

평면-인지 변형민감도 불확실성 기반 추론 제어를 통한 단안 깊이 추정의 기하 처리 성능 향상

순동현, 오민석, 이상철[†]
대구경북과학기술원

dhsoon@dgist.ac.kr, harrymark0@dgist.ac.kr, sangchul.lee@dgist.ac.kr

Improving Geometric Processing with Plane-aware Perturbation Uncertainty and Inference-time Control for Monocular Depth Estimation

Donghyeon Soon, Minseok Oh, SangChul Lee[†]
DGIST

요약

단안 깊이 추정은 단일 RGB 영상으로부터 깊이 맵을 예측하여 다양한 3D 비전 파이프라인의 입력으로 활용된다. 그러나 예측된 깊이를 3D 포인트로 역투영해 평면 피팅이나 표면 노말 계산에 사용하는 경우, 평면 내부의 미세한 깊이 요동과 국소적 오류가 3D 공간에서 outlier로 증폭되어 평면 적합 오차를 증가시키고 기하 처리의 안정성을 저하시킨다. 그럼에도 대부분의 단안 깊이 모델은 픽셀 단위의 신뢰도 정보를 제공하지 않아, 후속 기하 처리 관점에서 신뢰 가능한 입력을 선택적으로 구성하기 어렵다. 본 연구는 이러한 한계를 완화하기 위해 입력 영상에 대한 미세 변형(perturbation)에 따른 깊이 예측 변동성을 기반으로 한 불확실성 기반 신뢰도 표현인 PSU(Perturbation Sensitivity Uncertainty)를 제안한다. PSU는 입력의 작은 변화에 대한 예측 변동성을 통해 픽셀 단위 불확실성(또는 신뢰도) 정보를 추정하며, 이를 추론 단계에서 기하 처리 입력을 선택적으로 구성하는 제어 신호로 활용한다. NYU Depth V2[1] dataset에서의 실험 결과, PSU 기반 추론 제어는 불안정 픽셀의 영향을 완화하여 평면 기반 기하 처리의 안정성을 향상시키며, 별도 학습 없이도 downstream 성능을 개선할 수 있음을 시사한다.

I. 서론

단안 깊이 추정은 단일 RGB 영상으로부터 깊이를 예측해 로봇 내비게이션, AR, 3D 재구성 등 다양한 3D 비전 파이프라인의 입력으로 사용된다. 그러나 예측된 깊이를 3D 포인트로 역투영해 평면 피팅이나 표면 노말 계산에 활용할 때, 평면 내부의 미세한 깊이 요동과 국소적 오류가 3D 공간에서 outlier로 증폭되어 평면 적합을 불안정하게 만들고 노말의 공간적 일관성도 크게 저하시킨다. 따라서 단순한 깊이 오차뿐 아니라, 기하 처리 관점에서 신뢰할 수 있는 픽셀을 식별할 수 있는 기준이 필요하다. 본 연구는 입력에 가한 미세 변형(perturbation)에 대한 깊이 예측의 변동성을 이용해, 픽셀 단위 신뢰도 표현 PSU(Perturbation Sensitivity Uncertainty)를 제안한다. PSU는 불안정 픽셀을 정량적으로 드러내며, 이를 추론 단계 제어 신호로 사용해 3D 포인트 필터링과 불확실 영역 선택적 안정화를 수행한다. 또한 실내 장면 데이터셋을 통한 실험을 통해 PSU 기반 제어가 평면 피팅 성능(planarity, inlier ratio)과 평면 영역의 노말 안정성(normal instability)을 개선함을 보인다.

II. 본론

2.1 배경 및 문제 설정

단안 깊이 추정 모델 $f(\cdot)$ 은 입력 영상 I 로부터 깊이 맵 $D=f(I)$ 를 예측하며, 이를 카메라 파라미터를 통해 3D 포인트로 역투영되어 평면 피팅, 노말 계산, 포즈 추정 등 다양한 기하 처리 모듈의 입력으로 활용된다. 그러나 예측 깊이는 픽셀별 안정성이 불균일하며, 특히 평면-유사 영역의 미세한 깊이 요동이나 국소 오류는 3D 공간에서 outlier로 증폭되어 평면 적합을 불안정하게 만들고 inlier 비율을 저하시킨다. 또한 노말은 깊이 미분에 기반하므로 작은 깊이 변동도 노말의 공간적 불일

치를 유발할 수 있다. 결국 문제의 핵심은 깊이 오차의 크기 자체가 아니라, 후속 기하 처리 관점에서 신뢰 가능한 픽셀을 선별할 기준이 부족하다는 점이며, 이를 위해 픽셀 단위의 기하적 신뢰도 기준이 필요하다.

2.2 제안 기법: 변형-민감도 기반 불확실성 맵 (PSU)

본 연구는 단안 깊이 추정 결과의 기하적 신뢰도를 표현하기 위해, 입력 영상의 미세 변형에 대한 예측 변동성을 이용한 PSU(Perturbation Sensitivity Uncertainty)를 제안한다. PSU는 깊이의 절대 정확도를 직접 개선하기보다, 픽셀 단위로 예측 불안정성을 정량화하여 후속 기하 처리에서 신뢰 가능한 입력을 구성하도록 돋는다.

2.2.1 미세 변형 입력 생성

원본 영상 I 에 대해 K 개의 미세 변형 $T_i(\cdot)$ 를 적용하여 변형된 영상 집합을 생성한다. 본 연구는 구현이 단순하면서도 예측 민감도를 유발할 수 있는 변형으로 밝기·대비 변화, 약한 블러, JPEG 압축 등을 사용한다.

$$I_i = T_i(I), \quad i = 1, \dots, K$$

2.2.2 영상을 예측과 불확실성 정의

각 변형된 입력 I_i 에 대해 동일한 깊이 추정 모델 $f(\cdot)$ 를 적용하여 depth map D_i 를 얻는다. PSU는 동일 픽셀 (u, v) 에서 변형에 따른 깊이 예측의 변동성을 불확실성으로 정의하며, 픽셀 단위 불확실성 맵 U 는 K 회 예측 결과의 분산으로 계산한다.

$$U(u, v) = \text{Var}(\{D_i(u, v)\}_{i=1}^K)$$

$U(u, v)$ 가 큰 영역은 입력의 작은 변화에도 깊이 예측이 크게 달라지는 위치로, 후속 기하 처리에서 outlier로 작용할 가능성이 높다. 따라서 U 는 기하 처리에 부적합한 픽셀을 식별하는 기준으로 기능한다. 또한 불확실성 값은 해석이 직관적이지 않을 수 있으므로, 필요에 따라 이를

0~1 범위의 신뢰도로 변환해 사용한다.

$$C(u, v) = \exp\left(-\frac{U(u, v)}{\sigma}\right)$$

여기서 σ 는 스케일 조정 상수이며, $C(u, v)$ 가 클수록 해당 픽셀의 예측이 변형에 대해 안정적임을 의미한다. 이후 절에서는 이를 추론 단계 제어 신호로 사용하여 기하 처리 입력을 조정한다.

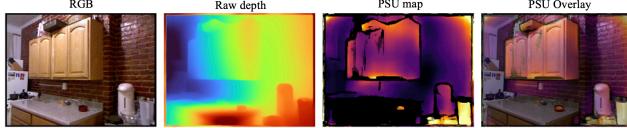


Figure 1. PSU(plane-aware) visualization

Figure 1은 PSU의 시각화 결과로, 입력 변형에 민감한 영역이 높은 값으로 강조됨을 보여준다.

2.3 PSU 기반 추론 제어와 downstream 성능 향상

PSU는 픽셀 단위 불확실성/신뢰도를 기반으로 후속 기하 처리 입력을 조정하기 위한 추론 단계 제어 신호로 사용된다. 본 연구는 별도의 추가 학습 없이, 추론 시 계산한 PSU를 이용해 다음 두 가지 전략으로 downstream 성능 향상을 도모한다.

2.3.1 PSU 기반 선택적 안정화(uncertainty-guided selective refinement)

전역적으로 동일한 스무딩을 적용하면 경계가 뭉개지거나 세부 구조가 손상될 수 있다. 따라서 본 연구는 PSU가 높은(불안정한) 영역에서만 안정화 효과가 크게 반영되도록 설계한다. 신뢰도 $C(u, v)$ 를 가중치로 사용하여 원본 깊이 D 와 스무딩 결과 Smooth(D)를 다음과 같이 결합한다.

$$\tilde{D}(u, v) = C(u, v) \cdot D(u, v) + (1 - C(u, v)) \cdot \text{Smooth}(D)(u, v)$$

즉, C 가 큰 픽셀은 원본 예측을 유지하고, C 가 작은 픽셀은 스무딩의 비중을 높여 국소적 요동을 완화한다.

2.3.2 PSU 기반 3D 포인트 선택

Point cloud 생성이나 PnP 기반 포즈 추정처럼 깊이를 직접 3D 기하 추정에 사용하는 단계에서는, 일부 불안정 픽셀이 outlier로 작용해 전체 추정을 불안정하게 만들 수 있다. 이를 완화하기 위해 불확실성 $U(u, v)$ 가 불확실성 임계값 τ 미만인 픽셀만 선택하여 3D 포인트를 구성한다.

$$\mathcal{M} = \{(u, v) \mid U(u, v) < \tau\}$$

이러한 포인트 필터링은 평면-유사 영역에서의 outlier 비중을 줄여 planarity 오차 감소 및 inlier ratio 증가와 같은 기하 처리 성능 향상으로 이어진다.

2.4 실험 결과

본 절에서는 PSU(Perturbation Sensitivity Uncertainty)가 후속 기하 처리에서 불안정 픽셀(outlier 후보)을 식별하는 기준으로 유효함을 보이고, 이를 활용한 추론 단계 제어의 평면 추정 성능을 평가한다. 실험은 NYU Depth V2[1] 실내 장면에서 수행하였으며, 사전학습 단안 깊이 예측을 3D로 역투영한 뒤 평면 적합을 적용하였다. PSU는 입력에 미세 변형을 가했을 때의 깊이 예측 분산으로 계산하였고, PointFilter는 불확실 픽셀을 제외해 3D 포인트를 구성하였다.

2.4.1 정성적 비교 결과: PSU 기반 포인트 필터링

Figure 2는 Raw depth를 역투영한 3D 포인트에 평면을 적합한 뒤, 각 포인트의 residual(평면으로부터의 거리

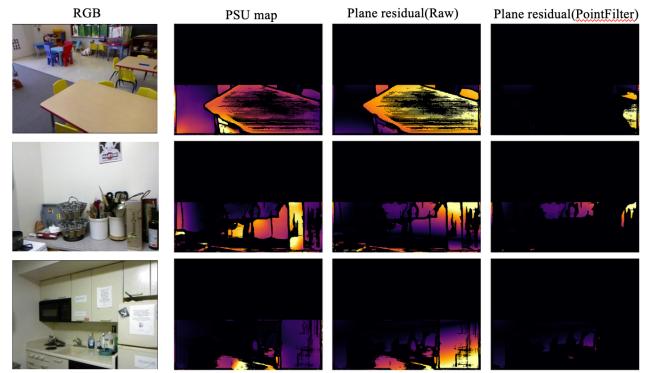


Figure 2. Plane residual result (Raw v PSU-guided PointFilter)

오차)을 색으로 시각화한 결과이다.

Raw에서는 평면-유사 영역에도 고잔차 포인트가 넓게 분포해 적합 안정성을 저하시킨다. 반면 PSU-guided PointFilter를 적용하면 불확실 픽셀에서 유래한 고잔차 포인트가 제거되어 residual tail이 줄고, 평면 적합이 더 안정적으로 수행된다. 이는 PSU 기반 포인트 필터링이 downstream 기하 처리에 유효한 3D 입력 구성에 기여함을 정성적으로 보여준다.

2.4.2 정량적 비교 결과: Planarity / Inlier ratio 비교

Table 1. Quantitative comparison

Method	Planarity ↓	Inlier ratio ↑	Normal instability ↓
Raw	0.2353 ± 0.1479	0.2730 ± 0.1039	0.2261 ± 0.0752
PointFilter	0.1915 ± 0.1815	0.3690 ± 0.1586	-
Refined	0.2268 ± 0.1503	0.2729 ± 0.1038	0.2229 ± 0.0757

표 1은 Raw, PointFilter, Refined 결과의 평면 적합 성능을 정량 비교한 것이다. Planarity error는 평면 잔차의 평균으로 낮을수록 좋고, Inlier ratio는 잔차가 임계값 이하인 포인트 비율로 높을수록 평면에 일관되게 부합하는 포인트가 많음을 의미한다. 실험 결과 PointFilter는 planarity error를 낮추고 inlier ratio를 높여 평면 추정 성능을 전반적으로 개선하였다. 또한 Refined는 normal instability를 감소시켜 평면 영역의 노말 방향 요동을 완화했다. 이는 PSU가 추론 단계에서 신뢰 포인트 선택과 선택적 보정을 가능하게 하는 제어 신호로 작동해 downstream 기하 처리 성능 향상에 기여함을 시사한다.

II. 결론

본 연구는 단안 깊이 추정 결과를 후속 기하 처리에 안정적으로 활용하기 위해, 입력 영상의 미세 변형에 대한 예측 변동성으로부터 픽셀 단위 기하적 신뢰도(PSU)를 추정하는 방법을 제안하였다. PSU는 불안정 픽셀을 식별해 추론 단계에서 신뢰 가능한 3D 입력을 구성하는 제어 신호로 활용된다. 실험에서 PSU 기반 포인트 필터링과 선택적 보정은 outlier 영향을 줄이고 평면 영역의 노말 요동을 완화하여, 추가 학습 없이도 기하 처리 성능을 개선함을 확인하였다. 이러한 결과는 별도의 추가 학습 없이도 PSU를 통해 downstream 기하 처리 성능을 개선할 수 있음을 시사한다.

ACKNOWLEDGMENT

본 연구성과는 산업통상자원부의 지원으로 한국산업기술진흥원의 지원을 받아 수행된 연구임(No. RS-2024-02633871).

참 고 문 현

- [1] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, "Indoor segmentation and support inference from RGBD images," in Proc. European Conf. on Computer Vision (ECCV), 2012.