

シーズン固有性に基づく主要撮影スポットの可視化

Visualizing Photo Spots Characteristic of Seasons

橋本和哉^{1*} 熊野雅仁² 木村昌弘²
Kazuya Hashimoto¹ Masahito Kumano² Masahiro Kimura²

¹ 龍谷大学理工学研究科電子情報学専攻

¹ Division of Electronics and Informatics, Ryukoku University

² 龍谷大学理工学部電子情報学科

² Department of Electronics and Informatics, Ryukoku University

Abstract: 近年, WEB 空間に共有された撮影位置と時間情報を持つ大量の写真データを用いて, 様々な都市の観光支援を行う研究が注目されている. 本研究では, 指定された都市に特有のシーズン群を検出し, さらに各シーズンにおける主要撮影スポット群を抽出して, 各主要撮影スポットのシーズン固有性とアノテーションを与える手法を構築し, それに基づいた観光マップ可視化法を提案する. Flickr データを用いた実験で有効性を検証する.

1 はじめに

近年, 様々な地域の施設や観光先を訪問した一般の人々により, Geo-tag が付随した訪問先の情報, テキスト情報や評価情報, 写真情報などが SNS やソーシャルメディアサイトを通じて WEB 空間に投稿され, ビッグデータとして成長し続けており, 新たな利活用が模索されている. ところで, TripAdvisor は, 様々な地域を訪問した人々のレビューを活用する口コミサイトであり, ユーザは, 訪問先の行動プランニングを行う際, 事前に観光名所や施設を検索することで, 注目した観光名所や施設に関する口コミから有益な情報を獲得できる可能性がある. しかし, 検索で一覧される観光名所は著名なものが多く, 施設という単位で観光先が扱われる傾向が高い. 近年, 訪日外国人の行動から, 意外な場所が観光スポットになり得ることが知られるようになったが, 人々は居住地と文化が異なる地域に訪れて実際に見たことのないものに興味を示す傾向がある. そのため, 街並み, 海や山, 公園, 道沿いや川沿いなど, 必ずしも施設が対応しない多様なスポットを観光スポットに含めることが望ましく, WEB 空間で共有化された人々の Geo-tag 付きの行動情報から観光スポットを見出す方法が注目されつつある.

Geo-tag が付随した Twitter のつぶやきというテキスト情報から観光地を特定する方法として, テキスト情報における地域特有の固有性を持つ単語に着目し, 居住地と移動先の地域を峻別し, さらに観光地との関

連度が高い地点の集合をカテゴリーとして抽出する研究 [1] がある. しかし, Twitter のつぶやきは, テキスト情報で扱う対象の地図上の位置と, つぶやいた位置が近傍である保証はないため, 推薦する地域が都市レベルになるなど, 空間解像度が粗くなる傾向がある.

一方, 写真を撮影する場合, 主に撮影地点の近傍に興味の対象が存在する可能性が高く, 仮に撮影対象が遠くの山や景色など遠い位置に存在しても, 写真は撮影者の心をつかむ対象に遭遇したとき撮影されることが多いことから, Geo-tag に記録された地点で撮影すること自体が観光スポットと密接に関連している可能性が高い. Crandall ら [2] は, 写真共有サイトに蓄積された大量の Geo-tag 付き写真を用いて年間を通じて人々が集まる主要撮影スポットを抽出した. しかし, 日本のように季節に応じて撮影スポットが変化し得る場合, 単純に年間を通じた撮影数の多さに着目すると, ある特定の時期のみ魅力的な撮影スポットを見逃す可能性がある.

我々は, 地域と期間のペアに着目して日常の撮影行動に対し, 異常性のある地域と期間を抽出する研究 [3] や, 主要撮影地域に対し, 特有の旬シーズン群を抽出し, 短い期間に多くの人々が訪れるというバースト性の観点から強制的に旬シーズンを可視化する研究 [4] を行ってきた. 本研究では, 旅行者の行先となる都市と時期が定まっているとき, その時期が含まれるシーズンにおいて, 年中人気のある主要撮影スポットだけでなく, そのシーズン特有の主要撮影スポットがどこにあるかを識別しながら複数の観光先に関する回遊を検討できるように, 都市全体の撮影スポットを効果的に地図上へマッピングするシーズン固有性に基づく主要撮影スポットの観光ダイジェスト可視化法を提案する.

*連絡先: 龍谷大学
滋賀県大津市 瀬田大江町横谷 1-5
E-mail: t15m020@mail.ryukoku.ac.jp

Flickr から収集した大量の Geo-tag 付き写真データを
を用いた実験で、提案法の有効性を検証する。

2 観光ダイジェストマップ

Flickr などの写真共有サイトにアップされた地理お
よび時間情報をもつ写真群に基づいて、指定された都
市に対する 1 年間の観光ダイジェストマップの構築を
考える。その都市で最近 y 年間に撮影された写真デー
タ全体の集合を $\mathcal{D} = \{d_n; n \in \mathcal{N}\}$ とし、これを用いて
その 1 年間の観光ダイジェストマップを構築する。こ
こに、 y は与えられた正の整数である。また、各写真デー
タ d_n には地理情報 $x_n = (x_{n,1}, x_{n,2})$ と時間情報 t_n が
付随している。ただし、 $x_{n,1}$ と $x_{n,2}$ はそれぞれ写真 d_n
が撮影された緯度と経度、 t_n は d_n が撮影された日であ
る。緯度と経度の情報を用いれば、対象とする都市内の
位置は 2 次元 Euclid 空間 $\mathbf{R}^2$ 内の領域 Ω 内の点と同
一視される¹ ことに注意しておく。

Crandall らの研究 [2] と同様、写真数よりも撮影者数
を重視するという観点から、領域 Ω を $5m$ 四方の単位
矩形領域の集まりとして離散化し、その各単位矩形領
域においては、1 人の撮影者が同じ日に複数の写真を撮
影していたとしてもそれを 1 枚とカウントすることに
する。このようにして得られる写真データ全体の集合
と撮影日全体の集合をそれぞれ、 $\mathcal{D}' = \{d_n; n \in \mathcal{N}'\}$ 、
 $\mathcal{T} = \{t_n; n \in \mathcal{N}'\}$ とする。本研究では、まず、対象と
する都市に対し、 $\mathcal{T}$ に基づいて、その都市に特有のシー
ズン群 $\{S_k; k = 1, \dots, K\}$ を検出する。次に、各シー
ズン S_k に対し、その時期に撮影された写真全体の集合
 $\mathcal{D}_k = \{d_n \in \mathcal{D}'; t_n \in S_k\}$ に基づいて、その都市のシー
ズン S_k における主要撮影領域群 $\{R_{k,j}; j = 1, \dots, J_k\}$
を抽出する。さらに、各主要撮影領域 $R_{k,j}$ のシーズン
固有度および各撮影スポット x_n のシーズン固有度を
定義し、それらに基づいて各シーズンの主要撮影領域
群を地図上に可視化することにより、1 年間の観光ダイ
ジェストマップを構築する。特に、固有度の高いものや
低いものを、 $\mathcal{D}$ に含まれている写真データを用いてア
ノテーションする。

3 提案法

3.1 シーズン検出

写真が撮影された日全体の集合 $\mathcal{T}$ に基づいて、対象
とする都市に特有のシーズン群 $\{S_k = [s_{k,0}, s_{k,1}]; k =$
 $1, \dots, K\}$ を検出することを考える。最近 y 年間のデー

¹本研究の実験においては、日本近郊を対象としているため、緯度
と経度情報を平面直角座標系 (平成十四年国土交通省告示第九号) の
手法を利用し、平面直角座標に変換している。

タから 1 年間の平均的な振る舞いを調べるので、12 月
31 日と 1 月 1 日は隣接するという時間軸における周期
性を導入する。そして、時間軸における周期性の下で
熊野らの手法 [4] を用いて、 $\mathcal{T}$ に基づき 1 月 1 日から
12 月 31 日をクラスタリングすることにより、シーズン
群 $\{S_k = [s_{k,0}, s_{k,1}]; k = 1, \dots, K\}$ を検出する。ここ
に、 $S_1 \cup \dots \cup S_K$ は 1 月 1 日から 12 月 31 日までのすべ
ての日であり、 $k \neq k'$ ならば $S_k \cap S_{k'} = \emptyset$ である。本手
法は、バンド幅を 15 日に固定した Epanechnikov カー
ネルによるミーンシフト法を用いる。

3.2 主要撮影領域の抽出

検出された各シーズン S_k に対し、期間 S_k に撮影さ
れた写真全体の集合 $\mathcal{D}_k$ が定義する Ω 上の点群に基づ
いて、その都市の S_k における主要撮影領域群を抽出す
ることを考える。まず、Crandall らの研究 [2] と同様、
ランドマークスケールとしてバンド幅を $100m$ に固定
し、Epanechnikov カーネルによるミーンシフト法を用
いて、 $\mathcal{D}_k$ が定義する Ω 上の点群をクラスタリングす
る。そして、与えられた値 N_0 以上の数の点 (すなわち、
写真) をもつクラスタ群 $\{C_{k,j}; j = 1, \dots, J_k\}$ を抽出
する。ここに $C_{k,j}$ の要素は、 Ω 上の点であり、写真の撮
影位置を表していることに注意する。次に、各 $C_{k,j}$ に
対して、 $C_{k,j}$ に属するすべての点を含む凸包 $R_{k,j}$ を構
築する。我々は、 $\{R_{k,j}; j = 1, \dots, J_k\}$ をその都市の S_k
における主要撮影領域群として抽出する。

3.3 シーズン固有性

シーズン S_k における主要撮影領域 $R_{k,j}$ の重要性を
示すための一つとして、それが他のシーズンにおける
主要撮影領域と比べてどのくらい固有の存在であるか、
すなわち、どのくらい珍しいものであるかを定量化す
ることを考える。我々は、シーズン S_k における主要撮
影領域 $R_{k,j}$ のシーズン固有度 (Season characteristic
degree) $E_{k,j}$ を、そのシーズンにおけるその領域での写
真の撮影位置全体 $C_{k,j}$ のうち、他のシーズンにおける
主要撮影領域に含まれないものの割合、すなわち、

$$E_{k,j} = \frac{\left| \left\{ x \in C_{k,j}; x \notin \bigcup_{k' \neq k} \bigcup_{j'=1}^{J_{k'}} R_{k',j'} \right\} \right|}{|C_{k,j}|}$$

で定義する。

また、シーズン S_k における主要撮影領域 $R_{k,j}$ 内に
ある撮影スポット x_n に対し、その重要性を示すための
一つとして、それがシーズン S_k にどのくらい固有であ
るかを定量化することを考える。我々は、そのような写
真の撮影位置 x_n のシーズン固有度 $e_{k,n}$ を、 S_k 以外の

表 1: Flickr から収集した写真データ

実験範囲	撮影期間	総撮影 写真枚数	離散化後 写真枚数
京都市	2013 ~ 2015 年	223,464	114,658
東京都	2013 ~ 2015 年	892,443	488,914

表 2: 各データセットにおけるシーズン検出結果

実験範囲	シーズン数	主要撮影領域数
京都市	11	1,540
東京都	15	6,746

シーズンのうち, x_n を主要撮影領域に含まないシーズンの割合, すなわち

$$e_{k,n} = \frac{K - b_n}{K - 1}$$

で定義する. ここに, b_n は x_n を主要撮影領域に含むシーズンの総数である. x_n が一つのシーズンの主要撮影領域にしか含まれないならば $e_n = 1$ であり, x_n がすべてのシーズンにおいて主要撮影領域に含まれるならば $e_n = 0$ である.

写真の撮影位置 x_n が主要撮影領域 $R_{k,j}$ に含まれるか否かの判定については, $R_{k,j}$ が凸包であることを利用し, Crossing Number Algorithm を用いて行う.

3.4 可視化法

指定された都市において検出された各シーズン S_k に含まれる主要撮影領域群 $\{R_{k,j}; j = 1, \dots, J_k\}$ を地図上に描画するうえで, 各 $R_{k,j}$ に含まれる写真の撮影位置 $x_n \in R_{k,j}$ に基づいて凸包を求め, その領域に対し, シーズン固有度 $E_{k,j}$ に応じて色づけを行う. シーズン固有度が最も低い 0 のときを青とし, 固有度が上がるにつれ水色, 緑, 黄色, 橙と変換し, シーズン固有度が最大 1 のときを赤とする. つまり, 凸包領域の色によりシーズン固有性をアノテーションする.

また, 撮影地点のうち, $e_{k,n}=1$ である各撮影位置 x_n は, それを含む凸包以外, 1 年間のすべてのシーズンで求められたどの凸包領域にも含まれなかった観点で固有性の観点から注目すべきであると考え, $E_{k,j}$ とは独立にシーズン固有度 $e_{k,n}$ が最大値 1 をとる主要撮影スポット x_n として紫色の小円で可視化する. また, 対比的に, すべてのシーズンで抽出された凸包に含まれるシーズン固有度 $e_{k,n}$ が最小値 0 をとる年中人気のある主要撮影地域のスポット x_n についても紺色的小円で可視化する.

ユーザは, 可視化された撮影スポットを示す小円を地図上でクリックすることにより, 写真を閲覧することが

できる. また, ユーザは, 観光ダイジェストマップ上でシーズンごとに主要撮影領域 R_k や主要撮影スポット x_n を視認できるだけでなく, 凸包の色に透明度を設定できるため, すべてのシーズンの領域を重ねて主要撮影領域や主要撮影スポットを見ることがもできる.

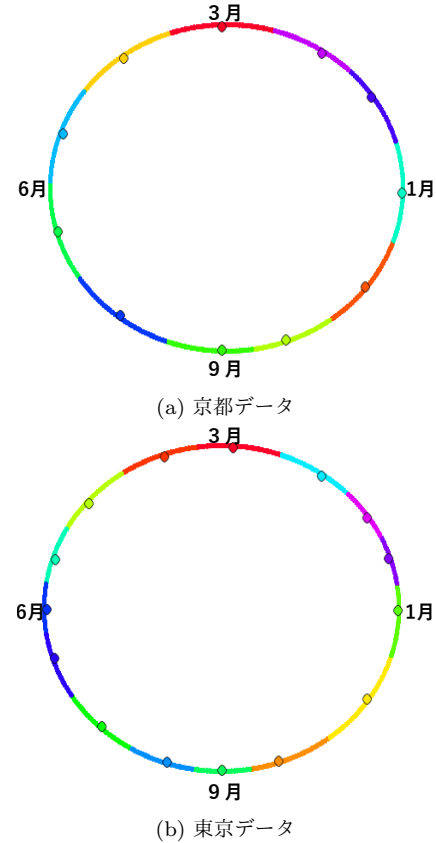


図 1: シーズン検出における各シーズンの期間

4 実験

4.1 データ

本提案法による可視化法の有効性を検証するため実データを用いた実験を行った. 本研究では, ユーザから指定される都市として, 日本でも有数の観光地である京都市と東京都を対象とし, 写真共有サイト Flickr から京都市 (WoID=15015372) と東京 (WoID=2345889) の写真データを収集した. また, シーズンの検出においては, 最近 y 年間のデータから 1 年間の平均的な振る舞いを調べるうえで, 表 1 に示すように 2013 年から 2015 年の $y=3$ 年間のデータを収集し実験を行った. ここで, 京都市に関するデータセットを京都データセットと呼び, 東京都に関するデータセットを東京データセットと呼ぶ.

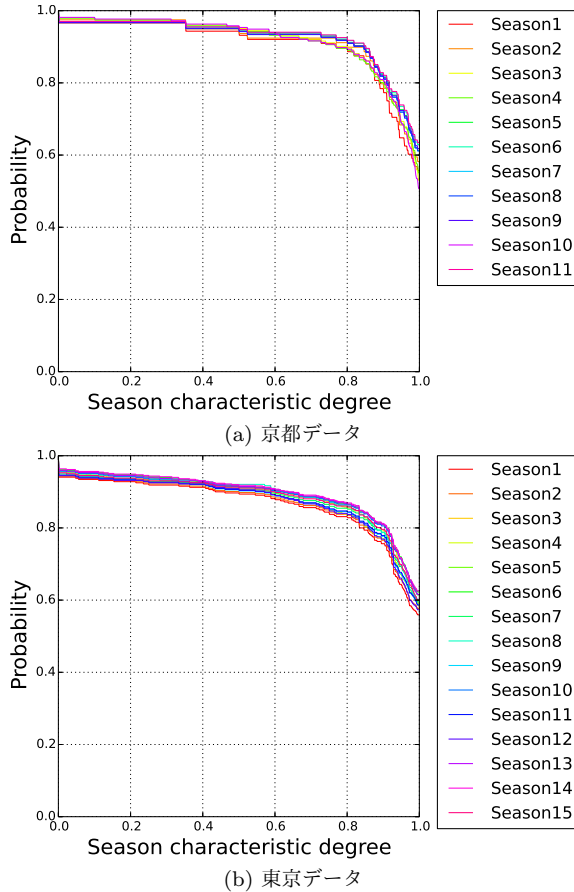


図 2: 主要撮影領域のシーズン固有度分布

4.2 シーズン検出の結果と固有度分布

京都データセットと東京データセットにおけるシーズン検出結果を表 2 に記す。また、図 1 に検出された各シーズンに異なる色を与えた結果を示す。京都データは 11 シーズン、東京データは 15 シーズン分解されているが、いずれも短めの期間や長めの期間が含まれていることがわかる。また、12 月 31 日と 1 月 1 日を連続化してシーズン検出を行った結果として、12 月末と 1 月初旬を含むシーズンが一つのシーズンとして検出されていることもわかる。

また、図 2 に、検出されたシーズンごとに抽出された主要撮影地域（凸包領域）のシーズン固有度分布を示す。横軸はシーズン固有度、縦軸はシーズン固有度が $E_{k,j}$ より大きい主要撮影領域の存在確率である。図 2 より、京都データセット（図 2(a)）におけるどのシーズンにおいてもシーズン固有度が最大となる主要撮影領域が 6 割程度存在することがわかる。この傾向は、東京データセット（図 2(b)）においても同様であった。

4.3 可視化結果

4.3.1 京都データセットの可視化結果

京都データセットの可視化例として、写真数が最も多かった Season3（3 月 17 日～4 月 20 日）を図 3 に示す。ここでは、撮影位置のシーズン固有度 $e_{k,n}$ が最大となる主要撮影スポットとして図 3 左の赤枠に示した地点に着目する。この地点は京都会館付近の川沿いを含む領域に含まれており、施設ではなく、川沿いに咲く桜の撮影スポットとして人気を得ている撮影スポットであることがわかる。つまり、この地点は、桜が咲くシーズンにのみ撮影が行われるという固有性を持つ主要撮影スポットと言える。一方、撮影位置のシーズン固有度 $e_{k,n}$ が最小となる図 3 左の黒枠に示した主要撮影領域内の撮影スポットに着目する。 $e_{k,n}$ が最小となる写真の例としては八坂神社の境内にある建物など、1 年を通じて変わらないものが撮影されている傾向があった。この地点は、言い換えれば年間を通して人気のある主要撮影スポットであると見なせる。

次に、京都データセットのシーズンのうち、2 番目に撮影写真数が多かった Season10（11 月 7 日～12 月 11 日）の可視化結果を図 4 に示す。図 4 左の地図上、二つの赤枠で示したシーズン固有度が最大で比較的大きい主要撮影領域に着目する。一つは吉田神社の一部、もう一つは京都御所境内の一部の領域であった。いずれも紅葉の写真が多いことから、これらの領域は紅葉のときにのみ撮影が行われることがわかる。つまり、施設が特定されただけではわからない、あるシーズンのみに固有の領域を捉えていることがわかる。

4.3.2 東京データセットの可視化結果

次に、都会である東京についても固有度の高い地点や領域としてどのようなところがあるかを調べるために、東京データセットの可視化例を図 5、図 6 に示す。図 5 は、東京データセットの可視化例として、写真数が最も多かった Season4（3 月 13 日～4 月 9 日）の可視化結果である。図 5 左の赤枠の地域に着目すると、撮影位置のシーズン固有度 $e_{k,n}$ が最大となる地点は、施設ではなく、千鳥ヶ淵戦没者墓苑付近の川周辺の領域であり、川沿いの桜と川を撮影した写真が多い主要撮影スポットであった。つまり、この地点は、桜のシーズンにのみ撮影が行われるという固有性を持つ地点であることがわかる。一方、撮影位置のシーズン固有度 $e_{k,n}$ が最小となる図 5 左の黒枠に示した主要撮影領域内の撮影スポットに着目する。 $e_{k,n}$ が最小となる写真の例としては東京駅の建物を撮影する傾向があった。つまり、この地点は、言い換えれば年間を通して人気のある主要撮影スポットであると見なせる。

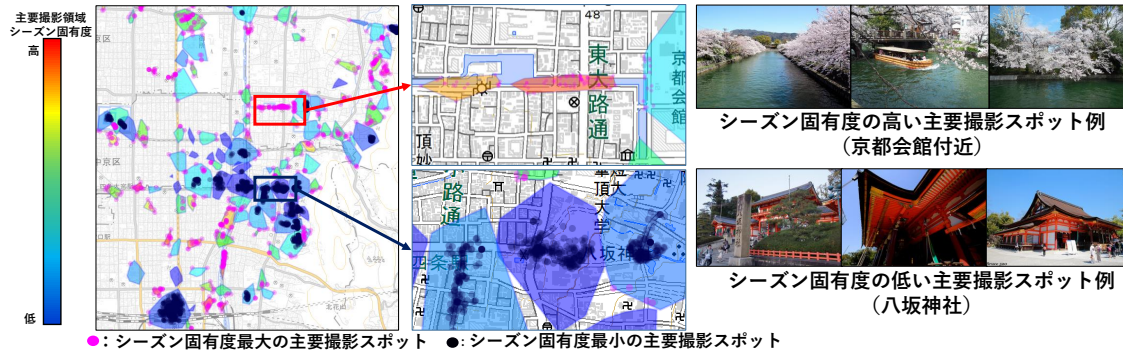


図 3: 京都データ (Season3[3月17日～4月20日]) の可視化結果



図 4: 京都データ (Season10[11月7日～12月11日]) の可視化結果

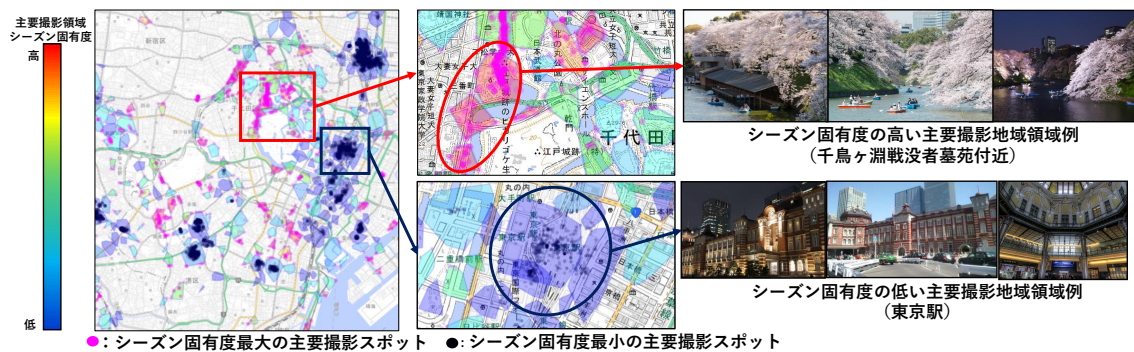


図 5: 東京データ (Season4[3月13日～4月9日]) の可視化結果

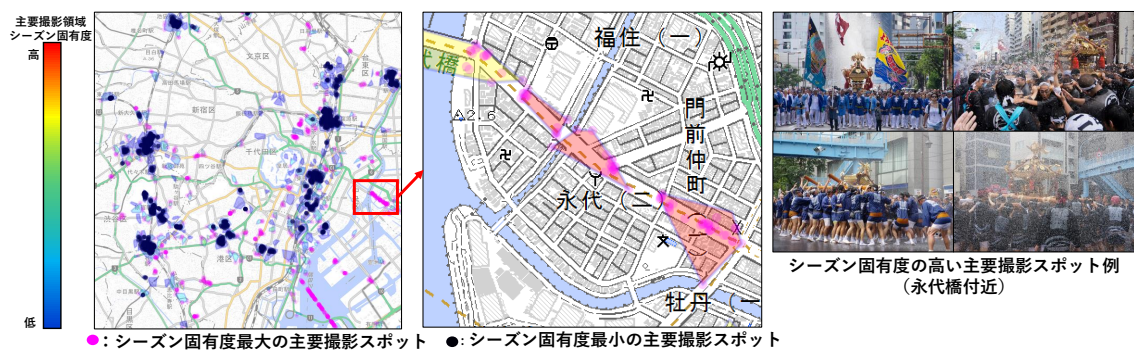


図 6: 東京データ (Season10[8月4日～8月30日]) の可視化結果

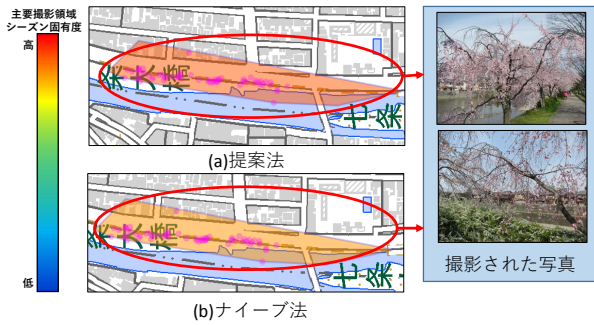


図 7: 各計算方法による主要撮影領域可視化の比較

また、図 6 は東京データにおける Season10 (8 月 4 日～8 月 30 日) の可視化結果である。図 6 より、道路沿いにシーズン固有度が最大の主要撮影領域が複数並んでいるのが見てとれる。これらの主要撮影領域内に含まれる撮影位置のシーズン固有度 $e_{k,n}$ が最大の主要撮影スポットに関する写真を調べると、深川八幡祭りを撮影する傾向があった。つまり、これらの領域は、祭りのときだけに撮影が行われるという固有性を持つ領域であることがわかる。

以上から、京都データセットおよび東京データセットのいずれにおいても、観光地域ごとシーズン固有度が高い主要撮影領域や、シーズン固有度が高い主要撮影スポットが、ある特定のシーズンにのみ見ることが出来る観光スポットと対応する傾向があることが分かった。

4.3.3 ナイーブ法と提案法の比較実験

本研究で提案したシーズン固有度に対し、凸包領域が重なる度合いをモンテカルロ法を適用して面積比で求めることでシーズン固有度を計算するナイーブな方法が考えられる。ここでは、ナイーブ法に対する提案法の有効性を検証するために、比較実験を行った。

図 7 に、提案法とナイーブ法で可視化結果に違いが現れた固有度の高い領域を示す。図 7 は京都データセットにおける Season3 において、五条大橋付近で撮影された鴨川沿いの写真が撮影されていた主要撮影領域である。ここで、図 7(a) の提案法による可視化と図 7(b) のナイーブ法による可視化の結果を見比べると図 7(a) の方が赤く、シーズン固有度が高い色を示しているが、図 7(b) のナイーブ法による領域は橙色に近づいていることがわかる。この領域内では、桜が咲いている様子が撮影されており、この時期のみに撮影が行われるという固有性をもった主要撮影領域であるため、顕著に赤い領域となることが望ましいと考えられるが、ナイーブ法で算出した場合、この領域と重なった異なるシーズンの領域の影響が強く表れたため、ナイーブ法によるこの領域の固有度が下がり、より橙色に近づいたものと思わ

れる。

以上の結果より、ナイーブ法によってシーズン固有度を決めるよりも提案法によりシーズン固有度を決める方がよりその主要撮影領域に適したシーズン固有度を与えることができることが示唆された。

5 まとめ

本研究では、指定された都市に対して、大量の Geo-tag 付き写真データを用いてシーズンを検出し、各主要撮影領域と各主要撮影スポットについてシーズン固有性を求め、観光ダイジェストマップとして可視化する手法を提案し、固有性の高い主要撮影領域や主要撮影スポットが、施設とはかぎらない様々な観光スポットと対応することを示した。また、ナイーブ法との比較により、提案法の有効性を示した。今後は、観光ダイジェストマップをより洗練させるため、主要撮影領域や主要撮影スポットのアノテーションをより強化する手法を探求する予定である。

謝辞

本研究は JSPS 科研費 26330352 の助成を受けたものである。

参考文献

- [1] 前田高志ニコラス, 吉田光男, 鳥海不二夫, 大橋弘忠. Twitter 位置情報・テキスト情報を用いた人の移動モデル構築と観光地推薦手法の提案. Technical report, 人工知能学会データ指向構成マイニングとシミュレーション研究会 (SIG-DOCMAS-9), 11 2015.
- [2] David J. Crandall, Lars Backstrom, Daniel Huttenlocher, and Jon Kleinberg. Mapping the world's photos. In *Proceedings of the 18th International Conference on World Wide Web*, pp. 761–770, 2009.
- [3] 熊野雅仁, 小関基徳, 小野景子, 木村昌弘. 地理および時間情報をもつ写真データに基づいたホット撮影スポットの抽出. 情報処理学会論文誌, Vol. 5, No. 3, pp. 41–53, September 2012.
- [4] 熊野雅仁, 岩渕聡, 小関基徳, 小野景子, 木村昌弘. 集合知に基づいたポピュラー撮影スポットに関する旬シーズンの可視化. 芸術科学会論文誌, Vol. 13, No. 4, pp. 218–228, 2014.