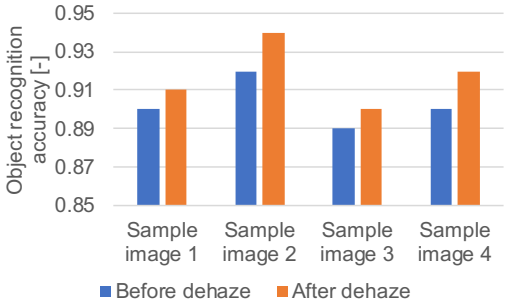
Fig. 8: Proposed system output with sample images.

TABLE IV: Specification of the applied USB camera for the experiments.

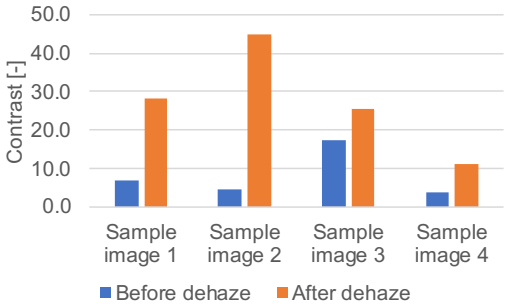
Pixel	400×10^6
View angle	126 deg
Maximum frame rate	60 FPS

the object recognition system proposed. For comparison purpose, two cases of DCP sub-system on/off were examined.

Fig. 10 shows an example result of the experiments. The figure shows that the proposed system with DCP (b) succeeded to detect the target object (pedestrian) in the hazy scene, although the system without DCP (a) failed the target detection. The frame rate of the proposed system (b) was faster than 50 FPS. This processing speed of the whole recognition system is considered fast enough for real-time computation, with respect to the speed of the cameras in common use for current ADS application [25].

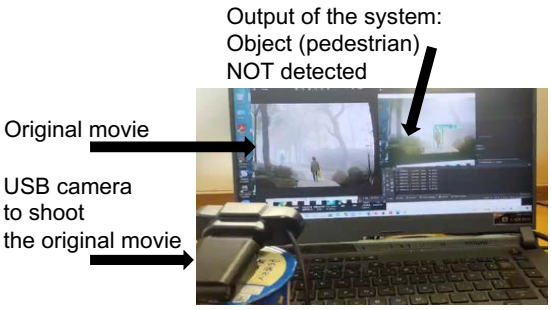


(a) Target recognition accuracy

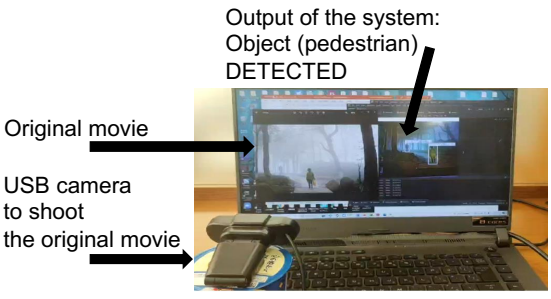


(b) Contrast

Fig. 9: Contrast and recognition accuracy with sample images.



(a) DCP off



(b) DCP on

Fig. 10: Experiments in quasi-real hazy condition.

E. Effective factor in hardware

The processing time and chip occupancies were assessed to confirm effective factors among the hardware with proposed software configuration. Table V describes the comparison between two sets of proposed systems. The upper and lower halves of the table show: the running results with the PC shown in table III and with another PC, respectively. In short, the CPU and GPU of the lower PC is about half speed comparing to those of the upper one. For each PC usage, CPU/GPU/Memory occupancy rate and frame rate were measured during the system is running with applying the identical USB camera shown in Table IV.

We see in the table that the on-board memory rates indicated only small fluctuation whether the GPU was on or off. This and the CPU/GPU occupancy rates confirm that the main part of processing is affirmatively performed in GPU, and that the specification of GPU directly affects the processing speed. Besides, GPU remained much spare capacity, and therefore software algorithm itself might have room for improvement to make full use of GPU's capability.

TABLE V: Processing time and chip occupancies (two laptop PCs).

	Processing time per frame [sec]	CPU occupancy rate [%]	GPU occupancy rate [%]	On-board memory occupancy rate [%]
CPU Ryzen7 only	0.130	73.0	(N/A)	55.0
Ryzen7 with RTX3070	0.018	17.0	4.0	57.0
CPU XeonE5 only	0.250	80.0	(N/A)	31.0
XeonE5 with Quadro K620	0.031	23.0	8.0	33.0

V. SUMMARY AND DISCUSSION

This paper discussed camera image dehazing algorithm and an object recognition system, aiming at the MRM strategy for ADS vehicles.

After reviewing related papers including the "dehaze" technology, a recognition system that works in hazy conditions was proposed. It consists of image processing and target detection sub-systems. A series of experiments were performed with sample images and movies recorded in hazy condition on real roads. The results of them validated affirmatively the real-time processing ability (50 FPS and faster) of the proposed system, which is fast enough for current in-vehicle usage. They also suggested that the dehazing effect was derived from the contrast enhancement. It was confirmed as well that GPU performs main role in the proposed system configuration, and that it still has spare capacity.

The future work of this research includes: to fine-tune other dehaze strategies, CLAHE for example, examined in Section IV for further comparison with the DCP, and to improve software structure to make full use of the GPU's capability.

REFERENCES

[1] Secretary of State for Transport and the Secretary of State for Business, Energy and Industrial Strategy, "Connected & automated mobility 2025," 2022.

[2] The European Road Transport Research Advisory Council (ERTRAC), "Connected automated driving roadmap," 2019.

[3] The US Department of Transportation (USDOT), "Automated vehicles comprehensive plan," 2021.

[4] Cabinet Office, Japanese Government, "Automated driving for universal services R&D plan," 2020.

[5] I. Altaf and A. Koul, "A survey on autonomous vehicles in the field of intelligent transport system," in *Applications of Networks, Sensors and Autonomous Systems Analytics*. Springer, 2022, pp. 11–31.

[6] Federal Ministry for Digital and Transport, German Government, "Minimum risk maneuvers (MRM)," 2020.

[7] H. A. Ignatious, Hesham-El-Sayed, and M. Khan, "An overview of sensors in autonomous vehicles," *Procedia Computer Science*, vol. 198, pp. 736–741, 2022.

[8] SAE International, "SAE j3016: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems," 2021.

[9] Y. Xu and K. Miyahara, "Object recognition system under hazy condition for automated driving systems," in *International Conference on Cybernetics and Intelligent Systems, Robotics, Automation and Mechatronics (CIS-RAM)*, no. 113, June 2023, pp. 1–6.

[10] E. H. Land and J. J. McCann, "Lightness and retinex theory," *Journal of the Optical Society of America*, vol. 61, no. 1, pp. 1–11, 1971.

[11] J. McCann, "Lessons learned from mondrians applied to real images and color gamuts," in *International Conference on Communications in Computing*, 1999, pp. 1–8.

[12] Z. U. Rahman, D. J. Jobson, and G. A. Woodell, "Retinex image processing," in *IEEE International Conference on Image Processing*, vol. 3, 1996, pp. 124–125.

[13] ——, "Multi-scale retinex for color image enhancement," in *IEEE International Conference on Image Processing*, vol. 3, 1996, pp. 1003–1006.

[14] ——, "A multi-scale retinex for bridging the gap between color images and human observation of scenes," pp. 965–976, 1997.

[15] B. Petro, C. Sbert, and J. M. Morel, "Multiscale retinex," *Image Processing On Line*, vol. 4, pp. 71–88, 2014.

[16] L. J. Wang and R. Zhu, "Image defogging algorithm of single color image based on wavelet transform and histogram equalization," *Applied mathematical sciences*, vol. 7, no. 79, pp. 3913–3921, 2013.

[17] J. Y. Kim, L. S. Kim, and S. H. Hwang, "An advanced contrast enhancement using partially overlapped sub-block histogram equalization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 4, pp. 475–484, 2001.

[18] C. Ramya and D. S. S. Rani, "Contrast enhancement for fog degraded video sequences using bpdfhe," *International Journal of Computer Applications*, vol. 3, no. 2, pp. 3463–3468, 2012.

[19] Z. Xu, X. Liu, and X. Chen, "Fog removal from video sequences using contrast limited adaptive histogram equalization," in *International Conference on Computational Intelligence and Software Engineering*, 2009, pp. 1–4.

[20] E. J. McCartney, *Optics of the atmosphere: scattering by molecules and particles*. John Wiley and Sons, 1976.

[21] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, p. 2341–2353, 2011.

[22] K. B. Gibson and T. Q. Nguyen, "On the effectiveness of the dark channel prior for single image dehazing by approximating with minimum volume ellipsoids," in *IEEE Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 1253–1256.

[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779–788.

[24] Ultralytics Inc., "yolov5," [Online] Available: https://github.com/ultralytics/yolov5.

[25] TIER IV, Inc., "Automotive camera," [Online] Available: https://sensor.tier4.jp/automotive-camera.