

技術・研究報告

アンドロイド・ロボットを用いた大学図書館司書のためのヘルプデスク遠隔対応システムの開発

菊地智也^a, 梅谷智弘^b, 才脇直樹^c

^a 甲南大学大学院 自然科学研究科 知能情報学専攻

神戸市東灘区岡本 8-9-1, 658-8501

^b 甲南大学 知能情報学部 知能情報学科

神戸市東灘区岡本 8-9-1, 658-8501

^c 奈良女子大学大学院 人間文化研究科 生活工学共同専攻

奈良市北魚屋西町, 630-8506

(受理日 2019 年 5 月 7 日)

概要

本稿では、アンドロイド・ロボットを用いた大学図書館司書のためのヘルプデスク遠隔対応システムの開発について述べる。大学図書館ヘルプデスクの業務には受付での接客業務だけでなく、図書館書架の整理や事務所での作業などがあり、受付から離れて業務を行う必要がある。その際には、受付が無人となってしまう、利用客にとって図書館の利便性を損なってしまう。また、専門知識が必要な場でロボットによる自動対話を行う場合、利用客の多様な要望に対応することが難しく、利用客に不快な感情を抱かせてしまう。本稿では、それらの問題点の解決に向けて、これまでに開発されたアンドロイド・ロボットを用いたビデオ通話システムを拡張し、ビデオ通話時のミュート機能と会話音声の録音、音声文字化による図書館司書のための記録機能を構築する。構築したシステムを図書館ヘルプデスクの実業務にて長時間運用することで本システムの可能性、有用性を示す。

キーワード: アンドロイド・ロボット, 遠隔コミュニケーション, 大学図書館, ヘルプデスク

1 はじめに

近年、人—ロボット間で対話するシステムの研究が盛んに行われており、ロボットを介した人へのサービス提供が期待されている。これまで、情報学研究のプラットフォームとしての受付案内ロボット[1]や、図書館案内ロボット[2],[3]が開発され、長時間にわたっての運用、実演がなされている。また、IBM Watson といった人工知能（拡張知能）技術の一般公開によって、人工知能技術をロボットへ搭載する際の開発難易度が下がり、ビジネス面などでの課題に人工知能（拡張知能）を搭載したロボットを用いて解決するといった期待感が高まっている[4]。

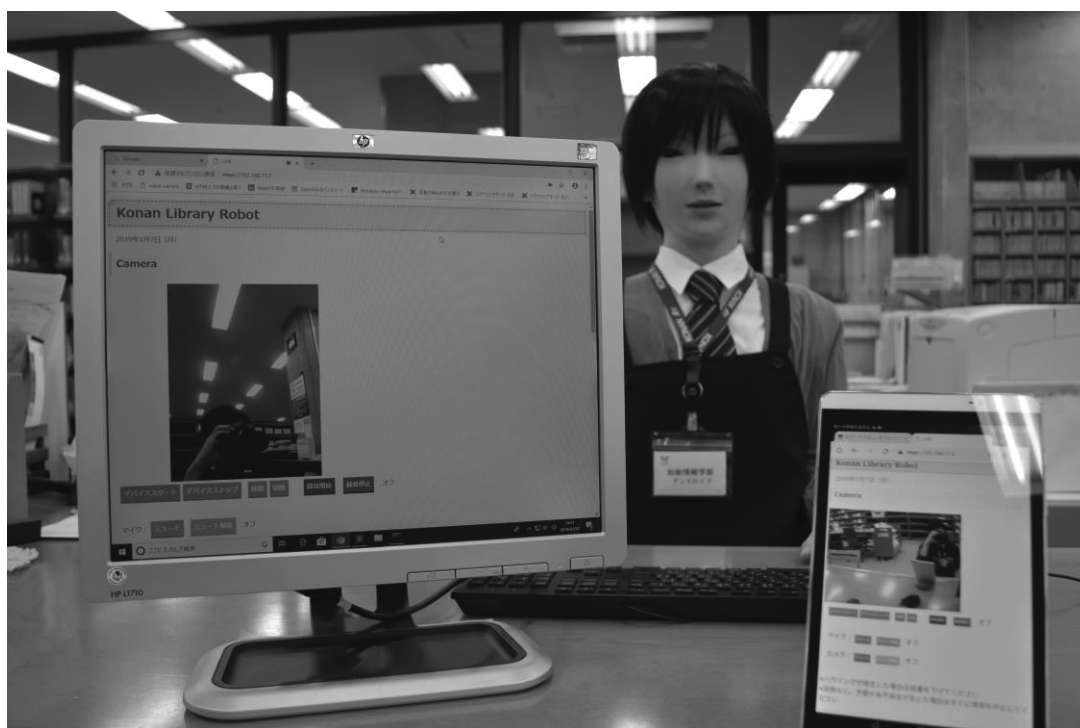


図1：アンドロイド・ロボットを用いたヘルプデスク遠隔対応システムの概観

ここで、図書館ヘルプデスクなど、専門知識を必要とする業務でのロボットの利活用を考える。このとき、ヘルプデスク職員は受付業務だけでなく、書架や事務所など受付から離れて業務を行う場合がある。その際、受付は無人となってしまい、利用者からの観点では機会損失に繋がり、図書館の利便性を損なってしまう。また、図書館司書が受付外での業務中は手が空いておらず記録を取りにくい状況が多々ある。さらに、利用者との自動対話を考えると、利用者の多様な要望があるため完全な自動化は難しく、専門知識の蓄積が課題となっている。

本稿では、アンドロイド・ロボットを用いた大学図書館司書のための遠隔対応システムについて述べる[5],[6]。図1に実際に図書館ヘルプデスク受付に設置されているアンドロイド・ロボットとシステム全体の概観を示す。本開発では、遠隔かつ無線でのビデオ通話によって来客がロボットと直接会話しているようなシステムを構築する。従来では音声通話サービス Skype を使用したシステムを構築していたが、拡張性の面で課題となっていた。

本研究ではこれまでに開発した基礎システム[5]を拡張し、ビデオ通話時のマイク、カメラのミュート機能、会話音声の録音機能、IBM Cloud Watson API を利用した音声文字化による記録機能を構築する。構築したシステムを図書館ヘルプデスクの実業務にて長時間運用することで本システムの有用性を示す。また、記録機能利用時の音声データおよび文字データの収集によって専門知識を必要とする場でのロボットの完全自動化への可能性を示す。

2 アンドロイド・ロボットを用いたヘルプデスク遠隔対応システム

本章ではアンドロイド・ロボットを用いたヘルプデスク遠隔対応システムについて説明する。まず、ヘルプデスク遠隔対応システムについて説明し、従来用いられていたビデオ通話システムの問題点を整理する。あわせて、遠隔対応システムにアンドロイド・ロボットを用いることの意義についてまとめる。

2.1 ヘルプデスク遠隔対応システム

ヘルプデスク遠隔対応システムについて説明する。ヘルプデスク遠隔対応システムとは、図書館職員が受付に在席せず、受付に来訪する利用客との対話を行うことを可能にするシステムである。図書館ヘルプデスクでの業務内容として図書館書架の整理、事務所での作業があり、受付から離れた場所で作業を行う必要がある。その際にヘルプデスク遠隔対応システムを使用する。職員側の持つデバイスには遠隔での対応時に持ち運びやすさの観点から、小型かつ軽量化する必要があるため、タブレット端末での遠隔かつ無線での対応が必要となる。また、遠隔での対話には利用客の有無の確認、音声による対話が必要のため、ビデオ通話によって互いのカメラ映像の取得、マイクおよびスピーカでの音声取得を行う。従来の遠隔対応システムでは、アンドロイド・ロボットに搭載されたカメラとマイクを用いて、ノートPCと図書館司書の持つタブレット端末との間にSkypeを用いたビデオ通話システムを構築していた。これによってロボットと話す感覚で職員と遠隔での対話を実現していた。本開発では、アンドロイド・ロボットに搭載されたカメラとマイクを用いて、デスクトップPCと図書館司書の持つタブレット端末との間にリアルタイムコミュニケーションシステムを構築する。具体的には、HTML上でのビデオ通話システム、音声およびカメラのミュートシステム、録音システム、録音音声文字化システムを構築する。

2.2 問題設定

従来のSkypeを用いたビデオ通話システムの問題点について説明する。Skypeではメールアドレスといった個人情報の登録やログインが必要となっている。この要因として、世界中に利用客がおり、セキュリティ面の強化が必須であることが考えられる。また、開発を行う際の動画像や音声のリアルタイムの加工が困難である。この点についてもセキュリティ面やプライバシー面での問題が関係しているからだと考えられる。他にも、機能がSkypeアプリケーションに依存しているという問題点がある。これは専門職など特定の条件や業務での利用者に対するアプリケーション開発が困難となり、開発時の拡張性について大きな問題点になる。これらからSkypeを用いたシステムは規模が大きく、専門知識を必要とする業務での問題点に対して、音声対話の自動化といった情報開発による柔軟な解決を行うには不向きであると考えられる。

また、現場でビデオ通話システムを使用する際の問題点として、ビデオ通話を常時接続していた場合には別業務での職員同士の会話内容が受付側で公開されてしまうことがある。他にも、職員側の映像が不必要なことや、別業務での作業中は利用客と対話しながら記録を取ることが困難

になる状況が多々ある。これらの問題に対する解決を本開発によって目指す。

2.3 アンドロイド・ロボットを用いる意義

ここで、アンドロイド・ロボットを用いる意義について説明する。対応システムにアンドロイド・ロボットと、タブレット端末や他ロボットを用いた場合の利便性、特徴を比較する。遠隔対応システムにおける他機器とアンドロイド・ロボットを用いることの特徴を図2に示す。

まず、カメラおよびマイクそのもの、タブレット端末や児童を対象としたロボットを用いた場合について説明する。カメラやマイクそのものを用いた場合には監視されている感覚が生じてしまい、受付が無人の場合は近づきにくく話しかけにくくなってしまう。結果として職員を探してしまうことになり利用者にとって不快感が生じてしまう。児童を対象としたロボットを用いた場合、丁寧な接客が必要な場にふさわしくないため、システムが利用されなくなってしまふ。これは図書館の利便性の低下に繋がってしまう。

これらの問題点に関して、アンドロイド・ロボットはヒトに似ているという特徴がある。これは、話しやすさの面で受付が無人となっても近づきやすく話しやすいことに繋がる。また、図書館司書といった専門知識が必要な職業や図書館という丁寧な接客が必要な場にふさわしいため、システム自体の利用が容易になる。システムの利用が容易になることに伴い、来客との対話の機会が増え、記録機能の使用回数の増加によって図書館内という静穏な場所で雑音の少ない音声データを多く収集することが可能となる。この収集した音声データは接客時における対話の自動化に活用することができる。他にもアンドロイド・ロボットを用いる利点として現代社会では風変わりであることがいえる。イベント時にはアンドロイド・ロボットに話しかけるために来館する利用客がいた。このことからアンドロイド・ロボットを用いることは、注意を引き、集客力を増加させる要因になり得る。以上より、アンドロイド・ロボットを用いる意義があると考えられる。

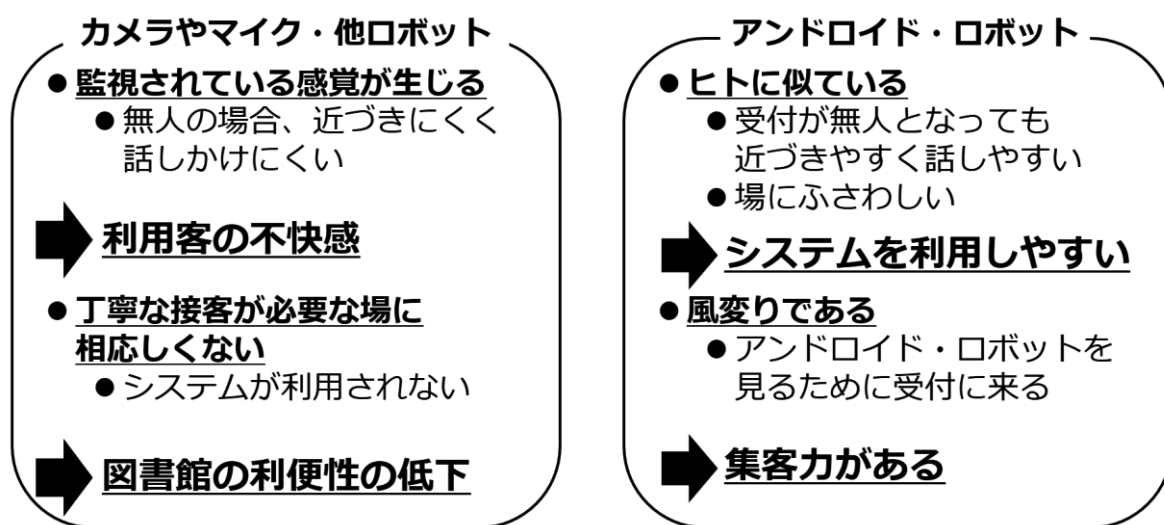


図2: アンドロイド・ロボットと他機器の比較

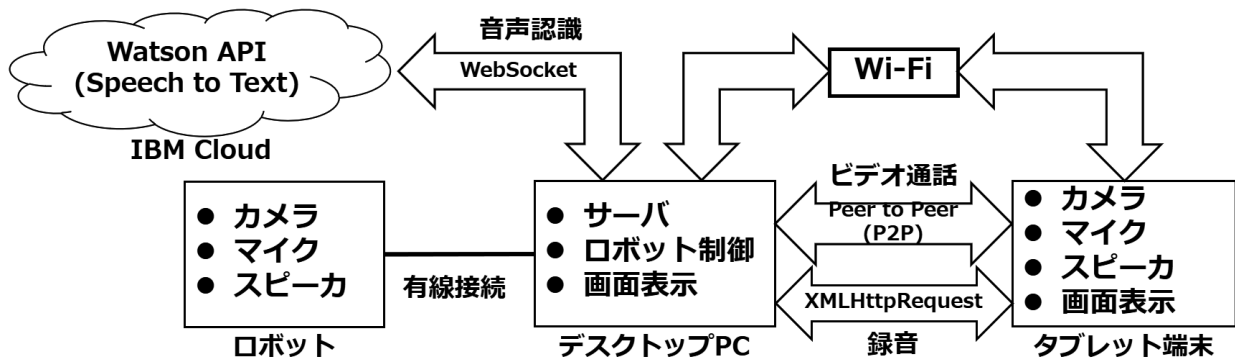


図 3：遠隔対応システムの全体図

3 システム構築

本章では、ヘルプデスク遠隔対応システムの構築について説明する。本開発は、大学図書館ヘルプデスクに備えたアンドロイド・ロボットに搭載されたカメラとマイクを用いて、デスクトップPCと図書館司書の持つタブレット端末との間にHTML上でのリアルタイムコミュニケーションシステムを構築する。システムの概観を図1に示す。本システムにはビデオ通話システム、音声およびカメラのミュートシステム、録音システム、録音音声文字化システムが含まれる。ハードウェア構成、ソフトウェア構成を説明し、図書館ヘルプデスクの実業務にて長時間利用することで本システムの可能性、有用性を示す。

3.1 ハードウェア構成

システムの全体図を図3に示す。カメラとマイクを搭載した1体のアンドロイド・ロボット、長時間運用を可能にするサーバ機としてWindows 10デスクトップPCを使用する。ロボット搭載のカメラ、マイク、スピーカはこのデスクトップPCに接続する。デスクトップPCと図書館司書とのリアルタイムビデオ通信用デバイスとしてAndroid 6.0.0タブレットを使用する。デスクトップPCとタブレット端末は同一LAN内にある。録音音声の文字化にIBM Cloudサービスを使用する。

3.2 ソフトウェア構成

3.2.1 ビデオ通話システム

ビデオ通話システムのタブレット端末からの接続の流れを図4に示す。HTML5上でのリアルタイムビデオ通話システムを実現するためにWeb Real-Time Communication (WebRTC)のAPIを使用する。WebRTCとはブラウザ間でビデオ通話システムやファイル共有などのリアルタイムコミュニケーションを可能にするHTML (Javascript)のAPIである。今回の開発では大学図書館の限定し

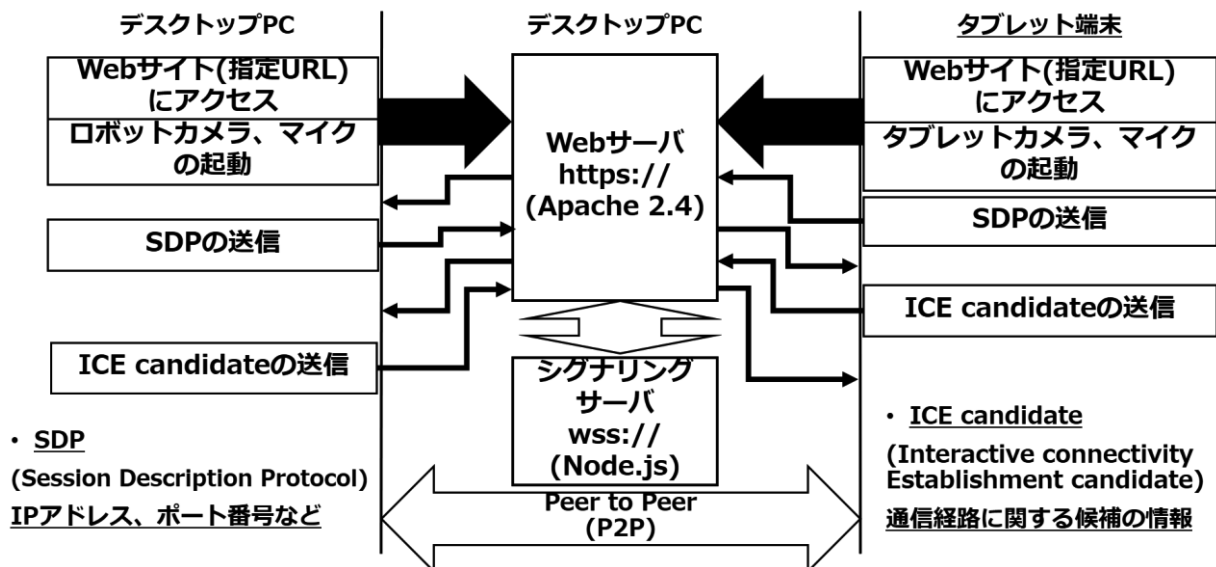


図4：ビデオ通話システム接続の流れ

た場所でWebRTCの機能を実現するためにデスクトップPCに2種類のサーバを構築する。1つは、Apache 2.4を使用したHTMLにブラウザからアクセスするためのWebサーバ、2つ目のサーバとして、Session Description Protocol (SDP)およびInteractive Connectivity Establishment candidate (ICE candidate)の相互交換を仲介するためにNode.js 8.9.4を利用したシグナリングサーバを構築する。このシグナリングサーバにはWebSocket通信用サーバモジュールwsを用いてjavascriptで実装する。サーバの実装にあたってGoogle Chrome上でのカメラ映像、マイク音声取得機能の実現にセキュアな接続通信が必要なため、opensslによってhttpsおよびwssでの実装を行う。また、本来のWebRTCの機能の使用ではLAN外からのサーバアクセスへの対応がある。その際にはICEサーバ (STUNサーバおよびTURNサーバ)が必要となるが、今回の実装では大学図書館内の同一LAN内のみで限定するため、構築しない。

接続の流れについて説明する。図4にビデオ通話システムのタブレット側からの接続の流れを示す。図4のように、はじめにデスクトップPCとタブレット端末の両方から同一のWebサイトにアクセスし、互いのカメラ、マイクを起動する。タブレット側からSDPを送信する場合、HTML (CSS)のボタン操作によって、ブラウザ、サーバの順に経由してタブレット端末のSDPをデスクトップPCに送信する。タブレット端末のSDPを受信後、デスクトップPCのSDPをタブレット端末へ送信する。その後、タブレット端末側からICE Candidateについても同様に相互交換される。ICE Candidateの交換中または交換後にデスクトップPCとタブレット端末側のPeer to Peer通信が確立され、ビデオ通話機能が実現される。

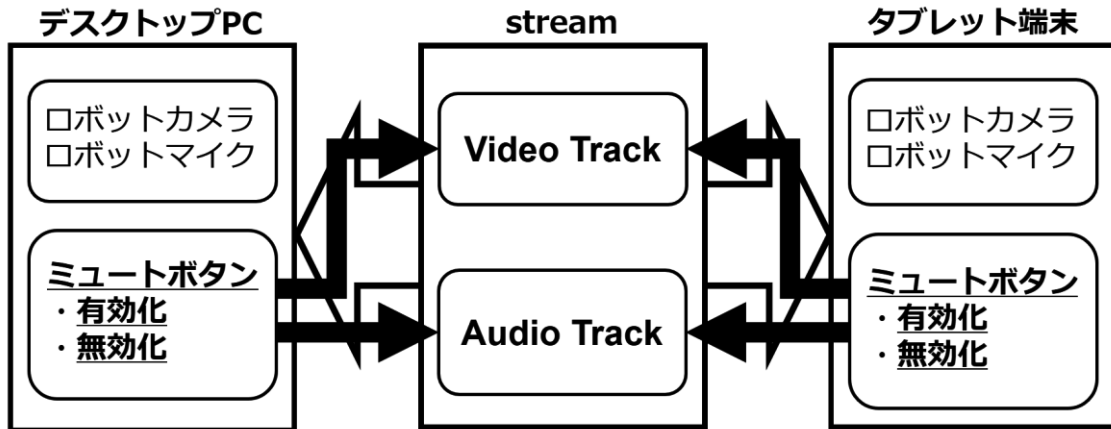


図 5 : ミュートシステムの仕組み

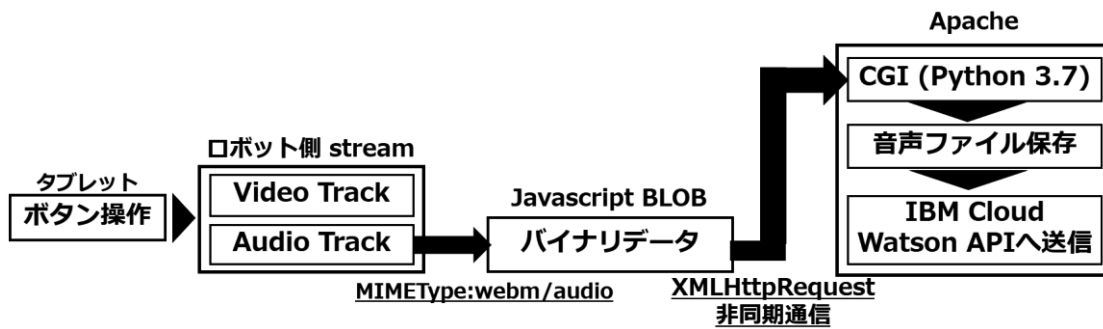


図 6 : 録音システムの流れ

3.2.2 ミュートシステム

ミュートシステムの仕組みを図5に示す。日常業務でビデオ通話システムを使用中、職員側の音声が必要な場合がある。その際にミュートシステムを活用する。ミュートシステムはビデオ通話システムが稼働中に使用できる。ビデオ通話システムが稼働中はビデオトラック（カメラ映像）、音声トラック（マイクから取得された音声）が1つのstreamとして常に相互交換されている。HTML (CSS)のボタン操作によってボタンを押した機器側のstreamからビデオトラックまたは音声トラックを取得し、無効化することでミュート状態とすることができる。開発したミュートシステムではカメラ映像、マイク音声のミュートができる。ミュートを解除する場合はstreamからトラックを取得し、有効化することでミュートの解除ができる。

3.2.3 録音システム

録音システムの流れを図6に示す。日常業務でビデオ通話システムの使用中には、どうしても

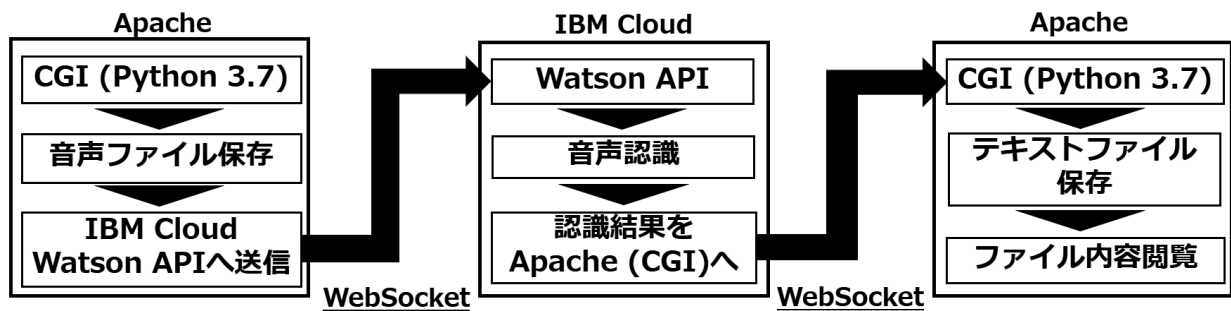


図 7 : 録音音声文字化システムの流れ

メモや記録を行えない状況がある。その際に録音システムを使用する。録音システムはビデオ通話システムが稼働中に使用できる。ビデオ通話システムの使用中にロボット側のマイク音声を保存し、再度聞けるよう録音システムを開発した。録音システムを実現するためにカメラ、マイクから取得したstreamを記録するMediaRecorder APIを使用する。今回の実装ではタブレット端末側からWebサイト内の録音ボタンを押下した場合、録音停止ボタンを押下するまでのロボット側から送信されるstream内の音声トラックを取得する。取得したトラックをMediaRecorder APIによってバイナリデータとしてタブレット端末に保持し、JavascriptのBLOB (Binary Large Object)オブジェクトに格納する。このBLOBオブジェクトをXMLHttpRequest APIの非同期通信によってApacheサーバ内で実行されるCGIプログラム(Python 3.7)に送信する。受け取ったCGIプログラムではBLOBオブジェクト内に格納されたバイナリデータをwebmファイルへ書き出し、デスクトップPC内の特定のディレクトリに保存する。

3.2.4 録音音声文字化システム

録音音声文字化システムの流れを図7に示す。日常業務での使用中には音声だけでなく文字としての記録も行う。録音音声文字化の処理は録音システムによって音声のバイナリデータがwebm形式のファイルへ書き出され、保存された後にApacheサーバ内で動作するCGIプログラムによって動作される。文字化の処理は既存システムを再利用し、システム開発を効率化するため、IBM Cloudサービス(旧Bluemix)によって公開されているWatson API Speech to Textを活用する。録音システムの処理終了後、CGIプログラムによって保存されたwebmファイルを取得し、Watson APIに対するURL (wss://)へWebSocket APIによる通信でIBM Cloud Watson APIに対して音声ファイルを送信する。Watson APIが音声ファイルを認識し、認識結果がCGIプログラムに返される。返された結果をテキストファイルへ書き込み、デスクトップPCの特定のディレクトリに保存する。

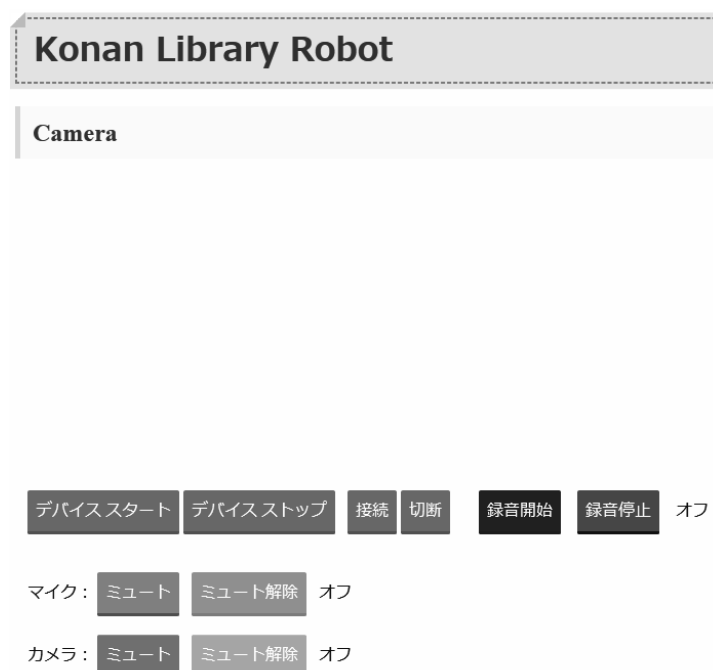


図 8 : システムのレイアウト. カメラ映像取得前の初期状態を示す.

4 実際の業務における使用実験

本章では，図書館ヘルプデスク遠隔対応システムを用いた実業務での検証について説明する．実験方法について述べ，図書館司書による運用結果，ヒアリングの結果を示し，本システムの可能性，有用性を示す．

4.1 実験方法

2018年10月4日から12月31日までの3か月間ヘルプデスクで勤務している4名の図書館司書が実際の業務で使用し，使用状況を記録する実験を行った．図書館の休館日，ヘルプデスクの休業日を除く平日9時から16時30分，土曜日9時から11時30分の営業中のシステム稼働状況を記録の対象とする．図書館司書が操作しやすいよう，2種類のサーバの起動，停止は実行に必要なコマンドをバッチファイルに保存することで任意に操作できるようにした．今回の実験ではロボットに搭載されたマイクに不具合があったため，小型の外付けマイクをデスクトップPCに接続し，音声取得部分を机の上に設置して代用した．使用状況の記録には，バッチファイルに現在日時を取得するコマンドを書き込み，バッチファイルを実行した際の時刻をテキストファイル，csvファイルへ保存することで起動開始時刻，停止時刻を記録する．その停止時刻から起動開始時刻の差を本システムの稼働時間とする．

実装したシステムのレイアウトを図8に示す．デバイススタートボタンの押下によってカメラ

による動画像とマイクによる音声の取得を行う。接続ボタンの押下によってSDPおよびICE candidateの情報が交換され、ビデオ通話機能が実現される。ビデオ通話機能の使用の際、職員の持つタブレット端末側の録音ボタン、録音停止ボタンの押下によってロボット側の音声のみがwebmファイルへ保存される。同時に、IBM Cloud Watson API (Speech to Text)へ保存されたwebmファイルが送信され、認識結果をテキストファイルとして保存する。生成された音声ファイル、音声の認識結果は図書館司書が任意に確認できる。

4.2 実験結果

システムの稼働状況を表1に示す。表のように日数における稼働率は全て1となり、ヘルプデスク営業日は必ずシステムを稼働させていることがいえる。また、10月、11月、12月、全体の使用時間は、156時間4分17秒、169時間43分6秒、127時間27分26秒、453時間14分49秒となり、稼働時間に関する稼働率はそれぞれ1.02、1.08、0.98、1.03と非常に高い。12月にはネットワーク停止が発生し、システムが一時的に使用できなかったものの、ヘルプデスクの総営業時間以上にシステムが稼働していることがいえる。

表 1：システム稼働状況

条件	10月	11月	12月	全体
受付営業日数	23	23	20	66
稼働日数	23	23	20	66
稼働率 (日数)	1	1	1	1
受付営業時間	152 h, 30 m, 00 s	157 h, 30 m, 00 s	130 h, 00 m, 00 s	440 h, 00 m, 00 s
稼働時間	156 h, 04 m, 17 s	169 h, 43 m, 06 s	127 h, 27 m, 26 s	453 h, 14 m, 49 s
稼働率 (時間)	1.02	1.08	0.98	1.03

使用状況の記録に加え、図書館司書に対して使用感に関するインタビューを行った。良かった点として、従来のSkypeシステムよりも映像の視認性とマイクの音声取得性能が向上し、ビデオ通話機能の性能が格段によくなった。サーバの操作とHTML (CSS)のボタン操作についても利用が容易になった。常に受付に在籍する必要がなく、受付外での業務に集中できるとの意見があった。また、ミュート機能の実現によって、業務中でのシステムの利用のしやすさが格段に上がったとの意見もあった。このことからシステムの使用感に関する全体的な評価として非常に良好である。

一方、気になった点としてボタンの数が多く、接客時には瞬時の操作が遅くなってしまう。システムが非常に便利で受付に在籍せず、受付外での業務に集中しすぎることや、不具合による一時的なシステム利用不可の際の不快感といったシステムへの依存性が生じた。

また、不具合として実業務中に長時間のビデオ通話を行っている際、通信が切れていることがあるという報告があった。これにはWebRTCのAPIを使用している際の通信負荷の増加、通信の不安定さが関係していると考えられる。さらに、音声文字化結果について音声と音声の間に空白

時間が長い場合、自動的に音声認識が終了してしまうとの報告があった。これにはWatson APIをパラメータ調整なしのデフォルトの状態で使用していることが考えられる。

アンドロイド・ロボットに関して、大学の生徒や教員など普段の業務で利用している側から、アンドロイド・ロボットに話し掛ける場合は恥ずかしいという抵抗があるものの、最初に話し掛けた後に続けて業務に必要な会話を行う場合はあまり抵抗がなさそうだとの声もあった。この話しかける際の抵抗感を払拭するための工夫として、利用客に対してアンドロイド・ロボット側から話し掛けることや動きを付け加えて注意を引くなど、アンドロイド・ロボット側からの行動を付け加える必要があると考える。

これらのシステムの稼働状況、図書館司書に対するインタビューの結果から提案システムの必要性は非常に高く、大学図書館司書の日常業務の負担軽減につながるため有用であると考えられる。

4.3 従来システムとの比較

提案システムでは、特定の限られた場所でのシステム構築によって通話機能の使用には個人情報の登録やログインが不要となっている。リアルタイムに動画像や音声データの取得、利用が可能となっており、実際にIBM Cloud Watson API Speech to Textの使用によって音声データの文字化が可能となっている。それに加え、HTML、CSS、JavaScript、Pythonといったプログラミング言語での実装によってレイアウトの変更や実業務で使用している中での問題点への早急な対応、新しい機能の追加など自由に機能を拡張することが可能となっている。実験中の使用感としてもSkypeを使用した場合よりもマイクの音声取得や動画像の視認性が上がったとの意見があった。スピーカから流れる音量をプログラムの変更によって変化させるような対応を行うことや、音声データを蓄積させることも可能となっている。このように提案システムではSkypeシステムよりも規模が小さく、専門知識が必要な業務での問題点に対して情報開発による柔軟な解決が可能であり拡張性が高いことがいえる。

5 おわりに

本稿では、アンドロイド・ロボットを用いた図書館ヘルプデスクにおける大学図書館司書のための遠隔対応システムを構築した。本システムはアンドロイド・ロボットを用いたHTML上でのリアルタイムコミュニケーションシステムを開発した。リアルタイムコミュニケーションシステムとして、ビデオ通話システム、ミュートシステム、録音システム、録音音声文字化システムを開発した。バッチファイルによるサーバの起動停止操作によって情報分野の知識量の少ない図書館司書にとって利用しやすい仕組みとなっている。従来ではSkypeを使用したシステムを構築していた。その点について、外部アプリケーションを使用した場合よりも拡張性が高く、専門業務での利用のしやすさ、柔軟な対応が可能となっている。構築したシステムを実業務で活用することでヘルプデスク遠隔対応システムの必要性、有用性を示した。

今後の課題として、図書館司書のための開発としてはHTML (CSS)画面のボタン数の減少、データベースを作成し、音声データや文字化したデータを確認する際の手軽さの向上、Watson API

Speech to Textの改良, 収集したデータを活用することによってロボット単体での接客システムの構築があげられる. 利用者のための開発としてはアンドロイド・ロボットが人間に対して受け入れられやすい状況を作り出すためにアンドロイド・ロボット側からの発声や動きを加えるなどといった注意誘導を行うようなシステムの実装があげられる.

謝辞

本研究の一部は, 科研費(JP17K06280), 私立大学経常費補助金特別補助「大学間連携等による共同研究」, 甲南大学KONANプレミア・プロジェクトの支援を受けました. また, アンドロイド・ロボットを用いた実験, 評価において, 甲南大学図書館ヘルプデスク職員の皆様より多大なご協力をいただきました. ここに謝意を表します.

参考文献

- [1] 怡土順一, 松本吉央, 西村竜一, 李晃伸, “情報科学研究の実環境プラットフォームとしての受付案内ロボット ASKA,” 日本機械学会ロボティクス・メカトロニクス講演会’02講演論文集, no. 1A1-D12, CD-ROM, 2 pages, 2002.
- [2] 岡田航大, 吉岡一樹, 和田昌浩, 田中雅博, “図書館入口における案内ロボット KoRo,” 日本機械学会ロボティクス・メカトロニクス部門講演会2016講演論文集, no. 1A1-14b1, DVD-ROM, 3 pages, 2016.
- [3] 岡田航大, 田中雅博, 和田昌浩, 楠和也, 大塚慶太, 中田悠貴, 中村歩, 高田尚真, “案内ロボット KoRoの開発と機能,” 甲南大学紀要知能情報学編, vol. 10, no. 2, pp. 155-184, 2017.
- [4] 立花隆輝, “Watsonとロボットの音声対話機能,” 日本ロボット学会誌, vol. 35, no. 3, pp. 199-202, 2017.
- [5] 菊地智也, 梅谷智弘, 才脇直樹, “アンドロイドロボットを用いた大学図書館のためのヘルプデスク遠隔対応システムの検討,” 第62回システム制御情報学会研究発表講演会講演論文集, no. 346-5, CD-ROM, 3 pages, 2018.
- [6] 菊地智也, 梅谷智弘, 才脇直樹, “アンドロイド・ロボットを用いた大学図書館ヘルプデスクの遠隔対応時記録システムの開発,” 第63回システム制御情報学会研究発表講演会講演論文集, no. TS09-2-2, pp. 835-838, 2019.