

令和 4 年 5 月 26 日現在

機関番号：14401

研究種目：基盤研究(C)（一般）

研究期間：2019～2021

課題番号：19K12019

研究課題名（和文）非視覚センサ情報を用いたリアルタイムな屋外画像の生成

研究課題名（英文）Real-Time Outdoor Image Generation using Non-visual Sensing Information

研究代表者

新田 直子 (Naoko, Nitta)

大阪大学・工学研究科・准教授

研究者番号：00379132

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究は、任意の地点において撮影された1枚の屋外画像を、非視覚センサである気象センサにより得られる信号値に応じた気象状況を表す画像に変換することにより、任意地点、時間におけるカメラ画像を疑似的に生成する技術の開発を目的とする。これを実現するため、少量の整合性の高い気象センサ信号値と屋外画像の対データの収集方法、及び大量の屋外画像や服装情報が付与された人物画像など他のより入手が容易な画像データセットを補完的に用いた、空模様や地面の様子を含む遠景画像、及び人物を含む近景画像を対象とした画像変換器の学習方法を提案し、その有効性の実験的検証を行った。

研究成果の学術的意義や社会的意義

本研究は、近年盛んに研究されている敵対的生成ネットワークを、視覚センサと非視覚センサという2つの異なる特性を持ったセンサから得られる生の観測情報に適用することにより、疑似的な視覚センサによる観測情報を生成する試みである。特に、ネットワークの学習データとして、誤った情報となるノイズを多く含むデータしか得られないという問題に対し、同じ環境を観測するセンサの観測情報に存在すべき整合性などを考慮した上で、他のより入手が容易なデータを補完的に用いることにより、学習が適切に行われることを実験的に検証し、画像生成技術のセンシング応用への可能性を示した。

研究成果の概要（英文）：The goal of this research is to develop a technique for generating pseudo outdoor images captured at arbitrary times and locations by transforming an outdoor image captured at a specific location according to the non-visual signals captured by a weather sensor. To achieve this, we have proposed a method for collecting a small number of consistent pairs of outdoor images and weather sensor signals, and methods for training networks for transforming a distant view containing sky and ground and a close-to-mid range view containing persons by utilizing more easily available large-scale datasets of outdoor images or fashion images.

研究分野：知覚情報処理

キーワード：画像変換 画像生成 非視覚センサ 屋外環境センシング 深層学習

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

近年、様々なセンサにより観測された実世界のリアルタイムな情報がインターネットを通して世界中の人々に提供されており、人々は遠隔地の現状を容易に知ることができる。提供される情報の代表例として、気象レーダや気象衛星により観測される、各地のリアルタイムな気温や雨量の分布、モバイル端末に搭載された GPS を用いて観測される世界中のユーザの位置情報に基づく、各地のリアルタイムな渋滞状況などが挙げられる。これらのセンサは世界中のあらゆる場所のリアルタイム情報を常時獲得できるが、各センサにより得られる情報は限定され、広域に対して俯瞰的な情報が提供されることが多い。

一方、センサの中でも特にライブカメラや個人が所有するモバイル端末に搭載されたカメラなど(以下、単にカメラと呼ぶ)により得られる映像や画像は、人間が普段目で見ると同様の視覚情報であるため、人間にとって直感的に分かりやすい形で、より局所的かつ多様な情報を提示できる。このため、各地の気象状況、河川や道路の状況、花の開花や紅葉状況を伝えるライブカメラや、各地のユーザが撮影した空の画像などを提供する様々なサービスが存在する。災害時にも、Twitter などのソーシャルネットワーキングサービス(SNS)には各地のユーザがスマートフォンなどで撮影した画像や動画が多数投稿され、リアルタイムな災害情報として活用されている。ライブカメラは設置された位置の情報を常時提示するが、画像が公開されているものは非常に少なく、空間的網羅性は低い。一方、SNS には世界中のユーザから投稿が集まるため、ライブカメラに比べ空間的網羅性は高いが、時間的には非常に局所的な情報しか得られない。

2. 研究の目的

カメラにより撮影される環境は、構造物など時間により変化しない静的な要素と、空模様、混雑状況、植物の生育状況など時間により変化する動的な要素により構成される。このうち静的な要素は、ある時点で撮影された画像が1枚あれば獲得でき、動的な要素は、カメラ以外の非視覚的センサから常時獲得され得る。そこで本研究は、任意の地点における画像が1枚入手できると仮定し、この画像が入力されたとき、カメラ以外の非視覚的なセンサにより観測された情報を用いて現在の状況を表す画像に変換することにより、任意地点、時間におけるカメラ画像を疑似的に生成する技術の開発を目的とする。

3. 研究の方法

機械学習手法の一つである深層学習の分野において、新しい画像を生成するモデルとして、GAN (Generative Adversarial Network: 敵対的生成ネットワーク) と呼ばれるネットワークが提案されて以降、実画像に近い画像を自動的に生成する技術が盛んに研究されている。GAN をベースに、文章や音など、与えられた条件に応じた画像を生成するメディア変換技術や、撮影された人物の年齢や、絵画を写真に変更するなど、入力された画像の構造を保持したまま属性やドメインを変換する画像変換技術も開発されている。GAN をベースとした画像変換技術では一般に、変換前と変換後の画像や、文章や音などの条件と画像などを学習用データセットとして予め大量に収集する必要がある。

本研究では、GAN をベースとした画像変換技術により非視覚センサが観測した信号値(以下、単にセンサ信号値と呼ぶ)に応じた画像変換を実現する。学習用データセットとしては、各種センサから自動的に集めたデータセットを用いるが、このとき、画像の属性を表すラベルに相当するセンサ信号値や画像の質、またそれらの整合性は保証されない。よって、センサから自動的に収集されたデータセットを用い、いかにセンサ信号値と画像間の関係を正しく学習するかという問題に取り組む。具体的な工程は、(1)センサ信号値と画像の収集、(2)画像変換モデルの学習、の2つとなる。

非視覚センサは、広域の情報を常時取得することが容易な気象センサとし、生成対象となる画像は気象センサが観測する屋外環境の画像とする。ここで、屋外環境の温度や雨量、雲量などは、例えば、景色であれば空模様や開花、紅葉情報、人物であれば服装情報など、画像の撮影対象に応じて異なる動的要素と関連し得る。よってまず、気象状況と画像の動的要素の関係が場所によらず類似したパターンを持ちやすいと考えられる、空模様や地面の様子を含む遠景の画像を対象に変換方法の具体化と実験的検証を進め、その後、人物や植物の状況が分かるような近景への拡張について検討する。

4. 研究成果

(1) センサ信号値と画像の収集

まず多様な時間、場所の遠景画像が投稿され得るソーシャルネットワーキングサービス(SNS)である Flickr (<https://www.flickr.com/>) から、2016、2017年の2年間にアメリカで撮影された画像を撮影時間、位置情報と共に収集した。このうち、既存のシーン認識器により屋外シーンと判定されたもの495,669枚を屋外画像として抽出した。これら各画像に対し、気象情報サービスである World Weather Online (<https://www.worldweatheronline.com/>) から、位置及び時間情報に基づき、同じ環境に対する気象センサ信号値を収集した。

ただし、画像を撮影したカメラや気象センサから得られる位置や時間情報の信頼性などの問題から、位置と時間に基づく対応付けのみでは、必ずしも同じ環境に対する画像と気象センサ信号値の対が得られない。そこで、約 18 万枚の屋外画像に対し、人手で天候クラスがラベル付けされた Image2weather データセット [1] を用いて、画像と気象センサ信号値それぞれに対する天候クラス推定器を学習し、収集した画像と気象センサ信号値それぞれに適用し、推定される天候クラスの整合性に基づき、信頼性の高い対を自動的に選別した。図 1 に示すように、気象センサから得られる天候クラスと整合性の低い画像が適切に除去され、126, 107 対の画像と気象センサ信号値が得られた。



図 1：気象センサ信号値との整合性を考慮して収集された屋外画像

(2) 画像変換モデルの学習

① 遠景画像を対象とした画像変換器の半教師あり学習

Step 1 学習データの収集

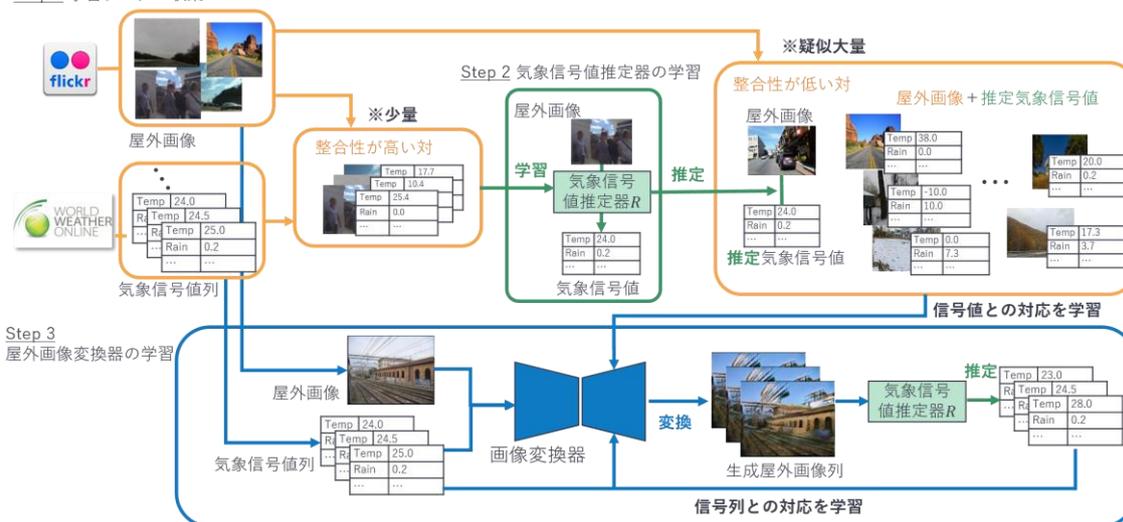


図 2：画像変換器の半教師あり学習手法の概要

(1) で収集した 126, 107 対の画像と気象センサ信号値は、画像変換器の学習には十分な量ではないが、画像から気象センサ信号値を推定する推定器の学習には十分と考えられる。そこで、図 2 に示すように、画像から気象センサ信号値を推定する推定器をまず学習し、この推定器を用いて、信頼性の低い対データに対しても画像から気象センサ信号値を推定することにより、疑似的に大量の対データを生成する。これを用いて、画像と気象センサ信号値を入力とし、条件となる信号値に応じて変換した画像を出力とする画像変換モデルを学習する半教師あり学習手法を考案した。

図 3 に示すように、条件を天候クラスとし、人手で収集した比較的少量の信頼性の高い対データである Image2Weather データセットを用いて教師あり学習した場合と比べ、疑似的な大量の

対データを用いた半教師あり学習を行った場合、より多様な環境を撮影した画像に対して晴れ (sunny)、雪 (snow) など条件として与えた天候クラスに応じた画像が生成できることを確認した。また、図 4 に温度 (tempC)、視程 (visibility)、雲量 (cloudcover) など各信号値を変化させたときの生成画像例を示す。温度が低くなると地面に雪、視程が低くなると全体に霧、雲量が高くなると空に雲といったように、各信号値に応じて、特定の領域に変化が生じることを確認した。さらに、時間的に連続した気象センサ信号値列を条件とすることにより、時系列的な気象変化を表す画像列が生成されることも確認した。

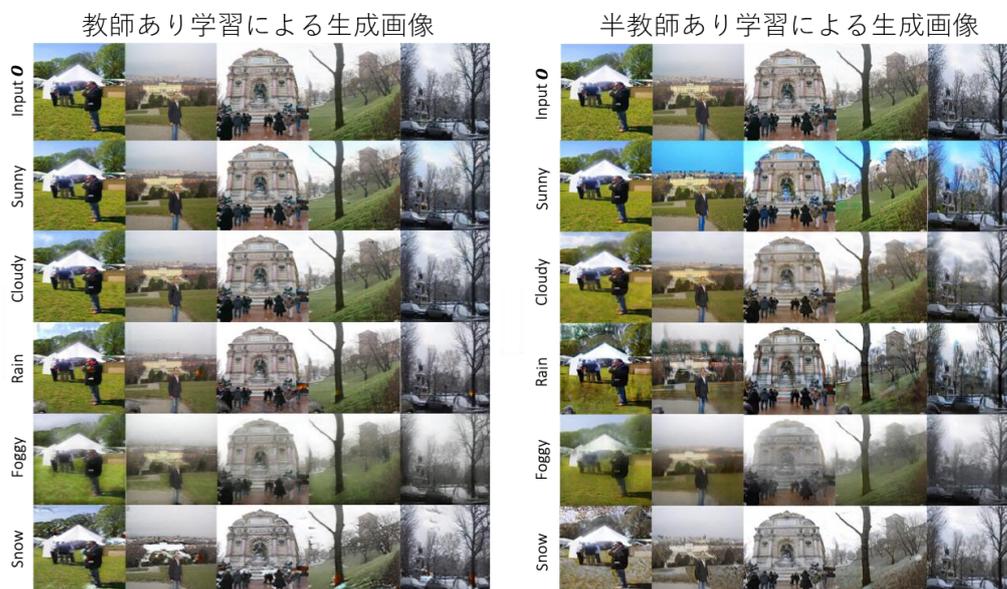


図 3：少量対データを用いた教師あり学習と大量疑似対データを用いた半教師あり学習の比較

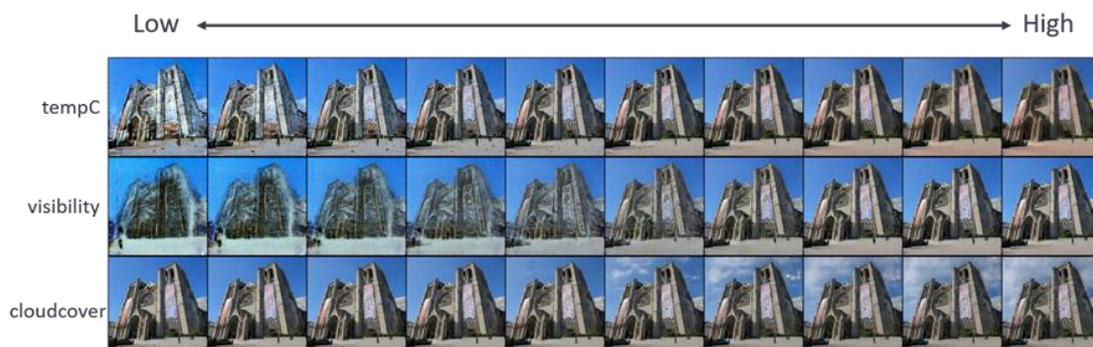


図 4：気象センサの各信号値を変化させたときの生成画像例

② 近景画像を対象とした人物の服装変換器の学習

近景画像を対象に、特に人物の服装に着目し、人物を含む屋外画像に対して気象センサ信号値を条件とした服装変換器への拡張方法を検討した。まず、(1)で収集し、(2)－①で疑似的な気象センサ信号値を付与した屋外画像のうち、既存の物体認識器により人物を含むと判定されたもののみを抽出した結果、24,571 対の屋外画像と気象センサ信号値の対となった。(2)－①の半教師あり学習を用いても、画像変換モデルの学習には不十分なデータ数である。そこで本研究は、気象状況に対し、人物の服装の種類が概ね一様に変化することに着目し、まず気象センサ信号値に応じた服装の種類を推定し、推定された服装の種類に応じて屋外画像中の人物の服装を変換するという 2 段階の処理を通して、気象センサ信号値に応じた人物の服装を変換する手法を提案した。図 5 に提案手法の概要を示す。これを実現するためまず、既存の服飾画像データセットである DeepFahion2[2] を利用し、気象状況と関係が深いと考えられる上衣の長袖 (long)、半袖 (short) に限定した服装の種類がラベル付けされた 111,317 枚の人物画像からなる人物画像データセットを用意した。これを用いて、人物画像から服装の種類を推定する服装推定器を学習し、収集した 24,571 枚の屋外画像に適用し、15,060 枚の服装の種類と気象センサ信号値の対に変換した。また、同じ人物画像データセットを用いて、入力された服装の種類に応じて人物画像の服装を変換する画像変換器を学習した。

この結果、図 6 に示すような変換が実現された。生成画像の質については改良の余地があるが、与えた気象信号値と変換画像に対する服装推定結果の一致度が、収集した気象センサ信号値と屋外画像の対に対する服装推定結果の一致度と同程度となったことより、気象状況に応じた画像に変換されたことを確認した。

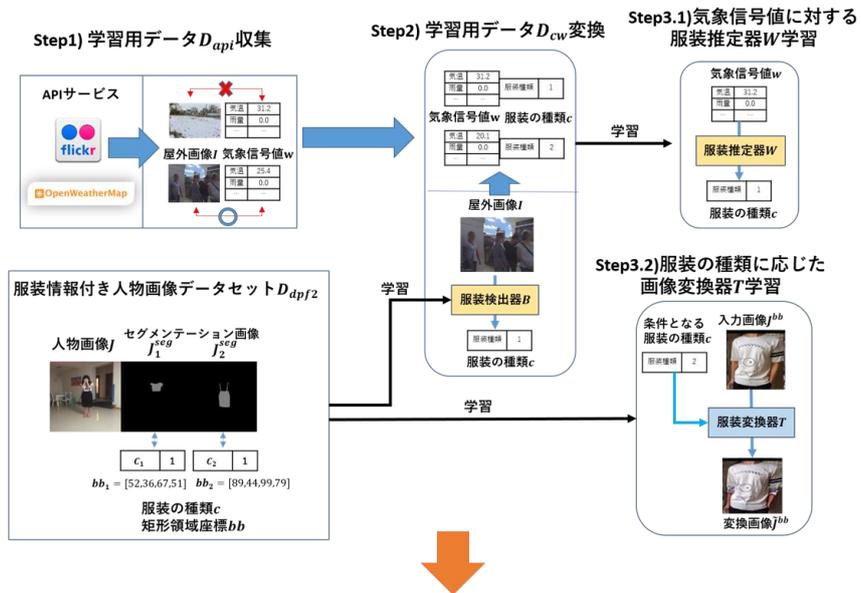


図 5 : 近景を対象とした人物の服装変換器の学習手法の概要



図 6 : 服装の種類に応じた画像変換の例

- [1] W.-T. Chu, X.-Y. Zheng, and D.-S. Ding, "Camera as Weather Sensor: Estimating Weather Information from Single Images," Journal of Visual Communications and Image Representation, vol.46, pp.233-249, 2017.
- [2] Y. Ge, R. Zhang, L. Wu, X. Wang, X. Tang, and P. Luo, "DeepFashion2: A Versatile Benchmark for Detection, Pose Estimation, Segmentation and Re-Identification of Clothing Images, CVPR, 2019.

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計8件（うち招待講演 1件 / うち国際学会 1件）

1. 発表者名 岡田 深、新田直子、中村和晃、馬場口登
2. 発表標題 気象センサ情報を用いた屋外画像における人物の服装変換
3. 学会等名 第14回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2022)
4. 発表年 2022年

1. 発表者名 内田祐生、新田直子、中村和晃、馬場口登
2. 発表標題 画像の印象操作のためのオブジェクトの外観変換
3. 学会等名 電子情報通信学会2021年総合大会
4. 発表年 2021年

1. 発表者名 川上蒼太、岡田 深、新田直子、中村和晃、馬場口登
2. 発表標題 半教師あり学習による非視覚センサ値を用いた時間軸をもつ画像列生成
3. 学会等名 電子情報通信学会パターン認識・メディア理解研究会 (PRMU)
4. 発表年 2021年

1. 発表者名 Sota Kawakami, Kei Okada, Naoko Nitta, Kazuaki Nakamura, and Noboru Babaguchi
2. 発表標題 Semi-Supervised Outdoor Image Generation Conditioned on Weather Signals
3. 学会等名 International Conference on Pattern Recognition (ICPR2021) (国際学会)
4. 発表年 2021年

1. 発表者名 新田直子
2. 発表標題 画像認識・生成へのソーシャルメディアの活用
3. 学会等名 電気関係学会関西連合大会（招待講演）
4. 発表年 2020年

1. 発表者名 川上蒼太, 岡田溪, 新田直子, 中村和晃, 馬場口登
2. 発表標題 半教師あり学習による非視覚センサ値を用いた画像生成
3. 学会等名 第23回画像の認識・理解シンポジウム(MIRU2020)
4. 発表年 2020年

1. 発表者名 岡田溪, 新田直子, 中村和晃, 馬場口登
2. 発表標題 気象センサとの整合性を考慮したソーシャルメディアからの屋外画像収集
3. 学会等名 電子情報通信学会2020年総合大会
4. 発表年 2020年

1. 発表者名 松崎大輔, 恒岡知生, 新田直子, 中村和晃, 馬場口登
2. 発表標題 半教師あり学習による気象条件に応じた屋外画像の生成
3. 学会等名 第12回データ工学と情報マネジメントに関するフォーラム (DEIM Forum 2020)
4. 発表年 2020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
---------	---------