

J-034

情報提示方法の違いがもたらすロボットサイネージの広告効果 Effects of Verbal or Nonverbal Information for a Robot Signage

塩見昌裕†, 篠沢一彦†, 中川賀史†, 宮下敬宏†, 坂本俊雄‡, 寺久保 寿光‡, 石黒浩‡, 萩田紀博†
Masahiro SHIOMI†, Kazuhiko SHINOZAWA†, Yoshifumi NAKAGAWA†, Takahiro MIYASHITA†, Toshio
SAKAMOTO‡, Toshimitsu TERAUBO‡, Hiroshi ISHIGURO‡, and Norihino HAGITA

1. まえがき

近年、ネットワークを介して複数種類のロボットとオペレータが連携し、人々の日常生活支援を目的とした「ネットワークロボット」の研究開発が進んでいる[1]。すでに、商業施設などの実環境内でネットワークロボットのコンセプトに基づいて開発されたロボットシステムを用いた実証実験[2~4]が行われており、将来のロボット利用に向けた研究開発が進んでいる。

ネットワークロボットのコンセプトと親和性が高く、将来の実環境で活動するロボットのための有望なアプリケーションの一つとして、ロボットを広告媒体として利用する「ロボットサイネージ」が注目されている[5]。ロボットサイネージとは、身体を持ったロボットが、対話などの人々にとって自然なコミュニケーションを通じてインタラクティブに情報を提示するシステムである。既に国内でも、実験的にはあるがロボットの広告利用に関する研究開発が進んでいる[6,7]。

我々は、ロボットサイネージにおける広告効果を向上させるためのロボットの振る舞いや情報提供の設計方法に興味を持つ。過去の研究から、実体を持つロボットという存在が、実体を持たないコンピュータエージェントに比べて、実環境下における情報提供ではより効果的であることが示されており[8, 9]、ロボットという存在を利用したロボットサイネージの有用性が示唆されている。しかし、ロボットサイネージは人々に情報を提供する際に、ロボットの発話や身体動作だけではなく、インタラクティブなディスプレイと連動して画像を提示するなど、様々な情報提示手法を持っており、どのような手法を用いることがより効果的に広告効果を向上させるかはいまだ明らかではない。

そこで本研究では、一般的なデジタルサイネージにおいて用いられる画像を用いた推薦行動と、ロボットの発話による推薦行動の、どちらが広告行動として効果的であるかを検証する。具体的には、実際の商業施設内で利用可能なクーポンの発行サービスを行うロボットサイネージシステムを開発し、商業施設内での検証実験を行った(図1)。実験を通じて、ロボットが発話による推薦行動を行った場合と、画像による推薦行動を行った場合の、広告効果の違いを検証した。



図1 商業施設におけるロボットサイネージ

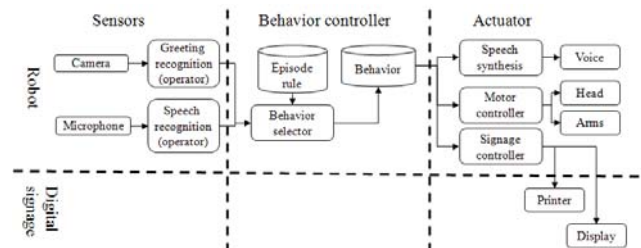


図2 システム構成

2. システム構成

我々は、ロボットとデジタルサイネージから構成されるロボットサイネージシステムの開発を行った。デジタルサイネージは、クーポンに関する画像を提示することと、クーポンを印刷することが可能である。

図2に、システム構成を示す。本研究において、音声認識はオペレータが行った。オペレータによる音声認識結果は、ロボットが次にどのようなビヘービアを選択するかの意志決定に利用された。さらに、ロボットはネットワークを通じてデジタルサイネージを制御することが可能であり、任意のタイミングでクーポンを印刷したり、特定の画像を表示したりすることが可能である。ビヘービアの遷移は、エピソードルールと呼ばれるあらかじめ決められたルールによって決定される。以下に、システムを構成する各要素の詳細を記述する。

2.1 ハードウェア

2.1.1 Robovie-miniR2

本研究では、小型ロボットとして、ATRで開発された”Robovie-miniR2”を用いた(図1)。このロボットは身長約30cmであり、指さしや頷きといった、人のような身体動作を実現するための上半身を持つ。首に3自由度と、両腕にそれぞれ4自由度を持つ。なお、音声発話を行うた

† (株) ATR 知能ロボティクス研究所, IRC

‡ 大阪大学, Osaka-Univ.

めに、本研究では音声合成システムの XIMERA[10] を用いた。

2.1.2 デジタルサイネージ

本研究で開発したデジタルサイネージは、50 インチディスプレイと、A7 サイズの小型プリンタ(Brother, MW140BT type-F) によって構成される。デジタルサイネージは、クーポンに関する情報を記載した画像を表示するために用いられる。なお、デジタルサイネージはロボットから制御信号をネットワーク越しに受け取ることで、表示する画像を変更したり、特定のクーポンを印刷したりすることが可能になっている。

2.2 ソフトウェア

2.2.1 サイネージコントローラ

ロボットは、自身の振る舞いや発話のタイミングに合わせてデジタルサイネージにクーポンに関する画像を提示し、またクーポンの印刷を行う。このような、ロボットとデジタルサイネージ間の同期した振る舞いを実現するために、過去に開発された複数ロボット間の連携動作を可能にするスクリプトを改良して利用した[11]。具体的には、デジタルサイネージに特定の画像を提示するための機能と、特定のクーポンを印刷するための機能を実装した。このスクリプトを用いることで、複数の異なるシステム間でタイミングを調整して情報提示などを行うことが可能になる。

2.2.2 ロボットのビヘービア

ロボットのビヘービアは、過去に ATR で開発されたコミュニケーションロボット Robovie-II の研究開発において提案されている、状況依存モジュールを参考に実装した[12]。状況依存モジュールとは、状況に応じてロボットによる握手・挨拶・道案内等の、意味をもつ一連の行動を実行するための機能を意味する。ロボットが人間と対話するためには、センサ入力や対話履歴に基づいて実装されたビヘービアを適切に遷移させる必要があるが、ビヘービアを遷移させるルールを記述したものをエピソードルールと呼ぶ[12]。例えば、ロボットは挨拶動作をした後に「こんにちは、僕はロボビーです」等と発話した後、マイクから得られた相手の発話音声内容に応じて次の振る舞いを決定する、といった動作が可能である。各ビヘービアは 5 秒から 20 秒程度の長さで実行される。なお本研究では、141 のビヘービアと 233 のエピソードルールを実装した。実装した主なビヘービアは、各店舗に関する情報提供と、道案内である。

道案内ビヘービアでは、ロボットは指さし動作と発話動作を組み合わせて情報提供を行う。例えば、ロボットはまず始めに進むべき方向を指さしながら、「この道をまっすぐ行って」等と発話し、その次に目印となる施設がある方向を指し示しながら、「2 つめのエスカレーターを左に曲がるとあるよ」等と発話する。ロボットには全ての店舗に関する位置と情報が記録されており、合計 101 の施設を案内することが可能である。なお、ロボットの指さし動作を生成するために、杉山らの提案した”three-layer attention-drawing model” [13]を利用した。

ロボットのビヘービアを設計するに当たって、我々は 2 つのデザインポリシーを利用した。1 つは、ロボットが対話のイニシアティブを取るように発話を行う点である。例えば、ロボットは単に対話相手の発話を待つわけではなく、自ら「僕は道案内が出来るんだ」「どこかに案内してあげようか？」などと発話することで、対話相手にロボットの持つ役割を明確に説明する。このような対話を行うことで、対話相手がロボットと円滑に対話を開始できるようになると考えた。もう 1 つは、ロボットの発話内容を人間が行うような、感情的な表現を行うように設計した。具体的には、「あのレストランはとってもおいしいよ」「イベントはとっても楽しかったよ」などと、より親和的な印象を与えるように、自身の経験を踏まえたように発話させた[2,14]。

2.3 オペレータ

本研究では、一人のオペレータがロボットの音声認識機能を代替する。音声認識機能は、ロボットが人々との自然な対話を実現するために必要不可欠な機能であるが、既存のシステムでは実環境下での利用において十分な性能を確保することがいまだ困難である。そのため本研究では、オペレータに環境情報と音声情報を提示し、効率的な音声認識機能代替を実現するためのインターフェースを利用した[15]。

このインターフェースを利用することで、オペレータはネットワークを通じて、ロボットのマイクから取得された対話相手の発話音声を聞き取ることが出来る。また、インターフェース上に表示される、音声認識結果を示すボタンを選択することで、音声認識結果がロボットに送信される。つまりオペレータは、ロボットの外部に存在する音声認識機能と同等の役割を行う。オペレータの作業をこのように設定した理由は、将来音声認識機能が高性能になった際に、容易に置き換えを可能にするためである。

3. 推薦方法の違いによる効果の比較検証実験

3.1 実験方法

3.1.1 実験環境

本実験は、大阪に存在するアジア太平洋トレードセンター(ATC)内における、ショッピングモール内の通路で行われた。実証実験は夏休み期間中に実施され、実験を行った通路には 1 時間あたり平均 1300 人の人々が通過した。モールを訪れるユーザは主に家族、カップル、および観光者であった。これらのユーザは、自由にロボットと相互作用を行うことが出来た。図 3 に、実験環境と周辺の店舗の配置図、およびデジタルサイネージの設置場所を示す。ロボットは、ディスプレイの横に配置した。ロボットの顔の高さとディスプレイの中央部分の高さを揃えるように、ロボットを高さ約 90cm の台に設置した。

なお、この実験は研究目的のみ利用するという条件の下で、ショッピングモールの管理者の承認を得て行われた。この手続きについては、所内の倫理審査委員会により承認済みである。

3.1.2 実験手順

実験は、昼間と夕方の2種類の時間帯で行われた。これは、実験環境において、昼間と夕方で人々の主な振る舞いに変化していたためである。具体的には、昼間は昼食や買い物、およびイベント会場に向かう人々が主であり、夕方には夕食に向かう人や帰宅する人が主であった。時間帯によって変化する人々の振る舞いの影響を考慮し、以下に示す各実験条件では、昼間2時間と夕方2時間の合計4時間を実験に割り当てた。なお、実験を行った曜日は、全て夏休み期間中の平日であった。

3.1.3 実験条件

条件1: 音声による推薦条件

本条件では、ロボットは特定の店舗に関する推薦を、音声情報のみを用いて行う。そのためこの条件では、ディスプレイに「ロボットがクーポンを印刷することができる」という情報のみを示した画像を常に提示していた。なお、推薦する店舗としては、昼間には定食などのランチメニューを提供している韓国レストランを、夕方には軽食やデザートなどを提供しているフードコートの推薦を行うこととした。

本条件では、ロボットは相互作用の開始時において自己紹介をした後、「僕は、お薦めのクーポン紹介と、道案内が出来るよ」などと話しかける。人々が道案内を望んだ場合には、指さし動作と発話動作を用いて道案内を行う。お薦めの紹介を求められた場合には、時間帯に応じて店舗の推薦行動を行う。推薦行動では、デジタルサイネージに画像を表示し、「僕のお薦めは韓国料理屋さんだよ。モダンな店内で、スパイシーな料理を楽しめるんだ！」等と、推薦する店舗に関する情報提供を行う。

さらに、発行するクーポンの割引内容を「20%引きで、お得だよ」などと音声で提示する。人々がクーポンの発行を望んだ場合には、クーポンの印刷を行う。人々がクーポンに興味を持たなかった場合には、「そっかー、残念だなあ。じゃあ、またね」等と発話し、対話を終了する。

条件2: 画像による推薦条件

本条件では、ロボットは特定の店舗に関する推薦を、画像情報を利用して行う。道案内時の発話内容や推薦する店舗は、全て条件1と統一した。

本条件では、ロボットが特定店舗の推薦を行う際に、「僕が発行できるクーポンは、これだよ」などと発話してディスプレイを指差し、同時にクーポンに関する画像情報を提示する。提示される画像情報は、特定店舗がWEB上に公開しているクーポンの画像をそのまま利用した。つまり、条件1との違いは、発行するクーポンの割引内容を音声で提示するか、画像で提示するかの違いとなる。

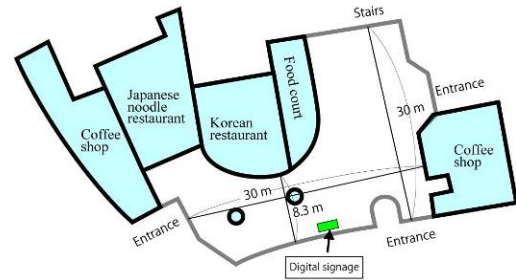


図3 実験環境

3.2 実験結果

3.2.1 クーポン印刷数の比較

実験において、音声による推薦条件では261人（109グループ）がシステムを利用し、画像による推薦条件では319人（118グループ）がシステムを利用した。なお、システムを同時に利用した1つの集団（家族や友人など）を、1つのグループとして計算した。

また、各条件において環境内を通行した人数は、音声による推薦条件が4369人、画像による推薦条件では5186人であった。システムを利用した人数と通行した人数を用いて、フィッシャーの正確確率検定を行った結果、条件間に有意な差は見られなかった。これは、条件間の違いがロボットの対話内容のみであり、外見や振る舞いに関しては条件間で統一しており、特に人々を引きつける影響に変化が起こらなかったため、有意差が発生しなかったと考える。

次に、各条件においてシステムを利用したグループを、1. クーポンを印刷した(Printing)、2. クーポンに関する情報を聞いたが、クーポンを印刷しなかった(Non-printing)、3. クーポンに関する情報を聞かず、単に挨拶や道案内のみを行った(Interaction only)の、3つのカテゴリに分類した(表1)。なお、実験において同じグループが複数回システムを利用した場合には、その分を除去して計算した。

分類した各カテゴリの値を用いて、カイ二乗検定を行った結果、PrintingカテゴリとNon-printingカテゴリにおいて条件間で有意な差があることが示された($p < .05$)。つまり、音声による推薦条件において、実際にクーポンを印刷した人々の数は、画像による推薦条件よりも多いことが示された。

表1 システム利用者数と通行人

	利用者数 (グループ)	通行人
音声のみ	261 (109)	4369
画像のみ	319 (118)	5186

表2 システム利用者のカテゴリ

	Printing	Non-Printing	Interaction only
音声のみ	19	6	84
画像のみ	10	28	80

3.2.2 アンケート結果

本実験において、我々はシステムを利用した人々に対してアンケート調査を実施した。その結果、音声のみ条件では21組が、画像のみ条件では24組がアンケートに回答した。アンケート項目は使いやすさ、有用性、楽しさの3項目であり、1(低評価)～7(高評価)の7段階評価とした。

図4に、アンケート項目の平均と標準誤差を示す。分散分析の結果、全ての項目間において有意な差は見られなかった。

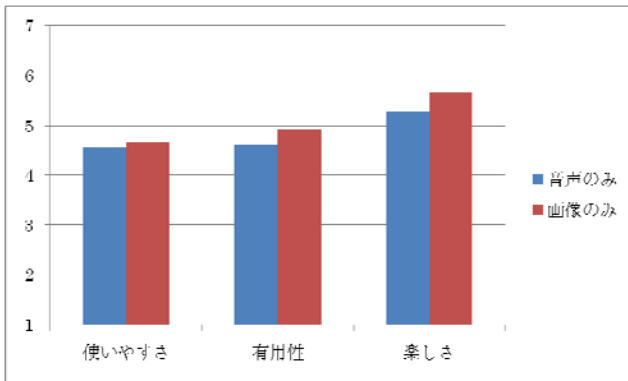


図4 アンケート結果

4. 考察

実験の結果から、音声による情報提示が画像による情報提示よりも有効であることが示された。一方で、たとえば複数の候補からユーザにクーポンを選んでもらう状況などを考慮した場合、音声による情報提示は画像で情報提示を行う場合に比べて時間がかかるなど、ユーザビリティの面で利用が困難になることが予測される。そのため、同時に複数の情報を提示する場合には、画像情報と音声情報を同時に利用するなど、ユーザによって利用しやすい情報提示手法が必要になると考える。

また今回の結果は、人々とのコミュニケーションを行うデバイスとして実体を持つあるロボットを利用して、ある一つの商業施設における特定店舗の推薦を行うという実験によって得られたものであるため、この知見の一般化には限界があることに注意が必要である。

5. おわりに

本研究では、ロボットが人とのコミュニケーションを通じて行う広告を行う「ロボットサイネージ」システムを開発した。開発したシステムは、実際の商業店舗内で利用できるクーポン印刷と、それらに関する情報提供、および商業施設内の道案内サービスを行うことが可能である。

開発したシステムを用いて、実環境下でロボットの情報提示手法の違いによる影響を検証した。具体的には、実際の商業施設で、ロボットが音声を用いてクーポンの内容を説明する場合と、画像を用いて説明する場合での、システム利用者およびクーポン発行率を比較した。

実験の結果、ロボットが音声を用いて情報提示を行うことで、ロボットが画像を用いて情報提示を行った場合よりも有意に多くの人々がクーポンを印刷することが示された。

謝辞

実験環境を提供していただき、様々なご支援をいただいたアジア太平洋トレードセンターの皆様へ厚く感謝申し上げます。本研究は、総務省の研究委託により実施したものである。

参考文献

- [1] A. Sanfeliu, N. Hagita, and A. Saffiotti, "Network Robot Systems," Special Issue: Network Robot Systems, Robotics and Autonomous Systems, 2008
- [2] T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, and N. Hagita, "An affective guide robot in a shopping mall," ACM/IEEE International Conference on Human-Robot Interaction (HRI2009), 2009.
- [3] T. Kanda et al., "Who will be the customer? A social robot that anticipates people's behavior from their trajectories," UbiComp2008.
- [4] M. Shiomi, et. al., "Field Trial of Networked Social Robots in a Shopping Mall," Proceedings of the 2009 IEEE International Conference on Intelligent Robots and Systems (IROS2009), 2009.
- [5] 宮下敬宏, ユビキタス特区における実証実験, 電子情報通信学会総合大会予稿集, DP-1-5, CD-ROM, 2010.
- [6] 村川 賀彦, ロボットによる販売促進活動の評価, HAI シンポジウム 2009.
- [7] 村川賀彦, HAI のロボット産業への応用, 人工知能学会誌, Vol. 24, No. 6, pp. 864-871, (2009)
- [8] K. Shinozawa, F. Naya, and J. Yamato, "Differences in Effect of Robot and Screen Agent Recommendations of Human Decision-Making," International Journal on Human-Computer Studies, Vol. 62, 2, Feb. 2005.
- [9] W. A. Bainbridge, J. Hart, E. S. Kim, and B. Scassellati, The Effect of Presence on Human-Robot Interaction, IEEE Int. Symposium on Robot and Human Interactive Communication (RO-MAN2008), pp. 701-706, 2008
- [10] H. Kawai, T. Toda, J. Ni, M. Tsuzaki, and K. Tokuda, "XIMERA: A New TTS from ATR Based on Corpus-Based Technologies," Proc. of Fifth ISCA Workshop on Speech Synthesis (SSW5), pp. 179-184,
- [11] D. Sakamoto, et al., "Humanoid Robots as a Broadcasting Communication Medium in Open Public Spaces," International Journal of Social Robotics, 1(2), pp. 157-169, 2009.
- [12] T. Kanda, H. Ishiguro, M. Imai, and T. Ono, "Development and Evaluation of Interactive Humanoid Robots," Proceedings of the IEEE, Vol. 92, No. 11, pp. 1839-1850, 2004.
- [13] 杉山治, 神田崇行, 今井倫太, 石黒浩, 萩田紀博, 安西祐一郎, 人とコミュニケーションロボットの直示的な会話の実現, ヒューマンインタフェース学会論文誌, 10(1), pp.73-86, 2008.
- [14] R. W. Picard, "Affective Computing," 1997.
- [15] D. F. Glas, T. Kanda, H. Ishiguro, and N. Hagita, Simultaneous Teleoperation of Multiple Social Robots, ACM/IEEE International Conference on Human-Robot Interaction (HRI2008), pp. 311-318, 2008.