

# Contents

<b>Spectral Encoding 상세 설계</b>	<b>1</b>
1. 개요	1
1.1 목적	1
1.2 설계 원칙	1
2. Range Image Projection	2
2.1 좌표 변환	2
2.2 파라미터	2
2.3 센서별 설정	2
3. FFT Transformation	2
3.1 Row-wise FFT	2
3.2 Rotation Invariance 원리	2
3.3 Frequency Selection	3
4. Per-Elevation Histogram	3
4.1 Log-scale Transformation	3
4.2 Histogram Binning	3
4.3 Global Normalization	3
5. 파라미터 영향	3
5.1 핵심 파라미터	3
6. Sensor Normalization	3
6.1 문제	3
6.2 해결: target_elevation_bins	4
7. Per-Elevation vs Sum	4
7.1 기존 방식 (Sum)의 문제	4
7.2 Per-Elevation 방식의 장점	4
8. 계산 복잡도	4
9. 요약	4

## Spectral Encoding 상세 설계

이 문서는 overall\_approach.md의 Stage 1-3 (Range Image + FFT + Histogram)의 상세 내용을 다룹니다.

### 1. 개요

#### 1.1 목적

3D LiDAR point cloud를 **rotation-invariant한 compact descriptor**로 변환합니다. FFT magnitude의 shift-invariance 특성을 활용하여 azimuth 방향 회전에 불변인 800차원 벡터를 생성합니다.

#### 1.2 설계 원칙

원칙	설명
Rotation Invariance	FFT magnitude로 azimuth shift 불변성 확보
Sensor Agnostic	elevation 해상도 정규화로 다양한 센서 지원
Height Preserving	Per-elevation histogram으로 높이 정보 보존

## 2. Range Image Projection

### 2.1 좌표 변환

Cartesian 좌표  $(x, y, z)$ 를 spherical 좌표  $(r, \theta, \phi)$ 로 변환:

$$r = \sqrt{x^2 + y^2 + z^2}$$

$$\theta = \arcsin\left(\frac{z}{r}\right) \quad (\text{elevation})$$

$$\phi = \arctan 2(y, x) \quad (\text{azimuth})$$

### 2.2 파라미터

파라미터	기본값	설명
Elevation bins ( $H$ )	16	수직 해상도
Azimuth bins ( $W$ )	360	수평 해상도 (1도 간격)
Max range	100m	최대 인식 거리

### 2.3 센서별 설정

센서	Elevation Range	Native Rings	Target Bins
VLP-16	$[-15^\circ, +15^\circ]$	16	16
HDL-64E	$[-24.8^\circ, +2^\circ]$	64	16
OS-128	$[-22.5^\circ, +22.5^\circ]$	128	16

## 3. FFT Transformation

### 3.1 Row-wise FFT

각 elevation row에 대해 1D FFT 적용:

$$F_i[k] = \sum_{n=0}^{W-1} R_i[n] \cdot e^{-j2\pi kn/W}$$

### 3.2 Rotation Invariance 원리

Azimuth 방향 회전  $\rightarrow$  row 방향 cyclic shift

$$R_i[n - \Delta n] \xrightarrow{\text{FFT}} F_i[k] \cdot e^{-j2\pi k \Delta n / W}$$

Magnitude는 shift에 불변:

$$|F_i[k] \cdot e^{-j2\pi k \Delta n / W}| = |F_i[k]|$$

### 3.3 Frequency Selection

컴포넌트	처리
DC (k=0)	제외 (노이즈에 민감)
Positive frequencies	사용
Negative frequencies	대칭이므로 제외

## 4. Per-Elevation Histogram

### 4.1 Log-scale Transformation

$$m_{\log} = \log(1 + m)$$

### 4.2 Histogram Binning

각 elevation별로 50개 bin의 histogram 생성:

$$\text{Output} = 16 \text{ elevations} \times 50 \text{ bins} = 800D$$

### 4.3 Global Normalization

전체 histogram이 sum=1이 되도록 정규화:

$$h_{\text{norm}} = \frac{h}{\sum_i h_i}$$

선택 이유: - 높이별 point 분포 정보 보존 - 지면 vs 건물 비율이 장소 특성 반영

## 5. 파라미터 영향

### 5.1 핵심 파라미터

파라미터	권장값	근거
n_elevation	16	센서 native 또는 정규화 타겟
n_azimuth	360	1도 간격 표준
n_bins	50	표현력과 차원의 균형
alpha	2.0	고주파 강조

## 6. Sensor Normalization

### 6.1 문제

다른 센서는 다른 elevation 해상도: - VLP-16: 16 rings - HDL-64E: 64 rings

## 6.2 해결: target\_elevation\_bins

모든 센서를 동일한 16개 elevation bin으로 정규화하여 sensor-agnostic 표현 생성

학습 센서	평가 센서	설정
VLP-16	HDL-64E	target=16
혼합	혼합	모두 target=16

## 7. Per-Elevation vs Sum

### 7.1 기존 방식 (Sum)의 문제

문제	설명
높이 정보 손실 Discriminative power 저하	ground vs building 구분 불가 다른 높이 구조가 유사한 descriptor 생성

### 7.2 Per-Elevation 방식의 장점

항목	Sum (50D)	Per-Elevation (800D)
높이 정보 Cross-sensor 일반화	손실 제한적	보존 개선

## 8. 계산 복잡도

단계	복잡도
Projection	$O(N)$
FFT (all rows)	$O(H \cdot W \log W)$
Histogram	$O(H \cdot W)$

실행 시간: ~5-10ms on CPU (single frame)

## 9. 요약

구성요소	역할
Range Image	3D → 2D projection
FFT Magnitude	Rotation invariance 확보
Per-Elevation Histogram	Height 정보 보존 (800D)
Global Normalization	Point 분포 정보 보존
Sensor Normalization	Sensor-agnostic 표현

문서 생성일: 2026-01-28