「マルチメディア，分散，協調とモバイル
(DICOMO2021) シンポジウム」 令和3年6月-7月

効果的な意思疎通を促すオンラインミーティングシステムの構築と評価

安永 貴之[1] 小口 寿明[1] 児矢野 友香[1] 土井 沙耶香[1]
三谷 洋之[1] 赤垣 慎吾[1] 飛田 博章[1]

概要： カメラを ON にして顔を見せなくても場の雰囲気や相手の表情，意見を伝えることのできるオンラインミーティングシステムの仕組みを Web アプリケーションとして構築した．既存の OSS のミーティングシステムである，BigBlueButton を実装の基盤として，話者の顔の動きに追従する 2D アバターと 3D アバターにより，視覚的な意思伝達を実現した．更に，発話内容から，自動投票を行い，発話内容のネガティブさやポジティブさ，会議の話題や，会議の活発さなどをレポートで表現することで，場の雰囲気や意見を可視化した．

Development and Evaluation of Online Meeting System to Promote Effective Communication

TAKAYUKI YASUNAGA[†1] TOSHIAKI OGUCHI[†1] YUKA KOYANO[†1] SAYAKA DOI[†1]
HIROYUKI MITANI[†1] SHINGO AKAGAKI[†1] HIROAKI TOBITA[†]

1. はじめに

近年，オンライン英会話やプログラミング教育をはじめ，ビジネスの場においてもクラウドソーシングやテレワークなどの普及により，オンラインミーティングシステムを利用する機会が増加している．

従来のオンラインミーティングシステムは，カメラを ON にして自身の顔を見せることを前提にしたデザインであり，参加者がカメラを ON にしない場合，相手の様子や表情が見えない．従って，発言後の無言の時間やレスポンスの悪さから発話者は壁に向かって話している感覚に陥りやすい．また，対面であれば，頷いたり首をかしげたりしている様子を見ることで，話の流れやその雰囲気を踏まえて相手が何を考えているかを読み取ることが可能だが，オンラインミーティングでは前述のようにカメラを ON にしていないことが原因で，その場の雰囲気を読み取ることが困難だと感じることが多い．

オンラインミーティングを支援するツールは，日程や場所の管理，ドキュメント共有や，ホワイトボード機能といったロジスティクス領域のサービスが中心であり，会議の途中で参加者の意思を確認する手法や議論の内容・会議状況を可視化するといった会議の中身そのものをサポートするサービスは少ない．そのため，対面でのミーティングと比較し，顔が見えないオンラインの場では，ファシリテーションや会議の振り返りの実施が難しい．

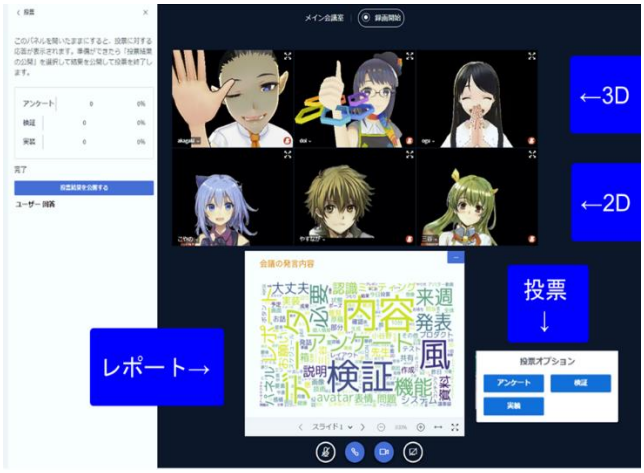

図 1：システム概要

Zoom [1] や Google Meet [2] などのオンラインミーティングシステムには上述の課題が存在する．また，FaceRig [3] のようにアバターにより参加者の表情を表現する手法は存在するが，ミーティングシステムと一体となっていないため機能の拡張に限界がある．特に，仮想カメラを使うアプローチは OS に依存するものや，複数のアプリの起動が必要など手軽に使うためには課題も多い．

そこで本研究では，カメラを ON にして顔を見せなくても場の雰囲気や相手の表情，意見を伝えることのできるオンラインミーティングシステムの仕組みを Web アプリケーションとして構築した．従って，指定された URL にアクセスするだけで手軽にシステムを利用することができる．

開発したプロダクトは，既存の OSS のミーティングシス

†1 東京都立産業技術大学院大学
Advanced Institute of industrial technology (AIIT)

© 2021 Information Processing Society of Japan

テムである，BigBlueButton (BBB) [4] を実装の基盤として，話者の顔の動きに追従する 2D アバターと 3D アバターにより，視覚的な意思伝達を実現した．更に，発話内容を自動的に集約することで，自動投票を行い，発話内容のネガティブさやポジティブさ，会議の話題や，会議の活発さなどをレポートで表現することで，場の雰囲気や意見を可視化した．

2. システム概要

システムはオープンソースである BBB をベースに，アバター機能と投票・レポート機能を実現した (図 1)． BBB をベースに作成したシステムに関し，アバター機能と投票・レポート機能について概要を述べる．

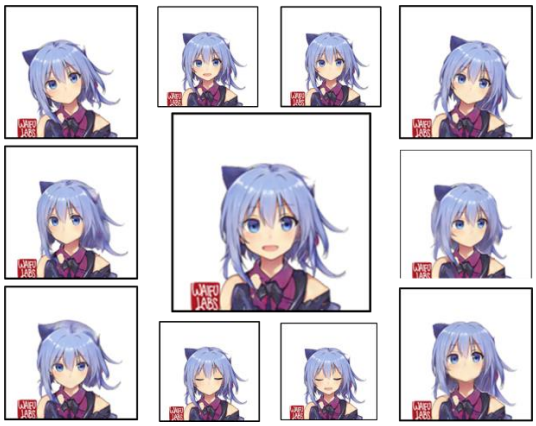

図 2：アニメ調アバターによる表情の表現

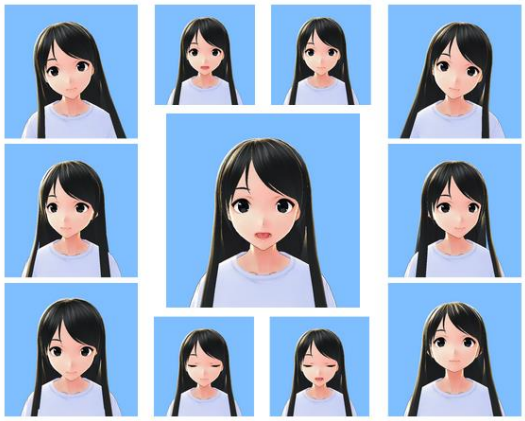

図 3：3D アバターによる表情の表現

2.1 アバター機能

参加者の多くが画像やカメラ機能を OFF にしてミーティングに参加するため，参加者の表情・様子・意思の読み取りが難しい．そこで，参加者の表情が直接反映されるアバターを使い，参加者全員の表情がアバターを介して分かる仕組みを実現した．システムは参加者の表情を分析し，リアルタイムでアバターの表情に反映させる．従って，参加者が首を傾げればアバターも首を傾げ，参加者が口を開ければアバターも口を開ける．本システムはグループワーク等での利用を考え，親しみやすさの点からアニメ調のアバターを提供する．

参加者は，システムにテンプレートとして用意されているアバターを選択し利用することに加え，自分に似たキャラクターを生成しアバターとして利用することができる．キャラクターの生成手法やそれを支援するソフトウェアは数多く存在し，2D 及び 3D のキャラクターを手軽に生成できる．提案手法では，こうした 2D 及び 3D のキャラクターを気軽に利用できる枠組みが提供されている．

図 2 は，2D のアニメ調アバターによる表情の表現である．図 2 の中心にあるアニメ調画像と参加者の首の姿勢などの状態パラメータから，その周辺に配置した図に示すとおり，様々な姿勢や表情をもつアニメ調画像が生成でき，この生成画像を基にアニメ調アバターの顔・目・口の追従動作をリアルタイムに実行できる．

図 3 は，3 次元ポリゴンデータによるヒューマノイドモデルを用いた例である．上記のアニメ調アバターと同様，話者の顔の動きに追従させ，アバターの動きを反映する．表情は予め設定したモードを手動で選定し，わからない，うれしいや，悲しいなどの表情の描写が選択できる．

また，自分に似たキャラクターを使うこともミーティングでは重要となる．単にテンプレートを使った場合には，顔表情は分かるが参加者の特定が難しくなる．こうした問題に対処するために，自分に似たキャラクターを生成する機能を提供している．

図 4：投票機能の表示

2.2 投票機能

BBB に備わっている投票 UI 機能を拡張し，発話とチャット欄への投稿の両方から投票できる機能を実現した．通常，BBB では明示的に投票機能を呼び出して使う必要があったが，より手軽に使えるように拡張した．投票の結果も BBB 内に表示されるので，この投票機能により，意見の相

© 2021 Information Processing Society of Japan

違をいつでも確認できる．また，投票内容をすぐに議論の材料として活かし，ファシリテ―ションに利用することも可能となる．

投票機能の実行例を図 4 に示す．ミーティングシステムのホストは，ミーティングの開始もしくは途中の任意のタイミングで，投票項目となるキーワードを複数個設定できる．ミーティングにおいて「リモート」,「対面」, または「両方」のいずれかを必要とするかについての模擬的な議論を行った場合の例を示す．図 4 は，これらの 3 つを投票項目のキーワードに設定し，投票機能を実行した結果である．オンラインミーティング中の意見から，ユーザが発する対話やチャット欄の書き込み内容を自然言語処理で処理し，設定した単語をカウントした．模擬した議論の投票結果は，「リモート」と「両方」が 40%であり，「対面」が 20%であった．このように，プレゼンターは意見の割れをリアルタイムで確認できる．また，投票内容をすぐに議論の材料として活かせるため，ファシリテ―ションを行うために便利な機能であると考える．

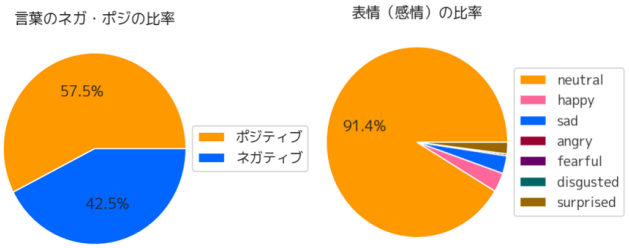

図 5：会議のネガポジ比率と表情の比率の表示

図 6：WordCloud による会議の発話内容の表示

2.3 レポート機能

会議の可視化は会議の質を上げるうえで重要となり，本システムでも可視化する機能を提供している．これらのレポート表示機能によって，会議終了時の振り返りやグループワークでの各チーム間での確認に利用できると考える．

レポート機能は，会議の全体の雰囲気を表示するために，表情の比率，会議の活発さ，発話内容，話題や，ネガポジ比率の画像を，PDF にまとめて BBB のプレゼンテーション領域にレポートとして出力する機能である．

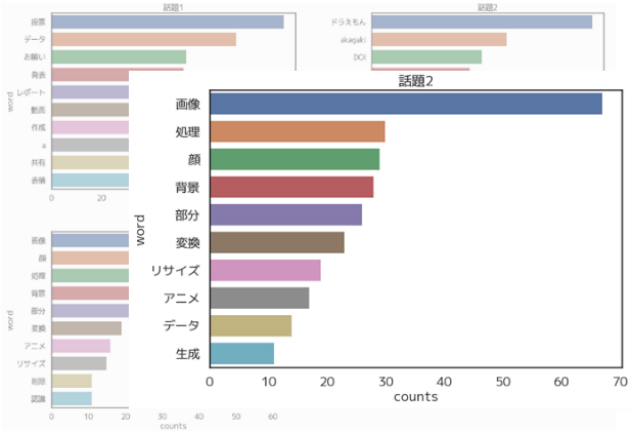

図 7：会議の話題（クラスタ分析）の表示

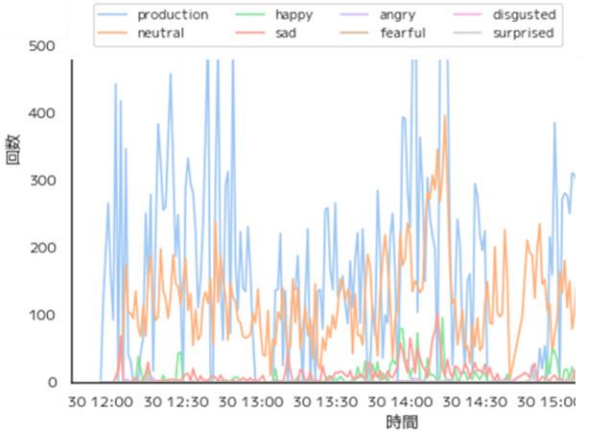

図 8：会議の活発さの表示

クラスタ分析により，会議中の発話量・表情 (感情) についても可視化されるので，会議の活発度を確認できるようになる．提案手法では，会議のネガポジ比率と表情の比率をレポートする機能が提供されている (図 5)．例えば，言葉のネガポジの比率から，この会議では，57.5%がポジティブであることを読み取ることができる．また，表情の比率から，ほとんどが neutral であるが，happy, sad の順に多いことが分かる．こうした会議のネガポジ比率・表情の比率の結果から，ポジティブな発言が多いことや，無表情で会議に参加しているかが分かる．

提案手法では，自然言語処理を用いて会議中の発話内容を解析し，その結果を WordCloud [8] とクラスタ分析により可視化する機能も提供されている．発話数の多い単語が大きく表示されるので，WordCloud により，会議途中や終了時に会議中の話題について振り返ることができる (図 6)．

WordCloud に表示される単語をカウントし横棒グラフで表示した結果が図 7 である．図に示すように，会議の話題から会議の中心的な話題について，どのようなものであったのかどうかをクラスタ分析の結果として見て取れる．

また，図 8 は，会議の活発さを，時系列折れ線グラフとして会議中の発話量と感情の動きの活発さが時系列で表示している．

© 2021 Information Processing Society of Japan

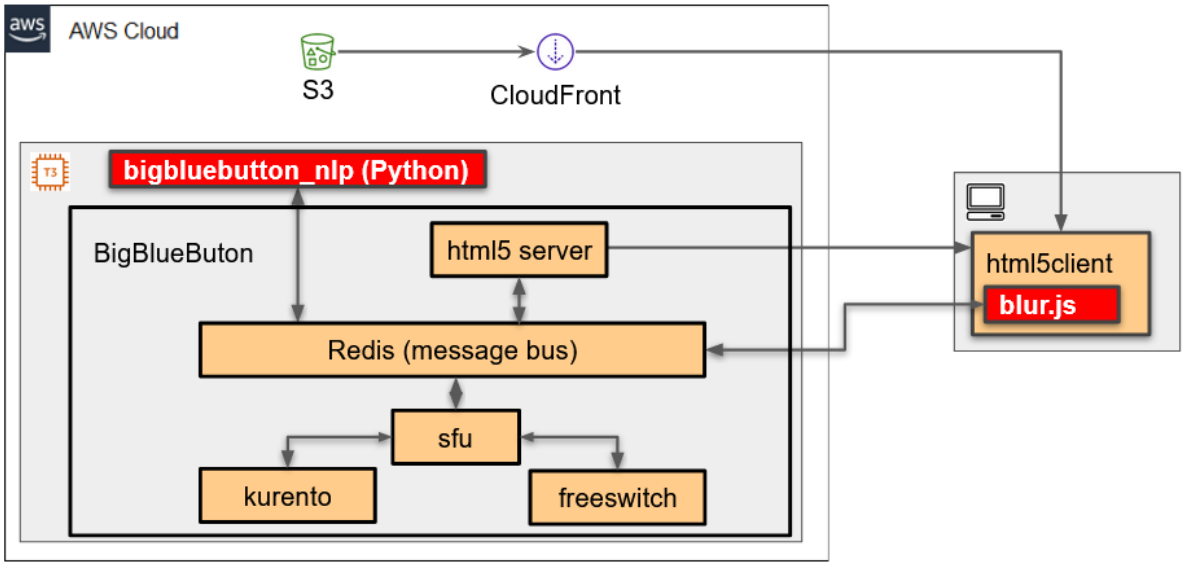

図 9：システムブロック図

図 10：アバター機能の実行例

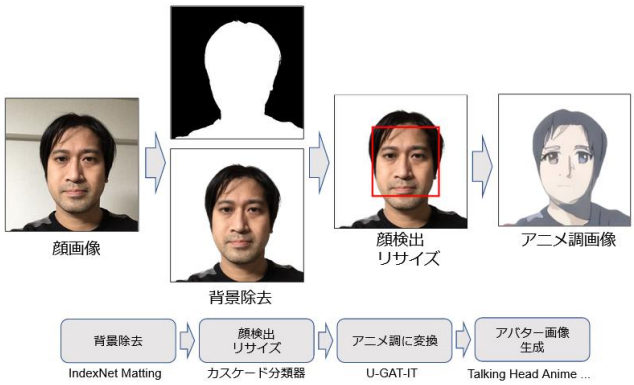

図 11：写真からアニメ調画像の生成

3. 実装

システムのブロック図を図 9 に示す．豊富な機能を持ち UI の完成度も高い BBB サーバを Amazon Web Service (AWS) 上に配置し，開発・評価環境として利用した．また，BBB に過度に依存した実装を避けるために実装上の工夫を行った．

3.1 アバター機能

アバター機能では，参加者の顔の動きを web カメラで検出する．ここで取得した顔画像は，Facemesh [5]を用いて，顔・口の動きや瞬きを解析し，これらの動作を本人の代わりとなるアバターに反映させる (図 10)．

2D アバターの画像生成には，Talking Head Anime from a Single Image [6] を利用している．この手法では，1 枚のアニメ調の画像と参加者の首の姿勢，目，口の開度などの状態をパラメータとして与えると，これらの状態に相当するアバターの画像を生成される．また，アニメ調画像のアバター生成は，深層学習を用いているため，リアルタイム表示が困難である．そのため動的に生成するのではなく，事前に 2D アバターを生成し，顔の動きにあったものを選択・表示させることにした．

また，自分に似たアバターを利用したい場合，自分に似たアニメ調画像をあらかじめ用意する必要がある．この画像生成の手間を省くため，顔写真をアップロードすると適切な加工を行いアニメ調画像に変換した上でアバターを生成する手法を実装した．アバターの生成機能について図 11 に示す．この図では，話者の 1 枚の実写の顔画像を元に本人に似たアニメ調画像を生成する一連のフローを示している．最初に顔画像に対して，背景除去やリサイズなどの前処理を行った後，U-GAT-IT [7] を利用することで，自分そっくりなアバターの元画像となるアニメ調画像 1 枚を生成する機能を実行できる．

3D アバターは，話者の顔のトラッキング情報から，3D CG アバターの顔・目・口の追従動作を再現した．2D アバターと異なり，3D アバターのモデルデータの対応する点を顔表情の特徴量の動作に合わせて移動させるだけで，リアルタイムでうれしい顔，悲しい顔などバラエティに富んだ表現が可能となる．

© 2021 Information Processing Society of Japan

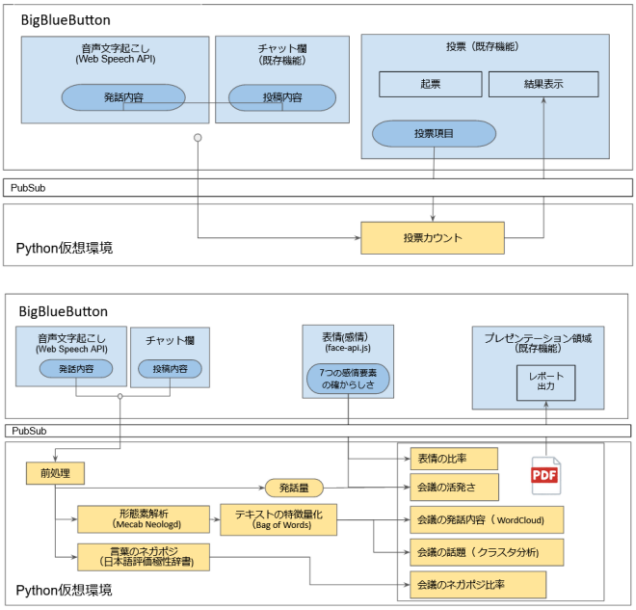

図 12：投票 (上) とレポート機能 (下) のブロック図

3.2 投票・レポート機能

投票機能は，Web Speech API [9] から発話内容，チャット欄から投稿内容と，投票機能から投票項目を受け取り，各自の発話内容・投稿内容の中に，どれだけ投票項目が含まれているかをカウントし，投票結果を BBB 上に表示する (図 12 (上)).

言葉のネガポジの比率は，python のライブラリ oseti [10] を用いて対話中の単語からネガポジをカウントして評価比率を表示した．また，ユーザの表情の比率については，2D・3D アバターでも使用している Face-API [11] から，neutral, happy, sad などの 7 つの感情要素の確からしさを取得した (図 12 (下)).

4. ユーザテスト

4.1 ユーザテスト概要

本システムの有効性を検証するために，簡易なユーザテストを実施した．システムをミーティングで利用している動画を被験者に視聴してもらい，有効性に関するアンケートに答えてもらった．アンケートは 8 問で構成され 5 分程度で終了し，各質問に対し 5 段階 (良く当てはまる=4 点，当てはまる=3 点，どちらとも言えない=2 点，当てはまらない=1 点，全く当てはまらない=0 点) で評価した．

動画視聴とアンケート入力で構成されるユーザテストは web 形式のオンラインで実施し，それぞれが所有する PC を使って実施した．オンラインミーティングを実施した場合，被験者が利用するネットワーク環境にストリーミングの質が依存してしまうことが想定されるので，利用している動画の視聴とした．アンケートの入力結果はユーザテスト実施者にメールで送信され，集計される．

表 1：アンケートの評価結果

有効性の視点	概要	評価※
目的適合性	プロダクトは問題点を解決していると言えるか？	★★
適時性	今プロダクトを導入することは適切なタイミングと言えるか？	★★
利便性	オンラインミーティングを行う上で、便利なツールとなっていると考えられるか？ 使ってみたいと思うか？	★★
有用性	スムーズに無駄なくオンラインミーティングを行うツールとなっていると言えるか？	★★
操作性	ユーザの操作が間違いを起こしにくく、操作しやすいツールであると考えられるか？	★★
機能性	オンラインミーティングに役立つ機能を有していると言えるか？	★★★
性能	オンラインミーティングを計画どおりに進めるうえで処理能力は十分だと言えるか？	★★
拡張性	オンラインミーティングの変化に容易に対応できるツールであると言えるか？ オンラインミーティングシステムの発展に寄与する要素があると言えるか？	★★★

※ 評価はアンケートの選択肢を下記の通りに数値換算し、４点満点とした平均点を算出
良く当てはまる＝４点、当てはまる＝３点、どちらとも言えない＝２点、当てはまらない＝１点、全く当てはまらない＝０点

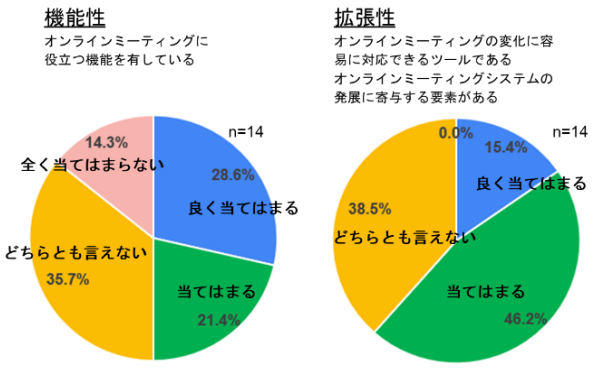

図 13：アンケート結果 (5 段階評価)

4.2 ユーザテストまとめ

提案手法と既存手法の比較ができることを想定し，リモートミーティングシステムを日常的に利用するとともに，その操作にも精通している大学院生 14 人を被験者とした．

表 1 は，アンケートの評価結果を示す．目的適合性，利便性，有用性，性能などの多くのアンケートの質問に対して，どちらともいえない，という回答が多く見られた．

一方で，機能性 (オンラインミーティングに役立つ機能を有しているといえるか?) と拡張性 (オンラインミーティングの変化に容易に対応できるツールであるといえるか?) に関しては，一定の評価を得ることができた．

図 13 に機能性と拡張性のアンケート結果の詳細を示す．機能性に関しては「良く当てはまる」と「当てはまる」で 50.0%を占め，拡張性に関しては「良く当てはまる」と「当てはまる」で 61.6%を占めた．

5. 考察

5.1 システム

実装基盤として BBB を使用し，そこにアバター機能および投票・レポート機能を追加することでシステムの実現を図った．

本システムのアバター機能の実現には顔パーツの位置

© 2021 Information Processing Society of Japan

検出が必要であり，特に 2D アバターの場合は 2D 画像の生成および選択，3D アバターの場合は 3D モデルの描画が必要になる．また，投票・レポート機能では，発話内容の語句解析および表情解析による感情取得が必要になる．本システムでは，これらの処理をブラウザで実行することで，特定のプラットフォームに依存せず，様々な環境で実行できるようにした．加えて，ブラウザで主要な処理を行うことでサーバ側の負荷が軽減され，比較的性能が低いサーバ環境でも動作させることが可能となった．

アンケート結果からの考察として，二つの課題が明らかとなった．１つ目は，参加者の表情・様子・意思の読み取りが難しい点であり，2 つ目は，表情や場の雰囲気から相手の意見・考えを読み取ることが難しく，議論をしている時に意見の割れ方をつかむのが難しいという点である．

まず，１つ目の課題に関連したアンケートでは，本システムのアバターの表現からは「微細な表情や意思が伝えられているのか分からない」,「表情の変化を目立つように，アバターならではの機能があるとよい」,「瞬きが気になる」という意見があった．これらの意見から，瞬きなどの不要な要素は表現せず，その代わりに必要な微細な表情の表現を加えると共に，喜怒哀楽の感情をオーバーな表情でアバターに表現させるなどの対策が効果的であるものと考える．

次に，二つ目の課題に関連して，アンケートでは，「意見の理由の可視化」に懐疑的な意見があった．投票だけでは相手の意見を知るのに不十分であり，発話内容に対して，より高度な構文解析・意味解析・文脈解析をすることで，「意見の理由」を可視化することが必要であるという課題が明らかになったと考える．

以上から，アンケートの結果などを含めて総合的に評価すると，色々な難しい課題はあるものの，役立つ機能があるとの回答もあり，オンラインミーティングの課題に対して，「アバターによる表情の見える化」と「発言の可視化」するアプローチは，ミーティングシステムの意思疎通を促すための有効的な手段であるといえる．

リモート会議システムが，利用される機会が増える中で，参加者同士の意思疎通が困難で意見の分かれ方を把握するために，相手の表情・動作や議論の内容から意見の読み取りを可視化するための解決手段として，アバター機能と投票・レポート機能を高度に実現することで，本ミーティングシステムが貢献できると考える．

5.2 課題

本システムで提案した 2D アバターの生成手法には深層学習が必要であり，GPU 等のアクセラレータが無い環境では主に処理速度の観点から実現が困難であった．そのため画像の事前生成および顔の動きにあった画像を選択・表示することで GPU 等が無い環境でも動作できるようにした．事前生成画像を減らすために，顔の動きに一致する画像を表示するのではなく，顔の動きをトリガとしてあらかじめ定義しておいたアニメーションを表示する手法にしたが，例えば，瞬きが気になる等の意見がアンケートから得られている．定義するアニメーションの内容は更なる検証が必要である．

投票・レポート機能では WebSpeech API による発話内容の語句解析および表情検出による感情の表示，頻出した単語を WordCloud で表示することで会議の話題を図示している．これらの情報を利用することで，会議終了時の振り返りやグループワークでの各チーム間での確認に利用できると考える．しかしながら，発話の速度や滑舌による誤認識，発話直後の単語が認識されない等の問題もあった．

6. まとめ

本論文ではカメラを ON にして顔を見せなくても場の雰囲気や相手の表情，意見を伝えることのできるオンラインミーティングシステムの仕組みを述べた．本人の顔の動きに追従する 2D アバターと 3D アバターにより，視覚的な意思伝達を実現した．また，投票機能とレポート機能により場の雰囲気や意見を可視化した．

今後は，参加者同士の意思疎通を促し，意見の分かれ方を把握するために，アバターの表現の高度化と構文解析・意味解析・文脈解析意見の可視化を進めていく必要がある．特に，効果的なアバター表現やミーティングの意見の可視化のバリエーションを増やすことで，ミーティングシステムの付加価値を高めることを目指す．

参考文献

[1] Zoom
https://zoom.us

[2] Google meet
https://apps.google.com/intl/us/meet/

[3] FaceRig
https://store.steampowered.com/app/1364390/Animaze_by_FaceRig/

[4] BigBlueButton
https://bigbluebutton.org/

[5] Facemesh
https://github.com/tensorflow/tfjs-models/tree/master/facemesh

[6] P. Khungurn. Talking Head Anime from a Single Image,
https://pkhungurn.github.io/talking-head-anime/

[7] J. Kim, M. Kim, H. Kang, and K. Lee. U-GAT-IT: Unsupervised Generative Attentional Networks with Adaptive Layer-Instance Normalization for Image-to-Image Translation, In Proceedings of ICLR ’20, 2020.

[8] F. Heimerl, S. Lohmann, S. Lange, and T. Ertl. Word Cloud Explorer: Text Analytics Based on Word Clouds, In Proceedings of IEEE HICSS ‘14, pp. 1833-1842, 2014.

[9] Web Speech API
https://wicg.github.io/speech-api/

[10] oseti
https://github.com/ikegami-yukino/oseti

[11] Face-API
https://github.com/justadudewhohacks/face-api.js/

© 2021 Information Processing Society of Japan