

視線情報を用いた労働状態推定手法の検討

出崎達也[†] 小坂隆浩[†]

概要：近年、リモートワークを取り入れる企業が急速に増えている。しかし、リモートワークには目の届かない所で働いている様子が分からないといった不安の声が多く、遠隔での労働状態を把握することが重要である。遠隔での労働状態を把握するための手段としてアイトラッカー等を設置し、視線に関する情報から労働状態を分析する研究が報告されているが、特別な機器を必要とするため、余剰な費用がかかり、大規模なリモートワーク環境での実用が困難であることが課題となる。本研究では、アイトラッカーを使用せず、視線情報を得る手法としてPC等に備わったwebカメラから視線情報を予測する技術に注目し、この技術を利用した労働状態推定手法について検討した。

キーワード：リモートワーク，視線検出，webgazer

1. はじめに

近年、リモートワークを取り入れる企業が急速に増えている。しかしリモートワークには「目の届かない所で集中して働いているかどうか分からないといった不安の声」^[1]があり、労働状態を把握する方法が必要とされている。本研究における労働状態とは、広義には集中して働いているかどうか、狭義にはどの程度集中して働いているかと定義する。リモートワークにおけるワーカーの労働状態の把握方法は、企業によって多様な方法が用いられてきた。既存のアプリ^[2]を例に挙げると、着席と退席を自動的に送信することで勤務時間を管理するアプリなどが存在するが、使用者の自己申告に基づくアプリであり、意図的に偽装し、働いているように見せることが可能である。またこのアプリでは、ほかのユーザーの着席と退席に関する画面を見ることが可能だが、画面内容の精査には人間による判断が必要であり、すべての状態を判断することは難しい。既存研究では、情報不足や偽装を解決するために、腕に装着するウェアラブルデバイスから心拍数や生体情報を利用したり^[3]、アイトラッカーを設置して視線情報を利用するなどといった方法が提案されている^[4]。しかし、特別な機器を必要とする方法は、機器の調達に多くの余剰費用が掛かり、実際のリモートワーク環境での利用が困難であることが課題となる。この課題の解決に対して、心拍数などに代表される生体情報は、特別な計測機器を用いなければ予測することが極めて困難なのに対して、視線情報に関してはノートPCなどに多く配備されているwebカメラからの映像から、視線情報を得て、視線予測が可能である。

本研究では、リモートワークにおける労働状態の把握と余剰費用の削減を目指して、webカメラからの視線予測を用いて労働状態を推定する手法について検討することを目的とする。

2. 予備実験

具体的な手法の検討に移る前に先行研究からwebカメラ

を用いた視線予測精度の調査と複数の作業において実際に視線を測定した結果を述べる。先行研究ではフレームワークとしてwebgazer^[5]を使用した。webgazerはブラウン大学の調査をもとに開発された、ユーザーが画面のどこを見ているかという視線を予測するアプリケーションであり、javascriptを利用してwebサイトのどこを見ているかを調査する。webgazerの特徴として、マウスをクリックした際、ユーザーがカーソルをクリックした点を注目していると仮定して学習を行い、予測をリアルタイムで改善していくことと、予測に際してカーソルの現在位置を大きな情報源としていることがあげられる。後者についてはカーソルを動かすと予測もそれを追従するような動きを見せること、またカーソルから遠い場所を見ている場合でもカーソルに近い位置を予測することがあげられる。

まず、webカメラを用いた視線予測精度の調査結果について述べる。調査方法は画面を図1のように3×3などの升目に分割した際に被験者がみている領域を正しく判別できるかといった方法で行われた。結果として表1示すように画面を3×3の升目に分解したときまで、9割を超える高い精度を保ち、4×4などのそれ以上の細かい升目に分割したときは、8割を下回る低い精度となった。

次に複数のアプリケーションにおいて実際に視線を測定した結果について簡潔に述べる。この実験の結果、作業の種類ごとに、縦の視線移動と横の視線移動に大きく変化があることがわかったwebサイト閲覧などの画面から情報を収集するような作業では横の移動より縦の移動が大きくなり、入力作業などの作業においては横の視線移動が増える傾向にあることがわかった。

1	4	7
2	5	8
3	6	9

図1 3×3の画面分割の例

[†] 同志社大学大学院
Graduate School of Doshisha University

表1 画面の分割ごとの平均

画面の分割	平均正解率(%)
2×2	99
3×3	95
4×4	78

3. 労働状態推定手法の検討

3.1 問題点

予備実験の結果をもとに労働状態推定手法について検討する。視線による労働状態推定手法においては、汎用性が重視される。社会規模での使用を考慮したときに、より多くの業種でより多くの個人が使用可能であるほうが望ましい。汎用性は、2つの要因によって阻害される。

1つは、個人ごとに存在する個人差である。些細な癖などから始まり、問題に直面した時における解決方法などといった点まで含まれる。具体例としては何かしら知らない情報に直面した際、PCで検索するかスマホで検索するかといった差異があげられる。この例において問題になるのがスマホで解決する側が画面外を見ることである。当初画面外を見るといった行為はよそ見、わき見といった行為に限定されると考えていたのだがこのような例を考慮するとその限りではなくなってしまう。これが個人差による汎用性の阻害の一例である。

もう1つが作業内容による汎用性の低下である。作業内容と、それによって使用するアプリケーションが視線の基本的動作に変化を与え、汎用性が低下してしまうことである。

労働状態推定において個人差と作業内容の2点が問題となる。この2つの問題に対して可能な限り汎用性を維持していくことが、労働状態推定手法を検討する過程において基本の方針となる。

3.2 対策の考察

個人差と作業内容の2つの問題点に対する対策について考察する。

まず、個人差に対する対策を考察する。個人差が存在することを前提とした場合、その対策は個人で変化しない動作を探していく方法と個人ごとに最適化していく方法の2つが方針として考えられる。個人ごとに最適化していく方法については膨大なデータ収集と機械学習を高度に組み合わせればあるいは可能ではあるが、先に挙げた汎用性を保つといった命題にはそぐわず、また機械学習用のデータを作成するためにもある程度型にはまったテンプレート、つまりは個人で変化しない動作しない動作を最初の段階として模索していく必要があるため本研究では深くは考慮しない。では個人で変化しない動作を探していく方針について

は予備実験での縦の移動と横の移動の変化から考察していくことができる。予備実験において縦と横の変化量の差が見受けられたのは現代のパソコンのデスクトップの仕様が影響していると考察することができる。ウィンドウは基本的に上から下にスクロールされるため1つのウィンドウに集中している場合横の動作が少なくなり、情報を読み取る場合必然縦の移動量が大きくなる。このように個人差には必然的に生じるであろう差を利用していくことが考えられる。

しかし、この論理は1つのアプリを全画面で使った作業の場合大きく変わることが予測される。2つ目の問題点である作業内容の差異による汎用性の阻害が大きく影響する。例えばwordやexcelなどの多くのアプリは上にツールバーを置きその下を作業場としていることが多い。画像編集ソフトであるgimpなどはウィンドウ左側にツールボックスが存在する。しかし、アプリケーションの画面構成は大別してツールボックスと主な作業領域の2つが主流であり作業内容の変化に対する対策としてはアプリケーションの作業領域と非作業領域を大別することができればある程度の場合分けが可能である。

3.3 具体的検討

次に労働状態推定手法の具体的検討を行う。3.2節で述べた考察に基づきアプリケーションの画面構成を事前取得しておくことを前提とし、労働者の視線移動について特に縦移動、および、横移動に着目しその特性を明かしていくことが必要となる。そのために、まずアプリケーションの画面構成について分析した。

現在のアプリケーションは、大別して図2のような縦長の画面設計が一般的である。ここでの主作業領域とは、アプリケーションにおける主題となる操作を行う領域、補助作業領域とは、例えばツールボックスやリボンバーなどに代表される作業を補助するためにUIなどが置かれている領域とここでは定義する。図2の例としてはwebブラウザ全般、Microsoft社のwordなどのアプリケーションが例に挙げられる。また専門性が高いアプリケーションなどでは必要なツールや表示する情報が増えた結果、図3のように画面左側にも補助作業領域が設置され横長の画面設計になる場合もある。しかし、どちらの場合にもアプリケーションは、主となる主作業領域とそれを補助する補助作業領域の二つの領域に大別することが可能であり、アプリケーション外などの画面の余白も含めれば3の領域に分割し、簡略化して考えることが可能である。



図2 アプリケーションの画面構成の模式図1



図3 アプリケーションの画面構成の模式図2

4. 実験

2つの画面構成についてその視線特性を調べるために評価実験を行った。被験者4人を対象にそれぞれ図2、図3で示されたアプリケーションの画面構成のアプリケーションに対して視線の計測を行いその特性を調査した。図2で示された画面構成の調査のための実験としてMicrosoft社のwordで作文する作業を2回に分けて行った。作業内容としては文章作成とし、1回目は簡単な課題を行うことでスムーズに書いてもらい、2回目は難度の高い課題を行い対して考えながら書いてもらうことを促し、作業状態を変化させた。これらの実験をそれぞれ実験1、実験2と呼称する。図3で示された画面構成の実験としてパワーポイントにて、簡単な図と難解な図を同様にそれぞれ作成してもらった。これらの実験をそれぞれ実験3、実験4とする。

4.1 結果

実験1から4について作業領域、補助作業領域、余白の3領域について、それぞれ1秒ごとに視線が遷移する確率について代表的に1人を表したものを、それぞれ図4から図7に示す。1人を代表した理由としては、基準となる数値の個人差が大きかったため、平均をとるとデータの意義が失われるためである。注目すべきは作業領域から作業領域に遷移する確率である。実験1は97%、実験2では96%、実験3では85%、実験4では88%となっており、powerpoint

を使った実験では、wordを使った実験より低下している。これは図形の作成という課題の特性上、補助作業領域との視線移動が増えた結果であると考えられる。また、実験1と実験2の値が差異3%、実験3と実験4も差異3%と近いことから、これらの値はアプリケーションごとにある程度固有の値を持つことが予想される。作業領域から作業領域に遷移する確率について、4人の被験者について表2に示す。被験者3については、被験者1と同様に実験1、実験2と実験3、実験4とある程度類似した値となった。しかし、被験者2は、被験者3の例に反するような結果となり、被験者4に関しては画面を見てすらいらないという結果となった。今回の実験はオンラインでの実施となり、オンラインによる説明不足や、間違いがあっても本人が気づかなければ、結果が出るまで分からないといった事例も見られた。本結果はその影響を受けた可能性があり、特に、個人ごとの視線予測の最適化が適切でなかった可能性もある。被験者2には、最低限の最適化は行われたが必要十分なレベルには達しておらず、被験者4は、誤った学習が行われた結果、あらぬ方向を常に見ているという結果となった。4人に対しての実験で半数が間違いを起こした以上、こうした事例は社会でも数多く起こると予想できる。労働推定手法にはこうした人的エラーに対応するため対策として、精度が低すぎると感知したならば、使用者に警告を出すといった機能も必要不可欠といえる。

次に、状態ごとの視線移動量の差について、特に縦移動に有益な結果が見られた。被験者1について実験1と実験2の縦の視線移動量の5秒ごとの移動平均をとったグラフを図8に示す。全体として実験2より実験1のほうが移動量が大きい結果となっており、全体の平均をとると実験1では193(単位はピクセル)、実験2では103と明確な差異が見られた。powerpointでの実験3、4の結果を図9に示す。図9は、図8と異なり分かりやすい特徴を読み取るとは難しいが、平均をとると実験3が85ピクセル、実験4が80ピクセルと、難度の高いほうが視線が落ち着いているといえる傾向が見られた。同様に高い視線予測精度となった被験者3でも、実験1にて179ピクセル実験2にて148ピクセル、といった関係と実験3にて189ピクセル、実験4にて142ピクセルといった結果になり、ある程度関係性が見受けられた。

実験結果のまとめとして個人差について、視線の縦移動量について共通した傾向を持つ可能性があるが、実数値に関して類似性はなく、個人ごとの基準値ともいえる数値について学習する必要があると考えられる。作業内容ごとの差に関しては、事前に作業領域についての割り振りが可能であれば、視線は主に作業領域に残留することから対応が可能であるといえる。

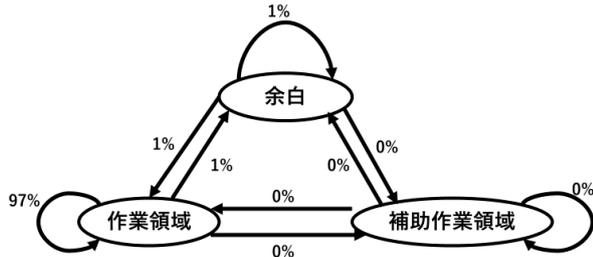


図4 実験1：word を簡単に行った場合の遷移モデル

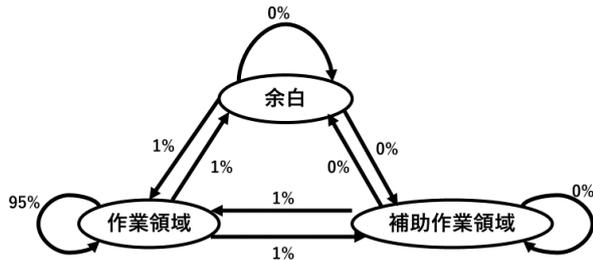


図5 実験2：word をじっくり行った場合の遷移モデル

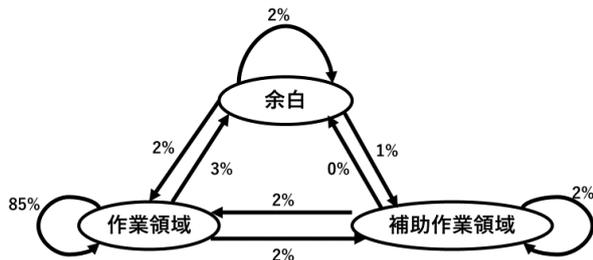


図6 実験3：powerpoint の簡単な場合の遷移モデル

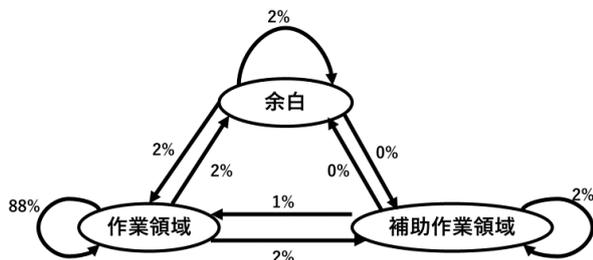


図7 実験4 powerpoint での難度の高い遷移モデル

表2 4人の被験者の実験結果

	実験1	実験2	実験3	実験4
被験者1	97	95	85	88
被験者2	31	33	15	43
被験者3	95	81	64	74
被験者4	0	0	0	0

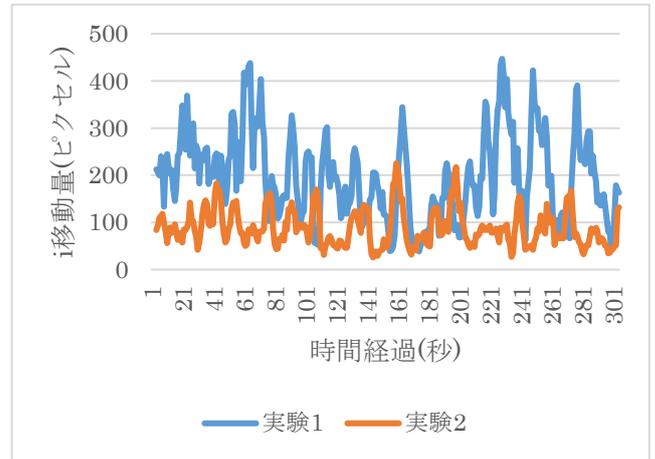


図8 視線の縦移動量 (実験1,2)

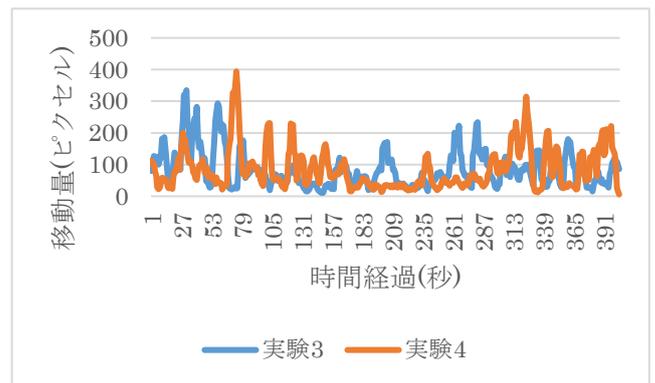


図9 視線の縦移動量 (実験3,4)

5. まとめ

本論文では視線を用いた労働状態推定手法について調査・検討した。先行研究の実験評価の結果から、アプリケーションの画面構成から主作業領域、非作業領域に分割し調査する手法を提案し、評価実験を行った。実験結果から、各領域の遷移確率はアプリケーションごとにある程度の特徴が見られる可能性があることと、難度の高い作業の場合、難度の低い場合より視線に落ち着きが見られる可能性があることを示唆した。

参考文献

- [1] 佐藤彰男, 国内における実証的テレワーク研究の展開, 大手大学人文科学部論集, 4巻, pp.165-180, 2003.
- [2] 業務効率化のための社内業務システムの操作性改善と定量評価手法の考案, 小川 晃司, デジタルプラクティス, Vol.10, No.4, 2019.

- [3] 林正幸, 吉川寛樹, 内山彰, 東野輝夫, 自律型生体データ収集における腕装着型センサの信頼性推定法の検討, マルチメディア, 分散, 協調とモバイル(DICOMO2019)シンポジウム, pp.1409-1414, 2019.
- [4] Tarmo Robal, Yue Zhao, Christoph Lof and Claudia Hauff, Webcam-based Attention Tracking in Online Learning: A Feasibility Study, In Proc. Of the 23rd International Conference on Intelligent User Interfaces, pp.189-197, 2018.
- [5] Alexandra Papoutsaki, Nediya Daskalova, PatsornSangkloy, Jeff Huang, James Laskey, and James Hays, WebGazer: scalable webcam eye tracking using user interactions, In Proc. of the IJCAI'16, pp.3839-3845, 2016.