

# 열화상 환경에서의 단안 깊이 추정을 위한 손실 함수 분석

장지웅, 신준하, 고현석\*  
한양대학교

wkdwldnd7487@hanyang.ac.kr, ipip0114@hanyang.ac.kr, \*hyunsuk@hanyang.ac.kr

## Loss Function Analysis for Monocular Depth Estimation in Thermal Imaging Environments

Ji Woong Jang, Jun Ha Shin, Hyun suck Ko\*  
Hanyang Univ.

### 요약

본 논문은 열화상 환경에서의 단안 깊이 추정에서 제한된 공간적·구조적 단서로 인해 안정적인 깊이 예측이 어려운 문제를 다룬다. 이를 위해 학습 단계에서만 사전 학습된 RGB 단안 깊이 추정 모델로부터 생성된 깊이 맵을 구조적 깊이의 보조 감독 신호로 활용하고, 추론 단계에서는 열화상 영상만을 입력으로 사용하는 학습 설정을 채택한다. 이때 깊이 표현을 구조적 깊이와 거리 기반 깊이로 분리하여 학습함으로써, 장면의 기하 구조 보존과 실제 거리 스케일 회귀를 각각 안정적으로 수행할 수 있음을 보인다. 다양한 손실 함수 조합에 대한 실험을 통해, 각 손실 항이 깊이 추정 결과의 구조 보존 특성과 거리 스케일 안정성에 미치는 영향을 정성적·정량적으로 분석한다.

### I. 서 론

열화상 환경에서의 단안 깊이 추정은 조명 변화에 개인하다는 장점이 있으나, 제한된 텍스처 정보로 인해 장면의 기하 구조와 거리 스케일을 동시에 안정적으로 추정하는 데 어려움이 존재한다 [1, 3]. 기존의 단안 깊이 추정 방법들은 주로 하나의 깊이 표현을 단일 감독 신호로 학습하는 경우가 많아, 특히 열화상 영상과 같이 시각적 단서가 제한된 환경에서는 구조적 특성과 거리 기반 특성을 충분히 반영하는 데 한계가 있다.

본 논문에서는 열화상 기반 단안 깊이 추정을 대상으로, 손실 함수 설계가 깊이 추정 특성에 미치는 영향을 분석하기 위해 깊이 표현을 구조적 깊이와 거리 기반 깊이로 분리하여 학습하는 분석 설정을 구성한다. 학습 단계에서는 RGB 기반 깊이 정보를 구조적 깊이 학습을 위한 보조 감독 신호로 활용하고, 추론 단계에서는 열화상 영상만을 입력으로 사용한다. 이러한 RGB-thermal 보조 학습 전략은 기존의 RGB- thermal 융합 기반 깊이 추정 연구에서 영감을 받아 구성되었다 [4].

또한, 다양한 손실 함수 조합에 따른 깊이 추정 특성 변화를 정량적·정성적으로 분석한다.

### II. 본론

본 논문에서는 깊이 정보가 가지는 서로 다른 성격을 고려하여 구조적 깊이와 거리 기반 깊이를 분리하여 학습하는 설정을 따른다. 구조적 깊이는 학습 단계에서 RGB 영상으로부터 얻은 깊이 정보를 기준으로 활용하며, 절대적인 거리 정확도보다는 장면의 기하 구조와 경계 정보를 안정적으로 표현하는 데 목적이 있다. 반면, 거리 기반 깊이는 실제 거리 단위의 깊이 값을 회귀하는 것을 목표로 하며, 다양한 거리 범위에서의 예측 안정성이 요구된다.

구조적 깊이 학습에서는 인접 화소 간의 상대적인 깊이 변화와 공간적 연속성을 고려한 손실 항을 적용하여, 깊이 맵의 기울기와 경계 구조가 유지되도록 한다 [2]. 이를 통해 열화상 환경에서 빈번하게 발생하는 경계 흐림이나 구조 붕괴 현상을 완화할 수 있다.

거리 기반 깊이 학습에서는 상대 오차 기반 손실과 스케일 불변 특성을 고려한 손실 항을 적용하여, 근거리와 원거리 영역 간의 불균형을 완화하고 전체

장면에 대한 거리 스케일의 일관성을 유지하도록 한다. 또한 구조적 깊이를 가이드로 활용하는 정규화 손실과 두 깊이 표현 간의 기울기 일관성 제약을 통해, 구조 불일치로 인한 거리 스케일 왜곡을 완화한다.

그림 1과 그림 2는 서로 다른 손실 함수 조합에 따른 구조적 깊이와 거리 기반 깊이 예측 결과를 정성적으로 비교한 것으로, 각 손실 항이 구조 보존과 거리 스케일 안정성에 미치는 영향을 확인할 수 있다.



[그림 1] Qualitative comparison of structural (raw) depth predictions under different loss configurations.



[그림 2] Qualitative comparison of metric depth predictions under different loss configurations.

From left to right: (a) full loss configuration, (b) without SiLog loss, (c) without smooth loss, (d) without consistency loss, (e) without smooth/consistency loss.

### III. 실험 결과

본 절에서는 손실 함수 조합에 따른 거리 기반 깊이 추정 성능을 AbsRel과 RMSE 지표로 비교하였다. 모든 실험은 동일한 학습 조건에서 특정 손실 항을 제거하는 방식으로 수행하였으며, 표 1에 그 결과를 제시한다.

Table 1: Ablation results of loss configurations

Configuration	AbsRel ↓	RMSE ↓
All losses	0.0809	0.1720
No SiLog	0.0820	0.1732
No Smooth	<b>0.0794</b>	<b>0.1700</b>
No Consistency	0.0816	0.2114
No Smooth & Consistency	0.0895	0.2328

[표 1] Quantitative depth estimation performance under different loss configurations

실험 결과, 모든 손실 항을 적용한 설정에서 전반적으로 안정적인 성능을 보였으며, 특히 구조적 깊이와 거리 기반 깊이 간의 일관성을 유도하는 consistency 손실을 제거한 경우 RMSE가 크게 증가하여 거리 스케일 예측이 불안정해지는 경향을 보였다. Smooth 손실을 제거한 경우 AbsRel과 RMSE 지표는 오히려 개선되었고, 예측 결과 또한 GT와 높은 시각적 유사도를 유지하였다. 반면, smooth와 consistency 손실을 동시에 제거할 경우 정량 지표가 급격히 저하되었으며, 예측된 깊이 맵에서도 구조적 붕괴가

관찰되었다. SiLog 손실은 제거 시 성능 변화 폭이 비교적 작게 나타났다. 이러한 결과는 손실 항 조합에 따라 깊이 추정 특성이 상이하게 변화함을 보여준다.

### IV. 결론

본 논문에서는 열화상 환경에서의 단안 깊이 추정을 대상으로, 손실 함수 조합이 깊이 추정 결과에 미치는 영향을 분석하였다. 구조적 깊이와 거리 기반 깊이를 분리하여 학습하는 설정 하에서, 각 손실 항이 구조 보존, 예측 안정성, 그리고 거리 스케일 정확도에 서로 다른 영향을 미침을 확인하였다. 실험 결과, consistency 손실을 제거한 경우 RMSE가 크게 증가하여 거리 스케일 예측이 불안정해지는 경향을 보였다. 반면, smooth 손실을 제거한 경우 정량 지표 측면에서는 성능 향상이 관찰되었으며, smooth 손실을 적용한 경우에는 깊이 맵의 공간적 변동성이 감소하는 경향이 나타났다. 또한 손실 항을 적절히 조합하지 않을 경우 깊이 추정 성능이 급격히 저하될 수 있음을 확인하였다. 본 연구는 열화상 기반 단안 깊이 추정 환경에서 손실 함수 조합이 깊이 예측 특성에 미치는 영향을 정량적·정성적으로 분석하였다는 점에서 의의를 가지며, 향후 보다 안정적인 깊이 추정 모델 설계를 위한 기초 자료로 활용될 수 있을 것이다.

### 참고 문헌

- [1] D. Eigen, C. Puhrsch, and R. Fergus, "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network," NeurIPS 2014, pp. 2366– 2374.
- [2] C. Godard, O. Mac Aodha, and G. J. Brostow, "Unsupervised Monocular Depth Estimation with Left-Right Consistency," CVPR 2017.
- [3] S. Shin, S. Kim, J. Kim, and S. Lee, "Deep Depth Estimation from Thermal Image," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 10450– 10459.
- [4] Z. Meng and T. Fukao, "RGB-Thermal Infrared Fusion for Robust Depth Estimation in Complex Environments," arXiv:2503.04821, 2025.