

音声対話型インターフェースのガイダンス出力 タイミング最適化に関する基礎検討

西澤 祐汰 (15816070)

ロペズ研究室

1. 背景

音声認識技術が急速に向上している[1]。マイクなどハードウェア面での精度向上も一つの要因であるが、SVM (support vector machine) などの学習モデルの洗練・大規模化とクラウドサーバ型のシステムの登場が大きい。このため、認識できるユーザの発話内容も飛躍的に増えた。

音声による操作を行うメリットは、主に 2 つ挙げられる。1 つ目は人と話すように操作ができる点である。パソコンなどの機械の操作に慣れていなくても、対話をするだけで操作ができる。2 つ目はアイズフリー、ハンズフリーで操作できる点である。他の作業をしても、機械を見ずに触ることなく操作ができる。

これらの音声操作のメリットは、運転時におけるカーナビなどのシステムの操作に適している。しかし、運転中に使用することを想定したシステムなので運転中の話すタイミングを考慮しなくてはならない。

本研究は「車載型音声対話システムにおけるガイダンス出力タイミングの最適化」を目的とし、具体的な目標は、以下の 2 つを掲げた。1 つ目は、多くの運転状況に対しての対話のし辛さを評価することである。2 つ目は、対話がし辛い運転シーンの特徴の明確化である。

2. 関連研究

内田らの研究以前は、運転状況のみ[2]や対話状況のみ[3]負荷を考慮する研究はあったが、運転状況および交通状況を踏まえた対話への心的負荷および作業負荷の研究はなかった。

本研究と同じく、「車載型音声対話システムにおけるガイダンス出力タイミングの最適化」を目的とした内田らの研究[4]では、運転状況および交通状況の違いから各運転シーン（車線変更、直進）を作成した後、対

話を評価する手法として「果物を 5 つ答えてください」といった課題を提示した。これは、運転シーン通過中にいくつ発話できたかに着目することで運転シーンごと対話のし辛さを指標化することを可能にした。

3. 各運転シーンに対しての対話のし辛さを評価

内田らの研究では、4 つの「車線変更」と 5 つのその他の運転シーンを比較することで、「車線変更」は、明らかに対話がし辛いシーンということを示した。本研究でも「車線変更」は、対話がし辛い運転シーンであるという結果に基づき進めていった。

まず運転状況および交通状況から、新たに多様な 19 個の運転シーンを定め、ドライビングシュミレーター上で作成した。運転シーンの数を増やすことでより多くの運転シーンの対話の負荷を考察できた。また、新たに聞き取り課題を設定したことで新たに、対話における聞き取りの負荷を考察できるようにした。

多くの運転シーンから「対話課題のみ」の発話終了時間と、「対話課題+運転課題」の発話終了時間の有意差検定を用いて運転状況を 3 グループに分けた後、3 グループ間の有意な差も示せた。これにより、新たな発話がし辛い運転シーンの抽出と運転シーンを、発話のし辛さで 3 グループに分けることができた（表 1）。また、聞き取り課題から発話ができても聞き取りがし辛い運転シーンを抽出できた。

4. 対話がし辛い運転シーンの特徴の抽出

多くの運転状況を比較して対話のし辛さを評価できたのは、時間によらない実数値に基づく指標を作ったことや課題のガイダンスを運転シーンの開始点と結びつけ、運転コースに埋め込むことで、自動で流れることにしたことにより運転シーン間を正しく比較できた点大きい。

2019 年度（令和元年度）卒業論文要旨

表 1 発話課題に基づく運転状況の分類

運転シーン名	発話 負荷
対向車ありトンネル内直進（追従）	低
T 字路左折待機時間	
交差点内直進	
左折指示後直進	
左折	中
目的地を探しながら直進	
T 字路右折待機時間	
右折	
歩行者あり左折	
対向車あり先の見えない左カーブ	
本線から右折待機時間	
前方車両減速	
車が本線に入ってくる	高
先の見えない右カーブ	
車線変更	
赤信号から交通量の多い交差点内直進	
交差点内右折待機時間	
十字路内直進	
合流	

本研究で、運転シーンの数を増やし、聞き取り課題の負荷を示せたことは、運転シーン間の比較による対話のし辛い負荷の原因の考察や、運転信号や NIRS 装置からとれる生体信号における対話のし辛い運転シーンの特徴の考察につながった。

運転信号をもちいた機械学習における対話のし辛いシーンの分類は、汎化性を示し、運転信号の特徴量が対話し辛い負荷の原因になっていることを示した。また、現時点でリアルタイムに分析できている上に、運転動作自体を範囲に機械学習のモデルを作成することで、運転信号の特徴量が与える対話し辛い負荷の分類精度を改善できると考える。

運転信号が対話のし辛さに関係しているという研究は神沼らが行っていたが[5]、本研究では、運転信号に

現れない負荷の存在を聞き取り課題と酸素化ヘモグロビンの濃度の変動量の絶対値の減少から示唆できた。

これらより本研究の 2 つ目の目標である「対話がし辛い運転シーンの特徴の明確化」まではできなかったが、対話がし辛い運転シーンの特徴を考察し、抽出することができた。

5. まとめ

1 つ目の目標である「多くの運転状況に対しての対話のし辛さを評価すること」を達成し、2 つ目の目標である「対話がし辛い運転シーンの特徴の明確化」は、達成できなかったが足掛かりはつかめた。よって本研究の目的である「車載型音声対話システムにおけるガイドダンス出力タイミングの最適化」に近づいた。

今後は、本研究で達成できなかった「対話がし辛い運転シーンの特徴の明確化」にむけて、研究を進める。

参考文献

- [1] 河原達也, 音声認識技術の現状と将来展望, 電気学会誌, 133(6):364-367, 2013
- [2] Yamamoto T. Kanamori R., Ando A. and Morikawa T. ive driving-stress and multiple physiological indicators in driving experiment, IEEE, 2017
- [3] 山岡将綺, 原直, 阿部匡伸, ほか, 車載用音声対話システムにおけるユーザ負荷を考慮した対話戦略の検討, 2014
- [4] 内田昂, 音声対話型インターフェースのガイドダンス出力タイミングの最適化に関する基礎検討. 青山学院大学修士論文, 2018
- [5] 神沼充伸, 内田昂, ロペズギヨーム, 運転信号を用いた車載型インタフェースのガイドタイミング制御