Multi-Echo LiDAR for 3D Object Detection

Yunze Man[1], **Xinshuo Weng**[1], **Prasanna Kumar Sivakumar**[2], **Matthew O'Toole**[1], **Kris Kitani**[1]

[1]Carnegie Mellon University, [2]DENSO

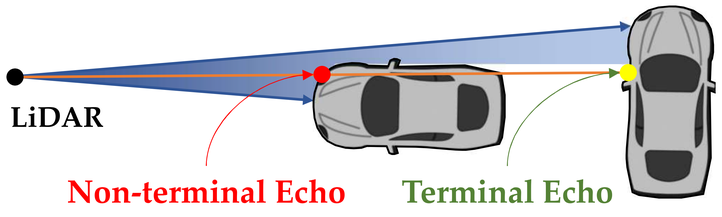
{yman, xinshuow, mpotoole, kkitani}@cs.cmu.edu, prasanna.kumar.sivakumar@na.denso.com

Background & Motivation

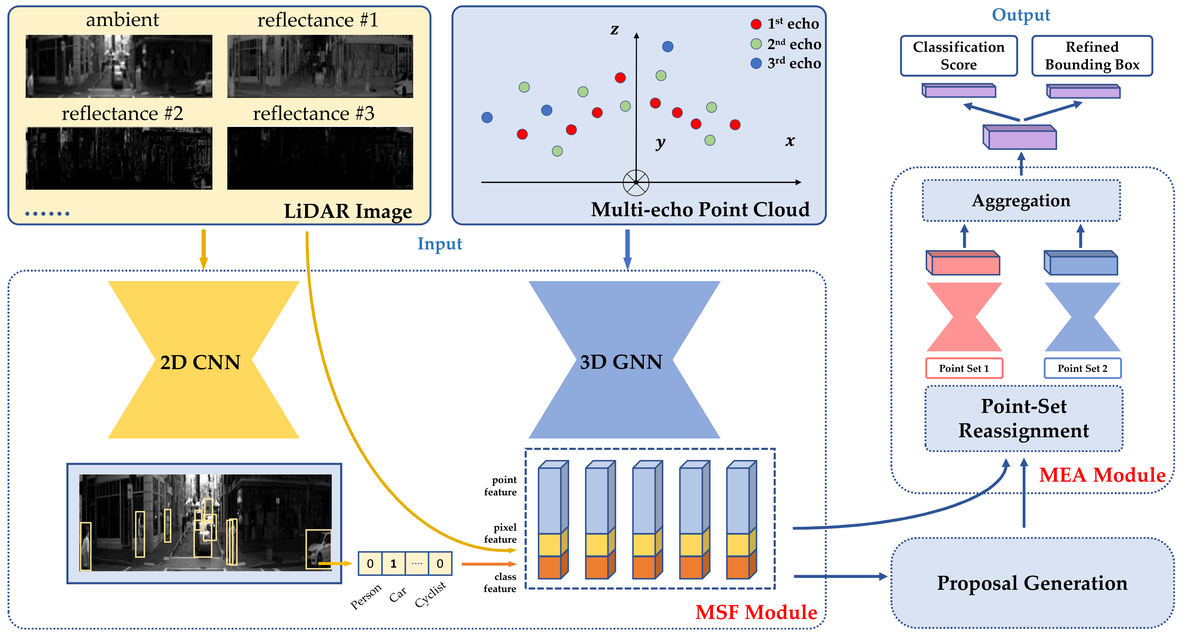
- LiDAR point clouds are becoming the most widely used signals in 3D perception tasks.
- However, LiDAR sensors can be used to obtain a wider range of measurement – more than single-echo 3D point clouds
- In fact, we found that a LiDAR sensor can provide these three modalities:
  - **MEPC**: Multi-Echo Point Cloud
  - **AS**: Ambient Intensity Signal
  - **RS**: Reflectance Intensity Signal

Modalities	What SOTA are using	What LiDAR can provide
SEPC: Single-Echo Point Cloud	√	√
MEPC: Multi-Echo Point Cloud		√
RS: Reflectance Intensity Signal	√	√
AS: Ambient Intensity Signal		√

- **Goal:** Leverage a wider spectrum of measurements provided by LiDAR sensor to improve 3D detection.
- **Challenges:**
  - How to learn from multiple echoes of points.
  - How to fuse AS, RS with MEPC.
- Therefore, we propose **MSLiD**: *Multi-Signal LiDAR Detector*, the first model that leverage multi-echo point cloud (MEPC), ambient signal (AS) and reflectance signal (RS) together from LiDAR sensor to help 3D object detection.
- We propose Multi-signal Fusion (**MSF**) module to merge AS, RS with point cloud, Multi-echo Aggregation (**MEA**) module to combine feature from MEPC.



MSLiD Pipeline & Results



Inputs: We rearrage the signals into two representations: A 2D image with multiple channels in range image space; And 3D point clouds in the classic 3D Cartesian space.
MSF Module: We adopt a two-stream structure to extract and fuse features from each input representation, i.e. a 2D detector branch and a 3D feature learning branch.
MEA Module: For each of the proposal, we first group the points inside it into two sets based on their ray-casting orders. Then we learn a feature vector from each group of point by a PointNet. And the two vectors are then aggregated together by concatenation.

Quantitative Results

Method		Car - IoU = 0.7			Car - IoU = 0.5			Person - IoU = 0.5			Person - IoU = 0.25		
		Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard	Easy	Moderate	Hard
SECOND	1st echo	67.9	37.1	27.3	79.9	57.0	56.2	42.1	25.2	19.8	54.8	35.5	23.9
	full echo	75.0	42.9	30.6	86.8	65.5	65.1	47.4	29.9	20.3	58.5	39.2	25.6
PointRCNN	1st echo	66.9	37.8	28.1	80.2	57.7	52.1	45.2	28.9	20.0	57.0	38.7	25.2
	full echo	73.6	41.9	28.9	85.0	65.6	62.7	51.6	31.4	20.2	61.2	40.7	25.9
3DSSD	1st echo	64.1	36.7	27.0	77.4	55.9	52.4	45.7	28.8	20.2	57.5	38.5	23.8
	full echo	72.4	40.6	28.1	83.9	65.1	63.9	51.9	30.7	19.9	60.8	40.5	26.2
SASSD	1st echo	68.8	37.9	28.6	81.2	58.3	56.8	42.5	26.7	18.2	54.7	36.6	23.2
	full echo	76.1	43.2	29.8	87.2	66.0	63.7	46.9	28.5	19.4	58.1	38.6	23.7
PV-RCNN	1st echo	69.1	38.3	28.4	81.7	59.1	57.4	44.9	28.2	19.7	56.3	38.2	24.9
	full echo	76.9	44.1	**31.2**	88.1	67.2	65.3	52.7	30.9	20.8	62.0	41.2	26.2
MSLiD		**79.5**	**45.3**	30.7	**89.7**	**68.1**	**65.3**	**57.5**	**34.2**	**21.5**	**66.5**	**43.2**	**27.7**
Improvement		*+2.6*	*+1.2*	*-0.5*	*+1.6*	*+0.9*	*+0.0*	*+4.8*	*+2.8*	*+0.7*	*+4.5*	*+2.0*	*+1.5*

Ablation Study

Input signals:

Point Cloud (PC)	AS	RS	Overall AP
SE			50.5
ME			53.6
ME	✓		53.7
ME		✓	55.1
ME	✓	✓	**55.5**

Conclusion: Our proposed method makes improvement on detection results by properly leveraging multiple modalities of input LiDAR signals.

MSF and MEA Modules:

Method	MSF		MEA	Overall AP
	class ft.	pixel ft.		
baseline				50.5
MSF only	✓	✓		52.1
MEA only			✓	53.6
w/o class ft.		✓	✓	54.0
w/o pixel ft.	✓		✓	55.3
MSLiD	✓	✓	✓	**55.5**

Conclusion: The MSF and MEA modules together improve the absolute AP by 5.0 from the baseline.

Grouping and Aggregation:

Aggregation Scheme			Grouping		Overall AP
Max.	Mean.	Concat.	Echo	Reassign	
✓			✓		51.1
	✓		✓		53.9
		✓	✓		54.5
		✓		✓	**55.5**

Conclusion: The point set reassignment module and the aggregation by concatenation help better learn the RoI features from MEPC.

IEEE International Conference on Computer Vision (ICCV), 2021.