

身体・頭部・視線の3方向のずれに着目した ライフログの振り返り

長嶺 和弥^{1,a)} 久米田 羽月^{1,b)} 角 康之^{1,c)}

概要：本研究では、ライフログ収集ツールとして身体・頭部・視線の3方向から一人称ライフログが閲覧できるビューアを提案する。一人称視点映像の中に頭部方向と視線方向の情報を書き込むことによって、身体方向、頭部方向、視線方向の関係性がわかりやすくなり、一人称ライフログ映像と同時に着用者の身体状況を記録できると考える。この記録から会話の発話内容だけでなく、その会話を取り込む周りの環境と身体状況に着目して、コミュニケーションの状況や会話の状況も分析できることが期待できる。本研究ではウェアラブルデバイスのみを用いる。着用者は広視野の胸部カメラを身に付け、さらに同時にメガネ型の視線計測装置を着用する。この2つのデバイスから得られた映像を使用して処理を行う。

1. はじめに

私たちの普段の生活の中では、立っている状態でのコミュニケーションが頻繁に行われている。例えば、友人と買い物にでかけた際に行う雑談や、ポスターを見かけた際に行う会話などが挙げられる。人は複数人で会話をする際、無意識で自分たちの身体配置を調整している。特に会話への参加や離脱の際や、会話内容に関する参照物が存在する場合、身体の位置やねじれは複雑に変化する。

また頭部方向や視線方向でもコミュニケーション中に様々な非言語のやり取りが行われる。例えば、会話中に一人がある対象に顔を向けると他の人もその方向を見るというような場合や、相手が自分の話をどれほど真剣に聞いているかを窺う時などに多い。また会話中は話し手は聞き手を見る際、眼球運動を主に用いるのに対し、聞き手は頭部方向で話し手の方を見ることが多い。このように人は体の向きや身体配置、頭部方向や視線方向などを重要なコミュニケーション手段として用いている。我々はこのような行動を直感的には理解できるが確認することは難しい。

そこで本研究では一人称ライフログ映像を用いて着用者の注視行動の変化を伴う身体配置やねじれに注目し、多人数インタラクション分析に関する新しい視座を提供する。

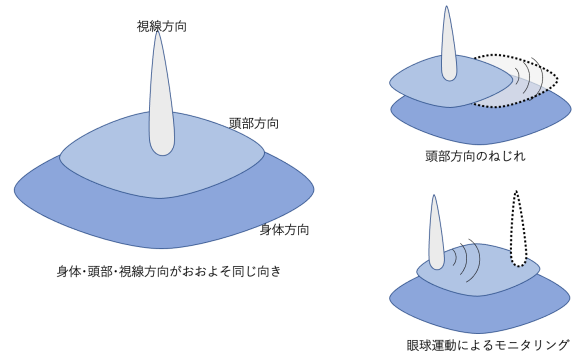


図 1 注視行動の階層モデル

2. 関連研究

2.1 一人称ライフログ映像からの社会的インタラクション理解

一人称ライフログ映像を用いた先行研究として奥野らの研究 [1] を挙げる。これは胸元で撮影した一人称ライフログ映像から映り込む顔の数を数え、さらに顔の大きさと顔が映り続ける時間を考慮することで社会活動量を計測を試みた。

また Aghaei ら [2] は低フレームレートカメラを装着し、長時間の撮影の中での交流の瞬間を自動で検出した。認知障害者の記憶訓練などへの応用できる価値があると説明している。しかし頭部方向や視線方向など着用者自身が注意を向けている方向は考慮されていない。

本研究では、頭部方向と視線方向の情報を一人称ライフログ映像に書き込むことでライフログの収集と同時に身体

¹ 公立はこだて未来大学

a) k-nagamine@sumilab.org

b) u-kumeta@sumilab.org

c) sumi@acm.org

状況、注視方向も記録できると考えている。

2.2 身体配置に注目した会話分析

身体配置によるコミュニケーション分析を試みた先行研究として Kendon の研究 [3] を挙げる。Kendon は人のコミュニケーションにおいて参加者が3人以上の場合、それらの身体配置は円形になると説明しこの身体の空間配置行動を F 陣形と呼んだ。また参加者によって作り出される O 空間、P 空間、R 空間という3つの空間を説明した。O 空間とは複数人で会話をしている時の参加者が身体を向けあった内側の空間のことであり、参加者が身体配置を調整することで相互に維持される。また P 空間とは参加者の身体を結んだ O 空間の外縁の空間のことであり、さらに R 空間とは P 空間よりもさらに外側の空間のことであり、R 空間では新たな会話参加者がとどまり、既存の F 陣形への参加の許可を待つ、あるいは通り過ぎるための空間である。このように複数の身体が関わり合う活動を観察するには F 陣形のような身体の空間配置に着目するべきである [4]。

Kendon の研究 [3] を踏まえて田村ら [5] は身体配置のみで会話中の役割を判断しお喋りをするロボットを開発した。ここでは話し手、聞き手、傍参加者、傍観者の4つに役割を分け、ユーザがどの立場なのか判別する。人が傍参加者や傍観者だとロボット同士で勝手に会話をし、人が会話に加わると人を交えて会話できるロボットである。

また Ricci ら [6] は三人称視点映像から身体方向と頭部方向とその会話グループの F 陣形の形の推定を試みた。

さらに Hedayati ら [7] は4つの定点カメラに加え、マイク・加速度計・Bluetooth・赤外線センサーの機能がついたデバイスを被験者に身につけてもらい、さらに人間同士、人間とロボットの対話を観察し、機械学習を用いて F 陣形の検出をした。

また Gaschler ら [8] はバーテンダーとの対話中の体と頭の非言語行動を記録して分析した。その結果から機会学習を用いて人間の顧客の社会的行動を認識するバーテンダーロボットを開発した。

しかしこれらの先行研究では3人称視点映像から解析をしているため場所の制約が大きい。本研究ではウェアラブルデバイスを用いて1人称視点映像を撮影するため、場所の制約が少なく、より日常に近いシーンでの観察ができることが期待できる。

2.3 社会的インタラクションにおける頭部運動と視線行動

視線行動に関する先行研究として伊藤らの研究 [9] では、実験室内においてポスターなどの展示物を見ながら会話をする際、話し手と聞き手で相手への視線の向け方が異なると指摘している。話し手は頭部方向を相手に向けて話す時間が多いのに対し、聞き手は多くの時間を展示物に向ける

事が多いことが分かった。

また大森ら [10] は4人が机に2人ずつ向かい合わせになるように座り、話している時の視線行動を観察した。その結果、話し手は斜め向かいの人を見る回数が多く、聞き手は話し手の方を見る回数が多いという結果になった。また男性より女性のほうが話し手を見ている時間があった。

本研究では、一人称ライフログ映像と身体・頭部・視線の方向を閲覧できる。そして先行研究の内容を踏まえて主に多人数でのコミュニケーションの状況などを分析することに適している一人称ライフログのビューアを提案する。

3. 提案手法

3.1 データ収集

本研究では身体の向きと頭部の向き、視線の向きのデータを取得する。実験参加者は広視野の胸部カメラ GoPro7 とメガネ型の視線計測装置である Pupil Invisible の2つの機材を装着し、同時に2つの映像を撮影する(図2)。この広視野の胸部カメラには身体方向のデータを取得する目的もあるが、画角が280°と広く実世界の状況を映すために効果的だと考えている。メガネ型の視線計測装置では、実世界を映すための外向きのカメラと、眼球の動きをセンシングするための内側のカメラがある。ここで撮れた映像から視線方向のデータ、そして映像の画角から頭部方向のデータが取得できると考えている。

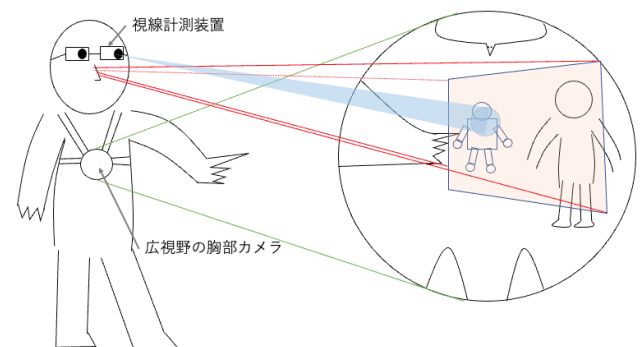


図2 身体・頭部・視線の3方向の記録と振り返り

3.2 データ加工

本研究では、広視野とメガネ型の視線計測装置で同時に撮られた2つの映像を扱う。図3の左側の映像は視線計測装置で撮られた映像で、右側の映像は広視野の胸部カメラで撮った映像である。広視野の胸部カメラで撮った映像は画角が広いのと引き換えに特に外側だと歪みが大きいという点が特徴である。図3は視線計測装置で撮られた映像の画角を四角で、視線方向を赤丸で示している。ここでは視線計測装置で捉えた映像の画角を頭部方向と考え、広視野の胸

部カメラで撮った映像に反映している。これは OpenCV の AKAZE を用いて frame-by-frame で 2 つの映像の特徴点をマッチングし、座標ペアの配列からホモグラフィ行列を推定する。この求めたホモグラフィ行列を用いて、視線計測装置で撮影された映像の四隅と視線の座標を広視野の胸部カメラで撮影した映像での座標に変換し、それぞれ四角と丸印を描画して表示している。

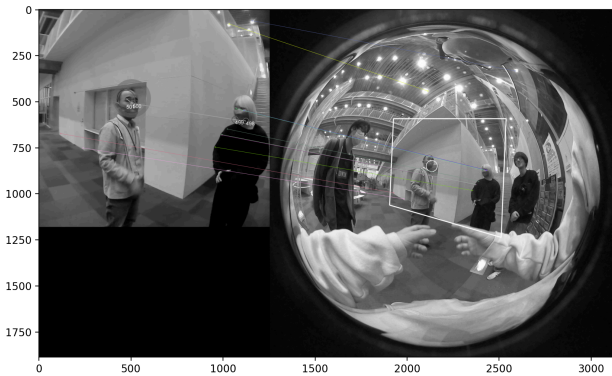


図 3 ビューアの画面例

4. 動作例

この章では本研究で提案したビューアの動作例を紹介する。ここでは、なにか対象のものに注目している場面と着用者が聞き手に回っている場面と話し手に回っている場面の 3 つを確認する。下記の例を見ると複数人での会話中は注意を様々な方向に向ける必要があり、我々はそれを頭部運動や眼球運動を用いることで対応していることが分かる。

4.1 対象物に注目している場面

図 4 は着用者がなにか対象物に注目している場面でのビューアの様子である。これは会話中に正面の奥の部屋についての話題になり、着用者がその部屋の様子を見ている場面である。このように人が対象物に注目する場合は、身体・頭部・視線方向はほとんど同じ方向を向くことが多い。本研究で提案したビューアで見ると、広視野の胸部カメラで撮影した映像のおおよそ中心辺りに頭部方向を表す四角が描かれており、さらにその中心辺りに視線方向を表す丸印が描かれる。この様子は図 4 のように本研究で提案したビューアで確認できた。

4.2 着用者が聞き手の場面

図 5 は着用者が聞き手に回っている場面でのビューアの様子である。これは複数人で会話をしている時に隣の人が話を始めたので、着用者は聞き手に回って発話者に視線を向けている場面である。人はこのように聞き手に回っている時は、頭部運動を主に用いることが多い。具体的には、身体方向に対して頭部方向を大きく動かし視線方向は頭部

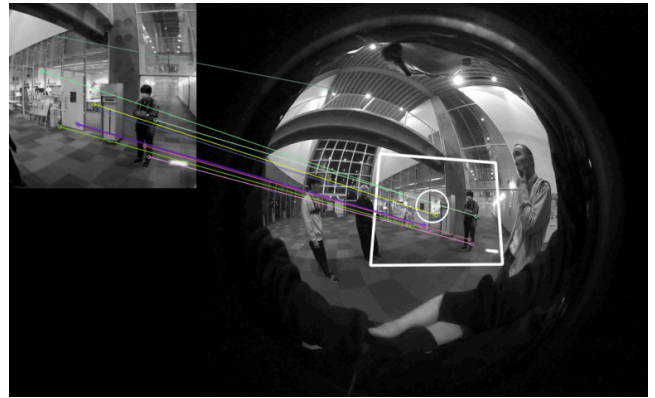


図 4 着用者が対象物に注目しているシーン：身体・頭部・視線のねじれが少ない

方向とおおよそ同じ方向を見ることが多い。本研究で提案したビューアでは、広視野の胸部カメラで撮影した映像の端のほうに頭部方向を示す四角とその四角の中心辺りに視線方向の丸印が描かれる。その様子は図 5 で確認できた。

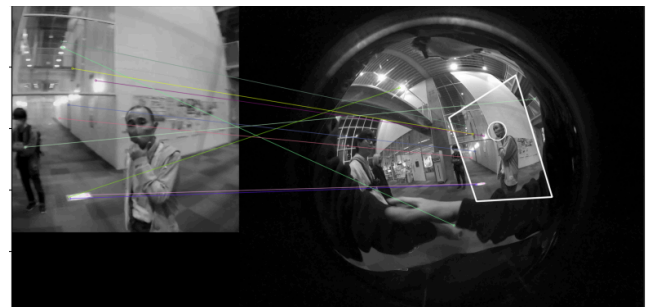


図 5 着用者が聞き手の時のシーン：頭部方向で話者に注意を向けている

4.3 着用者が話し手の場面

図 6 は着用者が会話に参加し、話し手に回っている場面でのビューアの様子である。人が話し手に回る場面では眼球運動を主に用いることが多い。具体的には身体方向と頭部方向はおおよそ一致していて、眼球運動を用いて周りを見渡す。図 6 の場合でも頭部方向と身体方向はおおよそ同じ方向を向いていて、視線方向が頭部方向を表す四角の端の方まで移動していることが分かる。

このように、複数人でのやりとりの中などでは役割によって身体・頭部・視線方向の動かし方が異なる。このことが本研究で提案した手法でも確認することができた。

5. 今後の展望

本研究では、ライフログの収集と同時に着用者自身の身体方向・頭部方向・視線方向などの身体状況も閲覧できるビューアを提案した。今後の展望としては、様々なデータを本研究で提案したビューアで観察し、先行研究では確かめられていなかった実験室外という日常でのシーンのコ

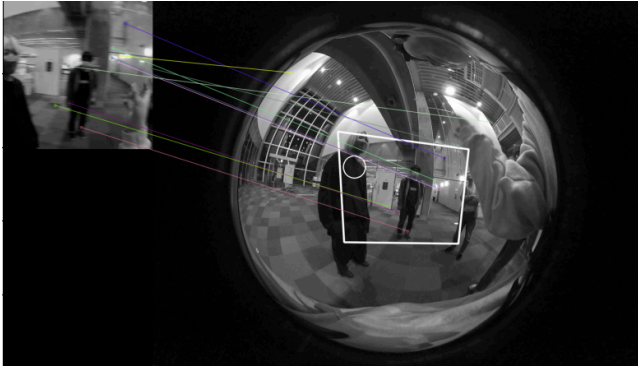


図 6 着用者が話し手のシーン：眼球運動で聞き手の反応をモニタリングしている

コミュニケーションの際の非言語行動を確かめたいと考えている。さらには身体方向・頭部方向・視線方向を観察することで発話情報を用いなくとも話者を推定できると考えている。また会話分析以外にも本手法を用いたビューアの可能性を探っていく。

謝辞

本研究は科研費（22H03634）の研究助成を受けた。

参考文献

- [1] 奥野茜, 角康之. 一人称ライフログ映像からの顔検出に基づいた社会活動計測. 情報処理学会論文誌, Vol. 62, No. 2, pp. 607–616, 2021.
- [2] Maedeh Aghaei, Mariella Dimiccoli, and Petia Radeva. With whom do I interact? Detecting social interactions in egocentric photo-streams. In *2016 23rd International Conference on Pattern Recognition (ICPR)*, pp. 2959–2964. IEEE, 2016.
- [3] Adam Kendon. *Conducting interaction: Patterns of behavior in focused encounters*, Vol. 7. CUP Archive, 1990.
- [4] 坊農真弓. 会話構造理解のための分析単位: F 陣形 (連載チュートリアル 多人数インタラクションの分析手法 [第 6 回]). 人工知能, Vol. 23, No. 4, pp. 545–551, 2008.
- [5] 田村真太郎, 伏木ももこ, 島崎景子, 因幡千尋, 那和一成, 岡田美智男. ”聞こえてくる” 会話を指向する多人数インタラクションエージェント. エンタテインメントコンピューティングシンポジウム 2017 論文集, Vol. 2017, pp. 422–424, 2017.
- [6] Elisa Ricci, Jagannadan Varadarajan, Ramanathan Subramanian, Samuel Rota Buló, Narendra Ahuja, and Oswald Lanz. Uncovering interactions and interactors: Joint estimation of head, body orientation and F-formations from surveillance videos. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4660–4668, 2015.
- [7] Hooman Hedayati, Annika Muehlbradt, Daniel J Szafir, and Sean Andrist. Reform: Recognizing F-formations for social robots. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11181–11188. IEEE, 2020.
- [8] Andre Gaschler, Sören Jentzsch, Manuel Giuliani, Kerstin Huth, Jan de Ruiter, and Alois Knoll. Social behavior recognition using body posture and head pose for human-robot interaction. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2128–2133. IEEE, 2012.
- [9] 伊藤禎宣, 岩澤昭一郎, 土川仁, 角康之, 間瀬健二, 片桐恭弘, 小暮潔, 萩田紀博. 装着型体験記録装置による対話インタラクションの判別機能実装と評価. ヒューマンインタフェース学会誌, Vol. 7, No. 1, pp. 167–178, 2005.
- [10] 大森慈子, 立平起子. 4 者間の自然会話場面における話し手と聞き手の視線行動-座席位置の違いによる検討. 日本心理学会大会発表論文集 日本心理学会第 72 回大会, pp. 1E157–1E157. 公益社団法人 日本心理学会, 2008.