

AI 特許紹介(38)
AI 特許を学ぶ！究める！
～3D 変換 AI 特許～

2022 年 3 月 10 日
河野特許事務所
所長弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第 4 次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

1.概要

特許出願人 Nvidia

出願日 2021 年 3 月 5 日

公開日 2021 年 9 月 9 日

公開番号 US2021/0279952

発明の名称 逆グラフィックス生成のためのニューラルレンダリング

952 特許は、2D 画像からマルチビューの 2D 画像を生成する StyleGAN 等の生成ネットワークと、生成したマルチビューの 2D 画像を用いてメッシュ、照明、テクスチャ等の 3D 情報を予測する逆グラフィックスネットワークとを用いたニューラルレンダリング技術に関する。

2.特許内容の説明

図 2 は、トレーニングパイプライン 200 を示す。

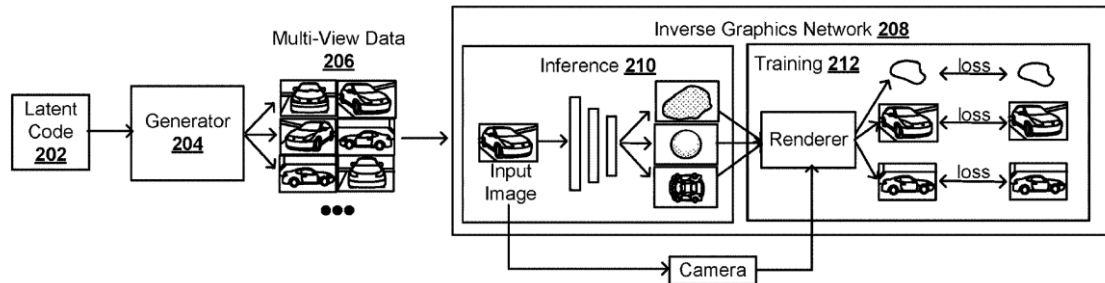


FIG. 2

パイプライン 200 は、2つの異なるレンダラーを含む。

第1のレンダラーはGAN (StyleGAN) 等の生成ネットワークであり、第2のレンダラーは補間ベースの微分可能レンダラー (DIB-R) 等の微分可能グラフィックスレンダラーである

DIB-R は、画像内の全てのピクセルのグラデーションを分析的に計算する微分可能なレンダリングフレームワークである。このフレームワークは、前景のラスタライズをローカルプロパティの加重補間として、背景のラスタライズをグローバルジオメトリの距離ベースの集約として表示できるため、様々な照明モデルを通じて、頂点の位置、色、法線、光の方向、テクスチャ座標を正確に最適化できる。

ジェネレータ 204 は、マルチビューデータセット 206 の効率的なアノテーションを備えた合成データジェネレータとして使用される。

次に、このデータセットを使用して、2D 画像から 3D プロパティを予測する逆グラフィックスネットワーク 212 をトレーニングする。このネットワークは、慎重に設計されたマッピングネットワークを介してジェネレータの潜在的なコードを解きほぐすために使用される。

マルチビューデータセットの入力画像は逆グラフィックスネットワーク 208 へ入力され、逆グラフィックスネットワークは、推論ネットワークを利用して、画像内のオブジェクトの形状、照明、およびテクスチャなどの 3D 情報を推論する。

トレーニングのために、3D 情報は、入力画像からのカメラ情報とともに、入力として

微分可能なグラフィックスレンダラーに供給される。レンダラーは、3D 情報を利用して、特定のカメラ情報の形状情報、3D モデル、または 1 つ以上の画像を生成する。

次に、これらのレンダリングは、損失値を決定するために、適切な損失関数を使用して関連するグラウンドトゥルースデータと比較される。当該損失は、ネットワークの重みまたはパラメーターを調整するために使用される。

逆グラフィックスネットワーク 208 の出力を使用して、ジェネレータ 204、すなわち StyleGAN をさらに訓練またはファインチューニングして、より正確な画像を生成する。

図 3 はパイプライン 300 を示す。

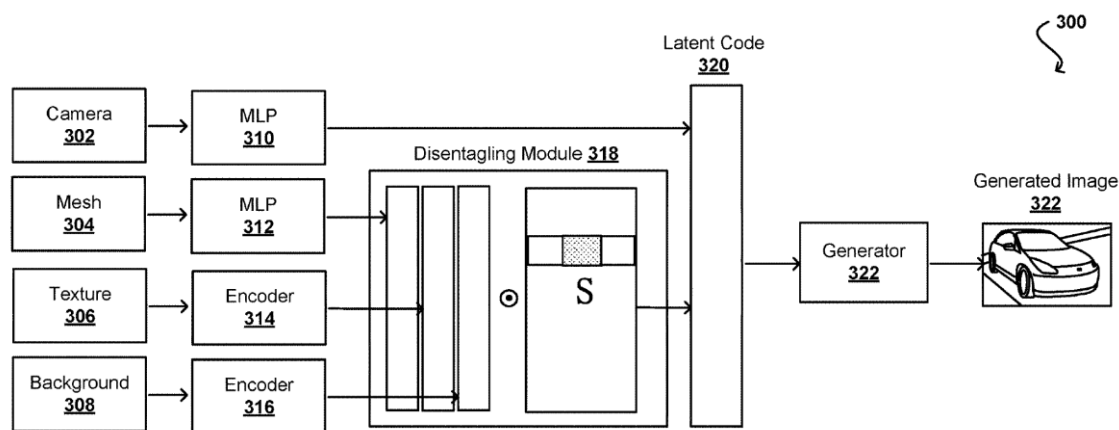


FIG. 3

図 2 で述べたように、画像内のオブジェクトに関する 3D 情報は、メッシュ 304、テクスチャ 306、および背景情報 308 を含んで推測することができる。また入力画像の 1 つまたは複数のアノテーションから抽出されたカメラ情報を利用することもできる。

この例では、メッシュ 304 は、機械学習プロセッサ (MLP) を使用して処理され、1 つまたは複数の次元または潜在的な特徴が抽出される。テクスチャ 306 および背景データ 308 は、関連する寸法または特徴を抽出するために、1 つまたは複数の畳み込みニューラルネットワーク (CNN) などのそれぞれのエンコーダ 314、316 で処理することができる。

次に、これらの特徴または寸法は、マッピングネットワークを含む解きほぐしモジュール 318 に渡される。マッピングネットワークは、出力画像 322 を生成するために、ジェネレータ 322 またはレンダラーに提供される単一の潜在コード 320 に様々なコー

ドの特徴をマッピングする。

潜在コードの一部はカメラ情報に対応し、残りはメッシュ、テクスチャ、および背景に対応する。潜在コード 320 を生成する前に、メッシュ、テクスチャ、および背景の 3 セットの特徴ではなく、1 セットの特徴にマージする試みを行う。選択行列 S を使用して、これらの特徴を選択し、カメラの特徴とともに潜在コード 320 にマージすることができる。この情報は、画像をレンダリングするために、StyleGAN などのジェネレータに提供される。

3. クレーム

952 特許のクレーム 1 は以下の通りである。

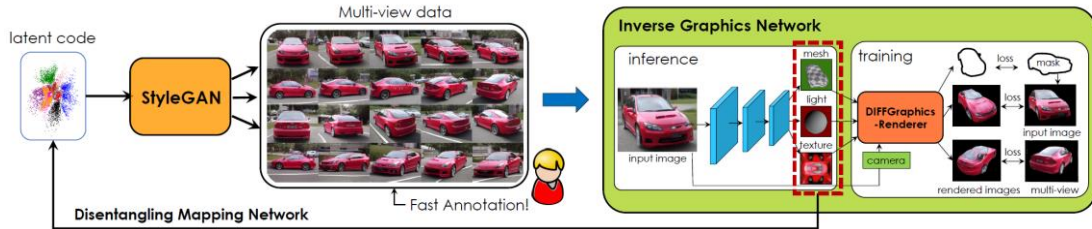
1. コンピュータで実装する方法において、
 - 生成ネットワークへの入力としてオブジェクトの 2 次元画像を提供し、
 - 生成ネットワークを使用して、様々なビューから表されたオブジェクトのビュー画像のセットを生成し、
 - 逆グラフィックスネットワークへの入力として、ビュー画像のセットと様々なビュー情報を提供し、
 - 逆グラフィックスネットワークを使用してセットの個々のビュー画像について、三次元情報のセットを決定し、
 - セットの個々のビュー画像について、三次元情報のセットおよびそれぞれのビュー情報を使用してオブジェクトの表現をレンダリングし、
 - 少なくとも 1 つの損失値を決定するために、レンダリングされた表現を対応するデータと比較し、
 - 少なくとも 1 つの損失値に少なくとも部分的に基づいて、逆グラフィックスネットワークの 1 つまたは複数のネットワークパラメータを調整する。

4. 本特許に関する論文

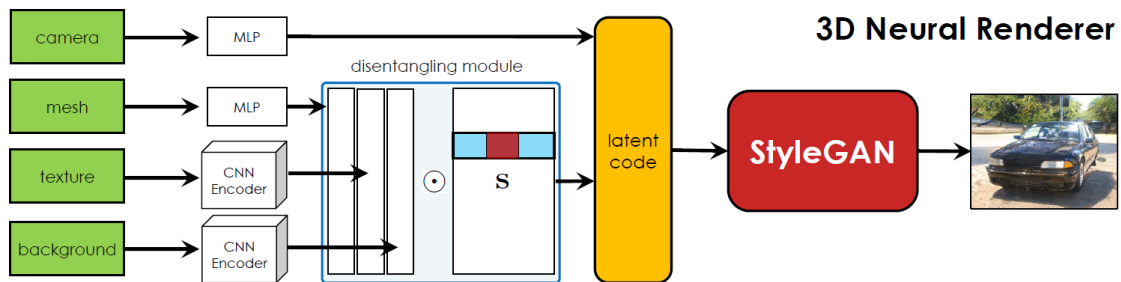
本特許に関する論文 “IMAGE GANS MEET DIFFERENTIABLE RENDERING FOR INVERSE GRAPHICS AND INTERPRETABLE 3D NEURAL RENDERING” が、Yuxuan Zhang 氏らにより公表されている¹。

¹ Yuxuan Zhang, Wenzheng Chen, Huan Ling, Jun Gao, Yinan Zhang, Antonio Torralba, Sanja Fidler “IMAGE GANS MEET DIFFERENTIABLE RENDERING FOR INVERSE GRAPHICS AND INTERPRETABLE 3D NEURAL RENDERING” arXiv:2010.09125v2 [cs.CV] 20 Apr 2021

下記図 1 はパイプラインを示す説明図である。

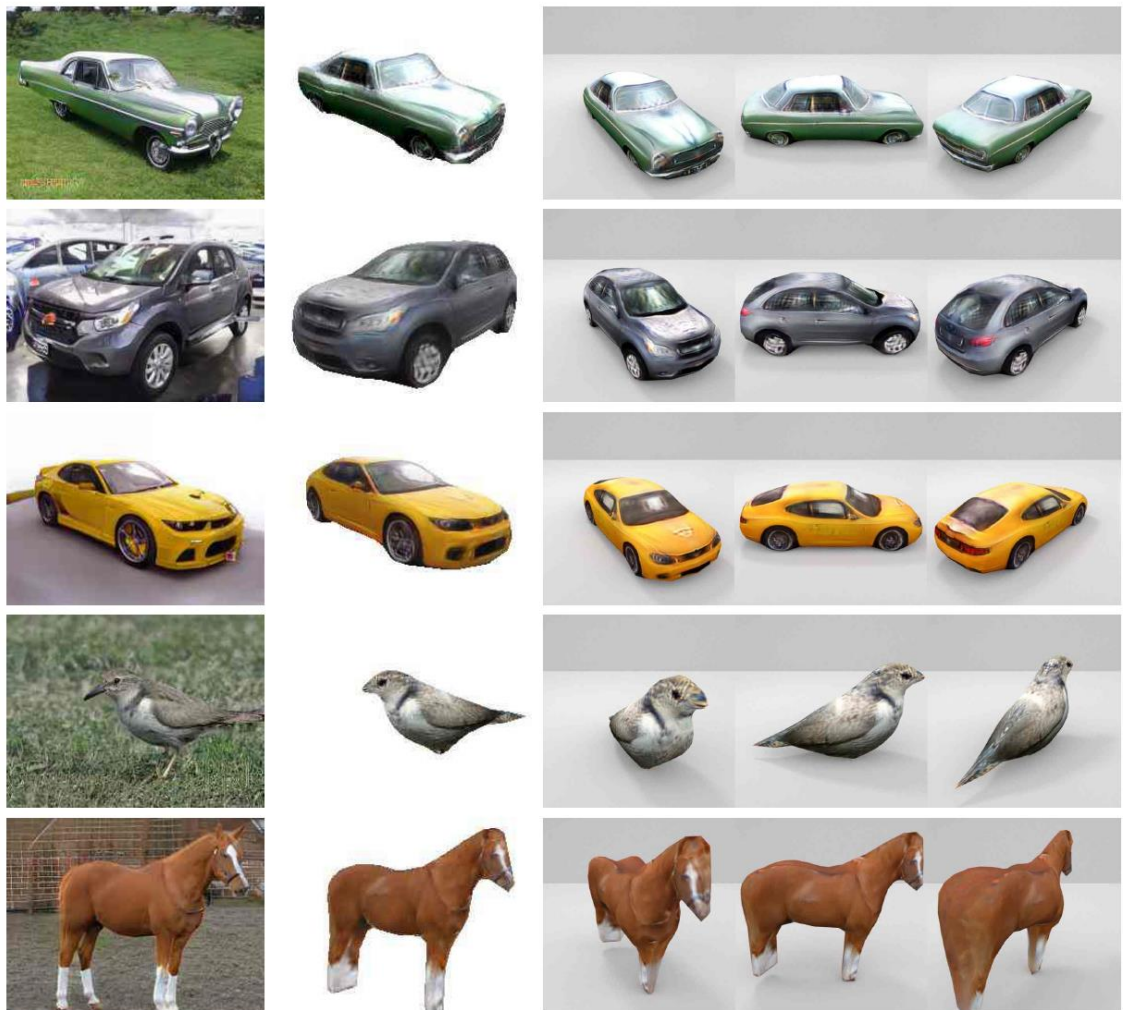


本論文では、2つの「レンダラー」を採用しており、前段に GAN (StyleGAN) が設けられ、その後段に微分可能なグラフィックスレンダラー (DIB-R) が設けられている。StyleGAN を合成データジェネレータとして活用し、このデータに効率的にラベル付けを行う。ラベル付けされた「データセット」は、画像から 3D プロパティ (メッシュ、照明、テクスチャ) を予測する逆グラフィックスネットワークをトレーニングするために使用される。このネットワークを使用して、下記図 2 に示すマッピングネットワークを通じて StyleGAN の潜在的コードを解きほぐす。



マッピングネットワークは、カメラ、メッシュ、テクスチャ、背景を、「レンダリング」するために、解きほぐされたコードにマッピングした上で、StyleGAN に出力する。

下記図に 2D 画像から 3D 画像への生成例を示す。



Input

Prediction

Multiple Views

入力画像（1 列目）が与えられた場合、3D 形状、テクスチャを予測し、それらを同じ視点（2 列目）にレンダリングする。また、3D 品質を示すために、残りの列に他の 3 つのビューでレンダリングを表示する。さまざまな形状、テクスチャ、視点で車を再構築することができる。また、より硬い（関節のある）オブジェクト、つまり鳥や馬にも同じアプローチで結果を得ることができる。

以上

著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI 研究所 AI コ

ース修了。

[AI 特許コンサルティング](#)、[医療 AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門 2.0](#)」、「[ブロックチェーン 3.0\(共著\)](#)」がある。