

# ビデオ会議中の表情ロギングシステム MeeMo の開発

志村 朋人<sup>†</sup> 徳原 耕亮<sup>†</sup> Makhlof Jihed<sup>†</sup> 石田 繁巳<sup>‡</sup> 荒川 豊<sup>†</sup>

<sup>†</sup>九州大学大学院システム情報科学研究府・研究院

<sup>‡</sup>公立はこだて未来大学システム情報科学部 / 九州大学システム LSI 研究センター

## 1 はじめに

本稿では、ブラウザ環境で実行できるビデオ会議システムに対して、リアルタイム解析とフィードバックを行うシステムについて報告する。遠隔会議支援の研究として、アイコンタクトを擬似的に再現する手法 [1] や発話衝突を防ぐ手法 [2] なども提案されているが、いずれも何らかの装置や環境が必要となる。特別な装置を使わずに頷きを表示するシステム [3] は存在するが、処理が重く長時間の運用に耐えられない、自身の頷きカウントが反転してしまう、ログが取れない、というデメリットが存在した。そこで、我々は、何らかの装置を使わず、上記デメリットを改善したビデオ会議中の表情ロギングシステム MeeMo を開発した。

MeeMo の実装には、Web ブラウザで動作する軽量顔検出フレームワーク MediaPipe Face Mesh を用い、ビデオ会議中の“頷き”と“発話”（以降、両者を合わせてマイクロ行動と呼ぶ）を参加者ごとにリアルタイムでカウントし、表示する。同時に、計測データはクラウド上のデータベースに記録される。今回 4 名の被験者に協力してもらい、MeeMo を利用しないビデオ会議と、MeeMo を利用したビデオ会議での体験の違いについて、アンケート形式でその効用について評価した。その結果、提案システム MeeMo によってリアルタイム認識とフィードバックを実現できたことを確認するとともに、参加者自身の振り返りや気づきに役立ったこと、さらにその後の行動やコミュニケーションにも影響を与えることを明らかにした。

## 2 表情ロギングシステム MeeMo

頷きと発話のリアルタイムフィードバックシステム MeeMo は、ビデオ会議中の参加者それぞれのパソコンにおいて、カメラ映像から頷きと発話のマイクロ行動をリアルタイムに認識し、各々のパソコンでビデオ会議システム上に表示する。このシステムの目的は、会議中のノンバーバルな表現を可視化することである。図 1 に示すように、リアルタイムマイクロ行動認識は、顔検出ブロック、頷き検出ブロック、音声グラフ検出ブ

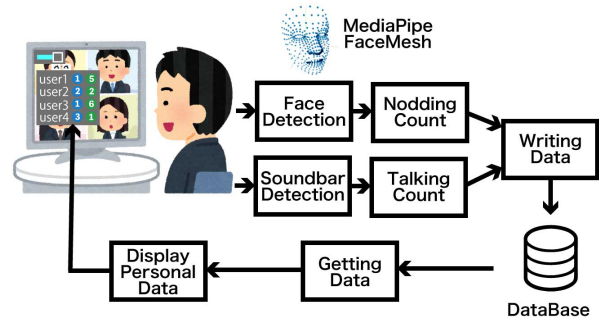


図 1: リアルタイム頷き・発話認識の概要

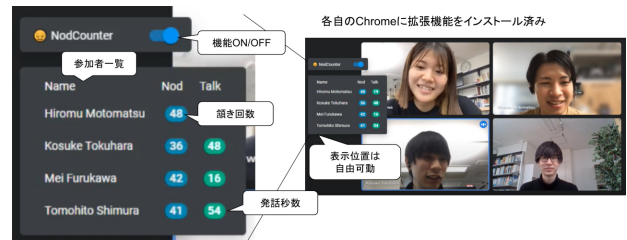


図 2: MeeMo 利用中のビデオ会議画面

ロック、発話検出ブロック、データ書き込みブロック、データ読み込みブロック、マイクロ行動表示ブロック、データベースの 8 つの要素から構成される。

顔検出ブロックでは、ビデオ会議の映像から MediaPipe Face Mesh を用いて、リアルタイムに顔の 3D 座標情報を取得する。頷きカウントブロックでは、3D 顔座標情報から x 座標成分のローテーションを計算し、機械学習を用いて頷きを検出する。音声グラフ検出ブロックでは、ビデオ会議の映像から自身のビデオ情報を検出し、音声グラフの変化を 0.1 秒毎に取得する。発話検出ブロックでは、音声グラフの 1 秒分の分散を特徴量として発話を検出する。データ書き込みブロックでは、頷きや発話などのイベント発生時に、イベントタイプとイベント発生時刻、ユーザーネーム、ブラウザのタブ URL から抽出した Meeting ID をデータベースに書き込む。データ読み込みブロックでは、Meeting ID に紐付いた情報をデータベースから取得する。マイクロ行動表示ブロックでは、データ読み込みブロックで読み込んだ情報を用いて、頷きの回数と発話の秒数

