

審査の結果の要旨

氏 名 シン アンドリュー

本論文は「Image Narrative Generation via Interactive Visual Question Generation and Answering」と題し、画像の表層的な説明文のみならず、感情的、推論的な要素を含むより幅の広い記述や、ユーザの好みに合わせた記述をおこなう画像物語文生成手法の提案及び構築をおこなったものであり、全10章からなる。従来の研究では、画像の大域的かつ表層的な要素の記述を主な目的としている。しかしながら、実際の人は大域的な事象のみならず、局所的な領域に注目して記述したり、もしくは表層的な要素ではなく、感情的な要素や推論的な要素に関しても記述することができる。さらに、同じ画像であっても、人によって着目点や興味を持つところは異なるので、それぞれ要求される記述の内容も異なるべきである。従って、画像からより人に近い文章を生成するためには、従来の研究のような大域的かつ表層的な記述では限界があり、感情的、推論的、ユーザの好みを反映可能な手法が要求される。本論文では、感情的要素を表現可能な画像特徴量と局所情報を保持する画像特徴量の構築と、自問自答モジュールにより、従来の表層的な記述を超えた「画像物語文」の生成ができることを示している。さらには、インタラクティブな画像質問生成と回答により、生成される画像物語文を各ユーザに合わせて適合する手法を提案及び構築している。

第1章「Introduction」では、従来の画像理解に関する研究の問題点や限界点を説明し、本論文の目的を提示している。続いて第2章「Defining Image Narrative Generation」では、画像物語文を構成する要素に関する議論を通じて、従来の研究との差別点を明確にし、本論文で扱うタスクの定義を与えている。

第3章「Previous Works on Vision and Language」では、従来の画像認識と自然言語処理の境界研究を説明することによって、本論文を理解するために必要な背景知識を伝えている。

第4章「Dense Image Representation for Locally Robust Captioning」では、画像中の小さい物体の情報も反映できるように局所情報を保持可能な画像特徴量を構築し、それが画像の局所的説明文生成として利用可能であることを実験的に示している。

第5章「Image Captioning with Sentiment Terms」では、ウェブ上で収集されたセンチメントデータセットとマルチラベル設定を通じて感情的要素を反映可能な画像特

微量を構築し、感情的用語の推定にもとづく画像説明文生成を通じて、提案特微量の有効性を確認している。

第6章「Visual Question Generation (VQG) and Answering (VQA)」では、画像に関した推論的要素の記述を実現するために、画像質問生成(VQG)と応答(VQA)を組み合わせた自問自答モデルを提案している。

第7章「Single Image Narrative Generation」では、第4章と第5章において構築された特微量と、第6章で構築されたVQGとVQAを組み合わせた自問自答モジュールにより、画像1枚から画像物語文の自動生成モデルを構築している。

第8章「Interactive Image Narrative Generation」では、第7章のモデルにユーザとのインタラクションを導入し、ユーザの好みを引き出す質問自動生成をおこない、ユーザからの回答にもとづいた学習を通じて、同じユーザが新規画像のどこに着目するかを予測するモデルを構築している。

第9章「Experiments」では、第7章で提案された画像物語文生成手法を大量の画像に適用し、人による評価によって提案手法が従来手法より有効であることを確認している。続いて、第8章で提案されたインタラクティブな画像物語文生成手法がユーザの好みを反映できることをクラウドソーシングによる実験を通じて実証している。

第10章「Conclusion & Future Works」では、以上を総括した上で、適切な画像特微量にもとづき、自問自答やインタラクションをおこなうことによって、感情的要素や推論的要素を含めると同時にユーザの好みを反映可能な画像物語文生成手法が本論文の学術的寄与であると結論付けている。

以上、これを要するに、本論文は、従来の画像説明文研究で見逃されてきた局所的な要素、感情的な要素や推論的な要素、そして各ユーザの好みを反映できるようなモデルを提案している。これにより、従来困難であった、表層的な記述を超えた画像の物語文生成の端緒が開かれたと言える。

以上の理由から、本論文は知能機械情報学上貢献するところ大である。よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。