

顔数計：一人称ライフログ映像からの社会活動計測

—社会的場における個や組織のウェルビーイングに向けて—
角 康之*

abstract

日々の社会活動量を測るために筆者らが研究開発している「顔数計（かおすうけい）」を紹介する。顔数計は、歩数計の類推から発想されたシステムで、胸に装着したカメラに映り込んだ対面者の顔の数を数えることで、一定時間内の社会活動量を計量するものである。本稿では、このような単純な仕組みで、社会的な場に参加している個々人の参与の深さを測ることが可能であることを示す。そして、筆者らが提案する顔数計と市販の身体活動計を併せて利用することで、日々のさまざまな活動における身体的・社会的な満足や疲労を多角的に理解することの可能性を議論する。

I はじめに

ウェルビーイングへの注目が高まり、情報処理学会誌でも特集が生まれ¹⁾、情報通信技術による人や社会のウェルビーイングの理解と支援について議論されている。本稿は、同様に情報通信技術の研究コミュニティに属する筆者が研究開発している「顔数計（かおすうけい）」^{2),3)}と呼ばれるシステムを紹介し、臨床現場での展開の可能性を提示するものである。

顔数計は日々の社会活動を計量することを目的としたシステムで、胸に装着したカメラに映り込んだ対面者の顔の数を数えることで、一定時間中の対面社会活動の量を計量する。歩数計が日々の運動量の目安として使われているように、顔数計は社会的な活動の量を直感的に把握するためのものである。「今日はたくさんの人と会って疲れた」、「今日は充実した日だった」といった感覚を、他者とのコミュニケー

ションの量を可視化することで自覚したり改善する手掛かりにしたい、というのが本研究の動機である。

図1に顔数計の利用イメージを示す。システム利用者は各々が胸に顔数計を装着して生活する。顔数計にはカメラが内蔵されており、一人称ライフログ映像を記録する。画像処理によって映像中に映り込んだ対面者の顔を検出し、それに基づいて時々刻々の参加度が計算され、その時間積分によってカメラ装着者の社会活動量が計算される。その結果は、一日の社会活動の総量と、各社会活動の割合を示すグラフの形でユーザーに提示される。

顔数計は、歩数計の類推から発想された。歩数計は元々、単純な体の揺れから歩数を数えるだけのものであった。しかし、身体動作の識別技術の向上により、歩数計は身体活動計に発展した。歩行、ジョギング、睡眠といった身体活動の見分けが可能になり、そのことによって、日々の運動量や睡眠量の振り返りや、ライフログのインデクス情報として活用

* はこだて未来大学教授



図1
対面者の顔の検出に基づいた
社会活動計測

されるようになった^{4),5)}。そこで筆者らも、対面者の顔の数を数えるだけの単純な仕組みから社会活動計測の検討を始めた。歩数計は、オフィスでの移動も家での家事も外での散歩も分け隔てなく、歩数の積み重ねで一日のおおよその身体活動量を定量化している。同様に、家族も知人も赤の他人も区別せず、他者との対面の量を機械的に数え、その積算によって一定時間中の社会活動を計量しよう、という割り切りが、筆者らの提案手法の特徴である。

以下、顔数計の仕組みと動作例を紹介し、顔数計が利用者の社会的場への参与の量と質を測る仕組みを説明する。また、顔数計で計測される社会活動量と、市販の活動計測デバイスから得られる身体活動量を組み合わせて、個と場の関わりを多角的に理解することへの可能性を議論する。

II 顔数計：対面者の顔検出に基づいた社会活動計測

本稿で紹介する顔数計が計量を試みる社会活動とは、実空間における対面状況において他者と何らかの関わり合いをもつ行為全般を指す。具体的には、2人から10人程度によって形成される立ち話、打ち合わせ、共同作業、共食などを想定している。計量対象者による、その社会的な場への関わり具合（つまり、発話したり積極的に共同作業に関与する度合い）を、本稿では対面的な社会活動への「参与度」と表現する。

参与度は時々刻々と変化すると思われる。社会学

者Goffman⁶⁾は会話に参加している人々を参与度の高さに応じた順序で、話し手、聞き手、傍参与者といった立ち位置に分類し、発話者交代と並行して動的にそれらの立ち位置を上ったり下りたりする現象を議論した。本研究ではその考え方を受け、動的に変化する参与度を、ある時点で計量対象者に向けられる他者の顔の数で表し、その時間積分を社会活動量と定義した。社会的な場への参与は、本来はその当事者の内面から生じるものであると考えるのが自然である。それにも拘わらず、当事者の周辺にいる人の反応（すなわち、当事者に顔を向けるという行為）によって、間接的に、当事者の参与度を測ろうとするところに、本手法の面白さがあると考えられる。

ここまでは基本的な考えを伝えるために、説明を単純化して「対面者の顔の数を数える」と記してきたが、本研究では重要な工夫をしている。それは、対面者の近接性と時間継続性を参与度計算に考慮することである。つまり、装着カメラの映像中で検出された顔ごとに、その大きさ（物理的近さに比例する）と、検出の時間継続性（カメラ装着者へ顔を向け続けた時間に対応する）を考慮して、値の重みづけを行う。そうすることで、例えば雑踏のなかで多くの異なる顔が入れ代わり立ち代わり検出される場合はその時間帯の社会活動の参与度を過大に加点することを避け、逆に、少人数だとしても同一人物と向き合って継続的な対話をしているような状況での参与度は重視して加点することが可能となる。

図2の具体例を見ながら、提案手法の効果を説明する。3種類のシーンについて、各フレームごとの

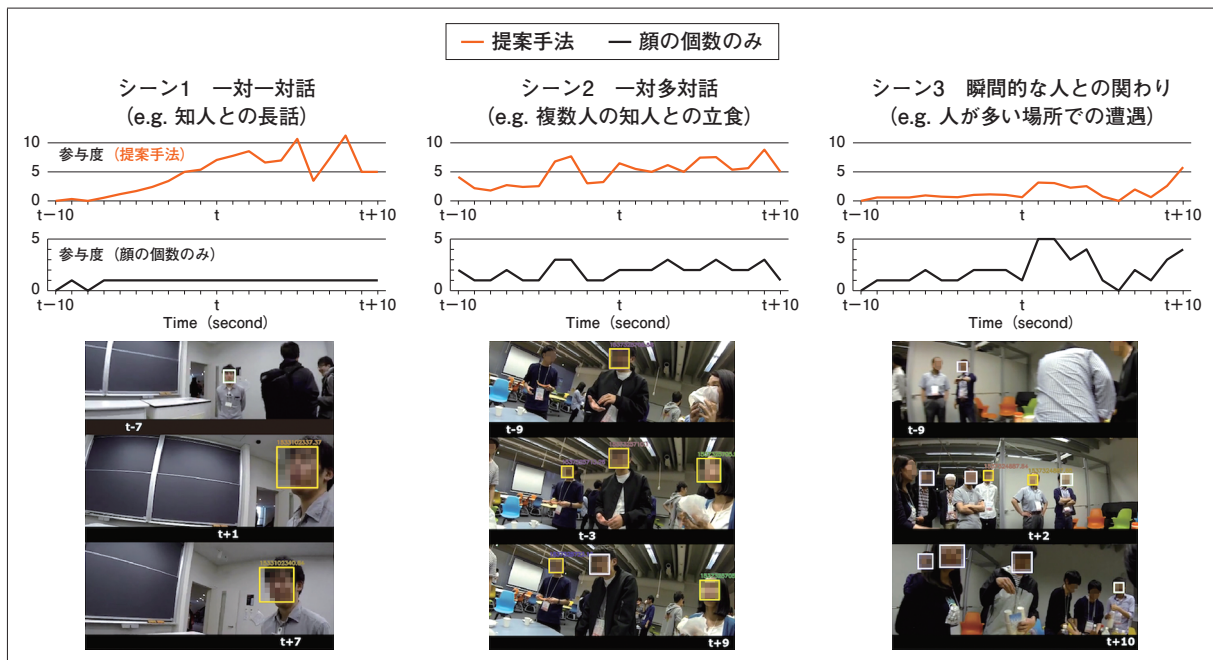


図2 参加度計算の例：一対一対話，一対多対話，瞬間的な関わり

参加度の変化をグラフにした^{注1)}。カメラ装着者は、シーン1では知人と一対一で長話をしており、シーン2では複数人が集まった場所で立食をしており、シーン3では大勢の人々が集まっている場所を渡り歩いている。提案手法による計算結果と併せて、検出された顔の個数をそのまま表示したグラフを示す。図の下部には、各シーンの代表フレームにおける顔検出結果を載せた。

顔の個数を単純に参加度として使う場合、カメラ視野内で検出された顔の数がそのまま使われるので、大勢が密集している場所に身を投じている場合に値が大きくなる。なので、シーン2とシーン3に比べてシーン1の参加度は低く評価される。一方、シーン2とシーン3の違いは顕著には現れない。

それに対して、提案手法では対面者との距離と対面継続時間を考慮しているため、シーン1のように対面者が一人の場合でも継続的な対話をしているシーンの参加度は徐々に高くなる。一方、シーン3のように各時刻で大勢の人と対面していたとしても、対面時間の継続性や対面者との距離が遠いまの場合、提案手法では参加度は低い値となる。また、同様に複数人と対面している場合、シーン2のよう

に同一の人々とししばらく対面時間が続く場合は参加度に寄与する。

注1) ここでは1秒ごとに1フレームとして顔検出の画像処理を施したので、合計20フレーム、約20秒間に相当する。

III 顔数計による出力結果の妥当性の議論

詳しくは文献³⁾をご覧いただきたいが、筆者らは、提案手法によって提示される社会活動量が、人の印象をどの程度再現できるかの実験を行った。具体的には、ライフログの当事者とその知人の合計8人を実験参加者とし、ライフログから抜粋した複数の映像を閲覧してもらって、それらのシーンにおけるカメラ装着者の社会活動量についての印象を答えてもらった。その主観的な印象が、当事者も含めて一定の傾向に収斂すること^{注2)}を確認したうえで、筆者らの顔数計が提示する値がその大小関係と対応することを確認した。つまり、少人数会話であっても当事者が積極的に発話していたシーンは、大勢の雑踏のなかに受動的に参加していたシーンよりも高い社会的活動量を示すことが再現できた。

注2) 例えば、少人数であってもカメラ装着者本人が積極的に発話

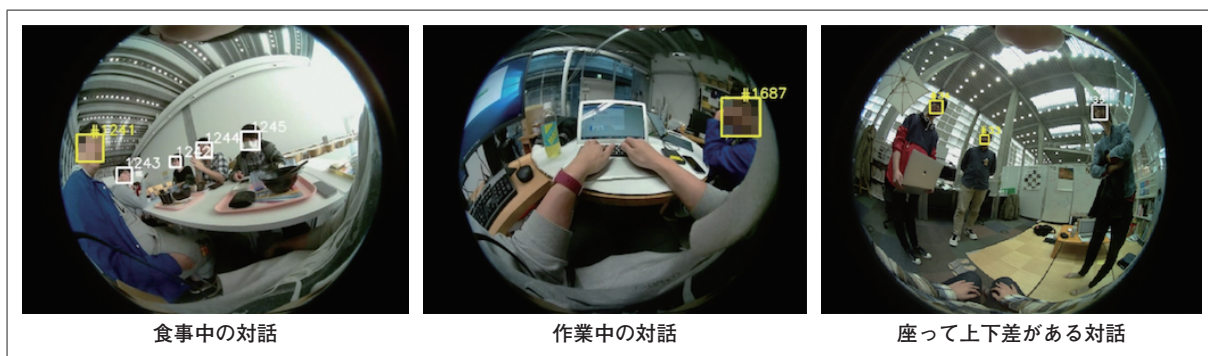


図3 広角カメラを採用した顔数計に映り込んだ対話場

参加している立ち話のシーンは社会的活動量が高く評価され、一方、大勢のなかの一聴衆者として雑踏のなかに参加しているだけのシーンは低く評価された。そしてその評価は、当事者本人とビデオ閲覧をした第三者でも同じ傾向を示した。

ここで注目すべきは、顔数計は装着者や周りの音声情報を利用してないにも拘わらず、当事者の会話参与（具体的には、発話や相槌による積極的な参加）の多寡を測ることができることである。会話場への参加者の参与構造を理解するには、各々の発話の有無、頭部運動、身振り手振り、立ち位置の変化といったさまざまな言語・非言語行動から総合的に理解する必要があると考えるのが常識的である。実際、筆者らもモーションキャプチャシステム、視線計測装置といった計測機器を用いた計測・分析^{7),8)}や、音場共有に着目した会話場検出^{9),10)}の研究を行ってきた。

しかし、そういった緻密なデータによる微視的な会話構造理解は多くの特殊なセンサを必要とし、日常的な運用には向かない。そこで顔数計では、胸に装着した小さなカメラデバイスで対面領域に他者の顔があるかどうかだけ判別することで社会的な場の状況検出を試みることにした。その結果は意外なことに、十分直感に近い傾向をとらえられることが確認された。

音声情報を利用することもなく装着者の参与度を測ることができることは、当初、筆者らにとっても不思議であった。が、よく考えると、装着者が発話したり重要な手作業によって社会的場に強く参与しているとき、周囲の人々の顔はその人の方に向くであろう。そして、それが結果的にその時間帯の当人

の顔数計の値の上昇として現れるのである。もちろん、細かい微視的な構造を観察するには顔数計は不向きであるが、数時間、数日といった巨視的な傾向を観察するには十分である。デバイス内で顔検出と参与度の計算が済めば映像を記録する必要がなく、また、音声情報も扱わないことを考えると、プライバシー保護の観点からも優れていると考える。

なお、顔数計試作の試行錯誤のなかで気づいたのだが、顔数計に利用するカメラの視野は十分に広いことが重要であった。なぜなら、人は少人数で密に社会的場を形成するとき、例えば、立ち話をしたり、一緒に作業をしたり、食事をするときは、真正面に対面することよりも、斜め横に並ぶことが多い(図3)。したがって、そういった社会的な場とそこへの関わりを計測するには、身体の前を広くとらえることができる広角(魚眼)カメラが必要であることがわかった。また、視界の中心部分に映り込んだ顔よりも、むしろ視界の周縁部分に写った顔を重視して加算するほうが、当事者の直感に近い社会活動計測が可能なのではないかと考え、現在、検討を進めている。

こういった経験から、オンラインビデオ会議だと、各々の参与度と関係なく常に複数人の顔が正面から向き合っていることが、オンライン生活の疲労の原因の一つなのではないかと想像している。顔数計を実世界の対面状況だけでなく、オンライン・コミュニケーションにおける疲労の計測にも利用してみたい。

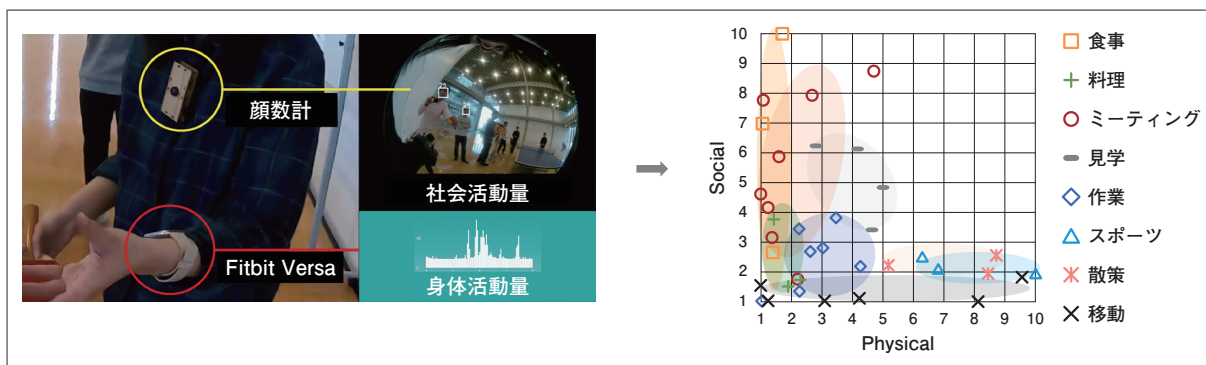


図4 身体活動量と社会活動量の関係の可視化

IV 身体活動と社会活動の統合的理解に向けて

近年、腕時計型の身体活動計測器が普及し、多くの人が日常的に自らの生活の振り返りや健康状態の把握に利用している。ここまでで紹介してきた顔数計による社会活動計測は、我々の生活の振り返りにもう一つの視座を与えるものであると考える。

図4に、日常生活のさまざまな活動シーンを身体活動量と社会活動量の2軸で可視化した例を紹介する¹¹⁾。グラフの横軸は身体活動量、縦軸は社会活動量に対応し、食事、ミーティング、スポーツといったさまざまなシーンをグラフ上にプロットした。身体活動量は、腕時計型の活動量計であるFitbit社のVersaを利用し、そのWebサービスから提供される消費カロリー量の10分間の総量を用いた。社会活動量は本稿で紹介してきた顔数計によって計測された社会活動量である。なお、いずれの軸も当人の計測期間中の活動量の最小値、最大値に基づいて正規化した値を用いた。

一見するとわかるように、移動、散策、スポーツといった活動は基本的には身体活動の一軸で測定可能なものであり、社会活動としての量は少ない。一方、ミーティングや食事といった活動は、状況に大きく依存するものの、身体活動よりも社会活動の側面が大きいことが確認できる。また、作業や見学といった活動は、適度に社会活動と身体活動を両立した活動であることがわかる。

この例はある一女子大学生の数日間のライフログ

データを用いた一例に過ぎないが、可視化結果は、当人にとっても、周りにいた友人にとっても直感に合うものであった。また、同じシーンの活動に参加していても、性別、年齢、立場によってこれらの値は異なることも確認された。したがって、こういったシステムは、個々人が自らの生活を振り返りつつ、組織のなかの構成員の個性の相違を理解する手掛かりになるのではないかと考える。

また、図4のグラフから読み取れるように、身体活動と社会活動のバランスという観点からは、ミーティングと共食という、まったく異なる活動が似たような位置づけにあることを知ることができる。例えば、「今日は一人で仕事をしている時間が長かった。とはいえ、ミーティングをする予定はないので誰かと一緒に食事をしよう」とか「今日は大勢の人と会って疲れたので、食事は一人でとろう」といったことについて、より自覚的に判断できるようになるであろう。こういう技術が、個人や組織の個性を定量的に知る視座を提供できればと考えている。

V おわりに

胸に装着したカメラによる一人称ライフログ映像に映り込んだ対面者の顔の数を数えることでカメラ装着者の社会活動量を計測するシステム「顔数計」を紹介した。社会的な場への参与の深さを測るために、検出された顔ごとの近接性（検出された顔画像の大きさ）と時間継続性（顔が検出された連続時間）

の重みづけをすることで、顔数計は利用者本人の主観評価をよく再現することが確認され、単純に顔の数を数えるだけの手法よりも明らかに適切な結果を提示できる。

これからの課題は、このようなシステムを日常利用することで、社会的場におかれた個人や組織の観察、理解、改善に役立てることである。本稿で紹介した顔数計は、同じ場にいる場合でも立場が異なる個々人の場への参与の違いを顕在化する。したがって、例えば、ミーティングに参加するメンバの参与度の違いや時間的変化を認識することや、臨床的現場でケアを提供する人と享受する人の間の参与関係を理解し改善する手段として利用することなどが考えられる。

参考文献

- 1) 湯浅 晃 (編) : 特集 ウェルビーイングとテクノロジーに関する動向. 情報処理学会誌 **65** : 292-295, e1-e33, 2024
- 2) Okuno A, Sumi Y : Social activity measurement by counting faces captured in first-person view lifelogging video. Proceedings of the 10th Augmented Human International Conference 2019 : 1-9, 2019
- 3) 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 **62** : 607-616, 2021
- 4) Guo F, Li Y, Kankanhalli MS, Brown MS : An evaluation of wearable activity monitoring devices. Proceedings of the 1st ACM International Workshop on Personal Data Meets Distributed Multimedia (PDM '13) : 31-34, 2013
- 5) 牧川方昭 : 加速度センサを用いた日常身体活動のモニタリング. 生体医工学 **54** : 96-103, 2016
- 6) Erving Goffman : Forms of talk. University of Pennsylvania Press, 1981
- 7) 角 康之, 伊藤禎宣, 松口哲也, シドニーフェルス, 間瀬健二 : 協調的なインタラクションの記録と解釈. 情報処理学会論文誌 **44** : 2628-2637, 2003
- 8) 角 康之, 矢野正治, 西田豊明 : マルチモーダルデータに基づいた多人数会話の構造理解. 社会言語科学会誌 **14** : 82-96, 2011
- 9) Nakakura T, Sumi Y, Nishida T : Neary : Conversational field detection based on situated sound similarity. IEICE Transactions on Information and Systems **E94-D** : 1164-1172, 2011
- 10) Toyama K, Sumi Y : Quick browsing of shared experience videos based on conversational field detection. International Conference on Mobile Computing, Applications, and Services (MobiCASE 2018) : 40-55, 2018
- 11) Okuno A, Sumi Y : Classification of daily activities based on the amount of social and physical activity for behavioral change toward wellbeing. 13th Augmented Human International Conference (AH2022) : 1-9, 2022