

オンラインミーティングを対象とした参加者のエンゲージメント計測システムの検討

Request of the Manuscript for IPSJ Kansai-Branch Convention 2018

曾根田 悠介[†] 荒川 豊[†] 安本 慶一[†]
Yusuke Soneda Yutaka Arakawa Keiichi Yasumoto

1. はじめに

日本では、働き方改革を行うため様々な制度が設けられている。勤務時間を減らす他にも、仕事の質の向上が重要視されている。株式会社ジェイアール東海エージェンシーの調査結果によると、日本の一般的な企業において、1日の勤務時間のうちミーティングが占める時間は平均で68.2分とミーティングは1日の勤務内容において大きな要素となっている[1]。ミーティングの質の定量的な評価が今後の働き方改革において重要になると考える。音声や映像のログを記録し、後から参照できるようなサービスや機能も多々存在するが、ミーティングに対するフィードバックを各自で行うのは、利用者にとって負担になる。また、ミーティングの評価に関するガイドラインも複数存在するが、定性的な指標がほとんどであり定量的にミーティングを評価することは困難と考えられている。ミーティングの定量的な評価が可能になれば、ミーティングはより活発なものになると期待できる。

また、リモートワーカーやテレワーカーの増加により、オンラインでのミーティングが一般的になりつつある。しかし、会議室等で机を囲い直接行うミーティングがまだ主流である[1]。働き方の多様性が広がるにつれ、今後さらにオンラインでのミーティングは重要になると考える。また、オンラインでのミーティングを積極的に導入している企業は業績が向上しているという報告もある。オンラインミーティングの質を向上させ、また対面でのミーティングよりもオンラインでのミーティングであることのメリットを付加させることが今後オンラインミーティングの普及に大きく貢献すると考える。

本研究では、PCを介したオンラインミーティングを対象とし、ミーティングの質の定量的な評価とミーティングに対するエンゲージメントの測定方法について結果を報告する。オンラインミーティングに限定することでユーザ毎の音声情報と映像情報を容易に取得できる。本研究の狙いは、これらの情報を解析することによって、どのようなデータがミーティングの定量的な評価に繋がる指標になるかを明らかにすることである。

本稿の構成は以下の通りである。第2章では、関連研究について述べる。第3章では、ミーティングデータの解析システムについて述べる。第4章では、ミーティングを評価する為の指標の検討について述べる。最後に、第5章で本稿のまとめと今後の課題を述べる。

2. 関連研究

Jonkerらの研究によると、ミーティングの討論内容が目的の内容に整合しているかを判定するために、ドメイン情報に基づくプロトコル解析の手法を提案している[2]。この手法では、発話の内容を解析すること、しかも、精度の高いテキスト解析を行う必要がある。本研究のミーティングの定量評価では、発話の内容には触れず、声の大きさや会話の間といったデータを指標とする点で異なっている。

堀らは、全方位カメラと指向性マイクを組み合わせ、会議参加者の位置推定や話者の特定を行なっている[3]。しかし、

複雑なアルゴリズムや学習済みのモデルが必要であったり、一般的な企業には導入が難しい設備の導入など実用上困難な部分が多い。また、ミーティングの質をどのように評価を行うかについてまでは触れられていない。会議室で行われるような一般的なミーティングでは解析自体も困難であり、様々な技術を組み合わせる必要がある。

宇佐美らは、音声入力を解析することによってミーティング中における参加者の関係やミーティングの活発度の定量化を試みている[4][5]。この研究は、スマートフォンを用いたミーティングを想定しており、話者の特定に複雑なアルゴリズムを用いる必要もない。また、豊田らは発話状態を解析し、対話における雰囲気の評価を試みている[6]。

本研究で提案するシステムでは、ミーティングへの参加者はそれぞれのパソコンを使用している前提のため、話者の位置を特定する解析であったり、話者の特定など複雑な音声分析を行う必要がない。また、PCを利用しているため音声のみでなくPCのフロントカメラから取得できる映像から、ミーティングの解析を行うことが可能であるため、導入が容易である。

3. ミーティングデータの解析システム

本章では、ミーティングをセンシングし、参加者の状態や会議の良さを数値化するシステムを提案する。

今回は、問題の簡略化のため、ミーティングの中でもオンラインミーティングに限定する。各参加者は、それぞれがカメラを備えたPCからアクセスするものとし、複数人が1つのPCからアクセスすることは想定しない。このシステムを、WebRTC[7]というオープンソースを利用し実装する。WebRTCはリアルタイムコミュニケーション用のAPIであり、Webブラウザ間のビデオチャットが可能になる。通信に関するAPIのみではなく、ブラウザが起動している端末のマイクやカメラの情報を取得することも可能である。

これらのデータを各自のPCもしくはクラウドサービスで解析を行い、ミーティングを数値化する。数値化されたデータはサーバを介してユーザ間で共有され、それぞれの画面に表示される。クラウドサービスの例としては、表情の分析を行うためにMicrosoft Corporationが開発しているAzureのCognitive ServicesのFaceAPI[8]などの使用を検討している。

ミーティング終了後にそのミーティングを解析したデータを表示する。また、あとからミーティングの評価を行うために、各端末の音声情報のログとひとつのPCの画面をキャプチャーした動画をサーバー上に保存する。

4. ミーティングを評価する為の指標の検討

ミーティングは、音声情報と映像情報から構成されている。これらを解析することで、参加者ごとの発話量や発話権の移行順序、領き回数など様々な定量的数値データを得ることができると考えられる。しかしながら、こういった数値データが考えられるのかは、既存研究において網羅的に検討されていない。また、それらの数値データがミーティングの良し悪しと関連しているのかも不明である。本研究では、前者にフォーカスを当て、ミーティングというイベントからどのような定量的数値データ（以降、指標と呼ぶ）が得られるのかを検討する。

[†] 奈良先端科学技術大学院大学, Nara Institute of Science and Technology

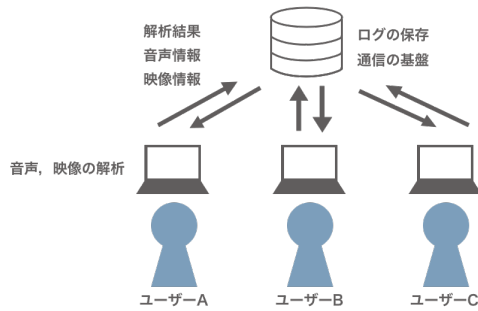


図 1: 提案システムの構成図

4.1 音声情報に基づいた指標の検討

音声情報は、発話量や声のトーンなどの外形的な数値データと、発話の内容という文章データから構成される。本研究では、プライバシーやセキュリティに配慮し、外形情報のみを取り扱う。

4.1.1 発話支配率

ミーティング全体を通しての総発話時間に対して、各ミーティング参加者の発話した時間の割合を発話支配率と定義する。この指標から、どのような人物がそのミーティングにおいて大きな影響力を持っているのかを解析することができると考える。

4.1.2 発話継続時間

一度の発話に対して途切れずどれだけ発話を継続したかを計測する。ミーティングの状況に応じて、一度の発話で長い時間発話した方が適切なのか、ファシリテーターの発話が長い方が適切なのかといった分析が可能と考える。

4.1.3 発話主導権の移り変わりによるミーティング参加者のネットワーク解析

同時に複数の参加者が発言しないという前提のもと、発話する権限の移り変わりからそのミーティングに参加している者の発話の移り変わりをグラフ化することが可能と考える。この様な解析によって、どのような人がミーティングにおいてどのような役割を果たしているのか、その人物のミーティングへのエンゲージメントの計測に役立つ指標になると考える。

4.2 映像情報に基づいた指標の検討

ミーティングにおいて、非言語的なコミュニケーションを解析することも重要になると考える。PCのフロントカメラから得られた画像からどのような非言語的な行動を解析できるかを提案する。

4.2.1 感情

3章で説明したように、Microsoft CorporationのFaceAPIから取得した顔の表情の解析結果から、ミーティングの雰囲気や推定を試みる。また、表情から会議の内容を理解できているのかについて推定を試みる。

4.2.2 傾き

大塚の研究によると、傾きや顔の向きといったパラメーターが非言語的なコミュニケーションとしてミーティングの活発度を向上させるものであることが明らかになった[9]。誰かの発言に対する肯定、理解の意味を示す傾きは、ミーティングの分析の上で重要な指標になると考えられる。

これらの指標から、ミーティングの質の評価を試みる。また、各指標ごとの組み合わせなども考慮していきたい。

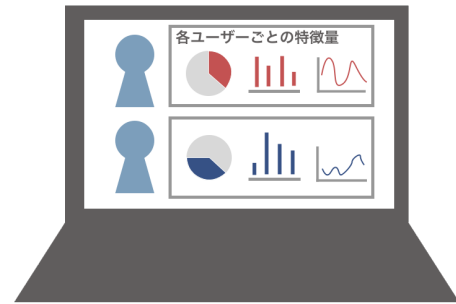


図 2: ミーティング終了後に表示される解析結果のイメージ図

4.3 ミーティングの状況の考慮

ミーティングの状況によって、4.1, 4.2で考案したような指標の評価や重要性は異なると考えられる。本論文では、ミーティングの状況を以下の三つの要素の組み合わせとして定義する。

- ・ミーティング参加者同士の関係
- ・ミーティングの目的
- ・ミーティングの参加人数

ミーティング参加者の関係とは、上司とその部下といった主従関係のもの、同僚同士といった対等な関係性、その場で初めて知り合った人同士など様々な状況が考えられる。またファシリテーターと呼ばれる、会議を中立的な立場で取りまとめより会議を活発なものにする役割で参加する人物も想定できる。

ミーティングの目的とは、アイデアの量を求めるブレインストーミングや2択の選択肢のどちらかをミーティングの総意として求めるミーティング、専門的な知識を他の人に伝達する状況などが考えられる。

ミーティングの参加人数は、大人数のものではなく、2人から4人を想定する。

例えば、「上司からその部下に専門的な事項を伝達する様なミーティング」と「同僚同士である事項について討論するミーティング」を考える。前者の場合、上司が専門的な知識を所有しているため、発話支配率は上司の方が部下よりも高くなるのが妥当と考えられる。一方後者の場合、一人の意見に偏ることがない様に意見を集約させるために一人の発話支配率が極端に高くなるのはあまり好ましくないと考えられる。この様に、ミーティングの状況が異なると発話支配率の評価の仕方が異なると考えられる。

5. おわりに

本研究では、今後より一般的になると考えられるオンラインミーティングを対象にし、ミーティング参加者のエンゲージメントをどのように計測できるのかを検討した。音声情報と映像情報からミーティング参加者の様々な情報を取得することができる。どのような情報がミーティングの解析において重要な指標になるかをさらに今後も検証して行く必要がある。

今後は、本研究で提案したシステムの実装を行い、実際に提案した指標がどれほどの精度で取得できるかを検証していく。さらに、指標の間にどのような関係があるのか調査を行う。

また、ミーティングの状況に応じて、抽出したデータからそのミーティングの質の評価を定量的に行うことが可能と考える。取得した指標の情報から、ミーティングの質を評価するモデルの生成も検討していく。

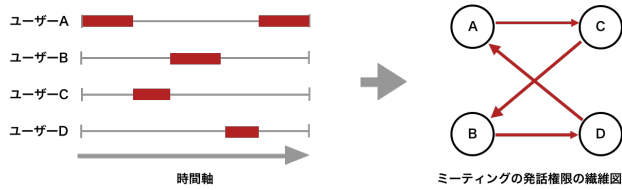


図 3: 発話主導権の遷移図

謝辞

本研究の一部は、JST さきがけ (16817861) の支援を受けて実施されたものである。

参考文献

- [1] 株式会社ジェイアール東海エージェンシー. ビジネスパーソンの「社内会議」に関する調査. ビジネスパーソンウォッチング調査, Vol. 14, , 2017.
- [2] Catholijn M. Jonker, Schut Martijn, Treur Jan, and Yolum POnar. Formal analysis of meeting protocols. *LANI*, Vol. 3415, pp. 114–129, 2005.
- [3] Hori Takaaki, Araki Shoko, Yoshioka Takuya, Fujimoto Masakiyo, Watanabe Shinji, Oba Takanobu, Ogawa Atsumori, Otsuka Kazuhiro, Mikami Dan, Kinoshita Keisuke, Nakatani Tomohiro, Nakamura Atsushi, and Yamato Junji. Low-latency real-time meeting recognition and understanding using distant microphones and omni-directional camera. *IEEE*, Vol. 20, No. 2, pp. 499–513, 2012.
- [4] 宇佐美格 王亜楠 高橋淳二 齊藤 裕樹 戸辺義人. 携帯端末を用いたミーティング定量評価システムの構築. 情報処理学会論文誌, Vol. 57, No. 2, pp. 553–561, 2016.
- [5] 王亜楠 宇佐美格 狐崎直文 戸辺義人. Kaihui: スマートフォンを用いたミーティング活発度定量化の試み. 情報処理学会研究報告, Vol. 2015-MBL-74, No. 42, 2015.
- [6] 豊田薫 宮越喜浩 山西良典 加藤昇平. 発話状態時間長に着目した対話雰囲気推定. 人工知能学会論文誌, Vol. 27, No. 2, pp. 16–21, 2012.
- [7] Google LLC. <https://webrtc.org/>.
- [8] Microsoft Corporation. <https://azure.microsoft.com/ja-jp/services/cognitive-services/face/>.
- [9] 大塚和弘. キネティックテレプレゼンス mmspace 小規模グループ対グループ遠隔会話場の構成と非言語行動解析. 電子情報通信学会, pp. 610–620, 2018.