

HappyMouth: マスク型デバイスによる 対面コミュニケーション能力の拡張

石井綾郁^{†1} 小松孝徳^{†1} 橋本直^{†1}

概要: 本研究では、人間の対面コミュニケーション能力を拡張するマスク型デバイス「HappyMouth」を提案する。本システムでは、マスクに小型ディスプレイが内蔵されており、口元での映像提示を行うことができる。映像提示の機能として、ユーザが自分の好みの口を選択して表示する機能、ユーザの発話をテキスト化して字幕表示する機能、ユーザの発したキーワードをインターネットで画像検索した結果を表示する機能がある。提案システムで口の画像を表示した装着者に対してどのような印象を持つのかを評価した結果、提案システムの装着者は一般的なマスクの装着者に比べて魅力的で積極的な印象を受け、あたたかさを感じさせるということがわかった。また、提案システムの装着者より一般的なマスクの装着者の方が素っ気ない印象を受けるということがわかった。字幕表示機能および画像検索・表示機能については著者自身の主観評価を行い、提案システムの可能性について検討した。

キーワード: コミュニケーション, 人間拡張, ウェアラブルデバイス

HappyMouth: Augmentation of Face-to-Face Communication Ability with a Mask-type Device

AYAKA ISHII^{†1} TAKANORI KOMATSU^{†1} SUNAO HASHIMOTO^{†1}

Abstract: In this research, we propose a mask-type device "HappyMouth" which extends human face-to-face communication ability. This device consists of a mask and a small display, and it is possible to present images at the mouth. As a function of video presentation, this system has a function that the user selects and displays his/her favorite mouth, a function to display the user's utterance as a text, and a function to display the result of image search by keyword issued by the user. As a result of evaluating the impression for the wearer of the proposed system, we found that the wearer of the proposed system received an attractive, positive and warm impression compared with the wearer of the general mask. We also found that the wearer of the general mask received a drier impression than the wearer of the proposed system. For the function to display subtitles and the function to search and display images, we made a subjective assessment of the author and examined the possibility of the proposed system.

Keywords: Communication, Augmented Human, Wearable Device

1. はじめに

会話の相手と直接対面して行われる対面コミュニケーションは、人間の基本的な行動である。対面コミュニケーションにおいては、音声言語だけではなく表情や視線・姿勢・しぐさなどを用いることで、円滑なコミュニケーションが可能である。音声言語を用いる場合を言語コミュニケーション、それ以外の表情や視線などによるものを非言語コミュニケーションと呼ぶ。コミュニケーションにおいて非言語コミュニケーションは重要な役割を担っているとされており、Mehrabianによると、発話による影響力を100%とした場合に、言語情報の占める割合は7%、音声や音質のような聴覚情報が占める割合は38%、表情のような視覚情報の占める割合は55%とされている[1]。これより、円滑にコミュニケーションを行うには、発話や口元の表情形成が重要であるということがわかる。しかし、日常生活において、さまざまな肉体的・精神的要因によって十分にパフォーマンスを発揮できないことがある。例えば、体調不良や緊張

によって思い通り声を出せない場合、接客やプレゼンテーションなどの活動に支障をきたす。また、気分が落ち込んでいる場合、口角を上げてにこやかな表情になることは難しい。外観の醜美もコミュニケーションに影響を与え、歯



図1 マスク型デバイス

Figure 1 Mask-type device.

^{†1} 明治大学
Meiji University

並びの悪い人は自分の口元に対してコンプレックスを感じ、対人関係に消極的になってしまう。

これらの問題を解決するために、我々は発話と口の表情を代替するマスク型デバイス「MouthOver」を開発した[2]。このシステムでは、マスクに内蔵されたスピーカから事前に録音された音声再生され、マスク前面のディスプレイ上に口の開閉アニメーションが表示される(図1)。これにより、ユーザが自身の精神や肉体的状態に関係なく、状況に応じた発話と表情形成を行うことが期待できる。

MouthOverの開発を通じて、我々は顔面での新たな情報提示の可能性について検討した。そして、マスク型デバイスを用いて音声情報の視覚化や会話の内容に関連した情報提示を行うことによって、会話をより楽しくしたり、会話の内容への理解を深めたりすることができるのではないかと考えた。

そこで本研究では、新たなマスク型デバイスとして、以下の三種類の特徴を有するHappyMouthを提案する。

- 自由に選択可能な口画像の表示による、口の外觀デザインの実現

- 発話内容の字幕表示による、会話内容の理解の支援
- 発話内容での画像検索による、会話内容の理解の支援および会話の活性化

本論文では、提案システムの機能とその実装について述べる。また、提案システムの機能の評価実験について報告する。

2. 提案システム

2.1 提案機能

マスク型デバイスで人間の対面コミュニケーション能力を拡張するために、三種類の機能を提案する。以下にそれぞれの機能について説明する。

2.1.1 口画像の表示・外觀デザイン

口の画像を表示させ、ユーザの発話の音声に応じて口の開閉アニメーションを行う。また、画面に表示する口をユーザが好みのものに変更できるようにしている。あらかじめさまざまな種類の口の画像が登録されており、ユーザはその中から自身の好みに合ったものを選択する(図2(a))。

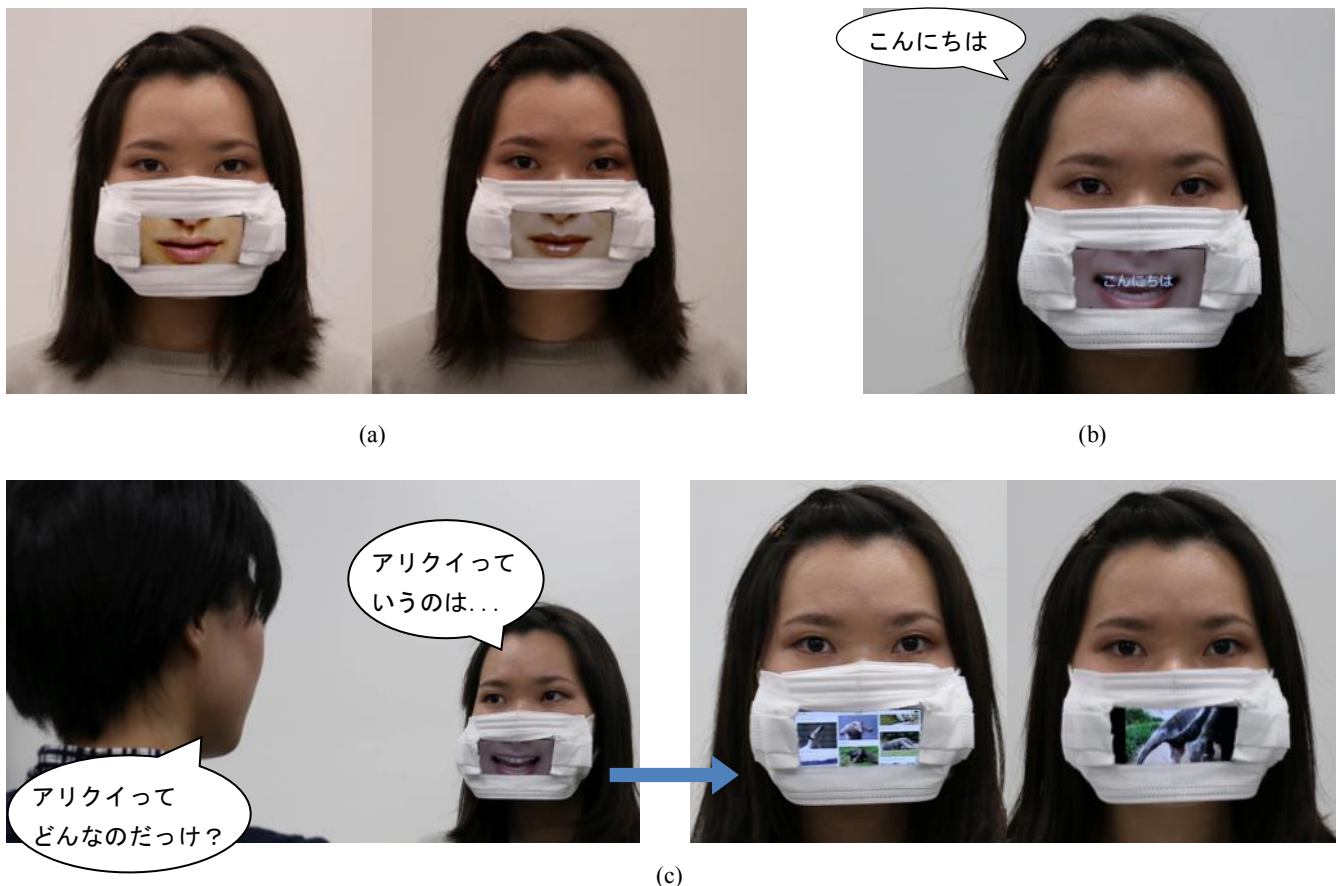


図2 提案システムの装着者
Figure 2 Wearer of the proposed system

2.1.2 字幕の表示

ユーザの発話をテキスト化し、画面に表示させる(図 2(b)). 具体的には、音声認識によってユーザの発話内容をテキスト化し、字幕のようにして口の画像の前に表示する。このとき、画面へのスクロール操作により過去の発話内容を見返すことができるようにすることで、ユーザとの対話者はユーザが話した内容を視覚的に確認することができる。

2.1.3 画像検索・表示

画面にインターネットでの画像検索の結果を表示させる(図 2(c)). 会話中にユーザの発した言葉を音声認識し、認識した文言で画像検索を行いその結果ページを画面に表示させる。通常状態では口の画像が表示されており、装着者が検索対象の語句を発するとウェブブラウザが表示され、検索結果のページを表示する。ウェブブラウザは一定時間が経過すると非表示になり、再度検索対象の語句を発すると表示される。

2.2 実装

提案システムを iPod touch 上で動く iOS アプリケーションとして実装した。口の開閉アニメーションは声の音量に応じて口画像を変形させることで実現した。iPod touch のマイクで取得した音声を入力とし、音量が一定値未満の場合は閉じた口の画像を、一定値以上の場合に開いた口の画像表示した。字幕の表示機能や画像検索・表示機能における音声認識には Speech Framework[3]を用いた。画像検索・表示機能については、ユーザが発するつもり語句をあらかじめシステムに登録しておき、音声認識でその語句の発話を検出したらウェブブラウザを表示して検索を行うようにした。また、ウェブブラウザ表示中は音声認識を停止させ、非表示になったら音声認識を再開するようにした。

3. 評価

今回提案した機能のうち、口画像表示機能に関する客観的評価を行った。また、字幕表示機能および画像検索・表示機能については、著者自身が主観評価を行った。

3.1 口画像表示機能の評価

提案システムで口の画像を表示した場合、装着者に対する印象にどのような影響を及ぼすのか検証した。

3.1.1 実験環境

明治大学中野キャンパスの実験室内で行った。実験中はビデオカメラで参加者の様子を撮影した。このとき、個人が特定できないよう参加者の斜め後ろから撮影した。

3.1.2 実験条件

実験には、理系学部在籍する、提案システムの装着者を知らない大学生および大学院生 12 名(20~23 歳の女性)が参加した。参加者を提案システムの装着者を知らない者としたのは、装着者を知っていることにより印象評価に影響を及ぼすと考えたためである。装着者の状態を下記の三条件に設定し、参加者内比較で実験を行った。

N 条件：何も装着しない

GM 条件：実験者が一般的なマスクを装着する

HM 条件：実験者が提案システムを装着する

参加者はそれぞれの条件を全て体験するため、各条件を体験する順序が偏ることのないようにカウンタバランスをとった。また、参加者はランダムにそれぞれの体験順序に振り分けられた。

3.1.3 実験手順

参加者を 1 名ずつ実験室に呼び出し、装着者と向かい合せて座らせる。その後、参加者には装着者に向かって質問をしてもらい、装着者はその質問に対して回答するというやりとりを行った。質問項目はあらかじめ用意しておき、参加者にはランダムに選んで質問してもらった。ひとつの条件につき 4 回質問してもらい、それぞれの条件終了後に装着者の印象をアンケートにより回答してもらった。アンケートで用いる形容詞を表 1 に示す。これらは、特性形容詞尺度[4]から本実験において有用であると考えた形容詞に、本実験において条件による差が見られるのではないかと予想した形容詞を追加したものである。本アンケートは 7 段階のリッカートスケールであり、1 が「全く同意できない」、7 が「非常に同意できる」というスコアになっている。また、自由記述欄も設けた。

3.1.4 結果と考察

本実験の結果を図 3 に示す。グラフは 7 段階のリッカートスケールで得た値の各条件における平均値と標準偏差を示す。なお、逆転項目については値を反転し、値が大きいくほど「非常に同意できる」を意味する。

アンケートの得点について、一要因参加者内分散分析を行った(独立変数：実験条件、三水準(N, GM, HM)、従属

表 1 実験に用いた形容詞

Table 1 Adjective used for experiment.

1	話しやすい	12	感じのわるい
2	信頼できる	13	積極的である
3	威圧的である	14	思いやりがある
4	社交的である	15	素っ気ない
5	魅力的である	16	なまいきでない
6	無気力である	17	近づきたい
7	怖い	18	親切である
8	自信がある	19	責任感がある
9	誠実である	20	分別のある
10	つめたい	21	恥ずかしがりである
11	親しみやすい	22	堂々としている

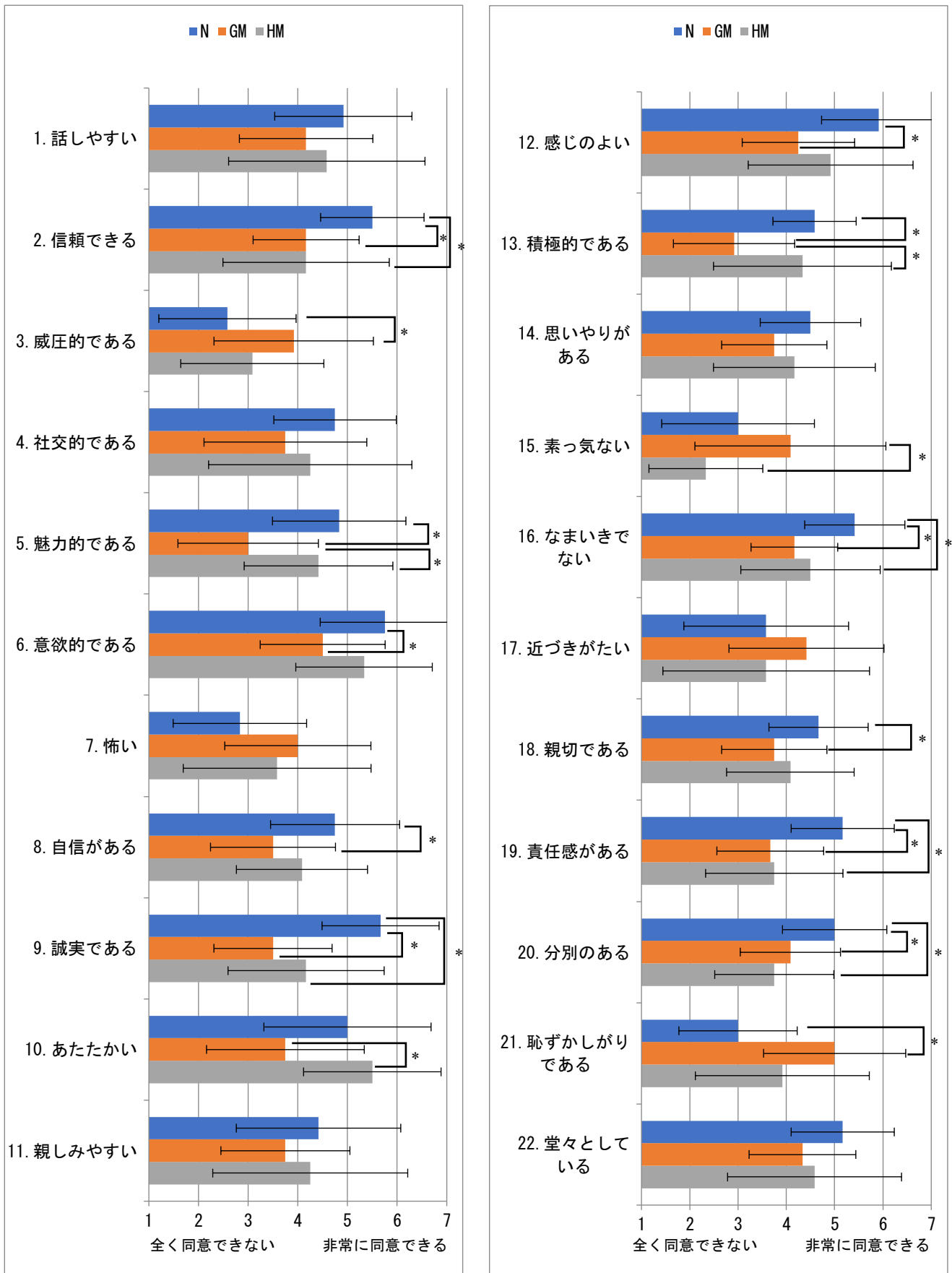


図 3 条件ごとの印象

Figure 3 Impression for each experimental conditions.

変数：アンケートの得点）。そのうち、群の効果に有意差が観察された項目について、LSD法による多重比較を行った。「5. 魅力的である」、「13. 積極的である」において、N条件およびHM条件の平均値がGM条件の平均値よりも有意に高いことが観察された ($p < .05$)。「10. あたたかい」において、HM条件の平均値がGM条件の平均値よりも有意に高いことが観察された ($p < .05$)。「15. 素っ気ない」において、HM条件の平均値がGM条件の平均値よりも有意に低いことが観察された ($p < .05$)。「2. 信頼できる」、「9. 誠実である」、「16. なまいきでない」、「19. 責任感がある」、「20. 分別のある」において、N条件の平均値がGM条件およびHM条件の平均値よりも有意に高いことが観察された ($p < .05$)。「6. 意欲的である」、「8. 自信がある」、「12. 感じのよい」、「18. 親切である」において、N条件の平均値がGM条件の平均値よりも有意に高いことが観察された ($p < .05$)。「3. 威圧的である」、「21. 恥ずかしがりである」において、N条件の平均値がGM条件の平均値よりも有意に低いことが観察された ($p < .05$)。その他の形容詞においては条件ごとに差は見られるものの、有意な差は見られなかった。

これらの結果により、提案システムの装着者は一般的なマスクの装着者に比べて魅力的で積極的な印象を受け、あたたかさを感じさせるということがわかった。また、提案システムの装着者より一般的なマスクの装着者の方が素っ気ない印象を受けるということがわかった。一方で、何も装着しない状態よりも提案システムを装着した方が有意に優れているという結果は得られなかった。

3.2 字幕表示機能および画像検索・表示機能の評価

著者の感じた提案システムの使用感を簡潔に記述する。字幕表示機能では、かなり正確にユーザの発話をテキスト化して表示することができていた。字幕表示機能を用いて会話していると、会話の相手が自分の発話内容が表示されている本システムに注目するため、話をきちんと聞いてくれていたような印象を受けた。画像検索・表示機能でも同様に、会話の相手が本システムに注目したり、画面操作を行ったりしたことで、自分の話に興味を持ってくれたような気がした。また、どちらの機能でも、会話の相手が本システムに興味を持って接してきた場合、会話の相手に対して親密さを感じ、安心感からリラックスして会話ができたと感じた。本システムは口の前に装着するため、装着者自身が画面の内容を確認しながら画面を操作するのは困難だった。このため、装着者が本システムの操作を簡単に行えるようなインターフェースが欲しいと感じた。

4. 議論

4.1 提案システムの装着者に対する印象

実験結果より、提案システムの装着者は一般的なマスクの装着者よりも良い印象を与えるということが示唆された。

自由記述アンケートおよび聞き取り調査によると、一般的なマスクの装着者にはかたい印象を持ち、口元の表情が見えないため怖さやつめたさを感じるという意見が得られた。また、提案システムの装着者には、表情が見えて安心した、笑顔が向けられていたのでポジティブな印象を抱いたという意見が得られた。提案システムで笑顔を表出することで、一般的なマスクの装着者に感じられるつめたさが緩和され、あたたかい印象になったのではないかと考えられる。

一方、現時点では、提案システムの装着者よりも何も装着していない人の方が好印象であることが示唆された。特に、「2. 信頼できる」、「9. 誠実である」、「16. なまいきでない」、「19. 責任感がある」、「20. 分別のある」においてはHM条件の平均値がN条件の平均値よりも有意に低いことが明らかになった。自由記述アンケートおよび聞き取り調査によると、提案システムは違和感があり怖い、口の開閉アニメーションが機械的で違和感があり話しづらい、口の位置や動きが想像しているのと違うため不自然である、ずっと笑顔なので不自然である、などの意見が得られた。したがって、口画像の表示位置や開閉アニメーションを改良したり、種類を増やしたりすることで不自然さや違和感を緩和できると考えられる。口の開閉アニメーションについては、母音ごとに口画像を用意し、装着者の発した単語の母音に合わせてそれぞれの画像を表示する方法が考えられる。また、現時点では常に笑顔の口画像を表示しているが、状況に応じて笑顔以外の口画像も表示することで、より自然な表情の提示ができると考える。このような改良を行うことで、本システムの装着者をより好印象にできると考えられる。

4.2 提案機能について

現在の実装では、あらかじめ登録されている口の画像の中から、ユーザが好みのものを選んで表示するという方式を採用している。ユーザがより自由に自身の外観をデザインするためには、より良い方法があると考えられる。例えば、ユーザがこの口になりたいと思う人の画像をシステムに登録するという方法である。この場合、人の画像から口の部分を切り抜き、口の開閉アニメーションのために画像を変形させる処理をする必要がある。顔の画像変形をする先行研究[5]をもとに実装可能であると思われる。

字幕表示機能では、ユーザの発話をそのままテキスト化して画面に表示するという実装にした。ユーザの発話を他言語に翻訳した上で画面に表示すれば、対話者同士に言語上の壁がある場合でも、気軽に会話することができると考える。

4.3 考えられる追加機能

本研究では、マスク型デバイスの一機能として、自分の好みの口元をデザインすることを提案したが、同様に自分の好みの発声をデザインすることも可能であると考えられる。具体的には、自分の肉声に対してリアルタイムに好きなエ

フェクトをかけたたり、自分の過去の発話をシステムに登録しておき再利用したりするなどが考えられる。また、自分の発話だけでなく、他人の発話を提示することも可能である。他人の口元と声を本システムで提示することで、自分の口に他者を寄生させることが可能になる。これにより、例えば子供に対する読み聞かせのような、物語のためのインターフェースとして本システムが利用できると考える。読み聞かせの際、本システムに登場人物の口元の外観とセリフの音声を登録し再生すれば、読み聞かせを行っている者が登場人物になりきることができる。このように、本システムは、エンターテインメントのためのインターフェースとしても利用できるであろう。

5. 関連研究

5.1 顔面装着型ディスプレイ

スマートフォンアプリでは、口のイラストが表示されたスマートフォンを口の前に掲げて体験するコンテンツがいくつ也存在する[6][7]。これらはさまざまな口の画像が選択でき、ユーザの発話の音量に応じて口の開閉アニメーションの表示を行うという点で本研究と共通している。また、梅澤らはユーザの顔をアバタの顔に変えられる仮面型ディスプレイを提案している[8]。システムの装着者の顔を変えることで、装着者の他者に与える印象や雰囲気进行操作し、対面コミュニケーションの支援を目指している。これらのシステムでは口や顔の画像のみを表示しているが、本研究では口の画像以外の画像やユーザの発話内容のテキストを表示し、人間のコミュニケーション能力をさらに拡張することを目指す。

ChameleonMask[9]は遠隔ユーザの顔が表示されたディスプレイを代理人が頭部に装着することで、遠隔ユーザの存在感を提示するテレプレゼンスシステムである。ディスプレイを顔面に装着することでユーザのコミュニケーション能力を拡張しているという点で本研究と共通している。ChameleonMaskの主な目的がテレプレゼンスであり、システムの装着者が遠隔ユーザの能力を借り遠隔ユーザとしてふるまうのに対し、本研究において装着者は能力が拡張された本人としてふるまう。

Emotional Cyborg[10]では、小型ディスプレイを搭載した眼鏡型のデバイスを用いて感情労働における目を使ったコミュニケーションを代替している。この研究と同様に感情労働の支援という目的で人間の口を対象としているものとして、Responsive Mouth[11]がある。本研究と同様に、マスク型のデバイスで、マスク上のディスプレイに口の画像を表示し口の外観を自由にデザインできるようにしている。萩原らは、他人と視線を合わせられない症状をもつ人々を社会福祉学的な観点から支援するためのシースルー型HMDを使用したシステムを提案した[12]。この研究では、他人と視線を合わせられないような人々の持つコミュニケ

ーション上の問題を緩和・改善する。我々が開発したMouthOver[2]は、人間の発話と表情形成を代替するマスク型デバイスである。本研究と同様に、ディスプレイ付きのマスク型デバイスを口の前に装着しコミュニケーションの支援を行っている。これらの研究はコミュニケーションにおいて十分なパフォーマンスを発揮できない人の負担軽減を目指している。本研究では負担の軽減だけではなく、さまざまな情報をシステムで提示することで人間の表現能力をさらに拡張することを目指す。

5.2 対面コミュニケーションの拡張

EyeFeel & EyeChime[13]やE-Gaze[14]では、本来視覚のみで確認する視線を触覚や聴覚で体感させることで、対面コミュニケーションの可能性を広げることを目指している。対面コミュニケーションの支援のために非言語コミュニケーションに着目するという点で本研究と共通しているが、これらの研究が非言語コミュニケーションの中でも視線に着目しているのに対し、本研究では表情に着目する。

5.3 表情・外観の操作

Yoshidaらは、リアルタイムにユーザの表情を変形させて表示させるシステムを提案した[5]。この研究の技術を用いて、ビデオ会議におけるユーザの顔の外観をリアルタイムに変形してフィードバックするシステムが提案されている[15]。また、このシステムにより、遠隔共同作業における創造力が向上したことが示されている。FaceShare[16]は、擬似的にミラーリングを発生させ画面に表示するビデオ電話システムである。ミラーリングとは他者の動作や表情を無意識的に真似てしまうことであり、ミラーリングをする相手には親近感を抱いたり、会話が弾んだりするとされている。このシステムでは、対話の相手が笑顔になった場合、自分の顔も笑顔になるように顔の画像を変形して相手の画面に表示している。FaceShareを用いることで、ビデオ電話中の会話が弾んだことが示されている。これらの研究では、コンピュータにより表情や外観を操作することで、ビデオチャットのような遠隔でのコミュニケーションが促進されるということが示されている。我々のシステムを用いることで、対話の相手と直接対面している状態でも同様にコミュニケーションを促進させることができると考えられる。

6. おわりに

本研究では、対面コミュニケーション能力を拡張するための機能を有するHappyMouthを提案した。提案デバイスの機能として、ユーザが自分の好みの口を選択して表示する機能、ユーザの発話をテキスト化して字幕表示する機能、ユーザの発したキーワードをインターネットで画像検索した結果を表示する機能の三種類を実装した。口画像を表示する機能を用いた場合、提案システムの装着者に対してどのような印象を持つのか調査するため評価実験を行った。その結果、提案システムの装着者は一般的なマスクの装着

者に比べて魅力的で積極的な印象を受け、あたたかさを感じさせるといったことがわかった。また、提案システムの装着者より一般的なマスクの装着者の方が素っ気ない印象を受けるといったことがわかった。一方で、何も装着しない状態よりも提案システムを装着した方が有意に優れているという結果は得られなかった。その他の機能については著者自身の主観評価を行い、提案システムの可能性について検討した。

今後は、口の開閉アニメーションを改良したり口画像の種類を増やしたりすることで、提案システムの装着者により良い印象を与えられるようにすることを目指す。また、表情以外の非言語コミュニケーションに着目することで、人間のコミュニケーション能力をさらに拡張したいと考えている。

参考文献

- [1] Mehrabian, A. Communication without words, Psychol. Today, 1968, vol. 2, no. 9, p. 52-55.
- [2] 石井綾郁, 橋本直. MouthOver: 発話と口の表情を代替するマスク型デバイス, インタラクシオン 2017 論文集, 2017, p. 907-912.
- [3] “Speech | Apple Developer Documentation”.
<https://developer.apple.com/documentation/speech>, (参照 2018-01-28).
- [4] 林文俊. 対人認知構造の基本次元についての一考察. 名古屋大学教育学部紀要 (教育心理学科), 1978, (25), p.233-247.
- [5] Yoshida, S. et al.. Manipulation of an emotional experience by real-time deformed facial feedback. Proc. AH. 2013, p. 35-42.
- [6] “MouthOff™を App Store で”.
<https://itunes.apple.com/jp/app/mouthoff/id306588353?mt=8>, (参照 2018-01-28).
- [7] “GearSprout”. http://gearsprout.com/mouth_mover.html, (参照 2018-01-28).
- [8] 梅澤章乃, 竹川佳成, 杉浦裕太, 平田圭二. e2-Mask から mime-Mask: 顔の印象を拡張する仮面型ディスプレイの提案, 第 176 回 HCI 研究会. 2018.
- [9] Misawa, K. and Rekimoto, J.. ChameleonMask: Embodied Physical and Social Telepresence using human surrogates. Proc. CHI. 2012, p. 401-411.
- [10] Osawa, H.. Emotional cyborg: Complementing Emotional Labor with Human-agent Interaction Technology. Proc. HAI. 2014, p. 51-57.
- [11] Osawa, H.. Responsive Mouth: Enhancing Your Emotional Skill with Partial Agency. Proc. HRI. 2015, p. 299-299.
- [12] 萩原早紀, 栗原一貴. シースルー型 HMD を用いた社会福祉学的アプローチに基づく”視線恐怖症的コミュ障”支援システムの開発と検証, WISS2014 予稿集, 2014, p. 103-108.
- [13] Hosobori, A. and Kakehi, Y.. Eyefeeel & EyeChime: a face to face communication environment by augmenting eye gaze information. Proc. AH. 2014, p. 7-10.
- [14] Qiu, S. et al.. E-Gaze: Create Gaze Communication for People with Visual Disability. Proc. HAI. 2015, p. 199-202.
- [15] Nakazato, N. et al.. Smart Face: enhancing creativity during video conferences using real-time facial deformation. Proc. CSCW. 2015, p. 75-83.
- [16] Suzuki, K. et al.. FaceShare: Mirroring with Pseudo-Smile Enriches Video Chat Communications. Proc. CHI. 2017, p. 5313-5317.