

# スマートフォン単体で動作する音声対話3Dエージェント 「スマートメイちゃん」の開発

山本大介<sup>†1†2</sup> 大浦圭一郎<sup>†1†2</sup> 西村良太<sup>†1†2</sup> 打矢隆弘<sup>†1†2</sup>  
内匠逸<sup>†1†2</sup> 李晃伸<sup>†1†2</sup> 徳田恵一<sup>†1†2</sup>

本研究では、スマートフォン単体で動作する音声対話3Dエージェントシステムを開発した。従来のスマートフォン向け音声対話システムとは違って、提案システムでは、スマートフォン内で音声認識や音声合成などの処理を完結することができるため、ネットワークに起因する遅延のない自然な音声対話が可能になった。さらに、スマートフォンに適した音声インタフェースを試作し、提案手法の評価を行った。

## Development of a Voice Interaction Agent using 3D Modeling for Stand-alone Smartphones

DAISUKE YAMAMOTO<sup>†1†2</sup> KEIICHIRO OURA<sup>†1†2</sup>  
RYOTA NISHIMURA<sup>†1†2</sup> TAKAHIRO UCHIYA<sup>†1†2</sup> ICHI TAKUMI<sup>†1†2</sup>  
AKINOBU LEE<sup>†1†2</sup> KEIICHI TOKUDA<sup>†1†2</sup>

In this research, we propose a voice interaction system using 3D modeling for stand-alone smartphones. Since this system can handle speech recognition and speech synthesis on a standalone smartphone differently from existing voice interaction systems for smartphones, this system enables us to talk naturally without delaying of network communications. Moreover, we develop the effective voice interface for smartphones, and discuss the proposed methods.

### 1. はじめに

近年、Apple の Siri [1]やNTT ドコモのしゃべってコンシェル[2]など、スマートフォン向けの音声対話システムが急速に普及しつつある。これらの音声対話システムでは、ユーザは仮定の秘書エージェントとの音声対話を通じて、経路案内や天気予報などの情報を得ることができると同時に、簡単な雑談なども行うことができる。このように、スマートフォンにおける音声対話技術に対するユーザの潜在的な需要は非常に大きい。

その一方で、これらのスマートフォン向けの音声対話システムでは、3D キャラクタなどが表示されていない。その理由はいくつか考えられるが、そもそも不要であるという考えと、何らかの技術的な問題があるという二点に集約できると考えている。特に、後述するように、既存のクラウド連携型スマートフォン向け音声対話システムはネットワーク通信に起因する対話の遅延が大きいため、3D キャラクタと遅延の無い自然な対話を実現することは容易ではない。

そこで、本論文の目的は、3D キャラクタと、より自然で、音声対話の遅延が少なく、スマートフォン単体で動作する音声対話3Dエージェントシステムの構築をするとともに、それを実現する上での問題点や、応用システムに関する検討を行う。

具体的には、ネットワーク通信に起因する遅延を無くすためには、スマートフォン内で音声認識や音声合成などの処理を完結させる必要がある。そこで、既存の PC 向け音声対話ツールキットである MMDAgent [3]をスマートフォン専用 OS である Android 向けに移植し、スマートフォンに適したインタフェースの構築を行う。

### 2. 問題点

#### 2.1 音声対話処理方式の検討

Siri やしゃべってコンシェルなど、既存のクラウド連携型スマートフォン向けの音声対話方式（以下、クラウド方式と呼ぶ）では、音声対話の応答時間が人間の対話に比べて長いという問題がある。なぜならば、クライアント側でユーザの音声を録音した後にサーバ（クラウド）側に転送し、さらに、サーバ上で音声認識処理と対話制御処理を施した後にクライアント側に認識結果を返す処理が必要になる。これらの処理は並列に実行できないため、音声対話に関する処理時間（遅延）が累積してしまい、結果として応答時間が長くなる。

一般に、自然な音声対話システムを実現するためには応答時間の短さは重要な要素[4]であり、許容される応答時間は2秒未満である[5][6]。現状のクラウド方式の音声対話システムでは数秒の遅延がかかることも多いが、これらのシステムのようなコマンド指令型インタフェースでは応答時間が数秒かかっても許容されるかもしれない。しかしながら、より自然な 3D キャラクタを表示した対話型インタフ

†1 名古屋工業大学

Nagoya Institute of Technology

†2 独立行政法人科学技術振興機構, CREST  
JST, CREST

ェースであればあるほど、遅延が際立って不自然に感じてしまう可能性がある。

そこで、音声対話に関する応答時間を最小化するために、以下の3つの方式を考えた。

- クラウド方式：クライアント側で音声録音し、サーバ側に転送して音声認識処理と対話制御処理を実施し、その結果をクライアントに返す方式。
- ストリーミング方式：クライアント側の音声をストリーミング形式でサーバ側に転送しつつ、サーバ側でリアルタイムに音声認識処理と対話制御処理を実施する方式。
- スタンドアロン方式：クライアント内で、リアルタイムに音声認識処理や対話制御処理、音声合成処理を行う方式。

図1と表1で示すように、それぞれの方式にはメリットとデメリットが存在する。

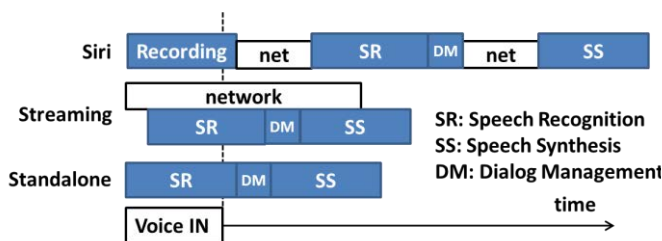


図1 音声対話方式の違いによる応答時間の比較. netはネットワーク通信に関わる遅延時間を表す。

Figure 1 Response time of voice interaction for each method. Net means the delay of network communication.

表1 スマートフォン向け音声対話方式

Table 1 Voice Interaction method for smartphones.

	Cloud	Streaming	Standalone
Cost of client	good	good	poor
Cost of network	fair	poor	good
Cost of server	poor	poor	nice
Delay of interaction	poor	good	nice

クラウド方式のメリットは、クライアント側の計算コストが小さいため、安価なスマートフォンなど、低速な端末に適している。その一方で、ユーザ数が増えるとサーバ側の負荷が大きく、また、応答時間も長くなる。Siri やしゃべってコンシェルはこれに該当する。

ストリーミング方式のメリットは、クライアント側の計算コストが小さく、応答時間も比較的短い。ただし、ネットワーク通信に起因する遅延は若干ある。その一方で、サーバ側の負荷が大きく、また、ストリーミング形式で常に音声データを転送する必要があるためネットワークの負荷も大きい。なお、我々は、ストリーミング方式の音声対話

システムとして、モバイルメイちゃん[7]を提案している。

スタンドアロン方式のメリットは、サーバ側の負荷が無く、ネットワークに起因する遅延が無い。その一方で、クライアント側の負荷が大きく、実時間での音声認識ができない場合は遅延が発生する。しかしながら、近年のスマートフォンの性能向上の速度は早いため、時間が解決する問題であると考えている。

## 2.2 音声対話ツールキット MMDAgent

MMDAgent は、音声対話のための高度な機能を備えたツールキットであり、音声認識、音声合成、3D モデルの描画や物理演算などを統合したシステムである。音声認識エンジンとして Julius [8]を、音声合成エンジンとして OpenJTalk [9]を、3D モデルとして MikuMikuDance 形式 [10]を、物理演算エンジンとして Bullet Physics [11]を採用している。MMDAgent は、音声対話を記述するため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などをトリガーとして、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。

MMDAgent は実時間かつ低遅延で音声対話を行うことができ、かつ、3D キャラクタの描画などが可能など、本研究の目的を満たすため、MMDAgent を採用した。しかしながら、MMDAgent は PC 向けのソフトウェアであり、Windows や Linux, Mac OS X での利用はできるが、スマートフォンでは利用できない。そこで、MMDAgent をスマートフォン(Android OS)に移植することを考えた。

## 3. 実装

MMDAgent をスマートフォンに移植するにあたって、いくつかの課題に直面した。

まず、MMDAgent は C++ で記載されているが、Android OS は基本的に Java、iOS は基本的に Objective C でプログラミングが記述されることを想定したシステムである。しかしながら、Android OS では Android NDK(Native Developer Kit) [12]を用いることによって、C++ でプログラムを記述することが可能である。また、Android OS は Linux ベースの OS であるため、Linux 向けのソフトウェアが移植しやすいという特徴がある。そこで、Android NDK を活用することによって、Linux 版の MMDAgent を Android OS へ移植した。ただし、Android OS の OS 機能と連携する一部の機能(メール、カレンダー、ライブ壁紙、アプリ起動など)は Java からしか利用できないため、図2に示すように、MMDAgent の内部メッセージを共用する方式で、Java で記述したプログラムと C++ で記述したプログラムを連携させている。また、音声入出力部分も、Android OS 独自の OpenSL ES [13]を用いた処理に変更した。

また、スマートフォンの処理能力は PC よりも劣るため、

いくつかの軽微な仕様変更を行った。音声認識エンジン Julius は、音響モデルをより簡易なもの(IPA 版 PTM モデル)に変更した。ただし、言語モデルは同一であり、認識できる語彙数は Web から学習された約 6 万語[14]と、PC 版 MMDAgent と同一である。音声合成エンジン Open JTalk は PC 版と同一の設定で動作させている。また、3D レンダリング時に影の処理を省略した。また、スマートフォンでは、スピーカで再生された音がマイクに回り込み、音声認識の誤動作となることが多いため、エージェントの音声再生されているときはマイク入力を遮断した。そのため、バージョン機能などは利用できない。

いくつかの仕様変更を行ったが、同一のスクリプトと素材を用いて、PC 版の MMDAgent と Android 版の MMDAgent で同じ動作を実現可能である。また、Java と連携することにより Android OS の機能を活用したシステムの実装が可能になった。

また、Android のホーム画面の壁紙として実装することにより、図 3 に示すように、Android OS のホーム画面と一体感あるシステムとして利用することも可能である。

なお、音声認識処理は電力を多く消費するため、電力消費量を抑えるために、他のアプリを実行中や端末のスリープ中にはこれらの機能を停止し、利用できないようにした。一般的な利用状況では、ホーム画面を見ている時間はそれほど多くないため、電力消費量の問題は顕在化しにくいと考えている。

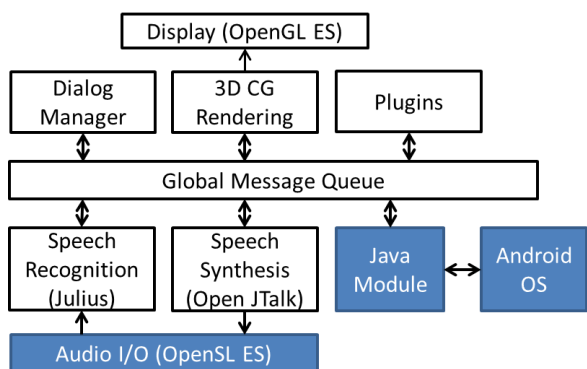


図 2 Android 版 MMDAgent のシステム構成図。Android OS と連携するため、Java Module と OpenGL ES を用いた Audio I/O 周りを新たに実装した。

Figure 2 System configuration of MMDAgent for Android. We implement the Java module and Audio I/O module using OpenGL ES in order to use the functions of Android OS.

#### 4. インタフェース

提案システムは 2 種類の動作形態を用意した。一つは、通常のアプリとして動作するモードであり、もう一つは、図 3 に示すように、ホーム画面のライブ壁紙として動作す

るライブ壁紙モードである。ライブ壁紙とは、Android OS のホーム画面の背景として動作するプログラムのことである。主に、提案システムは、ライブ壁紙モードとして利用されることを想定している。

提案システムは、2.5 頭身の女性エージェント「メイちゃん」を画面の中央下に配置し、画面中央に音声入力と音声認識の状況を表す表示領域を、エージェントの胴体周りに利用可能な機能を表すアイコンを表示している。2.5 頭身のエージェントを表示する理由は、小さい画面のスマートフォンにおいても、表情や体の動きが分かりやすいように配慮したためである。なお、全ての画像は 3D コンピュータグラフィックで描画されており、二本指で画面をなぞると、3D 空間におけるカメラの視点を変更することができる。実装した個別の機能を以下に述べる。

#### 4.1 音声対話機能

エージェントが表示されているとき、画面に向かって話しかけると同時に音声認識処理が開始し、音声認識処理が終了すると同時に、応答文の音声合成とその再生処理を開始する。これにより、応答時間の短い、スムーズな対話が可能である。音声対話のシナリオは MMDAgent と同じ FST 形式で記載した。



図 3 提案システム「スマートメイちゃん」のライブ壁紙モードでの動作例

Figure 3 Picture of proposed system in live wall mode.

#### 4.2 音声対話休止機能

常に、音声認識を実行していると外部ノイズや話し声などに反応して、意図しない反応や音声再生が再生されてしまう問題がある。

そこで、音声対話機能を利用したくない時は、「バイバイ」と話しかけると、図 4(中央)に示すように、カメラの視点を変えて待機状態に移す。待機状態になると、「こんにちは

は」という音声を認識するまでは無反応になる。「こんにちは」と音声認識すると、カメラの視点を元に戻し、通常状態へ遷移する。

ただし、音声対話休止状態であっても、音声認識機能を停止させることはできない。なぜならば、通常状態へ復帰するための認識キーワードである「こんにちは」を認識する必要があるためである。

### 4.3 Android OS 連携機能

Android OS が管理しているサービス、たとえばカレンダー機能と連携することにより、本日の予定などを聞くことができる。たとえば、「本日の予定は？」と聞くと、Android OS のカレンダーの状態を確認し、「本日は 2 件の予定があります。10 時から会議、13 時からゼミがあります。」などと音声で教えてくれる秘書機能を実現している。

### 4.4 その他の機能

インターネットの天気予報サイトと連携することにより、図 4(右)に示すように、音声とパネルで天気を提示することができる。たとえば、「今日の天気は？」と聞くと、「今日の天気は晴れです」と教えてくれる。同様に、現在時刻や、占い、学内案内図の表示や雑談などの機能も実現している。



図 4 インタラクションの例。左から、通常状態、待機状態、天気予報のパネルを用いた対話。

Figure 4 Example of voice interaction on proposed system.

Normal status, Sleep status, interaction with a weather cast panel.

## 5. 実験と考察

本章では、提案システムにおける対話の応答時間や、ユーザビリティ、電力消費量などに関する評価を行う。主に、クラウド方式、ストリーミング方式、提案手法であるスタンドアローン方式の三つの方式を比較した。なお、クラウド方式には、Apple の Siri を、ストリーミング方式にはモバイルメイちゃんを、スタンドアローン方式には提案システムを採用した。それぞれの方式で利用した主なアプリケーションと端末を表 2 に示す。なお、通信ネットワークは 3G 回線を用いた。

表 2 実験に利用したアプリケーションと端末

Table 2 Applications and smartphones for experiments

	Application	Smartphone
Cloud method	Apple Siri	iPhone 4S
Streaming method	Mobile Mei-chan ( Skype )	iPhone 4S
Standalone method	Proposed system	Galaxy S3

### 5.1 応答時間の評価

音声対話の応答時間について検証する。本稿における応答時間とは、ユーザによる音声発話終了してから、システムが音声合成された音声の再生を開始するまでの経過時間のことをいう。対話文は「今日の名古屋の天気は？」とした。応答時間の計測方法は、ユーザがシステムを利用している状況をマイクで録音し、ユーザ発話とシステム音声の音声波形の間の無音区間の長さを 5 回計測しその平均値を求めた。

応答時間を表 3 に示す。クラウド方式では応答時間が 4.1 秒であり、音声対話の遅延を感じる原因である。それに対して、ストリーミング方式は 1.1 秒、スタンドアローン方式は 0.82 秒と十分に早い。ただし、クラウド方式(Siri)は提案システムにくらべて認識できる語彙数が多く、対話制御機能もより複雑であるため、そのコストを考慮する必要があるが、その検証は今後の課題としたい。

スタンドアローン方式の方がストリーミング方式よりも早い理由は、後者はネットワークを経由することによる遅延があるのに対して前者はその遅延が発生しないからである。また、現状のストリーミング方式は 1 台のサーバで 1 つのセッションしか扱っていないが、1 台のサーバで処理する音声対話のセッション数を増やせば増やすほど、サーバ側の負荷が上がり、遅延が大きくなる可能性がある。

なお、自由な連続音声認識処理を実現するためには、入力音声の終端を検出する必要があるが、そのために、スタンドアローン方式とストリーミング方式(MMDAgent)は 0.24 秒の無音区間、クラウド方式(Siri)は 0.8 秒前後の無音区間を検出する必要があるが、その時間も応答時間に含んでいる。

表 3 音声対話の平均応答時間

Table 3 Response time of voice interaction

	Response time
Cloud method	4.1 sec
Streaming method	1.1 sec
Standalone method	0.82 sec

### 5.2 電力消費量の評価

提案システムは待機状態であっても、常に音声入力の監

視や、3D キャラクタの描画を更新する必要がある電力消費量が大きい。そこで、電力消費量に関する評価を行った。

利用した端末は Sony Xperia TX である。画面の輝度は 50% に固定し、常に画面を表示させ、Wi-Fi や 3G などの無線を切断した状態で放置した。提案システムを満充電から音声入力無しで 1 時間放置したときの電池残量を計測する。静止画の背景を表示させた状態と、提案システムをライブ壁紙モードで表示した状態を比較した。ただし、一般にスマートフォンに表示される残容量は目安であり誤差が大きいため正確な消費電力量を計測することはできない。正確な計測手法は今後の課題とする。

実験結果を表 4 に示す。静止画背景の残量が 93% であるのに対して、提案システムを利用した場合の残量は 58% であった。単純計算で、提案システムは 143 分の連続利用ができる計算になる。一見すると利用できる時間が短く感じるかもしれないが、一般的にユーザはホーム画面をみる時間よりも、Web ブラウザやメールを閲覧している時間やスリープ状態の方が長い。提案システムはホーム画面を見ている時間以外は電力を消費しないので、実用には問題ないと考えている。しかしながら、今後の課題として、電力消費量を抑制する仕組みを検討していかなければならない。

表 4 電力消費量の評価  
Table 4 Evaluation of Power Consumption

	Battery level (60 min)	Estimated time
Proposed System	58%	143 min
Wallpaper (Image)	93%	857 min

### 5.3 ユーザビリティの評価

被験者（大学生）16 人に対して、ユーザビリティの評価を実施した。被験者は、Siri、モバイルメイちゃん、提案システムの順にそれぞれ対話を実施し、以下の項目に対してアンケートによる 5 段階評価と自由コメントを記述した。

1. 音声対話の応答時間が早かったか？ (Response)
2. 合成音声の品質は良かったか？ (Speech Synthesis)
3. 音声認識の精度は良かったか？ (Speech Recognition)
4. エージェントに実在感を感じたか (Reality)
5. エージェントに魅力を感じたか (Charm)
6. 音声対話をして楽しかったか？ (Interesting)
7. 音声対話を自然に感じたか (Natural)
8. 映像の品質は良かったか？ (Video quality)
9. エージェントを表示させる必要性を感じたか？ (Character)

5 段階評価の平均結果を図 5 に示す。Response の項目において、提案システム(平均 4.7)は Siri(平均 2.7)だけでなくモバイルメイちゃん(平均 3.8)よりも良い結果になった。モバイルメイちゃんと提案システムの応答時間の差は、表 3

に示すように、わずか 0.3 秒であるにもかかわらず有意な差となった。応答時間の短さと映像品質の良さが、提案システムの Reality の項目(平均 4.7)や Charm の項目(平均 4.6)の良好な結果に繋がったと考えている。また、提案システムの Character の項目が平均 4.6 と非常に高く、エージェントを画面に表示させる必要性は十分にあるといえる。その一方で Natural の項目は平均 3.6-3.8 と手法の違いによる差は少ない。

提案システムに対する自由コメントとしては、エージェントが可愛い、映像がきれい、エージェントの存在感が高い、より人間味を感じた、音声が自然で分かりやすい、というポジティブな意見があった。その一方で、Siri と比べて応答パターン数が少ない、エージェントが表示されると周りの目が気になり恥ずかしいというネガティブな意見もあった。

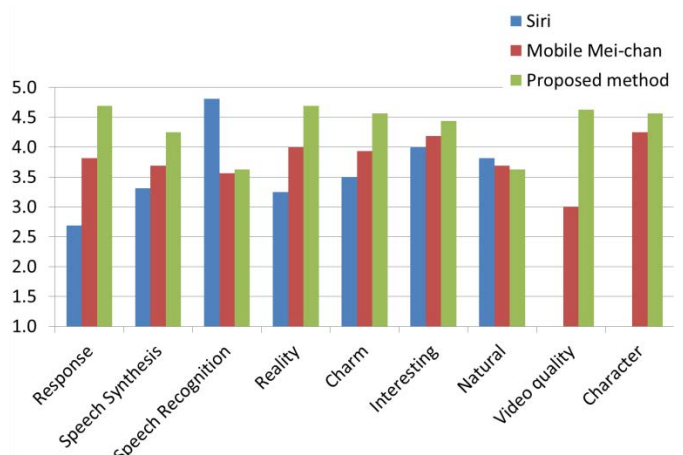


図 5 ユーザビリティに関するアンケート結果  
Figure 5 Result of questionnaire of usability for each system

## 6. 関連研究

MMDAgent を利用した音声対話システムとして、全天候型双方向音声案内デジタルサイネージ[15]とモバイルメイちゃんが挙げられる。

全天候型双方向音声案内デジタルサイネージは、名古屋工業大学の正門前に設置されたデジタルサイネージ[16]上に実装され、等身大 3D キャラクタ「メイちゃん」が表示されている。学生や教職員が Web ブラウザを用いてイベントパネルを投稿することが可能[17]であり、それらに基づいて動的にスクリプトを生成することが可能である。

モバイルメイちゃんは、スマートフォン等のビデオ通話機能を用いて、3D エージェントとの音声対話を実現したシステムである。サーバ上に、Skype API と MMDAgent を連携したシステムが構築されており、ユーザがビデオ通話を開始することによって、音声対話を実現している。

3D キャラクタを表示した初期の音声対話システムとしては、Talkman [18]がある。Talkman は、人間と自然なイン

タラクションを行うエージェントを目指して構築されており、人間的な顔と表情を持ち、音声認識技術による音声対話を実現している。キャラクタを用いた音声対話システムとしては他に、たけまるくんやキタちゃん[19]もある。

音声対話システムの実用化例としては、Let's GO バスシステム[20]、京都市バス運行情報案内システム[21]などがあり、携帯電話などから音声通話でバス運行情報を問い合わせることが可能である。

擬人化エージェントツールキットとしては、本研究で利用したMMDAgentの他に、Galatea ツールキット[22]が存在する。

## 7. おわりに

本稿では、スマートフォン単体で動作可能な音声対話3Dエージェントシステムを開発した。従来のスマートフォン向け音声対話システムとは違って、遅延の少ないより自然な音声対話が可能になった。また、ライブ壁紙として常時起動している形態での音声対話を可能にした。Android OSと連携することによってOS内部のカレンダー情報の提示やアプリの起動を可能にした。また、常時起動しているときに問題となる誤動作に対処するために、音声対話休止機能を実装した。

提案システムの応答時間や、ユーザビリティ、電力消費量などに関する評価を行い、提案手法の有意性を示した。特に、提案システムは、音声対話の応答時間の短さや映像品質の良さが、エージェントの実在感や魅力などを向上させることができた。

なお、提案システムのデモビデオ<sup>1</sup>も公開しているので参考にしてほしい。

今後の課題としては、まるでアプリを追加するかのように、ユーザが自由に音声対話の機能を追加できる仕組みを検討する。さらに、スマートフォンに適した音声インタフェースの考察や、高度な音声ナビゲーション手法についての検討、電力消費量を減少させる仕組みなどを考えている。また、より詳細な評価の実施も検討している。

**謝辞** 本研究は、独立行政法人科学技術振興機構 CREST「共生社会に向けた人間調和型情報技術の構築」による支援を受けた。

## 参考文献

- 1) Apple Inc., Siri, <http://www.apple.com/ios/siri/> (accessed October 26, 2012)
- 2) NTT ドコモ, シャベってコンシェル, [http://www.nttdocomo.co.jp/service/information/shabette\\_concier/](http://www.nttdocomo.co.jp/service/information/shabette_concier/) (accessed October 26, 2012)
- 3) 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, 情報処理学会研究報告, Vol. 2011-SL-89, pp. 1-6 (2011)

- 4) N. Ward, A. Rivera, K. Ward, and D. Novick. Root causes of lost time and user stress in a simple dialog system. In *Interspeech 2005*, Lisbon, Portugal. (2005)
- 5) 志和 敏之, 神田 崇行, 今井 倫太, 石黒 浩, 萩田 紀博, 安西 祐一郎, 対話ロボットの反応時間と反応遅延時における間投詞の効果, 日本ロボット学会誌, Vol. 27, No. 1, pp.87-95 (2009)
- 6) 伊藤 敏彦, 峯松 信明, 中川, 聖一, 間投詞の動きの分析とシステム応答生成における間投詞の利用と評価, 日本音響学会誌, Vol. 55, No. 5, pp. 333-342, (1999)
- 7) 打矢 隆弘, 山本 大介, 柴川 元宏, 吉田 真基, 西村 良太, 内匠 逸, ビデオ通話型音声対話サービス「モバイルメイちゃん」の開発, JAWS2012 講演論文集 (2012)
- 8) Lee, A. and Kawahara, T.: Recent Development of Open-Source Speech Recognition Engine Julius, APSIPA, pp. 131-137 (2009)
- 9) 大浦 圭一郎, 酒向 慎司, 徳田 恵一: 日本語テキスト音声合成システム Open JTalk, 日本音響学会春季講演集, Vol. 1, No.2-7-6, pp. 343-344 (2010)
- 10) MikuMikuDance, <http://www.geocities.jp/higuchuu4/index.htm>, (accessed October 26, 2012)
- 11) Bullet Physics, <http://bulletphysics.org>, (accessed October 26, 2012)
- 12) Google, Android NDK, <http://developer.android.com/tools/sdk/ndk/index.html> (accessed October 26, 2012)
- 13) Open SL ES, <http://www.khronos.org/opensles/> (accessed October 26, 2012)
- 14) Lee, A., Kawahara, T., Takeda, K., Mimura, M., Yamada, A., Ito, A., Itou, K., and no, K. S.: Continuous Speech Recognition Consortium – an Open Repository for CSR Tools and Models –, pp. 1438-1441 (2002)
- 15) 大浦 圭一郎, 山本 大介, 内匠 逸, 李 章伸, 徳田 恵一, キャンパスの公共空間におけるユーザ参加型双方向音声案内デジタルサイネージシステム, 特集「音声対話」, 人工知能学会誌, Vol.28, No.1 (2013) (in press)
- 16) 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)
- 17) 山本 大介, 大浦 圭一郎, 李 晃伸, 打矢 隆弘, 内匠 逸, 徳田 恵一, 松尾啓志, 双方向音声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, 大学ITC 推進協議会 2011 年度年次大会, CD-ROM, 2011 年 12 月.
- 18) Nagao, K., Takeuchi, A.: Speech dialogue with facial displays: Multimodal human-computer conversation, Proc. ACL-94, pp. 102-109 (1994)
- 19) 鹿野清宏, Tobias, C., 川波弘道, 西村竜一, 李 晃伸: 音声情報案内システム「たけまるくん」および「キタちゃん」の開発, 情報処理学会研究報告, 2006-SLP-63-7 (2007).
- 20) Raux, A., Bohus, D., Langner, B., Black, A.W. and Eskenazi, M.: Doing Research on a Deployed Spoken Dialogue System: One Year of Let's Go! Experience, Proc. Int'l Conf. Spoken Language Processing (INTERSPEECH) (2006)
- 21) 駒谷 和範, 上野 晋一, 河原 達也, 奥乃 博, ユーザモデルを導入したバス運行情報案内システムの実験的評価, 情報処理学会研究報告, Vol. 2003-SLP-75, pp. 59-64 (2003)
- 22) 川本 真一, 下平 博, 新田 恒雄, 西本 卓也, 中村 哲, 伊藤 克亘, 森島 繁生, 四倉 達夫, 甲斐 充彦, 李 晃伸, 山下 洋一, 小林 隆夫, 徳田 恵一, 広瀬 啓吉, 峯松 信明, 山田 篤, 伝康 晴, 宇津呂 武仁, 嵯峨山 茂樹, カスタマイズ性を考慮した擬人化音声対話エージェントツールキットの設計, 情報処理学会論文誌, Vol. 43, No. 7, pp.2249-2263 (2002)

<sup>1</sup> 【MMDAgent】スマートフォンで3D秘書エージェントを作ってみた, <http://www.youtube.com/watch?v=eR7aUj9RBio>