

深層生成モデルを用いた画像の周辺補完  
～通常画角から 360 度へ～

2022 年度

秋本 直郁

## 主 論 文 要 旨

報告番号	甲	第	号	氏 名	秋本 直郁
主論文題名：					
深層生成モデルを用いた画像の周辺補完 ～通常画角から 360 度へ～					
<p>(内容の要旨)</p> <p>アクションカメラやヘッドマウントディスプレイをはじめとして、多様な撮影デバイスと表示デバイスが普及している。そのため、撮影された画像の縦横比と表示画面のサイズが異なることが頻繁に発生しており、最適な表示のために画像のピクセル数を増減させる処理が行われる。本研究では、画像の周辺ピクセルを補完するという <b>Image outpainting</b> を用いて、画像拡張と 360 度画像の生成を行う。これらは画像の不足している領域を補うという処理であるため、多様な画面サイズに合わせた表示やコンテンツ制作にも有用である。しかしながら、外挿問題であるため、内挿を対象にした既存の <b>Image inpainting</b> 手法では望ましい結果を得ることができず、新たな手法の探求が必要である。そこで本研究では、問題を簡単化するための入力画像に対する前処理と、情報を広範囲から集約するためのネットワーク構造を持つ <b>Outpainting</b> 手法を検証し、その生成品質の高さで 3DCG 等のコンテンツ制作での実利用へ近づけることを目的とした。</p> <p>第1章に、本研究の背景と目的を述べた。</p> <p>第2章では、深層学習の基礎技術と、本研究に関連する先行研究について概説した。</p> <p>第3章では、<b>Image outpainting</b> による画像拡張に取り組んだ。既存手法の課題として、入力ピクセルと離れた位置の補完ピクセルの質が下がること、適用される画像のシーンが限定されることを挙げた。本研究での目標は、これらの改善に加え、生成画像のセマンティックのコントロール性の付与とした。これら目標に対し、入力画像の前処理の一種となる <b>Mirrored input</b> を提案した。そして、広範囲の情報を集約するための CNN 手法の検討を行った。<b>Mirrored input</b> は、補完領域の近くに入力ピクセルを配置することで、より補完の質が向上することを明らかにした。さらにそれは、多様なシーンに適用されてきた <b>Inpainting</b> 手法の性能を引き出せることと、生成をコントロールする条件付けとなることを実験で示した。一方、広範囲の情報を集約する CNN 手法でさえ、<b>Mirrored input</b> が有効であったことから、<b>Outpainting</b> にはより一層広範囲から情報を集約する方法が必要であるという課題を明らかにした。</p> <p>第4章では、<b>Image outpainting</b> による 360 度画像補完に取り組んだ。第3章で明らかにした広範囲から情報を集約すべきという知見を生かし、非局所的に情報を集約する目的で <b>Transformer</b> を導入した手法を提案した。結果として、既存手法よりも自然な見た目の生成が可能であることを示した。また、<b>Transformer</b> の出力からのサンプリングにより多様な出力結果を得ることで、出力結果を選択する機会を利用者に与えられることを示した。既存手法の課題の1つである、学習解像度への過適合に対して、補完と調整という2段階の生成方法を導入することで解決できることを示した。そして最後に応用として、3DCG ソフトウェア上で、生成された 360 度画像を遠景画像として利用可能であることをデモし、<b>Image outpainting</b> がコンテンツ制作での実利用へ近づいたことを示した。</p> <p>第5章では、結論として本研究で得られた成果を要約した。</p>					

Thesis Abstract

No. \_\_\_\_\_

Registration Number	<input checked="" type="checkbox"/> "KOU" <input type="checkbox"/> "OTSU" No. <span style="float: right;">*Office use only</span>	Name	AKIMOTO, Naofumi
Thesis Title			
<p>Image Outpainting with Deep Generative Model: Completion from Normal Field of View to 360-Degree</p>			
Thesis Summary			
<p>Many shooting and display devices, such as action cameras and head-mounted displays, have different aspect ratios of captured images and display screens, thus requiring processing to increase or decrease the number of pixels in the images. This study uses Image outpainting, which complements the surrounding pixels of an image, to expand the image and generate a 360-degree image. However, since this is an extrapolation problem, existing image interpolation methods do not produce desirable results. To achieve better performance, this work examines Outpainting methods that preprocess input images to simplify the problem and have network structures to aggregate information from a wide range. The goals are to achieve a high-quality generation by the proposed methods and bring it closer to real-world use in content creation, such as 3DCG.</p> <p>Chapter 1 describes the background and objectives of this work.</p> <p>Chapter 2 reviews the basic techniques of deep learning and related work.</p> <p>Chapter 3 describes image extension by Image outpainting. Existing methods suffer the low quality of completed pixels and limited scenes. This work aims to improve these issues and provide control over generated content. To that end, this paper proposed Mirrored input, which places the input pixels near the completion region. Moreover, this paper validated a CNN method for aggregating a wide range of information. Mirrored input improved completion quality and applied Inpainting methods, which have been applied to various scenes, to the Outpainting problem, and provided a conditioner to control the generation. On the other hand, the fact that mirrored input was effective means that Outpainting has the problem of requiring a way to aggregate information from a wider range of areas.</p> <p>Chapter 4 addresses 360-degree image completion by Image outpainting. Chapter 4 proposed a method that introduces a Transformer to aggregate information in a non-local manner, following the findings in Chapter 3. The experiments showed that the proposed method could generate more natural-looking images than existing methods. Also, diverse outputs could be obtained by sampling from the Transformer's output. Furthermore, the proposed method solved overfitting to the training resolution, which is one of the problems of existing methods, by introducing a two-step generation method: completion and adjustment. Finally, as an application, this work demonstrated that the generated 360-degree image could be used as a background image on 3DCG software, indicating that Image outpainting is getting closer to real-world use in content creation.</p> <p>Chapter 5 concludes this study.</p>			