

人間情報学会

Academy of Human Informatics



第34回 人間情報学会講演集

2019年12月11日

お茶の水女子大学 国際交流プラザ2階 多目的ホール

人間情報学会 オーラルセッションプログラム

◆日時:令和元年 12 月 11 日(水) 11:30-13:30

◆場所: 東京都文京区大塚 2-20-1

お茶の水女子大学 国際交流プラザ 2 階 多目的ホール

発表 7 分、質疑 3 分、切り替え 1 分

[11:30 - 12:18] オーラルセッション 1 (11 分×5 件)

1. BCI における電極間相関性を考慮した EEG 解析手法
田邊俊、田谷昭仁、戸辺義人
2. 案内ロボットの回避行動が利用者を与える心理的影響の調査
前原啓太、藤波香織
3. Evaluation of fragrance effects in emotional classification using biological signals
Runqing Zhang, Chen Feng, Peeraya Sripiyan, and Midori Sugaya
4. 事前・事後状態を考慮したヒューマンエラー予測手法の検討
吉田怜司、菅谷みどり
5. 遠隔者の多人数会話への参加を支援するテレプレゼンス仲介ロボット
堀江拓実、山口留実、桑原多瑛、古谷優樹、伊賀理心、飯森優斗、勢子雄大、高汐一紀

[12:25 - 12:35] 休憩

[12:35 - 13:19]オーラルセッション 2 (11 分×4 件)

6. OpenPose を利用した人の位置やベッド装置上での状態検知及び可視化システム
宇野健一、増田翔馬、岩井将行
7. 加速度センサを用いた包丁技術の判定
小林花菜乃、加藤岳大、横窪杏奈、Guillaume Lopez
8. 曲げセンサを用いた曲率推定による腹囲形状計測
大石征裕、藤波香織
9. 深層学習と単眼カメラを用いた骨格による人物識別手法の構築
浅見直人、高橋 正裕、橋本慧志、Alessandro Moro、池勇勲、梅田和昇

[13:20 - 13:30]企業セッション (10 分×1件)

10. サウナが生体情報に与える影響に関する研究と実用化について
ココヨ株式会社

BCIにおける電極間相関性を考慮した EEG 解析手法

田邊俊¹ 田谷昭仁² 戸辺義人²

¹ 青山学院大学 理工学研究科 理工学専攻 知能情報コース 〒252-5258 神奈川県相模原市中央区淵野辺 5 丁目 1 0-1

² 青山学院大学 理工学部 情報テクノロジー学科 〒252-5258 神奈川県相模原市中央区淵野辺 5 丁目 1 0-1

概要

EEG (Electroencephalogram) 信号から人間の意図する動作や指示を正確に特定するために、畳み込みニューラルネットワークを使用したシステムが増加している。これまで、EEG の電極配置に基づいて作成された行列や画像が入力データとされていたが、この手法が最適な入力である確証がなかった。本研究では、相互相関係数を用いて電極間の相関を算出し、相関が高い電極同士が隣接している場合に値が大きくなる指標を定義した。運動を想像させるタスクを分類する実験を行ったところ、本実験に関しては、定義した指標の値が大きくなるにつれて分類精度が向上した。

キーワード： BCI, EEG, 脳波

1 背景と目的

BCI (Brain Computer Interface) とは、人間の脳から得た情報を用いてコンピュータとの通信や制御を可能にするシステムである。EEG 信号から人間の意図する動作や指示を正確に特定するために、近年はニューラルネットワークを使用したシステムが増加してきている[1]。従来の研究では、EEG の電極配置に基づいて作成された行列や画像が入力データとされていたが、この手法が最適な入力である確証がなかった[2][3]。

そこで本研究では、従来の EEG ニューラルネットワークにおける入力行列を電極間の相関に基づいて再定義し、精度を評価した。

2 提案手法

・概要

本研究では、公開されている PhysioNet EEG データセットを用いて、被験者が想像した動きを分類するニューラルネットワークを構築し、従来の研究で使用されていた EEG の電極配置に基づいて作成された入力行列に代わる入力行列を提案する。

・PhysioNet EEG データセット

PhysioNet EEG データセットは、被験者が手を動かすといった運動や、運動の想像などのさまざまなタスクを実行している間に記録された脳波データをまとめたものである。このデータセットの被験者は 109 人おり、1 分から 2 分のタスクを計 1500 以上行い、被験者ごとに脳波データがまとめられ、公開されている。

本研究ではデータセットの中から、両足、両手、左手、右手を動かすことを想像する 4 種類のタスクに加え、目を閉じた状態を維持するタスクを含めた 5 種類のタスクを抽出し、脳波データから実行したタスクを推定する分類器を実装する。脳波計の電極数は 64 チャンネル、サンプリングレートは 160Hz である。図 1 に示すとおり、抽出された脳波データはウィンドウサイズで分割され、一つの学習データとする。

・ニューラルネットワーク

従来の研究で使用されていた、CNN (畳み込みニューラルネットワーク) と RNN (リカレントニューラルネットワーク) を用いてネットワークを構成する。ネットワーク構成を図 2 に示す。このネットワークは、64 チャンネルの時系列学習データから変換された時系列行

列を入力とし、CNN によって行列の特徴量を抽出した後、時系列における特徴を RNN によって抽出する。

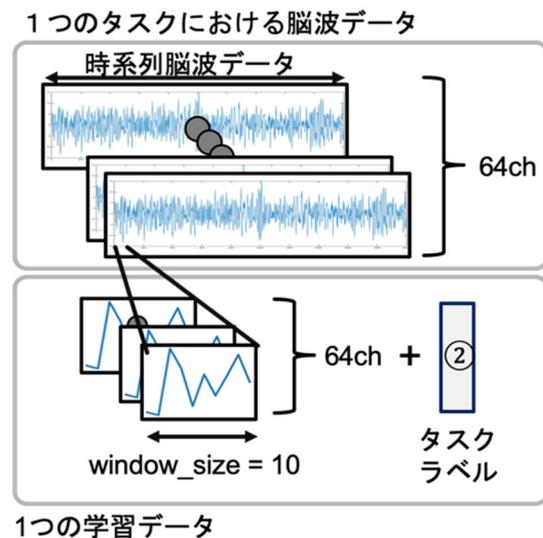


図1 ニューラルネットワークの構成

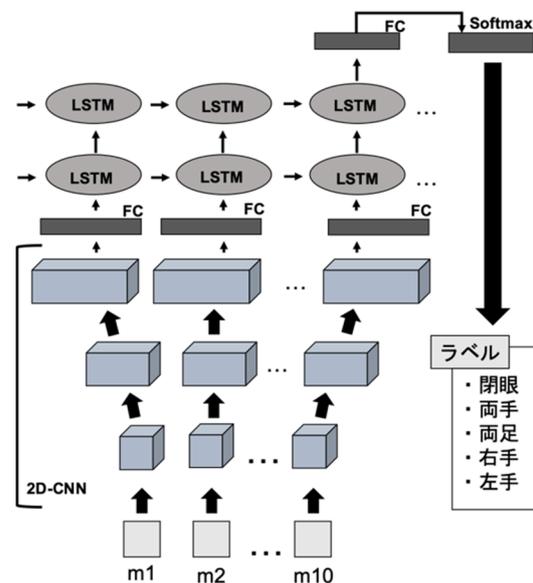


図2 ニューラルネットワークの構成

・電極配置を用いた行列

従来の研究は、脳波計の電極配置に基づいて、CNNの入力である時系列の入力行列を作成した。電極配置を2次元の行列に投影した際に、電極が配置されなかった要素に関しては、ゼロパディングを行う。

・電極間相関の算出

本研究では電極間の相関値を用いて、入力行列 E を提案する。電極間の相関値を算出するために相互相関係数を使用し、電極間の相関行列を生成する。相互相関係数の式は以下に示す。

$$C_{i,j} = \frac{1}{w} \sum_{t=0}^{w-1} x_i[t] \times x_j[t] \quad (1)$$

電極数が64であることから、 C は 64×64 の相関行列である。 x_i, x_j はそれぞれチャンネル i , チャンネル j の信号、 w はウィンドウサイズを示す。本研究では、64チャンネルの学習データを標準化したデータを用いて、相関行列を生成した。

・入力行列の評価

入力行列 E を定義するための指標として、相関が高い電極同士が隣接している場合に、値が大きくなる相関指標 s を以下に定義する。

$$s = \sum_{i=0}^{R-1} \sum_{j=0}^{C-1} \text{score}(i,j) \quad (2)$$

$$\text{score}(i,j) = \frac{1}{a_{i,j}} \sum_{r=-1}^1 \sum_{c=-1}^1 \text{correlation}(E_{i,j}, E_{i+r,j+c}) \quad (3)$$

$$\text{correlation}(e_1, e_2) = C_{e_1, e_2} \quad (4)$$

これらの式は、入力行列 E の各要素に対して、近接する要素との相関値の平均を算出し、総和したものである。 R と C はそれぞれ入力行列 E の行の長さ、列の長さを示し、 $a_{i,j}$ は入力行列の要素 $E_{i,j}$ に隣接する要素の個数である。

この相関指標 s の値が大きくなるような行列を探索し、新しい入力行列として提案する。

・行列の定義

定義した相関指標 s に基づいて定義した、3種類の行列選択方法を以下に示す。

(i) 無作為に抽出

相関指標 s に関わらず、 8×8 の行列に対して重複がないように無作為に電極を配置する。

(ii) 貪欲法

初期配置として、無作為に選択した電極4つを 8×8 の行列の中央に 2×2 の行列として配置する。その後、相関指標 s が最大となる電極と配置場所を一つずつ選択する。 8×8 の行列のすべて要素に電極が配置されるまで選択を続ける。

(iii) 局所探索法

初期配置として 8×8 の行列のすべての要素に対して、重複がないように無作為に電極を配置する。その後、 3×3 のウィンドウを用意し、ウィンドウ内の電極に対して相関指標 s が最大になるような解を求め、現

在の電極配置を上回る解が見つかった場合、行列の要素の更新を行う。解が更新できなくなるまで、ウィンドウを行列の一要素分移動し、解の探索・更新を繰り返す。

3 実験結果

本研究で定義した行列を用いて実際に学習、タスク推定を行った。学習データの90%を訓練データとし、残り10%を検証データとして精度を算出した。各行列の相関指標 s の値と、各行列を使用した学習における分類精度とを表1に示す。

表1 各行列における相関指標と分類精度

| | 無作為 | 貪欲法 | 局所探索法 |
|----------|-------|-------|-------|
| 相関指標 s | 34.29 | 47.87 | 39.69 |
| 分類精度 | 84.2% | 92.1% | 89.6% |

貪欲法を用いた行列が相関指標 s が高い行列となり、分類精度も高いものとなった。3種類の行列を比較すると、相関指標 s の値が大きくなるにつれて精度が向上した。

4 結論と今後の課題

本研究では、EEGニューラルネットワークにおける入力行列を電極間の相関に基づいて再定義し、従来の手法との精度を比較した。結果として、本実験に関しては相関指標 s の値が大きくなるにつれて精度が向上した。

本実験では、CNNの入力行列を 8×8 の正方行列に限定したが、ゼロパディングを使用することで 8×8 の正方行列に限らず様々な行列を生成することができる。今後は、入力行列の大きさを限定しなかった場合における電極間相関の高い行列について検討する必要がある。

また、信号の相関指標として相互相関係数を使用したが、信号の類似性を求める指標として挙げられるDTW(動的時間伸縮法)についても検討する必要がある。

文 献

- [1] Xiang Zhang, Lina Yao, Dalin Zhang, Xianzhi Wang, Quan Z. Sheng, Tao Gu "Multi-person brain activity recognition via comprehensive eeg signal analysis" In 14th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services, 2017.
- [2] Dalin Zhang, Lina Yao, Xiang Zhang, Sen Wang, Weitong Chen, Robert Boots, "Cascade and Parallel Convolutional Recurrent Neural Networks on EEG-Based Intention Recognition for Brain Computer Interface" The Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
- [3] Pouya Bashivan, Irina Rish, Mohammed Yeasin, Noel Codella, "Learning Representations from EEG with Deep Recurrent-Convolutional Neural Networks" International Conference on Learning Representations vol. abs/1511.06448, 2016

案内ロボットの回避行動が 利用者に与える心理的影響の調査

前原啓太¹ 藤波香織²

¹ 東京農工大学大学院 工学府 情報工学専攻 〒184-8588 東京都小金井市中町 2-24-16

² 東京農工大学大学院 工学研究院 先端情報科学部門 〒184-8588 東京都小金井市中町 2-24-16

概要

自律移動ロボットが人間の挙動を認識し誘導を行うことは、日常生活にロボットが浸透するために重要である。このような誘導ロボットの研究は数多く行われているが、誘導される人間の感情的作用に関しては明らかになっていない。本論文の目的は、誘導ロボットと利用者間の適切な距離感を明らかにし、障害物や通行人などが存在する状況において、ロボットの回避行動が誘導ロボットの利用者に与える心理的影響や距離感の変化について説明することである。

キーワード： 誘導ロボット 自律移動 誘導制御 人物検出 因子分析 ROS SLAM

1. はじめに

近年、人間の仕事や日常生活の支援を行うことを目的とした自律移動型誘導ロボットの開発が盛んに行われている。このような誘導ロボットに関する研究では、実地調査の報告[1]や独自アルゴリズムを評価する研究[2]が多く、誘導ロボット自体に焦点が置かれている。そのため、誘導される人間の感情的作用に関しては明らかになっていない。しかし、人同士のコミュニケーションにパーソナルスペースが存在するように、人とロボットのインタラクションに関しても快適に対話するための距離感が存在することが分かっている[3]。今後、人間の社会環境でロボットが共存するためには、ロボットと人との間に適切な身体的・心理的距離を保ち、障害物や通行人などの状況に応じて適切な回避行動を行う必要があると考えられる。

そこで本研究では、博物館や空港などの屋内公共空間における自律移動ロボットが人々を案内・誘導する状況を想定し、ロボットとの距離感や回避行動が利用者に及ぼす心理的影響を説明することを目的とする。

2. 提案システム

2.1 概要

本研究の目的を達成するために、公共空間に存在する障害物や通行人を回避し、利用者を目的地へと案内する自律移動型ロボットシステムを開発した。システムの概要を Fig.1 に示し、開発した誘導ロボットシステムを Fig.2 に示す。移動ロボットには YujinRobot 社の Turtlebot2 を使用し、接続したノート PC では ROS(Robot Operating System)と呼ばれるミドルウェアを利用してロボットを制御する。

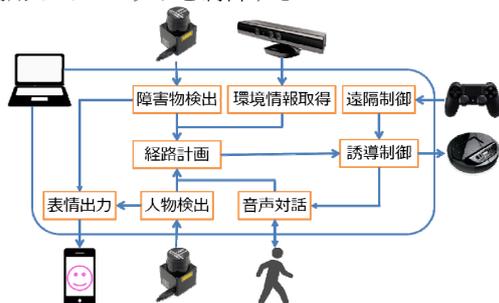


図.1 システム概要

2.2 人物誘導

自律誘導は環境地図に基づいて行われる。そこで移動ロボットシステムを遠隔操作し、測域センサの距離情報から生成した大学構内の廊下における環境地図を Fig.3 に示す。また目標地点を設定することで、現在位置からゴールまでの最短距離を高速で自動的に検出し、随時経路を修正しながらナビゲーションを行う。

利用者の追跡手法については、位置が連続的に移動することから、ある地点から数ミリ秒後のユーザが存在する範囲を制限することで、利用者を壁や障害物と区別してトラッキングすることが可能となる。

自律移動ロボットの通過行動がユーザに与える心理的影響に焦点を当てた研究では、ユーザの 91%がロボットの速度が可変であることを望んでいることが明らかになった[4]。これは、ユーザがロボットに慣れてくると歩行速度よりも遅いロボットに飽きてしまうことが原因である。そこで本システムでは、ユーザの歩行速度に合わせて移動速度を変え、一定の距離を保ちながら誘導することとする。



図.2 誘導ロボット

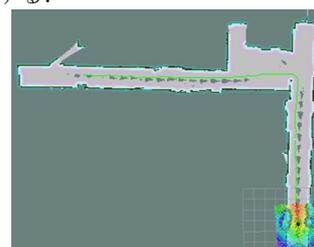


図.3 環境地図

2.3 ユーザとのインタラクション

音声合成システムを用いて音声発話を行う。これは誘導開始時、目的地の決定時や到着時にユーザに通知する目的で使用される。

また、表情変化を行うロボットはそうでないロボットと比較してより親密な印象を利用者に与えることが分かっている[5]。そこで、感情伝達の手段としてロボットシステムに設置したスマートフォンに顔画像を表示させる。世界的に共通の感情とされている6つの基本感情の内、負の感情である「嫌悪」「怒り」を除いた「喜び」「悲しみ」「驚き」「恐れ」を誘導ロボットの行動に対応付けて表示する。

3. 誘導ロボットの印象調査

開発した誘導ロボットを主観的に評価するため、ロボットの印象を表す形容詞群を用いたSD法による7段階評価のアンケート調査を行った。その後、得られた結果を用いて因子分析を行い、形容詞群をまとまりのあるグループに分類した。更にグループに共通する要素である「親和性」「利便性」「信頼性」という3つの代表因子を決定した。

4. 予備実験

誘導ロボットに案内される利用者が最も適切だと感じる距離を明らかにする。距離パラメータは0.5, 1.0, 1.5, 2.0[m]の4通りで、20代の被験者10名を対象として行った。利用者は7段階評価のアンケート調査を行い、20[m]の廊下を直進しながら誘導された際のロボットの印象を評価する。アンケート結果をTab.1に示す。実験結果より、距離1.0[m]での誘導が最も高い評価となったことから、評価実験では誘導時のロボットと利用者の基準距離を1.0[m]に決定する。

表.1 予備実験結果

| 距離[m] | 評価因子 | 平均値 | 標準偏差 | 合計 |
|-------|------|-----|------|------|
| 0.5 | 親和性 | 4.6 | 1.1 | 12.7 |
| | 利便性 | 4.3 | 1.4 | |
| | 信頼性 | 3.9 | 1.0 | |
| 1.0 | 親和性 | 5.3 | 0.8 | 15.3 |
| | 利便性 | 5.3 | 1.1 | |
| | 信頼性 | 4.8 | 0.6 | |
| 1.5 | 親和性 | 5.1 | 0.8 | 15.0 |
| | 利便性 | 5.0 | 1.2 | |
| | 信頼性 | 5.0 | 0.5 | |
| 2.0 | 親和性 | 4.5 | 1.2 | 14.1 |
| | 利便性 | 4.7 | 1.2 | |
| | 信頼性 | 5.0 | 0.6 | |

5. 評価実験

5.1 目的・方法

廊下を誘導する状況において、障害物となる人や歩行者を回避する行動が利用者に与える心理的影響を明らかにする。実験環境は予備実験と同じく20[m]の直線となっている廊下で行う。被験者は20代の男女12人を対象とする。まず初めに障害物となる静止した人を回避する実験では左回避と右回避の2通りで実験を行う。その後、歩行者を回避する実験は、左回避・右回避・待機の3種類の方法で行う。

5.2 実験結果

評価実験のアンケート結果をFig.4に示す。実験結果から、障害物と歩行人のどちらの状況においても左回避よりも右回避が評価が高くなることが分かった。それぞれ左右での回避方法に有意に違いがあるのか調べるために有意水準を5%としてt検定を行った。その結果、どちらの場合においても有意差は得られなかった。また、「左右の回避による違いを感じるか？」という質問に対してもほとんどの利用者が違いを感じなかったと回答している。更に、歩行人の回避においては、歩行人が通り過ぎるのを待つ場合、左右への回避と比べて利便性が低下する結果となった。

一方で、障害物と歩行人をそれぞれ左回避する場合と右回避する場合でt検定を行った結果、どちらの場合においても有意差が得られた。このことから、歩行人を回避するよりも障害物を回避する行動の方が利用者にとって印象が良いことが明らかになった。

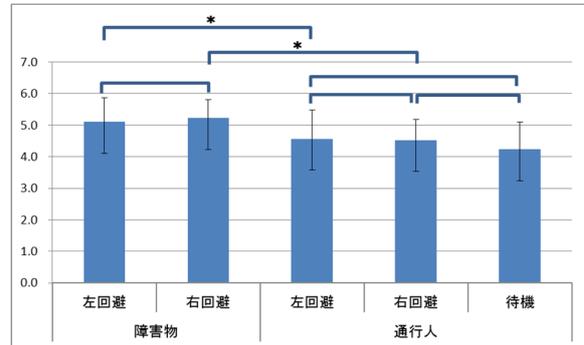


図.4 評価実験結果

5.3 考察

t検定や質問の結果から多くの利用者にとって左右の回避による誘導ロボットの印象に違いがないことが分かった。しかし一部の利用者はアンケートの自由記述に「ロボットが右側から回避するほうが自然なように感じる」という意見があったことから、どちらかといえば右回避が好ましいと感じる人がいることが明らかになった。その理由として普段道路などを歩く場合、右側を歩くことが多いため、ロボットも同様に右側から回避したほうが違和感がないからと考えられる。

6. 結論・今後の展望

本研究では、障害物や歩行人が存在する状況での誘導ロボットの回避行動が利用者に与える心理的影響について調査した。実験結果より、歩行人を回避するよりも障害物を回避する方が有意に誘導ロボットの印象が良くなることが分かった。また、障害物と歩行人どちらの条件においても左右の回避による違いは見られなかった。今後は収集した利用者との誘導ロボットの移動データを用いて利用者の行動を分析し、障害物回避時の移動ロボットとの相対位置関係の変化について調査する予定である。

文 献

- [1] Takayuki Kanda, et al. "An affective guide robot in a shopping mall". In Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, pp. 173-180, 2009.
- [2] Aniket Bera, et al. "Sociosense: Robot navigation amongst pedestrians with social and psychological constraints". In 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 7018-7025, 2017.
- [3] Koji Nakashima, et al. "Personal distance against mobile robot". The Japanese Journal of Ergonomics, Vol.35(2), pp. 87-95, 1999.
- [4] John Travis Butler, et al. "Psychological effects of behavior patterns of a mobile personal robot". Autonomous Robots, Vol.10(2), pp. 185-202, 2001.
- [5] 谷寄悠平, 他. "教育支援ロボットにおける身体動作と表情変化による共感表出法の印象効果". 知能と情報, Vol.30(5), pp. 700-708, 2018.

Evaluation of Aroma Effects in Emotional Classification using Biological Signals

Runqing Zhang¹, Chen Feng¹, Peeraya Sripian¹, Midori Sugaya¹

¹ Shibaura Institute of Technology Department of Computer Science and Engineering 〒135-8548 3-7-5 Toyosu, Koto City, Tokyo

Abstract

Aromatherapy is a natural way of healing a person's mind and body. The essential oils involved in aromatherapy are highly concentrated essences extracted from the plants through the process of the distillation. Each oil produces a predictable and reproducible effect on the user when its aroma is inhaled. There have been lots of researches that confirm the effects of inhaling aroma by using biological signals. However, most of these researches, only the heart rate or the pulse signals are used. And, EEG signal is rarely used in evaluating the aroma effects. So in order to increase the validity of the reputed effects of the previous research, we consider to used the combination of the EEG signal and the pulse signal to evaluate the effect of the aroma. This work is the preliminary experiment to evaluate the changes in this combination of biological signals after inhaling two aroma.

Keywords: EEG heart rate aroma

1. INTRODUCTION

Aromatherapy derived its name from the word aroma, which means fragrance or smell and therapy, or treatment. This therapy method is a natural way of healing a person's mind and body. Inhalation, local application, and baths are the major methods [1]. Today the popularity of aromas for pleasure, relaxation, and therapeutics is unabated and typified in the ever-popular application of aromatherapy.

In aromatherapy, the essential oils can produce a predictable and reproducible effect on the user when its aroma is inhaled. There have been lots of researches confirm that the different effect of inhaling each oil's aroma. In those works, the effects of the aroma were validated mainly by comparing the biological signals taken from the participant's before and after the aroma was inhaled. Furthermore, in those works, they also used a subjective questionnaire about how the participant is feeling to regard psychological changes. However, there are issues that the participants may not genuinely answer how they are exactly feeling, because they don't know their true feeling.

In order to objectively detect the genuine emotion, biological signals can be used. In these days, there are approaches that detect the emotion with the biological signals, such as brainwave signals and pulse signals [2]. These approaches provide the method of objectively detect the exact emotion of the people. So in order to increase the validity of the reputed effects of the previous research, we want to combine these two methods to evaluate the effects of aroma.

2. METHOD

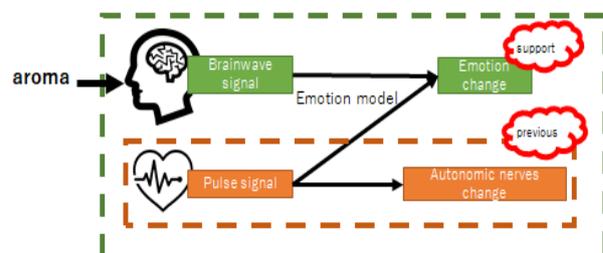


Fig. 1 The system of evaluation of aroma

Based on previous research, we increased the measurement of brainwave signals, as shown in Fig.1. And we combined brainwave signals and pulse signals with Russell's

Circumplex Model. In this way, we wish to find out how participants' emotion would change after inhaling the essential oil aroma.

In the experiment, participants are required to clear their hands, wear an EEG and a pulse sensor, and answer a questionnaire about personal information before starting the experiment. The EEG signals are obtained from Neurosky's mindwave, and the pulse signals are obtained by wearing pulse sensor at the participants' fingertips. After one minute, the olfactory of the participant is stimulated by a test strip with essential oil aroma for one minute. Peppermint oil and lavender oil from Ryohin Keikaku Co. Ltd. is used as the olfactory stimulation. Then, the participants are required to answer another questionnaire about the aroma. Between the two aroma, the olfactory of the participant can be reset by smelling the skin at wrist [3].

3. RESULT

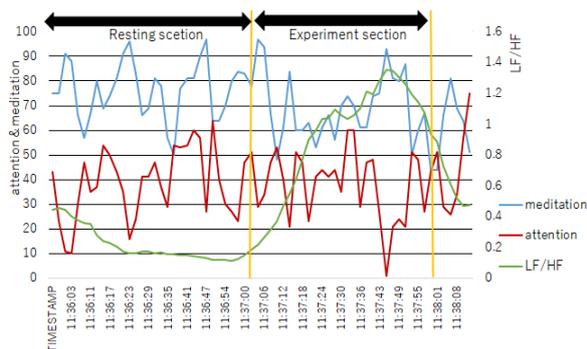


Fig. 2 Proposed methods for the combination in the evaluation of the effect of the lavender aroma

Figure 2 shows the value of meditation, attention, and LF/HF analyzed from the experiment when the participant inhaling the aroma of lavender. Attention value is derived from Beta waves, meditation value is derived from Alpha waves, while the LF/HF value is derived from LF and HF, which provided by the pulse sensor [2]. From Figure 2, there were some changes in the EEG signal and the pulse signal. The peak of the meditation and LF/HF reached the highest point at the same time, and the attention arrived at the lowest point. From these results, we can imply that the participant became more relaxed after inhaling the lavender aroma. This indicated that the lavender aroma has a relatively obvious

effect on the participant's status. We have confirmed the result of Lavender aroma oil [4] by analyzing biological signals in this work.

The lavender is widely regarded as possessing sedative and calmative properties, and is commonly found in products aimed at aiding relaxation. Such a proposition was supported by the increase in "Meditation" and the decrease in "Attention" reported here by participants in the lavender aroma condition compared to the peppermint aroma conditions. A significant difference in the "Attention" was isolated between the two aroma conditions, with a small increase in peppermint aroma and a decrease in lavender aroma.

4. DISCUSSION

The results of this work clearly support the contention that the aromas of essential oils can change the brainwave signals and the pulse signals at the same time in healthy participants. Furthermore, we want to find out if the participant's preferences affect the effect of essential oil aroma by analyzing the biological signals.

REFERENCES

- [1] Ali, B., Al-Wabel, N. A., Shams, S., Ahamad, A., Khan, S. A., & Anwar, F. (2015). Essential oils used in aromatherapy: A systemic review. *Asian Pacific Journal of Tropical Biomedicine*, 5(8), 601-611.
- [2] Ikeda, Y., & Sugaya, M. (2016, July). Estimate emotion method to use biological, symbolic information preliminary experiment. In *International Conference on Augmented Cognition* (pp. 332-340). Springer, Cham.
- [3] Groszofsky, A., Hauptert, M. L., & Versteeg, S. W. (2011). An exploratory investigation of coffee and lemon scents and odor identification. *Perceptual and motor skills*, 112(2), 536-538.
- [4] Moss, Mark, et al. "Modulation of cognitive performance and mood by aromas of peppermint and ylang-ylang." *International Journal of Neuroscience* 118.1 (2008): 59-77.

事前・事後状態を考慮したヒューマンエラー予測手法の検討

吉田 怜司¹ 菅谷 みどり²

¹ 芝浦工業大学 〒135-8548 東京都江東区豊洲 3-7-5 ma18108@shibaura-it.ac.jp

² 芝浦工業大学 〒135-8548 東京都江東区豊洲 3-7-5 doly@shibaura-it.ac.jp

概要

背景と目的：ヒューマンエラーの予防策・対処策は数多く研究されており、人の内部状態が原因であるとして脳波や心拍をモニタする研究がなされている。本研究ではこの2指標をも用いた分析を行い、エラー発生時のモデル化を目指す。

方法：負荷作業としてストループ課題を実施し、脳波・心拍・反応時間を計測しエラー発生時に特有の傾向がないか調査した。本課題は動画として提示し、発話による回答を指示した。未回答および誤回答をヒューマンエラーとした。

結果：1番目の未回答時には心拍変動指標の急激な低下が確認できたが、以降は未回答・誤回答を問わず、有意な特徴がみられなかった。反応時間に関しては、全体を通して常に変動していたため、エラーに関する有意な特徴はみられなかった。

考察：安静時・課題実施中の心拍変動指標はパターン1・パターン2の両方において同様の傾向が伺えた。初めのエラーの特徴は課題特有のものかエラーのものかは他の指標と合わせて確認する必要がある。

結論：負荷作業時の脳波・心拍・反応時間を計測しエラー特有の傾向を調査した。心拍に関しては一定の傾向が確認できた。脳波の分析と合わせて心拍の周波数分析を行い、エラー発生時の事前・事後を区別して機械学習することでモデル作成を目指す。

キーワード：ヒューマンエラー 生体情報

1 はじめに

近年、会計上のミスや AI によって自動で提示する会計ソフト[1]や、プログラムのミスや AI によって指摘する AI[2]など、ヒューマンエラーを AI によって指摘しその検知と修正にかかる時間を削減する取り組みがなされている。また、Robot Process Automation(RPA)[3]による業務の自動化・効率化などが進められている。このように、人が従事する作業は将来的により複雑度が高く高度なものとなることが予測される。そうした作業におけるヒューマンエラーは致命的となるため未然に防ぐことが求められる。

Fedota らは、エラーはそもそも複雑な現象であるとし、多くの異なる要因が積み重なり、人の事前の状態やエラー後の反応も関係すると述べている[4]。Lin らは、汎用的にこれらの人の行動に影響を与える要因を理解するためには人の認知状態に関わるエビデンスが必要であるとし、生体情報を用いたエラー分析を行った[6]。具体的にはエラー時の行動分析のモデリングと脳波測定を双方を学習させることで、エラー発生を予測できることを示した[5]。Lin らは脳波のみを人の内部の状態を示す情報として用いたが、和田らによると、心拍・精神性発汗といった生体情報も運転操作などの記録・分析と相関が得られると報告されている[6]。さらに、西川らによると、外部環境である温度設定を変化させた実験では不快な室温であるほど不快申告が多く、ヒューマンエラーが頻発し、快適な室温であるほど不快申告が少なくヒューマンエラーの頻度が下がったことが報告されている[7]。和田や西川らは、人に影響を与える外部環境とエラー発生時の人の内部状態データとして脳波以外の情報も有用であることを示した一方、Lin らのように人の行動モデルについては検討していない[5][6][7]。

2 負荷作業下におけるヒューマンエラー

Lin、和田、西川らの研究からヒューマンエラーを予測する上で脳波などの人の内部状態を観察することは有用であることが伺える。そこで、本研究では人の生体情報として、Lin らが行った脳波計測に加え、和田らによる心拍計測を行い、エラーの前後で発生する生体情報のパターンをデータからモデル化し、それによりエラーの予測が可能であるか検討した。これに当たり、本研究では負荷作業下のヒューマンエラー発生時の脳波と心拍を評価した。

3 内部状態としての生体情報

ヒューマンエラー発生時の人の内部状態を観察するため脳波計測から得られる各脳波指標を用いて負荷作業下における人の脳の活動状態を取得した。特に、日沖らによると、実施した作業の負荷に応じて作業実施者の脳活動が変化することが確認されている[8]。したがって、人の内部状態として脳波を観測することは妥当であると考えられる。また、自律神経はストレス指標として広く用いられている。本研究では心拍計測によりこの自律神経の活動を観測した。和田らによると、エラー発生時には心拍間隔の低下が確認されている[6]。したがって、脳波と同様に人の内部状態として心拍を計測することは妥当であると考えられる。

本研究では Emotiv 社の Epoc+[9]を用いて脳波計測を行った。Alarcao らの調査によると、本脳波計は脳波測定機器として多くの研究に用いられている[10]ため、実績のある脳波計といえる。データ計測としてサンプリング周波数は 128Hz に設定し、拡張 10-20 法に基づく 14 チャンネル (AF3、AF4、F3、F4、F7、F8、FC5、FC6、T7、T8、P7、P8、O1、O2) を使用した。

ストレス指標として心拍間隔 (RRI: R-R Interval) の差から得られる pNNx を用いた。pNNx は隣接する心拍間隔の差を取得し、30 個の差のうち x ミリ秒以上の

ものの割合であり、スペクトル解析によらない心拍変動性指標である。pNNx は主に自律神経系の副交感神経の活動指標であり、心拍間隔の揺らぎが小さいほど (pNN50 の値が小さいほど) ストレスを感じているとされ、一般的にストレス指標として用いられる。ユニオンツール社の myBeat[11]を用いて心拍間隔を測定した。森本らによると、認知負荷作業において刺激に対する反応時間は作業負荷によらず有意な差が確認できなかった[12]と述べている。一方、反応時間に応じたエラー率の変化については述べていない。本研究では2パターンの負荷作業に対する反応時間とエラー率について調査する。

4 負荷作業実験

本実験では実験協力者が負荷作業を行なっている間に脳波と心拍の計測を行い、作業中のヒューマンエラー前後の心拍と脳波の変化の特徴を把握することを目的とし、記録・検証を行なった。負荷作業としてストループ課題を実施し、未回答および誤回答をヒューマンエラーとした。本課題は動画として協力者に提示し、発話による回答を指示した。20代男性1人に対し以下の手順で実験を実施した。

- (1) 脳波計・心拍計を装着する
- (2) 1分間安静にする
- (3) 提示される漢字の読みを1秒間隔で100問答える (以降パターン1)
- (4) 1分間安静にする
- (5) 提示される漢字の色を1秒間隔で100問答える (以降パターン2)



図1 刺激提示動画

5 実験結果

実験結果は以下のとおりである。

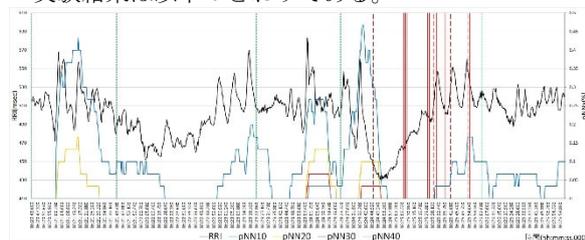


図2 作業中の心拍間隔及び pNNx

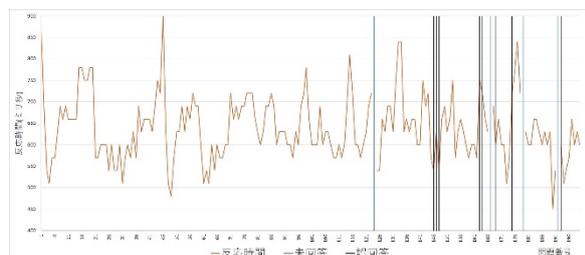


図3 作業中の刺激に対する反応時間

協力者はパターン1実施中にはエラーを起こさず、すべてのエラーはパターン2実施中に発生した。1番目のエラー発生時には RRI および pNN10 の急激な低下が確認できる。一方、以降のエラーでは未回答・誤回答を問わず、有意な特徴がみられなかった。また、安静時は RRI・pNNx とともに高い値を示した一方、作業中には一時的に下がった後、徐々に高くなるという傾向が確認できた。反応時間に関しては、安静時・作業中およびパターン1実施時・パターン2実施時において常に変動していたため、エラーに関する有意な特徴がみられなかった。

6 まとめと今後の課題

本研究では負荷作業中の生体情報を記録・分析しヒューマンエラー発生時の傾向のモデル化を目指した。脳波・心拍・反応時間を用いてエラー特有の傾向がないか調査した。今後は心拍変動指標の周波数解析や脳波の分析、エラー発生の事前・事後別々の分析を行い、有意な傾向がないか調査を続ける。

文献

- [1] freee 株式会社, <https://corp.freee.co.jp/news/freee-aigetsujikanns>, (2019).
- [2] K. Matt, <https://www.wired.co.uk/article/ubisoft-commit-assist-ai>, (2019).
- [3] 株式会社アイ・ティ・アール, <https://www.itr.co.jp/company/press/1710050102PR.html>, (2019).
- [4] J. R. Fedota and R. Parasuraman, “Neuro-ergonomics and human error,” *Theor. Issues Ergonom. Sci.*, vol. 11, pp.402–421, 2010.
- [5] C. Lin, C. Wu and W. A. Chaovalitwongse, “Integrating Human Behavior Modeling and Data Mining Techniques to Predict Human Errors in Numerical Typing,” *IEEE Transactions on Human-Machine Systems*, 45(1), pp.39-50, 2015.
- [6] 和田一成, 守屋祥明, 田所和孝, “ミスの連鎖の発生メカニズムに関する基礎的研究,” あんけん研究成果レポート, vol.5, pp.23–26, 2012.
- [7] 西川雅弥, 西原直枝, 田辺新一, “中程度の高温環境下の長時間作業が作業効率と疲労に与える影響に関する被検者実験,” *日本建築学会環境系論文集*, vol.74, no.638, pp.525–530, 2009.
- [8] 日沖求, 野澤昭雄, 水野総太, 井出英人, “時間圧迫下におけるメンタルワークロードの生理心理評価,” *IEEE Trans. EIS*, vol.127, no.7, pp.1000-1006, 2007.
- [9] Emotiv, <https://www.emotiv.com/>, (2019)
- [10] Soraia M. Alarcao, Manuel J. Fonseca, “Emotion Recognition Using EEG Signals: A Survey,” *IEEE Transaction on Affective Computing*, vol.10, no.3, pp.374-393, 2017.
- [11] ユニオンツール株式会社, <https://www.uniontool-mybeat.com/>, (2019).
- [12] 森本克彦, 久保田敏裕, 篠原一光, 木村貴彦, “運転士の指差・喚呼の実施方法に関する研究,” あんけん研究成果レポート, vol.2, pp.30–35, 2009.

遠隔者の多人数会話への参与を支援する テレプレゼンス仲介ロボット

堀江拓実*1, 山口留実*1, 桑原多瑛*2, 古谷優樹*3, 伊賀理心*3, 飯森優斗*3, 勢子雄大*3, 高汐一紀*3

*1 慶應義塾大学大学院 政策・メディア研究科

*2 慶應義塾大学 総合政策学部

*3 慶應義塾大学 環境情報学部

概要

多人数参加型のテレコミュニケーションでは、Fact-to-Face の場合と比較して遠隔者の非言語的な振る舞いが伝わりづらく、会話における発話権の取得が困難、すなわち会話に参加しにくいという問題を抱えていた。この問題に対し、テレプレゼンスロボットの身体性を強化しアバターロボット化することで、遠隔者の非言語的な振る舞いを直接的に他の参加者に伝え、発話権の取得を支援する手法が議論されている。しかし、これらの手法では実際に話者交代が起こるかは受け手の判断に依るため、不確実性が高い。本稿では、仲介者としてより積極的に遠隔者の発話権の取得を支援するエージェントソフトウェアを搭載したテレプレゼンスロボットの実現を目指し、その第一段階として、遠隔者の会話参与状況と被注視状況の関係性について分析する。

キーワード： テレプレゼンスロボット、メディアーション、非言語動作、話者交代

1. はじめに

テレビ電話システムと比較した際のテレプレゼンスロボットの利点は、遠隔者が自分で操作することで移動が可能であることだけでなく、その身体性を利用した遠隔者の会話参与支援が可能であることが挙げられる。これまで、テレプレゼンスロボットの身体性を用いて、Adalgeirsson らはテレプレゼンスロボットに移動以外のジェスチャを遠隔操作可能にすることで、会話参加者の Engagement や楽しさが向上することを示した[1]。一方で[1]のシステムのようにロボットのジェスチャをいちいち遠隔者が操作することは負担となるため、遠隔者の会話参与支援のためのロボットの挙動を自動的に生成する試みが研究されている[2,3,4]。これらのシステムは、非言語的振る舞いのビデオチャット越しでの伝わりづらさやレイテンシによって遠隔者が発話権を取得しづらい、という問題を解消することを目的としている。しかし既存研究では、テレプレゼンスロボットは遠隔者のアバターであり、遠隔者のボディランゲージをテレプレゼンスロボットが再現するなどの間接的な会話参与支援に限られている。そこで、本研究ではテレプレゼンスロボットが遠隔者の映像/音声を配信するだけでなく、独立したエージェントとしてヘルパーの役割を果たすことで、より効果的な遠隔者の会話参与支援を実現することを目指す。

2. 会話参与支援エージェント

本研究が目指すテレプレゼンスロボットは、自身が独立したキャラクタを持ち、遠隔者と共に行動をし、遠隔者の会話参与を支援する会話参与支援エージェントを持つ。例えば、遠隔者が発話する意図があるにも関わらず会話にうまく参加できていない場合、ロボットは図 1 (a) のように現地の参加者に掛け合い、遠隔者の話せるターンを作る。遠隔者側には図 1 (b) のように現地の映像音声の他に、ロボットのエージェントが表示されており、ロボット-遠隔者間のコミュニケーションによって、発話権の受け渡しが行われる。

本ロボットは会議におけるメディアータに近い機能を持つが、会話参与にディスアドバンテージを負っている遠隔者の付き添いとして会話に参加しており、公

平に場を仕切るメディアータとは立ち位置が異なる。



図 1. 本研究が目指すテレプレゼンスロボット

また、従来の会話参与支援を行うテレプレゼンスロボットとの大きな違いとしては、ロボットが現地の状況に基づいて自律的に行動するため、現地の会話の機微を捉えた行動が可能になるという点が挙げられる。

本システムの機能要件は以下のようである。

- R-1. 遠隔者の会話参与状況の認識
- R-2. 遠隔者の会話意図の認識
- R-3. 平時の会話への参与
- R-4. 現地の会話での発話権の奪取
- R-5. 遠隔者への発話権の譲渡

本論文では、R-1 の機能を実装するための予備実験について述べる。

3. 会話への参与状況の認識

3.1. 既存研究

「置いてけぼり」のユーザを会話に引き戻すという観点で多数の研究が行われており[5,6]、秋葉らと松山らは共に「発話量」に基づいて会話に「置いてけぼり」になっている参加者の検出を行なっている[5]。しかし、松山らは発話してなくても会話を楽しむタイプのユーザもおり、「置いてけぼり」の指標として「発話量」が不十分であることを指摘している。そこで、本論文では発話量の他に会話参加者からの注目時間も考慮することで、より正確に遠隔者の参与状況を認識することを目指す。

3.2. 仮説

会話の参加者は大きく話し手と聞き手に分けることができるが、Goffman は聞き手の役割をさらに細分化し「参与役割」という概念を提唱した[7]。Goffman の

区分に従えば、聞き手には、参与を承認されている(=話し手になることができる)聞き手と、承認されていない聞き手が存在している。また、話し手と承認された聞き手の間では、会話内容の理解状況を共有する責任を相互に負っているとされている[8]。したがって、話し手と承認された聞き手は、本来理解状況の確認のために相互に視線のやりとり等が行われていなければならない。一方、「置いてけぼり」にされた参加者は、理解状況の共有が相互に行われていない状態であり、そのほかの参加者からの注目が減少すると考えられる。本論文では「会話参加者の参与状況とその他の参加者からの注目時間に関連がある」を仮説とする。

4. 実験

4.1. 実験手順

実験参加者は学生4名である。各回の会話は10分とし会話は事前に参加者に提示した。実験参加者のうち1名が遠隔者として、Softbank Robotics社のPepperの胸部に取り付けられたタブレット上で動作しているgoogle hangoutを介して会話に参加し、残りの3名は会話中にロボットの前方の1m×1mの範囲内に入るように指示した。遠隔者を入れ替え、計4回実験を行なった。実験の様子を図2に示す。会話はPepperの背後に置かれたビデオカメラで記録を行なった。また、遠隔者が視聴しているgoogle hangoutの映像も記録した。実験後、記録した映像に対し、ELANを用いて発話区間および、遠隔者に注目していた区間のアノテーションを行なった。また、特に遠隔者の発話量が最も少なかった回では、遠隔者に話しづらかった箇所のアノテーションを行ってもらった。



図2. 実験環境(左)と実験の様子(右)

4.2. 結果と考察

遠隔者の発話量が最も多かった第2回と、最も少なかった第1回について、遠隔者の発話量と遠隔者が注目されていた区間のアノテーション結果を表1に示す。遠隔参加者の発話量が少ない場合には遠隔参加者が注目される時間も減少していることがわかる。また、遠隔者が発話量最大の際の現地参加者の発話量と被注目時間を表2に示す。なお、表2では遠隔者の注視方向については分析しておらず、被注目時間は2人分の注目時間の平均である。表1の第1回の遠隔者の場合と比較して、被注目時間に対して総発話量が多くなっている。したがって被注目情報だけで参与状態を推定することは困難と考えられる。図3は、第2回での遠隔者の被注目時間と遠隔者が話しづらかった箇所をタイムラインに示したものである。遠隔者が話しづらかった箇所は、いずれも遠隔者が視線を受けていないタイミングであり、視線が話し出すために重要な役割を果たしていると考えられる。

表1. 遠隔者の発話量と被注目時間

| | 総発話量 (s) | 被注目時間 (s/人) |
|-----|----------|-------------|
| 第1回 | 5.5 | 68.8 |
| 第2回 | 155.6 | 285.1 |

表2. 第2回での現地参加者の発話量と被注目時間

| 話者 ID | 総発話量 (s) | 被注目時間 (s/人) |
|-------|----------|-------------|
| A | 104.0 | 40.1 |
| C | 155.4 | 67.3 |
| D | 91.0 | 40.4 |



図3. 第1回のタイムラインとアノテーション

5. まとめ

我々は、会話参与支援エージェントを付与したテレプレゼンスロボットの実装を目指している。本論文では、ビデオチャットシステムを用いた実際の会話の事例における、遠隔参加者の参与状況と被注目時間の関係について分析した。結果、遠隔参加者が会話への参与が少ない場合に現地の参加者からの注視も減っている現象が見られ、会話の参与状況と被注目時間に関連性があることが示唆された。一方、被注目時間だけでは会話の参与状況を把握することが難しいことも示された。また、遠隔者が発話を切り出す上で視線を受けていることが重要であることが示された。

今後は、発話量および被注目時間からの遠隔参加者の会話の参与状況の推定を目指す。さらにその先の展望として、最終的な目標である独立したエージェントとしての機能を持ったテレプレゼンスロボットの実現を目指す。

文献

- [1] Adalgeirsson, Sigurdur O., and Cynthia Breazeal. "MeBot: a robotic platform for socially embodied presence." Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction. IEEE Press, pp. 15-22, 2010
- [2] Hasegawa, Komei, and Yasushi Nakauchi. "Preliminary Evaluation of a Telepresence Robot Conveying Pre-motions for Avoiding Speech Collisions." Proceedings of the Second International Conference on Human-Agent Interaction, Tsukuba, Japan, pp. 29-31, 2014.
- [3] Hasegawa, Komei, and Yasushi Nakauchi. "Facilitation of telepresence robot turn-takings by gesture exaggeration." 2014 IEEE/SICE International Symposium on System Integration. IEEE, pp. 650-654, 2014.
- [4] NT Fitter, Y Joung, M Demeter, Z Hu, MJ Mataric. "Design and Evaluation of Expressive Turn-Taking Hardware for a Telepresence Robot." 2019 IEEE International Symposium on Robot and Human Interactive Communication, 2019
- [5] 秋葉巖, 松山洋一, and 小林哲則. "多人数会話ファシリテーションロボットの主導権奪取手続き." 研究報告音声言語情報処理 (SLP) 2013.10 (2013): 1-8.
- [6] 松山洋一, et al. "「置いてけぼり」を救う多人数会話活性化ロボット." HAI シンポジウム (2012).
- [7] Goffman, Erving. Forms of talk. University of Pennsylvania Press, 1981.
- [8] Clark, Herbert H. Using language. Cambridge university press, 1996.

OpenPose を利用した人の位置やベッド装置上での状態検知及び可視化システム

宇野健一¹ 増田翔馬¹ 岩井将行¹

¹ 東京電機大学未来科学部情報メディア学科 〒120-8551 東京都足立区千住旭町5番1号館 11107A/B 実空間コンピューティング研究室

概要

近年、少子高齢化が深刻化している日本では、慢性的な介護施設での人手不足が問題となっている。特に、個々の被介護者を見回す時間が増加し、本来業務の従事する低下に繋がっている。本研究では、施設内の別の部屋から遠隔で被介護者の位置やベッド上での状態の確認を行えるプライバシーを考慮したシステムを開発することで、人手不足による介護パフォーマンスの低下の改善を試みた。本システムは、部屋全体を見渡せる位置とベッド上に USB カメラを設置し、カメラから得た映像を OpenPose で処理することにより得られた骨格情報のみから被観測者の位置や状態を Web のブラウザ上で可視化するものである。また、本ブラウザ可視化システムは、被観測者以外の「もの」の情報も可視化できるように機能を拡張した。実験の結果、部屋全体を見渡す USB カメラの映像から、人の座標推定は行うことはできたが足を OpenPose で認識できない状態があること、また誤差も確認された。さらに、ベッド上での位置検知も大部分では検知できていたが、寝返りの角度によっては人物を検出することが出来ず転落したと誤検知を起こすことがあった。可視化システムではラグは適正な範囲内であり、問題なく現在位置を表示することができたことを報告する。部屋内の座標推定の精度をさらに上げるためには、数式やカメラの位置などのさらなる調節が必要だと考えられる。OpenPose を用いることで、深度計測カメラを用いることなく室内の座標を求めるといった立体的な認識を行うことができた。一方で奥行き認識精度がやや低いことや、それまで違う形の部屋にカメラを設置する場合は部屋の形や大きさなどの情報の入力、カメラの位置調節を行う必要があるなど、実用には課題があることを明らかにした。

キーワード： OpenPose, Firebase, 位置推定, 可視化,

1. はじめに

近年、少子高齢化が深刻化している日本では、慢性的な高齢者介護施設での人手不足が問題となっている。それは、個々の被介護者を診る時間や回数が減ることによる介護パフォーマンスの低下に繋がっている。たとえば、ベッドからの転落や夜間の不用意な徘徊が発生した場合には、速やかな発見と対処が必要であるが、人手不足の問題などから、発見が遅れ被介護者が死亡してしまう事例も多々存在する[1]。

本研究は、様々な人手不足による問題の中でも、特にベッドからの転落や夜間の不用意な徘徊行動に着目し、それらを遠隔で検知、確認できるシステムを開発することで、介護パフォーマンス低下の改善を試みた。

2. 関連研究

画像処理を用いた骨格検知による被観測者の状態推定に関する研究は数多く存在する。吉武らは、プライバシー性の高い医療用患者監視システムの導入の低コスト化を目的に、Kinect センサを用いた研究を行った[2]。Kenny Chenらは、布団や医療道具などのノイズの多い実際の医療現場での利用にも対応する状態推定システムを単眼カメラと深層学習を用いて開発した[3]。

本研究では、被観測者の位置推定を行い、その情報を用いることで、ベッドからの転落などの状態も推定できるようにシステムを構築した。介護者は、そのシステムを通じて遠隔で被介護者の位置とベッド上での状態の確認が可能となる。

3. 位置検知システム

3.1 概要

このシステムでは OpenPose を用いてカメラからの映像を検証し、骨格のデータを取得し、そのデータを使って位置を検知する。カメラは部屋全体を見渡す1台と、ベッド1つにつき1台のカメラを用意する。部屋全体を見渡すカメラからの映像で部屋のどこに人がいるかという座標情報を検知し、ベッドに取り付けられたカメラからの映像でベッドからの転落を検知し、検知システムが Firebase 上のデータベースへ送信する。

3.2 部屋内での座標推定

部屋の中での座標推定は1つのカメラから得られた映像をもとに OpenPose で処理を行い、処理で得られた骨格データを用いて部屋の中での座標を推定するというシステムで行う。座標を推定する方法はまず、得られたデータから床のどの部分に立っているかを検出し、それを実際の x, y 座標への変換を行う数式や条件分岐で変換し部屋の中での座標とした。図1、図2に OpenPose で認識した際の画像を示す。



図1、図2 OpenPose での認識画像

3.3 ベッド上での転落検知

ベッド上での転落推定は、ベッドの頭上にカメラを設置しカメラからの映像を OpenPose で処理して得られたデータから転落したかどうかを判定してデータベースに通知するという流れで行われる。

4. 可視化システム

4.1 概要

今回開発した可視化システムは、位置情報を検知するシステムから送信される被介護者の情報や、あらかじめアプリ上で作成した点滴やベッド等の情報をクラウド上で記録し、それをリアルタイムで可視化するWebアプリケーションである。スクリーンショットを図3に示す。

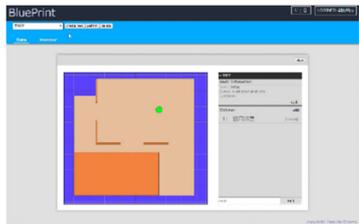


図3 Webアプリケーションのスクリーンショット

4.2 主な機能

クラウド上に記録されている被介護者や物の情報を基に3Dモデルを生成し、上部からの視点で可視化した。介護者は、図形の位置や色の変化から被介護者の異常を検知することが可能となる。また、被介護者の人数や介護施設、病室の大きさに合わせた可視化エリアの設定を可能にするため、部屋の大きさや数は任意に設定できるように実装した。

4.3 追加の機能

- 選択中の可視化エリアへのアクセスが可能なアカウント情報を表示する。
- 選択中の可視化エリアへのアクセス権を授与、剥奪ができる。
- 可視化エリアの追加、削除ができる。
- 選択中の可視化エリアへのオブジェクトの追加、編集、削除ができる。
- 選択中の可視化エリア内の被介護者や物の検索ができる。

4.4 システムの構成

- バックエンド

情報の記録には、FirebaseのRealtime Databaseを使用した。バックエンドの構成図を図4に示す。

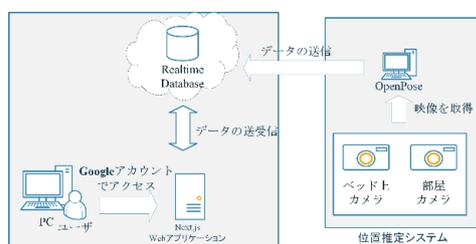


図4 バックエンドの構成図

- フロントエンド

被介護者や点滴、ベッド等の物を可視化するため、Next.jsフレームワークやReactライブラリのthree.jsを主に使用した。

5. 評価

本研究では、被験者2名を対象に、あらかじめ部屋内に定めた座標に沿って歩行してもらい、どれくらい位置を正確に推定できるか検証した。表1に、足を認識できる状態で歩いてもらった1人の被験者の結果を示す。

表1 実験結果の一部

| 観測場所 | 理想値 | | 結果 | | 誤差 | |
|------|-----|-----|-------|-----|-------|------|
| | x | y | x | y | x | y |
| 1 | 0.9 | 0.1 | 0.87 | 0.1 | -0.03 | 0 |
| 2 | 0.9 | 0.5 | 1.04 | 0.5 | +0.14 | 0 |
| 3 | 0.9 | 0.9 | 1.15 | 0.8 | +0.25 | -0.1 |
| 4 | 0.5 | 0.1 | 0.45 | 0.1 | -0.05 | 0 |
| 5 | 0.5 | 0.5 | 0.45 | 0.5 | -0.05 | 0 |
| 6 | 0.5 | 0.9 | 0.45 | 0.9 | -0.05 | 0 |
| 7 | 0.1 | 0.1 | 0.07 | 0.1 | -0.03 | 0 |
| 8 | 0.1 | 0.5 | -0.04 | 0.5 | -0.14 | 0 |
| 9 | 0.1 | 0.9 | 0.03 | 0.9 | -0.07 | 0 |

表1より0.25以下の誤差で移動の推定が可能であり、位置の確認には十分な精度であることが分かった。

被験者に仰向けの姿勢で横たわる、90度寝返りを打つ、うつ伏せの姿勢で横たわる、ベッドから転落するという動作を順番に行う実験を4回行った。その結果、4回ともベッドからの転落の検知に成功した。

6. まとめと展望

本研究では、介護現場を想定するOpenPoseを用いた位置推定がOpenPoseの3-D Reconstruction Moduleを用いず、また、深度カメラを用いることなく実現できることが分かった。一方、精度の低さやカメラの位置調節の手間、OpenPoseの商用ライセンス料が年間25000ドルと高額なことなど、実用に対して課題も多い。特に、精度面において実際の介護現場で用いるには現状のシステムからの一段の向上が望まれる。また、可視化システムのアイコンの種類の変現も不足する可能性がある。

今後の発展として、カメラとほかのセンサやデバイスを併用し、認識の精度を上げることや、個人を識別できるように顔認証と組み合わせるなどプライバシーを考慮して改善を進めていく。

文献

- [1] 医療事故調査・支援センター 一般社団法人 日本医療安全調査機構, “入院中に発生した転倒・転落による頭部外傷に係る死亡事例の分析”, 医療事故の再発防止に向けた提言第9号.
- [2] 吉武伸泰, 恒田晃完, 灘口尚大, 田中康一郎, “Kinect センサを用いた医療用患者監視システムにおける検知機能の実装”, 情報処理学会研究報告, 2013 B-5-3.
- [3] K. Chenet et al., “Patient-specific pose estimation in a clinical environment,” in Proc. Social Mach. Learn. Symp., Oct. 2017, pp. 1–9.

加速度センサを用いた包丁技術の判定

小林花菜乃¹ 加藤岳大² 横窪安奈¹ Guillaume LOPEZ¹

¹ 青山学院大学理工学部情報テクノロジー学科 〒252-0206 神奈川県相模原市中央区淵野辺 5-10-1

² 青山学院大学大学院理工学研究科知能情報コース 〒252-0206 神奈川県相模原市中央区淵野辺 5-10-1

概要

料理初心者にとって、包丁は怪我の恐れがある道具であり、食材を適切に切断することが難しい。包丁技術の向上支援の研究として、児童が安全に食材の切り方体験ができるシステムや、熟練者の包丁操作の可視化によって、学習者に所作を伝達するシステムが提案されてきた。しかし、いずれも調理中の使用が難しいことや、リアルタイムでのフィードバックが得られないことから、未だ多くの課題が残されている。本研究では、料理熟練者による包丁技術の特徴量を明らかにするために、包丁の柄先に装着可能な加速度センサを用いて包丁技術を判定する。実験結果により、料理熟練者は包丁を「前後」に動かしていることが明らかになり、包丁使用時にはリズムが影響することが示唆された。

キーワード： 料理支援 加速度センサ

1. はじめに

我々が健康に生きるために食生活習慣は重要である。食生活を充実させるためには、個人が自分の食生活を管理する能力を身に付け、自分自身で料理することが望ましい[1]。幅広い年齢層の人が基本的な調理技術を身につけるべきであると思っている一方、調理技術や知識のレベルにはばらつきや個人差が大きいことが指摘されている[2]。

調理技術力が低い場合、包丁の基本的な使い方を習得していない人が多く[3]、食材を適切に切断することが難しい。包丁技術を向上させる研究として、児童が安全に食材の切り方体験ができるシステム[4]や、熟練者の包丁操作の可視化によって、学習者に所作を伝達するシステムが提案されてきた[5]。しかし、いずれも調理中の使用が難しいことや、リアルタイムでのフィードバックが得られないことから未だ多くの課題が残されている。本研究では包丁の柄先に装着可能な加速度センサを用い、料理熟練者の特徴量を抽出し包丁技術を判定する。判定した値を元に、リアルタイムで包丁技術のフィードバックが可能となるシステムを提案する。

2. 提案システム

提案システムは、料理熟練者の特徴量を抽出するために、包丁及びタブレット端末からの情報を収集しPCを用いて分析するシステムである(図1)。包丁の柄先に加速度センサを装着し、タブレットをまな板に代替して使用する。今回は切断動作の妨げにならないように、包丁の柄先に装着可能なサイズである Arduino 互換マイコンである TinyDuino (TinyCircuits 社製)の加速度センサシールドを採用した。加速度センサから切断動作の加速度を取得し、切断対象となる食材、切断後の食材の位置をタブレット端末の座標から取得する。本システムでは、一連の切断動作の加速度データから料理熟練者(以下、熟練者)と料理初心者(以下、初心者)の動作の違いを分析する。

3. 実験

3.1 実験の概要

熟練者と初心者の動作違いを明らかにするために、切断動作の加速度データを収集する実験を行った。被

験者は男女15名(熟練者2名、初心者13名)で、今回は食材としてきゅうりを選択し、きゅうりの小口切りを計測した。

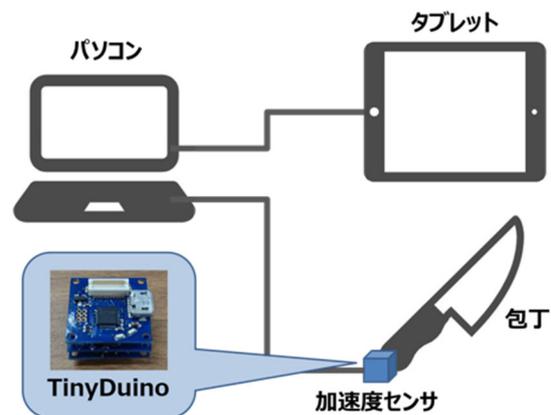


図1 提案システムの概要

3.2 実験の手順

以下の手順に従い、被験者にきゅうりを切断してもらった。

- (1) タブレットの上に長さ10cmのきゅうりを設置する。
- (2) 加速度センサを取り付けた包丁を用いて小口切りを行なう。
- (3) (1),(2)を被験者1人に対して5回行なう。

3.3 実験の結果

被験者1と被験者3は熟練者、それ以外の被験者が初心者であった。図2に被験者1のデータを示す。被験者1は日常的に調理を行っている熟練者であり、Z軸の変化が大きかったことがわかった。これは包丁できゅうりを切断する際に、包丁を垂直に下ろすのではなく、奥に押し込んでいることを意味すると考えられる。被験者1は熟練者の包丁技術の「包丁を押し切る」という特徴を再現していたと考えられる。

被験者1の結果を踏まえ、Z軸加速度の標準偏差の多重比較検定を行った(図3)。図3より、Z軸方向の標準偏差は、熟練者2名を除く初心者13人中11人に有意差があり、13人は被験者1よりも標準偏差の平均

値が低かった。これは、初心者は包丁使用時に、「前後」ではなく「上下」に包丁を動かしていることが多いと推察できる。以上から、Z軸方向の標準偏差が包丁技術の向上の指標の1つとして用いることができると考えられる。

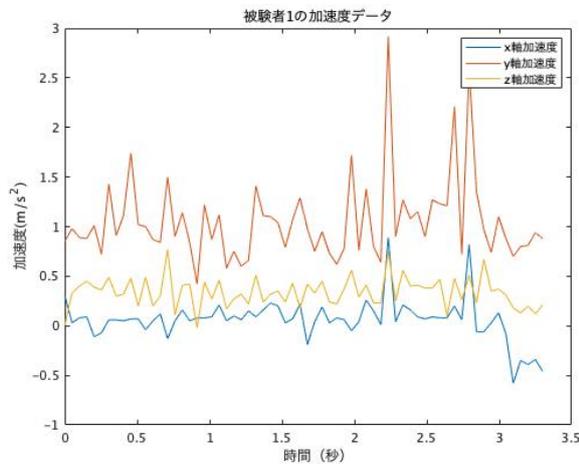


図2 被験者1の加速度データ

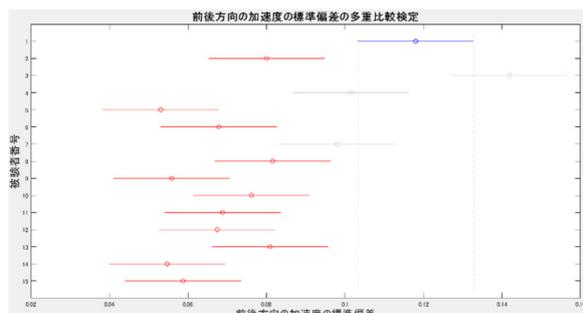


図3 Z軸加速度の標準偏差における多重比較検定

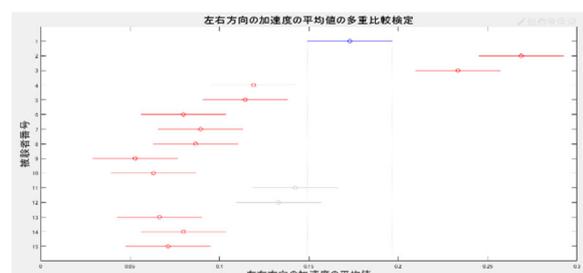


図4 X軸加速度の平均値の多重比較検定

Z軸加速度と同様に、左右方向の動きであるX軸加速度の平均値においても多重比較検定を行った(図4)。図4より、被験者1, 3は他の被験者に関して左右方向への動作が大きかった。これは食材を切断する際に、食材を抑える添え手に包丁が接触し、包丁が傾いているためだと考えられる。この結果から、X軸の加速度の値から添え手の有無を検出できる可能性が示唆された。

上下方向の動きであるY軸加速度に対して周波数解析を行い、最大周波数を算出した(表1)。被験者1, 3は最大周波数が初心者に比べて高かった。これは包丁技術が高く、きゅうりをリズムよく切断するためだと推察できる。被験者9は切断するリズムが緩やかであったにも関わらず、最大周波数が全被験者のうち

でも最大であった。この要因として、本システムで用いた包丁とTinyDuinoはテープで接着していたことで、包丁を動かす際にTinyDuinoが予期せず動いてしまったことが挙げられる。また被験者7, 8, 11, 13, 15は最大周波数が0.2以下であった。この要因は、加速度データから重力加速度を取り除いていないことが挙げられる。

表1 各被験者におけるY軸加速度の最大周波数(Hz)

| 被験者番号 | 最大周波数 | 被験者番号 | 最大周波数 |
|-------|-------|-------|-------|
| 被験者1 | 2.40 | 被験者9 | 5.79 |
| 被験者2 | 3.78 | 被験者10 | 0.96 |
| 被験者3 | 4.58 | 被験者11 | 0.05 |
| 被験者4 | 1.38 | 被験者12 | 0.40 |
| 被験者5 | 0.87 | 被験者13 | 0.13 |
| 被験者6 | 0.93 | 被験者14 | 2.00 |
| 被験者7 | 0.05 | 被験者15 | 0.04 |
| 被験者8 | 0.054 | | |

4. まとめと今後の展望

本実験を通して、初心者の多くは包丁を前後に動かせていないことが明らかになった。X軸方向の加速度の平均により、食材を抑える添え手を検出できる可能性が示唆された。また、熟練者はリズム良く食材を切断していることが明らかになり、熟練者と初心者の違いには、包丁使用時の動作リズムが影響すると考えられる。

今後の展望として、本実験では重力加速度を含めたデータが混在していたため、重力加速度を除去し、再度周波数解析を実施する必要がある。また、今回の分析は加速度データのみを対象とし、包丁の速度や変位における分析ができていなかった。今後はこれらのデータについても分析を行なっていきたい。また、本実験で収集したタブレット端末の座標を分析し、初心者に適切なフィードバックを提示するための適切な情報提示位置や提示コンテンツを検討し、初心者の包丁技術の向上支援システムを構築していきたい。

文献

- [1] 久保加織, 堀越昌子, 岸田恵津, 増澤康男, 細谷圭助, 中西洋子, 成瀬明子, “調理技術教育プログラムの構築に向けてのアンケート調査”, 日本調理科学会誌, Vol.40, No6, pp449-455, 2007.
- [2] 増澤 康男, 岸田 恵津, 久保 加織, 堀越 昌子, 細谷圭助, 中西 洋子, 成瀬 明子, “学習者の視点を取り入れた食物教育に向けてのアンケート調査 - 学校教育にのぞまれているもの -”, Vol.53, No1, pp65-77, 2002.
- [3] 堀 光代, 平島 円, 磯部由香, 長野宏子, “大学生の調理に対する意識調査”, 岐阜市立女子短期大学研究紀要第57輯(2008)
- [4] 斎藤志保, 野嶋琢也, 広田光一, 櫻井 翔, “KKse-包丁調理における児童の安全教育システム”, 電子情報通信学会技術研究報告, Vol.116, No.340, pp.39-44 (Nov.2016)
- [5] 由良 亮, 浜野 純, 萩原勇人, 楠瀬千春, “6軸モーションセンサーを用いた包丁技術の可視化装置の検討”, 中京学院大学中京短期大学部研究紀要第47巻第1号(2017)

曲げセンサを用いた曲率推定による腹囲形状計測

大石征裕¹ 藤波香織²

¹ 東京農工大学大学院工学府情報工学専攻 〒183-8588 東京都小金井市中町 2-24-16

² 東京農工大学大学院工学府先端情報科学部門 〒183-8588 東京都小金井市中町 2-24-16

概要

背景と目的: 近年、肥満が社会問題となっており、成人男性の約半数がメタボリックシンドロームの患者もしくは予備群と報告されている。そのため、定期健康診断では腹囲長が診断基準として用いられている。しかし、腹囲長のみでは発見不可能な肥満が存在し、内臓脂肪量の推定方法として腹囲形状を用いることが有効であると報告されている。そこで、本研究ではベルトの形状やベルトに加わる圧力を計測することで、装着者の腹囲形状を推定することを目的とする。

方法: 初めに、ベルトに装着した7つの曲げセンサを用いてベルトの形状を推定する。次に、7つの感圧センサを用いてベルトに加わる圧力を推定し、ベルトによる押下量を推定する。ベルトの形状と押下量を組み合わせることでデバイス装着者の腹囲形状を推定する。最後に、推定した腹囲形状を記録することで装着者が日々の腹囲形状の振り返りを可能にする。

キーワード: ベルト型デバイス 曲げセンサ 腹囲形状 形状推定

1. 背景と目的

近年、肥満者の増加が深刻な社会問題になっている。厚生労働省の報告書では、成人男性の約半数がメタボリックシンドロームの患者もしくは予備群であると述べられている[1]。そのため、内臓脂肪量を基準とした診断が行われ、メタボリックシンドロームの早期発見と患者および予備群への生活改善指導を行っている。しかし、内臓脂肪量の測定は、CTやMRIなどの特別な装置やそれらを取り扱う専門的な技術が必要であり、定期的な測定が困難である。そこで腹囲長に着目し、男性は85cm、女性においては90cmを基準値とした肥満の簡易的なスクリーニング手法が用いられている。しかし、腹囲長の計測のみでは発見できない肥満が存在するため、腹部断面積や腹囲形状の計測が肥満の判定に有効であると報告されている[2]。そこで、腹部断面形状の取得する手法として、X線やCTを用いた方法[3]や3Dスキャナを用いた方法[4]が挙げられる。しかし、これらの方法では大型な装置が必要であることやコスト面での問題がある。さらに、定期健康診断に加えて毎日腹囲を測定し、記録することが、メタボリックシンドロームによる病気リスクの減少に有効であると報告されている[5]。

そこで、本研究では日常的に装着可能なベルト型ウェアラブルデバイスを用いて、装着者の腹囲形状を計測するシステムの開発を目的とする。

2. 先行研究

我々の先行研究では、感圧センサ6個と曲げセンサ3個を用いて腹囲形状の推定を行う第一型プロトタイプシステムを製作した[6]。また、評価実験は計6名の被験者が参加した。被験者は条件を比較するために、日本肥満学会の判定基準であるBMIにより、「低体重」「普通体重」「肥満」に該当する各2名を選定した。

被験者A~Fの腹囲形状を推定した結果を図1に示す。各被験者において推定誤差は表1に示すようになり、平均3.6%となった。また、第一型プロトタイプシステムは一般的な腹囲長(79.7cm)を基準値として実装した。そのため、次式で与えられる腹囲長の差が精度に大きく影響することが分かった。

$$\text{腹囲長の差} = \left| \frac{\text{被験者の腹囲長} - \text{基準値}}{\text{基準値}} \right| \times 100 \quad (1)$$

そこで、本稿ではセンサ数や配置を変更した改良型プロトタイプシステムを開発する。また、装着者の腹囲長に応じた、腹囲形状推定に有効なセンサ位置について調査することで、推定精度向上を目的とする。

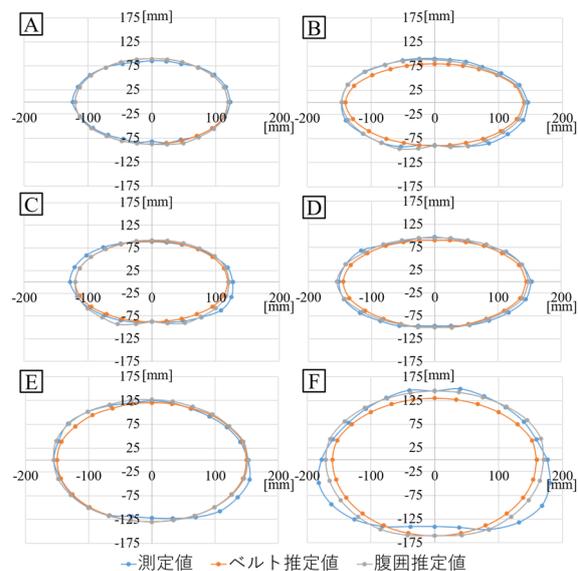


図1 第一型プロトタイプシステムの推定結果

表1 第一型プロトタイプシステムの推定精度

| 被験者 | 体重 | 推定誤差 | 腹囲長の差 |
|-----|------|------|-------|
| A | 低体重 | 3.5% | 16.9% |
| B | 低体重 | 2.9% | 5.1% |
| C | 普通体重 | 4.0% | 11.8% |
| D | 普通体重 | 2.4% | 0.4% |
| E | 肥満 | 3.2% | 10.7% |
| F | 肥満 | 5.6% | 34.4% |
| 平均 | | 3.6% | |

3. 改良型プロトタイプシステムの提案

3.1 改良型システム概要

改良型のシステム概要を図2に示す。事前に、曲げセンサのセンサ値と曲率半径、感圧センサのセンサ値と押下量の関係を調査する。次に、曲げセンサの値から調査した関係性を用いて曲率半径を推定し、ベルトの形状推定を行う。しかし、腹部はベルトによって締め付けられるため、腹囲形状とベルト形状は異なる。そこで、事前に調査した関係性を用いて感圧センサの値からベルトによる腹部の押下量を推定する。そして、ベルトによる押下量でベルト形状を補正することにより、ベルト未装着時の腹囲形状の推定を行う。

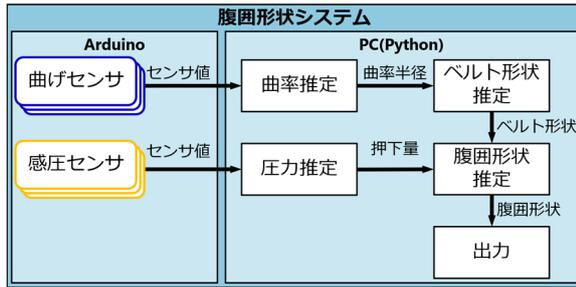


図2 改良型のシステム概要

3.2 センサ配置

第一型プロトタイプシステムでは、図3(a)に示すようにセンサを配置し、腹囲形状の推定を行った。改良型では、感圧センサと曲げセンサの数をベルトに装着可能な最大数の7個ずつに増加させることで、詳細にデータを取得する。また、センサ配置の間隔は、感圧センサは15°、曲げセンサは5°とした。これは、感圧センサは正面部や背面部のデータ取得が困難であり、曲げセンサは曲率が小さくなると測定が困難になるためである。さらに、押下量の推定は180°付近の感圧センサの方が有効であったので、感圧センサの配置は180°付近に集中して実装する。

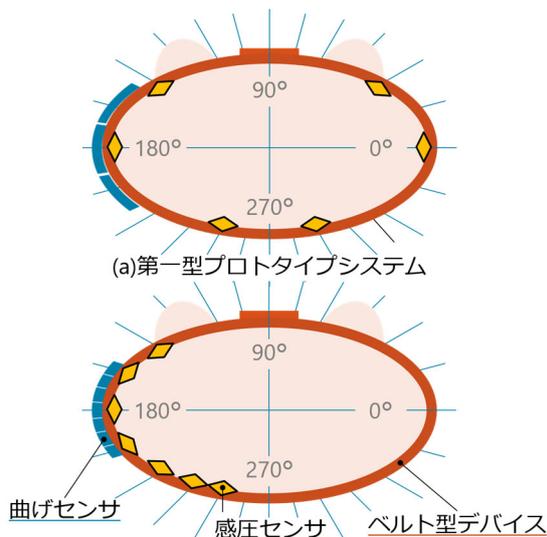


図3 システムのセンサ配置

3.3 提案システムの実装

現在、本研究では図3(b)の改良型プロトタイプシステムのセンサ配置に従って曲げセンサ、感圧センサをベルトに実装した。また、データ送信用のArduino Megaを使用した。実装したベルト型デバイスを図4に示す。図は、腹部模型にベルト型デバイスを装着させたときのデバイスである。さらに、図2に示した各機能も実装が完了した。そのため、今後は機能の統合と模型に対する評価実験を行っていく。

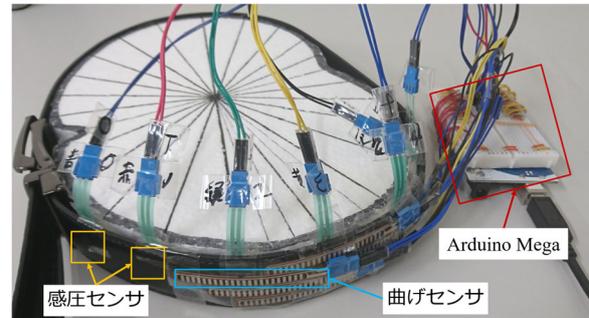


図4 ベルト型デバイス

4. おわりに

本稿では第一型プロトタイプシステムの結果から推定精度改善を目的とした改良型プロトタイプシステムを提案した。今後は実装した機能を統合し、改良型プロトタイプシステムを模型や人に対して腹囲形状推定の精度評価実験を行っていく。また、装着者の腹囲長に応じた、腹囲形状推定に有効なセンサ位置について調査を行い、最低限のセンサ数で高精度な推定が可能なベルト型デバイスを開発する。

さらに、リアルタイムに腹囲形状を装着者に伝えるシステムにすることで、ユーザにより好まれると考えられる。また、ベルト型デバイスを用いて装着者の腹囲形状を取得することで、装着者の行動の分類や食事量の計測が可能となると考えられる。

文献

- [1] 厚生労働省. 国民健康・栄養調査 (2017年). http://www.mhlw.go.jp/bunya/kenkou/kenkou_eiyouchousa.html. (2019/12/1 閲覧)
- [2] 齋藤仙一, 亀嶋英人, 佐藤幸男. “腹部断面形状からの内臓脂肪面積の推定法”. 電子情報通信学会論文誌 D, 92(11), pp.2059–2066, 2009.
- [3] 菊池信幸, 菅原雅弘, 米沢亮作. “メタボリックシンドローム検査の腹囲計測値と CT fat scan の内臓脂肪値”. 通信医学, 62(1), pp.41–49, 2010.
- [4] 河内まき子, 持丸正明. “形状スキャナによる人体寸法計測の誤差要因の検討”. Anthropological Science (Japanese Series), 113(1), pp. 63–75, 2005.
- [5] P. T. Katzmarzyk, et al. “The importance of waist circumference in the definition of metabolic syndrome: prospective analyses of mortality in men”. Diabetes care, vol.29(2), pp. 404–409, 2006.
- [6] 大石征裕, 前原啓太, 藤波香織, “ベルト型デバイスを用いた腹囲形状の推定”, 電子情報通信学会ヒューマンプロブ研究会第20回研究会, 2019

深層学習と単眼カメラを用いた骨格による人物識別手法の構築

浅見直人¹ 高橋正裕² 橋本慧志² Alessandro Moro³ 池勇勲¹ 梅田和昇¹

1 中央大学理工学部精密機械工学科 〒112-8551 東京都文京区春日 1-13-27

2 中央大学理工学研究科精密工学専攻 〒112-8551 東京都文京区春日 1-13-27

3 RITECS Inc. 〒190-0023 東京都立川市柴崎町 3-5-11

概要

近年、個人認証の中でも、認証者の負担とならない動画画像を用いた認証手法が研究されている。本研究では人物の歩行に着目し、歩行動画像から人物を識別するシステムの構築を目指す。天井に設置された魚眼カメラで撮影された動画画像から、骨格推定アルゴリズムである OpenPose によって骨格座標点を推定し、取得された時系列骨格座標データにより人物を識別する手法を提案する。さらに、潜在変数を考慮した深層生成モデルによる識別を行い、部分的には従来の識別モデルと同等の識別結果が得られたことを確認する。

キーワード：人物識別 深層学習 深層生成モデル

1. 序論

個人認証に用いられる特徴の中でも、骨格特徴や歩容特徴は他の多くの生体特徴と異なり、離れたセンサからでもデータを取得することが可能である。歩容を用いた認証には人物のシルエットに基づく特徴量が用いられることがある[1]が、服装や荷物の影響を受けてしまう問題がある。そこで、本研究では人物の骨格情報に着目する。骨格情報による人物識別に関してはいくつかの研究がされており、その有用性が示されている[2]。本研究では、一般的な単眼カメラを利用した骨格による個人識別手法の構築を目指す。

2. 提案手法

2.1 骨格特徴の抽出

まず、天井に設置された魚眼カメラから動画を取得し、YOLO[3]を用いて人物領域の推定を行う。推定された人物領域の画像を正像変換[4]し歪みを取り除いた後、OpenPose ライブラリ[5]を利用して骨格情報を取得する。

2.2 深層生成モデルの利用

一般的に、画像内での人物領域は部屋の様々な場所に存在し、人物は様々な方向に歩行するため、日常生活の歩行データを用いて識別を行う場合、骨格時系列データのクラス内分散は大きくなる。また、2次元に射影された画像内での骨格の長さ情報を正規化することは困難である。そのような骨格時系列データを大量に集めても、データはクラス内で多峰性の分布となってしまう。しかし、多峰性のデータが様々な潜在的関係性のもと生成されていることを考慮すれば、潜在変数を用いた生成モデルによる識別が可能である。本研

究では、深層生成モデルによる識別を考え、その第一段階として潜在変数が混合ガウス分布に従う変分オートエンコーダ(VAE: Variational autoencoder)を提案する。

2.2.1 混合ガウス分布による VAE

一般に VAE では、潜在変数 \mathbf{z} からデータ \mathbf{x} が生成される生成モデル $p_{\theta}(\mathbf{x}|\mathbf{z})p_{\theta}(\mathbf{z})$ を深層学習で表現し、以下の変分下限 $\mathcal{L}[\phi, \theta]$ を上げるように学習する。

$$\mathcal{L}[\phi, \theta] = \mathbb{E}_{q_{\phi}(\mathbf{z}|\mathbf{x})} \left[\log p_{\theta}(\mathbf{x}|\mathbf{z}) + \log \frac{q_{\phi}(\mathbf{z}|\mathbf{x})}{p_{\theta}(\mathbf{z})} \right] \quad (1)$$

ここで、 $q_{\phi}(\mathbf{z}|\mathbf{x})$ は近似事後分布であり、 θ, ϕ はニューラルネットワークを用いて表現されていることを意味する。変分下限のモンテカルロ推定を行う際、近似事後分布 $q_{\phi}(\mathbf{z}|\mathbf{x})$ からの勾配逆伝播可能なサンプリングと尤度関数の値が計算出来れば、任意の分布で VAE を作成できる。本研究では、カテゴリカル分布に従う潜在変数 \mathbf{s} と混合ガウス分布に従う潜在変数 \mathbf{z} を持つ生成モデルからデータが生成されると仮定する。カテゴリカル分布は各クラスへ分類される確率 $\boldsymbol{\pi}$ をパラメータとし、1つのみ1で他は0の要素を持つベクトル(one-of-K 表現)を生成する確率分布である。カテゴリカル分布は連続値のカテゴリに従う ExpRelaxedCategorical 分布[6]を用い、ガウス分布はクラス数分用意する。ExpGumbelSoftmax trick[6]を用いてサンプリングした実数値を持つ連続カテゴリ \mathbf{s}' を one-of-K 表現に強制して得られた離散カテゴリ \mathbf{s} と対応するガウス分布から \mathbf{z} をサンプリングする。

2.2.2 教師データ

VAE では単に教師データを与えることができないため、本研究では生成モデルの事前分布 $p_{\theta}(\mathbf{z})$ に教師

データを与える。ここで教師データを事前分布のカテゴリカル分布のパラメータ π として与える。 N 個の教師データ $\mathbf{s}_i (i = 1 \dots N)$ に微小なノイズを加え、温度付き Softmax 関数で正規化することで疑似的な π を作る。各クラスの平均については、近似事後分布の推論に用いる Arcface[7]から重みベクトルをクラスに対応させて取得する。Arcface は距離学習の識別器の一つで、クラスを代表するベクトル \mathbf{w}_k が、あるマージン以上に分離されるよう学習される。分散については単位行列を与える。結局、カテゴリ潜在変数 \mathbf{s} の事後分布を推論することがクラス分類となる。

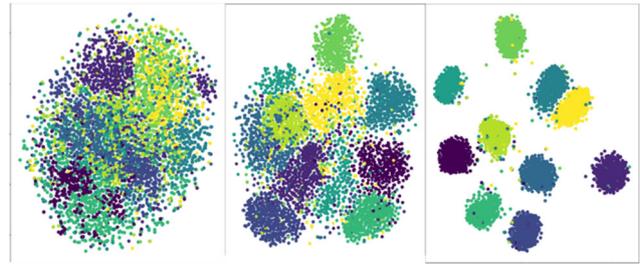
3. 実験

混合ガウス分布による VAE を用いて、手書き数字のデータセットである Mnist データセットと骨格時系列データについて識別を行った。Mnist データについて 10 クラス 60000 枚の訓練データで学習し、10000 枚のテストデータで識別を行い、提案手法と Arcface による識別のみとで比較した。10 epoch 学習させたときの識別結果の精度を表 1 に示した。Arcface による識別に比べて精度が落ちずに深層生成モデルによる識別が行えていることがわかる。図 1 は訓練データからランダムに 5000 枚画像を選び、その潜在変数を t-SNE[8]によって 2 次元へ変換し可視化したものである。各点の色は全部で 10 種類存在し、潜在変数に対応するクラスを表している。(a)はガウス分布の VAE による結果であり、(b)は事前分布の平均を $\mathbf{0}$ 、分散共分散行列を単位行列に固定した混合ガウス分布による VAE (GMM-VAE)による結果であり、(c)が提案手法の VAE による結果である。提案手法による結果では、潜在変数がクラスごとに分離していることが分かる。

骨格時系列データについて、5 人の人物の歩行動画から取得した一人当たり約 880 フレームのデータで 600 epoch 学習し、一人当たり約 240 枚のバリデーションデータで評価した。表 2 は識別結果の精度を示す。Arcface の識別のみの結果より、提案手法の結果の方が精度が低下してしまった。これは、骨格時系列データの再構成のための特徴抽出に時間がかかり、識別のための学習が適切に行えなかったためと考えられる。

4. 結論と今後の展望

本研究では、骨格情報から人物識別を行うシステムを構築するとともに、深層生成モデルによるクラス分類を提案した。今後の展望としては、人物以外例えば行動などのラベルを加えて潜在変数を増やした生成モデルでの識別を行う予定である。人物ラベルと行動ラベルのマルチラベル問題となり、深層生成モデルによって多峰性のデータ構造の再現を目指す。



(a)ガウス分布 VAE (b)GMM-VAE(平均 $\mathbf{0}$) (c)提案手法

図 1 t-SNE を用いた潜在変数の可視化

表 1 Mnist データの識別結果

| | Arcface のマージン | |
|----------------|---------------|----------|
| | 0.1[rad] | 0.5[rad] |
| 提案手法 | 97.34[%] | 97.02[%] |
| Arcface (識別のみ) | 97.72[%] | 97.79[%] |

表 2 骨格時系列データの識別結果

| | 精度 |
|----------------|----------|
| 提案手法 | 79.67[%] |
| Arcface (識別のみ) | 89.02[%] |

文献

- [1] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, "Gait Recognition Using a View Transformation Model in the Frequency Domain," Proc. Conf. on ECCV, Lecture Notes vol 3953, pp 151-163, 2006.
- [2] A. Sinha, K. Chakravarty, and B. Bhowmick, "Person Identification Using Skeleton information from Kinect," Proc. Intl. Conf. on ACHI, pp. 101-108, 2013.
- [3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," Proc. of the IEEE Conf. on CVPR, pp. 779-788, 2016.
- [4] 森隆寛, 外村元伸, 大住勇治, 池永剛, "キュービック補間を用いた魚眼レンズ画像の高画質補正アルゴリズム", 情報科学技術フォーラム一般講演論文集, 5-1, pp.7-8, 2006.
- [5] C. Zhe, T. Simon, and Y. Sheikh, "Realtime Multi-Person 2d Pose Estimation using Part Affinity Fields," Proc. of the 2017 IEEE Conf. on CVPR, pp. 1302-1310, 2017.
- [6] C.J. Maddison, A. Mnih, and Y.W. Teh, "The concrete distribution: A Continuous Relaxation of Discrete Random Variables," Workshop on BDL, NIPS, 2016
- [7] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive Angular Margin Loss for Deep Face Recognition," Proc. of the IEEE Conf. on CVPR, pp. 4690-4699, 2019.
- [8] L. Matten, G. Hinton, "Visualizing data using t-SNE," Journal of machine learning research vol. 9, pp. 2579-2605, 2008.

第34回 人間情報学会講演集

令和元年 12月 4日発行

発行所 人間情報学会

〒112-0012 東京都文京区大塚 2-20-1
お茶の水女子大学 国際交流プラザ407

Tel. 03-6912-1361 Fax. 03-6912-1362

E-mail: admin@npowin.org

URL: <http://www.ahi-soc.info>