

AI 特許紹介(41)

AI 特許を学ぶ！究める！

～ディープリフレクタンスフィールド特許～

2022年6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

特許権者 Google

出願日 2019年10月16日

登録日 2021年5月4日

登録番号 US10997457

発明の名称 予測されたディープリフレクタンスフィールドを使用して画像を再照明するための方法、システム、およびメディア

457 特許は、対象物の画像及び当該対象物に照射する光の照射方向の情報を入力した場合に、照射方向から光が照射されたかの如く陰影が形成された対象物画像を生成するディープリフレクタンスフィールド特許に関する。

2.特許内容の説明

人間の顔の写実的な再照明は、視覚効果からメタバース空間に至るまで、多くの用途で必要とされている。しかしながら、ユーザーは、実際の顔と合成レンダリングとを容易に区別することができる。人間の顔を写真のようにリアルに再照明することは、表面

下の散乱、血流のモデリング、光と個々の髪の毛との相互作用の推定など、多くの困難を伴う課題であり、その結果、計算量が多くなる。さらに、従来の多くの方法は、大量の入力画像のセットを必要とするため、動く被写体または動的な被写体での使用には適していないという問題がある。

そこで、457 特許は、予測されたディープリフレクタンスフィールドを使用して画像を再照明するための方法を提供する。

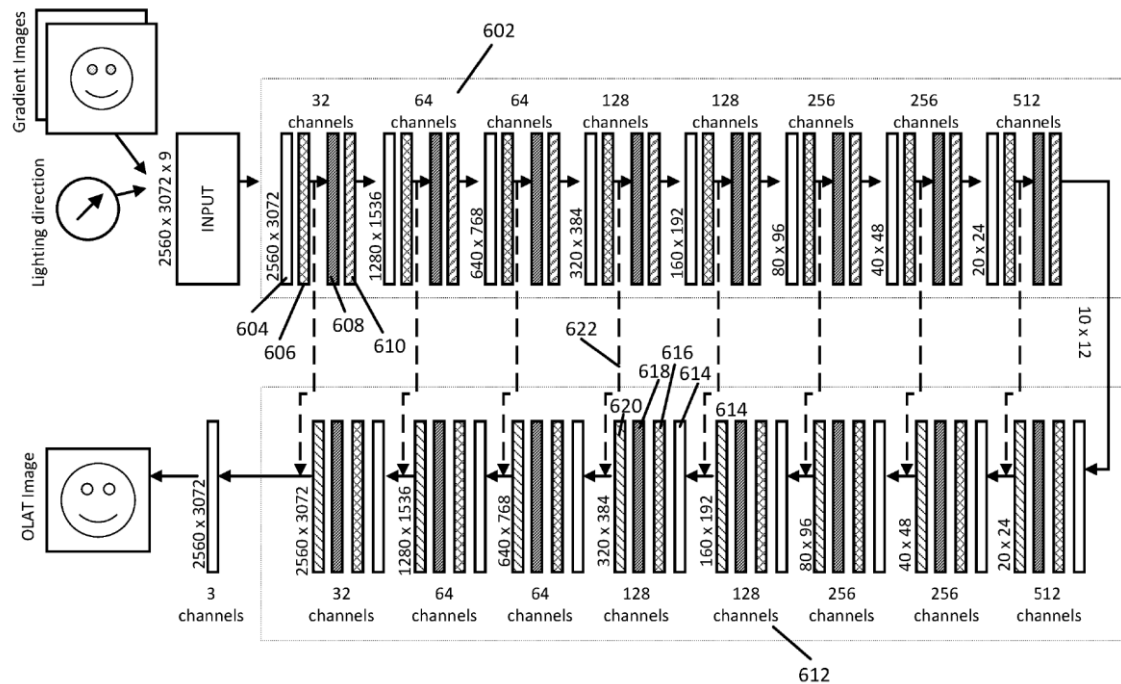
457 特許は、球面カラーグラデーション画像を使用して完全な 4D 反射フィールドを学習するように訓練された訓練済みニューラルネットワークを利用し、対象の球面カラーグラデーション画像を使用して対象の再照明画像を生成する。特に、ニューラルネットワークは、すべてアクティブ化された照明構造のライトのグループでキャプチャされた球状カラーグラデーション画像と、照明構造の各ライトが個別にアクティブ化された状態でキャプチャされる一度に 1 つのライト (OLAT : one-light-at-a-time) 画像のグループの両方を使用してトレーニングされる。

次に、ニューラルネットワークをトレーニングし、球面カラーグラデーション画像を使用して特定の照明方向から放射された光で特定の OLAT 画像を再構成する。

トレーニング後、ニューラルネットワークを使用して、球面カラーグラデーション画像のみを使用して被写体の再照明画像を生成する。

つまり、球形カラーグラデーション画像と OLAT 画像を使用してトレーニング中に特定の照明方向への OLAT 画像のマッピングを学習することにより、ニューラルネットワークは球形カラーグラデーション画像のみを使用して再照明画像を生成できる。

図 6 は、ディープリフレクタンスフィールドを使用して再照明画像を生成するために使用できる U-Net アーキテクチャ 600 を示す。



アーキテクチャ 600 は、図 6 に示されるように、入力として一連の画像および照明方向を受け取る。例えば、入力は、球形のカラーグラデーション画像のグループおよび OLAT 画像のグループを含む。各画像は、任意の適切なサイズ、 $W \times H$ である。照明方向は、RGB 値として球面カラーグラデーション画像の各ピクセルに追加される。

例えば 2 つの球形カラーグラデーション画像が使用される場合、2 つの球形カラーグラデーション画像を照明方向と組み合わせて、サイズ $W \times H \times 9$ のテンソルを生成する。つまり、各球形カラーグラデーションの各ピクセルは、そのピクセルの RGB 値に関連付けることができ、2 つの球形カラーグラデーション画像のサイズはそれぞれ $W \times H \times 3$ となる。また照明方向は、球面カラーグラデーション画像の各ピクセルに追加され、照明方向自体は、各ピクセルの RGB チャンネル値として指定される。トレーニングされたニューラルネットワークを使用する場合のテスト画像の場合、画像はそれぞれ第 1 の幅と高さ（たとえば、図 6 に示すように 2560 x 3072 ピクセル）にすることができる。

図 6 に示されるように、アーキテクチャ 600 は、エンコーダパス 602 を有する。エンコーダパス 602 は、一連のエンコーダ層を含むことができる。例えば、アーキテクチャ 600 は、エンコーダパス 602 に 8 つのエンコーダ層を示している。各エンコーダ層は、入力として画像を取得し、画像の特徴マップを生成する。生成された特徴マップは、次に、次のエンコーダ層に供給される入力画像となる。

入力画像の生成された各特徴マップは、エンコーダ層への入力に対してより小さな幅および高さを有することができるが、エンコーダ層への入力に対してより多くのチャンネル（例えば、より大きな深さ）を有する。たとえば、図 6 に示すように、第 1 のエンコーダ層はサイズ 2560 x 3072 の入力画像を取得し、サイズ 1280x1536 の特徴マップを生成する。

アーキテクチャ 600 は、デコーダパス 612 を有する。デコーダパス 612 は、一連のデコーダ層を含む。デコーダパス 612 に含まれるデコーダ層の数は、エンコーダパス 602 に含まれるエンコーダ層の数と同じである。例えば、図 6 に示されるように、デコーダパス 612 に含まれるデコーダ層は 8 つある。各デコーダ層は、入力として特徴マップを受信し、入力特徴マップのサイズよりも大きいサイズの出力画像を生成する。例えば、図 6 に示されるように、デコーダ層は、入力として、サイズ 1280×1536 の特徴マップを受信し、出力として、サイズ 256×3072 の画像を生成する。

入力画像のグループを一連のエンコーダ層に通すことにより、ニューラルネットワークは画像で表される特徴を学習でき、特徴マップを一連のデコーダ層に通すことにより、ニューラルネットワークは特徴マップを使用して画像を再構成することを学習する。

エンコーダパス 602 内の各エンコーダ層は、畳み込みブロック 604、整流器ブロック 606、正規化ブロック 608、および最大プーリングブロック 610 を有する。畳み込みブロック 604 は、エンコーダ層の入力での画像への任意の適切なサイズ（例えば、3×3、および/または任意の他の適切なサイズ）の任意の適切な畳み込みカーネルの適用を含むことができる。

損失関数が最適化されるため、各畳み込みブロックに適用される畳み込みカーネルは、ニューラルネットワークのトレーニング中に変更することができる。整流器ブロック 606 は、畳み込みブロック 604 の出力への任意の適切な整流器機能または活性化関数の適用を含む。ソフトプラス整流器、半波整流器、ロジスティック整流器、双曲線正接活性化関数、整流線形ユニット（ReLU）活性化関数、Leaky ReLU 活性化関数など、任意の適切なタイプの整流器または活性化関数を使用することができる。

正規化ブロック 608 は、整流器ブロック 606 の出力に対して任意の適切な正規化を実行することができる。正規化ブロック 608 は、ピクセル値を任意の適切な範囲内に正規化する。最大プーリングブロック 610 は、正規化ブロック 608 の出力で画像内の値をプールして、画像をダウンサンプリングする。

デコーダパス 612 のデコーダ層は、畳み込みブロック 614、整流器ブロック 616、正規化ブロック 618、およびアップサンプリングブロック 620 を含む。畳み込みブロック 614、整流器ブロック 616、および正規化ブロック 618 は、それぞれ、畳み込みブロック 604、整流器ブロック 606、および正規化ブロック 610 と同様である。

アップサンプリングブロック 620 は、特徴マップを入力として受け取り、入力特徴マップのサイズに比べてより大きなサイズの画像を生成する。例えば、図 6 に示されるように、アップサンプリングブロック 620 は、入力としてサイズ 160×192 の特徴マップを受け取り、出力として、サイズ 320×384 の画像を生成する。アップサンプリングブロック 620 は、アップサンプリング技術を使用して、アップサンプリングを実行する。例えば、アップサンプリングブロック 620 は、バイリニアアップサンプリングを使用することができる。その他、転置畳み込み（たとえば、2 x 2 畳み込みカーネルを使用）、アンパーリング、補間（たとえば、バイリニア補間、三次補間、最近隣法）など、他の適切なアップサンプリング手法を使用できる。

図 6 に示すように、アップサンプリングブロック 620 の出力は、スキップ接続 622 を使用して、同じ深さのエンコーダ層からの特徴マップと連結される。特徴マップをエンコーダ層から対応する深さのデコーダ層に伝播することにより、デコーダ層は、対応する深さでエンコーダ層によって生成された特徴マップを使用することができる。生成された出力画像およびグラウンドトゥルース OLAT 画像に基づいて損失を計算する。

3.クレーム

457 特許のクレーム 1 は以下の通りである。

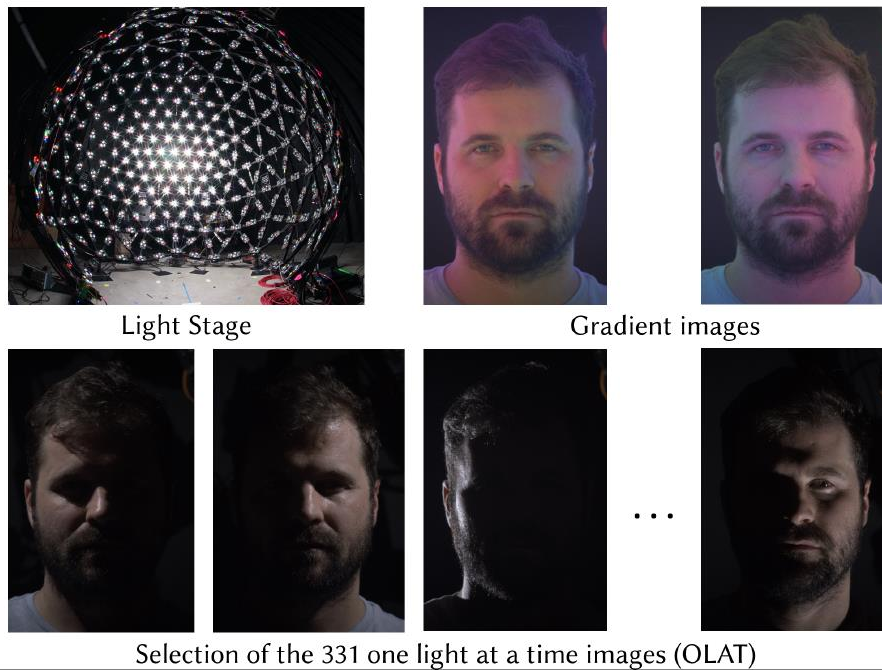
1. ディーププリフレクタンسフィールドを使用して画像を再照明する方法において、
トレーニングサンプルのグループを特定し、各トレーニングサンプルは、(i) 照明構造に配置された複数のライトの 1 つのライトがアクティブ化されたときにそれぞれがキャプチャされた一度に 1 つのライト(one-light-at-a-time (OLAT))の画像のグループ、(ii) 照明構造に配置された複数のライトがアクティブ化され、それぞれが特定の色を発生するときそれぞれがキャプチャされた球面カラーグラデーション画像のグループ、及び(iii)照明の方向を含み、OLAT 画像のグループ内の各画像と各球面カラーグラデーション画像は、被写体の画像であり、照明の方向は、被写体に対する光の相対的な向きを示し、
トレーニングサンプルのグループを使用して畳み込みニューラルネットワークをトレーニングし、

畳み込みニューラルネットワークのトレーニングには以下が含まれ、
一連のトレーニング反復の各トレーニング反復、およびトレーニングサンプルのグループの各トレーニングサンプルに関し、
トレーニングサンプルに関連付けられた照明方向からの照明を伴うトレーニングサンプルに関連付けられた対象の表現である出力予測画像を生成し、
トレーニングサンプルの照明方向に対応するトレーニングサンプルの OLAT 画像のグループに含まれるグラウンドトゥルース OLAT 画像を特定し、
出力予測画像と特定されたグラウンドトゥルース OLAT 画像間の知覚的相違を示す損失を計算し、
計算された損失に基づいて畳み込みニューラルネットワークのパラメータを更新し、
球面カラーグラデーション画像の第 2 のグループと第 2 の照明方向を含むテストサンプルを特定し、
訓練された畳み込みニューラルネットワークを使用して、第 2 の照明方向からの照明で球面カラーグラデーション画像の第 2 のグループのそれぞれに含まれる被写体の再照明画像を生成する。

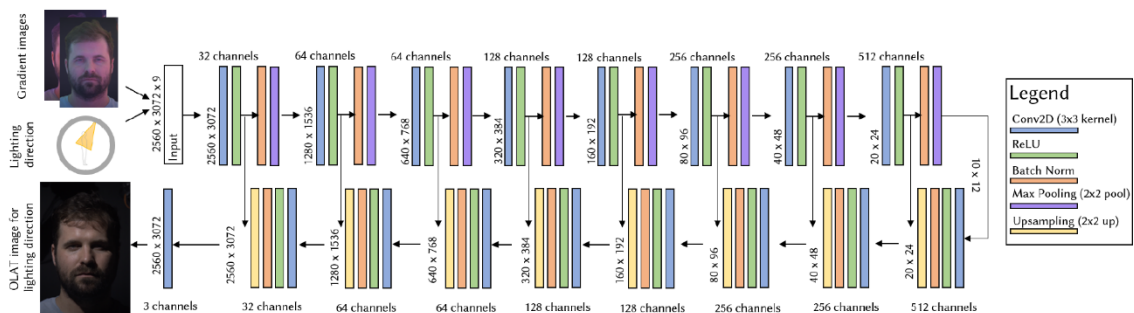
4. 本特許に関連する論文

本特許に関する論文 “Deep Reflectance Fields”¹が、ABHIMITRAMEKA 氏らにより公表されている。下記図は実験に用いた照明構造、RGB カラーグラデーション画像及び OLAT 画像を示す説明図である。

¹ ABHIMITRA MEKA, CHRISTIAN HÄNE, ROHIT PANDEY, MICHAEL ZOLLHÖFER, SEAN FANELLO, GRAHAM FYFFE, ADARSH KOWDLE, XUEMING YU, JAY BUSCH, JASON DOURGARIAN, PETER DENNY, SOFIEN BOUAZIZ, PETER LINCOLN, MATT WHALEN, GEOFF HARVEY, JONATHAN TAYLOR, SHAHRAM IZADI, ANDREA TAGLIASACCHI, and PAUL DEBEVEC, CHRISTIAN THEOBALT, JULIEN VALENTIN and CHRISTOPH RHEMANN, “Deep Reflectance Fields” ACM Trans. Graph., Vol. 38, No. 4, Article 77.

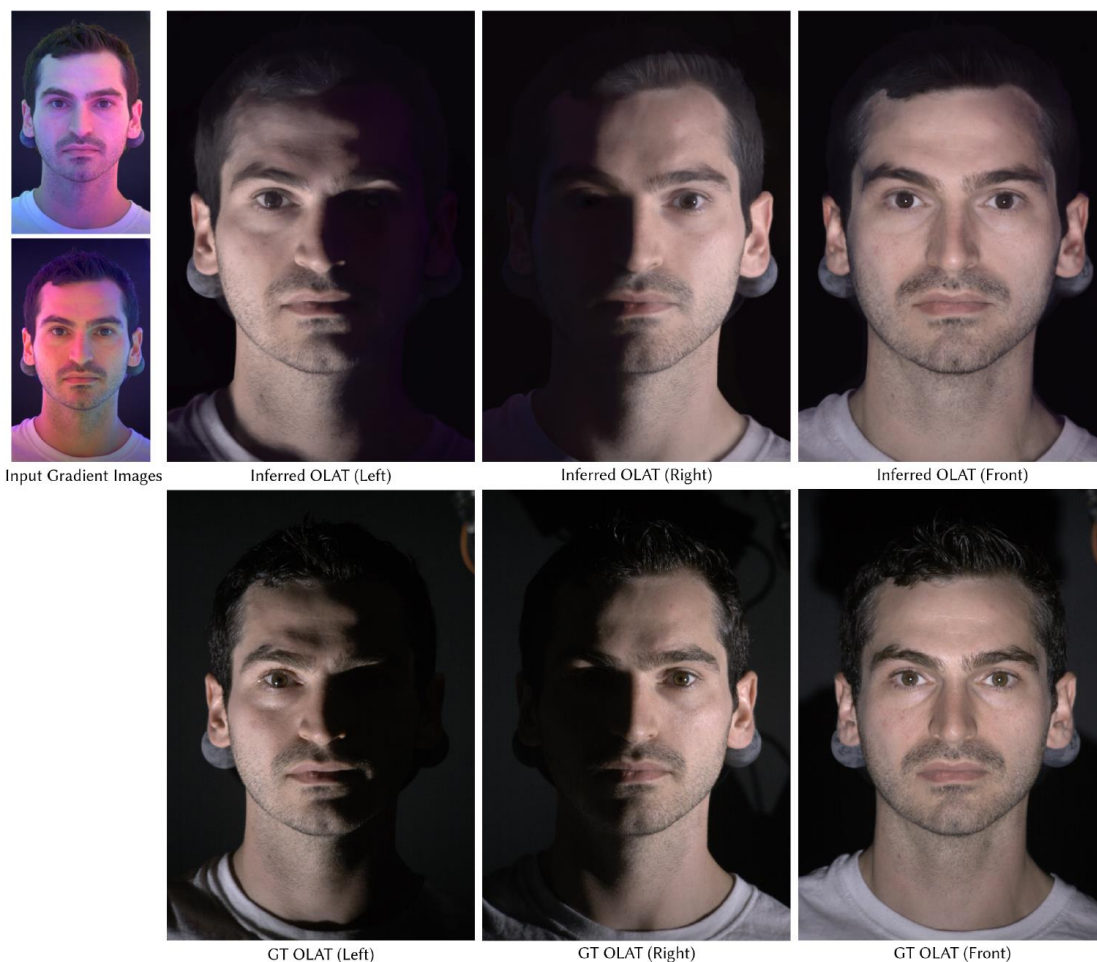


測地線ドームに取り付けられたプログラム可能な光源を使用して、トレーニング、推論、検証データ用の RGB カラーグラデーション画像と OLAT 画像の下で被写体を照らす。下記図は、ネットワーク構成図である。



本特許と同じく U-Net が用いられ、入力として 1 対のグラデーション画像と照明方向を受け取る。U-Net アーキテクチャを介して、その特定の照明構成に対応する OLAT イメージを回帰する。

下記図は、定性的結果を示す説明図である。左上側のグラデーション入力画像に対し、上側に推定された OLAT 画像を示し、下側に対比のためグラウンドトゥールズ画像を示す。



各照明方向に対する画像が適切に生成できていることが理解できる。鏡面反射性、影、肌の質感などの高周波の詳細がグラデーション画像から正しく抽出されている。

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。