

双方向音声案内デジタルサイネージのための学内イベント登録システム

山本 大介*1*2, 大浦 圭一郎*1*2, 李 晃伸*1*2, 打矢 隆弘*1*2,

内匠 逸*1*2, 徳田 恵一*1*2, 松尾 啓志*1

*1 名古屋工業大学 工学研究科 *2 CREST, JST

yamamoto.daisuke@nitech.ac.jp

概要：名古屋工業大学では、2011年4月から双方向音声案内デジタルサイネージ「メイちゃん」を校門前に設置した。「メイちゃん」は等身大のキャラクターを大型ディスプレイに表示させることによって、臨場感がある音声対話を行うことが可能なシステムである。我々は「メイちゃん」を用いて学内イベントを掲示・案内可能な Web システムを開発した。具体的には、対話用の掲載画像・掲示期間・応答文などを投稿可能であり、これらの手法や利用状況について報告する。

1 はじめに

近年、大型ディスプレイを用いたデジタルサイネージシステムが急速に普及している。デジタルサイネージは店舗や公共施設、学校内などに設置され、商品やサービスの広告だけでなく、列車の運行状況や学生の呼び出し情報など、社会的にも重要な情報も掲示している。また、タッチパネル技術を用いたインタラクティブなデジタルサイネージシステムも登場しつつあるが、不特定多数のユーザが利用することによって、ディスプレイが指紋で汚れやすく、美観や衛生面に抵抗感を感じる人も多い。また、必要な情報を検索する際に任意のキーワードを入力するためにはソフトウェアキーボード等が必要になるが、大型ディスプレイにソフトウェアキーボードが表示されるのは使いづらく、あまり実用化されていない。

そこで、音声対話技術を用いたデジタルサイネージ技術が注目されつつある[1]。音声対話は上記の問題が解決可能になるだけでなく、人間にとって最も原始的で潜在的に使いやすいインタフェースの一つである。その一方で、音声対話技術をよりよく使いこなすためには、注意すべき事項がいくつかある。本研究では、音声対話技術特有の注意すべき事項を考慮した、双方向音声案内デジタルサイネージのための学内イベント情報を登録するための Web システムに

ついて提案する。

2 双方向音声案内デジタルサイネージ

2.1 ハードウェア

本研究で利用する双方向音声案内デジタルサイネージを図1に示す。70インチの液晶ディスプレイを2台垂直に並べ、その前方にマイクブースが設置されている。また、マイクブースの下部には人感センサーが設置されており、ユーザがマイクの前に来たかどうかを感知することが出来る。さらに、デジタルサイネージの上部にカメラを備え、顔認識技術等を用いて複数のユーザの状況を認識することが可能になる。ディスプレイには等身大の仮想キャラクター「メイちゃん」が表示される。本ハードウェアは、本学の正門前の屋外に設置され、多くの学生や教職員、来校者が利用可能である。



図1 双方向音声案内デジタルサイネージ「メイちゃん」

2.2 ソフトウェア

音声対話エージェントシステムとして MMDAgent[2]が採用されている。MMDAgentとは、音声対話のための高度な機能を備えた基盤ソフトウェアであり、音声認識、音声合成、3Dモデルの描画や制御を統合したシステムである。MMDAgentは、音声対話を記述した専用のスクリプト言語に基づいて音声対話の処理を実行する。スクリプト言語は FST (Finite State Transducer)形式に基づいており、音声認識による特定のキーワードの認識や、センサーによるユーザの検出をトリガーとして、次の状態に遷移する有限オートマトン形式を記述可能である。

FSTは比較的低レベルでかつ詳細な情報を記述可能な形式であるため、詳細な表情やしぐさ、対話などを記述できる等の柔軟性が高い一方、スクリプトは人間にとって理解しにくく、一般ユーザが手書きで記述することは困難である。そこで、FSTをWebインタフェースを用いて容易に記述できる仕組みが必要であった。

3 音声対話システムのためのイベント登録システム

3.1 概要

名古屋工業大学では、従来からWebでイベント情報を共有する「イベントカレンダー」というシステムを運用してきた。提案システムは、イベントカレンダーを拡張する形で、双方向音声案内デジタルサイネージ「メイちゃん」にもイベント情報を掲示できるシステムである。これにより、ユーザは音声対話インタフェースを用いてイベント情報の獲得が可能になる。

イベントの登録方法の概要は以下のとおりである。まず、教職員は、Webインタフェースを用いてイベント情報を登録する。登録可能な情報はイベントの日時や内容、パネル画像・発話音声などである。登録されたイベント情報は、管理者によってその内容をチェックされた後に、掲示期間の間、自動的に双方向音声案内デジ

タルサイネージによって掲示される。ユーザは、その掲示に関連したキーワードを含む対話をすることによって、その内容を確認することができる。

3.2 システム構成

提案システムのシステム構成を図2に示す。Web Serverでは、ユーザがテキストや画像を用いてイベント情報を登録することが可能である。Web Serverは一時間毎に登録された情報に基づいて動的FSTファイルを生成する。動的FSTファイルはデジタルサイネージに送信され、デジタルサイネージ側で、一時間毎に動的FSTと一般的な動作を記載した静的FSTを結合し、FSTの再読み込みを行う。また、教員情報などが登録されたデータベースからも、必要な情報を獲得し、教員案内のためのFSTの動的生成も行っている。なお、Web Serverが置かれているセグメントとデジタルサイネージが置かれているセグメントは異なるのでVPNで接続している。

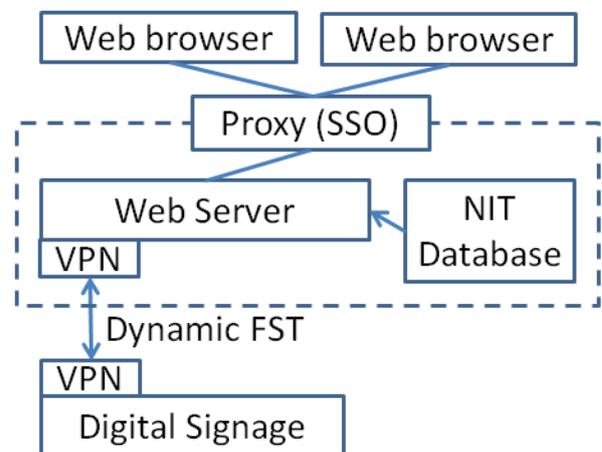


図2 システム構成図

3.3 音声対話インタフェースによる注意事項

学内イベントシステムにおいて、音声対話インタフェースをより効果的に扱う上で注意すべき事項をまとめる。

- 1 人間が期待する音声対話は本質的に自由度が高いため、それらの全ての場合に対応したFSTを用意することは極めて困難である。
- 2 音声認識の認識率を向上させるためには、事

前に認識すべきキーワードの辞書を作成する必要がある。

3 音声認識用キーワードに対して適切な「読み」と「品詞」を与えなければならない。

4 一般に、複雑な情報を音声のみで伝えることは難しい。

3.4 対話モデル

対話モデルとして、一問一答モデルを採用した。一問一答モデルでは、システムは初期状態として、音声入力の待ち状態になる。それぞれのイベント毎に用意された複数のキーワードのうちの一つ以上のキーワードのユーザ発話があった場合にそのイベントが選択されたと判断する。そして、そのイベント内容を示すパネル画像と「メイちゃん」の音声合成される。その後、初期状態である音声入力の待ち状態に戻り、その対話が終了する。

一般的な音声対話システムはエージェントとユーザの掛け合いによって会話が進展していくが、そのような状況を想定したスクリプトは複雑になりやすい(注意事項1に対応)。また、多種多様なスクリプトの記述によってスクリプト自体の矛盾や依存関係の破たんを引き起こす可能性がある。そのため、一問一答方式を採用した。

登録できるイベント情報は、タイトル、イベント期間、掲示期間、パネル画像、「メイちゃん」発話音声用文字列、キーワードである。音声だけでなくパネル画像も合わせて表示することによってユーザに理解を促す(注意事項4に対応)。

4 Web インタフェース

提案システムは、Web インタフェースを用いてイベントの登録が可能である。

4.1 イベント登録

イベント情報の登録画面を図3に示す。入力できる項目は大きく分けて二つある。一つは従来のイベントカレンダーが備えていたタイトルや本文、イベント期間などの情報である。もう一つは、「メイちゃん」に登録するための情報で

あり、パネル画像や音声発話情報、音声認識用のキーワードなどを登録することが可能である。パネル画像は既存の画像をアップロードする方式と、登録内容のテキスト情報から定型フォーマットで自動的に生成する方式の2種類がある。

図3 イベント登録インタフェース。上側が基本情報、下側がメイちゃん固有情報。

4.2 音声認識キーワード

次に、音声認識の認識率を向上させるために(注意事項2,3に対応)、音声認識用のキーワードを設定する必要がある。図4に音声認識用のキーワード設定画面を示す。4.1節で入力されたキーワードに対して、それぞれ、キーワードの読みと品詞を設定可能である。キーワードの読みと品詞としては、システムが辞書に基づいて自動的に認識可能であるが、認識不能な固有名詞等は、名前、人名、先生名、地名、組織名、イベント名、その他固有名詞から選択できる。これによって、音声認識の精度を向上させる。

キーワード	読み (カタカナ)	ジャンル
シンポジウム	シンポジウム	(システム辞書有)
都市水害	トスイガイ	その他固有名詞
秀島	ヒデシマ	先生名
日比	ヒビ	組織名

確定

Copyright (c) 名古屋工業大学 2011-, All Rights Reserved.

図4 音声認識キーワード設定インタフェース

4.3 管理者による承認

管理者は、認証画面において、イベント情報の承認と修正が可能である。タイトルや掲示期間などの他に、パネル画像やメイちゃんの発話音声や音声として確認することが可能である。またそれぞれのイベントに対して優先度の設定も可能である。

4.4 イベント情報の FST の自動生成

イベントカレンダーに登録された情報に基づいて FST ファイルを自動生成可能である。一定時間毎に、データベースの情報を参照して FST を生成し、サイネージに転送する。

4.5 教員検索性 FST の自動生成

教員用のデータベースを参照することによって、教員情報を検索するための FST を自動生成する。先生の名前を音声でユーザが発話すると、その先生の情報をパネルで提示するシステムである。一般に先生の名前をフルネームで呼ぶことは少なく、たとえば、「鈴木先生について教えてください」などと苗字のみで呼ばれる場合も多い。そのため、鈴木先生は複数いる場合が多く、本来ならば、複数の鈴木先生から目的の鈴木先生を探す対話をしなければならない。しかしながら、現状の一问一答方式の対話モデルではこれができないため、鈴木先生の一覧をあらかじめ表形式で作っておいて一覧表を提示する方式で対応した。

5. 運用と考察

4月1日から提案 Web システムの試験運用を開始し、6月1日からは全教職員にシステムを公開した。本稿では主に、Web インタフェースの側から提案システムを検証する。

イベントの開催月別(4月～11月)のイベント情報の登録状況を表1に示す。提案システムは従来から運用しているイベントカレンダーを拡張する形で「メイちゃん」に掲示可能なシステムである。2011年10月17日現在において、従来のイベントカレンダーと同様に Web のみに情報を掲示した数は、132個である。メイち

ちゃんにも掲示したイベント情報は、95個である。従来、教職員が主に利用しているイベントカレンダーを拡張する形で作ったため、教職員も積極的に利用していることがわかる。このように、提案システムは、イベント情報を掲示したい側の要求にこたえることが出来ているといえる。

表1 月別のイベント登録状況

	Web	サイネージ	合計
4月	---	7	7
5月	---	1	1
6月	17	9	26
7月	24	11	35
8月	16	23	39
9月	32	19	51
10月	24	14	38
11月	19	11	30

6 おわりに

本稿では、従来の Web ベースのイベントカレンダーを拡張する形で、双方向音声案内システムに掲示可能なイベント登録システムを開発した。音声対話モデルとしては、コンテンツ作成側にとって簡便な一问一答方式を採用した。さらに、実際に運用することによって、教職員が手軽に利用可能なシステムであることが示された。

今後の課題としては、学生によるイベントの登録や、より高度な音声対話スクリプトを登録できるようにする。また、携帯端末への対応や、学外からのアクセスにも対応できるように検討したい。

参考文献

- [1] 川出 陽一, 「双方向音声案内デジタルサイネージ」, 印刷雑誌, Vo.94, No.10, pp.25-29. 2011.
- [2] 李 晃伸, 大浦圭一郎, 徳田恵一. 「魅力ある音声インタラクションシステムを構築するためのオープンソースツールキット MMDAgent」, 第13回音声言語シンポジウム, to appear, 2011.12.