

令和 2 年 6 月 21 日現在

機関番号：57403

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K04966

研究課題名(和文) タブレット単体で利用可能なキャリブレーションフリー視線検出型意思表示支援アプリ

研究課題名(英文) Gaze region estimation algorithm without calibration available in a tablet

研究代表者

博多 哲也 (HAKATA, Tetsuya)

熊本高等専門学校・電子情報システム工学系CIグループ・教授

研究者番号：60237899

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では、特殊なセンサを使用することなくタブレットやスマートフォン単体で利用可能なキャリブレーションフリー視線検出型意思表示支援アプリを開発した。従来の視線検出装置は、キャリブレーションと呼ばれる初期操作が必要である。しかしながら、重度重複障害のある児童生徒は指定された数点を注視するキャリブレーション動作を苦手としている。本手法は、一般的な人の顔や目の特徴を事前に学習させておくことで、タブレット内蔵のカメラで撮影された画像からキャリブレーション操作することなく、特別支援教育の現場で使用可能な精度を有するアルゴリズムである。

研究成果の学術的意義や社会的意義

視線領域検出に機械学習を用いることで、キャリブレーションを不要な仕組みを構築した。高額で特殊なセンサ機器を使用することなくタブレット端末で児童生徒の意思表示の支援が可能になり有用である。例えば、食事の挨拶の場面で使用する際、児童生徒の画面注視による意思表示時に「いただきます」の音声再生されれば、周囲の児童生徒・教諭らに自立的な動作を行えたことが自然に伝わり、その場の空間を共有することができる。日常生活における児童生徒の自発的な意思表出を周囲が理解し、達成感や自己肯定感が得られることから、自立活動の拡大に繋がること期待される。

研究成果の概要(英文)：This study proposes gaze region estimation method that does not require an expensive device nor initial setting called calibration. Objective of this study is not only implementation of the application but also enhancement of utilization of the application by teachers in special needs education school who are not good at IT or computers, hence concept of this application is "simple, user friendly but effective". In order for challenged people to show their intentions, it is considered that detecting gaze with high precision is not necessary and it is sufficient to detect gaze region concisely. Therefore, a typical webcam is adopted to implement for the algorithm. Generally, smartphone or tablet computer have the built-in camera and the proposed method is available for these devices. To figure out of applicability of the proposed algorithm, some trial experiments were conducted and the accuracy and usability is evaluated.

研究分野：福祉・人間工学

キーワード：重度重複障害 人間福祉工学 社会福祉関係 アシスティブテクノロジー 特別支援教育

1. 研究開始当初の背景

今日、視線の推定に関しては数多くの手法が考案されており、その成果は様々な分野に利用されている。一例として、発話が困難な脳血管障害や ALS 患者の意思表示を目的とした視線入力装置がある。視線で画面上のアイコンを選択したり、文字を入力したりすることで文章を書いて読み上げるなどの機能を備えており、意思や気持ちの伝達が可能になるため、生活の質が改善されることが期待されている。

特別支援学校は、学習上または生活上の困難を克服し自立を図ることを目的とした学校であり、病院から通学する重度の重複障害のある児童生徒を対象とする課程も存在する。病気や障害の状況および教育的ニーズに応じた指導が行われることで、一人一人の能力を伸ばし主体的・自立的に生活する力を育てている。著者らはこれまでも特別支援学校と連携して、学習教材や支援機器の開発を行ってきた。教諭と児童生徒のコミュニケーションの手段は様々であるが、子どもたちの持つ力に応じて表情や目の動きによる意思表示が行われることもある。保護者や教諭らの、対面に限らず視線による意思表示を可能にすることで自発的な活動を広げたいという想いに触れた。

2. 研究の目的

一般に、非接触型の視線入力装置では、視線推定の精度を高めるため、使用開始前に複数回指定された点を注視する操作(キャリブレーション)を必要としている。しかし、身体機能の障害だけでなく知的障害もある児童生徒には、キャリブレーションの操作指示に従い視線を頻繁に移動させることは大変難しく、現行の視線入力装置を使いこなして意思表示を行うことは極めて困難である。また、EDINBURGH Univ.が作成した Eye Gaze Software Curve では、目の動きを利用したソフトウェアを操作性や対象レベルに合わせて 7 段階に分類している。この表におけるマウスオーバーやシングル左クリックなどは基礎的な機能と位置づけられているが、重度重複障害者や介助者にとって、視線を利用して手軽に意思の伝達が可能になることは実用面で意味がある。

本研究では、重度重複障害者の利用を想定して、視線の分解能は粗くとも特別な視線入力装置を必要とせず、キャリブレーションの負担もなく利用可能なリアルタイム視線領域推定システムの開発を行った。初めに、画面領域の分割、Convolutional Neural Network (CNN) モデルの選択に関する予備実験を行い、提案システムに最適な組み合わせを選択した。次に、学習済みモデルを用いて、カメラ映像からリアルタイムに視線領域推定を行うシステムを構築した。

3. 研究の方法

本研究において、図 1 に示されるように横になっている状態の対象者に対し、教諭や保護者がタブレット端末の画面を見せる、もしくは対象者がアーム等で固定されたタブレット端末の画面のアイコンを注視して意思表示をするといった場面を想定している。この利用場面では、「どの座標を注視しているのか」といった高精度な視線推定ではなく、「どの領域を見ているのか」といった視線領域推定ができれば意思表示が行える(図 2)。これを特殊なハードウェアを用いずに実現することができれば、利便性は著しく向上する。

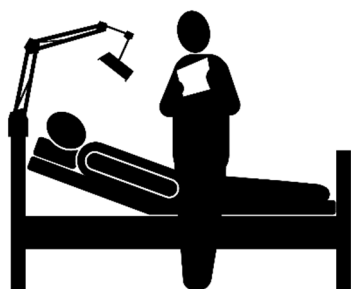


図 1 提案システムの利用シーン

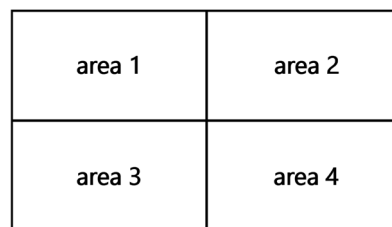


図 2 視線領域のラベリング

3. 1 AlexNet をベースとした視線領域推定モデル

はじめに、畳み込み層のベースに AlexNet を用い CNN モデルを構築した。その際、目領域に対しては畳み込み層を浅く、顔領域に対しては深く設計した。また、過学習を抑制するために、最終段の畳み込み演算及びプーリング処理の後にドロップアウトを施している。最適化手法には Momentum、活性化関数には Relu 関数を使用し、重みの初期値には He の初期値を用いた。しかしながら、AlexNet ベースの CNN モデルにおいて、画面 2 分割の教師データを 100epoch 学習させたときの認識精度は 47% から 53% の間で停滞した。この結果は、モデルが未学習に陥って

いることを示している。以上の予備実験を複数回行い、より抽象的な特徴量の抽出、過学習抑制の2点に問題を整理しモデルの改善を行った。提案手法では、視線を移した際に頭部が回転することも考慮し、視線推定時に、目領域画像に加えて頭部姿勢を得るための顔領域画像の学習も行う facial appearance を取得する。

3.2 VGG16 をベースとした視線領域推定モデル

予備実験の結果を踏まえ、より抽象的な特徴量を抽出できるように CNN モデルを AlexNet ベースから VGG16 ベースへと変更した(図3)。また、各領域の座標を1次元化した特徴マップに連結させる前に全結合層に通し、座標や顔の幅の学習及び特徴量の増加を図った。さらに、入力及び全結合層にドロップアウトを付加し、損失関数に係数 0.005 の L2 正則化項を加算することで過学習抑制を試みた。

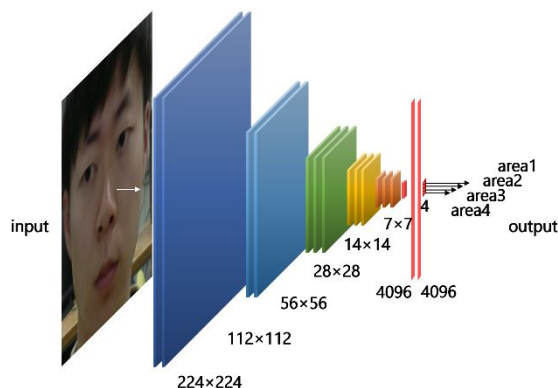


図3 VGG16 を用いた視線領域推定

4. 研究成果

最適化後の認識精度は、画面2分割で99.81%、4分割で95.18%、6分割で95.71%、9分割で87.88%となった。画面9分割においても高い認識精度が得られていることが確認できる。AlexNet ベースのモデルと比べると大幅に改善が図られた。図4は損失関数の推移を示しており、実線が訓練用データ、破線がテスト用データにおける結果である。画面2分割においては、テスト用データが訓練用データの損失関数を上回るものの極端な変化は見られず、適切な学習が行われていることが分かる。誤認識した画像を調べたところ、分割した画面の境界付近を注視しているものが多く含まれていたが、これらの画像は人が判断しても極めて難しいものばかりであった。また、正しい出力を得た画像の中には、頭部をほとんど動かさず目のみを動かしているもの、目をほとんど動かさず頭部のみを動かしているものの両方が含まれていた。このことから、CNNによって目領域の特徴量および facial appearance の学習が正しく行われ、視線領域推定に適切に反映されていることが確認された。

現在は、画面4分割の視線領域推定をリアルタイムで行えるシステムが完成しており、移動平均などの処理を加えることで、実用的に問題のない精度が維持されることを確認している。これにより、特殊なハードウェアを用いることなく、幅広い意思表示支援が可能になると考えられる。教育現場や家庭での集団データの蓄積が今後の課題である。

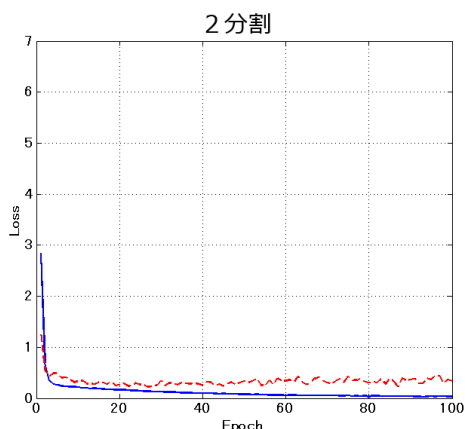


図4 損失関数 (学習:実線, テスト:破線)

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計5件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 Tatsuya Kato, Keigo JO, Koki SHIBASATO, Tetsuya HAKATA
2. 発表標題 Gaze Region Estimation Algorithm without Calibration Using Convolutional Neural Network
3. 学会等名 7th International Conference on Applied Computing and Information Technology (国際学会)
4. 発表年 2019年

1. 発表者名 城 啓悟, 柴里 弘毅, 大塚 弘文
2. 発表標題 Deep Learningを用いた視線領域推定 システムの構築と評価
3. 学会等名 電気学会九州支部 平成30年度（第9回）高専研究講演会
4. 発表年 2019年

1. 発表者名 城 啓悟, 大浦 稜平, 柴里 弘毅, 加藤 達也, 博多 哲也
2. 発表標題 キャリブレーションを必要としない視線領域推定アルゴリズムの構築
3. 学会等名 日本福祉工学会第22回学術講演会
4. 発表年 2018年

1. 発表者名 Masaki Kuwata, Koki Shibasato, Hirofumi Ohtsuka, Yasuyuki Shimada
2. 発表標題 A study on communication aid of face and eye motion
3. 学会等名 Proceedings of International Symposium on Innovative Engineering 2018 (国際学会)
4. 発表年 2018年

1. 発表者名 湊田雅輝, 博多哲也, 加藤達也, 大塚弘文, 柴里弘毅
2. 発表標題 単眼カメラを用いた視線方向推定のためのファジィ推論手法
3. 学会等名 日本福祉工学会 第21回学術講演会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	柴里 弘毅 (SHIBASATO KOKI) (60259968)	熊本高等専門学校・電子情報システム工学系C Iグループ・教授 (57403)	
研究分担者	加藤 達也 (KATO TATSUYA) (10707970)	熊本高等専門学校・電子情報システム工学系C Iグループ・助教 (57403)	