

人間型ロボットによる対話の人間らしさの向上 ～自由に移動できる子供型アンドロイドも開発～

ポイント

- カメラやマイクを駆使したマルチモーダル対話制御システムを開発し、アンドロイド「ERICA」の日常的な状況における自然な存在感を向上させた。
- 複数のロボットを用いたマルチロボット対話制御システムを開発し、社会的対話ロボット「CommU」による「対話しているという感覚（対話感）」を高めた。
- 車輪移動機構で自由に移動できる子供型アンドロイド「ibuki」を開発した。

JST 戦略的創造研究推進事業において、ERATO 石黒共生ヒューマンロボットインタラクショナルプロジェクトの石黒 浩 研究総括（大阪大学 大学院基礎工学研究科 教授、株式会社国際電気通信基礎技術研究所 石黒浩特別研究所 所長・ATRフェロー）、河原 達也 グループリーダー（京都大学 大学院情報学研究科 教授）らは、マルチモーダル^{注1)}対話制御システムとマルチロボット対話制御システムを開発し、人間らしい存在感や対話感^{注2)}を対話相手に与えるロボットを実現しました。また、車輪移動機構を持つ子供型アンドロイド「ibuki（イブキ）」を開発しました。

近年、対話ロボットの研究開発が盛んになりつつあります。しかし、従来のロボットとの対話では、人間との対話で得られる対話感や存在感、社会性を感じることができません。本プロジェクトは、カメラやマイクロフォンアレイ^{注3)}を用いたマルチモーダル認識システムや、意図や欲求に基づく対話制御システムにより、人間らしい存在感を実現しました。さらに、自然で多様な相づち生成や焦点語^{注4)}に基づく聞き返し技術により人間らしい対話感を実現し、アンドロイドが傾聴や面接を行える可能性を示しました。

また、対話感を演出する社会的対話ロボットの研究では、ロボット同士の掛け合いや役割交代をさせるなど、複数のロボットの発話や非言語的表現の表出タイミングを制御するマルチロボット対話制御システムを開発しました。ロボット同士の対話を人間に見せることで、ロボットが人間の発言や意図を認識できない場合でも、対話相手である人間に強い対話感を与えることを実現しました。

車輪で移動できる子供型アンドロイド「ibuki」は、移動を伴う人間との親和性を追求し、対話ロボットが活躍する場面を拡大するための研究基盤です。本プロジェクトは、これらの研究開発を通して、日常的な場面で人間とコミュニケーションできる自律対話型アンドロイド^{注5)}の実現を目指すとともに、対話ロボットの社会への普及を推進します。

本成果は、以下の事業・研究領域・研究課題によって得られました。

戦略的創造研究推進事業 総括実施型研究（ERATO）

研究プロジェクト：「石黒共生ヒューマンロボットインタラクショナルプロジェクト」

研究総括：石黒 浩（大阪大学 大学院基礎工学研究科 教授、株式会社国際電気通信基礎技術研究所 石黒浩特別研究所 所長・ATRフェロー）

研究期間：平成26年7月～平成32年3月

上記研究課題では、特定の状況と目的において自律的に対話できる機能、複数の情報伝達手段を用いて社会的状況で複数の人間と対話できる機能など、実社会において人間と親和的に関わり、人間と共生するための自律型ロボットの実現を目指しています。

<研究の背景と経緯>

現在、さまざまなロボットが一般家庭などの日常生活の場で活動しています。その中で、身体的なサポートだけではなく、対話などの関わりを通して心理的にも人間をサポートするロボットに不可欠なヒューマンロボットインタラクション技術の確立が重要になってきています。その中でも、人間型ロボット、特に外観が人間に酷似したロボットであるアンドロイドは、その姿形から人間と親密に関わり「共生」するロボットとして期待されています。

一方で、対話中の人間型ロボットの振る舞いや発話における「ちょっとした」違和感は、人間との親和性に大きな影響を与えます。この違和感を解消するためには、音声認識、発話生成、反応動作、ロボットの見た目の人間らしさなど、個々の要素技術を洗練することが必要不可欠であるとともに、これらの要素技術を統合する研究開発が必要です。本プロジェクトは複数の要素技術を統合し、日常的な状況において、人間型ロボットの自然な対話を実現する研究を進めています。

<研究の内容>

本プロジェクトでは、限定された状況や目的に限られるものの、日常的な場面において人間らしい存在感を感じさせる、アンドロイドのための対話システムを開発しました。具体的には、距離センサー（人の位置や頭部動作の認識）、カメラ（人の表情の認識）、マイクロフォンアレイ（発話者の位置や音声の認識）など多様なセンサーを用いるマルチモーダル認識システムを構築しました。また、意図や欲求に基づく発話、動作、視線や感情を制御することで、人間らしい存在感を感じさせる対話制御システムを開発し、既存成果であるアンドロイド「ERICA（エリカ）」との対話環境を構築しました（図1）。これらにより、研究所の待合室でアンドロイドが初対面の訪問者と対話するという状況において、人間らしい自然な対話を実現しました。その対話内容は人間と比較すれば非常に限定されたものであり、今後もさらなる研究開発が必要です。しかし、人間とロボットとのインタラクションをより円滑にする基盤技術として、従来のロボットにはない自然な対話や存在感を実現することができました。

また、人の話を聞く（傾聴）、面接官として質問する（面接）という役割をERICAに与え、人間と同程度の自然な音声対話を目指す研究にも取り組みました。本研究では、自然で多様な相づちを打つ技術、焦点語を解析して聞き返す技術、相手の反応を検出する技術などを用いて、人間らしい対話感を醸し出すことで、人間の発話を促進し、長く対話できるシステムを開発しました。スマートフォンやスマートスピーカーで実現されている一問一答形式の検索に基づく既存の対話システムとは異なり、アンドロイドが人間らしい対話を行うことで、将来的に面接の練習や語学学習における教師やカウンセラーなどの社会的役割を担うことが期待されます。

さらに、社会的対話ロボット「CommU（コミュニ）」を複数用いた研究として、ロボット同士の対話を人間に見せることで対話感を演出し、対話の印象を向上させるマルチロボット対話制御システムを開発しました（図2）。本システムは、人間の発話を認識できなかった場合でも、これまで発言していたロボットではなく、対話に同席している別のロボットに発言をさせることで、対話の継続性を保ちます。音声認識の不完全さなどから、

直前の対話の流れに沿った発話ができない場合や、急な話題展開が起きた場合に、もう1台のロボットの意図に沿って対話が進んでいるような想像を喚起し、対話が破綻していないように感じさせることができます（音声認識なし対話技術）。また、曖昧で多義的な解釈可能な発言を含んだ対話を複数のロボットに行わせ、人間の発話の意図に関わらず矛盾のない対話を作り出すことで、対話の破綻を防ぐことができます（意図認識なし対話技術）。すなわち、複数のロボットの連携によって社会的状況を作りだし、人間の想像力を引き出すことで、対話における違和感を軽減できるのです。

人間が暮らす日常的な場面で活動する対話ロボットの実現を目指して、人間と共生するロボットに必要な対話システム技術の開発とともに、移動機構を持つ子供型アンドロイド「i b u k i（イブキ）」を開発しました（図3）。i b u k iは車輪型の移動機構に加え、偏心した車輪と直動機構の組み合わせによる全身の揺動機構と、上半身を駆動する腰関節を持ち、人間の歩行と同様の躍動感を表現できます。さらに、表情を表出するために必要な顔や頭部の駆動機構だけでなく、ジェスチャーを生成するための手や腕にも多数の駆動関節を実装しました。従来注6)の空気圧アクチュエーターで駆動されるアンドロイドとは異なり、全ての自由度がバックドライバビリティのある減速機付電気モーター注7)で駆動されます。このためコンプライアンス注8)を持ったハードウェアであるにも関わらず、移動には適さなかった空気圧システムのためのエアーコンプレッサー注9)が不要になりました。これにより人間が生活する場で、オフィスビルでの道案内や店舗での商品説明など人間と共に移動したり、互いの位置関係を調整したりしながら、身体的なインタラクションを行うことも可能なアンドロイドとなっています。

<今後の展開>

今後はより多様な状況や目的において、ロボットの自然な対話を実現する研究開発に取り組んでいきます。また、単に移動が必要なタスクを遂行するロボットの開発に取り組むのではなく、移動の仕方や位置関係を利用した、移動を伴う人間との親和的なインタラクション技術を開発し、日常生活で活躍する自律対話型アンドロイドを実現します。さらに、これらの研究開発を通して得られた知見を活用することで、情報提供、生活支援、学習支援を目的とした社会的対話ロボットのアプリケーションの開発に取り組んでいきます。

<参考図>



図1 アンドロイド「ERICA（エリカ）」の外観とマルチモーダル対話制御システム

<アンドロイド「ERICA」の仕様>

・ロボット本体

身長：166センチメートル（立位時）

自由度：頭部、首、両腕、腰に合計44カ所

アクチュエーター：空気圧アクチュエーター

音声合成：VoiceText

センサー：CMOSカメラ2個（左右眼球）、マイクアレイ2個（左右外耳）

皮膚素材：シリコン樹脂

頭髪素材：人毛髪

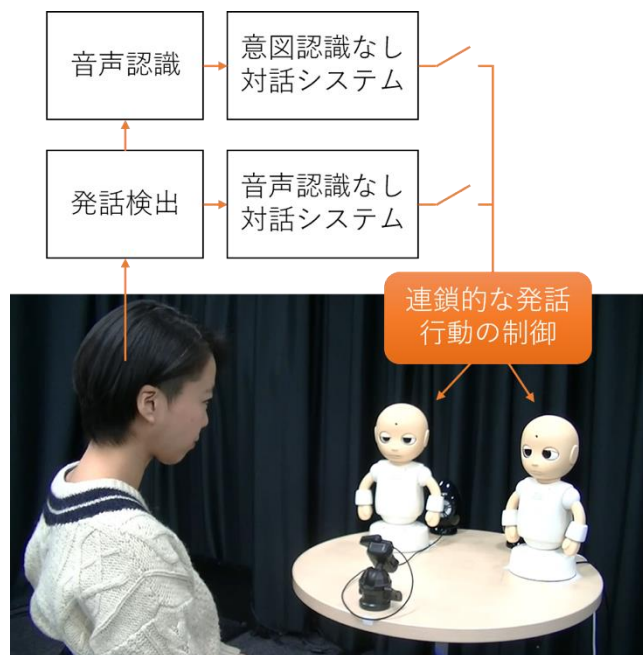
CPU：外部PC（Windows、Linuxなど）による制御

電源：空気圧レギュレーター用に外部電源（AC24V）

周辺機器：エアーコンプレッサー（AC100V）、制御用PC多数

・マルチモーダル認識システム

本体の周囲に配置した複数の距離画像センサー^{注10)}、16チャンネルマイクアレイ、CMOSカメラにより、人の位置、人の頭部動作、人の表情、発話者の位置、発話音声を認識する。



マルチロボット対話システム
(音声認識なし対話／意図認識なし対話)

図2 社会的対話ロボット「CommU (コミュニー)」によるマルチロボット対話システム

<社会的対話ロボット「CommU」の仕様>

- ・ロボット本体
 - 身長：30センチメートル
 - 体重：938グラム
 - 自由度：首部3、眼球部3、脛部1、口部1、腕部2×2、腰部2の合計14自由度
を用いた視線や表情、ジェスチャーの表出
 - アクチュエーター：サーボモーター
 - 音声合成：AITalk（日本語）、VoiceText（英語）
 - その他の出力モダリティ：スピーカー（胸部）、LED（両頬、胸部）
 - センサー：CMOSカメラ（額部）、マイクロフォン（胸部）
 - 制御システム：体内のRaspberryPI2/Edisonおよび外部PC（Windows、Linuxなど）による制御
 - 電源：外部電源（AC100V）
- ・マルチロボット対話システム
 - 本体付近に配置した、インテリジェントマイク（方向別分離機能を持った音声認識に適したマイクロフォン）、音声認識ソフトウェア、距離画像センサーにより、人の頭部位置、発話音声を認識し、複数台のロボットを同期して操作する。

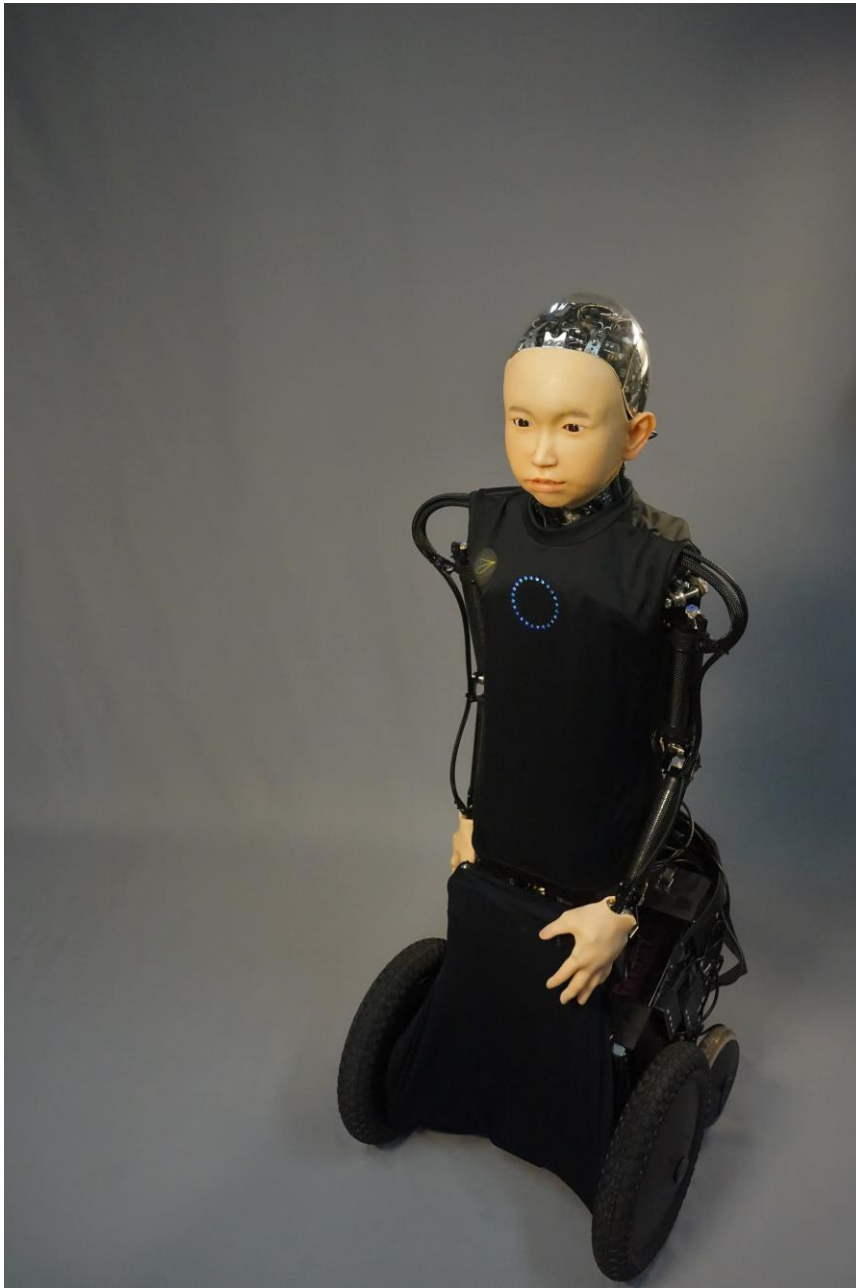


図3 移動型の子供型アンドロイド「ibuki」の外観

(開発中のアンドロイドの資料写真であり、外観は変更になる可能性があります。)

<アンドロイド「ibuki (イブキ)」の仕様>

・ロボット本体

身長：120センチメートル（立位姿勢、10歳程度の子供相当）

体重：約37キログラム（バッテリーを含む）

自由度：47カ所（頭部：15カ所、首：3カ所、腰：3カ所、腕：6カ所×2、手：
5カ所×2、移動機構：4カ所（左右独立駆動車輪+胴体上下動+ステアリング）

電動化によるスタンドアロンシステム^{注1)}

アクチュエーター：減速機付電気モーター（静音、低減速比、バックドライバブル）

センサー：CMOSカメラ2個（左右の眼球に1個ずつ）、2次元レーザー距離センサー、エンコーダーと電流センサーの組み合わせによるトルク推定

高密度実装モータードライブユニット：電流センサーによるトルク制御

皮膚素材：高柔軟シリコン樹脂

構造材料：繊維強化樹脂、カーボン素材による軽量化

電源：内蔵バッテリー（5.3Ah、25.2V）×4

・人の歩行時の全身の揺動を模擬する移動機構

左右の偏心車輪と上下直動機構を組み合わせ、人の歩行時の下半身の揺動を車輪駆動ロボットで模擬する。さらに、腰部の3軸の能動関節により人の上半身の揺動を模擬する。

・スタンドアロンシステムとネットワークシステムの統合

内蔵小型コンピューターによる自律動作およびROSを用いた外部高性能計算機とのシームレスな連携

<用語解説>

注1) マルチモーダル

インタラクションを行う際に複数の情報チャネルを利用する場合を示します。視覚（カメラ）や聴覚（マイク）といった複数種類の情報を統合して認識したり、発話だけでなく視線やジェスチャーを使って意図を伝えたり、人間が五感、身体や表情を使って行うような表現力の豊かな対話の実現できると期待されています。

注2) 対話感

本プロジェクトで注目している対話に関する指標の1つで、相手との対話に熱中したり、興味を持ったりすることで感じる「対話をしているという感覚」。人間との対話とは異なり、従来のロボットやCGエージェントとの対話では対話感が弱く感じられます。

注3) マイクロフォンアレイ

複数のマイクロフォンを備えた機器により、音源位置推定、音声分離や雑音除去などを行う技術。

注4) 焦点語

発話の主題となる単語。対話の流れを理解・認識する場合に重要となるキーワードです。例えば、「ゲームは気分転換にはもってこいです」という文は、ゲームについて述べています。この時の「ゲーム」がこの文の焦点語となっています。

注5) 自律対話型アンドロイド

遠隔操作型などとは異なり、操作者なしで人間と対話する機能を備えたアンドロイド。対話する機能には発話内容が適切であるという言語的側面に加え、視線方向、表情や振る舞いといった非言語的側面も含まれます。

注6) 空気圧アクチュエーター

圧縮空気を利用した駆動装置。アンドロイドの体の部位を動かすために用いられます。空気を用いた空気バネとして振る舞うことができるため、人間との身体的なインタラクションに必要な柔らかい動きを実現できます。

注7) バックドライバビリティのある減速機付電気モーター

ロボットの駆動によく用いられるサーボモーターは、ギア比の高い減速機を用いるため、外部から力が加わっても姿勢を変えずに剛体のように振る舞います。子供型アンドロイド *i b u k i* は、外部の力に応じて変位が変わる（バックドライバビリティのある）電気モーターで駆動されるため、柔軟な身体的インタラクションを行うことができます。

注8) コンプライアンス

ロボットに外力が加わった場合に、バネのように姿勢を変える特性。例えば、人間同士で握手する場合、相手の力の加わり方で動きが変わりますが、サーボモーターなどで駆動されるコンプライアンスのないロボットと握手する場合は人間がロボットの動きに完全に追従する必要が出てきます。このため、安全・安心な身体的なインタラクションを行うためにはロボットの持つコンプライアンスが重要になります。

注9) エアーコンプレッサー

空気圧アクチュエーターを駆動するための圧縮空気を供給する機器で、通常は小型冷蔵庫程度の大きさがあるため、移動型ロボットには適しません。空気圧アクチュエーターを駆動するためには可搬性の高い空気タンクも利用可能ですが、この場合は稼働時間が問題となります。

注10) 距離画像センサー

画像センサーのうち、各画素における色相や明度だけでなく、奥行方向の距離も同時に計測可能なセンサーで、RGB-Dセンサーとも呼ばれます。形状を計測することが可能なため、高度な認識が可能になります。なお、人間の視覚も両眼を用いたステレオ視による奥行方向の距離の知覚が可能です。

注11) スタンドアローンシステム

外部電源や通信ネットワークなしで単独で動くシステム。移動型アンドロイドの *i b u k i* は、小型の冷蔵庫と同程度のサイズのエアーコンプレッサーとの接続を必要とする従来の空有気圧アクチュエーター型のアンドロイドとは異なり、小型コンピューターとアクチュエーターをバッテリーで稼働することでスタンドアローンシステムとして稼働できます。また、無線ネットワークを介して高性能なサーバーの提供する高度な認識システムなどと統合した制御システムを構築することができます。