

乗り移るエージェントとスマートスピーカーの 親しみやすさの比較

Comparing Friendliness of Migration Agent and Smart Speaker

峯岸 七海^{1*} 内村 方哉² 大森 隆司³ 大澤 正彦³ 石井 健太郎¹

Nanami Minegishi¹, Masaya Uchimura², Takashi Omori³, Masahiko Osawa³, Kentaro Ishii¹

¹ 専修大学 ネットワーク情報学部

¹ School of Network and Information, Senshu University

² 日本大学 大学院 総合基礎科学研究科

² Graduate School of Integrated Basic Sciences, Nihon University

³ 日本大学 文理学部

³ College of Humanities and Sciences, Nihon University

Abstract: スマートスピーカーは音声で IoT 家電を操作できる。しかし人と人との会話で実現される関係性の構築はスマートスピーカーとの間では行われておらず、音声インターフェースを介しての機能の利用にとどまっている。本研究では、スマートスピーカーが親しみを持てるツールとして受け入れられることを目的として、デバイス間を乗り移るエージェントである ITACO を利用する効果を調査した。その結果、ITACO を利用することによって、スマートスピーカーは、より親しみが持てることが示唆された。

1 はじめに

スマートスピーカーは私たちの身近に存在する機器の一つであり、私たちの生活をサポートしてくれる便利な道具である。しかし、音声インターフェースを介しての機能を利用するにとどまっている。スマートスピーカーは住居内にある機器であるため、今よりも親しみやすさを感じられるようになると、ユーザーのより良い生活に繋がると考える。

市販のスマートスピーカーの利用に関しては、これまでも調査がなされてきた [1, 2, 3]。一方、スマートスピーカーをエージェント化することで、スマートスピーカーを改良しようという試みもある。川口らは、スマートスピーカーに指示を出す際に、「Alexa」や「OK Google」等のウェイクワードを使わず、注視によって指示待ち状態に移行するスマートスピーカーが提案した [4]。また、Bonfert らは、スマートスピーカーのディスプレイにバーチャルエージェントを表示することを提案した [5]。以上の研究は、スマートスピーカー本体を改良しようという試みである。

それに対して本研究では、周辺デバイスを巻き込む形で、より良いスマートスピーカーの実現を考える。そ

のアプローチとして、ITACO システム [6] の考え方を取り入れる。ITACO システムとは、周辺にあるデバイスにエージェントがアイデンティティを保ちつつ乗り移るような概念を持つシステムである。本研究では、スマートスピーカーに対するユーザの音声指示入力の後、スマートスピーカーからエージェントが操作対象の周辺デバイスに乗り移る映像音声表現を行い、操作対象デバイスを操作した後、再び元のスマートスピーカーに乗り移る映像音声表現を行うシステムを実装する。

本論文では、前述の操作対象デバイスへの乗り移りエージェントを用いたシステムと、通常のスマートスピーカーを模したシステムを比較する実験について述べる。実験参加者にそれぞれのシステムによる操作を体験してもらった後、質問紙調査とインタビュー調査を行い、親しみやすさに関する比較を行う。

2 関連研究

2.1 スマートスピーカーの利用に関する調査

Lee らは、スマートスピーカーを利用中の家族を対象に調査を行い、スマートスピーカーの継続利用が、同じスマートスピーカーを共同利用している家族間の調和を感じさせることを明らかにした [1]。一方、Voit らは、スマートスピーカーを利用したことがない家族や同居

*連絡先：専修大学ネットワーク情報学部
〒214-8580 神奈川県川崎市多摩区東三田 2-1-1
E-mail: ne201037@senshu-u.jp

世帯を対象に調査を行い、毎朝スマートスピーカーあいさつをする、期待通りに動作しなかった際にスマートスピーカーを責めるといった、スマートスピーカーを擬人化するような行動が現れることを明らかにした [2]。また、Bentley らは、Google Home の実際の利用ログをユーザに提供してもらう形で、音楽再生停止の指示が最も利用されているが、ヘビーユーザーはライトユーザーに比べて音楽再生停止の指示を利用していないことを明らかにした [3]。

2.2 機器のエージェント化による改良

川口らは、ウェイクワードではなく注視を用いることで、ユーザーとスマートスピーカーの対話を開始する手法を提案した [4]。ユーザーからの注視に対しては、スマートスピーカーも注視を返して、音声指示入力可能であることを示す。評価実験の結果、ウェイクワードを用いるよりも相互注視を用いる方が、システムの操作性とシステムへの関心が高まることが示された。

Bonfert らは、ディスプレイ付きのスマートスピーカーにバーチャルエージェントを表示することの効果を検査した [5]。調査は、エージェントを表示しない条件、人工的な見た目のエージェントを表示する条件、人間のような見た目のエージェントを表示する条件の比較にて行われた。調査実験の結果、定量的評価では有意な差は認められなかったが、インタビューによる定性的評価では人間のような見た目のエージェントを表示する条件を好む意見も得られている。

2.3 ITACO システム

ITACO システムとは、テーブルランプやタブレット等の様々なデバイスにエージェントが乗り移ったかのような表現をするシステムである [6]。もとななるアイデアは、ノート PC から自律ロボットへエージェントが乗り移って案内を継続する案内システムにて示されている [7]。ノート PC から自律ロボットへ案内を担当する機械が変わっても、前のデバイスで案内システムのユーザーが慣れたエージェントが乗り移って後のデバイスでも引き続き案内を行うことで、システムがコンテキストを理解していることを自然に示すことができる。

また、ITACO システムを利用して、人間の特性を理解する実験も行われている。Ono らは、実験中にロボットが現れて走行通路の「ごみ箱をどけてください」と依頼をした場合に、実験参加者がごみ箱をどけるかどうかを調査した [8]。この際に、事前に実験参加者とインタラクションを行っていたバーチャルエージェントがロボットに乗り移ってごみ箱をどける依頼を行った場

合には、乗り移らないでロボットがごみ箱をどける依頼を行った場合に比べて、より発言内容を聞き取ることができ、依頼に応じてごみ箱をどけることを明らかにした。小川らは、実験参加者と事前にインタラクションを行ったエージェントがテーブルランプを点けた後、実験参加者が第三者からの指示によってテーブルランプのスイッチを切るのかを調査した [6]。テーブルランプを点ける際に、エージェントが乗り移って点ける場合と乗り移らないで点ける場合を比較して、乗り移った場合の方が自分の手でキャラクターを消し去ってしまったと感じる度合いが高くなることを明らかにした。また、スイッチの切る際に躊躇があったかとの主観評価には有意差は認められなかったが、切ったあとに後悔を感じたかのようなフィードバックも得られている。Ogawa らは、「なんだか暑いな」という発言と共に手で顔を仰ぐような行動をとるロボットを用いた実験を行った [9]。この実験でも、事前に実験参加者と対話を行ったエージェントがロボットに乗り移る条件と乗り移らない条件で比較を行い、どちらの条件ともロボットの発言は理解できたが、乗り移る条件の実験参加者の方がより多くヒーターのスイッチを消した。この実験では、ヒーターを消してほしいと直接依頼していないにも関わらず実験参加者はヒーターを消しており、より自然な相互作用の実現でできていると言える。及川らは、非自然言語でコミュニケーションを行うロボットにおいても、適切な位置に乗り移りを行い発声することで、対話者の人間がコミュニケーションが成立したと感じ、ロボットの意図を感じ、ロボットの訴えかけに対して確信が持てたことを示唆する結果を示している [10]。

3 乗り移りによる機器の操作

本研究では、ITACO システムを応用して、スマートスピーカーから操作対象機器へのエージェントの乗り移りを実現する。このため、スマートスピーカーを模した球体の「拠点」を用意して、エージェントを表示できるような構成とした (図 1)。

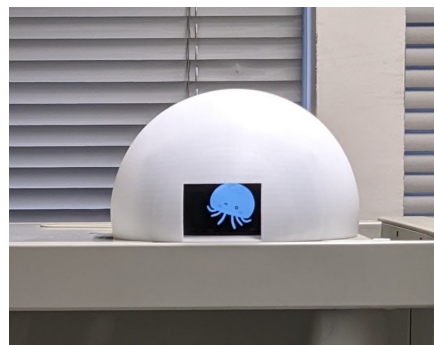


図 1: スマートスピーカーを模した「拠点」

ユーザーから見たシステムの動作は、音声指示を認識した後、拠点にいるエージェントがデバイスに乗り移り、デバイスを操作して、元の拠点に戻るというステップにて行われる（図2）。

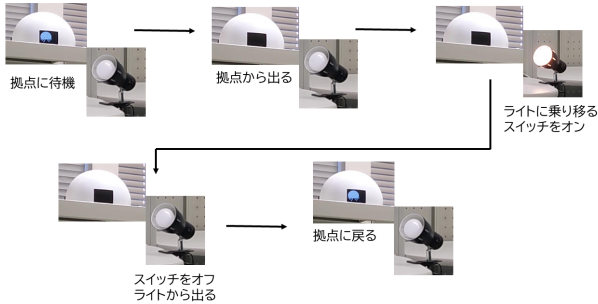


図 2: 乗り移り，操作して，拠点に戻る

一方，本論文の実装では，システムは乗り移りを表現する映像音声再生機能とデバイス操作機能を持つが，音声認識機能を持たない．そこで，ユーザーからの音声指示は，操作者が人手で操作をすることで，認識されたものとする．本論文の実装では，そのようなシステム操作者のための操作インタフェースも備えている．

3.1 システム構成

ITACO システムを応用した乗り移りエージェントの実装は，及川らの LEHU-IATCO のシステム [10] を参考にした．システムは，ITACO ユーザー，ITACO デバイス，ITACO アドミン，ITACO サーバーの 4 要素から成る．システムの構成を図 3 に示す．

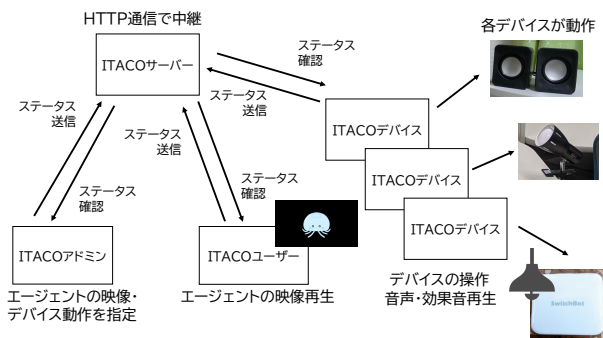


図 3: ITACO システムの構成

ITACO ユーザーは，ユーザー側のインタフェースを担当するソフトウェアで，エージェントの映像を映し出す．現在の実装では，クラゲを模した 2D キャラクターを用いた．ITACO ユーザーは React ライブラ

リを利用して TypeScript で実装されており，エージェントの映像はウェブブラウザ上で表示することができる．

ITACO デバイスは，エージェントの乗り移り先のデバイスの例えばライトをつけるような動作を実行するソフトウェアであり，Python で実装されている．1つのデバイスを操作するのに 1つの ITACO デバイスソフトウェアが必要となる．現在の実装では，デバイスの動作を実行する以外にも，エージェントからの自然言語によるユーザーへの呼びかけの音声や，デバイスや拠点に乗り移る際の効果音も，この ITACO デバイスで再生される．

ITACO アドミンは，エージェントの操作インタフェースを提供するソフトウェアである．現在の実装では，画面に表示されたボタンを操作者が押すことで，ITACO ユーザーの映像の切り替えや，ITACO デバイスの各種動作の実行を制御することができる．押したボタンの情報は HTTP リクエストとして後述の ITACO サーバーへ送られる．ITACO アドミンは React ライブラリを利用して TypeScript で実装されており，ウェブブラウザ上で操作ができる．

ITACO サーバーは，ITACO ユーザー，ITACO デバイス，ITACO アドミン間の通信を中継する HTTP サーバーである．ITACO ユーザー，ITACO デバイス，ITACO アドミンは一定周期ごとにお互いの状態を ITACO サーバーに問い合わせる．この周期的な問い合わせにより，ITACO ユーザーと ITACO デバイスは，操作者が ITACO アドミンのボタンで操作を行ったことを感知する．

実装上の制約により，ITACO アドミンの 1つのボタンで操作できる ITACO ユーザー・ITACO デバイスはいずれか 1つであるため，異なる ITACO ユーザー・ITACO デバイスを操作するには別のボタンを押す必要がある．例えば，エージェントが拠点から出ていく際には，ITACO ユーザーの映像の切り替えと ITACO デバイスからの効果音の再生をほぼ同時に行うことになるが，ITACO ユーザー用のボタンと ITACO デバイス用のボタンをそれぞれ押す必要がある．

現在の実装では，ITACO サーバーと ITACO アドミンは市販の PC 上で動作させ，ITACO ユーザーは市販のスマートフォンで動作させ，ITACO デバイスは Raspberry Pi 上で動作させた．

3.2 システム動作例

ここでは現在の実装における ITACO システムの操作と動作の例として，乗り移りエージェントがライトを点灯させる際の操作者の操作とシステムの動作を示す．この例では，操作者は以下の 4つの段階に分けて 6回のボタン操作が必要である（図4）．

1. 音声指示を受理したことをユーザーに示す「明かりをつけるよ」というエージェントの発話を再生するボタンを押す。
2. 拠点から出る映像を再生するボタンと拠点から出る効果音を再生するボタンをほぼ同時に押す。
3. ライトに入る効果音を再生する、ライトを点灯する、「明かりをつけたよ」という音声再生する、ライトから出る効果音を再生するという ITACO デバイスの一連の動作を実行するボタンを押す。
4. 拠点に入る効果音を再生するボタンと拠点に入る映像を再生するボタンをほぼ同時に押す。

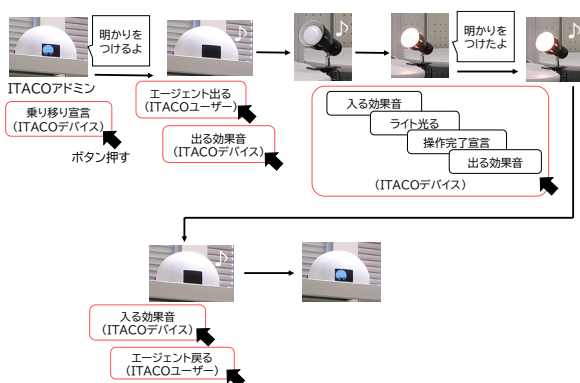


図 4: 乗り移りとデバイス操作の手順

4 評価実験

4.1 実験目的

既存のスマートスピーカーよりも ITACO システムを用いた乗り移りエージェントの方が、ユーザーがツールに対して親しみの気持ちを抱くことができるかを評価するための実験を行う。この目的のため、3 節で述べた乗り移りエージェントと、拠点をスマートスピーカーに見立てた乗り移り動作を行わないシステムを比較する。

実験において乗り移りエージェントが動作する条件を、ITACO あり条件と呼ぶこととする。一方、拠点をスマートスピーカーに見立てて乗り移り動作を行わない条件を、ITACO なし条件と呼ぶこととする。ITACO なし条件のシステム構成は、図 5 の通りである。ITACO なし条件では、拠点にエージェントは表示されず、したがって ITACO ユーザーは用いられない。また、拠点の位置からは音声再生されるので、その目的のためにスピーカーを繋いだ ITACO デバイスを用いるが、操作対象デバイスの位置から音声再生されることはない。

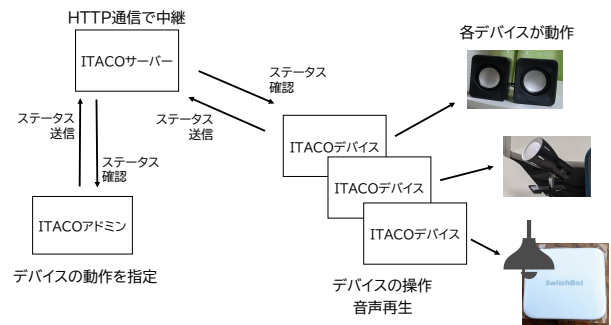


図 5: ITACO なし条件のシステム構成

4.2 実験方法

実験は、実験目的と乗り移りエージェントの説明、2 つの条件の体験と質問紙調査、全体を通したインタビュー調査の 3 つの段階に分かれて進行する。

実験目的と乗り移りエージェントの説明の段階では、2 つの条件の体験を評価してもらうことが実験目的であることを説明した後、乗り移りエージェントとはどのようなものかを実験参加者に知ってもらうため、次の 3 つのエージェントの動作を見てもらった。1 つ目は、エージェントが映し出された画面を触るとエージェントが喜ぶような動作である。これには、エージェントに慣れてもらう狙いがある。2 つ目は、エージェントが拠点からライトに乗り移り、ライトのオンオフを操作した後に拠点に戻る動作である。これには、乗り移りを理解してもらう狙いがある。3 つ目は、デバイスに隠れるエージェントを実験参加者が探し出すかくなればゲームである。これには、エージェントという存在と乗り移りの動作がどのようなものか実験参加者に伝える狙いがある。

2 つの条件の体験と質問紙調査の段階では、実験参加者に ITACO あり条件と ITACO なし条件の両方の動作を体験してもらう参加者内比較とし。奇数番目の参加者は ITACO あり条件から、偶数番目の参加者は ITACO なし条件からとする。それぞれの条件の体験では、一般的なスマートスピーカーのように、ユーザーからの音声指示に従いデバイスが動作するという流れを再現する。いずれの条件でも、実験参加者には、表 1 の音声指示を紙に書いて渡して指定する。実験参加者が音声指示を出すと、実験者は ITACO アドミンを操作してシステムを動かす。それぞれの条件の体験終了後、システムの印象を尋ねる質問紙調査を行う。質問紙調査は表 2 の 7 つの質問に対し、5 段階で評価してもらう形式にて行う。なお、Q2~Q6 の質問内容にはエージェント/システムと表記しているが、条件によって用語が異なることを意味する。ITACO あり条件では

エージェント、ITACOなし条件ではシステムという用語を用いる。

表 1: システムへの音声指示

指示順	指示内容
1	ライトの明かりをつけて
2	部屋の明かりをつけて
3	なにか音楽をかけて

表 2: 質問紙調査の質問内容

質問番号	質問内容
Q1	一緒に暮らしても良いと思ったか
Q2	エージェント/システムがライトの明かりをつけたとき、その行動に対して暖かさを感じたか、冷たさを感じたか
Q3	エージェント/システムが部屋の電気をつけたとき、その行動に対して暖かさを感じたか、冷たさを感じたか
Q4	エージェント/システムが音楽をかけたとき、その行動に対して暖かさを感じたか、冷たさを感じたか
Q5	エージェント/システムに親しみを感じたか
Q6	エージェント/システムに愛着を持てたか
Q7	もし一人暮らしをした時、このシステムがあれば、気分を明るく過ごせそうか

全体を通したインタビュー調査の段階では、両条件のシステムの印象についてのインタビュー調査を行う。

ITACOあり条件とITACOなし条件の動作の違いは5つある。第1はエージェントの映像の有無、第2は実験中の表1の1番目の指示での点灯の仕方、第3は表1の3番目の指示で再生する曲数、第4は実験参加者への呼びかけ時の声色とセリフ、第5はデバイス間の乗り移りの効果音の有無である。

4.3 実験結果

表2の質問内容の結果を表3に示す。5に近いほど質問に対して肯定的であることを意味する。また、各質問の結果に対してウィルコクソン符号順位検定を行った。

Q1は率直に自宅にあったら良いかを尋ね、有意差は見られなかった。Q2~Q4は実験で指示を出し、ITACOありとITACOなしの動きを見て感じたことを尋ねた。いずれも有意差が見られた。Q5~Q6は動作を全て見たうえで親しみや愛着は持てたか尋ねた。いずれも有意差が見られた。Q7は率直に寂しさを軽減し、明るく生活できそうか尋ね、有意差が見られた。

表 3: 質問紙調査の結果

質問番号	ITACO	1	2	3	4	5	平均	標準偏差
Q1	あり	1	0	5	8	2	3.63	0.93
	なし	0	4	7	3	2	3.16	0.95
Q2	あり	0	0	4	6	6	4.13	0.78
	なし	1	6	5	3	1	2.81	1.01
Q3	あり	0	0	4	5	7	4.19	0.81
	なし	0	6	5	4	1	3.00	0.94
Q4	あり	0	1	2	5	8	4.25	0.90
	なし	2	1	7	5	1	3.13	1.05
Q5	あり	0	2	0	7	7	4.19	0.95
	なし	6	2	4	4	0	2.38	1.22
Q6	あり	0	2	3	5	6	3.94	1.03
	なし	4	3	4	4	1	2.69	1.26
Q7	あり	0	1	2	7	6	4.13	0.86
	なし	1	2	8	3	2	3.19	1.01

5 考察

質問紙調査の結果とインタビュー調査の結果から、乗り移りエージェントの効果について考察する。Q2~Q4においてITACOあり条件が肯定的に見られた理由としては、インタビューでITACOあり条件に対して「エージェントの映像があることで仕事をしてくれたと感じる」、「エージェントの映像があることで存在を感じる」という意見が多かったことから、エージェントの映像があるだけで、ないときよりもエージェントがそこにいて、仕事をしてくれたと感じられるようになったと考える。ただし、Q4に関しては、ITACOあり条件での参加者への呼びかけ音声と再生する曲の数に好意的なインタビュー回答が多く、肯定的に見られる要因がはっきりしないため、追加で調査する必要がある。

Q5とQ6に関しては、ITACOあり条件に対して「エージェントの映像と乗り移り表現によって愛着が湧く」、「エージェントが可愛い」といったインタビュー回答があり、親しみや愛着の気持ちがポジティブに寄っているように考えられる。一方で、ITACOなし条件に対して「命令に従うだけ」という意見があり、ネガティブに寄っているように考えられる。

Q7でITACOあり条件が肯定的に見られた理由としては、ITACOあり条件に対して「エージェントの映像によって会話をしている感覚になる」、「ペットを飼っているようだ」、「部屋の中にもう一人いるようだ」というインタビュー回答があり、ITACOありは住居内にエージェントの存在を感じられることが一人暮らしでも明るく過ごせることに繋がるためと考える。

以上より、乗り移りエージェントを用いることで、既

存のスマートスピーカーよりも、スマートスピーカーに親しみを持つことができる可能性が示唆された。一方、その他のインタビュー回答として、ITACOあり条件における指示を出してからデバイスが動作するまでの時間が気になるという意見があった。親しみや愛着の感情を抱けたとしても、すぐにデバイスが動作しないことが気になるユーザーが存在する可能性が考えられる。

6 おわりに

本研究では、乗り移りエージェントを用いることで既存のスマートスピーカーよりも親しみが持てるようになるかを調査した。乗り移りエージェントの実現のために、エージェントがデバイスに乗り移る表現を行うITACOシステムを応用した。評価を行うにあたり、乗り移りエージェント用いたITACOあり条件と既存のスマートスピーカーに見立てたITACOなし条件を用意して、デバイス操作の指示を音声により行う実験を設計した。質問紙調査とインタビュー調査の結果、乗り移りエージェントを用いた方がユーザーがスマートスピーカーに親しみを持つことができる可能性が示唆された。

今後の課題は、ITACOあり条件とITACOなし条件の間で異なるように設定された5つのパラメータの効果をいくつかに分けて評価することである。例えば、動作による影響を確認するために、実験中のライトの点滅の仕方と再生する曲数をITACOあり条件とITACOなし条件で同じにし、その他は本実験と変えないという方法が考えられる。また、音による影響を確認するために、実験参加者への呼びかけの音声と効果音の有無をITACOあり条件とITACOなし条件で統一して実験を実施するという方法も考えられる。

参考文献

- [1] Lee, K., Lee, K.Y., Sheehan, L.: Hey Alexa! A Magic Spell of Social Glue?: Sharing a Smart Voice Assistant Speaker and Its Impact on Users' Perception of Group Harmony, *Information Systems Frontiers*, Vol. 22, pp. 563–583 (2020)
- [2] Voit, A., Niess, J., Eckerth, C., Ernst, M., Weingärtner, H., Woźniak, P.W.: 'It's not a romantic relationship': Stories of Adoption and Abandonment of Smart Speakers at Home, *International Conference on Mobile and Ubiquitous Multimedia*, pp. 71–82 (2020)
- [3] Bentley, F., Luvogt, C., Silverman, M., Wirasinghe, R., White, B., Lottridge, D.: Understanding the Long-Term Use of Smart Speaker Assistants, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No.3, Article No. 91 (2018)
- [4] 川口一画, 葛岡英明, ドナルドマクミラン: スマートスピーカーにおける注視の入出力を用いたインタラクションの効果, *ヒューマンインタフェース学会論文誌*, Vol. 21, No. 3, pp. 269–278 (2019)
- [5] Bonfert, M., Zargham, N., Saade, F., Porzel, R., Malaka, R.: An Evaluation of Visual Embodiment for Voice Assistants on Smart Displays, *Conference on Conversational User Interfaces*, Article No. 16 (2021)
- [6] 小川浩平, 小野哲雄: ITACO: メディア間を移動可能なエージェントによる遍在知の実現, *ヒューマンインタフェース学会論文誌*, Vol. 8, No. 3, pp. 373–380 (2006)
- [7] Imai, M., Ono, T., Etani, T.: Agent Migration: Communications between a Human and Robot, *IEEE International Conference on Systems, Man, and Cybernetics*, Vol. 4, pp. 1044–1048 (1999)
- [8] Ono, T., Imai, M.: Reading a Robot's Mind: A Model of Utterance Understanding based on the Theory of Mind Mechanism, *National Conference on Artificial Intelligence*, pp. 142–148 (2000)
- [9] Ogawa, K., Ono, T.: ITACO: Constructing an Emotional Relationship between Human and Robot, *International Symposium on Robot and Human Interactive Communication*, pp. 35–40 (2008)
- [10] 及川颯斗, 内村方哉, 小川裕太, 菊池華世, 福嶋稜, 板谷琴音, 大澤正彦: LEHU-ITACO: 非自然言語エージェントが乗り移ることで実現される意図伝達, P-9, HAI シンポジウム (2022)