

## Voicy:ロボットモーションを伴ったつぶやきシステムの提案

尾形 正 泰<sup>†1</sup> 大澤 博 隆<sup>†1</sup>  
篠 沢 一 彦<sup>†2</sup> 今 井 倫 太<sup>†1</sup>

本稿では、Twitter をつぶやきプラットフォームとして利用し、Twitter 上に集約されている実世界に関する情報をロボットアバターを用いて提示するエージェントシステムを提案する。ロボットアバターは近年エージェントとして利用される機会が増えているが、モバイル端末などから使用するにはシステムが大掛かりでユーザに負担であった。音声認識技術を用いて容易な入力方法と、肩に載せるタイプのアバターロボットを用いることで、従来と比べてカジュアルなアバターロボットの使用を実現した。また、つぶやきを発話と方向指示をアバターロボットに表現させることで、利用者に実世界の事象を指していることを理解を促すシステムを作成した。

### Voicy: Twitter system with robot motion

MASAYASU OGATA,<sup>†1</sup> HIROTAKA OSAWA,<sup>†1</sup>  
KAZUHIKO SHINOZAWA<sup>†2</sup> and MICHITA IMAI<sup>†1</sup>

In this paper, using Twitter as micro-blogging system, we propose agent system that presents information on twitter about real world with robot avatar. Though robot avatar is getting popular as agent recently, it is a burden for user to make use of agent through such as mobile device because of large system. Using easy input method with voice recognition and a type of avatar robot on shoulder, we realized more casual avatar robot than before. And to make avatar robot to present tweet as utterance and direction indication, system accelerate user to understand that information from avatar is about real world.

<sup>†1</sup> 慶應義塾大学  
Keio University

<sup>†2</sup> ATR  
Advanced Technology Research Laboratory

### 1. はじめに

アバターロボットは、親しみやすさとロボットの存在によって、人間のコミュニケーションにロボットが入り込むために有効な手段である。アバターコミュニケーションでは操作するユーザは操作端末を使用してアバターロボットを操作することで、インターネットを通じて遠隔地でのコミュニケーションを確立することができる。対して、インターネット上では人々の交流はソーシャルネットワークにおいて活発になってきている。その中でも特に短いテキストを用いてユーザ同士が会話を行うことができるサービスをマイクロブログ (micro-blogging) と分類する。マイクロブログではユーザからの投稿によって、ユーザの日常行動を記録するライフログや、特定のキーワードで検索することで得られるユーザのフィードバックなど、有益な情報がリアルタイムで配信される。ソーシャルサービスを使うことで得られる情報の価値は高まっているが、拡張現実感 (augmented reality) のようにデジタルデータと現実世界との融合が実現されれば、ユーザ自身が情報を受けとる利便性と視覚性が向上すると考えられる。ロボットを介したインタラクションシステムを構築することで、インターネット上のデータを用いて実世界の物事に言及することができる。

#### 1.1 関連研究

アバターロボットは、Anybots 社の QB Avatar Robot<sup>?)</sup> のように擬人的な姿をもち、人間が遠隔操作することで周囲の人間とコミュニケーションを行うものがある。また、現実の事例としてアバターロボットが病気の子どもが学校に行くことができるようにするよう、自分の分身として使用することができる。<sup>?)</sup> 我々の研究室ではブログロボット<sup>?)</sup> を作成し、ブログの投稿を音声発話とロボットモーションによる方向指示で説明するエージェントシステムを提案している。ブログはソーシャルサービスのような双方向性がないため、ブログロボットでは一方方向性の情報伝達と、エージェントとの記事の文脈に沿った会話インタラクションのみが可能であった。また、おもちゃとしてタカラトミーから「ついまる」<sup>?)</sup> が発売されている。スピーカーのみを備えた柔らかみのあるデザインのエージェントが、合成音声を用いて Twitter のタイムラインを読み上げてくれるなど、Twitter に関するアバターも登場している。

#### 1.2 課題

しかしながら、双方向性の実現の点から、ブログロボットを Twitter に適用することは難しい。実世界において人から人に情報が伝達される仕組みを Twitter でも実現するためには、2つの課題が存在する。1つは投稿するメッセージを作成する際に生じる遅延であり、

情報の受け手がメッセージを再生するのにかかるコストである。Twitter はブログよりは高速にコミュニケーションが可能であるが、通話よりは低速であり、遅延が生じてしまう。遅延は発信時にキーボードで文字を入力する時間と、受信時にそれまで行っていた作業を中断して画面を閲覧して文面を理解する時間である。次に、情報の受け手が端末の画面を参照してメッセージを読む必要があることである。ユーザが情報を受け取る際のメディアがモバイル端末などのデジタル画面であるため、本来人から人に伝えられるべき情報は画面を通じてユーザに配信される。

### 1.3 提 案

これらの課題は、音声ツイート機能を実装することで高速な入力と、実世界における人と人による会話の概念を取り入れることで解決する。ロボットを実世界エージェントとして使用することで、人と人が言葉を交わすような状況を作り出すことができる。音声ツイート機能により、ユーザが発話した音声は音声認識サーバに送信されて音声ファイルと認識結果のテキストが Twitter に投稿される。ツイートを受信するユーザは、発信側のユーザの肉声によってツイートの内容を受け取ることができる。このとき、エージェントであるロボットが人間と同じ空間においてツイートで言及されたオブジェクトや方向を共有することができるように GPS 位置情報をツイートに埋め込む。これらの要素を組み合わせ、インターネットから得られた情報を、実世界の情報として人に伝えることで、モーションを伴って音声でつぶやきを再生するエージェントシステムを提案する。

## 2. システム構成

我々は、Twitter のようなマイクロブログに投稿するときに、音声を用いてメッセージを伝送できるようにした。これは2段階のレイヤーで実現されている。1つは音声認識技術を用いた音声からテキストへの変換である。ここで変換されたテキストをマイクロブログに投稿することで、文字入力にかかる時間を省略している。2つ目は音声ファイルをサーバにアップロードした上で、メッセージ内にアップロードしたファイルに対する短縮 URL を埋め込むことである。これにより非同期的に高速な音声の伝達が可能になった。

次にエージェントに我々の研究室で制作している肩乗りロボットを用いることで、ロボットの利用者と同一方向を共有することができ、位置情報に基づいた利用者に対するサービスを行いやすくした。また、肩に乗っているためロボットの発話を受けたときに、ささやきの効果でメッセージの内容の信憑性が向上し、スピーカーの音声が聞きやすいという利点がある。それぞれをいかに詳しく説明する。

### 2.1 Android クライアント

Twitter を使うにはウェブから利用するかクライアントアプリケーションを用いる必要がある。本研究ではクライアントアプリケーションとして機能を追加する必要があったため、改変可能な Twitter のオープンソースプロジェクトが存在する Android プラットフォームを選択した。オープンソースの Twitter クライアントを拡張し、2つの機能を実装した。1つは「音声ツイート」と呼び、キーボードによるテキスト入力の代わりにマイクからの音声を録音し、音声認識を行うことでテキストを生成する方法である。また、音声ファイルは音声認識の処理の後、音声認識サーバにアップロードされるが、下記の音声認識サーバを構築して使用した。クライアントアプリケーションは音声認識サーバから返ってきた認識結果と音声ファイルへの URL をメッセージとして Twitter に投稿を行う。また、このときクライアントはツイートを投稿する際に端末の GPS から得られる位置情報タグを埋め込む。

### 2.2 音声認識サーバ

音声認識サーバに音声ファイルを送信すると、認識結果のテキストと音声ファイルの URL が返ってくる。Android では Google 音声認識技術を扱う API が用意されているが、音声を録音してファイルとして保存できないため別の方法を採用した。そのため、Android 端末外部に音声認識のためのサーバを設置し、さらに Android 端末から送信された音声ファイルを保存する。音声ファイルにはユニークな URL を対応させ、短縮 URL サービスと同様に音声ファイルにアクセスできるようにした。音声認識技術には京都大学で開発された Julius を使用し、ウェブアプリケーション側から待機状態のプロセスに対して Julius の処理を実行するようにソケット通信を用いて命令する。

### 2.3 アバターロボット

ロボットは実世界におけるエージェントであり、人間に対して存在感のあるメディアとして機能する。本研究の場合、端末の画面やスピーカーから得られるメッセージを受け取るよりも、人間が存在感を感じるロボットによって発話と発話に伴う動作を行わせることが、メッセージを投稿したユーザの存在感を伝達できると考えた。ロボットには、手に乗せてブログ記事を読み上げるブログロボットを肩に乗るように改良した TEROOS<sup>2)</sup> のハードウェアを使用した。TEROOS は肩にかけることができるため、手に乗せたり持ち歩くタイプのロボットよりもユーザに対する負荷が少なく、ユーザの耳に近い位置でツイートを読み上げることができるため親近感と存在感をユーザに感じさせやすいと考えた。TEROOS は Bluetooth で通信を通してコマンドを送ることでモーション動作を行うことができる。モーションは次の種類を用意し、下記のコントローラアプリケーションから、ツイートの再生に

あわせて実行できるようにした。

- 眼球 (左/中央/右)
- まぶた (上/下)
- 首 (左/中央/右/前)
- ささやき動作
- 首の自由動作

## 2.4 コントローラアプリケーション

本研究で使用する TEROOS は GPS と Bluetooth, スピーカーを備えているため, ロボットモーションと音声ツイートの読み上げを行うコントローラアプリケーションを作成し, Bluetooth を搭載した小型 PC で実行した. ロボットのマイクには, PC の音声出力を接続している. コントローラアプリケーションは TEROOS への Bluetooth 通信経路の作成と, Twitter タイムラインの取得, ツイートの位置情報タグの取得, 音声ファイルのダウンロード, そしてツイートの読み上げを行う. ロボットのいる位置情報を GPS により取得し, ツイートに組み込まれた位置情報タグとのロボットの距離を計算し, ユーザが移動する際に距離が近くなったツイートのメッセージをロボットモーションと音声でユーザに伝達する.

タイムラインの取得時にすべてのツイートを取得してしまうことがないように, ハッシュタグと呼ばれるタグ機能を用いて Voicy システム用に投稿されたデータのみを取得している. 例えば以下のようなツイートが Android クライアントから投稿され, コントローラアプリケーションが取得する.

このとき, **audio:**の部分は音声認識サーバのアドレスに変換され, HTTP 経由で音声ファイルのダウンロードが可能になる.

```
ようこそ私たちのキャンパスへ #voicy audio:qB7x0e
```

これは次のフォーマットで作られている. 太字が置き換えられる部分である.

```
(Text message) #voicy audio:(short url)
```

## 2.5 行動デザイン

ブログロボットでは, ブログ記事の中にモーションタグを用いて記事の読み上げ中にロボットが行う動作を記述することができる. しかしながら, Twitter のようなシステムではタグの埋め込みを行うことで可読性が低下し, 制限文字数に達してしまう可能性が高い. 本研究ではテキストから位置に関する情報と, 句読点に関する情報を抽出し, 音声ファイルの時間を文字数で除算することでおおまかなロボットモーションのタイミングを推定している.

位置情報と句読点に該当する文字列を表??に挙げ, それぞれのロボットモーションを対

表 1 文字列パターンとロボットモーションの対応

パターン文字列	対応するロボットモーション
「右」	首を右に傾ける
「左」	首を左に傾ける
「そこ」	首を前に傾ける
「ここ」	ユーザの方に頭を向ける
「あそこ」	対象の方向に首を向ける
「,」「。」	まぶたを開閉する

応づけた. 句読点に対する対応付けはロボットの行動に意味を持たせるものではなく, モータが動作する音をユーザに聞かせることで会話における自然なジェスチャを表現した.

## 3. アプリケーション

### 3.1 キャンパス案内

慶應義塾大学理工学部の矢上キャンパスでデモを行った. ロボットをユーザが装着し, 散策を行う途中でロボットがツイートを再生するデモを行った. ツイートを生成する場所は5カ所を選定し, 以下の内容を音声ツイートとともに投稿した.

- (1) ようこそ私たちのキャンパスへ, 坂を上ると建物があります
- (2) ここは新しい校舎です, 階段を上ると研究棟に行けます
- (3) 右に行くと図書館, 左に行くと売店があります
- (4) この建物の4階に安西・今井研究室があります
- (5) この胸像の人物は理工学部前身の大学の創始者です
- (6) グラウンドでは毎年ソフトボール大会が開かれています
- (7) ここから東海道新幹線が見えます

## 4. 将来の展望

### 4.1 システムの拡張

Voicy システムは音声ツイートとロボットモーションによる実世界つばやきシステムとして機能するだけでなく, ユーザに対してテキストから音声, 音声からロボットモーションへの推移的な通信手段の遷移を促すシステムとして拡張する予定である. これにより, 日常的にユーザが使用しているテキストによる投稿では表現しきれない感情や状況といったものを音声によって伝えることが可能になり, また音声では伝えきれない情報をロボットを用い

たジェスチャ動作で表現することで、より実世界の人同士のコミュニケーションに近い形にすることができると考えられる。

#### 4.2 フィールドでの応用

アプリケーションにおいて行われたデモから、フィールドにおける応用例が考えられる。広範囲の移動が必要であり、また説明する対象が多く存在する史跡公園や歴史遺跡では、パンフレットのように紙媒体の説明では移動しながら確認することは困難である。また、電子端末を持ちながら説明ポイントを回ることはユーザの負担が大きい。そこで、そのような場所における説明エージェントとして本研究のシステムが利用される価値があると考えられる。

ほかに、博物館などの訪問客が多く、展示物の説明が読みにくいような状況であっても、ユーザに対してささやき動作を行うことで周囲の訪問客に対する配慮を行いながら説明を行うことが可能である。

### 参 考 文 献

- 1) Geeky Gadgets : Anybots QB Avatar Robot, 入手先(<http://www.geeky-gadgets.com/anybots-qb-avatar-robot-19-05-2010/>) (参照 2010-05-19).
- 2) TODAY Health : Robot avatar allows sick boy to go to school, 入手先([http://today.msnbc.msn.com/id/41641984/ns/today-today\\_health/t/robot-avatar-allows-sick-boy-go-school/](http://today.msnbc.msn.com/id/41641984/ns/today-today_health/t/robot-avatar-allows-sick-boy-go-school/)) (参照 2011-02-17).
- 3) Moda, M. et al.: Blog robot: a new style for accessing location-based contents, *HRI '09 Proceedings*, New York, NY, USA, ACM, Addison-Wesley, DOI:<http://doi.acm.org/10.1145/1514095.1514136> (2009).
- 4) タカラトミー : ついまる, 入手先([www.takaratomy.co.jp/products/twimal/](http://www.takaratomy.co.jp/products/twimal/)) (参照 2011-3).
- 5) DigInfo : TEROOS, 入手先(<http://jp.diginfo.tv/2011/03/25/11-0071-r-jp.php>) (参照 2011-03-23).
- 6) 河原達也, 連続音声認識ソフトウェア Julius, 人工知能学会誌, vol 20, page 41 (2005).