

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	ソーシャル・コンピューティング (荒牧 英治 (教授))		
学籍番号	2211029	提出日	令和 6年 1月 17日
学生氏名	稲積 駿		
論文題目	Visual Question Answeringにおける視線情報を用いた質問の曖昧性解消		
要旨			
<p>画像に対する質問応答 (VQA: Visual Question Answering)のような、画像を参照する会話には指示語の利用により質問に曖昧さが生じる。また、言語によっては質問の中核となる情報を持つ項の省略が行われ、問題はさらに複雑になる。実際にこうした質問の曖昧さは、システムとユーザ(話者)の共同注視や話者の視線といった情報で補完することが可能である。本研究ではこうした視線情報の参照による質問の曖昧性解消に着目し、視線先の物体と質問の指示語や省略が対応づいた視線情報付きVQAデータセット (GazeVQA) を提案する。質問作成の際に物体名を含めないことで、視線情報を用いた補完が必要な曖昧な質問を収集した。また、VQAにおける回答の多義性・多様性を評価で考慮するため、GazeVQAのテストセットにおける各質問に10件の回答を割り付けた。本研究ではさらに、このGazeVQAの質問に対する回答の精度を高めるため、視線元からの注視対象推定を活用する質問応答モデルを提案した。GazeVQAを用いた実験の結果、提案モデルはGazeVQAにおける特定の質問タイプに精度良く回答ができ、注視対象推定を用いない既存モデルと比較して優れた性能を達成した。</p>			