

シーズン固有性に基づく撮影スポット群のダイジェスト可視化

熊野雅仁¹⁾(正会員) 橋本和哉²⁾(非会員) 木村昌弘¹⁾(非会員)

1) 龍谷大学 理工学部 2) ニッセイ情報テクノロジー (株)

Digest Visualization of Photo Spots in Terms of Seasonal Uniqueness

Masahito Kumano¹⁾ Kazuya Hashimoto²⁾ Masahiro Kimura¹⁾

1) Faculty of Science and Technology, Ryukoku University

2) Nissay Information Technology Co., Ltd.

kumano @ rins.ryukoku.ac.jp

概要

近年、位置情報を付加できる携帯端末が増加してきたことに伴い、膨大な量の写真データが Web 空間に共有されている。これらの写真データには撮影位置・撮影時間・撮影者の情報が含まれているものもあり、データマイニング技術によって観光支援に応用することが注目されている。本研究では、ユーザによって訪れる訪問都市が選択されたとき、大量の位置情報及び撮影時間情報が付随した写真データを用いて、訪問都市に特有のシーズン群を検出し、それぞれのシーズンにおける主要撮影スポットを抽出することを考える。本論文では、これらの手法を適用した訪問都市の観光ダイジェストマップ構築法を提案する。そして、写真共有サイト Flickr の実データを用いた実験により、提案法の有効性を検証する。

キーワード ビッグデータ、シーズン固有性、ダイジェスト可視化、視覚的データマイニング、時空間データマイニング

Abstract

Recently, an enormous number of geo-tagged photos with time-stamps have been shared in the web space. Attention has been devoted to their analysis using data mining techniques and its application to tourism business. When a visit city is given by a user, from a large amount of data for those geo-tagged photos with time-stamps, we consider detecting the set of seasons associated with the visit city and extracting its major photo-spots proper to each of the seasons detected. In this paper, we propose such a method and apply it to the visualization of sightseeing digest maps for the visit city. Using real data from Flickr, we verify the effectiveness of the proposed method.

Keyword Big data, Seasonal uniqueness, Digest visualization, Visual data mining, Spatio-temporal data mining

1 はじめに

写真共有サイトや Geo-tag を付与できるカメラの登場以来、写真を通じて人々の関心の対象を分析する研究 [1] や、多くの人々が観光時に撮影した写真が含まれる Flickr などの写真共有サイトに蓄えられた写真データを用い、撮影地点群の空間的密集性に着目して人気のある撮影領域内のランドマークを抽出する研究 [2][3] が注目されてきた。近年、その撮影地点群の空間的密集性を旅行や観光に応用した空間ベースの研究が注目されている。日本のような時季に依存した観光先を数多く含む地域を訪れる場合、どのような観光先を訪れるかだけではなく、いつ頃訪れるかも大事な観点と考えられる。本研究では、人々の撮影行動を活用する空間ベースの研究をさらに進めて、ユーザが訪れたい都市を指定したとき、Flickr などの写真共有サイトの撮影位置や撮影時間を持つ大量の写真データを活用し、観光プランニングのためのシーズンを考慮した可視化システムの構築を目指す。

ところで、観光プランを検討するユーザ、特に外国人で何度も訪問先を訪れることが困難な人々にとって、限られた時間の中で可能なかぎり多くの観光先を巡ろうと考えることは自然である。ただし、日本では都市を観光するとき、時季によって見どころとなる観光先が異なり得る。実際、都市ごとに観光地の旬の時季やイベントの開催時季が異なるため、時季を無視すると観光時に時季外れとなり得る。また、時季依存性の高い同じ対象でも都市ごとに旬の時季が異なり得る。例えば桜を考えると、都市が違えば時季がずれる。これは、都市ごとに観光客数の日毎の変動パターンが異なり得ることを示唆している。つまり、都市ごとに独自のシーズン構造が存在し得ると思える。日本の都市への観光プランニングを考えると、そのようなシーズン構造を考慮することが重要である。また、Flickr のような写真共有サイトに投稿された大規模な写真データを利用すると、その都市への観光客数の変動パターンは、そこでの撮影者数の変動パターンで近似できると考えられる。つまり、各都市のシーズン構造は、その都市への日毎の撮影者数の変動に関する時系列において、撮影日の密集性における特徴的なパターンとして現れると考えられる。本研究では、都市全体での日々の撮影数の変動を数年にわたって調べることで、1 年間の（日単位）時間領域での確率密度

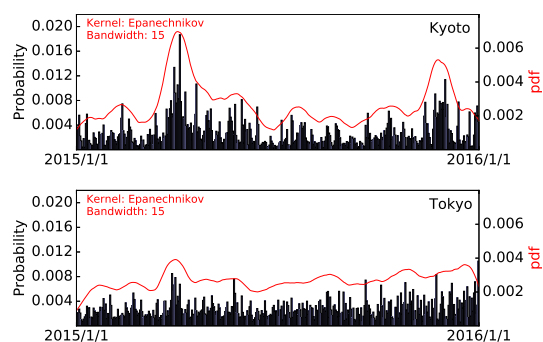


図 1: 各都市の撮影数に関する時系列パターン

関数を推定し（図 1 参照）、その極大値の吸引領域をシーズンと呼ぶ。図 1 は、京都、東京それぞれの都市に関する撮影数の変動を棒グラフで示した 1 年間の時系列パターンであり、赤い曲線は推定されたその確率密度関数 (pdf) を表している。各都市の確率密度関数を見ると、撮影日の密集性に沿って特徴的な大小の山がいくつも存在し、山の数や幅、つまりシーズン数や期間長が異なり得ることや、各都市で異なるパターンを示すことがわかる。つまり、どの都市も等しく四季で強制的にシーズンを分けるよりも都市独自のシーズンを捉えることが望ましいと考える。また、年の変わり目で撮影数が増えている。つまり、年末で強制的にシーズンを分けずにシーズンを捉えることが望ましい。もし、都市ごとに 1 年を特徴的なシーズンにうまく分解し、シーズンごとに撮影領域を抽出できれば、訪問時期に合わせて時季外れではない複数の撮影領域がどこにあるかを容易に把握し得る。また、各撮影領域のシーズン依存性の程度がわかれば、注目したシーズンの撮影領域がそのシーズンに限って人気を得るのか、それ以外のシーズンでも見どころを含み得るか、もしくは年中人気を得るのかを容易に把握し得る。さらに、各撮影領域内の撮影地点群は、それぞれ異なるシーズン依存性を持ち得ることから、各撮影地点のシーズン依存性の程度がわかれば、撮影領域内の撮影対象をシーズンをふまえて効率良く調べ得る。したがって本研究では、ユーザが訪問したい都市について、

- (1) 訪問都市のシーズンはいくつあり、どの時期か、
- (2) 各シーズンの主要な撮影領域はどこか、
- (3) その主要撮影領域はどの程度のシーズン依存性があるか、
- (4) その主要撮影領域に含まれるそれぞれの撮影地点はどこにあり、どの程度のシーズン依存性があるか、

などを効果的に可視化するシーズンを考慮した観光プランニング支援システムの構築を目指す。

本論文では、このようなシステムを構築するために次のようなアプローチを提案する。まず、(1) に対しては、ある年のみに撮影が行われる対象よりも数年を通じて安定して存在する撮影領域に着目し、さらに年末と年始の連続性と撮影時系列パターンの年単位の周期性を考える。そこで、ミーンシフトを周期性の観点から拡張を行うことでシーズン検出を行う手法を提案する。(2) に対しては、検出されたシーズンごとに主要な撮影領域（空間的に密集した撮影地点群）を地図上で視認可能にするため、Crandall らの撮影領域抽出法 [2] と点群から凸包を構築する手法を応用した主要撮影領域構築法を提案し、シーズンごとに主要撮影領域を調べることが可能にする。(3) に対しては、シーズン固有性という新たな概念を定義し、主要撮影領域があるシーズン特有であるか、複数のシーズンで人気を得るか、全シーズンで人気を得るかなど、主要撮影領域のシーズン依存性を把握可能とするため、領域のシーズン固有度を色で表現して視認可能にする手法を提案する。そして (4) に対し、各主要撮影領域に含まれる撮影地点群にもシーズン固有性という新たな概念を定義し、撮影地点に関するシーズン依存性を色で判別可能にする手法を提案する。さらに、その都市における平均的な 1 年の撮影状況を上記の手法に基づいてダイジェスト可視化することにより、シーズンごとに主要撮影領域や撮影地点の特徴が容易に把握できるような観光プランニング支援システムを提案する。そして、写真共有サイト Flickr から収集した実データを用いて提案法の有効性を示す。

まず 2 章で関連研究について述べ、3 章で提案法を述べ、4 章で人工データを用いたシーズン検出実験と評価を行い、5 章で大量の写真データを用いた可視化実験結果とシステムの評価および議論を行い、6 章でまとめる。

2 関連研究

写真の撮影行動を観光に応用する空間ベースの研究としては、旅行行動の分析 [4][5]、魅力的な地域の探索支援 [6][7]、略地図の作成 [8][9]、推薦 [10][11][12][13]、個人化推薦 [14]、旅行プランニングの支援 [15][16] などがある。本研究では空間ベースの着眼点だけではなく、シーズン固有性の観点からシーズンに依存する撮影領域の特

徴を把握可能にすることを目指す。

時系列データの時季特異性に着目した可視化の研究としては、電力消費などの時系列データ上で時季の変動を考慮しつつピークパターンの分析を行う研究 [17] がある。一方、施設やイベントの開催地などの個別の撮影領域ごとの時季特異性に着目する研究としては最盛期を持つ Prime spots を適切な時季に推薦する研究 [18] や、撮影日時の変動の時季異常性を捉える研究 [19]、周期的・非周期的なイベントを検出する研究 [20]、非日常的なイベント検出を行う研究 [21]、より短い期間での撮影行動の時季異常性を可視化する研究 [22] がある。いずれも個々の撮影領域ごとに時季を捉えることから、同じシーズンの複数の撮影領域どうしをシーズン依存性の観点から比較することが必ずしも容易ではないため、訪問時期ごとに観光プランニングを支援する問題には向いていない。本研究では、抽出されたシーズンごとに同じシーズンの撮影領域どうしの比較を容易にすることを目指す。

ソーシャルメディアの Geo-tag が付随する大量のユーザ行動データから観光先を推薦する研究として固有性という概念を導入した研究 [23] が行われている。ただし、どの地域にも存在するような地域固有ではない商業施設に対して地域固有の施設を抽出し、観光先として推薦を行う研究である。以上のような研究とは異なり、提案システムでは、シーズン固有性という新たな観点から、抽出されたシーズンごとに複数の主要撮影領域の特徴を比較しながら調べ得る環境をユーザに提供する。

3 提案法

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。そこで、まず提案法の概要について述べ、シーズン検出法、主要撮影領域の抽出法、撮影領域および撮影地点それぞれのシーズン固有度の定義、そして可視化システムについて述べる。

3.1 概要

指定された都市で最近 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

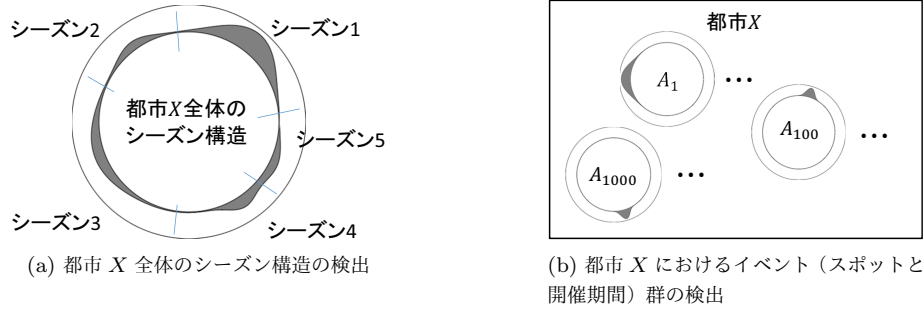


図 2: 地理および時間情報をもつ写真群に基づいた指定都市 X におけるシーズン検出とイベント検出の例. (a) は提案法による都市 X のシーズン構造の検出例であり, (b) は時空間バースト検出を土台とした従来法による都市 X 内のイベント群の検出例である. ここに, 1 年を円周で表し, 各日の写真数を半径方向に棒グラフで表している. (a) では都市 X 全体での写真数を示し, (b) では都市 X 内の各スポット A_j での写真数を示している. スポット A_j の場所はイベントの地理上の位置に対応し, その円上の棒グラフにおいて写真数がバーストしている部分はイベントの開催期間に対応している. 数多くのイベント群が検出されうるため, それらから都市 X のシーズン構造を認識するのは一般に容易ではない.

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 を単位矩形領域 (小領域) の集まりとして離散化し, その各单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地点 $x_{k,n}$ 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と撮影地点群を地図上に可視化し, ユーザが焦点を当てる区域を指定することにより 1 年間の観光ダイジェストマップを構築する. また, 各主要撮影領域や各撮影地点に対

して $\mathcal{D}$ に含まれている写真データを閲覧可能にする.

3.2 シーズン検出

本研究の目的は, 観光したい都市全体をグローバルに見てトップダウンにそのシーズンを検出し, ユーザがいつ頃その都市を訪問するか立案を支援することである. ところで, 時空間におけるバースト検出を土台とした, 「地理および時間情報をもつ写真群に基づいたスポットと期間の検出」に関する従来研究 [18], [19], [20], [21], [22] では, 都市内のスポットで行われるイベントとその開催期間の検出が主目的であるため, それらを単純に本研究のシーズン構造検出の問題に適用するには限界がある. 図 2 に, 本研究と従来研究の差異を例示する. 以下に, 提案するシーズン群の検出法について詳述する.

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, 円周上に写像した $\mathcal{T}$ 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, バンド幅を固定した Epanechnikov

カーネルによるミーンシフト法 [24] を用いる*1。

3.3 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 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、 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 L_0 以上の数の点（すなわち、写真）をもつクラスター群 $\{C_{k,j}; j = 1, \dots, J_k\}$ 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 $C_{k,j}$ が地図上の空間で占める領域を視認し易くすることを考える。各 $C_{k,j}$ には、複数の撮影対象が含まれ得るが、撮影対象から近い撮影地点であってもその距離は撮影者の意向により様々である。本研究では、主要撮影領域内の撮影対象近傍の撮影地点を含みつつ可能な限り撮影されない領域を排除する上で各 $C_{k,j}$ ごとにそれに属するすべての点を含む凸包（Convex hull[25]）を主要撮影領域 $R_{k,j}$ として構築する。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。

3.4 主要撮影領域のシーズン固有度

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、どのくらいそのシーズンならではのものであるかを定量化することを考える。これにより、シーズン固有度が高い場合にそのシーズンに限って見どころを持つ領域であることや、低い場合に年中人気を得る領域であることを表す。また、春は桜、秋は紅葉、その他のシーズンもお祭りやイベントなど、異なる対象を異なるシーズンのほぼ同じ区域で撮影する場合が存在し得る。そこで、異なるシーズンの主要撮影領域どうしが部分領域を共有する程度をシーズン固有度の中間の程度として示すことにより、複数のシーズンで見どころを持つ区域の主要撮影領域の発見を促すようにする。つまり、主要撮影領域がどのようなシーズン依存性を示すかを容易に視

認できるようにする。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 (Degree of seasonal uniqueness) $E_{k,j}$ 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。

3.5 撮影地点のシーズン固有度

また、シーズン S_k における主要撮影領域 $R_{k,j}$ 内にある撮影地点 $x_{k,n} \in R_{k,j}$ 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 $x_{k,n}$ のシーズン固有度 $e_{k,n}$ を、 S_k 以外のシーズンのうち、 $x_{k,n}$ 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 $x_{k,n}$ を主要撮影領域に含むシーズンの総数である。 $x_{k,n}$ 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 $e_{k,n} = 1$ であり、 $x_{k,n}$ 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 $e_{k,n} = 0$ である。

写真の撮影位置 $x_{k,n}$ 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 [26] を用いて行う。

3.6 可視化システム

図 3 は、提案可視化システムの概要である。ここでは、主な三つの機能について説明する。

3.6.1 地図上への可視化対象の描画機能

指定された都市の撮影地点群（図 3(a)）に対し、検出されたシーズン S_k ごとに、 S_k に含まれる主要撮影領域 $\{R_{k,j}; j = 1, \dots, J_k\}$ と撮影地点 $x_{k,n} \in R_{k,j}$ を描画する。まず、主要撮影領域のみもしくは撮影地点のみを描画する機能がある。このとき、主要撮影領域はシーズン固有度 $E_{k,j}$ と図 3 右端のカラーバーとの対応に従って着色した凸包として描画する。また、撮影地点は $e_{k,n}$ と図 3 右端のカラーバーとの対応に従って着色した小円として描画する。それぞれの可視化対象の描画色を見るときで、あるシーズンの主要撮影領域や撮影地点が赤に近

*1 提案法のシーズン検出および主要撮影スポット抽出で用いるミーンシフト法は、各撮影地点を一つのクラスターに所属させるハードクラスタリングとなることに注意しておく

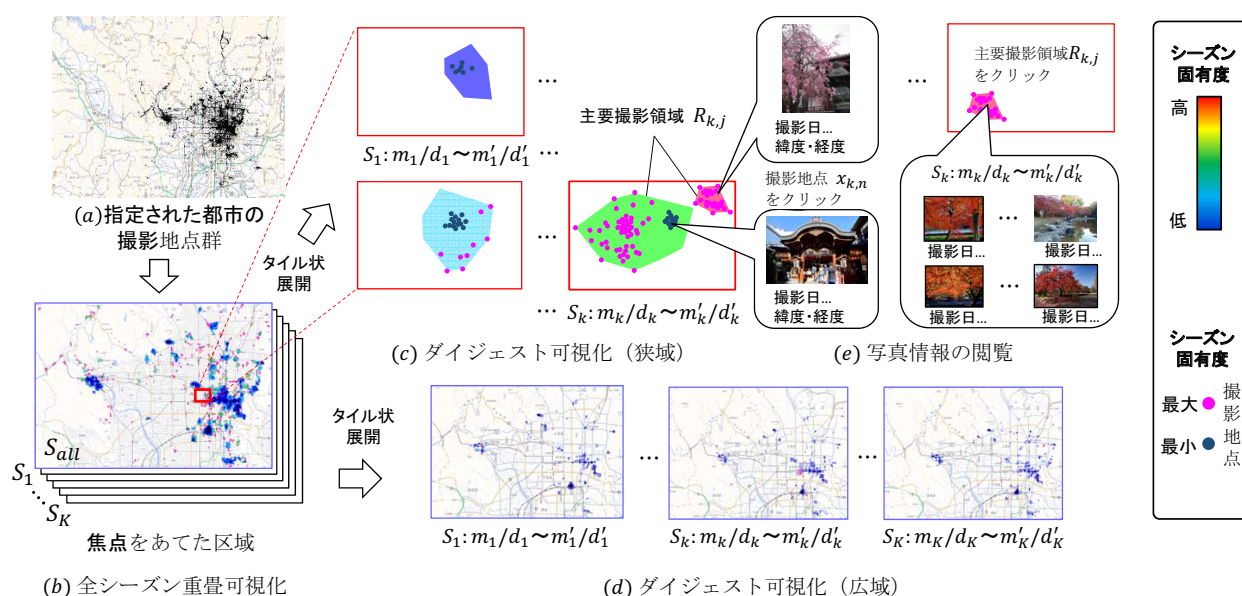


図 3: 提案可視化システム概要

ければ、そのシーズン特有の見どころをもつ可能性が高まり、青に近ければどのシーズンでも人気を得る年中人気を持つ可能性が高まる。つまり、カラーバーの描画色によってシーズン固有度をアノテーションする。

また、主要撮影領域と撮影地点の両方を重ねて可視化する機能がある^{*2}。その場合、撮影地点に対応する小円の円周を白くすることで撮影領域の色と撮影地点を区別する機能がある。さらに、撮影地点のシーズン固有度が最大と最小のものに限定して描画する機能がある。撮影地点のシーズン固有度が最大とは、 $e_{k,n}=1$ である各撮影位置 $x_{k,n}$ が、それを含む主要撮影領域以外、1 年間のすべてのシーズンで求められたどの主要撮影領域にも含まれなかったことを意味する。この機能では、シーズン固有度 $e_{k,n}$ が最大値 1 をとる非交差領域中の主要撮影地点 $x_{k,n}$ として、カラーバーに無い色として赤紫色の小円 (図 3 右端) で強調的に描画する。一方、対比的に、抽出されたすべてのシーズンの主要撮影領域を対象として、全シーズンで重なる交差領域に含まれ、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撮影地点 $x_{k,n}$ についても紺色の小円で描画する。なお、主要撮影領域と撮影地点を重ねて可視化する場合、主要撮影領域内に可視化された撮影地点の分布から、シーズン

固有度が同じ撮影地点群がそれぞれどのあたりに分布するかを容易に把握することができる。図 3(c) に示された緑色の主要撮影領域の例では、領域内にシーズン固有度 $e_{k,n}$ 最大と最小の撮影地点のみを描画した様子を示しており、それぞれ赤紫色と紺色を割り当てる機能を適用した例ともなっている。

3.6.2 ダイジェスト可視化機能

本研究では、指定された都市に対し、図 3(b) の全シーズンを重畳可視化した地図上でユーザが焦点を当てた区域について全シーズンの同じ区域をタイル状に展開して並べ、数年の平均的な 1 年の傾向を一望して視認可能にする地図提示法をダイジェスト可視化法と呼ぶ。図 3(c) は、図 3(b) の赤い枠の区域を縦横二次元にタイル展開した例であり、図 3(d) は図 3(b) の広域地図全体を横一列にタイル状展開した例^{*3}である。なお、各シーズンの地図ごとにシーズンの期間 (m_k 月 d_k 日 $\sim$ m'_k 月 d'_k 日) を描画する。この手法により、タイル群を一望することで、訪問都市のシーズンはいくつあり、どの時期かを容易に視認することができる。また、各タイルが一つのシーズンに対応しているため、シーズンごとに主要撮影領域がどこにあるかを容易に視認することができるため、ユー

^{*2} 全シーズンの主要撮影領域を重ねて可視化する場合、シーズン固有度の高いものを最後に描くよう実装した

^{*3} 広域に焦点を当てる場合、解像度が懸念されるが、近年、高解像度で大型の 4K や 8K ディスプレイも登場しつつあり、VR や、AR を有効に用いた提示法も考えられるため、全シーズンを並べて可視化することは大きな問題ではないことに注意しておく

ザが訪問を検討するシーズンに集中して撮影対象を調べることができる。また、シーズンどうしを比較することで、各シーズンの特徴を把握することができる。

3.6.3 写真情報の閲覧機能

図 3(e) に示すように、ユーザは、可視化された撮影地点 $x_{k,n}$ を示す小円をシステム上でクリックすることにより、撮影地点に対応する写真画像や撮影日時などの写真情報を閲覧することができる。また、主要撮影領域上をクリックすることにより、所属するシーズンの情報や、領域内の写真をまとめて閲覧することもできる。

以上の三つの機能により、ユーザは年始から年末までの全シーズンのダイジェストを通じて地図上で各シーズンや各区域を比較することができ、注目した主要撮影領域や、撮影地点に集中して効率良く撮影スポットの特徴を調べることができる。

3.6.4 可視化システムの実装とユーザインタフェース

本研究では、Web ブラウザ上で動作する地図システム Leaflet を基盤とし、Geojson 形式で表現した描画対象を地図上に可視化する試作システムを Javascript で実装した。なお、全シーズン重畳可視化において矩形区域を選択する機能の実装は Leaflet のプラグイン leaflet-areaselect を活用した。ユーザは地図システム上の都市をマウスでクリックすることにより全シーズン重畳可視化ページへ移り、注目する区域を定めてクリックすることでダイジェスト可視化ページへ移ることができる。また、ダイジェスト可視化においてマウスを用いて地図上を平行移動もしくは拡大・縮小した際、すべてのシーズンの地図が連動して同じ区域を表示するよう実装した。なお、撮影地点に対応する写真は地図上の円をクリックすることでポップアップウィンドウ上に表示し、撮影領域内の写真一覧は、凸包領域をクリックすることで Web ブラウザ上の新規ページに表示するよう実装した。

4 人工データによるシーズン検出法の評価実験

イベント検出のような従来研究の場合と異なり、本論文で対象とする「指定された都市におけるシーズン構造検出の問題」に対しては、明確な正解を定めることが一般に困難である。また、図 2 に例示したように、イベント検出の従来法を単純に本問題に適用するには限界がありうる。したがって本節では、提案するシーズン検出法

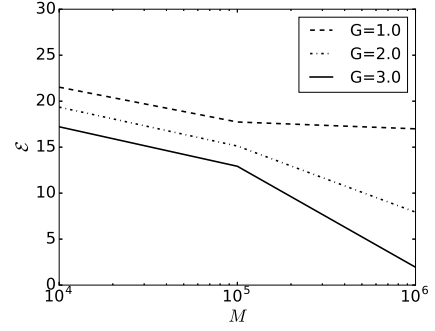


図 4: 人工データによるシーズン検出法の評価結果

の性能を人工データを用いて定量的に評価する。

4.1 人工データ

$\{S_k = [s_{k,0}, s_{k,1}]; k = 1, \dots, K\}$ を対象とする都市のシーズン構造とする。各シーズン S_k においては、 S_k に特有な写真が M_1 枚あり、それらの撮影日は、平均が S_k の中心で標準偏差が S_k の長さの $1/4$ である正規分布にしたがって分布しているとモデル化する。ただし、 S_k 内のどの日も少なくとも 1 枚の写真が存在するようにする。ここに、 k ごとに S_k の長さが異なりうるので、一般に、シーズンごとに撮影される写真の密度は異なることに注意しておく。

実験では、まず K を指定し、次に各 $s_{k,j}$ をランダムに設定することにより、その都市のシーズン構造を与えた。また、シーズン構造とは無関係にこの都市で 1 年間を通して撮影される写真の総数を $M_0 = KM_1/G$ とし、それらの撮影日を一様乱数を用いてランダムに与えた。ただし、正数 G はパラメータである。この都市の 1 年間の総写真数は $M = KM_1 + M_0$ であり、そのうち、 $\tilde{M} = KM_1$ 枚がシーズン構造と関係のある写真数であることに注意しておく。

4.2 評価尺度

$\{\hat{S}_\ell = [\hat{s}_{\ell,0}, \hat{s}_{\ell,1}]; \ell = 1, \dots, \hat{K}\}$ を提案法によるシーズン構造の検出結果とする。このとき、その推定誤差を、

$$\mathcal{E} = \frac{1}{K} \sum_{k=1}^K \mathcal{E}_k$$

で評価する。ただし、各 k に対して、

$$\mathcal{E}_k = \frac{1}{2} \min_{1 \leq \ell \leq \hat{K}} (|s_{k,0} - \hat{s}_{\ell,0}| + |s_{k,1} - \hat{s}_{\ell,1}|)$$

である。

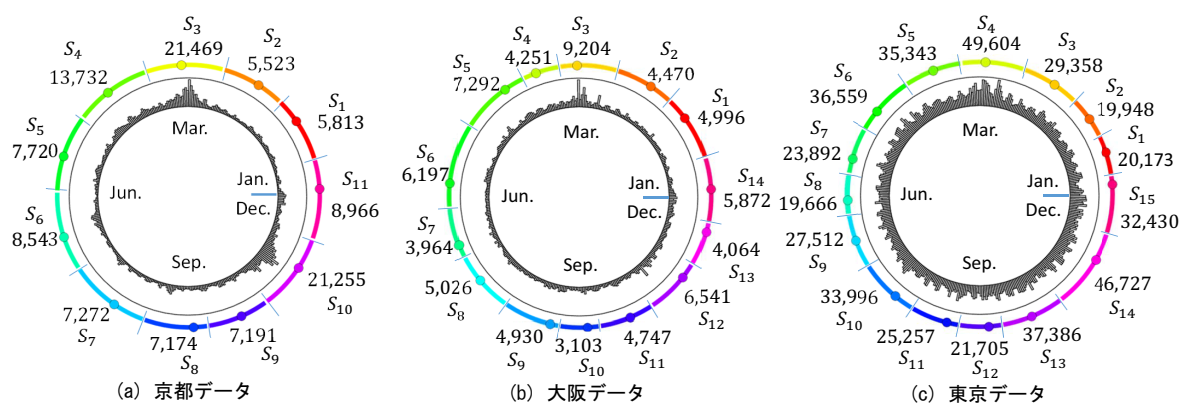


図 5: 検出されたシーズン期間とその写真数

4.3 人工データ実験の結果

$K=4$ で $G = 1, 2, 3$ の場合の結果を図 4 に示す. ここに, 各 M_1 の値に対して, 人工データを生成し提案法の評価指標 $\mathcal{E}$ を計算するという試行を 1000 回行った. 図 4 には, それらの平均値が総写真数 M の関数として示されている. $\mathcal{E}$ は, M が大きいほど小さくなり, また, G が大きくなるほど小さいことが観察される. 特に, 写真総数が十分にあるとき, すなわちサンプルが十分にあるとき, さらに M に対する M_0 の値が小さいとき, すなわちノイズレベルが十分低い場合にはほぼ完全にシーズン構造が復元できていることが見てとれる. なお, 推定シーズン数 $\hat{K}$ は常に正しく K の値を同定できていたことに注意しておく. 以上より, シーズン構造と関係のある十分な写真数があれば提案法は精度良くシーズン構造を同定できることがわかる.

5 実データによる実験

5.1 実験設定

提案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. シーズン検出において, 最近 y 年間に關する一年間の平均的な振る舞いを調べるうえで, 2013 年から 2015 年の $y=3$ 年間を対象とした. また, ユーザから指定される都市として, 本研究では, 日本有数の観光地である京都市 (WOEID=15015372), 大阪府 (WOEID=2345881), 東京都 (WOEID=2345889) を対象とした. 以上の条件のもと, 写真共有サイト Flickr から Geo-tag 付き写真データを収集した. 各都市の写真データ数は京都市が 223,464 枚, 大阪府が 142,432 枚, 東京都は 892,443 枚であった. 離散化後, 各都市の写真

表 1: 検出されたシーズン数と抽出された主要撮影領域総数

データセット	シーズン数	主要撮影領域総数
京都データ	11	1,540
大阪データ	14	1,177
東京データ	15	6,746

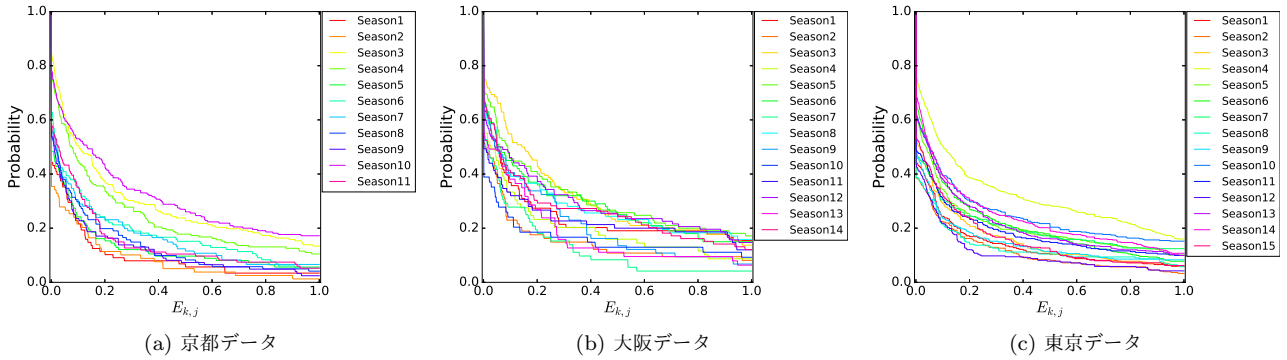
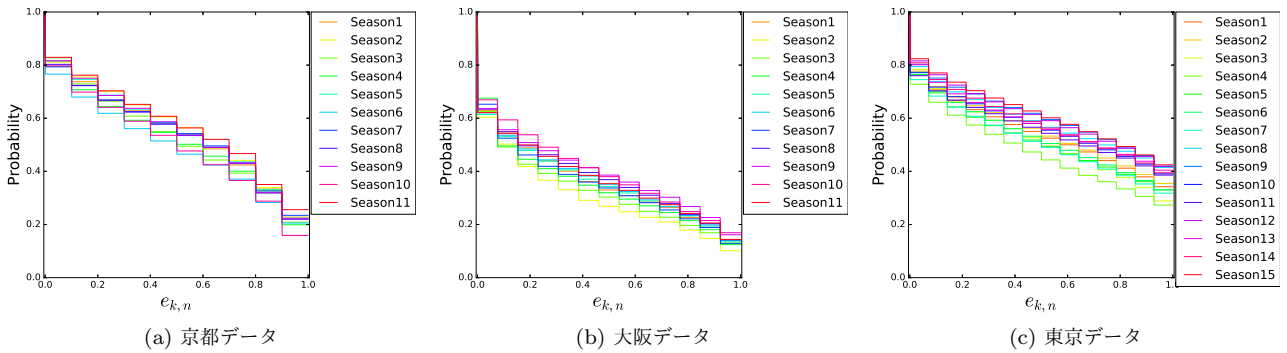
データ数は京都市が 114,658 枚, 大阪府が 74,657 枚, 東京都は 488,914 枚となった. ここで, 京都市に関するデータセットを京都データ, 大阪府に関するデータセットを大阪データ, 東京都に関するデータセットを東京データと呼ぶ. なお, シーズン検出におけるバンド幅は 15 とした. また, 撮影スポットの抽出におけるバンド幅は Crandall らに従い 100m とし, シーズン k ごとに抽出されたクラスタのうち, $N_0 > 10$ を満たす各シーズンのクラスタ集合を $\{C_{k,j}\}_{j=1}^{J_k}$ とした.

5.2 シーズン検出の結果

京都データ, 大阪データ, 東京データそれぞれについて, 検出されたシーズン数 K および全シーズンの主要撮影領域総数 $\sum_{k=1}^K J_k$ を表 1 に示す. 京都データは 11 シーズン, 大阪データは 14 シーズン, 東京データは 15 シーズンがそれぞれ検出されており, 地域に応じてシーズン数が変わり得ることがわかる.

図 5 に検出されたシーズン期間と写真数の対応を示す^{*4}. 各データセットの内円上の棒グラフは正規化された日毎の写真数時系列パターンである. また, 外円上の色で別れた区域が検出されたシーズン S_k の期間であり, ミーンシフトのセンターを小円として描画している. さ

^{*4} 図 5 および図 6 と図 7 は, 可視化システムの機能ではないこと, また, 図 5 の外周および図 6 と図 7 に用いた線の色は, 可視化システム上のカラーバーとは関係ないことに注意しておく


 図 6: 検出された各シーズン S_k における主要撮影領域 $R_{k,j}$ についてのシーズン固有度 $E_{k,j}$ の分布

 図 7: 検出された各シーズン S_k における撮影地点 $x_{k,n}$ についてのシーズン固有度 $e_{k,n}$ の分布

らに、各シーズン名 S_k の近傍にそのシーズンの写真数（離散化により撮影者数に相当）を示している。京都データでは S_3 (4/21-5/25)，大阪データでは S_3 (4/11-4/27)，東京データでは S_4 (4/10-5/5) が 1 年で最も写真数が多いシーズンであるが、円周上の期間の位置からいずれもほぼ同じ時季であることがわかる。また、写真数が最も少ないシーズンは、京都データでは S_2 (3/17-4/20)，大阪データでは S_{10} (10/10-11/3)，東京データでは S_8 (7/11-8/3) と円周上異なるシーズンになっており、異なる都市の地域性の違いが存在する傾向が示唆された。

5.3 シーズン固有度の基礎統計

図 6 は、主要撮影地域を対象とした全シーズンの固有度分布、すなわち、 $f_k(a) = |\{j; E_{k,j} \geq a\}|$ とするとき、 a の関数 $f_k(a)$ をプロットしたものである。いずれのデータセットにおいてもロングテール分布の傾向を示し、 $E_{k,j}$ が最大となる領域の割合がシーズンにより異なる傾向を示すことがわかる。全シーズンで $E_{k,j}=1$ となる領域の割合が最も大きいシーズンは、京都データで S_{10} ，大阪データで S_5 ，東京データで S_4 であった。また、図 7 は、同様に撮影地点を対象とした全シーズンの

固有度分布を示したものである。京都データや大阪データに対し、東京データの分布が緩やかに下がり、また、シーズン固有度 $e_{k,n}$ が最大となる撮影地点が京都データや大阪データに対し、多い割合で残る傾向があることがわかる。

5.4 可視化結果

提案する可視化法の有効性を検証する上で、紙面の都合から、京都データに焦点を当てる。なお、可視化機能のうち、主要撮影領域と撮影地点を重ねて可視化する機能を用いた結果を示す。ただし、撮影地点については、シーズン固有度が最大および最小を示すものに限って可視化対象とする機能を用いる。さらに、撮影地点のシーズン固有度が最大のものには赤紫、最小のものには紺色で描画する機能を用いる。最初に京都データ全シーズンの重畳可視化の結果を示し、次に、焦点を当てた二つのエリアを例として提案法の有効性を検証する。

5.4.1 概要

図 8 は、地図上に全シーズンの可視化対象をすべて重ねた重畳可視化の結果である。図 8 を見渡すと、シーズン固有度 $E_{k,j}$ が低い青い主要撮影領域が多重に重なり

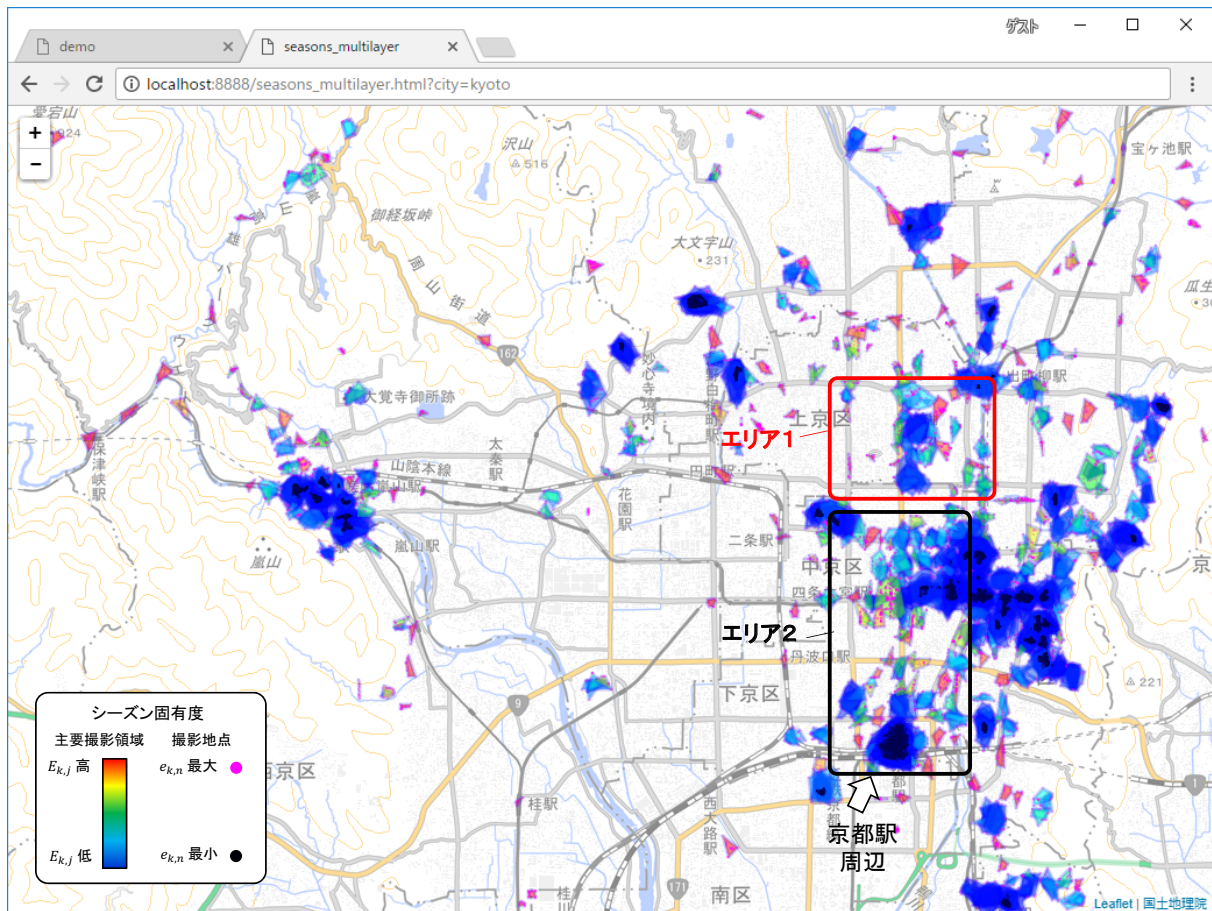


図 8: 全シーズン重畳可視化の結果

密集し、全シーズンの交差領域に紺色の小円群が視認できることで、年中人気のある地域を容易に把握し得ることがわかる。また、 $E_{k,j}$ が高い赤い主要撮影領域が市内各地で多数存在することも容易にわかる。ただし、重畳可視化では、特にシーズン固有度が高い様々な主要撮影領域がどのシーズンに含まれるかについては、個々にクリックして調べる必要がある。これはシーズン固有度が高い主要撮影領域のシーズンを調べる上で非効率であると考えられる。そこで、まず、図 8 のエリア 1 に焦点を当てたダイジェスト可視化を行い、その効果を検証する。また、単一および複数のシーズンで見どころを持つ主要撮影領域と撮影地点の事例を通じてシーズン固有度の評価を行う。次にエリア 2 で提案法の有効性を検証する。

5.4.2 エリア 1 のダイジェスト可視化結果

図 9 は、図 8 のエリア 1 で著名な観光スポットである京都御所周辺の区域の全シーズンをタイル状に展開したダイジェスト可視化の結果である。説明上、京都御所の

境内を図 9 の S_1 上に破線で示している。

まず、京都御所境内に着目すると、主要撮影領域が全く無いシーズン (S_1) や、多いシーズン (S_3 , S_4 や S_{10} など) を一目で視認できる。また、領域色が青に近いものに注目することで、多くのシーズンで人気のある領域がどこにあるかを容易に視認することができる。次に領域のシーズン固有度 $E_{k,j}$ が高く赤い領域のうち最も大きいシーズン S_{10} の主要撮影領域 $R_{10,1}$ に着目し、その領域内のシーズン固有度最大の撮影地点（赤紫の小円）をクリックして写真を閲覧すると、図 10(a) のように紅葉が撮影されていた。同様に S_4 の主要撮影領域 $R_{4,1}$ では図 10(b) のように池と庭園が撮影されていた。つまり、この S_{10} や S_4 でシーズン固有度が高い領域が、そのシーズンにかぎって人気を得ることがわかる。次に S_3 に着目すると、青に近い領域が多い中、 $E_{3,1}$ が低い水色の主要撮影領域 $R_{3,1}$ では、赤紫の多数の小円が中央寄りの一部に分布していることから、どのシーズンとも重

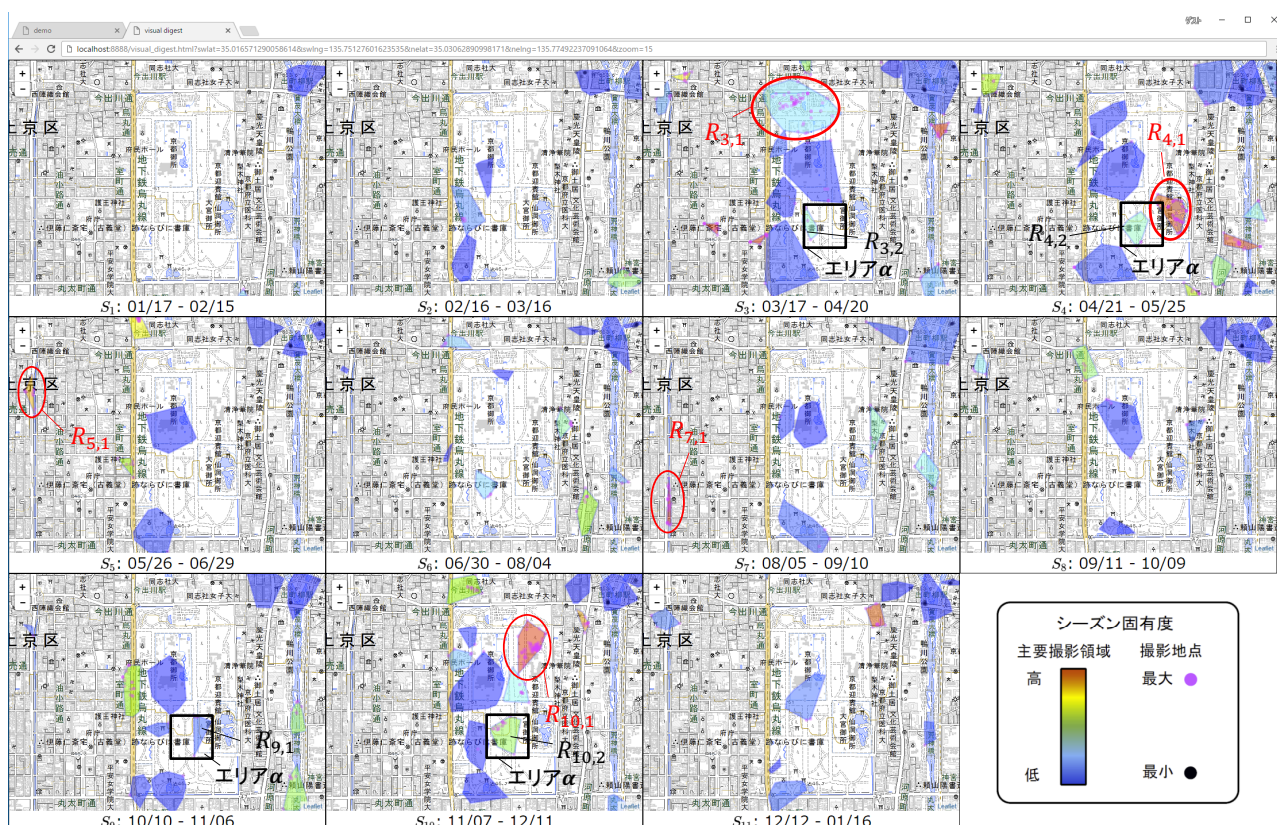


図 9: エリア 1 のダイジェスト可視化結果

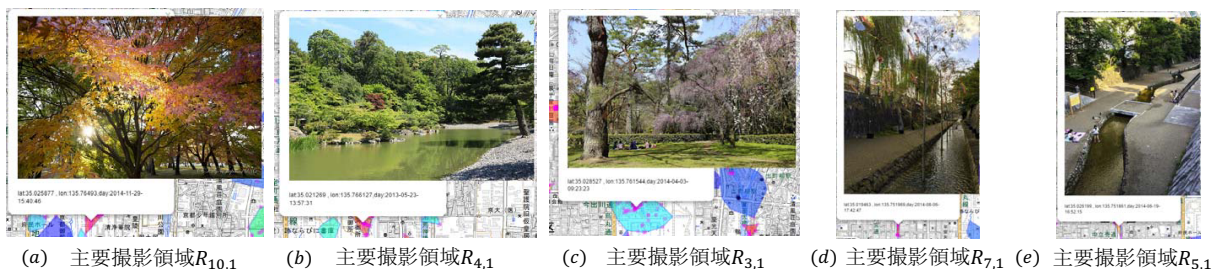


図 10: エリア 1 の主要撮影領域におけるシーズン固有度最大の撮影地点の写真

ならない非交差領域中の撮影地点が多い、つまりシーズン固有度が最大の撮影地点を多く含むことが容易に把握できる。図 10(c) に、主要撮影領域 $R_{3,1}$ 内で赤紫の小円をクリックして写真を閲覧した様子を示している。写真から、このシーズンに特有の桜が撮影されていた。さらに、 S_7 や S_5 に着目すると、御所境内にはシーズン固有性の高い領域が存在しないものの、境内の外に目を向けるとシーズン固有度が高い主要撮影領域がある。主要撮影領域 $R_{7,1}$ の場合、図 10(d) のように飾り付きの笹が水路沿いに並ぶ笹祭りの様子が撮影されていた。その

シーズンに御所内では季節ならではの撮影スポットが乏しいものの、近場で立ち寄り先として検討に値する撮影スポットになり得ると考えられる。また、主要撮影領域 $R_{5,1}$ の場合、図 10(e) のように水遊びの場がこのシーズンのみに撮影されていた。寒いシーズンにはむしろ避けられるが、暑くなり始めるシーズンに人気を得る地域と思われる。つまり、シーズンに依存するもしくは依存しない観光先内部やその周辺の主要撮影スポットについて、効率良く地域の特徴を調べ得ることがわかる。

次に、複数のシーズンで見どころを持つ地域の発見を

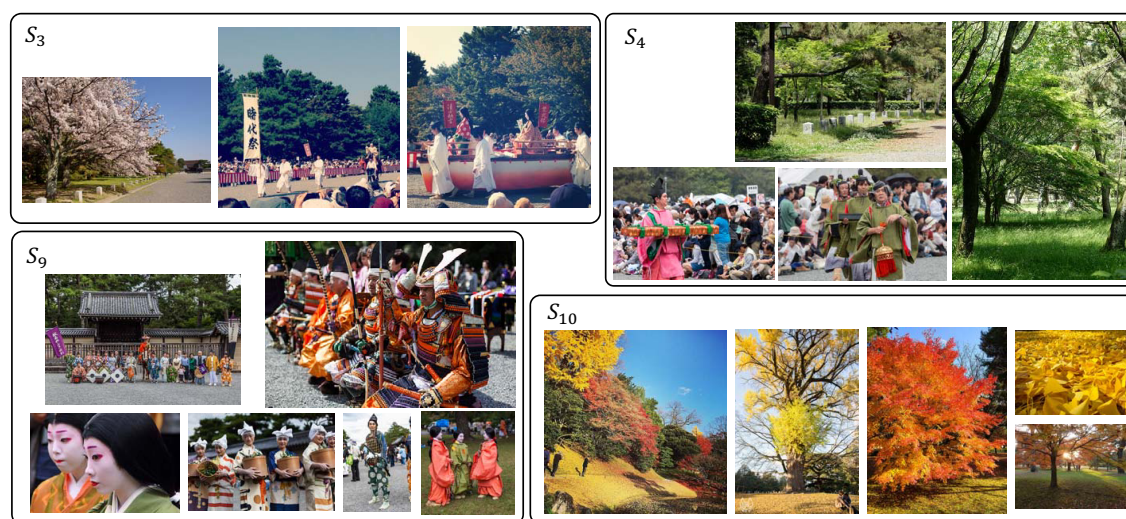


図 11: エリア α における異なるシーズンの主要撮影領域の写真

促し得る例を示す．ほぼ同じ区域でシーズンごとに主要撮影領域が現れたり消えたりする場合，各シーズン特有の撮影対象が存在する可能性がある．もし，そのような主要撮影領域どうしが交差領域を持つ場合，シーズン固有度は下がるものの交差領域が少なければシーズン固有度は下がりすぎず，提案法では薄い青から橙色までの色として発見を促し得る．そのような事例として四つのシーズン S_3 で水色， S_4 で水色， S_9 で薄い青， S_{10} で緑の主要撮影領域が現れる図 9 のエリア α がある．図 11 にそれぞれのシーズンの主要撮影領域に含まれる写真を抜粋すると， S_3 の場合は桜と時代祭り， S_4 の場合は別の時代祭りおよび森林内や緑地， S_9 の場合はさらに別の時代祭り， S_{10} の場合は紅葉がそれぞれ撮影されており，いずれもシーズン特有の撮影対象であった．つまり，ダイジェスト可視化を通じてほぼ同じ区域で複数のシーズンに現れる主要撮影領域の発見を促し，また，主要撮影領域のシーズン固有性を捉えることで，複数のシーズンに異なる見どころを持つ主要撮影領域の発見を促し得ることがわかる．

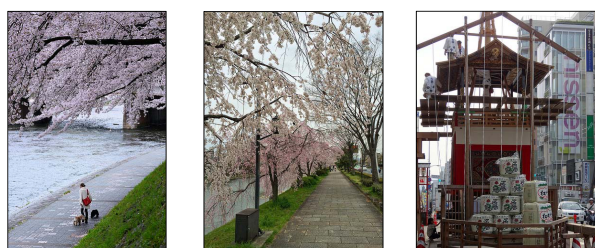
5.4.3 エリア 2 のダイジェスト可視化結果

図 8 のエリア 2 に焦点を当ててタイル状に展開したダイジェスト可視化が図 12 である．提案するダイジェスト可視化では，次のようなことを容易に把握できる．冒頭に示した (1) に対して可視化結果からいくつかのシーズンがあり，各シーズンの期間を容易に把握できる．(2) に対して，各シーズンを見ると，シーズンごとに主要撮影

スポットがどこに存在するかが容易にわかる．(3) に対してあるシーズンの主要撮影領域がどの程度他のシーズンの主要撮影領域と重なるかが容易にわかる．例えば，京都駅周辺は領域の色が青いことから描画色で多くのシーズンと重なることがわかり，全シーズンのほぼ同じ地域に領域があることで年中撮影される地域であることを明瞭に把握できる．一方， $E_{k,j}$ が高い赤い主要撮影領域があることで，他のシーズンでは撮影が行われない領域であることが容易にわかる．そのシーズン特有の対象を調べると，例えば， S_3 に焦点をあて京都駅に近いシーズン特有の訪問先として川沿いの互いに近い赤い楕円の主要撮影領域 $R_{3,1}$ と $R_{3,2}$ に着目し，それぞれの写真を閲覧したところ，図 13 の $R_{3,1}$ と $R_{3,2}$ のように，いずれも川沿いで桜が楽しめる撮影領域であった．(4) に対しては，京都駅周辺の領域内に $e_{k,j}$ が最小の紺色の小円が広く分布することから，年中撮影される撮影地点群の位置を明瞭に把握し得ることがわかる．また， S_6 の $R_{6,1}$ に着目すると， $e_{k,j}$ が最大の赤紫色の小円の分布により，このシーズン特有の見どころとして特にどこが人気があるのかも容易にわかる．その撮影地点をクリックすると，祇園祭の山鉦が数多く撮影されていたが，図 13 の $R_{6,1}$ の写真のように山鉦を組み立てる写真が多く含まれていた．なお，図 13 の $R_{6,1}$ は，他のシーズンと重なる領域があることで $E_{6,1}$ が低くなり緑色となっているが，その領域の中心部に赤紫色の撮影地点が数多く分布しているため，非交差領域はシーズン固有性が高い地



図 12: エリア 2 のダイジェスト可視化結果



(a) 主要撮影領域 $R_{3,1}$ (b) 主要撮影領域 $R_{3,2}$ (c) 主要撮影領域 $R_{6,1}$

図 13: エリア 2 においてシーズン固有度が高い主要撮影地点の写真

域であり、この S_6 シーズンのみに撮影が多くなる、祇園祭でのみ人気の高い撮影スポットであることがわかる。

以上のように、ユーザが焦点をあてる区域を自由に選り、ダイジェスト可視化を通じてシーズン特有の見どころや年中人気のある地域を効率的に調べ得る環境を提供

している点で、提案法の有効性が示唆された。

5.5 被験者実験による評価

提案するダイジェスト可視化は、指定された都市について、まず、都市全体の撮影動向からその大局的なシーズン群を検出し、次に、シーズンごとに主要撮影領域群を抽出し、各主要撮影領域をシーズン固有度に基づいて色付けて描画するという手法である。ところで、ジオタグつき写真群に基づく都市の観光ダイジェストマップ作成では、その都市におけるすべての撮影地点群データを用いて主要撮影領域群を抽出する Crandall らの手法 [2] が、一般に土台となっている。したがって、本論文ではこのような手法をベースラインとする。ただし、都市のシーズン検出と主要撮影領域のシーズン固有度の提示が提案可視化の主要な要素であるので、それら以外の要

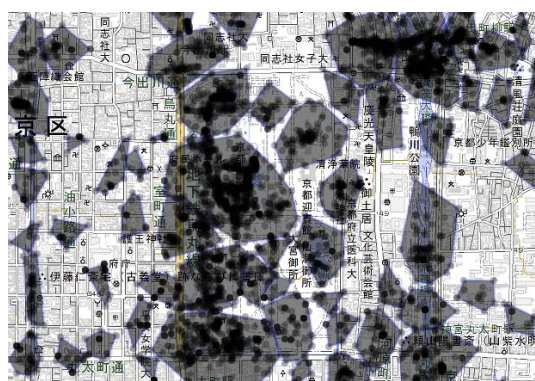


図 14: 被験者実験のベースライン法

素については基本的には提案可視化と同様にした。つまり、図 14 のように、ベースライン可視化法においても、主要撮影領域は凸包で描画し、また、撮影地点をクリックすればその写真が閲覧でき、主要撮影領域をクリックすればその中の写真一覧を閲覧できるようにした。

次に、被験者実験の評価法を述べる。本論文で取り組む問題は、1 章の第 2 段落の最後に示した 4 つのポイントを効果的に可視化する観光プランニング支援システムの構築である。したがって、被験者には提案法とベースライン法の両システムを使用してもらったうえで、それら 4 つのポイントに関して、有用性、ユーザ受容性の観点から、提案可視化法がベースライン法に比べてどのくらい改善しているかを被験者実験により評価した。

以下に、4 つのポイント (1) から (4) に対して被験者に質問した Q1 から Q4-3 までの質問内容を示す。

- (1) 訪問都市のシーズンはいくつあり、どの時期であるか
Q1. この情報が旅行計画の立案に役に立ちますか？
- (2) 各シーズンの主要な撮影領域はどこか
Q2. この情報が旅行計画の立案に役に立ちますか？
- (3) その主要撮影領域はどの程度のシーズン依存性があるか？
Q3-1. シーズンにかかわらず見れるものが見つけやすいですか？
Q3-2. シーズン特有のものがたくさん見つけやすいですか？
Q3-3. シーズン特有のものが見れるシーズンは把握しやすいですか？
Q3-4. シーズンに特有の主要撮影領域が見つかることで旅行計画の立案に役立ちますか？
- (4) その主要撮影領域が含まれるそれぞれの撮影地点はどこにあり、どの程度のシーズン依存性があるか
Q4-1. シーズンにかかわらず見れるものが見つけやすいですか？
Q4-2. シーズン特有のものがたくさん見つけやすいですか？
Q4-3. シーズンに特有の撮影地点が見つかることで旅行計画の立案に役に立ちますか？

被験者実験においては、20 代の男性 20 名の被験者が参加した。なお、被験者には、6 段階の Score (0 から 5) による評価を要求した。例えば、旅行計画の立案に役

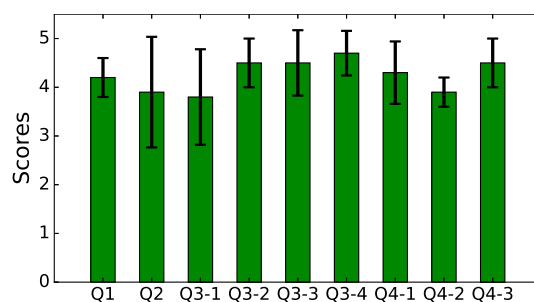


図 15: 被験者実験の結果

立つかというユーザ受容性の観点では 5 は非常に役立った、0 は全く役立たなかったと位置づけた。また、見つけやすいかもしくは把握しやすいかという有用性の観点では、5 は非常に見つけやすいもしくは非常に把握しやすいこと、0 は見つけにくい、把握しにくいと位置づけた。また、各質問ごとに意見やコメントを求めた。

図 15 は、各質問に関する平均 Score と標準偏差を Error bar として表示した被験者実験の結果である。Q1 については、評価が 4 程度であり、他のものと比べて相対的に小さいことから、評価が揃う傾向があることがわかる。Q2 や Q3-1 については、Error bar の長さが他のものと比べて長いことから、シーズンごとに主要撮影領域が確認できることや年中撮影が行われる撮影領域の見つけやすさについては評価が分かれる傾向が観察される。また、Q3-2, Q3-3 などシーズンに特有な撮影領域の提示についても、ユーザ受容性や有用性に関する質問において 4 程度の Score を獲得しており、評価の分散も小さい傾向があることもわかる。また、Q4-1 や Q4-2 の撮影地点の可視化については、シーズン特有のものよりシーズンに依存しないものが見つけやすいという評価結果となった。これは、年中撮影される撮影スポットの撮影地点数が相対的に多いことが原因と考えられるものの、評価値は 4 程度となっていることがわかる。そして、Q1, Q2, Q3-4, Q4-3 におけるユーザ受容性においては平均で 4.13, Q3-1, Q3-2, Q3-3, Q4-1, Q4-2 の有用性においては平均で 4.36, 全体平均で 4.26 の評価を得ていることから、ベースライン法に対して提案法が改善している評価を得ている点での有効性が示唆された。

5.6 議論

5.6.1 被験者実験による提案システムの効果

ここでは、被験者実験における各質問に対して被験者に求めた意見やコメントから、提案システムの効果につ

いて議論する。被験者の意見やコメントでは、提案法で期待した通り、「訪問時期の参考になる」、「旅行計画のおおまかな指針になる」、「どこに行こうか迷ったときに役立つ」、「シーズンを見てその時期の訪問を考えることができる」、「どの場所が人気を得ているかがわかる」、「シーズンごとに何が见れるかがわかる」、「訪問予定時期および訪問スポットの決定に役立つ」、「月の区切りではわからないシーズンの違いが見つけれられる」、「せっかくの旅行ならばシーズン特有のものを撮って自慢したいので見つけやすいのはありがたい」など、提案法の効果を示す意見を多く獲得することができた。ただし、被験者の中には、人が集まっている様子はわかるが、なぜ集まっているかはわからないという意見があった。撮影領域や撮影地点の効果的なアノテーション法の構築が望まれるが、その手法については今後の課題とする。また、提案システムがベースライン法から改善している点での評価を示しているものの、個人的な見解として、旅行時には、サービスやシステムに頼るより、現地で偶然興味を持ったものを気ままに見る旅行の仕方を好んでいるという意見もあった。また、他の被験者として、観光情報として公開されている有名なものを見るだけで満足するという意見もあった。ユーザの嗜好は多様であるため、誰にとっても受け入れられるシステムの構築は難しいものの、提案システムは、本研究の目的を達成していると考えられる。

5.6.2 シーズン固有度算出法の比較

主要撮影領域のシーズン固有度 $E_{k,j}$ を算出する方法には、提案法以外に、主要撮影領域の面積に対する領域どうしの交差領域面積の比をモンテカルロ法などを用いて求めるナイーブな方法が考えられる。つまり、非交差領域の面積が大きいほどシーズン固有度が高くなるという方法である。図 16 は、京都データにおける鴨川の五条大橋周辺の同区域に対して、図 16(a) が面積比法、図 16(b) が提案法を適用した可視化結果である。なお、図 16(a) と図 16(b) の領域 $R_{1,j}$, $R_{2,\ell}$, $R_{3,m}$ は、説明上、各主要撮影領域の境界を黒い実線で描画している。

図 16 の区域中央にある主要撮影領域 $R_{1,j}$ の領域色において、提案法による図 16(b) の橙色と比べ、面積比法による図 16(a) の $E_{1,j}$ が低く黄色になっている。これは、面積比法の場合、 $R_{1,j}$ に重なる別シーズンの $R_{2,\ell}$ や $R_{3,m}$ との交差領域が存在し、交差領域の面積が少し

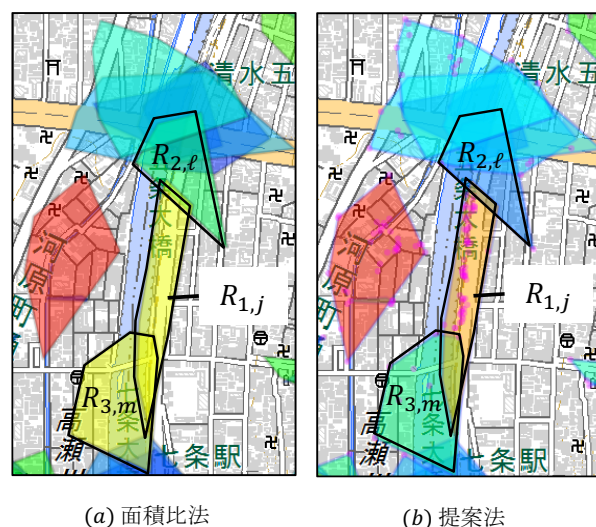


図 16: シーズン固有度 $E_{k,j}$ 算出法の比較

大きいことが影響してシーズン固有度が下がるためである。一方、面積比法による $R_{1,j}$ の色よりも提案法による $R_{1,j}$ の色がより赤に近く見えるのは、 $R_{2,\ell}$ や $R_{3,m}$ との交差領域よりも、相対的にどのシーズンの主要撮影領域とも重ならない非交差領域の写真数がかかなり多いからである。図 16(b) において、 $R_{1,j}$ の非交差領域に、 $e_{1,n}$ が最大の $x_{1,n}$ が多数存在することがわかる。つまり、 $R_{1,j}$ の撮影地点群は非交差領域が主要領域と見なせる。この場合、シーズン固有性の高さを強調する観点において、 $R_{1,j}$ は非交差領域の撮影行動を強く反映して $E_{1,j}$ が定まることが望ましいが、面積比法では $R_{k,1}$ のシーズン固有性を過小評価する結果となっているのに対し、提案法はその状況を良好に反映していると言える。

また、面積比法では $R_{3,m}$ のシーズン固有性を過大評価する結果ともなっている。 $R_{3,m}$ 内の撮影地点群の 4 割程度が $R_{1,j}$ との交差領域に存在していたため、 $R_{3,m}$ のシーズン固有性は黄色よりもさらに下がることが望ましい。その観点において、提案法の $R_{3,m}$ は緑色となっており、面積比法よりもその状況をより良好に反映していると言える。

以上より、面積比法に対して非交差領域中の撮影地点群の多さに注目する提案法の有効性が示唆される。

6 まとめ

本研究では、撮影位置・撮影時間情報が付随する一般の人々により撮影された大量の写真データを用い、観光

プランを検討するユーザから指定された都市に対して、ユーザが訪れる都市に関するシーズンに依存した特徴を効率的に調べ得る環境を提供する問題に取り組んだ。そして、数年間の平均的な傾向として、その都市における独自のシーズンを検出し、シーズンごとに主要撮影領域群を抽出し、主要撮影領域と撮影地点それぞれのシーズン固有性の観点から主要撮影領域と撮影地点を地図上に可視化し、全シーズンをタイル状に並べ1年のダイジェスト可視化を行う手法を提案し、さらに、人工データを用いたシーズン検出の特性評価、実データを用いた実験で可視化結果の評価、また、被験者実験およびナイーブ法との比較を行い、提案法の有効性を検証した。そして、提案法により、年中人気を得る地域や、シーズン特有の見どころを効率よく調べ得る事例を示し、その有用性も示した。

人々の実世界における行動がビッグデータとして蓄積され続ける時代、さらなる効果や拡張性を求め、より効率的に観光地の特徴を探索できるようダイジェスト可視化をより洗練させるための手法を探求する予定である。

参考文献

- [1] Lyndon Kennedy, Mor Naaman, Shane Ahern, Rahul Nair, and Tye Rattenbury. How flickr helps us make sense of the world: Context and content in community contributed media collections. In *Proceedings of the 15th ACM international conference on Multimedia (MM'07)*, pp. 631–640, 2007.
- [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] Pavel Serdyukov, Vanessa Murdock, and Roelof van Zwol. Placing flickr photos on a map. In *Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval*, pp. 484–491, 2009.
- [4] Yuki Arase, Xing Xie, Takahiro Hara, and Shojiro Nishio. Mining people's trips from large scale geo-tagged photos. In *Proceedings of the 18th International Conference on Multimedia*, pp. 133–142, 2010.
- [5] Zahra Farzanyar and Nick Cercone. Trip pattern mining using large scale geo-tagged photos. In *Proceedings of the International Conference on Computer and Information Science and Technology*, pp. 113:1–8, 2015.
- [6] Slava Kisilevich, Florian Mansmann, Peter Bak, Daniel A. Keim, and Alexander Tchaikin. Where Would You Go on Your Next Vacation? - A Framework for Visual Exploration of Attractive Places. In *GeoProcessing 2010*, pp. 21–26, February 2010.
- [7] Davide Eynard, Alessandro Inversini, and Leonardo Gentile. Finding similar destinations with flickr geotags. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing (SAC'12)*, pp. 733–736, 2012.
- [8] Wei-Chao Chen, Agathe Battestini, Natasha Gelfand, and Vidya Sethur. Visual summaries of popular landmarks from community photo collections. In *Proceedings of the 17th ACM international conference on multimedia*, pp. 789–792, 2009.
- [9] 王佳な, 野田雅文, 高橋友和, 出口大輔, 井手一郎, 村瀬洋. Web 上の大量の写真に対する画像分類による観光マップの作成. 情報処理学会論文誌, Vol. 52, No. 12, pp. 3588–3592, 2011.
- [10] Abdul Majid, Ling Chen, Gencai Chen, Hamid Turab Mirza, and Ibrar Hussain. Gothere: travel suggestions using geotagged photos. In *Proceedings of the 21st International Conference on World Wide Web (WWW'12)*, pp. 557–578, 2012.
- [11] Jinpeng Chen, Yu Liu, Zhenyu Wu, Ming Zou, and Deyi Li. Recommending interesting landmarks in photo sharing sites. In *Proceedings of Web Information Systems Engineering (WISE 2013)*, pp. 151–159, 2013.
- [12] Yue Shi, Pavel Serdyukov, Alan Hanjalic, and Martha Larson. Nontrivial landmark recommen-

- dation using geotagged photos. *ACM Transactions on Intelligent Systems and Technology*, Vol. 4, No. 3, pp. 47:1–27, 2013.
- [13] Imran Memon, Ling Chen, Abdul Majid, Mingqi Lv, Ibrar Hussain, and Gencai Chen. Travel recommendation using geo-tagged photos in social media for tourist. *Wireless Personal Communications*, Vol. 80, No. 4, pp. 1347–1362, 2015.
- [14] Thomas Phan, Jiayu Zhou, Shiyu Chang, Junling Hu, and Juhan Lee. Collaborative recommendation of photo-taking geolocations. In *Proceedings of the 3rd ACM Multimedia Workshop on Geotagging and Its Applications in Multimedia(GeoMM’14)*, pp. 11–16, 2014.
- [15] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *Proceedings of the 18th International Conference on Multimedia*, pp. 143–152, 2010.
- [16] Huagang Yin, Xin Lu, Changhu Wang, Nenghai Yu, and Lei Zhang. Photo2trip: an interactive trip planning system based on geo-tagged photos. In *Proceedings of the 18th International Conference on Multimedia*, pp. 1579–1582, 2010.
- [17] M. C. Hao, H. Janetzko, S. Mittelstadt, W. Hill, U. Dayal, D. A. Keim, M. Marwah, and R. K. Sharma. A visual analytics approach for peak-preserving prediction of large seasonal time series. *Computer Graphics Forum*, 2011.
- [18] Vishal Sharma, Kyumin Lee, and Jinwook Chung. Recommending prime spots of a destination and time to visit from geo-tagged social data. In *Proceedings of International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom’14)*. IEEE, 11 2014.
- [19] 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.
- [20] Ling Chen and Abhishek Roy. Event detection from flickr data through wavelet-based spatial analysis. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management, CIKM ’09*, pp. 523–532, 2009.
- [21] J. Chae, D. Thom, H. Bosch, Y. Jang, R. Maciejewski, D. S. Ebert, and T. Ertl. Spatiotemporal social media analytics for abnormal event detection and examination using seasonal-trend decomposition. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 143–152, Oct 2012.
- [22] 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.
- [23] 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 (SIG-DOCMAS-9), 11 2015.
- [24] Dorin Comaniciu and Peter Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 5, pp. 603–619, 2002.
- [25] Kai Hormann and Alexander Agathos. An algorithm for finding convex hulls of planar point sets. In *Proceedings of 2012 2nd International Conference on Computer Science and Network Technology*, pp. 888–891, 2012.
- [26] C. Bradford Barber, David P. Dobkin, and Hannu Huhdanpaa. The point in polygon problem for arbitrary polygons. *Computational Geometry: Theory and Applications*, Vol. 20, No. 3, pp. 47:1–27, 2001.

熊野 雅仁



1991 年立命館大学理工学部基礎工学科卒業。1991 年龍谷大学理工学部実験助手。2008 年龍谷大学理工学部実験講師。現在に至る。メディア論，知能情報学の研究と教育に従事。博士 (工学)[神戸大学大学院工学研究科情報知能学専攻]。芸術科学会，認知科学会，情報処理学会，ACM，IEEE-Computer Science 他会員。

橋本 和哉



2015 年龍谷大学理工学部電子情報学科卒業。2017 年龍谷大学大学院理工学研究科電子情報学専攻修士課程修了。2017 年ニッセイ情報テクノロジー（株）に入社。現在，個人保険システム事業部 WEB サービスブロックの業務に従事。現在に至る。

木村 昌弘



1987 年大阪大学理学部数学科卒業。1989 年同大学院理工学研究科数学専攻修士課程修了。同年，日本電信電話株式会社入社。NTT コミュニケーション科学基礎研究所を経て，現在，龍谷大学理工学部電子情報学科教授。複雑ネットワーク科学，データマイニングおよび機械学習の研究と教育に従事。博士 (理学)。日本数学会，日本応用数理学会，人工知能学会，電子情報通信学会各会員。