

地域の特徴を考慮したメリハリ・エリア型観光プランニング

森竹壮汰郎[†] 笠原 秀一^{††} 馬 強^{†††}

[†] 京都大学大学院 情報学研究科 〒606-8501 京都市左京区吉田本町

^{††} 大阪成蹊大学 〒533-0007 大阪市東淀川区相川3丁目10-62

E-mail: [†]moritake@db.soc.i.kyoto-u.ac.jp, ^{††}kasahara.hidekazu.k13@kyoto-u.jp, ^{†††}qiang@i.kyoto-u.ac.jp

あらまし 物見遊山や街歩きなど様々な観光のスタイルがあり、これに派生する多様なニーズに合わせた支援が求められる。既存手法の多くは、単一の観光スタイルに特化し時間や距離の最適化を行う。そこで本研究は、観光地を複数のエリアに分割し、エリアごとに観光スポット選択の粒度に緩急をつけ、観光都市の「らしさ」を味わう「メリハリ・エリア型観光」を提唱し、これを支援するプランニング手法を提案する。提案手法では、訪問候補のスポットをクラスタリングし、観光都市を複数のエリア（クラスタ）に分割する。次に、軌跡データと画像データから得た特徴を比較・分析し、エリアらしさを象徴する代表地を算出し、それを含む複数のエリア訪問プラン（コース）を生成する。それらをユーザの時間等の条件から最適化し緩急のついた観光プランを生成する。生成されるプランは、エリアごとに異なる観光スタイルを選択可能であり、同じエリアでもユーザごとに異なる観光スタイルを選択可能である。そのため、ユーザの目的やスタイルに最適化したプランを提供できる。

キーワード 観光 (Sightseeing), 観光プランニング (Sightseeing Planning), POI(Point Of Interests), クラスタリング (Clustering), 特徴抽出 (Feature Extraction)

1 はじめに

観光には様々なスタイルが存在する。例えば、有名・著名な観光スポットをできるだけ多く訪問することを目的とする「物見遊山型」や、現地を周辺散策することで街の雰囲気を味わうことを目的とする「まち歩き型」がある。これらのスタイルは、スポットの選択や滞在時間、移動範囲などの点で異なる。しかし実際の観光客の行動を注視すると、団体やツアー客を除いて旅行者が旅程を計画をする場合は、複数の観光スタイルを組み合わせていることが多い。そのため、複数の観光スタイルを組み合わせる取り組みを行うことで、実際の観光客のニーズや行動をより深く捉えた満足度の高い観光支援ができると考えられる。

既存の研究では、主に単一の観光スタイルに特化して取り組みが行われている。例えば、訪問スポットをクラスタリングし、時間や距離で最適化することで観光計画を行うもの[1]や、主要な観光地に対する周辺の散策地の数を評価することでまち歩きに適したスポットを分析するもの[2]がある。また、個々のPOIをその他のPOIと比較・分析した上での特徴（外見や知名度など）を詳細に考慮をしない。POIの特徴を比較・分析することで、都市・地域におけるPOIの位置付けが明らかになり、より都市・地域の魅力を味わい、異なる特色を持つ観光スポットを幅広く訪問できる観光計画が可能となると考えられる。

そこで本研究では、観光地を複数のエリアに分割し、エリアごとに観光スポット選択の粒度にメリハリ（緩急）をつけながら、観光都市の「らしさ」を味わう「メリハリ・エリア型観光」を提唱し、これを支援する。メリハリ・エリア型観光の概要を図1に示す。

この観光スタイルを実現するため提案手法は、ユーザが観光

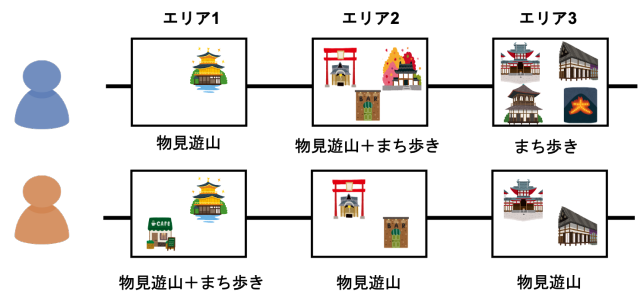


図1: メリハリ・エリア型観光の概要

都市の中で訪問候補としたPOI集合に対して、クラスタリング、代表地の抽出、コースの生成、プランニングの大きく4つの処理を行うことで旅程（メリハリ・エリアプラン）を組み立てる。生成されるプランは、エリアごとに異なる観光スタイルを選択可能であり、同じエリアでもユーザごとに異なる観光スタイルを選択可能である。

本研究における貢献は以下の通りである。

- 新たな観光スタイル「メリハリ・エリア観光」を提唱している。複数の観光スタイルを併用する手法を提案している。代表地という概念を導入することで、観光地におけるPOIの位置付けを考慮し、それぞれの観光スタイルに対応する複数の訪問プランを生成し、それらをユーザの条件に応じて最適化した旅程を生成する手法を提案している。[3節]
- 観光都市やエリアにおける代表地の抽出手法の提案をしている。代表地は Normalized Google Distance, Vision Transformer から得られた、認知の度合いや見た目の特徴を評価することで抽出する。[3.2節]

- 評価実験を行い、提案手法の有用性を検証した。京都府内の観光客の軌跡・画像データを適用し、代表地の妥当性や生成されたコース・プランの妥当性や緩急がつけられているかを主に評価した。[4 節]

2 関連研究

観光情報学の分野の研究は「どのようなスポットなのか」「どうやって行くのか」などの様々な切り口から取り組まれている。特に、SNS(Social Networking Services：ソーシャルネットワーキングサービス)やブログなどで生成されるUGC(User Generated Content：ユーザ生成コンテンツ)を対象に、利用者視点から観光地に関する観光集合知をマイニングし、観光計画に活用する試みが多数行われている[3][4]。

2.1 ツアープランニングについて

D.Gavalas ら[1]は、ユーザの時間予算や移動の効率を最適化したプランニングを行うため、POIをクラスタリングした後、時間予算に応じて複数のクラスタから順次POIを選択することでルート生成を行っている。しかし、POIをデータ点とみなし、個人の好みや地域の特性を考慮していない。本研究では、地域の特性を考慮したPOI選択を行い、エリアごとにコースを設定するため、エリアやPOIに対して滞在時間の重みづけを可能にしている。さらに我々の手法は、観光地(エリア・POI)の違いを考慮することで、異なる特徴をもつ場所をより多く訪問できるようにできる点が新規性として挙げられる。

X.Lu ら[5]は、ジオタグ付き写真から回収したジオタグ付き写真から、訪問スポットとスポット内の移動経路、スポットの滞在時間、スポットを訪問する順序、スポット間の移動時間などの軌跡情報を利用、ユーザーの好みに応じてカスタマイズした旅行ルート計画を提案している。また、J.Sun ら[6]は、軌跡データから観光客の移動パターンを分析し、場所と遷移の知識を組み合わせ、その両方に対して報酬を与える移動経路推薦手法を提案している。

2.2 まち歩き型観光について

原ら[2]は、事前設定した主要な観光スポットに対し、周辺散策スポットの総数に応じてまち歩きに適した度合いを算出する取り組みを行なった。しかし、複数の観光スタイルの切り替えはユーザの意思決定支援にとどめられており、観光スポットにおける地域や都市の特徴を考慮しない。我々が提案する手法は、「代表地」と呼ばれる地域や都市の特徴を考慮したスポット選択を行うことや、それをもとに生成された複数のコースを条件に応じて最適化することで、観光スタイルの切り替えを手法の中で行う点が新規性である。

2.3 POI分析について

Y. Shen や M.Ge ら[7]は、観光の価値を自然の価値と文化の価値の側面から推定し、ジオソーシャル画像の分析を行なった。J.Sun ら[8]は、観光客から収集した軌跡データからPOIを抽出し、重み付き HITS ベースのアルゴリズム[9]を適用さ

せることで、POIの性質を分類する試みを行なった。また、J.Sun と木上ら[10]は、異なる粒度のPOIを発見するための都市適応型クラスタリングフレームワークを考案した。フレームワークでは、データ点をDBSCANでクラスタリングしたのち、生成されたクラスタと、既知の主要スポットとの被覆度に基づいたスコアを組み合わせることで、異なる都市に適応的に対応する取り組みを行なった。

3 メリハリ・エリア型観光プランニング

この節では具体的な提案手法について説明する。概要を図2に示す。大きな流れは、軌跡データ、画像データ、ユーザの時間予算、訪問候補POI集合を入力とし、観光プラン(メリハリ・エリアプラン)を出力としている。まず、軌跡データ・画像データから構築されるPOIデータベースを構築する。その中からユーザは訪問候補POI集合を作成する。訪問候補POI集合に対して、実際の移動距離を基準としたクラスタリングを行いエリア(クラスタ)を生成する。次にエリアごとに、観光都市・エリアを象徴する「代表地」を抽出し、地域の特徴を考慮する。エリアごとに設定された代表地をもとに観光スタイルを表現する複数のコース(訪問プラン)を生成し、それらをユーザの時間予算に応じてコースを組み合わせ最適化することで、メリハリ・エリアプランを生成する。

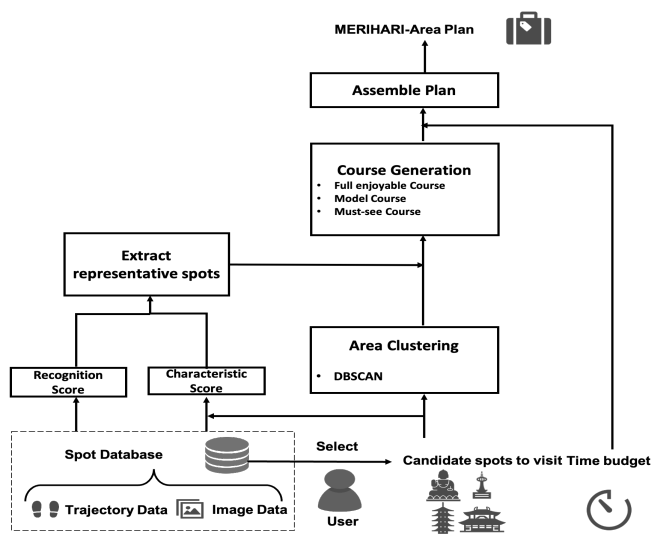


図2: メリハリ・エリア観光フレームワークの概要

3.1 エリアの生成

DBSCANを用いて、訪問候補POI集合からエリアを生成する。DBSCANでは、Open Route Service¹(以下、ORS)の経路検索機能で算出された徒歩での移動時間を用いる。ORSを用いることで、実際に通行可能な経路を通った場合のコストを考慮できる。そのため、直線距離では算出が難しい標高差があるケース[11]などに対応が可能である。今回は、クラスタの半径をORSで徒歩8～9分の移動範囲とし、データ点(POI)が

¹ : <https://openrouteservice.org>

1つのクラスタ(エリア)の存在を認める。

3.2 代表地の抽出

この節では、3.1節で生成されたエリアにおいて、特にその都市・エリアの特徴や知名度を表す代表地を抽出する方法について述べる。定義を以下に示す。

代表地

生成されたエリアの中で、特にその都市・エリアの特徴や知名度を十分にあらわす POI

代表地は、同じエリアで普遍的だが、他のエリアでは普遍的でない特徴や、都市全体の中でもそのスポットでしか見ることができない特異的な特徴を持つ POI を抽出することを目的としている。本研究では「代表」という性質を、認知度と固有度の2つの尺度から測る。

● 認知度

- 多くの観光客が訪れ、訪れた人が様々な媒体で発信し、広まるという循環に着目する。
- 例えば、「京都といえば XX」と呼ばれている等、広く認知されている程度を表す。

● 固有度

- 観光のモチベーションでもある「その都市・エリアらしさを感じる」という「らしさ」に着目する。
- 他の都市・エリアでは見られない、「ならではの(固有の)」特徴を表す。

3.2.1 認知度の算出

Normalized Google Distance(以下, NGD) は2つの単語の意味の関わり方の強さ、共起度合いを示す指標である[12]。2つの単語を Google 検索にかけた際の、検索結果の総数を式(1)に代入し、意味の関わり方の強さを算出する。なお、ここで N は Google が検索した web ページの総数、 $f(x)$ 、 $f(y)$ はそれぞれの検索語 x 、 y を検索した際のヒット数、 $f(x, y)$ は x 、 y 両方を同時に検索した際のヒット数を表す。本手法では、「(観光都市の地名) sightseeing」とそれぞれの POI の名称をクエリとし(例: “kyoto sightseeing” gion)、結果を算出することで、都市・エリアと POI の認知度を求める。なお今回は、各言語で検索した際の結果数の偏りを緩和するため、日本語・英語それぞれの言語でクエリを作成し、NGD を算出し、それらを式(2)のように調和平均をスコアとして用いる。

$$NGD(x, y) = \frac{\max\{\log f(x), \log f(y)\} - \log f(x, y)}{\log N - \min\{\log f(x), \log f(y)\}} \quad (1)$$

$$Score(x, y) = \frac{2(NGD_{Japanese}(x, y) \times NGD_{English}(x, y))}{NGD_{Japanese}(x, y) + NGD_{English}(x, y)} \quad (2)$$

3.2.2 固有度の算出

Vision Transformer(ViT) は、画像認識タスクに使用されている Transformer で、画像を小さなパッチに分割し単語のように扱うことでベクトルとして出力をしている[13]。ViT を用い

て画像からエリア・POI らしさを特徴ベクトルとして抽出し、特徴ベクトル間の類似度をエリア内・外で比較することで、そのエリアにおける特徴を抽出する。エリア内の類似度は、該当する POI の特徴ベクトルと同エリアに属する POI それぞれの特徴ベクトルに対して比較し、エリア外の類似度は、該当する POI の特徴ベクトルとエリアに属する POI の平均特徴ベクトルをエリアごとに比較を行う。

それらの比較の結果は、表1のように分類できる。エリア外類似度・エリア内類似度が共に低い特異的な特徴を持つ POI と、エリア外類似度が低く・エリア内類似度が高いエリア固有の特徴を持つ POI は固有度が高いと判定する。

3.2.3 代表地の決定フロー

認知度および固有度で得た POI ごとの特徴を代表スコアとして表現する。代表スコアは以下の式(3)で表現される。

$$RepresentativeScore = 0.6(1 - NGD) + 0.3(1 - Sim_{Areaout}) + 0.1(1 - Sim_{Areain}) \quad (3)$$

ここで、NGD は Normalized Google Distance のスコア、 $Sim_{Areaout}$ はエリア外の類似度、 Sim_{Areain} はエリア内の類似度の結果をそれぞれ表す。スコアの係数は、ハイパーパラメータのような位置付けで、著名・有名かつその都市・エリアのみでみられる特徴を持つ POI のスコアが高くなるよう設定している。

その後、代表地をエリアごとに図3に示すフローに沿って決定する。

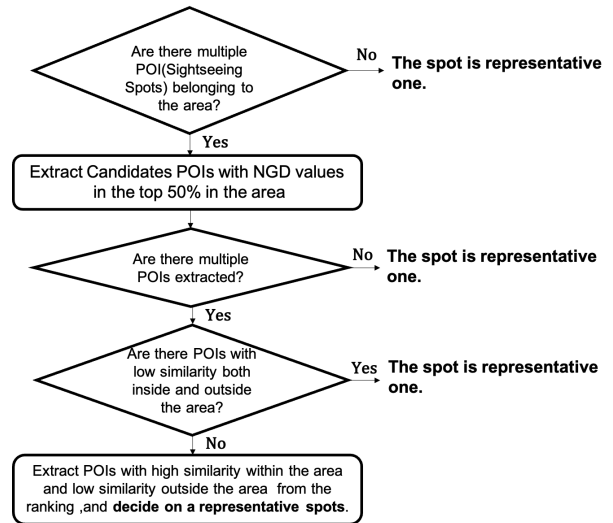


図 3: 代表地決定フロー

3.3 コースの生成

この節では、3.1節で生成された各エリアから、観光スポット選択の粒度に緩急をつけた3種類のエリア訪問経路(コース)を生成する手法について述べる。コースは訪問スポットの数や滞在時間が異なり、それぞれ以下のように定義する。

表 1: 特徴ベクトルのエリア内・外の比較

		エリア内類似度	
		高い	低い
エリア外類似度	高い	その都市において普遍的特徴を持つ POI	該当エリア以外の場所で同様の特徴を見ることができる POI
	低い	該当エリアの特徴や知名度を示す POI	その都市全体の中でも特異的で、都市・エリアを象徴する POI

満喫コース エリアに属する全ての POI を平均的な滞在時間で訪問するコース

モデルコース エリアに属する人気度の高い POI と代表地を平均的な滞在時間で訪問するコース

必見コース エリアに属する代表地のみを最小の滞在時間 (平均的な滞在時間での訪問も可能) で訪問するコース

生成にあたり、3.2 節で抽出された代表地と滞在時間の情報を用いる。各 POI の滞在時間は、軌跡データから滞在点を算出した際に滞在点に関連づけられている滞在時間を参照する方法 [8] を用いることで算出する。また、それぞれのコースの出力は、コースに属する POI 集合、コースに費やす時間 (移動時間・訪問時間)、コースごとの満足度の情報である。満足度は、式 (4) のようにコースに属する POI における 3.2.3 節で定義した代表スコアの線形和として定義する。

$$C_i^x = \sum_{k=1}^K RepScore_{xik} \quad (4)$$

ここで、 C_i^x はエリア x における i コースの満足度を指し、エリア x 、コース i に属する POI の数を K とする。

既存手法では、クラスタリングによるエリア生成のみを行うものが多く、観光プラン出力時には、クラスターに属する POI をすべて訪問することが想定された観光計画が出力される。本手法では、エリアごとに複数のコースを生成することで、ユーザーの条件に応じて観光スポット選択の粒度に緩急をつけた観光プランを組み立てることが可能となる。3 種類のコースはエリアごとに決定するため、満喫コースと必見コースを使い分けるなど、ユーザーの目的や好みに応じて抑揚をつけたメリハリのある観光プランの生成が可能となる。

3.4 メリハリ・エリアプランの組み立て

この節では、各エリアで生成された 3 つのコースを、ユーザーの条件内で選択、組み合わせることでメリハリ・エリアプランを生成する手法について述べる。

3.3 節で生成したコースを、観光における整数計画法の文脈 [14] [15] に基づいて、以下の式 (5) を定義し、ユーザーの時間予算 $Timebudget_u$ の範囲内でこれを最大化するメリハリ・エリアプラン $I = (Area_1 Course_i, \dots, Area_N Course_i)$ を生成する。このとき、各エリア、コースのスコアは式 (4) で定義される C_i^x を用いる。この式では、ユーザーの時間予算に合わせて「エリアらしさ」と「移動の効率性」を最適化することで緩急のあるプラン生成を可能にしている。そのため、全てのエリアから 1 つずつコースを選択するのではなく、よりスコアの高い

エリアおよびコースをユーザーの条件内で選択し、組み合わせることでプランを生成する。また、エリアごとに異なる観光スタイルを選択可能であり、同じエリアでもユーザーごとに異なる観光スタイルを選択可能である。

$$\text{maximize} \sum_{x=1}^N \sum_{y=1}^N \sum_{i=1}^3 \sum_{j=1}^3 (A_{x_i y_j} \times C_i^x) \quad (5)$$

ここで、 $A_{x_i y_j}$ はユーザーがエリア x の i コースに訪問した後、エリア y の j コースに連続して訪問した場合に 1 とするバイナリ変数である。この式 (5) を次のような制約のもと解く。

$$\sum_{i=1}^3 \sum_{y=1}^{N-1} \sum_{j=1}^3 A_{0_i y_j} = \sum_{x=1}^{N-1} \sum_{i=1}^3 \sum_{j=1}^3 A_{x_i N_j} = 1 \quad (6)$$

$$\sum_{x=1}^N \sum_{i=1}^3 A_{x_i w_k} = \sum_{y=1}^N \sum_{j=1}^3 A_{w_k y_j} \leq 1 \quad (7)$$

$$\forall w = 2 \dots N - 1, \forall k = 1, 2, 3$$

$$\sum_{x=1}^{N-1} \sum_{y=1}^N \sum_{i=1}^3 \sum_{j=1}^3 Time_{x_i y_j} \times A_{x_i y_j} \leq Timebudget_u \quad (8)$$

$$u_i - u_j + 1 \leq (N - 1) \times (1 - A_{x_i y_j}) \quad (9)$$

$$2 \leq u_i, u_j \leq N, \forall i, j = 2 \dots N$$

[14] [15] 等の既存の取り組みでは、スポット・エリア間の移動を表現する式を定義しているが、提案手法では各エリアで複数のコースを生成し、それらを選択・組み合わせる。そのため、本研究ではコースに関する制約 (総和計算の部分) をそれぞれ式 (6)、式 (7)、式 (8) で新たに考案し追加した。式 (6) は旅程はエリア 1 で開始し、エリア N で終了するというを、式 (7) は旅程の経路が接続されており、同じエリアを 1 度以上訪問しないことおよび、各エリアからは 3 つのコースのうち高々 1 つのコースが選ばれる制約を表現している。式 (8) は旅程に費やす全ての時間がユーザー u の時間予算 $Timebudget_u$ を超えないことを、式 (9) はサブツアーの排除を表現している。

4 評価実験

この節では、提案手法の有用性を立証するために実験について述べる。

4.1 データセット

京都府内の 91 箇所のスポットとその軌跡が含まれたデータセット [6] を用いた。データセットに含まれる各カラムの値の例を表 2 に示す。用いた軌跡データの中には、観光客が撮影した写真から推定したデータが含まれる場合がある。軌跡データ

表 2: データセットに含まれる各カラムの値の例

POI id	名称	経度	緯度	滞在時間 (min)
1	河原町	135.7654815	35.00701	60.00
2	二条城	135.749473	35.0139545	50.00
3	清水寺	135.784286	34.994452	42.41

から得られた平均滞在時間と Google Map や POI の公式 Web サイトに示されている平均滞在時間を比較し、2 倍以上の乖離がある場合は、後者の滞在時間を採用し修正を行った。表 2 において滞在時間のカラムの値が太字になっている箇所は、それを表す。また、POI の画像データについては、POI の画像を Google Custom Search API 経由で取得し、1 スポットあたり 10~30 枚程度使用した。

4.2 データセット適用結果：エリアクラスタリング

DBSCAN の結果を図 4 に示す。なお、データ点から探索を行う半径を表すパラメータを $\epsilon = 550$ (徒歩 9 分圏内) とし、クラスタ (エリア) を形成するために必要となるデータ点の最小数を表すパラメータを $minPts = 1$ とした。

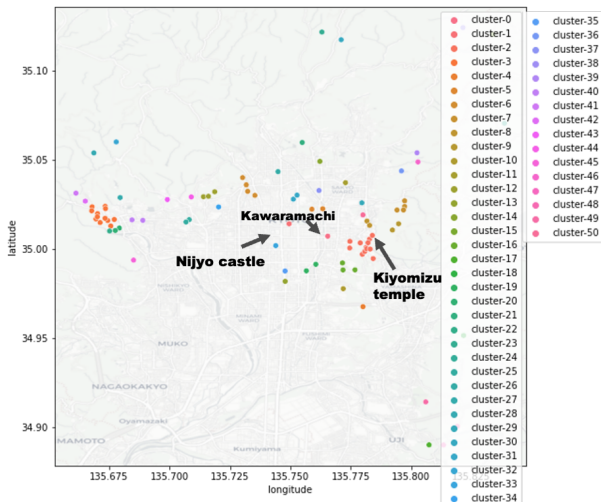


図 4: エリアクラスタリングの結果

4.3 データセット適用結果：代表地の抽出

4.2 節で得られたエリアにおいて、3.2.3 節の決定フローに従って代表地を抽出することができた。結果の一部を表 3 に示す。各行のすべての値が太字となっている POI が、抽出された代表地である。

4.4 データセット適用結果：コースの生成

4.2 節で得られたエリアにおいて、それぞれ 3 つのコースを生成することができた。結果の一部を表 4 に示す。太字で表示されている POI は、4.7.1 節で得られた、このエリアでの代表地を指す。今回、コースの作成では必見コースの POI 滞在時間を平均滞在時間と設定した。

表 3: 代表地の抽出例 (東山エリア)

POI 名	NGD	エリア外類似度	エリア内類似度
三年坂	0.2811	0.4114	0.4310
二年坂	0.3426	0.2980	0.3514
清水寺	0.3552	0.2400	0.2656
八坂神社	0.3567	0.6982	0.6138
高台寺	0.4248	0.7339	0.6260
建仁寺	0.4464	0.7282	0.6342
祇園	0.4801	0.5927	0.5504
知恩院	0.5219	0.7359	0.6410
青蓮院	0.5302	0.5099	0.4983
霊山観音	0.5689	0.4580	0.4327
円山公園	0.5692	0.3491	0.3808
京都霊山護国神社	0.5951	0.6992	0.6177

表 4: 生成コース例 (東山エリア)

満喫コース (7.94 時間)	モデルコース (3.33 時間)	必見コース (2.11 時間)
三年坂	三年坂	三年坂
二年坂	二年坂	二年坂
清水寺	清水寺	清水寺
八坂神社	八坂神社	
祇園	祇園	
建仁寺	建仁寺	
高台寺		
知恩院		
青蓮院		
霊山観音		
円山公園		
京都霊山護国神社		

表 5: メリハリ・エリアプラン例 (出発地：京都タワー，到着地：河原町，時間予算：8 時間)

訪問順	コース	訪問するスポット	所要時間 (時間)	コースの満足度
1	必見	京都タワー	0.32	0.793
2	モデル	清水寺，二年坂，三年坂 祇園，八坂神社	4.10	3.51
3	満喫	平安神宮，岡崎公園	1.76	1.10
4	モデル	河原町	1.77	0.634

4.5 データセット適用結果：メリハリ・エリアプランの生成

前節で得られたエリアおよびそれに付随するコースから、ユーザの時間予算に対して最適化されたプランを生成した。出発点を「京都タワー」(京都駅周辺)，到着点を「河原町」，時間予算を 8 時間とした時の提案手法の結果を表 5 に示す。なお、コースの満足度は 3.3 節の式 (4) から得られた値を指す。複数の観光スタイルを条件内で組み合わせることができている。

4.6 評価方法

大きく、手法自体の評価とユーザ実験による評価を行なった。手法自体の評価では代表地抽出指標の Ablation Study を行い、代表地の抽出指標を nDCG で評価した。ユーザ実験では開始点・到着点・時間予算から構成されるクエリを設定し、クエリ

表 6: 評価に用いたクエリ

出発地	目的地	時間予算 (時間)
京都タワー	河原町	4, 6, 8
京都タワー	嵐山駅	4, 6, 8
京都タワー	金閣寺	4, 6, 8

から生成されるコース・プランについて評価を行った。

4.6.1 手法自体の評価

提案手法の代表地抽出（認知度と固有度のスコアを組み合わせ）手法を評価する。複数のスポットから構成される全てのエリアにおいて、筆者による代表地の妥当順と組み合わせたスコアおよびその Ablation Study の順位を比較し、 $nDCG@k$ ($k=1, 3, 5, 10$) を用いて代表地の妥当性を評価した。

4.6.2 ユーザ実験による評価

ユーザアンケートを実施し、提案手法を評価した。参加者は 20 代 3 名で、いずれも京都府内の観光をした経験がある。開始点・到着点・時間予算から構成されるクエリ (式 (10)) を設定し、各クエリから生成される 9 種類のプラン (時間予算 3 種類×プラン 3 種類) に対して行なった。評価に用いたクエリを、表 6 に示す。

$$Query = (Startpoint, Endpoint, Timebudget) \quad (10)$$

プランは、提案手法を含む 3 つの手法を比較した。

提案手法 エリア・コースどちらも作成される。エリア、コースに対して最適化を行う。

エリア手法 エリアは作成されるが、コースは作成されない。エリアに対して最適化を行う。

スポット手法 エリア・コースどちらも作成されない。スポット集合に対して最適化を行う。

ユーザアンケートではクエリから生成されるコース・プランにおける、代表地の妥当性・コース・プランの妥当性（非効率な移動が少なく、時間内に訪問を完了可能か）、メリハリ度合いについて評価をもらった。

4.7 評価結果

4.7.1 手法自体の評価・ユーザ実験結果：代表地の抽出

代表地決定に用いた指標、式 (3) の Ablation Study については表 7 のようになった。この評価では、複数スポットから構成される全てのエリアに対して評価を行なった。ここから、代表地決定には提案手法が最も優れていることがいえる。

次に、ユーザアンケートの評価結果について述べる。提案手法が示した代表地と、被験者がつけた代表地に相応しい度合いを $nDCG$ を用いて評価した結果、以下ようになった。表 8 が評価対象のエリアとそれに属するスポットの名称、表 9 が各エリアの $nDCG$ の結果である。表 8 について、太字で示されるスポット名は提案手法が抽出した代表地を指す。この結果から、提案手法が生成する代表地には高い妥当性があることがいえる。

表 7: 代表地決定スコアの実験結果

	$nDCG@1$	$nDCG@3$	$nDCG@5$	$nDCG@10$	$nDCG$
提案手法	0.9028	0.9699	0.8700	0.9032	0.9869
NGD	0.8958	0.8950	0.8211	0.8599	0.9807
エリア内類似度	0.7014	0.8470	0.7196	0.8123	0.9779
エリア外類似度	0.7153	0.8141	0.8686	0.8804	0.9679
NGD+エリア内類似度	0.9028	0.9699	0.8700	0.9001	0.9864
NGD+エリア外類似度	0.9028	0.9589	0.8249	0.8889	0.9810
エリア外類似度+エリア内類似度	0.7569	0.8937	0.8686	0.9062	0.9788
NGD+エリア外類似度+エリア内類似度	0.8056	0.9588	0.8308	0.9204	0.9816

表 8: ユーザ実験で用いた評価対象のエリアとその代表地

エリア番号	スポット
1	京都タワー、東本願寺
2	平安神宮、岡崎公園
3	清水寺、三年坂、二年坂、祇園、八坂神社、知恩院、建仁寺、高台寺、円山公園、青蓮院、霊山観音、京都霊山護国神社
4	竹林の小径、渡月橋、嵐山駅前、天龍寺、嵐山公園、トロッコ嵐山駅、大河内山荘庭園、常寂光寺、清涼寺、祇王寺、二尊院、宝篋院
5	金閣寺、北野天満宮、平野神社、わら天神

表 9: ユーザ実験における代表地に関する評価結果

	エリア 1	エリア 2	エリア 3	エリア 4	エリア 5	平均
$nDCG$	1.000	1.000	0.761	1.000	1.000	0.952

ユーザアンケートにおける結果から、エリアに属するスポット数が少数で、それぞれの知名度・アクセスの良さに差がある場合、提案手法の抽出結果と被験者による代表地の抽出結果は一致することが示された。しかし、エリアに属するスポットの数が多い場合、提案手法の結果とユーザの結果に差が見られた。具体的に、表 8 の東山エリア (エリア 3) において、提案手法は抽出しなかったが被験者が代表地に相応しいと高い順位をつけたスポットに、祇園や八坂神社が挙げられる。なお、各 POI のスコアは表 3 で示した通りである。これについて、認知度が検索結果数に大きく依存することが考えられる。祇園では、認知度を 2022 年 10 月頃に算出したため、紅葉に関する記事が通常より多く検出され、紅葉スポットとしても名高い高台寺や建仁寺が祇園より認知度が高く算出されたと考えられる。また、八坂神社については固有度の類似性のため、提案手法は代表地から外したことが考えられる。八坂神社の朱塗りの西楼門は、同じエリアに属する清水寺の仁王門、別のエリアに属する下鴨神社の楼門等と色合いや形状が似ている。そのため、ViT が示す特徴からエリア内類似度・エリア外類似度がどちらも高く算出されたと考えられる。

また、認知度・固有度以外に代表地を選出する指標として、スポットのジャンルや訪問人数などが考えられる。代表という性質を多面的に考えることで、より納得感のある代表地を抽出できると考えられる。

4.7.2 ユーザ実験結果：コースの生成

エリアから生成されたコースの妥当性を、移動が理にかなっており、提示された時間内に訪問を完了できるかという観点から評価した。評価対象のエリアは、表 8 である。ユーザアンケートで、被験者に 5 段階で評価してもらった平均値を表 10 に示す。結果から、コースの妥当性について東山周辺の地区 (エ

表 10: コースの妥当性に関するアンケート結果

	満喫コース	モデルコース	必見コース	平均値
エリア 1	4.67	4.67	4.33	4.55
エリア 2	4.67	4.67	4.33	4.55
エリア 3	2.67	2.67	4.00	3.11
エリア 4	3.00	3.33	4.67	3.66
エリア 5	3.33	4.00	4.67	4.00

表 11: エリアプランニング例 (出発地：京都タワー，到着地：河原町，時間予算：8 時間)

訪問順	エリア内の訪問スポット	所要時間 (時間)
1	京都タワー，東本願寺	0.94
2	京都水族館	1.20
3	東寺	0.67
4	平安神宮，岡崎公園	1.81
5	聖護院	0.73
6	京都大学	0.60
7	河原町	1.80

表 12: スポットプランニング例 (出発地：京都タワー，到着地：河原町，時間予算：8 時間)

訪問順	エリア内の訪問スポット	所要時間 (時間)
1	京都タワー	0.31
2	東本願寺	0.46
3	平安神宮	1.12
4	岡崎公園	0.58
5	青蓮院	0.60
6	八坂神社	0.49
7	高台寺	0.58
8	二年坂	0.54
9	三年坂	0.71
10	建仁寺	0.64
11	祇園	0.54
12	河原町	1.23

リア 3) を除いては平均値が 3.5 以上を得られており，概ね妥当性のあるコースを生成できているといえる．この結果の考察として，被験者はピーク時におけるスポット・移動の混雑を考慮していることが考えられる．本実験は実際の軌跡データを用いてスポットの滞在時間を算出しているが，スポット間の移動は一定速度で移動が可能なることを想定しており，混雑を考慮していない．被験者は東山周辺の道路の混雑を理解した上で評価したため，コースの妥当性の評価が低くなったとされる．

4.7.3 ユーザ実験結果：プラン生成

各手法が生成したコースの妥当性を，非効率な移動が少なく時間内に訪問を完了可能かという観点で評価した．また，スポットの選択に緩急をつけられているかどうかを 5 段階で評価した．なお，「京都タワー ～ 嵐山駅」のエリア手法については，クエリを満たす時間予算が最低 8 時間のため，時間予算 4・6 時間は評価を行わなかった．なお，4.5 節と同じクエリに対して，比較対象として用いたエリア手法の結果例を表 11 に，スポット手法の結果例を表 12 に示す．

ユーザアンケートの結果，妥当性については比較手法のスポット手法（個々の POI 間の移動を最適化したもの）が最も

表 13: 各プランの妥当性に関するアンケート結果

	時間予算 (h)	提案手法	エリア手法	スポット手法
京都タワー ～ 河原町	4	3.67	3.67	4.67
	6	4.00	3.00	4.67
	8	3.00	3.00	4.33
京都タワー ～ 嵐山駅	4	4.67	-	4.67
	6	4.00	-	4.33
	8	3.33	3.00	4.00
京都タワー ～ 金閣寺	4	4.33	5.00	5.00
	6	3.67	4.33	4.00
	8	2.67	2.67	4.33

表 14: 各プランのメリハリ度合いに関するアンケート結果

	時間予算 (h)	提案手法	エリア手法	スポット手法
京都タワー ～ 河原町	4	3.00	3.33	2.00
	6	3.67	4.33	3.33
	8	4.33	4.33	2.67
京都タワー ～ 嵐山駅	4	2.67	-	4.33
	6	3.67	-	3.67
	8	4.33	4.33	2.67
京都タワー ～ 金閣寺	4	3.67	3.33	4.33
	6	4.00	4.33	4.00
	8	4.00	4.33	3.00

表 15: NGD ランキング上位 10 スポット

順位	スポット名	NGD 値
1	二条城	0.248
2	竹林の小径	0.277
3	三年坂	0.281
4	鞍馬寺	0.294
5	京都御所	0.301
6	貴船神社	0.304
7	岡崎公園	0.313
8	二年坂	0.343
9	金閣寺	0.347
10	京都タワー	0.347

妥当性が高いという結果が得られた．スポット手法は，個々の POI 集合に対して最適化を行うため，提案手法・エリア手法と比較すると POI 間の移動距離が短くなるように算出される．特に，河原町，東山，嵐山駅周辺は POI の数が豊富であり，それらを順に訪問する移動距離の短いプランが生成されたため，被験者が妥当性が高い（移動に無理がない）と判断した可能性が高い．また，スポットの選択に緩急をつけられているかに関するメリハリ度合いについて，提案手法は，時間予算が大きくなるほどメリハリ度合いが高い結果を得ることができた．実際に提案手法が生成したプランは，時間予算が大きくなるほどコースが切り替わる回数が増加する傾向が見られた．そのため，提案手法ならびにメリハリ・エリア型観光は長期間・広範囲の観光に適している可能性が考えられる．

また，式 (3) で使用した NGD，エリア外類似度ならびにコースのスコアのランキングを表 15，16，17 に示す．認知度・固有度をスコアリングする際の重みを調整することで，パーソナライズされたスコアリング，コース生成が可能になると考えられる．

表 16: エリア外類似度ランキング上位 10 スポット

順位	スポット名	エリア外類似度
1	嵐山モンキーパーク	0.101
2	河原町	0.150
3	渡月橋	0.191
4	京都水族館	0.204
5	京都タワー	0.226
6	清水寺	0.241
7	京都大学	0.264
8	竹林の小径	0.296
9	二年坂	0.296
10	嵐山公園 亀山地区	0.310

表 17: 代表スコアランキング上位 10 スポット

順位	スポット名	代表スコア
1	竹林の小径	0.867
2	清水寺	0.828
3	二年坂	0.804
4	三年坂	0.801
5	京都タワー	0.793
6	渡月橋	0.729
7	京都水族館	0.721
8	嵐山モンキーパーク	0.657
9	岡崎公園	0.644
10	河原町	0.634

4.8 スポットに関する考察

提案手法は、複数のスポットをクラスタリングしエリアを生成し、コース、プランを出力する。提案手法が十分に効果を発揮するためには、スポットが一定数以上あることが望ましい。また、スポット間の距離関係についても考察が必要である。異なる都市についても提案手法を適用させることで、提案手法を十分な精度で適用できるスポットの数について定量的に分析する必要がある。

5 まとめおよび今後の展望

本研究は複数の観光スタイルを組み合わせ、観光都市の「らしさ」を味わう「メリハリ・エリア型観光」を提唱し、それを支援するために時間予算などユーザの条件に応じてスポット選択の粒度に緩急（メリハリ）をつける観光プランを生成する手法を提案した。提案手法は、エリアらしさを象徴する代表地を抽出する点、エリアごとに複数の周遊ルートを生成し、それをエリアごとに切り替え・組み合わせることでユーザの目的や好みに合った緩急のあるプランが生成できる点で独自性がある。提案手法が、実際の観光客の行動をより捉えた満足度の高い観光支援につながることで、ならびに地域の特徴を味わう観光ができることを想定している。

今後の展望として、まず大規模なユーザ評価を行い提案手法の有用性を広く検証する。その後、提案手法の各部分についてさらに工夫を重ね、個々のニーズに対応し、よりメリハリの効いた満足度の高いプランを生成できるようになること、また長期間・広範囲の観光に対しても適応可能にすることが挙げら

れる。

謝辞 本研究の一部は総務省 SCOPE（201607008）による。

文 献

- [1] Damianos Gavalas, Vlasios Kasapakis, Charalampos Konstantopoulos, Grammati Pantziou, Nikolaos Vathis, and Christos Zaroliagis. The ecompass multimodal tourist tour planner. *Expert systems with Applications*, Vol. 42-21, pp. 7303–7316, 2015.
- [2] 原辰徳, ホーバック, 宮本瞭, 青地孝, 太田順, 倉田陽平. 周辺散策の見どころ情報の提示によるまち歩き観光プランニングの支援. 観光と情報: 観光情報学会誌, Vol. 16, No. 1, pp. 33–48, 2020.
- [3] 馬強. UGCを利用した観光資源の発見と推薦. 情報処理, Vol. 62, No. 11, pp. e12–e17, 2021.
- [4] 馬強. 観光情報学の最前線-観光の分散化と個人化を促進する集合知活用情報技術. 情報処理, Vol. 58, No. 3, pp. 220–226, 2017.
- [5] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. Photo2trip: generating travel routes from geo-tagged photos for trip planning. *ACMMM2010*, pp. 143–152, 2010.
- [6] Junjie Sun, Chenyi Zhuang, and Qiang Ma. User transition pattern analysis for travel route recommendation. *IEICE Trans. Inf. Syst.*, Vol. 102-D, No. 12, pp. 2472–2484, 2019.
- [7] Yizhu Shen, Min Ge, Chenyi Zhuang, and Qiang Ma. Sight-seeing value estimation by analysing geosocial images. *Int. J. Big Data Intell.*, Vol. 5, No. 1/2, pp. 31–48, 2018.
- [8] Junjie Sun, Yuta Matsushima, and Qiang Ma. Property analysis of stay points for POI recommendation. *DEXA 2021*, pp. 374–379, 2021.
- [9] Longzhuang Li, Yi Shang, and Wei Zhang. Improvement of hits-based algorithms on web documents. *WWW2002*, pp. 527–535, 2002.
- [10] Junjie Sun, Tomoki Kinoue, and Qiang Ma. A city adaptive clustering framework for discovering pois with different granularities. In *DEXA 2020*, pp. 425–434. Springer, 2020.
- [11] 森竹壮太郎, 笠原秀一, 馬強. Multi-period 観光にむけた POI クラスタリング手法. *DEIM Forum 2022*, pp. B43–4, 2022.
- [12] Saket Karve, Vasisht Shende, and Swaroop Hopale. Semantic relatedness measurement from wikipedia and wordnet using modified normalized google distance. In *Data Analytics and Learning*, pp. 143–154. Springer, 2019.
- [13] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [14] Clair E Miller, Albert W Tucker, and Richard A Zemlin. Integer programming formulation of traveling salesman problems. *Journal of the ACM (JACM)*, Vol. 7, No. 4, pp. 326–329, 1960.
- [15] Kwan Hui Lim, Jeffrey Chan, Christopher Leckie, and Shanika Karunasekera. Personalized trip recommendation for tourists based on user interests, points of interest visit durations and visit recency. *Knowledge and Information Systems*, Vol. 54, No. 2, pp. 375–406, 2018.