

# 音声によって運転指示可能な自動運転車両の開発

## Driving Self-Driving Cars by Spoken Language Instructions

塚原裕史<sup>\*1</sup>    日垣博考<sup>\*2</sup>    大前学<sup>\*2</sup>    大田原菜々<sup>\*3</sup>    稲子明里<sup>\*3</sup>    小林一郎<sup>\*3</sup>  
 Hiroshi Tsukahara    Hirotaka Higaki    Manabu Oomae    Nana Ootawara    Akari Inago    Ichiro Kobayashi

<sup>\*1</sup>株式会社デンソーアイティラボラトリ  
 Denso IT Laboratory, Inc.

<sup>\*2</sup>慶応義塾大学大学院政策・メディア研究科  
 Graduate School of Media and Governance, Keio University

<sup>\*3</sup>お茶の水女子大学  
 Ochanomizu University

Self-driving cars running on public roads are getting to be able to see recently, even though their operation is limited in certain areas and roads. Currently, those self-driving cars automatically maneuver themselves in 'from a point to a point' manner on a route that is planned by an equipped navigation system. However, the driving situation changes dynamically and there should be some unexpected things happen that is not scheduled. Thus, a user interface of self-driving cars should be devised for sharing the dynamical environmental situation around it with human passengers, i.e., the grounding, and for understanding their intentions to control it in response to what is occurring. In this study, we developed an experimental self-driving car with a spoken language interface through which passengers can give their driving instructions and can intentionally step in the operation of its control system. We evaluated the usability and the issues of such interfaces by using it on a test driving course in a private area.

### 1. はじめに

現在、非常に限られたエリアあるいは運転シーンのみではあるが、公道を走る自動運転車を見かけることができるようになってきている。ただし、現在の自動運転車は、基本的にはナビゲーションシステムによって、あらかじめ設定されたルート上を自動的に移動するものであって、途中で予期しない事故などの動的な道路状況において、安全に経路を変更し、問題を回避するためには、人間のドライバーが操作を引き継ぎ、運転する必要がある。そもそもハンドルが設置されていない自動運転車も現れてきており、運転ができないような人たちが自動運転車を利用するようになって行くと考えられる。そうなれば、運転代行者を必要とせずに自律的な移動ができるようになり、生活の利便性やクオリティの向上に貢献できると考えられる。

仮にこのような自動運転車が市場に現れ、実際に日常生活の中で利用されるようになるとすると、単に始点と終点とを結ぶ経路を走るのみでは十分な利便性が得られるとは言い難い。人は移動中に急にトイレに立ち寄りたくなったり、周辺に興味あるものを見つけたら、一時的に停車させたり、あるいは予定の経路を部分的に変更したくなるものである。このような人の意図をいかに自動運転システムへ伝達すれば良いかということが、新たな情報インタフェースの課題として上がってくる。

従来からある情報インタフェースは、情報の世界に閉じているものであったが、人が自動運転システムとの間で音声言語を介して意思疎通を行うことができるようにするためには、人が視覚的に認識している自車周辺の環境が自然言語で表現された情報とシステムがセンサーを通じて認識している周辺環境とがお互いに共有され、かつ正しく対応付けられていなくてはならない。すなわち、自然言語と実世界とのグラウンディングが重要となる。

以上のような課題を鑑みて、本研究では実際に音声対話に

よって運転指示が可能な自動運転車両を構築し、音声による運転操作を行う体験を通して、このような情報インタフェースの自動運転車における嬉しさや問題点に関する初期的な検討を行うことを目的とする。

### 2. 関連研究

コマンドレベルではなく、自然に発せられる音声対話によって実世界を自律に移動可能なシステムを操作する研究としては、貨物車ロボットを自然言語によって操作可能にする Kollar らの研究がある [1]。この研究では、指示文で表現されている空間的意味を Spatial Description Clause (SDC) という階層的な構造化データへ変換し、このデータ構造を手掛かりとして、言語情報と環境にあるオブジェクトや場所とを関係づける確率モデルを生成し、与えられた言語情報に対して、確率を最大化するオブジェクトや場所の組み合わせを探索することで、グラウンディングを行っている。

Kollar らの手法を日本語のような語順が自由に入れ替え可能で、ゼロ代名詞のように省略が頻出するような言語へも適用可能にするために、自然言語から直接的に SDC を抽出するのではなく、空間的意味をカテゴリとして定義する特殊な組合せ範疇文法 (CCG) を定義し、言語のどの部分がどのような空間的意味を表し、それらの間の依存関係を表す構文木を中間的なデータ構造として抽出し、そこから SDC を生成するという手法が稲子らによって提案されている [2]。

### 3. システム構成

#### 3.1 実験車両

本研究では、図 1a に示すような市販のミニバンを改造した実験車両を構築した。またコックピットから見た車内ディスプレイの表示を図 1b に示す。ディスプレイには、自動運転経路上の自車位置から前方の経路情報や LiDAR によって検出されている自車周辺の障害物の情報が表示されている。また、一時

連絡先: 塚原裕史, 株式会社デンソーアイティラボラトリ,  
 〒150-0002 東京都渋谷区渋谷 2-15-1, 渋谷クロスタワー  
 28F, htsukahara@d-itlab.co.jp

停止するポイントでは、標識マークを表示するなどの情報が提供される。



図 1: 実験車両

### 3.2 階層的制御アーキテクチャ

この実験車両における音声による運転指示から車両制御までを繋ぐ概念的なアーキテクチャーを図 2 に示す。このように、3 つのレベルの制御がある程度、お互いに独立性を持って制御を行うようになっている。

一番上位の目的制御層は、音声対話で与えられた自然言語情報を入力とし、それを意図された運転指示内容へ変換し、下位のリスク制御層で、実際の車両制御レベルの情報へ解釈され、下位の車両制御層へ渡され、実際に車両のアクチュエーションがなされる。図中の  $T_E$  は各層における入力情報を処理し、その内容を理解するために必要な処理時間であり、 $T_S$  は、その内容を元に制御目標を生成するために必要な処理時間、 $T_A$  は、制御を実行し、そのフィードバックを得るのに必要な時間を表している。上位層ほど制御周期が長くなっている。また、最下層は車両近傍の物理的な世界の情報のみによる制御であり、経路に沿って移動する、障害物がある場合には衝突をさせるために停車するなどの反射的な制御を行うが、中間層や最上位層では、過去の履歴や知識に基づく知的な制御を行うようになっている。

この階層的な制御アーキテクチャーは、Engström らによる 4 階層からなる Extended Control Model (ECOM) [3] の monitoring, regulating, tracking 層からなるモデルとみなすことができる。但し、我々の制御モデルでは上位層から下位層へのトップダウン的な相互作用のみでなく、下位層から上位層へのボトムアップ的な相互作用も考慮されている。また、各層が主に管理するデータについても示している。これらのデータは他の層とも交換し合うことは可能とする。

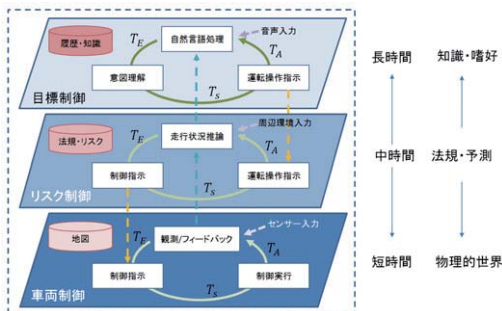


図 2: 実験車両の階層的制御アーキテクチャー

以下では図 2 の上位 1 層を担う部分を運転指示理解部、下位 2 層を担う部分を車両制御部と呼ぶ。(今回は実験用のコースのみを走行するため、第 2 層を車両制御部に含めたが、公道を走行する場合には、第 2 層の役割が増大するため、第 3 層とは独立に制御を行うべきであろう。)

### 3.3 各層間の連携インタフェース

図 3 に、運転指示理解部と車両制御部との連携インタフェースを示す。これらの制御部は、それぞれ独立に動作し、お互いにソケット通信 (TCP) によって接続し、シンプルな制御文字列の交換によって、お互いに連携することができるようになっている。運転指示理解部から車両制御部へは運転指示以外に定期的に車両制御部から GPS 情報、車両速度、グリッドマップなどの情報送信を要求する制御メッセージを送信し、車両制御部は、その要求を受けたタイミングで要求された情報を運転指示理解部へ送信する制御メッセージも定義している。グリッドマップ情報については、データ通信量を削減するため、路面に相当するグロッドの情報は送信しないように制限している。

運転指示理解部は、さらに音声認識部と音声合成部とソケット通信を介した XML メッセージの交換で連携することができるようになっている。音声認識部は、Bluetooth マイクからの音声信号を常時受信し、発話区間を認識すると、クラウド上の音声認識サーバへ発話音声を送信し、認識結果を運転指示理解部へ送信する。音声認識部は、WiFi ルータを介して、モバイルネットワーク回線に常時接続されている。今回、音声合成部は、クラウドを介さずローカルで音声合成を行うようにした。

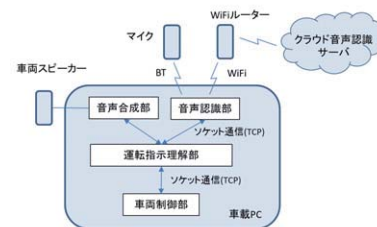


図 3: 運転指示理解部と車両制御部との連携インタフェース

### 3.4 車両制御部

図 4 に実験車両の車両制御部のシステム構成図を示す。車両制御部では、自動走行制御システムが、ハンドル角や速度等の車両状態量を CAN 経由で受信し、目標速度や目標ハンドル角を実現するよう、車両 ECU に車両センサ (ペダルストロークセンサ、トルクセンサ) の疑似信号を与えることで、車両の速度制御、操舵制御を可能としている。自動走行制御システムは、予め絶対位置情報で記述された走行パス情報と、後述する自車位置推定により得られた現在の車両の位置・向きとの相対位置の関係から、走行パスに追従するための操舵角を計算し、目標ハンドル角を決定している。目標速度については音声処理によって決定した目標速度に応じて設定している。障害物回避や大回り指令など左右に寄る操舵の場合は、走行パスから横方向にオフセットした位置を目標コースとすることによって目標ハンドル角を決定している。左折指令などで経路を変更する場合は走行パス情報を切り替えることで対応している。自車位置の推定については、速度、ヨーレート、RTK-GNSS の測位情報を用いた自車位置推定と、速度、ヨーレート、LiDAR による測域情報と二次元反射強度地図のスキャンマッチングを用いた自車位置推定 [4] の二重系の自車位置推定を行っている。

### 3.5 運転指示理解部

運転指示理解部では、大きく運転指示と駐車位置指示との 2 つを扱う。前者は車両自身の制御で閉じた指示であり、後者は環境とのグランディングが必要となる指示である。以下、それぞれについて説明する。



図 4: 車両制御部のシステム構成

### 3.5.1 運転指示理解

音声によって基本的な運転操作が行えるようにするために最低限必要なものとして、(1) 停車状態からの発進、(2) 加減速(変化量、変化速度指定可能)、(3) 停車指示(停止までの距離指定可能)、(4) 横位置変位(3段階のレベル指定)、(5) 右左折(直近の交差点)、(6) 速度指定(目標速度、速度変化速度指定可能)、(7) 徐行、(8) 巡行状態への復帰、(9) 駐車位置指示開始指示の運転指示ができるようにした。これらはグラウンディングを行わずとも指示可能なレベルであると考えられるので、ルールベースによるマッチング処理で実装する。但し、コマンドの音声認識とは異なり、様々な言い方に対応できるように、さまざまな運転操作の発話パターンにも対応できるように設計した。

なお、もし特定の場所を指定した右左折や停止操作を行うことができるようにするためには、以下の駐車位置指示のように場所を特定するためのグラウンディング処理が必要となる。

### 3.5.2 駐車位置指示理解

駐車位置指示では、同乗者が自由な発話形式で駐車場所をシステムへ指定することができるようにする。但し、今回は実験コース内の特定の駐車場で固定されたシーンの中で、空いている駐車場所への駐車位置指定を理解可能なようにすることを目指す。

図 5 に示すように、シーンが固定され、駐車できる場所も限られていたとしても、人が表現する指示文の内容には、非常に多様性がある。また、駐車位置指示を出す場所が変わると見え方が大きく変わり、指示の出し方が変わる。



(a) 駐車指示地点 1



(b) 駐車指示地点 2

図 5: 駐車シーンにおける別視点からの見え方の違い

図 6 における「データ」で示したグラフは、クラウドソーシングを利用して 100 人の一般のワーカーから収集した 1000 件の駐車指示文の文字数に対する頻度(規格化してある)をプロットしたものである。駐車指示文は 2 つの地点それぞれにおいて「あなたが助手席に乗っていると、図 5 に見える 5 つの駐車可能な各スペースに駐車する場合に、ドライバーへどのような指示を出しますか?」という質問の回答として収集した

ものある。最短で 5 文字、最長のものは 76 文字からなり、平均値は 19.7 で、分散は 78.5 であり、かなり幅が広いと考えられる。実際、指示文の多様性に対して、データ数がまだ少ない為、分布の揺らぎが大きい。ポアソン分布、負の二項分布、ガンマ分布でフィッティングを行うと、ポアソン分布はあまり一致が良くないことが分かる。一方、負の二項分布とガンマ分布は比較的良く分布の傾向を表しているように見える。このような分布を取った場合、ポアソン分布が得られるのではないかとまず予想されるが、比較的長い指示文も多いことから、分散が大きく、ポアソン分布よりも幅が広い分布になっていると考えられる。仮にポアソン分布に従うと考え、そのパラメータがガンマ分布を事前分布として揺らいでいるとすると、その予測分布は負の二項分布となるため、場所や人ごとにパラメータが揺らいでいることを反映しているとも考えられる。

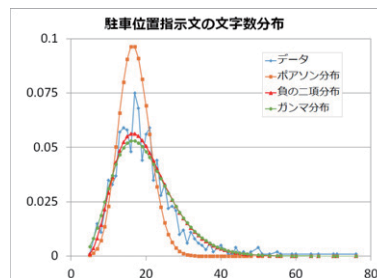


図 6: 駐車位置指示文における文字数分布

このように駐車位置指示の表現には多様性が高い為、ルールベースによるパターンマッチングで対応することは困難である。しかし、今回は、駐車位置が固定されているため、与えられた駐車指示文に対するクラス分類の問題として扱うことが可能である。そこで、運転指示文から単語に関するユニグラム、バイグラム、トライグラム素性として、リッジロジスティック回帰によって学習した。10-分割交差検定による評価で、リッジパラメータを 0.1 から 100 の間でグリッド探索し、リッジパラメータが 10 付近で最も良い汎化性能が得られ、Cohen の  $\kappa$  値で 0.68 であった。

実験車両では、駐車位置が正しくシステムに認識できたかを確認するために、図 8b に示すように、認識された駐車位置に黄色い車両マークを表示し、指示者が確認できるようにした。(白い車両マークは自車位置) システムが音声合成により、「この場所で良いですか?」と質問した後に、同意する旨の応答を返すと、自動駐車が開始されるようにした。また、同意せずに別の駐車位置を指定し直したり、駐車指示を中止することも可能なようにしてある。

## 4. 実車実験

### 4.1 実験コース

図 7 に、実車実験を行った慶大新川崎キャンパス構内の走行コースを示す。このコースは一方通行になっていて、矢印で示した方向へ進行することができる。一番外側は周回できるコースで、途中に左折ができる交差点が 3 箇所ある。(右折して、このコースから外へ出る道路は除いている。) 図中のアルファベットは、以下の運転操作を行うことを想定した地点を示している: (A) 直進(デフォルト)か左折か、(B) カーブを内側寄りに曲がるか、外側寄りに曲がるか、(C) 徐行するか否か、(D) 障害物回避、(E) 駐車位置指示。但し、衝突したり、逆走する恐れがない限り、任意の場所で任意の運転指示を音声で指示すること

が可能である。



図 7: 実験コース

## 4.2 実験結果

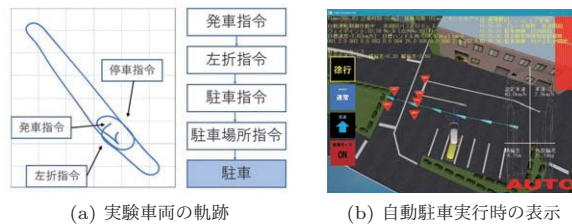
実際に図 7 における駐車場 E にある所定の駐車場所に停車している状態から出発して、音声によって運転指示を行い実験コースを回って、最後にまた駐車場所 E へ戻り、正しく指定した駐車場所へ停車させることができるかを実験した。図 8a に、実験車両の運転操作順序とその時の移動軌跡の一例を示した。このように、図 2 に示したようなアーキテクチャの元、図 3 に示したようなソケット通信を介した制御メッセージのやり取りによって、人が音声で指示した運転操作に従って移動できていることが確認できる。また、車両制御部から運転指示理解部へは要求に応じて一定周期で GPS 情報、車両速度、LiDAR によるグリッドマップ情報を送信しているが、特に大きな遅延などなく送信できていた。

駐車シーンにおいては、自車位置と駐車された車両との相対的位置関係に関する表現から、指定した駐車位置が認識され、意図した通りに自動駐車できることが確認できた。駐車位置指示後の車内ディスプレイの表示を図 8b に示す。音声指示で認識された駐車位置には、画面上で黄色の車両マークを表示し、駐車位置が意図した通りであるか確認する音声を流し、確認する音声入力を確認後に、自動駐車を実行する。車両マークは、駐車向き（前向き、後ろ向き）が指定された場合には、その向きが分かるように合わせて表示する。なお、この確認操作中に、画面上で車両マークをタッチして駐車のを切り替えるなどの操作も可能にしても良いだろう。

このように音声指示により意図にあった運転指示が概ね行えることが確認されたが、一方で、本実験から以下のような課題も観測された: (1) 主にクラウドサイドでの音声認識による遅延により、音声で運転指示を出してから、音声認識結果が得られるまでの時間（図 2 における  $T_E$ ）に、その運転操作を行うタイミングを逸してしまうことがある。(2) 駐車位置指示以外のグラウンディングを必要としないと仮定した操作であったとしても、ある地点にきたタイミングで運転操作を実行したいというようなグラウンディングが必要となる場合がある。(3) 障害物を回避したり、徐行をした後に元の巡行状態に戻りたい時に、その運転指示が単純には出せないことがある。例えば、3.5.1 節で示した運転指示 (8) で復帰させたい場合、単に「復帰して」

などと言っても、システム的には徐行のみを解除するのか、経路上の横位置のみを元に戻すのか、あるいはその両方を行えば良いのか、人とシステムとの間での齟齬がある場合がある。(4) 人が出した指示に従うと障害物と近づき過ぎ、そのままでは車両制御部で車両が停止されたままになってしまうことがある。

(1) については、ローカルの音声認識とのハイブリッド構成を利用するなどによって対処することが考えられる。また、(1-3) については操作履歴を収集し、あらかじめ場所や走行状況において、どのような運転操作指示が出やすいかなどを学習し、あらかじめ制御計画の準備をしておく、または実行タイミングや制御量を個人適合させるなどの対処が考えられる。



(a) 実験車両の軌跡

(b) 自動駐車実行時の表示

図 8: 実験結果

## 5. まとめ

本論文では、音声対話による運転指示に従って車両を制御することができる自動運転車両を構築し、実験コースにおいて実際に音声指示による運転操作を行う実験を行い、設計の妥当性を確認すると同時に、概ね意図に合った制御ができることを確認した。一方、実験から本システムでの音声指示における課題も得られた。今後、これらの課題を解決する取り組みと共に、動的な環境でのグラウンディング、リクス制御層での事故リスク知識 [5] などを取り入れた制御などについても対応して行きたい。

## 参考文献

- [1] T.Kollar, S.Tellex, D.Roy, and N.Roy, Toward Understanding Natural Language Directions, ACM/IEEE Int'l Conf. on Human-Robot Interaction (HRI), pp.259-266, 2010.
- [2] 稲子明里, 塚原裕史, 小林一郎, 「自動運転の言葉による指示を対象にした空間的意味表現の構造化への取り組み」, 2B3-OS-07a-5in2, 愛知, 2017 年度人工知能学会全国大会 (第 31 回), 2017.
- [3] Johan Engström and Erik Hollnagel, A General Conceptual Framework for Modelling Behavioural Effects of Driver Support Functions, in "Modelling Driver Behaviour in Automotive Environments", ch. 4, pp.61-84, 2007.
- [4] 大前学, 松下寛治, 久松亮史, 佐藤周也, 岡田成弘, 反射強度グリッドマップを活用した自動車の自動運転システム, 第 13 回 ITS シンポジウム 2015 講演論文集, 2015.
- [5] 西村悟史, 岩田麻希, 黒川美和, 丸田峻也, 梶大介, 丹羽伸二, 西村拓一, 江原遥, 「法例・判例ベース自動運転システムの実現へ向けて」, 2F4-NFC-03b-4in1, 愛知, 2017 年度人工知能学会全国大会 (第 31 回), 2017.