

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 15 日現在

機関番号：94301

研究種目：若手研究(B)

研究期間：2013～2014

課題番号：25730165

研究課題名(和文)人の行動を引き込むことで指示物体認識性能を向上させるロボットシステムの実現

研究課題名(英文)Object reference recognition utilizing entrainment in human-robot interaction

研究代表者

飯尾 尊優 (Iio, Takamasa)

株式会社国際電気通信基礎技術研究所・その他部局等・研究員

研究者番号：70642958

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：本研究では、人間が指示した物体をロボットが認識するという対話において、人間からロボットが認識しやすい曖昧性のない指示を引き出すための方法を研究した。研究の結果、人間は指示するときにロボットが確認で使った言葉と同じ言葉をする傾向があることを明らかにした。また、その傾向を利用して、曖昧性のない指示をさせやすくするロボットの確認行動の生成手法を開発した。評価実験により、この手法を用いて人間が指示された物体を認識する精度が向上する可能性が示唆された。

研究成果の概要(英文)：This research investigated how a robot elicited references to an object which is unambiguous and recognizable for the robot from humans in object reference conversation. The results revealed that humans tended to use same terms or same categories of terms in their references as the robot used in its confirmations. Based on the results, I developed the generating algorithm of confirmation behaviors of robots to entrain humans to use unambiguous terms. Through an evaluation experiment, it was suggested that the success rate of recognizing object references may be better by using the algorithm.

研究分野：知能ロボティクス

キーワード：知能ロボティクス ヒューマンロボットインタラクション

## 1. 研究開始当初の背景

ロボットによる指示物体認識において、人の指示行動の曖昧性による指示物体の誤認識率の増大という問題がある。ここで、指示行動の曖昧性とは、物体を一意に特定できない指示行動のことをいう。たとえば、環境に複数の雑誌がある場合でも、人は「雑誌をとってきて」と言うことがある。この発話には、物体に関して「雑誌」という情報しか含まれないため、環境に存在する全ての雑誌が指示物体候補となってしまう、指示物体を誤認識する可能性が増大する。実際、Shinozawa らの実験によれば、人がロボットに本を取ってくるように依頼するというタスクにおいて、全指示発話のうち、言葉だけで一意に本を特定できない指示発話は約 49% もあった[1]。

このような問題に対し音声対話システムの研究では、誤認識が生じた場合の対処や、さらなる誤認識を生じないようにするための対話戦略について研究されている[2]。特にシステム側から人への確認の有効性が指摘されている。これらの研究では、確認発話のみを考慮しているが、物体指示対話では、発話と指さしを含むマルチモーダルな確認行動に関して考慮すべきだと考える。

これまで申請者は、人とロボットの単純な物体指示対話において、ロボットの確認行動が人の指示行動に及ぼす影響を予備的に検証してきた[3]。その結果、ロボットが確認行動で使用した言葉を人は次回の指示行動で使用しやすくなるという傾向や、ロボットが確認行動で視線と指さしを使う場合、人も指示行動に指さしを付随しやすくなるという傾向を見いだした。この傾向は、人と他者との対話において、人がその他者の行う発話や身体動作に自分の発話や身体動作を自然に合わせようとする引き込み現象の知見と一致している。これらの予備的な検証と引き込み現象の知見から、申請者は、「ロボットの確認行動を使って、人から指示物体を一意に特定できる言葉や物体の位置を示す指さしを引きこみ、自然な対話の中で指示物体認識の誤認識を低減させられるのではないか？」という着想に至った。

## 2. 研究の目的

本研究の目的は、引き込み現象に関する知見から、ロボットの確認行動を使って、人から曖昧性のない指示行動を引きだすことによって、指示物体認識の誤認識を低減させることである。この目的を達成するために、以下の2つの研究課題に取り組んだ。

**研究課題1：物体指示対話における引き込み現象の発生の検証**

人間がロボットに対して環境中の物体を指示するという具体的なタスクを物体指示対話と定義し、物体指示対話において引き込み現象がどのように生じてくるかを理解す

ることを目的とした。具体的には、従来研究の知見に基づき、当該状況で発生すると考えられる引き込み現象に関して、「人間はある物体を指示するとき、前にロボットがその物体を確認するときに使用した言葉を採用しやすい」、そして「人間は物体を指示するとき、それまでロボットが物体を確認するときに使用してきた言葉のカテゴリに属する言葉を採用しやすい」という2つの仮説をたて、その仮説を検証する被験者実験を実施し、仮説の妥当性を示した。

**研究課題2：引き込み現象の知見を利用した指示物体認識システムの有効性検証**

人から曖昧性のない指示行動を引きだすために、物体の特徴や位置などの静的な環境情報と指示する人の位置やロボットの位置などの動的な環境情報、さらに指示行動の音声認識結果と指さし認識結果に基づいて、指示物体候補を算出し、その中から確信度が最も高いものを一意に特定する確認行動を生成する確認行動生成アルゴリズムを構築した。確認行動生成アルゴリズムを音声認識と指さし認識を組み合わせ、指示物体認識システムに実装し、提案アルゴリズムで生成されたロボットの確認行動が、曖昧性のない指示行動を引きだすかどうか、それが指示物体認識の誤認識率の低減に貢献するかどうか、を検証する被験者実験を実施した。

## 3. 研究の方法

それぞれの研究課題について研究方法を説明する。

**研究課題1：物体指示対話における引き込み現象の発生の検証**

物体指示対話における引き込み現象の発生を検証するため、人間とロボットが物体指示対話を行うときにおこりうる仮説をたて、それを検証する被験者実験を実施した。

我々は物体指示対話における語彙の引き込み現象に関して2つの仮説を立てた：

**仮説1**：被験者は物体を参照するため、ロボットが物体を確認するために使用したものと同じ言葉を使用する傾向がある (図1)

**仮説2**：被験者は物体を参照するため、ロボットが物体を確認するために使用したものと同一カテゴリを使用する傾向がある (図2)

これらの仮説を検証するために、物体指示対話を用いて2つの実験を実施した。

**仮説1の検証実験**

**【実験計画】**本実験は確認発話要因とセッション要因の2×2の混合計画を採用した。

確認発話要因では、被験者間で、ロボットが確認発話を行うか行わないかを操作した。具体的に、確認発話ありの場合では、ロボットは物体ごとに決められた確認発話を行った。一方、確認発話なしの場合では、ロボットは確認発話を行わず、「わかりました」と発話した。

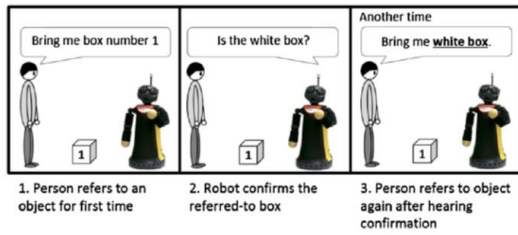


図 1 言葉の単位での引き込み現象

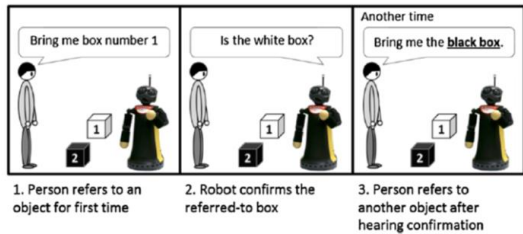


図 2 カテゴリ単位での引き込み現象

セッション要因について、本実験では、被験者は物体指示対話を 2 回実施し、前半をセッション 1、後半をセッション 2 とした。セッションとは、被験者が 5 冊の本を指示し終わるまでの区切りを意味する。セッション 1 では、被験者は 5 つの本を初めて指示するため、ロボットが各本をどのような言葉を使って確認するのか知らない状態であり、セッション 2 では、被験者はすでにセッション 1 に参加しているので、ロボットが各本をどのような言葉を使って確認するのか知っている状態であった。つまり、セッション要因は被験者内で確認発話を聞く前後の状態を操作したことになる。

【被験者】被験者数は 24 人で、女性 5 名男性 7 名を確認発話あり群に、女性 6 名男性 6 名を確認発話なし群に割り当てた。

【環境】実験の様子を図 3 に示す。本は被験者とロボットの間に配置された。各本の位置は全被験者で同じであった。本実験では、ATR 知能ロボティクス研究所で開発されたヒューマノイドロボット Robovie-R ver. 2 を用いた。

【オペレータ】現在の音声認識や画像認識技術では、自由な指示発話が可能な環境で被験者の音声を確認することが困難であり、実験を円滑に実施できない可能性があったため、オペレータが、マイクから被験者の発話を聞き、ロボットの右後方に設置されたビデオカメラを見て、指示された物体を認識した。

【手順】物体指示対話では、まず、ロボットが被験者に挨拶と自己紹介をした後、被験者に物体を指示するように依頼する。被験者は実験環境中に配置された複数の物体から自由に 1 つを指示した。ロボットは指示された物体を被験者に確認した。この確認は、その



図 3 実験の様子

物体に視線を向け、指差しをしながら、物体ごとに用意された確認用のセリフを発話するという方法で行った。確認した物体が正しければ被験者は次の物体を指示し、そうでなければ再度同じ物体を指示した。

【計測対象】ロボットが確認発話で使用する単語をターゲット語と定義し、ターゲット語を含む被験者の指示発話の数を計測した。

### 仮説 2 の検証実験

【実験計画】本実験は確認発話要因のみを含む  $1 \times 2$  の被験者間計画を採用した。

確認発話要因では、ロボットが確認発話で用いるカテゴリ、具体的には、色カテゴリ確認発話と確認発話なしを操作した。色カテゴリ確認発話では、ロボットは本の表紙の色に基づく確認発話を行った。一方、確認発話なしでは、仮説 1 の検証実験と同様に、ロボットは確認発話を行わず、「わかりました」と発話した。

【被験者】被験者数は 12 名であった。男性 4 名、女性 8 名であった。

【環境・ロボット・オペレータ・実験手順】実験環境と使用した本、ロボット、オペレータの操作規則、実験手順は仮説 1 の検証実験と同じであった。

【計測対象】ロボットが確認発話で使った色カテゴリの言葉をターゲット語と定義し、そのターゲット語を含む被験者の指示発話の数を計測した。

## 研究課題 2：引き込み現象の知見を利用した指示物体認識システムの有効性検証

研究課題 1 の結果を踏まえ、曖昧性のない指示行動を引き出すための確認行動生成アルゴリズムを構築し、有効性を検証した。

### 確認行動のデザイン

指示物体認識性能を向上させるという観点から見て望ましい指示発話は、物体を特定するのに十分な情報が含まれている指示発話である。こうした指示発話を人間から引き出すためのロボットの確認行動について、研究課題 1 の結果から、次のようになる。

- ・ 語彙の引き込みの観点から、ロボットは物体を特定できる言語表現を使うべき。
- ・ 動作の引き込みの観点から、ロボットは指さしを使うべき。
- ・ 人間はロボットの提供する情報が冗長だと、自身の使用する言葉を減らす傾向があるため、ロボットは冗長な言語表現を避けるべき。
- ・ 情報が冗長だと、自身の使用する言葉を減らす傾向があるため、ロボットは物体を特定するのに意味のない指さしは避けるべきである。

### 確認行動生成アルゴリズム

前述の考察を踏まえ、指示されたと推定した物体の確認するため、ロボットはその物体を特定するのに十分かつ冗長ではない確認行動を出力する。確認行動は指さしと物体を特定する言葉を含む発話で構成される。

【指さしの有無】指さしの有無は、指さし動作が、確認対象の物体を、その近傍に存在する複数の物体からどれくらい特定することができるか、という観点で選択される。たとえば、確認対象の物体の近傍にたくさんの物体がある状況では、指さしをしても近傍の物体から確認対象の物体を識別するのは容易ではない。この場合、指さし動作は確認対象を特定するのに有用な情報とはなりえないので、ロボットは指さしを使う必要はない。逆に、確認対象の物体の近傍に物体がほとんどない状況では、指さしをすることによって、確認対象の物体をほぼ一意に特定することが可能である。この場合、指さし動作は確認対象を特定するのに有用な情報となるので、ロボットは指さしを行う。

【言葉の選択】確認発話で用いる言葉の選択について、ロボットは物体を特定できる最小の情報を用いる。具体的には、物体を特定するのに使える固有の情報、たとえば「名称」「色」「大きさ」「形」などの最小の組み合わせを用いる。

#### 有効性検証実験

【指示物体認識システム】提案した確認行動生成アルゴリズムによる指示物体認識性能向上の有効性を検証する実験を行うため、指示物体認識システムを構築した(図4)。人間が発話し、物体を指さすとき、発話認識モジュールが人間の発話からその物体を特定するための言語表現を抽出し、指さし認識モジュールは指さしジェスチャを検出し、その方向を計算する。各モジュールの結果は統合モジュールに集約され、各物体について指示されたもっともらしさ(物体指示尤度)を計算する。物体指示尤度の最も高かった物体のIDを確認行動生成モジュールに送り、確認行動生成アルゴリズムで生成された確認行動をロボットに送信することによって、確認行動が実行される。

#### ハードウェア

**ロボット:** 研究課題1でも利用した Robovie-R ver.2 を用いた。

**センサ:** 音声を取得するため、身体装着型の小型マイクロフォンを用いた。また指さし動作を検出するため距離画像センサを利用した。さらに、環境中の物体の位置を計測するため、Webカメラを用いた。位置検出は、物体に付与されたARマーカをWebカメラで検出することによって実現した。

#### ソフトウェア

システムのソフトウェアは4つのモジュールで構成される: 音声認識、指さし認識、統

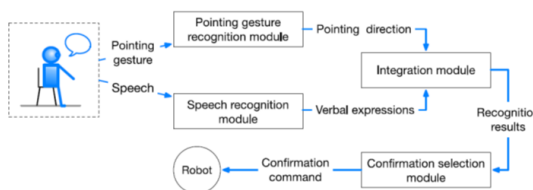


図4 システムアーキテクチャ

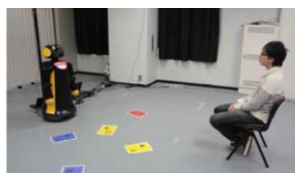


図5 実験の様子

合、そして確認生成である。最初に、音声と距離画像が音声認識モジュールと指さし認識モジュールに送信される。これらのモジュールはそれぞれの物体の指示尤度を計算する。各モジュールで計算された指示尤度は統合モジュールに送られ、最終的な指示尤度を計算し、指示尤度の最も高い物体のIDを確認行動生成モジュールに送信する。確認行動生成モジュールは確認行動生成アルゴリズムによって、環境情報から確認対象の物体の確認行動を生成し、ロボットに送信する。

【実験計画】開発したシステムの効果をいくつかのありうる状況において調査するため、ロボットの確認行動と環境における物体の配置という2つの要因を統制した。

#### 確認要因

確認要因は、被験者内要因で、確認行動生成アルゴリズムによる適切な情報量の確認発話をする(提案条件)か、すべての情報を用いた冗長な確認発話をする(冗長条件)か、を操作した。冗長条件では、ロボットは物体の持つすべての情報と指さしを使って確認動作を生成した。確認の発話の形式は次のような物体の属性の系列とした。

#### 【指示語】【図】【記号】【色】物体名

たとえば、ロボットは「あの丸い赤の本ですか?」や「あの三角とBの青い本ですか?」と確認した。

#### 配置要因

物体指示認識において、物体の配置はどんな種類の言語表現と指さし動作が選択されるかに影響を与える可能性があるため配置要因を設定した。具体的に、密集、2グループ、疎状態の3つの条件を設定した。

【被験者】8名の被験者(女性2名、男性6名、全員20代)が実験に参加した。

【環境】実験環境を図5に示す。各物体は、被験者からみて、0.6mから2.6mの距離に配置された。物体として異なる色と記号と文字の組み合わせが印刷されたB6判の本を用いた。色は赤と青と黄の3種類、記号は丸と三角と四角の3種類、文字はAとBの2種類であった。本は属性の全ての組み合わせを網羅するように18冊用意された。

【手順】最初に被験者は18種類の本の中からの自由に5冊を選び、配置要因に従い本を配置した。被験者は、ロボットの「本を指示してください」という発話の後、自由に本を1冊指示した。被験者が本を指示した後、ロボットは指示物体認識機能によって推定された本を、確認発話選択機能によって選ばれた確認発話を用いて確認した。被験者は、ロボットの確認に対して、「そうです」もしくは

は「違います」と答えた。確認の正否にかかわらず、ロボットは被験者に次の本の指示を促し、被験者は再び自由に本を1冊指示した。本を10回指示した後、被験者は本を戻し、そして再び5冊を自由に選び、実験者の指示に従って配置を変更し、再び10回の物体指示を行った。この対話は3つの配置に関して行われたので、最終的に合計30回の本の指示が行われた。どの配置条件から実験を始めるかについて、被験者間でカウンターバランスをとった。

【測定対象】本実験では次の項目を測定した。  
**認識性能**

人間が指示した物体を正しく認識した数を計算した。

**情報量**

物体の属性と物体を識別するのに有効な指さしを情報としてみなし、参照に含まれる情報を計算した。

#### 4. 研究成果

##### 研究課題1：物体指示対話における引き込み現象の発生を検証

物体指示対話における引き込み現象に関する仮説1と仮説2の検証実験の結果と示す。

###### 仮説1の検証実験

ターゲット語を含む指示発話の回数の平均と標準偏差を図6に示す。確認発話要因(対応なし)とセッション要因(対応あり)に関して、2要因の分散分析を行った結果、要因間の交互作用が認められた[F(1, 22) = 4.425, p < 0.05]。

交互作用が認められたため、各要因について下位検定を行った。その結果、セッション2水準における確認発話要因の単純主効果[F(1, 22) = 17.760, p < 0.01]が認められた。

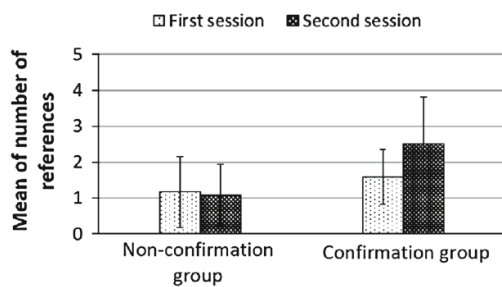


図6 ターゲット語を含む指示の数

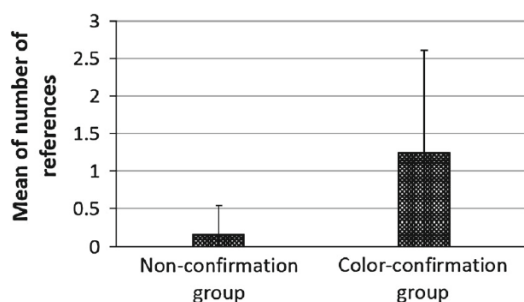


図7 色カテゴリーの単語を含む指示発話の数

表1 冗長条件(Excess)と提案条件(Decent)における認識性能

	Sparse set	Two groups	Congestion	Total
Excess	73.8%	62.5%	61.3%	65.8%
Decent	72.5%	63.8%	70.0%	68.8%

表2 冗長条件(Excess)と提案条件(Decent)における情報量

	Sparse set	Two groups	Congestion	Total
Excess	1.50 (0.596)	1.59 (0.478)	1.91 (0.670)	1.67 (0.581)
Decent	1.56 (0.452)	1.78 (0.526)	1.94 (0.693)	1.76 (0.557)

a. Means with standard error in brackets

これは、セッション2水準において、確認発話を聞いた被験者は、確認発話を聞かなかった被験者よりも、ターゲット語を含む指示発話を多く行ったことを示している。

さらに、確認発話あり水準におけるセッション要因の単純主効果[F(1, 22) = 7.436, p < 0.05]も認められた。確認発話あり水準において、確認発話を聞いた後のほうが、確認発話を聞く前よりも、ターゲット語を含む指示発話が多かったことを示している。

以上の結果は、人間はある物体を指示するとき、前にロボットがその物体を確認するときを使用した言葉を採用する傾向がある、という仮説1の妥当性を支持している。

###### 仮説2の検証実験

【結果】ターゲット語を含む指示発話の回数の平均と標準偏差を図7に示す。確認発話要因(対応なし)に関して、1要因の分散分析を行った結果、水準間で有意差が認められた[F(1, 22) = 6.477, p < 0.05]。これは、色カテゴリー確認発話水準の被験者は、確認発話なし水準の被験者よりも、ターゲット語を含む指示発話を多く行ったことを示している。

以上の結果は、人間は物体を指示するとき、それまでロボットが物体を確認するときを使用してきた言葉のカテゴリーに属する言葉を採用する傾向がある、という仮説2の妥当性を支持している。

##### 研究課題2：引き込み現象の知見を利用した指示物体認識システムの有効性検証

構築した指示物体認識システムにおいて、提案した確認発話生成アルゴリズムの有効性を検証するために行った実験結果について述べる。

【認識性能】表1に各条件での認識成功率を示す。認識成功率は平均的に提案条件(Decent)において高かったが、各要因に対してt検定を実施したところ、有意差は認められなかった。

【情報量】表2に各条件での情報量を示す。情報量はすべての配置条件について超過条件より適切条件において高かったが、各要因に対してt検定を実施したところ、有意差は認められなかった。

【考察】実験の結果として、認識性能と情報量ともに条件間での有意差は認められなかったが、提案確認行動を行う適切条件では、少しばかり精度の向上が認められた。この傾向を検証するためには、認識性能のレベルを上げる必要がある。

## 各研究課題の成果のまとめ

研究課題1では、人間同士や人間とコンピュータでの言語の同調傾向に関する従来研究を基に、人間とロボットの対面状況での物体指示対話において、言語の同調傾向が生じるかどうかを検証した。実験の結果、人間は物体を指示する時、前にロボットがその物体を確認する時に使用した言葉を採用する傾向があることと、人間は物体を指示する時、それまでロボットが物体を確認する時に使用してきた色のカテゴリに属する言葉を採用する傾向があることが明らかになった。この結果は、人間の語彙の多様性と指示発話の曖昧性がロボットの発話によって自然に解消される可能性を示している。

研究課題2では、人間から曖昧性のない言葉を引き出すためのロボットの確認行動の生成方法を提案し、その有効性の検証を行った。具体的には、人間に物体の識別に利用可能な情報が多く含まれる指示発話を行ってもらうために、ロボットが指示物体の識別に必要な最小の情報で構成される確認発話を行うという手法を提案した。この提案手法と、ロボットが指示物体の識別に利用できる情報の全てを含む確認発話を行う手法とで、指示物体認識システムの認識性能に変化がどうかどうかを検証した。その結果、指示物体認識の成功率は提案手法で68.8%、全ての情報を含む手法で65.8%、また、人間の指示に含まれる情報も提案手法で1.76、全ての情報を含む手法で1.67となり、ロボットの確認発話を利用して、人間から物体を特定するのに有用な情報を引き出し、指示物体認識性能を向上させられる可能性が示唆された。

以上の結果は、ロボットの指示物体認識技術に関する先駆的かつ実用的な研究であり、この成果はこの分野の発展に貢献をなすものである。

### <引用文献>

- ① Shinozawa, K. et al. User specification method and humanoid confirmation behavior. In Proceedings of International Conference on Humanoid Robots, 2, 558-563, 2007.
- ② 河原達也. 話し言葉による音声対話システム, 情報処理 45(10), 1027-1031, 2004.
- ③ 飯尾尊優. et al. 語彙の引き込み: ロボットは人間の語彙を引き込めるか?, 情報処理学会論文誌, 51(2), 277-289, 2010.

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

#### [雑誌論文] (計 1 件)

- ① Takamasa Iio, Masahiro Shiomi, Kazuhiko Shinozawa, Katsunori Shimohara, Mitsunori Miki, Norihiro Hagita, Lexical Entrainment in Human Robot Interaction, International

Journal of Social Robotics, 2014, Vol.7(2), pp.253-263, 2015.

#### [学会発表] (計 5 件)

- ① Masahiro Shiomi, Takamasa Iio, Koji Kamei, Chandraprakash Sharma, Norihiro Hagita, "User-friendly Autonomous Wheelchair for Elderly Care Using Ubiquitous Network Robot Platform" The 2nd Int. Conf. on Human Agent Interaction, 2014.
- ② Takamasa Iio, Masahiro Shiomi, Koji Kamei, Chandraprakash Sharma, Norihiro Hagita "Social Acceptance by Elderly People of a Fall-detection System with Range Sensors in a Nursing Home," The 2nd Int. Conf. on Human Agent Interaction, poster session, 2014.
- ③ 飯尾尊優, 塩見昌裕, 亀井剛次, Chandraprakash Sharma, 萩田紀博 "距離画像センサを用いた高齢者の見守りに関する意識調査" 日本ロボット学会第32回学術講演会講演論文集, RSJ2014, 2E2-02, 2014.
- ④ 塩見昌裕, 亀井剛次, Chandraprakash Sharma, 飯尾尊優, 宮下敬宏, 萩田紀博 "ユビキタスネットワークロボット・プラットフォームを用いた高齢者の見守り・移動支援サービスの実現" 日本ロボット学会第32回学術講演会講演論文集, RSJ2014, 2Q1-01, 2014
- ⑤ 塩見昌裕, 飯尾尊優, 亀井剛次, Chandraprakash Sharma, 萩田紀博 "高齢者を対象とした車いす型ロボットによる移動支援" 日本ロボット学会第32回学術講演会講演論文集, RSJ2014, 2E2-01, 2014.

#### [図書] (計 0 件)

#### [産業財産権]

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

#### [その他]

ホームページ等 なし

### 6. 研究組織

#### (1) 研究代表者

飯尾 尊優 (IIO, Takamasa)

株式会社国際電気通信基礎技術研究所・知能ロボティクス研究所・連携研究院  
研究者番号: 70642958

(2) 研究分担者 なし

(3) 連携研究者 なし