

# 訪問客のエンゲージメントに基づいた接客ロボットの社会的応答

## Social Responses of a Robotic Salesperson Based on the Engagement of Visitors

周 剣<sup>\*1</sup>  
Jian ZHOU

岩崎 雅矢<sup>\*1</sup>  
Masaya IWASAKI

金 延儒<sup>\*1</sup>  
Yen-Ru CHIN

ミシェル ガイレンベルグ<sup>\*1</sup>  
Michelle GEILENBERG

池田 瑞<sup>\*1</sup>  
Mizuki IKEDA

河村 竜幸<sup>\*2</sup>  
Tatsuyuki KAWAMURA

中西 英之<sup>\*1</sup>  
Hideyuki NAKANISHI

<sup>\*1</sup> 大阪大学大学院工学研究科 知能・機能創成工学専攻    <sup>\*2</sup> 京都イノベーション株式会社  
Department of Adaptive Machine Systems, Osaka University    Kyoto Innovation, Inc.

In recent years, customer service robots have been rapidly deployed. However, it is still unclear what kind of behavior is effective for sale. A conversation cannot be initiated, if the customers are not interested in the robot. Therefore, in this paper, we focus on engagement in the field of Human-Robot Interaction (HRI). In order to produce socially intelligent responses, we introduce a machine learning method to calculate a degree of engagement by analyzing the data which we conducted a field experiment in a real shop. As a result, we develop a model that the robot can hence decide on a socially intelligent response in the start of the conversation.

### 1. はじめに

近年、店舗や博物館、学校などの公共の場所においてヒューマノイドロボットが頻繁に用いられている[Stock 17][Niemelä 17]. また、博物館を通して訪問者を案内するために予め定義されたシーケンスを実行するというような研究も存在している[Kuno 07][Shiomi 06]. しかし、そのほとんどの場合、ロボットが行う動作は単純なものに限られている。たとえば、ロボットは、タブレットなどの画面に何度も同じ情報を表示することや、何度も同じ動作を繰り返すことによって、訪問者の注目を集めようとする。このようなロボットは主に情報を提供するために使用されているが、人々と社会的に交流することはないため注意を引くことができず、それらの効果は低いままである。これらはロボットの機能の欠如が原因となっているのではなく、ロボットの社会性の欠如が原因となり引き起こされている。したがって、ロボットは人々からの注目の度合いを理解し、ある一定以上の興味がある時に話しかける必要がある。このような他の参加者と同じ場所に留まり、インタラクションを続けることに起因する値としてエンゲージメントがある[Christopher 05]. そこで我々の研究では接客ロボットが最初に訪問客に対して発言を行う際に、その発言の前に会話の対象となる人のロボットに対するエンゲージメントの程度を分析し、会話を開始するのに十分なエンゲージメントの程度になった時にロボットが呼びかけを行うための反応モデルを設計することを目的とする。

本研究では、実際の店舗での実験によるデータを用いるということに重点を置く。そこで、SoftBank Robotics のヒューマノイドロボット Pepper を用いて京都にある実際の店舗において実験を行った。実験中の訪問客を観察し、得られたデータに基づいて反応モデルを設計する。本稿は、まず実際の店舗における実験方法について説明した後、エンゲージメントを測定するための適切な訪問客の非言語的な要因を選択し、Pepper によってこれらを分析する方法を説明する。そして、その要因の評価とエンゲージメントの計算方法、ロボットの社会的応答モデルの設計について記述する。

### 2. 関連研究

エンゲージメントの定義とモデルは複数存在している。Christopher によると、エンゲージメントは、あるインタラクションにおける参加者が他の参加者と同じ場所に留まり、インタラクションを継続するという目的に帰着する値として定義される[Christopher 05]. また、Sidner らは直接エンゲージメントを注意に関連付けた[Sidner 05]. 彼らは、エンゲージメントは認識された結びつきを確立し、維持し、終了させるプロセスであると述べた。さらに、エンゲージメントは注意だけではなく、それ以外の要因の影響をも受けると考えられてきた。他のアプローチとして、Corrigan らは、3 つの異なる種類のエンゲージメントを定義している[Corrigan 13]. すなわち、ソーシャルエンゲージメント、タスクエンゲージメント、ソーシャル-タスクエンゲージメントである。また、Salam らは、エンゲージメントはインタラクションの前後関係に応じて定義する必要があると考えた[Salam 15]. そして、会話の状況や話題に基づいて、エンゲージメントが分岐すると定義した。これらの研究が行われてきたが、エンゲージメントに関する明確な定義はまだ存在していない。しかし、これらから得られる有益な情報は多い。

また、実験室実験では実験室にいるということが人とロボット間のインタラクションに影響を与えるということが考えられる。例えば、実験室では、カメラなどで録画されているという意識が



図1 ロボットと握手をする訪問客

強いため、ロボットに対して失礼な行動などはとらない人が多いが実際のフィールドではそのようなことも観察される可能性がある。このように、実験室実験では観察することができないロボットとの自然なインタラクションを観察するため、実際の博物館での実験[川口 09]や実際の教室での実験[神田 04]が行われている。接客ロボットに関しても様々なその目的に応じた実際のフィールドにおける実験が行われてきた[宮下 08]。しかし、接客ロボットが人のエンゲージメントの程度を分析し、それに応じて行動するような研究はされていない。

### 3. 実際の店舗における実験

実験中の様子を図 1 に示す。本実験では接客ロボットとして Pepper を用いた。Pepper は、2014 年にソフトバンクモバイル株式会社が開発した人型パーソナルロボットである。Pepper を採用した理由の 1 つ目は Pepper の大きさが、設置しても商品を購入することの妨げにならず、店の外からも目立程度であり、客とコミュニケーションするにはロボット搭載のタブレットを用いて商品の説明や紹介ができるからである。2 つ目の理由は Pepper には赤外線センサ、ソナーセンサなどの多くのセンサを用いた安全装置が搭載されており、非常に安全であるからである。実験をする際、Pepper は周辺にいる人との衝突を回避する機能が備わっている。したがって、至近距離に人がいる場合、手振り身振りの動作における腕の動きが自動的に停止または低速化するようになっている。3 つ目の理由はその頭を動かすことによって視線を制御することができ、その腕を使ってジェスチャーとポインタを行うことができるからである。一般販売向けの Pepper は日本語と英語を話すことができ、実験では接客にこの 2 ヶ国語を用いた。

また、今回は京都の商店街にある七味専門店「ちんくら」において実験を行った。この店舗の見取り図を図 2 に示す。この店舗を入口側から見た図を図 3 に示す。実験中、訪問客の顔がわからないようにするためピントをずらした 3 つの不鮮明カメラを使用した。1 つ目はロボットの正面側から映したもの、2 つ目はロボットの背面側から映したもの、3 つ目は全体の様子を側面側から映したものである。また、訪問客がロボットに近づくとすぐにロボットに搭載したタブレット端末にタッチ操作によって回答可能な同意書を表示し、訪問客から撮影の同意が与えられた場合のみ、ロボット側面に設置した鮮明カメラによって、顔を認識可能な映像を取得した。さらに、実験者が実験中にロボットの遠隔操作を行う際店内の様子を観察するために、ビデオ通話をつないだタブレットを店内のロボットの後方に設置した。

ロボットの操作はスマートフォンを使って Pepper を遠隔操作できるアプリケーションを開発し、実験者が Wizard of Oz 法[Kelley 84]を使用して行った。Wizard of Oz 法とは、対話システムの開発において効果的なシミュレーション手法である。システムのふりをした人がシステムのかわりにユーザと会話をする。この手法

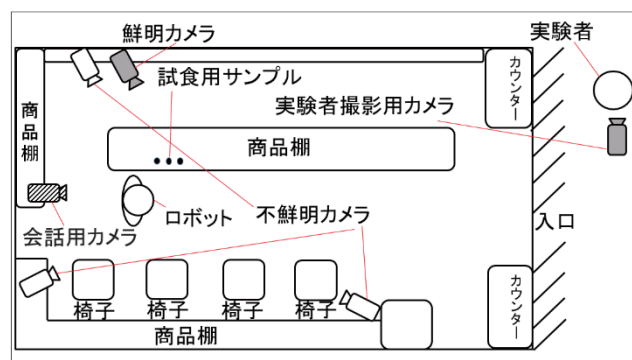


図 2 店舗の見取り図



図 3 店舗を入口側から見た図



図 4 ロボット操作ユーザインタフェース

はシステムに必要な要求や起こりうる問題をあらかじめ知ることができるという利点がある。また、遠隔操作アプリケーションはスマートフォンを Pepper と同じ無線 LAN 内に接続していれば、Pepper にあらかじめインストールされた行動をとり、入力した言葉を話させることができる。図 4 はそのアプリケーションを示したものである。実験中、Pepper を操作する際にすべての言葉を入力すると、時間がかかり沈黙ができてしまい会話が成り立たなくなる。そこでよく使う一連の言動をあらかじめ Pepper にインストールした。それらの行動は、Pepper のアプリケーションを開発できるソフトウェアである Choregraphe を用いて開発した。実験者は店舗の前にある道の反対側立ち、ロボットと訪問客との間のインタラクションを目視とビデオ通話の映像の両方を用いることによって確認し、訪問客とのインタラクションにおいて適切なロボットの行動を選択した。

実験は 2017 年 4 月 10 日 15:15-16:22, 13 日 15:21-16:22, 14 日 16:01-17:21, 2017 年 8 月 14 日 16:21-17:58, 15 日 15:20-16:22, 16 日 14:15-15:58 の 2017 年 11 月 27 日 16:08-17:34, 12 月 02 日 14:07-16:58, 04 日 14:05-17:13, 05 日 14:13-16:45 の 10 日間行った。今回の実験では合計 164 組がこの店舗を訪れた。このうち 76 組がロボットと会話をを行った。1 組の人数の平均は 2.2 人であり、約 370 人の訪問客が実験中この店舗を訪れた。



4. エンゲージメントに起因する非言語要因の選択

広く認められた明確なエンゲージメントの定義はまだ提案されていないが、エンゲージメントをモデル化するために複数の先行研究で使用される非言語要因を収集することは可能である。前述したように、エンゲージメントにはさまざまな要因が含まれている。1つの一般的な要因は、アイコンタクトである。これは、会話相手に対する注目度における強いサインである[Castellano 14] [Knapp 13]。アイコンタクトや視線の他にも、要因を選択することができる。アイコンタクトをすることで、顔の表情を分析することもできる。また、ロボットに対する反応、例えばうなずきや姿勢なども重要な情報を与えることができる[Knapp 13][Yanco 12]。さらに、ロボットと訪問客の間の距離もエンゲージメントにとって重要な要因である[Michalowski 06] [神田 08]。

今回実験で用いた Pepper は、店舗や家庭での日常生活における人型ロボットコンパニオンとして設計されたものである。このため、レーザーセンサ、超音波ソナーセンサ、タッチセンサ、2つの 2D カメラと 1 つの 3D カメラが搭載されており、位置や行動、声を検知することが可能である。また、それらだけでなく、幸福感、怒り、恐怖感などの人間の感情をも認識することが可能である。実際の店舗において訪問客を分析するためには、その訪問客を区別し、個々に関するデータを収集する必要がある。これを行うために、Pepper の顔認識機能を使用する。これはそれぞれの People ID によって固有の顔の特性が保存される。まず人について収集されたすべての情報を、People ID とともにロボットのメモリに保存する。そして、そこから People ID に対応する特定の要因の情報を取得し、特定の状況が観察された場合のみ特定の行動をロボットが行うように設計する。これらより、我々はエンゲージメントの程度を計算するモデルを設計するため、Pepper が提供するツールを用いて分析できるアイコンタクト、アイコンタクト継続時間、距離、接近、笑顔という合計 5 つの非言語要因を選択した。次の章でこれらの非言語要因の選択の妥当性を検証するため、エンゲージメントと各要因の相関を計算する。

5. データ分析と考察

実際の店舗における実験のビデオを使用してデータを収集した。この実験の際に 164 組の訪問客が店を訪れた。今回はロボットから発言する場合を考えるため、我々は、ロボットが最初に発言を行った 153 人の訪問客について、訪問客が来店してからロボットが訪問客に対して最初の発言を行うまでの訪問客の行動をデータとして用いて統計を行う。

アイコンタクトは X1 と表すバイナリデータである。例えば、ロボットが訪問客に挨拶したときに訪問客がロボットを見ていた場合、X1 の値は 1 となる。アイコンタクト継続時間は X2 として表す線形データであり、入店から挨拶までのロボットと訪問客がアイコンタクトを行った秒数である。距離は、X3 として表す線形データである。ここでは、ロボットから店舗の入り口までの範囲を 5 つに分割し、ロボットが挨拶をしたときに訪問客がどの範囲に位置するかを距離として定義する。例えば、訪問客がロボットの横

表 1 訪問客の各データの例

訪問客のId	Y	X1	X2	X3	X4	X5
41001	1	1	15	3	1	1
41002	1	1	5	1	1	0
41003	1	1	20	1	0	1
41004	1	1	6	1	1	0
41005	0	0	0	1	0	0

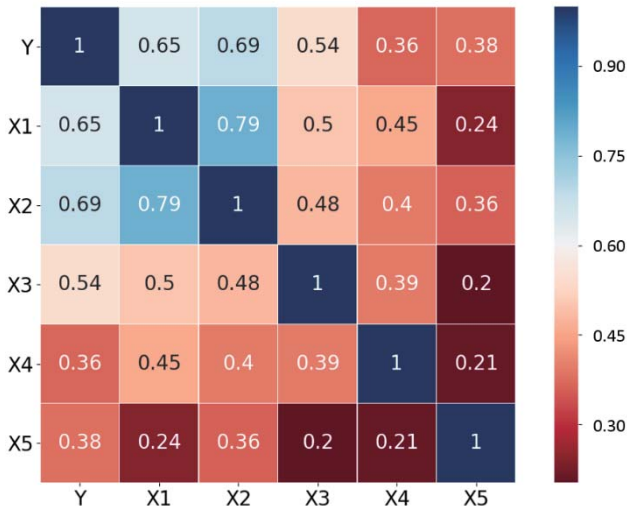


図 5 Y と各 X の相関係数 (Y:エンゲージメント, X1:アイコンタクト, X2:アイコンタクト継続時間, X3:距離, X4:接近, X5:笑顔)

に立っている場合、距離は 0 であり、訪問客が入り口に立っている場合、距離は 4 である。接近の定義は、ロボットが初めて挨拶をしたときに訪問客がロボットに接近するかどうかであり、X4 と表すバイナリデータである。訪問客がロボットに接近した場合、X4 の値は 1 である。笑顔は X5 と表すバイナリデータである。これは、ロボットが挨拶をしているときに訪問客が笑っているかどうかとして定義した。ロボットが挨拶をしているとき、訪問客が笑っている場合、X5 の値は 1 である。最後に、訪問客がロボットとコミュニケーションを開始したかどうかを Y としてバイナリデータで表す。訪問客がロボットと会話を開始した場合、Y の値は 1 である。以上のパラメータを用いて、表 1 に例として 5 つのグループ訪問客のデータを示す。

図 5 はそれぞれの変数の間の相関係数を示しており、横軸と縦軸が各変数を表している。相関係数は-1 から 1 の間の値になり、1 に近い方が強い正の相関があることを示し、-1 に近い方が強い負の相関があることを示す。これらを見ると、全ての変数において負の相関がなく、互いにある程度の正の相関があることが分かる。したがって、エンゲージメントに起因する非言語要因としてのこれらの選択は妥当であると言える。また、アイコンタクトとアイコンタクト継続時間がエンゲージメントにとって他の要素に比べて、より重要な要因であることが確認できる。

6. 社会的反応モデルの開発

データを分析するため、今回は簡単のため機械学習手法、ロジスティック回帰を使用する。式 5.1 および式 5.2 はロジスティック回帰における数式を示している。

$$Y = \frac{1}{1+e^{-z}}$$
 (5.1)

$$Z = W^T X + b$$
 (5.2)

W が重みであり、b はバイアスである。コスト関数を最小化することによって、最適化された重みを得ることができる。表 2 は最適化した重みのパラメータを示している。また、バイアスのパラメータは-0.19654401 であった。

このデータに用いた 153 人の訪問客に対して 5 分割交差検定を行った。5 分割交差検定では、データを 5 つに分割し、訓練 4、テスト 1 の比率で行った。交差検定の結果、88.9%の正確率で訪問客のエンゲージメントを測定できた。

表 2 重みの最適化パラメータ

要因	重み
アイコンタクト	0.53
アイコンタクトの継続時間	1.35
距離	0.99
接近	-0.17
笑顔	1.20

したがって、これらのパラメータを使用して、ロボットの社会的反応モデルを開発することができる。このモデルをロボットにインストールし、タイマーを設定してモデルを何度も繰り返し使用することで、ロボットは訪問客のエンゲージメントをリアルタイムに測定することができ、ある一定以上のエンゲージメントを持つ訪問客に対して、呼びかけるなどの接客方法を選択することができる。今回はロボットが最初に発言する場合のみを考えたため、訪問客から話しかけられる場合を考慮していない。実際の店舗においてこのモデルを使用する際にはこの訪問客が最初に発言を行う場合も考慮しなければならない。しかし、最初に訪問客が発言した場合は、ロボットはその発言に対して返答すればよい。このモデルと組み合わせることは十分可能であると考えられる。

今後の研究では、このモデルを実店舗で試験し、再現性を調査しなければならない。そして、より多くの訪問客のデータを用いて評価し、また他の要因についても調査する必要がある。

## 7. おわりに

今回は接客ロボットが訪問客に最初に挨拶などの発言をする際、訪問客のエンゲージメントに基づいて社会的に発言するための反応モデルを設計することを目的とした。そして、実際の店舗における訪問客を観察することによって得られたデータに基づき、ロジスティック回帰を用いてロボットの社会的反応モデルを開発した。このモデルは実際の店舗における実験データにおいて高い正解率を持っており、このモデルを用いれば、ロボットは会話を開始するのに十分なレベルのエンゲージメントを持つ訪問客に対して挨拶を行うことができ、ロボットが行う様々な接客方法に対する訪問客の注意を引くことができると考えられる。本研究は、訪問客のエンゲージメントに応じて返答する接客ロボットの反応モデルを実際の店舗における訪問客の観察データに基づいて設計したという点で HRI 分野に対して大きく貢献していると考えられる。

## 謝辞

本研究を遂行するにあたり、京七味専門店ちんとの皆様には実験場所の提供および実験における多大な協力を頂きました。心より厚く御礼申し上げます。

## 参考文献

- [Stock 17] R. Stock, and M. Merkle, “A Service Robot Acceptance Model: User Acceptance of Humanoid Robots During Service Encounters.” Pervasive Computing and Communications Workshops (PerCom Workshops), 2017 IEEE International Conference on. IEEE, pp. 339-344, 2017.
- [Niemelä 17] M. Niemelä, P. Heikkilä, and H. Lammi, “A Social Service Robot in a Shopping Mall: Expectations of the Management, Retailers and Consumers.” Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. ACM, pp. 227–228, 2017.
- [Kuno 07] Y. Kuno, K. Sadazuka, M. Kawashima, and K. Yamazaki, “Museum Guide Robot Based on Sociological Interaction Analysis.” Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, pp. 1191-1194, 2007.
- [Christopher 05] P. Christopher, et al. “Engagement capabilities for ecas.” AAMAS’05 workshop Creating Bonds with ECAs, 2005.
- [Sidner 05] C. L. Sidner, C. Lee, C. D. Kidd, and N. Lesh, “Explorations in Engagement for humans and robots.” Artificial Intelligence, 166(1-2), pp. 140-164, 2005.
- [Shiomi 06] M. Shiomi, T. Kanda, H. Ishiguro, and N. Hagita: “Interactive humanoid robots for a science museum.” HRI’06 Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, pp. 305-312, 2006.
- [Castellano 14] G. Castellano, I. Leite, A. Pereira et al.: “Context-Sensitive Affect Recognition for a Robotic Game Companion.” ACM Transactions on Interactive Intelligent Systems, vol. 4, no. 2, pp. 1–25, 2014.
- [Corrigan 13] L. J. Corrigan, C. Peters, and G. Castellano, “Social-Task Engagement - Striking a Balance between the Robot and the Task.” Embodied Commun Goals Intentions Workshop ICSR, Vol. 13, pp. 1-7, 2013.
- [Salam 15] H. Salam and M. Chetouani: “A multi-level context-based modeling of engagement in Human-Robot Interaction.” in 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1–6, 2015.
- [Knapp 13] M. L. Knapp, J. A. Hall, and T. G. Horgan, “Nonverbal communication in human interaction.” Cengage Learning, 2013.
- [Yanco 12] H. Yanco, “Personalization in HRI: A Longitudinal Field Experiment.” Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, pp. 319–326, 2012.
- [Kelley 84] J. F. Kelley, “An iterative design methodology for user-friendly natural language office information applications.” ACM Transactions on Information Systems (TOIS), Vol. 2 No. 1, pp. 26-41, 1984
- [Michalowski 06] M. P. Michalowski, S. Sabanovic, and R. Simmons, “A Spatial Model of Engagement for a Social Robot.” Advanced Motion Control, 9th IEEE International Workshop on, pp. 762-767, 2006.
- [川口 09] 川口 一画, 葛岡 英明, 鈴木 祐也, 中尾 誉, 山下 淳, カローラピッチ, 山崎 敬一: ロボットの発話途中の沈黙と言い直しによる人の注意誘導, 日本バーチャルリアリティ学会論文誌, Vol. 14 No. 3. pp. 257-264, 2009.
- [宮下 08] 宮下 善太, 神田 崇行, 塩見 昌裕, 石黒 浩, 萩田 紀博, “顧客と顔見知りになるショッピングモール案内ロボット.” 日本ロボット学会誌, Vol. 26 No.7, pp.821-832, 2008.
- [神田 04] 神田 崇行, 石黒 浩, “対話型ヒューマノイドロボットからの日常生活の中の友達関係の推定.” 情報処理学会論文誌, Vol. 45 No.8, pp. 2098-2104, 2004.
- [神田 08] 神田 崇行, “ロボットメディアによる人間情報処理研究: 4. コミュニケーションロボットと人間との距離.” 情報処理, Vol. 49 No.1, pp. 24-29, 2008.