

令和 3 年 5 月 14 日現在

機関番号：12601

研究種目：研究活動スタート支援

研究期間：2019～2020

課題番号：19K24354

研究課題名（和文）3次元強化学習による新たなデザインの理解と生成

研究課題名（英文）Understanding and Generation of Designs using 3D Reinforcement Learning

研究代表者

古田 諒佑（Furuta, Ryosuke）

東京大学・生産技術研究所・助教

研究者番号：20843535

交付決定額（研究期間全体）：（直接経費） 2,200,000円

研究成果の概要（和文）：2019年度は、本研究のベースとなる画素強化学習の発展に主に取り組んだ。saliency-driven image enhancementという新たな応用を提案した。マルチメディアの分野において最難関国際論文誌であるIEEE Transactions on Multimediaに投稿し採択された。3次元的な形状を扱うために、2020年度は3次元形状推定に必要な技術である特徴点マッチングの教師なし学習に取り組んだ。本研究の成果はパターン認識の旗艦国際会議であるInternational Conference on Pattern Recognition (ICPR2020)に採択された。

研究成果の学術的意義や社会的意義

IEEE TMMに掲載された画素強化学習は、各画素でどのアクションが取られたかを可視化することができるため、どの画素値がどう変更されたかを人が見て理解することができる。そのため、ブラックボックスとしての画像処理技術の使用が敬遠されることの多い医療画像処理の応用も期待される。ICPRにて発表した教師なし特徴点マッチングは、深層学習に基づく特徴点マッチングによる3次元形状推定が、大量の正解付き学習データを必要とせず可能となるため、実応用上でボトルネックとなることのアノテーションの労力を必要としない。そのため実応用可能性の向上に貢献する。

研究成果の概要（英文）：In FY2019, I worked on reinforcement learning with pixel-wise rewards, which is the base of this study, and proposed its novel application: saliency-driven image enhancement. This work was accepted to IEEE Transactions on Multimedia (TMM), which is the top journal in the field of multimedia.

In order to deal with 3D shape, in FY2020, I worked on unsupervised learning of feature point matching, which is useful for 3D reconstruction. This work was accepted to International Conference on Pattern Recognition (ICPR), which is a flagship conference in the field of pattern recognition.

研究分野：コンピュータビジョン

キーワード：強化学習 教師なし学習 特徴点マッチング

1. 研究開始当初の背景

人間は昔に作られた車や建物、洋服などを目にした際に、(汚れや劣化具合からではなく)デザインから“なんとなくの古さ”を感じることができる。また時が経っても同じデザイナーやブランドに共通するデザイン(例えば、昔から変わらないフェラーリっぽさ)をなんとなく感じることができるだろう。しかし、人間がデザインに対して奥底に持つ主観的な感覚は数学的、または客観的に記述することができない。また、プロのデザイナーでもない限り、フェラーリっぽい車をデザインすると言われてもそれは容易ではない。そこで本研究では、人間が3次元的设计に対して抱く感覚をニューラルネットワークに理解させ、新たなデザインを創造させることを目的とする。本研究は、3次元的设计全てを対象としているため、車や建築、服のデザインなどあらゆる物の創作に利用することができ、世の中に非常に大きな影響を与えると考えられる。

2. 研究の目的

申請当初の検討としては、3次元的设计に対する感覚を人工知能(ニューラルネット)に理解させ、新たなデザインを創造させることを目的としていた。一般的に、新しいデザインの創造は難しいとされるため、人工知能にデザインを理解・創造させることができれば、人々の創作活動支援に有用である。また、囲碁や将棋、ゲームなどでトッププロのプレーヤーを超える性能を見せ、人間が思いもよらないような手を用いることもある強化学習を利用してデザインを創造させることで、創作活動支援のみにとどまらず、デザイナーも想像もしないような全く新しい芸術性を生み出す可能性も秘めている。また、そのために3次元強化学習という新たな学術的テーマに取り組む。

3. 研究の方法

まず提案手法である3次元強化学習を定式化し実験によりその性能を確認する。申請時に検討していた提案手法の流れは以下ようになる。まず事前学習としてWebから集めた画像を使い識別ネットワークを学習させる(例えば様々な年代の車の画像を集め、入力画像の車の年代を識別できるようにネットワークを学習させる)。次にベースとなる3Dモデルを入力とし、それを変形させる3次元アクション(頂点の移動、テクスチャ、色の変更)を出力するネットワーク(変形ネットワーク)を強化学習により学習させる。このとき、変形させた3Dモデルをレンダリングした画像を識別ネットワークに入力し、その識別スコアを報酬とする(例えばベースとなる車の3Dモデルを1980年代風のデザインにしたい場合、識別スコアの1980年代の部分を報酬とし、そのスコアが高くなるように変形ネットワークを強化学習する)。他にも、例えばベースとなる車の3Dモデルをフェラーリ風のデザインにしたい場合、識別ネットワークにはフェラーリの画像とそれ以外の車の画像の判別を学習させ、フェラーリとしての識別スコアが高くなるように変形ネットワークを強化学習する。

4. 研究成果

2019年度は、本研究のベースとなる画素強化学習の発展に主に取り組んだ。画素強化学習は私が以前に提案した、2次元の画像の各画素にエージェントを配置し、それぞれのエージェントが局所的なアクションを取ることで様々な画像処理(ノイズ除去や画像復元、色調変換)を行うことができる手法である。画素強化学習の技術といくつかの応用は申請書でも述べたように既に発表済みであったが、2019年度はこれを発展させ、saliency-driven image enhancementという新たな応用を提案した。saliency-driven image enhancementは、人が画像のどこに注目するかを表すsaliencyという指標を基に画像の変換を行う技術である。提案手法を用いると、画像中の注目させたい部分や注目させたくない領域を指定することで、それらの領域の見た目を自然に強調や抑制することができる。saliencyの指標の計算式は、これまで様々な研究で提案されてきたが、それらの多くは微分不可能なものであり、深層学習でsaliencyに基づいて画像の強調や抑制を行う手法はこれまで提案されていなかった。そこで私は微分不可能な目的関数を扱うことができる画素強化学習を用いることでこの問題を解決した。実験では、これまでに提案されていたhand-craftedな手法と比較して、より自然なだけでなく効率的に強調や抑制が行えることを示した。図1はSaliency-driven image enhancementにおける結果の比較である。左から強調したい領域を表した入力マスク、入力画像、比較手法(OHA, HAG, WSR, SDIM)の強調結果、提案手法の強調結果である。t_maxは提案手法の取るアクションの回数を表す。図1から、提案手法は他の手法と比較して自然な協調ができており、またアクションの回数を増やすごとに強調度合いを強めていることがわかる。3次元的设计ではなく2次元的设计である画像が対象ではあるが、どのように変換すれば指定領域に人の注目を引き付けられるかをエージェントが学習できたと捉えることができる。2019年度は、この内容を含めた画素強化学習の論文をマルチメディアの分野において最難関国際論文誌であるIEEE Transactions on Multimediaに投稿し採択された。

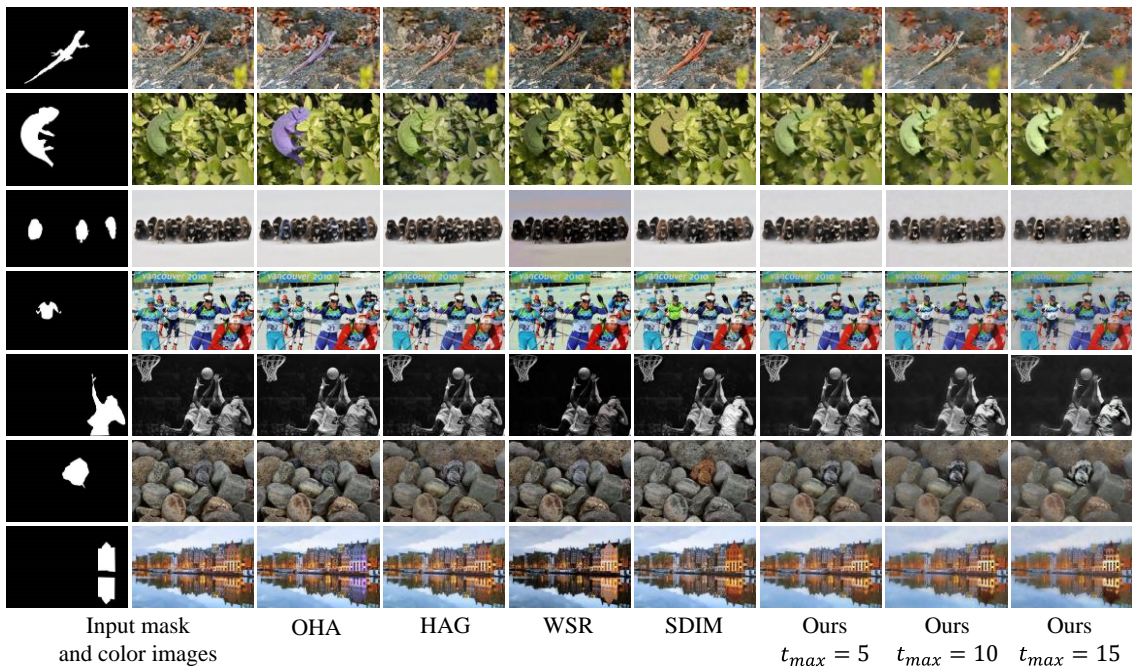


図1 Saliency-driven image enhancement の結果の比較

申請当初は 3 次元物体のデザインの変形を検討していたが、強化学習による形状変形は技術的に困難であったため、2020 年度は方針を転換し、その 1 つ前のステップである 3 次元形状の取得に注力した。具体的には、3 次元形状推定に必要な技術である特徴点マッチングの教師なし学習に取り組んだ。画像から 3 次元形状推定を行う方法の 1 つとして、対象物体を複数カメラで撮影した画像間で対応点を見つける多視点ステレオ法がある。対応点を見つけるためには特徴的な点を特徴量ベクトルを用いて記述する必要があるが、従来は hand-crafted 特徴量が用いられることが多かった。近年は高精度は対応点探索のために深層学習を用いた特徴量ベクトルの記述手法も提案されているが、大量の教師データを集める必要があることが課題である。そこで本研究では、cycle-consistency を考慮することで教師データ無しで深層学習における対応点探索のための特徴量ベクトルの学習手法を提案した。例えば入力画像が 2 枚の場合、画像 A 上の点から画像 B 上の点に対して対応点を探索し、また画像 B 上から画像 A 上に対応点を探索した場合、対応をたどると画像 A の元の点に戻るべきだという性質を利用した学習方法である。実験では 2 枚の画像から奥行きを推定するステレオマッチングに応用し、制度の面で hand-crafted な特徴量を大きく上回る性能を示し、2 枚の画像のスペクトルが異なる場合（クロススペクトルマッチング）においても優れた性能を発揮することを確認した。図 2 は 2 枚の入力ステレオ画像のうち、片方が R チャンネル、もう一方が GB チャンネルしか値を持たないアナグリフのクロススペクトルステレオマッチングにおける奥行き推定結果の比較である。図 2 から、提案手法はクロススペクトルステレオマッチングにおいて比較手法よりも優れた結果を取得していることがわかる。一般に、RGB 画像と赤外画像のようなクロススペクトルのステレオ画像において、正解データを大量に取得することは難しい。

本研究の成果はパターン認識の旗艦国際会議である International Conference on Pattern Recognition (ICPR2020) に採択された。

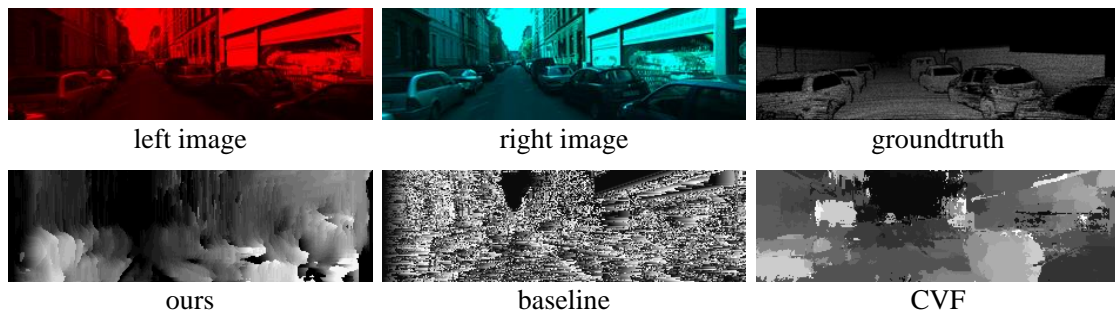


図2 アナグリフステレオにおける奥行き推定結果の比較

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 0件）

1. 著者名 Ryosuke Furuta, Naoto Inoue, and Toshihiko Yamasaki	4. 巻 -
2. 論文標題 PixeIRL: Fully Convolutional Network with Reinforcement Learning for Image Processing	5. 発行年 2019年
3. 雑誌名 IEEE Transactions on Multimedia	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1109/TMM.2019.2960636	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件/うち国際学会 1件）

1. 発表者名 古田諒佑, 井上直人, 山崎俊彦
2. 発表標題 [特別講演] 近傍を考慮した画素ラベリング
3. 学会等名 画像工学研究会
4. 発表年 2020年

1. 発表者名 野口尚晃, 古田諒佑, Xueting Wang, 山崎俊彦, 相澤清晴
2. 発表標題 Cycle Consistencyを用いた画像間の特徴点マッチング
3. 学会等名 パターン認識・メディア理解研究会 (PRMU)
4. 発表年 2020年

1. 発表者名 Ryosuke Furuta, Naoaki Noguchi, Xueting Wang, and Toshihiko Yamasaki
2. 発表標題 Feature Point Matching in Cross-Spectral Images with Cycle Consistency Learning
3. 学会等名 25th International Conference on Pattern Recognition (ICPR2020) (国際学会)
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
--	---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------