

浴室向け音声コントローラの開発

馬場朗[†], 日比谷新平[†], 奈木野豪秀[‡], 小笹詩織[‡], 庄境誠[‡], 竹原清隆[†]

[†] パナソニック 電工株式会社 新規商品創出技術開発部

[‡] 旭化成株式会社 音声ソリューションビジネス推進部

E-mail: {baba.a, hibiya, takehara_kiyotaka}@panasonic-denko.co.jp

E-mail: {nagino.gb,kozasa.sb,shozai.mb}@om.asahi-kasei.co.jp

あらまし 住宅の利便性と快適性の向上を目的に、住宅設備機器の高機能化や機器数の増加が進んでおり、その操作手段も複雑化している。本稿では、住宅設備機器の操作性改善に対する当社の取組み事例として、浴室向け音声コントローラの開発事例を紹介し、試作を通じて得られた知見について報告する。

キーワード 音声認識, 浴室, 音響モデル, 非定常雑音, 残響

Development of Speech Interface System for Bathroom Environment

Akira Baba[†], Shinpei Hibiya[†], Goshu Nagino[‡], Shiori Kozasa[‡],
Makoto Shozakai[‡], Kiyotaka Takehara[†]

[†]New Product Technologies Development Dep., Panasonic Electric Works Co., Ltd.

[‡]New Business Development, Asahi Kasei Corporation

Abstract

Recently, in order for people to live in convenient and comfortable home, highly-functional household equipments increase in number. This forces users to operate complicated tasks to control such equipments. This paper presents our approach to improve the utility of the equipments utilizing speech recognition system. As an example, an experimental speech recognition system is developed and embedded in a bathroom environment, and evaluation results are described.

Key words: Speech recognition, Speech interface, Bathroom, Acoustic model

1. はじめに

本稿では、浴室を統合制御するコントローラへの音声認識機能の搭載事例について、音声認識システムの性能改善結果を報告する。

近年、車載ナビゲーション機器、携帯電話、家庭用ゲーム機器等で音声認識機能を搭載した商品の実用化が進んでいる。車載ナビゲーション機器では、運転中の視線移動や手動操作による危険を回避するなどの目的で音声認識機能が搭載されており、従来からの目

的地設定における検索タスクに加え、オーディオ操作や楽曲検索、エアコン制御などタスクに広がりが見られる。同様に携帯電話においても、従来からの音声ダイヤルに加え、ヘルプシステムでのキーワード検索や、家庭用ゲーム機器と同様に音声認識機能を利用して操作性を高めたゲームコンテンツなどが商用化されている。

しかしながら住宅分野においては、一部家電製品を除くと商品化事例も少なく、普及している状況ではない。特に、インターホンシ

システム、システムバス等、住宅引渡し時に既設されている住宅設備機器分野では音声認識機能の搭載事例は極めて少ない。

今回、我々はアイズフリー、ハンズフリーなどの音声認識の特長が、住宅設備機器においても操作性の改善や、利用シーンの快適化につながると考えた。特に、近年入浴目的の多様化により多機能化され、機器操作を容易に行うために音声操作の必然性の高いと考えられる浴室空間に着目し、音声認識機能の搭載を検討した。浴室空間は、住環境全体を想定した場合、高残響で、かつ高騒音という特徴があり、技術的にハードルの高い空間と考えている。

以下、第2章では浴室コントローラの試作事例を説明し、第3章に試作品評価によって得られた知見について報告する。

2. 浴室コントローラの試作

2.1 浴室システム

近年、浴室空間では暖房換気乾燥機、床暖房設備、給湯コントローラ等に加え、バスタテレビやジェットバス等、入浴を楽しむための設備も設置されることがある。しかしながら、機器ごとにコントローラが設置されるために、見栄えが悪い、操作したい機器のコントローラが分かりづらい、などの問題が生じつつある。

これらの問題への対策の一つとして、複数の設備機器を統合制御する浴室コントローラの検討を行った。統合化することにより操作項目が増加し、GUIでは階層構造が複雑になる可能性があること、湯船でくつろいだ姿勢のまま、眼鏡を外していても操作できることを目指し、音声インターフェースを用いたコントローラの試作を行った。

浴室コントローラと制御対象となる浴室空間の概要を説明する。

図1に、全体システム構成を示す。浴室内に設置した調光照明器具等の設備部材は、浴室外に設置した制御ユニットと制御信号線で接続されている。音声コントローラは、ユーザインターフェース部のみを搭載した表

示部と、音声認識部及び音声認識結果から各制御部への制御信号を生成する制御部からなる。

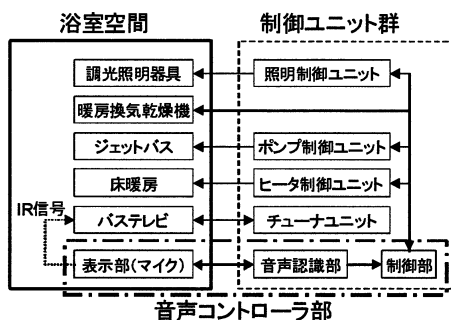


図1: 浴室システムの構成

浴室の仕様を表1に示す。浴室は、床材、壁材等の部材まで含む形態で販売・施工されるシステムバスを使用した。

表1:浴室仕様

内寸(m)	1.6(W)×2.4(D)×2.2(H)
材質	浴槽部:人造大理石 床:FRP 壁、天井:化粧鋼板
残響時間(秒)	0.7秒

2.2 音声コントローラ的设计

浴室内で音声認識を利用するにあたって利点を以下のように仮説設定し、コントローラ操作部と認識タスクを設計した。

- 小さなコントローラにて、操作性を向上できる
- 湯船でくつろいだ姿勢のまま操作できる
- 洗髪中など手が塞がった状態でも操作できる
- 眼鏡を外していても操作できる

コントローラ操作部と認識タスクの特徴を以下に示す。

- 音声操作モードとスイッチ操作モードの併用
ユーザビリティの観点からは、複数の操作モードを設けるのは好ましくはないが、入

浴者が複数名いる場合や、歌を歌う場合などを想定し、音声操作が無効となるスイッチ操作モードも搭載し、切り替え可能とした。

- 発話スイッチレス
音声操作モードでは、連続待ち受け状態となる。このため、水音などを誤受理しないことが重要である。
- コマンド&コントロールタスク
簡単な機器操作が主であるため、孤立単語認識によるコマンド&コントロールタスクで、語彙数は81語彙である。

音声コントローラの全体構成を図2に、音声認識処理を担当するコントローラ本体部の仕様を表2に示す。

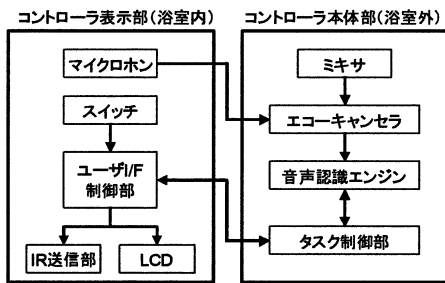


図2: 音声コントローラの構成

表2: コントローラ本体部仕様

項目	仕様
CPU	富士通 FR450(360MHz)
音声認識エンジン	旭化成 VORERO Ver.7.0
エコーキャンセラ	旭化成 VOCLE Ver.4.0

2.3 音声認識システムの改善

話者が浴室内のどこにいても認識できることが望ましいが、高残響な環境では話者とマイクとの距離が離れるにつれて、床、壁、天井における反射成分の遅延が大きくなり、認識が困難になる。また、騒音環境下でのSNRも低下する。そこで話者位置と認識性能について予備検討を行い、最も性能のよい条件として、浴槽に入った状態で顔をコントローラ側へ向けている場合を推奨仕様条件とし、本条件に対して音響モデルの残響と騒音適応を主とした最適化を行った。また、騒音の誤受理対策としては、浴室環境騒音から作成した雑音棄却モデルの追加を行った。その他の改善策としては、エコーキャンセラのパラメータの最適化、語彙の選別、文法表記における音素表記の最適化、音響モデル学習コーパスの最適化等を実施した。本報告では、改善効果の大きかった音響モデルの最適化と雑音棄却モデル[1]について改善結果を報告する。

改善前後の性能を表3に示す。音声認識率の評価は、2916発話(成人男女合計36名×81語彙)を防音室で収録した後、浴室環境で測定した伝達特性の畳込みと騒音の重畳により、浴室での発音音声を模擬して行った。誤受理頻度は、実入浴騒音とテレビ騒音に対して、1時間あたりの誤受理数と定義し、45分間のテレビトーク番組と、一般家庭6軒で収録した実際の入浴騒音を評価データとした。但し、表中のSNRについて、テレビ音声はエコーキャンセル処理前のSNRであり、認識処理を行った音声は、エコーキャンセラにより12dB程度SNRが改善されている。

表3: 音響モデル適応化前後の性能

	騒音種別	改良前	音響モデル適応化	音響モデル適応化+雑音棄却モデル
認識率(%)	騒音なし(SNR=36dB)	55.1	97.1	96.5
	ジェットバス(SNR=22dB)	55.1	94.4	91.8
	テレビ音声(トーク番組, SNR=8dB)	46.5	91.8	88.5
誤受理頻度(回/時間)	入浴騒音	0.7	19.1	2.8
	テレビ音声(トーク番組)	0.0	27.7	0.0

実験の結果、改良前の認識率は55.1%と低く、認識誤りの内訳は、2.2ポイントの置換え誤り（他の語彙に認識）と42.6ポイントの誤棄却である。浴室環境の残響時間は0.7秒であり、VOREROエンジンの想定している自動車空間や、通常の居住空間より大幅に長いために、音響的なミスマッチを生じ、その結果として誤棄却が生じたものと考えられる。音響モデルの浴室音響環境への適応を行うことで、認識率が大幅に改善するが、一方で誤受理も増加している。この原因は、音響モデルの再学習に用いた音声コーパスに、環境騒音も重畳したために、騒音にマッチしやすくなったためと考えられる。さらに、雑音棄却モデルを追加した場合、正しい発話音声が入って雑音棄却モデルにマッチすることによる認識率低下が僅かながら生じるものの、誤受理頻度を大幅に低下させ、総合的には高い認識性能と雑音棄却性能を得ることができた。

3. 試作機の動作確認

試作機の動作確認として、入浴実験を行った。

実際に入浴した場合の認識率を測定したところ、表3の結果よりも低くなる傾向があった。語彙外発話や言いよどみの影響を除くため、仕様を熟知した開発関係者8名により、自由タスクでの認識率を集計したが、騒音のない条件でも認識率は82.8%であった。ある程度の認識性能が得られているときには、正確な発話を行わない傾向や発話パワーの低下があり、認識性能が低下しているものと考えられる。また、騒音の誤受理は8名の入浴の総計で6回発生し、うち5回が体を洗う音などの入浴騒音で、残りの1回はテレビ音声であった。総入浴時間は約2時間であったので、表3の結果に比較的近い頻度である。

一般のモニター被験者4名でも入浴実験を行った。実験補助者が浴室内で付き添った上で、タスクと操作可能語彙の一覧を提示し、主に使用感に関する感想を収集した。得られた被験者意見の傾向を以下にまとめる。

- 音声認識に対する否定的な意見（声を出すのは恥ずかしい、等）は殆ど見られなかった。
- 便利である。特に、音声認識によって環境が自分にあわせてくれる感覚が好評であった。被験者によっては、「王様気分」という表現もあった。
- 洗い場などもっと遠くから使いたい。シャンプーをしながら使いたいなどの、使用制限の緩和を求める声があった。

ユーザインターフェースとしての使用性を評価するため、ユーザビリティ評価の専門家6名による入浴評価を行った。慣れていないユーザを想定し、認識語彙等の詳細な仕様は事前知識としては与えてはいないが、主要な機能は全て使用することができた。しかしながら、操作方法の学習が十分に進まない場合もあり、今後の改善が必要である。

4. おわりに

現在研究開発中の浴室向け音声コントローラの概要と入浴実験から得られた知見について述べた。ユーザインターフェースとして音声認識を利用する際には、ある程度認識性能が向上すると、認識率のさらなる改善よりも、「コントローラではなくテレビに向かって発話したい」などのように使用条件の制約緩和をユーザは期待する傾向も確認された。今後は、さらに遠距離からの発話、騒音環境下での発話に対応するために認識性能の改善と、利用者の納得性の高い使用条件について検討を進める必要がある。

本報告では、浴室環境をターゲットとして、試作・評価を実施したが、距離等の制約緩和とローコスト化を進めることにより、他の住宅内環境でも様々な利用シーンが広がると思われる。

参考文献

- [1] 奈木野, 小笹, 馬場, 日比谷, 竹原, 庄境, “浴室向け音声コントローラに用いる音響モデルの開発,” SLP-73-5, Oct, 2008.