

ビデオ会議における会話の流れを可視化するシステムの検討

今井 廉¹ 呉 健朗² 富永 詩音³ 木村 悠児¹ 酒井 知尋² 小島 一憲² 宮田 章裕^{1,a)}

概要: ビデオ会議システムが普及してきており、リモートワークなどオンラインで議論する機会が以前より増加してきている。ファシリテータによる会議の円滑化は対面会議のみならず、ビデオ会議でも重要である。ところが、ビデオ会議では視線方向・頷きなどの情報が伝達されにくいいため、会議参加者にとって他者の発言・反応が誰に対するものなのか把握しにくく、記憶にも残りにくいと考えられる。こうした背景から対面での会議に比べビデオ会議の円滑化はファシリテータにとって困難である。我々はファシリテータによる会議の円滑化を支援するため、会議参加者の会話の流れに着目し、有向グラフを用いて会話の流れをリアルタイムに可視化する手法を提案している。本稿ではビデオ会議における会話の流れを可視化するシステムの実装の改善案について検討を行った。

1. はじめに

ビデオ会議システムが普及してきており、リモートワークなどオンラインで議論する機会が以前より増加してきている。一般的な会議では、参加者に発言を促したり話の流れをまとめたりして会議全体の円滑化を行う、ファシリテータという役割が設けられることが多い。ファシリテータによる会議の円滑化は、対面会議・ビデオ会議問わず行われているが、一般的にビデオ会議のような分散環境で行われる会議の円滑化は対面会議と比べて難しい [1]。ビデオ会議の円滑化が難しい一因として、ノンバーバル情報の欠落が考えられる。特にビデオ会議では視線方向・頷きなどの情報が伝達されにくいいため、会議参加者にとって他者の発言・反応が誰に対するものなのか把握しにくく、記憶にも残りにくいと考えられる。これではファシリテータは会議全体の流れを追ったり、発言をまとめたりすることは難しく、会議の円滑化は困難である。そこで我々は会議参加者の会話の流れに着目し、これを有向グラフを用いてリアルタイムに可視化する手法を提案している [2]。本稿では、会話の流れを可視化するシステムの実装の改善案に関する検討の報告をする。

2. 関連研究

本研究は会議の支援を行う研究領域に属する。関連研究は次のように大別できる。

2.1 会議参加者の行動を可視化する研究

会議における会議参加者の行動を可視化することで会議の支援を行う研究が行われている。これらの研究では、会議参加者の発言時間 [3][4][5]、発言回数 [6]、発言速度 [7][8]、発言内の単語 [8][9]、発言率、無音率、ターンテイク数 [6]、発言内の単語などが可視化されている。会議参加者の行動を可視化する研究は概ね全会議参加者の会議参加度の均等化や、遠隔会議をより対面環境での会議に近づけることを目的としている。会議参加者の行動の可視化により会議参加者の参加度の均等化がなされたことを検証実験で明らかにした研究もある [3][4]。同様に、遠隔会議をより対面環境での会議に近づける効果 [5] や会議参加者の主観的評価と議論状況の相関 [6]、自身の発言速度を意識させる効果 [7] も確認されている。また、会議参加者の行動の可視化をリアルタイムに行う研究だけでなく、会議後にシステムが会議参加者に対し可視化内容を提示する研究 [10][11] も行われており、対話プロセスの評価が可能であることが示唆されている。他にもフィードバックを行う対象とその形式について調査 [12] が行われており、視覚によるフィードバックよりも触覚によるフィードバックの方がターンテイキングを促すこと、潜在的話者に対してフィードバックを行うよりも、全ての会議参加者、現行話者に対してフィードバックを行う方がターンテイキングを促すことが示唆されている。

2.2 システムによる会議の円滑化を目指す研究

システムによる適切な会議の円滑化が行えた場合、ビデオ会議などの人間のファシリテータによる会議の円滑化が

¹ 日本大学文理学部

² ソフトバンク株式会社

³ 日本大学大学院総合基礎科学研究科

a) miyata.akihiro@acm.org

困難であるシーンにおいて優れた手段となり得ると考えられる。システムによる会議の円滑化を目指すために、会議参加者の発言の意図を推定する試み [13] やファシリテータの行動の分析・分類 [14] が行われている。これらの研究をもとにシステムによる人と同程度あるいは人以上の会議の円滑化がなされるようになることが期待される。実際にシステムがファシリテータを務めている研究 [15][16] も存在しており、システムによる会議参加者に向けた指示によって、会議参加者の参加度に影響を与えたことが示唆されている [15]。Limayem らの研究 [16] では、学生を対象として人間のファシリテータによる会議の円滑化とシステムによる会議の円滑化の比較を行っている。実験結果からはシステムも人と同程度の会議の円滑化が行えるということが示唆された。

3. 研究課題

ファシリテータにとって会議が滞ることは望ましくない。会議が滞らないように、ファシリテータは会議の円滑化を行うが、それには何らかの判断材料が必要である。例えば、オフラインの対面会議であれば、会議参加者はお互いに実空間を共有しているため、視線、姿勢、ジェスチャーといったノンバーバル情報から、ファシリテータは参加者の参加度の把握、参加意欲の推測ができる。このため会議の円滑化も比較的行いやすいと思われる。しかし、ビデオ会議となると、視線方向などのノンバーバル情報の欠落が生じてしまう。ファシリテータが会議参加者の発言、行動が誰に向けたものなのか把握することは容易でなく、記憶にも残りにくいと考えられる。結果的にファシリテータは判断材料が乏しい状態で会議の円滑化に努めなければならない。

この問題に対して、関連研究で行われているように、会議参加者の行動の可視化 [3][4][5][6][7][8][9][10] を行うことでファシリテータに判断材料を提供することも可能である。研究 [3][4][5] では、可視化内容を見ることでファシリテータは会議参加者の会議への参加度を把握し、これに基づいて参加者の参加度の均等化が行える。しかし、参加者の発言関係は把握しにくいいため、一部の会議参加者間のみ発言のやり取りが多くなってしまふことが危惧される。研究 [6] では、発言量の他に無音率、話者交代回数を可視化している。話者交代の可視化により会議の流れを把握できると思われるが、ビデオ会議における会議参加者の行動、発言が誰に向けたものなのかは依然把握しにくく、ファシリテータの記憶にも残りにくいままであると考えられる。研究 [7][8] では、発言速度を可視化しており、ファシリテータは会議参加者の発言速度の把握はできるが、会議参加者の発言関係は把握しにくい。研究 [9] はチャットベースの会議を想定したシステムであり、音声が多く用いられるビデオ会議の場での効果は定かでない。研究 [10] では発言

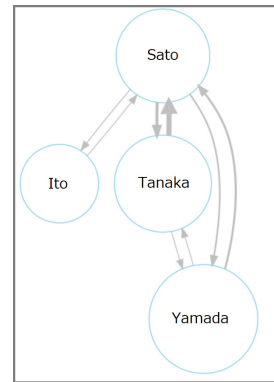


図 1 提案手法により生成されたグラフ

量、発話順に領きといった身体動作を含めて可視化しているが、会議の場でのリアルタイムの運用には至っていない。システムによる会議の円滑化を目指す研究 [13][14][15] をビデオ会議の場で用いることも考えられる。しかし、研究 [13][14] では準備段階であり会議の円滑化をリアルタイムで行うシステムの作成には至っていない。研究 [15] では、システムと人それぞれによる会議の円滑化の比較はなされていない。そのためシステムが人と同程度の会議の円滑化が行えるか明らかにされていない。研究 [16] では、システムによる人と同程度の会議の円滑化が可能であることが示唆されたが、論文内でも述べられているようにシーンが限定的であり、この知見が一般化できるかは定かでない。

本章の冒頭で示したように、ビデオ会議では一部のノンバーバル情報が欠落するため、ファシリテータは会議の円滑化を容易に行えないという問題が存在する。そこで我々は、以下の要件を満たすファシリテータの会議の円滑化を支援するシステムの構築を研究課題とする。

- ビデオ会議参加者の発言関係を可視化する。
- ビデオ会議参加者の発言量を可視化する。

4. 提案手法

本稿ではユーザの発言方向・発言回数・発言時間を総じて会話の流れと定義する。3章で設定した研究課題を達成するため、我々は会話の流れを可視化する方法として有向グラフに着目した。有向グラフでは、エッジが方向性を保持している。また、エッジの重みによってエッジの表現する対象の量を表すことができる。このため、有向グラフは会議参加者の発言方向、発言回数、発言時間を一度に可視化することができると考えられる。以上より、我々は会話の流れを有向グラフを用いてリアルタイムに可視化する手法を提案している [2]。有向グラフでは、発言者からその発言の相手に対してエッジを描く。この手法により、会話の流れの可視化が行われ、ファシリテータは会議参加者間での程度会話が行われているか、誰が発言過多・発言過少であるか、といった会議の円滑化のための判断材料を獲得できると考えられる。

5. 実装

4章の提案手法の効果を検証するため、システムの実装を行う。このシステムはサーバクライアントシステムとして実装する。様々なビデオ会議システムと併用できるよう、Webブラウザ上で動作するようにする。システムは、ユーザの発言によりグラフを生成する。生成したグラフをシステムがWebブラウザ上に表示することで、ユーザに可視化内容を提示する。提案手法により生成されたグラフを図1に示す。グラフのノードは参加者を表しており、エッジの向きは誰に対する発言かを、エッジの太さは発言量を表している。

システムの構成は次のとおりである。クライアント部では、音声の取得、発言開始終了時刻の送信、グラフの出力を行う。サーバ部では、発言開始終了時刻の受信、グラフの生成、グラフの送信を行う。5.1節にてクライアント部、5.2節にてサーバ部の詳細な説明を行う。

5.1 クライアント部

Webブラウザ上でマイクからユーザの発言開始時刻・発言終了時刻を取得する。音声を取得した際に、無音区間が200 msec以下の場合には連続音声区間として判定する。連続音声区間の開始時刻を発言開始時刻、終了時刻を発言終了時刻、長さを発言時間とする。取得した発言開始時刻・発言終了時刻・発言時間をサーバへと送信する。サーバへの送信後、レスポンスとしてサーバからグラフ画像が送信される。受信したグラフ画像をWebブラウザ上に表示する。

5.2 サーバ部

サーバはクライアントから発言開始時刻・発言終了時刻を受信する。受信後、グラフの生成を行う。グラフには、サーバに接続しているクライアントをノードとして描画する。各ノードがどのクライアントのものか区別するため、ノードの中にはクライアントの名前を表示する。会話の流れを表現するため、グラフに有向エッジを描画する。有向エッジの描画アルゴリズムは次のとおりである。

エッジの起点

エッジの起点は現在の発言者の名前が表示されているノードとする。

エッジの終点

先行研究では現在の発言者が直前の発言者に対して発言を行ったものとみなしていた。このため、直前の発言者の名前が表示されているノードをエッジの終点としていた。しかしながら、実際の会議における会議参加者の発言は直前の発言者のみでなく、過去の一定時間内の発言者に向けた場合もあると考えられる。このことから現在の発言者が過去の一定時間内の発言者全員に向けて発言を行ったものとみなすことにする。以

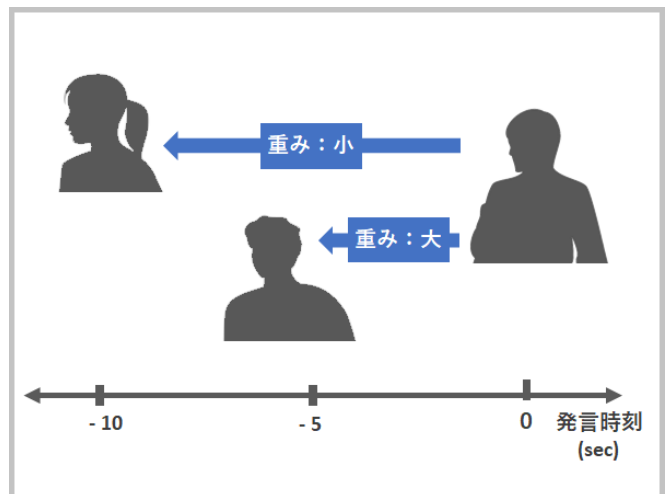


図2 発言時刻による重みづけの一例

上を踏まえ過去の一定時間内の発言者の名前が表示されている全てのノードをエッジの終点とする。

エッジの太さ

有向エッジで発言量を表現するため、エッジの太さに重みづけを行う。重みづけの際の基準は次のとおりである。

発言回数・発言時間

発言の回数が多いほど、発言量も多いとみなせると考えられる。このため発言回数による重みづけを行う。また、長い発言と短い発言の会議の場への影響力が等価なものであるとは考えにくいことから、発言時間に比例する重みづけを行う。

発言時刻

サーバは会議参加者の発言終了時刻を記録しておく。サーバはクライアントから新たに発言開始時刻・発言終了時刻を受信した際、記録してある過去の発言終了時刻と受信した発言開始時刻の差(sec)を求める。この差が大きくなるほど減衰する重みを付与する。この概念を図2に示す。

最後に、サーバは生成したグラフをクライアントへ送信する。

6. おわりに

実際の会議における会議参加者の発言は直前の発言者のみでなく過去の一定時間内の発言者に向けた場合もあることから、従来のビデオ会議における会話の流れを可視化するシステムのエッジ描画処理には改善の余地があった。このため、本稿ではビデオ会議における会話の流れを可視化するシステムの実装の改善案について検討を行った。今後は検証実験を行い、ビデオ会議における会話の流れを可視化するシステムがビデオ会議の円滑化に与える影響を明らかにしていく予定である。

参考文献

- [1] Niederman, F., Beise, C. M., and Beranek, P. M.: Facilitation issues in distributed group support systems, Proc. *SIGCPR'93*, pp.299–312 (1993).
- [2] 今井 廉, 呉 健朗, 富永詩音, 木村悠児, 酒井知尋, 小島一憲, 宮田章裕: 会話の流れの可視化によるビデオ会議ファシリテーションへの影響, ワークショップ 2020 (GN Workshop 2020) 論文集, Vol.2020, pp.1–7 (2020).
- [3] Samrose, S., Zhao, R., White, J., et al.: CoCo: Collaboration Coach for Understanding Team Dynamics during Video Conferencing, Proc. the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol.1, No.4, pp.1–24 (2018).
- [4] Adachi, H., Myojin, S., and Shimada, N.: ScoringTalk: a tablet system scoring and visualizing conversation for balancing of participation, Proc. *SA'15*, pp.1–5 (2015).
- [5] Kim, T., Chang, A., Holland, L., et al.: Meeting Mediator: Enhancing Group Collaboration with Sociometric Feedback, Proc. *CSCW'08*, pp.357–466 (2008).
- [6] 岡澤大志, 大山涼太, 石川誠彬, 望月俊男, 江木啓訓: 発言状況のリアルタイム可視化が議論への参加意欲に及ぼす影響, GN Workshop 2018 論文集, Vol.2018, pp.1–6 (2018).
- [7] Duan, W., Yamashita, N., and Fussell, S. R.: Increasing Native Speakers' Awareness of the Need to Slow Down in Multilingual Conversations Using a Real-Time Speech Speedometer, Proc. ACM on Human-Computer Interaction, Vol.3, No.CSCW, Article No.171 (2019).
- [8] Mathur, P., and Karahalios, K.: Visualizing Remote Voice Conversations, Proc. *CHI EA'09*, pp.4675–4680 (2009).
- [9] Leshed, G., Perez, D., Hancock, J. T., et al.: Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups, Proc. *CHI'09*, pp.537–546 (2009).
- [10] 根本啓一, 高橋正道, 堀田竜士, 林 直樹: 時間区間ごとの発話順序ネットワーク構造に着目したワールドカフェ型のダイアログの計測と評価, 情報処理学会論文誌, Vol.57, No.1, pp.248–259 (2016).
- [11] 宮田章裕, 林 剛史, 福井健太郎, 重野 寛, 岡田謙一: 思考状態と発話停止点を利用した会議の動画ダイジェスト生成支援, 情報処理学会論文誌, Vol.47, No.3, pp.906–914 (2006).
- [12] 市野順子, 八木佳子, 西野哲生, 小澤 照: グループディスカッション支援のための振動によるフィードバックの提示, 情報処理学会論文誌, Vol.60, No.4, pp.1171–1183 (2019).
- [13] 河野進, 相原健郎: グループ会話における発話意図の推定システム, 情報処理学会論文誌, Vol.58, No.5, pp.1113–1123 (2017).
- [14] 大本義正, 戸田泰史, 植田一博, 西田豊明: 議論への参加態度と非言語情報に基づくファシリテーションの分析, 情報処理学会論文誌, Vol.52, No.12, pp.3659–3670 (2011).
- [15] Tausczik, Y. R., and Pannebaker, J. W.: Improving teamwork using real-time language feedback, Proc. *CHI'13*, pp.359–468 (2013).
- [16] Limayem, M.: Human versus automated facilitation in the GSS context, ACM SIGMIS Database, Vol.37, No.2–3, pp.156–166 (2006).