慶應義塾大学学術情報リポジトリ
Keio Associated Repository of Academic resouces

Title	人間の行動支援のためのIoTスピーカを用いたモノが喋る環境の構築
Sub Title	Development of talking objects environment using IoT speakers for human behavior
Author	榎本, 敬太(Enomoto, Keita) 小木, 哲朗(Ogi, Tetsurō)
Publisher	慶應義塾大学大学院システムデザイン・ マネジメント研究科
Publication year	2020
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2020年度システムエンジニアリング学 第318号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40002001-00002020-0023

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

修士論文　　2020 年度

人間の行動支援のための IoT スピーカを用いたモノが喋る環境の構築

榎本敬太

（学籍番号：81933082）

指導教員　教授　小木哲朗

2021 年 3 月

慶應義塾大学大学院システムデザイン・マネジメント研究科

システムデザイン・マネジメント専攻

Master's Dissertation 2020

Development of Talking Objects Environment Using IoT Speakers for Human Behavior

Keita Enomoto

(Student ID Number：81933082)

Supervisor Testuro Ogi

March 2021

Graduate School of System Design and Management,
Keio University
Major in System Design and Management

論　　文　　要　　旨

学籍番号	81933082	氏　名	榎本敬太

論 文 題 目：

人間の行動支援のための IoT スピーカを
用いたモノが喋る環境の構築

（内容の要旨）

現代の情報化社会において、人々は多くの情報を基に生活していて、生活は多くの情報と結びついている。しかし情報を得る為には、自ら探し出し取りに行かなくてはならないという問題点が挙げられる。一方で、音声認識の精度が向上したことで音声インタフェースが普及、注目されている。

そのため本研究では、周りの環境が音声によって情報を与えることで、どんな人に対しても行動を支援することを研究目的とする。音声によって情報を提供するデバイス開発を行い、その有効性を検証していく。

ユーザに対して周りの環境が情報を与えるために、IoT スピーカを用いた「モノが喋る環境」の構築を目指した。開発した IoT スピーカは、あらゆるモノに取り付けられるように小型であり、インターネット経由で情報を取得し、その情報をインタラクティブに音声出力するデバイスである。

IoT スピーカを消毒液に取り付け、消毒液使用を促す音声出力を行うことによって消毒液使用率の変化を観測する検証、つまり行動支援を行うことができるかどうかの検証実験を行った。

結果は、適切な音声出力を行うことで消毒液が喋っている状態を再現することができ、消毒液が喋ることで消毒液使用率が向上した結果、「モノが音声により情報を与えることで行動支援する」ことの一端を担うことができた。

キーワード（5 語）

IoT スピーカ、音声コミュニケーション、コミュニケーションサポート、音声インタフェース、ヒューマンインタラクション

SUMMARY OF MASTER'S DISSERTATION

Student Identification Number	81933082	Name	Keita Enomoto

Title

Development of Talking Objects Environment
Using IoT Speakers for Human Behavior

Abstract

In today's information society, people live their lives based on a lot of information, and their lives are connected to a lot of information. However, the problem is that in order to obtain information, we have to search for it and go get it ourselves. On the other hand, as the accuracy of voice recognition has improved, voice interfaces have become popular and are attracting attention.

Therefore, the purpose of this research is to support the behavior of any person by providing information about the surrounding environment through voice. We will develop a device that provides information by voice, and verify its effectiveness.

In order to provide the user with information about the surrounding environment, we aimed to build an "environment where things talk" using IoT speakers. The IoT speaker we developed is small enough to be attached to any object, and is a device that obtains information via the Internet and outputs that information interactively.

We conducted a verification experiment to observe the change in disinfectant usage rate by attaching an IoT speaker to the disinfectant and providing voice output to encourage disinfectant usage, in other words, whether it is possible to provide behavioral support. We were able to reproduce the state in which the disinfectant was talking by providing appropriate voice output, and the disinfectant usage rate increased as the disinfectant talked. In addition, we were able to play a part in "behavioral support by giving information to objects through voice.

Key Word(5 words)

IoT speaker, Voice information, Communication support, Voice interface, human interaction

目次

第 1 章　序章 1
1.1　研究の背景 2
1.1.1　情報化社会の現状 2
1.1.2　音声インタフェース 4
1.2　研究の目的 6
1.3　先行研究 7
1.3.1　デジタルサイネージ 7
1.3.2　スマートフォンアプリケーション 9
1.3.3　公共空間の音声案内 11
1.3.4　スマートスピーカ 11
1.3.5　関連製品 14
1.4　本論文の構成 18
第 2 章　システム提案 19
2.1　モノが喋る環境 20
2.2　システム設計 22
2.2.1　IoT とは 22
2.2.2　本研究での IoT スピーカ 23
2.2.3　モノが喋る環境のための IoT スピーカ設計 24
2.2.4　KiCad を用いたプリント基板の作成 27
2.2.5　プリント基板に部品を配置 33
2.2.6　Arduino IDE を用いたプログラム書き込み 34
2.2.7　音声記号列の作り方 36
第 3 章　提案システムの基礎実験 39
3.1　人感センサの性能検証 40
3.1.1　検証方法 40
3.1.2　結果と考察 41
3.2　スピーカの音量の検証と結果 42
3.2.1　検証方法 42

3.2.2　結果と考察 42
3.3　インターネット接続速度の検証と結果 43
3.3.1　検証方法 43
3.3.2　結果と考察 44
3.4　スピーカの向きによる音の聞きやすさ検証と結果 45
3.4.1　検証方法 45
3.4.2　結果と考察 46
3.5　モノが喋っている感覚の要因検証 48
3.5.1　検証方法 49
3.5.2　結果と考察 50
第 4 章　提案システムのシステム検証と評価 53
4.1　妥当性検証の実験準備 54
4.2　音声出力によるモノの擬人化検証と結果 58
4.2.1　検証方法 58
4.2.2　結果と考察 59
4.3　提案システムのシステム検証と評価 61
4.3.1　計測方法 61
4.3.2　結果と考察 64
4.4　複数の IoT スピーカの連携 65
第 5 章　結論と今後の展開 67
5.1　結論 68
5.2　今後について 70
5.2.1　今後の課題 70
5.2.2　今後の展望 70
謝辞 71
外部発表 72
参考文献 73
付録 76

図目次

図 1-1　通勤時にスマートフォンを操作する人 2
図 1-2　グルメサイトを検索する人々 3
図 1-3　インタフェースの進化 4
図 1-4　駅にある音声による情報案内サイネージ 5
図 1-5　音声による操作が可能なエアコン 5
図 1-6　松田らの提案システムの概念図 8
図 1-7　張らのインタラクティブサイネージのシステム図 9
図 1-8　野村らの行動モデルによる行動支援システム図 10
図 1-9　坂宮らの津波避難行動支援システム図 10
図 1-10　田中らの服薬支援システムのシステム図 12
図 1-11　音声ガイドシステム外形図 12
図 1-12　Chappet (Pechat Inc.，2017) 14
図 1-13　TALKINGPOP(博報堂アイ・スタジオ，2017) 15
図 1-14　西村らの無電源小型通信端末を用いた情報支援システム 16
図 2-1　IoT スピーカの概念図 1 21
図 2-2　IoT スピーカの概念図 2 21
図 2-3　IoT スピーカのコンセプト 23
図 2-4　ESP-WROOM-02 の外観図 24
図 2-5　音声合成 LSI ATP3011F4-PU の外観図 25
図 2-6　焦電型赤外線センサーモジュール SB412A の概略図 25
図 2-7　オーディオアンプ IC NJM386BD とダイナミックスピーカ 26
図 2-8　IoT スピーカのシステム構成 26
図 2-9　「ESP-WROOM-02 ピッチ変換キット」のコンポーネント図 29
図 2-10　電気回路図 29
図 2-11　動作モードの選択 30
図 2-12　通信モードの選択 31
図 2-13　「ESP-WROOM-02 ピッチ変換キット」のフットプリント図 31
図 2-14　完成基板のレイヤー図 32
図 2-15　プリント基板の外観 33

図 2-16　IoT スピーカの外観.....33
図 2-17　ESP-WROOM-02 単体にプログラムを書き込む回路図.....34
図 2-18　USB シリアル変換モジュール（AE-TTL-232R）.....35
図 2-19　プログラムのフローチャート図.....35
図 2-20　Arduino IDE による IoT スピーカのプログラム（付録付き）.....36
図 2-21　IoT スピーカの音声出力の書き換え方法.....37
図 3-1　人感センサの反応距離検証の様子.....40
図 3-2　人感センサの方向による反応距離.....41
図 3-3　スピーカからの距離に応じた音の大きさ.....42
図 3-4　音声出力のタイムラグ.....43
図 3-5　人感センサの向きに応じた聞こえやすさ検証実験.....45
図 3-6　歩く速さ毎の発声距離の長さ.....46
図 3-7　歩く速さに応じた音の聞こえやすさ 5 段階評価.....47
図 3-8　実験の様子.....49
図 3-9　アンケート調査表.....50
図 3-10　音源と対象物が分離する感覚が生じた平均距離.....51
図 3-11　音源と対象物が分離する感覚が生じた平均角度.....51
図 3-12　アンケート調査の結果.....52
図 4-1　IoT スピーカのケース用の 3DCAD データ.....54
図 4-2　スライスデータの概念.....55
図 4-3　Prusa Slicer における消毒液ケースのスライスデータ.....55
図 4-4　Original Prusa i3 MK3S 3D プリンター.....56
図 4-5　消毒液専用ケース.....57
図 4-6　アンケート調査表.....58
図 4-7　アンケート調査の結果.....59
図 4-8　システム検証の実験場所.....61
図 4-9　システム検証の実験の様子.....61
図 4-10　システム検証の記録メモ.....62
図 4-11　IoT スピーカを取り付けた消毒液.....63
図 4-12　音声出力による消毒液の使用率比較.....64

図 4-13　IoT スピーカの連携その 1 65
図 4-14　IoT スピーカの連携その 2 65
図 4-15　IoT スピーカの連携その 3 66

第1章
序章

本章では、本研究の背景、研究目的、および本論文の構成について述べる。1.1 節で情報化社会の現状、音声インタフェースについて説明し、本研究を研究対象とするに至った背景について述べる。1.2 節では、1.1 節で述べた背景に基づき、本研究の目的とするところについて述べる。1.3 節では、本研究に関する先行研究、関連製品について述べる。そして、1.4 節では、本研究の構成について述べる。

1.1 研究の背景

1.1.1 情報化社会の現状

近年のインターネットの普及に伴い、現在は多くの情報に溢れた情報化社会となってきた。特にコンピュータ・ネットワークの普及により、生活に大きな変化をもたらした。世界中のどこの国や地域にもコンピュータがあり、衛星やケーブルの通信回線によって、地球的規模で接続されている。また、身近な電化製品にはコンピュータの機能が利用されたり、学校生活においても成績処理や名簿管理がされたりしている。街には、銀行の ATM、各種の自動販売機、コンビニエンスストアにあるレジ、コンサートのチケット予約、レストランの注文、などがある。また、コンピュータ・ネットワークの普及は、コミュニケーションの形態を変化させた。とくに、携帯電話での電子メールを読むために、携帯電話機画面を見るという光景を生んでいる[1]。

図 1-1　通勤時にスマートフォンを操作する人

例えば、図 1-1 のように、多くの人が通勤時にスマートフォンを操作して今日の天気やニュース、または株価の情報など多くの情報を得る。検索して得たその情報は、その日の行動に大きな影響を与える。また図 1-2 のように、情報端末を用いてレストランを決める際にはグルメサイトを閲覧し、口コミ情報によってお店を決める人が多い現状がある。

これらのように人々は多くの情報を基に生活していて、生活は多くの情報と結びついていると言える。

図 1-2　グルメサイトを検索する人々

しかしながら情報が誰でも簡単に手に入れられるようになってきた半面、様々な問題が生じてくる。デジタルデバイドである[2][3]。デジタルデバイドとは、パソコンやスマートフォンなどでネットを活用できる人と、まったく IT 機器に触れることができない人の間で生じる貧富や社会的地位の格差のことである。この問題を引き起こす原因の一つとして、「情報を得る為には、自ら探し出し取りに行かなくてはならない」という問題点を挙げる。情報端末を簡単操作できる若者と困難と感じる年配の方との差、情報収集能力の高い人と低い人との差が、情報を取りに行く際に差を生んでしまう。つまり情報を入手できるハードルが人によって異なり、その情報取得の差がダイレクトに生活の差に影響を及ぼすということである。

そのため、誰もが平等に情報を受け取る環境を構築することができれば、情報を取りに行く必要もなくなり、情報を取りに行く際の差が解消される。これに伴い、生活の格差の解消になることが期待される。

1.1.2 音声インタフェース

ユーザインタフェースとは、コンピュータシステムと使用者（ユーザ）の間で情報をやり取りするためのインタフェースである。

近年、音声インタフェースを備えたデバイスを用い、音声コミュニケーションによる操作でユーザビリティ向上を図る動向がある。一例として AI アシスタント機能を持つスマートスピーカの普及・利用が進んでおり、全世界での出荷台数は 900 万台を超えている[4]。また、スマートスピーカと連携させることで音声による操作を可能とした家電も見本市等に出品されるようになっている[5]。モノの操作をより簡便にする、低下してしまったユーザビリティを改善するために音声コミュニケーションによるデバイス操作が注目されている[6]。このように音声インタフェースが普及、注目されてきているのである。

いま、ユーザとシステムとのインタフェースはキーボードとマウスの時代からタッチパネルへ、そして音声の利用へと移り変わりつつある。

図 1-3　インタフェースの進化

図 1-3 は、技術の進化により時代と共にインタフェースが進化していることを表している。1970~80 年代はキーボードとマウスによるボタン操作が主であった。2000 年付近からマルチタッチ操作によりタッチパネルで操作することが可能となり、より感覚的に操作することができるようになった。そして、現在、音声やジェスチャー、視線、脳波など複数の技術を組み合わせることで画面操作も不要

なインタフェースが登場しようとしている。中でも音声によるインタフェースは言語解析機能の向上によりさらに普及している。

音声情報提示の例としては、例えば、歩行者支援システムやスマートフォン、音声による情報伝達が可能なデジタルサイネージ（図 1-4）、音声で操作が可能な家電（図 1-5）、あるいはスマートスピーカなどがあげられる [7][8][9]。これらは、音声による情報提示を可能としている。実際にこれら機器では音声案内によって、知りたい情報を容易に知ることができるようになってきた。

図 1-4　駅にある音声による情報案内サイネージ

図 1-5　音声による操作が可能なエアコン

このように、ユーザに対する情報提示に関しても、視覚情報だけではなく音声情報を利用する時代へと変わろうとしているのである。

1.2 研究の目的

現代の情報化社会から生ずる“情報は自ら探し出し、取りに行かなくてはいけない”という問題意識と、時代とともに進化するインタフェースの背景から“音声情報を利用する時代へと変わってきている”という現状がある。これを踏まえて、本研究の目的は「周りの環境が音声によって情報を与えることで、どんな人に対しても行動を支援すること」とする。

各個人が情報端末を用いて情報を探し出す必要をなくすために、情報を提供できる環境を整える。更に、視覚情報に対する音声（聴覚）情報の特徴として、「音」という刺激を感じるためユーザのアテンションを引き付けやすいことや、「人間味」を抱きやすいことが挙げられる。例えば、美術館や博物館では、壁などに書かれた説明を読むのではなく音声情報によって説明を提供すれば、小さい文字を読むことなく作品を見ながら作品の詳細情報を聴くことができる。更に、デジタルサイネージなどで使えば、目の前のユーザに音声を届けるため、必要な人にしっかりと商品情報を届けることが可能である。よって情報提供の手段は視覚によるものではなく、音声によって提供する。

そして提示情報による行動支援という要求を満たすために、音声によって情報を提供するデバイス開発を行い、その有効性を検証していく。

1.3　先行研究

本研究に関連した先行研究を挙げ、本研究の位置づけを示す。周囲環境からの情報提示に関連する研究として、視覚による情報提示の先行研究を紹介する。また、次に音声による情報提示の先行研究を紹介する。例えば視覚による情報提示では、ポスターやデジタルサイネージ、スマートフォンアプリやタブレットが挙げられる。聴覚による情報提示では、アナウンス放送や、エレベータの音声案内、スマートスピーカやスマート家電、アンドロイドロボットなどが挙げられる。その後、製品化されている情報提示デバイスについて紹介し、これらとの違いについて説明を行っていく。

1.3.1 デジタルサイネージ

周囲環境からの情報提示システムの関連研究として、視覚による情報提示システムとして、デジタルサイネージを紹介する。

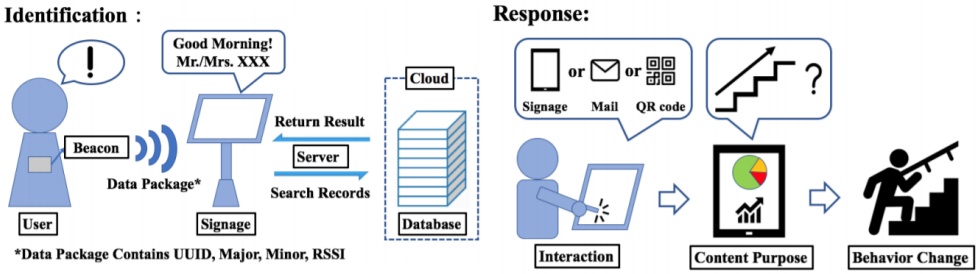
松田らは、デジタルサイネージを用いて個人の興味に対応した広告提示システムを構築した。図 1-6 は松田らの提案システムの概念図である。デジタルサイネージとは、屋外や店頭、交通機関など、一般家庭以外の場所においてディスプレイなどの電子的な表示機器を使って、情報を発信するものである。販売促進を目的とした広告以外にも、案内板や提示板といったパブリックディスプレイとして様々な場所で利用されている[10]。松田らが提案したシステムは、カクテルパーティ効果に基づき、サイネージの前を通った人の興味・関心をシステムが察知し、その人に応じたコンテンツを提示することで興味をひくことのできるデジタルサイネージシステムである。これはコンテンツの興味の強さに応じて歩く人の反応に大きな有意差が現れ、より興味の引くデジタルサイネージとして有効なシステムである。この研究は、通行する人の注目を集めるために個人の興味関心のある情報を適切なタイミングで提示を行っている。

図 1-6　松田らの提案システムの概念図

また、張らの「行動変容のためのインタラクティブサイネージの検討」について紹介する。近年の、人の健康状態を管理・把握するためのツールやアプリケーションが数多く登場しているが、アプリケーションを起動しなくては、自身の情報を知ることができず、この操作が不便につながっていると考え、デジタルサイネージで情報提示を行っている。この研究では、インタラクティブサイネージのユーザに特化した情報の提示とユーザへの印象の残りやすさという特徴を利用し、ユーザの行動変容を誘発させるためのデジタルサイネージの提案と検討を行っている。

提案するインタラクティブサイネージは、Bluetooth Low Energy（BLE）ビーコン、ビーコンの受信と情報提示を行うデジタルサイネージ、ユーザのデータの照合と提示する情報を決定するクラウドサーバから構成され、BLE ビーコンには iBeacon を使用し、ビーコンの Universally Unique Identifier（UUID）と Major と Minor の二つの識別子をビーコンの所有者の情報に紐づけることで、デジタルサイネージに近づいたユーザを特定する（図 1-7）。しかしながら、デジタルサイネージの問題の一つに Display Blindness が挙げられる。Display Blindness はユーザがサイネージをほとんど見ず、見たとしても遠距離で短時間であるというもので、デジタルサイネージの課題としては、ユーザに如何にサイネージを見ても

らうかが重要となる。BLE ビーコンによりユーザを特定し、そのユーザに特化した情報を、音声などを用いて提示することにより、ユーザの注意や興味を喚起する。提示した情報によりユーザの行動変容を誘発し、ユーザの健康の支援を行うシステムである[11]。



(a) ユーザ認識フェーズ　　(b) 行動介入フェーズ

図 1-7　張らのインタラクティブサイネージのシステム図

1.3.2 スマートフォンアプリケーション

情報提示による行動支援に関する関連研究として、野村らの「スマートフォンを用いた行動モデルによる行動支援システム」について紹介する。スマートフォンの普及により、多くの人が高機能な端末を携帯できるようになり、スマートフォンを利用した行動支援を行うことで、ユーザの日常を快適にすることができるという考えのもと、スマートフォンを用いたシステム提案をしている[12]。Android OS を搭載したスマートフォンを用いてユーザの行動履歴やプロファイルなどのデータ取得を行い、端末ディスプレイ上のアバタにより情報の提示を行う。Android 端末で取得するユーザの行動履歴はユーザのスケジュールやプロファイル、位置情報、加速度やジャイロセンサから得られるセンサデータである。

これら情報は一旦、端末内の簡易データベースに蓄積され、ホスト PC に転送され、ユーザの行動や習慣を抽出した行動モデルと活動モデルの生成を行う。これら、モデルの情報を再度 Android 端末に転送し、ユーザの状況やデータベースで蓄積された情報を基にシステムがユーザに適した情報を選択し、ユーザへの情報提供が行われる仕組みである（図 1-8）。

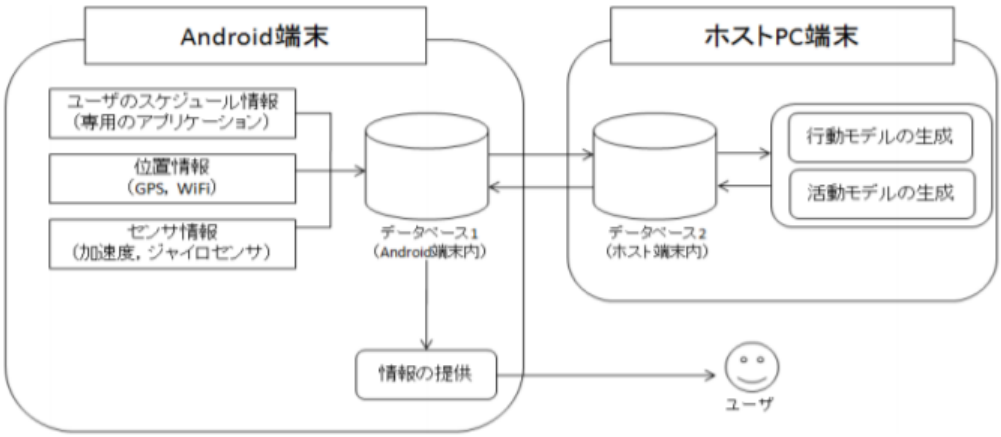


図 1-8　野村らの行動モデルによる行動支援システム図

また、板宮らの「津波避難行動支援スマートフォンアプリの開発と実証実験」について紹介する。東日本大地震の教訓から、津波発生時の迅速かつ安全な非難を可能にするための情報伝達手段の強化と多様性が求められている背景のもと、スマートフォンの「デジタル皆助ナビ」システムを提案している。このシステムは、身近な情報表示機器に、災害時の避難経路上における危険な方向を明示し、よりリスクの少ない経路による避難誘導を支援するものである（図 1-9）。また、災害発生時に携帯電話等の通信網が利用不可になる場合を想定して、現在位置情報に基づいた最適な避難経路、危険情報をパターン化して各端末に保存しておくことでオフラインでも作動可能である。カーナビやナビアプリの従来の「行くべき経路」を示すのではなく、「行ってはいけない経路」を示すことで、リスクの少ない経路の誘導非難を実現し、行動支援するものである [13]。

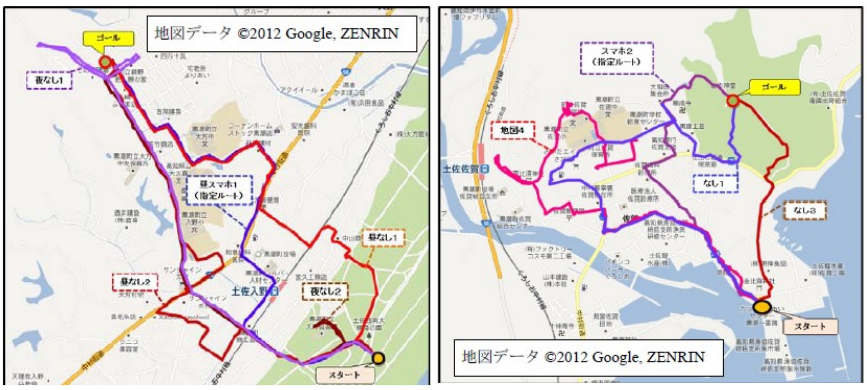


図 1-9　坂宮らの津波避難行動支援システム図

1.3.3 公共空間の音声案内

音声情報システムは、駅や公共空間でのアナウンスやトイレまでの経路の音声案内、エレベータの「開く扉は反対側です」などの音声案内がある。これらのシステムの特徴は、事前に録音した音声を出力することで行動支援を行っているといえる。国土交通省の「旅客施設における音による移動支援方策ガイドラインの策定について」では、コンコースからのエスカレーター、またトイレに音声案内が用いられている。特に視覚障がい者の音声案内に対するニーズに応えたものである[14]。

1.3.4 スマートスピーカ

関連研究として、田中らの「スマートスピーカを使った独居高齢者のための服薬支援システムの開発」について紹介する。

処方された薬を飲み忘れる問題は、患者の健康状態を悪化させるだけでなく医療費の増大を招くという背景のもと、この問題を解決するために、スマートスピーカを用いた服薬支援システムを開発している。システムは、QR コードから服薬情報を読み取るスマートフォンアプリと服薬時刻になるとユーザに服薬を促すスマートスピーカから構成されている。服薬情報はクラウド上のカレンダーで管理され、服薬状況を確認できる。また、遠くに離れた家族に服薬状況を通知する機能も備え、見守りシステムとしての利用もできる。3 回にわたる運用テストの結果、不具合は改善され、問題なく作動するようになった一方、服薬アドヒアランスの向上には関係者間の連携が求められるが、これは情報通信技術だけで解決できるものではない。しかしながら、このシステムは、関係者間の連携と情報交換を支援するツールとしての可能性を秘めている[15]。図 1-10 は、スマートスピーカを用いた高齢者に優しい服薬支援システムのシステム図である。このシステムは、音声と視覚による情報提示によって行動を支援する。

図 1-10　田中らの服薬支援システムのシステム図

また博物館の音声ガイド案内の研究として、加藤の「博物館のサウンドスケープ・デザインにおける音声ガイドの考察」を紹介する。ここでは、音声ガイドの実例を紹介している。「よみがえる黄金文明展：ブルガリアに眠る古代トラキアの秘宝」では、音声ガイド（図 1-11）によって 170 点ほどの作品の中から主要な作品 26 点の解説を聞くことができる。その音声は舞台声優の声を事前に録音されたものを再生するシステムである。展示品の見所やより深い鑑賞のために知っておきたい事柄・時代背景などをわかりやすくまとめてあり、壁の解説パネルを読む必要がなく、解説を聞きながら同時に作成を鑑賞することができる。
このシステムの利点として、聴きたいところを聴きたいときに何度でも聴けるように、作品番号と再生ボタンを押すだけの簡単な仕組みになっている。これによって、展示室のソファに座ってゆっくり聴き直すことができるだけでなく、歩きながらや食べ物を食べながら聴くことが可能となる[16]。

図 1-11　音声ガイドシステム外形図

先行研究では、適切な情報を適切なタイミングで提供するために、様々なシステム提案がなされてきた。デジタルサイネージの関連研究では、システムに近づく人を特定し興味・関心のある情報を提示して、行動変容を誘発してユーザの健康を支援することや、通行する人の注目を集めようとしたシステム提案である。スマートフォンの関連研究では、スマートフォンを用いてユーザの行動履歴やプロファイルなどのデータ取得を行い、それを基にレコメンド情報を提案するシステムや、津波避難経路を取得するために最適な情報を提案してくれるシステム提案である。音声案内のシステムは、事前に録音している音声を公共空間にて出力することで人々の行動を支援する。スマートスピーカの関連研究では、音声情報によって薬の服用を支援するシステムである。そのシステムは、スマートフォンアプリと連携しているので設定時刻になったら音声による情報が提示される。博物館の研究では、事前に録音した音声を、情報端末を用いて音声提示することで多くの人へ効果的な展示物の解説を提供している。

今までの音声提示を行うシステムは、予め音声を発するための仕組みが内蔵され、音声情報提示を前提に設計されたものである。これに対し、予め話すことを前提に設計されていないモノを含め、あらゆるモノに音声提示機能を付加することができれば、ユーザの興味を引き効果的な情報伝達ができることが期待される。例えば、植物のような有機物、柱や壁等の建造物、これらが喋ることでユーザは対象物への興味を増し、注意を向けることが考えられる。また音声に対してポジティブに感じるようになり、音声インタフェース利用の促進に貢献すると考えられる。本来であれば、IoT 化できるはずのないモノに情報出力機能を持たせて IoT 化することで、モノからのダイレクトな情報提示を行い、注目を集め情報提示を効率的に行うシステムを提案する。

1.3.5 関連製品

次に、あらかじめ音声機能が内蔵されていないモノに発音機能を与えるシステムに関して、幾つかの関連製品を紹介する。

【Chappet】

図 1-12 にある Pechat 社が開発した Chappet は、ぬいぐるみに取り付け可能なボタン形状の小型スピーカである[17]。このシステムはスマートフォンの専用アプリで発話内容を操作することが可能で、子供に対してぬいぐるみと喋っている感覚を与えることができる。スマートフォンのアプリが Bluetooth 通信で会話内容を小型スピーカへ送信し、スピーカが音声出力するという仕組みである。

図 1-12　Chappet (Pechat Inc.,　2017)

専用アプリケーションには、既存のライブラリがあり「一緒にあーそぼ！」や「ありがとう」のフレーズが多々含まれる。また、セリフを自由に入力することが可能で好きなフレーズを音声出力可能であり、子供の名前を呼ぶ等で親密性を向上させる等が考えられる。更に歌の歌詞も含まれており、歌を歌うことを実現できる。これによって小さい子供は、眠りにつくことができるであろう。この Chappet の特徴は非常に小型であることである。故に、ぬいぐるみに取り付けても単なるボタンである Chappet ではなく、ぬいぐるみが言葉を持って喋っているように感じさせることができる。

【TALKINGPOP】

図 1-13 は、株式会社博報堂が開発した TALKINGPOP というプロモーションツールである[18]。これは、お店の棚に喋る商品を作り出す新しい形のプロモーションツールであり、野菜などの商品に触れた際の静電容量の変化を感知し音声を提示することで、その商品から音声を出力しているような状況を作り出す。具体的には、静電容量センサがあり、野菜に間接的に接触していて、人が野菜に触れることで野菜の静電容量が変化し、静電容量センサが反応して事前に録音してある音声が出力される仕組みである。これにより、来店客の注意喚起や商品理解の促進、また店員と来店客とのコミュニケーションの増加がみられ、店舗の雰囲気の変化に貢献している。

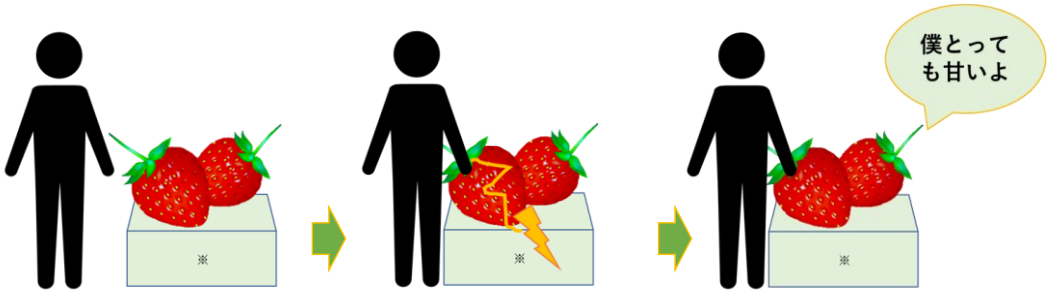


図 1-13　TALKINGPOP(博報堂アイ・スタジオ，2017)

TALKINGPOP の特徴は、本来喋るはずのない野菜に声を出す機能を与えたことである。静電容量センサを用いたことで、適切なタイミング、つまりユーザが野菜に近づき音声が聞こえるタイミングで音声が出力できる点である。

【CoBIT】

更に、西村らは、環境やユーザが提供するエネルギーのみで、環境側の装置およびユーザとの情報の送受信を実現する小型通信端末(Compact Battery-maintenance-free InformationTerminal)（以下、CoBIT）を用いた位置に基づく情報支援システムを提案している[19]。環境装置の構成は、赤外 LED 付きカメラ、ガルバのミラー、指向性赤外 LED などであり、CoBIT 側はクリスタルイヤホン、太陽電池、反射板等で構成される。

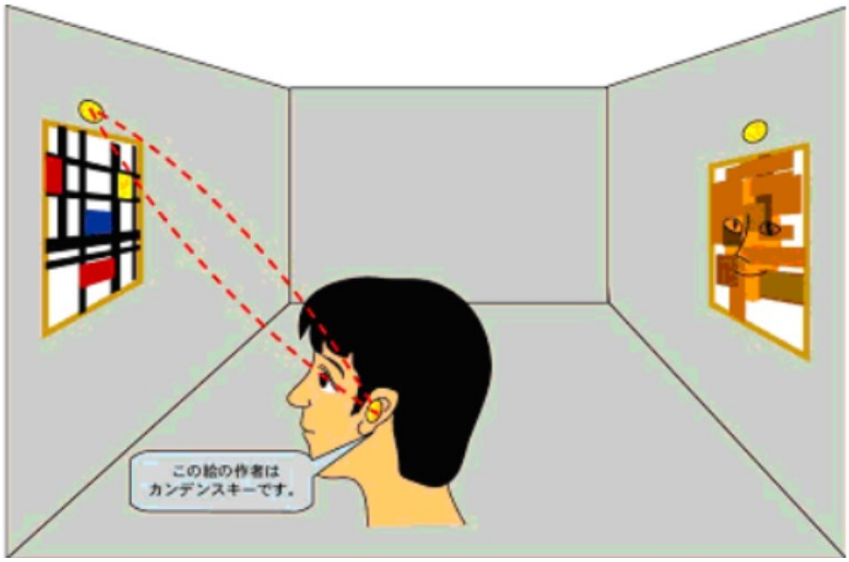


図 1-14　西村らの無電源小型通信端末を用いた情報支援システム

システムのフローとしては、環境装置側がビームにより光エネルギーを発射する。CoBIT は、その光エネルギーを光電変換素子で電気エネルギーに変換し、それを電源として用いる。得られた電力でイヤホンの音声出力や、振動子による振動が可能になり、ユーザがこれを感知するという流れである。また、環境の光エネルギーを反射して、反射光の強度を音声や手の動きで変調することができる。つまり、ある場所である方向を向いたときに見えるものについて、適切な情報を提供するのである。図 1-14 は、左側の壁の赤黄青の絵の方向を向いた場合に、この絵に関する情報が端末から音声出力される様子を表している。

CoBIT の特徴は、誰もが自然に端末を利用可能にするために無電源化となっている。これによって、バッテリーが必要なくなりメインテナンスが不要となり超長時間の使用が可能であること。また、持ち歩くことを想定して大変小型である。また環境システムと端末が連携して、適切な情報を、適切な場所、適切なタイミングで提供することができる点である。

Chappet や TALKINGPOP は、小型デバイスを取り付けることで様々なモノがユーザの興味を喚起し情報提供を行い、新しい価値を生み出している。Chappet は小型のスピーカで、スピーカからの音声出力によって、モノが喋る状態を再現している。また TALKINGPOP はセンサを用いて適切なタイミングで情報を提供し、事前に録音された音声をいいタイミングで提供することが可能である。更に CoBIT では、人が端末を耳に装着して、その端末が環境システムと連携することで適切な情報を適切なタイミングで提供している。

本研究では、従来の音声提示の特徴であった事前に録音された音声を出力するのではなく、「ネットワークを介してより柔軟に喋る機能」を付加することで、音声を使用した効果的なコミュニケーション環境を構築するものである。更に第三者が行動支援を目的に情報を提示する音声情報ではなく、モノに音声提示の機能を与え、モノが自分の意志で喋るかのように情報を提供することを目標とする。具体的に、博物館の音声案内というのは、展示物に詳しい人が説明を行っている音声内容であるが、これとは異なり展示物自身が意思を持ちが自分のことについて、自己紹介するかのように、展示物の情報を提供する音声案内を目指すということである。

つまり、ユーザに特定の操作を求めず、喋るはずのないモノに音声出力機能を与えることで「周りの環境が情報を与える」状態を目指すものである。

1.4 本論文の構成

本論文は全 5 章で構成されている。

第 1 章では、情報化社会の現状、音声インタフェースについての研究背景を述べ、研究目的、先行研究、そしてここで本論文の構成について述べる。
第 2 章では、提案するシステムの概念「モノが喋る環境」を述べ、これに必要な機能や定義、それに基づいたシステム設計を詳細に述べていく。
第 3 章では、第 2 章で構築した提案システムの基礎検証について述べる。
第 4 章では、第 2 章で構築した提案システムの妥当性確認の検証について、実験準備内容と検証実験を述べる。
第 5 章では、結論と今後の展開について述べる。

第2章
システム提案

本章では、モノが喋る環境の構築のための IoT スピーカシステムの概要について述べる。2.1 節ではモノが喋る環境の概念を述べ、2.2 節では IoT スピーカの説明、提案するシステム構成、また設計について述べる。

2.1 モノが喋る環境

モノが喋る環境とは、ユーザに対し状況に応じた音声サービスを、モノを通して提供する環境の実現である。様々なモノに取り付けられ、ユーザに情報発信する小型のスピーカを取り付けることでユーザとモノの情報のやり取りを実現する。従来の機械システムの多機能化や複雑化の問題に対しては、システム自体をシンプル化やアップデートすることでユーザビリティを向上させることで対処してきたが、このモノが喋る環境は、ユーザに対して周りの環境が情報を与えてくれることで、ユーザの行動を支援する。またモノとは、必ずしも機械システムだけではない。本来であれば、IoT 化できるはずのないモノに情報出力機能を持たせることで、モノからのダイレクトな情報提示を行うシステムを提案する。

例えば、図 2-1 のように、スーパーマーケットでリンゴが自分自身のセール情報や産地情報を伝えてくれる。これによって、ユーザは情報を受け取ることができる。リンゴが「青森産のリンゴだよ！」と音声情報発信することで、通り過ぎるユーザはその情報に関心を持てば、いい宣伝効果になるというわけである。これのメリットとしては、スーパーマーケットでは多くの人が野菜や果物の産地情報などを、スマートフォンをポケットから取り出し、ウェブブラウザを開き、その情報について検索することは考えにくい。そのため、簡単に知りえない情報を対象物からダイレクトに、かつ特定の操作なしに、スマートフォンなしに、情報を知ることができる点である。

また図 2-2 のように小型のスピーカを壁やドアに取り付けることで、壁やドアが道案内してくれる等である。対象物は何でもよいが、例えば時計や花瓶、絵など公共空間にあるものであればよく、それらに IoT スピーカを取り付ける。これによって、ユーザは知りたい情報を、スマートフォンを操作することなく対象物からダイレクトに情報が発信されるので目的地に行くためには比較的簡単に到着できると考える。

これらのように、音声を発することがあらかじめ内蔵されていないものから

音声が出力されることで、目新しさが生まれ注目を集めることができ、音声情報の内容に注目することが期待される。モノが喋ることによる、モノからのダイレクトな情報提示によって、状況に応じた適切な情報を音声によりユーザに提供することが可能になると考える。

図 2-1　IoT スピーカの概念図 1

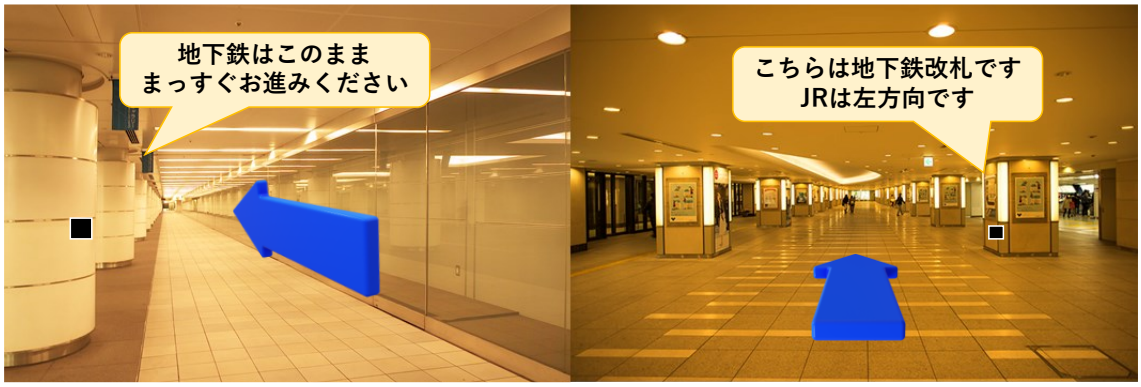


図 2-2　IoT スピーカの概念図 2

2.2 システム設計

前章で述べた目的と先行研究から、本研究で提案する IoT スピーカについて検討を行った。

2.2.1 IoT とは

IoT とは「Internet of Things」の略であり、「モノのインターネット」と呼ばれ、建物、電化製品、自動車、医療機器などパーソナルコンピュータやサーバ以外のものがインターネットに接続されて、相互に情報をやり取りすることである。かつてインターネットはコンピュータ同士をつなぐものであったため、従来はパソコンやサーバなどの IT 関連機器が接続されていた。しかし、現在は PC 等だけでなくカメラ、イヤホン、スピーカや家電など新しいモノがインターネットに接続されてきている。このようにいままで、インターネットに接続されなかったものをインターネットにつなぐことを「IoT」と呼ぶ。

IoT によって実現できることは、IoT でモノを操作する、IoT でモノの状態を知る、IoT でモノ同士の対話をするである。IoT によってセンサ機能とインターネット接続機能を持つモノが、インターネットを介して離れている場所で情報を活用することができるようになるのである。またネットワークが届く場所であれば、特殊な場所を除いてどこでも適用場所となる。

IoT で使用する通信手段は、いくつかある。現在も、WiFi(無線 LAN)や 3G、LTE、5G、WiMAX といった無線方式がありインターネットの接続利便性を向上させる。データ伝送には高速の通信網（ブロードバンド）と低速の通信網（ナローバンド）がある。パソコン類は映像や音楽など、大きなデータ通信を扱う必要があるので高速通信網が必要である。一方で、IoT によって新たにインターネットに接続される機械やモノが扱うデータ量は非常に小さいので低速通信網で充分である。IoT の場合、接続されるモノの数が多くなり同一エリア内に多数のノード接続が必要とされ、それに対応できる無線技術が必要である。多くのモノが無線通信をするためには低速通信網（ナローバンド）の無線を使用し、使用帯域を小さくして通信が込み合うことを避ける必要があるのだ[20]。

2.2.2 本研究での IoT スピーカ

本研究で考える IoT スピーカとは、以下の 4 つの特性を備えているものと考える。

まず、あらゆるモノに取り付けられるように小型である。次にインターネット経由で情報を取得するためにインターネットに接続する機能を持つこと。次に、情報提示するために、音声情報を出力する機能を持つこと。最後にインタラクティブに情報提供を行う機能を持つこと。音声出力の機能を持たせた理由については、音声インタフェースを備えたデバイスを用い、音声コミュニケーションによる操作でユーザビリティ向上を図る動向が実際にあり、音声による家電操作や AI スピーカのやり取りから、音声はちょっとした情報提供に向いていると言えるためだ。インタラクティブな情報提供のためには人感センサを用いることができる。これによりユーザが特別な操作なしに、いつでもどこでも利用できる。

また、今までの研究や製品において IoT 化は電子製品が対象であり、喋る冷蔵庫、喋るエレベータなど様々な IoT 製品が生まれてきた。しかし、本研究では野菜や果物、植物、壁に IoT スピーカを取り付けることで本来喋るはずのないモノに声を与えるデバイスを提案し開発する。

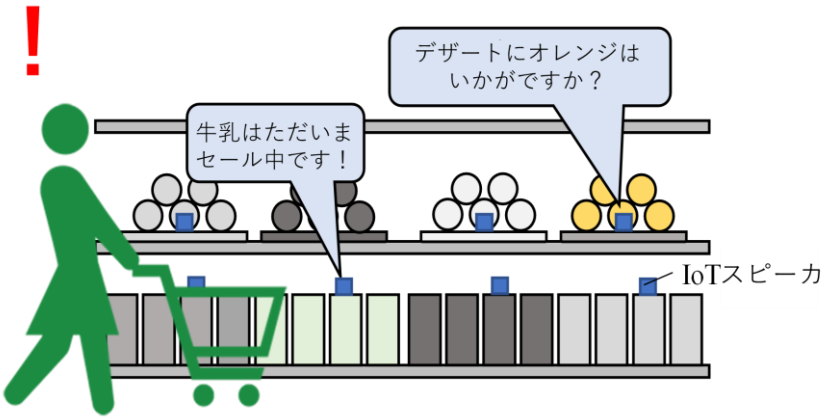


図 2-3　IoT スピーカのコンセプト

図 2-3 は、IoT スピーカの使い方のコンセプトである。この図は、スーパーマーケットで、近づいてきたユーザに対して、商品に取り付けられた IoT スピーカがデザートや食品の情報を音声によって発信することで、ユーザは音声に反応し、商品に注目する状態を表している。

2.2.3 モノが喋る環境のための IoT スピーカ設計

本研究ではいろいろなモノに取り付けられる小型の IoT スピーカの開発を行った。IoT スピーカに求められる要素としては、あらゆるモノに取り付けられるように小型で、人感センサを備えることで人の動作に反応し、ネットワーク経由で情報を取得し、音声情報の出力を行うことで、対象物がユーザに対してインタラクティブに情報提供を行うことができることである。

用いた部品は、インターネット経由で情報を取得するための WiFi 機能を有する ESP-WROOM-02、音声出力のための音声合成 LSI ATP3011F4-PU にオーディオアンプ IC NJM386BD とダイナミックスピーカ、インタラクティブに情報提供を行うための焦電型赤外線センサーモジュール SB412A である。

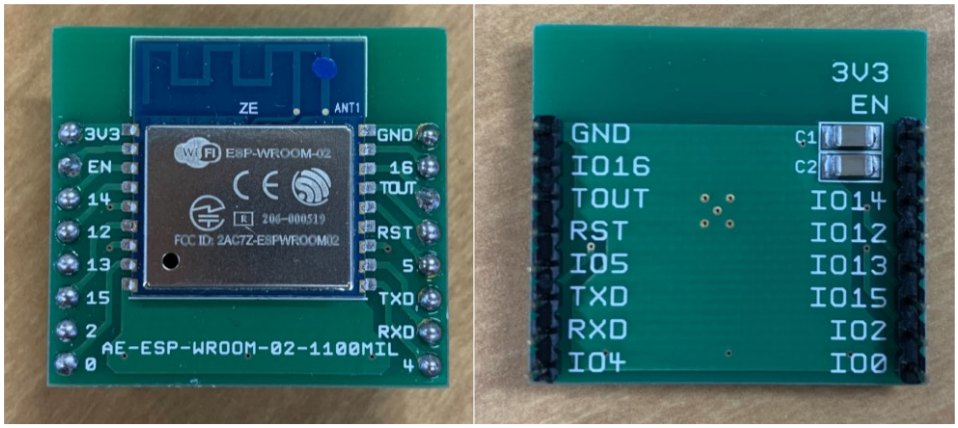


図 2-4　ESP-WROOM-02 の外観図

ESP-WROOM-02 とは、ESP8266EX というチップを搭載した WiFi 通信が可能な小型モジュールのことである（図 2-4）。ESP-WROOM-02 の特徴としては、小さいサイズながらも Arduino のプログラムを書き込め（Arduino として利用出来る）、また WiFi 通信が可能、そして安価というメリットがある。これらメリットにより電子工作に多く用いられている[21]。本研究でこの ESP-WROOM-02 を用いた理由としては、インターネット通信が可能で、小型である点を考慮して最も適切であると判断したためだ。

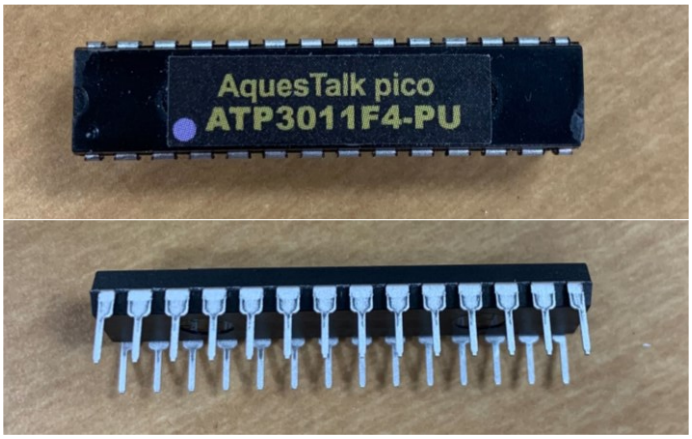


図 2-5　音声合成 LSI ATP3011F4-PU の外観図

音声合成 LSI ATP3011F4-PU（以下、音声合成 LSI）とは、音声合成エンジン AquesTalk pico を用いた小型サイズのパッケージ 1 つに、音声合成機能のすべてを実装している小型モジュールである（図 2-5）。1 チップでローマ字表記のテキストを自然言語による音声出力に変換する処理を行う音声合成 LSI である[22]。LSI とは、Large-Scale Integration の略であり、IC に比べてより集積度の高い複雑な回路をおさめた集積回路という意味である。大規模集積回路、また高密度集積回路と呼ばれる。本研究で用いる音声合成 LSI は LCD モジュールを用いるような感覚で、簡単に音声合成機能を実現可能で、使い方も簡単であり、ローマ字表記の文字列を、シリアルインターフェースを介して送るだけである。あとは AquesTalk pico LSI がリアルタイムに読み上げる。

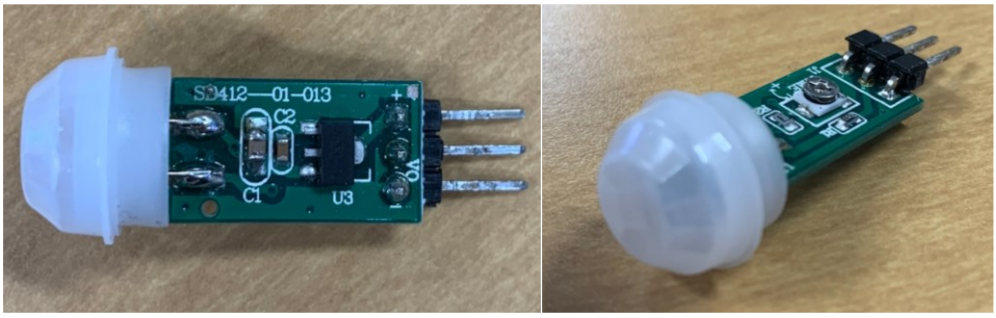


図 2-6　焦電型赤外線センサーモジュール SB412A の概略図

焦電型赤外線センサーモジュール SB412A（以下、人感センサ）とは、自分から検出の為の赤外線を発光するのではなく、人体から放射される赤外線を受け

ることによって検出をするモジュールである（図 2-6）。センサ自身が人体検出のために常時レンズから光を出力しているのではなく、入射赤外線の変化を検出することにより人体を検出するためパッシブタイプと呼ばれている。このため低消費電力というメリットがある。しかしながら、熱源である人体と床や壁などの背景との温度差に応じて動作しまうこともあるため、人以外の熱源によっての温度差が生じた場合でも動作するので注意が必要である[23]。

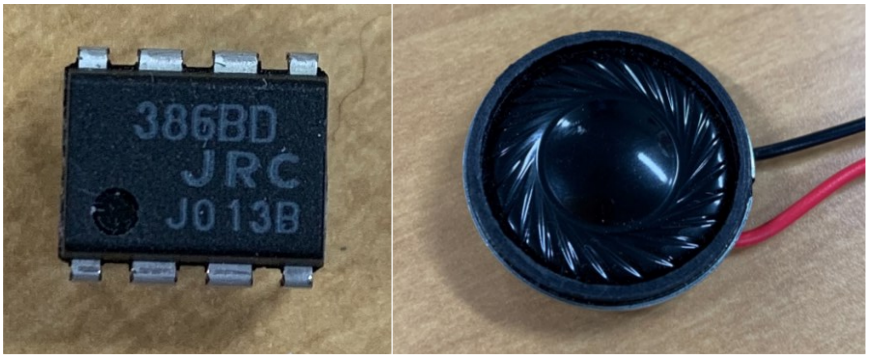


図 2-7　オーディオアンプ IC NJM386BD とダイナミックスピーカ

図 2-7 は、オーディオアンプ IC NJM386BD（以下、オーディオアンプ IC）とダイナミックスピーカである。アンプとはアンプリファイヤー（Amprifier）の略で、増幅器という意味である。オーディオアンプ IC とは、オーディオ（本研究では、音声出力）信号を増幅または、制御を行い、サウンドを拡張することや音質を変える能力を持ったデバイスである。

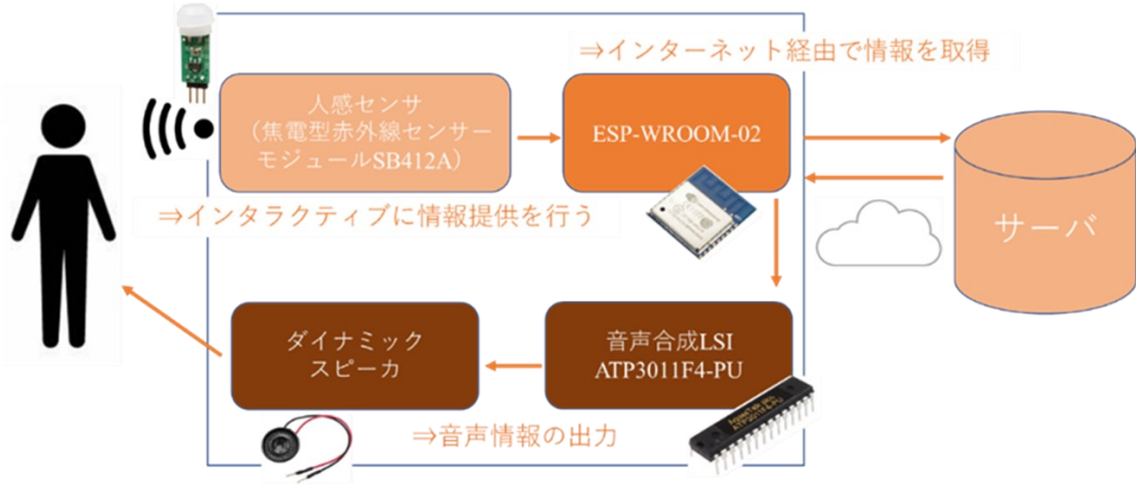


図 2-8　IoT スピーカのシステム構成

そして、これら部品を用いて作成した IoT スピーカのシステム構成とシステムのフローを図 2-8 に示す。インタラクティブに情報提供を行うための人感センサは、人の動きを感知すると、感知した情報を ESP-WROOM-02 に送る。ESP-WROOM-02 は、インターネット経由でサーバ上の指定されたファイルにアクセスし、文字列を取得する。その取得した文字列情報を音声合成 LSI に送り、音声合成 LSI は文字列情報を、音声情報に変換しダイナミックスピーカで出力するという仕組みである。

人感センサが、ユーザの動きを検知すると用意されたサーバ上の文字列を読み込み、スピーカがそれを発声情報として出力する。この時、サーバ上の文字列は任意に書き換えが可能で様々な言葉を発声させることができる。IoT スピーカのソフトウエアは、Arduino IDE で開発を行い、ESP-WROOM-02 に書き込みを行うことができる。この際、IoT スピーカは毎回サーバにアクセスしてその都度用意された音声情報を提示するサーバ音声の方法、事前にサーバにアクセスしデバイス内にプリセットされた音声情報を提示するプリセット音声の方法を切り替えることが可能である。この IoT スピーカをいろいろなモノに取り付けることで、あらゆるモノが喋る IoT スピーカ環境を構築することができる。

2.2.4 KiCad を用いたプリント基板の作成

IoT スピーカを作成していく段階で、小型化を目指すために電子部品を取り付けるプリント基板を、KiCad を用いて作成する。プリント基板は、電子部品を 1 カ所にまとめてスペースを有効活用できるため、製品自体の小型化も可能となる。

KiCad は、オープンソースの EDA 用のソフトウエアスイートであり、電子回路設計とプリント基板への変換を容易にする。 KiCad は、Jean-Pierre Charras により開発され、回路図エディタと PCB 配置設計のための統合環境を特徴としているものだ。KiCad のコンテンツは、大まかに Eeschema、Pcbnew、GerbView、の 3 つがある。Eeschema（以下、回路図エディタ)は、部品シンボルの表示、部品と部品を配線する機能を持つ。基本的な部品は既存のライブラリに掲載されているが、必ずしもすべての部品が掲載されているわけではなく、シンボルを自

作しなければならない。Pcbnew は（以下、PCB エディタ）、回路基板レイアウトエディタであり、プリント基板を作成するためのデータを作成する機能を持つ。回路図エディタで作成した回路図から基板外形データに変更するために、実際の部品のサイズや形状のデータであるフットプリントを表示する。しかし、シンボル図と同様、ライブラリに基本的な部品は掲載されているが、必ずしもすべての部品が掲載されているわけではなく、自分でフットプリントを作成しなければならない。GerbView は、ガーバービューアと呼ばれ、PCB エディタで作成した基板製造データ（発注する規格のデータ）を確認することができる。民間の企業へのプリント基板発注は、データの規格がありそれらに対応するために変換可能である。

次に、KiCad を用いてプリント基板のデータ作成を行う。

—回路図エディタの操作—

1. シンボルを配置

電源関係のシンボルは、電源 VCC シンボル、GND シンボルである。電源は 3V、5V 等必要に応じて様々な選択肢がある。本研究で用いる電子部品の対応電圧は、それぞれのデータシートから ESP-WROOM-02：3.3V~5.0V また、音声合成 LSI：2.5V~5.5V、またオーディオアンプ IC：4.0V~18V、人感センサ：3.3V~12V より、適切な範囲の電圧は 4.0V~5.5V である。よって、電源シンボルは 5.0V を選択した。

また今回作成する IoT スピーカは、人感センサが人の動きを検知したタイミングを可視化するために、LED を取り付ける。ESP-WROOM-02 には、電源と GND 以外に先ほどの LED と、LED に直列に接続する抵抗、更に別の足に抵抗 4 個をそれぞれ接続し、プルアップとプルダウンさせる。オーディオアンプ IC には、電源と GND 以外にコンデンサを接続する。

音声合成 LSI に、オーディオアンプ IC、ダイナミックスピーカ、人感センサに最適なシンボルを見つけ、シンボルを配置した。ここで本研究では、ESP-WROOM-02 単体を使用するために「ESP-WROOM-02 ピッチ変換キット」（秋月電子通商）を使用した。

2. 使用する部品シンボルを作成（既存ライブラリにないものは自作する）

KiCad の既存ライブラリには ESP-WROOM-02 変換キットのシンボル図が存在しない為、専用のシンボルデータを自作した(図 2-9)。秋月電子通商の専用のページのデータシートを参考に作成した。

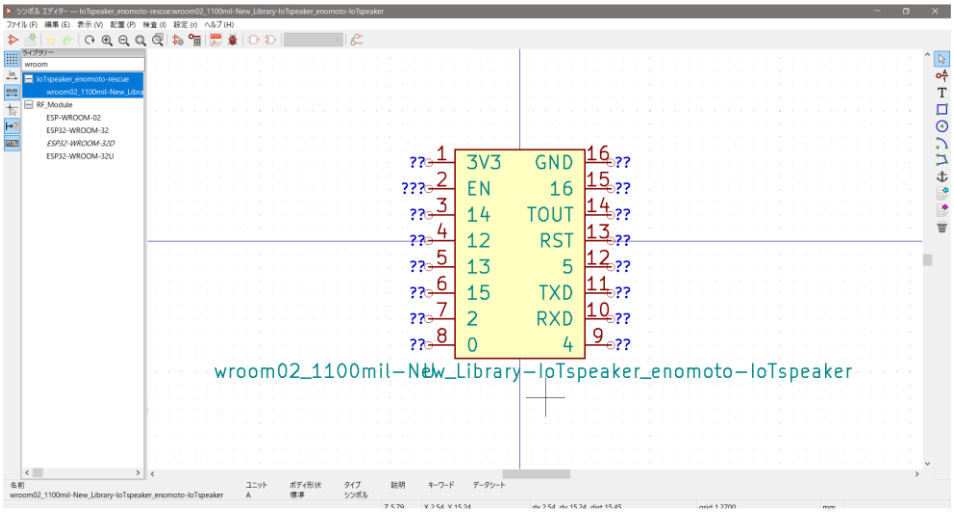


図 2-9 「ESP-WROOM-02 ピッチ変換キット」のコンポーネント図

3. 抵抗器、コンデンサの値を入力して部品同士を配線

ESP-WROOM-02 に接続する抵抗は、100Ω の抵抗 1 つと 10kΩ の抵抗 4 個である。また、オーディオアンプ IC には 100μF のコンデンサを配置する。図 2-10 は電気回路の完成図である。

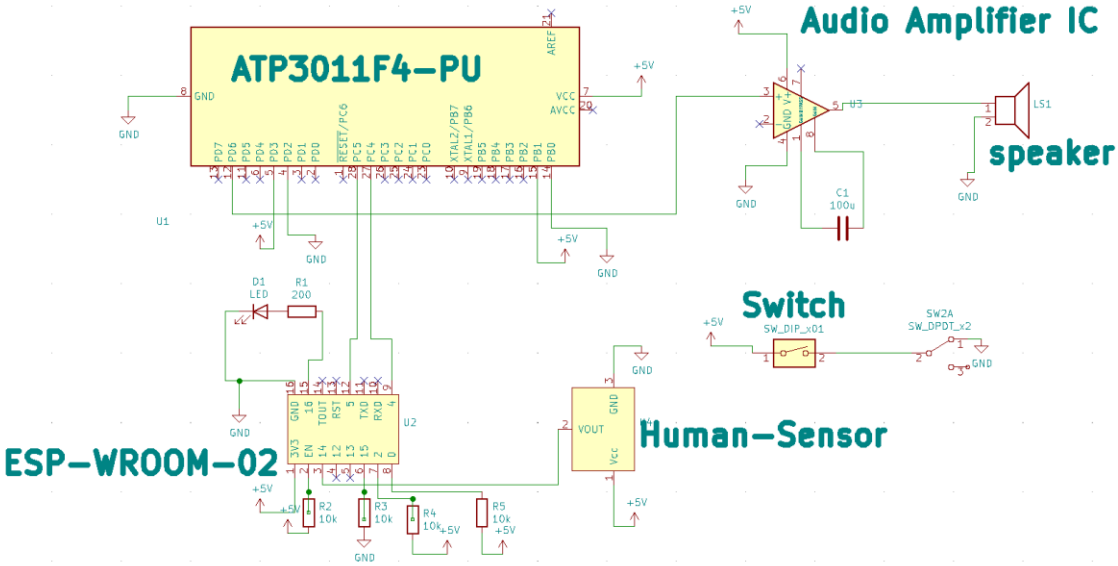


図 2-10 電気回路図

まず、部品間の配線を、データシートを参考に行い以下に示していく。

- 音声合成 LSI とオーディオアンプ IC

  12.AOUT（音声出力端子）　－　3.+INPUT

- 音声合成 LSI と ESP-WROOM-02

  27.SDA(I2C Serial Data)　－　4.GPIO4

  28.SCL(I2C Serial Clock)　－　5.GPIO5

- ESP-WROOM-02 と人感センサ

  GPIO14　－　2.VOUT

- オーディオアンプ IC とダイナミックスピーカ

  5.OUTPUT　－　1.INPUT

次に、データシートを参考に各部品の GND と VCC の配線を完了させる。ここで、GND 接続とは電源のマイナスに接続することであり、VCC 接続とは電源のプラスに接続することである。

次は、音声合成 LSI の通信モードと動作モード選択について説明を行う。音声合成 LSI には、PMOD1、PMOD0 端子を GND または VCC に接続することで、様々な動作モードが選択可能である。本研究では、ネットワークを介して音声出力を行うので 15.PMOD1 を VCC(=1)に接続し、14.PMOD0 を GND(=0)に接続し、セーフモードを選択した（図 2-11）。セーフモードとは、UART/I2C/SPI のいずれかのシリアル通信でホストからコマンドを送ることで、任意の音声メッセージを出力することができるモードである。

PMOD1	PMOD0	動作	説明
1	1	コマンド入力モード	基本モードであり、シリアル通信で任意のメッセージを音声出力する。
1	0	セーフモード	USARTのボーレートが9600bps、I2Cのスーレブアドレスが2EHとなる。それ以外はコマンド入力モードと同じ。
0	1	スタンドアロンモード	プリセットメッセージをPC0-PC3の端子の操作で出力する。
0	0	デモモード	全てのプリセットメッセージを順に繰り返して発声

図 2-11　動作モードの選択

SMOD1、SMOD0 の端子は GND または VCC に接続することで、様々な通信モードが選択する。本研究では“I2C 通信”を行うために、5.SMOD1 を VCC に接続し、4.SMOD0 を GND に接続した（図 2-12）。I2C 通信とは、SDA と SCL 端子 2 線でホストと通信する。

SMOD1	SMOD0	通信モード
1	1	UART
1	0	I2C
0	1	SPI(MODE3)
0	0	SPI(MODE0)

図 2-12　通信モードの選択

4. 関連付けるフットプリント図を作成（既存ライブラリにないものは自作する）

シンボルにフットプリント図を対応させていくが、電気回路同様「ESP-WROOM-02 ピッチ変換キット」のフットプリントデータが既存ライブラリに掲載されていないので、変換キットのデータシートを参考にデータを自作した(図 2-13)。

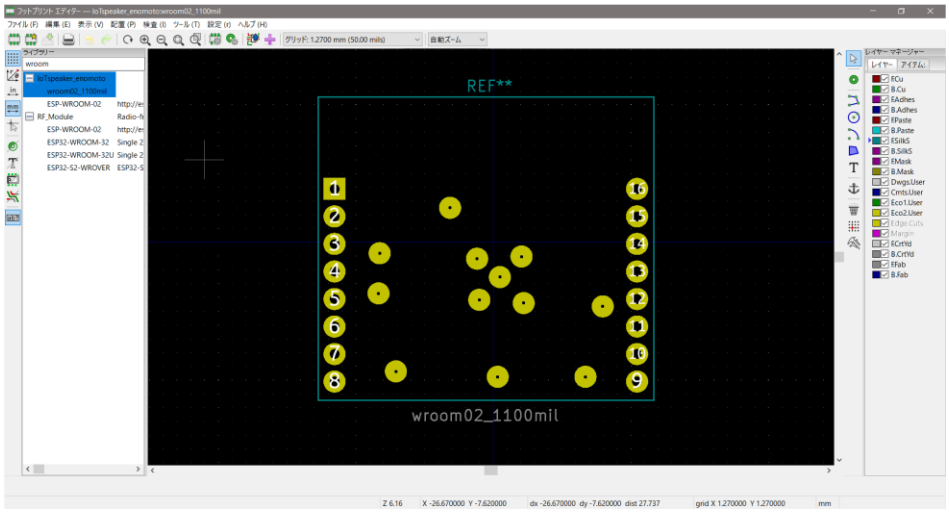


図 2-13　「ESP-WROOM-02 ピッチ変換キット」のフットプリント図

—PCB エディタの操作－

5. PCB エディタの設定と基板外形を描書と部品配置

PCB エディタでは、シンボルに対応するフットプリントの配置を決めていくが自由に配置することができる。まず決めることは、基板の外形である。本研究では、電源を単 5 電池 3 本と決めたので電池 3 本の大きさ程度の基板であれば問題ないと考えたため、縦 39.3mm×横 51.0mm の外形を描書した。外形が決まったので、次の段階に移行する。フットプリントの配置を決めていく。配置する部品のフットプリントは、ESP-WROOM-02 ピッチ変換キット、音声合成 LSI に、オーディオアンプ IC 、ダイナミックスピーカ、人感センサ、抵抗×5、コンデンサ、電源、スイッチ、である。注意することは、人感センサ付近にはスペースを空けることである。これは、人の動きを検知する際の赤外線センサを妨害しないためである。また、はんだ付けを行う際に必要なスペースを確保する。

6. 基板の完成

ガーバーデータを出力することで、製造データが完成し、発注をかけることが可能になる。ガーバービューアでこのデータを読み込み、スノーホールの位置確認や基板の外形の最終確認を行う。図 2-14 は完成したプリント基板の 3D レイヤー図である。

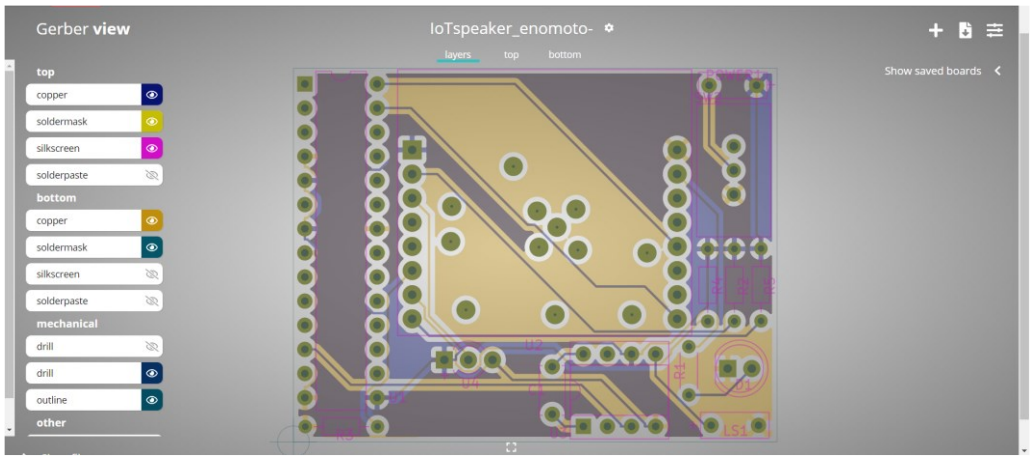


図 2-14　完成基板のレイヤー図

図 2-15 は、プリント基板製造会社に発注をかけて、製造した基板の完成外観である。できるだけ小型化するための部品を集積する配置を設計した。

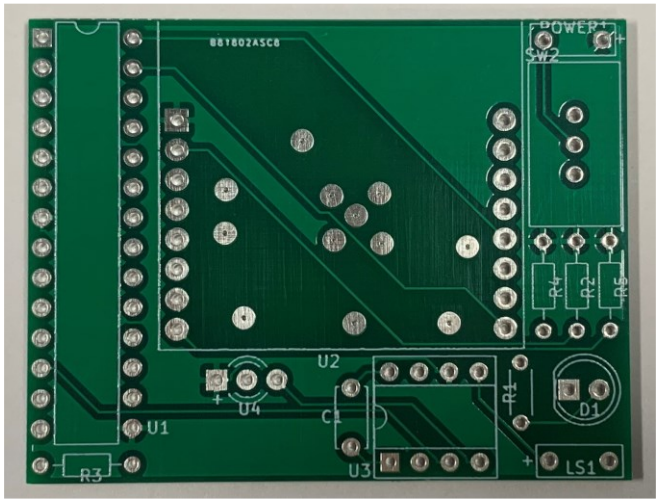


図 2-15　プリント基板の外観

2.2.5 プリント基板に部品を配置

完成した自作のプリント基板(52.7×41ｍｍ)上に WiFi 機能を有する ESP-WROOM-02 と音声合成 LSI に、オーディオアンプ IC 、ダイナミックスピーカ、人感センサ、スイッチ、抵抗、LED、コンデンサを配置し、はんだ付けすることで IoT スピーカが完成した（図 2-16)。電源は単 5 乾電池を使用した。

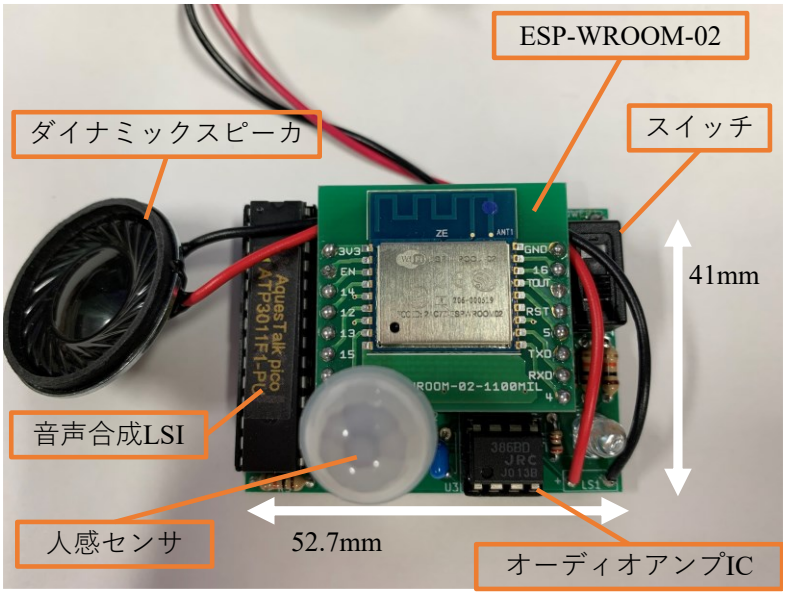


図 2-16　IoT スピーカの外観

2.2.6 Arduino IDE を用いたプログラム書き込み

IoT スピーカにプログラムを書き込む、つまり ESP-WROOM-02 単体にプログラムを書き込むためには、別途専用回路を必要とする。この回路に ESP-WROOM-02 単体を接続することで Arduino IDE を通してプログラムを書き込むことが可能となる。その回路図を図 2-17 に示す。ESP-WROOM-02 は単体ではパソコン等によるプログラムを書き込むことができない。そこで、専用回路を作成しプログラム書き込み可能な状態にするが、この回路には、AE-TTL-232R という USB シリアル変換モジュール（図 2-18）を使用しパソコンと ESP-WROOM-02 の懸け橋となる。また、ESP-WROOM-02 の②⑥⑦⑧はプルアップ抵抗を挟んで＋と－に接続され、プルアップ、プルダウンされている。プルアップ抵抗とは電子回路においてなんらかの信号線への入力が切り離されたような状態になっても、その信号線を内部で 5V などの電位が高い電圧源にその抵抗を通して接続しておくことにより、電位が不安定な状態にさせず、充分に高い電位に保っておくための抵抗である。

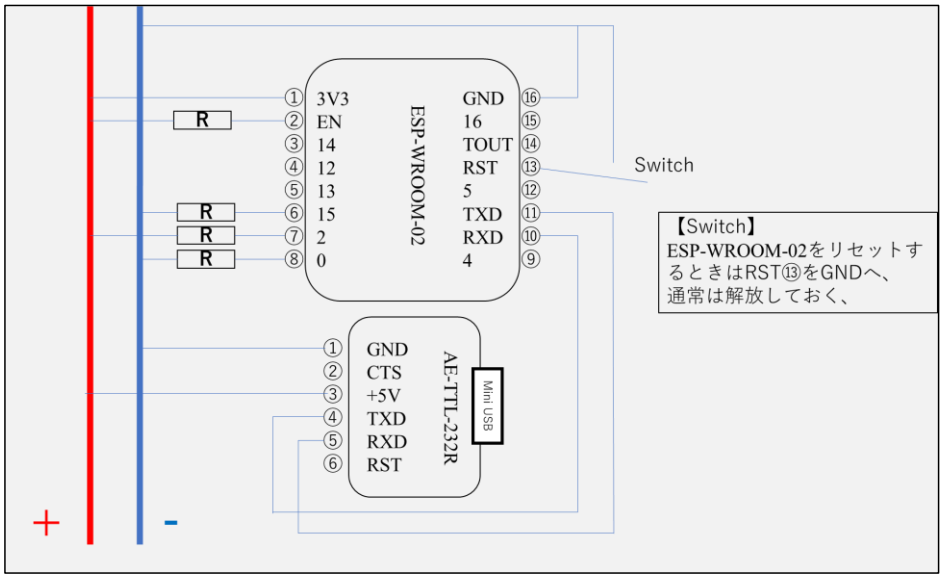


図 2-17　ESP-WROOM-02 単体にプログラムを書き込む回路図

注意すべきポイントは、ESP-WROOM-02 と AE-TTL-232R の「TXD」「RXD」がクロスして接続されている点である。AE-TTL-232R に PC から Mini USB で接続を行い、Arduino IDE を通してプログラムを書き込む。

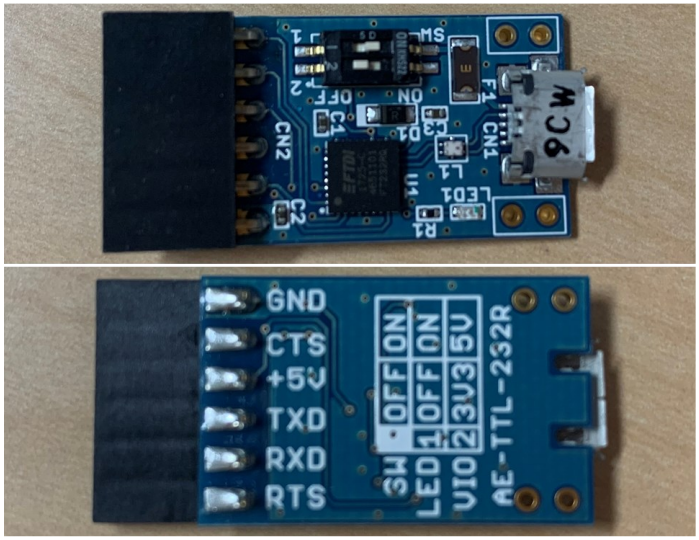


図 2-18　USB シリアル変換モジュール（AE-TTL-232R）

IoT スピーカのソフトウエアは、Arduino IDE で開発を行い、ESP-WROOM-02 に書き込みを行うことができるが、本研究で作成したプログラムは大きく分けてセットアッププログラムと、ループプログラムの 2 つがある（図 2-19）。

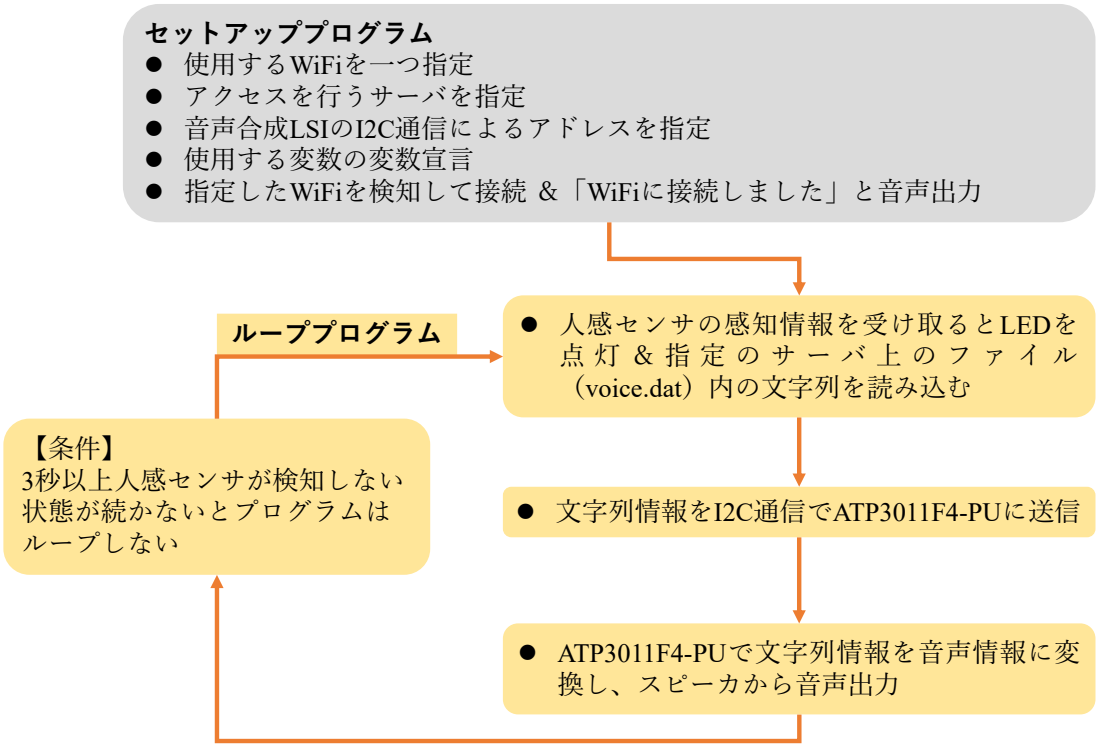


図 2-19　プログラムのフローチャート図

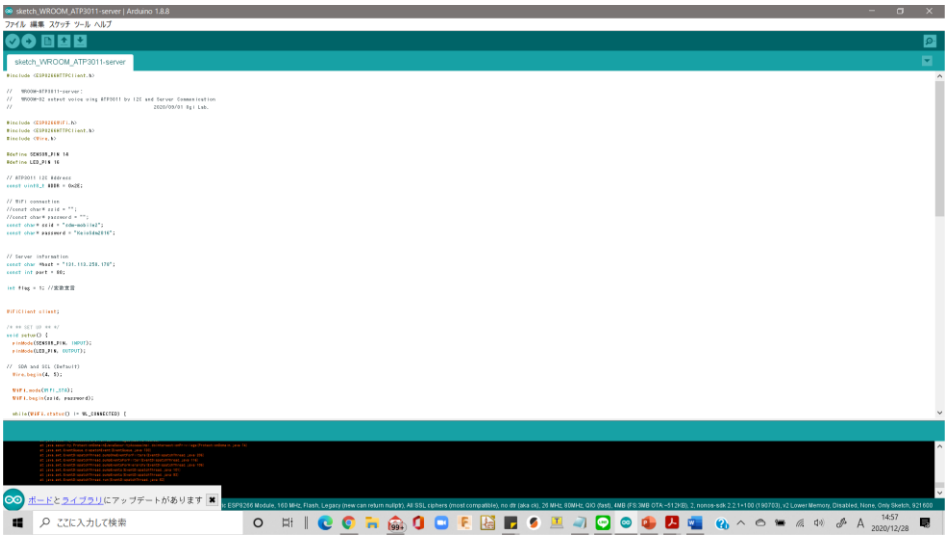


図 2-20　Arduino IDE による IoT スピーカのプログラム（付録付き）

図 2-20 は Arduino IDE 上でのプログラム図である。セットアッププログラムとループプログラムをそれぞれ書き込んでいくが、全角入力はエラーのもとであるため半角で文字を打ち込んでいく。

2.2.7 音声記号列の作り方

サーバ上の指定したファイル内の文字列はローマ字入力である。それらを変換することで音声出力内容を変更することが可能であり、Tera Term を通してファイルの文字列を書き換えることで実現できる。Tera Term とは、リモートログオンクライアントである。現在では TeraTerm Project によりオープンソースとして開発されている。 SSH・telnet・シリアルの各通信プロトコルに対応し、Microsoft Windows で使用できる。

図 2-21 は Tera Term でサーバ上のファイルにアクセスして、ファイル内の文字列を書き換えようとしている図である。サーバ上（サーバ名：mlp）にある voice1.dat ファイル（~/public_html/voice.dat）を「vi」で編集する。各コマンドの意味を述べる。

ls：ファイルを一覧表示する他、ファイルの詳細情報を表示するコマンド

cd：ディレクトリを移動するコマンド

vi：エディタを起動するコマンド

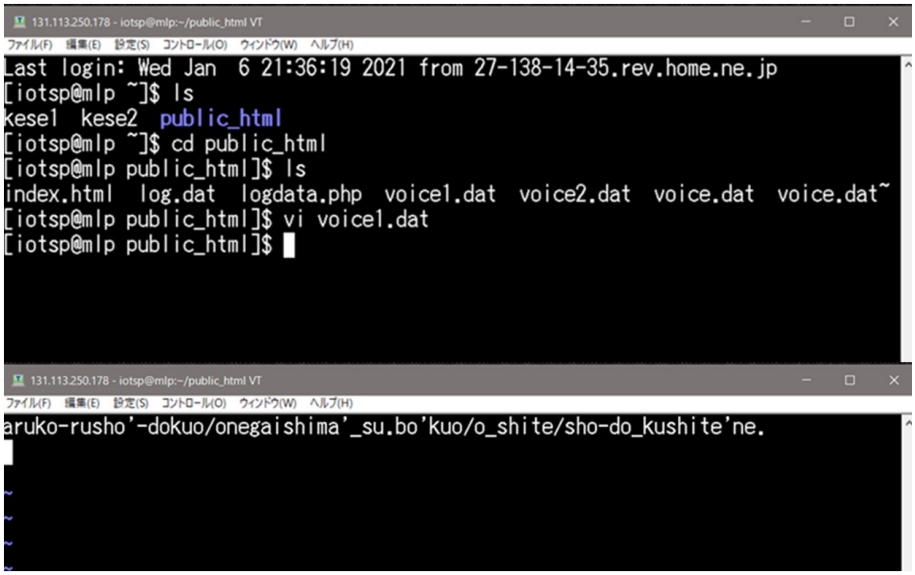


図 2-21　IoT スピーカの音声出力の書き換え方法

音声記号列とは、生成する音声を文字コードで表現したものである。音声記号列は、読み記号、アクセント記号、区切り記号、タグ記号を並べて構成される。読み記号とは、読みをローマ字で指定し、一つ一つが日本語の音節にほぼ該当する。アクセント記号は、アクセントの位置を「'」で指定し、アクセント核の読み記号の直後に指定する。アクセント記号の数は、1 つのアクセント句に 1 つ、あるいはなしとする。区切り記号は、アクセントの境界を示し、様々な種類がある。「.」：文の終わりに指定、「?」：文の終わりに指定する（疑問形）、「空白」：呼気段落の境界に指定（ポーズ有）、「,」：呼気段落の境界に指定（ポーズ有、空白より短いポーズ）、「;」：声立てを行う場合に指定（ポーズ無）、「/」：通常のアクセント句の句切り（ポーズ無）、「+」：後ろのアクセント句が副次アクセントの場合に指定（ポーズ無）と多様な区切り記号がある。タグ記号は、棒読み、桁読み、英数読みの 3 つの定義があり、棒読み <NUM VAL=(数値)>、桁読み <NUMK VAL=(数値)[COUNTER=(助数詞)]>、英数読み <ALPHA VAL=(英数)> または、<ALPHA VAL="(英数)">である。

音声記号列の作り方の例を以下に示す。単なるローマ字による読み記号でも音声記号列は作製可能であるが、ぎこちない日本語が出力される。そこで、「/」や「’」などのアクセント記号、区切り記号、タグ記号を用いることで滑らかな発声が可能となる。

例１：「音声記号列の作り方を紹介します！」
onnse-kigo-'retsuno/_tsukurikata'o/sho-kaishima'_su

例２：「僕の名前は、榎本敬太。」
bo'kuno/namaewa eno'moto/ke'-ta.

例３：「手と指の消毒をお願いします。」
te'to/yubino/sho-dokuo/onegaishima'_su.

例４：「僕はアルコール！僕を押して消毒してね！」
bo'kuwa/aruko-ru bo'kuo/o_shite/sho-do_kushite'ne

例５：「協力ありがとう！」
kyo-ryoku/ari'gato-

第3章
提案システムの基礎実験

本章では、前章で構築した IoT スピーカの性能検証について述べる。人感センサの性能検証について述べ、スピーカの音量について述べ、インターネット接続スピードについて述べ、最後にスピーカの向きにおける音の聞きやすさ検証について述べる。

3.1　人感センサの性能検証

3.1.1 検証方法

人感センサの反応距離やどのようにセンサが反応するのかを計測した。

使用した人感センサ（焦電型赤外線センサーモジュール SB412A）は、仕様上は検知距離 3～5m、検知角度 100 度となっているが、IoT スピーカに実装した状態での反応精度を計測した。計測方法としては、人感センサの正面方向を 0 度とし 10 度ごとの角度に対して、10cm 刻みで人間が近づきながら手を振る動作を行い、反応距離を 10 回ずつ計測した。環境や温度によって人感センサの感度は変化するため、計測は室温 26℃の建物内において行った（図 3-1）。

図 3-1　人感センサの反応距離検証の様子

3.1.2 結果と考察

計測結果の角度と反応距離の関係を図 3-2 に示す。

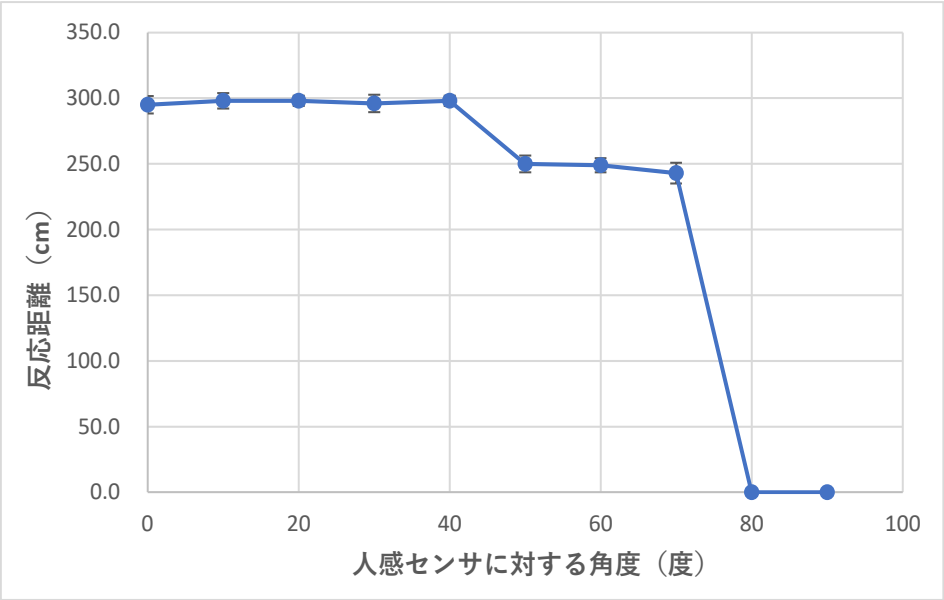


図 3-2　人感センサの方向による反応距離

計測結果は、センサ正面では反応距離が平均 300cm であったのに対し、角度が大きくなるに伴い 40 度付近で反応距離が短くなり、70 度付近で感知不能となることが分かった。10 回ずつ計測する中で、反応するときと反応しない時に分かれた。IoT スピーカのプログラムでは 1ms 毎に人感センサの感知情報を取得しているが、このセンサは人の動きを検知するため、人が近づく場合には検知距離はもう少し短くなることが考えられる。

人感センサは、赤外線などを利用して周囲温度と温度差のあるものが検知範囲内で動いたときに、その温度変化を検知する仕組みになっているため、人感センサから見た縦方向の動きよりも、横方向の検知の方が敏感に反応したと考察する。

この結果は、おおよそ予想通りであり期待した結果を得られた。

3.2　スピーカの音量の検証と結果

3.2.1 検証方法

小形デジタル騒音計（サンコー）を用いて、IoT スピーカの出力音の大きさが距離に対してどのように変化するのかを計測した。計測方法としては、「こんにちは」の音声を出力した際の音の大きさの平均値を、計測距離を 50cm 刻みでそれぞれ 10 回ずつ測定を行った。

3.2.2 結果と考察

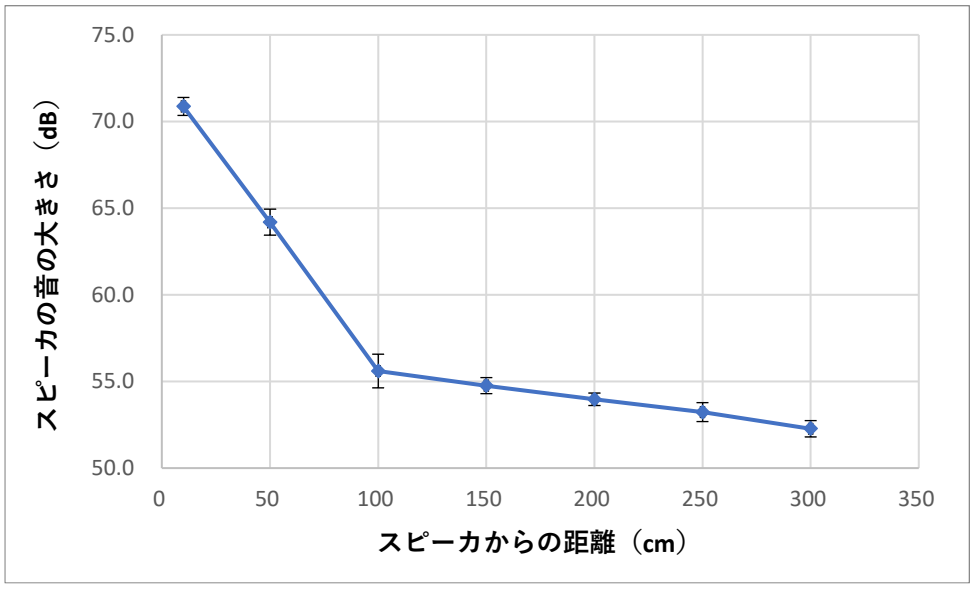


図 3-3　スピーカからの距離に応じた音の大きさ

計測結果から、100 ㎝付近以内と外では音の減衰に違いがあることが分かった。音の大きさの目安としては、50dB が小さい声で人が話す大きさと同等であり、60dB が普通の声で話すのと同等、70dB が大きい声で話すのと同等である。よって人がセンシングされる 300cm 以内では、スピーカから発せられる音はユーザに十分届く大きさと考えられる。スピーカからの距離と音の大きさの関係を図 3-3 に示す。

3.3　インターネット接続速度の検証と結果

3.3.1 検証方法

IoT スピーカが事前に音声データを読み込んでいるプリセット状態において、人感センサが人の動きを感知してから音声データをスピーカから提示するまでの遅延時間の計測を行った。具体的には、iPhone のカメラで撮影した 60fps の動画を動画音声解析ソフトによってタイムラグを測定した。IoT スピーカでは、人感センサが感知したことを可視化するために LED を点灯させているため、人感センサの感知タイミングは LED の点灯開始時間、音声提示は音声の出力開始タイミングとして、10 回計測を行った。

次に、IoT スピーカが毎回サーバにアクセスし、発音データを読み込みながら提示する形態において、人感センサが感知してからサーバを経由してスピーカが発音するまでの時間計測を行った。計測方法としてはプリセット音声と同様に、iPhone のカメラで撮影した 60fps の動画を動画音声解析ソフトによってタイムラグを測定した。10 回計測を行った

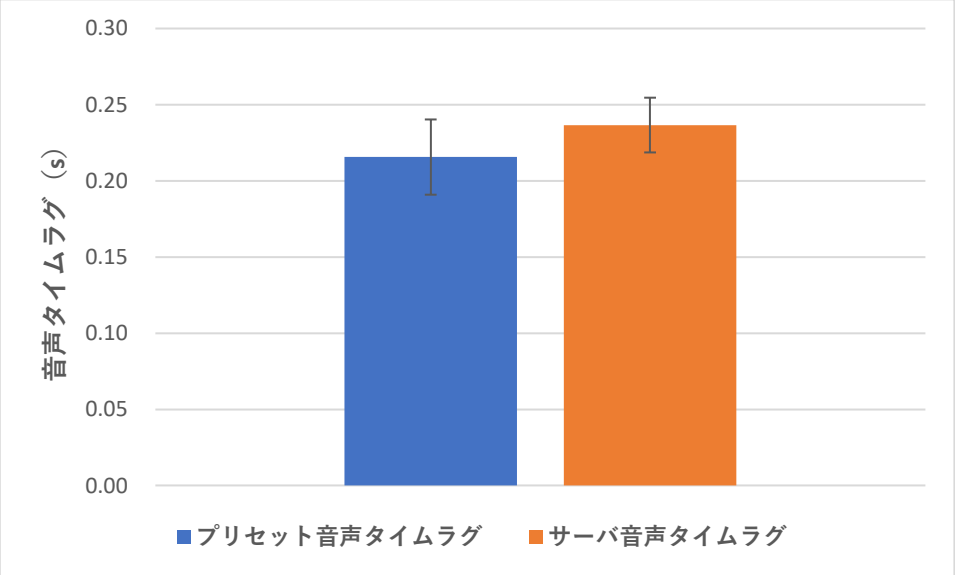


図 3-4　音声出力のタイムラグ

3.3.2 結果と考察

IoT スピーカが事前に音声データを読み込んでいるプリセット状態のタイムラグの平均は 0.22 秒（標準偏差 0.03）であった。また、IoT スピーカが毎回サーバにアクセスし、発音データを読み込みながら提示する形態のタイムラグの平均値は 0.24 秒（標準偏差 0.02）であった。結果を図 3-4 に示す。これらタイムラグを比較することで、サーバアクセスにおおよそ 0.02 秒要することが読み取れる。このサーバアクセス時間はネットワーク環境によって影響を受けるため、プリセット音声を使用するかサーバ音声を使用するかは、使用環境と目的によって使い分けることが必要である。

当初、プロトタイプでのネットワーク接続はかなりの時間を要した。4~5 秒程度時間を要したので、本結果の 0.30 秒以下のタイムラグは予想以上に高速であった。ネットワーク環境によって時間遅れが発生していて、WiFiClient のライブラリが安定していない為であったので、HTTPClient に書き換えた。つまりサーバアクセスの部分を HTTPClient のライブラリで書き直すことで、本結果の通りほとんど時間遅れ無しでサーバのデータを発話できるようになった。

3.4 スピーカの向きによる音の聞きやすさ検証と結果

3.4.1 検証方法

成人男性の歩く速さは、一般に約 4km/h（≒111cm/s）であり 300cm 歩くのに約 2.7 秒要する。そのため、歩行中のユーザに IoT スピーカが音声情報提示を行う場合、ユーザがどの程度スピーカの音を聞き取ることができるか、ユーザの歩行速さに対する聞き取りやすさについて検証実験を行った。

図 3-5　人感センサの向きに応じた聞こえやすさ検証実験

人感センサの向き縦方向（青矢印）と横方向（橙色矢印）の場合において、近くを歩く人は、どちらの方がスピーカから出力される音声を聞きやすいかを検証した実験である（図 3-5）。音声出力内容は「山手線はこちらです。右手に男子トイレ。」である。

まず、スピーカに向かって歩いていく場合を想定した。モノに設置された IoT スピーカに向かって歩いていくユースケースである。実験では 3 人の被験者に歩くスピードを、早歩き・普通の速度・ゆっくり、の 3 段階の歩き方で 5 回ずつ歩行してもらい、それぞれどのタイミングでスピーカから音が発生されたかを音声発声時の被験者の位置をビデオ映像から計測し、その際の音の聞こえ方を

5 段階評価で回答してもらった。それぞれの歩き方をした時の歩行速度は、早歩き：5~6km/h、普通の速度：4km/h 程度、ゆっくり：2~3km/h であった。買い物などで何かを探しながら歩くような場合には、歩くスピードは更に遅くなることが考えられる。音の聞こえ方は、5：非常によく聞きとれる、4：よく聞きとれる、3：聞きとれる、2：あまりよく聞きとれない、1：よく聞きとれない、の評価項目から選択してもらった。

次に、壁などに取り付けられた IoT スピーカの横 1m の距離を通り過ぎる場合を想定した実験を行った。この場合、人感センサの正面は被験者に対して 90 度の方向を向いているが、IoT スピーカの横を通り過ぎるまでにスピーカから音が発声されなければ、ユーザは対象物に関心を示さない可能性がある。上記と同じ条件で 3 人の被験者に各速度で 5 回ずつ歩いてもらい計測を行った。

3.4.2 結果と考察

歩く速さと発声距離の結果を図 3-6、および歩く速さと音の聞こえ方に関する評価の結果を図 3-7 に示す。

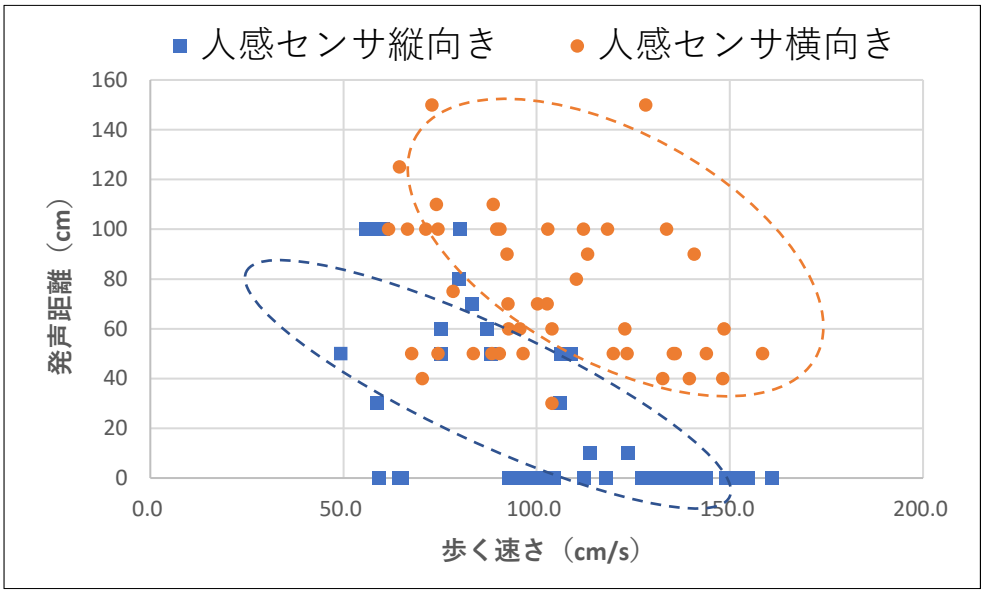


図 3-6　歩く速さ毎の発声距離の長さ

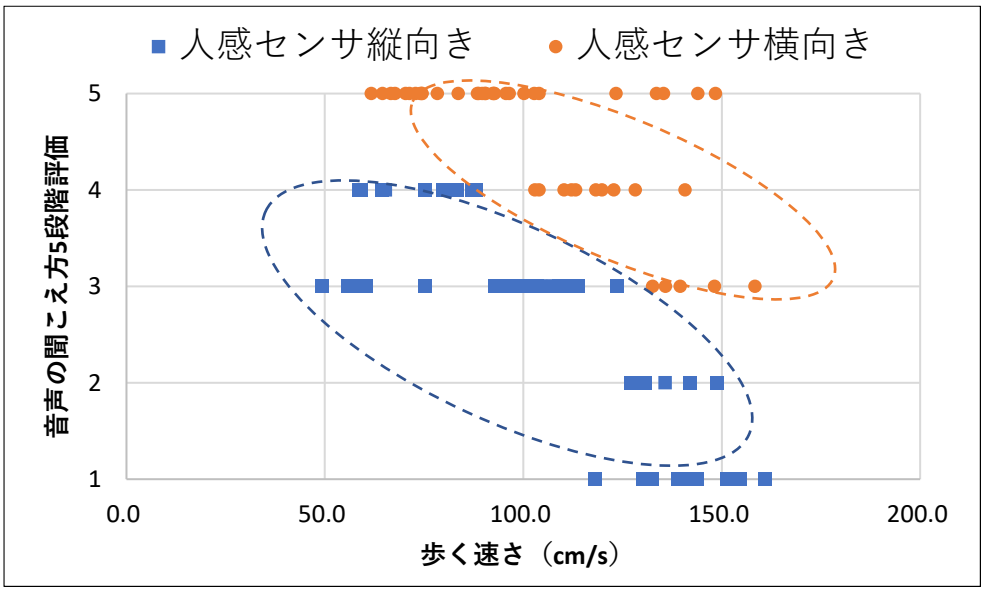


図 3-7　歩く速さに応じた音の聞こえやすさ 5 段階評価

上記の実験から、ユーザが立ち止まっている状況では 300cm 程度が人感センサの検出範囲であったが、ユーザが歩行している場合は、動きの感知やセンシングにおけるタイムラグのため、検出範囲が縮小されることが分かった。これは、人感センサは、人体検出のために常時レンズから光を出力しているのではなく、人体から放射される赤外線を受けることによって検出を行うため、正面方向に近づくよりも、横方向に動く物体の方が検知しやすいためだと考えた（図 3-6）。また検知が早くなる分、音声出力に使える時間が確保でき、結果聞き取りやすいという結果になったと考える（図 3-7）。

また人感センサの方向を変えた 2 つの実験結果からは、IoT スピーカの正面方向に向かってユーザが歩行するよりも 90 度方向から横切る方が、人感センサの感知が早く音声提示のタイミングが早まり、音声情報を伝達しやすいことが確認された。このことから IoT スピーカの実装において、歩行者を対象とした情報提示では人感センサを横切る方向に設置すべきであることがわかる。

3.5 モノが喋っている感覚の要因検証

モノ（消毒液）と音源（スピーカ）を対象に、IoT スピーカを取り付けたモノが音声出力を行い、喋っていると感じさせるか検証を行った。

① モノと音源の距離・位置関係

② 音源の発声内容

の２点が要因としてそれぞれ検証を行った。

聴覚による音源の力向定位は、視覚的な手がかりによって影響を受けることがよく知られている。日常生活の中で私たちは、聴覚によって知覚する環境と視覚によって知覚する環境が一致しているように思っている。例えば、何もないところから音が聞こえてくれば非常に不思議に思うし、腹話術のように視覚的に音源と考えられるものが存在すると、そこから音が聞こえてくるように錯覚する。視覚情報と聴覚情報は統合されて音源定位の知覚が生じているのである。このことは、視覚が音源方向の知覚に大きく影響を与えるというように言うこともできる。音源定位に関する視覚と聴覚の相互作用については、様々な研究がなされている[24]。

視覚が音源方向の知覚に影響を与える「腹話術効果」について、Klemm の実験では評定者の前に電話機を置き、隠した別の電話機で音を鳴らした。その結果、音と電話機が分離したのは 12°以上離したときという[25]。つまり、見えている電話機の位置に音源が知覚されるのである。人間は無意識のうちに、音を発するものとそうでないものを学習し、区別している。この場合、電話機はベル音を発するものと学習しているため、電話機が音源の視覚情報となっていると考えられる。すなわち、その音を発していると考えられるものが視覚情報となり、視覚が聴覚に影響を及ぼしていると言える。

中林らはテレビ映像を用いた、音像と映像の相互作用に関する基礎実験を報告している[26]。この報告では、一つのモニタに提示した人物像を視覚情報とし、提示された音像と映像の閾隔が 10°以内であると、音像は映像に方向支配され、10°～20°の間では、10°以内のときよりも音像は映像に方向支配されにくく、20°以上であると音像と映像が分離して知覚されている。このことから、 音像は映

像に影響されていることが実験的に確認されている。

また藤井らは、音源定位課題における検討として、視覚情報と聴覚情報の対応関係について報告している[27]。ここでは、音声と話者像の発話動作の対応関係について検討し、視聴覚情報間の対応がとれる対象のみ、つまり発話動作と音声が一致するものに対して定位の誘引効果が存在することを示唆している。

3.5.1 検証方法

① 消毒液とスピーカがどの程度の距離離れている場合に、モノからの音声出力と感じなくなってしまうのかを、10 人を対象に実際の距離を計測した。計測方法として、被験者は、消毒液とスピーカセットの対象物から 150cm 離れた場所にて、消毒液とスピーカの距離を 0cm から 5cm ずつ距離を増やし、音源と対象物が分離する感覚が生じた距離 L と角度 θ を計測する（図 3-8）。
$L = R\theta$の関係から、R=150cm なので、距離 L を求め更に角度 θ を求めることができる。

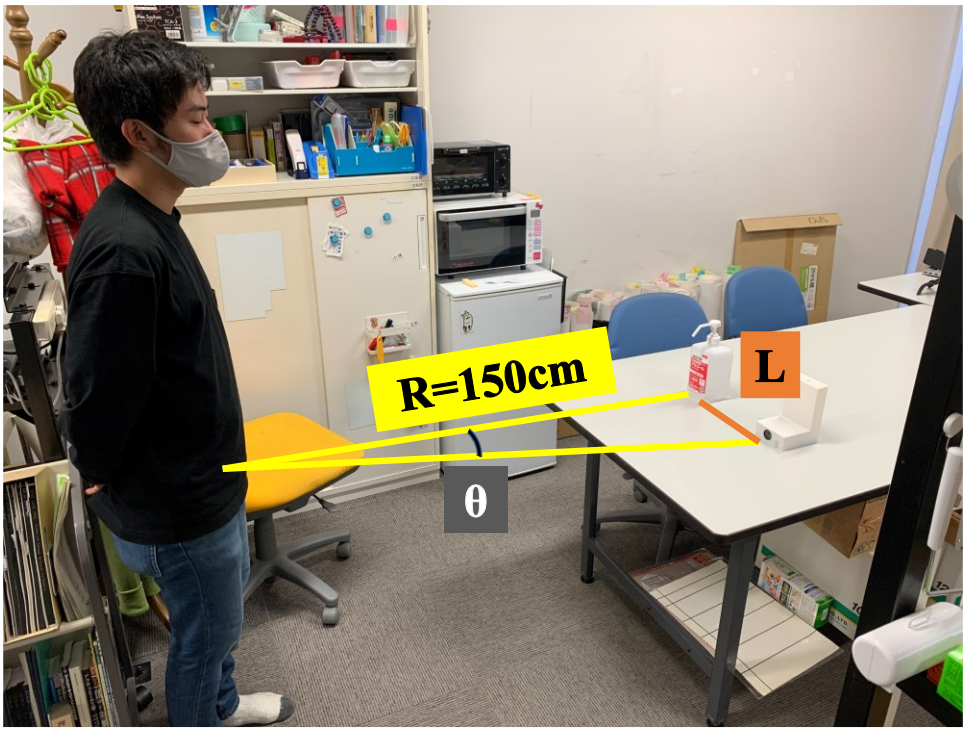


図 3-8　実験の様子

② 発声内容をポスターや公共アナウンスのような文語で音声出力した場合と、会話に用いられる口語で音声出力した場合の比較検証を、10 人を対象に行いヒアリング調査を行った。文語出力は、“手指の消毒にご協力をお願いします”であり、口語出力は“僕はアルコール！僕を押して消毒してね”である。

ヒアリングに関するアンケート

以下の質問に答えて下さい。

Q１．口語による音声出力で、消毒液が喋っている感覚を感じましたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

Q２．文語による音声出力で、消毒液が喋っている感覚を感じましたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

ご回答ありがとうございました。

慶應義塾大学

図 3-9　アンケート調査表

図 3-9 のヒアリング項目は、Q1：口語による音声出力で消毒液が喋っている感覚を感じたか？Q2：文語による音声出力で消毒液が喋っている感覚を感じたか？である。また、文語の音声出力は、「手指のアルコール消毒にご協力お願いします」である。また、口語の音声出力は「僕はアルコール！僕を押して消毒してね！」である。

3.5.2 結果と考察

① 検証結果図 3-10 から、平均距離は、14.5cm（標準偏差 4.7）であった。また結果図 3-11 から、平均角度は 17.4°（標準偏差 5.7）であった。まず平均距離に関してだが、この結果から消毒液とスピーカが多少離れていてもスピーカからの音声出力は、消毒液に注目することで、消毒液がメッセージを伝えていると感じやすい結果となった。次に平均角度に関しては、音と対象物が分離した感覚が生じたのは、12°以上離したときという先行研究の結果からこれは先行研究とやや近い結果である。これら結果から、消毒液の使用に

関する音声情報は、視覚で消毒液をとらえているためにこれらが結び付けられると考察する。よって、多少スピーカが離れていても消毒液とスピーカに一体感を感じるという結果を得た。

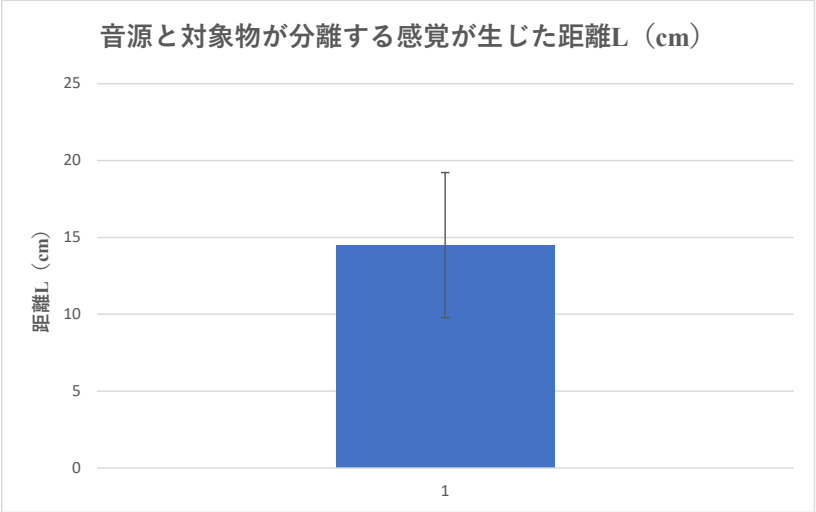


図 3-10　音源と対象物が分離する感覚が生じた平均距離

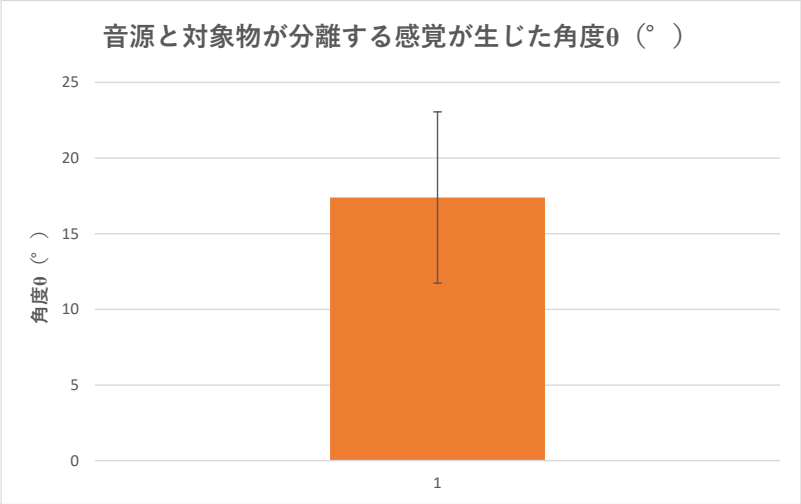


図 3-11　音源と対象物が分離する感覚が生じた平均角度

② ヒアリング回答選択肢に 1:全くそう思わない、2:そう思わない、3:どちらともいえない、4:そう思う、5:非常にそう思うという番号を振り当て回答を数値化した結果を、図 3-12 に示す。Q1：口語で消毒液が喋っている感覚を感じましたか？の平均スコアは、4.9（標準偏差 0.3）であり、Q2：文語で消毒液が喋っている感覚を感じましたか？の平均スコアは 2.8（標準偏差 0.7）であった。文語については、ポスターに掲載されている文章を音声出力することで表現した。よって単なる音声出力に感じ、喋りかけてくるようには感じない結果となった。また口語については、より人間が喋りかける際の言葉を選択した。これによって、音声出力の言葉の背景に人が話しかけてくる感覚を感じた人が多かった。よってこの結果から、出力音声によって、モノが喋りかける感覚を表現できることが分かった。

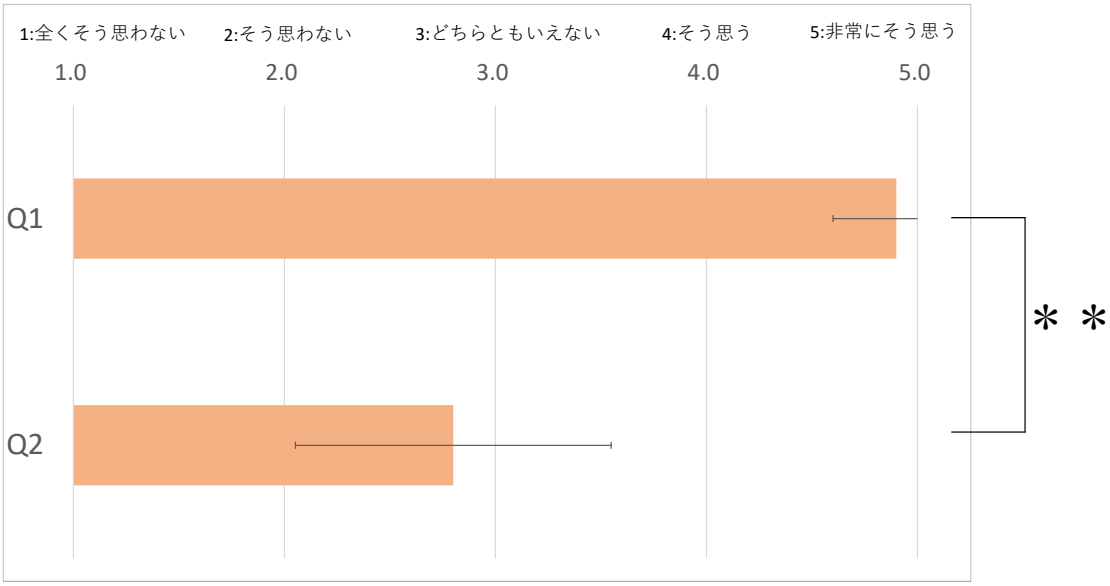


図 3-12　アンケート調査の結果

第4章
提案システムのシステム検証と評価

本章では、前々章で構築した IoT スピーカのシステム検証について述べる。モノに取り付けた IoT スピーカは音声出力をすることで、人間の行動支援を行うことができるかどうかの検証についてと、結果と考察を述べる。

4.1 妥当性検証の実験準備

妥当性確認の本実験では、取り付ける“モノ”は消毒液を選択した。理由としては、身近にあるものかつ、今年流行した新型ウイルスの影響で多くの場所に消毒液が設置された。一定間隔で置かれた消毒液は多くに人が身近に感じるモノとなったためだ。そこで消毒液に IoT スピーカを装着するために消毒液専用カバーを 3D プリンターで作製する。3D データ作成のために Autodesk Fusion360CAD ツールを用いて消毒液ケースの 3D データを作成した。

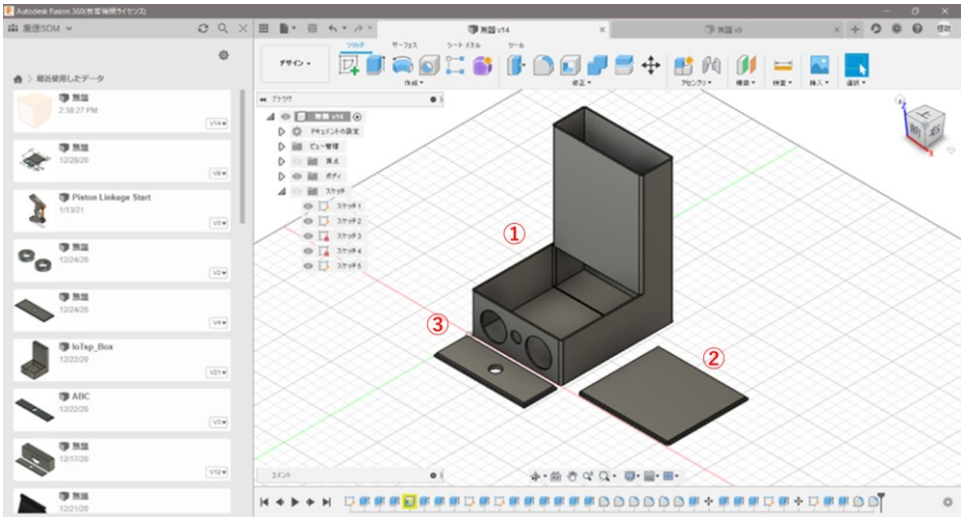


図 4-1　IoT スピーカのケース用の 3DCAD データ

図 4-1 は、IoT スピーカの消毒液用ケースの 3DCAD データであり、消毒液の大きさに合わせケースのサイジングを行った。ここで、消毒液に近づく人と、消毒液を使用してくれた人に対して、異なるメッセージを発信することを考えた。近づく人には、消毒液使用を催促するメッセージを、使用してくれた人に対しては、ありがとうのメッセージを出力する。これによって、よりモノが喋っているように感じさせることが可能と考えた。①は消毒液ケースの本体であり、②は消毒液を乗せる部分の取り付け部品である。③は①本体の最上部分に取り付ける

ケースであるが、人感センサを取り付ける穴を空けている。

ここで、CAD データ作成における注意点を述べる。部品の高さの基準面を揃えることである。3 つの部品が互いに違った面に存在していると、3D プリンターで作成する際に空中の途中からオブジェクトを作成してしまい、失敗の原因となる。

次に、作成したデータを STL ファイルに変換する。STL とは Standard Triangulated Language の略であり、三次元形状を表現するデータを保存するファイルフォーマットのひとつである。

「Prusa Slicer」で STL ファイルを G-code ファイルに変換する。スライサーソフトは、STL などの 3D データを 3D プリンター用のデータに変換できるソフトウエアである。3D プリンターにデータを認識させるためには、G-code という形式に変換する必要がある。これは、図 4-2 のようにオブジェクトをレイヤー状にスライスする必要がある。断面ごとのデータにすることで 3D プリンターはオブジェクト作成を行うためだ。

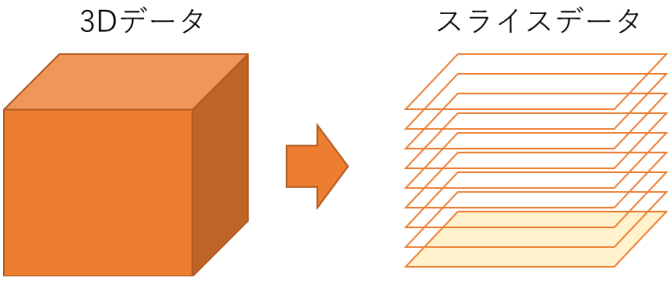


図 4-2　スライスデータの概念

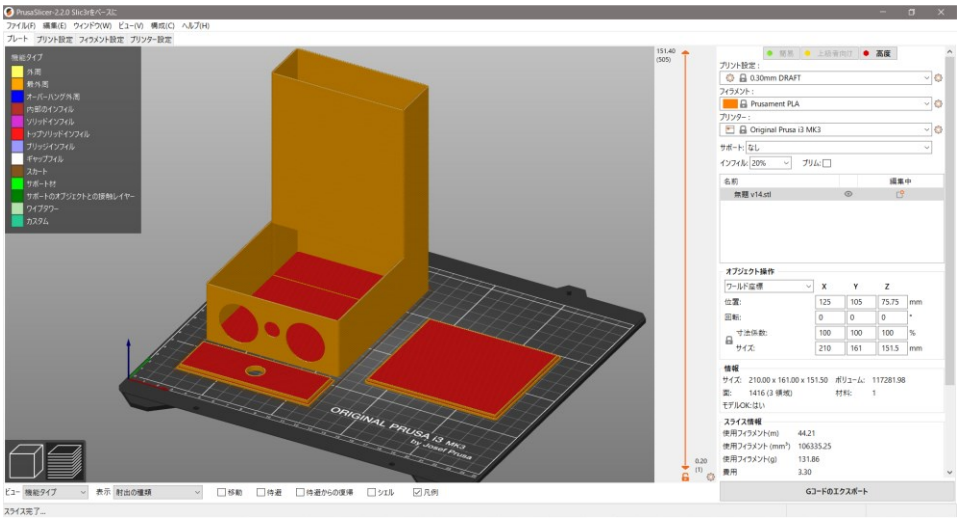


図 4-3　Prusa Slicer における消毒液ケースのスライスデータ

図 4-3 は Autodesk Fusion360CAD ツールを用いて作成した STL ファイルを G-code ファイルに変換したデータである。この Prusa Slicer では設定をカスタマイズでき、精巧に創り上げるため、フィラメントの細さまたサポートを完成物に合わせて設定可能である。G-code ファイルに変換できたら、3D プリンターで実際に製造に取り掛かる。使用した 3D プリンターは、Original Prusa i3 MK3S 3D プリンターである（図 4-4）。

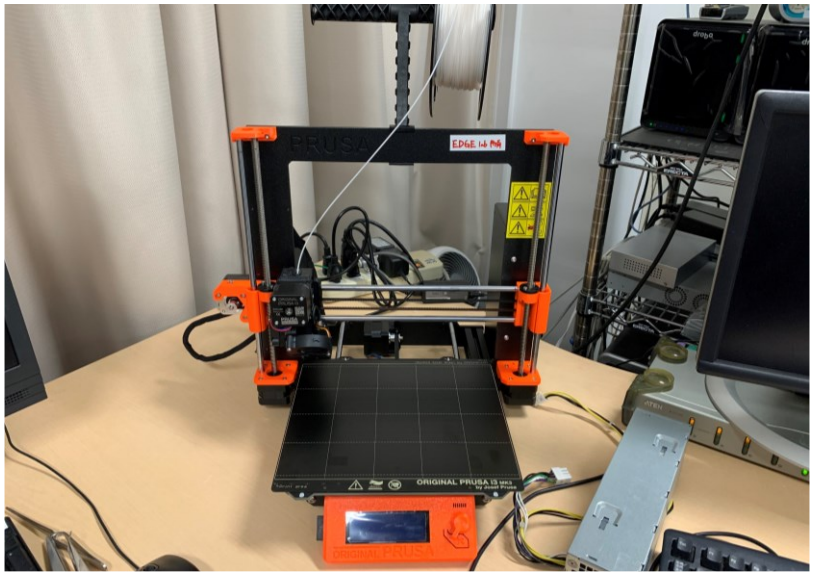


図 4-4　Original Prusa i3 MK3S 3D プリンター

使い方はシンプルで、専用の SD カードに G-code ファイルを保存して本体に差し込むだけである。SD カードを差し込むと、プリンター本体の画面にファイルデータ一覧が表示されるので選択するだけで、自動で完成まで仕上げてくれる。ここで、失敗のための原因と対策を述べる。3D プリンターは使いやすく、便利である一方で初期設定を怠ると失敗が続いてしまう。これは、ノズル（フィラメントの出口）と土台の間の距離が重要なポイントであるためだ。理想的なノズルと土台の Z 方向の距離は、0.1mm と言われており、これは普通用紙一枚分の厚さである。初期設定のキャリブレーションで Z 方向の高さ調節をすることで失敗を防ぐことができる。その際は、紙一枚をかませることで理想的な高さに調節可能だ。

図 4-5 は実際に完成した消毒液ケースである。この図ではケースの中に IoT スピーカ 2 つを既に設置している状態である。

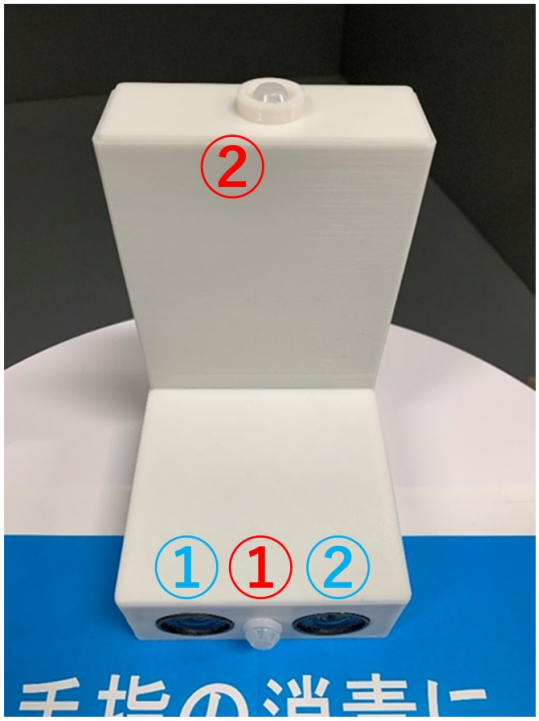


図 4-5　消毒液専用ケース

このケースは、IoT スピーカは 2 つ組み込まれる設計にしてある。赤の数字は人感センサであり、青の数字はスピーカの数字である。IoT スピーカは計 2 つあるので、それぞれが対応したセットとなっている。人感センサ①が外から近づく人に対して検知を行い、スピーカ①から音声出力し、人感センサ②が消毒液を使用してくれた人の手を検知し、スピーカ②から音声出力するものである。つまり、それら人感センサに対応したスピーカから音声が出力される仕組みである。

4.2　音声出力によるモノの擬人化検証と結果

IoT スピーカで音声出力が可能となった消毒液が、出力する音声によって、モノが喋っているように感じるかどうかの検証を行った。

4.2.1 検証方法

検証方法は、観察により特徴的な行動を取ったユーザに対して、本人の了解を得た上で、使用性についてのアンケートまたはインタビューを 21 人に行った(図 4-6)。インタビューは録音等の記録は行わず、実験者がメモをとる形で実施した。項目は、全くそう思わない、そう思わない、どちらともいえない、そう思う、非常にそう思う、からそれぞれ一つを選んでもらいチェックを付けてもらう。基本的に個人情報は扱わないが、一般人に対して許可を取らずに観察を行うため、観察者は腕章をつけることで、周囲の人に不信感を与えないように注意した。音声出力内容は近づく人に対して口語出力の、「僕はアルコール！僕を押して消毒してね！」と、使用してくれた人に対して「協力ありがとう！」である。

IoT スピーカシステムの検証に関するアンケート

以下の質問に答えて下さい。

Q1．消毒液が話しかけるタイミングは適切でしたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

Q2．消毒液の音声内容に自然さは感じましたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

Q3．消毒液の音声内容を理解できましたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

Q4．消毒液が喋っているように感じたか？

□全くそう思わない　□そう思わない　□どちらとも言えない　□そう思う　□非常にそう思う

Q5．あなたの属性について教えて下さい。

性別　□男性　□女性

年齢　□10 代　□20 代　□30 代　□40 代　□50 代　□60 代以上

ご回答ありがとうございました。

慶應義塾大学

図 4-6　アンケート調査表

4.2.2 結果と考察

インタビューの回答選択肢に 1:全くそう思わない、2:そう思わない、3:どちらともいえない、4:そう思う、5:非常にそう思うという番号を振り当て、リッカート尺度により回答を数値化した結果を、図 4-7 に示す。

「Q1:消毒液が話しかけるタイミングは適切でしたか？」に対する回答平均は、3.6（標準偏差 0.9）であり、「Q2:消毒液の音声内容に自然さは感じましたか？」の平均回答は、3.9（標準偏差 0.8）、「Q3:消毒液の音声内容を理解できましたか？」の回答平均は、4.3（標準偏差 0.8）、「Q4:消毒液が喋っているように感じたか？」の回答平均は、3.6（標準偏差 0.8）であった。

この選択肢の中で、一番平均が高かったのは、消毒液の話す音声内容を理解することができたか？の回答であり、多くの人が音声内容を理解してくれた結果である。次に、音声内容に自然さを感じたか？の回答であった。消毒液が話しかけるタイミングと、消毒液が喋っているように感じるかどうかの、回答はほぼ同じであったが、3 を上回ったことで、過半数以上の人が、消毒液が喋っているように感じる結果であった。

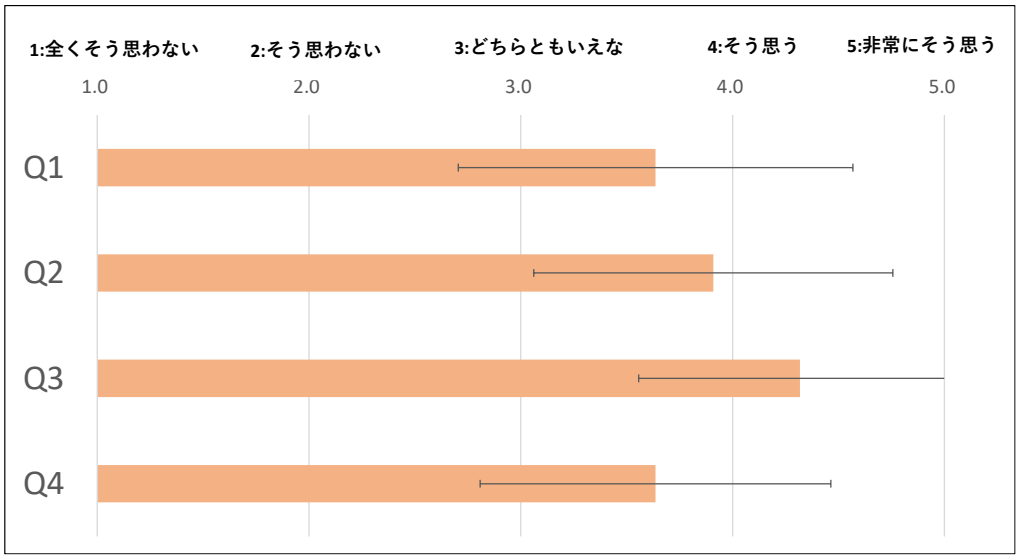


図 4-7　アンケート調査の結果

検証の結果に対して考察を述べる。

話しかけるタイミングに関しては、ゆっくり歩いている人が特に音声に反応している傾向があった。消毒液の近くをゆっくり歩いている人は、最初は素通りしても音声に反応して振り返り、来た道を戻って、音声を聞き消毒をする人が多かった。逆に、イヤホンをしている人や目的をもって早歩きしている人、また両手に荷物を持っている人はあまり、消毒液に注目をしなかった。また、雨の日があったが傘を持っている人も同様、なかなか消毒液に関心を示さなかった。改善点は、人が通過しないタイミングでも人感センサが反応して、音声出力していたので更に正確な人の感知を目指す。

音声出力に自然さを感じたかどうかについては、音声出力を文語ではなく口語にしたことで、このようなポジティブな結果を得られたと考える。当初は文語で音声出力していたが堅い音声出力と感じ、結果口語にすることで親近感を感じてもらえるように変更を行った結果であると考察した。

消毒液の内容を理解できたかどうかの回答が最もよかった理由は、スピーカの位置であると考える。作成段階でスピーカの位置を正面にしていなかった為に、聞きにくいと指摘を受けたことがあり実際に自分でも感じていた。ゆえに、2 つのスピーカを正面方向に取り付けたことにより、音が聞きやすくなったためであると考えた。

最後に、消毒液が喋っているように感じるかどうかについては、回答者の過半数以上の方がそう感じてくれたが、「そう思わない」の回答が一番多い項目でもあった。想定以上の回答スコアであったが、この項目は本研究において重要項目であるので更なる飛躍を目指したい。実験前は、消毒液に「顔」がないために音声出力のみで消毒液が喋っている感覚を演出できずに、「そう思わない」の項目が多くなりスコアが下がることを想定していたが、実際は複数の音声出力により、消毒液が喋っている演出ができ、高スコアに繋がったと考える。

4.3　提案システムのシステム検証と評価

IoT スピーカを消毒液に取り付けることで、人の注意を引き、人に対する使用率の変化が現れるかどうかを実験により評価する。音声による注意喚起が可能な消毒液と音声出力させない普通の消毒液を一般の通行客に対してどの程度、消毒液の使用率に効果があるのかの比較検証である。

4.3.1 計測方法

計測方法は、図 4-8 のように、ローソン前の開けた空間の一角に IoT スピーカを取り付けた消毒液を設置し、平日午後の 1 時間で一般の人々を対象に、メモをとる形で実施した。

図 4-8　システム検証の実験場所

図 4-9　システム検証の実験の様子

図 4-9 は、音声出力する消毒液に反応して、消毒液に注目し、使用する様子である。ローソンに向かって歩いてくる人とローソンから出てくる人が対象である。実験とは事前に知らせずに、観察することで客観的に使用率を把握する。更に、人感センサはサーバにアクセスを行うので人感センサが反応した時刻はサーバ上に記録される。よって、目視と記録の両方を参考にカウントを行った。

日付	時間	性別	年齢	人数	特徴（例えばイヤホン聞いている）
1月12日	15:01	男性	10代	ひとり	高校生
	15:02	女性	60代	ひとり	セントラル利用のおばさん
	15:09	女性	40代	ひとり	
	15:17	女性	30代	ひとり	あったから使ったような感じ
	15:20	女性	40代	ひとり	ローソンの店員
	15:21	女性	40代	ひとり	ローソンから出てきた人
	15:32	女性	30代	複数	子供連れの親子
	15:36	男性	20代	ひとり	SDMの学生
	15:41	男性	30代	ひとり	
	15:44	女性	50代	ひとり	ローソンに入っていった人
	15:50	男性	20代	ひとり	

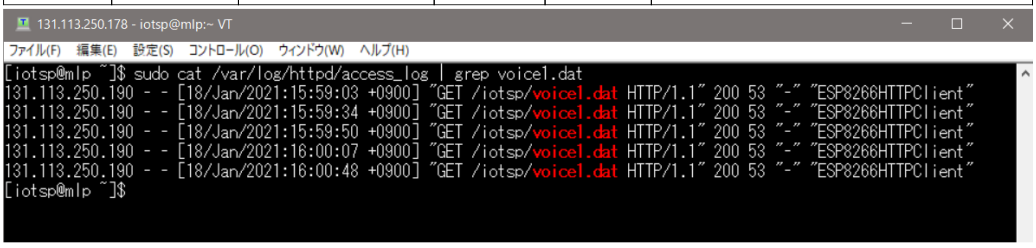


図 4-10　システム検証の記録メモ

基本的に個人情報は扱わないが、一般人に対して許可を取らずに観察を行うため、観察者は腕章をつけることで、周囲の人に不信感を与えないように注意した。データの取り扱いに関して、記録として収集されるデータは、実験期間中に前を通った人数と、アルコール消毒を行ったかどうかの目視によるカウント数であり、個人情報は一切含まれない（図 4-10）。人感センサのカウントデータは、学内のサーバ内に記録・保存され、パスワードで管理されるため、実験者以外はアクセスすることができない。また実験結果のデータに関しては、修士論文や学会発表等における基礎データとして使用するが、研究以外の目的では使用しないという条件で学内倫理員会の承認を得て、実験を行った（承認番号：SDM-2020-E044）。

図 4-11　IoT スピーカを取り付けた消毒液

音声出力させる場合は、前方の人感センサが反応して「僕はアルコール！僕を押して消毒してね！」と音声出力を行う。消毒を行うために消毒液上部のノズルを押した場合、二つ目の人感センサが反応して「協力ありがとう！」が出力される。音声出力させない場合では、IoT スピーカの人感センサは作動せるが、音声出力をオフにすることで音声出力をしない。スピーカの前方に「手指の消毒にご協力をお願いします」のポスターをセットする（図 4-11）。

4.3.2 結果と考察

結果は図 4-12 に示す通りであった。音声出力なしの場合、通行人数は 84 人、使用人数は 20 人となりその使用率は、23.8%であった。一方で IoT スピーカを取り付け、音声出力した場合は、通行人数は 72 人、使用人数 32 人で使用率は 44.4%であった。

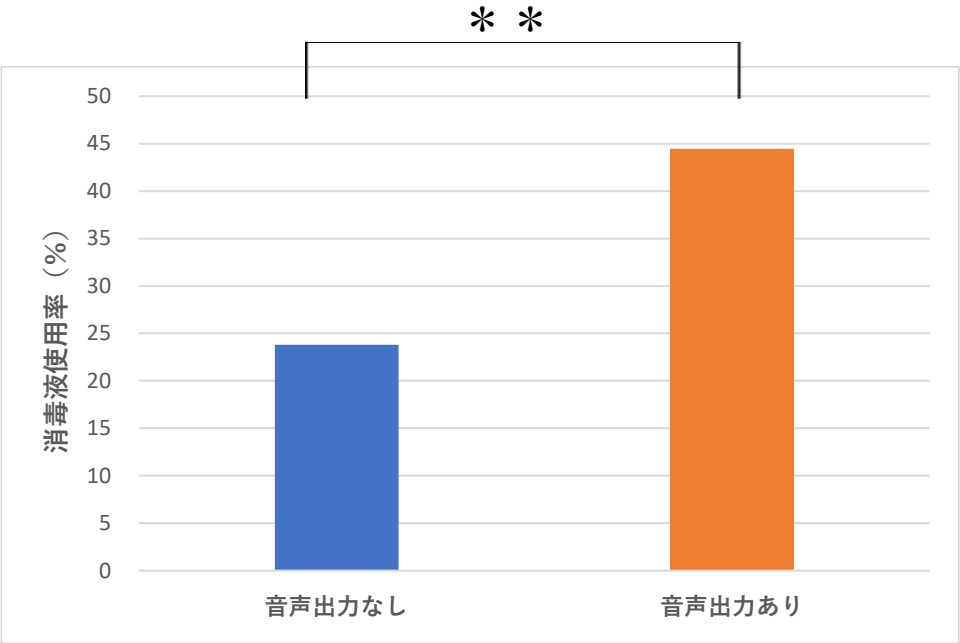


図 4-12　音声出力による消毒液の使用率比較

観察による客観的な見解は、検証 3.5 と同様にゆっくり歩いている人が特に音声に反応して消毒液を使用する傾向があった。消毒液の近くをゆっくり歩いている人は、最初は素通りしても音声に反応して振り返り、来た道を戻って、音声を聞き消毒をした。また、人が近づく段階で音声出力されるため、消毒液に注目しながら近づき、結果使用する人が多かった。しかし音声なしの場合は、消毒液に気づいても使用しない人が多かった。消毒液に注目するも関心を示さず、使わないで素通りしていた。これら結果から、音声出力による注目度アップにより、消毒液の使用率が向上したといえる。

以上のことから、本研究の目的である IoT スピーカによる音声情報提示がユーザの行動を変える支援効果があることが確認された。

4.4 複数の IoT スピーカの連携

複数の IoT スピーカを用いた連携による行動支援の例を紹介する。ここでの例は、協生館の 3F において、廊下から実験室への入室までを目的とした行動支援である。まず額縁に入ったポスター①が喋ることで実験室の場所を案内する。次に、実験室入り口で消毒液②が消毒液使用を勧める。最後に入室する際のスリッパ使用をぬいぐるみ③が促す音声出力を行うものである。

図 4-13　IoT スピーカの連携その 1

図 4-13 は、廊下突き当りを右手に進み階段を登っていく様子である。研究室は正面右手に進んだ先にあるため、額縁ポスターに付けた IoT スピーカが「実験室へようこそ！右に進んでね！」と発声することで、額縁ポスターが実験室へいくには、右手に進めばいいことを教えくれる。

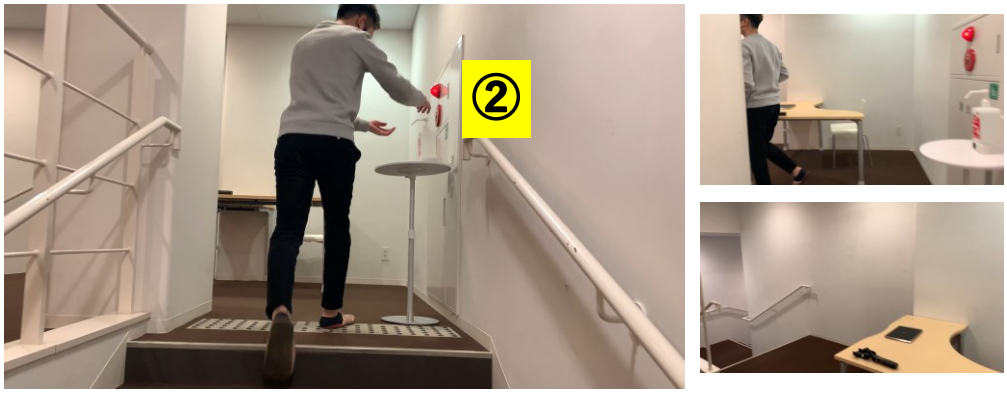


図 4-14　IoT スピーカの連携その 2

図 4-14 は、階段を進み実験室に向かう際に消毒液使用を促す消毒液を使用している様子である。入室する際に、消毒液の使用を音声出力するので通過する人は手を消毒する。

図 4-15　IoT スピーカの連携その 3

図 4-15 は、実験室に入室する際の様子である。入室には、外履きを脱いでスリッパを履き入室することになっているので、ぬいぐるみに取り付けたスピーカが「スリッパを履いて中に入ってね！」と発声する。これによって、訪問者はスリッパを履いて中に入っていけばよいことが理解できる。

今回のデモンストレーションでは、IoT スピーカは独立してそれぞれ実験室案内、消毒液使用の促進、スリッパ使用という行動支援を行った。これは IoT スピーカはそれぞれ独立した行動支援であり、互いにインターネット通信で連携しているわけではない。しかしながら、勿論プログラム次第では、三つの IoT スピーカを互いにインターネット通信により、システム連携させ多段的に行動支援を行うプログラムを書くこむことができる。

第5章
結論と今後の展開

本章では、本研究の結論及び今後について述べる。5.1 章では 1 章から 4 章をまとめ、本研究によって得られた結論について述べる。5.2 節では、本研究の今後の課題と発展について述べる。

5.1 結論

本研究では、いろいろなモノが喋る環境構築を目標に、様々なモノに取りつけられる小型 IoT スピーカの開発を行い、開発した装置の機能を評価するために基礎実験を行った。またその提案システムの実証実験を実施した。

提案システムの IoT スピーカに求められる要素は、あらゆるモノに取り付けられるように小型で、人感センサを備えることで人の動作に反応し、ネットワーク経由で情報を取得し、音声情報の出力を行うことで、対象物がユーザに対してインタラクティブに情報提供を行うことである。柔軟な情報提示を行うために、ネットワーク接続可能な ESP-WROOM-02 を用いた。また、音声記号列を音声情報に変換する音声合成 LSI、情報提示のために音声出力のダイナミックスピーカを用いた。更に、ユーザに特別な行動を求めないインタラクティブな情報提供のために人感センサを用いた。

人感センサの精度とスピーカの音の出力の基礎実験では、300cm の感知範囲において 50dB 以上の出力があり、実際にユーザに聞こえるレベルであることを確認した。また人感センサがユーザを感知してから音声が発せられるまでのタイムラグは 0.30 秒以下であり充分な速さであることも確認できた。

歩行者に対する音声提示実験を行った結果は、スピーカに向かって正面に歩いていく場合も、真横を通り過ぎる場合もサーバ音声を十分に再生する時間が確保できた。しかし、実験結果から IoT スピーカの正面に向かって歩行するよりも 90 度方向から横切る方が、人感センサの検知精度が向上し、音声情報を伝達しやすいことが分かった。

モノが喋る感覚の要因検証では、消毒液とスピーカの距離が全く一緒でなくても、ある程度近い距離であれば、消毒液の音声出力と感じる結果を得た。
また、文語による音声出力よりも口語による砕けた表現を用いることで、より消毒液が喋りかける感覚を与えること府ができた。

IoT スピーカを消毒液に取り付け、消毒液が喋っているように感じるかどうかの検証においては、回答者の過半数以上が消毒液からの音声出力に自然さを感じ、消毒液が喋っているように感じる結果となった。特に音声内容の理解に関しては、多くの人が音声内容を理解した結果となった。この結果から、モノが喋っている状態を実現することができたといえる。

IoT スピーカを消毒液に取り付けることで、人に対する使用率の変化が現れるかどうかの検証では、スピーカを取り付けることで使用率が向上した。音声に反応して、振り返り消毒液を使用する人を多数確認し、音声提示の有効性を確認することができた。この結果から、モノが喋ることでユーザの行動を変えることができた。

以上のまとめから、本研究の結論として次のことが言える。

- IoT スピーカ開発を行い、「モノが喋り情報を与える環境」を構築することで提案システムの有効性を検証した。
- IoT スピーカを消毒液に取り付け、適切な音声出力を行うことで「モノが喋っている状態」を再現することができた。
- 消毒液が喋ることで、消毒液使用率が向上した結果、「モノが音声により情報を与えることで行動支援する」ことの一端を担うことができた

5.2 今後について

5.2.1 今後の課題

本研究の今後の課題としては、次の点が挙げられる。

- IoT スピーカの更なる小型化
- 対象ユーザ個人の特定
- IoT スピーカ同士の連携機能の構築

本研究では、公共空間において特定の人のみでなく誰に対しても聞こえ、かつ役に立つ情報を提供することで一般の人々の行動支援を目指した。そのためにモノに取り付け可能な小型スピーカを搭載したデバイスの開発を行い、検証を行った。学校施設内で実験を行ったが、公共空間では騒音が大きいため音声出力が聞こえにくくなる可能性が考えられる。そのため、音の大きさを大きくすることも考えられるが、そうなると逆に宣伝広告のようになり、ユーザが関心を持たない可能性が出てくるため、適度な音の大きさが求められる。

また「モノが喋る」という概念を実現するために、あるユーザの特定と、IoT スピーカの連携が今後の大きな課題である。不特定多数の誰にでも役に立つという情報は限られてくるので、個人にカスタマイズされた情報を提供することが考えられる。ターゲットとするユーザを特定し、複数の IoT スピーカがそのターゲットに何回もアクセスするという連携を実現させることで「モノが喋る環境」を実現できるであろう。

5.2.2 今後の展望

本研究では、公共空間において特定の人だけでなく、誰に対しても聞こえ、かつ役に立つ情報を使用することで一般の人々の行動支援を目指した。今後の展望として、一定の範囲の空間にマイクを設置することで、モーションのみでなく音声によるインタラクションも追加する。これによって、音声とモーションの 2 つのインタラクションが実現でき、モノと会話するという概念を更に進展させることができる。更に、イヤホンを使用するユーザへのアプローチの為に既存のアプリケーションとの連携をすることでイヤホンユーザへの音声提供も可能になると考える。

謝辞

この研究は、周りの方々のサポートなしに完成することはできませんでした。研究にかかわってくれたすべての方々に、心より感謝いたします。
研究の考え方から、研究の進め方、提案システムの概念に至るまで親身にご指導くださった小木哲朗教授に感謝いたします。学会発表では、様々な研究に触れることができ、いい刺激になりました。

副査を担当してくださった春山真一郎教授にも感謝いたします。お忙しい中でも時間を割いていただき、研究についてアドバイスをいただき、研究の糧となるご指導をくださいました。

小木研究室の松岡慧特任助教にも感謝いたします。幅広い知見から研究のアドバイスをいただきました。また研究テーマを決める最初の段階でもご指導をいただきました。冷静な姿勢で研究を正しい方向へ導いてくださいました。

研究室の同期の木田勇輝さん、王鳳超さん、陳祐蕎さんにも感謝いたします。研究に関して、講義をしたり励ましあったりしながら、大学院への生活を送ることができました。

実験の被験者として協力して下さった皆様にも感謝いたします。お忙しい時にも関わらず、快くご協力をいただき実験を行うことができました。
最後に、これほど長い期間にわたって私に学ぶ機会を与えてくれた両親に感謝の意を表します。これまで学んできた多くのことを糧として、社会のために貢献できるよう努力してまいります。

2021 年 2 月 12 日
慶應義塾大学大学院
システムデザイン・マネジメント研究科
システムデザイン・マネジメント専攻
榎本 敬太

外部発表

１．榎本敬太、松岡慧、小木 哲朗：マルチ IoT スピーカを用いたモノとのインタラクション、ヒューマンインターフェース学会、ヒューマンインターフェースサイバーコロキウム、pp.473-476、2020.

２. Tetsuro Ogi、Keita Enomoto、Shohei Akita、Toshifumi Miyahashi、Takehito Akima、Takumi Sano：Construction of Information Presentation Environment Where Various Objects Talk Using Multiple IoT Speakers、日本バーチャルリアリティ学会テレイマ―ジョン技術研究委員会、第 41 回テレイマージョン技術研究会、2020.

参考文献

[1]　水越敏行、久保田賢一：ICT 教育のデザイン、日本文教出版株式会社、Vol.3、No.1、pp.12、2008.

[2]　上嶋智大、中平勝子、三上喜貴：デジタルデバイドの評価指標についての一提案、情報科学技術レターズ、No.6、pp.481-484、2007.

[3]　今川拓郎：デジタルデバイドの実証分析、OSIPP Discussion Paper、DP-2002-J-015、2002.

(https://pmall.gpoint.co.jp/g-ranking/ranking.php?themeid=12314)

[4]　Canalys. ： Smart Speakers Analysis Analysis.

(Canalys Newsroom- Google beats Amazon to first place in smart speaker market)

[5]　CES Topics - Home &Family Smart Home Home.

(https://www.ces.tech/Topics/Home-Family/Smart-Home.aspx)

[6]　秋田祥平、秋間建人、佐野拓海、宮橋理史、小木哲朗：人間の行動支援のためのマルチ IoT スピーカを用いた音声コミュニケーション環境の構築、日本機械学会第 28 回設計工学・システム部門講演会講演論文集、No.18-11、2018.

[7]　大久保紘彦、久良知國雄、藤澤正一郎、末田統：赤外線音声情報案内システムを利用する視覚障害者の方向特定特性、福祉のまちづくり研究、Vol.11、No.2、pp.21-28、2010.

[8]　伊藤亮介、駒谷和範、河原達也：機器操作マニュアルの知識と構造を利用した音声対話ヘルプシステム、情報処理学会論文誌、Vol.43、No.7、pp.2147—2154、2002.

[9]　辻野孝輔、栄藤稔、礒田佳徳、飯塚真也：実サービスにおける音声認識と自然言語インタフェース技術、人工知能学会、Vol.28、No.1、pp.75-81.2013.

[10]　松田侑己、小木哲朗：デジタルサイネージによる個人の興味に対応した広告提示システム、修士学位論文 2013 年度システムデザイン・マネジメント学、Vol.148、2013.

[11]　張志華、高橋雄太、藤本まなと、荒川豊、安本慶一：行動変容を誘発するためのインタラクティブサイネージの検討、2017 年度情報処理学会関西支部支部大会講演論文集、2017.

[12]　野村篤史、外嵜一登、小川均、林勇吾：スマートフォンを用いた行動モデルによる行動支援システム、情報処理学会全国大会講演論文集、Vol.74、No.3、pp.3469-3470、2012.
[13]　板宮朋基、吉岡英樹、大山昌彦、小川高志：津波避難行動支援スマートフォンアプリ「デジタル皆助ナビ」の開発と高知県黒潮町における実証実験、情報処理学会全国大会講演論文集、Vol.75、No.4、pp.4515-4516、2013.
[14]　Ministry of Land、Infrastructure and Transport
(https://www.mlit.go.jp/kisha/kisha02/01/011017_.html)
[15]　田中昌昭、平松生光、小山真生、谷口菜々子、勇裕加里、内田直希：スマートスピーカを使った独居高齢者のための服薬支援システムの開発、川崎医療福祉学会誌、Vol. 30、No. 1、pp.257-263、2020.
[16] 加藤修子：博物館のサウンドスケープ・デザインにおける「音声ガイド」の考察、文化情報学 ： 駿河台大学文化情報学部紀要、Vol16、No.2、pp.1-19、2009.
[17]　Chappet
(https://chappet.com/)
[18]　TALKINGPOP
(https://www.i-studio.co.jp/news/2017/03/sxsw-talkingpop.html)
[19]　西村拓一、伊藤日出男、山本吉伸、中島秀之：無電源小型通信端末を用いた位置に基づく情報支援システム、人工知能学会全国大会論文集、第 16 回全国大会、2002.
[20]　モノを無線でつなぐ、モノワイヤレス株式会社
(https://mono-wireless.com/jp/tech/Internet_of_Things.html)
[21]　SWITCHSCIENCE：OPEN SOURCE HARDWARE SHOP
(https://www.switch-science.com/catalog/2346/)
[22]　AQUEST：Text To Speech Technology
(https://www.a-quest.com/products/aquestalkpicolsi.html)
[23]　秋月電子通商
(https://akizukidenshi.com/catalog/g/gM-09002/)

[24]　西田鶴代、筧一彦、穂刈治英、島田正治：音源定位における視覚情報の影響－FLMP による視覚情報の影響の定量化－、日本音響学会誌、Vol.55、No.11、pp.735-741、1999.
[25]　O.Klemm：Lokalisation von Sinneseindrücken bei disparaten Nebenreizen、Psychol、Stud.5、pp.73-162、1909.
[26]　中林克巳、二階堂誠也：音像、映像の相互作用に関する基礎案験、音響学会聴覚研資、pp.1-4、1979.
[27]　H.Fujii、K.Kakehi：Effect of visual information on Speaker localization among several speakers、pp825-828、1996.

付録

—Arduino IDE のプログラム—

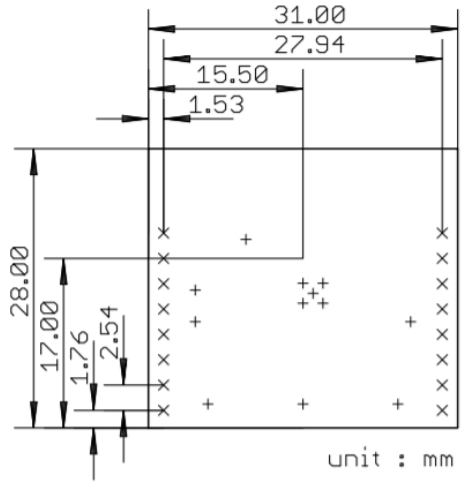
```
#include <ESP8266HTTPClient.h>

//      WROOM-ATP3011-server:
//      WROOM-02 output voice uing ATP3011 by I2C and Server Communication
//                                                              2020/09/01 Ogi Lab.

#include <ESP8266WiFi.h>
#include <ESP8266HTTPClient.h>
#include <Wire.h>

#define SENSOR_PIN 14
#define LED_PIN 16

// ATP3011 I2C Address
const uint8_t ADDR = 0x2E;

// WiFi connection
//const char* ssid = "";
//const char* password = "";
const char* ssid = "sdm-mobile2";
const char* password = "KeioSdm2016";


// Server information
const char *host = "131.113.250.178";
const int port = 80;
```

```
int flag = 1; //変数宣言


WiFiClient client;

/* ** SET UP ** */
void setup() {
  pinMode(SENSOR_PIN, INPUT);
  pinMode(LED_PIN, OUTPUT);

//   SDA and SCL (Default)
  Wire.begin(4, 5);

  WiFi.mode(WIFI_STA);
  WiFi.begin(ssid, password);

  while(WiFi.status() != WL_CONNECTED) {
    delay(50);
  }
  speak("waifaini/ setuzokushimashitayoo¥r");
}

/* ** LOOP ** */
void loop() {
  String line;

  if( digitalRead(SENSOR_PIN)==HIGH ){
    if(flag>=3){                                    //条件追加 flag
      digitalWrite(LED_PIN, HIGH);
```

```
        HTTPClient http;
        http.begin("http://131.113.250.178/iotsp/voice1.dat");

        int httpCode = http.GET();
        String line = "";
        line = http.getString();
        http.end();
        speak(line);

        delay(8000);
    }
        flag=0; //条件追加
      } else if( digitalRead(SENSOR_PIN)==LOW) {
        digitalWrite(LED_PIN, LOW);
        delay(1000);     // check sensor every 0.5 sec
        flag=flag+1; //条件追加



      }
}

// Speak voice by ATP3011
void speak(String buff) {
    uint8_t *data = (uint8_t *)buff.c_str();
    write_data(ADDR, data, buff.length());
    uint8_t cmd[] = {0x0D};    // send CR to ATP3011
    write_data(ADDR, cmd, 1);
    delay(100);
}
```

```
// Devide 32 bytes for I2C communication
void write_data(uint8_t addr, uint8_t *data, size_t len) {
    const size_t max = 32;
    size_t n = len / max;
    size_t m = len % max;
    size_t i = 0;
    for (; i < n; i++) {
        write(addr, &data[i*max], max);
    }
    write(addr, &data[i*max], m);
    delay(500);
}

// Send voice data to ATP3011 by I2C
void write(uint8_t addr, uint8_t *data, size_t len) {
    Wire.beginTransmission(addr);
    Wire.write(data, len);
    Wire.endTransmission();
}
```

—「ESP-WROOM-02 ピッチ変換キット」の基板寸法図—



Sym	TN	mm	Qty	TH/NTH
+	1	0.30	12	TH
×	2	0.90	16	TH

AE-ESP-WROOM-02-1100mil