

# ユーザによる案内情報の付与とそれに基づく モバイル音声道案内システム

柳 倫浩<sup>1</sup> 山本 大介<sup>1</sup> 高橋 直久<sup>1</sup>

**概要:** 従来、我々は、管理者が入力した地図情報に基づいて案内文を自動生成し、それに基づくモバイル音声道案内システムを提案してきた。しかしながら、音声のみでは必ずしも分かりやすい案内を実現することは出来ず、また、全ての経路について管理者が詳細な案内情報を入力することは人的コストが大きくなるという問題点があった。そこで、本稿では、ユーザによる案内情報が付与可能なモバイル音声道案内システムを提案する。具体的には、案内場所に詳しい任意の複数のユーザに、スマートフォンを用いて写真やランドマークなどを含む案内情報を付与してもらう仕組みを実現することで、より多くの詳細な案内情報を収集する。また、ランドマークは眺める位置により見え方が異なるため、ノード・ランドマークモデルを採用することにより、任意の視点（交差点・曲がり角）から適切な案内を実現する。さらに、収集された案内情報に基づいて、音声だけでなく画像等も用いたより分かりやすいモバイル音声道案内を実現した。提案システムを試作し、評価実験を行った。

## 1. はじめに

近年、Google Maps[1] や NAVITIME[2] など、モバイル音声道案内システムが普及している。我々は、管理者が地図上に案内情報を入力することで、案内地図を作成し、作成した案内地図に基づいて、自動的に音声案内文を生成することで、音声による道案内を行うシステム [3] を提案してきた。本システムでは、スマートフォン上の音声対話システム構築ツールキットとして、MMDAgent[4], [5] を採用した。音声のみによる道案内は、端末の画面を見ることなしに、移動しながら道案内を受けることが可能である利点がある一方で、以下に示す問題があった。

**問題点 1** 管理者が、任意の経路に対応した詳細な案内情報を入力することは、人的コストが大きくなる問題がある。

**問題点 2** 音声のみによる道案内では、複雑な経路を正確に伝えることが難しい。より分かりやすい道案内を実現するためには、音声だけでなく、必要に応じてランドマークの写真などを併用した道案内を実現することが望ましい。

**問題点 3** ランドマークは眺める場所や方向によって見え方が異なるため、ランドマークを用いた道案内をする場合、ユーザの位置（交差点・曲がり角）

に応じて写真や説明文を変えるなど、適切な案内を実現する必要がある。

そこで、本研究では管理者以外の任意のユーザによる案内情報の付与が可能なモバイル音声道案内システムを実現する。また、そのための必要となる要件を以下に示す。

**要件 1** 任意のユーザが、スマートフォンを用いて、地図上に写真やランドマークを含む案内情報の付与ができる。

**要件 2** 音声だけでなく、写真や地図を併用した道案内が、スマートフォン上で行える。

**要件 3** 同じランドマークに対して、交差点や曲がり角毎に、異なる写真や説明の付与が行える。

## 2. 音声対話構築ツールキット MMDAgent

MMDAgent は、名古屋工業大学国際音声技術研究所で開発された音声対話システム構築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のモジュールを統合したシステムである。音声認識エンジンとして Julius[6] を、音声合成エンジンとして OpenJTalk[7] を、3D モデルとして MikuMikuDance 形式 [8] を、物理演算エンジンとして Bullet Physics [9] を採用している。また図 1 に示す、スマートフォン単体でも動作可能な Android 版 MMDAgent[10] の開発も行われている。

<sup>1</sup> 名古屋工業大学大学院工学研究科



図 1 Android 版 MMDAgent

MMDAgent は FST ファイルと呼ばれるスクリプトファイルで設定された状態遷移 (Finite State Transition) に基づいて動作する。FST ファイルの例を図 2 に示す。FST ファイルは、状態番号、遷移先の状態番号、遷移条件 (イベント)、コマンドの 4 つの組から成る。図 2 の例では、状態番号が 1 であるとき、「こんにちは」という単語を認識し、状態番号 10 へ遷移することで、「こんにちは」としゃべる動作を行い、状態番号 11 へ遷移する。FST ファイルは任意に書き換えることが可能であるため、音声案内における対話をスクリプト化させることにより、MMDAgent を用いた音声案内を実現する。

FST ファイルは 1 つのメイン FST ファイルに対し、複数のサブ FST ファイルを設定できる。サブ FST ファイルを設定することで、メイン FST ファイルとは別の状態遷移を設定できる。サブ FST の状態遷移はメイン FST の状態遷移とは独立した状態遷移となっており、これにより複数の状態遷移に従って MMDAgent を動作させることができる。

また、Android 版 MMDAgent の特徴としてレスポンスの速さがある。スマートフォン単独での、音声合成・音声認識に基づく対話をリアルタイムに実現しているため、ネットワークに起因する遅延が無いために、スムーズな案内を実現することができる。

```

1 10 RECOG_EVENT_STOP|こんにちは <esp>
10 11 <eps> SYNTH_START|mei|こんにちは.
11 1 SYNTH_EVENT_STOP|mei <eps>

```

図 2 FST ファイルの例

### 3. 提案システムの概要

#### 3.1 提案システムの特徴

本稿では、以下の特徴をもつユーザによる案内情報の付与が可能なモバイル音声案内システムを提案する。

特徴 1 案内情報を収集するサーバを作成し、任意のユーザがスマートフォンを用いて写真やランドマーク

を自由に地図上に付与できるインターフェースを実現することで、案内情報をより多く収集する。

特徴 2 プラグインを作成し MMDAgent の機能を拡張することで、MMDAgent 上で写真を利用した音声案内を実現する。また FST テンプレートを用いて、案内 FST の生成を行うことで、より汎用性のある案内 FST の生成が可能となる。

特徴 3 ノード・ランドマークモデルを導入することにより、同じランドマークに対して、交差点毎に異なる案内情報の付与を可能とする。

#### 3.2 提案システムの構成

提案システムは大きく分け、ユーザ生成型案内情報収集システム、案内 FST 生成システム、及び、案内データベースから構成される。案内データベースはノード、ランドマーク座標、名前、アノテーションデータなどの案内に必要なデータを格納している。ユーザ生成型案内情報収集システムと案内 FST 生成システムの構成を以下に示す。

##### 3.2.1 ユーザ生成型案内情報収集システム

ユーザ生成型案内情報収集システムは図 3 に示すように、スマートフォンアプリケーションと案内情報収集サーバから構成される。スマートフォンアプリケーションは案内情報が付与された案内地図をユーザに提示し、ユーザは提示された地図上に写真やランドマークなどの案内情報を付与する。案内情報収集サーバはスマートフォンアプリケーションに対し、案内データベース内の案内情報を送受信する役割をもつ。ユーザ生成型案内情報収集システムでは、スマートフォンアプリケーションと案内情報収集サーバがインターネットを介して案内情報の送受信を行うことで、管理者以外の任意のユーザによる案内情報の付与を可能とする。

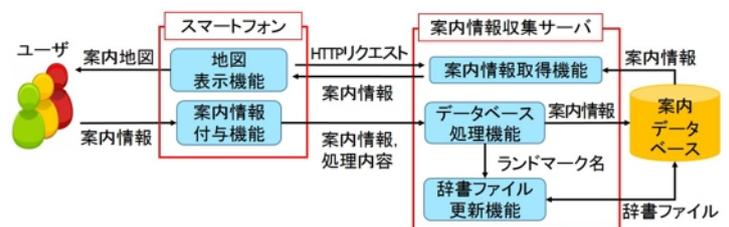


図 3 案内情報収集システムの構成図

##### 3.2.2 案内 FST 生成システム

案内 FST 生成システムは、図 4 に示すように、経路探索機能、案内 FST 生成機能から構成される。最初に出発地点と目的地点を入力として経路探索機能において 2 点間の案内経路を探索する。そして案内 FST 生成機能において、案内経路上の案内情報、FST テンプレート、案内生成テンプレートを用いることで 2 点間の経路案内を行う案内

FST を出力する。

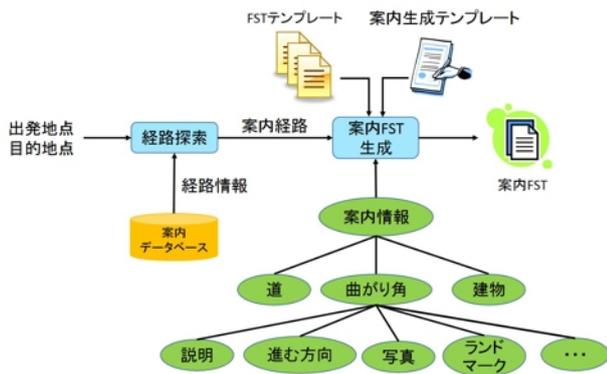


図4 案内FST生成システムの構成図

## 4. 提案システムの実現法

### 4.1 ユーザによる案内情報の付与

提案システムでは、スマートフォン上で動作するアプリケーションが情報収集サーバとインターネット介して通信することで任意ユーザによる案内情報の付与を実現する。通信方法はREST方式を採用し、アプリケーションがサーバへHTTPリクエストを送信することで、サーバに対してデータベースの更新や情報の取得などの様々な処理を要求する。またアプリケーションとサーバ間の案内情報の受渡しには、JSONフォーマット[12]を利用している。

#### 4.1.1 案内情報収集サーバ

案内情報収集サーバが有する機能を以下に示す。

- データベース処理機能  
スマートフォンアプリケーションから受け取った案内情報の処理を行う。HTTPリクエストのパラメータの値によって指定された追加、更新、削除のいずれかの処理をデータベースに対して実行する。
- 案内情報送信機能  
HTTPリクエストにより指定された案内情報をデータベースより取得し、スマートフォンへアプリケーションへと送信する。
- 単語認識辞書ファイル更新機能

MMDAgentでは、辞書ファイルに単語を登録することで、新たな単語の認識が可能となる。MMDAgentの辞書ファイルを更新することで、ユーザが新たに追加したランドマークをMMDAgent上で認識できるようにする。情報収集サーバは、ユーザがデータベースに対して新たなランドマークを追加した際に、登録単語リストを参照し登録されていない単語であった場合、次に示す手順に沿って辞書ファイルを更新する。

手順1 KAKASI/Javaライブラリ[13]を用いて、ランドマーク名の読みをカタカナと音素列へと変換する。

手順2 図5に示す形式に基づいて、辞書ファイルへ追加出

力する。

手順3 登録単語リストへ今回登録した単語を追加する。

東京:トーキョー:東京:515 @1.0 鶴舞:ツルマイ:鶴舞:515 [鶴舞] tsurumai

図5 辞書ファイル登録の例

### 4.1.2 スマートフォンアプリケーション

- 地図表示機能  
起動時に情報収集サーバよりデータベース内の案内情報を取得し、地図上に案内情報を付与することでユーザに構内地図を提示する。
- 案内情報付与機能  
ユーザが付与した案内情報を情報収集サーバへと送信することで、案内データベースへ新たな案内情報を追加する。

## 4.2 案内FSTの生成

### 4.2.1 FSTテンプレートを利用した案内FSTの生成

写真の表示や道の説明、曲がる方向の指示などの案内の動作をそれぞれ一つの状態として考え、案内生成の流れを図6,7に示すような、複合状態を含む状態遷移[14]として定義する。そして、それぞれの動作の状態に対するFSTテンプレートを用意することで、より柔軟な案内FSTの生成が可能となる。案内生成における状態の遷移は主に案内情報の有無により行われる。

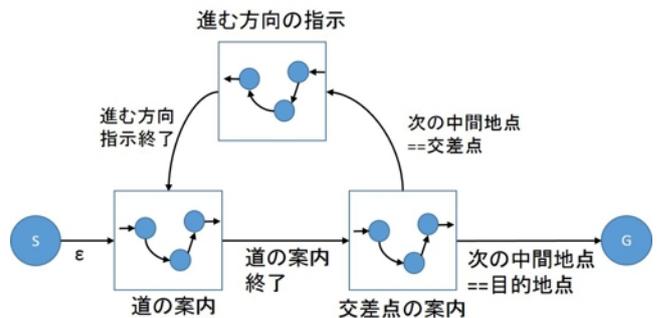


図6 案内FST生成の状態遷移モデルの例

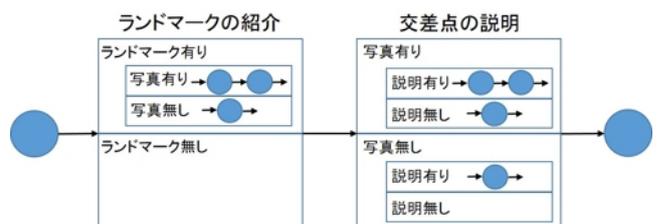


図7 交差点の案内の状態遷移モデルの例

01 <eps> SYNTH\_START|mei|voice|<direction>に<landmark\_name>が見えます。



01 <eps> SYNTH\_START|mei|voice|右側に大きな看板が見えます。

図 8 ランドマークを紹介する FST テンプレートの例

#### 4.2.2 利用する写真の選択

同じ曲がり角を撮影した写真でも、撮影した場所からの曲がり角に対しての方角によって全く異なる場所に見えてしまうという問題がある。そこで FST 生成の際に、案内における移動方向と写真の撮影方向を比較することで、複数の候補の中から適切な写真の選択を行う。また複数の案内に利用可能な写真がある場合には、一番新しい写真を選択する。またランドマークについては、ノードとランドマークの関係に基づいて写真の選択を行う。

#### 4.3 プラグインによる写真表示機能の追加

新たに写真画像の表示 (図 9) を行うプラグインを作成し以下に示すコマンドとイベントを追加することで、MMDAgent 上での写真を利用した案内を可能とした。

- PHOTO\_ADD|ailias|file\_path|apearing\_time  
表示するに写真画像に対して ailias で指定した変数名を与え、file\_path で指定された写真画像を apearing\_time だけ MMDAgent 上へ表示する
- PHOTO\_DELETE|ailias  
表示中の変数名が ailias である写真画像を MMDAgent 上から消去する。
- PHOTO\_EVENT\_ADD|ailias  
表示中の変数名が ailias である写真画像が MMDAgent 上から消えた時に、発生するイベントメッセージ。
- PHOTO\_EVENT\_DELETE|ailias  
表示中の変数名が ailias である写真画像が MMDAgent 上に表示された時に、発生するイベントメッセージ。



図 9 MMDAgent 上での写真画像の表示

#### 4.4 案内データベースの構造

提案システムで用いる案内データベースについて説明

する。本研究ではデータベース管理システムとして PostgreSQL[11] を利用している。データベース内のノード・ランドマーク・写真についてのデータ構造を表 1 から表 5 に示す。ノード・ランドマーク・写真についての基本情報は、表 1,3, 4 に示す形式で案内データベースに格納されている。またユーザが付与する案内情報は、表 2 に示す形式で管理されている。node\_id はアノテーションが付与されているノードを一意に識別するために用いられる。主キーとしてアノテーションを管理する id を用いることで、一つのノードに対して複数の説明や写真を付与することが可能となる。ノードとランドマークの関係は、表 5 に示すテーブルにより管理される。landmark\_id が参照されるランドマークを示し、node\_id がそのランドマークを参照するノードを示す。このような互いの二つの id の組み合わせにより、ノードとランドマークの関係を決定する、ノード・ランドマークモデルを導入することで、同じランドマークに対してノード毎に異なる案内情報を付与することが可能となる。

表 1 ノードテーブル

フィールド名	型	説明
id	int	ノードの識別 ID
name	text	ノードの名前
latitude	double	ノードの緯度
longitude	double	ノードの経度

表 2 ノードアノテーションテーブル

フィールド名	型	説明
id	int	ノードに付与されるアノテーションの管理 ID
node_id	int	ノードの識別 ID
explanation	text	ノードについての説明
photo_id	int	ノードの写真

表 3 ランドマークテーブル

フィールド名	型	説明
id	int	ランドマークの識別 ID
name	text	ランドマークの名前
latitude	double	ランドマークの緯度
longitude	double	ランドマークの経度

表 4 写真テーブル

フィールド名	型	説明
id	int	写真の識別 ID
file_name	text	画像ファイルの名前
file_path	text	サーバ内の画像ファイルへのパス
takenTime	timestamp	写真が撮影された時間
orientation	text	被写体に対する方位

フィールド名	型	説明
id	int	ノードとランドマークの関係アノテーションを管理する ID
landmark_id	int	ランドマークの識別 ID
node_id	int	ランドマークを参照するノードの識別 ID
explanation	text	ランドマークが特定のノードから参照されるとき説明
photo_id	text	ランドマークが特定のノードから参照されるとき写真

## 5. プロトタイプシステム

提案システムのプロトタイプを Java 言語 [15] を用いて実装した。プロトタイプシステムは地図画面とそれぞれの情報入力画面からなる。各インターフェースの説明と動作を以下に示す。

システム起動時に情報収集サーバーより案内情報を取得し、案内情報が付与された構内地図 (図 10) をスマートフォン上へ表示する。緑色のマーカーが曲がり角、赤色のマーカーが建物、赤色のピンのマーカーがランドマーク、水色の線が道を表している。左下のボタンよりモードを切り替えることで案内情報の追加と入力を行う。



図 10 地図画面

### 5.1 案内情報の追加

追加モードを選択することで案内情報の追加を行うことができようになる。画面上の任意の場所をタップすることで追加する案内情報の種類選択画面が表示される。この中から追加を行いたい案内情報の種類を選択することで、案内情報を追加することができる。ランドマークや建物を追加する場合は名前を入力する必要がある、名前が入力されていない場合は追加操作が完了しない。また道を追加する場合は、2つのマーカーをそれぞれタップすることで道を追加することができる。

### 5.2 案内情報の入力

入力モードを選択することで地図上のノードやランド

マークについての情報を入力することができるようになる。地図上の情報入力を行いたい、マーカーをタップすることで、操作選択画面 (図 11) が表示される。各操作について以下に示す。

- 情報入力 (マーカーの種類)  
情報入力画面へ遷移し選択したマーカーについての情報の入力を行うことが可能となる。
- 情報入力 (エッジ)  
地図上の線を形成する両端のマーカーを選択することで、道についての情報入力画面へ遷移する。
- 写真一覧  
マーカーに登録されている写真の一覧を見ることが出来る。
- 写真撮影  
写真を撮影して、マーカーへ新たに写真を追加することができる。

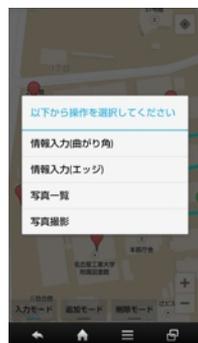


図 11 操作選択画面

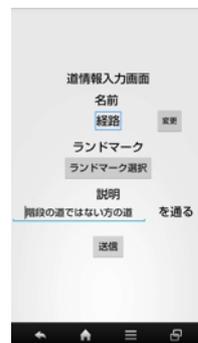


図 12 情報入力画面

### 5.3 写真, ランドマークの選択

曲がり角と道はそれぞれの情報入力画面上のランドマーク選択ボタンを選ぶことで、ランドマーク選択画面 (図 13) へと遷移し一覧から目印となるランドマークを選択することができる。写真選択ボタンを選ぶことで、写真選択画面 (図 14) よりそのランドマークに登録されている写真一覧から案内に利用する写真を曲がり角や交差点毎に選択することができる。また、詳細ボタンを押すことでランドマークの情報入力画面へと遷移する。

## 6. 評価実験

### 6.1 実験手法

被験者 (名古屋工業大学の学生 12 名) に対して、4 人毎 3 つのグループ A,B,C に分け、指定した 2 種類の経路に対して、以下のタスクを与えた。ただし、実験場所は大学近くの公園であり、グループ B,C における道案内の際には目的地点を伝えず、案内に沿って目的地点まで移動する。

- グループ A (案内情報の付与)  
実際に指定した 2 種類の経路を歩いてもらい、スマー



図 13 ランドマーク選択画面



図 14 写真選択画面



図 16 経路 2 の案内に利用された情報一覧

トフォン端末を用いてその場で案内情報の付与を行う。

- グループ B(従来手法による案内)  
管理者が事前に入力した情報に基づいて、生成される案内に従って経路を移動する。
  - グループ C(提案手法による案内)  
グループ A が付与した情報も利用して、生成される案内に従って経路を移動する。
- その後グループ B,C の被験者は、2 種類の経路それぞれに対して以下のアンケート項目に 7 段階評価で回答した。

- アンケート項目
  - 案内は分かりやすいか
  - 案内の内容は適切か
  - 一つの案内文の長さは適切か

す。経路 1,2 共に分かりやすさが向上していることが分かる。しかしながら、経路 2 に関してはあまり大きな変化が見られていない。そこで、経路 2 に関してグループ B, グループ C の被験者に分かりにくかった案内について調査を行った結果以下のような意見が挙げられた。

- グループ B
  - 指定されたランドマークが同じようなものが多くありどれか特定しにくい
- グループ C
  - 一部案内をややこしくしているランドマークがあった

## 6.2 実験結果と考察

図 15,16 に実際に付与された情報を示す。得られた結果の特徴として、付与される案内情報はテキストのみの説明文よりも実際の写真の画像が付与される傾向が見られた。また、すでに案内情報が付与されている場所に対して他の被験者が新たに別の案内情報を付与することがほとんど無かった。

グループ B における問題点は、グループ A において情報の付与が行われたことにより改善された。しかし、グループ A で新たに追加された他のランドマークが案内を分かりにくくしてしまったことが分かった。このことから、どのように入力された情報の取捨選択をして、案内を生成するかが今後の課題となる。



図 15 経路 1 の案内に利用された情報一覧

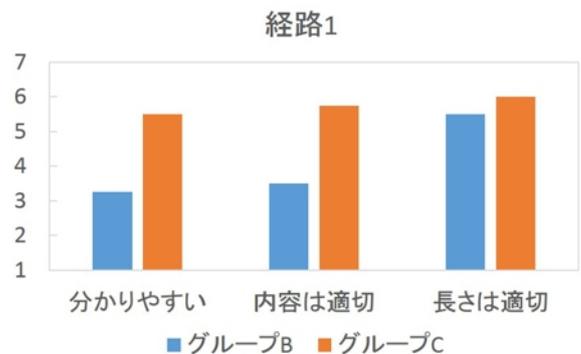


図 17 経路 1 のグループ別のアンケート結果比較

経路 1,2 の案内に関するアンケート結果を図 17, 18 に示

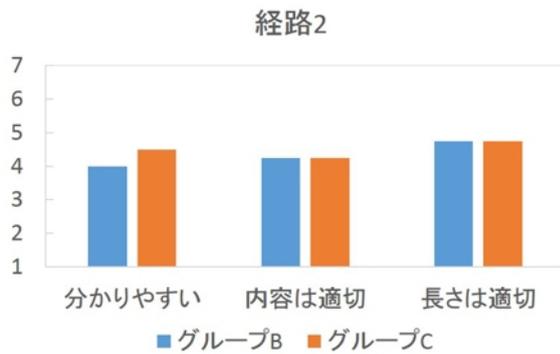


図 18 経路 2 のグループ別のアンケート結果比較

## 7. 関連研究

歩行者向けの道案内を行う手法として、経路より案内文を生成し案内を行う手法 [16], [17] が提案されている。これらは、次に向かうべき目的地点の説明として目的地点周辺のランドマークが利用されている。しかし、与えられる情報はランドマークの名称だけであり外観の情報を利用していないために、ランドマークの特定が困難な場合が考えられる。そこで、文献 [18] では外観情報も利用することが有効であるとし、Web や Blog の文書群の Web マイニングに着目し、案内のための外観情報を収集する手法を提案している。提案システムでは、ユーザ生成型に着目し、案内情報を収集する。また文献 [19] において、風景画像と携帯電話を用いた歩行者ナビゲーションシステムを提案している。目的地点周辺の関係者が最寄駅から目的地点までといった特定の経路の案内を案内文と風景画像を用いて作成することで、歩行者のユーザの立場に立った適切な情報の提供を可能としている。またプロトタイプシステムを用いた実験において、風景画像を用いた案内の有用性を示している。提案システムでは、特定の経路に対してでなく場所毎に地図へのアノテーションという形で情報の付与を行う。

## 8. おわりに

本稿では、ユーザによる案内情報の付与に基づくモバイル音声道案内システムについて提案した。ノード・ランドマークモデルを導入することで、同じランドマークに対してノード毎に異なる案内情報の付与を可能とした。またプロトタイプとして任意ユーザがスマートフォン端末上から案内情報の付与が可能な、ユーザ生成型案内情報収集システム、及び、MMDAgent 上での写真表示を可能とする機能を実装し、システムの検証実験を行った。

今後の課題として、案内生成時に複数の案内情報が付与されているときに、適切な案内情報の選択をどのように行うかを考える必要がある。

## 謝辞

本研究は JSPS 科研費 26330136, 25700009, 及び、科学技術振興機構 CREST のにより助成を受けたものです。

この場を借りて、感謝いたします。

## 参考文献

- [1] Google Maps, 入手先 (<https://maps.google.co.jp/>) (2014.5.1)
- [2] NAVITIME, 入手先 (<http://www.navitime.co.jp/>) (2014.5.1)
- [3] 尾前 貴則, 山本 大介, 高橋 直久, 音声対話を用いた構内道案内システムの実現, DEIM Forum 2013 D1-4.
- [4] 李晃伸, 大浦圭一郎, 徳田恵一, 魅力ある音声インタラクティブ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, 電子情報通信学会技術研究報告 SP, 音声, 2011.
- [5] Akinobu Lee, Keiichiro Oura, Keiichi Tokuda, MMDAgent - A fully open-source toolkit for voice interaction systems, Proceedings of the ICASSP 2013, pp. 8382-8385, 2013.5.
- [6] Lee Akinobu, and Tatsuya Kawahara, Recent Development of Open-Source Speech Recognition Engine Julius, Proceedings of the APSIPA ASC, pp. 131137, 2009.
- [7] 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論集, Vol. 1, No.2-7-6, pp. 343344, 2010.
- [8] MikuMikuDance, 入手先 (<http://www.geocities.jp/higuchuu4/>) (2014.5.1)
- [9] BulletPhysics, 入手先 (<http://bulletphysics.org/>) (accessed February 3, 2014)
- [10] 山本 大介, 大浦 圭一郎, 西村 良太, 打矢 隆弘, 内匠 逸, 李 晃伸, 徳田 恵一, 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, インタラクティブ 2013, IPSJ Symposium Series Vol.2013, No. 1, pp. 675- 680, 東京, 2013.
- [11] PostgreSQL, 入手先 (<http://www.postgresql.org/>)
- [12] JSON, 入手先 (<http://www.json.org/>) (2014.5.1)
- [13] KAKASI/Java, 入手先 (<http://nicolas-raoul.github.io/kakasi-java/>) (2014.5.1)
- [14] 若林 敬太郎, タブレット端末のための状態遷移図を用いた音声対話コンテンツ編集システムに関する研究, 平成 25 年度 名古屋工業大学 卒業論文, 2014.
- [15] Java, 入手先 (<http://java.sun.com/>) (2014.5.1)
- [16] 長谷川 保, 松田 三恵子, 久保田 浩明, 歩行者向け対話型道案内サービスのための案内テキスト生成, 映像情報メディア学会技術報告, vol.25, no.48, pp.43-48, 2001.
- [17] 藤井 憲作, 杉山 和弘, 歩行者ナビゲーション支援のための場所案内文生成手法, 電子情報通信学会論文誌, vol.J82-D-, no.11, pp.2026-2034, 1999.
- [18] 服部 峻, 手塚 太郎, 田中 克己, サーチエンジンのメタデータを用いた道案内補助のための外観情報の抽出, 電子情報通信学会技術研究報告, vol.106,no.149, pp.131-136, 2006.
- [19] 福井 良太郎, 白川 洋, 歌川 由香, 重野 寛, 重野 , 岡田 謙一, 松下 温, 携帯電話における歩行者ナビゲーション情報の表示方法に関する提案と評価, 情報処理学会論文誌, vol.44, no.12, pp.2968-2978, 2003.