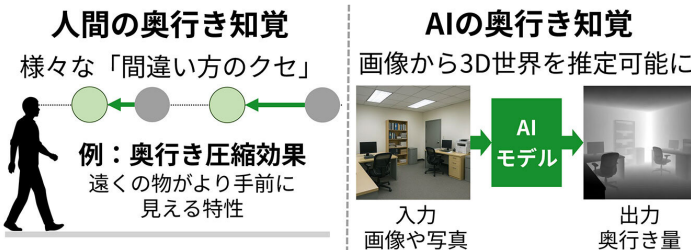


<p>どんな研究</p>	<p>私たちは写真から自然に立体を知覚でき、今やAIもその能力を獲得しつつあります。では、そうしたAIは、人間と同じような手がかりを使って3D世界をとらえているのでしょうか？本研究では、大規模データを用いて人間とAIの奥行き知覚における「間違い方のクセ」を比較しました。</p>
<p>どこが凄い</p>	<p>屋内外の多様な写真に対する人間の評価データを収集し、多数のAIと比較しました。その結果、AIは人間と共通のバイアスを持つ（例：遠くのものほどより手前に見える）一方、高精度なAIが、かえって人間とは異なる傾向を示す「精度と人間らしさの間のトレードオフ」を発見しました。</p>
<p>めざす未来</p>	<p>人間が3D世界をとらえる際の「間違い方のクセ」をAIで精密にモデル化することで、人間に寄り添う情報提示の実現に貢献します。画面越しのロボットの遠隔操作のような、錯覚や見間違いが起こりうる状況でAIが注意を促し、私たちの安全・安心をサポートする未来をめざします。</p>

研究の背景

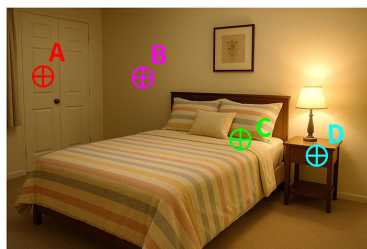
人間は、網膜に映った情報から、**様々な手がかりを使って、3D世界を推定**しています。



問：人間とAIの奥行き知覚を比べたら、「**間違い方のクセ**」はどのくらい似ている？
なぜ重要？：人間が3D世界を捉える際の見間違いを精密に予測し、人間に寄り添った情報提示を実現したい。

人間の奥行き評価データの収集

オンライン実験により、**人間の奥行き評価データ**を大規模に収集しました。

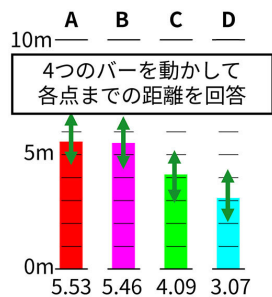


※実験刺激を模した画像をGPT-4.0により生成

■方法

- ・画像の撮影位置から各点までの距離をメートル単位で回答
- ・1画像あたり8点または16点の回答を収集して分析

- 画像 屋内写真（NYU Depth V2データセット）
- 屋外写真（KITTIデータセット）



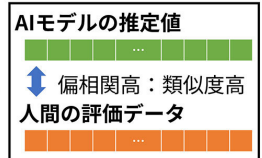
人間とAIの「間違い方のクセ」の比較

■使ったAIモデル：69種類の既存AIモデル

- ・学習方法、学習に用いるデータ、モデルの構造などが多様なAIモデルを準備しました。
- ・ただし全てのAIモデルは、**物理的に正しい奥行きを推定することを目標に訓練**されています。

■比較方法：偏相関による解析

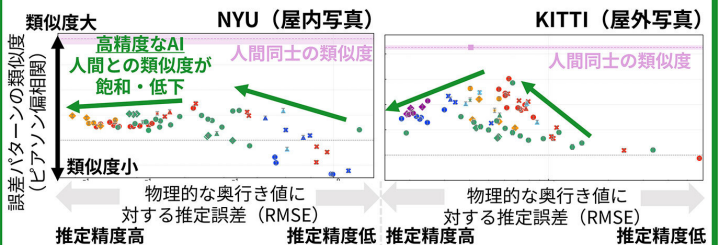
- ・物理的な奥行き値の影響を取り除き、**人間とAIの間違い方のパターンを比較**しました。



AIモデルの精度と「人間らしさ」の関係

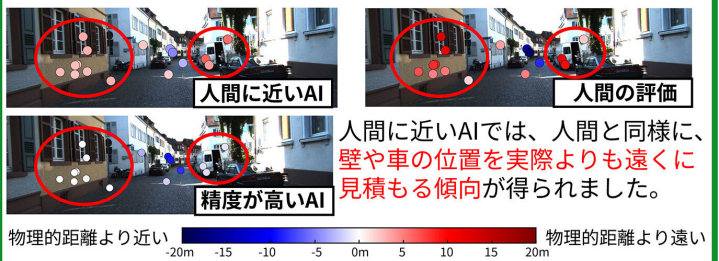
屋内写真（NYU）と屋外写真（KITTI）について、

- ・多くのAIモデルが、**人間と共通のバイアスを持つ**ことを発見しました。（例：奥行き圧縮効果）
- ・高精度なAIモデルが、人間と異なる傾向を持つ「**精度と人間らしさのトレードオフ**」を発見しました。（例：壁や車の位置）



※図の各点が、各AIモデルの推定精度（横軸）と人間の誤差パターンとの類似度（縦軸）の程度を示します。（色：学習に用いるデータ、形状：学習方法の違い）

■人間とAIの奥行き推定を比較した画像例（屋外写真）



関連文献

[1] Y. Kubota, T. Fukiage, "Human-like monocular depth biases in deep neural networks," *PLOS Computational Biology*, Vol. 21, No. 8, e1013020, 2025.
 [2] Y. Kubota, T. Fukiage, "Accuracy does not guarantee human-likeness in monocular depth estimators," *arXiv*, 2512.08163, 2025.
 [3] Y. Kubota, T. Fukiage, "Benchmarking human and DNN biases in monocular depth estimation," under review, 2026.

連絡先

久保田 祐貴 (Yuki Kubota) 人間情報研究部 感覚表現研究グループ

共同研究先・外部資金

本研究は、日本学術振興会 (JSPS) 科研費 (23KJ2226) の助成を受けて実施しました。