

MouthOver: 発話と口の表情を代替するマスク型デバイス

石井綾郁[†] 橋本直[†]

概要: 発話や口元の表情形成は、コミュニケーションにおいて重要な要素である。しかし、本人の性格や体調、生まれつきの外観などの精神的・肉体的要因によって、十分なパフォーマンスを発揮できないことがある。このような問題を解決するために、本研究では発話と口の表情を代替するマスク型デバイス「MouthOver」を提案する。提案システムでは、マスクにスピーカとディスプレイが搭載されており、事前に録音された音声と、コンピュータで加工された口の画像が提示される。これにより、ユーザは自身の精神や肉体の状態に関係なく、状況に応じた発話と表情形成を実現できる。本稿では、システムの概要と想定されるアプリケーション例について説明し、作成したプロトタイプを示す。

MouthOver: A Wearable Device to Substitute Voice and Mouth Expression

AYAKA ISHII[†] SUNAO HASHIMOTO[†]

Abstract: The ability to utter and create expression of the mouth is an important element in communication. However, it may not be possible to demonstrate sufficient performance owing to mental and physical factors such as the person's personality, physical condition and natural appearance. In order to solve these problems, we propose a wearable device called "MouthOver" that substitutes voice and mouth expression. In the proposed system, a speaker and a display are mounted on the mask, and it presents the recorded sound and the image of the mouth processed by the computer. This allows the user to utter and create expression that suits the situation, regardless of their own mental and physical conditions. This paper describes the overview of the system, examples of assumed applications, and created prototype.

1. はじめに

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。例えば、相手に情報を正確に伝えるためには、聞き取りやすい声で説明する必要がある。また、口元の外観は相手に与える印象に影響し、口角の上がった口元は信頼度が高くなり、魅力的に見えるとされている[1]。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。

そこで本研究では、これらの問題を解決するために、発話と口の表情を代替するマスク型デバイス「MouthOver」を提案する(図1)。本システムでは、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画像がインタラクティブに表示される。これにより、ユーザは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実現できる。提

案システムを用いて、コミュニケーションの質の向上と負担の軽減を目指す。

2. 関連研究

人間の顔のコミュニケーション能力をコンピュータで代替するような研究として、ChameleonMask[2]や Emotional Cyborg[3]がある。ChameleonMask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。このシステムが、リアルタイムの遠隔通信によって他人との代替を行うのに対し、我々のシステムは事前に記録されたリソースを用いるため、他人だけでなく過去の自分



図1 MouthOverのプロトタイプ

Figure 1 Prototype of MouthOver.

[†] 明治大学総合数理学部先端メディアサイエンス学科
Department of Frontier Media Science, Faculty of Interdisciplinary Mathematic
Science at Meiji University

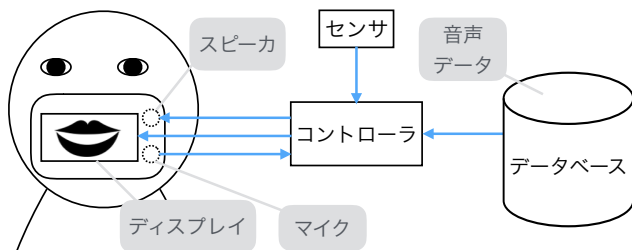


図 2 システムの構成
Figure 2 System configuration.

との代替もできるという点で異なる。Emotional Cyborg では、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感情表現をコンピュータで代替することを提案している。この研究が人間の目を対象にしているのに対し、我々は人間の口を対象にしている。

3. MouthOver

3.1 システム構成

MouthOver のシステム構成を図 2 に示す。マスク内部に搭載されたスピーカからは、事前に録音された、装着者本人または他人の音声再生される。また、マスク前面に取り付けられたディスプレイには、装着者本人または他人の口の画像が表示される。この画像は事前に撮影されたものを使用し、ソフトウェアによって、音声と連動した口の開閉アニメーションの表示が行われる。なお、ユーザが自身の声で直接発話する際にも口の開閉アニメーションを行うために、マスクに内蔵されたマイクを用いて音声信号を取得する。

音声の再生制御（どの音声をいつ再生し、いつ停止するか）は、ユーザ自身の行動や環境の状態と連動させる。このため、ユーザの目的や利用シチュエーションに応じて、適切なセンシング方法を選択する。例えば、再生する音声ボタンごとに割り当てられた小型リモコンをユーザが直接操作する方法や、環境側に複数の近接センサを設置し、それぞれのセンサに人間や車両が近づいた際に割り当てられた音声を自動再生する方法などを用いる。

3.2 想定されるアプリケーション例

接客業従事者は、労働中に自身の感情を制御する必要があり、精神的負担を伴う。例えば、気分が落ち込んでいる時にも笑顔で労働を行う事が要求される。本システムを用いて笑顔の口の画像を表示すれば、周囲からは常に笑顔に見えるように見える。また、接客において「いらっしゃいませ」のような定型句を連呼しなければならない場合がある。あらかじめ本システムにこのような音声を登録しておけば、接客中の発話に関する負担を軽減することが期待できる。さらに、レジと本システムを連動させれば、金額に関する読み上げを自動化し、言い間違いをなくすことができる。

プレゼンテーションでは、聴衆に対して正確に情報を伝

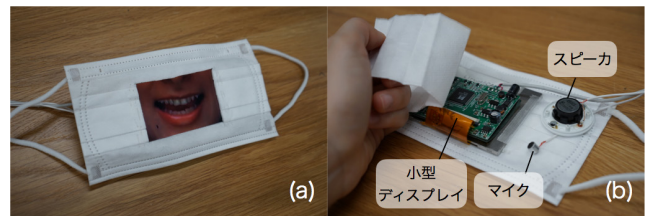


図 3 プロトタイプの実装 (a: 前面, b: 背面)
Figure 3 Prototype implementation (a: front, b: back).

える必要がある。しかし、体調不良や緊張によって思い通りに声が出せない場合や、焦りから早口になってしまう場合がある。あらかじめ本システムに伝えたい文言を登録し、プレゼンテーション時に再生すれば、ユーザは失敗することなく正確にプレゼンテーションを行うことができる。また、原稿を暗記する必要もなくなり作業負担が軽減される。

3.3 プロトタイプの実装

作成したプロトタイプを図 3 に示す。ディスプレイ、スピーカ、マイクは PC に有線接続されている。口画像の表示を行うソフトウェアは Processing で作成した。このソフトウェアは、PC に録音された音声およびマイクで取得した音声を入力とし、音量が一定値以上の場合に開いた口の画像を表示し、音量に応じて口画像を変形させる。

録音された音声の再生制御は、小型の Bluetooth キーボードを操作することにより実現している。複数のキーに再生する音声割り当てられており、ユーザは状況に応じて再生したい音声に対応するキーを操作する。

4. まとめと今後の課題

本研究では、発話と口の表情を代替するマスク型デバイス「MouthOver」を提案した。本システムを使用することで、ユーザは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実現できる。

現在のプロトタイプでは、録音された音声をそのまま再生しているが、直接発話された音声との聞こえ方の違いの低減や、状況に合わせた音量やトーンの調整を行う必要がある。今後は、センシングによるインタラクティブな音声再生制御について実装を行い、ユーザテストを通じて、装着者のユーザビリティや対話相手の印象などを検証していく予定である。

参考文献

- [1] Kleisner, K., Priplatova, J., Frost, P., and Flegr, J. Trustworthy-Looking Face Meets Brown Eyes. PLoS One, 8(1), e53285, 2013, doi:10.1371/journal.pone.0053285.
- [2] Osawa, H. Emotional cyborg: Complementing Emotional Labor with Human-agent Interaction Technology. Proc. HAI. 2014, p. 51-57.
- [3] Misawa, K., and Rekimoto, J. ChameleonMask: Embodied Physical and Social Telepresence using human surrogates. Proc. CHI. 2012, p. 401-411.