PLA-LiDAR :

Physical Laser Attacks against LiDAR-based 3D Object Detection in Autonomous Vehicle

Zizhi Jin[1], Xiaoyu Ji[1], Yushi Cheng[1,2], Bo Yang[1], Chen Yan[1], Wenyuan Xu[1]

[1]Zhejiang University, [2]Tsinghua University,

智能系统安全实验室 UBIQUITOUS SYSTEM SECURITY LAB.

浙江大學 ZHEJIANG UNIVERSITY

清華大學 Tsinghua University

LiDAR

- LiDAR (Light Detection and Ranging) is widely used for perception.
- Correct LiDAR perception provides the foundation for safety in self driving.

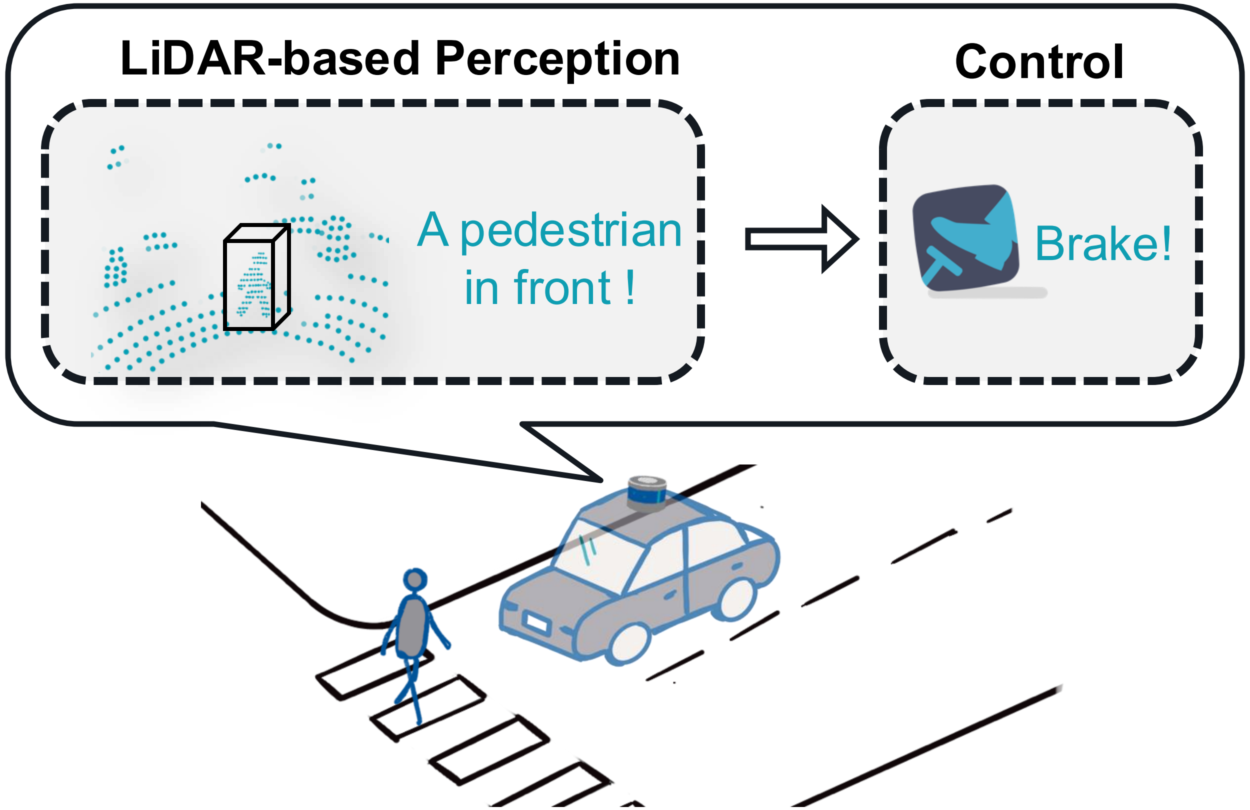
Self-driving Car

CVIS

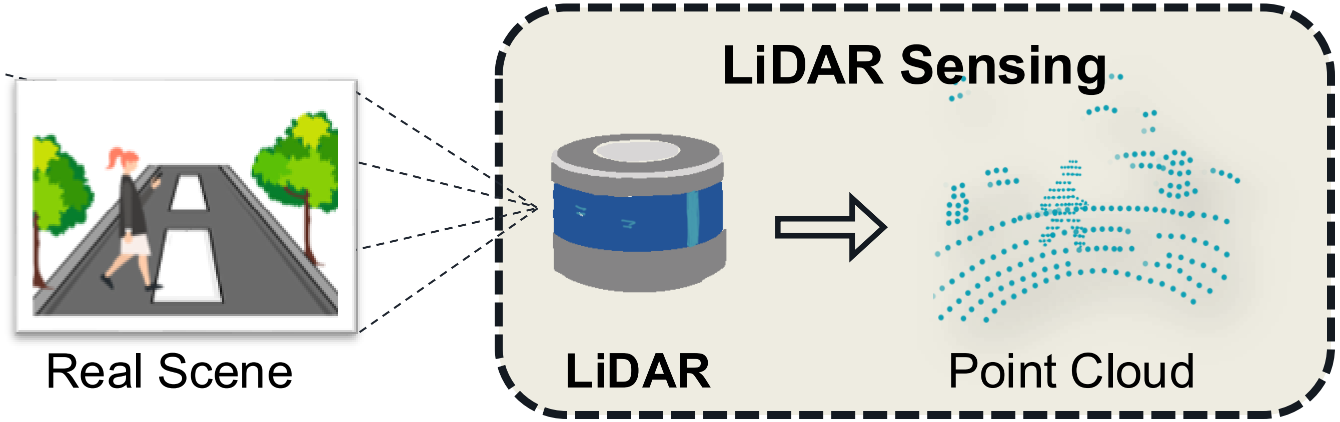
Robots

Drones

Source: www.velodynelidar.com

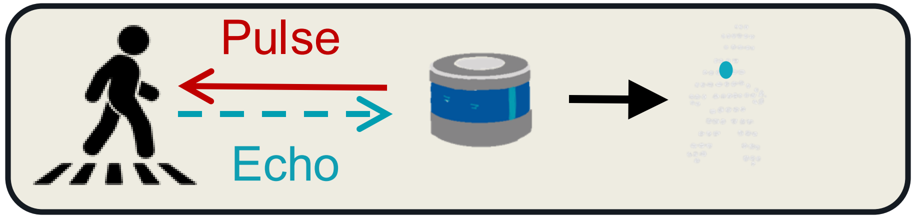


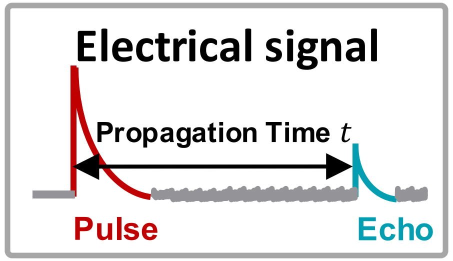
How Does LiDAR work?



LiDAR generates point cloud by laser ranging and laser beam steering.

Laser Ranging

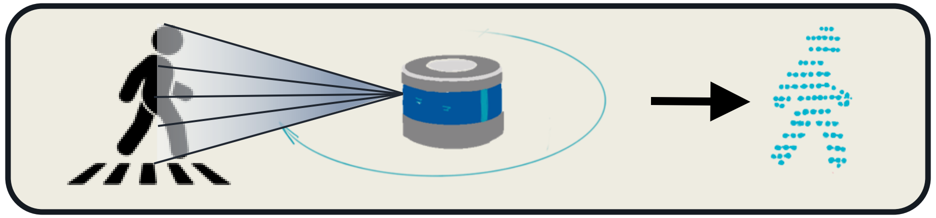




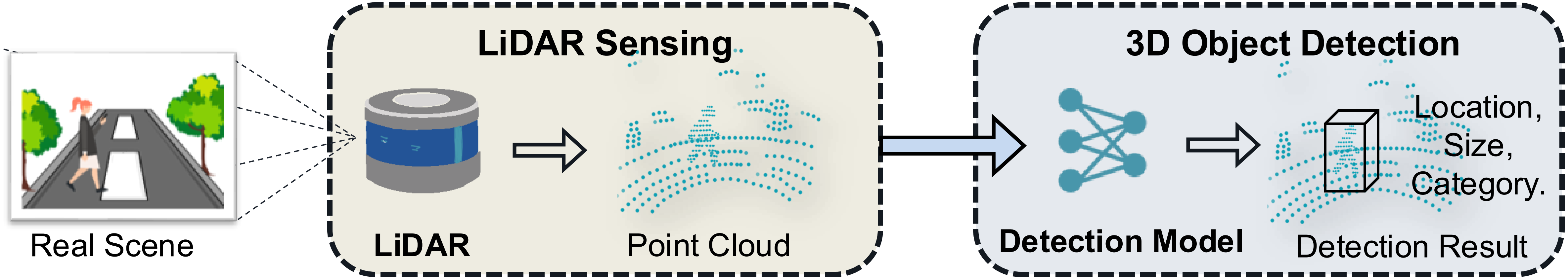
- **Direction** (θ, φ)
- **Distance**

  $d = 0.5 * t * c$

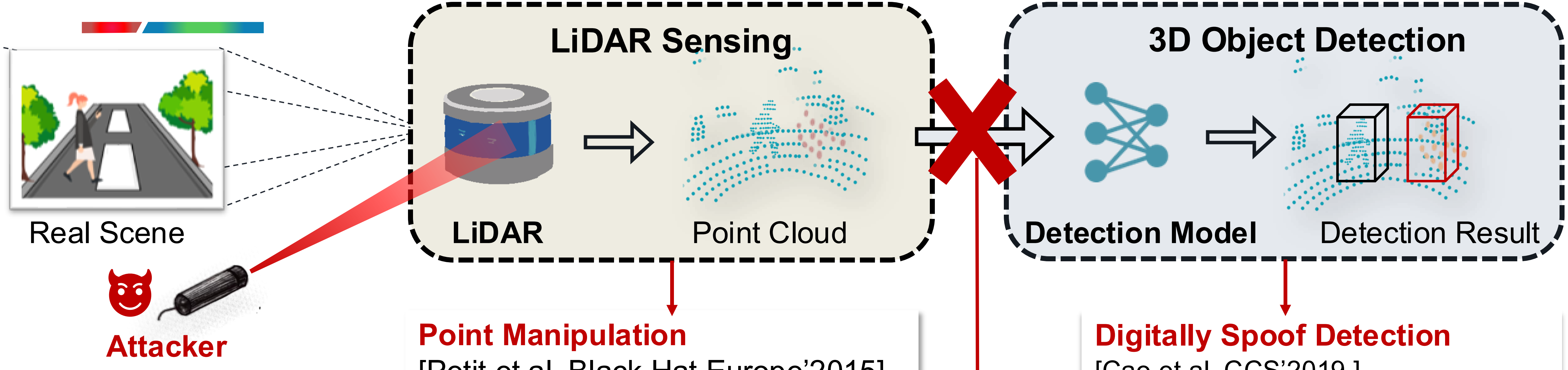
Laser Beam Steering



How Does LiDAR-based Perception work?



Security of LiDAR-based Perception

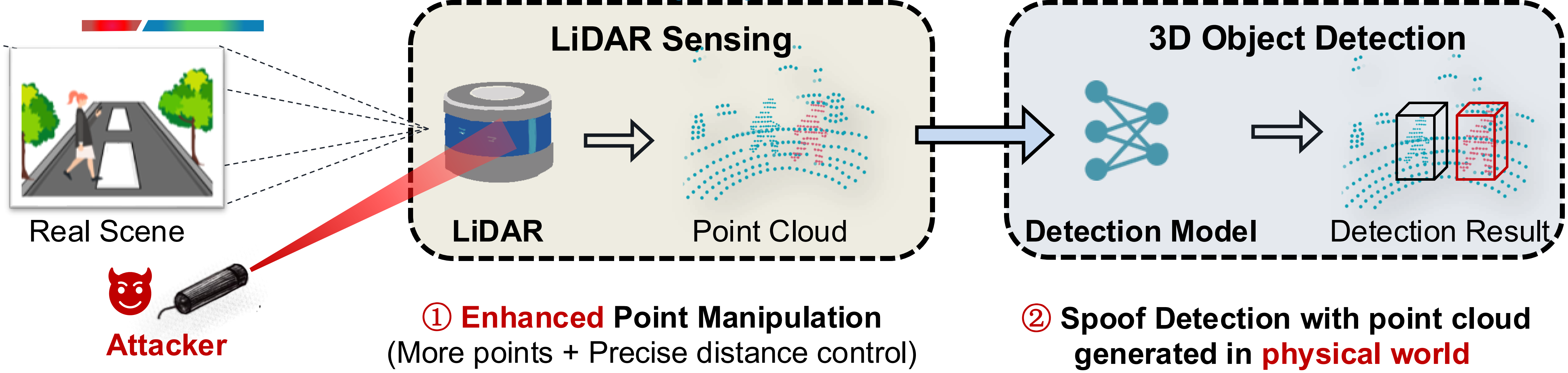


Point Manipulation
[Petit et al. Black Hat Europe'2015]
[Shin et al. CHES'2017]
…

Digitally Spoof Detection
[Cao et al. CCS'2019]
[Sun et al., USENIX'2020]
…

Is it possible to fool 3D object detection using **lasers** in the **physical world**?

PLA-LiDAR can achieve physical laser attacks!



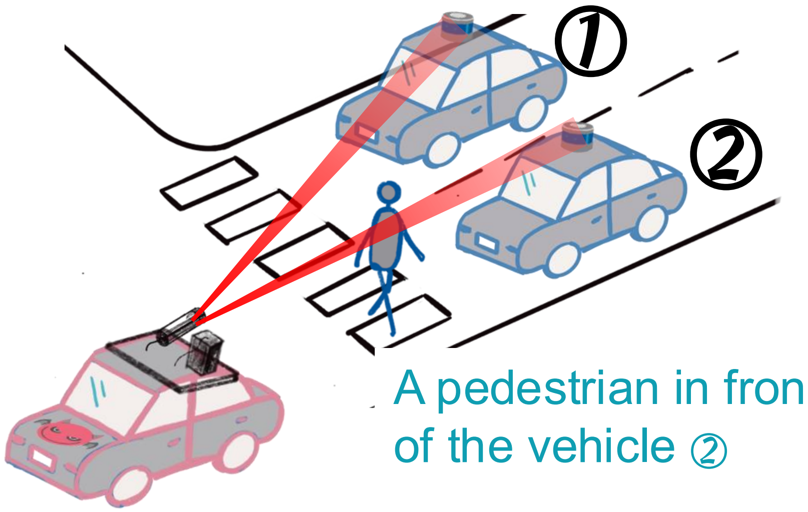
Contributions of PLA-LiDAR:

1. Enhanced Point Manipulation Capability
2. Physical-world Laser Attacks against 3D Object Detection

Threat Model and Attack Goal

Attack Scenarios

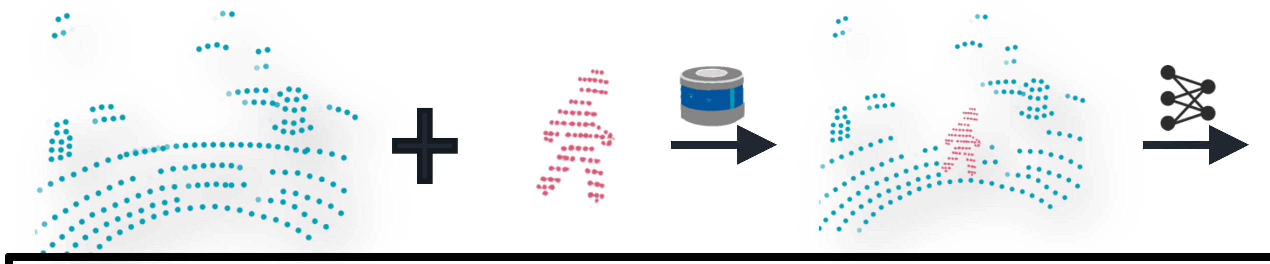
No obstacles in front of the vehicle ①



A pedestrian in front of the vehicle ②

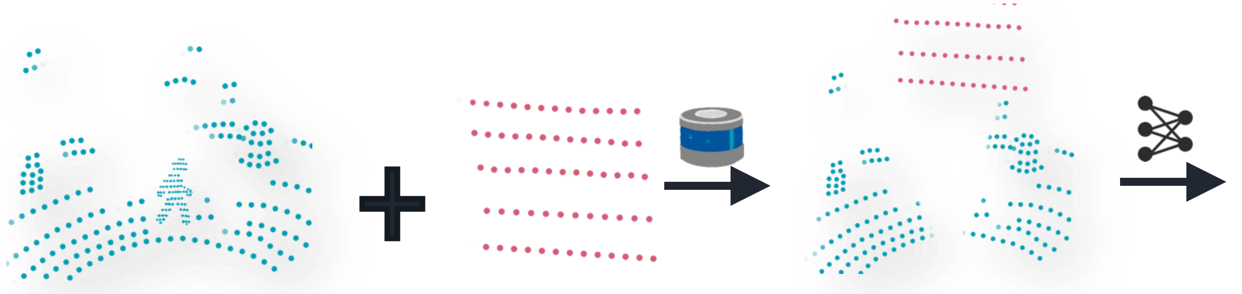
Attacker
Road-side

Benign Point Cloud	Injected Points	Captured Point Cloud	Detection Result



A Pedestrian in Front

① Creating : the victim LiDAR perceives a non-existing object

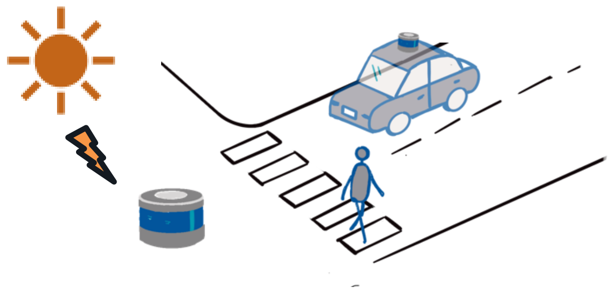


No Obstacle in Front

② Hiding : the victim LiDAR fails to perceive an existing object

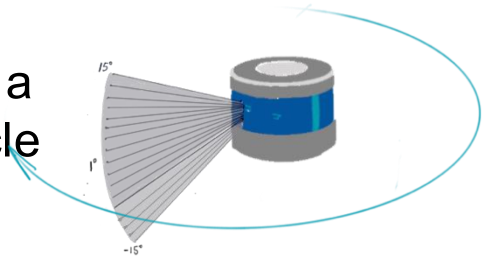
Challenges of Physical Laser Attack

1. Interference of surrounding environment



2. Scanning and rotating in high speed!

Scanning a single cycle in ~2.3μs



Rotating in 300~1200 RPM

3. "Curse of Light Speed"

small time error * **light speed** = large distance error

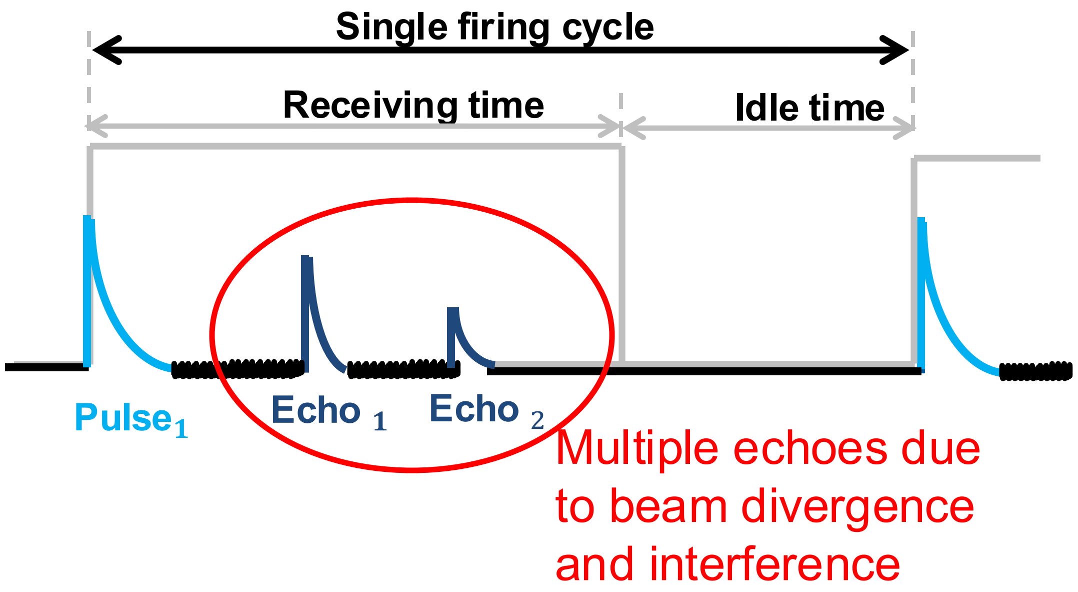
$e.g.\ 1ns * 3 * 10^8\, {}^{m}/_{s} = 30cm$

C1: **How to make the attack signal be considered as a valid echo?**

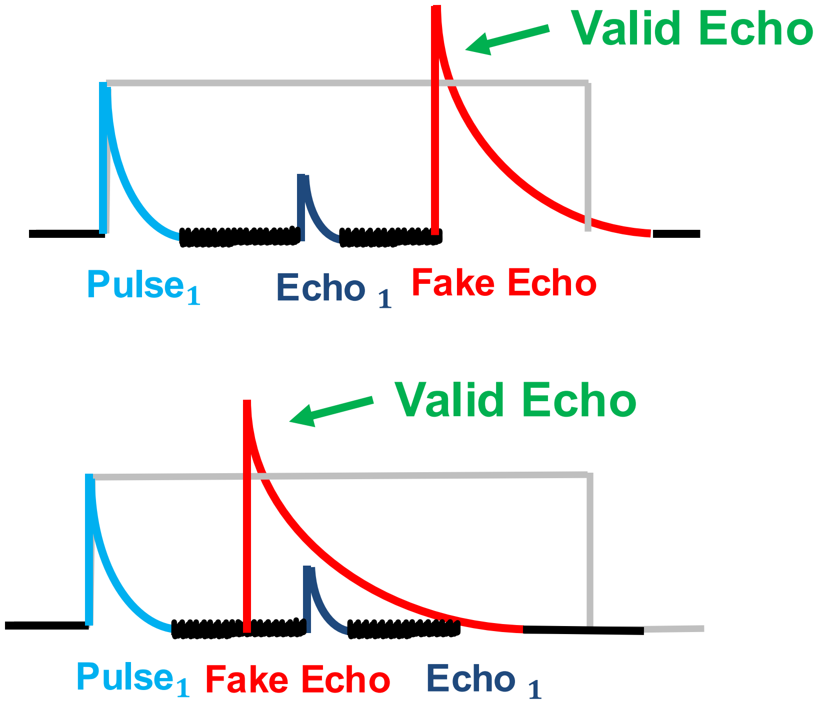
C2: **How to have a fine control of the injected point clouds?**

C1: What is valid echo?

- Return Mode:



Strongest — the strongest echo. ✓
Last —the last (temporally) detected echo.
Dual —both the strongest and last echo.



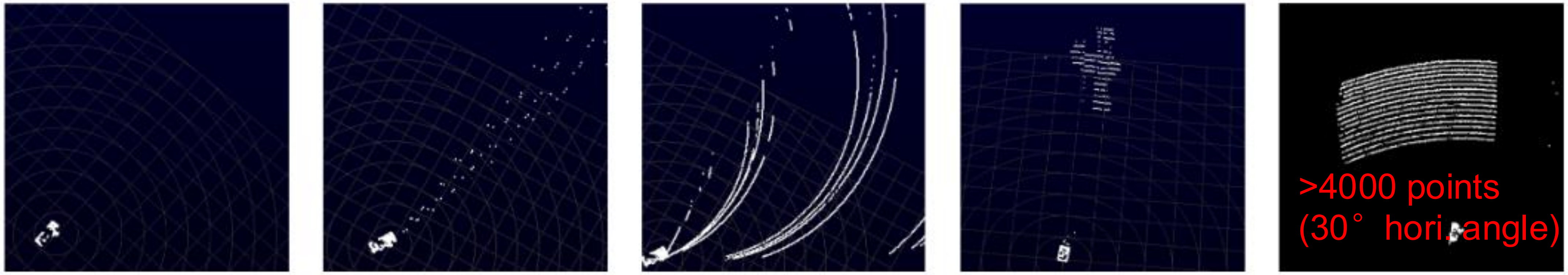
Vulnerability: Echo can be forged.

Insights:
- A **high-power** fake echo can be recognized as a valid echo.
- The real echo will be **ignored**.

Preliminary: Point Injection Experiment

Critical Laser Parameters：

Parameters	Requirements
Wavelength	Appropriate: Same as LiDAR laser
Power	High: The higher the better
Frequency	Precise: According to LiDAR scanning sequence



(a) Oiginal (b) Scatter (c) Wave (d) Broken Wall (e) Wall

SOTA:
Cao. CCS'19.
200 points
(8° hori. angle)

C2. Have a fine control of the point clouds: PLA-LiDAR

Step1. Point Cloud Design

- *Design the point cloud to either hide or create objects.*
- **Requirement:** Injectable point cloud.

Step2. Laser Signal Design

- *Convert the point cloud into laser signals*
- **Requirement:** Fully the 3D information of the point cloud.

Step3. Points Injection

- *Inject the signals into the lidar to generate desired point cloud.*
- **Requirement:** Precise synchronization.

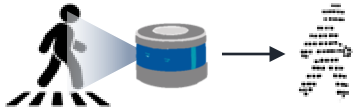
Step1: Injectable Point Cloud Design

Physical Constraints *:*

a. Every generated point can only locate on one of the LiDAR's laser rays;

b. Each laser ray can generate at most one point.

1. Record-based



Pros: 1) Black box. 2) Naturally satisfy the physical constraint.✓

2. Optimization-based

Question formulation:

Constrained Search Space.

$$\min_{P'} \ \mathcal{L}(P')$$

$$\text{s.t.} \ (R_i', \alpha_i', \omega_i') \in \mathrm{Loc}^{exp}, i \in [1, n] \quad \longrightarrow \text{Satisfy } \textit{constrain a.}$$

$$|\alpha_i' - \alpha_j'| + |\omega_i' - \omega_j'| \neq 0, i, j \in [1, n] \quad \longrightarrow \text{Satisfy } \textit{constrain b.}$$

Step2: Laser Signal Design.

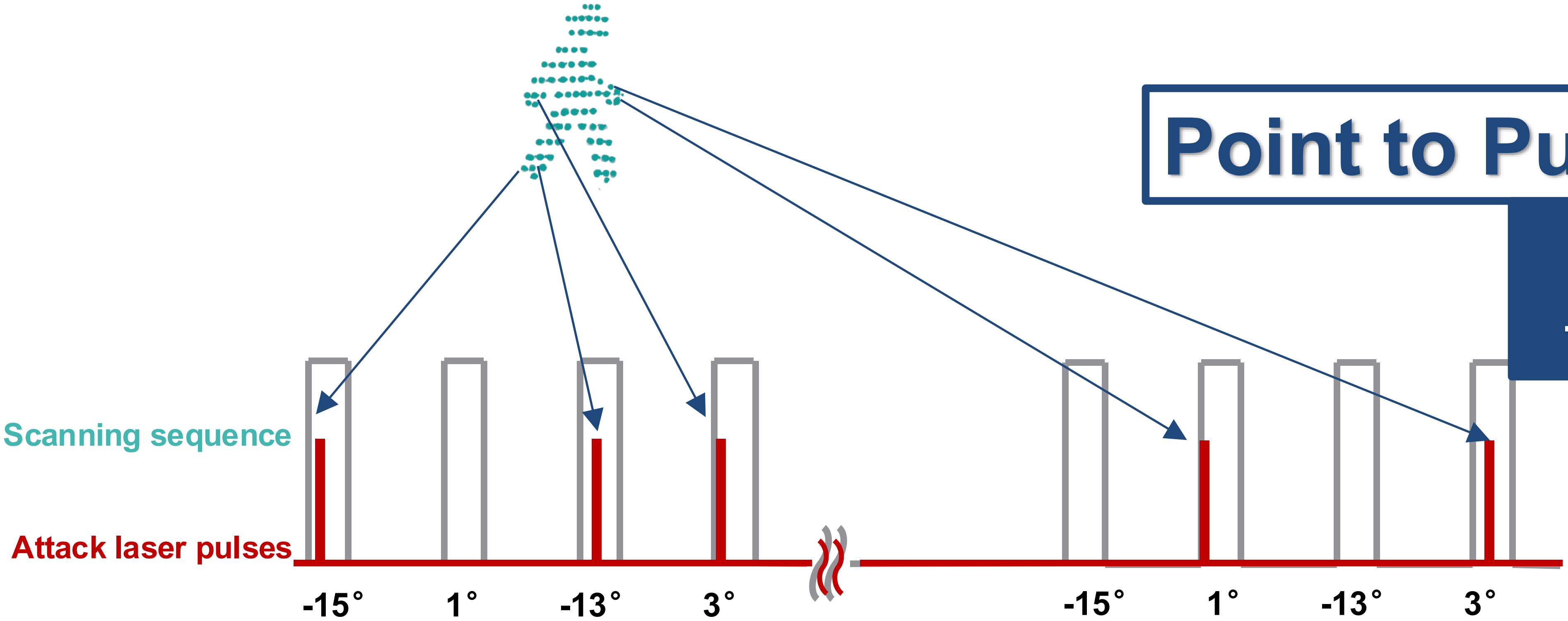
Point Cloud:
Spatial coordinates(X,Y,Z)

\+

Scanning Sequence:
Time sequence of how LiDAR scans

→ **Pre-Processing** → **Point Cloud Parameters:** ***Fullcycle_ID*** ***Singlecycle_ID*** ***ToF*** → **Signal Design** →

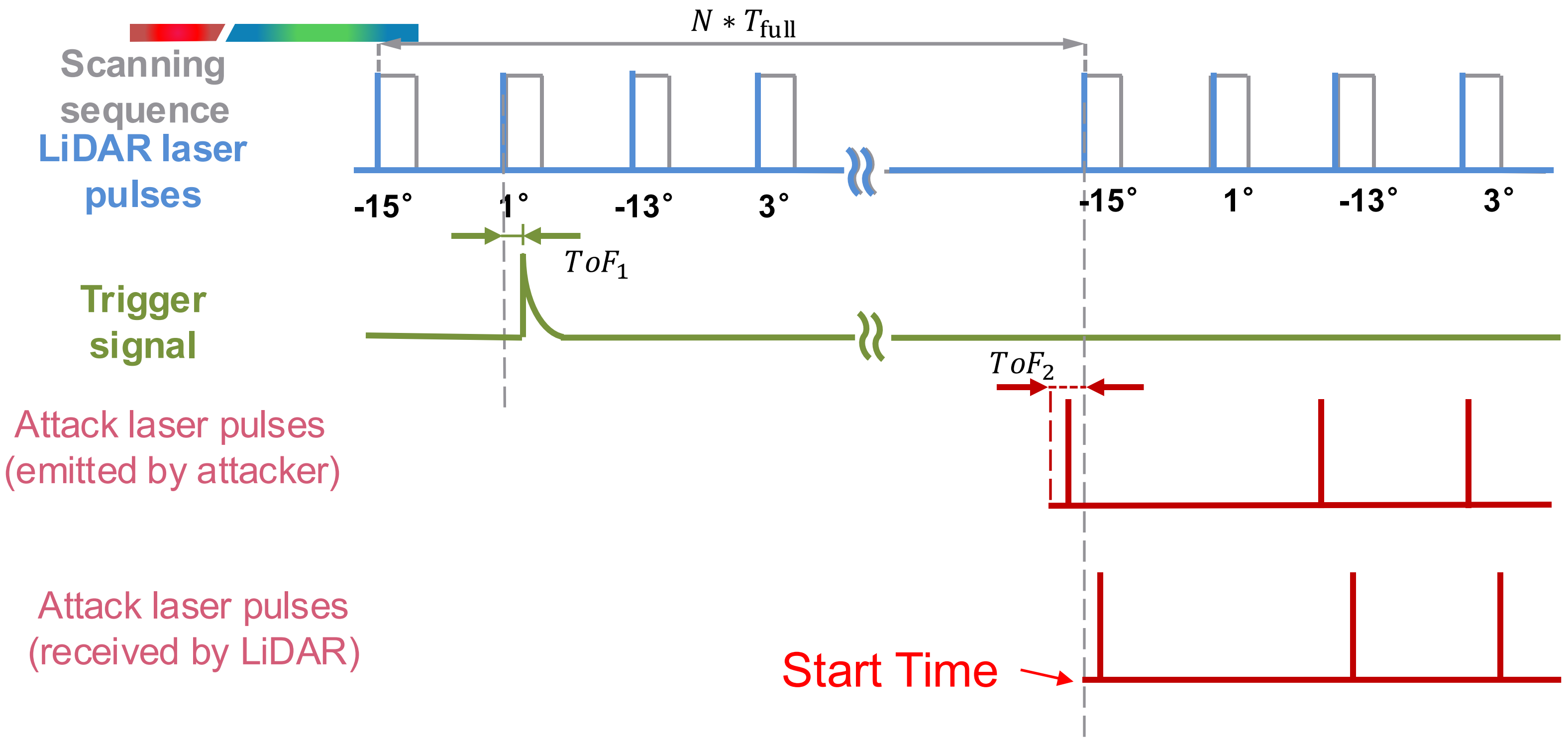
Attack Signal
Laser Pulses
Timestamp, *Pulse width*

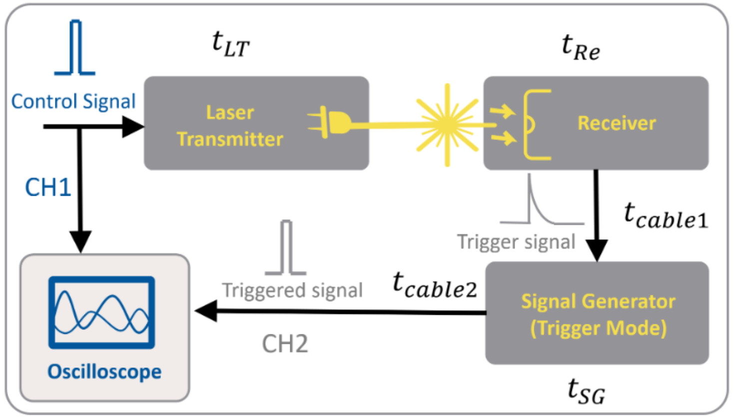


Point to Pulse Mapping

3D information
to
Temporal information

Step3: Points Injection with Precise Sychronization.





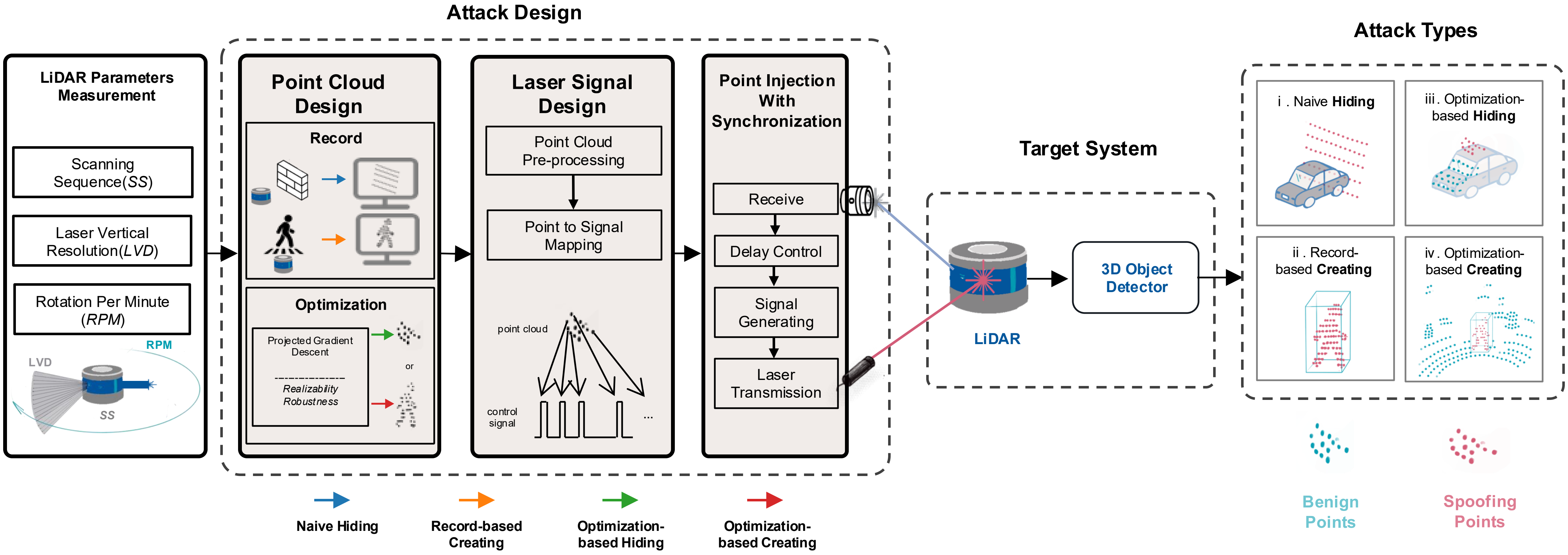
The measurement method of t_{device}

Empirically, time error should be within **3 nanosecond**.

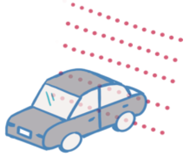
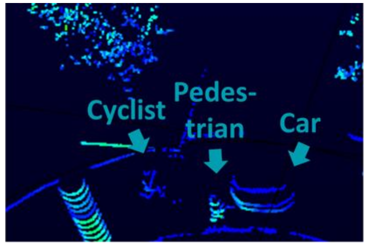
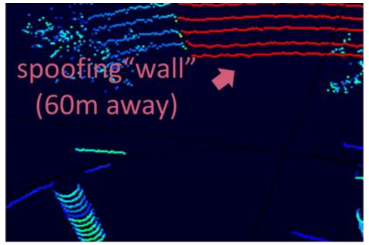
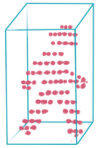
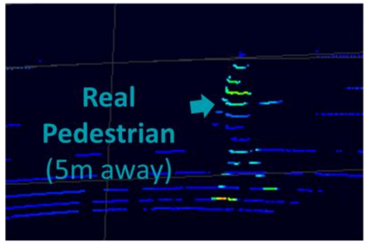
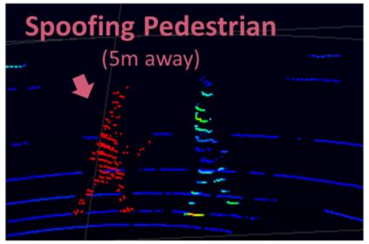
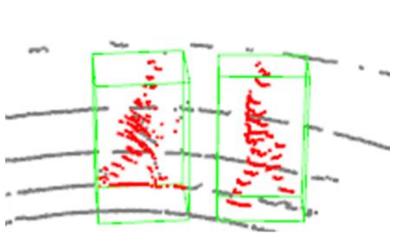
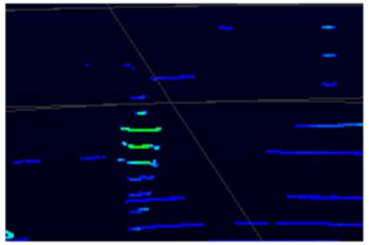
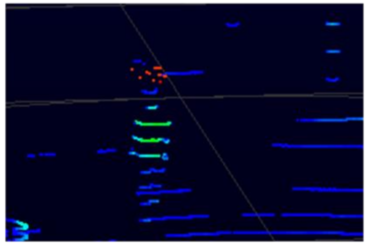
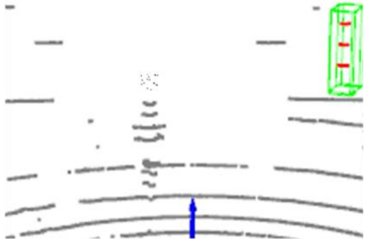
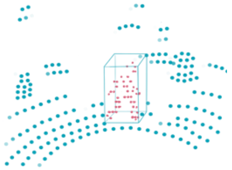
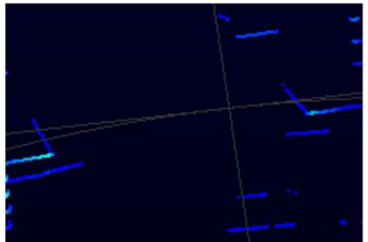
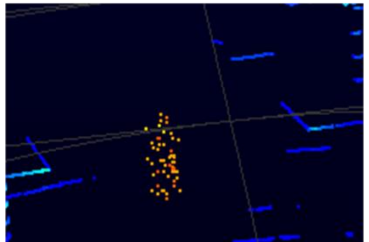
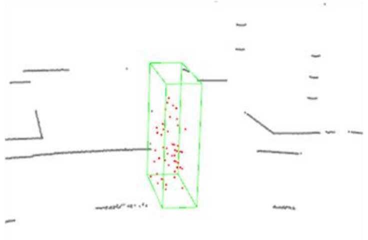
The delay should be set: $t_{delay} = t_{align} - ToF_1 - ToF_2 - t_{device}$

where $t_{align} = \mathrm{N} * T_{full} - T_{sfc}$

PLA-LiDAR: System Design.



Physical-World Attacks

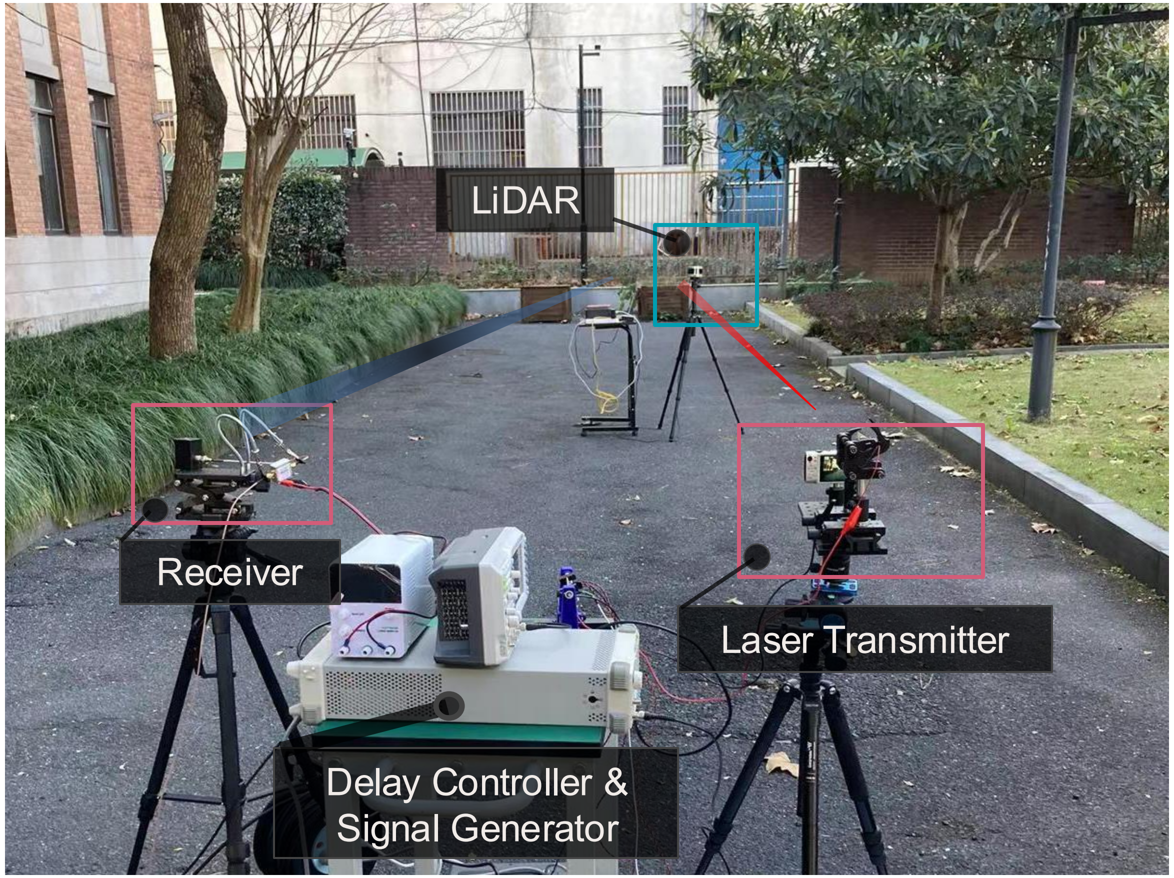
		Ground Truth	Benign point cloud	Point Cloud under attack	Detection under attack
	**Naïve Hiding**				
	**Record-based Creating**				
	**Optimization Hiding**				
	**Optimization Creating**				

Evaluation

- Physical-world attack Evaluation
  - **Four Attacks**
  - **Two LiDARs:** VLP-16, RS-16

  - **Three Models**: Second, Pointpillar, Apollo

Physical Attack Setup



Physical-World Attacks

- **Overall Performance**

90.5% 48.8%

Detector	LiDAR Model	Attack Types			
		Nai-Hide	Rec-Create	Opt-Hide	Opt-Create
SECOND	VLP-16	100%	98%	38%	72%
	RS-16	100%	86%	33%	61%
PoinPillar	VLP-16	100%	64%	79%	15%
	RS-16	100%	51%	68%	12%
Apollo	VLP-16	100%	98%	77%	37%
	RS-16	100%	89%	73%	21%

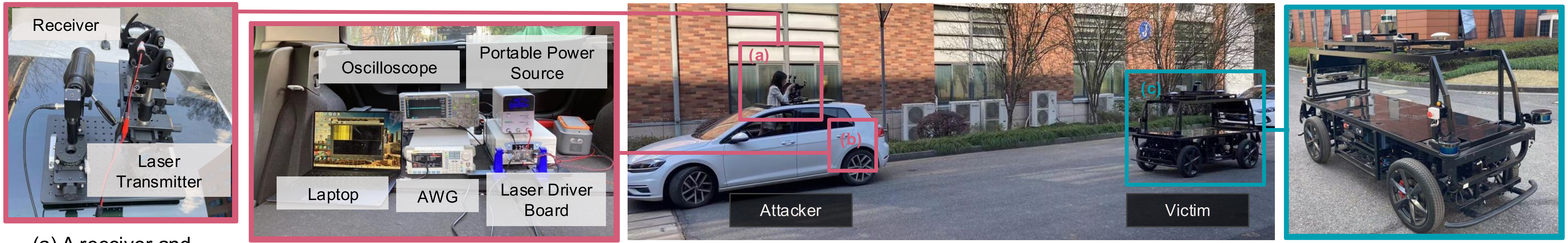
VLP-16: 73.2%

RS-16: 66.2%

Obersavations:

1. **VLP-16 (73.2%) is more vulnerable than RS-16 (66.2%).**
   - Period randomization can mitigates our attacks.
2. **Naive attack (90.5%) is better than optimization-based attack (48.8%).**
   - Optimization attack has a higher requirement on timing.

Feasibility Study on Moving Vehicle



(a) A receiver and a laser transmitter on the car roof.

(b) The attack equipment in the car trunk.

Setup of moving experiment .

(c) The victim car (Apollo KiT) with a VLP-16.

- Improvement to conquer Jitter when moving

➢ **A large-diameter telescope** (Φ = 50 mm) to expand the receiver's receiving area

➢ **Large spot diameter** (8 cm), and use a **high-power laser** diode (P_{peak} = 300 W)

- Attack success rate:

➢ Hiding Attacks **94.1%** (16/17 trials)

➢ Creating Attacks **78.9%** ASR (15/19 trials)

Feasibility Study on Moving Vehicle - Demo

Hiding Attacks

Creating Attacks

Potential Mitigation

1. **LiDAR Improvement**
   - Pulse Encoding
   - Pulse Randomizing and Scanning Period Randomizing
2. **Security Redundancy**
   - Multi-LiDAR Fusion
   - Multi-Sensor Fusion

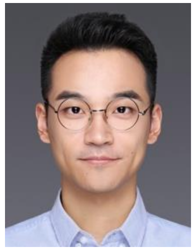
Summary

- **Proposed the PLA-LiDAR attack**
  - ➢ 20 times more points than prior works.
  - ➢ 4 types of attacks.
- **Extensive physical experiments**
  - ➢ 2 LiDARs + 3 Detection Models
  - ➢ Stationary + Moving
- **Show the physical threats of lasers against LiDAR-based object detection!**

PLA-LiDAR : Physical Laser Attacks against LiDAR-based 3D Object Detection in Autonomous Vehicle





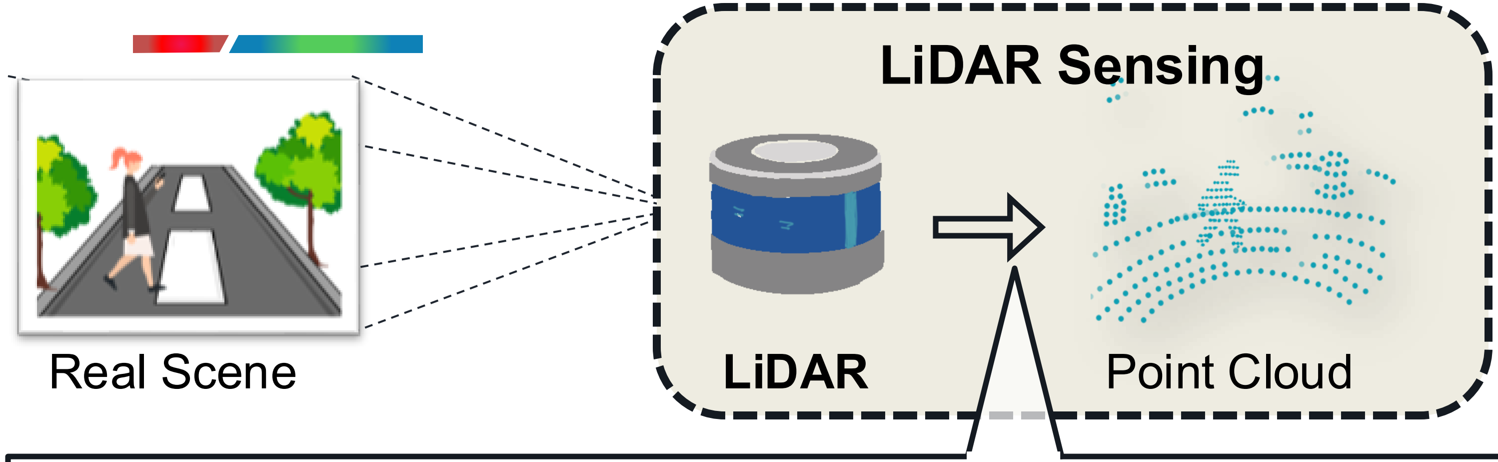



Demo Website:
https://sites.google.com/view/physical-lidar-attack

Corresponding Authors:
xji@zju.edu.cn, wyxu@zju.edu.cn

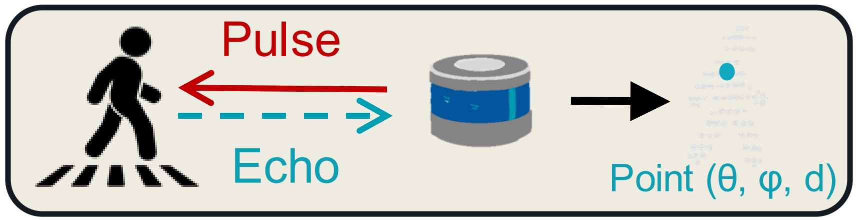
USSLAB Website: www.usslab.org

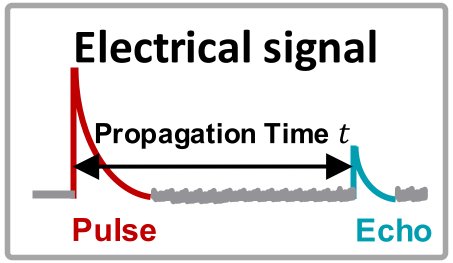
How Does Mechanical LiDAR work?



LiDAR generates point cloud by **laser ranging** and **laser beam steering.**

Laser Ranging

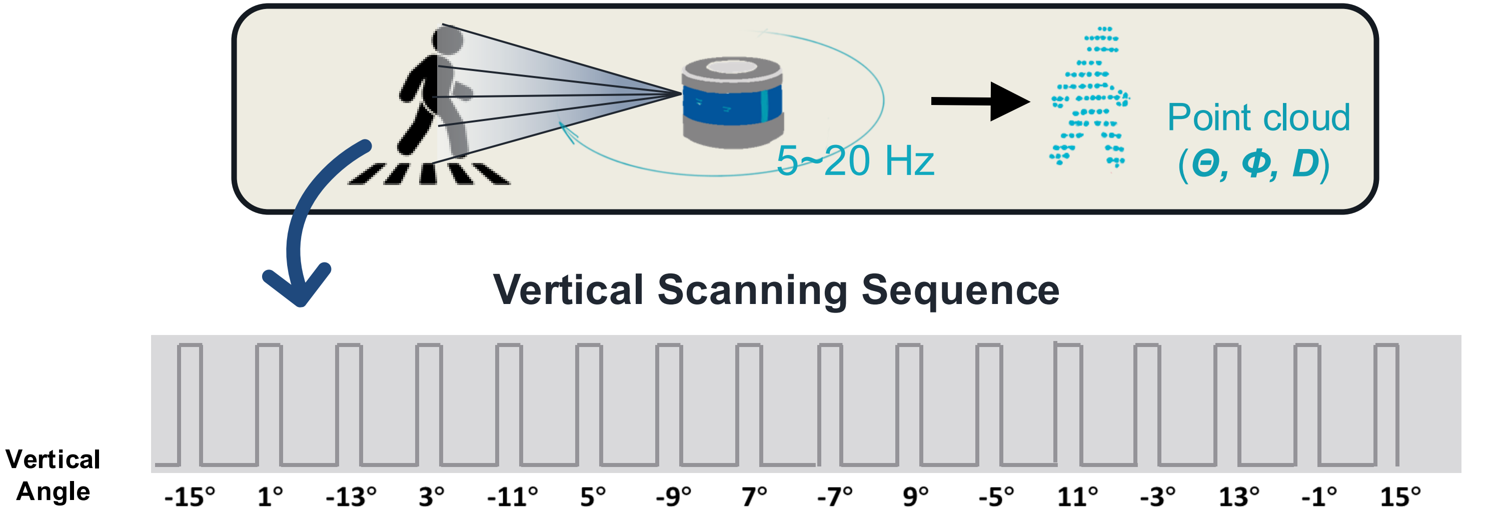




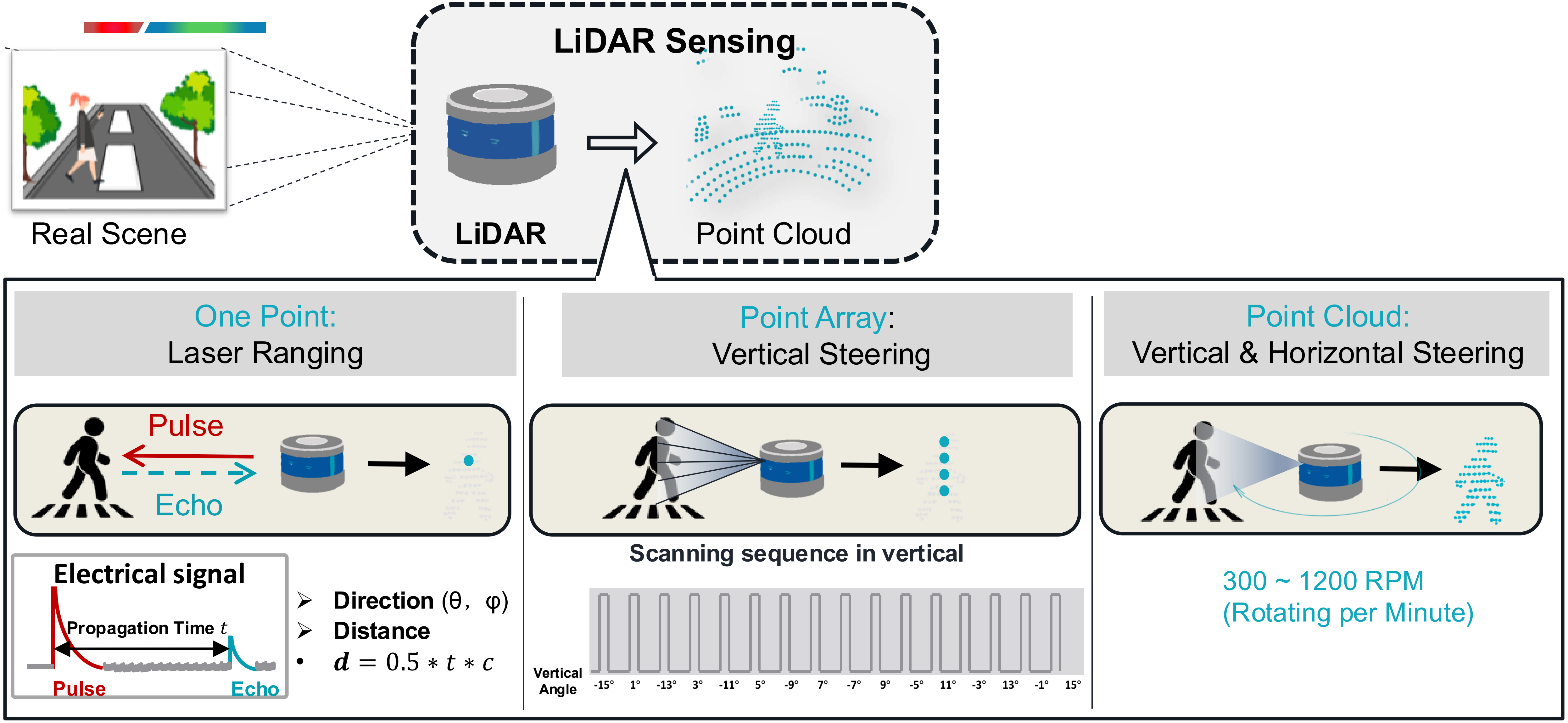
- **Direction** (θ, φ)
- **Distance**

$d = 0.5 * t * c$

Laser Beam Steering: Vertical Scanning + Horizontal Rotation

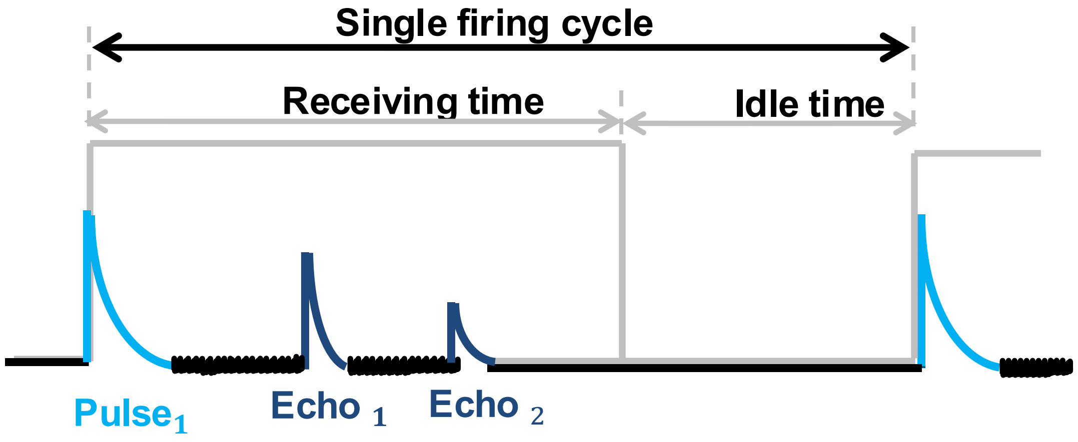


How Does LiDAR-based Perception work?

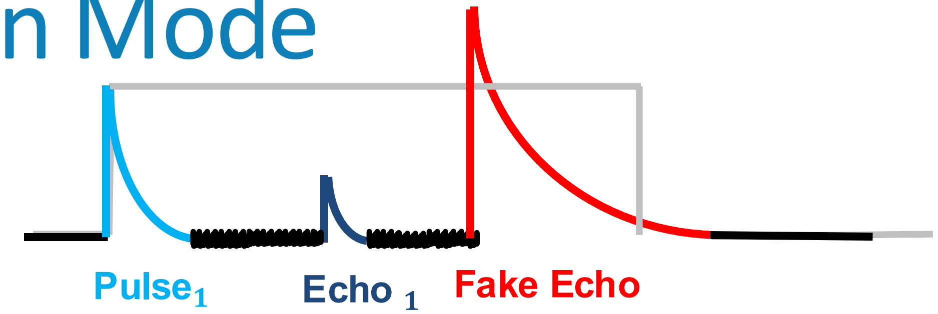


Dive into Mechanical LiDAR - Return Mode

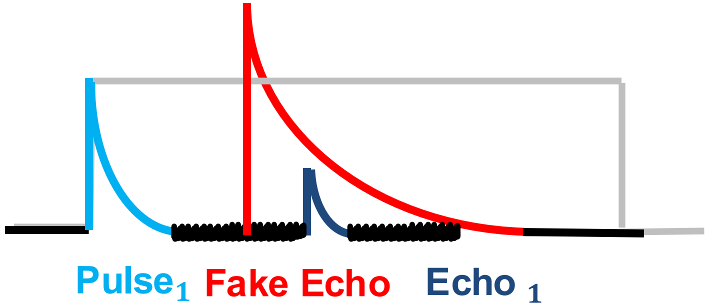
☐ Return Mode: **Strongest or Last or Dual**



Return Mode	Valid Echo	Point Number
Strongest (Default)	Echo $_1$	1
Last	Echo $_2$	1
Dual	Echo $_1$ & Echo $_2$	2



Return Mode	Valid Echo
Strongest	**Fake Echo**
Last	**Fake Echo**
Dual	Echo $_1$ & **Fake Echo**



Return Mode	Valid Echo
Strongest	**Fake Echo**
Last	Echo $_1$ **/ Fake Echo (saturation)**
Dual	Echo $_1$ & **Fake Echo**