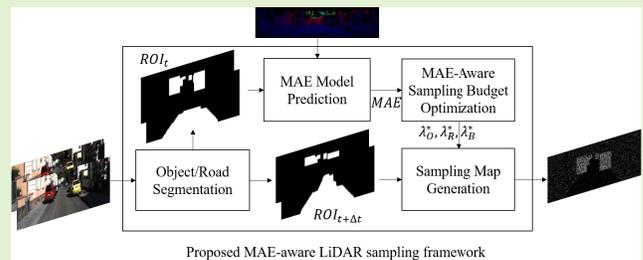


MLS: An MAE-Aware LiDAR Sampling Framework for On-Road Environments Using Spatio-Temporal Information

Quan-Dung Pham, Xuan Truong Nguyen¹, Khac-Thai Nguyen, Hyun Kim², *Member, IEEE*,
and Hyuk-Jae Lee³, *Member, IEEE*

Abstract—In recent years, light detection and ranging (LiDAR) sensors have been widely utilized in various applications, including robotics and autonomous driving. However, LiDAR sensors have relatively low resolutions, take considerable time to acquire laser range measurements, and require significant resources to process and store large-scale point clouds. To tackle these issues, many depth image sampling algorithms have been proposed, but their performances are unsatisfactory in complex on-road environments, especially when the sampling rate of measuring equipment is relatively low. Although region-of-interest (ROI)-based sampling has achieved some promising results for LiDAR sampling in on-road environments, the rate of ROI sampling has not been thoroughly investigated, which has limited reconstruction performance. To address this problem, this article proposes a solution to the budget distribution optimization problem to find optimal sampling rates according to the characteristics of each region. A simple yet effective mean absolute error (MAE)-aware model of reconstruction errors was developed and employed to analytically derive optimal sampling rates. In addition, a practical LiDAR sampling framework for autonomous driving was developed. Experimental results demonstrate that the proposed method outperforms all previous approaches in terms of both the object and overall scene reconstruction performances.

Index Terms—Autonomous driving, LiDAR sampling, on-road environment, ROI-based sampling.



I. INTRODUCTION

AUTONOMOUS driving has recently attracted attention; with the goal of reducing traffic accidents, congestion, and pollution. Moreover, when autonomous driving is

Manuscript received December 31, 2020; revised January 21, 2021; accepted January 31, 2021. Date of publication February 5, 2021; date of current version March 5, 2021. This work was supported in part by the MOTIE/KEIT through the Research and Development Program (Development of Deep Learning-Based Low-Power HW IP Design Technology for Image Processing of CMOS Image Sensors) under Grant 20010582 and in part by the National Research Foundation of Korea (NRF) funded by the Ministry of Education through the Basic Science Research Program under Grant NRF-2019R1A6A1A03032119. The associate editor coordinating the review of this article and approving it for publication was Dr. Amitava Chatterjee. (Corresponding author: Hyun Kim.)

Quan-Dung Pham, Xuan Truong Nguyen, Khac-Thai Nguyen, and Hyuk-Jae Lee are with the Department of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea (e-mail: pqdung@snu.ac.kr; truongnx@capp.snu.ac.kr; haink@snu.ac.kr; hyuk_jae_lee@capp.snu.ac.kr).

Hyun Kim is with the Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul 01811, South Korea, and also with the Research Center for Electrical and Information Technology, Seoul National University of Science and Technology, Seoul 01811, South Korea (e-mail: hyunkim@seoultech.ac.kr).

Digital Object Identifier 10.1109/JSEN.2021.3057383

integrated with emerging drive-sharing services, it eliminates the enormous cost of owning personal vehicles [1]–[3]. Autonomous driving, which is considered to be the future of driving, is actively being developed in both academia and industry [4]–[6]. In autonomous driving, sensors, such as grayscale/color camera sensors, inertial and GPS navigation sensors, radio detection and ranging (RADAR) sensors, and light detection and ranging (LiDAR) sensors are installed in the vehicle to imitate the complex natural sensing system of humans [7], [8]. Among these sensors, the LiDAR sensor is based on range sensing, which measures the time interval between the emission of light from the sensor and the arrival of light reflected from distant objects to estimate the distance to them. Therefore, a LiDAR sensor can provide rich information over a wide field of view (FOV) [9]–[12]. Since LiDARs are becoming the most popular sensor in environment perception [13], development of a real-time simultaneous localization and mapping (SLAM) solution based on 3D LiDAR measurements that can work in large-scale environments, is of great importance. By scanning at specified spatial intervals and directions within a surface through a series of observations, a topography that represents a 3D data cloud can be readily

obtained. Various scanning patterns have been designed and used; for example, the uniform raster scan or circular scan trajectories [14]. Although a LiDAR sensor detects obstacles on the road in autonomous driving, there are many challenges to real-time application. First, the quality of LiDAR measurements strongly depends on the reflective properties and angle of the reflecting surface. In practice, measurements may be lost because the reflected laser beams are not returned during the distance measurement process. As a result, it is necessary to increase the laser power, which leads to a reduction in the number of measurements for safety. Second, although a LiDAR sensor is capable of constructing a high-definition map of objects, it requires considerable resources to process and store large-scale point clouds. Therefore, it is critical to limit the number of measurements while maintaining the captured information. Many applications such as SLAM have modern 3D laser-range scanners, such as LiDAR, that are computationally intensive owing to their high data rates [15], [16]; this may be a barrier for real time processing while preserving precision up to a certain level. Third, LiDAR sensors require significant time to acquire laser range measurements.

A. Related Work

To address the issues of quality of measurement from LiDAR and the limitation of resources to process and store large-scale point clouds, fast and accurate sampling methods have recently been proposed [12], [17]–[21]. However, these methods have some shortcomings. Howe *et al.*'s [17] approach, which is motivated by the property of wavelet transforms, i.e., relevant coefficients coincide with discontinuities, suggested that a data acquisition system should pick samples at the discontinuities or along the gradients. However, this method is not practical owing to the lack of gradients in the disparity map prior to sampling and the difference between the gradient of color images and that of disparity maps. Therefore, it is challenging to infer the disparity gradient from the color image. Schwartz *et al.* [18], [20] proposed a saliency-guided sampling approach to perform sampling in a two-stage manner. First, approximately 10% of the samples are randomly sampled, and an approximate depth map is derived from the sampled data. Following this, object information or saliency is extracted from the estimated depth to select better locations with the remaining sample budget. The approach proposed by Liu *et al.* [19] is analogous to two-stage sampling. In the first stage, during the pilot stage, half of the sample budget is allocated randomly or along the gradients of a color image. Second, the refinement stage is used to estimate a round disparity map and then compute locations for the remaining sample budgets. Liu and Nguyen [21] proposed a motion compensation-assisted sampling (MCAS) scheme for reconstructing depth video sequences from a subset of samples. By exploiting the temporal information and corresponding RGB images, the proposed MCAS scheme achieves efficient one-stage sampling. However, this approach relies on motion compensation and motion vector estimation, which entails heavy computational loads, and is accompanied by problems related to fast-moving and thin objects, such as occlusion and disocclusion, brightness change, and motion blur [22].

In conclusion, these approaches require time-consuming rough disparity estimations and are consequently inappropriate for autonomous driving in on-road environments owing to the complexity of the background in an outdoor environment, the reliability of the gradient image of a scene in outdoor scenarios, and the reduced focus on object and road areas, which are significantly important for autonomous driving in on-road environments.

Nguyen *et al.* [12] proposed a LiDAR sampling method to distribute samples in an on-road environment based on semantic segmentation information from the region of interest (ROI). Given a fixed number of samples, sampling budgets are moved from the background into road areas, significantly enhancing the quality of object reconstruction. However, the ratio between the sampling rates of objects, roads, and background areas has not been thoroughly investigated. Therefore, the overall reconstruction quality is degraded.

B. Contributions

To complement the drawbacks of existing studies and develop the most suitable sampling method for autonomous driving, this article presents a LiDAR sampling framework to minimize the mean average error (MAE) of the object and overall regions. The key concepts are to build an MAE model reflecting the characteristics of the object, road, and background regions in an autonomous driving environment and to find optimal sampling rates for each region based on its characteristics (i.e., MAE model). Because MAE models are not available prior to sampling, the proposed framework exploits temporal information to efficiently predict MAE functions.

The three contributions of this article are as follows:

- 1) *ROI-Based Sampling* (Section III-A). An efficient sampling strategy to maximize reconstruction performance is proposed. We demonstrate that for a given sampling budget, a sampling pattern can be obtained by allocating random samples in object, road, and background regions with different sampling rates. In particular, each region's sampling optimization problem is solved independently by allocating random samples with probabilities proportional to the magnitudes of the depth gradients. This makes it possible to solve the problem that the sampling process of existing studies is time-consuming.
- 2) *MAE-aware Budget Distribution Optimization* (Section III-B). While the sampling rates for the ROI were not well-defined in the previous research done by Nguyen *et al.* [12], we propose an optimization scheme that determines the optimal sampling rates for the object, road, and background regions by minimizing the overall reconstruction error. Both hard and soft optimization schemes are presented. This enables the most accurate sampling of the proposed method compared to existing state-of-the-art studies.
- 3) *MAE Modelling and Practical Considerations* (Section IV). A simple yet efficient MAE model is presented. Moreover, the sampling budget optimization is simplified, and its analytical solution is derived.

In practice, MAE models depend on sample data; thus, prior sampling cannot be used, which is the main drawback of the existing studies in [18] and [20]. To address this problem, the MAE models were predicted by exploiting temporal samples from a previous frame. Moreover, by predicting MAE models, the time-consuming process of estimating disparity maps in the existing studies [17]–[21] can be eliminated. Finally, the efficiency and robustness of the proposed framework are demonstrated.

The remainder of this article is organized as follows. Section II describes the background of this study, the probabilistic model of the sampling problem, and ROI-based LiDAR sampling. Sections III and IV introduce the proposed MAE-aware sampling framework and its practical considerations, respectively. The experimental results are presented in Section V, and conclusions are presented in Section VI.

II. BACKGROUND

A. Probabilistic Model of Sampling Problem

In [19], a probabilistic model was used to represent the sampling problem. For N locations in an FOV, a diagonal matrix $S \in R^{N \times N}$ represents the sampling operation with the $(i, i)^{th}$ entry of S as follows:

$$S_{ii} = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases} \quad (1)$$

where $\{p_i\}_{i=1}^N$ is a sequence of predefined probabilities, and for each $i = 1, \dots, N$, p_i must be bounded such that $0 \leq p_i \leq 1$. Given S , the sampled data $b \in R^{N \times 1}$ is defined as follows:

$$b = Sx \quad (2)$$

where the i^{th} entry b_i is zero if $S_{ii} = 0$. The target budget is defined by the target sampling ratio ζ as an alternative to minimizing the number of sampled data with $0 < \zeta < 1$, which is used to represent the average sampling frequency. Then, the following constraint is derived:

$$\frac{1}{N} \sum_i^N p_i = \zeta \quad (3)$$

For large N , the standard concentration inequality guarantees that the average number of entries in S is approximately ζN (i.e., $\zeta N = M$) [19].

Let $a = [a_1, \dots, a_N]^T$ be a vector representing the magnitude of the gradient of the depth map. The intuition underlying Liu *et al.*'s [19] method is that the average gradient computed from all N samples is similar to the average gradient computed from a subset of ζN samples.

Let $\{p_i\}_{i=1}^N$ be the optimal sampling probability, which can be determined by solving the following optimization problem:

$$\min_{p_1, \dots, p_N} \frac{1}{N} \sum_{i=1}^N \frac{a_i^2}{p_i} \quad (4)$$

subject to $\frac{1}{N} \sum_{i=1}^N p_i = \zeta$ and $0 \leq p_j \leq 1$. The solution for this optimization is formulated as follows:

$$p_i = \min(\tau a_i, 1) \quad (5)$$

where τ is the root of the equation $g(\tau) = \sum_{i=1}^N \min(\tau a_i, 1) - \zeta N$.

B. ROI-Based LiDAR Sampling

In [12], Nguyen *et al.* proposed an ROI-based sampling algorithm for an on-road environment for autonomous driving. With the aid of a state-of-the-art object and road detection method based on convolutional neural networks [5], [23], this sampling method utilizes semantic information, as shown in Fig. 1. A scene is assumed to be segmented into three regions: road, object, and background. Let S_O , S_R , and S_B be the index sets of points in the object, road, and background areas, respectively. The union of the three sets is the set of positions in the FOV, and the intersection of any two sets is an empty set as follows:

$$S_O \cup S_R \cup S_B = \{1, 2, \dots, N\} \quad (6)$$

$$S_O \cap S_R = S_O \cap S_B = S_R \cap S_B = \emptyset \quad (7)$$

Because the road, object, and background areas have different characteristics, as shown in Fig. 2, the sampling problem can be reformulated with two weighting parameters:

$$\min_{p_1, \dots, p_N} \frac{1}{N} \left(\sum_{i \in S_B} \frac{a_i^2}{p_i} + \alpha \sum_{j \in S_R} \frac{a_j^2}{p_j} + \beta \sum_{k \in S_O} \frac{a_k^2}{p_k} \right) \quad (8)$$

subject to $\frac{1}{N} \sum_{j=1}^N p_j = \zeta$ and $0 \leq p_j \leq 1$, where α and β are the weights of the road and object areas, respectively. Although Nguyen *et al.* [12] significantly enhanced the quality of object reconstruction, the ratio between the sampling rates of the object, road, and background areas has not been thoroughly investigated, and solving this problem was an important motivation for this study.

C. Point Cloud and Memory Space in a LiDAR System

1) *Point Cloud*: Point clouds produced by LiDAR are visualized for the ease of measurement. A point cloud is basically a set of data points in a 3D coordinate system, commonly defined by x, y, and z coordinates. They are used to represent the surface of an object and do not contain data of any internal features, color, materials, and so on [24]. Normally, in a conventional LiDAR system, LiDAR acquires N locations in the FOV, but if the sampling algorithm is applied, the number of locations acquired can be reduced to $M = \zeta N$, as mentioned in Section II-A.

2) *Memory Space*: In [25], a $G()$ function was used to represent the amount of stored/transferred data. Therefore, $G()$ depends on the sampling budget ζ , depth resolution n , and sampling pattern SP , and can be expressed as follows:

$$G(\zeta, n, SP) = G(M, n, i_1, \dots, i_M) \quad (9)$$

If the sampling pattern is non-uniform and irregular because each pixel consists of n -bits, the amount of data becomes $n \times \zeta \times N$ -bits. In addition, the sampling pattern SP is stored and/or transmitted. Because one bit is necessary for each pixel in the input image of size N , the amount of data for SP is N .

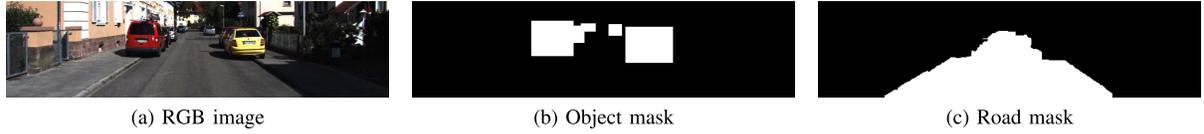
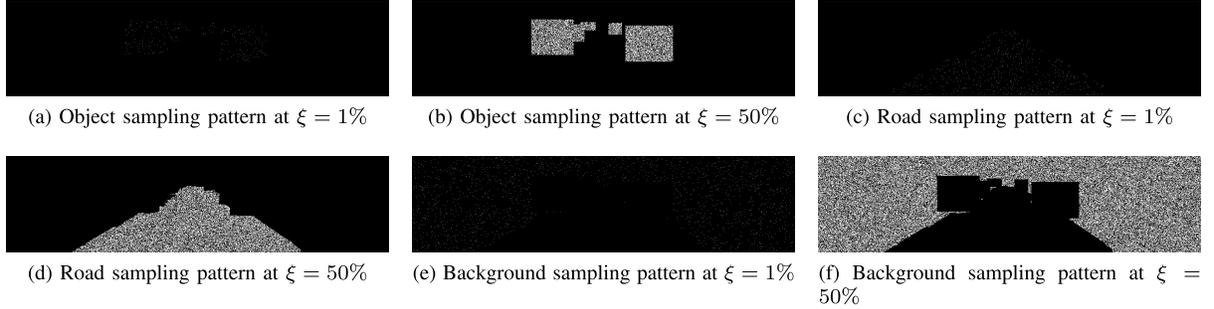


Fig. 1. RGB image and segmentation results.

Fig. 2. Sampling patterns at sampling rates $\xi = 1\%$ and $\xi = 50\%$.

Therefore, the total amount of data $G(\xi, n, SP)$ is derived as follows:

$$G(\xi, n, SP) = n \times \xi \times N + N \quad (10)$$

Uniform grid sampling does not require the sampling map to be stored because the pattern is fixed. Therefore, the required memory space $G(\xi, n, SP)$ is modified as follows:

$$G(\xi, n, SP) = n \times \xi \times N \quad (11)$$

III. ROI-BASED SAMPLING SCHEME AND MAE-AWARE SAMPLING BUDGET DISTRIBUTION

This section presents an ROI sampling scheme. Let $\{p_m\}_{m \in S_O}$, $\{p_n\}_{n \in S_R}$, and $\{p_k\}_{k \in S_B}$ be the optimal sampling probabilities in the object, road, and background areas, respectively. $\{p_m\}$, $\{p_n\}$, and $\{p_k\}$ must be bounded such that $0 \leq p_m \leq 1$, $0 \leq p_n \leq 1$, and $0 \leq p_k \leq 1$. In contrast to [12], three separate sampling ratios, λ_R , λ_O , and λ_B ($0 < \lambda_O, \lambda_R, \lambda_B < 1$) are defined for the road, object, and background areas, respectively. Eventually, the average of probabilities in the road, object, and background areas must reach their target sampling ratios:

$$\frac{\sum_{m \in S_O} p_m}{N_O} = \lambda_O, \quad \frac{\sum_{n \in S_R} p_n}{N_R} = \lambda_R, \quad \frac{\sum_{k \in S_B} p_k}{N_B} = \lambda_B \quad (12)$$

where N_O , N_R , and N_B are the number of sampling points in the road, object, and background areas, respectively. The total sampling budget in the ROI ξN consisting of the sampling budget in object $\lambda_O N_O$, road $\lambda_R N_R$, and background $\lambda_B N_B$ must achieve the sampling budget of the scene ξN :

$$\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \xi N \quad (13)$$

A. ROI-Based Random Sampling Scheme

Inspired by [19], the proposed sampling scheme is derived by minimizing the variance of the average gradient computed

for the object, road, and background areas, where the variance of the average gradient for each area is as follows:¹

$$\mathbb{E}[(Y_O - \mu_O)^2] = \frac{1}{N_O^2} \sum_{m \in S_O} a_m^2 \left(\frac{1 - p_m}{p_m} \right) \quad (14)$$

$$\mathbb{E}[(Y_R - \mu_R)^2] = \frac{1}{N_R^2} \sum_{n \in S_R} a_n^2 \left(\frac{1 - p_n}{p_n} \right) \quad (15)$$

$$\mathbb{E}[(Y_B - \mu_B)^2] = \frac{1}{N_B^2} \sum_{k \in S_B} a_k^2 \left(\frac{1 - p_k}{p_k} \right) \quad (16)$$

Similar to [19], the intuition underlying this sampling method is that the average gradient computed for all N samples is similar to the average gradient computed for a subset of ξN samples. Moreover, in our proposed method, the overall scene is divided into three different regions. Therefore, the optimal sampling probabilities $\{p_i\}_{i=1}^N = \{p_m\}_{m \in S_O} \cup \{p_n\}_{n \in S_R} \cup \{p_k\}_{k \in S_B}$ can be determined by minimizing the sum of the variance of the average gradient computed for the object, road, and background areas $\mathbb{E}[(Y_O - \mu_O)^2]$, $\mathbb{E}[(Y_R - \mu_R)^2]$, and $\mathbb{E}[(Y_B - \mu_B)^2]$ as follows:

$$\min_{p_1, \dots, p_N} \left(\frac{1}{N_O^2} \sum_{m \in S_O} \frac{a_m^2}{p_m} + \frac{1}{N_R^2} \sum_{n \in S_R} \frac{a_n^2}{p_n} + \frac{1}{N_B^2} \sum_{k \in S_B} \frac{a_k^2}{p_k} \right) \quad (17)$$

subject to $\frac{1}{N_O} \sum_{m \in S_O} p_m = \lambda_O$, $\frac{1}{N_R} \sum_{n \in S_R} p_n = \lambda_R$, $\frac{1}{N_B} \sum_{k \in S_B} p_k = \lambda_B$, $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \xi N$, $0 \leq p_m \leq 1$, $0 \leq p_n \leq 1$, and $0 \leq p_k \leq 1$.

The optimal sampling probabilities $\{p_i\}_{i=1}^N$ can be determined by solving the above optimization problem. From (6), (7), (12), and (13), it can be deduced that the optimization problem (17) is equivalent to the optimization problem in each of the regions.

¹For consistency, in the variance $\mathbb{E}[(Y - \mu)^2]$, we use $\frac{1}{N^2}$ instead of $\frac{1}{N}$, as in [19]. See Appendix for derivation.

First, the optimization problem for the object area is as follows:

$$\min_{p_1, \dots, p_{N_O}} \frac{1}{N_O^2} \sum_{m \in S_O} \frac{a_m^2}{p_m} \quad (18)$$

subject to $\frac{1}{N_O} \sum_{m \in S_O} p_m = \lambda_O$ and $0 \leq p_m \leq 1$. The solution to this optimization is

$$p_m = \min(\tau a_m, 1) \quad (19)$$

where τ is the root of the equation $g(\tau) = \sum_{m \in S_O} \min(\tau a_m, 1) - \lambda_O N_O$.

Second, the optimization problem for the road area is as follows:

$$\min_{p_1, \dots, p_{N_R}} \frac{1}{N_R^2} \sum_{n \in S_R} \frac{a_n^2}{p_n} \quad (20)$$

subject to $\frac{1}{N_R} \sum_{n \in S_R} p_n = \lambda_R$ and $0 \leq p_n \leq 1$. The solution to this optimization is

$$p_n = \min(\tau a_n, 1) \quad (21)$$

where τ is the root of the equation $g(\tau) = \sum_{n \in S_R} \min(\tau a_n, 1) - \lambda_R N_R$.

Third, the optimization problem for the background area is as follows:

$$\min_{p_1, \dots, p_{N_B}} \frac{1}{N_B^2} \sum_{k \in S_B} \frac{a_k^2}{p_k} \quad (22)$$

subject to $\frac{1}{N_B} \sum_{k \in S_B} p_k = \lambda_B$ and $0 \leq p_k \leq 1$. The solution to this optimization is

$$p_k = \min(\tau a_k, 1) \quad (23)$$

where τ is the root of the equation $g(\tau) = \sum_{k \in S_B} \min(\tau a_k, 1) - \lambda_B N_B$. See Appendix A-A and A-B for the derivation.

B. Optimal ROI-Based Sampling Budget Distribution

1) Hard Optimal ROI-Based Sampling Budget Distribution:

Given the information about the characteristics of the MAEs of the ROI, as indicated in Section IV-A, the sampling budget distribution can be formulated as an optimization problem whose goal is to minimize the total MAEs of the ROI. Mathematically, we consider the following problem:

$$\min_{\lambda_O, \lambda_R, \lambda_B} (MAE_O + MAE_R + MAE_B) \quad (24)$$

subject to $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \xi N$, $0 \leq \lambda_O \leq 1$, $0 \leq \lambda_R \leq 1$, and $0 \leq \lambda_B \leq 1$.

2) Soft Optimal ROI-Based Sampling Budget Distribution:

Thus far, we have assumed that the object, road, and background regions contribute equally to the optimization problem (24). In reality, however, object and road areas are more important in on-road environments. Moreover, in different scenarios, the object and road regions will have different levels of importance. Therefore, we must modify the optimization problem (24) to reflect the different roles of the object, road, and background areas in minimizing MAEs in the ROI and

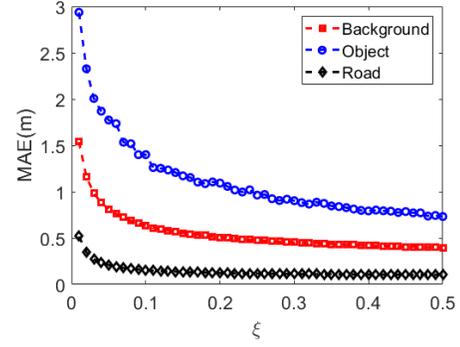


Fig. 3. Average MAEs according to sampling rates.

the trade-off between MAE_O , MAE_R , and MAE_B . To do so, we introduce $\Phi_O > 0$, $\Phi_R > 0$, and $\Phi_B > 0$, which are the weighted parameters denoting priorities for the object, road, and background regions. Consequently, we modify the optimization problem as follows:

$$\min_{\lambda_O, \lambda_R, \lambda_B} (\Phi_O MAE_O + \Phi_R MAE_R + \Phi_B MAE_B) \quad (25)$$

subject to $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \xi N$, $0 \leq \lambda_O \leq 1$, $0 \leq \lambda_R \leq 1$, and $0 \leq \lambda_B \leq 1$.

IV. PRACTICAL MAE MODELING

A. Practical Models of MAEs

Unlike in [12], this article presents a simple yet effective method to model MAEs. Fig. 3 demonstrates the average MAEs of the object, road, and background areas at sampling rates from 0.01 (=1%) to 0.5 (=50%). First, it is clearly shown that the MAEs gradually decrease as the sampling rates increase. In particular, the MAEs decrease steeply in the range of $\xi = (0, 0.1)$ and decrease gradually in the range of $\xi = (0.1, 0.5)$. This numerical example demonstrates that MAEs are likely to be modeled as rational functions with similar behaviors. In particular, the MAEs of the object, road, and background are modeled with three separate models because they have different characteristics. By applying numerical interpolation to these sets of data points, it can be deduced that the MAEs are modeled as functions $f_O(\lambda_O)$, $f_R(\lambda_R)$, and $f_B(\lambda_B)$ such that $f_O : [0, 1] \mapsto \mathbb{R}_+$, $f_R : [0, 1] \mapsto \mathbb{R}_+$, and $f_B : [0, 1] \mapsto \mathbb{R}_+$.

$$MAE_O = f_O(\lambda_O) = a_O + \frac{b_O}{c_O + \lambda_O} \quad (26)$$

$$MAE_R = f_R(\lambda_R) = a_R + \frac{b_R}{c_R + \lambda_R} \quad (27)$$

$$MAE_B = f_B(\lambda_B) = a_B + \frac{b_B}{c_B + \lambda_B} \quad (28)$$

where λ_O , λ_R , and λ_B are the sampling rates for the object, road, and background regions, respectively.

Reconsider Fig. 3, where the MAEs of the object, road, and background areas are shown by blue circle, red square, and black diamond lines, respectively, and each point represents an MAE at a sampling rate. Meanwhile, the interpolated points estimated by rational functions are represented by dashed blue,

red, and black lines, respectively. It is clearly shown that the rational functions well represent the MAE characteristics of the object, road, and background areas.

B. Analytical Solution of Sampling Budget Distribution Problem

To solve the optimization problem in (25) with the practical models shown in IV-A, the Lagrangian of the above equation can be expressed as follows:

$$\begin{aligned} \mathcal{L}(\lambda_O, \lambda_R, \lambda_B, \gamma, \eta_O, \eta_R, \eta_B, \nu_O, \nu_R, \nu_B) \\ = \Phi_O a_O + \Phi_O \frac{b_O}{c_O + \lambda_O} \\ + \Phi_R a_R + \Phi_R \frac{b_R}{c_R + \lambda_R}, \\ + \Phi_B a_B + \Phi_B \frac{b_B}{c_B + \lambda_B} \\ + \gamma (\lambda_O N_O + \lambda_R N_R + \lambda_B N_B - \zeta N) \\ - \eta_O \lambda_O - \eta_R \lambda_R - \eta_B \lambda_B \\ + \nu_O (\lambda_O - 1) + \nu_R (\lambda_R - 1) + \nu_B (\lambda_B - 1), \quad (29) \end{aligned}$$

By solving (29), the optimal solution of (25) with the practical models in IV-A can be expressed as follows:

$$\lambda_O = \max(\min(\sqrt{\frac{\Phi_O b_O}{\gamma N_O}} - c_O, 1), 0) \quad (30)$$

$$\lambda_R = \max(\min(\sqrt{\frac{\Phi_R b_R}{\gamma N_R}} - c_R, 1), 0) \quad (31)$$

$$\lambda_B = \max(\min(\sqrt{\frac{\Phi_B b_B}{\gamma N_B}} - c_B, 1), 0) \quad (32)$$

where γ is chosen based on the primal feasibility $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \zeta N$. See Appendix B-A for the derivation.

Fig. 4 shows the objective function on various planes and the minimum value of the objective function projected on various planes. Figs. 4a, 4b, and 4c illustrate the objective function $\Phi_O MAE_O + \Phi_R MAE_R + \Phi_B MAE_B$ on coordinates (λ_O, λ_R) , (λ_B, λ_O) , and (λ_B, λ_R) , respectively. In Fig. 4a, λ_B is formulated as a function of (λ_O, λ_R) . Similarly, in Fig. 4b, λ_O is formulated as a function of (λ_B, λ_R) , and in Fig. 4c, λ_R is formulated as a function of (λ_B, λ_R) . Figs. 4d, 4e, and 4f indicate the optimal solutions of λ_O , λ_R , and λ_B , respectively. In fact, these figures show the projections of objective functions, λ_O , λ_R , and λ_B , on the plane $(\Phi_O MAE_O + \Phi_R MAE_R + \Phi_B MAE_B)$, respectively.

C. Prediction MAEs From Two Consecutive Frames

To utilize the temporal information from the previous frame, we examine the characteristics of two consecutive frames. The input data of RGB and LiDAR images at time t and $t + \Delta t$ are shown in Fig. 5. It can be seen that within Δt , the scenes at two times t and $t + \Delta t$ are slightly different. The MAEs of the object, road, and background areas at time $t + \Delta t$ are

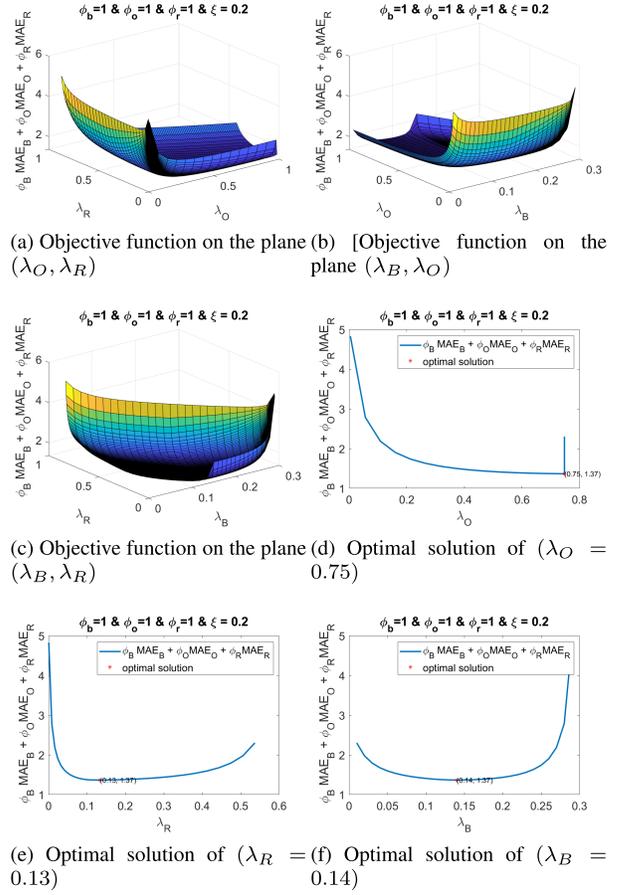


Fig. 4. Objective function on various planes and its optimal solution.

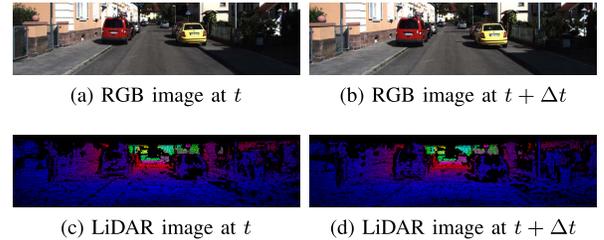


Fig. 5. Input data for the proposed sampling scheme.

defined as follows:

$$\begin{aligned} MAE_{O,t+\Delta t} &= f_{O,t+\Delta t}(\lambda_{O,t+\Delta t}) \\ &= a_{O,t+\Delta t} + \frac{b_{O,t+\Delta t}}{c_{O,t+\Delta t} + \lambda_{O,t+\Delta t}}, \quad (33) \end{aligned}$$

$$\begin{aligned} MAE_{R,t+\Delta t} &= f_{R,t+\Delta t}(\lambda_{R,t+\Delta t}) \\ &= a_{R,t+\Delta t} + \frac{b_{R,t+\Delta t}}{c_{R,t+\Delta t} + \lambda_{R,t+\Delta t}}, \quad (34) \end{aligned}$$

$$\begin{aligned} MAE_{B,t+\Delta t} &= f_{B,t+\Delta t}(\lambda_{B,t+\Delta t}) \\ &= a_{B,t+\Delta t} + \frac{b_{B,t+\Delta t}}{c_{B,t+\Delta t} + \lambda_{B,t+\Delta t}}, \quad (35) \end{aligned}$$

Fig. 6 shows the MAE difference between time t and $t + \Delta t$, and TABLE I lists the parameters of these characteristics in detail. It is evident that the characteristics of the MAEs at times t and $t + \Delta t$ are likely to be similar [26]. The profiling

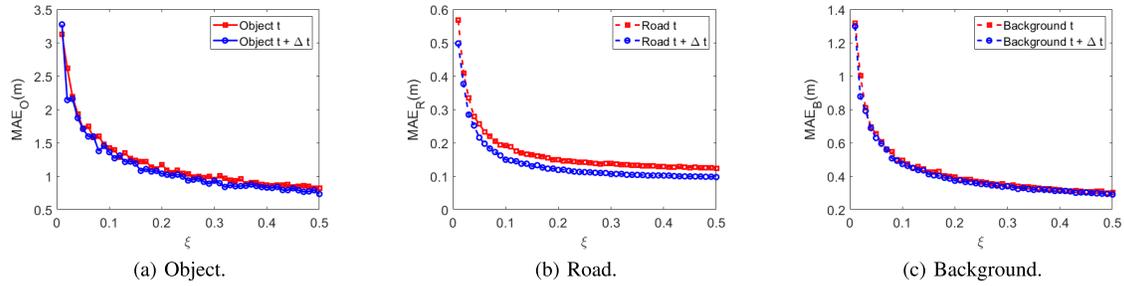


Fig. 6. Example of prediction MAEs of the object, road, and background areas from two consecutive frames with $MAE_{O,t}$ and $MAE_{O,t+\Delta t}$, $MAE_{R,t}$ and $MAE_{R,t+\Delta t}$, and $MAE_{B,t}$ and $MAE_{B,t+\Delta t}$.

TABLE I
PARAMETERS OF $MAE_{O,t}$, $MAE_{O,t+\Delta t}$, $MAE_{R,t}$,
 $MAE_{R,t+\Delta t}$, $MAE_{B,t}$, AND $MAE_{B,t+\Delta t}$

	O_t	$O_{t+\Delta t}$	R_t	$R_{t+\Delta t}$	B_t	$B_{t+\Delta t}$
a	0.79	0.75	0.11	0.09	0.30	0.31
b	0.39	0.36	0.03	0.03	0.10	0.10
c	0.16	0.14	0.07	0.05	0.09	0.08

results in Fig. 6 and TABLE I indicate a useful fact: for a certain time interval Δt , the characteristics of the MAEs at time t can be used as the characteristics of the MAEs at time $t + \Delta t$ without the need to examine the characteristics of the MAEs at time $t + \Delta t$. In other words, it can be expressed as follows:

$$MAE_{O,t+\Delta t} \approx MAE_{O,t} \quad (36)$$

$$MAE_{R,t+\Delta t} \approx MAE_{R,t} \quad (37)$$

$$MAE_{B,t+\Delta t} \approx MAE_{B,t} \quad (38)$$

Section IV-E demonstrates how this fact plays a pivotal role in dealing with the time constraint issue of sampling.

D. Prediction MAEs Using Partial Sampling

As discussed in Subsection IV-A, the MAE functions are estimated from the sample data, which are not available prior to sampling. Obviously, the accuracy of the MAE estimation depends on the number of acquired data or the sampling rate. The more data points are acquired, the more accurate the prediction MAEs are. However, acquiring many data certainly results in significant resource requirements to process these data. This subsection further considers the trade-off between the sampling rate for prediction and its MAEs.

Fig. 7 illustrates the sampling data points up to 10%, 15%, 20%, and 50%. Figs. 7b, 7d, and 7f are zoomed-out views of the prediction MAEs for the object, road, and background areas, respectively. These figures show that there are differences in the MAEs functions when using different sampling rates for the prediction stage. These differences in the MAE functions may affect the optimal solutions. In other words, the more sampling points in the prediction stage lead to the more accurate MAE functions. However, by using a small sampling rate to interpolate the MAEs of the object, road, and background areas, the resource requirements for data storage and processing can be considerably reduced.

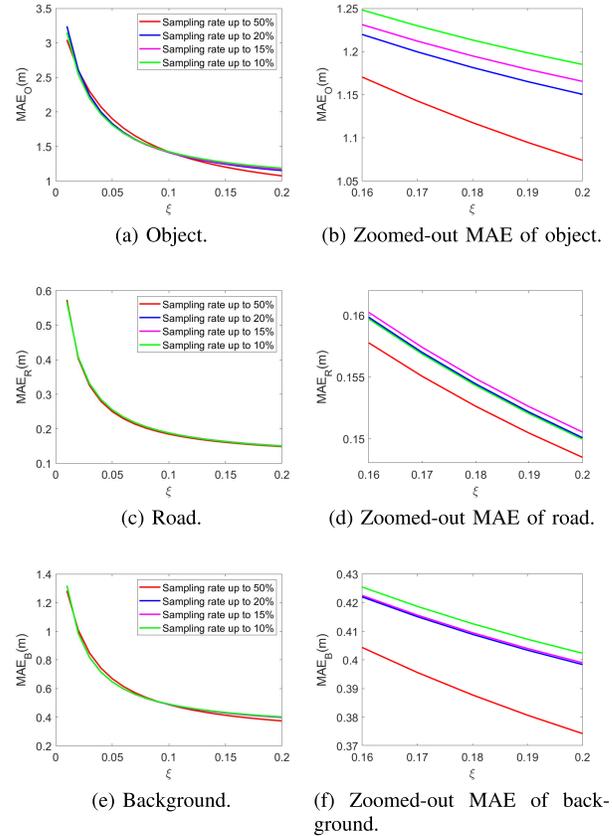


Fig. 7. Prediction MAEs of object, road, and background regions for sampling rates of 50%, 20%, 15%, and 10%.

E. MAE-Aware Sampling Framework

Fig. 8 illustrates the proposed sampling framework consisting of four stages. First, object and road detection algorithms are performed. Using a convolutional neural network, the object and road areas can be extracted fast and accurately, resulting in masks of the object, road, and background areas. Secondly, the MAE functions are estimated by utilizing the sample data from the previous frame. In particular, equations (36), (37), and (38) are applied. Thirdly, the optimal sampling budget optimization is solved, and the sampling rates of the object, road, and background areas are derived. Finally, the sampling pattern is derived by solving the sampling optimization problems for the object, road, and background regions separately, as in (19), (21), and (23), respectively.

TABLE II

MAE COMPARISON (M) BETWEEN THE PREVIOUS SAMPLING ALGORITHMS AND THE PROPOSED METHOD USING SAMPLING RATES IN THE PREDICTION STEP ON THE OBJECT AREA, ROAD AREA, AND OVERALL SCENE, 10%, 15%, AND 20%, RESPECTIVELY WITH 64 CHANNELS

	Object			Road			Overall		
	10%	15%	20%	10%	15%	20%	10%	15%	20%
Two-stage sampling [19]	1.229	0.989	0.841	0.154	0.147	0.116	0.631	0.517	0.482
ROI-based [12]	0.767	0.678	0.631	0.223	0.188	0.164	0.704	0.518	0.485
MCAS [21]	2.000	1.853	1.816	0.225	0.224	0.223	1.026	1.000	0.989
Proposed ($\Phi_O = 1, \Phi_R = 1, \Phi_B = 1$)	0.713	0.673	0.654	0.253	0.182	0.159	0.576	0.514	0.477
Proposed ($\Phi_O = 2, \Phi_R = 1, \Phi_B = 1$)	0.669	0.647	0.630	0.286	0.195	0.164	0.640	0.541	0.479
Proposed ($\Phi_O = 1, \Phi_R = 2, \Phi_B = 1$)	0.736	0.680	0.656	0.148	0.132	0.124	0.579	0.518	0.483

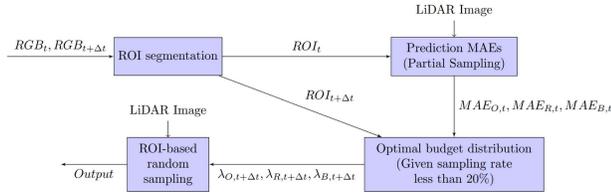


Fig. 8. Practical sampling scheme.

V. EXPERIMENTAL RESULTS

In this section, we compare the proposed sampling scheme with existing sampling methods from [12], [19], and [21], and examine the proposed method using different sampling rates at the prediction MAEs stage in the framework. To evaluate the various schemes, we utilize the well-known public KITTI [7] dataset.

A. Data Preparation

The **KITTI dataset** [7] was recorded from a moving vehicle using a variety of sensor modalities, such as high-resolution color and grayscale stereo cameras, a Velodyne 3D laser scanner, and a high-precision GPS/IMU inertial navigation system while driving in and around Karlsruhe, Germany. According to [7], the cameras are triggered by a LiDAR 3D scanner at 10 frames per second, so the time interval Δt between two consecutive frames Q and $Q+1$ is 0.1 s, as shown in Fig. 5. We selected nine pairs of 2D images that correspond to nine pairs of LiDAR images. Each pair has two images captured at t and $t + \Delta t$, where $\Delta t = 0.1$ s.

Since the spatial resolution of LiDAR applications apparently plays the most critical role in the MAE and quality of 3D images, we create several LiDAR images with various spatial resolution of 64, 32, and 16 channels. Based on the public KITTI dataset [7], the Velodyne HDL-64E is set up as a rotating 3D laser scanner with the following settings: 10 Hz, 64 beams, 0.09° angular resolution, 2 cm distance accuracy, 360° horizontal, and 26.8° vertical FOV, 120 m range. In addition, with reference to Velodyne HDL-32E [27] and Velodyne PUCK [28], to acquire the LiDAR datasets with different spatial resolutions, we modified the KITTI dataset with the following settings: 1) 32 beams, 0.09° angular resolution, 360° horizontal and 26.8° vertical FOV, range: 100 m, and 2) 16 beams, 1° angular resolution, 360° horizontal and 26.8° vertical FOV, range: 100 m. Fig. 9 shows the point cloud images of LiDAR with the different spatial resolutions.

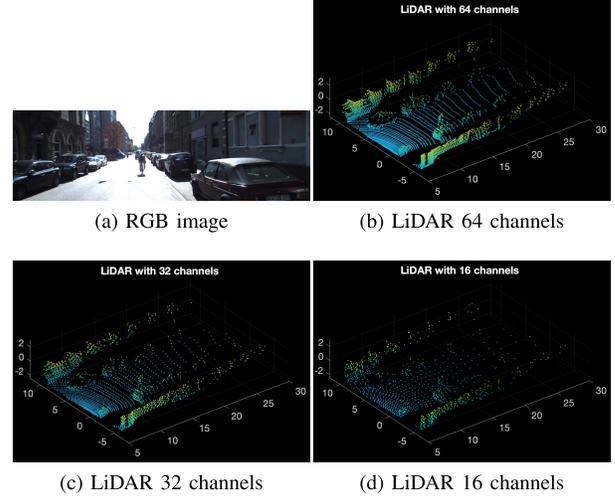


Fig. 9. RGB image and point cloud images with the different resolution.

B. Performance Comparison

1) **Mean Absolute Error:** This subsection presents comparison results between the proposed method and three previous approaches: two-stage [19], ROI-based [12], and MCAS [21]. In addition, three variations of the proposed method, ($\Phi_O = 1, \Phi_R = 1, \Phi_B = 1$), ($\Phi_O = 2, \Phi_R = 1, \Phi_B = 1$), and ($\Phi_O = 1, \Phi_R = 2, \Phi_B = 1$), are presented. In particular, the first variation considers the case in which the MAEs of the object, road, and background regions can be considered equally by setting $\Phi_O = 1, \Phi_R = 1$, and $\Phi_B = 1$. This constraint is suitable for achieving a good overall performance. Meanwhile, the MAEs of object or road areas can be emphasized in the objective function by setting their weights to larger than unity (i.e., $\Phi_O = 2$ or $\Phi_R = 2$) in order to achieve optimized performance for a specific area.

The results in Table II show that the proposed method outperformed the previous approaches in the object areas and overall scenes when using LiDAR with 64 channels. The proposed method performed much better than those from [12], [19], and [21] in terms of object area, where the proposed method reduces the error by up to 45.6%, 20.9%, and 66.6%, respectively. For the overall image, the proposed method's MAEs are reduced by up to 8.7%, 9.9%, and 43.9%, respectively. With the setting of 32 channels in Table III, the proposed method outperforms those from [12], [19], and [21] in terms of object area, where the proposed method reduces the error by up to 48.5%, 27.8%, and 67.3%, respectively.

TABLE III

MAE COMPARISON (M) BETWEEN THE PREVIOUS SAMPLING ALGORITHMS AND THE PROPOSED METHOD USING SAMPLING RATES IN THE PREDICTION STEP ON THE OBJECT AREA, ROAD AREA, AND OVERALL SCENE, 10%, 15%, AND 20%, RESPECTIVELY WITH 32 CHANNELS

	Object			Road			Overall		
	10%	15%	20%	10%	15%	20%	10%	15%	20%
Two-stage sampling [19]	2.083	1.572	1.358	0.291	0.263	0.242	0.893	0.754	0.679
ROI-based [12]	1.484	1.267	1.057	0.332	0.279	0.244	0.926	0.806	0.726
MCAS [21]	3.278	2.935	2.823	0.425	0.397	0.465	1.356	1.485	1.339
Proposed ($\Phi_O = 1, \Phi_R = 1, \Phi_B = 1$)	1.147	1.053	1.049	0.352	0.275	0.245	0.869	0.728	0.667
Proposed ($\Phi_O = 2, \Phi_R = 1, \Phi_B = 1$)	1.072	1.045	1.019	0.392	0.297	0.253	0.940	0.764	0.682
Proposed ($\Phi_O = 1, \Phi_R = 2, \Phi_B = 1$)	1.363	1.206	1.156	0.225	0.205	0.190	1.021	0.822	0.738

TABLE IV

MAE COMPARISON (M) BETWEEN THE PREVIOUS SAMPLING ALGORITHMS AND THE PROPOSED METHOD USING SAMPLING RATES IN THE PREDICTION STEP ON THE OBJECT AREA, ROAD AREA, AND OVERALL SCENE, 10%, 15%, AND 20%, RESPECTIVELY WITH 16 CHANNELS

	Object			Road			Overall		
	10%	15%	20%	10%	15%	20%	10%	15%	20%
Two-stage sampling [19]	3.009	2.828	2.786	0.603	0.569	0.469	1.448	1.302	1.249
ROI-based [12]	2.983	2.472	2.057	0.659	0.583	0.536	1.466	1.331	1.261
MCAS [21]	4.735	5.536	5.346	0.881	0.858	0.901	2.198	2.564	2.321
Proposed ($\Phi_O = 1, \Phi_R = 1, \Phi_B = 1$)	2.065	2.027	1.997	0.714	0.611	0.574	1.432	1.299	1.186
Proposed ($\Phi_O = 2, \Phi_R = 1, \Phi_B = 1$)	2.022	2.006	1.979	0.768	0.645	0.591	1.551	1.352	1.243
Proposed ($\Phi_O = 1, \Phi_R = 2, \Phi_B = 1$)	2.098	2.017	2.000	0.552	0.534	0.504	1.444	1.288	1.235

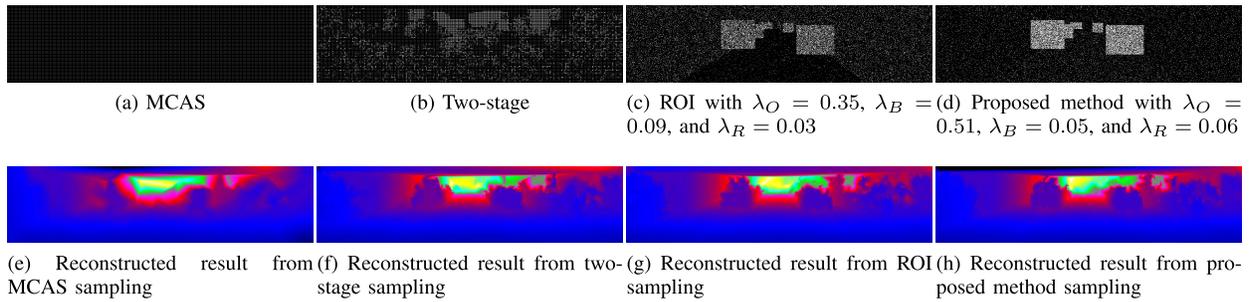


Fig. 10. Sampling maps and reconstructed results by various sampling methods with sampling rate $\xi = 0.1$.

For the overall image, the MAEs of the proposed method are reduced by up to 2.7%, 8.1%, and 50.2%. With the setting of 16 channels in Table IV, the proposed method outperforms those from [12], [19], and [21] in terms of object area, where the proposed method reduces the error by up to 32.8%, 32.2%, and 57.2%, respectively. For the overall image, the proposed method's MAEs are reduced by up to 1.1%, 5.9%, and 50.2%. It is notable that the higher weights assigned to the specific areas lead to a better performance in these areas. In other words, with increasing Φ_O and Φ_R , the MAEs in the object and road areas decrease, respectively.

Figs. 10a, 10b, 10c, and 10d show the sampling maps determined by the MCAS [21], two-stage [19], ROI-based [12], and proposed method, respectively. It is notable that the proposed method distributes more sampling budget to the object and road areas than the background area. By using the ROI technique, the proposed method can know the area with more important information with certainty. Moreover, with MAE optimization, the proposed method also performed better in sampling in the road area and overall compared to the conventional ROI method [12]. Figs. 10e, 10f, 10g, and 10h show the maps reconstructed by the MCAS [21], two-stage [19], ROI-based [12], and proposed method, respectively. It is

obvious that the proposed sampling method yields the best performance in the object area. Meanwhile, the reconstruction errors for the road areas are visually similar among all methods.

2) Memory Consumption: We compare the memory consumption when the proposed method is applied and the sampling algorithm is not used. In LiDAR applications that do not use a sampling algorithm, a depth image with an image input size of N and a resolution of n -bits can be considered as having a sampling rate of 100% and a uniform sampling pattern. Therefore, the required memory space can be calculated as follows:

$$G(\zeta = 100\%, n, S) = n \times \zeta \times N = n \times N \quad (39)$$

On the other hand, for the proposed method, the required memory space is as follows:

$$\begin{aligned} G(\xi, n, S) &= n \times (\lambda_O N_O + \lambda_R N_R + \lambda_B N_B) + N \\ &= n \times \xi \times N + N \end{aligned} \quad (40)$$

By comparing (39) and (40), the maximum memory savings achieved by the proposed method can be calculated as follows:

$$\chi = 1 - \frac{n\xi + 1}{n} \quad (41)$$

TABLE V
MEMORY SAVING OF THE PROPOSED METHOD
ACCORDING TO VARIOUS SAMPLING RATES

Sampling rate	Saving memory ratio
0.1	77.5%
0.15	72.5%
0.2	67.5%

TABLE V shows the memory saving ratio according to the various sampling rates when a depth image of size N with 8-bit resolution is sampled and stored in memory. We can see that our method saves up to 77.5% memory as compared to conventional LiDAR systems.

VI. CONCLUSION

In this article, a sampling framework to minimize the MAEs of the object and overall regions is presented with three contributions. First, we provide empirical results for modeling the characteristics of MAEs in the object, road, background, and overall areas. Second, by assuming that the characteristics of each area are known prior to the sampling process, we propose a method to minimize the MAE in each region by using weighted parameters for each area. The weighted parameters can be efficiently selected according to the environment in which autonomous driving is conducted. Third, we propose an efficient LiDAR sampling scheme. The current framework only works for offline processes, but it can further accelerate the proposed framework by using GPU implementation to achieve real-time operation and support LiDAR sampling in on-road environments. Therefore, the proposed framework can contribute significantly to the commercialization of autonomous driving systems.

ACKNOWLEDGMENT

This work also received useful comments on mathematical proofs by Tri Phu Vu (Department of Economics, Seoul National University) and Ha Young Shin (Department of Statistics, Seoul National University).

APPENDIX A

RANDOM SAMPLING SCHEME

A. Proof of the Variance $\mathbb{E}[(Y - \mu)^2]$

$$\begin{aligned}
 \text{Var}[Y] &= \text{Var} \left[\frac{1}{N} \sum_{j=1}^N \frac{a_j}{p_j} I_j \right] \\
 &= \frac{1}{N^2} \sum_{j=1}^N \text{Var} \left[\frac{a_j}{p_j} I_j \right] \\
 &= \frac{1}{N^2} \sum_{j=1}^N \frac{a_j^2}{p_j^2} \text{Var}[I_j] \quad (42)
 \end{aligned}$$

$\{I_j\}_{j=1}^N$ is a sequence of Bernoulli random variables with probabilities $\text{Pr}[I_j = 1] = p_j$, so $\text{Var}[I_j] = p_j(1 - p_j)$.

Therefore, (42) is as follows:

$$\text{Var}[Y] = \frac{1}{N^2} \sum_{j=1}^N a_j^2 \left(\frac{1 - p_j}{p_j} \right) \quad (43)$$

Also, $\mathbb{E}[(Y - \mu)^2] = \text{Var}[Y]$.

Thus,

$$\mathbb{E}[(Y - \mu)^2] = \frac{1}{N^2} \sum_{j=1}^N a_j^2 \left(\frac{1 - p_j}{p_j} \right) \quad (44)$$

B. Auxiliary Results for Optimization Problem

$$\min_{p_1, \dots, p_N} \frac{1}{N^2} \sum_{j=1}^N \frac{a_j^2}{p_j} \quad (45a)$$

$$\text{subject to } \frac{1}{N} \sum_{j=1}^N p_j = \zeta, \quad (45b)$$

$$\delta_j \leq p_j \leq 1. \quad (45c)$$

The Lagrangian of (45) is

$$\begin{aligned}
 \mathcal{L}(\mathbf{p}, \lambda, \eta, \nu) &= \frac{1}{N^2} \sum_{j=1}^N \frac{a_j^2}{p_j} + \nu \left(\sum_{j=1}^N p_j - \zeta N \right) \\
 &\quad + \sum_{j=1}^N \lambda_j (p_j - 1) + \sum_{j=1}^N \eta_j (\delta_j - p_j) \quad (46)
 \end{aligned}$$

where $\mathbf{p} = [p_1, \dots, p_N]^T$ are the primal variables, $\lambda = [\lambda_1, \dots, \lambda_N]$, $\eta = [\eta_1, \dots, \eta_N]$, and ν are the Lagrange multipliers associated with the constraints $p_j \leq 1$, $p_j \geq \delta_j$, and $\sum_{j=1}^N p_j = \zeta N$, respectively.

The first order optimality conditions imply the following:

- *Stationarity*: $\nabla_p \mathcal{L} = 0$. That is, $-\frac{1}{N^2} \frac{a_j^2}{p_j^2} + \lambda_j - \eta_j + \nu = 0$.
- *Primal feasibility*: $\sum_{j=1}^N p_j = \zeta N$, $p_j \leq 1$, and $p_j \geq \delta_j$.
- *Dual feasibility*: $\lambda_j \geq 0$, $\eta_j \geq 0$, and $\nu \geq 0$.
- *Complementary slackness*: $\lambda_j (p_j - 1) = 0$, $\eta_j (\delta_j - p_j) = 0$.

$\lambda_j (p_j - 1) = 0$ implies that for each j , one of the following cases always holds: $\lambda_j = 0$ or $p_j = 1$.

- Case 1: $\lambda_j = 0$:
 - If $\eta_j = 0$, then $p_j \geq \delta_j$. Substituting $\lambda_j = \eta_j = 0$ into the stationarity condition: $p_j = \frac{a_j}{N\sqrt{\nu}}$. Because $\delta_j \leq p_j \leq 1$, we must have $\frac{a_j}{N} \leq \sqrt{\nu} \leq \frac{a_j}{N\delta_j}$.
 - If $p_j = \delta_j$, then $\eta_j > 0$. Substituting $p_j = \delta_j$ and $\lambda_j = 0$ into the stationarity condition, $\eta_j = \nu - \frac{1}{N^2} \frac{a_j^2}{\delta_j^2}$. Because $\eta_j > 0$, we have $\sqrt{\nu} > \frac{1}{N} \frac{a_j}{\delta_j}$.
- Case 2: $p_j = 1$:
 - $\eta_j = 0$. Substituting $p_j = 1$, $\eta_j = 0$ into the stationarity condition $\lambda_j = \frac{a_j^2}{N^2} - \nu$. Because $\lambda_j > 0$, we have that $\sqrt{\nu} < \frac{a_j}{N}$.

1) *Combining These Two Cases, We Obtain:*

$$p_j = \begin{cases} \delta_j, & \text{when } a_j < N\delta_j\sqrt{v} \\ \frac{a_j}{N\sqrt{v}} & \text{when } N\delta_j\sqrt{v} \leq a_j \leq N\sqrt{v}. \\ 1, & \text{when } a_j > N\sqrt{v} \end{cases} \quad (47)$$

The optimal solution is

$$p_j = \max(\min(\frac{a_j}{N\sqrt{v}}, 1), \delta_j) \quad (48)$$

Let $\tau = \frac{1}{N\sqrt{v}}$, $p_j = \max(\min(\tau a_j, 1), \delta_j)$. Thus, the desired value of τ can be obtained by finding the root of the equation $g(\tau) = \sum_{j=1}^N \max(\min(\tau a_j, 1), \delta_j) - N\zeta$.

In case $\delta_j = 0$, (48) is:

$$p_j = \min(\tau a_j, 1) \quad (49)$$

where τ is the root of the equation $g(\tau) = \sum_{j=1}^N \min(\tau a_j, 1) - \zeta N$,

APPENDIX B MAE OPTIMIZATION

A. Results for Proposed Optimization Problem

$$\min_{\lambda_O, \lambda_R, \lambda_B} (\Phi_O MAE_O + \Phi_R MAE_R + \Phi_B MAE_B) \quad (50a)$$

$$\text{subject to } \lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \zeta N, \quad (50b)$$

$$0 \leq \lambda_O \leq 1. \quad (50c)$$

$$0 \leq \lambda_R \leq 1. \quad (50d)$$

$$0 \leq \lambda_B \leq 1. \quad (50e)$$

The Lagrangian of (50) is

$$\begin{aligned} \mathcal{L}(\lambda_O, \lambda_R, \lambda_B, \gamma, \eta_O, \eta_R, \eta_B, \nu_O, \nu_R, \nu_B) &= \Phi_O a_O + \Phi_O \frac{b_O}{c_O + \lambda_O} \\ &+ \Phi_R a_R + \Phi_R \frac{b_R}{c_R + \lambda_R} + \lambda_R, \\ &+ \Phi_B a_B + \Phi_B \frac{b_B}{c_B + \lambda_B} \\ &+ \gamma (\lambda_O N_O + \lambda_R N_R + \lambda_B N_B - \zeta N) \\ &- \eta_O \lambda_O - \eta_R \lambda_R - \eta_B \lambda_B \\ &+ \nu_O (\lambda_O - 1) + \nu_R (\lambda_R - 1) + \nu_B (\lambda_B - 1). \end{aligned} \quad (51)$$

where $\lambda_O, \lambda_R, \lambda_B$ are the primal variables, $\eta_O, \eta_R, \eta_B, \nu_O, \nu_R, \nu_B$ are the Lagrange multipliers associated with the constraints $0 \leq \lambda_O \leq 1$, $0 \leq \lambda_R \leq 1$, $0 \leq \lambda_B \leq 1$, and $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \zeta N$, respectively.

The first order optimality conditions imply the following:

• *Stationarity:*

$$- \nabla_{\lambda_O} \mathcal{L} = 0. \text{ That is, } -\frac{\Phi_O b_O}{(c_O + \lambda_O)^2} + \gamma N_O - \eta_O + \nu_O = 0.$$

$$- \nabla_{\lambda_R} \mathcal{L} = 0. \text{ That is, } -\frac{\Phi_R b_R}{(c_R + \lambda_R)^2} + \gamma N_R - \eta_R + \nu_R = 0.$$

$$- \nabla_{\lambda_B} \mathcal{L} = 0. \text{ That is, } -\frac{\Phi_B b_B}{(c_B + \lambda_B)^2} + \gamma N_B - \eta_B + \nu_B = 0.$$

• *Primal feasibility:*

$$\begin{aligned} - \lambda_O N_O + \lambda_R N_R + \lambda_B N_B &= \zeta N \\ - \lambda_O &\leq 1, \text{ and } \lambda_O \geq 0 \\ - \lambda_R &\leq 1, \text{ and } \lambda_R \geq 0 \\ - \lambda_B &\leq 1, \text{ and } \lambda_B \geq 0 \end{aligned}$$

• *Dual feasibility:* $\eta_O > 0, \eta_R > 0, \eta_B > 0, \nu_O > 0, \nu_R > 0, \nu_B > 0$

• *Complementary slackness:*

$$\begin{aligned} - \eta_O \lambda_O &= 0, \nu_O (\lambda_O - 1) = 0 \\ - \eta_R \lambda_R &= 0, \nu_R (\lambda_R - 1) = 0 \\ - \eta_B \lambda_B &= 0, \nu_B (\lambda_B - 1) = 0 \end{aligned}$$

$\eta_O \lambda_O = 0$ implies that one of the following cases always holds: $\eta_O = 0$ or $\lambda_O = 1$.

• Case 1: $\eta_O = 0$:

- If $\nu_O = 0$, then $\lambda_O \leq 0$. Substituting $\eta_O = \nu_O = 0$ in the stationarity condition: $\lambda_O = \sqrt{\frac{\Phi_O b_O}{\gamma N_O}} - c_O$, because $0 \leq \lambda_O \leq 1$, we must have $\frac{\Phi_O b_O}{N_O (c_O + 1)^2} \leq$

$$\gamma \leq \frac{\Phi_O b_O}{N_O c_O^2}.$$

- If $\lambda_O = 1$, then $\nu_O > 0$. Substituting $\eta_O = 0$ and $\lambda_O = 1$ into the stationarity condition, $\nu_O = \frac{\Phi_O b_O}{(c_O + 1)^2} - \gamma N_O$. Because $\nu_O > 0$, we have $\gamma < \frac{\Phi_O b_O}{N_O (c_O + 1)^2}$.

• Case 2: $\lambda_O = 0$:

- $\nu_O = 0$. Substituting $\lambda_O = 0$ and $\nu_O = 0$ into the stationarity condition $\eta_O = \gamma N_O - \frac{\Phi_O b_O}{c_O^2}$. Because $\eta_O > 0$, we have $\gamma > \frac{\Phi_O b_O}{N_O c_O^2}$.

1) *Combining These Two Cases, We Obtain:*

$$\lambda_O = \begin{cases} 0, & \text{when } \gamma > \frac{\Phi_O b_O}{N_O c_O^2} \\ \sqrt{\frac{\Phi_O b_O}{\gamma N_O}} - c_O & \text{when } \frac{\Phi_O b_O}{N_O (c_O + 1)^2} \leq \gamma \leq \frac{\Phi_O b_O}{N_O c_O^2} \\ 1, & \text{when } \gamma < \frac{\Phi_O b_O}{N_O (c_O + 1)^2} \end{cases} \quad (53)$$

The optimal solution is

$$\lambda_O = \max(\min(\sqrt{\frac{\Phi_O b_O}{\gamma N_O}} - c_O, 1), 0) \quad (54)$$

$\eta_R \lambda_R = 0$ implies that one of the following cases always holds: $\eta_R = 0$ or $\lambda_R = 1$.

• Case 1: $\eta_R = 0$:

- If $\nu_R = 0$, then $\lambda_R \leq 0$. Substituting $\eta_R = \nu_R = 0$ in the stationarity condition: $\lambda_R = \sqrt{\frac{\Phi_R b_R}{\gamma N_R}} - c_R$, because $0 \leq \lambda_R \leq 1$, we must have $\frac{\Phi_R b_R}{N_R (c_R + 1)^2} \leq \gamma \leq \frac{\Phi_R b_R}{N_R c_R^2}$.

- If $\lambda_R = 1$, then $\nu_R > 0$. Substituting $\eta_R = 0$ and $\lambda_R = 1$ into the stationarity condition, $\nu_R = \frac{\Phi_R b_R}{(c_R + 1)^2} - \gamma N_R$. Because $\nu_R > 0$, we have $\gamma < \frac{\Phi_R b_R}{N_R (c_R + 1)^2}$.

- Case 2: $\lambda_R = 0$:
 - $v_R = 0$. Substituting $\lambda_R = 0$ and $v_R = 0$ into the stationarity condition $\eta_R = \gamma N_R - \frac{\Phi_R b_R}{c_R^2}$. Because $\eta_R > 0$, we have $\gamma > \frac{\Phi_R b_R}{N_R c_R^2}$.

2) *Combining These Two Cases, We Obtain:*

$$\lambda_R = \begin{cases} 0, & \text{when } \gamma > \frac{\Phi_R b_R}{N_R c_R^2} \\ \sqrt{\frac{\Phi_R b_R}{\gamma N_R}} - c_R & \text{when } \frac{\Phi_R b_R}{N_R (c_R + 1)^2} \\ \leq \gamma \leq \frac{\Phi_R b_R}{N_R c_R^2} \\ 1, & \text{when } \gamma < \frac{\Phi_R b_R}{N_R (c_R + 1)^2} \end{cases} \quad (55)$$

The optimal solution is

$$\lambda_R = \max(\min(\sqrt{\frac{\Phi_R b_R}{\gamma N_R}} - c_R, 1), 0) \quad (56)$$

$\eta_B \lambda_B = \mathbf{0}$ implies that one of the following cases always holds: $\eta_B = \mathbf{0}$ or $\lambda_B = 1$.

- Case 1: $\eta_B = 0$:
 - If $v_B = 0$, then $\lambda_B \leq 0$. Substituting $\eta_B = v_B = 0$ in the stationarity condition: $\lambda_B = \sqrt{\frac{\Phi_B b_B}{\gamma N_B}} - c_B$, because $0 \leq \lambda_B \leq 1$, we must have $\frac{\Phi_B b_B}{N_B (c_B + 1)^2} \leq \gamma \leq \frac{\Phi_B b_B}{N_B c_B^2}$.
 - If $\lambda_B = 1$, then $v_B > 0$. Substituting $\eta_B = 0$ and $\lambda_B = 1$ into the stationarity condition, $v_B = \frac{\Phi_B b_B}{(c_B + 1)^2} - \gamma N_B$. Because $v_B > 0$, we have $\gamma < \frac{\Phi_B b_B}{N_B (c_B + 1)^2}$.
- Case 2: $\lambda_B = 0$:
 - $v_B = 0$. Substituting $\lambda_B = 0$ and $v_B = 0$ into the stationarity condition $\eta_B = \gamma N_B - \frac{\Phi_B b_B}{c_B^2}$. Because $\eta_B > 0$, we have $\gamma > \frac{\Phi_B b_B}{N_B c_B^2}$.

3) *Combining These Two Cases, We Obtain:*

$$\lambda_B = \begin{cases} 0, & \text{when } \gamma > \frac{\Phi_B b_B}{N_B c_B^2} \\ \sqrt{\frac{\Phi_B b_B}{\gamma N_B}} - c_B & \text{when } \frac{\Phi_B b_B}{N_B (c_B + 1)^2} \\ \leq \gamma \leq \frac{\Phi_B b_B}{N_B c_B^2} \\ 1, & \text{when } \gamma < \frac{\Phi_B b_B}{N_B (c_B + 1)^2} \end{cases} \quad (57)$$

The optimal solution is

$$\lambda_B = \max(\min(\sqrt{\frac{\Phi_B b_B}{\gamma N_B}} - c_B, 1), 0) \quad (58)$$

γ is chosen based on the primal feasibility $\lambda_O N_O + \lambda_R N_R + \lambda_B N_B = \xi N$

REFERENCES

- [1] J. Kellett, R. Barreto, A. V. D. Hengel, and N. Vogiatzis, "How might autonomous vehicles impact the city? The case of commuting to central adelaide," *Urban Policy Res.*, vol. 37, no. 4, pp. 442–457, Oct. 2019.
- [2] T. Litman, "Autonomous vehicle implementation predictions: Implications for transport planning," Victoria Transp. Policy Inst., Victoria, BC, Canada, Tech. Rep. 15-3326, 2020.
- [3] M. A. Schreurs and S. D. Steuwer, "Autonomous driving-political, legal, social, and sustainability dimensions," in *Autonomes Fahren*. Berlin, Germany: Springer, 2015, pp. 151–173.
- [4] K. Bengler, K. Dietmayer, B. Farber, M. Maurer, C. Stiller, and H. Winner, "Three decades of driver assistance systems: Review and future perspectives," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 4, pp. 6–22, winter 2014.
- [5] J. Choi, D. Chun, H. Kim, and H.-J. Lee, "Gaussian YOLOv3: An accurate and fast object detector using localization uncertainty for autonomous driving," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 502–511.
- [6] M. Liang, B. Yang, Y. Chen, R. Hu, and R. Urtasun, "Multi-task multi-sensor fusion for 3D object detection," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 7337–7345.
- [7] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231–1237, Sep. 2013.
- [8] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3354–3361.
- [9] B. Prabhakaran, Y.-G. Jiang, H. Kalva, and S.-F. Chang, "Editorial IEEE transactions on multimedia special section on video analytics: Challenges, algorithms, and applications," *IEEE Trans. Multimedia*, vol. 20, no. 5, p. 1037, May 2018.
- [10] VelodyneLiDARInc. (2017). *Velodyne LiDAR HDL-64ES3*. Accessed: Dec. 12, 2020. [Online]. Available: <http://velodynelidar.com/hdl-64e>
- [11] X. T. Nguyen, V. L. Dinh, H.-J. Lee, and H. Kim, "A high-definition LIDAR system based on two-mirror deflection scanners," *IEEE Sensors J.*, vol. 18, no. 2, pp. 559–568, Jan. 2018.
- [12] X. T. Nguyen, K.-T. Nguyen, H.-J. Lee, and H. Kim, "ROI-based LiDAR sampling algorithm in on-road environment for autonomous driving," *IEEE Access*, vol. 7, pp. 90243–90253, 2019.
- [13] K. Lenac, A. Kitanov, R. Cupec, and I. Petrović, "Fast planar surface 3D SLAM using LIDAR," *Robot. Auto. Syst.*, vol. 92, pp. 197–220, Jun. 2017.
- [14] J. Castorena and C. D. Creusere, "Random impulsive scan for LIDAR sampling," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 365–368.
- [15] D. Droschel and S. Behnke, "Efficient continuous-time SLAM for 3D LIDAR-based online mapping," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1–9.
- [16] X. Chen *et al.*, "OverlapNet: Loop closing for LiDAR-based SLAM," in *Proc. Robot., Sci. Syst. Robot., Sci. Syst. Found.*, Jul. 2020, pp. 12–16.
- [17] S. Hawe, M. Kleinsteuber, and K. Diepold, "Dense disparity maps from sparse disparity measurements," in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2126–2133.
- [18] S. Schwartz, A. Wong, and D. A. Clausi, "Multi-scale saliency-guided compressive sensing approach to efficient robotic laser range measurements," in *Proc. 9th Conf. Comput. Robot Vis.*, May 2012, pp. 1–8.
- [19] L.-K. Liu, S. H. Chan, and T. Q. Nguyen, "Depth reconstruction from sparse samples: Representation, algorithm, and sampling," *IEEE Trans. Image Process.*, vol. 24, no. 6, pp. 1983–1996, Jun. 2015.
- [20] S. Schwartz, A. Wong, and D. A. Clausi, "Saliency-guided compressive sensing approach to efficient laser range measurement," *J. Vis. Commun. Image Represent.*, vol. 24, no. 2, pp. 160–170, Feb. 2013.
- [21] L.-K. Liu and T. Q. Nguyen, "A framework for depth video reconstruction from a subset of samples and its applications," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4873–4887, Oct. 2016.
- [22] W. Bao, W.-S. Lai, X. Zhang, Z. Gao, and M.-H. Yang, "MEMC-Net: Motion estimation and motion compensation driven neural network for video interpolation and enhancement," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Sep. 17, 2019, doi: [10.1109/TPAMI.2019.2941941](https://doi.org/10.1109/TPAMI.2019.2941941).
- [23] D. Bolya, C. Zhou, F. Xiao, and Y. J. Lee, "YOLACT: Real-time instance segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 9156–9165.
- [24] C. K. Chua, C. H. Wong, and W. Y. Yeong, *Standards, Quality Control, and Measurement Sciences in 3D Printing and Additive Manufacturing*. New York, NY, USA: Academic, 2017.
- [25] X. T. Nguyen, H. Kim, and H.-J. Lee, "An efficient sampling algorithm with a K-NN expanding operator for depth data acquisition in a LiDAR system," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 12, pp. 4700–4714, Dec. 2020.

- [26] H. Kim, A. No, and H.-J. Lee, "SPIHT algorithm with adaptive selection of compression ratio depending on DWT coefficients," *IEEE Trans. Multimedia*, vol. 20, no. 12, pp. 3200–3211, Dec. 2018.
- [27] VelodyneLiDARInc. (2017). *Velodyne LiDAR HDL-32E*. Accessed: Dec. Jan. 12, 2020. [Online]. Available: <https://velodynelidar.com/products/hdl-32e>
- [28] VelodyneLiDARIncPuckTM. (2017). *PUCK*. Accessed: Dec. Jan. 12, 2020. [Online]. Available: <https://velodynelidar.com/products/puck>



Quan-Dung Pham received the B.S. degree in electronics and telecommunications engineering from Vietnam National University, Ho Chi Minh City, Vietnam, and the University of Technology, Ho Chi Minh City, in 2017, and the M.S. degree in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2020. He is currently the Applied Scientist of VinBrain, Vietnam. His current research interests include digital signal processing, applied machine learning, deep neural networks, and SoC design for

low-complexity multimedia applications. He was a recipient of the Korean Government Scholarship Program (KGSP) in 2017.



Xuan Truong Nguyen received the B.S. degree in electrical engineering from the Hanoi University of Science and Technology, Hanoi, Vietnam, in 2011, and the M.S. and Ph.D. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2015 and 2019, respectively. He is currently a Postdoctoral Fellow with the BK21+ of Electrical and Computer Engineering Department, Seoul National University. His research interests include the algorithm and SoC design for

low-complexity computer vision and multimedia applications.



Khac-Thai Nguyen received the B.S. degree in electronics and telecommunications from the Hanoi University of Science and Technology, Vietnam, in 2014, and the M.S. degree in electrical engineering and computer science from Seoul National University, South Korea, in 2018. He is currently a System Software Engineer with Google.



Hyun Kim (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from Seoul National University, Seoul, South Korea, in 2009, 2011, and 2015, respectively. From 2015 to 2018, he was a BK Assistant Professor with the BK21 Creative Research Engineer Development for IT, Seoul National University. In 2018, he joined the Department of Electrical and Information Engineering, Seoul National University of Science and Technology, Seoul, where he is currently an

Assistant Professor. His research interests include algorithm, computer architecture, memory, and SoC design for low-complexity multimedia applications and deep neural networks.



Hyuk-Jae Lee (Member, IEEE) received the B.S. and M.S. degrees in electronics engineering from Seoul National University, South Korea, in 1987 and 1989, respectively, and the Ph.D. degree in electrical and computer engineering from Purdue University, West Lafayette, IN, in 1996. From 1996 to 1998, he was with the Faculty of the Department of Computer Science, Louisiana Tech University, Ruston, LS. From 1998 to 2001, he was a Senior Component Design Engineer with the Server and Workstation

Chipset Division, Intel Corporation, Hillsboro, OR. In 2001, he joined the School of Electrical Engineering and Computer Science, Seoul National University, where he is currently a Professor. He is also the Founder of Mamurian Design, Inc., a fabless SoC design house for multimedia applications. His research interests include computer architecture and SoC design for multimedia applications.