

# Auto-adjusting Camera Exposure for Outdoor Robotics using Gradient Information

Inwook Shim<sup>1</sup>, Joon-Young Lee<sup>2</sup>, and In So Kweon<sup>3</sup>

**Abstract**—We present a new method to auto-adjust camera exposure for outdoor robotics. In outdoor environments, scene dynamic range may be wider than the dynamic range of the cameras due to sunlight and skylight. This can result in failures of vision-based algorithms because important image features are missing due to under-/over-saturation. To solve the problem, we adjust camera exposure to maximize image features in the gradient domain. By exploiting the gradient domain, our method naturally determines the proper exposure needed to capture important image features in a manner that is robust against illumination conditions. The proposed method is implemented using an off-the-shelf machine vision camera and is evaluated using outdoor robotics applications. Experimental results demonstrate the effectiveness of our method, which improves the performance of robot vision algorithms.

## I. INTRODUCTION

For robot vision algorithms, capturing well-exposed images is an essential prerequisite for the success of the algorithms, therefore, determining proper exposure is a fundamental problem in vision-based robotics research. Despite the importance of camera exposure, most robot vision research has relied on either a camera’s built-in auto-exposure algorithm [1], [2], [3] or a fixed exposure setting that is manually tuned by a user. While these methods work fine in controlled environments that have stationary illumination, they become reasons for failures of vision algorithms when the algorithms work in uncontrolled environments. Especially in outdoor environments, scene radiance is unpredictable due to large illumination changes over space and time; this radiance may have a much wider dynamic range than the dynamic range of the cameras in question in the presence of sunlight, skylight, and cast shadows.

Figure 1 shows examples of common failure cases in attempts to capture well-exposed images using a built-in auto-exposure method and a fixed exposure setting in an outdoor environment. Both methods can capture well-exposed images when proper parameters are given and the dynamic range of the scene radiance is relatively narrow. However, both methods fail to capture well-exposed images under varying illumination conditions, especially when the dynamic range of the scene radiance is wide. In applications working under controlled environments, a manually tuned



Fig. 1: Images are captured in different illumination conditions. From the left to the right, a camera built-in auto-exposure method, a manually tuned fixed exposure setting, and our method are used. Both the built-in auto-exposure method and the manual setting fail to capture well-exposed images for vision-based algorithms, while our method captures images suitable for the processing of vision algorithms.

fixed exposure setting may be preferred to prevent undesired exposure changes; however, such an exposure setting clearly has a problem in handling illumination changes in uncontrolled environments. Compared to a fixed exposure setting, a built-in auto-exposure method automatically adjusts camera exposure based on scene radiance and is very robust against illumination changes. In outdoor environments, however the built-in automatic exposure setting wipes out important image features in low radiance regions when the dynamic range of the camera is insufficient to capture details of a whole image area in high contrast outdoor scenes.

In this paper, we present a new method to auto-adjust camera exposure; this method is designed in particular for outdoor robotics applications. To handle severe illumination changes and a wide dynamic range of scene radiance, we evaluate the proper exposure of a scene in the gradient domain. Since the gradient domain is robust against illumination changes and is preferred for many robot vision algorithms, our method enables us to capture well-exposed images for robot vision algorithms by determining the proper exposure needed to maximize useful image features. We estimate the gradient variations according to exposure changes using a  $\gamma$ -correction technique and develop a feedback system to auto-adjust camera exposure. The overall framework is shown in Figure 2. The proposed method is implemented on a machine vision camera; the effectiveness is validated with extensive experiments.

<sup>1</sup>I. Shim is with the Robotics and Computer Vision Laboratory, Division of Future Vehicle, KAIST, Daejeon, 305-701, Korea iwshim@rcv.kaist.ac.kr

<sup>2</sup>J.-Y. Lee is with the Robotics and Computer Vision Laboratory, Department of Electrical Engineering, KAIST, Daejeon, 305-701, Korea jyilee@rcv.kaist.ac.kr

<sup>3</sup>I.S. Kweon is with the Department of Electrical Engineering, KAIST, Daejeon, 307-701, Korea iskweon@kaist.ac.kr

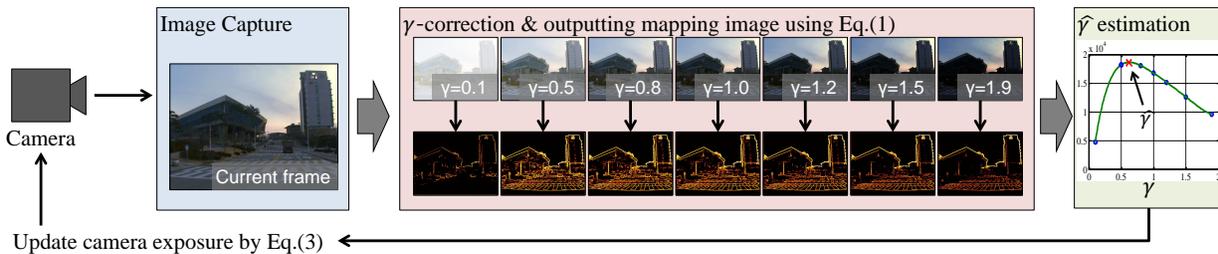


Fig. 2: The overall framework for auto-adjusting camera exposure using gradient information.

## II. RELATED WORK

In general, auto-exposure methods adjust camera exposure by evaluating the average brightness of an image. However, auto-exposure algorithms fail to capture well-exposed images when there is a large luminance difference between an interest area and the background. This often happens in outdoor environments, for example, when both skylight and cast shadow regions co-exist in scenes.

To directly account for the exposure problem in outdoor robotics, the capturing of a high dynamic range (HDR) image is considered. An HDR image is produced by combining a set of images taken with difference exposures. Nuske *et al.* [4] captured three images in order to reconstruct one HDR image. To maximize the visual information in the HDR image, they determined the exposure levels of the three images by exploiting intensity statistics. Hrabar *et al.* [5] captured HDR images using a stereo camera. They used disparities from the stereo image pair to determine a set of exposure times for HDR imaging. While the HDR approach recovers scene radiance well, it requires the capturing of multiple images, which may cause an additional alignment problem in mobile robotics.

Several researchers have exploited prior scene information to determine proper exposure. Kao *et al.* [6] detected moving objects and human faces in real-time; they used the detection results to guide auto-exposure in a surveillance scenario. In [7], Vatani and Roberts utilized a pre-defined mask to calculate an intensity histogram instead of using a whole image; the intensity histogram was used for exposure control. Lu *et al.* [8] quantified illumination changes by calculating a mean brightness value of a reference area in a captured image; they then used image entropy to optimize the exposure level in a manner that was robust against illumination changes. Neves *et al.* [9] set camera exposure according to an intensity histogram with known black and white regions. Montalvo *et al.* [10] controlled camera exposure by analyzing a color histogram; additionally, they refined the exposure level using histogram matching between a reference image and an input image. While the approach based on prior information works well with a known environment, it is difficult to assume prior information for a scene in most outdoor robotics applications. Therefore, those methods that employ prior information are suitable only for specific applications.

There is an heuristic approach that can divide an image into several regions and apply different weightings to determine the camera exposure. To emphasize the luminance of

a main object against back-lighting, Lee *et al.* [11] divided an image into five blocks and calculated the total luminance by adding the weighted luminances of each block. Murakami and Honda [12] divided an image into object and background regions and derived the importance of the background by detecting the *hue* and *chromaticity* values of the pixels. Exposure was adjusted by fuzzy logic in order to retain useful information about the backgrounds.

In comparison with previous works, our method determines proper exposure by evaluating the gradient information within an image. By maximizing the gradient magnitude, our method can automatically adjust the camera exposure to capture important image features whilst being robust against illumination conditions.

## III. IMAGE QUALITY FROM A ROBOT PERSPECTIVE

In robot vision, intensity gradient is one of the most important cues for image processing and scene understanding. Most image features, such as edges, corners, SIFT [13], and HOG [14], are computed in the gradient domain because intensity gradient is robust against illumination changes and can characterize object appearance and shape well. Many vision algorithms extract such image features for high level processing such as object detection, tracking, recognition, and SLAM. Therefore, capturing images with rich gradient information is an important first step toward the success of many vision algorithms. In this section, we explain how to evaluate the exposure of a scene from a robot vision perspective.

In order to evaluate image quality from a robot vision perspective, it is natural to exploit the gradient domain because gradient is a dominant source of visual information for many robot vision algorithms [15]. In this perspective, we regard well-exposed images as images that have rich gradient information, and therefore we evaluate exposure as the total sum of the gradient information in an image.

It is difficult to quantify gradient information explicitly because the importance of the gradient may differ for certain tasks and applications. To quantify the gradient information in a statistical way, we have designed a mapping function that correspond the *gradient magnitude* to the *amount of gradient information* by reflecting general tendencies of the gradient; then, we compute the gradient information of an image by adding all the outputs of the function.

It is well known that the gradient magnitude of a natural scene has a heavy-tailed distribution [16]; therefore, most gradients have relatively small values in comparison to the

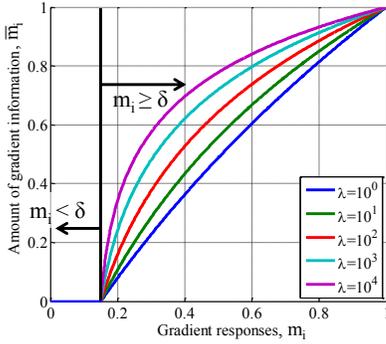


Fig. 3: Our mapping function between gradient magnitude and amount of gradient information according to the control parameters,  $\delta$  and  $\lambda$ . We designed the mapping function by reflecting the general characteristics of the gradient.

maximum possible gradient value. On the other hand, large gradients are usually observed around object boundaries, which have higher probabilities of encoding important information. To balance the importance of small and large gradients, we use a non-linear logarithm function to map the relationship between the gradient magnitude and the amount of gradient information.

Since gradient is sensitive to subtle intensity variations, image noise should be considered in the mapping function. Especially, when we apply the logarithm mapping, there is an overemphasis on small gradients that occur due to image noise. For robust mapping against image noise, we modify the mapping function so that it is activated after a certain gradient level is satisfied. With this modification, the mapping function is defined as

$$\bar{m}_i = \begin{cases} \frac{1}{N} \log(\lambda(m_i - \delta) + 1) & \text{for } m_i \geq \delta \\ 0 & \text{for } m_i < \delta \end{cases} \quad (1)$$

s.t.  $N = \log(\lambda(1 - \delta) + 1)$ ,

where  $m_i^1$  is the gradient magnitude at pixel location  $i$ ,  $\delta$  is the activation threshold value,  $\lambda$  is a control parameter to adjust the mapping tendencies, and  $\bar{m}_i$  represents the amount of gradient information corresponding to the gradient magnitude.  $N$  is a normalization factor to bound the output range of the function to  $[0, 1]$ .

Figure 3 shows the mapping function. In the function, there are two user control parameters,  $\delta$  and  $\lambda$ .  $\delta$  determines the activation threshold value; therefore, the mapping function regards a gradient value smaller than  $\delta$  as a noise response and ignores it.  $\lambda$  determines the tendencies of the mapping. We can emphasize strong intensity variations by setting  $\lambda$  to a small value; we can vitalize subtle texture variations by setting  $\lambda$  to a large value.

Using Eq. (1), we can calculate the total amount of gradient information in an image as in

$$M = \sum \bar{m}_i. \quad (2)$$

In our method, we regard images having larger  $M$  as better exposed images that contain rich gradient information of a scene.

<sup>1</sup>We assume that gradient magnitude  $m_i$  is in the range of  $[0, 1]$ .

#### IV. CAMERA EXPOSURE AUTO-ADJUSTING

For auto-adjusting of the camera exposure, we need to make a feedback system that can update the camera exposure according to an evaluation of the current exposure. In conventional methods that evaluate exposure using brightness, camera exposure is directly related to brightness: the longer the exposure time, the brighter the image intensity. Auto-adjusting of camera exposure can be done by comparing current brightness to a certain reference brightness. For our case, which evaluates exposure in the gradient domain, we compute the amount of gradient information need to relate the camera exposure to the gradient. However, we cannot determine the reference amount because our objective is to maximize the gradient information, which is scene dependent. As an alternative, we use the  $\gamma$ -correction technique to estimate the reference, which process maximizes the gradient information.

We generate  $\gamma$ -corrected images  $I_{out} = I_{in}^\gamma$  from a current image  $I_{in}$ <sup>2</sup>. The  $\gamma$ -correction process makes an image darker when  $\gamma$  is larger than one; it makes an image brighter when  $\gamma$  is smaller than one. Using this characteristic, we simulate exposure changes using the  $\gamma$ -correction and estimate variations of the gradient information according to exposure changes using  $\gamma$ -corrected images.

It is time-consuming to compute the total amount of gradient information from all possible  $\gamma$ -corrected images. For efficiency, we calculate the total amount of gradient information from seven anchor images given by  $\gamma \in [0.1, 0.5, 0.8, 1.0, 1.2, 1.5, 1.9]$ ; we estimate fifth-order polynomial fitting using the outputs of the anchor images. We take the maximum output of the polynomial function in the range of  $\gamma = [0.1, 1.9]$  as a reference; its corresponding  $\gamma$  is assigned to  $\hat{\gamma}$ .

After determining  $\hat{\gamma}$ , we update the camera exposure according to

$$E_{t+1} = E_t(1 + \alpha K_p(1 - \hat{\gamma}))$$

s.t.  $\alpha = \begin{cases} 1/2 & \text{for } \hat{\gamma} \geq 1 \\ 1 & \text{for } \hat{\gamma} < 1, \end{cases} \quad (3)$

where  $E_t$  is the exposure level at time  $t$  and  $K_p$  is the proportional gain required to control the convergence speed. According to  $K_p$ , there is a trade-off between convergence speed and stability of the system. A high value of  $K_p$  makes the feedback system settle quickly but may cause oscillation. In our implementation, we manually tune  $K_p$ , and set it to 0.2 because this gives quick convergence with slight oscillation.

#### V. EXPERIMENTAL RESULTS

To validate the performance of the proposed method, we performed experiments using two outdoor robot applications: surveillance and automotive visual odometry. We compared our method to two conventional methods: camera built-in auto-exposure (AE, henceforth) and manually tuned fixed exposure setting (ME, henceforth). In presenting the experimental results, we begin by describing the implementation details.

<sup>2</sup>We assume that intensity is in the range of  $[0, 1]$ .



Fig. 4: Camera system for experimental validations. The three cameras had the same hardware specifications and were synchronized using an internal software trigger. Each camera determined exposure parameters using a built-in auto-exposure algorithm, our algorithm, and a manually fixed exposure setting

#### A. Implementation

Figure 4 shows our camera system for the experimental validation. For the comparative evaluation, we used three Flea3 cameras, which use a Sony ICX424 CCD  $1/3''$  sensor that has a  $640 \times 480$  resolution with 58.72 dB dynamic range. The three cameras were placed in parallel and were synchronized using an internal software trigger. Each camera determined the exposure parameters using AE, ME, and our algorithm, respectively.

There are two camera parameters, *shutter speed* and *gain*, that determine the image brightness. We adjusted both parameters using the exposure level  $E$  in Eq. (3). Since increasing the gain amplifies the image noise, we initially set the gain to a small value and controlled the shutter speed according to the exposure level  $E$  until the shutter speed reached the pre-defined maximum value. After the shutter speed reached its maximum value, we adjusted the gain accordingly. The maximum shutter speed is usually determined according to a frame rate.

In our implementation, we set the maximum shutter speed value to  $25.51ms$  for the whole experiment. We used the *Sobel* operator. All other edge operators, of course, were also available. The activation threshold,  $\delta$ , was set to 0.11. This value was empirically selected. The whole processing time was around  $30 \sim 35ms$  per frame in C++.

#### B. Surveillance application

**Dataset.** To validate our method in a surveillance application, we recorded image sequences every two hours from 8:00 to 18:00 in a single day. We collected two types of dataset.

One dataset was collected from three cameras after the cameras reached steady-states (SURVEILLANCE-A). These sequences were recorded for around 10 minutes for every time step. This dataset was used to compare the performance of pedestrian detection as a surveillance application; approximately 2500 pedestrians were found to appear in total. For the ME method, we used the same parameter in the dataset; this parameter was initially set to 8:00 by manual tuning.

The other dataset was taken by sweeping for a full range of possible exposure levels in order to validate our algorithm (SURVEILLANCE-B). We sample 210 exposure parameters that cover the exposure range; therefore, the dataset consists of 1260 ( $= 210 \times 6$  time steps) images.

**Comparison of steady-state exposures.** Figure 5 provides thumbnail images of the SURVEILLANCE-B dataset. In the figure, we indicate images with rectangle markers if an image has nearest exposure levels with one of the steady-state images from the three methods at each time step. From the thumbnail, we can easily see that the ME method suffers from illumination changes and that the AE method wipes image details out due to the high luminance of sky regions. On the other hand, our method adjusts the camera parameters to capture properly exposed images with much better details than those obtained using other methods.

**Validation of our feedback system.** In Figure 6, we show the progress of our feedback system, which auto-adjusts camera exposure, as is explained in Section IV. To simulate extreme cases in which illumination conditions rapidly vary, we applied our auto-adjusting algorithm to SURVEILLANCE-B. In the figure, output images of our feedback system are presented from the left to the right. We first put both an under-exposed image (the leftmost image on the top row) and an over-exposed image (the leftmost image on the bottom row) into our feedback system. From each image, our algorithm estimated  $\hat{\gamma}$  and updated the camera exposure according to Eq. (3). Using the updated camera exposure, we were able to obtain an image that matched the exposure parameters in the dataset and iteratively apply our feedback system until  $\hat{\gamma}$  converges. Our system converged to the rightmost image; this shows that our method can reliably adjust camera exposure even in extreme cases. The numbers in the figure indicate the  $\hat{\gamma}$  that our algorithm estimates; in the rightmost image, we can observe that  $\hat{\gamma}$  converges to one as the output images converge to the proper exposure level.

It is also noteworthy that the scene has much wider dynamic range than the dynamic range of our camera. In this situation, the AE method wipes out important details in low radiance areas to prevent saturation of the sky region, as shown in Figure 7. Our method naturally adjusts camera exposure to emphasize important details by evaluating the camera exposure in the gradient domain.

**Pedestrian detection.** We compared the performance of pedestrian detection according to the different exposure methods. Pedestrian detection is one of the most important tasks for surveillance. In this experiment, we performed pedestrian detection using the SURVEILLANCE-A dataset; we used one of the state-of-the-art pedestrian detectors from [17], [18].

Figure 7 shows example results for pedestrian detection. Images obtained the AE method are under-exposed images; images obtained using the ME method are over-/under-exposed according to the illumination changes. According to the image quality, the pedestrian detector fails to detect humans in badly exposed images.

We used a HOGgles [15] for the visualizing of images from a perspective of robot vision. The HOGgles inverts HOG feature spaces back to a natural image; therefore, it is useful to understand how robot vision sees visual images. In Figure 7, HOGgles visualizations of our method show consistently detailed features in spite of the large illumination

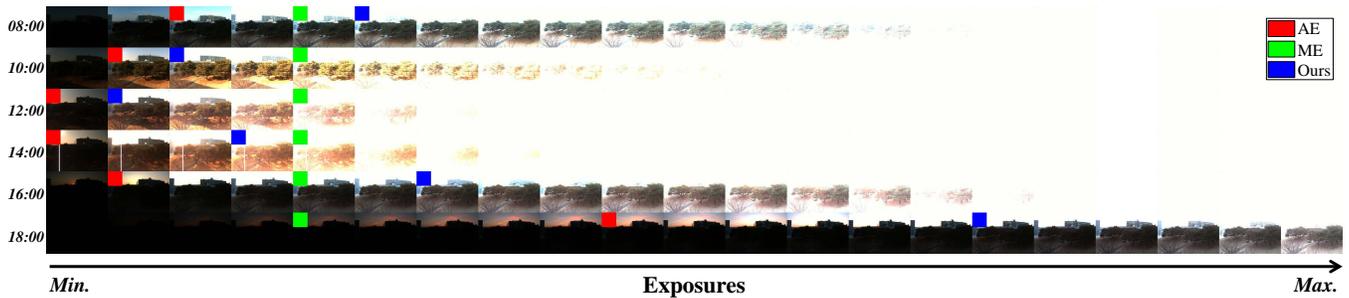


Fig. 5: Thumbnail images of the SURVEILLANCE-B dataset. Each row image was captured every two hours from 8:00 to 18:00. In every row, we display 21 images sampled from 210 images at every time step. Images in red, green, and blue markers indicate that the images have the nearest exposure levels for the steady-state exposure parameters of three methods. We recommend readers zoom-in to see the details clearly.

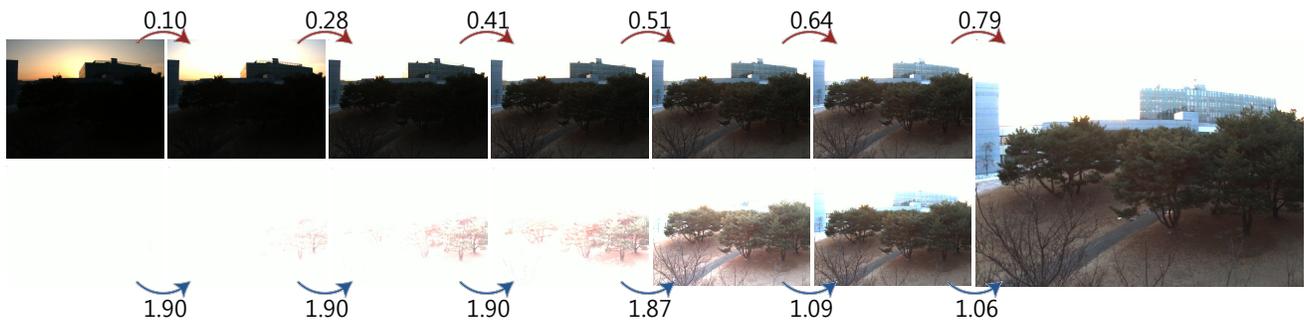


Fig. 6: Progress of our feedback system that auto-adjusts camera exposure. We first put the leftmost images into our feedback system; the system then iteratively converged to the rightmost image. The numbers indicate the  $\hat{\gamma}$  that our algorithm estimated.

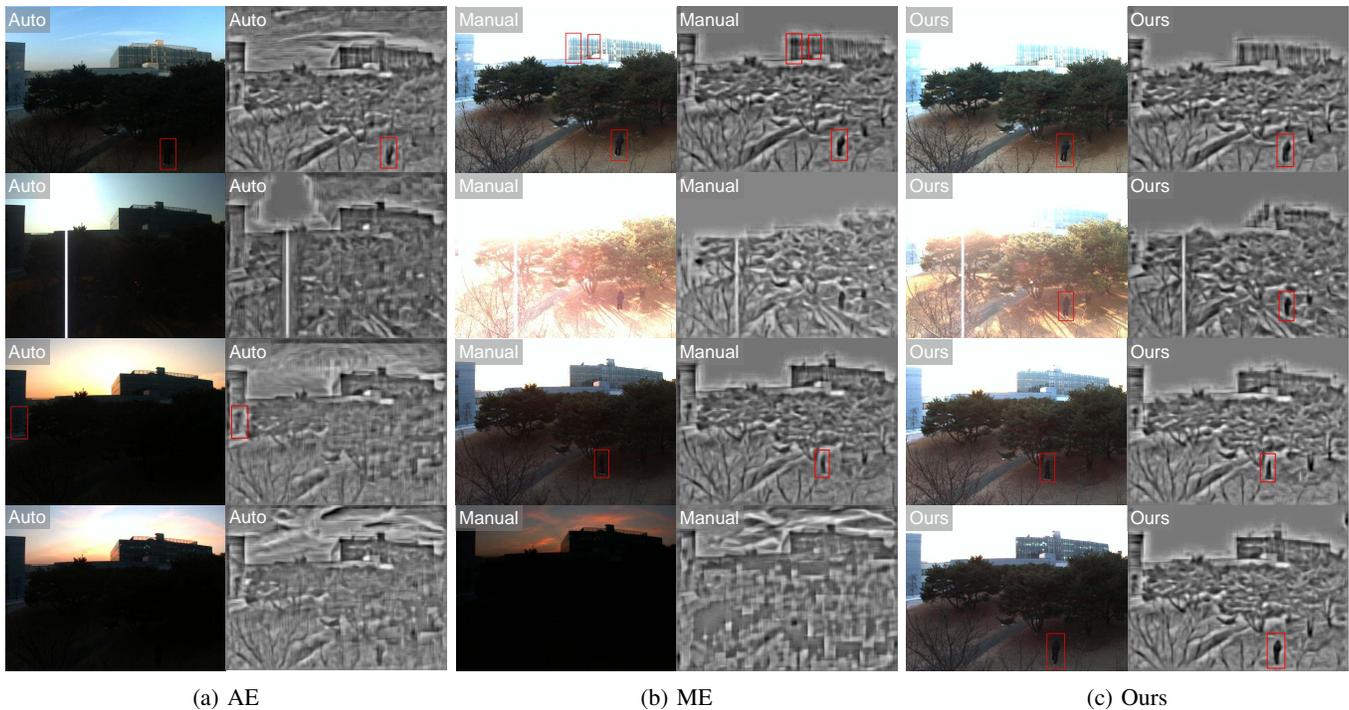


Fig. 7: The images show example results for pedestrian detection and visualizations of feature spaces using HOGgles [15]. All input images are contained in the SURVEILLANCE-A dataset; images in each row were captured at the same time. From the top to the bottom, images were captured around at 8:00, 14:00, 16:00, and 18:00.

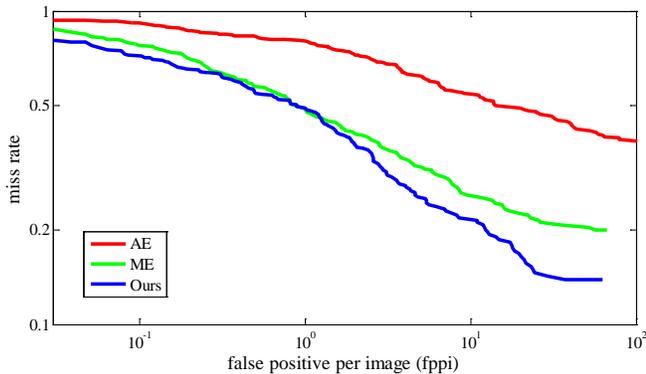


Fig. 8: Quantitative evaluation results for the pedestrian detection experiment. In the figure, it can be seen that there is better performance as results get close to the bottom left.

changes; on the other hand, the AE and ME methods cannot preserve visual information in the low radiance regions due to poor exposure. Visualization using the HOGgles method clearly shows that our method is much more suitable than the conventional AE and ME methods for outdoor vision tasks.

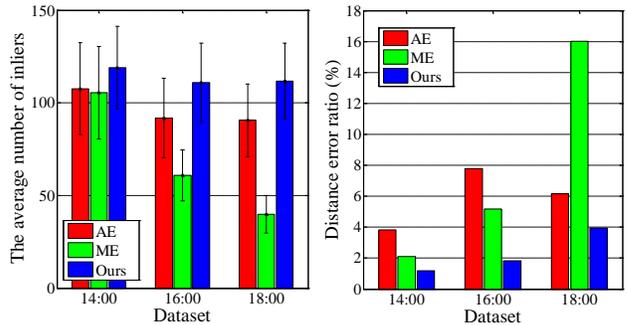
Figure 8 shows the result of quantitative evaluation for the pedestrian detection experiment. For the evaluation, the ground-truth is manually labeled for all the data; we depict the miss rate against false positives per image (fppi) as an evaluation metric according to [20]. In the evaluation metric, this means better performance as results get closer to the bottom left; therefore, the figure shows that our method outperforms the two other methods.

### C. Automotive visual odometry application

**Dataset.** For automotive visual odometry, we collected images taken from a vehicle driving through a campus. To validate the performance under different illumination conditions, we drove the same path three times at 14:00, 16:00, and 18:00. We made the path a closed-loop in order to allow us to easily measure the translation error between the starting point and the ending point of the path. The exposure parameter for the ME method was initialized at 14:00.

**Visual odometry.** We performed visual odometry with the automotive driving dataset. Visual odometry is the process of incrementally estimating the pose of a vehicle by analyzing the images of a camera mounted on the vehicle. It is an essential technique for autonomous navigation of vehicles and robots. In order to conduct visual odometry, we calibrated the intrinsic camera parameters; we also calibrated the pitch angle and the height from the ground using the vanishing points of the image. We used the monocular slam algorithm from [19] for this experiment.

Figure 9 shows the trajectories of the vehicle for qualitative comparison. Since the vehicle drove a closed-loop path, the ground-truth of the starting points and ending points of all trajectories can be seen in the figure to co-exist at  $(0, 0)$ . We were able to observe that the results of our method have more similar trajectories among different time results and smaller distance error between starting points and ending points than do the results of the other methods.



(a) The number of inliers

(b) Distance error ratio

Fig. 11: Quantitative evaluation results for the visual odometry experiment. (a) Mean and standard deviation of the number of inlier features, (b) relative distance errors.

Images at two locations on the path are shown in Figure 10. In the figure, we can observe that the illumination conditions excessively varied according to space and time, and that both the AE and the ME methods failed at feature tracking due to under-exposure, while our method captured well-exposed images and successfully tracked the image features.

In Figure 11, we present the quantitative evaluation results. The statistics on the number of inlier features are depicted in (a) and the distance error ratios for each trajectory are presented in (b). Distance error ratio is computed by dividing the distance error between the starting point and the ending point of a closed-loop trajectory by the length of the estimated trajectory. Distance error ratio is used as an evaluation metric because an estimated trajectory has a scale ambiguity in monocular visual odometry. In the figure, the results of our method can be seen to have consistently extracted larger numbers of inlier features and shown smaller errors compared to those of the AE and the ME methods. To allow for a robust estimation, a larger number of inlier features is preferred and this is the main reason that our method shows results that are better those of the other methods.

## VI. CONCLUSIONS

We have presented a novel auto-exposure method that is designed in particular for outdoor robotics applications. In outdoor environments, conventional methods fail to capture well-exposed images; this failure degrades the performance of vision-based algorithms. We adjusted the camera exposure to maximize the gradient information of a scene; therefore, our method was able to determine the proper exposure and was robust against severe illumination changes even when scene radiance had a much wider dynamic range than the possible range of the camera. We evaluated our method with extensive outdoor experiments and have proven that it is suitable for outdoor robotics.

### ACKNOWLEDGMENT

The Authors gratefully acknowledge the support from UTRC(Unmanned Technology Research Center) at KAIST

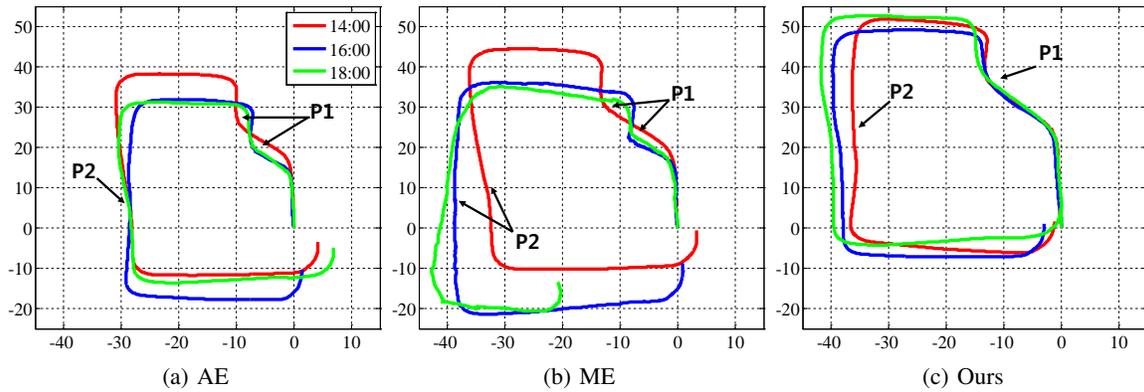


Fig. 9: Trajectories estimated using visual odometry [19]. Since the vehicle drove a closed-loop path, the ground-truth of the starting points and the ending points of all the trajectories are laid on (0,0). The results obtained using our method have smaller estimation errors than do those of the other methods.

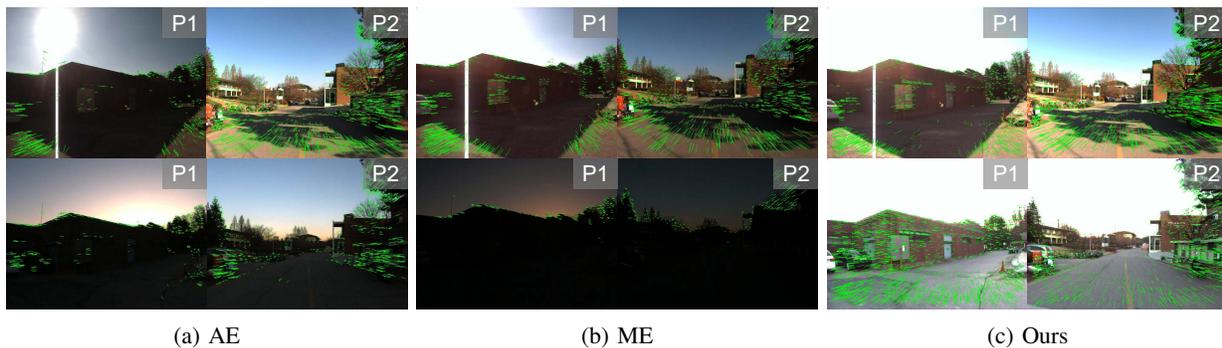


Fig. 10: Images at two locations in the path are shown. The locations are indicated in Figure 9. The top row shows images at 14:00; the bottom row shows images at 18:00. In the images, green lines indicate tracked image features between adjacent frames.

(Korea Advanced Institute of Science and Technology), originally funded by DAPA, ADD.

#### REFERENCES

- [1] M. Muramatsu, "Photometry device for a camera," Jan. 7 1997, uS Patent 5,592,256.
- [2] B. K. Johnson, "Photographic exposure control system and method," Jan. 3 1984, uS Patent 4,423,936.
- [3] N. Sampat, S. Venkataraman, T. Yeh, and R. L. Kremens, "System implications of implementing auto-exposure on consumer digital cameras," in *Proceedings of the SPIE Electronics Imaging Conference*, 1999, pp. 100–107.
- [4] S. Nuske, J. Roberts, and G. Wyeth, "Extending the dynamic range of robotic vision," in *IEEE International Conference on Robotics and Automation*, 2006, pp. 162–167.
- [5] S. Hrabar, P. Corke, and M. Bosse, "High dynamic range stereo vision for outdoor mobile robotics," in *IEEE International Conference on Robotics and Automation*, 2009, pp. 430–435.
- [6] W.-C. Kao, C.-C. Hsu, C.-C. Kao, and S.-H. Chen, "Adaptive exposure control and real-time image fusion for surveillance systems," in *IEEE International Symposium on Circuits and Systems*, 2006, pp. 4–pp.
- [7] N. Nourani-Vatani and J. Roberts, "Automatic camera exposure control," in *Proc. of the 2007 Australasian Conference on Robotics and Automation, Brisbane, Australia*, 2007.
- [8] H. Lu, H. Zhang, S. Yang, and Z. Zheng, "Camera parameters auto-adjusting technique for robust robot vision," in *IEEE International Conference on Robotics and Automation*, 2010, pp. 1518–1523.
- [9] A. J. Neves, B. Cunha, A. J. Pinho, and I. Pinheiro, "Autonomous configuration of parameters in robotic digital cameras," in *Pattern Recognition and Image Analysis*. Springer, 2009, pp. 80–87.
- [10] M. Montalvo, J. M. Guerrero, J. Romeo, M. Guijarro, M. Jesús, and G. Pajares, "Acquisition of agronomic images with sufficient quality by automatic exposure time control and histogram matching," in *Advanced Concepts for Intelligent Vision Systems*. Springer, 2013, pp. 37–48.
- [11] J.-S. Lee, Y.-Y. Jung, B.-S. Kim, and S.-J. Ko, "An advanced video camera system with robust af, ae, and awb control," *IEEE Transactions on Consumer Electronics*, vol. 47, no. 3, pp. 694–699, 2001.
- [12] M. Murakami and N. Honda, "An exposure control system of video cameras based on fuzzy logic using color information," in *IEEE the Fifth International Conference on Fuzzy Systems*, vol. 3, 1996, pp. 2181–2187.
- [13] D. G. Lowe, "Object recognition from local scale-invariant features," in *IEEE International Conference on Computer Vision*, vol. 2. IEEE, 1999, pp. 1150–1157.
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1. IEEE, 2005, pp. 886–893.
- [15] C. Vondrick, A. Khosla, T. Malisiewicz, and A. Torralba, "HOGgles: Visualizing Object Detection Features," *IEEE International Conference on Computer Vision*, 2013.
- [16] D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-laplacian priors," in *Advances in Neural Information Processing Systems*, 2009, pp. 1033–1041.
- [17] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2008, pp. 1–8.
- [18] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [19] A. Geiger, J. Ziegler, and C. Stillér, "Stereoscan: Dense 3d reconstruction in real-time," in *Intelligent Vehicles Symposium*, 2011.
- [20] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 4, pp. 743–761, 2012.