

## 実走行環境下における車内音声データベース

SD-2-1

## In-Car Speech Database

河口 信夫† 松原 茂樹† 岩 博之‡ 梶田 将司†  
武田 一哉† 板倉 文忠† 稲垣 康善\*

Nobuo KAWAGUCHI† Shigeki MATSUBARA† Hiroyuki IWA‡† Shoji KAJITA†  
Kazuya TAKEDA† Fumitada ITAKURA† Yasuyoshi INAGAKI\*

†名古屋大学統合音響情報研究拠点 ‡小島プレス(株) \*名古屋大学工学研究科

†Center for Integrated Acoustic Information Research (CIAIR), Nagoya University

‡Kojima Press Co.Ltd. \*Graduate School of Engeneering, Nagoya University

## 1 はじめに

音声認識技術の発展とともに、実環境において利用可能なさまざまな音声対話システムの実現の要求が高まりつつある。特に車室内における高度情報システムの実現は、近年高い期待が寄せられている [3]。走行中の車室内は、ロードノイズや風切り音といった様々な騒音に加え、加速音やパンプ音といった非定常雑音が重畳される環境である。例えば時速 90km で走行している車内のダッシュボードからドライバーの音声を収録した場合、SN 比は -5dB 以下になる [1]。実走行車内において操作可能な情報システムを実現するためには、高騒音下でのロバスト音声認識技術と、自然発話を受理可能な音声対話技術が必要である。

名古屋大学統合音響情報研究拠点 (CIAIR) では、ロバストな車内音声対話システムの実現を一つの目標と定め、その要素技術の研究開発を進めている。実走行車内環境での音声処理技術の研究開発を支援するために、大規模な車内音声データベースの構築を行ってきた [5]。本データベースは、ドライバーを被験者として、車内における発話を収集しており、大きく 2 種類のデータベースが構築されている。

音響モデル用のデータベースでは、音素バランス文、連続数字、および離散単語の読み上げを行い、高ノイズ環境下での音響モデルの学習・評価用のデータを収集している。車内音声対話データベースでは、レストランやコンビニ等の情報検索が可能な車内情報システムとドライバー間の対話を収録し、書き起しによって対話コーパスを構築している。収集したデータの概要を表 1 に示す。

また、運転状況や画像等の収録を、音声と同期して行うために、車内音声収集用に専用の車両 (図 4) を開発した。以下、収集車およびデータベースについて述べる。

## 2 音響モデル用データベース

ロバスト音声認識、音響モデル学習・評価用の音声として、音素バランス文、連続数字、離散単語の読み上げ音声を収録している。音素バランス文については、停車アイドリング中、および走行中、および運転席と助手席において読み上げ音声を収録している。運転中に音素バランス文の読み上げを行うために、被験者のヘッドセットに適宜音声プロンプトを流す手法を用いている。

音声収録は複数のマイクロホンを用い、口元だけでなく、車内のさまざまな位置の音を取り込んでいる。これは、実環境下におけるマルチチャンネル信号音響処理 [4] の研究・評価にも利用可能なデータを収集するためである。同期音声データの一例を図 1 に示す。この図では、ドライバーとナビゲータの 2 本の接話マイクに加え、6 箇所の分散マイクの音声、制御情報等を表示している。

表 1: 収集した車内音声データ (一人あたり)

| 停車アイドリング中    |            |
|--------------|------------|
| 離散単語の読み上げ    | 30 単語      |
| 連続数字の読み上げ    | 4 桁 × 40 回 |
| 音素バランス文の読み上げ | 50 文       |
| 走行中          |            |
| 音素バランス文の読み上げ | 25 文 (運転席) |
| 音素バランス文の読み上げ | 50 文 (助手席) |
| 情報検索タスク      | 約 10 タスク   |

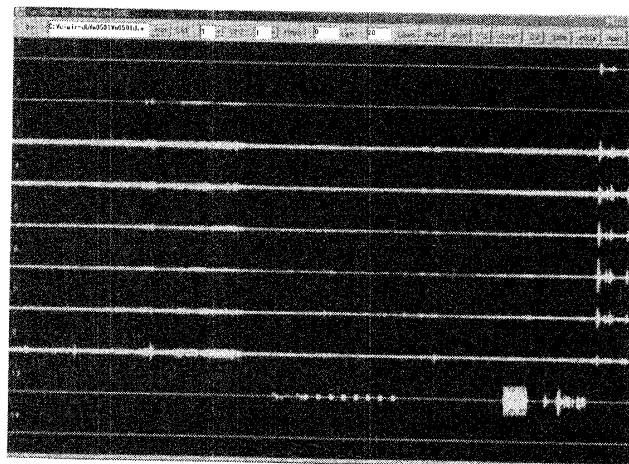


図 1: 同期音声データ

## 3 車内音声対話データベース

自然発話を受理可能な音声対話技術の研究のためにレストランやコンビニ等の情報検索が可能な車内情報システムとドライバー間の対話を収録し、書き起しによって対話コーパスを構築している。

情報検索タスクでは、ドライバーに対するタスクの指示のために「和食」「中華」等の単語や「喉が乾きました。ちょっと一休みしたいですね。」といった状況を表す文章が書かれたプレートを表示している。今回の収録では、車内情報システムを利用している環境を模擬的に実現するために、訓練されたナビゲータが情報システムの役割を果たしている。ナビゲータは後部座席に座り、ドライバーとは直接視線を合わせることはない。また、雑音環境下でも、対話を可能にするために、音声はヘッドセットを用いてやりとりをしている。将来的には、ナビゲータを Wizard of OZ のシステムや実際

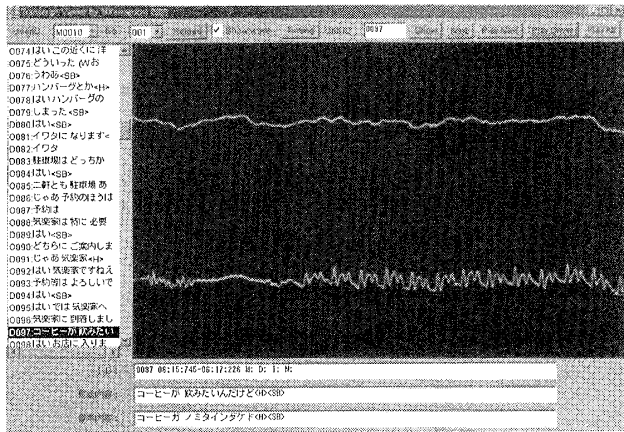


図 2: 書き起しテキストと音声

表 2: 収録機器仕様

| メディア  | 仕様                    |
|-------|-----------------------|
| 音声入力  | 16ch, 16bit, 16kHz    |
| 音声出力  | 16ch, 16bit, 16kHz    |
| 画像入力  | 3ch, MPEG1            |
| 踏力    | アクセル・ブレーキ 16bit, 1kHz |
| D-GPS | ジャイロ, 加速度センサ付き        |

の対話システムと置き換えて収録していくことを計画している。

収集されたデータは、発話の開始時間と終了時間の時間タグをつけ、基本形と発音形が併記された形に書き起しが行われている。図2に書き起しテキスト、および対応する音声を示す。

#### 4 実走行データ収集車

音声、画像、他の情報を収集するために、本収集実験専用のデータ収集車両を作成した。本車両は音声入出力、画像入力、アクセル、ブレーキ、車両位置などの情報をすべて同期しながら90分間の収録が可能である。表2に収録機器の仕様を示す。また、図3にブレーキ、アクセル、速度、エンジン回転数等の車両制御情報を示す。収集車両は6台の収録用PC及び、1台の制御用PCを搭載しており、これらは100Base-Tのネットワークで接続されている。収集車両の概要を図4に示す。本車両のダッシュボードや天井には、マイクを自由な位置に固定するためのネットが取り付けられている。また、ビデオカメラの取り付け位置を変更可能にするために、天井にはレールが取り付けられている。後席には無停電電源装置やマイクアンプ、TVモニター等が設置され、オペレータが収集の管理をするためのPCが設置されている。

本車両は、運転状況の収集について、多目的に利用可能であるため、今後の活用が期待されている。

#### 5 まとめ

CIAIRでは、ロバストな音声対話認識技術確立のために、高騒音下である実走行車室内の音声データベース、および音声対話コーパスの構築を行なっている。収録した音声や対話データは、整理を行って順次公開していく予定である。また、今後も様々な車内音声データベースを収集していく予定である。本データベースのサンプルデータは、ホームページ[7]を通して公開している。



図 3: 車両制御情報

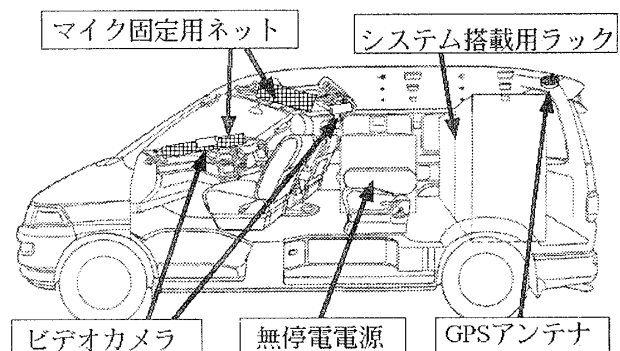


図 4: 収集用車両の概要

御意見、御要望等があればホームページの連絡先にお知らせ願いたい。

#### 謝辞

本研究は文部省科学研究費補助金 COE 形成基礎研究費 (課題番号 11CE2005) の補助を受けて行われた。

#### 参考文献

- [1] 中村 哲, 鹿野清宏: 認識技術の進展, 情報処理, Vol.38, No.11, pp.998-1006(1995).
- [2] 松原茂樹, 河口信夫, 外山勝彦, 稲垣康善: 音声対話インタフェースを備えた車内秘書システムの提案, 人工知能学会全国大会 (第13回) 論文集, 16-02, pp.231-234(1999).
- [3] 松原茂樹, 河口信夫, 外山勝彦, 稲垣康善: 発話の同時理解・同時生成に基づく車内音声対話秘書システムの提案, 人工知能学会研究会, 言語・音声理解と対話処理, SIG-SLUD-9902, pp.1-6(1999).
- [4] 清水泰博, 梶田将司, 武田一哉, 板倉文忠: 空間音響特性依存HMMによるスペースダイバーシチ型音声認識手法の評価, 音講論集, 1-8-9, pp.17-18(2000).
- [5] 河口信夫, 松原茂樹, 岩 博之, 梶田将司, 武田一哉, 板倉文忠: 実走行車内における音声データベースの構築, 情報処理学会研究会, 音声言語情報処理, 99-SLP-30-12, pp.57-62 (2000).
- [6] 岩 博之, 河口信夫, 武田一哉, 板倉文忠: 実走行環境下における車内音声対話・音響データ収録装置, 音講論, 1-Q-29, pp.190-191(2000).
- [7] CIAIR: <http://www.ciair.coe.nagoya-u.ac.jp/>