

音声情報案内システム「キタロボ」と「キタちゃん」の運用

川波 弘道 ツィンツアレク トビアス 竹内 翔大 猿渡 洋 鹿野 清宏

奈良先端科学技術大学院大学 情報科学研究科

kawanami@is.naist.jp

1. はじめに

2002 年 11 月より奈良県生駒市市民センター内で音声情報案内システム「たけまるくん」[1]を運用している。この開発と運用の成果を活かし、2006 年 3 月より鉄道駅構内に 2 種類のシステム「キタロボ」「キタちゃん」を運用している。

駅構内は雑音も多様であり、幅広いユーザ層の利用が期待できる。また「キタロボ」にはロボット型ボディを与えている。従来の CG エージェントを持つ「キタちゃん」とユーザ発話の比較を行うことでインタフェイスが与える影響の比較分析が可能である。本稿ではシステムの概要とこれまでの運用結果の概略を報告する。

2. 「キタロボ」と「キタちゃん」

「キタちゃん」(図 1 右)は学研北生駒駅(図 1 左、近畿日本鉄道、奈良県生駒市)のマスコットキャラクターである。これを CG エージェントとした音声案内システム「キタちゃん」と、高さ約 1.3 m で三次元化したロボット型ボディを持つ「キタロボ(キタちゃんロボット)」を開発した(図 2)。これらは駅構内の改札口の横に並べて運用している[2]。

2.1 システム構成およびサービス

基本的な構成および音声認識処理、応答生成処理は両システムとも「たけまるくん」に準ずる。ただし「キタロボ」は人と話すようにハンズフリー対話が目標であるためマウスなどの入力デバイスは使用しない(図 3)。音声認識部は「たけまるくん」で有効であった GMM に基づく雑音棄却処理と大人用と子ども用の音響・言語モデルを用いた並列デコーディングによる大人子ども判別処理も導入している。また「たけまるくん」では記述文法による認識は特定タスク(市内名勝に関する特設ページの案内)のために補助的に用いていたが、新システムでは主要なサービスとして乗換案内(train)、周辺地図表示(local)、Web ブラウザ操作(web-navi)を記述文法で実現している。駅や周辺施設の案内、観光案内、ニュース、「キタちゃん」プロフィールや挨拶などの一般的な質問は N-gram による認識結果を質問用例データベースと照合することで対応する。

2.2 設置環境

設置場所は外部とのしきりが無い。雑音レベルは約 60dBa であり、風雨の音の影響を受ける。またロータリーに面しているためバスのクラクションや車の排気音が入力される。しかしながら、これらの雑音も「たけまるくん」収集データから作成した雑音 GMM で棄却可能であった。また指向性マイクを使用した結果、振幅と零交叉

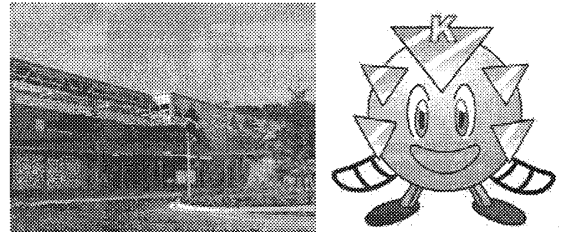


図 1. 学研北生駒駅とマスコットキャラクター「キタちゃん」

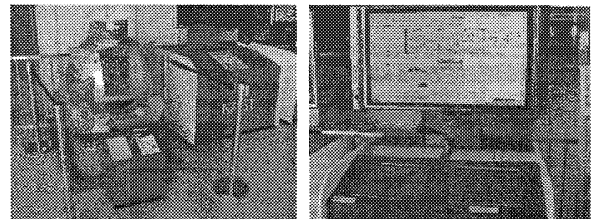


図 2. 外観(左:キタロボ, 右:キタちゃん)

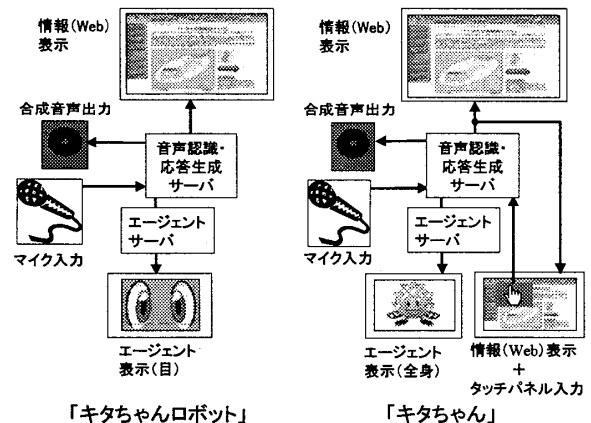


図 3 インタフェイス(左:キタロボ, 右:キタちゃん)

数の閾値を用いた従来の発話区間検出が有効に機能した。そのため運用の初期段階においても周囲の雑音に対してシステムが応答してしまうという問題はほとんど起こっていない。

3. データベース

3.1 システム入力データ

図 4, 5 に「キタロボ」と「キタちゃん」へのシステム入力数とその自動判別結果による内訳を月ごとに示す。「noise」は非音声として棄却された数、「child」、「adult」の数は並列認識による判別結果に基づく。識別誤りを含むが大局的な傾向を観察することができる。「キタロボ」の入力数はほぼ「キタちゃん」の 2 倍で推移している。「noise」の絶対数も「キタロボ」に多く見

Operation of speech-oriented information guidance systems, "Kita-robot" and "Kita-chan" by Hiromichi Kawanami, Tobias Cincarek, Shota Takeuchi, Hiroshi Saruwatari and Kiyohiro Shikano. (Nara Institute of Science and Technology, Graduate School of Information Science)

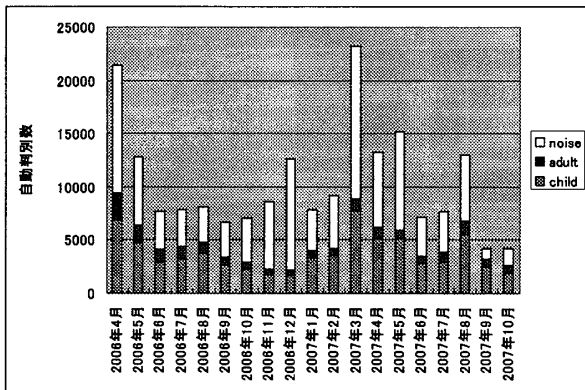


図4. 「キタロボ」への入力数と自動判別結果

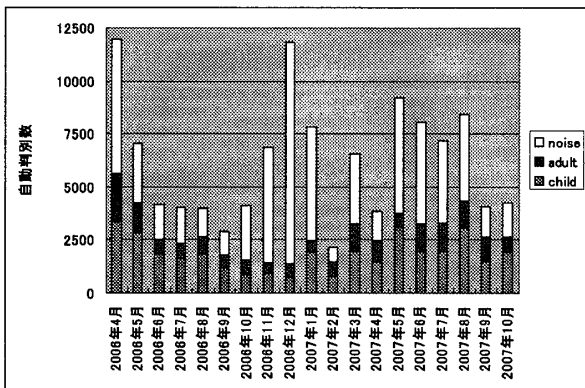


図5. 「キタちゃん」への入力数と自動判別結果

られるが、これは2つのシステムの配置の関係上風雨の影響を受けやすいためと考えられる。

3. 2 書き起こしデータベース

運用開始から2007年1月までの10ヶ月間の入力データに対して5名の技術補佐員が聴取により書き起こしと年齢層および性別ラベルと雑音タグの付与を行った。年齢層のカテゴリは、<幼児/低学年子ども/高学年子ども/大人/高齢者/不明>である。図6, 7に両システムのユーザ年齢層の割合を示す。聴取によりユーザ発話から、ひとり言や歌声などシステム応答が不要と判断した発話(無効発話)を除いた有効発話における割合である。有効発話の総和は「キタロボ」が32,350発話、「キタちゃん」が16,935発話である。「キタロボ」に比べて「キタちゃん」の発話は年齢層が相対的に高いが、発話数の絶対数はどの層でもほぼ同等(高学年子ども)か「キタロボ」への入力が多く、ロボット型がユーザの興味を引いていることが分かる。

4. おわりに

駅での運用開始から2年弱が経過した「キタロボ」「キタちゃん」の概要とその運用状況について紹介した。音声案内システムが広く普及することは、誰でも分け隔てなく情報にアクセスできる環境を実現する手段のひとつとして期待できる。そこで筆者らは既存システム(ここでは「たけまるくん」)のデータの再利用性の定量的評価を行っている[3]。新規にシステムを構築する際の性能とコストとの関係を示すことによりシステム導入のバリアを低くできる。また、ユーザがシステムとより気軽に親しみやすく対話できるようにインタフェイスの

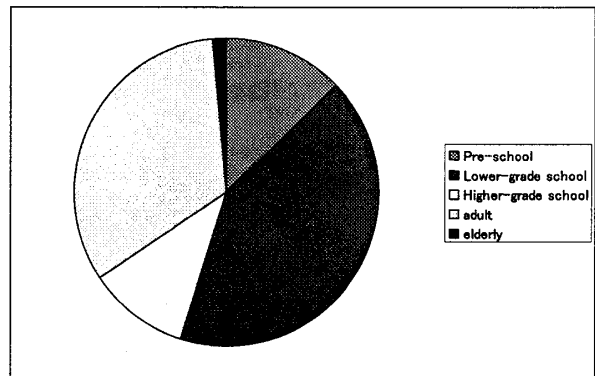


図6. 「キタロボ」のユーザ年齢層の割合(聴取による)

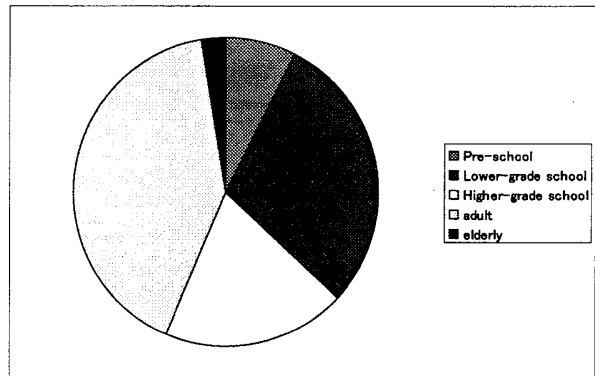


図7. 「キタちゃん」のユーザ年齢層の割合(聴取による)

改善も必要である。そこで指向性マイクを必要としない音声区間検出技術[4]、ロボット型ボディの「キタロボ」にマイクアレイを導入したハンズフリー入力対話システムの開発も進めている[5]。

謝辞

本研究の一部は文部科学省「e-Society 基盤ソフトウェアの統合開発プロジェクト」の援助を受けた。

参考文献

- [1] 西村竜一 他, “実環境プラットフォームとしての音声情報案内システムの運用,” 信学論, vol. J87-D-II, no. 3, pp. 789-798 (2004-3)
- [2] 川波弘道 他, “駅構内音声案内システム「キタちゃん」「キタちゃんロボット」の開発,” 信学技報, vol. 106, no. 123, pp. 19-24 (2006-6)
- [3] T. Cincarek, et al., “Development, Long Term Operation and Portability of a Real-Environment Speech-Oriented Guidance System”, IEIE Trans Inf. & Sys. (to be published in Mar. 2008)
- [4] 酒井啓行 他, “音響モデルと言語モデルに基づく音声区間検出を用いたハンズフリー音声認識アルゴリズムの評価,” 信学技報, vol. 107, no. 406, pp. 13-18 (2007-12)
- [5] 猿渡洋 他, “ハンズフリーロボット対話実験システムの構築”(本予稿集収録)