

音声コミュニケーションを可能とする 音声 Twitter クライアントシステムの開発

嶋田 浩希[†] 西崎 博光[†] 関口 芳廣[†]

[†] 山梨大学大学院医学工学総合教育部

1 はじめに

情報通信技術の発展により、インターネットを通じた様々なサービスが提供されている。特に、近年、ソーシャルメディアの利用者は、スマートフォン等の普及も相まって、Facebook や、Twitter などを利用するユーザが爆発的に増加している。

2011年3月11日に発生した東日本大震災においても、これらのソーシャルメディアの果たした役割は大きい。例えば、Twitter は、ユーザ同士の安否確認、公共交通機関の運行状況などの情報交換に使われ、有効活用した利用者が多かったという報道もなされた。

Twitter は、140文字以内の短文(つぶやき)を投稿・共有するサービスであり、情報伝播力やリアルタイム性が高く、最近では、マスメディアと連動して視聴者とのつながりを創り出すといった利用もされている。

現在、Twitter のサービスをより便利に利用可能にしたものや、音声の読み上げ機能を搭載したもの、提供するプラットフォームの操作を最適化したものなど、様々な特徴を持ったクライアントソフトウェアが多く提供されている。しかし、どのソフトウェアも投稿・共有可能なのは、Twitter にて提供されているサービスと同様にテキスト情報が主体で、感情などの表現を絵文字などで伝えるほかない。生の声そのものを投稿できれば、感情を直接的に反映させたつぶやきになると考えた。そこで、筆者らは、言語的な情報を持つ「文字」に加え、非言語的な情報を持つ「音声」を投稿・共有することができるクライアントシステムを開発した。

これまでに、パソコンの Web ブラウザ上で動作するアプリケーションとして、Twitter クライアントを構築し、公開した [1]。しかし、PC 上の利用では、音声認識を用いるよりもキーボード入力の方が速く入力できること、PC 用のマイクを用意しなくてはならないことなどから、携帯性が高く、別途マイクを用意する必要のないスマートフォン (Android OS) で動作する Twitter クライアントを開発した*。

本稿では、開発した音声 Twitter クライアントシステムについて紹介する。

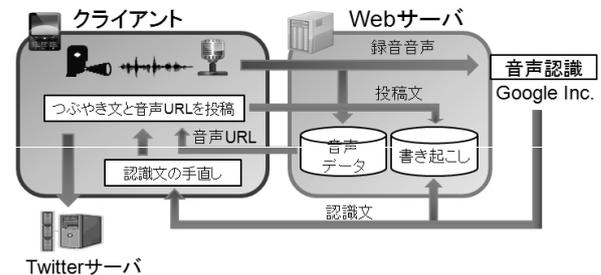


図1: 音声 Twitter クライアントシステムの処理の流れ

2 音声 Twitter クライアントシステム

2.1 システムの特徴

本システムは、主に3つの特徴を持つ。

1. 生の音声を Twitter に投稿できる
2. シームレスな音声再生が可能である
3. 音声認識を用いることで素早い文字入力を実現

既に、音声インタフェースを備えた Twitter クライアントは複数提供されているが、それらは音声認識を利用した文字入力支援や音声読み上げなどを行うのみで、音声そのものを投稿したり、同じクライアント上で投稿された音声の再生を行うことはできない。本システムでは、音声データをサーバに保存することで、それへのリンク情報 (URL) を投稿でき、システムに音声再生プレイヤーを付与することで音声の双方向コミュニケーションを可能とした。

音声でコミュニケーションを行えるため、テキスト情報だけでは困難だった感情や臨場感などが伝えやすくなる。また、災害などの緊急時に時間のかかる文字入力を行うことなく、音声を投稿することで、正確で素早い情報伝達が可能になる。このように、音声情報の伝達を容易に行える本システムを利用することで、様々なシーンで応用することができる。

2.2 システムの処理の概要

本システムの基本的な処理の流れについて説明する。図1に処理の概要を示す。

ユーザは端末のマイクに向かって「つぶやき」たい音声を発声する。発声された音声は、Google 社が提供する大語彙連続音声認識 (ASR) エンジンによって音声認識されると同時に、音声データとして、筆者らが提供する Web サーバにも保存される。音声認識の結果も、ユーザの携帯端末に返されると同時に、書き起こしとして、Web サーバにも保存される。

ユーザは、返された音声認識結果もしくは、Web サーバに保存されている音声へのリンク情報を編集後、

Development of a voice Twitter client system for speech communication

[†] Hiroki Shimada (roki@alps-lab.org)

[†] Hiromitsu NISHIZAKI (hnishi@yamanashi.ac.jp)

[†] Yoshihiro SEKIGUCHI (sekiguti@yamanashi.ac.jp)

Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi ([†])

* 現在, <http://www.alps-lab.org/alpstweet/>にて公開中。



図 2: 録音(左図)・再生(右図) インタフェース

Twitterへ投稿することができる。この際、Twitterに投稿されたテキストも書き起こしとして、Webサーバに保存される。なお、音声データへのリンク情報を投稿しない場合（音声データをTwitterにポストしない）でも、音声データはWebサーバに保存されており、かつ、その音声認識結果も紐付けられている。最終的に、発話された音声データとその音声認識されたテキスト（修正されたテキストを含む）をWebサーバ上で関連付けておくことで、タグ付きの音声データとして収集してゆく。

2.3 音声インタフェースの概要

本システムの音声インタフェースの使用例を説明する。図1の左側に録音インタフェースを、右側に再生インタフェースを示す。

● 録音インタフェース

つぶやき編集画面において、音声の録音を行う。音声の録音時には、ダイアログが表示され、進捗テキスト、録音時間、音量バーなどにより、ユーザは音声入力を目視で確認することができる。

録音が終了すると、バックグラウンドで音声認識および認識結果の取得が行われ、ユーザは録音後すぐにつぶやき編集作業が再開できる。

認識結果の取得が完了すると、画面下部のメニューの音声再生、認識結果・リンク情報挿入が選択可能になる。音声再生メニューにより、自身の録音した音声を投稿前に聞くことができる。認識結果・リンク情報挿入メニューにより、得られた認識結果や録音音声のURLをつぶやき編集フォームに挿入することができる。

得られた認識結果のままTwitterに投稿されるのではなく、ユーザは得られた結果から自身の判断で修正作業等が行える。また、録音音声が入らなければ、何度でも音声録音・認識を行うことも可能であり、各音声と対応した認識結果・リンク情報をいつでも参照することもできる。

● 再生インタフェース

ユーザのフォローしたアカウントのつぶやきが表示されるタイムラインと呼ばれる表示画面において、投稿された音声の再生を行う。

表示されているつぶやきから、音声のリンク情報とハッシュタグ「#alpstweet」が含まれたものを長押しタップすると、音声の再生ができる。

画面上部に、選択したつぶやきの音声再生プレイヤーが表示され、音声のダウンロード完了後、自動で音声の再生が行われる。なお、音声のダウンロード、音声の再生ともにバックグラウンドで行われるため、ユーザの操作を妨げることはなく、処理中につぶやきの閲覧を行うことができる。

3 音声データの収集

音声・自然言語処理の研究では、大量の書き起こし付きの音声データが必要となる。インターネットの普及によりWebを通じて大量のテキストデータを入手するのは容易になってきた。また音声データも以前より容易に集めることができる環境が整ってきた一方で、それらの書き起こしやタグ等を付与するには非常に手間とコストがかかることから、既存の整備された音声コーパスを使わざるを得ないのが現状である。

本システムでは、音声認識を行ったユーザの音声データをサーバに保存するため、実環境の音声データを大量に入手することが可能である。この音声データとユーザがTwitterに投稿したテキストデータを収集し、対応付けを行うことで、容易に書き起こし付きの音声データを作成できる。この際の書き起こしは、Twitterに投稿したテキストを利用するため、ユーザ自身が認識誤りを訂正する可能性が高く、精度の高い書き起こしの収集が期待できる。

4 おわりに

本稿では、開発とした音声Twitterクライアントシステムを紹介した。

本システムは、ユーザが音声認識を行う音声をサーバに保存し、リンク情報を用いた音声投稿機能と投稿された音声をタイムライン上でスムーズに再生する音声再生機能を提供することで、ユーザ間の双方向コミュニケーションを実現した。また、投稿されたテキスト情報を収集することで、実環境の音声データとそれに対応する書き起こしを容易に作成することができるため、音声・自然言語処理の研究におけるコーパスの作成に寄与すると考えている。

今後は、インタフェースの評価を行い、使い勝手の向上を目指す。

参考文献

- [1] 嶋田浩希ほか, “ソーシャルメディア「Twitter」を利用した音声データ収集の試み”, 電子情報通信学会, 言語理解とコミュニケーション研究会, 電子情報通信学会技術報告, NLC2012-61, pp.31-36, 2012