

# 言語情報と地理情報を融合した魅力的な経路案内

大滝啓介 吉村貴克 徳久良子

株式会社豊田中央研究所

{otaki, yoshimura, tokuhisa}@mosk.tytlabs.co.jp

## 概要

経路案内アプリを用いる際、移動時間などの客観的な指標だけではなく、経路景観の「楽しさ」や「美しさ」といった主観的な指標も用いることで、より魅力的な経路案内が実現できると考えられる。本研究では、事前に収集された経路上の景観画像、景観画像の説明文、基盤モデルを用いて魅力度スコアを計算する枠組みを構築し、経路探索に組み込むことで、魅力的な経路を案内する手法を提案する。本稿では提案手法を実装したプロトタイプについて説明し、アンケート調査を行った結果を報告した上で、今後の課題や展望について述べる。

## 1 はじめに

移動行動は、人との交流や、情報・モノの獲得のための基本的な行動であり、健康維持や幸福感とも関連する [1, 2]。しかし、オンライン会議やフードデリバリーなどのサービスの発展にともない、人々の移動行動が減少しつつある。移動行動の減少は、人々が集う街の持続的な発展にも悪影響を与えるため、重要な社会課題として認識されている。そこで我々は、人々の移動を活性化させる手段の一つとして、スマートフォンなどのデバイスが発展した現代の街歩きや移動に欠かせないナビアプリに介入することを考えている。

ナビアプリに実装される経路案内や経路推薦を通じて移動を活性化するためには、移動時間や距離などの客観的な情報だけではなく、街が持つ「楽しさ」や「美しさ」などの主観的な情報を考慮すると効果的であろう。しかし、これらの客観的な情報を統一的に扱うことは難しい。また景観に関する評価が個人の主観的な評価・感覚に依存するため、大規模な実験を行うことや、様々な地域で詳細なユーザー調査を行うことは、いずれも容易ではない。

近年発展している基盤モデルを活用することで、言語や画像・地理情報などの多様なデータを活用

した情報処理が可能になることが期待されている [3, 4]。また地理情報処理の分野において、街の(知覚的な)歩行可能性 (Walkability) を、深層学習技術を用いて評価する研究が注目されている [5, 6]。本研究では、これらの技術を拡張して、個人の好みに適合した魅力的な経路を案内する経路探索手法を実現することを目指している。

個人に適合した経路を案内するためには、ユーザーの経路選択を調査したデータセットが欠かせない。本稿では大規模なデータセットの構築に向け、知覚的な歩行可能性スコアを拡張した魅力度スコアを定義する。計算したスコアを経路探索アルゴリズムに埋め込んで複数の経路を計算し、データ収集を行うためのプロトタイプ環境を整備する。用意した環境を用いて小規模なアンケート調査を行い、経路の選択傾向と風景画像の関係性を調べる。最後に得られた結果や今後の展望について述べる。

## 2 手法

提案手法は以下の2段階の処理から構成される。

- 対象地域から抽出した風景画像と風景を説明する文の類似度を、CLIP<sup>1)</sup> [7] により学習されたモデルを用いて測定し、魅力度スコアを計算する。
- 魅力度スコアと移動距離を用いて多目的経路探索を行い、経路の候補を複数取得する。

以降ではそれぞれの構成要素について述べる。

### 2.1 景観の魅力度評価

本稿では知覚的な歩行可能性スコア [6] を拡張した魅力度スコアを定義し、経路の魅力度を評価する。以下、知覚的な歩行可能性スコアについて概説し、それを拡張した魅力度スコアについて述べる。

**知覚的な歩行可能性スコア** 画像を説明する正と負のプロンプトを用意し、プロンプトと風景画像の類似度を測定する。具体的には、正と負のプロンプト

1) Contrastive Language-Image Pre-training



図 1: 知覚的歩行可能性スコア  $S(I) = 1.274$  を持つ画像の例。類似度の高い 3 つの文として、正のプロンプト「There are open green zone beside the road」、  
「There are playground beside the road」と負のプロンプト「There is a narrow sidewalk」が対応した。

の集合を  $X_{pos}, X_{neg}$  とするとき、ある画像  $I$  との関連度を確率値  $P(x_i)$  ( $x_i \in X := X_{pos} \cup X_{neg}$ ) で計算し、画像  $I$  の知覚的歩行可能性スコアとして

$$S(I) := \sum_{x_i \in X} d_i p(x_i) + H(X_{pos}) - H(X_{neg}) \quad (1)$$

を定義する。ただし  $d_i$  は  $x_i \in X_{pos}$  であれば +1、そうでなければ -1 となる係数、 $H(X_{pos})$  と  $H(X_{neg})$  は正と負のプロンプト集合上で計算されたエントロピーを表す。例えば、図 1 の画像を入力として、風景を説明する文集合との類似度を計算すると、類似度の高い文として「There are open green zone beside the road」および「There are playground beside the road」が検索される。逆に、類似度が低い文として「There is a narrow sidewalk」が検索される。式 (1) によってスコアを計算すると、知覚的歩行可能性スコアは  $S(I) = 1.274$  となる。

**魅力度定量化のための提案手法** 上記の知覚的歩行可能性スコアの考え方に基づき、プロンプトを設計することで魅力度スコアを定量化する。具体的には、知覚的歩行可能性スコアのプロンプト  $X_{pos}, X_{neg}$  から、それぞれ魅力度 (Attractiveness) に相当する部分を抽出して拡張したプロンプトを表 1 のように定義して用いた。なお、プロンプトを最適化する手法 ([8] など) を適用したり、専門家の知見を考慮したりして、より詳細なプロンプトを構築することは今後の課題である。

## 2.2 多目的経路探索

辺コストが複数付与された無向グラフ  $G = (V, E, \vec{c})$  における多目的経路探索は、2 点  $v_s, v_g \in V$

表 1: 用いたプロンプトの一覧。ただし下線つきのインデクス  $i$  は [6] のプロンプトをそのまま利用したものである。

$i$	$x_i$	$d_i$
<u>1</u>	There are vehicles parked on the sidewalk	-1
<u>2</u>	There are scooters parked on the sidewalk	-1
<u>3</u>	There are bicycles, and motorcycles parked on the sidewalk	-1
4	There is a wide sidewalk	1
5	There is a narrow sidewalk	-1
6	There is a heightened sidewalk	1
7	There is a tile pavement	1
8	There are cracks, depressions, and flooded sidewalks	-1
9	There are skyscrapers	-1
10	We can see the sky clearly	1
11	There are various heights of buildings	1
12	The heights of the buildings are the same	1
13	There is graffiti on the walls	-1
14	The walls of buildings share the same strong theme	1
15	There are billboards or advertising signs on the pavement	-1
16	There are garbages on the road	-1
17	Roads are clean and free of litter	1
18	There are trees	1
19	There are shadow areas	1
20	There are benches along the sidewalks	1
21	There are unleashed dogs on the sidewalk	-1
22	There are shops along the road	1
23	There are no stores along the road	-1
24	There are cafes along the road	1
25	The road is deserted	-1
26	There are public buildings such as hospitals, schools, libraries, and office complexes	1
27	There are open green zone beside the road	1
28	There are playground beside the road	1
29	The road is crowded and there are no playground beside the road	-1
30	There are bus or underground stations	1
31	There are landmarks	1
32	We cannot see any landmarks	-1

をつなぐ経路  $\pi_{v_s, v_g}$  のうち、コストベクトルが非支配な経路の集合を求める問題である (技術的な詳細は [9] などを参照)。本稿では紙幅に関係から詳細は述べないが、距離  $c_{\text{dist}}(e)$  と魅力度スコア  $c_{\text{attr}}(e)$  を用いて、2 次元の多目的経路探索を BOA\* アルゴリズムで求める。ただし  $c_{\text{attr}}(e)$  は大きい方が望ましい最大化問題となるため、定数シフトを行って最小化問題に変換した。このとき BOA\* アルゴリズム単体では厳密なパレート最適解は求まらないため、後処理をして近似解を求めた。図 2 に、本稿の提案手法 (手法 2) と 3 章で用いるグラフ上で計算した多目的経路を可視化した例を示す。また図 2 中の最短経



図 2: 2 地点 (S と G) を結ぶ最短経路 (赤) と魅力的な経路 (青) を Mapbox Static Image API で可視化した。



図 3: 最短経路と抽出された風景画像の例。

路 (赤) に紐づいて選択された風景画像と経路の例を図 3 に示す。

**魅力度スコアのグラフ構造への埋め込み** 一般に、ある道路ネットワークの辺上に複数枚の画像が対応する場合と、位置情報付きの画像データのスパース性から画像が存在しない場合が考えられる。本稿では以下の手法ですべての辺に魅力度スコアを付与する。まず複数画像が存在する場合、辺  $e$  に対応する画像の集合  $\mathcal{I}_e$  についての魅力度スコアの平均値  $\frac{1}{|\mathcal{I}_e|} \sum_{I \in \mathcal{I}_e} S(I)$  を辺  $e$  の魅力度スコアと  $c_{\text{attr}}(e)$  として設定する。次に頂点の魅力度スコアを、接続する辺の魅力度スコアの平均値に設定した上で、辺の更新した魅力度スコアを両端点の魅力度スコアの平均とする。この操作をすべての辺に魅力度スコアが付与されるまで繰り返し、魅力度スコアが埋め込まれたグラフ構造を用意する。

### 3 実験と考察

本稿では 2 章で構築した手法を用いて図 3 に示す経路案内を複数作成し、風景画像が経路選択に与える影響を調査するアンケート調査を行った。実験の対象地域として東京都新宿区の東側を用いた<sup>2)</sup>。図 4 は Mapillary から収集したデータと OpenStreetMap



図 4: グラフ構造と収集画像の地点。赤色の辺は画像から魅力度スコアが最初に付与された辺を示す。このグラフを入力として、2.2 節で述べた手法を用いて全ての辺にスコアを付与して用いる。

から作成した新宿区のグラフ構造を可視化したものを示す。得られたグラフ構造  $G$  は、頂点数  $|V| = 1,509$ 、辺数  $|E| = 1,850$  であった。

#### 3.1 準備

**データの準備** 表示する 2 つの経路 ( $P, P'$ ) と付随する景観画像の情報を以下の手法で求めた。

1. 提案手法を用いて、対象地域から取得したグラフ  $G$  に  $c_{\text{attr}}(e)$  を付与した。
2. 歩行距離が約 500[m]、750[m]、1000[m] となる出発地・目的地のペアについて多目的経路を検索し、経路に対応する画像集合を取得した。
3. 最短経路  $P$  と風景画像集合  $\mathcal{I}(P)$  を取得した。
4. 風景画像の影響を調べるため、 $P$  以外の経路  $P'$  について、集合  $\mathcal{I}(P)$  と  $\mathcal{I}(P')$  の Jaccard 係数が 0.3 未満となる経路のうち、歩行距離が最短のものを比較対象  $P^\dagger$  として設定した。

**アンケート調査** クラウドワークス<sup>3)</sup> のタスク機能と JotForm<sup>4)</sup> を用いて、2 経路  $P$  と  $P^\dagger$  のうちどちらを歩いて移動するかをアンケート調査した。アンケート調査を実施する前に、弊社の倫理審査委員会の審査を受けた (番号: 23A-29)。アンケートページでは実験同意書を表示した後に、図 2 のように 2 地点 (S と G) を結ぶ 2 経路を先に表示し、次に図 3 のように経路 A・経路 B のそれぞれについて経路上の風景画像を表示し、最後にどちらの経路を歩いて移動するかを聞いた。またアンケート中にはダミー

2) 緯度経度ではおおよそ [35.6859, 35.6989] と [139.6987, 139.7090] の領域を用いた

3) <https://crowdworks.jp/>  
4) <https://www.jotform.com/>

表 2: 各経路を選んだ人数と表示された経路の距離

設定距離 [m]	経路 A(距離 [m])	経路 B(距離 [m])
500	20 (530)	10 (617)
750	9 (780)	21 (796)
1000[	11 (1030)	19 (1109)

の質問を 2 つ追加し、両方に正答した回答者の回答結果のみを抽出した。

### 3.2 アンケート結果と考察

**選択結果** 関東圏に居住する 30 人の回答者が、移動距離を 500[m]、750[m]、1000[m] 前後で作成した経路選択のアンケートに回答した。30 人の回答者の内、男性が 14 名、女性が 16 名であった。回答者の年代別の人数は、20 代が 3 名、30 代が 14 名、40 代が 10 名、50 代が 3 名であった。また 30 人の中で 16 名が東京在住であり、残りの 14 名は東京以外の関東地方の県在住であった。

3 つの経路選択の問題設定で、経路 A (最短経路) と経路 B (魅力的な経路) のどちらを選んで移動するかを聞いた結果を表 2 に示す。実験結果より、街歩きの際に、歩行距離が極端に伸びない場合には、景観を考慮した経路が受け入れられる余地があることが分かった。

**自由記述** 最後に自由記述で得られた結果を述べる。経路 B (魅力的な経路) に対するポジティブな意見として、「道が真っ直ぐのほうが分かりやすい感じがします」「歩くのであれば、車が少なく静かそうな道を選びたい」「数分くらいの違いであれば距離の近さより景色を優先したい」という意見が得られた。

一方で経路 B (魅力的な経路) に対するネガティブな意見として、「経路 B は遠回りだし、大通り沿いをただ歩く感じになると画像から感じるので選択したくないです」「階段があるので避けたいと思いました。」「全体的にうるさくない経路だと思い選びました。」という意見が得られた。

**考察** 目的地到着までの余暇時間に街歩きする状況を設定しているため、景観の良い経路を提示することや、風景画像を表示した上で経路を選択するタスクに抵抗が少なかったと考えられる。回答全体を通じて、遠回りすることへの拒否感は少なかった。この結果より、適切な指示と誘導によって、最短経路以外の経路へ誘導したり、街歩きに貢献しうる魅

力的な経路を提示する余地があると考えられる。一方で景観が良い経路を選ぶにあたり無駄な遠回りや、人間が不自然に感じる迂回路が発生することがあり、ネガティブな回答が得られる場合もあることが分かった。

実験の設定や経路計算の都合として、画像がない場所の扱いが難しく、アンケート調査に提示できる事例に限られた。より高精度なデータを用意することや、モデルのファインチューニングを通じてより差のつくペア  $P, P'$  を提示することで、一般的な知見を調べられる可能性がある。また世の中の都市の景観評価・街歩きの評価などとの指標との関連性や整合性を調べることも重要である。先に述べたようにプロンプト自身を改良することで既存の手法と相関が高い魅力度スコアを自動評価できれば、マルチモーダル深層学習を活用した有用な技術となることが期待される。

## 4 まとめ

マルチモーダル深層学習モデルを活用し、魅力的な経路を案内するタスクに取り組んだ。具体的には、画像を説明する自然言語文を用いて魅力度スコアを計算し、経路案内に組み込み、経路とともに画像を提示するシステムを構築した上で、アンケート調査を実施した。結果より、景観の良い経路を提示する技術が受け入れられる余地があることを明らかにした。一方で、より自然な経路を計算したり、データ不足を補う必要があることも明らかになった。今後はプロンプトの最適化やデータ増強、経路探索アルゴリズムの改良などを通じて、個人化経路探索手法を深化させるとともに、より大規模なユーザ調査や UX 検証などを行う。

## 参考文献

- [1] Aaron S Heller, Tracey C Shi, CE Chiemeka Ezie, Travis R Reneau, Lara M Baez, Conor J Gibbons, and Catherine A Hartley. Association between real-world experiential diversity and positive affect relates to hippocampal–striatal functional connectivity. **Nature Neuroscience**, Vol. 23, No. 7, pp. 800–804, 2020.
- [2] Richard Layard and Jan-Emmanuel De Neve. **Wellbeing: Science and Policy**. Cambridge University Press, 2023.
- [3] Rishi Bommasani, Drew A Hudson, Ehsan Adeli, Russ Altman, Simran Arora, Sydney von Arx, Michael S Bernstein, Jeannette Bohg, Antoine Bosselut, Emma Brunskill, et al. On the opportunities and risks of foundation models. **arXiv preprint arXiv:2108.07258**, 2021.
- [4] Ce Zhou, Qian Li, Chen Li, Jun Yu, Yixin Liu, Guangjing Wang, Kai Zhang, Cheng Ji, Qiben Yan, Lifang He, et al. A comprehensive survey on pretrained foundation models: A history from BERT to ChatGPT. **arXiv preprint arXiv:2302.09419**, 2023.
- [5] Filip Biljecki and Koichi Ito. Street view imagery in urban analytics and GIS: A review. **Landscape and Urban Planning**, Vol. 215, p. 104217, 2021.
- [6] Xinyi Liu, James Haworth, and Meihui Wang. A new approach to assessing perceived walkability: Combining street view imagery with multimodal contrastive learning model. In **Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Spatial Big Data and AI for Industrial Applications**, pp. 16–21, 2023.
- [7] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In **Proceedings of the ICML2021**, pp. 8748–8763, 2021.
- [8] Chrisantha Fernando, Dylan Banarse, Henryk Michalewski, Simon Osindero, and Tim Rocktäschel. Promptbreeder: Self-referential self-improvement via prompt evolution. **arXiv preprint arXiv:2309.16797**, 2023.
- [9] Zhongqiang Ren, Richard Zhan, Sivakumar Rathinam, Maxim Likhachev, and Howie Choset. Enhanced multi-objective a\* using balanced binary search trees. In **Proc. of SOCS2022**, pp. 162–170, 2022.