

物体にキャラクターイメージを想起させる 情報提示手法のオーサリングツール

An Authoring Tool for Interaction Method Evoking Character Imagery on an Object

石井 健太郎*¹ 大澤 博隆*² 山田 誠二*³
Kentarō Ishii Hirotaka Osawa Seiji Yamada

*¹東京大学大学院 情報学環

Interfaculty Initiative in Information Studies, the University of Tokyo

*²慶應義塾大学 理工学部

Faculty of Science and Technology, Keio University

*³国立情報学研究所 / 総合研究大学院大学 / 東京工業大学

National Institute of Informatics / SOKENDAI / Tokyo Institute of Technology

We propose an authoring tool that helps to create communication contents of an anthropomorphic home appliance. Previous work showed that self-introduction by an anthropomorphic printer was more memorable and understandable than introduction of the printer by a humanoid robot. This result can motivate salespeople at retail stores to use the anthropomorphic product performing self-promotion. The proposed authoring tool named EmbEdit provides a feasible way for on-site creation of such communication contents. Three features for easy content creation are demonstration-based condition registration for content switching, automatic position calculation of the components by visual pattern recognition, and on-site text-to-speech synthesis. This paper describes our design and implementation of the proposed system.

1. はじめに

家庭電化製品など、市場では様々な特長を持った製品が販売されており、一見しては製品の特長がわからないということがある。そのため、製品の販売店では、製品の特長を記した貼り紙をして来店客に提示するといった方法がとられる。購入後の製品の利用時とは異なり、販売店での購入製品の検討は時間が限られている。また、事前に決めた1つの製品を検討するのではなく、多くの製品を比較して検討する場合も多い。

本研究は、既存の貼り紙による情報提示に代わるキャラクターイメージを想起させる情報提示手法について扱う。Osawaらは、目や腕のパーツを取り付けることで物体を擬人化させ情報提示する手法を提案した [Osawa 09]。この手法の有効性は、プリンタの機能を説明するタスクにおいて、人間型ロボットが物体の脇で説明する手法との比較において、説明を受けた者がより対話に集中し、より多くプリンタの機能を覚えたことによって示された。本研究では、情報提示を受けた者がキャラクターイメージを想起させていると考える (図 1)。

プリンタの機能の情報提示を受けた者がプリンタの機能を多く覚えたことから、Osawaらの情報提示手法は冒頭に述べた製品の販売店で用いることが有用であると考えられる。しかし、この手法を製品の販売店で実現するためには、販売店の店員がその場で対話コンテンツを作成しなければならない。なぜならば、擬人化させる物体や作成者によって対話コンテンツを変更することが考えられるためである (図 2)。

本論文では、上記の情報提示手法を行う対話コンテンツを容易に作成するためのオーサリングツール EmbEdit を提案する。EmbEdit による対話コンテンツ作成は、物体へのパーツの取り付け・動作内容の入力の2つの手順からなる。動作内容は、動作が切り替わるための条件と発話文字列・パーツの動作を示すタグ文字列によって記述される。パーツの位置関係はカメラにより自動的に取得し計算するため、任意のパーツ配置に



図 1: キャラクターイメージを想起させる情報提示手法

において、目や腕によるポインティングも動作内容の入力時にタグ文字列を入力するだけで実現できる。また、情報提供を受ける者の物体に対する操作を検知するためのセンサを用意しており、例えば扉を開いた・ボタンを押したといった場合にインタラクティブに実行される動作を用意することができる。

2. 関連研究

人間のような外見により、愛着がわき理解を促進することが報告されている [DiSalvo 03]。このため、擬人化エージェントによる製品の説明は、購買を促進することが期待される。しかし、ロボットのような独立した擬人化エージェントによる説明は、説明する物体への注目を妨げる場合がある。物体に擬人化パーツを取り付け、物体自身が説明する手法は、この問題を避け、対話者の注目を物体に向けることに有効であることが示されている [Osawa 09]。また、擬人化された物体による主観的な表現が、特に子どもやお年寄りに好まれるという報告もある [Osawa 07]。

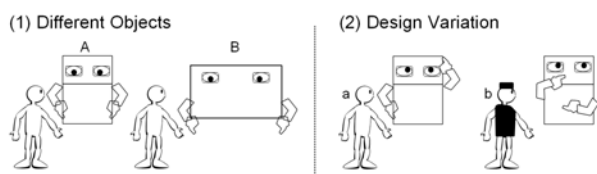


図 2: 本研究が対象とする状況

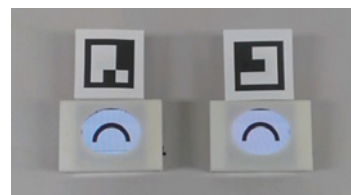


図 4: 目を模した身体パーツの外観

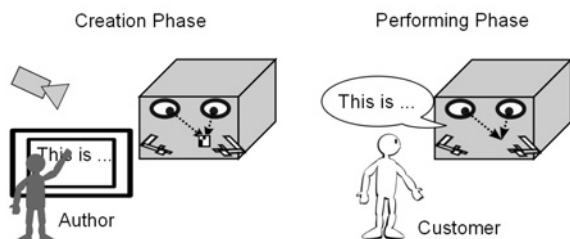


図 3: 本研究の情報提示手法の利用手順

一方、ウェブページをはじめとして様々な用途で、コンテンツの作成・編集を容易に行うためのいわゆるオーサリングツールが開発されている。例えば、専門家でなくとも時系列センサーデータのパターン認識のフィルタを容易に作成できるオーサリングツールが提案されている [Hartmann 07]。ロボットやエージェントの対話シナリオを作成するためのオーサリングツールも存在する [Hayashi 05, Cassell 99]。しかし、既存のオーサリングツールでは固定された身体を仮定している。本研究で用いる身体パーツを利用したエージェントでは、身体パーツの位置関係は任意に変更できるためそのまま適用することができない。任意の位置関係の身体パーツに対応できるオーサリングツールが必要がある。

3. EmbEdit

本研究の情報提示手法は、コンテンツ作成・コンテンツ再生の2つのフェーズに分けられる(図3)。コンテンツ作成フェーズでは、擬人化パーツを取り付け、本論文で提案するオーサリングツール EmbEdit を用いて、物体に対話コンテンツを作り込む。本論文では、EmbEdit のユーザに対話コンテンツの作成者を単に作成者と呼ぶ。本研究で想定する製品の販売店では、店員が作成者にあたる。コンテンツ再生フェーズでは、擬人化パーツは自律的に動作を繰り返し、情報提示を行う。本論文では、情報提示を受ける者を対話者と呼ぶ。本研究で想定する製品の販売店では、来店客が対話者にあたる。対話者の反応を検出し、特有のコンテンツを再生することも可能であり、そのような対話シナリオを EmbEdit を用いて作成者は実現することができる。

3.1 システム構成

ハードウェアとしては、カメラ付きのコンピュータ・身体パーツ・センサパーツ・マーカパーツからなる。対話コンテンツの作成者が主に扱うのはカメラ付きのコンピュータで、オーサリングソフトウェアが動作する。このオーサリングソフトウェアにより作成者は対話コンテンツを入力する。身体パーツは擬人化を実現するためのデバイスであり、人間の目と腕を模したロボットパーツである。センサパーツは、扉を開ける・ボ

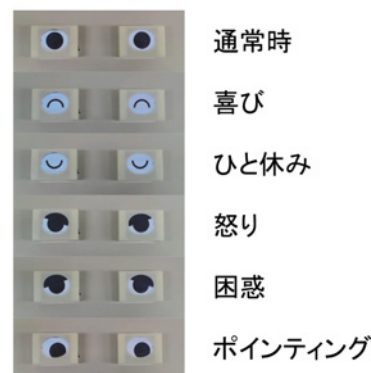


図 5: 目を模した身体パーツの表示パターン

タンを押すといった対話者の反応を検出するデバイスである。コンテンツ再生時には、作成者が入力した対話コンテンツとセンサパーツから得られる対話者の反応に基づき身体パーツを制御する。マーカパーツは、コンテンツ作成時に目・腕の身体パーツによりポインティングをする点に取り付けることで、キャラクターの発話中にポインティングの動作を行うために用いられる。コンピュータ・身体パーツ・センサパーツは ZigBee により、無線で接続される。

3.2 ハードウェア

カメラ付きのコンピュータには、市販のコンピュータと USB カメラを用いた。身体パーツ・センサパーツ・マーカパーツは、独自に開発したものであり、接着ゴムを用いて物体に取り付けることができる。身体パーツ・マーカパーツには、カメラで識別するためのビジョンマーカが取り付けられている。身体パーツのビジョンマーカは差し込み式になっており、コンテンツ再生時には取り外すことができる。

3.2.1 身体パーツ

目を模した身体パーツは、小型の有機 EL ディスプレイを内部に持ち、ディスプレイに目のパターンを出力することによって、表情を表すことやポインティングを行うことができる。図4に外観を、図5に表示可能なパターンを示す。

腕を模した身体パーツは、4自由度の関節を持ち、目を模した身体パーツの表情に合わせてポーズをとることや先端を任意の方向に向けてポインティングを行うことができる。図6に外観を示す。

3.2.2 センサパーツ

センサパーツには3種類あり、照度センサ・曲げセンサ・接触センサがそれぞれ取り付けられている。照度センサは、扉の開閉のように物体の形状を変えることで照度が増減するような状況で利用することを想定している(図7)。曲げセンサは、電話の受話器のように物体が重ね合わさる箇所において、その有



図 6: 腕を模した身体パーツの外観

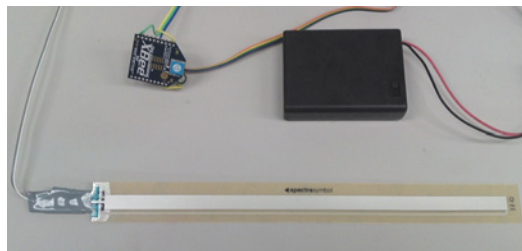


図 9: 接触センサのセンサパーツ

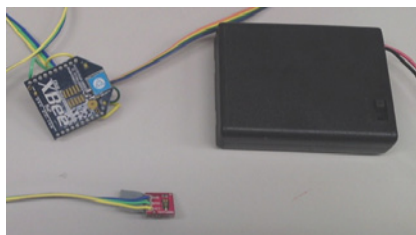


図 7: 照度センサのセンサパーツ

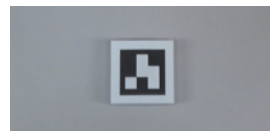


図 10: マーカーパーツ

無や度合いを検出するために利用することを想定している (図 8)。接触センサは、物体についているボタンの押されたかどうかのように、対話者が物体に触れたことを検出するために利用することを想定している (図 9)。

3.2.3 マーカーパーツ

マーカーパーツは、ビジョンマーカのみからなるパーツで、コンテンツ作成時にポインティング点を指定するために用いられる (図 10)。

3.3 ソフトウェア

ソフトウェアは、コンテンツ入力モジュール・コンテンツ再生モジュール・ビジョンマーカ認識モジュールからなる。コンテンツ入力モジュールは、コンテンツ作成時に作成者が主に利用するモジュールで、対話コンテンツを入力するために用いられる。コンテンツ再生モジュールは、センサパーツの値を読み取り、入力された対話コンテンツに基づき、発話と身体パーツを制御する。ビジョンマーカ認識モジュールは、コンテンツ作成時に物体に取り付けた身体パーツ・センサパーツ・マーカーパーツの位置関係を認識するために用いられる。

3.3.1 コンテンツ入力モジュール

作成者は、コンテンツの切り替えを行う条件・エージェントが発話する言葉・身体パーツの動作をコンテンツ入力モジュールを用いて入力する。図 11 は、コンテンツ入力モジュールの

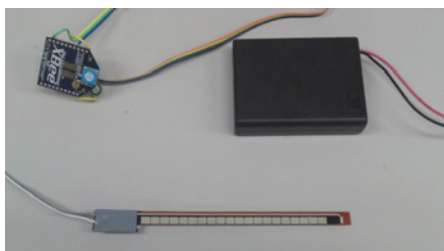


図 8: 曲げセンサのセンサパーツ

スクリーンキャプチャである。左側のペインにはコンテンツの切り替えを行う条件がリストされており、選択されている条件時に再生される動作コンテンツが右側のペインに表示されている。

コンテンツの切り替えを行う条件は、センサパーツから得られる値により特定される。作成者はコンテンツが切り替わるべきセンサ状態にして「条件登録」ボタンを押すことで、コンテンツ切り替えの条件を登録する。例えば、扉が開いた状態で「条件登録」ボタンを押すことで、扉が開いたことで切り替えを行う条件を登録することができる。

右側のペインの動作コンテンツには、エージェントが発話する言葉と身体パーツの動作を記述する。エージェントが発話する言葉は、テキスト読み上げエンジンが解釈できるようひらがなで入力する。身体パーツの動作は、タグ文字列として記述する。記述された身体パーツの動作は、タグ文字列の位置に発話が達したときに実行される。解釈可能なタグ文字列として、<smile>、<sleep>、<angry>、<strand>、<point:#>の 5 つのタグを用意しており、それぞれ喜び・ひと休み・怒り・困惑・ポインティングの動作に対応している。<point:#>は、#で示す箇所にパラメータを 1 つとり、ポインティング点を示す識別子を記述する。ポインティング点を示す識別子はビジョンマー

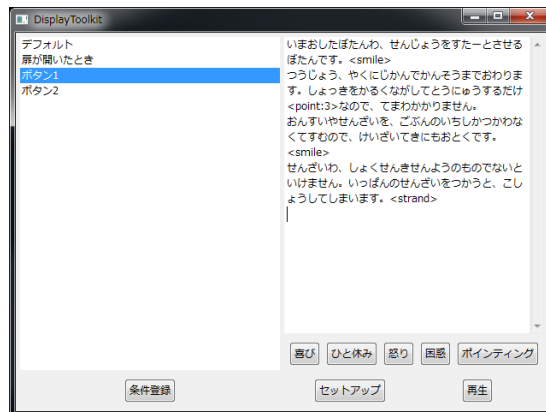


図 11: コンテンツ入力モジュールのスクリーンキャプチャ

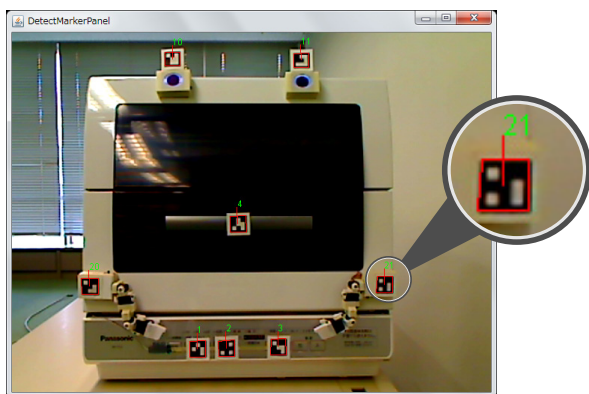


図 12: ビジョンマーカ認識モジュール

力認識モジュールに表示される数値が対応しており，作成者は後述のビジョンマーカ認識モジュールのウィンドウを確認して識別子を入力する．

「喜び」「ひと休み」「怒り」「困惑」「ポインティング」のそれぞれのボタンは，タグ文字列を自動挿入するためのボタンで，ボタンを押すとカーソルの位置にタグ文字列が挿入される．「セットアップ」ボタンは，パーツの位置関係を記憶するために用いる．作成者は，後述のビジョンマーカ認識モジュールにてビジョンマーカが認識されていることを確認して「セットアップ」ボタンを押す．「再生」ボタンは，作成中または作成後のコンテンツを再生するために用いる．作成者は再生されるコンテンツを確認して，反復的にコンテンツを修正することができる．

3.3.2 コンテンツ再生モジュール

コンテンツ再生モジュールは，コンテンツの再生を制御する．入力されたコンテンツ文字列をもとに，テキストを音声合成して再生する．身体パーツの動作を表すタグ文字列が現れた場合には，身体パーツに対応する信号を送る．センサパーツから得られた値がコンテンツの切り替えを行う条件に合致した場合には，再生するコンテンツを切り替える．音声合成エンジンには，規則音声合成ライブラリ AquesTalk を用いた [AquesTalk] ．

3.3.3 ビジョンマーカ認識モジュール

ビジョンマーカ認識モジュールは，身体パーツ・センサパーツ・マーカパーツに付属するビジョンマーカを認識する (図 12) ．ビジョンマーカは， 3×3 の 2 次元パターンで，パターンの 4 隅を正方形の透視変換とみなすことによって，3 次元的位置関係を計算する．

4. 今後の課題

まず改善点と考えられるのが，コンテンツ入力モジュールの発話文字列をひらがなではなく，漢字かな混在の文字列とすることである．作成者は通常の漢字かな混在の文字列を入力し，システムが漢字かな変換を行ってテキスト読み上げを行う．漢字かな変換が理想的に機能すれば，作成者の利便性は高まると考えられる．ただし，漢字かな変換が完全に誤りなく動作することは考えにくく，ひらがな入力の手間と漢字かな変換の誤り修正の手間のトレードオフ・修正する場合の方法とユーザインタフェースを，よく検討することが必要となる．

上記も含め，ツールとしてのシステムデザインを改善する

ため，作成者を招きオーサリングツールを利用してもらったユーザテストによる反復的改善のプロセスが必要であると考えられる．今後そのようなユーザテストを行い，ツールを段階的に改善していくことを計画している．その後，現在一般的に行われている貼り紙による情報提示と比較してツールとしての評価を行う予定である．

5. まとめ

本論文では，物体に目や腕を模した身体パーツと対話者の反応を検出するセンサパーツを取り付け，キャラクターイメージを想起させることで，効果的な情報提示を行う手法のオーサリングツール EmbEdit を提案した．製品の販売店で人間の代わりに製品自体が自らの宣伝をすることで，分かりやすい説明を実現すること・注目を集めることが期待できる．しかし，従来は物体ごとに対話コンテンツは作り込みがなされていた．EmbEdit は，センサの値に応じたコンテンツ切り替え・カメラによるパーツの位置関係の自動認識・動的な音声合成による任意のテキストの読み上げを組み合わせることで，物体によって異なるパーツ構成・作成者によって異なる要求に応じて，対話コンテンツの作成を容易に行うことができる．

参考文献

- [AquesTalk] 株式会社アクエスト: AquesTalk, <http://www.a-quest.com/products/aquestalk.html>
- [Cassell 99] Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., and Yan, H.: Embodiment in Conversational Interfaces: Rea, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 520–527 (1999)
- [DiSalvo 03] DiSalvo C. and Gemperle, F.: From Seduction to Fulfillment: The Use of Anthropomorphic Form in Design, *Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces*, pp. 67–72 (2003)
- [Hartmann 07] Hartmann B., Abdulla, L., Mittal, M., and Klemmer, S. R.: Authoring Sensor-based Interactions by Demonstration with Direct Manipulation and Pattern Recognition, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 145–154 (2007)
- [Hayashi 05] Hayashi, K., Kanda, T., Miyashita, T., Ishiguro, H., and Hagita, N.: Robot *Manzai* – Robots' conversation as a passive social medium–, *Proceedings of IEEE-RAS International Conference on Humanoid Robots*, pp. 456–462 (2005)
- [Osawa 07] Osawa, H., Mukai, J., and Imai, M.: "Display Robot" – Interaction between Humans and Anthropomorphized Objects, *Proceedings of the 16th International Symposium on Robot and Human Interactive Communication*, pp. 451–456 (2007)
- [Osawa 09] Osawa, H., Ohmura, R., and Imai, M.: Using Attachable Humanoid Parts for Realizing Imaginary Intention and Body Image, *International Journal of Social Robotics*, Vol.1, No.1, pp. 109–123 (2009)