Uncertainty-Aware Interactive LiDAR Sampling for Deep Depth Completion

Kensuke Taguchi[1] Shogo Morita[1] Yusuke Hayashi[1] Wataru Imaeda[1] Hironobu Fujiyoshi[2]
[1]KYOCERA Corporation [2]Chubu University
[1]{kensuke.taguchi.xm,shogo.morita.nf,yuusuke.hayashi.zs,wataru.imaeda.xm}@kyocera.jp
[2]fujiyoshi@isc.chubu.ac.jp

Abstract

Programmable scan LiDAR is able to measure arbitrary areas and is expected to be used in various applications. In this paper, we study a LiDAR sampling strategy for deep depth completion of a programmable scan LiDAR with an RGB camera. General data sampling strategies include adaptive approaches such as active learning, in which candidate data are assessed through a task model for data selection and then the selected data pool is updated sequentially. Although it is an effective approach, the adaptive approach requires many iterations involving the inference process to assess the candidate data, which is not suitable for LiDAR systems. Therefore, we propose a novel interactive LiDAR sampling method without each inference process. Our key insights are that we assess sampling candidates by depth estimation uncertainty and virtually update the uncertainty by an approximation of the candidate assessment. This enables us to add interactivity to the model state without requiring each inference process. We demonstrate the effectiveness of our method on the KITTI dataset and the generalization performance on the NYU-Depth-v2 dataset in comparison with a conventional adaptive LiDAR sampling method, and we find superior results in the depth completion task. We also show ablation studies to analyze our approach.

1. Introduction

Dense and high frame rate depth information is important for understanding scenes as required by advanced applications such as autonomous driving. Light detection and ranging (LiDAR) is a key sensor for 3D sensing that can be used outdoors, has a long measurement range, and has high measurement accuracy.

The scanning area cannot be freely adjusted in conventional LiDAR because the scan position is mechanically determined. However, with the evolution of LiDAR technology, solid-state LiDAR[18, 24, 16] that is cheap and capable of programmable scanning has appeared owing to advances in phased array and micro electro mechanical systems (MEMS) technology. Programmable scan is a method that can measure an arbitrarily configured area. In recent years, there have been several studies on adaptive sampling for depth completion with programmable scan LiDAR, such as end-to-end, in which the sampling algorithms and upstream task are optimized together[3], and non-learning-based methods[26]. In addition, the adaptive sampling method[9], which can sequentially sample LiDAR depth points based on assessment through the model output, is proposed. This is effective because LiDAR points can be assessed in model output that contains complementarity with the already selected data. However, it is very costly in terms of operation speed owing to the need to perform many iterations involving the inference process to assess the candidates because it is necessary to select comparatively many samples in LiDAR sampling.

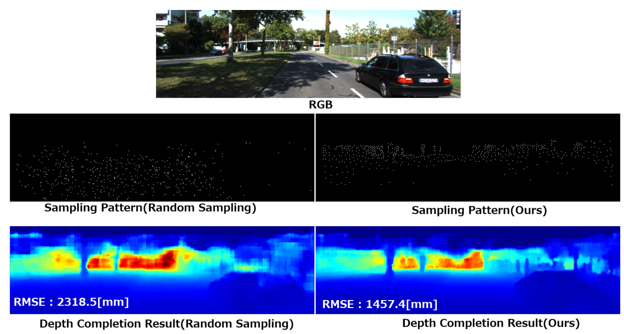

Figure 1: Sampling results by naive sampling (random) and our method. Top to bottom: RGB image, sampling patterns, estimated dense depth map

In this paper, we propose an uncertainty-aware interactive LiDAR sampling method that does not require each inference process to query the candidate assessment. Our key insights are that we assess the sampling candidate by

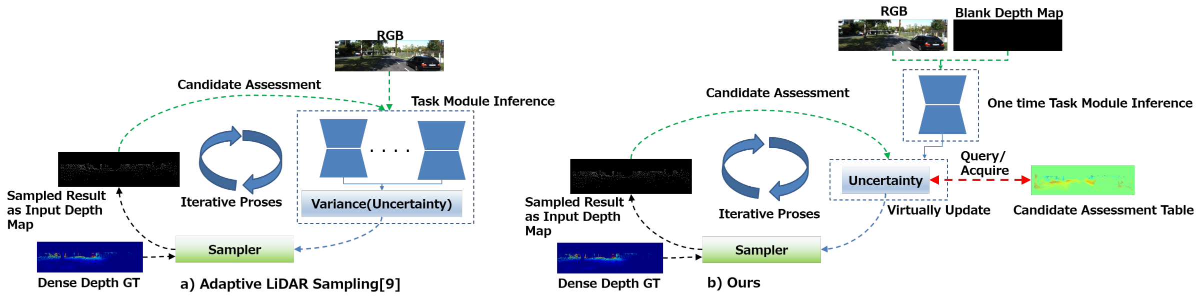

Figure 2: Comparison between adaptive LiDAR sampling[9] and our method. a) Adaptive LiDAR sampling performs the task model inference process for every sampling process to update the uncertainty of the task model output. b) Our method enables interactive sampling without performing the inference process by virtually updating the uncertainty using a candidate assessment table.

the depth estimation uncertainty, and we design a feed-forward module which can directly predict the assessment of depth estimation uncertainty with respect to each candidate. This module is called the candidate assessment module. We also refer to its output as the candidate assessment table, which is an efficient approximation of the candidate assessment. Since the candidate assessment table can help to virtually update the depth estimation uncertainty with respect to selected LiDAR points, it enables interactive LiDAR sampling that considers already selected data without an inference process. Fig 1 shows examples of sampling results from using our method, which is based on the uncertainty of task output, like active learning, and can create a high-quality sampled depth map for estimating dense depth. Fig 2 shows a comparison between adaptive LiDAR sampling[9] and our method. Our contributions are as follows:

- We propose a novel LiDAR sampling method for deep depth completion on programmable scan LiDAR. We assess sampling candidates based on depth estimation uncertainty and virtually update the uncertainty. This enables us to add interactivity to the depth completion module state.

- We demonstrate the effectiveness of our method on the KITTI dataset[8] and the generalization performance on NYU-Depth-v2[22] in comparison with the conventional adaptive LiDAR sampling method and various ablation studies.

2. Related Work

2.1. Depth Completion

Since high-end LiDAR that can output dense depth is very expensive, depth completion that generates dense depth from sparse depth and possibly image data is a suitable solution for obtaining dense depth with low spec LiDAR. Traditional depth completion methods usually employ handcrafted features or specific kernels to predict missing values. The performance of these methods is often limited. Recently, deep learning-based approaches have been proposed that have achieved promising results. Uhrig et al.[23] proposed sparsity invariant CNNs to extract better representation from sparse input only. Ma et al.[14] proposed feeding the concatenation of the sparse depth and the color image into an encoder-decoder deep network. A similar approach was also applied to the self-supervised setting[13]. Instead of using a CNN, Cheng et al.[5]used a recurrent convolution to estimate the affinity matrix for depth completion. Zhang et al.[29] proposed estimating surface normal and solving for depth via a global optimization. Based on Zhang et al. [29], Qiu et al.[19] integrated surface normal and image-based estimation results on an attention-based method to perform on outdoor scenes targeting longer distance ranges. Since the LiDAR sampling problem assumes the performance of cheap solid-state LiDAR, it deals with lower density problems compared to these depth completion works and it is very challenging.

2.2. Spatial Sampling Strategy

Spatial sampling strategies have been studied in various fields. Furthest point sampling[6] is a technique often used in 3D reconstruction problems. In the field of computer graphics, Poisson disc sampling[21, 4] has been introduced, and is also applied in the field of LiDAR sampling[27]. Recently, adaptive sampling algorithms for depth completion have been introduced. Wolff et al.[26] proposed an image-driven sampling and reconstruction strategy based on dividing the image into approximately piecewise segments. Bergman et al.[3] introduced a deep neural network for end-to-end sampling and reconstruction.

The sampling problem is also related to active learning in terms of data selection. The core idea of active learning is that the system being trained actively selects samples according to a policy and queries their labels surveyed in [20]. Uncertainty sampling is one of the most popular

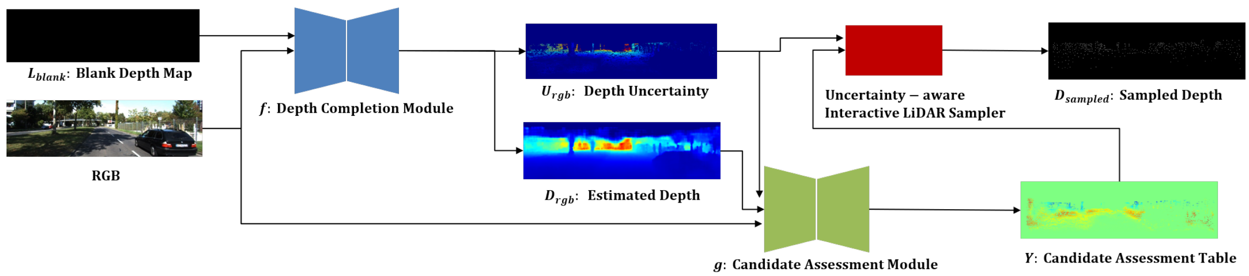

Figure 3: Overall schematic of our method. First, we obtain the depth estimation and its uncertainty from rgb. Next, the candidate assessment module predicts the candidate assessment table from rgb and the first step results. In the end, the uncertainty-aware interactive LiDAR sampler performs depth sampling and virtually updates the depth estimation uncertainty using the candidate assessment table.

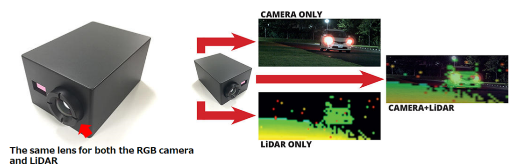

Figure 4: Our prototype of a programmable scan LiDAR, which uses a MEMS mirror with an RGB camera. Because both devices use the same lens, the LiDAR and the RGB camera have identical optical axes.

strategies in active learning for determining which samples to request labels for[2, 25, 10, 21]. As an evolution of these approaches, Yoo et al.[28] proposed a loss prediction module that directly evaluates loss of sampling candidate data for the target task. Gofer et al.[9] applied an active learning approach to the LiDAR sampling for depth completion that showed good performance. Inspired by this work, our method can perform LiDAR sampling without inference processing every time to query the candidate assessment.

3. Proposed Method

In this section, we introduce our adaptive LiDAR sampling method. Our method consists of a depth completion module that estimates dense depth from sparse depth obtained by LiDAR sampling and RGB images, a candidate assessment module that directly predicts the assessment with respect to candidate LiDAR points, and an uncertainty-aware interactive LiDAR sampler. Fig 3 shows an overall schematic of our method. More details of each component are shown below.

3.1. Problem Formulation

In this paper, LiDAR sampling is used to create a sparse depth map including an arbitrary amount of depth information as the input for deep depth completion on programmable scan LiDAR. As the source for the LiDAR sampling, we use the dense depth map as ground truth of deep depth completion. Note that the representation of depth is a 2D depth map that projects LIDAR depth points onto the RGB image plane so that the coordinate system of the depth map can be shared with the RGB image. In other words, we can find the coordinates for sampling LiDAR depth points using the RGB image coordinate system.

In a hardware implementation, the sampling coordinates can be used as the scan position by assuming that RGB and LiDAR are aligned on the optical axis. As an example, Fig 4 shows our prototype of a programmable scan LiDAR, which uses a MEMS mirror and an RGB camera that are aligned on the same optical axis.

3.2. Our Approach

Our method is an interactive approach which uses sequential data selection based on the uncertainty of depth completion module output such as active learning. Unlike general active learning, the LiDAR sampling aims to generate a sparse depth map as the input for the depth completion module and does not require a re-training process of the module. However, the iterative inference process to query a model output is big issue in terms of operation speed because LiDAR sampling problems require real-time performance and many sampling points.

Our solution is to approximate all the candidate assessment processes using a deep neural network(DNN) designed pixel-wise task to avoid iterative inference process. The module that realizes this approximation is called the candidate assessment module. This module can directly predict the candidate assessment by training on the already measured assessment value. The key idea is how we achieve this approximation. We explain more details in the next section.

3.3. Candidate Assessment Module

3.3.1 Definition of Candidate Assessment

The candidate assessment module aims to assess all the sampling candidates in one process. First, we explain the definition of the sampling candidate assessment. We take the depth completion module as f, input RGB image as rgb, a single LiDAR point projected onto the RGB image plane coordinates(x, y) as $L(x, y)$ and zero-filled depth map as L_{blank}. Note that $L(x, y)$ is zero-filled except at (x, y) and has the same resolution as rgb for input to f. Depth estimation and its uncertainty with and without $L(x, y)$ as input to f are described as follows

$$U_{L(x,y)}, D_{L(x,y)} = f(rgb, L(x, y)) \tag{1}$$

$$U_{rgb}, D_{rgb} = f(rgb, L_{blank}) \tag{2}$$

The difference in the uncertainty with respect to $L(x, y)$ is

$$A_{L(x,y)} = U_{rgb} - U_{L(x,y)} \tag{3}$$

This is very simple and directly express the assessment for $L(x, y)$ based on the uncertainty of f. Note that $A_{L(x,y)}$ is the same resolution as $U_{L(x,y)}$ and $D_{L(x,y)}$. Since f estimates the dense depth from sparse depth by exploiting the local and global context mainly about rgb, each LiDAR point affects a wide area around the sampling coordinates. Thus, the entire size of uncertainty map is required to express the effect of a LiDAR point naively. This is not realistic because a large amount of memory is required to express $A_{L(x,y)}$ for all LiDAR candidates. Therefore, to handle it more efficiently, we encode $A_{L(x,y)}$ as the average value in a local patch of size H, W centered on (x, y) as

$$\hat{Y}(x, y) = \frac{1}{HW} \sum_{i=x-\lfloor W/2 \rfloor}^{x+\lfloor W/2 \rfloor} \sum_{j=y-\lfloor H/2 \rfloor}^{y+\lfloor H/2 \rfloor} A_{L(x,y)}(i, j) \tag{4}$$

By using this encoding, $A_{L(x,y)}$ can be expressed by one element included local area influence, and the assessment of all candidate coordinates can then be expressed in a 2D table. We call this the candidate assessment table and this is the output format of the candidate assessment module as g . We show an example of this in Fig 5. We create the candidate assessment table measured by already trained f as the ground truth for training the candidate assessment module.

3.3.2 Module Design

We design the candidate assessment module as a basic encoder-decoder model for predicting the candidate assessment table as a pixel-wise task. An important design factor is deciding what to input to this module. The candidate assessment table is defined based on the output of f in Eq. 4. Thus, we select rgb , U_{rgb} and D_{rgb} as inputs to the module for considering information about f. Indeed, U_{rgb} is a bias factor of the candidate assessment table, and it is effective to estimate it.

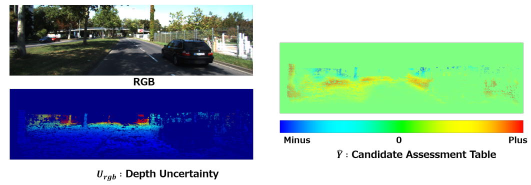

Figure 5: Example of depth estimation uncertainty and candidate assessment table. (Top-Left) RGB image. (Second row-left) Depth estimation uncertainty with only RGB. (Right) Candidate assessment table defined by equation4

Algorithm 1 Pseudo-code of the uncertainty-aware interactive LiDAR sampler

Input: Sampling Budget N, Predicted Candidate Assessment Table Y, Dense Depth Map $\hat{D}$, Depth Uncertainty U_{rgb} as Eq(2)
Initialize:
  $U_i = U_{rgb}$
for $i \leftarrow 1, N$ **do**
  $x, y = \arg\max(U_i)$ ▷ This is an acquisition function
  $U_{i+1} = U_i + \mathrm{decode}(\mathrm{Y(x, y)})$
  $D_{\text{sampled}}(x, y) = \hat{D}(x, y)$
end for
Output: Depth map generated by sampler D_{sampled}

The loss function of the candidate assessment module is the simple mean squared error as

$$Loss(Y, \hat{Y}) = (Y - \hat{Y})^2 \tag{5}$$

where Y is the prediction of the candidate assessment module and $\hat{Y}$ is the ground truth as given by Eq. 4. Since $\hat{Y}$ is created from the depth estimation uncertainty of f, the candidate assessment module can be considered as a form of the depth estimation uncertainty distillation.

3.4. Uncertainty-aware Interactive LiDAR Sampler

In this section, we describe the uncertainty-aware interactive LiDAR sampler. The uncertainty-aware interactive LiDAR sampler performs LiDAR sampling based on the depth estimation uncertainty and then virtually updates it with respect to LiDAR sampling sequentially. The sampling process is written in algorithm 1.

In the data sampling problem, the informativeness of new points is assessed by an acquisition function. Our acquisition function selects the largest area in U_i, which is the virtually updated uncertainty map initialized by U_{rgb}. The value from candidate assessment table decoded to the patch size H, W for considering local context is applied to virtually update the uncertainty map. Our sampling policy reduces the depth estimation uncertainty by sampling the LiDAR points in high uncertainty areas. The uncertainty can

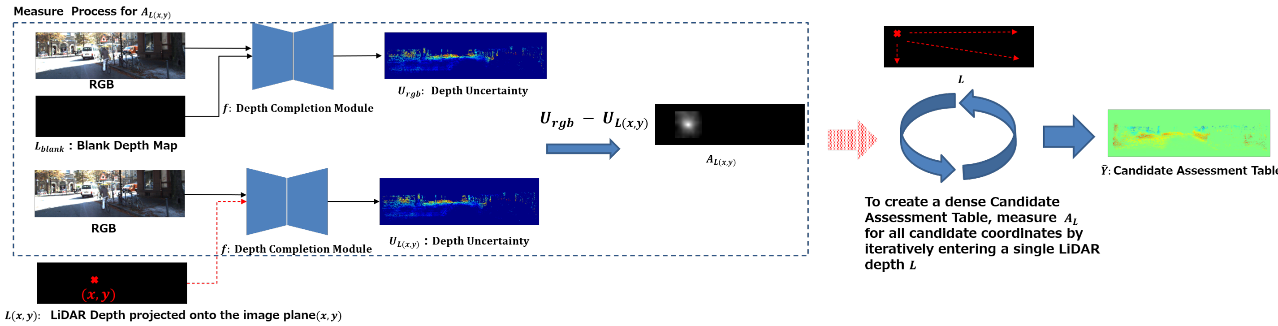

Figure 6: Creation process of our candidate assessment table.

be reduced by sampling the LiDAR depth because LiDAR depth is more precise and helpful for dense depth estimation. Like active learning, reducing the uncertainty of the output is effective for improving the task performance.

The key point of our sampler is interactivity by sampling LiDAR while querying the state of the model as expected uncertainty. Our sampler enables candidate data to be assessed while considering already sampled data through the uncertainty and this is useful for efficient sampling over a scene. Moreover, since we can determine the uncertainty with respect to the sampling sequentially, it is possible to set a sampling number according to the difficulty of depth estimation for the scene. For example, in easy scenes for depth estimation, the uncertainty is low enough with even a small number of points and we can determine this during the sampling process, so the depth estimation performance can be achieved with a small number of LiDAR points.

3.5. Modeling Uncertainty

As a method of uncertainty modeling, we train the depth completion module to infer the mean $\mu(d)$ and variance $\sigma^2(d)$ parameters for the distribution of the output by log-likelihood maximization[15, 17]. The loss function of the depth completion module, which can model the uncertainty, is expressed as

$$Loss_{depth}(d, \hat{d}) = \frac{|\mu(d) - \hat{d}|}{\sigma^2(d)} + \frac{1}{2} log\sigma(d) \qquad (6)$$

where d is the predicted depth of the depth completion module and $\hat{d}$ is the dense depth map as ground truth for the depth completion module. We treat $\sigma^2(d)$ as the uncertainty. This method can capture the uncertainty caused by noise inherent in the observations, which is called Aleatoric uncertainty[11]. In our case, the region where uncertainty is large implies that depth estimation is difficult due to less information, for example, in the far region. The uncertainty can then be reduced by sampling LiDAR points to add effective information about this kind of region.

3.6. Training Method

In our method, the modules that need to be trained are the depth completion module and candidate assessment module. Because we need the candidate assessment table, which is measured using the trained depth completion module to train the candidate assessment module, our training strategy is two-step. First, we train the depth completion module using Eq. 6 taking the sparse depth map as the input randomly sampled from dense depth ground truth. Note that amount of sampled depth in this sparse depth map is the target LiDAR sampling number. Next, we create the candidate assessment table $\hat{Y}$ using Eq. 4 with the already trained f. Fig 6 shows the creation process of dense candidate assessment table by iteratively entering a single LiDAR depth L to f for all candidate coordinates. We then train the candidate assessment module by Eq. 5 with $\hat{Y}$. Note that although our training strategy is two-step, there is no need to split the training data.

4. Experiments

To show the effectiveness of the proposed method, we evaluated our LiDAR sampling method on a depth completion task in comparison with the conventional LiDAR sampling method and ablation study. Since our method is aimed at applications such as autonomous driving that require high-frame rate operation, we mainly perform the evaluation on the KITTI dataset composed of outdoor in-vehicle scenes. We also evaluate NYU-Depth-v2 dataset and our prototype for indoor scenes to confirm the generalization.

4.1. Datasets

KITTI:

We first evaluate our approach on the KITTI dataset[8, 23], which includes RGB images and LiDAR depth transformed between camera and LiDAR in outdoor scenes. The ground truth is created from multi-frame LiDAR and a stereo camera to remove outliers in the

Table 1: Performance of depth completion on KITTI validation set. We compare our method with other LiDAR sampling and depth completion methods. Our method outperforms the others in terms of depth completion performance. Lower is better.

Method	Sample	Number of module executions	RMSE[mm]	MAE[mm]
Sparse-to-Dense[13]	Random	1	2290.5	957.7
NConv-CNN-L2[7]	Random	1	2379.6	829.0
Deep-Adaptive-LiDAR[3]	Poisson-disc	2	1767.7	613.6
Deep-Adaptive-LiDAR[3]	Adaptive	3	1753.1	642.0
Adaptive-LiDAR-Sampling(phase1)[9]	Adaptive	4	1895.7	823.3
Adaptive-LiDAR-Sampling(phase4)[9]	Adaptive	13	1500.1	623.5
Ours	Adaptive	3	**1463.5**	**592.5**

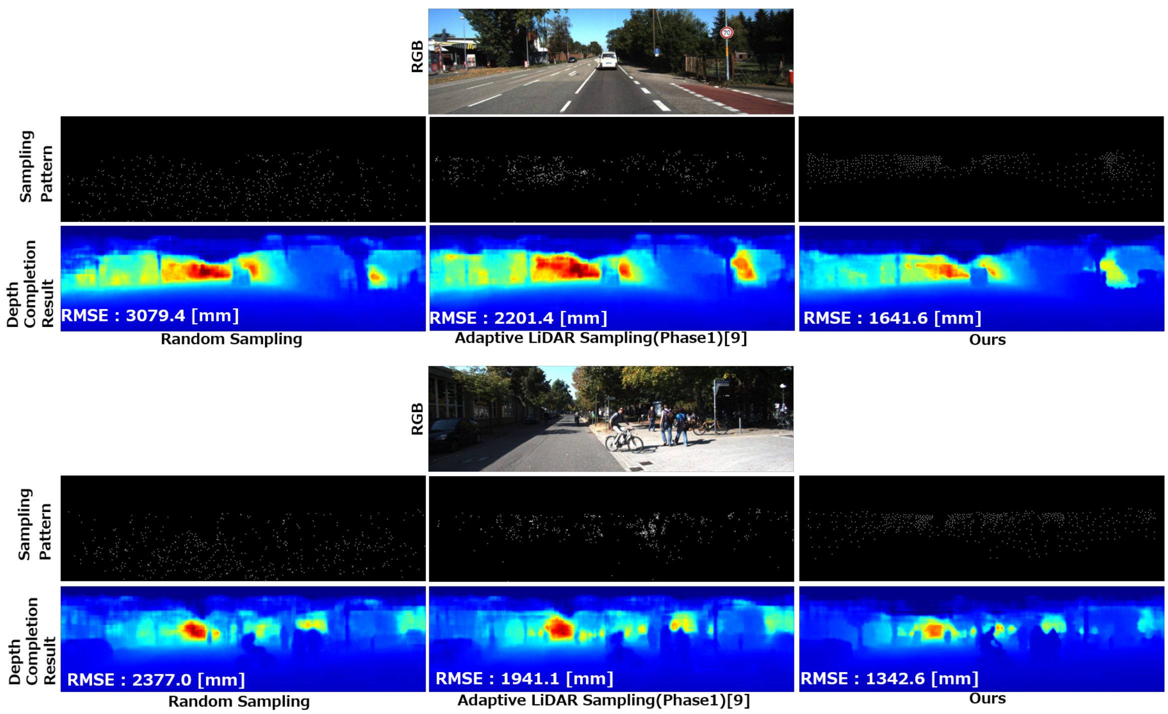

Figure 7: Qualitative results using the KITTI dataset.

laser scans by comparing the scanned depth to results from stereo reconstruction. It is not 100 percent dense and only about 30 of points in the ground truth depth map have valid depth values. The dataset consists of 85,898 training data, 1,000 selected validation data, and 1,000 test data without ground truth comparison with state-of-the-art methods. Our evaluation setting for the LiDAR sampling number is 512 which is the same as [9, 3].

NYU-Depth-v2:

The NYU-Depth-v2 dataset[22] provides RGB images and dense depth maps captured by flash LiDAR in indoor scenes. The dataset contains 120K training samples and 654 testing samples. We train our method on a 50K subset and missing depth values are filled in using the inpainting method[12], which is the same setting as [1, 3]. Note that we have down-sampled the original images from the original resolution of 640×480 to half-resolution and center-cropped to dimensions of 320×256 following [14, 29]. Our evaluation setting of the LiDAR sampling number is 200.

4.2. Implementation Details

The model architecture of the depth completion module and training parameters is based on Ma et al.[13]. We set the model layer size to 18 and add another final layer to model the depth estimation uncertainty. The model architecture of

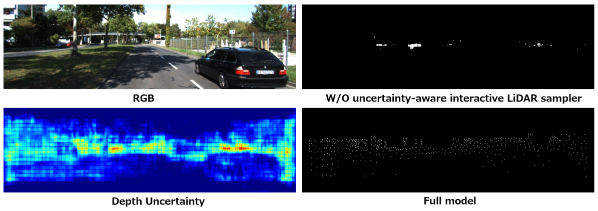

Figure 8: Comparison between the sampling pattern difference with and without the interactive LiDAR sampler.

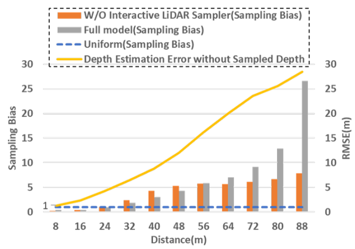

Figure 9: Comparison of sampling bias versus distance. The number of sampled points computed by our interactive method tends to increase with distance, as does the depth estimation error without sampled depth.

the candidate assessment module uses the same settings as the depth completion module. Note that to handle the three elements U_{rgb}, D_{rgb}, and rgb as inputs, we add an initial convolution layer. The patch size H, W with the candidate assessment table are set to 30 respectively.

4.3. Comparison with the State of the Art

We first compare our method against several state-of-the-art methods which treat very sparse depths as input on the KITTI validation dataset. Table 1 shows the results. It can be observed that our method achieves much better results than related works. Our method has better adaptability to a scene by performing point by point sampling shown in algorithm 1. Although the performance difference between our method and adaptive LiDAR sampling(Phase4)[9] is small, adaptive LiDAR sampling requires multiple inference processes with multiple modules for effectiveness, which is not suitable for operation speed in real scenarios. In addition, the comparison with adaptive LiDAR sampling(Phase1) which is few iterations, our method has great advantages. Fig 7 shows the qualitative results from using our method. Note that phases represent the number of iterations, and each phase execute at least 3 modules for computing the variance.

4.4. Ablation study

Importance of Interactive Approach: To understand the impact of each module on the final output, we conduct ablation studies on the KITTI dataset. We compare the depth estimation performance with and without the uncertainty-aware interactive LiDAR sampler. When the uncertainty-aware interactive LiDAR sampler is not used, the LiDAR sampling candidates are automatically selected in order from the largest regions on the depth estimation uncertainty. This means that we can consider each candidate only in terms of depth estimation uncertainty and loss of interactivity. Table 2 shows the results. Performance without the uncertainty-aware interactive LiDAR sampler obviously degrades, which shows the large impact of the interactive approach. Fig 8 shows an example of qualitative results. It seems unbalanced because the sampling pattern without the uncertainty-aware interactive LiDAR sampler is too dense in areas of high uncertainty. Fig 9 shows the sampling bias versus the distance, which indicates that our sampler provides good bias toward areas that require LiDAR sampling, such as distant areas with large depth estimation errors.

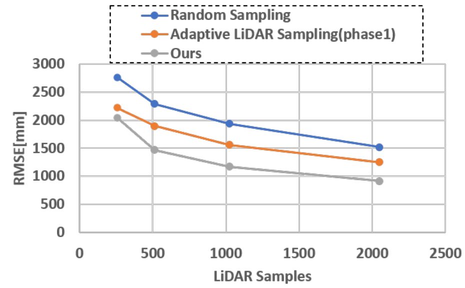

Figure 10: Depth completion performance for different sparsities.

Against Input Sparsity: Since the LiDAR sampling problem mainly aims to create a depth map suitable for the target task based on a very sparse setting, we evaluate the performance of our method with various densities of sampled depth. We set the LiDAR sampling numbers to 256, 512, 1024, and 2048, which correspond to 0.06%, 0.12%, 0.24%, and 0.48% of the density, repsectively. Fig 10 shows the depth estimation performance for various input densities of the KITTI dataset. Since our method exploits the uncertainty to sample more important areas, it works well under both very sparse and more dense settings.

4.5. Generalization Capability

NYU-Depth-v2: We also evaluate our method on NYU-Depth-v2 to investigate the generalization capability under different kind of datasets. NYU-Depth-v2 was created in an indoor environment and is different from the KITTI dataset in terms of sensor range and sensor position variety. NYU-Depth-v2 is captured by flash LiDAR which has more short range than KITTI and various sensor positions. Table 3 show the quantitative results. It shows that our method is superior to the others and has the generality to extend to different datasets. We show examples of the results in Fig 11.

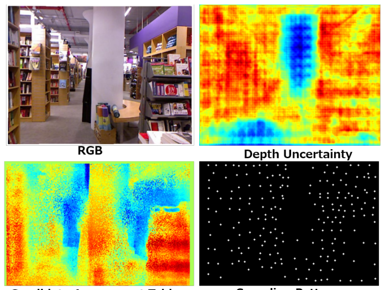

Figure 11: Qualitative results on NYU-Depth-v2 dataset.

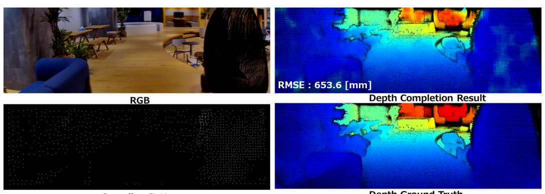

Figure 12: Qualitative results from using our prototype.

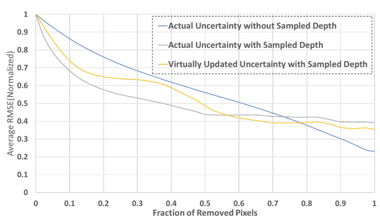

Figure 13: Sparsification plots of both actual and virtually updated uncertainties monotonically decrease against the removed pixels with the uncertainty on KITTI validation set. This indicates good modeling of the uncertainty.

Our prototype: We show some qualitative results from using our LiDAR sampling method on the prototype shown in Figure 4. Since this prototype cannot yet perform programmable scans sequentially, we perform a virtual scan as a sampling problem, using a dense depth that integrates multiple LiDAR frames. Figure 12 shows examples of the results in our lab. The density of the sampled depth map is 0.10%, and this shows good performance at the same level as other datasets.

5. Discussion

We analyze our key idea which is to virtually update the uncertainty by the uncertainty-aware interactive LiDAR sampler with the candidate assessment table. Thus, we compare the difference between the virtually updated uncertainty and the actual uncertainty of the depth estimation by sparsification plots. Figure 13 shows the result on the KITTI dataset. There are small gaps between these uncertainties. In addition, since our method assumes that the effect of the uncertainty with respect to each LiDAR point is independent of the other LiDAR points, the candidate assessment table is measured by inputting only one LiDAR point to each candidate coordinate. However, multiple LiDAR points are in fact sampled as the sparse depth map and input to the depth completion module at the same time. Therefore, even if the predicted candidate assessment table is perfect, it is thought that the behavior of the uncertainty differs from the assumption that the other LiDAR points occur as errors. However, since the error is very small as shown in the evaluation results, we think that our assumption and approach are reasonable.

Table 2: Ablation study for an interactive approach on the KITTI validation set. Our full model can perform interactive sampling, thereby achieving the best performance.

Module	RMSE[mm]
W/O uncertainty-aware interactive LiDAR sampler	2780.3
Full model	**1463.5**

Table 3: Performance of depth completion on NYU-Depth-v2 dataset. Note that the RMSE metric is in meters.

Method	Sample	RMSE[m]
Bilateral	Random	0.479
Sparse-to-Dense[13]	Random	0.230
NConv-CNN-L2[7]	Random	0.209
Deep-Adaptive-LiDAR[3]	Poisson-disc	0.207
Deep-Adaptive-LiDAR[3]	Adaptive	0.193
Super-Pixel-Sampler[26]	Adaptive	0.211
Ours	Adaptive	**0.179**

6. Conclusion

In this paper, we study LiDAR sampling for deep depth completion of programmable scanning LiDAR with RGB camera. Our method performs sampling candidate assessment as depth completion uncertainty and uncertainty-aware interactive sampling for depth completion module state. We design a candidate assessment module which can assess all the sampling candidate assessments in a single process. This enables us to add interactivity to the model state without each inference process. We conduct comparison with the state of the art and various ablation studies to show the effectiveness of our LiDAR sampling.

References

[1] Ibraheem Alhashim and Peter Wonka. High quality monocular depth estimation via transfer learning. *arXiv preprint arXiv:1812.11941*, 2018.

[2] William H Beluch, Tim Genewein, Andreas Nürnberger, and Jan M Köhler. The power of ensembles for active learning in image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[3] Alexander W Bergman, David B Lindell, and Wetzstein Gordon. Deep adaptive lidar: End-to-end optimization of sampling and depth completion at low sampling rates. In *IEEE International Conference on Computational Photography (ICCP)*, 2020.

[4] Robert Bridson. Fast poisson disk sampling in arbitrary dimensions. *SIGGRAPH sketches*, 10:1278780—-1278807, 2007.

[5] Xinjing Cheng, Peng Wang, and Ruigang Yang. Depth estimation via affinity learned with convolutional spatial propagation network. In *the European Conference on Computer Vision (ECCV)*, 2018.

[6] Yuval Eldar, Michael Lindenbaum, Moshe Porat, and Yehoshua Y Zeevi. The farthest point strategy for progressive image sampling. *IEEE Transactions on Image Processing*, 6:1305–1315, 1997.

[7] Abdelrahman Eldesokey, Michael Felsberg, and Fahad Shahbaz Khan. Confidence propagation through cnns for guided sparse depth regression. *IEEE transactions on pattern analysis and machine intelligence*, 42:2423–2436, 2019.

[8] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[9] Eyal Gofer, Shachar Praisler, and Guy Gilboa. Adaptive lidar sampling and depth completion using ensemble variance. *IEEE Transactions on Image Processing*, 30:8900–8912, 2021.

[10] Ajay J Joshi, Fatih Porikli, and Nikolaos Papanikolopoulos. Multi-class active learning for image classification. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2009.

[11] Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in Neural Information Processing Systems (NIPS)*, 2017.

[12] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. In *ACM SIGGRAPH 2004 Papers*. 2004.

[13] Fangchang Ma, Guilherme Venturelli Cavalheiro, and Karaman Sertac. Self-supervised sparse-to-dense: Self-supervised depth completion from lidar and monocular camera. In *IEEE international conference on robotics and automation (ICRA)*, 2019.

[14] Fangchang Ma and Karaman Sertac. Sparse-to-dense: Depth prediction from sparse depth samples and a single image. In *IEEE international conference on robotics and automation (ICRA)*, 2018.

[15] David A Nix and Andreas S Weigend. Estimating the mean and variance of the target probability distribution. In *IEEE international conference on neural networks (ICNN'94)*, 1994.

[16] Francesco Pittaluga, Zaid Tasneem, Justin Folden, Brevin Tilmon, Ayan Chakrabarti, and Sanjeev J Koppal. Towards a mems-based adaptive lidar. In *International Conference on 3D Vision (3DV)*, 2020.

[17] Matteo Poggi, Filippo Aleotti, Fabio Tosi, and Stefano Mattoccia. On the uncertainty of self-supervised monocular depth estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.

[18] Christopher V Poulton, Ami Yaacobi, David B Cole, Matthew J Byrd, Manan Raval, Diedrik Vermeulen, and Michael R Watts. Coherent solid-state lidar with silicon photonic optical phased arrays. *Optics letters*, 42:4091–4094, 2017.

[19] J Qiu, Z Cui, Y Zhang, X Zhang, S Liu, and M Zeng, B.and Pollefeys. Deeplidar: Deep surface normal guided depth prediction for outdoor scene from sparse lidar data and single color image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[20] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.

[21] H Sebastian Seung, Manfred Opper, and Haim Sompolinsky. Query by committee. In *the fifth annual workshop on Computational learning theory*, 1992.

[22] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. Indoor segmentation and support inference from rgbd images. In *European conference on computer vision (ECCV)*, 2012.

[23] N Uhrig, J.and Schneider, L Schneider, U. Franke, T Brox, and A Geiger. Sparsity invariant cnns. In *international conference on 3D Vision (3DV)*, 2017.

[24] Dingkang Wang, Conor Watkins, Sanjeev Koppal, Mengyuan Li, Yingtao Ding, and Huikai Xie. A compact omnidirectional laser scanner based on an electrothermal tripod mems mirror for lidar. In *20th International Conference on Solid-State Sensors, Actuators and Microsystems & Eurosensors XXXIII (TRANSDUCERS & EUROSENSORS XXXIII)*, 2019.

[25] Keze Wang, Dongyu Zhang, Ya Li, Ruimao Zhang, and Liang Lin. Cost-effective active learning for deep image classification. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(12):2591–2600, 2016.

[26] Adam Wolff, Shachar Praisler, Ilya Tcenov, and Guy Gilboa. Super-pixel sampler: a data-driven approach for depth sampling and reconstruction. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2020.

[27] Xin Xiong, Haipeng Xiong, Ke Xian, Chen Zhao, Zhiguo Cao, and Xin Li. Sparse-to-dense depth completion revisited: Sampling strategy and graph construction. In *the European Conference on Computer Vision (ECCV)*, 2020.

[28] Donggeun Yoo and In So Kweon. Learning loss for active learning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[29] Yinda Zhang and Thomas Funkhouser. Deep depth completion of a single rgb-d image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.