

マルチモーダル対話型自動運転車の誕生！ ～ドライバーと自動運転車の新しい人間—機械協奏関係の構築～

この度、国立大学法人名古屋大学（愛知県名古屋市、総長：松尾 清一、以下 名古屋大学）は、国立大学法人徳島大学（徳島県徳島市、学長：野地 澄晴、以下 徳島大学）、アイシン精機株式会社（愛知県刈谷市、取締役社長：伊勢 清貴、以下 アイシン精機）、との共同で音声・視線・ジェスチャを用いて操作する自動運転車を開発しました。

今回の研究は、名古屋大学の 武田 一哉 教授、竹内 栄二郎 准教授、石黒 祥生 特任准教授らの研究グループが、徳島大学、アイシン精機（協力：株式会社豊田中央研究所）と共同で、自動運転オープンソースソフトウェア^{注1)}『Autoware』を利用し、音声認識・顔画像認識・ジェスチャ認識を組み合わせたマルチモーダルインタフェース^{注2)}で、自動車を直観的に操作するシステムを開発したものです。

これにより、例えば「そこを曲がって」と指をさせば、ドライバーの意図するところで曲がってくれるなど、コミュニケーションをはかりながら操作できる新たな自動運転車が誕生しました。今後、ドライバーと自動運転車との新しい人間—機械協奏関係が拓かれていくことが期待できます。

なお、この研究は平成 28 年度から始まった科学技術振興機構『産学共創プラットフォーム共同研究推進プログラム』の支援のもとで行われたものです。



自動運転車両外観



実験の様子

特徴

- ・ 一般ユーザでも自動運転車を操作できるマルチモーダルインタフェースを開発
- ・ 音声・視線・ジェスチャを用いて「タクシー運転手に指示するように」自動運転車を操作可能
- ・ 音声認識技術、視線・顔向き・口動作認識技術、ジェスチャ認識技術を結集
- ・ 将来の「機械—人間」協奏の一つの在り方を提案

研究の背景と経緯

どのように「自動運転車を操作する」のか？

近年、自動運転車の研究開発は急速に進展し、日米で公道において走行実験が行われるようになってきました。2020年の東京オリンピックまでには一部で実用化を目指す動アジア各国の交通事情が劇的に改善されうる可能性もあります。

しかし、自動運転のための周囲のセンシングや車の制御などの技術は進展しているものの、実際に一般の人が自動運転車に乗り込み、目的の場所に移動するために、どのように「自動運転車を操作する」のかについての検討は、あまり行われてこなかったのが実情です。自動運転車が真に社会に浸透し、一般の人が容易に利用できるようにするためには、自動運転車と人とのコミュニケーションが必要になります。すなわち、自動運転車のインタフェースが必要です。この一つの理想形がタクシーではないでしょうか。乗客は運転手に行き先を告げ、必要に応じて曲がる場所などを伝え、行き先に来たら停まってもらう、こんなやり取りは、音声による対話を中心に、ジェスチャなどを交えながら交わされます。こんなやり取りを自動運転車との間でできないか、そんな期待が持たれてきました。

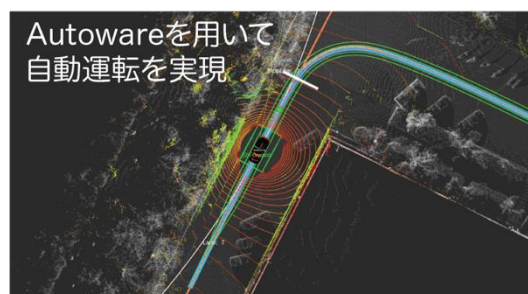
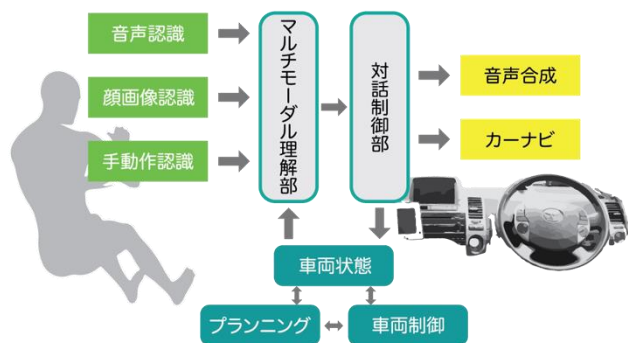
研究の内容

マルチモーダルインタフェースを備えた自動運転車

研究グループは、音声対話を中心とし、ジェスチャや視線も情報伝達手段として用いることのできる、「人対人のコミュニケーション」から発想を得た『自動運転車用マルチモーダルインタフェースシステム』の開発に取り組みました。

このシステムは、人（ユーザ）の音声、ジェスチャ、視線などはそれぞれ音声認識、深度センサ、映像処理を用いて認識をします。例えば、ユーザが「右に曲がって」と発声すれば、システムはユーザの意図に従って、自動車に右に曲がるように制御を依頼し、実際に自動車は右に曲がります。一方、音声とジェスチャや視線を同時に用いる場合、それらの入力は並行して行われるので、ユーザが自動運転車に伝えようとする「意図」はそれらを統一して理解する必要があります。例えば、ある建物を見ながら「あれは何？」と尋ねれば、「あれ」と言っている時にユーザが見ている建物の名称を答えてほしい、という意図があります。このように、各入力手段（モダリティ）の時間的關係も考慮しながら、マルチモーダル理解を行います。そして、その意図に応じて応答し、車を制御します。

このように、複数のモダリティを統合的に理解し、実際に自動車の制御まで行いながら応答するマルチモーダルインタフェースを備えた自動運転車を世界で初めて開発・実現しました。



今後の展開

機械と人間の協調・共生

自動車はカーナビなどを含めて情報化が進んでいますが、今後、自動運転も含めて様々な車載機器の操作を容易に行うための一つの方向性として、こうしたマルチモーダルインタフェースが進展することが考えられます。また、自動運転車に限らず、機械と人間の協調・共生を考えた場合に、マルチモーダルインタフェースは自然で使いやすいインタフェースとして、取り入れられていくと期待されます。

成果の意義

自動運転車を自然に操作するシステム

自動運転車を音声・ジェスチャ・視線を用いて自然に操作するシステムを開発しました。自動運転車に象徴される将来の「機械」と「人間」との協調・共生の在り方を示す一つの形です。また、音声操作、ジェスチャ認識、顔画像認識（視線認識、顔向き認識、口動作認識）といった個々の技術それぞれが、高度化した自動車、あるいは情報機器とのインタラクションの手段として有効であることを提案します。

【用語説明】

【オープンソースソフトウェア】

ソフトウェアのソースコードを一般に公開し、ソフトウェア利用者自身による利用・修正・再頒布を許すことによって、ソフトウェア開発の発展を目指したもの。

【マルチモーダルインタフェース】

音声やジェスチャなど、人間のさまざまな入力を受け付け、機器の操作などに反映するインタフェース。音声など個々の入力手段を「モダリティ」と呼び、それが複数あることから「マルチモーダル」と呼ばれる。

■デモンストレーションビデオ

<https://youtu.be/Mesx4qgONqs> （英語）

<https://youtu.be/EkVdZUEEFwM> （日本語）