

遠隔会議における円滑な議論を支援するための 興味推定モデルに関する検討

鳥山 英峻^{1,a)} 石田 繁巳² 白石 陽²

概要：遠隔会議では議論相手の興味・理解度・納得度を推定することが難しく、会議の長時間化や理解度・関与度の低下につながる要因となっている。このような遠隔会議特有の問題解決に向け、参加者が議論中にシステムを操作することで興味などを示させる研究が複数行われているが、会話への注意の逸脱など、円滑な議論を停滞させる要因となるため主体的操作を必要としないものが望ましい。一方、対面環境において、対象への興味、理解度、納得度の推定として、身体動作特徴や発話特徴を用いた研究が複数存在する。そこで本研究では、遠隔会議の議論においてもこれらが特徴として表れると考え、ユーザ操作が不要な意思疎通支援システムを実現する。本稿では、まず始めに、遠隔会議における参加者の議論内容に対する興味の推定に着目し、推定モデルに関する検討を行った。実験結果として、遠隔会議中の参加者の顔の水平方向の動き、姿勢の変化、感情の変化、発話頻度については議論内容への興味が特徴に表れやすく、推定に有効であることが示唆された。一方、本稿の調査では発話内容特徴については議論内容への興味が特徴に現れにくい結果となった。

キーワード：遠隔会議、身体動作特徴、発話特徴、予測モデル

1. はじめに

近年、リモートワークやオンライン授業などの普及による遠隔会議システムの利用が急激に増加しており、今後もユーザの使用率が増加していくことが予想される [1, 2].

遠隔会議では、相手の表情や仕草を汲み取りにくく、他の会議参加者の興味、理解度、納得度などの意思を把握することが難しい [3]. この問題は、議論の長時間化や会議参加者の理解度・関与度の低下につながり、円滑な議論を妨げる要因となっている [4, 5].

このような遠隔会議特有の問題解決に向けて、議論中、会議参加者にシステムを操作させて自ら興味などの意思を示させる研究 [6–11] が複数報告されている。しかし、議論中のシステム操作は会話への注意の逸脱など、円滑な議論を停滞させる要因となるため、主体的なシステム操作を必要としないものが望ましい。

一方、対面環境において、対象への興味、理解度、納得度を推定する研究が複数行われている。文献 [12–15] では姿勢、頭部運動、表情などの身体動作特徴から、文献 [16, 17] では発話頻度、発話内容などの発話特徴から興味などを推

定可能なことを報告している。

そこで本研究では、遠隔会議の議論においても興味、理解度、納得度が身体動作や発話に特徴として表れると考え、これらの特徴を議論中に利用した会議円滑化のための意思疎通支援システムを実現する。システムは遠隔会議参加者の身体動作特徴、および発話特徴に基づいて議論内容への興味などを推定し、他の会議参加者に共有・可視化を行う。また、支援を受けるにあたり、操作を不要にすることで議論の妨げにならないシステムを実現する。

本稿では、まず始めに、遠隔会議における参加者の議論内容に対する興味の推定に着目し、推定モデルに関する検討を行う。興味推定に有効と思われる特徴を検討し、実際の議論環境において、身体動作特徴と発話特徴を抽出して会議参加者の興味評価との相関関係を確認した。

実験結果として、遠隔会議中の参加者の顔の水平方向の動き、姿勢の変化、感情の変化、発話頻度については議論内容への興味が特徴として表れやすく、推定に有効な特徴であることが示唆された。一方、発話内容については、議論内容への興味が特徴として表れにくい結果となった。

本稿の構成は以下の通りである。2. では関連研究を基に本研究の位置づけを述べ、3. で提案する手法の全体像を詳細に述べる。そして、4. で評価実験について述べ、5. で評価実験の結果から提案システムの実現のため、考察を行い、

¹ 公立はこだて未来大学大学院 システム情報科学研究科

² 公立はこだて未来大学 システム情報科学部

a) g2122043@fun.ac.jp

6. でまとめとする。

2. 関連研究

関連研究として、まず 2.1 節では、人の行動から対象への興味を推定する研究事例について述べる。次に、2.2 節では、円滑な議論を実現するための複数の研究事例について述べる。

2.1 対象への興味推定に関する研究

身体動作から対象への興味を推定する研究は、複数報告されている [12-15]。文献 [14] では、タブレット端末で映像を見る被験者の様子を端末上のフロントカメラで記録したところ、流れる映像に対して、表れる興味の変化による顔の動きと姿勢の変化を観測している。文献 [15] では、公共の場に大型ディスプレイや、プロジェクタなどを設置し、情報を不特定多数に向けて発信を行うデジタル広告上で具体的なユーザの興味情報を収集するため、体の向きと注視情報を取得し効果測定に活かしている。文献 [13] では、雑談対話システムの研究・開発において、言語情報のみではユーザ興味の有無を把握することは困難として、表情に着目することで予測の精度を向上させている。

発話特徴から対象への興味を推定する研究は、複数報告されている [16, 17]。文献 [16] では、参加者が対話を通じて合意形成をする際、その対話に満足しているのか、合意結果に納得しているのかは、個人で異なると述べており、対話の満足度の評価を推定することを目指している。特に、発話内容から積極性と影響性の観点で評価しており、それぞれ議論の中心となる単語を含めた発話を積極性、合意形成への貢献を影響性としてそれぞれ評価している。文献 [17] では、発話頻度にも着目しており、議論内の重要単語が切り替わるまでの個人の発話総数は、議論への興味や結論への満足度に対して影響があると述べている。

2.2 遠隔会議の円滑化に関する研究

意思疎通の補助に着目した研究 [6-8] が報告されている。主体的なシステム操作による意思疎通の補助に着目した会議円滑化は「使用時に注目を浴びてしまう」など、匿名性の低さについての意見が被験者により述べられていることが多くある。文献 [7] では、ボタンひとつで簡単に意思を示すことが可能な、気持ち可視化ボタンを提案している。しかし、発言につながるボタンは使用頻度が平均して低かったことが報告されている。

様々な外部デバイスを用いた研究 [9-11] が報告されている。外部デバイスを用いた主体的なシステム操作による会議円滑化は「会話の注意を逸らされる」など、注意の逸脱についての意見が被験者により述べられていることが多くある。文献 [11] では、ジェスチャー表現を搭載した VR 空間での遠隔会議システムに着目している。しかし、対面で

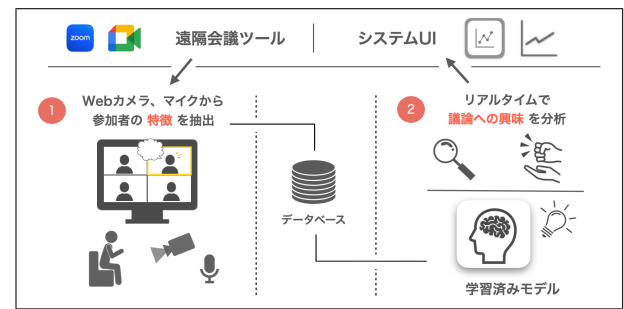


図 1: 提案システムの概要図

の会議と比較して、会話への注意の逸脱についての意見が被験者により述べられている。

議論中に提案システムを操作させる行為は会議を停滞させる要因となりうるため、本研究では操作を要求せず、会議中にストレスを生じさせないシステム使用感を実現する必要がある。興味・理解度・納得度をフィードバックするための UI に関しては検討の余地があるが、本稿では議論せず、遠隔会議において参加者特徴をインプットする UI、および興味推定手法の調査・検討を行う。

3. 提案手法

本章では、会議円滑化のための意思疎通支援を行うため、議論内容への興味を推定する提案手法を詳細に述べる。3.1 では、提案するシステムについての全体像を述べ、3.2, 3.3 では、提案システムの各要素である、興味推定システム、共有・可視化システムについて述べる。

3.1 システム概要

図 1 に提案手法のシステム概要を示す。提案システムは、興味推定システム、共有・可視化システムの 2 つから構成される。興味推定システムは、会議参加者の特徴検出を行い、特徴から現在の議論内容に対して、個人ごとの興味度合いを推定する。共有・可視化システムは、個人ごとの興味推定結果を集約し、会議参加者全体に向けて共有・可視化する。以降では、各処理の詳細について述べる。

3.2 興味推定システム

興味推定システムは、議論中の話題ごとに興味度合いを推定する。議論中の話題は文献 [16] に示された定義で特定する。議論の中で時間あたり最頻出となる単語を重要単語と定義し、重要単語の出現から別の重要単語に切り替わるまでをその重要単語に対応する 1 つの話題とする (図 2)。

図 3 に興味推定システムの概要図を示す。議論中のリアルタイム性とシステム使用時の手軽さを重視するため、興味推定システムでは Web カメラを用いる。議論フェーズ毎の興味を推定するため、興味推定システムは議論中の映像と発話音声から身体動作特徴、発話特徴を検出し、分析を行う。具体的には関連研究 [12-17] より、Web カメラの

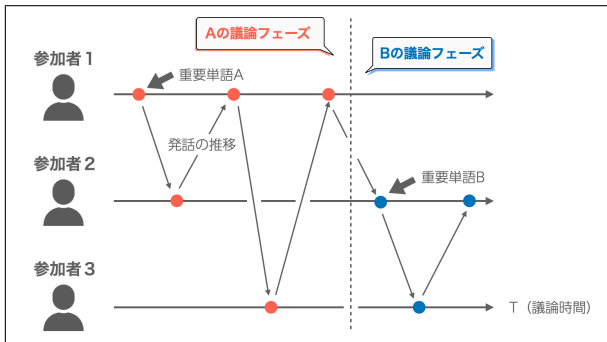


図 2: 議論フェーズについて

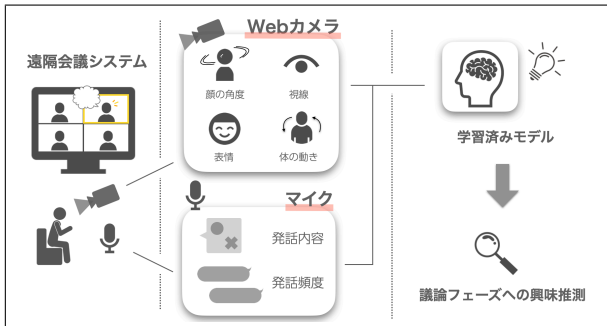


図 3: 興味推定システムの概要図

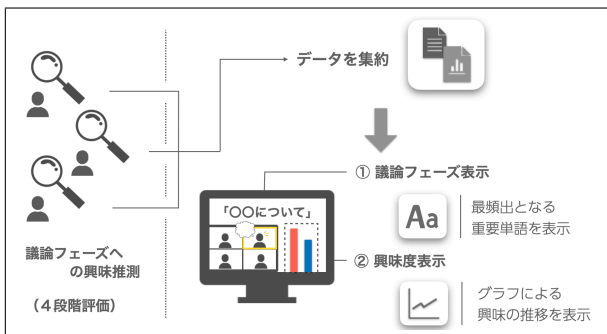


図 4: 共有・可視化システムの概要図

映像から、顔の角度、姿勢、表情（感情）、発話音声から、発話内容、発話頻度の特徴分析を行う。また、個人ごとの議論フェーズに対する興味分析の結果は「興味がある」「やや興味がある」「あまり興味を持ってない」「興味を持ってない」の4段階を想定している。

3.3 共有・可視化システム

図 4 に共有・可視化システムの概要図を示す。文献 [8] において意思表示の匿名性が大きなメリットを持つことが示されていることから、共有・可視化システムでは個人ごとの興味推定結果は表示せず、全員に対して共通のフィードバックを行うことを検討している。

4. 評価実験

本章では、会議円滑化のための意思疎通支援を行うため、興味推定システムの実装に向けた評価実験について述べる。

表 1: 特徴抽出に使用したフレームワーク

取得する特徴	使用するフレームワーク
1. 顔の角度	Mediapipe Facemesh
2. 姿勢	OpenPose
3. 表情（感情）	Amazon Rekognition
4. 発話内容	OpenAI Whisper
5. 発話頻度	OpenAI Whisper

4.1 では特徴量を抽出するシステムについて、4.2 では実験を行う上での条件について述べ、4.3 では実験結果について述べる。

4.1 実装

実装に用いたフレームワークを表 1 に示す。顔の角度については Facemesh [18] を用いることで会議中の顔の角度を抽出する。具体的には 432 個の顔特徴から線形ベクトルを作成し、顔の角度に換算する。表情については Amazon Rekognition [19] を用いて、特徴を感情に変換し推定する。API は感情を 8 種類に分類し、感情としては、CONFUSED（困惑）、CALM（平静）、SURPRISED（驚き）などが挙げられる。姿勢は OpenPose [20] を用いて推定する。具体的には、会議中のカメラに映る、胸部、首、両目、両耳、両肩、両肘の特徴を抽出する。身体動作の分析頻度として顔の角度、姿勢は 0.5 秒に 1 回、表情は 2.5 秒に 1 回としている。発話内容分析及び発話頻度分析に必要な会話の文字起こしは OpenAI Whisper [21] を用いる。

4.2 実験条件

被験者は 20 代の 4 名を 1 グループとして「若者の投票率を上げるためにはどのような政策が有効か」というテーマを設定し議論させた。議論は条件として以下の 4 つを与えて議論させた。議論時は、補助ツールとして WEB ホワイトボードアプリケーションの miro を使用させた (図 5)。

- (1) 若者の定義をする
- (2) 主要アプローチは 1 つ
- (3) 実現可能性を考える
- (4) 最終的にグループ全員が共通意見を持つことを目指す

議論時間に関して、予備実験時に発散・収束合わせて 15 分の議論時間を設けて議論させたが、実験設定の議論時間が短いため集中力が途切れず特徴が大きく変化しなかった。このことから、意見発散フェーズと収束フェーズに分けて 20 分、10 分の計 30 分の議論を行わせた。

身体動作特徴・発話特徴と興味の関連性の調査するため、各会議参加者に各フェーズの興味度合いの相対評価と 5 秒ごとの興味評価の 2 種類の評価を行ってもらった。各フェーズの興味度合いの相対評価では、議論フェーズを提示して興味度の高い順に並べ替えてもらった。5 秒ごとの興味評価では、会議の録画を見ながら、5 秒ごとに興味評

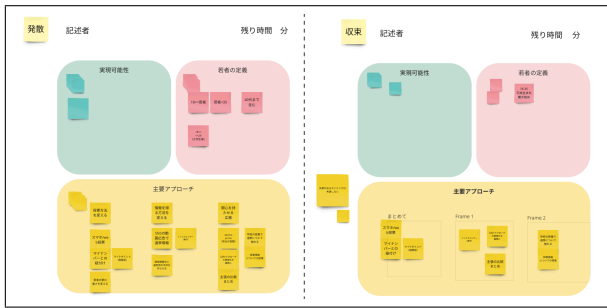


図 5: WEB ホワイトボード miro 使用時の様子

価を -4 から +4 までの 8 段階で入力してもらった。興味評価の値は興味度合いが高いほど高い数値とした。

4.3 実験結果

会議中の顔の角度、表情、姿勢と 5 秒ごとの興味評価を行わせた際の興味評価の相関係数を表 2 に示す。5 秒ごとの興味評価を行わせた際の各被験者ごとのデータ数として、議論中に取得した顔の角度特徴、姿勢特徴が 1 次元あたり 4,345 件、表情特徴が 1 次元あたり 867 件となり、興味評価データは 360 件となった。相関係数は、各被験者の特徴と興味評価データを全て集約したのち、算出した。

発散議論時、収束議論時の議論フェーズ推移と各被験者のフェーズに対する興味度合いを表 3、表 4 にそれぞれ示す。数字は被験者が他フェーズとの相対評価で回答した興味の強い順から数字を低く表し、括弧内には動画を見せながら回答させた 5 秒刻みの興味評価の平均値を示す。また、表下部に被験者が回答した興味上位、下位それぞれ 3 件の合計発話回数、および総発話回数を示す。

発散議論終了後、各議論フェーズに対しての興味度合いの相対評価を各被験者に回答させた。結果として被験者 F, G, H は「主要アプローチの発散」、「主要アプローチの深掘り」、「再度アプローチの発散」、被験者 I は「主要アプローチの発散」、「深掘り」に加えて、「投票における一票の格差について」の議論フェーズに強い興味を示した。また、相対評価における興味上位、下位 3 件の議論フェーズを抽出し、それぞれの発話回数を算出したところ、被験者 F が 44 回と 5 回、被験者 G が 28 回と 10 回、被験者 H が 24 回と 9 回、被験者 I が 13 回と 3 回という結果になった。収束議論終了後、各議論フェーズに対しての興味度合いの相対評価を各被験者に回答させた。結果として各被験者は「主要アプローチの選択」、「決定」、「決定したアプローチのマイナスポイント」、「インセンティブ報酬政策の効果」、「インフルエンサーの政策効果」、「授業などでの政策について」の議論フェーズに強い興味を示した。また、相対評価における興味上位、下位 3 件の議論フェーズを抽出し、それぞれの発話回数を算出したところ、被験者 F が 10 回と 4 回、被験者 G が 26 回と 28 回、被験者 H が 1 回と 5 回、被験者 I が 9 回と 2 回という結果になった。

表 2: 身体動作特徴と興味評価の相関係数

身体動作特徴	特徴名	相関係数
顔の角度	水平方向 (rad_x)	-0.510
	垂直方向 (rad_y)	-0.036
	前後方向 (rad_z)	-0.170
表情特徴	ANGRY (怒り)	0.170
	CALM (平静)	-0.440
	CONFUSED (困惑)	0.310
	DISGUSTED (嫌悪)	0.220
	FEAR (恐怖)	0.100
	HAPPY (喜び)	0.160
	SAD (悲しみ)	0.110
	SURPRISED (驚き)	0.150
姿勢特徴	胸部-x	-0.250
	胸部-y	-0.052
	右肩-x	-0.330
	右肩-y	-0.031
	右肘-x	-0.190
	右肘-y	-0.190
	左肩-x	-0.190
	左肩-y	-0.120
	左肘-x	0.110
	左肘-y	0.120
	右目-x	-0.510
	右目-y	-0.130
	左目-x	-0.460
	左目-y	-0.150
右耳-x	-0.400	
右耳-y	-0.100	
左耳-x	0.027	
左耳-y	0.096	

5. 考察

5.1 身体動作特徴

身体動作特徴に関しては、5 秒ごとの興味評価データを用いて、興味推定に有効であるかの考察を行った。

本稿で想定した、Facemesh を用いた顔の角度分析の中でも、顔の水平方向の動きが興味推定に有効であると考えられる。表 2 に示したように顔の水平方向の動きと興味評価との相関係数が -0.510 と比較的、関連性が高い数値となった。回帰直線を含めた顔の水平方向の特徴と興味評価の散布図を図 6(a) に示す。図では負の相関を示しており、顔の水平方向を示す rad_x が、左向きである 0.607 を示した際にその議論への興味評価が低い傾向にある。rad_x はカメラに対して正面を向いている時に 0.5 付近を数値として示すため、興味評価が低い際は画面の左端を見ているか、体勢を崩しながら議論をしていることがあると考察できる。実際、興味評価が低い際、被験者が右肘をつくように体勢を崩して顔を左向きにして画面を見るような場面が多く見られた。しかし、参加者が体勢を崩す際は片側だけに偏るとは考えにくく、正規分布のような形になることが考えら

表 3: 発散議論時の各参加者の議論フェーズに対する興味度合い順と発話回数, 重要単語初発話者

議論フェーズ	被験者 F	被験者 G	被験者 H	被験者 I	重要単語初発話者
1. 若者の定義の確認	4 (3.77)	6 (1.45)	7 (2.08)	7 (0.35)	被験者 H
2. 若者の定義の発散	6 (3.93)	5 (1.23)	8 (1.62)	6 (-0.31)	被験者 H
3. 若者の定義の発散確認	5 (3.57)	7 (1.79)	4 (1.64)	8 (-1.21)	被験者 H
4. 他の定義の確認	10 (2.00)	4 (-2.00)	10 (1.00)	9 (-1.25)	被験者 H
5. 主要アプローチの発散	1 (3.70)	3 (0.36)	3 (1.58)	2 (1.82)	被験者 H
6. 主要アプローチの深掘り	2 (3.84)	1 (1.39)	1 (1.94)	1 (3.28)	被験者 I
7. 主要アプローチの実現可能性	9 (2.83)	8 (0.50)	9 (-0.33)	10 (2.00)	被験者 G
8. 再度アプローチの発散	3 (3.79)	2 (1.11)	2 (1.47)	5 (3.32)	被験者 F
9. 若者の動向	8 (4.00)	9 (-1.00)	6 (1.00)	4 (3.67)	被験者 F
10. 投票における一票の格差	7 (2.82)	10 (1.00)	5 (1.36)	3 (3.55)	被験者 G
興味上位 3 件の総発話数 (回)	44	28	24	13	
興味下位 3 件の総発話数 (回)	5	10	9	3	
総発話回数	46	48	25	13	

表 4: 収束議論時の各参加者のフェーズに対する興味度合い順と発話回数, 重要単語初発話者

議論フェーズ	被験者 F	被験者 G	被験者 H	被験者 I	重要単語初発話者
1. 若者の定義の決定	4 (3.25)	6 (0.81)	5 (1.00)	7 (1.00)	被験者 G
2. 主要アプローチの選択	6 (2.90)	7 (2.05)	1 (1.38)	6 (2.19)	被験者 G
3. インセンティブ報酬政策の効果	3 (3.38)	4 (2.54)	6 (1.62)	3 (2.54)	被験者 G
4. インフルエンサーの政策効果について	5 (1.57)	2 (1.64)	3 (-0.43)	2 (3.64)	被験者 I
5. 授業などでの政策	7 (3.47)	3 (1.77)	4 (-1.00)	1 (2.80)	被験者 G
6. 主要アプローチの決定	1 (4.00)	1 (2.89)	2 (1.22)	5 (3.67)	被験者 F
7. 決定したアプローチのマイナスポイント	2 (4.00)	5 (2.00)	7 (1.00)	4 (2.80)	被験者 F
興味上位 3 件の総発話数 (回)	10	26	1	9	
興味下位 3 件の総発話数 (回)	4	28	5	2	
総発話回数	19	61	8	11	

れる。本稿の評価実験ではそのような結果は得られなかったが、画面を正面から、もしくは左右寄りで見ているのか、特徴分析が可能であるため顔の水平方向の動きは興味推定に有効であると考えられる。

本稿で想定した、Amazon Rekognition を用いた表情分析の中でも、平静と困惑の推定が興味の推定に有効である。表 2 に示したように平静の推定と興味評価の相関係数が -0.440 、困惑の推定と興味評価の相関係数が 0.310 、と比較的関連性が高い数値となった。平静と推定される精度が高まるにつれて、議論への興味が低くなりやすくなる一方、何らかの感情が発生している際は議論への興味が高くなる可能性が示されている。実際、図 7 に示すように平静は他の表情特徴全てに対して負の相関関係が表れている。また、困惑は比較的他の表情特徴より高い正の相関が表れている (図 6(b))。

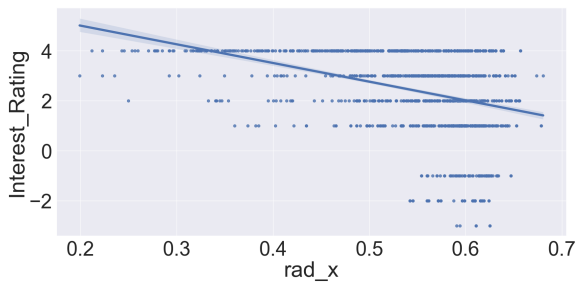
本稿で想定した、OpenPose を用いた姿勢分析の中でも、肩と目の座標推定が興味の推定に有効であると考えられる。表 2 に示したように右肩の x 座標 (Bodyparts2- x) と興味評価の相関係数が -0.330 、右目の x 座標 (Bodyparts14- x) と興味評価の相関係数が -0.510 、左目の x 座標 (Bodyparts15- x) と興味評価の相関係数が -0.460 、右耳の x 座標 (Bodyparts16-

x) と興味評価の相関係数が -0.400 と比較的、関連性が高い数値となった。これらは Facemesh を用いた顔の角度分析と同じく、右に体勢を崩しながら議論をしている際は議論への興味評価が低いことが考察できる。実際、図 8 に示すように顔の水平方向と各姿勢特徴は互いに正の相関があることから考察することができる。この結果は体勢を崩して議論を行う被験者は興味を抱いていない仮説を後押しするものであり、興味推定に用いることが可能であると考えられる。

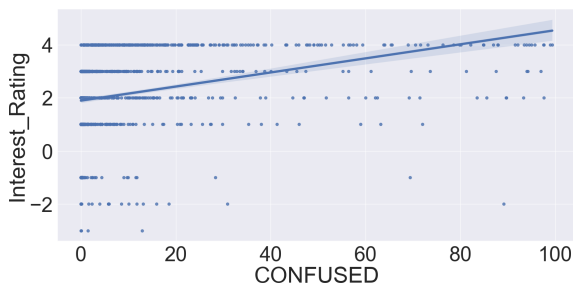
5.2 発話特徴

発話特徴に関しては、各議論フェーズの興味度合いの相対評価を用いて、興味推定に有効であるかの考察を行った。

本稿で想定した、OpenAI Whisper を用いた発話頻度分析は、議論フェーズに対する興味推定に有効であると考えられる。発散、収束議論において相対的に興味の強い議論フェーズ上位 3 つ、下位 3 つの発話回数をグループ全体と比較すると 155 対 66 となった。また、「興味度合いごとに発話量の差はない」という帰無仮説を立て、Mann-Whitney U test にて優位差検証を行った。各被験者から相対的に興味のあるフェーズ上位 3 つ、下位 3 つの発話回数を抽出し



(a) 顔の水平方向 (rad_x)



(b) 表情特徴 (困惑)

図 6: 興味評価の散布図と回帰線

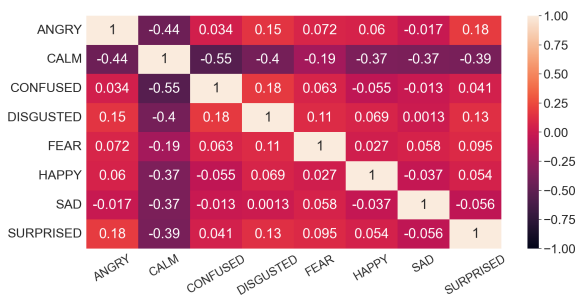


図 7: 表情特徴の相関係数

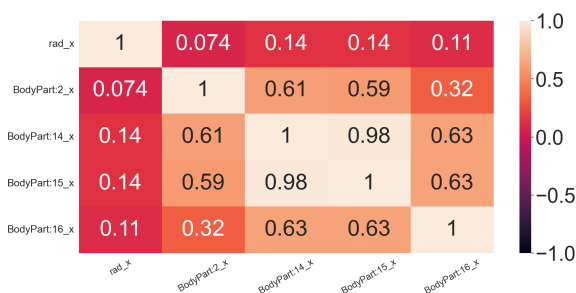


図 8: 顔の水平方向と姿勢特徴の相関係数

たところ、2 群の観測データ数はそれぞれ計 21 件となったため、U 値を標準化して z 検定を行った。結果として、 $zU = 2.091\dots$, p 値は $0.05 \geq p(0.036\dots)$ となり、帰無仮説は棄却され、興味度合いによる発話量変化の優位性が検証された。よって、被験者が興味のある議論フェーズに対しては相対的に発話量が多くなる。

本稿で想定した、OpenAI Whisper を用いた発話内容分析は、議論フェーズに対する興味推定が困難であると考えられる。

重要単語の初発話を行った被験者がそのフェーズに強い興味 (相対評価上位 3 つ以内) を示した割合は 7/17 パターンとなった。また、「重要単語の初発話を行っても、興味度合いに変化はない」という帰無仮説を立て、Mann-Whitney U test にて優位差検証を行った。重要単語を初発話者したフェーズに対する相対的な興味評価とそれ以外を抽出したところ 2 群の観測データ数はそれぞれ 17 件、51 件となったため、U 値を標準化して z 検定を行った。結果として、 $zU = 0.334\dots$, p 値は $0.05 \geq p(0.737\dots)$ となり、帰無仮説は棄却されず、重要単語の初発話と興味度合いの関係性が検証できなかった。

この要因としてファシリテータの存在が考えられる。発散議論では被験者 H、収束議論では被験者 G がそれぞれファシリテータを担当していたため重要単語初発話の半数を占める結果となった。結果、ファシリテータが初発話を行ったフェーズへの相対興味評価は被験者 H が平均 6.4、被験者 G が平均 5.0 と強い興味を持っていたとは言えない結果となっている。一方、ファシリテータの評価を除いた場合の初発話を行ったフェーズへの相対興味評価は被験者 I が平均 1.5、被験者 F が 3.5 と比較的高い評価となっている。このことから、ファシリテータなど、会議内の役職を考慮して再評価をする必要がある。

5.3 議論

意思疎通支援システムを実現するため、今後検討すべき点が 3 つある。

1 つ目に、興味推定を行うための学習済みモデルの作成である。本稿では、興味推定システムを実現するにあたっての有効な特徴調査を行った。今後は、議論中にリアルタイムで興味推定を行えるようにするため、学習済みモデルを作成する必要がある。特に、議論の円滑化にはどの程度の興味推定精度が必要かという点に関して、深く検討する必要がある。

2 つ目に、興味・理解度・納得度をフィードバックするための最適な UI の検討である。参加者映像に画像処理を用いた研究 [22-24] や発話支援に着目した研究 [25-27] が報告されている。これらの研究は主体的なシステム操作が不必要なように、参加者映像に画像処理を用いたり、発話タイミングを自動支援するシステムを提案しているが、利便性に関して不満を感じる会議参加者がいたことが報告されている。中でも発話支援などは、積極的な発話者は自身で発話のタイミングを掴むことに抵抗がないために不満が出ていたことが報告されているため、システムが適切なタイミングで適当な参加者を支援することは重要な要素であることが考えられる。

3 つ目に、興味・理解度・納得度をフィードバックするための最適な会議環境の検討である。この点に関して検討の余地があるため、予備実験時に「議論中、興味推定の可

表 5: 「興味推定の可視化が必要と感じるか」アンケート結果

回答内容	回答人数 (n=15)	主な回答理由
1. 必要	8 名	不要議論の区別が容易
2. 場合により必要	6 名	意思表示が苦手なため
3. 場合により不要	1 名	知人同士では不要
4. 不要	0 名	-

視化が必要と感じるか」と「どんな場面でそう思うか」の 2 つを計 15 名にアンケートにより調査した (表 5)。結果として必要と答えた人数は 8 名、場合により必要と答えた人数は 7 名、場合により不要と答えた人数は 1 名、不要と答えた人はいなかった。必要派の理由として「多数が興味のない議論を不要な議論として時間が短縮できる」という回答が多くみられた一方で、場合により不要派の理由として「知人同士の議論では有用性を感じない」という回答がみられた。この結果から、「初めて話す他人同士が、短い時間で、意見を収束する必要がある」などの、明確な会議目的を持った議論環境には本システムが有効であると考察できる。今後は、本システムを使用するにあたっての適切な会議環境を検討し、評価する必要がある。

6. おわりに

本稿では遠隔会議円滑化のための意思疎通支援を行うため、議論内容への興味を推定・可視化するシステムを提案し、興味推定に有効と思われる特徴を考察して評価を行った。議論中の会議参加者の身体動作特徴と興味評価の相関係数を求めた結果、顔の水平方向の動きと興味評価の相関係数が -0.510 、平静の感情推定が -0.440 、困惑の感情推定が 0.310 、右目の x 座標が -0.510 、左目の x 座標が -0.460 となった。また、発話頻度に関して、興味のある議論フェーズ上位 3 つ、下位 3 つの発話回数に対して、「興味度合いごとに発話量の差はない」という帰無仮説を立てて Mann-Whitney U test にて優位差検証を行った。結果として p 値は $0.05 \geq p(0.036\dots)$ となり、帰無仮説は棄却され、興味度合いによる発話量変化の優位性が検証された。一方、本稿の調査では発話内容について、議論内容への興味が特徴に現れにくい結果となった。

今後の課題として、3 つのことが上げられる。興味推定を行うための学習モデルの作成と興味・理解度・納得度をフィードバックするための UI、および最適な会議環境の検討である。本稿では会議参加者の主体的な操作が不必要なシステムを提案したが、関連研究よりシステムが適切なタイミングで適当な参加者を支援することは重要な要素であることが考えられる。同じく、議論の円滑化にはどのぐらいの興味推定精度が必要かという点に関して、様々な会議目的を持った議論環境で評価する必要がある。

参考文献

- [1] 総務省 : 新型コロナウイルス感染症が社会にもたらす影響:5G が促すデジタル変革と新たな日常の構築, 令和 2 年版 情報通信白書 第 1 部, pp. 138-166 (2020).
- [2] 総務省 : 令和 3 年通信利用動向調査の結果, 情報通信統計データベース, pp. 1-27 (2021).
- [3] 宮内佑実, 遠藤正之: オンライン会議とオフライン会議の意思疎通の比較, 経営情報学会全国研究発表大会要旨集, pp. 144-147 (2020).
- [4] Yang, L. and et al.: The Effects of Remote Work on Collaboration among Information Workers, *Nature Human Behaviour* 6, No. 1, pp. 43-54 (2021).
- [5] Cao, H., Lee, C., Iqbal, S., Czerwinski, M., Wong, P., Rintel, S., Hecht, B., Teevan, J. and Yang, L.: Large Scale Analysis of Multitasking Behavior During Remote-Meetings, *Proc. the CHI Conference on Human Factors in Computing Systems*, pp. 1-13 (2021).
- [6] 小松真子, 澤村三奈, 敷田幹文: 組織間テレビ会議の円滑な進行を支援する意思表示方式の提案と評価, 情報処理学会研究報告, Vol. 2020-GN-110, No. 2, pp. 1-8 (2020).
- [7] 阿部花南, 築館多藍, 桑宮 陽, 小林 稔: 会議円滑化支援を目的とした気持ち可視化ボタンの提案, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO), pp. 774-783 (2021).
- [8] 杉本知佳, 又吉康綱, 古市冨江, 中村聡史: オンラインミーティングでの発言障壁を低減するカードによる匿名での意思表示支援手法, 情報処理学会研究報告, Vol. 2019-GN-106, No. 1, pp. 1-8 (2021).
- [9] 大野直紀, 徳久弘樹, 中村聡史: 自身のみ聴取可能な音楽を用いたコミュニケーション円滑化手法の提案, 情報処理学会研究報告, Vol. 2021-GN-113, No. 22, pp. 1-7 (2019).
- [10] 田之頭吾音, 川口一画: 視線情報および姿勢を用いたウェアネス情報提示による遠隔対話の生起支援, 情報処理学会研究報告, Vol. 2021-HCI-195, No. 28, pp. 1-8 (2021).
- [11] 井出将弘, 大島昇時, 森 真吾, 市野順子, 田野俊一: VR 会議におけるアバターを介したジェスチャー表現の影響, 情報処理学会研究報告, Vol. 2020-GN-110, No. 1, pp. 1-8 (2020).
- [12] 富増紗也華, 荒木雅弘: 雑談対話におけるマルチモーダル情報からの興味の有無の判定, 人工知能学会全国大会論文集 JSAI, pp. 1-4 (2016).
- [13] 小関 健, 小坂哲夫: ユーザの身体動作から興味を推定するマルチモーダル対話システム, 情報処理学会 インタラクション 2018 IPSJ Interaction, pp. 646-650 (2018).
- [14] 倉野大二郎, 松村耕平, 角 康之: マルチモーダルデータを用いた映像閲覧者の興味推定, 情報処理学会 インタラクション 2013 IPSJ Interaction, pp. 435-439 (2013).
- [15] 南竹俊介: 歩行者の顔向き情報と移動軌跡を用いたデジタルサイネージの効果測定と情報表示, 筑波大学大学院博士課程 システム情報工学修士論文, pp. 1-53 (2010).
- [16] 西條涼平, 永陽子, 山口大地, リドウィナアンダリニ, 松尾翔平, 戸嶋巖樹, 倉橋孝雄, 小澤史朗: 感性コミュニケーション技術の実現に向けた熟練度と対話満足度に関する取り組み, NTT 技術ジャーナル / 日本電信電話株式会社編, No. 34, pp. 32-35 (2022).
- [17] 鳥羽望海, 藤本まこと, 諏訪博彦, 酒井元気, 酒造正樹, 安本慶一: マルチモーダルデータを用いたオンラインミーティング参加者の感情推定, 研究報告知能システム (ICS) 2022-ICS-206, No. 2, pp. 1-7 (2022).
- [18] Kartyannik, Y., Ablavatski, A., Grishchenko, I. and Grundmann, M.: Real-time Facial Surface Geometry from Monocular Video on Mobile GPUs, *The Workshop on Computer Vision for Augmented and Virtual Real-*

- ity in conjunction with *Conference on Computer Vision and Pattern Recognition*, pp. 953–959 (2019).
- [19] Amazon Web Services, I.: Amazon Rekognition, Amazon com. (online), available from (<https://aws.amazon.com/jp/rekognition/>) (accessed 2023/04/11).
- [20] Cao, Z., Hidalgo, G., Simon, T., Wei, S.-E. and Sheikh, Y.: OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields, *arXiv* <https://doi.org/10.48550/arXiv.1812.08008>, pp. 1–14 (2019).
- [21] Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C. and Sutskever, I.: Robust Speech Recognition via Large-Scale Weak Supervision, *arXiv* <https://arxiv.org/abs/2212.04356>, pp. 1–28 (2022).
- [22] 安永貴之, 小口寿明, 見矢野友香, 土井沙耶香, 三谷洋之, 赤垣慎吾, 飛田博章: 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, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO), pp. 760–765 (2021).
- [23] Ziting, G., 金井秀明: 遠隔会議における相手の顔を変換することによる緊張緩和手法に関する研究, 情報処理学会研究報告, Vol. 2021-GN-113, No. 3, pp. 1–8 (2021).
- [24] 鈴木啓太, 横山正典, 吉田成朗, 望月崇由, 布引純史, 鳴海拓志, 谷川智洋, 廣瀬通孝: 同調的な表情変形技術を用いた遠隔コミュニケーションの拡張, 情報処理学会論文誌, Vol. 59, No. 1, pp. 52–60 (2017).
- [25] Samiha, S.: Automated Collaboration Coach for Videoconferencing based Group Discussions, *Proc. the ACM Int. Joint Conference and Int. Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*, pp. 510–515 (2018).
- [26] 鍋谷航平, 村岡泰成, 石川誠彬, 江木啓訓: 消極的発話者の発言率向上を目的とした音声による個別指示議論支援システムの開発, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO), pp. 776–773 (2021).
- [27] 山田楓也, 白石 陽, 石田繁巳: Web 会議における予備動作を用いた発話欲求推定手法の提案, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム (DICOMO), pp. 395–403 (2021).