

空間キーワードと人名に基づく空間演算による スポット検索手法の提案

上坂 佳[†] Panote Siriaraya[†] 王 元元^{††} 河合由起子^{†,†††}

[†] 京都産業大学コンピュータ理工学部 〒603-8555 京都府京都市北区上賀茂本山

^{††} 山口大学大学院創成科学研究科 〒755-8611 山口県宇部市常盤台 2-16-1

^{†††} 大阪大学サイバーメディアセンター 〒567-0047 大阪府茨木市美穂ヶ丘 5 番 1 号

E-mail: [†]{g1644123,kawai}@cc.kyoto-su.ac.jp, ^{††}spanote@gmail.com, ^{†††}y.wang@yamaguchi-u.ac.jp

あらまし POI やレストランなどのスポット検索では、店舗名や観光名所、地名といった空間的なキーワードの想起が容易であるが、それら空間キーワードとゆかりのない人物名をキーワードとする検索は困難である。本研究では、先行研究の空間演算を用いることで、空間キーワードと人名によるスポット検索手法を提案する。まず、Wikipedia やニュースなどのオープンデータからユーザが入力した人物の特徴語（名詞）を算出し、人物特性として抽出する。次に、空間演算に基づき空間キーワードを起点とした閾値以上の特徴語のペアを含む地物を検索し、最後に検索結果の積集合からスポットを抽出する。これにより、例えば「ハリウッド 大谷翔平」といった空間と人名キーワードからハリウッド周辺の大谷選手の好みのクレープ店舗や和菓子屋などが検索でき、人物とゆかりのない場所でもその人物と関連あるスポットを発見できる。本論文では、空間演算による空間および人名キーワードに基づく人名と関連性の高いスポット検索手法について述べる。

キーワード 時空間制約, 空間演算, スポット検索

1 はじめに

文書や地図などの検索サービスの需要は増しているが、POI やレストランといった空間や時間制約を含めた地図における地物オブジェクト（スポット）の検索は可能であるが、人物など動的なオブジェクトに対する地図検索は困難である。例えば、空間の問い合わせとして代表的な地図検索では「ハリウッド」レストラン」といったハリウッドという空間制約におけるレストランなどのスポット検索は可能であるが、「ハリウッド」大谷翔平」の検索では、大谷翔平とハリウッドのキーワードを含む文書検索は可能であるが、ハリウッドの地図空間に大谷翔平と関連するスポットを検索することはできない。要因の一つとして、空間キーワードのハリウッドと大谷翔平という人物との関連性が発見できないことがあげられる。

そこで、本研究では、動的なオブジェクトとして代表的な人物と関連する地図空間上のオブジェクトを発見することで、空間キーワードと人名による地図空間におけるスポット検索手法を提案する。これにより、例えば「ハリウッド」大谷翔平」といった空間と人名キーワードからハリウッド周辺の大谷選手の好みのクレープ店舗や和菓子屋などが検索でき、人物とゆかりのない場所でもその人物と関連あるスポットを発見できる。

提案手法では、先行研究の空間演算に基づく地図のスポット検索を用いる [1]。先行研究の空間演算では従来のテキスト検索で用いられている「空白 ()」に着目し、これまで考慮されていなかった空白の長さを、空間を表現する演算子として定義することで、2つのキーワードを起点と検索対象オブジェクトと



図 1 空間演算による地図検索結果例

して検索できる。例えば、既存手法では「A」B」も「A」B」も「A」とBが同時に含まれる」という問合せ文が生成されるため、検索結果は同じになる。ここで2つのワード間の空白を考慮することで「A」を起点とした「B」の問い合わせを実現でき、「A」B」はAから最も近いB、「A」B」はAから近い5件のBとなる空間制約の検索が可能になる。また、空白長を距離として用いることが可能で、例えば、Aを起点として5km離れたBの空間検索の問い合わせは「A」5km」B」となる。さらに、空間演算子として集合演算子と新たな範囲、方向、時間などを含む10種類を定義、実装しており、これらを用いることで方向検索（演算子:「*」「^」「_」「@」）、範囲検索（「□」または差演算）などの集合演算による問い合わせができる。図3に各空間演算による地図検索結果例を示す。(a)は方向演算「^」による北方向の問い合わせ結果例を示しており、(b)は差演算「-」による周囲検索結果、(c)は積演算「*」による2地点間で共通する店舗検索例である。本研究では起点の場所における地物オブジェクト検索に空間演算を用いることで、人物の好みの地物や苦手な料理を提供するレストランを外した検索が可能に

なる。

人物の特徴抽出は、Wikipedia やニュースなどのオープンデータからコーパスを構築し、ネガティブワードとポジティブワードを機械学習より特徴語として抽出する。地名と特徴語をポジティブな場合は和集合演算、ネガティブは差集合演算式に変換し、空間演算よりスポットを抽出する。

本論文では、空間演算による空間および人名キーワードに基づく人名と関連性の高いスポット検索手法の概要を提案する。

2 関連研究

2.1 複数ユーザに対する観光案内サービス

観光情報では複数人の旅行での様々なサービスが提供されている。検索作業を伴わずに、複数人の嗜好を反映させた観光地を推薦するシステムでは、まず各参加者の嗜好を抽出し、その結果を集団意思決定手法を用いて統合し、嗜好に適した観光地の推薦を行う [2]。また、スポットごとに最も有効な環境要因の組合せを用いる環境要因モデルと、協調フィルタリングを用いて旅行者の嗜好を記述する人的要因モデルを融合させた観光行動モデルを用いて、旅行ルートを推薦する手法が提案されている [3]。

2.2 Web および位置情報 SNS 分析によるスポット推薦

スポット抽出手法は主に Web 情報を用いる方法と SNS 情報を用いる方法に分類される。位置情報サービスがユーザの行動に与える影響を分析し、その影響を活用した POI 推薦手法の提案。チェックイン回数、チェックイン場所、チェックイン間の移動距離などを分析し行動への影響を評価し、位置情報サービスの利用者に対する POI 推薦手法が提案されている [4]。また、画像テキスト間対応を利用したスポット情報拡張手法を提案として、抽出した画像特徴量 BoVW として利用することで、ユーザの嗜好に合わせた画像ベースのスポット情報推薦を実現し、画像情報とテキスト情報が投稿されているスポットから、画像特徴語の対応を構成する手法が提案されている [5]。

POI の所在地を表すために用いられる場所の表現である POI の関連地名を、ブログなどの Web 上の文書から抽出することによって、POI に関連地名を付与する手法が提案されている [6]。

2.3 地図およびルート検索

地図検索ではユーザの入力した検索クエリによって検索精度や効率は大きく影響される。しかし、既存の地図検索サービス (Google Maps や Bing Maps など) の検索クエリは、単純なロケーションを含めたキーワードや文章で構成されている。例えば、ユーザが京都駅にあるレストランを検索する場合に Restaurants_in_kyoto station」の文章で検索可能だが、距離や件数といった明示的な空間制約を含めた検索は困難である。本論文で提案するスポット検索は、それらのワードの空間の関係を演算式の形式としてクエリ生成でき、直感的で精度の高い検索を実現できる。

地図検索のうち、特にルート検索の研究も活発に行われている。景色のよい、心地の良い、安全性などの基準を考慮し、二地

点間の最適なルートを検索するシステムが提案された [7] [8] [9]。また、最短なルートだけでなく、方向転換 (曲がる) 回数をできるだけ少なくすることで、記憶しやすいルートを検索する手段も提案されている [10]。さらに、スマートフォンといったモバイル端末における地図検索を利用する際の画面制約の解消や操作性の向上を目指したインターフェイスの研究も存在する [11]。

2.4 意味的空間データ検索

近接演算子を用いた検索対象となる項目内での検索語間の距離や位置関係を指定して検索する近接演算がある。例えば、ScienceDirect や Scopus では近接演算子 (W/n) を用いて「microscopy W/3 gfp」のように「microscopy」と「gfp」を語順を問わず 3 語以内に含む文献を検索できる [12]。Oracle Text では近接演算 (near) を用いて「near((dog, cat), 6)」のように「cat」の 6 ワード以内に「dog」を含むすべてのドキュメントを検索できる [13]。また、テキスト検索では、Mikolov ら [14] は、ニューラルネットワークを用いて、Word2Vec という語のベクトル表現を獲得する手法を提案している。この手法では、類似語を抽出するのみではなく、「king - man + woman = queen」のように語の意味演算を行うことが可能とされている。本研究では、起点に対する距離や範囲などのワード間の空間制約を用いた集合演算を可能にしており、空間演算によりキーワードのみで地図上の空間的かつ意味的な近さを含めた検索を実現できる点が特徴となる。

3 空間キーワードと人名に基づく空間演算

本研究では、動的なオブジェクトとして代表的な人物と関連する地図空間上のオブジェクトを発見することで、空間キーワードと人名による地図空間におけるスポット検索手法を提案する。提案手法では、先行研究の空間演算に基づく地図のスポット検索を用いる。次に、Wikipedia やニュースなどのオープンデータからユーザが入力した人物の特徴語 (名詞) を算出し、人物特性として抽出する。最後に、空間演算に基づき空間キーワードを起点とした閾値以上の特徴語のペアを含む地物を検索し、最後に検索結果の積集合からスポットを抽出する。本章では、空間演算と人物特徴抽出方法について述べる。

3.1 空間演算の定義

地図上の複数の地点を起点とした同心円上の集合演算を可能とする空間演算について述べる。提案する空間演算は下記を前提とする。

定義 1: 最小の空間演算の問い合わせ (SQ) は「 $A_{\perp}$ [空間演算子] 空間長 $_{\perp}\alpha$ 」の形式とする。

A と α はキーワードであり、A を起点とした空間上の α の検索を意味する。先行研究より提案した空間演算子 [1] に続いて数値と単位を記述する。

(例) 「 $A_{\perp}1km_{\perp}\alpha$ 」は「A から 1km 圏内の α 」

定義 2: 最小の空間演算 SQ の問合せを集合演算子の組合せ「 $SQ^1 \cup SQ^2 \cup \dots = SQ^+$ 」で記述できる。

(例) 「 $(A_{\perp}80m_{\perp}\alpha) + (B_{\perp}60m_{\perp}\beta)$ 」は「A から 80m

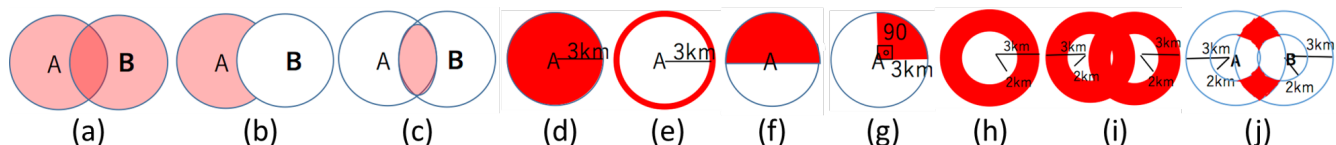


図2 (a) 和 (b) 差 (c) 積 (d) 周辺 (e) 周囲 (f) 方向 (北) (g) 角度 (90) (h) 範囲 (i) 和と範囲 (j) 積と範囲

圏内の α と B から 60m 圏内の β の和集合

3.2 集合演算子

上記の定義に基づき、待ち合わせスポットや遭遇率を低下させるスポット検索を集合演算の和「+」、差「-」、積「*」より実現する (図 2)。

和演算 (例 1) 「 $(A_{\perp 3km} \cup \alpha) + (B_{\perp 3km} \cup \alpha)$ 」

A と B の 3km 圏内の和演算 (図 2 (a))

差演算 (例 2) 「 $(A_{\perp 3km} \cup \alpha) - (B_{\perp 3km} \cup \alpha)$ 」

A と B の 3km 圏内の差演算 (図 2 (b))

積演算 (例 3) 「 $(A_{\perp 3km} \cup \alpha) * (B_{\perp 3km} \cup \alpha)$ 」

A と B から 3km 圏内の積演算 (図 2 (c))

3.3 人物からの特徴語算出

人物の特徴抽出は、Wikipedia やニュースなどのオープンデータからコーパスを構築し、極性となるネガティブワードとポジティブワードを機械学習より特徴語として抽出する。地名と特徴語をポジティブな場合は和集合演算、ネガティブは差集合演算式に変換し、空間演算よりスポットを抽出する。以下に処理の流れを示す。

- (1) 人物名 q から wikipedia のドキュメント D を取得
- (2) ドキュメント D からコーパス構築
- (3) 人物名 q と関連単語を特徴語として抽出
- (4) 特徴語の極性値を先行研究の感情辞書より抽出

3.4 空間演算処理

空間演算式は先行研究 [1] の Web リクエスト¹によって空間演算処理される。これにより、各起点から空間長の指定範囲内にあるスポット情報 (位置や店舗名、評判など) のリストを得ることができる。最後に、取得したリストから集合演算を行う。

3.5 Web 入出力処理と空間演算式変換

ユーザの入力は起点の地点名と人物名とする。前節より抽出した人物名から特徴語と極性を抽出し、地点名と特徴語、極性に基づいて空間演算式に変換する。出力結果は、地図に地点を起点とする半径 d の円が描かれ、検索されたスポットがプロット推薦される。

3.6 複数地点間のスポット推薦システム

図 3 にスポット推薦結果の例を示す。「空間キーワード」人名として「ハリウッド」大谷翔平」を入力すると、大谷翔平に関して極性をもつ特徴語が抽出され、それらをポジティブ極

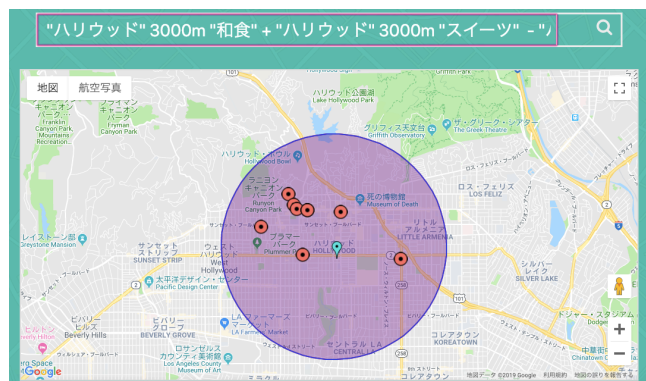


図3 スポット検索結果の一例

性の場合は和演算、ネガティブ極性の場合は差演算として演算式を作成する。作成された演算式は図 3 の上部に表示される。なお、例ではハリウッドを中心とする 3km 以内のスポットとしている。

スポット検索結果は、地図上にプロットされる。プロットされたマーカーをクリックするとスポット情報を確認できる。

4 まとめ

本論文では、動的なオブジェクトとして代表的な人物と関連する地図空間上のオブジェクトを発見することで、空間キーワードと人名による地図空間におけるスポット検索手法を提案した。スポット検索により、人物に関連するスポットを地図の実空間上で検索できることを確認した。今後、提案手法によるプロトタイプを構築し、空間キーワードと人名による地図検索より提示されたスポットに対して、定量的評価および定性的評価を行う。また、スポット検索の利活用についても検証する。

謝辞

本研究の一部は、JSPS 科研費 16H01722, 17K12686, 17H01822 の助成を受けたものである。ここに記して謝意を表す。

文献

- [1] 阪田晴香, Panote Siriraya, 王元元, 河合由起子. 検索ワード間の空間演算の提案と地図検索への応用. 情報処理学会論文誌データベース (IPSJ-TOD), Vol. 12, No. 1, pp. 1-5, 2019.
- [2] 奥蘭基, 牟田将史, 平野廣美, 益子宗, 星野准一. 複数人での旅行における嗜好分析による観光地推薦システムの提案. 研究報告ヒューマンコンピュータインタラクション (HCI), Vol. 2015, No. 19, pp. 1-8, 2015.
- [3] 笠原秀一, 田村和範, 飯山将晃, 棕木雅之, 美濃導彦. 行動履歴に

¹: <http://yklab.kyoto-su.ac.jp/~sakata/spatialQuery/>

基づく地域の環境要因を考慮した観光行動モデルの構築とその応用. 情報処理学会論文誌, Vol. 57, No. 5, pp. 1411-1420, 2016.

- [4] 落合桂一, 深澤佑介, 山田渉, 松尾豊. 位置情報サービスの利用状況を活用した poi 推薦手法. 人工知能学会全国大会論文集, Vol. 2017, p. 2, 2017.
- [5] 大東祐太, 有山俊一郎, 延原肇. 位置情報 sns 上の画像-テキスト間対応を利用したユーザ嗜好抽出と推薦スポット候補拡張. 情報処理学会論文誌, Vol. 58, No. 12, pp. 2006-2014, 2017.
- [6] 今井良太, 数原良彦, 戸田浩之, 鷺崎誠司. Web 文書を利用した poi に関連する地名の抽出. 情報アクセスシンポジウム 2013, Vol. 2012, pp. 9-13, 2013.
- [7] K. Fu, Y. C. Lu, C. T. Lu. *Treads: A safe Route Recommender using Social Media Mining and Text Summarization*. In *Proc. of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2014)*. pages 557-560, 2014.
- [8] J. Kim, M. Cha, T. Sandholm. *SocRoutes: Safe Routes Based on Tweet Sentiments*. In *Proc. of the 23rd International Conference on World Wide Web (WWW 2014)*. pages 179-182, 2014.
- [9] D. Quercia, R. Schifanella, L. M. Aiello. *The Shortest Path to Happiness: Recommending Beautiful, Quiet, and Happy Routes in the City*. In *Proc. of the 25th ACM conference on Hypertext and social media (HT 2014)*, pages 116-125. 2014. DOI=<http://dx.doi.org/10.1145/2631775.2631799>.
- [10] M. Duckham, L. Kulik. "Simplest" Paths: Automated Route Selection for Navigation. In *COSIT*, Vol. 2003. pages 169-185, 2003.
- [11] K. Church, J. Neumann, M. Cherubini, N. Oliver. *The "Map Trap"? An Evaluation of Map Versus Text-based Interfaces for Location-based Mobile Search Services*. In *Proc. of the 19th international conference on World wide Web (WWW 2010)*. pages 261-270, 2010.
- [12] ノイズを抑えた文献検索を行うには? ScienceDirect & Scopus での“近接演算子”の活用. エルゼビア・ニュースレター:2015 年 8 月 19 日号, 2015.
- [13] Oracle Text リファレンス, 11g リリース 2(11.2). B61357-06. 2015 年 10 月.
- [14] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, Jeff Dean. *Distributed Representations of Words and Phrases and their Compositionality*. In *Proc. of the 26th International Conference on Neural Information Processing Systems (NIPS '13)*. pages 3111-3119, 2013.