

令和 6 年 6 月 13 日現在

機関番号：32657

研究種目：基盤研究(C)（一般）

研究期間：2021～2023

課題番号：21K12801

研究課題名（和文）視線と瞬目を用いたキャリブレーションフリー入力インタフェース

研究課題名（英文）Calibration-free Input Interface Using Eye-gaze and blink

研究代表者

阿部 清彦（Abe, Kiyohiko）

東京電機大学・システムデザイン工学部・教授

研究者番号：40408646

交付決定額（研究期間全体）：（直接経費） 2,900,000円

研究成果の概要（和文）：ユーザの視線や瞬目の情報によりパソコンを操作する視線入力、重度肢体不自由者など一般的な入力インタフェースの使用が困難な人たちのコミュニケーション支援に使われる。従来の視線入力インタフェースの多くは、使用前にユーザごとにキャリブレーション（校正）を行なう必要があり、使用に煩雑さがあった。本研究では、画像識別の一手法である畳み込みニューラルネットワーク（CNN）を利用することにより、ノートパソコンのインカメラで撮影されたユーザの眼球近傍画像から視線と瞬目の情報をリアルタイムで捉え、パソコンに文字を入力しWebブラウザを操作する新しいキャリブレーションフリーの視線入力インタフェースを開発した。

研究成果の学術的意義や社会的意義

本研究課題で開発した新しい視線入力インタフェースは、キャリブレーションフリーであり、利用に際してユーザごとの校正を一切必要としないため、使用したいときにすぐに使うことができる。またこの視線入力インタフェースは、廉価なノートパソコンのみで使用が可能であり、視線と瞬目の情報のみで文字入力を行ったりWebブラウザを操作するなど、重度肢体不自由者のコミュニケーション支援に利用できる。

研究成果の概要（英文）：Eye-gaze input interface is able to operate a PC using information from the user's gaze or blinking of the eye. Therefore, it is used to support communication for people who have difficulty using general input interfaces, such as people with severe physical disabilities. Most conventional eye-gaze input interfaces require calibration for each user before use, which is cumbersome. In this research, by using a convolutional neural network (CNN), which is a method of image identification, we have developed a new method to input text into a PC and operate a Web browser by capturing information on eye gaze and eye blinks in real time from a user's near-eye image captured by an in-camera camera of a notebook PC. A new calibration-free gaze input interface has been developed.

研究分野：福祉工学，ヒューマンインタフェース，ソフトウェア工学

キーワード：視線入力 キャリブレーションフリー 畳み込みニューラルネットワーク 視線計測 瞬目計測 コミュニケーション支援

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1. 研究開始当初の背景

近年、ユーザの視線や瞬目の情報を捉え、パソコンを操作したり文字を入力したりする視線入力の研究がさかに行なわれている。視線入力は眼球やまぶた以外の運動を必要としないため、ALS（筋萎縮性側索硬化症）などの重度肢体不自由者のコミュニケーション支援に利用されているが、その多くはユーザの視線を計測するために赤外線照明など特殊な機器を使用していた。また、これらの視線計測装置はその計測手法の性質上、システム使用前にキャリブレーションが必要であり煩雑さがあった。

これらの従来の視線計測装置は比較的高価であるものが多く、十分に普及しているとは言えない。技術革新により、安価な視線計測装置も販売されるようにはなったが、販売後数年で生産終了となってしまい、同じ機器を長年使うことが難しい状況もあった。

### 2. 研究の目的

本研究では、1. で述べた弱点を克服するため、ユーザの眼球近傍画像から畳み込みニューラルネットワーク（CNN）を用い視線方向と随意性瞬目（意識的な瞬き）を計測する新しい入力インタフェースを開発した。この入力インタフェースは、あらかじめ用意しておいた学習モデルから注視方向を識別するためキャリブレーションを必要としない大きな利点がある。また、この入力インタフェースは、ノートパソコンで撮影されたユーザの顔画像から視線方向と随意性瞬目の情報をリアルタイムで捉えることにより、文字入力や Web ブラウザの操作が可能である。

### 3. 研究の方法

#### (1) 視線方向識別モデルの生成と評価

視線入力インタフェースを実現するためには、ユーザの視線方向（注視方向）をリアルタイムで識別・計測する必要がある。本研究課題ではキャリブレーションを必要としない視線方向識別法として、画像認識に利用されるディープラーニング手法である CNN と、Vision Transformer (ViT) により、上下左右正面計 5 方向の視線を識別するモデルを生成した。これらの学習モデルを用い、学習データに顔画像を含むユーザ（特定ユーザ）と含まないユーザ（非特定ユーザ）における視線方向識別実験を行ない、その識別精度を評価した。図 1 に学習画像の一例を示す。



図 1 視線方向識別モデルの学習画像の一例

#### (2) 随意性瞬目識別モデルの生成と評価

視線入力インタフェースにおいて、カーソルの移動にはユーザの視線方向の情報をを用いるが、アイコンを選択し入力を決する際にはマウスのクリックに相当する情報が必要となる。本研究課題では、この目的としてユーザの随意性瞬目の情報を使用する。そのために、三次元畳み込みニューラルネットワーク（3D CNN）によるキャリブレーションフリーな随意性瞬目識別モデルを開発したが、通常の CNN よりも処理コストが高く、研究期間中にリアルタイム計測を実現できなかった。そのため、代替案として CNN を用いてユーザの閉眼と開眼の状態を眼球近傍画像から識別し、閉眼状態があるしきい値以上の時間継続した場合に、随意性瞬目として識別する簡易的な手法を開発した。このときに使用する学習モデルを用い、特定ユーザ非特定ユーザにおける視線方向識別実験を行ない、その識別精度を評価した。

#### (3) 視線入力インタフェースの開発と評価

(1) と (2) で述べた手法を用い、視線でカーソルを移動させ随意瞬目を行なうことによりカーソルで指示されたアイコンを選択、入力決定をする視線入力インタフェースを開発した。このシステムは、一般的なノートパソコンのみで使用することができ、日本語（50音）を入力し漢字変換を行えるほか、Web ブラウザを起動しユーザが入力したキーワードを基に Web ページの検索と閲覧が行えるインタフェースを開発した。開発したインタフェースに関して、操作時間に関する評価実験を行なった。

### 4. 研究成果

#### (1) 畳み込みニューラルネットワークによる視線方向識別

本研究課題において、当初開発した CNN による視線方向識別学習モデルは学習に使用した顔画像が十分ではなく、実用的な計測精度を確保することが難しかった。この点を改善するため、以下の (2) で述べる ViT による視線方向識別モデルの開発も並行して行ったが、CNN に比べ ViT は計算処理コストが高く、一般的なパソコンでは 1 回の識別処理に約 120ms 程度の処理時間を

要した。本研究課題で開発した入力インタフェースでは、視線方向に加え瞬目の識別も行わなければならない。一般的に、スムーズな視線入力を実現するためには、計測のサンプリング時間は100ms以下であることが望ましい。そのため、より識別の計算コストが低いCNNの学習画像を増加させることにより識別精度の向上を図った。

この新しいCNNによる視線方向識別モデルの学習には、被験者15名×顔画像180枚×分類クラス5方向×水増し16倍の合計216,000枚の画像を使用した。バッチ数は128、エポック数は6として学習を行った。被験者11名による本モデルの識別精度検証を行った。検証用画像は学習用画像と同様の手法で撮影し、被験者11名×顔画像180枚×分類クラス5方向の合計9,900枚の画像を収集した。本モデルのマクロ平均F値は71.7%であった。

## (2) Vision Transformerによる視線方向識別

(1)で示したモデルと同様に上下左右正面の計5方向の視線を識別する、ViTによる学習モデルを作成した。視線方向識別モデルの学習には、被験者15名×顔画像180枚×分類クラス5方向×水増し16倍の合計216,000枚の画像を使用した。バッチ数は128、エポック数は14としてファインチューニングを行った。被験者11名による本モデルの識別精度検証を行った。検証用画像は学習用画像と同様の手法で撮影し、被験者11名×顔画像180枚×分類クラス5方向の合計9,900枚の画像を収集した。本モデルのマクロ平均F値は90.7%であった。ViTの視線入力への応用は世界でも報告がなく、国際会議に採択され発表し高評価を得た。また、電気学会論文誌に研究開発レターとして掲載が決定した。

## (3) 3次元畳み込みニューラルネットワークによる随意性瞬目識別

視線入力インタフェースにおける入力決定のために、随意性瞬目を利用する。当初は瞬目過程の動画像から時間方向の情報を含み、3D CNNで随意性瞬目を識別する手法を開発した。本手法で構築した学習モデルのために、瞬目過程を撮影したデータセットを準備した。データセットは、随意性瞬目、非随意性瞬目、瞬目でない（開眼状態）画像サンプルからなっており、学習データは合計で19,606サンプルであり、検証データは8,406サンプルである。これらは、被験者5名のデータを学習用と検証用に分割したものである。またテスト用として、学習および検証データとはことなる被験者7名のデータ38,843サンプル用意した。本モデルのマクロ平均F値は87.7%であった。従来の随意性瞬目識別の手法は、いずれもキャリブレーションを必要とした。3D CNNによるキャリブレーションフリーの随意性瞬目の識別は世界初の成果であり、国際会議での発表のほか電気学会論文誌に原著論文として掲載された。

## (4) 畳み込みニューラルネットワークによる開眼閉眼識別モデルによる随意性瞬目識別

(3)で述べた3D CNNによる随意性瞬目の識別法は、識別精度が比較的高いものの、本研究課題内ではリアルタイム計測を実現していない。そのため、CNNによりユーザの開眼と閉眼の状態を識別し、閉眼の状態が400ms以上持続すれば随意性瞬目と判定する、簡易的な手法を採用した。

瞬目識別モデルの学習には、被験者4名×顔画像600枚×分類クラス2種類×水増し16倍の合計76,800枚の画像を使用した。バッチ数は128、エポック数は20として学習を行なった。被験者11名による本モデルの識別精度検証を行った。検証用画像は学習用画像と同様の手法で撮影し、被験者11名×顔画像180枚×分類クラス2種類の合計3,960枚の画像を収集した。本モデルのマクロ平均F値は93.4%であった。

## (5) 視線と瞬目の情報を用いたキャリブレーションフリー視線入力インタフェース

(1)と(4)で述べた手法を用い、視線入力インタフェースを構築した。開発したインタフェースのスクリーンショットを図2に示す。被験者が上下左右の方向を注視すると、それに従い入力カーソルが移動する。そして、随意性瞬目を行なうことにより入力が決定される。本インタフェースは50音の入力だけでなく漢字変換を行なうことができ、Webブラウザを操作して検索をし、マウスカーソルを移動させて閲覧を行なうことができる。

被験者9名による本インタフェースの評価実験では「電大」と入力し、東京電機大学のWebサイト開いたのちに、「東京電機大学について」→「学部」→「情報システム工学科」とページを表示させる。実験は各被験者とも2回実施し、平均の操作時間は2分12秒、1分35秒であり、被験者全員が操作を完了させることができた。



図2 開発した視線入力インタフェース

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 0件）

|  |                         |
|--|-------------------------|
| 1. 著者名<br>佐藤 寛修, 阿部 清彦, 松野 省吾, 大山 実                  | 4. 巻<br>143             |
| 2. 論文標題<br>3次元畳み込みニューラルネットワークをもちいた瞬目種類識別の詳細分析        | 5. 発行年<br>2023年         |
| 3. 雑誌名<br>電気学会論文誌C (電子・情報・システム部門誌)                   | 6. 最初と最後の頁<br>971 ~ 978 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>10.1541/ieejeiss.143.971 | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難               | 国際共著<br>-               |

|  |                         |
|--|-------------------------|
| 1. 著者名<br>佐藤 寛修, 阿部 清彦, 松野 省吾, 大山 実                  | 4. 巻<br>144             |
| 2. 論文標題<br>3D CNNをもちいた瞬目種類識別のデータ拡張による性能向上            | 5. 発行年<br>2024年         |
| 3. 雑誌名<br>電気学会論文誌C (電子・情報・システム部門誌)                   | 6. 最初と最後の頁<br>328 ~ 329 |
| 掲載論文のDOI (デジタルオブジェクト識別子)<br>10.1541/ieejeiss.144.328 | 査読の有無<br>有              |
| オープンアクセス<br>オープンアクセスではない、又はオープンアクセスが困難               | 国際共著<br>-               |

〔学会発表〕 計14件（うち招待講演 0件 / うち国際学会 4件）

|   |
|---|
| 1. 発表者名<br>Shogo Matsuno, Daiki Niikura, Kiyohiko Abe   |
| 2. 発表標題<br>Gaze Direction Classification Using Vision Transforme  |
| 3. 学会等名<br>2023 15th International Congress on Advanced Applied Informatics Winter (IIAI-AAI-Winter) (国際学会) |
| 4. 発表年<br>2023年   |

|  |
|--|
| 1. 発表者名<br>佐藤 寛修, 阿部 清彦, 松野 省吾, 大山 実     |
| 2. 発表標題<br>最頻値をもちいた3D CNN瞬目種類識別の出力補正法の検討 |
| 3. 学会等名<br>電気学会電子・情報・システム部門大会            |
| 4. 発表年<br>2023年                          |

|  |
|--|
| 1. 発表者名<br>新倉 大希, 阿部 清彦                |
| 2. 発表標題<br>Vision Transformerによる視線方向識別 |
| 3. 学会等名<br>電気学会電子・情報・システム部門大会          |
| 4. 発表年<br>2023年                        |

|                                    |
|------------------------------------|
| 1. 発表者名<br>新倉 大希, 阿部 清彦            |
| 2. 発表標題<br>視線方向識別による文字入力とウェブ検索システム |
| 3. 学会等名<br>電子情報通信学会総合大会            |
| 4. 発表年<br>2023年                    |

|                                    |
|------------------------------------|
| 1. 発表者名<br>張 超凡, 阿部清彦              |
| 2. 発表標題<br>3D-CNNによる視線方向の識別        |
| 3. 学会等名<br>2022年電気学会電子・情報・システム部門大会 |
| 4. 発表年<br>2022年                    |

|                                    |
|------------------------------------|
| 1. 発表者名<br>新倉大希, 阿部清彦              |
| 2. 発表標題<br>視線方向推定による文字入力とウェブ検索システム |
| 3. 学会等名<br>2022年電気学会電子・情報・システム部門大会 |
| 4. 発表年<br>2022年                    |

|   |
|---|
| 1. 発表者名<br>佐藤寛修, 阿部清彦, 松野省吾, 大山 実                 |
| 2. 発表標題<br>瞬目種類識別のための3次元畳み込みニューラルネットワーク向けデータ拡張の検討 |
| 3. 学会等名<br>2022年電気学会電子・情報・システム部門大会                |
| 4. 発表年<br>2022年                                   |

|  |
|--|
| 1. 発表者名<br>Chaofan Zhang, Kiyohiko Abe   |
| 2. 発表標題<br>Eye-gaze Detection based on 3D Convolutional Neural Network                           |
| 3. 学会等名<br>The Twenty-Eighth International Symposium on Artificial Life and Robotics 2023 (国際学会) |
| 4. 発表年<br>2023年  |

|  |
|--|
| 1. 発表者名<br>Hironobu Sato, Kiyohiko Abe, Shogo Matsuno, Minoru Ohyama                             |
| 2. 発表標題<br>Performance Improvement of 3D-CNN for Blink Types Classification by Data Augmentation |
| 3. 学会等名<br>The Twenty-Eighth International Symposium on Artificial Life and Robotics 2023 (国際学会) |
| 4. 発表年<br>2023年  |

|                                    |
|------------------------------------|
| 1. 発表者名<br>新倉大希, 阿部清彦              |
| 2. 発表標題<br>視線方向識別による文字入力とウェブ検索システム |
| 3. 学会等名<br>2023年電子情報通信学会総合大会       |
| 4. 発表年<br>2023年                    |

|   |
|---|
| 1. 発表者名<br>Hironobu Sato, Kiyohiko Abe, Shogo Matsuno, and Minoru Ohyama                          |
| 2. 発表標題<br>Blink State Classification Using 3D Convolutional Neural Network                       |
| 3. 学会等名<br>The Twenty-Seventh International Symposium on Artificial Life and Robotics 2022 (国際学会) |
| 4. 発表年<br>2022年   |

|                                    |
|------------------------------------|
| 1. 発表者名<br>阿部清彦, 佐藤寛修, 松野省吾, 大山実   |
| 2. 発表標題<br>視線入力インタフェース向け遠隔実験システム   |
| 3. 学会等名<br>2021年電気学会電子・情報・システム部門大会 |
| 4. 発表年<br>2021年                    |

|   |
|---|
| 1. 発表者名<br>佐藤寛修, 阿部清彦, 松野省吾, 大山実          |
| 2. 発表標題<br>3次元畳み込みニューラルネットワークによる瞬目種類識別の検討 |
| 3. 学会等名<br>2021年電気学会電子・情報・システム部門大会        |
| 4. 発表年<br>2021年                           |

|  |
|--|
| 1. 発表者名<br>張超凡, 阿部清彦                                   |
| 2. 発表標題<br>畳み込みニューラルネットワークによる 視線方向の推定                  |
| 3. 学会等名<br>電子情報通信学会 情報・システムソサイエティ特別企画 ジュニア&学生ポスターセッション |
| 4. 発表年<br>2022年  |

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

|           | 氏名<br>(ローマ字氏名)<br>(研究者番号)                      | 所属研究機関・部局・職<br>(機関番号)                     | 備考 |
|-----------|--|---|----|
| 研究<br>分担者 | 佐藤 寛修<br><br>(Sato Hironobu)<br><br>(90813741) | 関東学院大学・理工学部・助手<br><br><br>(32704)         |    |
| 研究<br>分担者 | 松野 省吾<br><br>(Matsuno Shogo)<br><br>(60836245) | 電気通信大学大学院・情報理工学研究科・准教授<br><br><br>(12301) |    |

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

| 共同研究相手国 | 相手方研究機関 |
|---------|---------|
|---------|---------|