

N2 脳情報科学

屋内環境理解のための3次元質問応答 ～3次元空間の内容を理解して質問に答えるAIシステム～

概要

本研究では、視覚的質問応答技術と近年発展が著しい3次元空間認識技術を統合することで、3次元の空間情報を自然言語で問い合わせできる3次元空間質問応答技術を開発しました。また、複数の屋内環境をRGB-Dスキャンした3次元空間情報の質問応答データセットを作成し、実空間の意味内容の問い合わせに3次元情報が有用であることを実証しました。

特徴

- 本研究課題では、3次元空間データに対して質問応答できる実世界質問応答技術の確立に挑戦し、それを評価するための大規模な3次元空間質問応答データセットを作成しました。
- 従来の視覚的質問応答技術は、2次元の画像情報を対象としてきたため、実空間の3次元空間に情報アクセスする方法がありませんでした。この問題を3次元空間認識技術と質問応答技術を組み合わせることで解決しました。
- 人工的に合成した3次元データではなく、実際の屋内環境をRGB-Dスキャンした3次元空間データに対して質疑応答できるので、本技術を事務用・ホームロボットなどに転用することができます。

今後の展開

- 本研究で作成した3次元空間質問応答データセットを拡張し、3次元空間内を言語指示に従って移動できるエージェントに対して、質疑応答できる手法とデータセットを作成する予定です。

テーマ「Society5.0への貢献～サイバーとフィジカルの融合に向けて～」との関連

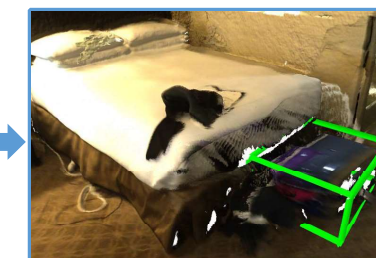
- 本技術で実現される実空間を質疑応答する機能は、実空間の意味内容を理解して対話できるロボットや、サイバー・フィジカル空間の3次元空間情報を検索できる検索エンジンへの展開など、幅広い分野への応用が期待できます。

質問 + 3Dスキャン



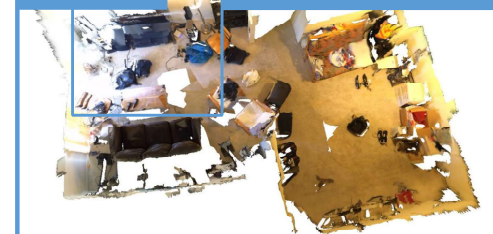
質問：中くらいのサイズの青いスーツケースはどこに置いてありますか？

回答 + 3D矩形



回答：ベッドの右横にあります

3Dスキャン



質問：テレビと木の椅子の間の床には何が置いてありますか？



回答：黒いバックパックが2つあります

