

一人称映像からの顔識別に基づいた 社会ネットワークの可視化

東 爵亜久^{1,a)} 角 康之^{1,b)}

概要：一人称映像記録のデータを用いた社会ネットワークの可視化の試みについて報告する。我々は、アートユニット明和電機によるコンサートの6日間にわたる準備期間に参加し、プロジェクトに参加した17名の共同作業を多角的に映像記録する機会に恵まれた。参加者各々にカメラ装着をお願いし、プロジェクト参加者の6日間にわたる一人称視点映像を記録することができた。画像処理により参加者各々が胸につけたカメラに映りこんだ人物との対面回数を数え上げ、その蓄積データに基づいて社会ネットワークの可視化を試みた。本稿では、社会ネットワークの可視化によってプロジェクト期間中の人間関係の変化をどの程度読み取れるか議論する。

1. はじめに

本稿では、一人称映像記録のデータを用いた社会ネットワークの可視化の試みについて報告する。具体的には、参加者各々が胸につけたカメラで記録される映像から、画像処理技術によって対面の人物を逐一特定し、その判別結果の蓄積データを利用して参加者間の社会的インタラクションの頻度を定量化した。任意時間における複数参加者間の「社会的インタラクションの頻度」を可視化したものを、本稿では「社会ネットワーク」と呼ぶ。

図1はその一例である。社会ネットワークは参加者に対

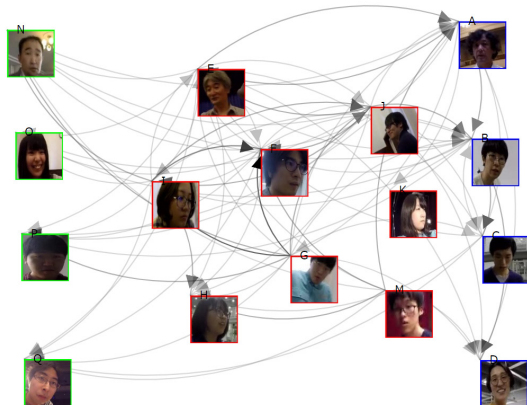


図1 可視化された社会ネットワークの例

れる。リンクは、始点がカメラ装着者を示し、終点はそのカメラで撮影された映像内に発見された参加者を示す。リンクの濃さは、その時間帯の発見回数の大きさを示す。

社会ネットワークは、計測する時間幅と時間帯を変えることで、参加者間の離合集散の変化を見ることに役立つと期待される。参加者間の社会的インタラクションの頻度とその変化は、作業の種類、各人の役割、相互の信頼や感情など、様々なことを反映したものであると考えられる。グループ活動の質は、本来は、それらの様々な要因を区別しながら議論すべきであると考えられる。しかし本研究では、それらの集約が「人物同士の対面」という現象に現れると仮定し、その量の変化を可視化するだけでも十分に社会的インタラクションの質的議論を行うことに役立つのではないかと、という仮説を立てる。つまり、本研究が解明すべき問いは以下の通りである。

- 各人の一人称カメラに映り込んだ人物を数えるという単純な方法によって可視化された社会ネットワークが、グループ活動の参加者間の大まかな人間関係や活動種類の変化をとらえることに役立つのか？
- 可視化された社会ネットワークを見ることは、当事者自身が活動を振り返り、その後の改善につながる気付きを得ることに役立つのか？

我々は、2018年5月に函館市で実施された「明和電機はこだてプロジェクト」に参加し、多角的に映像記録を行うという機会に恵まれた。このプロジェクトは、アートユニット明和電機のメンバー(4名)と、公立はこだて未来大学の学生(教員1名を加えた9名)が協力し、2018年5

応した複数ノードと、それらをつなぐ有向リンクで構成さ

¹ 公立はこだて未来大学

^{a)} k-azuma@sumilab.org

^{b)} sumi@acm.org

月 19 日に開催された「ナンセンス楽器コンサート」*1 を作り上げるというプロジェクトであった。我々はその準備期間と当日を加えた 6 日間に密着取材を行い、参加者各々にカメラ装着をお願いし、カメラ 9 台による合計 38 時間の一人称視点映像のデータを取得した。その 6 日間では、機材の移動や組み立て、楽器演奏や踊りの練習、ステージ演出、ミーティング、リハーサルといった様々な種類の作業が行われた。その過程を参与観察した我々は以下のような印象を持った。

- 期間前半では既存組織内のインタラクションが多かったが、期間後半になるにつれて、組織をまたがるインタラクションが増えていた。
- 組織やステージ上の役割が、自由作業におけるサブグループ形成にも影響していた。
- それらのサブグループ間を取り持つハブになるメンバーがいた。
- ステージの主演となる明和電機の土佐社長は、少人数から形成される共同作業は少なく、一人で黙々と作業するか、あるいは、チーム全体に対して対面・指示するような場面が多かった。

本稿の目標は、こういった印象を社会ネットワークの可視化でどの程度再現できるのかを議論することと、これらの印象とは異なる気付きを得られるかどうかを検討することである。

2. 関連研究

本研究は、胸に装着したカメラで取得された一人称映像の単純な画像処理から、人間関係を定量化する試みである。ウェアラブルセンサによって社会的インタラクションを計測しようとする試みは数多くなされてきた。

例えば Pentland らのグループは、赤外線センサやマイクから任意の参加者間の近接性を特定し、数日単位でグループメンバー間の社会ネットワークを取得する技術を開発してきた [1], [2]。矢野らの「ビジネス顕微鏡」[3], [4], [5] は赤外線センサで大まかな近接性を特定した上で、参加者間の身体動作の同期・非同期を加速度センサで計測し、それらのデータから参加者間のグループ活動の種類と質の特定を試みた。そして、それらの長期的運用から、組織内活動の質評価や問題点の特定・改善への適用を行った。我々の試みは、身体装着型のカメラによって得られる一人称映像を用いて社会的インタラクションを計測し、映像中に映り込む他者の顔画像の数え上げといった単純な方法で社会的インタラクションの定量的な議論が可能かどうかを見極めることである。

一人称映像から社会的インタラクションを読み取ろうとする試みとして Fathi ら [6] の試みがある。彼らは、映像

中の顔の抽出・人物特定だけでなく、顔の方向や立ち位置まで特定し、その時間的変化パターンから社会的インタラクションの状況を推定することまで試みている。技術的には、彼らの試みは本稿で導入した手法を完全に包含するものである。本稿の目的は、もっと単純な（つまりコストが低く頑健である）方法で、実用的な社会的インタラクションの測定や分析が可能かどうかを議論することである。

3. 一人称映像記録からの対面者の抽出

本研究では、プロジェクト参加者のグループ活動が行われている様子をウェアラブルカメラで撮影し、その映像内に映り込んだ対面者の顔識別を行うことで、参加者間のインタラクションの量を数えることとした。機械的になされた顔識別の結果を基に参加者の社会ネットワークの変化を検討したい。しかし顔識別結果のデータは膨大であるため、カメラ映像から抽出された顔の数を人ごとに数え分け、任意の参加者間の特定の時間帯のインタラクションの量と見なすこととした。そして、社会ネットワークをグラフとして可視化し、特定の時間帯の社会ネットワークの変化をインタラクティブに閲覧できるビューワを作成した。つまり、各参加者をノードとし、任意のノード間のリンクの濃淡をその時間帯のインタラクションの量として表現することとした。そうすることで、プロジェクト中の作業の変化や参加者の役割に応じた社会ネットワークの変化が、直感的に閲覧可能になり、膨大な作業記録映像の振り返りの手掛かりとして有益であると考える。本章の以下では、「明和電機はこだてプロジェクト」の概要を説明すると共に、提案システムに必要な準備作業について説明する。

3.1 分析対象となるプロジェクトの概要

今回記録した「明和電機はこだてプロジェクト」はアートユニット明和電機のメンバーだけではなく、公立はこだて未来大学（以下、未来大学とする）の学生もステージに上がり、6 日間の短期間でステージを作り上げるプロジェクトであった。記録するにあたって参加者計 17 名について、表 1 にまとめた。明和電機グループ 4 名（A, B, C, D）のうち、A, B, C はステージに上がる主演であると共に、楽器のメンテナンスや演奏データ打ち込みなどの創作全般を担当する。D はマネージャーであり、ステージ周辺のこまごまとした渉外を担当している。未来大学の原田研究室の教員と学生 9 名（E, F, G, H, I, J, K, L, M）については K, L, M 以外の 6 名はステージに上がり、演奏内容も明和電機メンバーと一緒に作る立場であった。F, G（いずれも男性学生）は工員 B, C とほぼ同じ立場であり、ステージ上では工員を演じる。H, I, J の 3 名（いずれも女性学生）は楽器を持って演奏に参加する。未来大学の角研究室の教員と学生 4 名（N, O, P, Q）は参与観察者として、映像記録のみを行い、基本的にはステージ作業には関

*1 <https://bit.ly/2AhLAUT>

わっていない。日程については

- 1日目:未来大学から楽器等の撤収作業
- 2日目:舞台設置および練習
- 3日目:舞台作成および練習
- 4日目:舞台作成および練習
- 5日目:通しリハーサル
- 6日目:本番

となる。

表 1 参加者一覧

参加者	所属	性別	属性	カメラ装着
A	明和電機	男性	社長	○
B	明和電機	男性	工員	○
C	明和電機	男性	工員	○
D	明和電機	女性	マネージャー	○
E	未来大, 原田 G	男性	大学教員	○
F	未来大, 原田 G	男性	大学生	○
G	未来大, 原田 G	男性	大学生	○
H	未来大, 原田 G	女性	大学生	○
I	未来大, 原田 G	女性	大学生	○
J	未来大, 原田 G	男性	大学生	○
K	未来大, 原田 G	女性	大学生	○
L	未来大, 原田 G	女性	大学生	
M	未来大, 原田 G	男性	大学生	○
N	未来大, 角 G	男性	大学教員	○
O	未来大, 角 G	女性	大学生	○
P	未来大, 角 G	男性	大学生	
Q	未来大, 角 G	男性	大学生	

3.2 ウェアラブルカメラでの記録

今回は函館で開催された「明和電機はこだてプロジェクト」の準備期間から本番直前までの6日間を記録した。長時間の一人称視点映像を記録してもらうために、ウェアラブルカメラを胸部に装着した。使用するウェアラブルカメラはGoPro4であり、画角は垂直画角が69.5度、水平画角が118.2度である。作業をしている日の開始時間から終了時間までの間、装着してもらう。図2のような一人称視点映像による作業風景の記録から、装着者がどのような作業を行っていたか、誰と対面してコミュニケーションをとっていたかを把握することができる。

3.3 一人称視点映像による顔識別

ウェアラブルカメラを用いることで一人称視点映像から装着者がコミュニケーションをとっていた状況を記録している。しかし、コミュニケーションをしている人物を特定できない。社会ネットワークを可視化するためには、人物が誰であるかという情報が必要である。そこで顔識別を行い、誰と対面しているかという情報を取得する。

今回行った手順は、



図 2 一人称視点映像による作業風景の記録

- (1) 記録された映像を1秒ごとに切り出す
- (2) 画像に装着者の情報と日時を付与
- (3) Google フォトのフェイスグルーピング機能を用いて顔識別

となっている。今回記録した映像だと、切り出す画像の解像度は640×360ピクセルである。切り出した画像の総数は約70万枚である。その後、装着者ごとに分けていくと1人あたり約8万枚であった。この手順を踏んで出来上がったデータをもとにノード値の計算を行い、社会ネットワークをグラフとして可視化する。

3.4 社会ネットワークの可視化

3.3節で行った顔識別のデータをもとに社会ネットワークをグラフとして可視化する。各ノードを参加者として、インタラクション、つまり顔検出がなされた人物同士をリンクでつなぐ。参加者間の社会ネットワークのつながりが強いと、色やリンクの濃さが変化する。濃さの変化からノード間のつながりの変化を捉えることができる。また、スクロールバーを作成し、スクロールバーの幅を変えていくことで、任意の時間幅のグラフが表示される。この実装をすることにより、ノイズを無くし、時間経過するごとに、社会ネットワークの変化をグラフで捉えていけるようにする。

4. 社会ネットワークの変化の可視化

4.1 顔識別結果と全体の社会ネットワーク

今回の「明和電機はこだてプロジェクト」を記録した一人称映像を用いて、3.3節で提示した顔識別を行った。6日間で検出された顔を可視化すると図3のようになる。

全体で見るとステージに上がる人同士のリンクが濃く見える。また明和電機のB、Cに対して多くの学生から入るリンクが見える。しかし社会ネットワークを全体で捉えると、時間幅が大きすぎるため、異なるグループを見つけること、異なるグループ間をつなぐ役割の人を見つけることが難しい。したがって時間経過ごとに分けて社会ネット

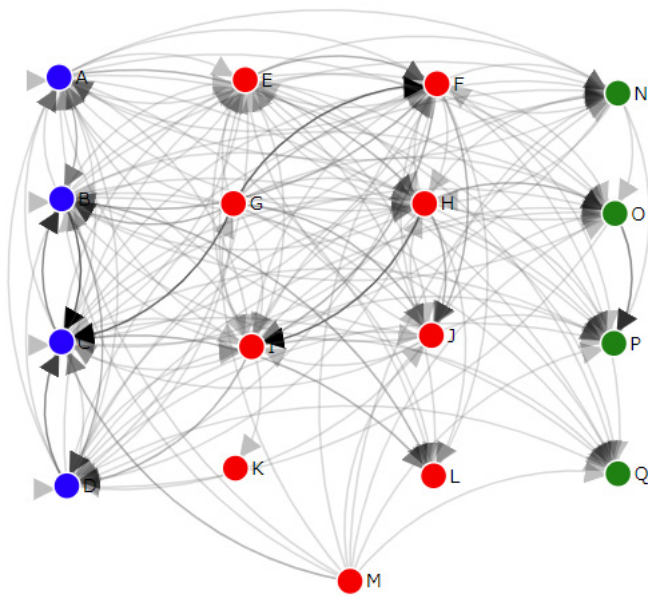


図 3 6日間を通して蓄積されたデータによって可視化された社会ネットワーク

ワークを見ていく。

4.2 日数経過による社会ネットワークの変化

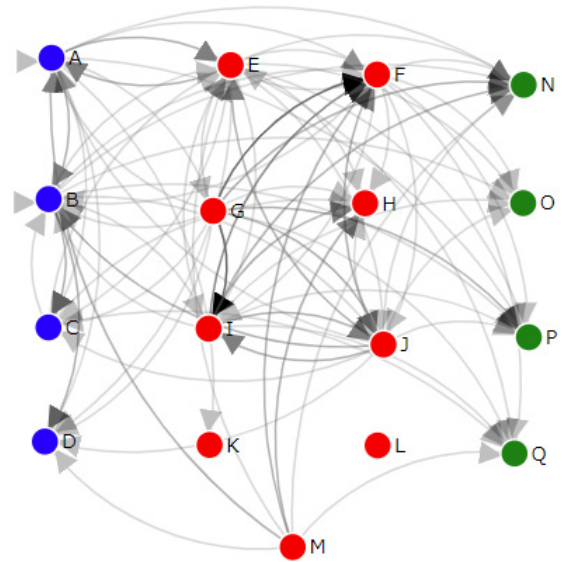
4.1 節では、顔識別の結果と全体の社会ネットワークの可視化について述べた。社会ネットワークを全体で見てしまうと、時間幅が大きく、役割や変化を捉えづらいという問題点がわかったため、社会ネットワークを日数ごとに分けて考えた。その結果図4のようになった。

これらの結果でもほぼすべての人につながりが見えた。しかしつながり方に変化が見られている。例えば1日目では学生同士のつながりが強く見られる。この日は撤収作業を行っていた。1日目ということから、明和電機と未来大学の原田 G でつながりが見られなかった。しかし5日目になると学生と明和電機の工員がつながっている変化が見られる。この日はリハーサルである。そのため工員役を務めている明和電機の B, C と未来大学の F, G は、ステージに上がる共通の役割を持っていた。よって互いにつながりがある社会ネットワークに変化した。

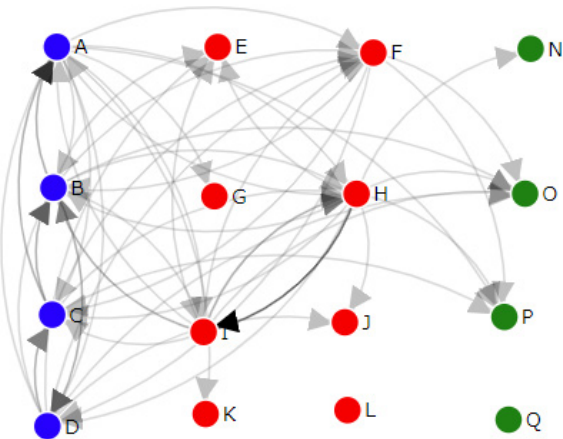
3日目の社会ネットワークを見るとは未来大学の原田 G に所属している H, I といった女性同士でリンクが結ばれている。それぞれの活動日を見ると、明和電機社長の A は全体を統括するリーダーという役割を持ちながら、全体的にリンクの少なさが見て取れる。そのため、他の人や学生と関わらずに作業していることがわかる。

5日目に関して言えば、リハーサルであったため、明和電機社長の A は全体とつながっていることが社会ネットワークから読み取れる。しかし実際の記録では、暗く距離が離れていたため、印象とは異なる結果であった。

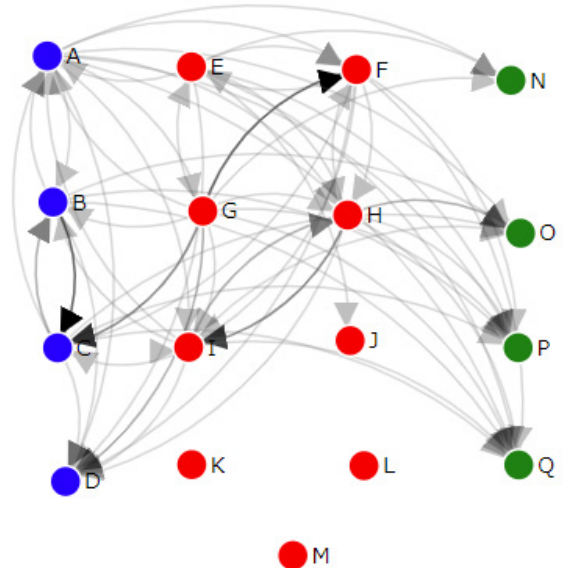
また原田 G の女子学生 3 名がステージ上で演奏する役割



(a) 1 日目



(b) 3 日目



(c) 5 日目

図 4 日にち単位での社会ネットワークの変化

を持っていたため、その役割同士で練習を行うことでグループが生まれたここでは図5のように女性で集まり、演奏を練習している場面で女性のグループが形成されていた。



図5 3日目の女子グループが形成されている場面

異なるグループが形成や関係が変化することを分析することができた。しかし日数ごとに社会ネットワークでは見ている時間幅が広すぎるので、時間ごとでどのような変化を捉えることで、どのタイミングで社会ネットワークの変化が起こるかをみることができる。また誰から誰へのリンクの変化を細かく見ることができる。

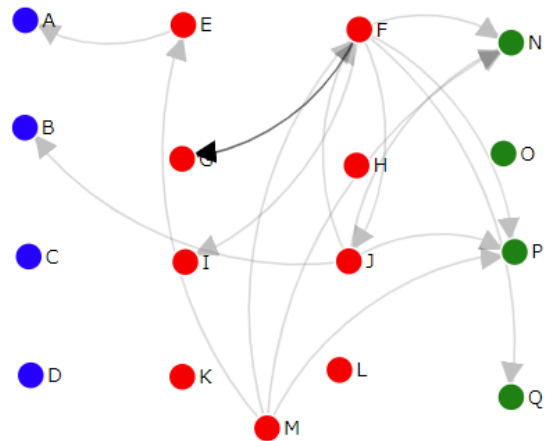
4.3 時間経過による社会ネットワークの変化

4.2節での結果をもとに組織間のハブになった一例として1日目の作業のうち、17:59からの10分間の社会ネットワーク(図6(a))と18:49から10分間の社会ネットワーク(図6(b))と19:49から10分間の社会ネットワーク(図6(c))を可視化した。

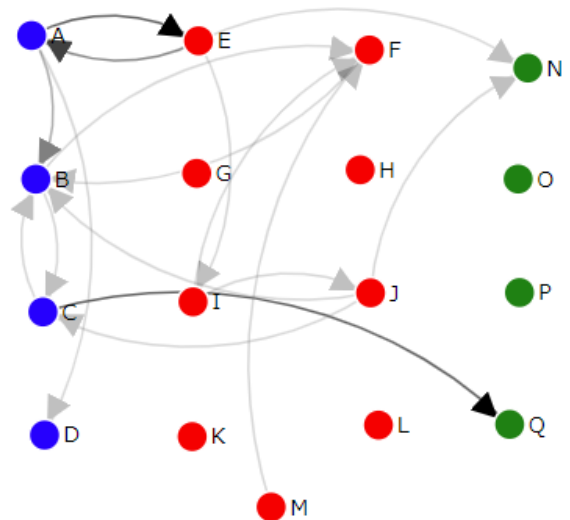
作業開始時は参加者全員が集まっていたわけではなく、作業自体もまだ行われておらず、学生同士でそれぞれグループが形成されていた。

図6(b)の18:49分から10分間の社会ネットワークを見ていくと、明和電機の人同士のグループと学生同士のグループでそれぞれ作業していることがわかる。またそれぞれのグループをつなぐハブの役割を持った人も出てきているのがわかる。特に未来大学の原田Gから教員Eと学生Fがハブとなり、学生につないでいる。この2人はもともと明和電機ともつながりがあり、顔馴染みであったため、早い段階でリンクがつながっていた。

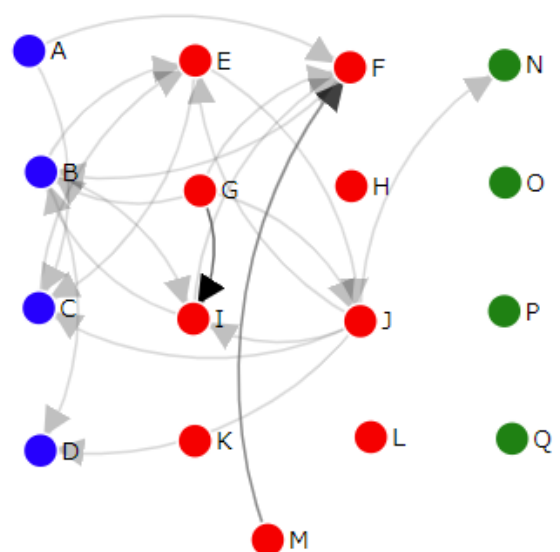
図6(c)の19:49から10分間の社会ネットワークを見るとやはり図6(b)の18:49の場合と同じような社会ネットワークが見える。未来大学の原田Gから教員Eと学生Iがハブとなり、学生につないでいる。さらに原田GのIが明和電機と未来大学の学生をつなぐハブとなる。学生Iも明和電機の方々と面識があったため、明和電機のグループに入っていき、また学生ともつながっているのが社会ネットワークから読み取れる。それ以外の参加者は、主に同一のグループ同士でつながりあって作業しているのが



(a) 17:59 から 10 分間の社会ネットワーク



(b) 18:49 から 10 分間の社会ネットワーク



(c) 19:49 から 10 分間の社会ネットワーク

図6 10分単位での社会ネットワークの変化の一例

読み取れる。作業 1 日目なため、同じ組織内の人同士でつながりあい、互いのグループのつなぐハブの人が若干見られた。そして異なるグループをつなぐハブの役割を持った人が現れたことを時間経過によって確認できた。今後は日数経過と時間経過を合わせ、日数経過や時間経過していくと同じグループの人から異なるグループへと関係の変化を捉えていけると考えている。

5. おわりに

本研究では、函館で開催された「明和電機はこだてプロジェクト」の準備期間から本番直前までの 6 日間をウェアラブルカメラを用いて記録した。記録した一人称視点映像から顔識別を用いて社会ネットワークの可視化するシステムの提案を行った。本稿では、一人称映像に映り込んだ顔情報を使い、顔識別を行った。顔識別では、一人称映像から画像に切り出し、切り出された画像を Google フォトで顔識別を行った。顔識別の結果を用いて社会ネットワークを可視化した。最初に全日程分のデータから、社会ネットワークを可視化した。その際に時間幅が大きいため、すべての人とのつながりが見えた。そのためグループ間をつなぐ役割の人、異なるグループ自体を見つけるのが困難であった。次に日数経過で社会ネットワークの可視化を行った。すると、ある程度の日数経過で社会ネットワークに変化が見られた。またステージに上がるといった役割に応じて社会ネットワークの動きが見られた。さらに細かい変化を捉えるために時間経過で社会ネットワークの可視化を行った。すると初日だけで明和電機と原田 G の異なるグループをつなぐハブの役割を持った人が現れた。これらの結果から日数経過や時間経過に合わせて、グループ間の変化を捉えていけると考える。

本稿はまだ道半ばな状態であるが、本研究の本来の目的はライフログデータを高速・効率的に見返すことで、自らの活動を振り返ったり、その後の改善につながる気付きを見出すツールを提供することである。これまでも、画像処理、音響処理、赤外線センサなどを用いて社会的インタラクションの自動インデキシングを行い、それをビデオ閲覧の手掛かりとする試みがなされてきた（例えば [7], [8], [9]）。今後、これらの知見を活用してツールを実装し、振り返り作業の効果を議論したい。

謝辞 「明和電機はこだてプロジェクト」への参加と映像記録へのご協力を頂いた土佐信道氏をはじめとする明和電機の皆様、原田泰氏をはじめとする公立はこだて未来大学の原田研究室の皆様、映像記録とデータ整備にご協力いただいた奥野茜、渡邊悠一、村上聖将、高崎昌宏の諸氏に深く感謝いたします。

参考文献

- [1] Tanzeem Choudhury and Alex Pentland. Sensing and modeling human networks using the sociometer. In *The 7th IEEE International Symposium on Wearable Computers (ISWC 2003)*, pp. 216–222, Washington, DC, USA, October 2003. IEEE Computer Society.
- [2] Alex Pentland. *Social Physics: How Good Ideas Spread – The Lesson From a New Science*. The Penguin Press, 2014.
- [3] 矢野和男. センサは Web を超える. 情報処理学会誌, Vol. 48, No. 2, pp. 160–170, 2007 年 2 月.
- [4] 早川幹, 大久保教夫, 脇坂義博. ビジネス顕微鏡：実用的人間行動計測システムの開発. 電子情報通信学会論文誌, Vol. J96-D, No. 10, pp. 2359–2370, 2013 年 10 月.
- [5] 佐藤信夫, 浅原彰規. ウェアラブルセンサ「ビジネス顕微鏡」を用いた展示施設における顧客満足度に関連する顧客行動分析. 第 14 回情報科学技術フォーラム (FIT2015), 第 4 巻, pp. 119–122. 情報処理学会, 電子情報通信学会, 2015 年 9 月.
- [6] Alirza Fathi. Social interactions: A first-person perspective. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1226–1233, Washington, DC, USA, 2012. IEEE Computer Society.
- [7] Keita Higuchi, Ryo Yonetani, and Yoichi Sato. EgoScanning: Quickly scanning first-person videos with egocentric elastic timelines. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*, pp. 6536–6546, New York, NY, USA, 2017. ACM.
- [8] Kai Toyama and Yasuyuki Sumi. Quick browsing of shared experience videos based on conversational field detection. In *9th International Conference on Mobile Computing, Applications, and Services (MobiCASE 2018)*, pp. 40–55, February 2018.
- [9] Yasuyuki Sumi, Masaki Suwa, and Koichi Hanaue. Effects of viewing multiple viewpoint videos on metacognition of collaborative experiences. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, pp. 648:1–648:13, New York, NY, USA, 2018. ACM.