

Recognition, Control and
Learning Algorithm Lab.

超音波センサーを用いた 単眼深度推定器のスケール補正

西嶋 拓哉¹, 池畠 諭^{2,3}, 佐藤 育郎^{1,2}

¹東京科学大学, ²デンソーITラボラトリ, ³国立情報学研究所

Institute of
SCIENCE TOKYO

**DENSO
IT LAB**

NII 大学共同利用機関法人 情報・システム研究機構
国立情報学研究所
National Institute of Informatics

1. 背景と目的

背景

- 駐車支援における鳥瞰ビューア生成のために、自車両近傍の物体（駐車車両など）の3次元情報が必要。



<https://toyota.jp/safety/scene/parking/index5.html>

- 近年、一般物体に対する密な単眼深度推定を可能にする基盤モデルが利用可能。（Depth Anything V2 [L. Yang, et al., arXiv]等）
- 低コストで、既に多くの自動車に搭載されている超音波センサーは、物体の最近傍点までの距離計測が可能。

目的

1 : 単眼深度推定器と超音波センシングを用いた高精度かつ密な絶対深度推定手法の開発

- 単眼深度推定器と超音波センシングの欠点をフュージョンにより克服。

表. 単眼深度推定器と超音波センシングの比較

| 技術 | 単眼深度推定器 | 超音波センシング |
|------|-----------|----------------|
| 動作原理 | パターン認識 | Time-of-Flight |
| 利点 | ・方位分解能が高い | ・距離精度が比較的高い |
| 欠点 | ・絶対距離が不正確 | ・方位分解能が極めて低い |

2 : 評価用データセット構築

- 単眼カメラの画像と超音波センサーによる距離出力と真の密な深度を取得。

2. 関連研究(単眼深度推定)

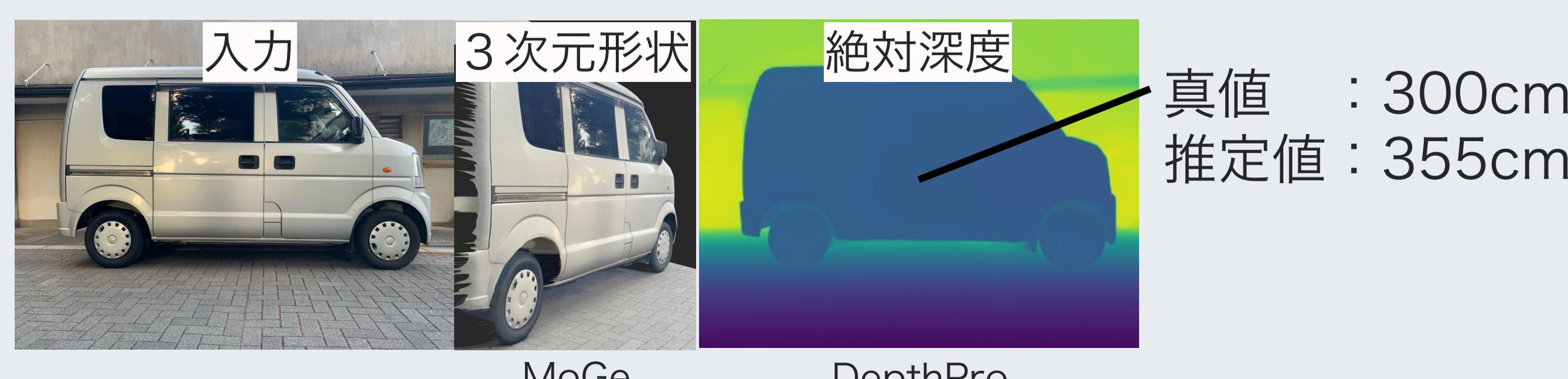
概要

- 単一のカメラ画像から密な深度を推測する技術。
 - 多数のデータセットで学習された基盤モデルが利用可能。

表. 単眼深度推定の代表的な手法

| 手法 | 出典 | 識別ベース | 生成ベース | データセット数 |
|----------------|------------------------|-------|-------|---------|
| MoGe | Wang+, CVPR2025 | ✓ | | 21 |
| Depth Pro | Bochkovskii+, ICLR2025 | ✓ | | 27 |
| GeoWizard | Fu+, ECCV2024 | | ✓ | 5 |
| Depth Anything | Yang+, CVPR2024 | ✓ | | 14 |
| Marigold | Ke+, CVPR2024 | | ✓ | 2 |
| DepthGen | Saxena+, arXiv2023 | | ✓ | 5 |
| ZoeDepth | Bhat+, arXiv2023 | ✓ | | 12 |
| MiDaS | Ranftl+, arXiv2019 | ✓ | | 5 |

- 課題
- 3次元形状は比較的高精度だが、絶対深度は不正確。



3. 提案手法

入力

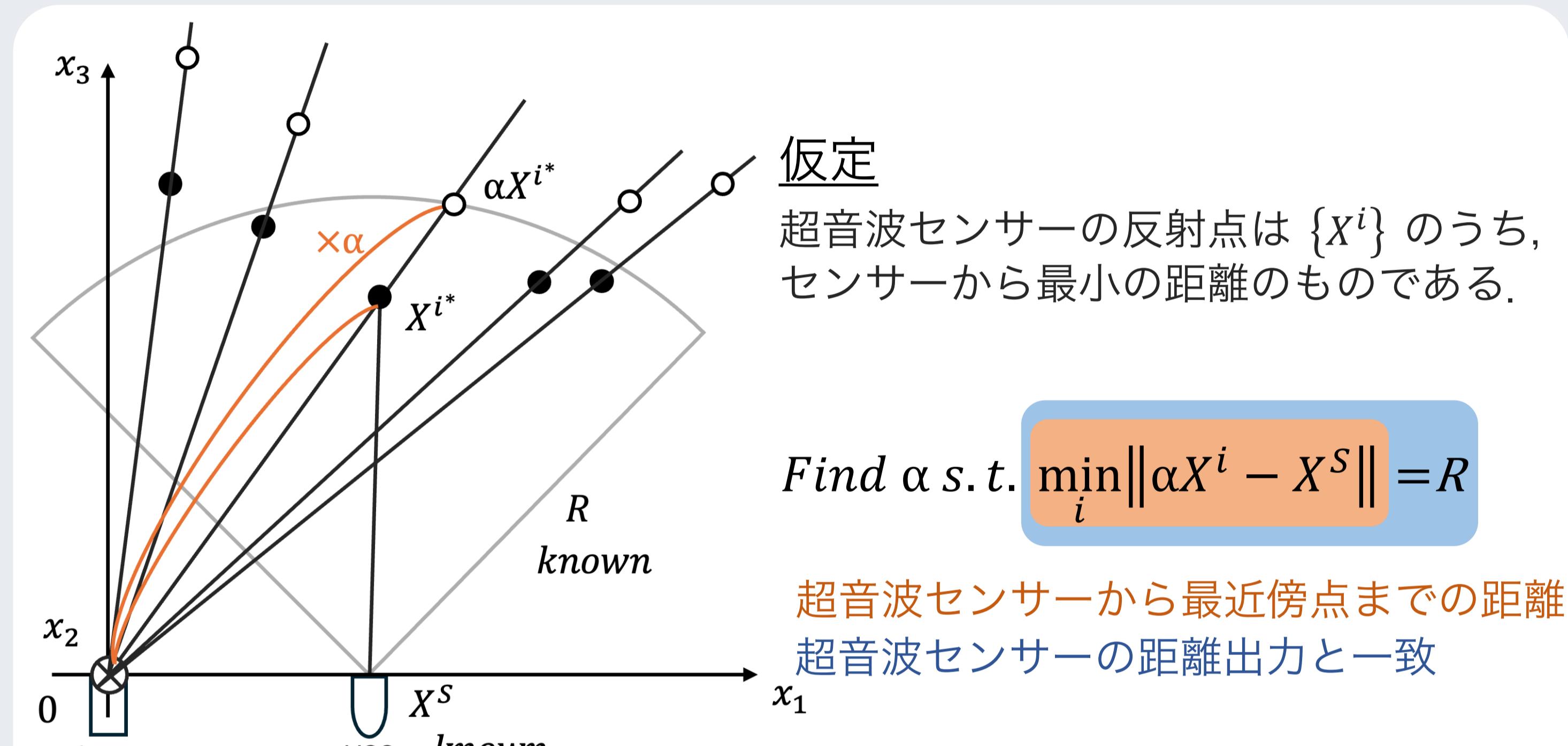
- $\{X^i\}_{i=1,\dots,N}$: 単眼深度推定器に画像 I を入力したときに outputされる 3 次元点群
- $R > 0$: 超音波センサーの距離出力

点群 $\{X^i\}$ のスケール補正のモデリング

$$\bar{X}^i := \alpha X^i, \alpha > 0 \quad s.t. \quad \frac{1}{N} \sum_i \|\bar{X}^i - X^{i,GT}\| \rightarrow \min$$

この時、 α をスケール補正因子と呼ぶ。

フュージョンによる超音波センサーの反射点 $X^{i,*}$ の位置、および密な点群のスケール補正因子 α の同時推定



4. 初期評価

評価用データ

- 屋外の静止した自動車に対し、デプスカメラ (RealSense D455f) と超音波センサー (JSN-SR04T-3.0) を各 1 台ずつ用いて RGB 画像、距離出力を 1 組取得した。



+ $R=157.7\text{cm}$

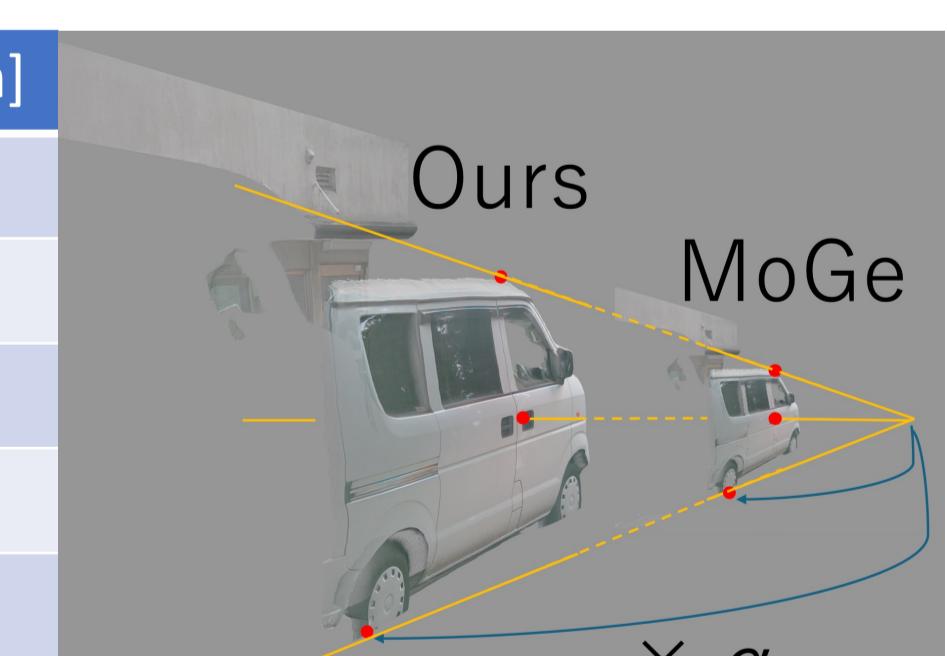
超音波センサーの
距離出力

評価結果

- 提案手法 (MoGe+スケール補正) を、既存の 4 つの単眼絶対深度推定手法と反射点画素 i^* の深度の誤差と RMSE を比較した。

表. 定量評価結果：反射点画素の誤差と有効画素 RMSE

| 手法 | i^* 深度 [cm] | 誤差 [cm] | 相対誤差 [%] | RMSE [cm] |
|-------------------|---------------|---------|----------|-----------|
| GT(Real Sense) | 158.9 | - | - | - |
| 提案手法 | 155.8 | 3.1 | 2.0 | 34.6 |
| UniK3D | 326.3 | 167.4 | 105.3 | 210.8 |
| DepthPro | 324.0 | 165.1 | 103.9 | 177.0 |
| Metric3D v2 | 276.2 | 117.3 | 73.8 | 136.8 |
| Depth Anything V2 | 489.3 | 330.4 | 207.9 | 399.7 |



- 提案手法は i^* の深度誤差を 3.1cm (2.0%) に、RMSE を 34.6cm に抑え、他手法を大きく上回った。
- 単眼深度推定モデルに超音波センサーの距離情報を 1 点追加するだけで、絶対深度推定の精度が飛躍的に向上することを確認した。

5. 今後の計画

- 多数のシーンによる統計的検証
- 超音波センサー特性の考慮による高精度化
- 混合整数二次計画問題等の最適化手法による高速化