

# ランドマーク着目傾向を考慮したルートマップ型 地図検索インタフェースの検討

田中 智大<sup>\*1</sup> 北村 尊義<sup>\*2</sup> 泉 朋子<sup>\*3</sup> 仲谷 善雄<sup>\*2</sup>

## Examination of Route map type map search interface considering landmark attention trend

Chihiro Tanaka<sup>\*1</sup>, Takayosi Kitamura<sup>\*2</sup>, Tomoko Izumi<sup>\*3</sup> and Yosio Nakatani<sup>\*2</sup>

**Abstract** - The format used for the location identification system is mainly based on survey map type space recognition such as a bird's eye view. However, those who can not use the map well can not understand their position with route map type spatial cognition and it is difficult to convert to survey map type spatial cognition. The purpose of this research is to develop a system that reduces the cognitive load of people who are not good at survey map type spatial cognition. In this method, the tendency of the landmark that the user focus on is clarified, and a search interface based on the result is studied.

**Keywords:** landmark, map search interface

### 1. はじめに

位置情報の把握の手法としてこれまでは紙媒体の地図が主に用いられてきた。今日、パソコンやスマートフォンなどの情報機器の登場により地理情報システム (Geographic Information Systems : GIS) が急速に発展・普及を遂げている。GIS は地理的位置を手掛りとして、位置や空間に関する情報を持ったデータを重ね合わせ、視覚的 (2次元・3次元) に表示するシステムである。これまでの紙媒体での地図は特性上 1 ページに 1 エリアしか情報を表示できないが、GIS は画面上でスクロールを行うことで、ページの遷移なくして任意のエリアを 1 画面にて表示することが可能であるため、普及が進んでいる。

近年の地図利用で特徴的な機能は、現在位置情報に基づく目的地までのルート探索機能である。この機能を可能にしている技術が GPS (Global Positioning System : 全地球測位システム) である。これによって現在位置と目的地の位置関係の把握、最適ルートの検索や所要時間の予測、現在地周辺の地物の検索、道路状況などの現在地付近のリアルタイムな情報検索が可能となっている。

しかし、地図検索システムの発達によって地図を見る者(以下、ユーザ)はより幅広い情報を得ることが可能となったものの、位置情報を把握するという点においては紙媒体の地図と同様に、ユーザは地図と実際に見えている風景とを照らし合わせる作業が必要である。また、現在

の GPS による位置情報は数メートルから数十メートルの誤差を生じる場合があり、電子コンパスによるユーザが向いている方向を特定する機能も高頻度で乱れを生じる。そのため、ユーザの空間認知能力によっては、地図が正確に読めない可能性がある。

本研究では、ユーザが目前に見る風景中のランドマークを検索キーとして音声で入力する地図検索方式において、ユーザ側の検索の認知的負荷の軽減につながる入力インタフェースとしてどのような形式が有効であるのか調査することを目的とする。

### 2. 研究動向

#### 2.1 人間の空間認知に関する研究

Shemyakin[1]は、人の空間認知方法を「サーベイマップ型認知」と「ルートマップ型認知」の 2 つに大別している。サーベイマップ型認知とは、一般的に市販されている地図のような、地形や街並みを上空から平面的にとらえた空間認知のことであり、自宅と学校などといった建物間の相互関係と直線距離を理解しやすいという特徴が指摘されている。一方、ルートマップ型認知とは、あるルートを頭の中で移動する際に浮かぶ特徴的な地物などの系列から構成されている空間認知のことである。近年の一般的な位置情報検索システムに用いられている形式は鳥瞰図などのサーベイマップ型空間認知に基づいて設計されているが、地図をうまく利用できない人はルートマップ型の空間認知で自身の位置を把握しており、サーベイマップ型の空間認知への変換が困難であるとされている。

新垣ら[2][3]は、方向音痴の原因として考えられる要因を紹介している。

① 空間能力説 : 空間認識能力が低いと、街を移動する際に建物同士の位置関係のイメージ操作が難しく、

\*1: 立命館大学大学院 情報理工学研究科

\*2: 立命館大学 情報理工学部 情報コミュニケーション学科

\*3: 大阪工業大学 情報科学部 情報システム学科

\*1: Graduate School of Information Science and Engineering, Ritsumeikan University

\*2: College of Information Science and Engineering, Ritsumeikan University

\*3: College of Information Science and Technology, Osaka Institute of technology

自分の位置や進むべき方向がわからなくなる。

- ② 用いる知識の違い説：道順を考える際に、ルートマップ型で考える場合の方が、サーベイマップ型で考える場合に比べて道に迷う可能性が高い。
- ③ 問題解決方略説：道順を学習する際に、走行車や歩行者といった固定的ではない物に注意が行ってしまう傾向の強い人は、建物や道路標識などの固定的なランドマークに注意する人より道に迷う可能性が高い。
- ④ 社会的ラベリング説：「自分は方向音痴である」といったラベル付けを自らに行う結果、他人に道案内などを頼ることになり、空間認知能力が低下する。

これらの要因が絡み合って、方向音痴になると述べられている。これらの研究成果を踏まえると、道に迷わないためには、ルートマップ型空間認知においても利用されているランドマークを有効に利用するなどにより、サーベイマップ型の空間のイメージを構築しやすい状況を工夫して作り出すことが望ましいと言える。

## 2.2 都市のイメージ

Kevin Lynch[4]は都市のパブリックイメージに影響するイメージアビリティのうち、物理的な知覚できる物体についてパス、エッジ、ディストリクト、ノード、ランドマークの5つのエレメントに分類している。以下にそれぞれの概要を示す。

1. パス(道路)： 観測者が通りうる道筋
2. エッジ(縁)： 2つの局面の境界であり、パスでないもの
3. ディストリクト(地域)： 中から大の大きさを持つ都市の部分で2次元の広がりをもつもの
4. ノード(接合点)： 都市内部にある主要な地点。
5. ランドマーク(目印)： 建物、看板、山などの単純な物理的要素から成り立っているもの

実際に道に立って風景を見る、あるいは地図を読む際、主に着目するものは、パス、ノード、ランドマークなどが考えられる。ノードとはおおよそ2種類存在し、一つは接合によるノードで例としては交差点や鉄道の駅などが挙げられる。もう一つは、特徴の集中によるノードで例としては広場や公園などが挙げられる。

さらにこれらのノードの周辺にはひときわ目立ち覚えられやすい何らかの特徴を持ったもの（ランドマーク）が存在している可能性が高い。ランドマークの例としては建物が挙げられる。また、ランドマークの性質を強める要因として、配置、築年齢、規模などの対比、色、音、匂いなどが挙げられる。これらの要素を複数満たしている場合は関係性が生まれ性質が強化される。

本研究では、見たままの風景から着目した特徴物（パスあるいはランドマークあるいはノード）を選択し、検索キーとして検索を行うシステムを目指している。しか

しノードやその周辺に存在するランドマークやパスは多種多様であり、入力に時間を要する可能性が考えられる。そのため、本研究ではまず初めに風景の表現方法について実験を行いエレメントの着目傾向について考察を行う。

## 2.3 地理情報システムのための地図描画高速化と地図検索に関する研究

坂入[5]による「地理情報システムのための地図描画高速化と地図検索に関する研究」は、GISの高度化を図ることを目的としている。地図検索に焦点を当て、空間認知の特性を考慮した地図検索の実現を目指している。坂入の提案するLSAS (Location Search by Actual Scenery)は風景中あるいは記憶中のランドマークの名称や種類と、視野内での位置関係を検索キーとする地図検索手法である。入力方法は、図1に示すように、道路沿いに位置するユーザの現在視点からみて右、左にある建物をそれぞれ3つまで入力する。

図1 LSAS アプリケーションの入力画面

LSASはルートマップ的視点そのままを入力するため、ルートマップ的視点からサーベイマップ的視点への変換が不要である。しかし、左右3つランドマークを入力するといった入力インタフェースについては試作段階であり、具体的な検討がなされていないため、前節にて述べたエレメントの着目傾向を明らかにする必要がある。

## 3. 風景表現方法の調査と入力インタフェースの設計

### 3.1 調査の方法

本研究では、2章にて述べたエレメントの着目傾向を検討するにあたり、見ている風景を自由発話形式でどのように表現するのかを調査する。調査のために用いる場所は大阪駅周辺(駅前第1ビル周辺、第4ビル御堂筋付近)、新大阪駅周辺の計3か所とする。調査対象者は大学生10名(第4ビル周辺のみ、残り2か所は4名ずつ)とする。調査場所(第4ビル)の風景を図2に示す。調査対象者の目の前には道路(パス)や交差点(ノード)、SMBC日興証券、KINSANなどの看板、JTBやみずほ銀行等の窓ガラスなどに企業名が書かれたビル、青の世界時計、黄色の特設宝くじ売り場などの特徴的な地物といった様々なラ

ランドマークが存在する。



図2 大阪駅近くの第4ビル周辺の風景

調査対象者はひとりずつ図2の風景が見える場所に立つ。入力の際、調査対象者には「待ち合わせなどで別の場所にいる誰かに自分のいる場所を伝える状況を想定して、注視した物やそれらの位置関係を自由に表現してほしい。」と依頼する。また、入力するランドマークの条件として、歩行者や車などの固定的でない物についての入力は禁止した。なお、調査中の発話についてはすべてボイスレコーダーで録音を行い、調査後に録音したデータの逐語録を作成して分析する。

### 3.2 調査の結果

1人あたりの調査に要した時間は、被験者によって多少のばらつきはあったが、最短で42秒、最長で94秒であった。被験者10名の平均は59.7秒であった。

逐語録を分析した結果、入力されたエレメントの大半はランドマークであった。ノードについての入力が見られたのは被験者10名中1名、パスについては10名中5名である。ディストリクトやエッジに該当する入力は得られなかった。また、ランドマークを入力する順番については以下の3種類に分類された。

1. 一番近くに見えるランドマークを最初に入力し、そこから横に視点をずらし順番に入力を行う（例：みずほ銀行、世界時計、OSビル、JTB、SMBC日興証券、KINSAN）
2. 一番近くに見えるランドマークを最初に入力し、そこから横に視点をずらししていくが、途中から位置関係がランダムになる（例：第四ビル、世界時計、SMBC日興証券、KINSAN、OSホテル）
3. 完全にランダムな順番に入力を行う（例：世界時計、SMBC日興証券、KINSAN、みずほ銀行（梅田第四ビル））

人数分布に関しては、1の傾向が見られた被験者が6名、2が2名、3が2名であった。結果の分析から明らかになった事項を以下に整理する。

- a) 3の傾向が見られた被験者の2名はともに、被験

者から見て最も近いランドマーク（世界時計、みずほ銀行）に最初に着目していた。つまり、すべての被験者が最も近いランドマーク（みずほ銀行、第四ビル）に最初に着目するという傾向が見られた。

- b) 初めに入力したランドマークの位置から少しずつ横に視点を移動させながら、見えるランドマークをそれぞれ入力するという傾向が非常に強いと言える。調査では最も近いランドマークはみずほ銀行、第4ビル、世界時計であり、被験者から見て右側に存在していたため、右から左に視点を少しずつずらしながら入力された。最も近いランドマークが左側にある場合は左から右に順に入力されると推測されるが、別途検証が必要であると考えられる。
- c) 基本的にはb)に述べたように右から左の方向に目立つランドマークを述べているが、ランドマークの周辺に存在する目立ちにくい地物について追加情報として述べる場合に、順番が左右すると言える（例「〇〇の隣に□□や△△もあります」）。
- d) 入力されたランドマークの種類に関しては、文字が見える看板が多数を占めており、その他には道路や、特徴的な色、大きさ、形をした建造物が挙げられた。ただ、注視すべきランドマークが定まらず、どのような地物をランドマークとするべきか、またその属性やランドマーク間の位置関係をどこまで表現するべきかに悩む被験者が見られた。

### 3.3 入力インタフェースの設計

まず、入力時間に関し、調査では入力の順番や内容に制約を設けない自由発話形式にて行なったため、注視するランドマークが定まらず、また、それらの位置関係をどこまで表現するのかに悩む様子が被験者のうちの数人に見られた。そのため、入力に要する時間が42秒から94秒まで幅が広く、分散が大きかった。調査後のインタビューでは、入力に要した時間が長く、煩わしかったという意見が少なからず出た。

以上のことから、提案するシステムでは対話形式にて質問を行うこととする。対話形式にすることで、入力の内容や順番を誘導できるため、自然発話形式に比べてより円滑な入力が可能になると期待される。

作成した質問の基本的なフローは以下の通りである。

質問1：近くに何が見えますか？

質問2：それは左右どちらに見えますか？

質問3：（最初に入力したランドマーク）から見て（右・左）に見える物を順番に入力してください

質問4：ほかに特徴的な物がありますか？

以上の質問手順に関するフローチャートを図3に示す。

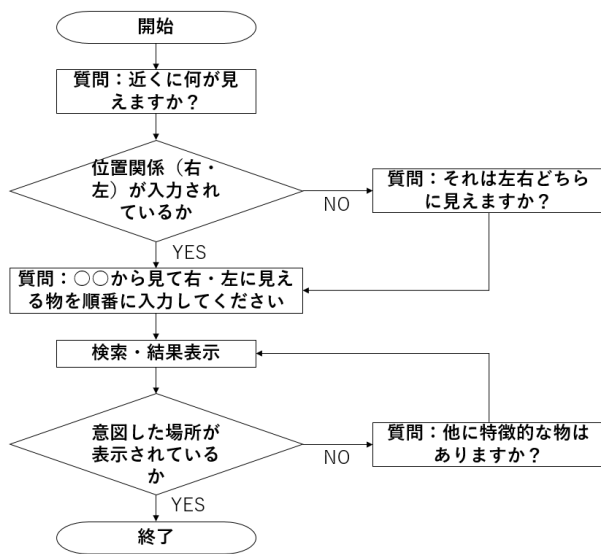


図 3：質問フローチャート

## 4. システムの試作

### 4.1 システムの概要

3 章にて設計を行った質問の手順に基づいて、システムを構築した。本システムは開発言語に PHP, JavaScript を使用し、Web アプリケーションとして実装した。音声処理のため、Google 社が提供する Web Speech API を使用しており、対応ブラウザの都合上、Google Chrome にてシステムを動作させている。プログラムの実行端末には Galaxy S7 edge を使用した。

### 4.2 システム利用の流れ

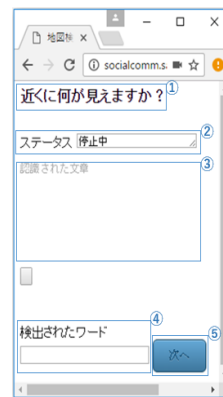
本章をまとめると、ユーザは以下のような流れでシステムを利用する。

- ① ユーザはシステムを起動する。
- ② 開始ボタンをタップする。
- ③ システムが提示する質問に対し、ユーザは見えている風景や建物をそれぞれ入力する。

### 4.3 システムの機能

本システムは、ユーザによって入力された自然言語を、音声認識によって文章として記憶する。次に、認識された文章に形態素解析を行うことで、品詞ごとに単語を分類し、名詞のみを抽出する。ユーザがシステムに「右にマクドナルドが見えます。」と入力した場合、システムが認識を行い、「右にマクドナルドが見えます」という文字列が配列に格納される。次にこの文字列に形態素解析を行い、「右 マクドナルド」のような、名詞のみが抽出された文字列を生成し、変数として保存する。

システム画面は図 4 のとおりである



- ① 質問文
- ② 音声認識ステータス
- ③ 認識された文章
- ④ 検出されたワード
- ⑤ 次へ進むボタン

図 4：システムメイン画面

## 5. 試作システムの評価

### 5.1 実験の目的

本研究の目標は、自由発話形式による音声入力への傾向を明らかにし、それに基づいた対話形式の質問を行うことで、ユーザの入力の負担を削減することである。そのため、評価実験では 3 章にて説明した自由発話形式と提案する対話形式の比較を行い、使用感や入力時間の評価をするとともに、上記した目的に本システムが有効かどうかの検証を行った。

### 5.2 実験場所

本研究の評価実験は滋賀県草津市の JR 南草津駅付近の国道一号線沿いにて行った。実際に実験場所にて撮影した風景の画像を図 5 に示す。被験者は図 5 の風景が見える場所に立ち音声入力を行う。

実験場所から見て、左手にマクドナルドが見える。その奥には極楽湯の看板である「ゆ」の文字が確認できる。画像では確認しづらいが、さらに奥にはラーメン屋、ハンバーグレストラン Big Boy やアミューズメント施設の BB などが見える。道路の向かい側にはファミリーマート、ラーメン店にぼ次郎、SEIYU の看板、肉バル&ダイニングの看板などが存在する。なお、これらの看板はライトアップされているため、どんな時間帯であっても、看板の内容を視認することが可能である。



図 5：実験場所から見える風景

### 5.3 実験の手順

自由発話形式と提案する対話形式の比較を行うため、

次の2種類の実験をそれぞれ実施する。

実験1：自由発話形式による入力実験+提案システムの評価

実験2：対話形式による入力実験+提案システムの評価

被験者は、実験1、実験2、それぞれに男性18名、女性2名の20名ずつとする。年齢の分布は、実験1では19歳以下：2名、20~24歳：15名、25歳~29歳：2名、50~54歳：1名である。実験2については19歳以下：2名、20~24歳：16名、35~39歳：2名である。以下に、実験1の手順を示す。

- ・ 実験の目的、概要の説明
- ・ 被験者は前節にて説明した実験場所に立ち、スマートフォンのボイスレコーダーを準備
- ・ 被験者は入力開始の際、ボイスレコーダーを起動させ、録音を開始
- ・ 見えている風景を自由発話形式で表現してもらう。
- ・ ランドマークの入力が終了した場合、被験者はボイスレコーダーの録音を停止
- ・ 本研究で作成した対話形式のシステムの説明
- ・ 対話形式のシステムの利用
- ・ アンケート記入、インタビューの実施

以下に実験2の手順を以下に示す。

- ・ 実験の目的、概要の説明
- ・ 被験者は前節にて説明した実験場所に立ち、スマートフォンのボイスレコーダーを準備
- ・ 実験者は、対話形式の質問開始の際、ボイスレコーダーを起動させ、録音を開始する。
- ・ 被験者は、見えている風景の中から、注視したランドマークを入力
- ・ 作成したシステムを被験者に使用してもらう。
- ・ アンケート記入、インタビューを行う。

なお、実験1、実験2におけるランドマークの入力について、3章にて実施した調査と同様に、車やバイクなどの固定されていない地物ものについての入力は禁止する。

各実験の最後に行ったアンケート項目を表6に示す。項目番号6~11は音声入力の形式の評価に関する項目である。項目番号12~15は構築したシステムの評価に関する項目である。なお、項目番号8と9は対話形式の質問に関する項目であるため、実験1では削除している。これらの質問の内、項目番号5~9,12は5段階評価で、数値が大きいほど肯定的な回答を示す。項目番号10,14は5段階評価で、数値が小さいほど肯定的な回答を示す。また、項目番号11,15~17は自由記述方式とする。

## 5.4 実験結果

実験1および2の結果を以下にまとめる。

表6：実験アンケート

1. あなたの性別を教えてください
2. あなたの年齢はどれに該当しますか？
3. あなたは普段どれくらいの頻度で Siri や Google 音声検索などの音声認識システムを使用しますか？
4. あなたはこの場所に訪れる機会がよくありますか？
5. 「正面に銀行が見える」など、あなたは待ち合わせなどで自分がいる場所を別の場所にいる誰かに伝えることは得意ですか？
6. (入力形式について) 質問に対する回答はしやすかったですか？
7. あなたが質問の回答に要した時間はどう感じましたか？
8. 質問の流れ、順序は自然でしたか？ (実験2のみ)
9. 質問の数についてどう感じましたか？ (実験2のみ)
10. この入力形式に対しストレスを感じましたか？
11. ストレスを感じた方はその理由をお聞かせください
12. (システムについて) 質問に対する回答はしやすかったですか？
13. 正常に音声認識は作動しましたか？
14. システムでの入力実験に対しストレスを感じましたか？
15. ストレスを感じた方はその理由をお聞かせください
16. 本システムに今後期待することは何ですか？
17. 今回の実験に関し、感想があればよろしく願います

- 入力時間の範囲について、実験1では12秒~125秒、実験2では11秒から78秒であった。入力時間の平均は実験1が44.8秒、実験2が29.45秒（質問に要した時間を含めると38.45秒）であった。また、入力されたランドマーク(RM)の数の平均では、実験1が6.3個、実験2が5.8個であった。これより、実験1に比べて実験2の方が、入力時間および入力するランドマークの数について、より少ない結果となった。
- ランドマークを1つ入力するのに要する時間は、実験1で約7.11秒、実験2で約5.16秒（質問時間を含めた場合、6.74秒）であり、実験2のほうが実験1より効率的にランドマークを入力していた。
- 実験1では、入力時間に60秒以上を要した実験協力者の数は5名であり、そのうち1名は、120秒以上要した。実験2では、質問時間を含めた場合でも、入力に60秒以上を要した実験協力者は1名のみであった。これより、実験2の方が入力時間にバラつきが少なかった。

対話形式による入力方式の有効性を統計的に示すため、実験 1、実験 2 における入力時間、ランドマークの数、ランドマーク 1 つを入力するのに要する時間について、それぞれの平均値に差があるのかを検定したところ、下記の結果が得られた。

#### ■入力時間について

独立 2 群の差の検定であるため、まず 2 群の分散が等しいことを帰無仮説として、F 検定を行った。有意水準 5%で行ったところ、P 値は 0.0094 となった。P<0.05 であるため、帰無仮説は棄却され、等分散でないと言える。次に分散が等しくないとして仮定した 2 標本による検定 (Welch 法) を行ったところ、t 値は 2.2 となり、P 値は両側で 0.029 となり、このことから、実験 1 と実験 2 とでは入力時間に統計的な有意差があると言える。

#### ■ランドマークひとつあたりの入力時間について

上記の入力時間の分析と同様に F 検定を行ったところ、P 値は 0.017 となり、等分散ではないため、分散が等しいと仮定した 2 標本による検定 (Welch 法) を行った。t 値は 2.1 となり、P 値は両側で 0.039 となった。この結果、実験 1 と実験 2 ではランドマークひとつあたりの入力時間に有意な差が見られた。

以上のことから、実験 2 で行った対話形式による質問は、実験 1 で行った自由発話形式による質問より優れている可能性が考えられる。

### 5.5 アンケート結果

アンケートに設けた質問項目のうち、音声入力方式の評価に関する項目は 6「質問に対する回答はしやすかったですか?」、7「あなたが質問の回答に要した時間はどの感じましたか?」、8「この入力形式に対しストレスを感じましたか?」3 項目である。これらについても平均値について F 検定と t 検定 (Welch 法) を行った。検定の結果をまとめたものを表 7 に示す。

質問 6 について P<0.05 となるため、実験 1 と実験 2 では質問に対する回答のしやすさに有意な差が見られたといえる。質問 7、質問 8 については P>0.1 であり、有意な差は見られなかった。

表 7: アンケート項目に関する平均値の検定の結果

	6	7	8
F検定P値	0.26	0.49	0.026
等分散の判定	等分散	等分散	等分散でない
t	2.1	-0.5	1.2
P(T<=t)両側	0.04	0.62	0.23

以上の結果から、本研究の実験では提案手法は自由発話形式と比べて、入力効率(入力時間、ランドマーク当りの入力時間)と入力のしやすさにおいて優れている可能性がある。

## 6. まとめ

本研究では、ユーザが目前に見る風景中のランドマークを検索キーとして音声で入力する地図検索方式において、ユーザ側の検索の認知的負荷の軽減につながる入力インタフェースとしてどのような形式が有効であるのか調査した。そのために、まず風景表現方法および入力インタフェースについて検討した。次に、その検討結果をもとに自由発話形式と対話発話形式の 2 つを比較するシステムを構築し、評価した。その結果、入力の効率性や容易さの面で、対話発話による入力方式が優れている可能性を確認した。

## 謝辞

本研究は JSPS 科研費 JP16K21484 の助成を受けたものです。

## 参考文献

- [1] Shemyakin, F.N: "Orientation in space.", Psychological Science in the USSR 1, pp.186-255 (1962)松村真宏: 仕掛学への誘い, 人工知能学会誌, Vol.28, No.4, p.583, 2013.
- [2] 新垣紀子, 野島久雄: 方向音痴の科学, 講談社 (2001)
- [3] 日本心理学会:心理学Q&A 心理学ふしぎふしぎ (新垣紀子回答)  
<http://www.psych.or.jp/interest/ff-08.html> (参照 2016/8/25)
- [4] Kevin Lynch,丹下健三,富田玲子(訳):都市のイメージ, 岩波書店(2007)
- [5] 坂入威郎:地理情報システムのための地図描画高速化と地図検索に関する研究, 立命館大学博士論文 (2016)