

位置情報付き写真の撮影内容に基づいたイラストマップの自動生成

周 文トウ[†] 牛尼 剛聰[‡]

[†]九州大学大学院芸術工学府 〒815-8540 福岡県福岡市南区塩原 4-9-1

[‡]九州大学大学院芸術工学研究院 〒815-8540 福岡県福岡市南区塩原 4-9-1

E-mail: [†]zhou.wentong.904@s.kyushu-u.ac.jp, [‡]ushimama@design.kyushu-u.ac.jp

あらまし イラストマップとは、イラストを利用して地域の特徴を視覚的にわかりやすく表現した地図である。旅行者はイラストマップを通じて、旅行先の魅力を直感的に理解できるようになる。しかし、従来の手書きイラストマップでは、制作コストが高いため、ユーザが興味のある場所やトピックに基づいたイラストマップを利用したいという個別のニーズに応じることは難しい。一方、SNSに投稿された位置情報写真には、撮影者の関心や地域の特徴が反映されていると考えられる。そこで、本研究では、SNSに投稿された写真の位置情報と写真の撮影内容に基づき、地域内でユーザが関心を持っている地点を発見して、ユーザが関心のある地点における代表写真を抽出し、イラストマップを構成する手法を提案する。提案システムにより生成されたイラストマップは観光等様々な目的に応用可能であると考えられる。

キーワード SNS, 地理情報, 画像認識, 可視化

1. はじめに

「イラストマップ」とは、「場所や建物などを絵を用いて大まかに表した、観光案内などの地図」である[1]。図1は手書きで作成されたイラストマップの例である[2]。このようなイラストマップを利用して、旅行者は、この地域の面白さや魅力などを直感的に理解できるようになる。しかし、このようなイラストマップを手書きで作成するためには、多くの時間と労力が必要である。イラストマップを制作するために、イラストレーターは、地域の情報を収集・把握した上で、地域の魅力を絞り込み、そして適切なスタイルを活用して、最後にマップを描く。そのため、ユーザがそれぞれ欲しい地域に関して、特定の興味や関心に基づいたイラストマップを作成することは、現実的ではない。

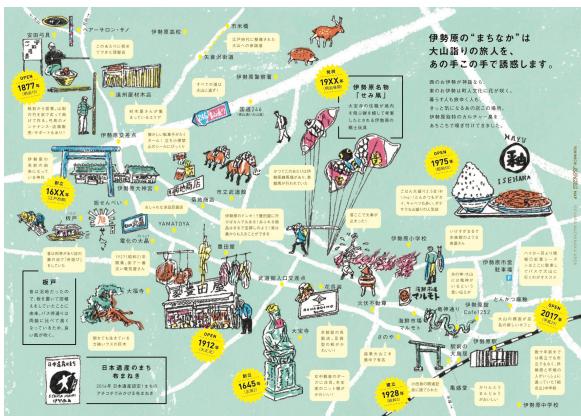


図 1 「ウワサで巡る! よりみち伊勢原 MAP」¹

一方、近年、スマートフォンとSNSの普及に伴い、スマートフォンで撮影した写真をSNS上でシェアす

る行為が日常的に行われるようになった。例えば、代表的な写真SNSであるInstagram[3]では、2020年1月の時点で、月間アクティブアカウント数は10億人を突破し、デイリーアクティブユーザー数も500万を超える[4]。

こうした中で、SNSに投稿された膨大な写真データを分析・可視化することで、写真から反映された潜在的な内容に関する研究が活発に行われている。Redi[5]らはInstagramに投稿された写真の内容と写真のスタイルを分析して可視化することで、全世界有名な五つの大都市間の文化的類似点と相違点を明らかにしている。Manovich[6]らはInstagramに投稿する自撮り写真の内容を分析して可視化することで、世界中の5都市間のユーザの特性を明らかにしている。

また、SNSの写真データを利用して、地域の特徴を反映されるピクトリアルマップの自動生成にする手法が提案されている。その手法の一つとして、高ら[7]は投稿された写真の色情報を分析して、代表写真を抽出して地図にマッピングする手法を提案している。しかし、この手法では、写真の撮影対象が考慮されていないため、適切に代表的な写真を選ぶことができない場合がある。また、代表的な写真として選択された写真が、視覚的に美しい写真であるとは限らないため、得られたマップのクオリティが低くなるという問題が存在する。

そこで、我々は、写真の撮影対象を認識し、それに基づいて代表写真を抽出し、抽出した写真を用いて、適切なマップを自動生成するシステムを開発することを目標に研究を行っている[8, 9]。本論文では、写真の

¹ <http://www.tegakimap.jp/map/3045>

撮影対象に基づく代表写真の選択に基づいて、美的な視点から高品質なイラストマップを自動的に生成する手法を提案する。

本論文の構成は以下の通りである。第2章では我々のアプローチについて述べる。第3章では提案手法について述べ、第4章では実験と結果を述べる。第5章では評価実験を述べ、第6章では本研究のまとめと今後の展望を述べる。

2. アプローチ

本研究の趣旨は、SNSに投稿された写真の位置情報と写真の撮影対象に基づいて、ユーザの関心地点における代表写真を抽出することである。図2に提案手法の流れを示す。本手法では、まず、写真共有SNSであるFlickrから指定した地域の位置情報つき投稿写真をフィルタリングする。次に、写真全体の美しさについて、自動的に評価を行う。そして、写真に付けられたジオタグ情報に基づいて、クラスタリングを行って、ユーザが関心を持っている地点(Point of Interest: POI)を抽出する。その後に、認識された写真の視覚属性に基づいて、写真とPOIを特徴ベクトルに変換する。最後に、写真の特徴ベクトルと地点の特徴ベクトル間の類似度と写真の美しさ、両方を考慮して、写真の適合度を計算する。

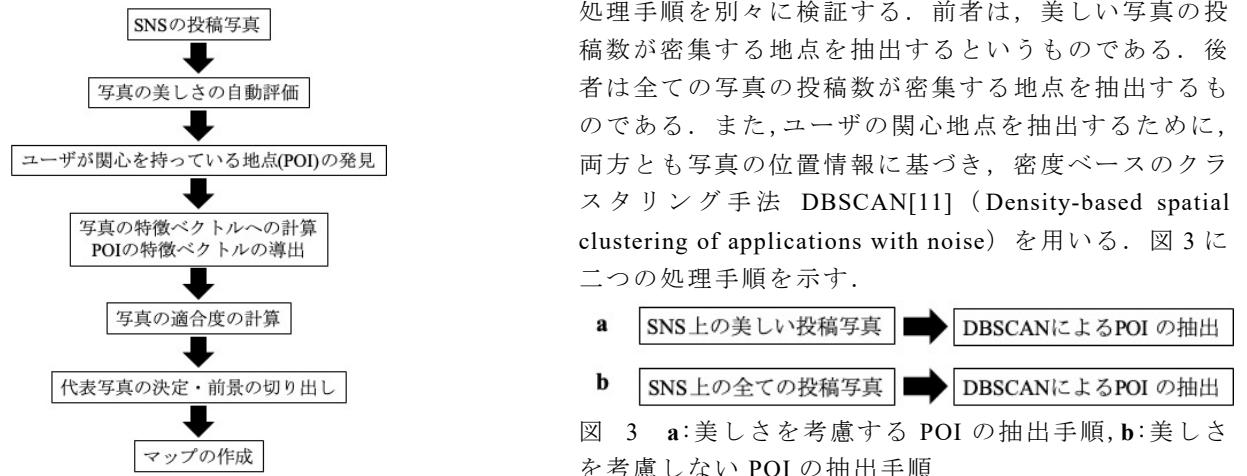


図2 提案手法の流れ

3. 提案手法

3.1. 写真の美しさの自動評価

本研究では、美的な質を有するイラストマップの自動生成システムを開発することを目標にする。そのため、本研究では、写真の美しさと写真の内容という2つの観点から写真の分析を行う。本節では、写真の美しさの自動評価について述べる。

写真が美しいかどうかは人それぞれの基準や価値

観によって決まる。しかし、機械学習の手法の一つであるディープラーニング (Deep Learning) を通して、写真の美しさといった高度な特徴を自動的に抽出できるようになってきた。本研究は、人手で作成された教師データを利用して訓練された画像の美しさ識別するためのディープラーニングモデルであるILGNet[10]を利用して、投稿写真の美しさを自動的に評価する。ILGNetを用いることにより、画像の美しさをスコア化して推定することができる。本手法では、ILGNetを利用して、投稿写真の美しさを表すスコアを求め、そのスコアを写真の美しさの尺度とする。そして、閾値を設定することで、全ての画像を美しく撮る写真と単に記録するため撮られた写真を2種類に分類する。

3.2. POIの抽出

POIは、人々が興味を持つ場所を指す。SNS上で話題になるスポットや、大勢の人が訪ね、チェックインする場所などがPOIの候補となる。POIの付近では写真を撮影・投稿する人数が多くなると考えられる。そこで、本研究では、SNS写真の投稿数が密集する地点を抽出すると、ユーザの関心地点が発見できると考える。

本研究では、クラスタリング手法によりPOIを決定する。クラスタリングする前に、前節に述べた写真の美しさを考慮するか、考慮しないかによって、二つの処理手順を別々に検証する。前者は、美しい写真の投稿数が密集する地点を抽出するというものである。後者は全ての写真の投稿数が密集する地点を抽出するものである。また、ユーザの関心地点を抽出するために、両方とも写真の位置情報に基づき、密度ベースのクラスタリング手法 DBSCAN[11] (Density-based spatial clustering of applications with noise) を用いる。図3に二つの処理手順を示す。

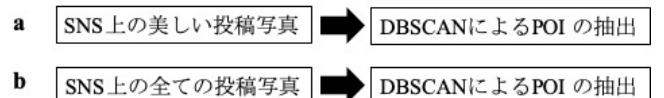


図3 a:美しさを考慮するPOIの抽出手順, b:美しさを考慮しないPOIの抽出手順

DBSCANは、クラスタ間の距離の閾値 Eps とクラスタを構成する最小データ数の閾値 $MinPts$ との2つの閾値を持つ。ある点 x から、距離 Eps 内にある点集合を近傍 $N_{Eps}(x)$ と定義し、以下の接続関係を満たす時、同じクラスタに分類する。

$$(1) y \in N_{Eps}(x)$$

$$(2) |N_{Eps}(x)| \geq MinPts(\text{コア点})$$

つまり、ある座標から半径 Eps 以内に $MinPts$ 以上の座標集合が存在するなら、同じクラスタに分類する。

また、生成されたクラスタに対して、それぞれのクラスタの中心点を求める。クラスタ内のスポットに属するそれぞれの写真の撮影位置の緯度と経度を (ϕ_i, λ_i) 、クラスタの重心を (ϕ_o, λ_o) とする。 $\Delta\phi$, $\Delta\lambda$ がその差の絶対値とした際に、2 点間の中心角 $\Delta\sigma$ は球面余弦定理より以下の式で求める。

$$\Delta\sigma = \arccos(\sin \phi_i \cdot \sin \phi_o + \cos \phi_i \cdot \cos \phi_o \cdot \cos(\Delta\lambda))$$

大円距離 d 、すなわち円弧長は、球の半径 r 、弧度で表された $\Delta\sigma$ を用いて以下の式で求める。

$$d = r\Delta\sigma$$

クラスタ C の中心点 (x_c, y_c) を以下の式で求める。

$$(x_c, y_c) = Min(d_1, d_2, \dots, d_i, \dots, d_n)$$

つまり、スポットと重心の大円距離をそれぞれに計算して、距離は最小となるスポットをクラスの中心点として判断する。この中心点は指定地域のマップに提示する POI とする。

ただ、本手法では、以下の問題点を解決していないため、これらに対する対応は今後の課題である。

- (1) 少数のユーザが短期間で大量かつ同じ内容の写真を投稿することに対する対応
- (2) 写真に付与された撮影日時情報を考慮すること

3.3. 代表的な写真の抽出

3.3.1. 画像認識による画像の視覚属性の識別

Hwang[12]らは、画像の視覚属性 (visual attribute) を、物体カテゴリ間で共有される人間が理解可能な属性であると定義している。本研究では、画像認識技術を用いて、画像の視覚属性を識別する。近年、ディープラーニングを用いて高精度に画像認識を行うことができるようになった。そして、いくつかの代表的な画像認識アルゴリズムが WebAPI として利用できるようになっている。本研究では、代表的な画像認識 API の一つである Google の Cloud Vision API[13]を利用して、対象となる全ての画像に関して画像認識を行う。Cloud Vision API によって得られたタグ情報 (label annotations) の例を、図 4 に示す。本研究では、score が 0.8 以上のタグ情報を画像の視覚属性とする。

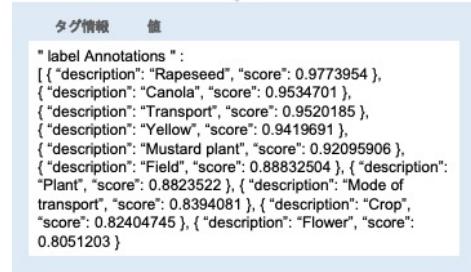


図 4 Cloud Vision API による取得されたタグ情報 (label annotations) の例

3.3.2. 写真の特徴ベクトルの計算

各画像を得られた全ての視覚属性 (単語) において、どの特徴が地域の特徴を決めるのか明らかにするため、それぞれの視覚属性の単語の重要度を計算する。ここで、キーワードの重要度を表現する TF-IDF と同じアプローチを利用する。本研究では、各視覚属性を一つのキーワードとし、各写真を一つの文書とし、式(1) のように算出する。 af はある写真に出現する視覚属性の頻度、 isf はある POI に全ての写真数/視覚属性が出現する写真の総数の対数 (Inverted Spot Frequency) であり、 $afisf$ はその積となる。

$$afisf_{i,j} = af_{i,j} \cdot isf_i \quad (1)$$

$$af_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}}$$

$$isf_i = \log \frac{|S|}{sf_i} + 1$$

$$sf_i = |\{s: attr(s) \ni a_i\}|$$

つまり、ある POI で、どの写真でもよく出現する視覚属性の単語は、例えば木や空など、より一般的な特徴と認められるため、重要度を低くする。一方、特定の写真でしか出現しない視覚属性の単語は、重要度を高くする。そこで、式(2) のように、計算された単語の重さ $afisf$ を使って、画像を特徴ベクトルに変換する。

$$V_k^{image} = (afisf_{1,k}, afisf_{2,k}, \dots, afisf_{n,k}) \quad (2)$$

表 1 に、写真の特徴語と各特徴の $afisf$ 値の一部の結果

の例を示す。

表 1 写真の特徴語と各特徴の *afisf* 値の一部の結果

画像	特徴語	<i>afisf</i> 値
	Lavender	0.367758986
	Purple	0.297781691
	Wisteria	0.444284834
	Tree	0.239813676
	Plant	0.392780745
	Flower	0.282643273

	Fireworks	0.344897105
	New Year's Day	0.298454239
	Night	0.466857812
	Midnight	0.302096124
	Sky	0.351812205
	Event	0.326701577

	Waterway	0.339595604
	Canal	0.38320493
	Water	0.334651182
	Bank	0.239056573
	Vehicle	0.38320493
	Tree	0.448958463

	Flower	0.329225172
	Flowering plant	0.232030533
	Plant	0.25055317
	Petal	0.295596221
	Purple	0.287154732
	Violet	0.226303094

3.3.3. POI の特徴ベクトルの導出

前節ではそれぞれの POI における全ての写真を特徴ベクトルに変換して、写真の特徴ベクトルを得る手法を説明した。本節では、POI の特徴ベクトルの導出方法を述べる。式(2) を示すように、POI の特徴ベクトルは、POI における各写真の特徴ベクトルを平均したものとして定義する。

$$V_j^{spot} = \frac{1}{|I(s_j)|} \sum_{k \in I(s_j)} V_k^{image} \quad (3)$$

3.3.4. 代表的な写真の選択

ここでは、代表的な写真の選択方法について述べる。まず、コサイン類似度を用いて写真の特徴ベクトルと POI の特徴ベクトル間の類似度を算出する。写真の特徴ベクトルを V_j^{spot} 、POI の特徴ベクトルは V_k^{image} とした場合、写真と POI の類似度を以下のように定義する。

$$\text{cosine_similarity}(V_j^{spot}, V_k^{image}) = \frac{V_j^{spot} \cdot V_k^{image}}{|V_j^{spot}| |V_k^{image}|} \quad (4)$$

また、本研究では、一定の美的質を有するイラストマップを生成するため、写真の視覚属性だけでなく、写真の美しさも考慮すると考える。そこで、3.1 で自動評価された写真のスコアを写真の美しさ尺度として使う。もう一つの尺度は、コサイン類似度とする。以下に示す式で写真の適合度を算出する。

$$\text{total score}_{image} = \alpha \text{ cosine_similarity} + \beta \text{ aesthetic_score}$$

適切な α と β を決定するため、検証を行う。表 2 に示した五つのパターンで各 POI 内の全ての写真の適合度をそれぞれに計算する。

表 2 五つのパターン

	α	β
パターン 1	0	1
パターン 2	0.25	0.75
パターン 3	0.5	0.5
パターン 4	0.75	0.25
パターン 5	1	0

最後に、計算した $\text{total score}_{image}$ をランクインすることで、順位一枚の写真を代表的な写真として決定する。

4. イラストマップの自動生成

本章では、3 章で述べた提案手法に基づき、イラストマップの自動生成システムのデータ収集、生成されたイラストマップについて説明する。

4.1. データセット

ここでは Flickr にアップロードされた画像から構成された全種類の写真のデータセット YFCC100M を使用した。YFCC100M[14]は、世界中で撮影された約 1 億枚の写真及び動画を含んでおり、その中の約 5,000 万件は位置情報付き写真である。本研究では、福岡県と京都府を対象とするピクトリアルマップを生成するため、写真の位置情報（緯度）を用い、23,073 枚の福岡県内の写真と 11,769 枚の京都府の写真をフィルタリングした。また、写真の美しさの自動評価を行って、美しさスコア 0.5 以上の写真 5579 枚を美しい写真として分類した。図 5 と図 6 に ILGNet によって美しさの自動評価を行った結果の例を示す。



図 5 美しさスコア 0.85 以上の分類した美しい写真の例



図 6 美しさスコア 0.05 以下の分類した美しくない写真の例

4.2. イラストマップの作成

DBSCAN を用いて POI を抽出した時、パラメータの値は表 3 のように実装した。

表 3 パラメータ

5,579 枚の美しい写真から POI の抽出	$Eps = 2\text{km}/6371.0088^2, MinPts = 20$
23,073 枚の写真から POI の抽出	$Eps = 2\text{km}/6371.0088, MinPts = 50$
117,629 枚の写真から POI の抽出	$Eps = 0.5\text{km}/6371.0088, MinPts = 100$

また、各 POI における代表的な写真を算出した。そして、Background Removal API[15] を用い、代表写真の前景を分離する処理を行った。図 7 に処理手順を示す。

図 10 と図 11 には写真の撮影対象を地図にマッピングすることで生成された福岡県のイラストマップの例を示す。図 12 には京都府のイラストマップの例を示す。

写真

撮影対象を抽出

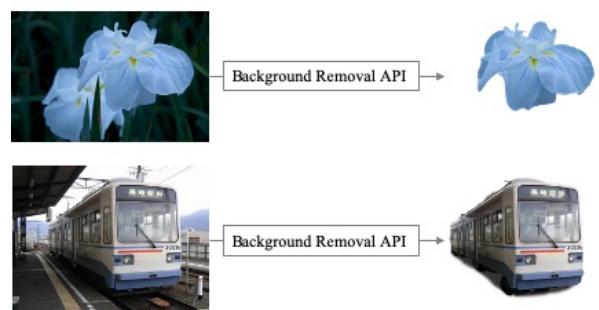


図 7 代表写真の前景を切り出す処理手順

5. 評価実験

本章では、提案手法により生成されたイラストマップの有効性を検証するため、評価実験を行う。

5.1. 抽出された POI の有効性の検証

抽出された POI の有効性を検証するため、提案手法に基づいて抽出した POI と人間によって手動で抽出された“絶景スポット”（インターネットで公開されている二つの記事「福岡県の絶景スポット」[16]、「福岡県の写真撮影スポット」[17]に載っている 51 個のスポット）と比較し、適合率と再現率を以下の式を用いて算出した。

$$\text{適合率} = \frac{\text{Web サイトに載っている POI と一致した POI の総数}}{\text{POI の総数}}$$

$$\text{再現率} = \frac{\text{Web サイトに載っている POI と一致した POI の総数}}{\text{Web サイトに載っている POI の総数}}$$

4.2. イラストマップの作成

DBSCAN を用いて POI を抽出した時、パラメータの値は表 3 のように実装した。

表 3 パラメータ

図 8 には本研究により生成された二つのマップ「美しい写真に基づいたイラストマップ」と「美しさにかかわらず写真に基づいたイラストマップマップ」の適合率と再現率の結果を示す。

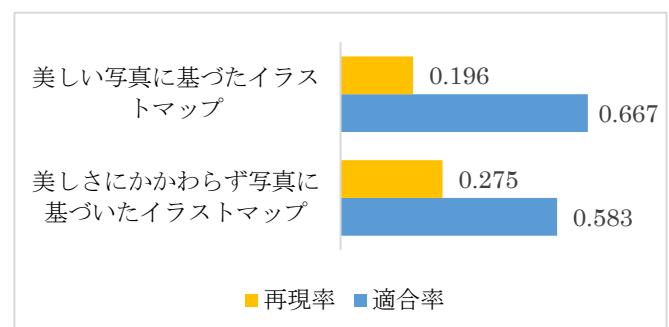


図 8 POI の再現率と適合率

結果を見ると、美しい写真から抽出した POI の方が、

² 1 ラジアンのキロメートル数

適合率が高く、再現率は低かった。美しさにかかわらず全ての写真データから抽出したPOIの方が再現率がより高く、適合率は少し低かった。両方とも抽出したPOIの数が少ないため、再現率が良くないことがわかった。今後の課題として、SNS上の投稿写真の分析だけでなく、テキスト文の分析に加えて、最適なPOIの抽出方法を検討する必要がある。

5.2. 生成されたイラストマップの有効性の検証

生成された3つのイラストマップの有効性を検証するため、被験者実験を行った。

被験者は福岡県福岡市に在住する6人である。生成された3つのイラストマップ（地図1は図10に示す美しい写真を利用して生成された福岡県のイラストマップであり、地図2は図11に示す全ての写真を利用して生成された福岡県のイラストマップであり、地図3は図12に示す全ての写真を利用して生成された京都府のイラストマップである）を被験者に提示し、被験者に以下の2つの質問について5段階（「全くそう思わない」、「あまりそう思わない」、「どちらとも言えない」、「ややそう思う」、「とてもそう思う」）の尺度を用いた回答欄を設け、以下の質問に対して回答してもらった。

質問1: 地図1は福岡県の魅力が反映されているか？

質問2: 地図2は福岡県の魅力が反映されているか？

質問3: 地図2は京都府の魅力が反映されているか？
実験結果を図9に示す。

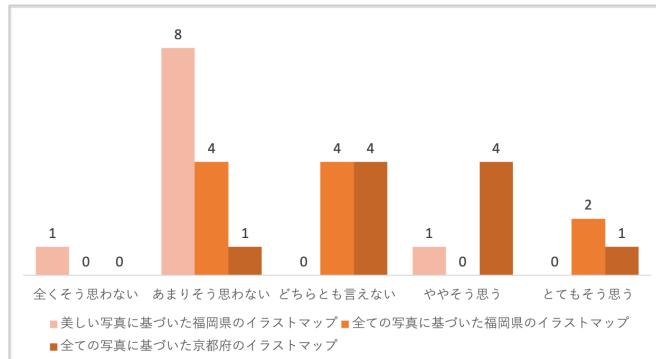


図9 被験者実験の結果

被験者実験の結果を見ると、まず、福岡県の2つのイラストマップを比較すると、「美しい写真に基づいたイラストマップ」よりも「全ての写真に基づいたイラストマップ」の方が地域の魅力がより反映されていると感じる被験者が多い傾向が見られた。これは、被験者はマップを構成する写真の美しさより内容の豊富さを重視することがわかった。つまり、提案手法により美しさに基づいて写真をフィルタリングすることにより、対象となる写真の数が減少してしまうことにより、

地域の特徴が失われてしまうことが明らかになった。本手法において、写真の美しさによって対象写真をフィルタリングすることは、対象とする写真からノイズを除去することを本気としていたが、ノイズ以外の写真もフィルタリングてしまっていると考えられる。今回の提案では、美しさに基づいた写真のフィルタリングは適切に機能しているとは言えないが、写真の質の判断に深層学習が利用できる可能性があることがわかったため、今後、美しさ以外の基準でのフィルタリングを検討していく予定である。全体の傾向から見ると、提案手法により生成されたイラストマップは地域の魅力を示すことが可能であると考えられる。

6. まとめと今後の展望

SNSに投稿された位置情報付き写真からイラストマップを自動生成する手法を提案し、評価実験により有効性を検証した。

今回は、出力されたマップでは、それぞれのPOIにおける代表写真の一部分をそのまま表示しているが、抽出されたオブジェクトのスタイルを変換することにより、表現に統一性があるより魅力的なイラストマップを構成できると考えられる。

参考文献

- [1] 松村明,『大辞林第三版』, 1988.
- [2] 「ウワサで巡る！よりみち伊勢原 MAP」: <http://www.tegakimap.jp/map/3045>.
- [3] “Instagram”, <http://instagram.com/>.
- [4] “Instagram Statistics”, <https://instagram-press.com/our-story/>.
- [5] Miriam Redi, Damon Crockett, Lev Manovich, Simon Osindero, “What Makes Photo Cultures Different?”, Proceedings of the 24th ACM international conference on Multimedia, pp. 287-291, 2016.
- [6] L. Manovich, M. Stefaner, M. Yazdani, D. Baur, D. Goddemeyer, A. Tifentale, N. Hochman, and J. Chow, “Selfiecity”, <http://selfiecity.net/#intro>.
- [7] 高尚暉, 牛尼剛聰, “SNSを利用したピクトリアルマップの自動生成”, DEIM Forum 2017, D7-4, 2017.
- [8] 周文彤, 牛尼剛聰, “投稿写真の撮影内容に基づく地域の特徴を表すピクトリアルマップの自動生成”, 第6回ADADA Japan 学術大会(ADADA Japan 2019), pp. 8-11, 2019.
- [9] Wentong Zhou, Taketoshi Ushijima, “Automatic Generation of Pictorial Maps from Photos on Social Media to Represent Regional Features”, The 14th International Conference on Ubiquitous Information Management and Communication, 2020.
- [10] Xin Jin and Jingying Chi and Siwei Peng and Yulu Tian and Chaochen Ye and Xiaodong Li, “Deep image aesthetics classification using inception modules and fine-tuning connected layer”, 8th International Conference on Wireless Communications & Signal Processing, pp.1-6, 2016.

- [11] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu, "A density-Based Algorithm for Discovering Clusters", KDD-96, pp. 226-231, 1996.
- [12] S. J. Hwang, F. Sha, K. Grauman, "Sharing Features Between Objects and Their Attributes", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 1761-1768, 2011.
- [13] "Cloud Vision API",
<https://cloud.google.com/vision/docs/?hl=ja>.
- [14] B. Thomee, D.A. Shamma, G. Friedland, B. Elizalde, K.Ni, D. Poland, D. Borth and L. J. Li, YFCC100M: The new data in multimedia research, Communications of the ACM, Vol.59, No.2, pp.64-73, 2016.
- [15] "Background Removal API",
<https://www.remove.bg/api>.
- [16] "福岡県の絶景スポット",
<https://zekkeiproject.com/areas/countries/153/prefectures/44/spot>.
- [17] "福岡県の写真スポット",
<https://ganref.jp/spot/photo/jpn/fukuoka.html>.



図 10 美しい写真に基づいた福岡県のイラストマップ

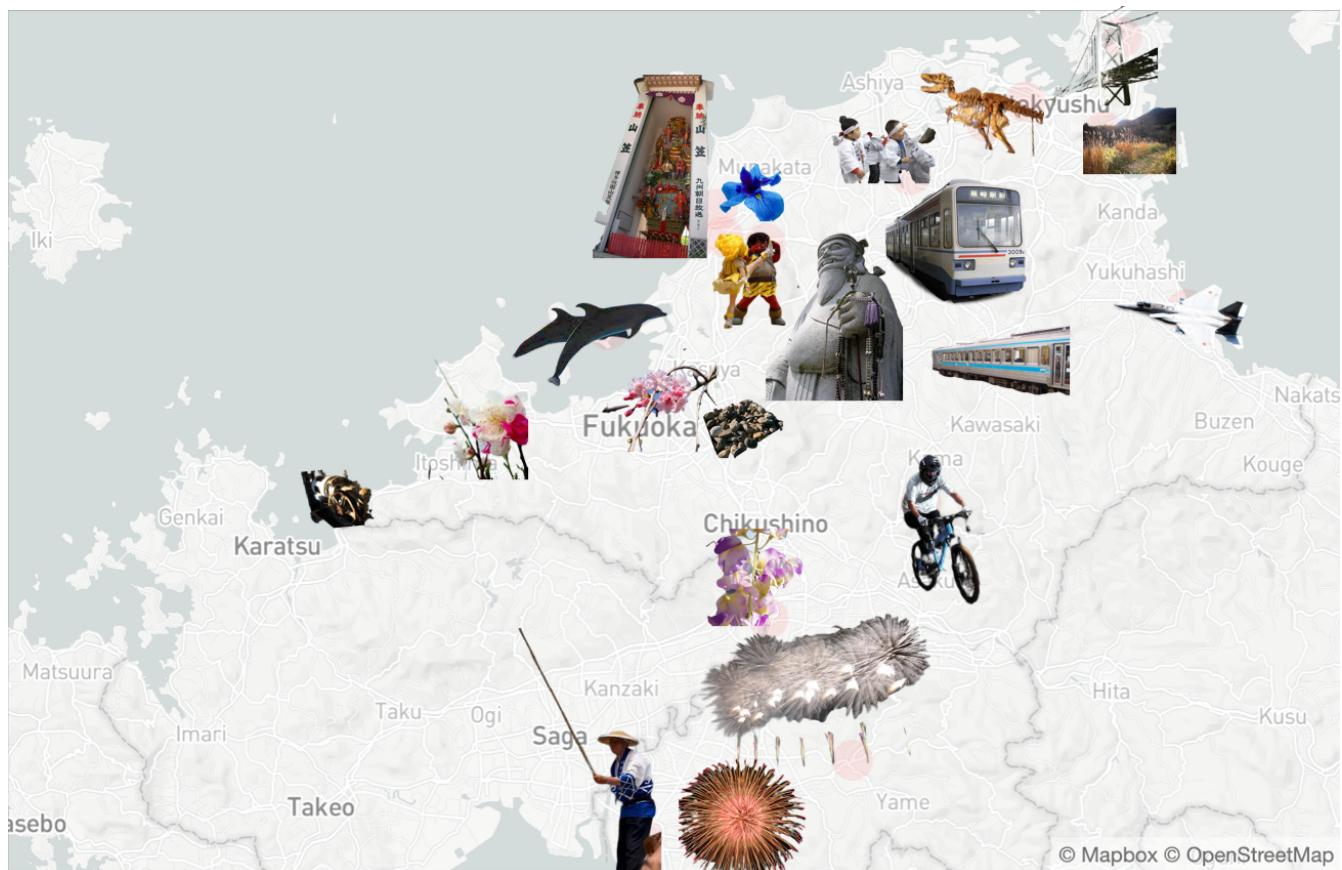


図 11 美しさにかかわらず写真に基づいた福岡県のイラストマップマップ

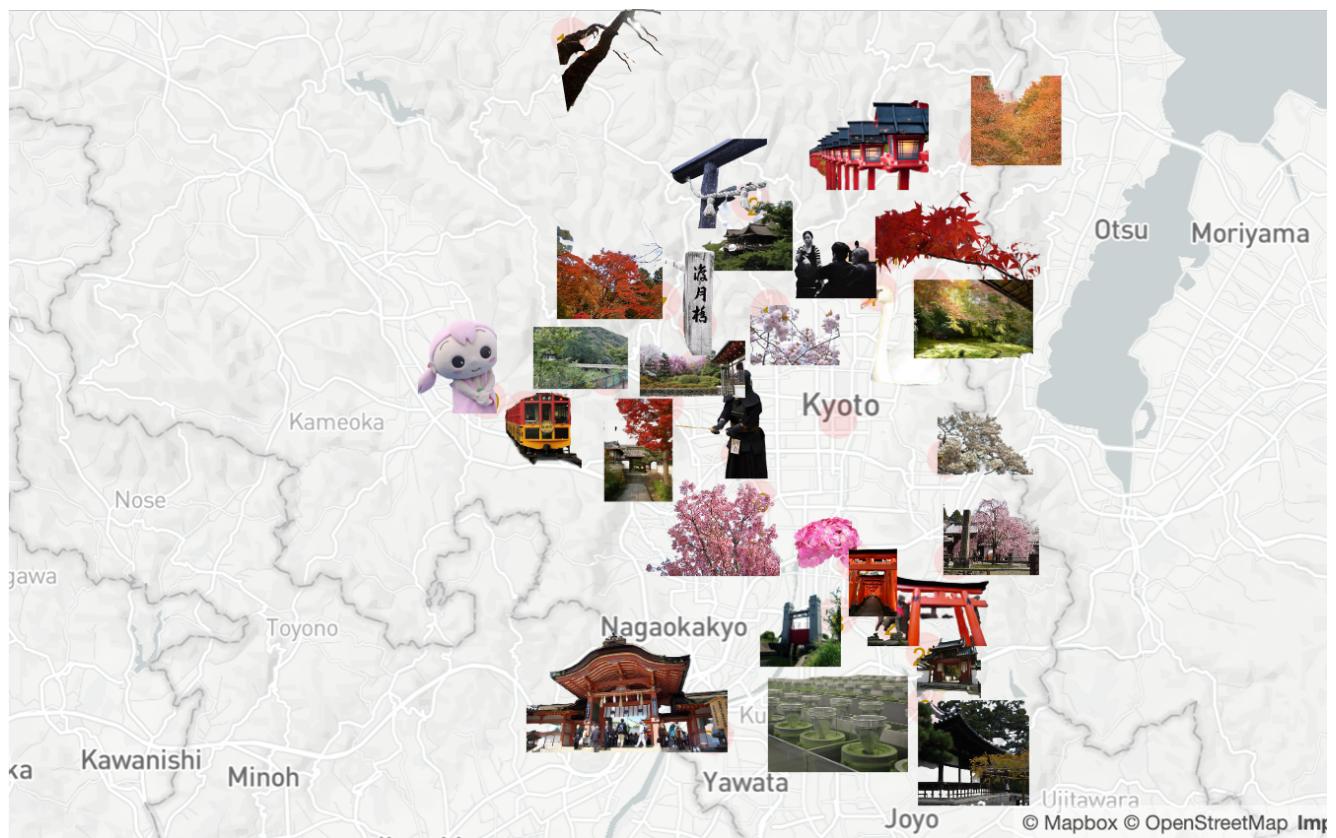


図 12 美しさにかかわらず写真に基づいた京都府のイラストマップマップ