

## 潜在的意味解析を用いた車内発話の話題推定に基づく目的地提案

矢澤 大心<sup>†</sup> 西崎 博光<sup>†</sup> 関口 芳廣<sup>†</sup> 北岡 教英<sup>††</sup><sup>†</sup> 山梨大学大学院医学工学総合教育部 <sup>††</sup> 名古屋大学大学院情報科学研究科

## 1 はじめに

近年の音声認識技術の進歩により、音声認識のインタフェースを搭載したカーナビが普及している。音声入力にはハンズフリーで操作ができる点などからカーナビには有用な機能であるが、あまり利用されていないのが現状である。その理由として、誤認識の多さや機械に話しかける心理的負担が考えられる。

車内音声認識に関する研究は多く行われている。河口ら [1] は実走行車内における音声や対話を収録した車内音声データベースを作成し、車内対話における特徴の分析を行っている。また、加藤ら [2] の対話構造木コーパスを用いた手法なども提案されている。これらはユーザの発話を正しく理解するための研究であるが、いずれも目的地を設定する発話やエアコンの操作など、車載情報機器としてのカーナビを操作するためのドメイン発話を主な対象にしている。

しかし、ドメイン発話以外の発話も正しく音声認識が行えるのであれば、ユーザがどのようなことに興味を持っているかなどが推定できるはずである。雑談などの車内発話からユーザが興味を持ちそうな物件が自動的に紹介できれば機械に話しかけるストレスを軽減できると考えられる。

そこで本研究では、ドメイン発話以外のユーザの発話に焦点を当て、ユーザが求めているような目的地提案を行う手法について検討した。物件の紹介記事を潜在的意味解析 [3] (以下、LSA) を用いクラスタリングし、ユーザ発話との類似度を求めることで目的地提示を行った。

本稿では、提案手法でドライバーの発話から目的地提示を行い、アンケートにより評価したことを報告する。ドライバー発話の書き起こしを利用した予備実験において 69 % の物件を正しくユーザに提示し、音声入力による実験では正解率 49 % という結果を得た。

## 2 LSA を用いた話題推定手法

## 2.1 LSA を用いた話題の定義方法

車内でリアルタイムに進行する会話がどのような話題であるかを判断するには、まずその話題というものをどのように定義するかが重要になる。Web サービスや観光ガイドなどは、物件毎に手動でタグを付けてジャンル分けされていることが多い。しかし、手動でタグ

を付けるためには多大な手間がかかる。また、手動で分類するために定義した話題というものがユーザの感性和合致しないことも考えられる。

そこで、本研究ではこの話題を定義する際に LSA によるクラスタリングを用いた。LSA には文書×単語行列が必要になる。本研究では「マップルガイドデータ」(以下、MGD) を用いて文書×単語行列を作成した。

具体的には、まずクラスタリングしたい物件それぞれを 1 文書として扱い、各物件の紹介記事を形態素解析し単語出現頻度を求めることで文書×単語行列を作成した。この際、普通名詞、形容詞、形容動詞のみを単語として用い、助詞などのクラスタリングのノイズとなる単語は除去した。文書×単語行列を作成した後、そこから特異値分解を行い、その後次元の縮約を行う。本研究では 200 次元に次元縮約を行った。その後、K-means クラスタリングにより 200 クラスタに分類した。このクラスタ 1 つずつを話題と定義して扱った。

## 2.2 発話の話題分類

ユーザの発話が、どの話題に当てはまるかにはコサイン尺度による類似度計算を用いた。MGD の 7836 件それぞれの紹介記事とユーザの発話の間でコサイン類似度を求める。その後、コサイン類似度上位 3 位までの物件がそれぞれどのクラスタに属しているかで話題分類を行う。例えば、上位 3 件の属するクラスタがそれぞれ *cluster\_A*, *cluster\_B*, *cluster\_C* であれば、その 3 つのクラスタのいずれかのクラスタに属する物件から目的地提示を行う。

## 3 評価実験

提案手法により適切な目的地提示が可能であるかは、アンケート実験によって評価した。

## 3.1 アンケート内容

評価を行うために実施したアンケートの内容について説明する。アンケート回答者は、ある運転者と車に同乗しているという設定で、

1. 車内で運転手が行った発話
2. その発話をもとに MGD の物件情報から提示された物件の紹介記事

を読む。2. は何件か提示されている。そして、この物件の中から運転者の発話を考慮して、目的地にするのが適当な物件にチェックを付ける。チェックを付ける基準は直感や個人の嗜好で判断してもらった。チェックが付いた物件を正解として扱った。そして、提示された物件のうちチェックされた物件の割合を正解率として評価した。

## 3.2 予備実験

まずは、LSA の有効性を確認するために、誤認識を含まないドライバー発話の書き起こしを入力に用いて

Destination Proposal Based on Topic Estimation from Conversation in Car Using Latent Semantic Analysis.

<sup>†</sup> Daishin YAZAWA (daicn@alps.cs.yamanashi.ac.jp)

<sup>†</sup> Hiromitsu NISHIZAKI (hnishi@yamanashi.ac.jp)

<sup>†</sup> Yoshihiro SEKIGUCHI (sekiguti@yamanashi.ac.jp)

<sup>††</sup> Norihide KITAOKA (kitaoka@nagoya-u.jp)

Department of Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi (<sup>†</sup>)

Graduate School of Information Science, Nagoya University

(<sup>††</sup>)

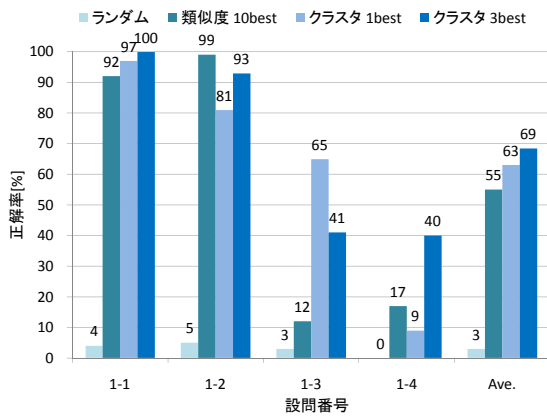


図 1: 書き起こしデータによる目的地提示実験の正解率

話題分類を行う予備実験を行った。ドライバー発話には CIAIR-HCC の書き起こしから 4 パターンの車内対話を用い、設問を作成した。

提案手法とその性能比較対象を以下に示す。以下の 4 種類の条件で被験者に目的地提示を行った。

**ランダム** MGD 全 7836 物件から無作為に抽出した 10 件。

**類似度 10best** ドライバー発話と MGD 全物件紹介記事のそれぞれをコサイン尺度で類似度計算し、類似度が高かった物件上位 10 件。

**クラスタ 1best** LSA を用いた提案手法。ドライバー発話と MGD 全物件紹介記事のそれぞれをコサイン尺度で類似度計算し、最も類似度が高かった物件が属するクラスタから無作為に抽出した物件 10 件。

**クラスタ 3best** LSA を用いた提案手法。ドライバー発話と MGD 全物件紹介記事のそれぞれをコサイン尺度で類似度計算し、類似度が高かった物件上位 3 件がそれぞれ属するクラスタから無作為に抽出した物件 10 件。

このうち、ランダムとクラスタ 1best の実験では 14 名からアンケートの回答を得た。また、類似度 10best とクラスタ 3best では 10 名から回答を得た。

実験結果を図 1 に示す。コサイン類似度で高かった物件を提示した場合には、ドライバー発話によっては非常に高い正解率を示した。しかし、曖昧な表現の発話に弱く、極端に正解率が落ちる場合がある。

一方で提案手法のクラスタ 3best を用いた場合には、ドライバー発話に近い物件と似たような特徴を持つ物件群から目的地を提示できた。そのため、最低でも 40 % 程度の正解率を示すことができ、平均で 69 % の正解率となった。

### 3.3 音声認識結果を利用した実験

入力にドライバー発話を音声認識した結果を用いて提案手法の評価を行った。音声認識には、音声認識エンジンに Julius4.1.5 を使用した。言語モデルは、CIAIR-HCC から学習した語彙サイズ 20k の trigram を用い、音響モデルは日本語話し言葉コーパスの全講演から学習した triphone モデルを使用した。なお、音声認識率は、単語正解率が 72 %、単語正解精度が 52 % である。目的地候補の提示には、3.2 節のクラスタ 3best と同様

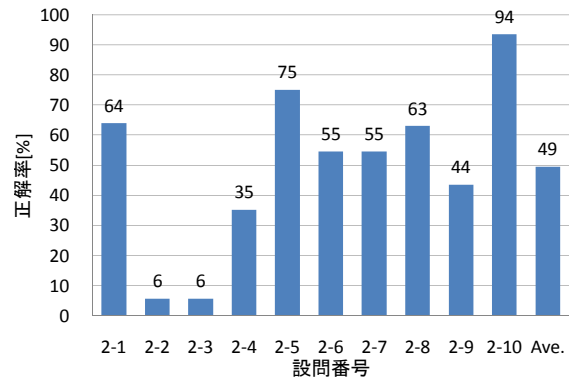


図 2: 音声認識結果を用いた目的地提示実験の正解率

の手法を用いた。アンケートは予備実験とは異なる 10 個の設問を設け、12 名から回答を得た。

実験結果を図 2 に示す。94 % という高い正解率を示す設問も存在したが、10 設問の平均では 49 % という予備実験を 20 % 程度下回る値となった。

この原因の 1 つ目は入力に誤認識を含むようになったためである。必要な語彙の脱落や、不要な語彙の挿入によりドライバー発話から関連がありそうな物件を探索することが困難になった。

2 つ目は LSA による話題の定義が不十分なことである。ドライバー発話と類似度が高く、内容も適切な物件が存在しても、その物件が属するクラスタは必ずしも適切ではない。例えば 2-2 の設問では海水浴に行きたいという発話に対し、最も類似度が高い物件は海水浴場に近い旅館だった。この物件が属するクラスタは温泉宿に関連した物件が中心であり、被験者が求めている目的地提示になってしまう問題が起きた。

### 4 おわりに

本稿では、車内発話の話題分類に基づく目的地提案手法について論じた。評価実験の結果、発話の誤認識が少なく、話題分類に有用な単語が取得しやすい場合には適切な目的地を提示できることを確認した。しかし、LSA により話題を定義する精度が不十分であることや、車内発話の音声認識率の改善が課題である。

今後は、LSA に用いる単語に含まれてしまう不用語を除去することや、話題の定義を階層的に行うことによる、話題の定義の改善が考えられる。また、車内音声認識率を引き上げることが必要である。

謝辞

この研究は、新エネルギー・産業技術総合開発機構の産業技術研究助成事業（若手研究 Grant）の助成によって行われた。

### 参考文献

- [1] 河口他, “CIAIR 実走行車内音声データベース”, 情報処理学会研究報告. 2003-SLP-49, pp.139-144, 2003.
- [2] 加藤他, “対話構造木コーパスを用いた発話意図予測”, 情報処理学会研究報告. 2005-SLP-59, pp.73-77, 2005.
- [3] T.K.Landauer, et. al, Handbook of Latent Semantic Analysis, Lawrence Erlbaum Assoc Inc, 2007.