

物体を擬人化するディスプレイロボットの開発

Development and evaluation of displaying robot for anthropomorphization

慶應義塾大学理工学研究科 大澤博隆 (Hirotaka Osawa)*

慶應義塾大学工学部 向井淳 (Jun Mukai)†

慶應義塾大学工学部 今井倫太 (Michita Imai)‡

概要

本研究では、身体部品を取り付けて物体を直接擬人化し、情報提示を行うディスプレイロボットを提案する。ディスプレイロボットは目や手などの身体部品を持ち、取り付けられた位置や、取り付け状況にあわせた擬人化を行う。これにより、ユーザに対し取り付け対象の身体イメージを想起させつつ、周囲の環境情報や、物体があらかじめ備えた機能を身体イメージを通して変換し、ユーザにとってより理解し易い形で伝えることが出来る。本論文では、現在までに行ったディスプレイロボットの実装および、実装したロボットを実際の展示場で展示した際の、ユーザの評価結果について述べ、物体擬人化手法の効用について検討する。

1 はじめに

近年の家庭内家電の高機能化、ネットワーク化に伴い、ユーザが身の回りの機器から受け取る情報は爆発的に増えつつある。たとえば、ネットワーク化された冷蔵庫では、RFIDのついた食品と連携して賞味期限をユーザに通知することが出来る。

従来、このように複雑な機器の各機能をユーザにわかりやすく伝えるため、デザインによる解決やインターフェース設計による解決などが行われて来た。しかしながら、複数の機器同士が連携したり、各機器が環境中のセンサと連携を行って情報を提示する

次世代の家電ネットワーク社会では、単一機器のインターフェース設計の変更だけで、爆発的に増加する情報に対処するのは限界がある。このため、各機器自体をの改良を行うのではなく、機器とユーザの間にエージェントを介し、ユーザとエージェントがインタラクションを行うことで、ユーザに情報を伝える方法が注目されている。

本研究ではエージェントによる情報伝達の一手法として、物体に人間の身体部品を取り付けて、物体の擬人化を行うディスプレイロボットを提案する。ディスプレイロボットの目標は、物体を擬人化することで、環境中に置かれた無線アドホックセンサやRFIDなどの外部情報取得デバイスから得られる情報を物体から見た相対的な情報に変換し、人間と環境中の物体との直観的なインタラクションを可能にすることである。

擬人化エージェントを利用して、人間へ指示を行う方法としては、従来では図1上のように、擬人的なCGエージェント [3] やロボット [1] が主に使用されてきた。しかしながらこの手法では、エージェント自体の身体イメージと物体の身体イメージが異なり、物体と別に擬人的なエージェントが存在する形になるため、物体とのインタラクション中に、ユーザの認識する対象が増える。その結果、図1上のようにユーザの興味対象が、説明対象よりも説明を行うエージェントに向かってしまい、円滑なインタラクションが進まない可能性が考えられる。一方、ディスプレイロボットを使用し、擬人的な目や腕を物体に取り付ける手法では、図1下のように物体が直接擬人化される。これにより、エージェント自体の身体イメージと物体の身体イメージが統一され、ユー

*osawa@ayu.ics.keio.ac.jp

†mukai@ayu.ics.keio.ac.jp

‡michita@ayu.ics.keio.ac.jp

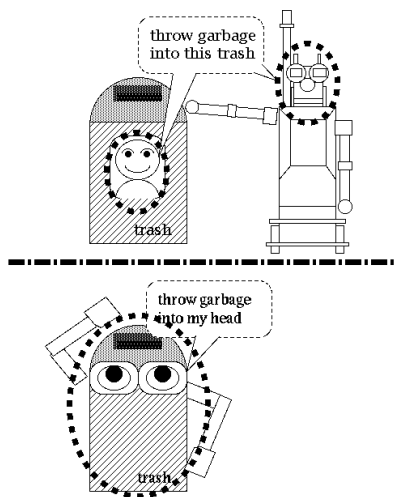


図 1: 擬人的なエージェントとディスプレイロボットの違い

ザの注意が他に向かうことがなくなると考えられる。また、従来手法と比較して、擬人化によって生じる「頭」や「腹」などの身体イメージを使用した会話が行えるようになり、従来より直観的な指示・インタラクションが可能になると考えられる。

ディスプレイロボットのうち、目を模したデバイス「Iris-board」を使用したディスプレイロボットに対する評価は、筆者らの以前の研究で行われている [4] [5] [6]。しかしながら、これら研究では、目を模したデバイスによる擬人化効果しか検討されておらず、複数の身体部品を取り付けたときの、具体的な身体イメージの変化が分からない。またこの実験の際には、被験者があらかじめ実験の内容を強く意識していたため、ユーザが実使用環境化でディスプレイロボットにどのような印象を受けるか、さだかではなかった。本研究では擬人化によって、ユーザがどのように物体の身体イメージを受け取るか調べるため、腕を模したデバイス「Arm-bot」を開発し、Iris-board と Arm-bot を物体に取り付け、大学の学内展示にて、目と腕を利用したユーザとのインタラクションを行った。その結果、ディスプレイロボットによる擬人化が、主に女性に受け入れられることを確認した。また、Iris-board や Arm-bot の取り付け位置を変化させることで、ユーザの受け取る身体イメージが変わることを確認した。

2 ディスプレイロボットの設計

2.1 理論的背景

Reeves らの Media Equation に関する研究 [7] では、人間は状況によって、ただの機器にパーソナリティを感じ、あたかも物体自体がコミュニケーション主体であるように反応し、物体を操作することが示されている。

ディスプレイロボットでは、この考え方を押し進め、物体に対し直接身体部品を取り付けることで物体に内在する身体イメージを拡張し、積極的に主体を持たせる。たとえば、洗濯機がディスプレイロボットによって擬人化されていれば、ユーザはその蓋を「口」として扱うことが出来るようになると考えられる。

このようにして得られる身体イメージを利用することで、従来物体がユーザに対し発信していた指示情報を変換することができる。たとえば洋服を洗濯機の口に入れて欲しいときには「その服を上蓋の部分に入れてね」と指示を行う代わりに「その服を僕の口に入れてね」と発話を行うことが可能となる。これによって、ユーザに対しより直観的な指示が行えるとともに、ユーザの物体に対する親近感を増すことが出来る。

2.2 システム構成

ディスプレイロボットのシステム図を、図 2 に示す。ディスプレイロボットはまず、取り付けられた座標を元に自らの身体イメージの座標を計算し、これを元に基本モーションや音声を決定する。ディスプレイロボットの主要な処理はシナリオサーバで行われる (図 2 中央)。シナリオサーバはディスプレイロボットを使用して行うシナリオを選択し、ユーザの顔座標、Iris-board や Arm-bot の取り付け位置から求めた身体イメージ座標を利用して、内部のシナリオにしたがって発話内容や目、腕のモーションを選択する。

2.3 ハードウェア構成

ディスプレイロボットのハードウェアは以下の部品より構成されている。

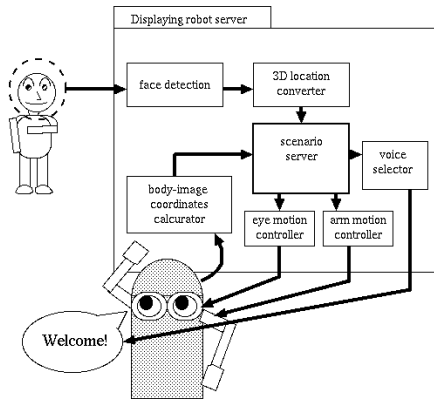


図 2: システム図

Iris-board は液晶を使用した仮想的な目であり、取り付けた 3 次元座標と向きベクトルを元に、瞳の表示位置を決定している。Iris-board の厚さは 2cm である。これにより、物体自体に埋め込まれた目が注視するかのように、任意の座標を指し示すことが可能である。

Arm-bot は片腕 6 軸の自由度を持った腕型デバイスであり、手の指先を動かした細かいモーションまで行うことができる。Arm-bot はサーボモータとアルミフレームで構成されているが、外装を布で隠すことで、物体に取り付けたときの違和感を少なくしている。

また、ユーザの顔の追跡には、フレーム速度の高い ieee1394 カメラ及び、haar-like feature [8] を利用した顔画像検索手法を用いた。これにより、毎秒 10 フレームで人間に追従して動く、自然なインタラクションが可能となる。

この他、ディスプレイロボットの発する声は、小型のスピーカから再生される。スピーカは隠して設置しており、ユーザはあたかも物体自身から声が聞こえるように感じる事が可能である。

3 調査実験

本章では、ディスプレイロボットの効果を検討するため、ディスプレイロボットを実際の機器に取り付け、ユーザとのインタラクションを行い、ディスプレイロボットをユーザがどう取扱うか、およびユー

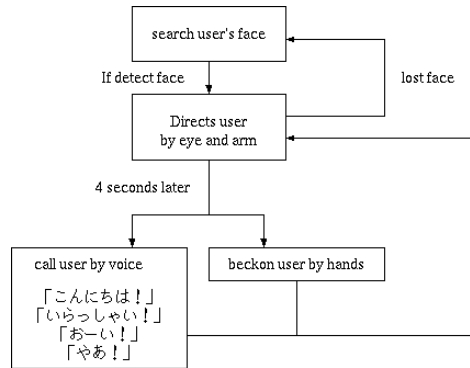


図 3: インタラクションの流れ

ザが物体の身体イメージをどう受け取るか、調査実験を行った。ユーザへのインタラクションシナリオとして今回は、ディスプレイロボットがユーザに呼びかけを行うことで、ディスプレイロボットを取り付けた対象へユーザを呼び寄せる、呼び込みタスクを選択した。

調査は、慶應義塾大学矢上キャンパスで行われた矢上祭(文化祭)の研究室展示にて行った。調査期間は 2 日間(10/14,10/15)合わせて 12 時間であり、初日が 12:15-17:15 の 5 時間、2 日目が 10:15-17:15 の 7 時間であった。

ディスプレイロボットによる、ユーザとのインタラクションの流れを図 3 に示す。まず、インタラクション前に対象となる物体に対し、目と腕となる部品、およびカメラとスピーカを取り付け、身体イメージの座標を初期化する。その後、まずカメラによって顔画像の検出を行い、検出が行われた場合には、その顔の大きさと位置から実世界上の顔の位置を求め、そちらに向けて目による注視、および腕の移動を行った。さらに、一定時間(4 秒間)検知が行われた際に、あらかじめ用意した「こんにちは」「いらっしゃい」「おーい」「やあ」という 4 つの発話からランダムに 1 語を選び、ユーザへの呼びかけを行うとともに、手先を動かして手招きを行った。これによって、ユーザを物体に呼び寄せるシナリオを実現した。

ディスプレイロボットの実装対象は、卓上サイズのゴミ箱(初日。図 4)、室内用運動器具のエアロバイク(2日目。図 5)であった。本研究では、目や腕、

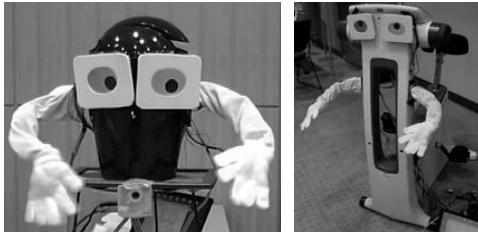


図 4: ゴミ箱への取り付け 図 5: エアロバイクへの取り付け図

カメラの取り付け位置はあらかじめ手動で入力した。

3.1 評価手法

ディスプレイロボットの評価のため、来場者に対し、展示の見学・解説を行ったあと、アンケートによる評価を行ってもらった。アンケートは、17 対の形容詞対を用いた 7 段階の数値評価 (表 1) とディスプレイロボットが注視・呼びかけを行った際の印象についての記述評価からなる。なお、来場者に対するアンケートは強制ではなく任意であった。

また、取り付け位置による身体イメージの変化を調べるため、アンケートでは、表 1 の形容詞対に加え各々体験してもらった物体の写真を見せ、来場者が「頭」「お腹」と感じるところにそれぞれ丸をつけてもらい、取り付け対象による身体イメージの変化を調べた。

4 結果

アンケートの有効回答数は、初日 17 人、2 日目 35 人の合計 52 人であった。このうち、男性は 31 人、女性は 16 人、性別の無回答者は 5 人であった。また、ユーザの年齢は 10 代毎に区切って来場者に解答してもらった。年齢の有効回答数は 46 人であった。アンケートに回答したユーザの年齢は、10 歳以下から 50 歳以上まで、幅広い年代にわたっていた。来場者のほとんどが、ロボットを見たことが無く、また全ての来場者にとってディスプレイロボットを見るのは初めての体験であった。

表 1: 各形容詞対

堅苦しい	うちとけた
柔軟な	こわばった
新しい	古い
こわい	やさしい
つまらない	面白い
冷たい	暖かい
近づきやすい	近づきたい
不愉快な	愉快な
陽気な	陰気な
愚かな	賢い
地味な	派手な
のろい	すばやい
わがままな	思いやりのある
単純な	複雑な
わかりにくい	わかりやすい
弱気な	強気な
気持ちいい	気持ち悪い

4.1 主成分分析によるによる社交性値結果

来場者のアンケート回答は任意であったため、形容詞対による評価を、各形容詞対単独で見えていくだけでは正確な評価が行えない。よって、まずこの結果に対して主成分分析を行って次元の削減を試みた。その結果、評価値が 1 以上となった軸が 6 つ存在した。結果を表 2 に示す。

このうち、情報をあらかずディスプレイロボットの評価にもっとも効果的である軸は、3 割弱の影響が見られる PC1 の社交性であると考えられる。よって、この社交性の軸を元に、各来場者の社交性値を求め、性別毎の、社交性値の平均値を求めた。その結果、男性の平均値が -0.377851 となり、女性の平均値が 0.433807 となった。結果を図 6 に示す。また、年齢回答があった各個人に対して、年代ごとの社交性値の変化を求めた。この結果を図 7 に示す。さらに、社交性を正と感じた来場者と負と感じた来場者に対し、見られたとき、呼びかけられたとき、それぞれの状況での有効な印象評価記述を分類した。分類結果は表 3,4 の通りである

4.2 身体図による身体イメージの比較

ディスプレイロボットを付けたそれぞれの物体 (ゴミ箱, エアロバイク) に対し、写真に丸が付けられた領域を、図 8 のそれぞれの区画に従って分類した。

表 2: 主成分分析結果

PC1: 社交性 (28.8%)	
暖かい←冷たい	0.793
うちとけた←堅苦しい	0.680
すばやい←のろい	0.657
派手な←地味な	0.613
賢い←愚かな	0.598
PC2: 新奇性 (11.26%)	
気持ちいい←気持ち悪い	0.611
新しい←古い	0.600
地味な←派手な	0.526
柔軟な←こわばった	0.451
PC3: 直観性 (8.30%)	
気持ちいい←気持ち悪い	0.458
わかりやすい←わかりにくい	0.443
こわい←やさしい	0.438
PC4: 単純さ (7.96%)	
わかりやすい←わかりにくい	0.490
単純な←複雑な	0.475
陽気な←陰気な	0.391
PC5: やわらかさ (7.06%)	
気持ちいい←気持ち悪い	0.480
やさしい←こわい	0.422
柔軟な←こわばった	0.404
PC6: 近付き易さ (6.40%)	
近づきやすい←近づきたい	0.679
わがままな←思いやりのある	0.353
地味な←派手な	0.321

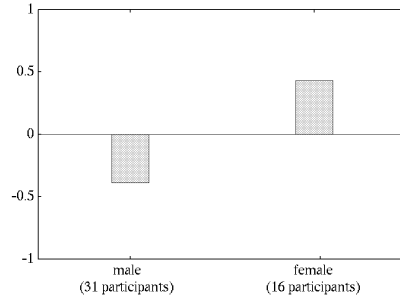


図 6: 性別による社交性値の差

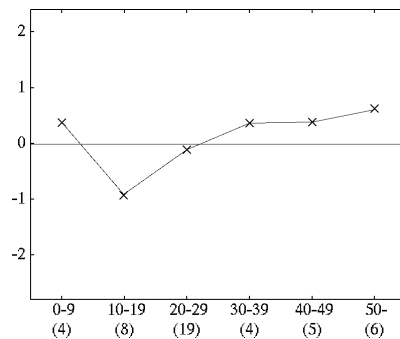


図 7: 年代による社交性値の差

「頭」と「お腹」の分類結果を図 9, 10 に示す。スラッシュ記号の前のアルファベットが「頭」と受け取られた領域を示し、スラッシュ記号の後ろのアルファベットが「お腹」と受け取られた領域を示す。

5 考察

5.1 性別による評価の差

ディスプレイロボットの取り付けられた物体に対する、男性、女性の来場者が評価した社交性値に対して、F 検定を行ったところ、分散が等しいことが仮定できたため、等分散を仮定した t 検定を行った。その結果、 $p = 0.070$ となり、男性と女性の評価した社交性値の違いに対して、有意傾向が見られた。

この結果より、女性の方が男性よりも、ディスプレイロボット及びそれを取り付けた物体に対し、好意的な印象を得ていることがわかる。アンケートの自由記述に注目してみると、女性の回答では、「かわいくてユニーク」など、ディスプレイロボットと

物体を一体化して見た直観的な回答が多いのに対し、男性の方が「構造が簡素」など各 부품の構造に注目した回答が多い(ディスプレイロボットと取り付け対象の物体が別であることを説明すると、驚く女性の来場者が多かった)。よって、上記の有意傾向が起きた原因として、女性の方がディスプレイロボットの取り付けられた物体をキャラクターとして捉えることが出来たのに対し、男性の方は、ディスプレイロボットの各機能に注目してしまい、その結果、ディスプレイロボットと取り付け対象物体を含めた身体イメージが受け取れず、各機能の不備さが目に付き、不気味に感じてしまったという理由が考えられる。

したがって今後は、男性に対し、ディスプレイロボットではなくディスプレイロボットと取り付け対象物体を包含して、身体イメージを感じてもらうためには、ディスプレイロボット各機能の精度を高めるとともに、男性にとっても不自然さを感じさせないシナリオを設計する必要があると考えられる。

表 3: 見られたときの評価

評価+	うれしい! ビックリしました かわいい とても不思議な気がした かわいい。なんかめがねをかけてるみたい とても面白い
評価-	きどった 恐怖 こわい ちゃんと見てるな 本当に目が合ったのか確信が持てなかった。 わからなかった。 目をそらしたら負けだと思ってる。 何されるのかと驚きました 今回の高さだと常に上目遣いで怖かった 見てる

表 4: 呼びかけられた時の評価

評価+	不思議なきもち 目がさめた 親しみを感じた ビックリ! 気持ちが良い かわいい。なんかめがねをかけてるみたい おどろく。でも、誰もいない時は安心する
評価-	おどろいた きもい。はい、と返事するかなくてわけわからない びっくりしました。タイミングが... チョット機械的。 困った。 どうしていいかわからなかった 聞き取りづらい 仰天した 話す機能があると思っていなかったので驚いた

5.2 年代による評価の差

ディスプレイロボットの取り付けられた物体に対する、実験者の年代ごとの社交性値は図7の通りである。これより、特に0代から10代にかけて大きく評価が下がり、そのあと緩やかに評価が良くなっているのが見て取れる。

この原因として、10代以下の年齢であれば、たとえ物体に目や腕がついていてもあまり抵抗感無く受け入れられるが、10歳を過ぎると、目と腕をつけた物体とのインタラクションに気恥ずかしさを感じるようになり、その結果ディスプレイロボットに対する社交性値が低く感じられるようになるのではないかと考えられる。展示中に観察したところ、10歳以下の子供は物怖じせずにロボットに向かい、腕や目

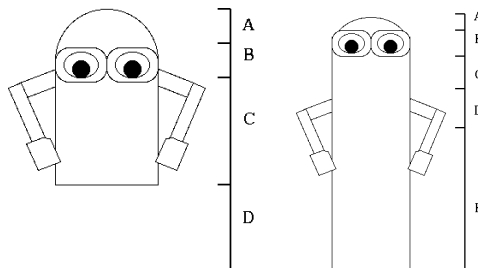


図 8: 身体イメージ分割図

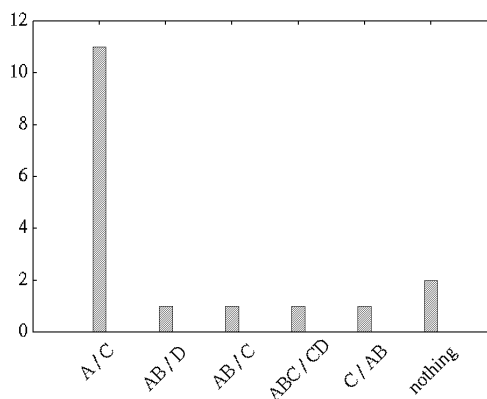


図 9: ゴミ箱に対する身体イメージ

を引っ張って積極的に遊ぶ子供が多かったが、10歳を過ぎた子供から20歳の大学生にかけては、主に展示を遠巻きに眺めていることが多かった。

また、30代から上にかけて、ディスプレイロボットに対する社交性の評価値が上がっていくが、これは年齢があがるにつれ、ディスプレイロボットとのインタラクションを客観的に観察することが出来、気恥ずかしさを感じなくなるからではないかと考えられる。また、特に30代より上の女性は、子供を連れた親子の来場者も多く、子供のポジティブな反応を見て、社交性値が上がった可能性も考えられる。

以上より、ディスプレイロボットは特に、10代から20代にかけて気恥ずかしさを感じさせる傾向があると考えられる。よって、今後は気恥ずかしさを感じさせず、インタラクションに集中させるようなディスプレイロボットの設計、およびシナリオの設計が必要になると考えられる。

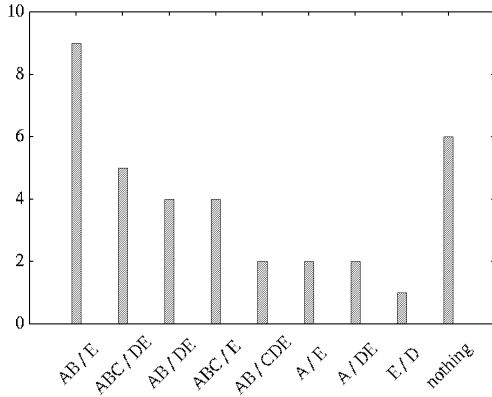


図 10: エアロバイクに対する身体イメージ

5.3 視線を受けたときの評価

社交性が負と感じたの意見には、目が合っておりそのため恐怖感を感じているものが多い。一方、社交性が正と感じた来場者は、目の印象そのものを高評価しているものが多い。したがって、顔認識の精度とカメラの検知幅を広げ、ディスプレイロボットの視線をより良くユーザに合わせる必要があると考えられる。

5.4 呼びかけを受けたときの評価

社交性が負と感じた来場者の中では、呼びかけられた後の反応に戸惑っている人が多い。これは、呼びかけを行った対象の意図が読み取れなかったためと思われる。今回の調査では呼びかけを行うことで、来場者を物体の近くに呼び寄せることを期待していたが、取り付け対象であったゴミ箱やエアロバイクでは、手招きや呼びかけを行っても、来場者が意図を適切に読み取れなかったものと思われる。よって、今後は機器の形状や機能を生かし、ゴミ箱であればごみを捨てるような発話、エアロバイクであれば運動を促す対話など、呼びかけ対象の意図をはっきりさせるようなインタラクションデザインが必要であると考えられる。一方で、社交性が正と感じた来場者の意見には、かわいい、きもちが良い、の他安心する、というものもあった。これは、ディスプレイロボットによって来場者がより多く存在感を感じる

ことが出来、それが好評価につながったものと考えられる。

5.5 身体イメージの評価

ディスプレイロボットをゴミ箱に取り付けた状況では、正面から見たときに目の付いている場所と、腕の生えている場所の垂直位置が同じになる(図8左のB)。この状況下では、ほとんどの来場者が図9のように、腕と目がついた垂直位置よりも上の領域(図8左のA)を「頭」と受け取り、腕と目がついた垂直位置よりも下の領域(図8左のC)を「お腹」と受け取っている。

一方、ディスプレイロボットをエアロバイクに取り付けた状況では、正面から見たときに目の付いている場所(図8右のB)と、腕の生えている場所(図8右のD)の垂直位置が異なる。図9より、この状況下ではほとんどの来場者が目をふくめた図8右のAとBの領域を「頭」と捉えている。これは、目デバイスを取り付けた領域が含まれるため、ゴミ箱の場合とは異なる。また、「お腹」の身体イメージは、図8右のCからEまでと受け取る者と、図8右のDからEまでと受け取る者の2種類に分かれている。

このような変化がおこった原因として、目と腕を近づけて取り付けることで、その部分が頭やお腹の一部と受け取りづらくなり、ユーザがその領域を明確な境界として、頭やお腹のどちらかに組み入れることなく、物体の上部を頭、下部をお腹と知覚した可能性が考えられる。

一方、エアロバイクでは、目と腕の取り付け位置が離れている。これによって、目や腕をそれぞれ、頭やお腹の一部として受け取り易くなり、その結果、多数のユーザが目を含めた部分を頭と受け取り、また半数のユーザが、腕を含んだ部分をお腹と受け取ったと考えられる。

以上より、ディスプレイロボットを使って、身体イメージを物体に付与する場合に、ユーザが設計者の意図を間違えず、頭とお腹を認識する必要があるれば、目と腕の取り付け位置を近づけるのが良いと考えられる。

ただし、本研究ではゴミ箱とエアロバイクへの取り付け例の、2例しか扱っていない。今回の取り付け対象となったゴミ箱とエアロバイクは、それぞれ物体が連続しており、物体がいくつかの部分に分かれ

ていない。物体がいくつかの部分から構成される場合には、身体イメージがそれらの部分から影響を受けることが筆者らの過去の研究で示されている [5]。よって、ディスプレイロボットによって得られる身体イメージのモデルを作るためには、さらなる調査が必要であると考えられる。

6 結論

本研究では物体に人間の身体部品を取り付けて、物体の擬人化を行うディスプレイロボットを提案した。本研究ではディスプレイロボットを開発し、物体に取り付けユーザとのインタラクションを行うことで、ディスプレイロボットによる擬人化という手法が、どの性別、どの世代に受け入れられるかを調査した。その結果、ディスプレイロボットは、主に女性にとって評価が高く、また年代別に見ると、思春期の時代を除いて、各世代間にまんべんなく好印象をもって受け入れられていることが確認できた。また本研究では、取り付け部位による身体イメージの変化に付いても、調査を行った。その結果、目と腕の部品を近接させて取り付けることで、身体イメージがより強固に形作られることがわかった。

今後は、ディスプレイロボットの印象を高めるため、個々の調整を行うとともに、調査結果として得られた身体イメージを利用して、物体とユーザ間でのどのようなインタラクションが可能であるか、それぞれ検討していく必要があると思われる。

7 謝辞

本研究で使用したディスプレイロボットは、独立行政法人情報処理推進機構未踏ソフトウェア創造事業「未踏ユース 2006」の援助を受けて開発されました。ここに感謝の意を表します。

参考文献

[1] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and R. Nakatsu, "Development and evaluation of an interactive humanoid robot "Robovie."," in *Proceedings of IEEE International Conference*

on Robotics and Automation (ICRA2002), Washington, DC, USA, May 2002, pp. 4166–4173.

- [2] H. Kozima, C. Nakagawa, and H. Yano, "Attention coupling as a prerequisite for social interaction," in *IEEE International Workshop on Robot and Human Interactive Communication*, San Francisco, CA, USA, Oct. 2003, pp. 109–114.
- [3] N. Mukawa, A. Fukayama, T. Ohno, M. Sawaki, and N. Hagita, "Gaze Communication between Human and Anthropomorphic Agent," in *Proceedings of IEEE International Symposium on Robot and Human Interactive Communication*, Bordeaux-Paris, France, Sept. 2001, pp. 366–370.
- [4] H. Osawa, J. Mukai, and M. Imai, "Anthropomorphization of an object by displaying robot," in *Proceedings of IEEE International Symposium on Robot and Human Interactive Communication*, Hatfield, United Kingdom, Sept. 2006.
- [5] —, "Acquisition of Body Image by Anthropomorphization Framework," in *Joint 3rd International Conference on Soft Computing and Intelligent Systems and 7th International Symposium on advanced intelligent Systems*, Tokyo, Japan, Sept. 2006.
- [6] —, "Acquisition of subjective representation using body parts by an object," in *IEEE Workshop on International Conference on Cognitive Science 2006*, Vancouver, Canada, July 2006.
- [7] B. Reeves and C. Nass, *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Univ. of Chicago Press, 1996.
- [8] P. Viola, and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," in *Proc. of CVPR*, 2001.