

店舗情報と Tweet の影響を考慮した時間帯別路上 GSV 生成手法の提案

松石 侑樹[†] 羽倉 輝[†] 山口 琉太[†]

Panote Siriaraya^{††} 河合 由起子^{†,†††}

[†] 京都産業大学情報理工学部 [〒] 603-8555 京都府京都市北区上賀茂本山

^{††} 京都工芸繊維大学 [〒] 606-8585 京都府京都市左京区松ヶ崎橋上町

^{†††} 大阪大学サイバーメディアセンター [〒] 567-0047 大阪府茨木市美穂ヶ丘 5 – 1

E-mail: [†]{g2054254, kawai}@cc.kyoto-su.ac.jp, ^{††spanote@kit.ac.jp}

あらまし 近年、交通事故による死傷者数の割合は減少傾向にあるが、歩行者横断中の自動車接触事故は最も多く、特に、薄暮時の 17 時～22 時が最多とされている。本研究は、路上の明暗を考慮した安全な経路推薦を目指し、時間帯別の路上画像生成手法を提案する。具体的には、まず、昼間の明るい時間帯の路上画像を Google Street View (GSV) から取得する。次に、GSV の緯度経度から位置情報付き SNS の投稿画像 (Flickr) を取得する。さらに、同地点情報として周辺の店舗等の地物情報 (google Maps や Open Street Maps)，位置情報付き SNS のテキスト情報 (Twitter) を取得する。これら取得した地点情報を時間帯別に分類し、学習する。これにより、新たに入力された GSV 画像から時刻帯別の GSV を生成する。本稿では、特に、学習器として敵対的生成ネットワークである GAN を用い、昼間の GSV 画像から生成された時間帯別の GSV 画像を検証する。

キーワード UI・UX, Web, ユーザ支援, twitter, GSV, 機械学習, 画像処理

1 はじめに

近年、地図プラットフォームや Street View の発展によって、道路情報や各地点の景色を把握することが可能である。しかし、夕方や夜間の時間帯では同じ地点であっても明るさが異なるため、ユーザが見る景色は異なってしまう。また、Google Street View では、時間帯別の景色を見ることはできず、ユーザが時間帯別の景色や地点の明るさを事前に把握することは困難である。また、夕方や夜間に自動車や自転車の運転時に、周囲が次第に暗くなり、物体や歩行者などが見えにくくなる。また、歩行者にとって昼間は明るく見通しの良い道路や場所でも、夕方や夜になると急激に暗くなり、ライトなしでは歩行が難しく危険な道もある。さらに、公表されている交通事故の統計データでは、一日における 2 時間ごとの時間帯別の交通事故発生件数が明らかにされている。この統計より、16 時以降の時間帯に交通事故が多く発生していることが明らかになっている。要因として、仕事の終業にともなう交通量の増加が考えられるが、それ以外にも薄暮における周囲の環境の変化が影響していると考えられる。

本研究では、路上の明暗を考慮した安全な経路推薦を目指し、時間帯別の路上画像生成手法を提案する。時間帯別の路上画像生成では、まず、路上の画像を Flickr から取得し、Flickr の緯度経度から同地点情報として周辺の店舗情報、位置情報付き Tweet を取得する。次に、取得した地点情報を時間帯別に分類し、学習する。これにより、入力された GSV 画像から時刻帯別の GSV が生成される。本稿では、特に、学習器として敵対的生成ネットワークである GAN を用い、昼間の GSV 画像か

ら生成された時間帯別の GSV 画像を検証する。

以下、2 章では、Generative Adversarial Network の関連研究を紹介する。3 章と 4 章で学習モデル生成の手法および画像生成の手法を説明し、5 章で作成したシステムの実装を行い、最後に 6 章で本研究で提案した時間帯別 GSV 表示システムのまとめと今後の課題について述べる。

2 関連研究

本章では、GAN および GAN を応用した自動画像生成についての研究を紹介する。

2.1 Generative Adversarial Network

GAN (Generative Adversarial Network) [1] [2] とは、画像生成や音声生成など、様々なタスクにおいて提案された深層学習モデルである。GAN は、生成器 (Generator) と識別器 (Discriminator) の 2 つのモデルからなるニューラルネットワークで構成されている。Generator は、ランダムなノイズから入力データを生成するモデルであり、Discriminator は、入力されたデータが生成器で生成されたものか、本物のものかを判別するモデルである。GAN は、これらのネットワークを交互に競合させ学習を進めることで、より本物に近い画像を生成することが可能である。

2.2 GAN を応用した画像生成手法

GAN による生成モデルは、高い表現力を持ち、人間が生成したものと区別がつかないほどの高品質の画像や文章を作成することができる。しかし、GAN は学習が安定しないという欠点があり、様々なネットワークの拡張や学習方法に関する研究

が行われている。

例えば, Denton らによる CNN (Convolutional Neural Network) を用いた画像生成タスクに特化したモデルである LAPGAN (Laplacian Pyramid of Generative Adversarial Networks) [3] がある。この LAPGAN では, 各周波数帯の画像特徴ごとに GAN を学習し, 周波数帯ごとの画像特徴を繰り返し生成・合成することで, 低解像度の画像から高解像度の画像を生成することが可能となる。また, LAPGAN の発展として, Mirza らによる DCGAN (Deep Convolutional Generative Adversarial Networks) [4] がある。この DCGAN では, 一度に高解像度の画像を生成することができ, LAPGAN の発展形となっている。

また, zhu らによる CycleGAN [5] は, 従来の Pix2Pix が学習に 1 対 1 で対応したペア画像が対象に必要だったのに対して, 大量のペア画像を用意しなくとも 2 つの違う画像データセットからその関係を学習することで画像変換を行うことができる。しかし, CycleGAN では, 学習モデルに夜画像を用意し, 識別器を学習させている。そのため, 時間帯別の景色画像, 特に地点別の輝度を重みとして画像を生成することは容易ではない。

本研究では, 緯度経度情報及び位置情報付き Tweet, 撮影時間などのアノテーション画像における学習モデルを作成し, 時間帯別 GSV のシステム実装を行う点が特異点である。なお本研究では, GAN の応用研究である StarGAN-V2 [6] を用いることで, 生成する画像のコントロールを行う。

3 GSV 画像取得と時間帯別画像生成モデルおよび時間帯別 GSV システム

本章では, 提案手法の利用例, GSV 画像の取得方法, 時間帯別画像生成モデル, 時間帯別 GSV 表示システムについて説明する。

3.1 時間帯別 GSV システムの利用例

提案するシステムの利用場面の例として 2 つ紹介する。1 つ目は, 自動車や自転車での運転前の利用である。慣れていない地域で運転する場合, 予め通る道や周辺の景色を把握したい場合が多いが, 現状それができるのは, GSV のみである。しかし, 現状の GSV では, 昼間の景色しか把握することができず,

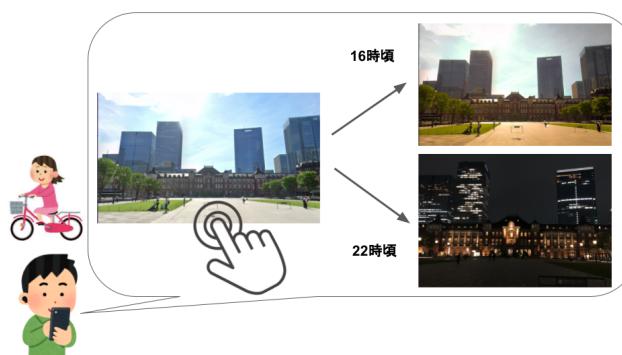


図 1 リアルタイム時間帯別 GSV による利便性

昼間と夜間や夕方では, 周辺の景色は大きく異なる可能性が高い。そこで, 本システムを活用することで, 事前に道の特徴や街灯の有無, 時間帯別の景色を把握することが可能となり, 安全性を向上できる。2 つ目は, 図 1 のような賃貸物件等の下見での活用である。土地勘のない地域に引っ越す際には, 周辺施設の有無や安全性等の周辺環境を事前に知りたい場合が多い。しかし, 時間的・経済的制約があり, 全ての人が事前に下見することは難しい。そこで, 提案するシステムを活用することで, 事前にスマホから, 物件周辺の環境情報を知ることができるので, 制約に縛られず, 簡単に時間帯別の物件周辺の環境状況を知ることができる。

3.2 GSV 画像取得と時間帯別画像生成

本研究では, 位置情報付き Tweet および地点別情報に基づく学習モデルの作成と時間帯別路上 GSV 生成を実現するため, 昼間の画像データをユーザが操作する GSV より取得する。本節では, 時間帯別 GSV のシステム構成に基づいた処理の流れについて概要を述べる。

図 2 に, 画像取得と分析システムの構成を示す。まず, ユーザはモバイル端末上で表示された GSV を自身で操作し, GSV 上で表示されている地点の位置情報を取得する。次に取得した位置情報をもとに, 位置情報付き Tweet および地点の周辺施設を取得する。具体的には, 店舗の営業終了時刻, 店舗の収容人数, 店舗カテゴリ, 隣の店舗との距離がある。ここで, 位置情報に基づいた画像に, 取得した地点別情報による重み付けを行う。続いて, 独自に作成した学習モデルによって学習させた GAN の Generator が画像の生成を行う。最後に生成した画像をモバイル端末に表示する。

3.3 時間帯別 GSV 表示システム

図 3 に本研究で実装するモバイルアプリのプロトタイプを示す。それぞれ夕方 (17 時頃) と夜 (21 時頃) の景色をアプリに表示している。アプリ左側の画面で GSV を表示し, ユーザは右端のプログレスバーで任意の時間を操作することで, アプリ右側の画面に任意の時間の景色を表示する。

4 位置情報に基づく Tweet と地点別情報の取得と学習モデル生成

本章では, 位置情報に基づく Tweet と地点別情報における学習モデル生成手法について説明する。

4.1 画像の収集

本研究では, 画像生成に GAN を用いるため, 学習モデル作成における画像を用意する必要がある。そこで, 学習用の画像の効率的に収集するために, Flickr API を使用する。Flickr は, 写真共有 Web サイトおよびオンラインコミュニティプラットフォームである。ユーザは自分の撮影した写真をアップロードおよびタグ付けし, 共有することができる。この Flickr API を使用し, 位置情報および撮影日時付き画像を収集する。

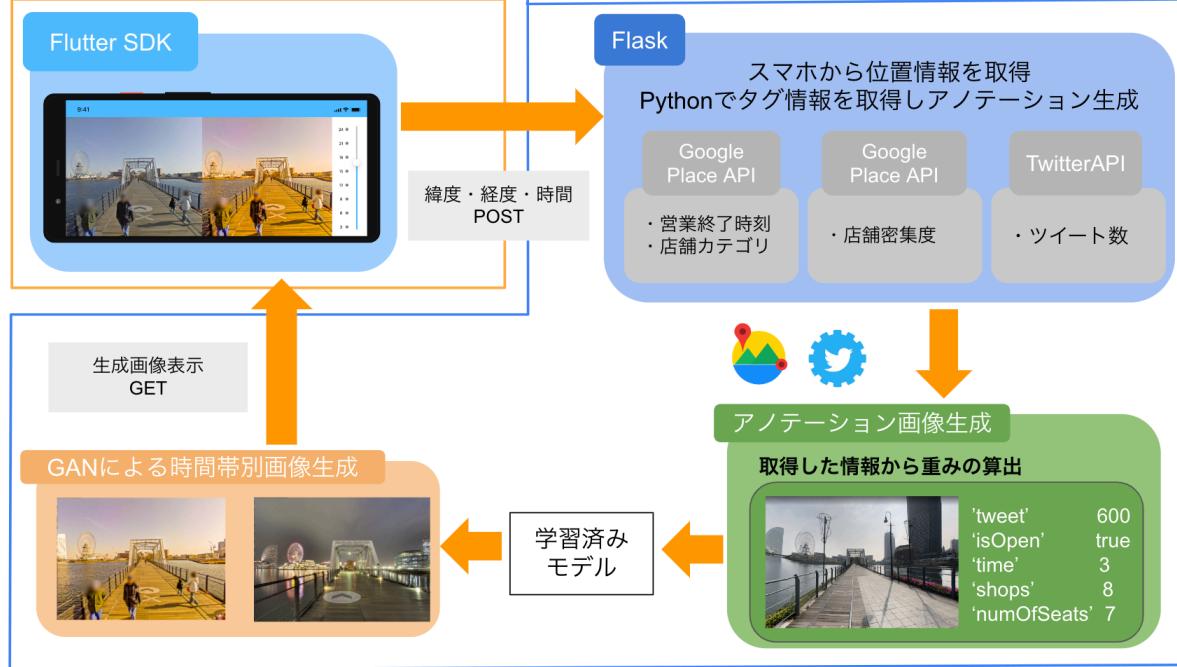


図 2 Flickr と地物情報を用いた GAN による時間帯別 GSV 生成システムの処理の流れ



図 3 時間帯別 GSV プロトタイプ：オリジナル GSV（左）16 時間帯と 21 時間帯として生成された GSV（右）

4.2 地点別情報および Tweet の取得

図 4 は、Flickr API で取得した画像についている位置情報をもとに、地点別情報および Tweet データを取得する際の流れである。地点別情報には、店舗の営業終了時刻、店舗の収容人数、店舗カテゴリ、隣の店舗との距離がある。これらの情報は、Google が提供する Place API および Distance API を用いて取得を行う。この取得したデータをもとに画像に対してラベル付けをする。

4.3 地点別情報による輝度の重み付け

収集した地点別情報による重み付けの手法を説明する。地点

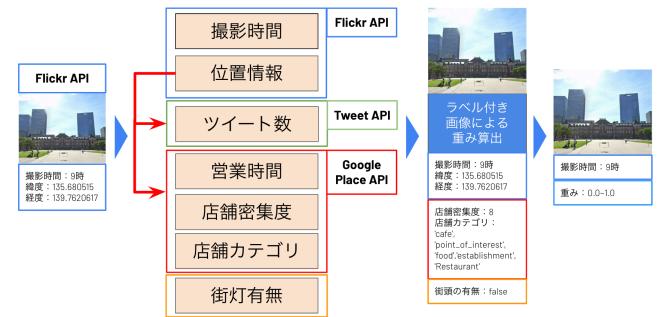


図 4 地点別情報による輝度の重み付け

別情報による輝度の重みは 1~10 までの 10 段階で評価する。重みは、暗いほど 10 に近づき、明るいほど 1 に近づく。本研究では重み付けの定義を次のようにした。ある地点において周辺の店舗の営業終了時刻が 22 時の場合、22 時以降の地点は暗い可能性が高いので重みは 9 とする。続いて、周辺の店舗カテゴリについて、ある地点において、周辺にあるお店のカテゴリが、ショッピングモールの場合は、人が多いので周辺は明るい可能性が高い。よって、重みは 4 程度となる。一方で、周辺にお店がない場合は暗い可能性が高いため、重みは 10 になると考える。上記の定義に基づいて、取得した画像に対して、重み付けを行う。

4.4 学習モデルの作成手法

学習モデルの作成手法について説明する。本研究では、StarGAN-v2 を用いて学習を行う。学習には、12 時、18 時、21 時の 3 ドメインを用いる。また、1 ドメイン 2500 枚の時間帯別画像を学習データとし、テストデータを 300 枚として学習を行った。学習に用いたパラメータとしては、バッチサイズ 4,

学習回数 1000 回から 5000 回までの学習を行う。

5 時間帯別 GSV 表示システムの実装

本研究では、12 時、18 時、21 時をドメインとして学習をおこなった。また、本研究では、実験データとして、FlickrAPI から収集した 1980 年 11 月から 2023 年 1 月までの約 80000 枚の撮影時間付き画像を用いた。実際に使用した画像としては、欠損値を除去した時間帯別の画像 2500 枚を使用した。

5.1 データ入力と出力画面

これらの画像を用いて、通常の StarGAN-V2 を適用して、特定の時間の景色の生成を行った。図 3 に作成したアプリケーションのアプリ起動から地点における時間帯別画像表示までの流れを示す。アプリを起動すると、現在地を中心とする GSV が表示されるので、GSV を用いて移動を行う。地点移動後、プログレスバーを操作し、時間を変更する。この時点で、地点別の画像がサーバに送信され、地点における重みを自動で付与する。その後、予め学習をおこなった識別器を用いて、昼間の画像から時間帯別の画像を生成し、モバイル端末の画面に画像を表示する。

5.2 学習回数による精度検証と考察

本研究では、学習回数による生成画像の精度検証を目的として検証を行った。具体的には、学習回数 1000 回から 5000 回まで学習を行い生成画像を比較する。

図 5 に任意の地点の GSV 画像を入力として、提案手法がら生成された画像結果を示す。結果から、学習回数が 5000 回に着目すると、12 時の時には若干明るく、18 時の時には薄暗く、21 時の時には暗くなっていることが確認でき、時間帯別の特徴をもとに変換できることが確認できた。しかし、本研究における時間帯別の地点別画像を生成するという目的において、期待する精度の画像を生成することはできておらず、改善の余地があると考えている。具体的には、識別器の改善による精度向上の可能性や輝度に対する重み付与による更なる精度向上の可能

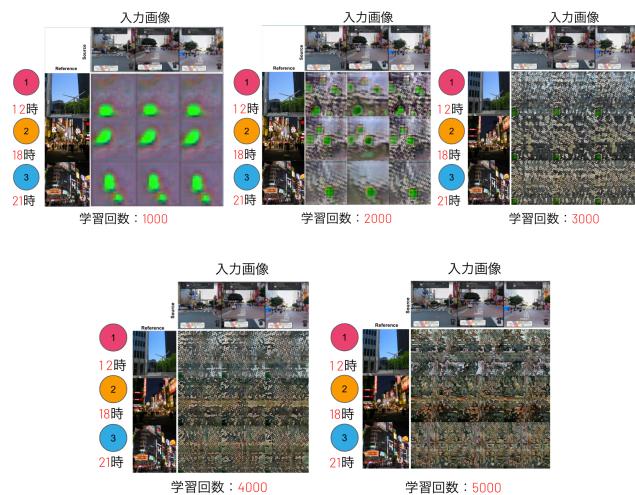


図 5 3 地点の GSV 入力画像と出力結果画像

性があると考えている。高精度な画像を生成できなかった要因としては、不十分な学習回数、学習に使用するデータセットの枚数、バッチサイズなどが考えられる。

今後の課題としては、十分な学習を行うことに加え、学習データセットの増加及びバッチサイズの増加が必要であると考えている。さらに、輝度に対する重みの付与、季節を考慮した時間帯別画像の生成が必要であると考えている。

6 終わりに

本研究では、路上の明暗を考慮した安全な経路推薦を目指し、時間帯別の路上画像生成手法を提案した。

実験より、独自に生成した識別器を用いることで、時間帯別の画像を生成することはできなかった。今後は、より精度の高い識別器を作成することで、精度の高い時間帯別 GSV 画像を生成し、経路推薦に活用する予定である。

謝 辞

本研究の一部は、JSPS 科研費 JP19K12240, JP20H00584 および京都産業大学先端科学技術研究所（HMD 共生科学研究センター）の助成を受けたものである。ここに記して謝意を表す。

文 献

- [1] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.
- [2] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks. *Commun. ACM*, 63(11):139–144, oct 2020.
- [3] Emily Denton, Soumith Chintala, Arthur Szlam, and Rob Fergus. Deep generative image models using a laplacian pyramid of adversarial networks, 2015.
- [4] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks, 2015.
- [5] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks, 2017.
- [6] Yunjey Choi, Youngjung Uh, Jaejun Yoo, and Jung-Woo Ha. Stargan v2: Diverse image synthesis for multiple domains, 2019.