テンプレート画像とコンテキスト情報を用いた観光地の現況画像の自動生成法の提案

河中 昌樹[1]　中村 優吾[2]　諏訪 博彦[1,3]　安本 慶一[1,3]

概要：スマートフォンを用いて観光地の情報を提供する観光支援システムが多く提案されているが，これらのシステムの多くは，観光地の最新の状況（現況）を提供できていない．リアルタイムに観光地の情報を取得するために，観光している人に観光地の現在の写真を共有してもらう方法があるが，提供者のプライバシー情報が漏洩するリスクが伴う．そこで我々は，ユーザが撮影した写真を直接共有するのではく，オブジェクト認識モデルを写真に適用して取得可能な観光地のコンテキスト情報（天気や混雑，動植物などのオブジェクト等）のみを共有するというアプローチを検討している．本研究では，観光地の最新のコンテキスト情報が共有されるという前提のもと，観光地のテンプレート画像とコンテキスト情報から観光地の現況画像を自動生成する手法を提案する．提案手法は，テンプレート画像にセマンティックセグメンテーションを適用し，各セグメントごとにコンテキスト情報を考慮して画像変換を行うことで，観光地の現況画像の自動生成を行う．評価実験より，提案手法により生成された画像は観光地の現況をユーザに適切に伝えることが可能であることが示された．

Automatic generation of up-to-date tourist spot images using template images and context information

Masaki KAWANAKA[1]　Yugo NAKAMURA[2]　Hirohiko SUWA[1,3]　Keiichi YASUMOTO[1,3]

1. はじめに

近年，スマートフォンを用いて情報を提供する観光支援システムが提案されている [1, 2]．磯田ら [1] は，観光スポットごとに時間や天気などを考慮した評価値から複数の次訪問スポットを推薦する観光客の意思決定支援システムを提案している．このシステムでは，推薦の際に，ユーザが次訪問スポットについて知るためにリアルタイムの観光情報が必要となる．リアルタイムの観光情報を取得する方法としてカメラを各観光地に設置することが考えられるが，設置・維持コストの点から現実的ではない．

この問題に対し，河中ら [3] は，観光地の情報収集を行うゲーミフィケーションを用いた参加型センシング手法を提案している．しかし，この手法では，個人が撮影した写真を直接共有するため，写真提供者のプライバシーが漏洩するリスクが生じる．プライバシー保護を行いながらリアルタイムに情報を共有するために，ユーザが撮影した写真を直接共有するのではなく，画像に含まれるコンテキストのみを共有する方法が考えられる．

画像に含まれるコンテキストを抽出するためにオブジェクト認識モデルが利用可能である．観光や地域に特化したオブジェクト認識モデルを構築するには多数の写真を使った学習が必要であるが，学習のために写真を提供させるとプライバシー情報が漏洩してしまう．そこで，プライバシーに配慮してオブジェクト認識モデルを学習する方法として，プライバシーを含むデータを直接学習データとして用いずに，各端末で学習された機械学習モデルの重みパラメータを交換・統合する連合学習（Federated Learning）が注目されている [4]．冨田ら [5] は，観光客端末間での近距離直接通信を活用した連合学習に基づくユーザ参加型オブジェクト認識モデル構築手法を提案している．この手法では, すれ違った他の観光客が持つモデルの重みパラメタを受信し，自分のモデルに統合することで，ユーザが撮影した観光地の写真を提供することなく，観光地のコンテキスト情報を抽出するオブジェクト認識モデルの学習を可能とする．しかし，この研究では，抽出したコンテキスト情報をどのように観光地の現況を伝えるために共有するのかま

1 奈良先端科学技術大学院大学, Nara Institute of Science and Technology
2 九州大学, Kyushu University
3 理化学研究所, RIKEN

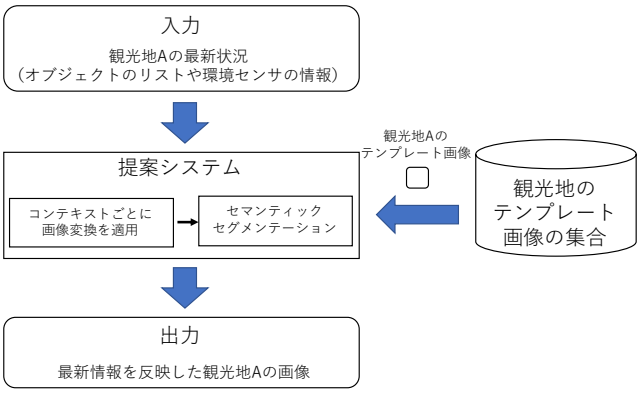


図 1　システムの概略図

では考慮していない．

そこで，我々は，観光客が撮影した観光地の現在の状況（現況）を示す画像からコンテキストを抽出し，各観光地のテンプレート写真に反映させるシステムを提案している [6]. 本システムの概略図を**図 1** に示す．本システムでは，観光地を代表するテンプレート画像にセマンティックセグメンテーションを適用し，反映したいコンテキストの各要素（混雑や天気，季節を反映する樹木など）に対応するセグメントの画像変換を行い，観光地の現況を説明する画像を生成する．

本稿では，実際にテンプレート画像に対してコンテキスト情報に基づく画像変換を行い，コンテキストを反映した画像を共有する具体的な手法を提案する．提案システムでは，テンプレート画像の特徴や意味が類似した部分領域ごとにユーザから収集したコンテキスト情報を用いて画像変換を行い，観光地の現況画像を自動生成する．

提案システムの有効性を評価するために，生成された観光地の現況画像を見た時に，観光地の状況がユーザにどのように認識されるのか評価した．その結果，概ねシステムが提供したい情報をユーザに伝えることが可能であることが明らかになった．

2.　関連研究

本章では，観光地コンテキストの共有に関する従来手法を整理したうえで，提案手法で利用する技術について述べる．

2.1　観光地コンテキストの共有

現在地から離れた場所の状況を説明する際には，状況を文章で説明する方法やその場所の写真を提供する方法が一般的である．一度も訪問したことがない場所に対しては，文章で説明する方法では，状況を理解することが困難である一方，画像を用いて説明する方法では，状況を理解することが容易であると考えられる．本論文で対象としている観光分野のシステムでは，一度も訪問したことがない場所について説明する場合が多いと考えられるため，画像で状況を説明する方法を検討する．

観光支援システムの多くは，観光地のホームページの写真や前もって撮影した観光地を代表する写真を用いて観光地を紹介する．しかし，これらの写真では観光地の現在の状況は考慮されていないため，季節感や混雑度などの動的な情報をユーザに伝えることができない．市民参加型写真収集アプリ [7] や観光地の情報収集を行うゲーミフィケーションを用いた参加型センシング手法 [3] を用いることで，最新の観光地の情報を反映した写真を収集することが可能であるが，ユーザが撮影した写真を直接共有することはプライバシー情報漏えいのリスクが伴う．

そこで我々は，観光客が撮影した観光地の現在の状況を示す画像からコンテキストを抽出し，コンテキストのみを収集することで，プライバシーを保護しながら観光地の最新状況の収集を行う．その後，テンプレート画像の特徴や意味が類似した部分領域ごとにコンテキスト情報を用いて画像変換を行い，観光地の現況画像を自動生成する．そのため，以下の節では，特徴や意味が類似した部分領域に分割するセグメンテーションと画像変換技術であるスタイル変換について述べる．

2.2　セグメンテーション

画像を特徴や意味が類似した部分領域に分割する処理はセグメンテーションと呼ばれる．画像の分割方法によって様々なセグメンテーション法が提案されている [8–10]. 本研究では，同じ物体には同様の画像変換を行うため，物体のクラスごとに画像の分割を行うセマンティックセグメンテーション [8] を利用する．機械学習を用いたセマンティックセグメンテーションとして最も広く認知されているものに Fuuly Convolutional Network (FCN) が提案されている．FCN は，畳み込み層のみで構成されており，入力画像のピクセル単位でクラス分類を行う手法である．その後，FCN で有効性が示された skip connection を更に導入した U-net [11] や Encoder-Decoder 構造を用いた FCN である segnet [12] を用いたセマンティックセグメンテーションが提案されている．

2.3　スタイル変換

スタイル変換とは，物体の位置情報などの構造を担保するコンテンツ画像と画風を示すスタイル画像を入力として，コンテンツ画像の構造とスタイル画像の画風を持った画像を生成する手法である [13]．この手法では，CNN を用いてコンテンツ画像の構造とスタイル画像の画風を持った画像に近くなるようにホワイトノイズを更新する．この時，物体の形状を保つための損失関数であるコンテンツ誤差 L_{content} とコンテンツ画像に画風を反映させるための損失関数であるスタイル誤差 L_{style} を用いて以下の損失関数を最小化するように学習を行う．

$$L_{\text{total}} = \alpha L_{\text{content}} + \beta L_{\text{style}} \tag{1}$$

ここで，α と β は重み係数であり，コンテンツ誤差 L_{content} とスタイル誤差 L_{style} の比重を決定する．例えば，α を β と比較して大きくすると物体の位置や形状が頑健になるが，画風が反映されにくくなる．一方，β を α 比較して大きくすると画風が強く反映されるが，物体の位置や形状が失われる．コンテンツ誤差 L_{content} は，l 層目の i 番目のチャネルの j に位置するコンテンツ画像の特徴マップ P^l_{ij} と生成画像の特徴マップ F^l_{ij} を用いて以下のように表される．

$$\mathcal{L}_{\text{content}}\left(\vec{p}, \vec{x}, l\right) = \frac{1}{2}\sum_{i,j}\left(F^l_{ij} - P^l_{ij}\right)^2 \tag{2}$$

ここで，$\vec{p}$ はコンテンツ画像，$\vec{x}$ は生成画像を示す．また，スタイル誤差 L_{style} は，チャネル間の相関を示すグラム行列を用いて以下のように表される．

$$L_{\text{style}}(\vec{a}, \vec{x}) = \sum_{l=0}^{L} w_l E_l \tag{3}$$

$$E_l = \frac{1}{4N_l^2 M_l^2}\sum_{i,j}\left(G^l_{ij} - A^l_{ij}\right)^2 \tag{4}$$

ここで，$\vec{a}$ はスタイル画像，w_l は各層の重み，N_l と M_l はそれぞれ特徴マップの数とサイズ，G^l_{ij} と A^l_{ij} はそれぞれ生成画像とスタイル画像のグラム行列を示す．

入力画像に対応する出力を生成する Generator とその入出力のペアが本物かの判別を行う Discriminator を競わせながら学習を行う手法である GAN [14] などを活用した手法も提案されており，質の高いスタイル変換が可能になっている [15,16]．

3. 提案システム

本章では，提案システムと想定環境について述べる．提案するシステムの概要図を**図 2** に示す．

3.1 提案システムの概要

提案システムの目的は，観光地の現況を観光者に伝えることである．この目的を達成するため，提案手法では，観光者が撮影した写真を直接共有するのではなく，オブジェクト認識モデルを用いて画像に含まれるコンテキストのみを共有する．

オブジェクト認識モデルによって出力される観光地のコンテキスト情報と対象とする観光地のテンプレート画像を用いて観光地の現況画像を自動生成することで，観光地の現況を観光者に伝える．提案システムは，図 2 左に示すように事前に撮影したテンプレート画像にセマンティックセグメンテーションを適用して，特徴や意味が類似した複数のクラス（セグメント）に画像を分割する．その後，図 2 中央に示すように，それぞれのセグメントに観光地のコンテキスト情報を用いて画像変換を適用する．最後に図 2 右に示すように，適用結果の画像を 1 枚に統合し，明るさとコントラストを変換することで観光地の現況画像を生成する．

表 1 東大寺のオブジェクトリスト

クラス分類	サブクラス	要素
空クラス	空の色	紺碧, 天色, 空色
	雲の色	白い雲, 黒い雲
	雲量	多い, 普通, 少ない
樹木クラス	葉の状態	桜, 黄葉, 新緑, 紅葉
芝生クラス	芝生の状態	新緑, 黄化
混雑クラス	混雑状態	非常に多い, 多い, 普通, 少ない, 非常に少ない
全てのクラス	明るさ	明るい, 少し明るい, 普通, 少し暗い, 暗い

3.2 オブジェクトリスト

提案手法では，観光地を代表するテンプレート画像を観光地のコンテキスト情報を用いて画像変換することで観光地の現況画像を自動生成を行う．そのため，テンプレート画像に対して，観光地のコンテキストに対応するオブジェクトのリストが必要である．**図 3** に異なる時期に撮影した東大寺の写真を示す．左は 2021 年 7 月に撮影したもの，右は 2022 年 4 月に撮影したものである．左右の写真の違いとして，天気，樹木や芝生の状態，混雑度，時間帯が挙げられる．そのため，これらのコンテキスト情報を反映している**表 1** のような観光地のオブジェクトリストと考えられる．本論文では，それぞれのサブクラスが持つ要素は One-hot ベクトルであると仮定する．そのため，空の色においては，空色，天色，紺碧が重複することは考えない．

3.3 セグメンテーション

提案手法は，観光地を代表するテンプレート画像を使って現況画像を生成する．この時，テンプレート画像に対してセマンティックセグメンテーションを適用し，特徴や意味が類似した複数のクラスに分割する．

セマンティックセグメンテーションは，DNN を用いた方法と手作業で行う方法がある．DNN を用いた方法は，画像を入力するだけで自動的に特徴や意味が類似した複数のクラスに分割することが可能であるが，それぞれの想定環境に適した正解ラベルを大量に用意する必要がある．本研究では，テンプレート画像は頻繁に更新されるものではないため，高精度に特徴や意味が類似したクラスに分割が可能な手作業のセマンティックセグメンテーションを用いることとした．東大寺のテンプレート画像に手作業でセマンティックセグメンテーションを適用した結果を**図 4** に示す．セマンティックセグメンーションは，観光地のオブジェクトリストのクラス分類と同様に空クラス，樹木クラ

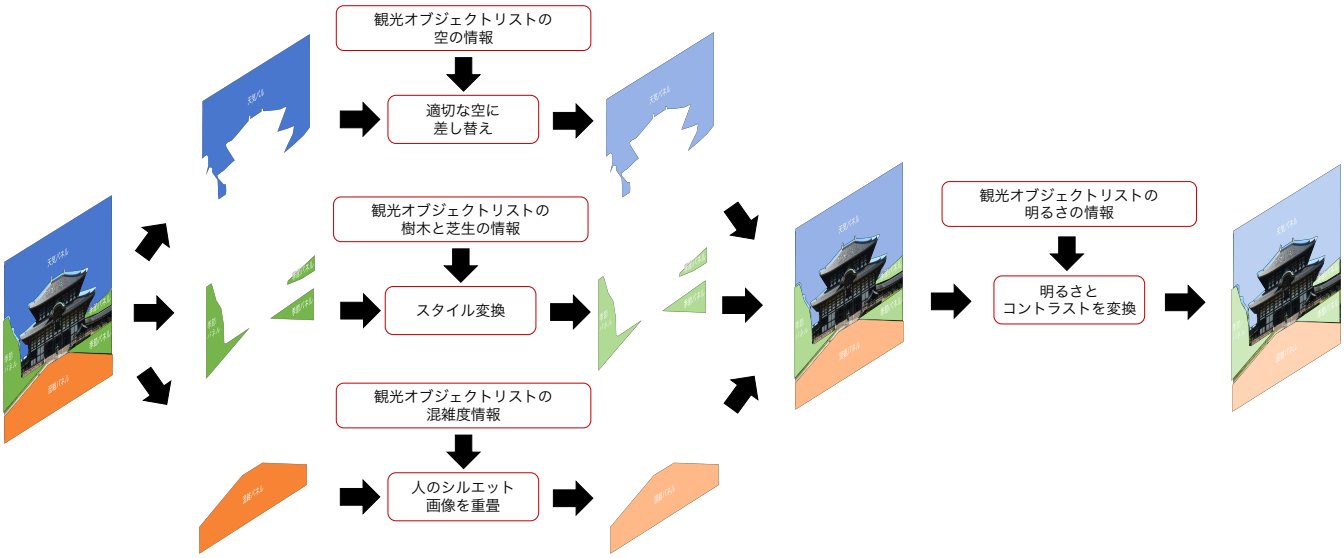


図 2　本システムの概要図

図 3　時期による観光地の違い

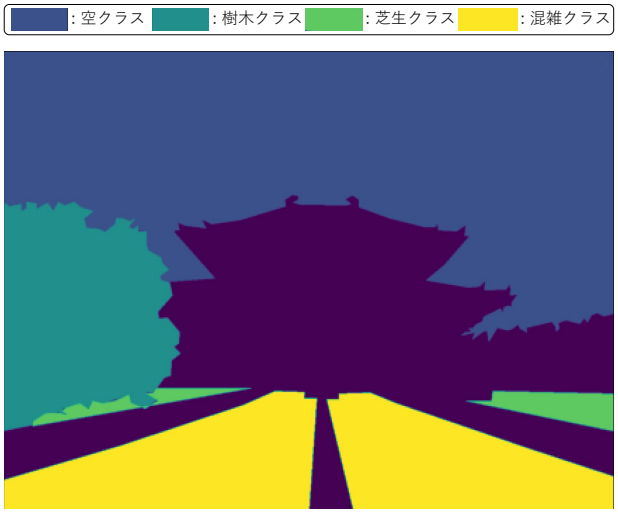


図 4　セマンティックセグメンテーションの結果

ス，芝生クラス，混雑クラスに分類を行った．

提案手法では，オブジェクト認識モデルによって，観光地のコンテキスト情報が取得可能であると仮定し，観光者は，自身が撮影した写真をオブジェクト認識モデルに入力して観光地のオブジェクトリストを取得する．オブジェクト認識モデルを用いて，写真に含まれるオブジェクトを検出する研究は盛んに行われており，高精度な検出が可能である [17,18]．

テンプレート画像と観光地のコンテキスト情報（オブジェクトリスト）を用いてセグメントごとに画像変換を行う．画像変換手法については，次章で詳しく述べる．

4.　コンテキスト情報に基づく画像変換

本章では，テンプレート画像をセマンティックセグメンテーションによって分類したクラスごとに適用する画像変換について詳しく述べる．本システムを用いて観光地のオブジェクトリストを反映させた結果を**図 5** に示す．図 5 を生成するために，どのような画像変換を各クラスごとに適用するのかを以下に述べる．

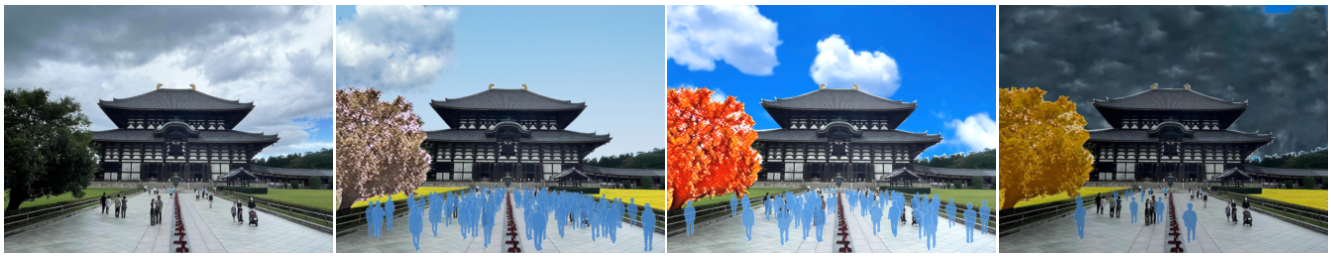
4.1　空クラス

空クラスでは，オブジェクト認識モデルによって得られる空と雲の状態によって，適切な空に置き換える．そのため，空の色，雲の色，雲量の全ての組み合わせを含む 18 枚の空画像を用意する．

空クラスのオブジェクトリストに基づいてテンプレート画像の空を置き換えた結果を**図 6** に示す．画像は左から，元画像，天色+白い雲+雲量少し，紺碧+白い雲+雲量普通，空色+雨雲+雲量多いの結果を示す．空クラスを変換することによって, 天気の印象が変わっていることが分かる．観光地のより詳細な情報を考慮するためには，用意する空画像を増やしたり，空クラスのオブジェクトリストに基づいて，GAN などを使って空画像を生成する方法が考えられる．

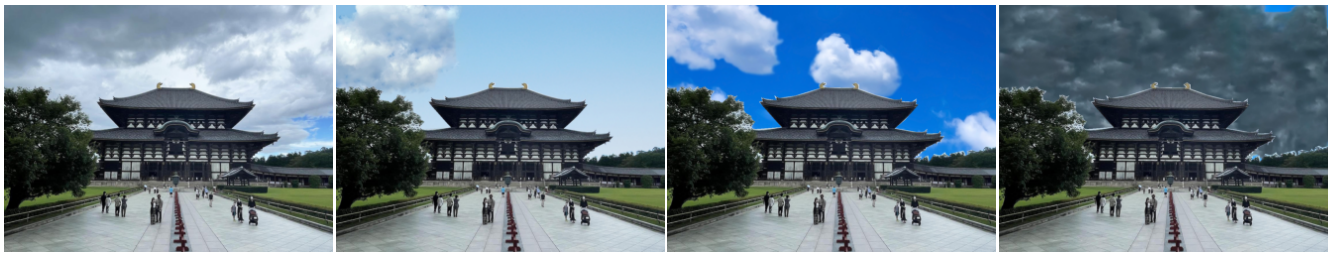
4.2　樹木と芝生クラス

樹木クラスと芝生クラスでは，オブジェクト認識モデルによって得られる葉や芝生の状態に近づくように，スタイル変換を行う．スタイル変換には，create ML の Style



元画像　空色+白い雲+雲少ない+桜+黄化+非常に混雑の結果　紺碧+白い雲+雲普通+紅葉+混雑度普通＋少し明るいの結果　天色+黒い雲+雲多い+黄葉+黄化+非常に少ない+少し暗いの結果

図 5　提案システムを用いた画像変換の結果



元画像　空色+白い雲+雲少ないの変換結果　紺碧+白い雲+雲普通の変換結果　天色＋黒い雲＋雲多いの変換結果

図 6　空クラスの画像変換の結果

Transfer を用いてスタイル変換を行う．学習には 500 回パラメータの更新を行い，ハイパーパラメータである Style Strength と Style Density はデフォルトのまま使用した．

樹木クラスと芝生クラスのオブジェクトリストに基づいてテンプレート画像の葉と芝生の状態を変化させた結果を**図 7** に示す．画像は左から，元画像，桜への変換，紅葉への変換，芝生を黄化させた結果を示す．スタイル変換を用いることで，テンプレート画像の木や芝生の形状を保持したまま，任意のスタイルに変換できていることが分かる．

4.3　混雑クラス

混雑クラスでは，オブジェクト認識モデルによって得られる混雑状態に近づくように画像変換を行う．混雑度の表現方法として，道がどのくらい混んでいるかを人の増減で表す方法が考えられる．そこで，オブジェクト認識モデルによって得られる 5 段階の混雑度情報 (非常に多い，多い，普通，少ない，非常に少ない) に応じて，人のシルエットを混雑クラスに重畳する．そのため，テンプレート画像に人のシルエットを適切なサイズで重畳する必要がある．そこで，テンプレート画像に写っている人のサイズを用いて，座標ごとの適切なサイズの計算を行った．テンプレート画像に写っている人のサイズは，YOLOv3 [18] を用いて出力される bounding box の大きさを用いる．異なる y 座標に位置する人の bounding box のサイズの関係から y 座標と人の大きさの関係を一次関数で表現することが可能である．

混雑クラスのオブジェクトリストと bounding box のサイズと y 座標の関係を示す一次関数を用いてテンプレート画像の混雑状態を変化させた結果を**図 8** に示す．画像は左から，150 人の追加，80 人の追加, 10 人の追加した場合の結果を示す．結果より，人を配置する場所によって適切なサイズで重畳できていることが分かる．

4.4　全体クラス

全体クラスでは，オブジェクト認識モデルによって得られる明るさ情報から現在の時間帯に近づくように画像変換を行う．明るさの変換方法として，openCV の convertScaleAbs を用いて明るさとコントラストを調整した．

全体クラスのオブジェクトリストを用いてテンプレート画像の明るさを変化させた結果を**図 9** に示す．画像は左から，明るい，少し明るい，少し暗いの結果を示す．明るさを暗くすることで日が落ちて夕方や夜に近づいているような表現が可能である．

5.　評価実験

本実験の目的は，提案システムで作成した画像をユーザが見た時に，伝えたい情報が正しく伝わっているかを評価することである．提案システムを用いて生成した画像について，ユーザがどのように感じるかを Yahoo! クラウドソーシングを利用して調査した．本実験では，それぞれのコンテキストによって，目的とする情報が正しく伝わっているかを確認するために，テンプレート画像に対して 1 種類のコンテキストのみ変化させたものをユーザに提示する．そのため，表 1 より，空クラス 18 種類（空の色の総数 × 雲の色の総数 × 雲量），樹木クラス 4 種類，混雑状態

元画像　桜への変換結果　紅葉への変換結果　芝生の黄化の変換結果

図 7　樹木クラスと芝生クラスの画像変換の結果

元画像　150人のシルエットを追加した結果　80人のシルエットを追加した結果　10人のシルエットを追加した結果

図 8　混雑クラスの画像変換の結果

5 種類，明るさ 5 種類のテンプレート画像から 1 つの要素が変化した計 32 枚の画像を用意した．400 人に対して，テンプレート画像と比較して 1 つの要素が変化した画像を 1 枚提示し，空クラスでは天気，樹木クラスでは季節，混雑状態では混雑度，明るさでは時間帯についての質問を答えるタスクを 9 枚の画像が終わるまで行った．9 枚の画像のうち 1 枚はチェック設問であり，写真の右上に記載した文字を正しく回答しなかった場合には，回答としての信頼度がないとして，その人のデータは全て破棄した．空クラスの場合には，天気として晴れ，曇り，雨の選択肢，樹木クラスでは，季節として春，夏，秋，冬の選択肢，混雑クラスでは混雑度として非常に多い，多い，普通，少ない，非常に少ないの選択肢，明るさでは，時間帯として朝，昼，夕方，夜の選択肢を用意した．

空クラスに対する実験結果を**図 10** に示す．空の色に対する結果（図 10 左）より，空の色によって天気が変わると回答する人は少ないことが分かった．しかし，雲の色に対する結果（図 10 真ん中）より，雲の情報は天気に与える影響が大きく，雲の色が白色の場合には，多くの画像で晴れであると回答することが多い．一方で，雲の色が黒色の場合には，晴れと回答する回数が減少し，曇りや雨などと回答することが多い．このことから，黒い雲は雨雲などを連想し，曇りや雨の回答が増加したと考えられる．また，雲量に対する結果（図 10 右）より，雲の量については，雲が少ない時は晴れが多く，雲と空の比率が同じぐらいの場合には，晴れと曇りで意見が割れることが分かった．これは，人の感性によって晴れと曇りだと判断基準が異なるためだと考えられる．

樹木に対する結果（**図 11**）より，春を代表する桜の場合には，ほぼ全ての人が春と回答していた．また，同様に秋を代表する黄葉と紅葉の場合には，ほぼ全ての人が秋と回答していた．しかし，新緑については想定していた春以外の回答が多かった．これは，一般的な樹木の葉の色と新緑の色が似ているため，夏などの季節の回答が増えたと考えられる．この問題を回避するためには，新緑についてはより高精度な画像変換を適用し，若葉のみずみずしい葉を表現する必要がある．

混雑度に対する結果（**図 12**）より，混雑クラスに重畳する人数が最も少ない 10 人の場合には，普通や少ないの回答が多い．一方，混雑クラスに重畳する人数が最も多い 150 人の場合には，非常に多いや多いの回答が多く，人数の増加によって混雑していると認識する人が増えることを確認した．また，混雑の評価として「非常に少ない」の回答が少ない点については，テンプレート画像に写っている人物の影響があると考える．そのため，「非常に少ない」を表現するためには，テンプレート画像にあまり人が写らないものを使用する方法や機械学習を用いて人物を削除する方法が考えられる．

全体クラスの明るさに対する結果（**図 13**）より，明るさが最も明るい変換の結果では，朝や昼などの回答が多くなっている．一方，明るさが最も暗い変換の結果では，夕方と夜の回答が増える傾向だった．しかし，明るさを変化させるだけでは，朝と昼の違いを表現することが難しいことが分かった．そのため，朝と昼を区別するためには，太陽の位置などの明るさ以外の要素を取り入れる必要があると考えられる．また，観光地によっては，街灯などの点灯

元画像　明るくした結果　少し明るくした結果　少し暗くした結果

図 9　全体クラスの画像変換の結果

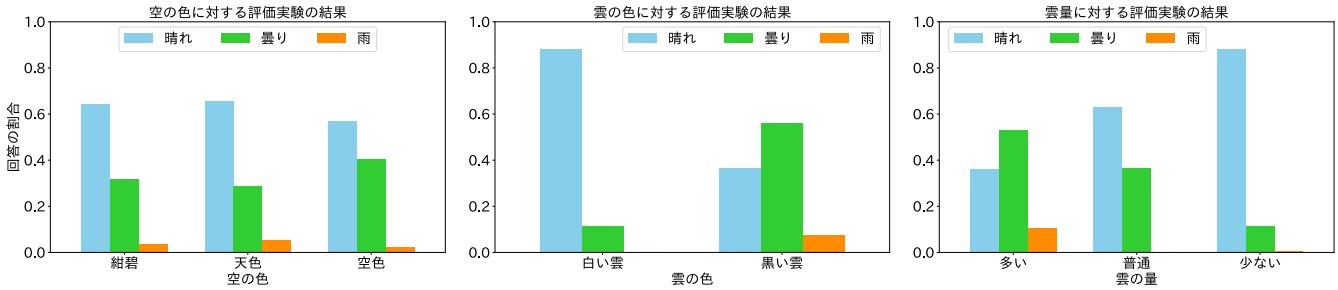


図 10　空クラスに対するクラウドソーシングの結果

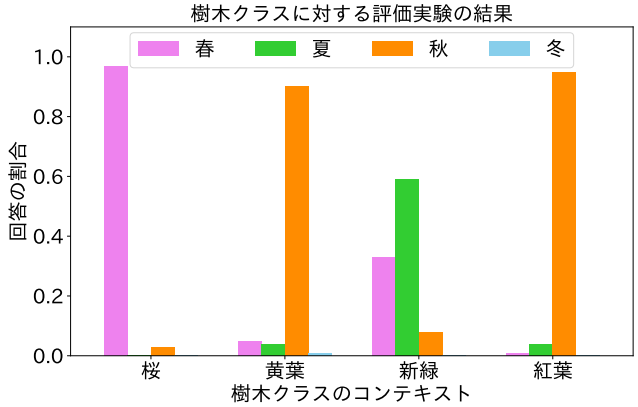


図 11　樹木クラスに対するクラウドソーシングの結果

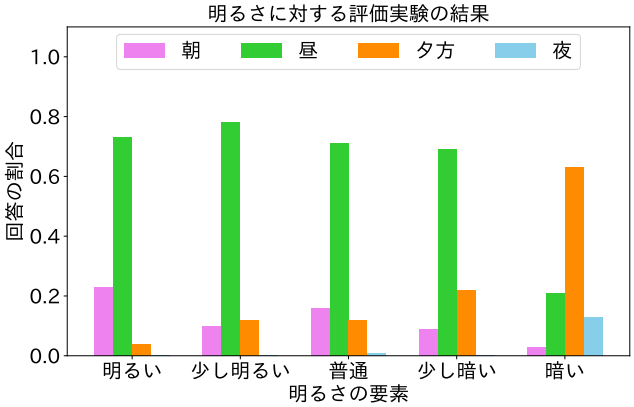


図 13　全体クラスに対するクラウドソーシングの結果

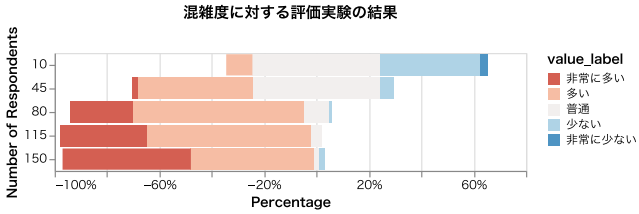


図 12　混雑クラスに対するクラウドソーシングの結果

場所がある場合には，セマンティックセグメンテーションによって事前に点灯場所の位置を特定しておいて，点灯場所付近は明るくするなどの工夫が必要であると考える．

6.　おわりに

本研究では，テンプレート画像とコンテキスト情報を用いた観光地の現況画像の自動生成法の提案を行った．テンプレート画像にセマンティックセグメンテーションを適用し，各クラスごとにコンテキスト情報を考慮して画像変換を行うことで，観光地の現況画像の自動生成を行った．評価実験より，概ねシステムが提供したい情報をユーザに伝えることが可能であることが明らかになった．今後の課題として，本手法を様々な観光地への適用し評価することや，実際の観光客の写真にオブジェクト認識モデルを適用して抽出したコンテキスト情報から画像を生成し評価することなどが挙げられる．

謝辞

本研究は JSPS 科研費 JP21H03431 の助成を受けたものです．

参考文献

[1]　Shogo Isoda, Masato Hidaka, Yuki Matsuda, Hirohiko Suwa, and Keiichi Yasumoto. Timeliness-aware on-site

planning method for tour navigation. *Smart Cities*, Vol. 3, No. 4, pp. 1383–1404, 2020.

[2] Masato Hidaka, Yuki Kanaya, Shogo Kawanaka, Yuki Matsuda, Yugo Nakamura, Hirohiko Suwa, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. On-site trip planning support system based on dynamic information on tourism spots. *Smart Cities*, Vol. 3, No. 2, pp. 212–231, 2020.

[3] Shogo Kawanaka, Yuki Matsuda, Hirohiko Suwa, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. Gamified participatory sensing in tourism: An experimental study of the effects on tourist behavior and satisfaction. *Smart Cities*, Vol. 3, No. 3, pp. 736–757, 2020.

[4] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pp. 1273–1282. PMLR, 2017.

[5] Shusaku Tomita, Jose Paolo Talusan, Yugo Nakamura, Hirohiko Suwa, and Keiichi Yasumoto. Fedtour: Participatory federated learning of tourism object recognition models with minimal parameter exchanges between user devices. In *2022 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)*, pp. 667–673. IEEE, 2022.

[6] 河中昌樹, 冨田周作, 中村優吾, 諏訪博彦, 安本慶一. 観光地のリアルタイム状況説明システムの検討. 2021 年度 情報処理学会関西支部 支部大会 講演論文集, No. 2021, 09 2021.

[7] 松田裕貴, 河中祥吾. Web ブラウザ上で動作する市民参加型写真収集アプリの開発と運用. 第 27 回社会情報システム学シンポジウム（ISS27）, No. 2021, pp. 1–5, 01 2021.

[8] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 4, pp. 640–651, 2017.

[9] Yi Li, Haozhi Qi, Jifeng Dai, Xiangyang Ji, and Yichen Wei. Fully convolutional instance-aware semantic segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4438–4446, 2017.

[10] Alexander Kirillov, Kaiming He, Ross Girshick, Carsten Rother, and Piotr Dollar. Panoptic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241. Springer, 2015.

[12] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 39, No. 12, pp. 2481–2495, 2017.

[13] Leon A Gatys, Alexander S Ecker, and Matthias Bethge. Image style transfer using convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2414–2423, 2016.

[14] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, Vol. 27, , 2014.

[15] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125–1134, 2017.

[16] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.

[17] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779–788, 2016.

[18] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.