

音を用いたカジュアルな遠隔コミュニケーションの拡張

綿谷 晃太郎^{1,a)} 塚田 浩二¹

概要: Skype や Zoom 等に代表されるビデオ会議システムは、会議等のフォーマルな場面だけでなくオンライン飲み会などのカジュアルな場面で利用されることが増えてきた。こうした場面では、細かな感情や雰囲気が共有されにくい、自宅から参加できることから辞め時が難しいといった遠隔コミュニケーションの特性や社会的な問題が絡み合って様々な課題が生じてくる。本研究では、オンライン飲み会のような複数人でのカジュアルな遠隔コミュニケーション場面において、参加者の身振りや行動、雰囲気をきっかけに効果音や BGM などのアンビエントな情報を付加することで、上述したような問題を軽減しつつ、魅力的な体験を生み出すことを目指す。こうしたカジュアルな遠隔コミュニケーション場面での効果音付与の可能性を探るために、Zoom 等のビデオ会議中に事前登録した任意の音を再生できる GUI アプリケーションを実装した。このアプリケーションを用いた評価実験を通して、効果音等の印象やコミュニケーションへの影響を調査した。さらに、評価実験の結果に基づいて、システムの課題や展望について議論する。

Extending Casual Remote Communication with Sound

WATAYA KOUTARO^{1,a)} KOUJI TSUKADA¹

1. 序論

1.1 背景と目的

昨今のコロナウィルスの流行や技術の発展などから、Skype や Zoom 等に代表されるビデオ会議システムは、会議等のフォーマルな場面だけでなくオンライン飲み会などのカジュアルな場面で利用されることが増えてきた。こうした場面では、議題等がなく進行が不規則/少数グループが頻繁に入れ替わるといったカジュアルなコミュニケーションの特性に加えて、複数人が同時に発言するのが難しい/細かな感情や雰囲気が共有されにくいといった遠隔コミュニケーションの問題や自宅から参加できることから辞め時が難しい/参加を断りにくいといった社会的な問題が絡み合ってくる。

一方、カジュアルな遠隔コミュニケーションを支援する手法として、センサやディスプレイを内包したインタフェースを活用する研究が行われてきた。こうした研究では、一対一、又は少人数で相手の存在感やアウェアネスを

さりげなく共有することに焦点が置かれることが多かった。

また、本研究では音とコミュニケーションの関係に着目した。これまでの研究では、音楽を聴くことによって人間の感情や行動が変容することなどが知られており、それを利用することで音楽がコミュニケーションの促進に効果があると考察されている [1][2]。音楽が利用される身近な例としては、カフェなどの飲食店では雰囲気づくりのために BGM をかけているところも多くみられる。

本研究では、オンライン飲み会のような複数人でのカジュアルな遠隔コミュニケーション場面での問題点を軽減することを目的とする。参加者の身振りや行動、場の雰囲気を、効果音や BGM といったアンビエントな情報を付加し強調することで、問題の軽減を図る。

1.2 論文構成

ここでは本論文の構成を示す。第 2 章「関連研究」では、本研究と関連のある研究として、「音が人の行動や感情に与える影響に関する研究」、「遠隔コミュニケーションを支援する研究」について紹介する。第 3 章「提案システム」では、本研究で提案するシステムのコンセプトと構成につ

¹ 公立はこだて未来大学
Future University Hakodate

^{a)} g2121063@fun.ac.jp

いて述べる。第4章「実装」では、提案システムとして作成した音提示プログラムの構成や実装について述べる。第5章「評価実験」では、本研究の提案システムに実装する音の候補を選定し、遠隔コミュニケーション中に実際に提示してその時の印象等を調査し考察する。第6章「議論」では、評価実験の内容を踏まえ、本研究のシステムの課題について述べる。第7章「結論と展望」では、本研究のまとめと展望について述べる。

2. 関連研究

本章では、本研究に関連する研究事例として、「音が人の行動や感情に与える影響に関する研究」、「遠隔コミュニケーションを支援する研究」の2つの観点から説明する。

2.1 音が人の行動や感情に与える影響に関する研究

高口 [1] は、テンポの違う二つ BGM を使った実験により、BGM の違いが人の行動にどのような違いを生むかを検討した。実験結果から、テンポの遅い BGM はテンポの速い BGM の時に比べゆっくりとした行動を促すことを確認した。

大野ら [2] は音楽のもつリラックス効果を利用して、会話者それぞれが独立して音楽を聴取しながらコミュニケーションをはかるといった緊張緩和手法の提案を行い、実験によってコミュニケーションに及ぼす影響を調査した。その結果、音楽を聴取することによってコミュニケーション時の緊張が緩和され、さらに会話がしやすくなる傾向にあることを明らかにした。

高橋ら [3] は、音楽が情動にどのような影響を及ぼすかを明らかにするために、主観的な情動の指標である POMS (Profile of Mood State) を用いた実験を行った。その結果、実験者側で選択した鎮静的音楽を聴取した場合と、リラックスするために被験者が選んだ音楽 (好みの音楽) を聴取した場合のどちらでも情動が変化することを明らかにしている。

2.2 遠隔コミュニケーションを支援する研究

InPhase[4] は、遠隔地にいる相手とたまたま同じ時間に同じ行動をする「偶然の一致」をチャイム音などの効果音をならすことで相互に知ることができるシステムである。これによって、離れた場所にいながら同調効果を与え合うことや話題のきっかけを生むことができる。

また、Lover's cups[5] は遠隔地にいる相手がコップ型インターフェースを使って飲み物を飲むとき、自分側のコップ型インターフェースが光ることによってその瞬間に把握できるというものである。これによって離れた場所にいる相手と「飲み物を飲む」という時間を共有することができる。

LumiTouch[6] では、一方の写真立てのふちに触れると、

もう一方の写真立てのふちが点灯する。これによって、離れた場所にいるカップルの感情を交換し、共有できる。

Limonect[7] はマットに人が乗ったことがきっかけとなり、使用者の足跡をセンシングし、周囲の環境音を録音し始める。そして、足跡をその場の生活音と共に遠隔に伝え始める。実際にデバイスを使用し検証したところ、遠隔地の家族とまるでともに暮らしているかのような感覚を得ることができたとしている。

SyncDecor[8] は、家具や調度品を遠隔地に設置し、その動きを連動させることで相手の行動や雰囲気さをさりげなく伝えるものである。SyncDecor の中でも、SyncLamp は遠隔地にある2つの調光電気スタンドをネットワークで接続し、一方の明るさを調整した時に、もう一方のものの明るさも同じ明るさに調節される。これによって、帰ってきたことや寝るといった情報を互いにさりげなく伝えることができる。

SakeEffects[9] は、アルコールセンサを内蔵したマイクを試作し、オンライン飲み会中にアルコールセンサからの情報により酔い度を評価しユーザーの顔にビデオエフェクトを合成するというシステムを提案した。試用の結果、ビデオエフェクトの変化によって話題が広がったり盛り上がりが増すことができていた。

これらの研究の共通点は、遠隔地にいる相手の行動を光や音といったアンビエントな情報によって提示することで、さりげなく相手が今どのような状態であるかを認識でき、物理的に離れている相手とのコミュニケーションを補助しているという点である。

2.3 本研究の特徴

本研究ではこれらの研究のように、アンビエントな情報として BGM や効果音を用いて遠隔コミュニケーションの円滑化を図るが、オンライン飲み会等のカジュアルなコミュニケーションを対象として、ビデオ会議システムなどと併用する点がこれらの研究と異なっている。

3. 提案システム

本章では、本研究で提案するシステムのコンセプトと構成について説明する。

3.1 コンセプト

本研究では、カジュアルな遠隔コミュニケーション中に参加者が音を任意に提示できるシステムを作成し、細やかな感情や雰囲気が共有されにくい問題と話者の交代ややめ時が難しいといったきっかけの問題を軽減したいと考える。

細やかな感情や雰囲気が伝わりにくい問題を軽減するために、特定の動きやジェスチャ、表情に合わせて適切な効果音を挿入する。提案例として、コミュニケーションの最初に乾杯をしたいとき、画面に向かってグラスを突き出す

動きとともにグラスのぶつかり合う音を挿入することで、乾杯をしたときの雰囲気をより感じられるような効果があると考えられる。

話者の交代ややめ時が難しいといったきっかけの問題を軽減するために、きっかけとしてBGMを挿入することで自然とアクションの流れを作ったり、会話の盛り上がり具合でBGMを挿入・変更することで、コミュニケーションの盛り上がり具合を変化させたいと考えた。

将来的には、マイコンでのセンシングや画像・音声認識を利用し、参加者の身振りなどをきっかけに音を挿入するシステムに拡張する予定であるが、本稿ではまず基本的な音提示システムを構築する。

3.2 システム構成

システム構成図を図1に示す。提案システムはzoomなどの一般的なビデオ会議システムに対して、Botとして参加させて使用する。システムは大きく、認識部と生成部から構成される。認識部はユーザーの操作を認識する。現在実装が完了しているのは、Bot 役のユーザーを想定した「手動入力アプリ」のみであるが、議論の章で後述するように、今後音声認識システム／画像認識システム等の実装も検討している。生成部では、認識部からの入力や認識結果に基づき、事前に設定した効果音やBGMを生成する。生成された効果音等は、後述するようなループバック処理を通して、Bot のマイク入力としてビデオ会議システム中で再生される。

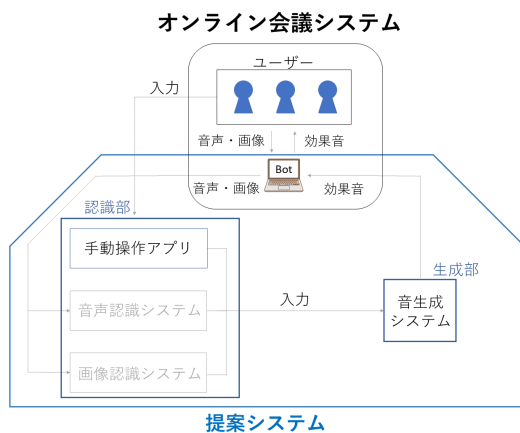


図1 システム構成図

前提として、ビデオ会議システムの使用時に各参加者の声（マイク入力）は共有されるが、コンピュータの音は共有されない。使用するビデオ会議システムが画面共有機能を持つ場合は、コンピュータの音を共有できる場合があるが、複数人同時に共有を行うことはできない。しかし、本システムの使用時は参加者全員がコンピュータの音を共有している必要がある。この問題は、Bot を用意し遠隔コミュニ

ケーションに参加させ、Bot を動かすコンピュータで音のルーティングを行うことで解決できる。そこで、コンピュータの出力音をマイク入力に統合する手法として、ループバック処理を行う。ループバック処理のイメージを図2に示す。出力先をビデオ会議システムに設定することで参加者それぞれのコンピュータの音を同時に共有することが可能になる。今回は開発環境が Windows 10 であったため、Windows 環境に対応している VoiceMeeter Banana*1 というフリーのソフトウェアミキサーを使用する。

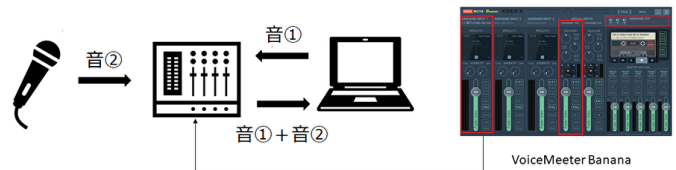


図2 ループバック処理のイメージ

4. 実装

4.1 ループバック処理の事前設定

ループバック処理を行うための、Windows や VoiceMeeter Banana, ビデオ会議システムの設定について紹介する。VoiceMeeter Banana のインストール時は特に設定項目がないため割愛する。

まず初めに、Windows のサウンド設定が必要である。「設定→システム→サウンド」の設定画面から、「出力」の部分を「VoiceMeeter Input」に変更し、同様に「入力」の部分を「VoiceMeeter Output」に変更する。

次に、サウンド設定画面の関連設定から「サウンドコントロールパネル」を開き、「既定の通信デバイス」を設定する。具体的には、「再生」タブから「VoiceMeeter Aux Input」を、「録音」タブから「VoiceMeeter Aux Output」を選択する。

次に VoiceMeeter Banana の設定を行う。VoiceMeeter Banana 起動時の画面を図3に示す。まず、画面右上の「HARDWARE OUT」の「A1」（図3赤枠部）をクリックし、遠隔コミュニケーション中に音を再生する機器（スピーカーやイヤホン）を選択する。次に、画面左上の「HARDWARE INPUT 1」（図3青枠部）をクリックし遠隔コミュニケーション中に使用するマイクを選択する。次に、遠隔コミュニケーション中に流す音のバランス等を設定する（図3黄枠部）。図3の黄色枠部ではそれぞれ、左から自分のマイクの音、コンピュータの音、他の参加者の音の音量やルーティングを設定できる。スライダーで音量を調節し、「A1」から「B2」のボタンでルーティングを設定する。「A1」は自分を、「B2」は他の参加者を表している。今回は

*1 <https://vb-audio.com/Voicemeeter/banana.htm>

図3のように自分の声を「B2」、コンピュータの音を「A1」と「B2」、他の参加者の音を「A1」に設定する。



図3 VoiceMeeter Banana の説明

最後にビデオ会議システムの設定を行う。使用するマイクとして、「VoiceMeeter Aux Output」を選択し、使用するスピーカーとして「VoiceMeeter Aux Input」を選択する。

4.2 音提示プログラム

本システムは Processing を用いて作成した。プログラム起動時の状態を図4に示す。このバージョンだと手動入力と音生成が一体化しているが、今後切り分けていく予定である。



図4 音提示プログラム起動時

画面の構成として、上部左側から、画面の拡大・縮小ボタン、セットを切り替えるボタン、音声ファイル設定ボタン、音を止めるボタンとなっている。その下に効果音ボタンが10個、BGMボタンが5個あり、さらに下にアラームとタイマーを配置している。以下各機能の詳細について述べる。

画面の拡大・縮小ボタンは、押すごとにウィンドウの大きさが切り替わる。本システムはビデオ会議システムと同時に使用することを想定しているため、起動時のサイズでは大きすぎると考え、画面の占有率を下げるためにこの機

能を実装した。縮小後のサイズは Processing で表示できる最小サイズの100×100ピクセルとし、音は後述するようにキーボード操作で再生できる。

「SET1」「SET2」「SET3」の3つのボタンを押すことで、効果音が10種類、BGMが5種類の計15種類の音声セットを3セット作れるように設計した。コミュニケーション場面（オンライン飲み会、懇親会など）に合わせて複数の音声のセットを組めるように設計した。音声セットを切り替えた状態を図5に示す。



図5 セット2へ遷移後の状態. セット1にに対し、背景色と赤枠内部が変化している

CONFIG ボタンを押すと音ボタンに任意の音を割り当てる設定画面に切り替わる。設定画面で各音ボタンを押すとエクスプローラーが開かれ、音声ファイルを選択することでボタンに音を設定することができる。この時、ボタンの文字が選択したファイル名に切り替わる。各音ボタンに音を設定した後、もう一度 CONFIG ボタンを押すこと通常画面に戻る。

STOP ボタンを押すことで、鳴っている音を止めることができる。音ボタンで鳴らした効果音やBGMだけでなく、タイマー・アラーム機能で鳴らした音も STOP ボタンで止めるよう想定している。キーボードの「s」キーを押すことでも同じく音を止められる。

効果音、BGMの各音ボタンを押すことで、それぞれのボタンに設定した音を鳴らすことができる。なお、ボタン名が無いものを押しても音はならない。システム起動時は前回の設定がそのまま反映される。効果音ボタンの左側の数字はキーボードの0から9キーに対応しており、キーボードからでも音を鳴らすことができる。

システム画面の下部にはタイマーとアラームをそれぞれ実装した。タイマーはセットした時間分カウントダウンし、数字が0になった時に音が鳴る。アラームはセットした時刻になった瞬間音が鳴る。「タイマー」「アラーム」という表記の横のボタンを押すことで、それぞれの機能で鳴らす音を設定するための画面に切り替わる。設定画面に切り替わった後、効果音ボタン/BGMボタンを押すことで、該当する音がタイマーの音になる。音の設定後、もう一度「タイマー」「アラーム」という表記の横のボタンを押すと通常画面に切り替わる。それぞれの機能を開始する場

合は、時間表示部分横の ON/OFF ボタンを ON に切り替えることで開始でき、途中で止めたい場合は OFF に切り替えることで止められる。

5. 評価実験

本章では、音提示の印象評価の目的、手法、結果及び考察について述べる。

5.1 目的

ビデオ会議システムを用いたカジュアルな遠隔コミュニケーション中に、本システムを用いて効果音等を提示し、効果音等の印象やコミュニケーションへの影響を調査した。

5.2 手法

4名を1グループとし、合計12人、20分程度ビデオ会議システムを用いた遠隔コミュニケーションを行ってもらい、実験者が様子を確認しながら提案システムを使用して音を提示した。提示した効果音は「拍手」、「乾杯」、「正解」、「不正解」、「挙手」、「クイズ出題」の6種類とした。提示したBGMは、「テンポが速めで活発な印象のもの（以降『活発なBGM』）」、「テンポが遅めでゆったりとした印象のもの（以降、『ゆったりBGM』）」、「蛍の光」の3種類とした。被験者には行動等に合わせて効果音が鳴ることを知らせなかった。なお、効果音の提示は全て実験者がシステムを用いて行った。また、実験者はコミュニケーション自体には参加しなかった。

遠隔コミュニケーションのシチュエーションとしては、オンライン飲み会や懇親会等のカジュアルな場面を想定し、参加者に以下のようなシナリオを伝えた。

前半約10分はお互いの近況報告会として、一人ずつ近況報告（最近取り組んでいること、面白かったこと、頑張っていることなど）をし、それについて質問等があれば自由に行うように指示した。この時、前半の2人が発話している間は「活発なBGM」を、後半の2人が発話している間は「ゆったりBGM」を提示した。1人の近況報告が終わった時点で、全員が「拍手」をしてから次の参加者に交代するよう指示し、「拍手」の動作に合わせて効果音を提示した。

後半約10分はレクリエーションとしてクイズ大会をするよう指示し、クイズの出題者を1人指定し残りの3人を回答者とした。この時、出題者には各回答者の回答が正しいか間違っているかをジェスチャー付きで答えるよう指示し、正解のジェスチャー時に「正解」の効果音を、不正解のジェスチャー時に「不正解」の効果音をそれぞれ提示した。また、出題前に「問題」と発言してから出題するよう指示し、出題者が「問題」と発言した直後に「クイズ出題」の効果音を提示した。回答者には、解答する前に手を挙げるよう指示し、手を挙げる動作に合わせて「挙手」の効果音を提示した。

表 1 効果音についてのアンケート調査結果（平均値（標準偏差））。

| | 音量の適切さ | 場面の適切さ | 必要性 |
|-------|------------|------------|------------|
| 拍手 | 3.42(0.95) | 4.08(1.04) | 3.58(0.86) |
| 乾杯 | 2.83(1.21) | 4.25(1.30) | 3.50(1.44) |
| 正解 | 3.42(0.49) | 4.75(0.60) | 4.00(1.22) |
| 不正解 | 3.58(0.76) | 4.58(0.64) | 4.08(0.86) |
| 挙手 | 3.33(0.62) | 3.67(1.18) | 3.25(1.23) |
| クイズ出題 | 3.83(0.69) | 4.67(0.47) | 4.17(0.80) |
| 全体 | 3.40 | 4.33 | 3.76 |

さらに、遠隔コミュニケーション開始時から19分程たった時点で「蛍の光」を提示し、コミュニケーションの終了を暗に促すようにした。被験者には実験の説明の時点で、終了のタイミングの合図を何かしらの形で提示するとだけ伝えた。

遠隔コミュニケーション終了後、アンケートを行い、提示した音やその効果等の印象について調査した。

5.3 結果と考察

効果音についての結果を表1に示す。音量の適切さを(1. 小さすぎる～5. 大きすぎる)、場面の適切さと必要性を(1. 当てはまらない～5. 当てはまる)の5段階でそれぞれ評価してもらった。

音量の適切さを調査した結果、全体の平均値は3.40となった。この結果から、全体としてはやや音量が大きかったと感じていたことがわかる。各効果音ごとに見ると、ほぼすべての効果音について平均値が3以上であったのに対し、「乾杯」の効果音のみ2.83と3を下回った。この結果の要因としては、「乾杯」の効果音のみ著者が複数の効果音を組み合わせで作成したものであったため、加工時に音量が下がってしまった可能性が考えられる。

場面の適切さを調査した結果、全体としての平均値は4.33となった。この結果から、全体としては場面にそぐわない音はなかったと言えるが、他の効果音の平均値が4を超えているのに対し、「挙手」の効果音のみ3.67であった。また、「挙手」に対する自由記述に「音自体は聞こえたがなったときはなんのための音か分からなかった」「正解ともう少し区別をつけたほうが良いと感じた」「正解の音を間違えて鳴らしたのかと思った。」などの記述があった。これらのことを踏まえると、本実験で「挙手」の効果音として選定した音は、「挙手」のイメージに合っておらずわかりにくかったと考える。

必要性について調査した結果、全体としての平均値は3.76となった。全体としては、効果音があった方がいいという意見が多かったと言える。中でも「正解」「不正解」「クイズ出題」の3つが4を超えた。評価が高かった3つの効果音の理由として、「クイズ番組みたいで楽しかった」「楽しい雰囲気を作れていた」といった意見が見られた。しかし、標準偏差を見ると「乾杯」「正解」「挙手」の効果音

表 2 「正解」「不正解」「クイズ出題」のみの結果の抜粋 (平均値 (標準偏差)).

| | 場面の適切さ | 必要性 |
|-------|------------|------------|
| 正解 | 4.75(0.60) | 4.00(1.22) |
| 不正解 | 4.58(0.64) | 4.08(0.86) |
| クイズ出題 | 4.67(0.47) | 4.17(0.80) |
| 全体 | 4.67 | 4.08 |

表 3 BGM についてのアンケート調査結果 (平均 (標準偏差)).

| | 活発 | ゆったり |
|------------|------------|------------|
| どれくらい聞こえたか | 4.67(0.62) | 4.67(0.62) |
| 楽しさの変化 | 3.92(0.76) | 3.75(0.72) |
| 緊張の具合の変化 | 3.33(1.03) | 3.92(0.86) |
| 盛り上がりの変化 | 3.67(0.85) | 3.08(0.76) |

は値が 1.22~1.44 と高い値であり、個人差が大きい結果であった。

ここで「正解」「不正解」「クイズ出題」について「場面の適切さ」「必要性」のみ抽出した結果を表 2 に示す。場面の適切さの平均値は全体で 4.67、必要性は 4.08 とどちらも高い値となった。この 3 つの効果音の共通点として、「クイズ大会」で使用する効果音として用意した音という点が挙げられる。このことから、限定されたシチュエーションで、その場面に適した効果音を鳴らすことができたため、他に比べて評価が高かったのではないかと考える。

BGM についての結果を表 3 に示す。どれくらい聞こえていたかを (1. 全く聞こえなかった~5. しっかり聞こえた)、楽しさに変化があったかを (1. つまらなくなった~5. 楽しくなった)、緊張の具合に変化を感じたかを (1. 緊張した~5. リラックスした)、雰囲気 of 盛り上がりに変化を感じたかを (1. 盛り下がった~5. 盛り上がった) の 5 段階でそれぞれで評価してもらった。

BGM についてのアンケート結果について述べる。

「活発な BGM」に対し、楽しさの変化が平均 3.92 (標準偏差 0.76)、盛り上がりの変化が平均 3.67 (標準偏差 0.85) であったことから、楽しさや盛り上りを演出することができたこと考える。また、自由記述からも「聞いてて楽しくなった」「話しているときに (BGM が) かかると、面白い話ができると感じてしまう」といったように、楽しさや盛り上がりを感じるような記述が見られた。

一方、「ゆったり BGM」に対し、楽しさの変化が平均 3.75 (標準偏差 0.72)、緊張具合の変化が平均 3.92 (標準偏差 0.86) であったことから、楽しさを増しつつ、ゆったりした雰囲気を演出しリラックスさせることができたこと考える。自由記述では、「ふわふわした雰囲気が出た」、「前の (活発な BGM が) かかっていた) 時よりも会話のテンポがゆっくりになったように感じた」「気の抜けた感じ」といったものが見られ、ゆったりした雰囲気を感じるような記述が見られた。また、どれくらい聞こえていたかに対する評価の平均値がどちらの BGM でも 4.67 であった。BGM

表 4 「蛍の光」に対するアンケート調査結果。

| | 平均 (標準偏差) |
|--------------|------------|
| どれくらい聞こえていたか | 4.33(1.18) |
| 終了のきっかけになるか | 4.67(0.85) |

は十分聞こえていた一方で、「話が入ってきづらかった」、「音量が少し大きかった」といった記述が見られたため、BGM としては音量が大きすぎた可能性がある。

「蛍の光」についてのアンケート調査結果を表 4 に示す。アンケートでは、どれくらい聞こえていたかを (1. 全く聞こえなかった~5. しっかり聞こえた)、終了のきっかけになるかを (1. ならない~5. なる) の 5 段階で評価してもらった。

どれくらい聞こえていたかに対する評価の平均値が 4.33 (標準偏差 1.18) から、個人差はあったもののしっかり聞こえていたとわかった。また、終了のきっかけになるかに対する評価の平均値が 4.67 (標準偏差 0.85) から、肯定的な意見が多いことが分かった。

「蛍の光」を聞いたときの感想として、「終わる感じ」「この会はおしまいだと感じた」「びっくりしたあと、終了なんだと気づいた」といったものがあった。

これらの結果から、「蛍の光」は遠隔コミュニケーションを終わらせるきっかけとして十分な効果を期待できるのではないかと考えた。

5.4 まとめ

効果音に関する結果についてまとめる。場面の適切さの項目に対し評価が高かった効果音は、「拍手」、「乾杯」、「正解」、「不正解」、「クイズ出題」であった。必要性の項目に対し評価が高かった効果音は、「正解」、「不正解」、「クイズ出題」であった。評価の高かった効果音に対し、「楽しい」、「嬉しい」といったポジティブな意見が多くみられたことも踏まえると、特定の場面で楽しさを増すことができたこと考える。

BGM についてまとめると、楽しさという項目で評価が高かったものは「活発な BGM」、「ゆったり BGM」の両方であった。緊張具合の項目に対し評価が高かったものは「ゆったり BGM」であった。盛り上がり具合の項目に対し評価が高かったものは「活発な BGM」であった。これらの結果や、自由記述でのポジティブな意見、「蛍の光」が終わりのきっかけになりうることに多くの肯定的な意見があったことから、コミュニケーションの雰囲気に変化をつけることができていたと考える。一方で、「活発な BGM」「ゆったり BGM」の両方で「音量が大きい」のような意見も見られたため、音量の調節を今後の課題とする。

6. 議論

本章では、本システムの課題と改善点について述べる。

6.1 音量設定

遠隔コミュニケーションを音で支援するにあたって、提示する音の大きさは重要である。今回の調査でも、それぞれの音の大きさを調整したいと感じる場面が多々あった。しかし、現状のシステムでは音の大きさの設定を個々に行うことはできない。したがって、各効果音・BGMに対し個々に音量設定ができる機能を実装する。

6.2 Bot 機能の実装

今回の実験を行うにあたって、数名に対しループバック処理を行ってもらおうと予備的な検証を行ったが、正しく設定したつもりでも音が出ないといったアクシデントが多々発生したため、Bot を用意する必要性が改めて分かった。Bot を介してユーザーと提案システム間で音情報をやり取りすることで、Bot を動かしている端末のみがループバック処理を行うだけで、参加者全員が効果音を聞くことができる。

6.3 ジェスチャーなどをきっかけとした音の挿入

3章でも述べたように、将来的には画像・音声認識を利用し、参加者の身振りなどをきっかけに音を挿入するシステムに拡張する予定である。

手動入力と自動挿入機能は両方の機能を用意し、認識対象の動作等の発生頻度や、(システムでの)検出精度、音の性質等を考慮し、運用を通して調整していきたい。

音認識と画像認識については、ビデオ会議システムとの連携の関係上、音情報の方が扱いやすいため、音から検出可能な情報(例:拍手や笑い声)は音認識を利用し、音が生じないジェスチャー(例:両腕で丸を作る動作や拳手)等に対して画像認識を利用する予定である。

こうした画像・音声認識を利用したシステム拡張として、Google が提供している機械学習ツールである Teachable Machine を利用した予備的な試作を行っている。

Teachable Machine は画像と音声の双方で、機械学習の分類モデルを手軽に構築できる。また、作成したモデルを外部プログラム(例:p5.js)に組み込むことで、Web 上で動作する音/画像認識システムを構築できる。作成したプロトタイプを Bot を動かしている端末内で起動し、遠隔コミュニケーション中の音声・画像情報をアプリに取り込み効果音を挿入する。

現在、オンライン会議中の音(拍手等)を認識して効果音を発生させる機能が(技術的に)実現できることを確認したが、精度の点等でまだ課題がある。今後は多様な音声/音響認識ライブラリ(例:Google Speech to Text, SpeechRecognition)の利用も検討していきたい。

一方、ビデオ会議の映像を抽出することが出来ていないため、画像認識システムはまだ未実装である。今後、Zoom SDK といった各ビデオ会議システムの拡張機能などにつ

いて調査を進めていく。

7. 結論と展望

本研究では、zoom などの一般的なビデオ会議システムと同時にしようすることを想定して効果音やBGMといったコンピュータの音を複数人同時に共有することを可能にし、画面またはキーボード入力で任意の音を提示できるシステムを作成した。さらに、提案システムを使用して遠隔コミュニケーション中に数種類の効果音・BGM を提示し、その印象などについて調査を行った。その結果、効果音は特定の場面での楽しさを増すことができ、BGM はコミュニケーションの雰囲気を変化させることができた。

展望として、音の自動挿入機能をシステムに搭載し、参加者が意識しなくてもその場に合った音をコミュニケーション中に鳴らせるように、現状のシステムを発展させていきたい。また、カジュアルな場面だけでなく、会議や授業といったフォーマルな場面でも活用できるように調査、改善を進めていくことも検討したい。

参考文献

- [1] 高口 央. BGM は購買意欲をかき立てるか., 流通経済大学社会学部論叢, 2011, 22.1: p.89-100.
- [2] 大野 直紀, 徳久 弘樹, 中村 聡史. 自身のみ聴取可能な音楽を用いたコミュニケーション円滑化手法の提案, 情報処理学会 研究報告グループウェアとネットワークサービス (GN), Vol.2019-GN-106, No.22, pp.1 - 7, 2019.
- [3] 高橋幸子, 山本賢司, 松浦信典, 伊賀富栄, 志水哲雄, 白倉克之. 音楽聴取が情動に与える変化について:音楽聴取前後の POMS スコアの変化を中心として. 心身医学, Vol.39, No.2, p. 167-175, 1999.
- [4] 辻田 眸, 塚田 浩二, 椎尾 一郎. InPhase: 日常の偶然の一致に着目したコミュニケーションシステムの提案, コンピュータ ソフトウェア, 2010, 27 巻, 1 号, p. 1.18-1.28,
- [5] Hyemin Chung, Chia Hsun Jackie Lee, Ted Selker, Lover's cups: drinking interfaces as new communication channels, CHI EA '06: CHI '06 Extended Abstracts on Human Factors in Computing Systems, April 2006, Pages 375-380.
- [6] A. Chang, B. Resner, B. Koerner, X. Wang, and H. Ishii. LumiTouch: an emotional communication device. CHI EA '01: CHI '01 Extended Abstracts on Human Factors in Computing Systems, pp. 313-314, New York, NY, USA, 2001.
- [7] 郡山和彦, 小泉真理子, 大澤公美子, 奥出直人. Limonect: 離れて暮らす家族のアンビエントコミュニケーション, インタラクション 2007, 2007
- [8] 辻田眸, 塚田浩二, 椎尾一郎:遠距離恋愛者間のコミュニケーションを支援する日用品” SyncDecor” の提案; 日本ソフトウェア科学会論文誌 (コンピュータソフトウェア), Vol.26, No.1, pp.25-37 (2009)
- [9] 長井彩乃, 相澤裕貴, 須田翔悟, 渡邊恵太. SakeEffects: アルコールセンサを用いたオンライン飲み会を盛り上げるための顔エフェクトシステム, インタラクション 2021 論文集, 2021