

待機中の人の状況理解に向けた姿勢変動推定 - 画像解析による - Pose estimation for understanding situation of person in waiting - Based on image analysis -

井手 美優 長谷川 雄史 洲鎌 康 虹川 雅浩
Miyu Ide Takehumi Hasegawa Yasushi Sugama Masahiro Abukawa

三菱電機株式会社 情報技術総合研究所
Information Technology R&D Center, Mitsubishi Electric Corporation

1. はじめに

現状の機器やシステムは、音声コマンドやボタン、テキスト情報等の明示的な入力に基づいて動作を行うものが多い。しかし近年、機器やシステムの付加価値向上手法の一つとして、機器自身はその場の状況に応じて、明示的な入力なしで最適な動作・制御を行うことが求められている。そのためには、機器がユーザの状況をセンシングし判断する技術が必要である。

本稿では、昇降機やタクシーの待機時等、立位状態で機器の応答を待機中の人の状況理解について取り上げ、映像情報から待機時のユーザ状況を推定する技術開発を行った。

2. 研究目的

待機中の人の状況に応じた機器制御に有用な要素として、人の「イライラ度」のセンシングが挙げられる。イライラ度の定量化手法はいくつかある。例えば、生体情報を用いた手法では、心拍、脳波等の生体情報や、室温等の環境情報をもとに、疲労及び快・不快感情を推定する方式がある[1]。また、非接触で感情を推定する方法として、音響分析による音声から感情量の定量化の取り組みもある[2]。しかし、生体情報を用いる手法は接触センサが必要となり測定が困難であること、音声を用いる手法はユーザが発声を行う必要があることから、上記手法は待機中の人の状況理解への適用に向いていない。

そこで本稿では、待機中の人に適用可能な新たなイライラ度の客観指標のために用いる情報として、非接触で取得できる監視カメラ等の画像・映像情報に着目した。非接触で情報を得られるため情報収集が容易であり、また情報取得のための新たな機器の増設が不要であるため、実用化の際にコストを抑えられる。画像・映像から得られる情報として、人の表情、動作（姿勢）等が挙げられるが、待機中の人は感情の表情表出や表情の変化が非常に小さいため、感情の検出が難しい。加えて、表情検出可能な撮影範囲が狭く、カメラと顔の距離が離れている場合や、カメラ以外の方向を向いている場合は表情検出を行えないため、表情情報を用いたイライラ度の客観評価を機器の増設なく実用化することは困難である。一方、姿勢・動作は、検出可能な撮影範囲が広く、待機中のユーザであっても通常のカメラ1台で抽出可能なため、姿勢情報を用いたイライラ度の客観評価を検討することとした。

3. 感情推定手法の提案

3.1 観察実験

待機時のユーザの感情変化について観察・検証するため、機器の応答を立位状態で待機する実験を行った。被験者は20～50代の男女18名で、事前に実験内容の説明を行い、

実験参加への同意を得た。被験者はロビーに立って機器の応答を待機し、その様子をビデオカメラで撮影する。機器応答までの待機時間を長く設定することでイライラ感情を誘発させるタスクを実施し、撮影した被験者の様子を観察することで特徴を抽出する。待機時間 T [s]は $T = 0, 90, 150, 180$ の4種類とし、各被験者に対し計4回待機タスクを実施した。各タスク実施前後には十分な安静時間を設け、前に行ったタスクによる感情変化の影響が小さくなるようにした。さらに、待機時間が異なるタスクを複数回行うため、順序効果の影響が小さくなるよう、被験者ごとにランダムな順番でタスクを実施した。タスクごとの感情の変化を、感情に関する主観評価アンケートの記入により取得する。取得対象感情は「怒り」の1項目とし、取得対象の感情を推察できないよう別の6項目も追加した。各感情項目に対し、その感情の該当度を7段階で示してもらい、その度合いごとの被験者の様子の違いを観察した。

3.2 行動特徴の抽出

各タスク中の被験者の様子を撮影した動画について、アンケートで取得した怒りの該当度ごとに、(A) イライラ度が低いタスク（怒りの該当度1, 2）、(B) イライラ度が中程度でのタスク（怒りの該当度3, 4）、(C) イライラ度が高いタスク（怒りの該当度が5, 6, 7）の3グループに分けて観察した結果、(C)群のイライラ度が高いタスクでの被験者は、何度も姿勢を入れ替える、体をゆらゆらと動かす等、じっとせず動きが多い傾向が見られた。そこで、待機中の姿勢変化回数とイライラとの関係性について分析するため、タスク(3)、(4)において各被験者が待機中に姿勢を入れ替えた回数を目視で計測した。ただし、体をゆらゆらと動かす等で継続的に姿勢が変化する場合は、一連の動作で1回とカウントした。測定した姿勢変化回数について、イライラ度に応じた(A)～(C)のグループごとに平均値と標準偏差を求めた結果を図1に示す。図1のグラフより、待機中に感じたイライラ度が高かったタスクほど、被験者の姿勢変化回数の値が大きくなっており、待機に対するイライラ度と姿勢の変化回数に相関があると推察される。

この結果を踏まえ、立位状態での待機時のイライラ度をユーザの姿勢変化から求められるかを検討するため、ユーザの姿勢変化を検出するソフトウェアを開発した。

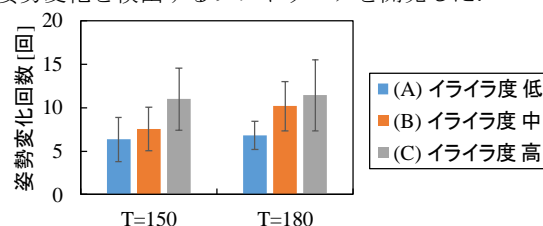


図1 姿勢変化回数の観察実験結果

4. 状況理解に向けた姿勢変化検出 SW の開発

4.1 姿勢検出ソフトウェアの概要

姿勢変化検出ソフトウェアの概要を示す。まず、機器待機中のユーザを撮影した動画データから、骨格抽出ライブラリ OpenPose[3]を用いてユーザの骨格特徴点 18 点を検出し、骨格座標データを得る。この得られた骨格データはノイズや外れ値が含まれることや、人物とカメラの距離によりスケールが異なることから、外れ値処理や平滑化、値の正規化といったデータの前処理を行う。次に、一定時間内における各骨格座標点の移動量を算出し、図 2 に示すようにその移動量が事前に設定した閾値以上となると、動画中の人物の姿勢が変化していると判定する。

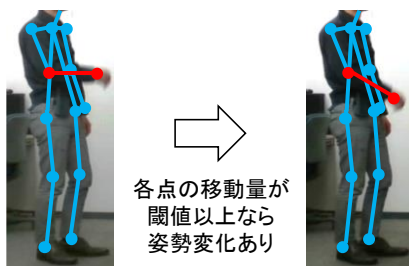


図 2 人物の姿勢変化判定

4.2 骨格データの前処理

動画から得られた骨格データは、オクルージョンによるデータ欠損やノイズが存在するため、これらの補正を行った上で姿勢変化検出処理を行う。また、動画内の人物が動いているか否かを各骨格座標点の時間変化から求める際、動画内での人物の大きさはカメラからの距離や角度によって異なる。そこで、複数の動画データ内の人物のスケールを揃えるため、正規化を行う。

4.3 閾値設定

姿勢が変化すると判定するための骨格座標点の移動量の閾値を設定する。全被験者のタスク中の様子を観察した結果、機器応答待機時における動作や姿勢は約 0.4 秒以上継続していたため、0.4 秒間の骨格座標点の移動量に着目することとする。体を動かしている場合の移動量を知るため、今回の実験で待機時にゆっくりと身体を動かしていた被験者のデータについて、静止中と姿勢変化中それぞれの 0.4 秒間の骨格 18 点の正規化した移動量を求めた。その移動量と、実際に動画データ中の人が動いているかどうかを照らし合わせることで、骨格座標点ごとに移動量の閾値を求めた。いずれかの骨格座標点の移動量が閾値を超えれば「姿勢が変化している」、全ての骨格座標点の移動量が閾値以下であれば「静止している」と判定する。この判定結果を用いて、タスク時の機器応答待機中に姿勢が変化していた時間の割合を求め、姿勢変化検出ソフトウェアの評価と、イライラ度との関連性の検証を行った。

5. 姿勢変化検出精度の評価と検証

5.1 検出精度の評価

4 章で作成した姿勢変化検出ソフトウェアを評価するため、精度検証を行った。

閾値設定に用いていない被験者の動画のうちの 3 つのデータを目視で確認し、体を動かしている時刻の正解データ

を与えた。この結果と、姿勢変化検出ソフトウェアで求めた結果を比較することで精度検証した結果を表 1 に示す。表 1 より、姿勢変化検出の正解率は 90%以上であり、精度よく人の姿勢変化を検出可能なことを確認した。そこで、この姿勢変化検出ソフトウェアを用いて、待機中の人の姿勢変化とイライラとの関連性を検証する。

表 1 姿勢変化検出ソフトウェアの精度検証結果

| No | データ数 | 正解率 | 再現率 | 適合率 | F 値 |
|----|------|------|------|------|------|
| 1 | 1250 | 94.8 | 80.4 | 92.6 | 86.1 |
| 2 | 3782 | 90.9 | 96.6 | 91.3 | 93.9 |
| 3 | 4016 | 99.2 | 94.0 | 85.9 | 89.7 |

5.2 姿勢変化とイライラとの関連性の検証

4 章で作成したソフトウェアを用いて、3 章で行った観察実験において各タスク中の昇降機待機時間のうち被験者の姿勢が変化している時間の割合 (=R とする) を求めた。全被験者が行ったタスクのうち、データ分析が可能であったのべ 62 回のタスクを、3.2 節と同様の方法で、(A) イライラ度が低いタスク (22 回)、(B) イライラ度が中程度のタスク (23 回)、(C) イライラ度が高いタスク (17 回) の 3 グループに分け、姿勢変化時間の割合 R を求めた。この結果を図 3 に示す。図 3 より、 $R \geq 0.4$ のタスク (黄色で示した棒グラフ) 7 回のうち、5 回 (71%) のタスクでイライラ度が高く、 $R < 0.05$ のタスク (青色で示した棒グラフ) 7 回のうち、6 回 (85%) のタスクでイライラ度が低かった。姿勢変化が非常に多いまたは非常に少ない場合には、姿勢変化とイライラ度に強い相関が見られた。これらの結果より、イライラ度が非常に高い、または非常に低いユーザに対しては、姿勢変化検出ソフトウェアを用いてイライラ度を推定可能であることが示唆された。

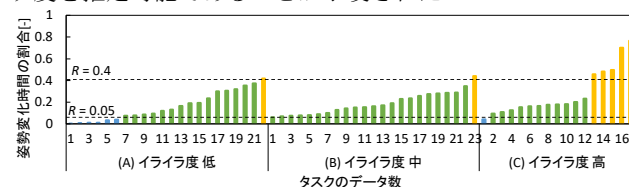


図 3 イライラ度ごとの姿勢変化時間の割合の検証結果

6. おわりに

本稿では、立位中の機器応答待機時のユーザの状況を映像情報から推定する技術開発として、イライラ度と相関のあったユーザの姿勢変化を映像情報から取得することで、待機中のユーザのイライラ度を推定するシステムを開発した。今後は、具体的な姿勢や行動の特徴に着目し、イライラ度との関連性の探索を行っていく。

文献

- [1] James A. Russell. A Circumplex Model of Affect. *Journal of Personality and Social Psychology*. 1980, vol. 39, no. 6, p. 1161-1178
- [2] 川村剛, 松澤直之, 政倉裕子, 大野澄雄, “音声に含まれる感情量の定量化のための検討”, FIT2011 (第 10 回情報科学技術フォーラム), pp.227-228 (第 2 分冊), 2011
- [3] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaser Sheikh, “OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields”, CVPR 2017.