



視線操作の使いやすさ向上に関する研究

-第 2 報 瞬目の判別と個人差-

Research on improving usability of gaze control
- 2hd report: Blink detection and individual differences -

原田帆七海¹⁾, 高橋秀智¹⁾

Honami HARADA, Hidato TAKAHASHI

1) 東京工科大学 工学研究科 (〒192-0982 東京都八王子市片倉町 1404-1, takashihdtm@stf.teu.ac.jp)

概要: 視線操作の動作としてよく使用される瞬目(瞬き)は, 自然性瞬目と随意性瞬目の判別がつかないため, 誤操作や操作遅延に繋がるケースがある. これは, 瞬目を判別できる仕組みを導入することで解決することができ, 現状としてもその判別手段はある. ただ, 瞬きは個人差が大きい動作であるため, 瞬きの速さや開閉の大きさ等を考慮した個人に最適化された判別モデルを構築することで, 瞬目の判別精度を向上させる.

キーワード: ユーザインタフェース 瞬き

1. はじめに

アイトラッキング技術はその発展によって, 私達の身近な機器にも搭載されるようになり, VR や AR に限らず幅広いフィールドで視線操作という形で使用されている. その中で, 瞬きという動作は, 誰でも扱いやすく疲れにくいという特徴を持ち[1], 瞬き単独や手指の動きとの併用等と様々な操作手法として用いられている.

しかし, 瞬きを動作検出として利用する場合, 無意識的に行われる瞬きである自然性瞬目と意識的に行われる瞬きである随意性瞬目の判別がつかない場合, 誤操作や操作遅延に繋がる. それは, 私が以前行った実験でも確認され, やりやすさに対しても悪影響を及ぼしていた. これは, 瞬目を判別できる仕組みを導入することで解決することができ, 現状としてその判別手段はある. 例えば, 関連研究として, 佐藤らによる 2 種類の随意性瞬目と自発性瞬目を瞬目時波形から判別するというものがある[2]. そこでは, 約 7 万 5 千枚の画像を, CNN の学習データとして用い, モデルのマイクロ F 値平均は 90%以上だった. ただ, 学習データの量が膨大である事が懸念点の一つとして考えられた. 加えて, 瞬きは個人差が大きい動作である. その要因の例としては, 性差や年齢が確認されており[3], 動作としての違いとしても瞬きの速さや開閉の大きさ等と多岐にわたって挙がる. それを踏まえ本研究では, 少ない学習データで判別できるような学習方法の提案と個人の瞬きの特性を考慮し, より個人に最適化された判別モデルを構築することで, 瞬目の判別精度の向上を狙う.

2. 実験概要

スマートフォンを用いて日常的に使用するアプリケーションを操作してもらい, その様子から瞬目のリアルタイム判別を行う. 実験では, 瞬目の収集とモデルによる判別, 自然性瞬目か随意性瞬目かのラベル付けを同時進行で行い, 上記を一連とし, ①汎用モデルとなる共通モデル②対象のデータのみで学習した個人モデル③共通モデルと個人モデルを掛け合わせた複合モデルの 3 種のモデルを構築し, それぞれ実験を行っていく. また, 今回は学習に用いたデータと実験者の瞬きの特性が乖離したとしても機能するよう補正機能を実装する. そのため, モデル別と補正有無の計 6 つのパターンのデータを収集する.

2.1 実験環境

以下の機材・モデルを使用し, 実験様子を図 1 に示す.

- ・スマートフォン(被験者所有のもの) ・PC(Windows)
- ・web カメラ(DFK 33UX273) ・Dlib(瞬き検出用)[4]
- ・瞬目判別モデル(LightGBM ; 機械学習アルゴリズム)[5]



図 1. 実験の様子

2.2 評価指標・計測項目

実験で使用する3つのモデルには、共通した表1にあるa~hの8つの特徴量を用いて判別を行っている。

表1. 特徴量

項目名	意味	詳細説明
a duration_ms	瞬きの持続時間	瞬きの開始から終了までの所要時間
b ear_min	最小EAR	瞬きの閉じ具合
c ear_diff	EARの変化幅	EARの最大値-最小値で表される開閉の大きさ
d after_saccade	サッケード直後かどうか	サッケード(視線移動)があったかどうか
e gaze_shift	視線移動の距離	どのくらい視線移動を行ったか
f ear_integrated	EAR変動量の積分	瞬きの滑らかさ
g blink_speed	瞬き速度	EARの変化幅を瞬き時間で割ったもの
h closure_ratio	瞬きの閉じ具合の割合	1-(ear_min/ear_max)で表される瞬きの閉じ具合

表1にあるEARは目の縦幅の合計を目の横幅で割ったものであり、目の開き具合の定量指標となる。これは、Dlibライブラリ[4]で、図2のように目の端から順に輪郭に沿った点 p_1 p_2 p_3 p_4 p_5 p_6 を6点取得し、

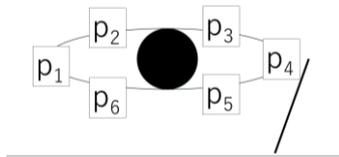


図2. 目のランドマーク

それらを用い、式(1)で求められる。

$$EAR = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2 \times \|p_1 - p_4\|} \quad (1)$$

また、モデルに適用する補正は、計算式(2)のように、使用するモデルの学習に使用した duration_ms の平均(M_0)を基準として掛けたものとなっている。実験を開始して20秒間で被験者から duration_ms の平均(M_1)を算出し、モデル学習時の平均と比較することで補正值(H)を決定する。実験時の方が大きければ値を小さく補正し、実験時の方が小さければ値を大きく補正する。

$$H = \frac{M_1}{M_0} \quad (2)$$

評価指標としては、自然性瞬目と随意性瞬目の判別精度を用いる。これは、瞬目全体の数をA、自然性瞬目の数を B_0 、自然性瞬目誤判断の数を B_1 、随意性瞬目の数を C_0 、随意性瞬目誤判断の数を C_1 とすると計算式(3)(4)(5)で求める。

$$accuracy_{nature} = \frac{(B_0 - B_1)}{B_0} \times 100 = D_0 \quad (3)$$

$$accuracy_{intended} = \frac{(C_0 - C_1)}{C_0} \times 100 = D_1 \quad (4)$$

$$accuracy_{entirety} = \frac{(D_0 \times B_0) + (D_1 \times C_0)}{2A} = D_2 \quad (5)$$

2.3 実験手順

[事前準備]

①姿勢調整

被験者はカメラが顔の正面に来るように椅子に腰掛け、顔の輪郭を損ねないようにスマホを利き手で持ち操作をする姿勢を整える。判定の手は、パソコンのスペースキーに指を置き、いつでも押せるよう待機する。

②微調整

実験プログラムを実行させ、瞬きが取得できるように、微調整を行う。その後、プログラムを一度終了させる。

[実験1]

実験プログラムを実行させ、実験を開始する(12人行う)。

①動作内容

被験者は、決定動作と定義した動作を行う際に、その決定動作と随意性瞬き、スペースキー(ラベル付けのための動作)、計3つの動作を同時に行う。

②時間

①を10分間行い続け、その後プログラムを終了させる。

[モデルの構築]

①共通モデルの構築

収集した瞬目の全ての学習データから汎用である共通モデルを構築。

②個人モデルの構築

一個人の学習データのみから個人モデルを構築。

③複合モデルの構築。

共通モデルを土台に再度一個人の学習データのみを学習させ複合モデルを構築。

[実験2]

瞬目データの収集とラベル付けに加えてリアルタイム判別をしていくプログラムを用いて実験を行っていく。基本的な流れは実験1と同様となる。ただし、実験時間は2分、試行回数3回とする。

3. 実験結果

20代前半男性5名の被験者に本実験を行った所、次の結果が得られた。

3.1 被験者特性

被験者らの実験における瞬目特性は、図3~図7で示す。a~hは、2.2 評価指標・計測項目の表1を参照し、青色の線が被験者の特性を、オレンジ色の線が被験者全体の特性の平均を表している。(ただし、図では比較するために各値のスケールを調整している。)

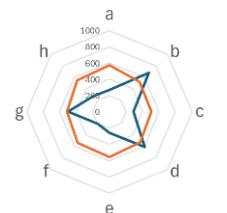


図3. 被験者Aの特性

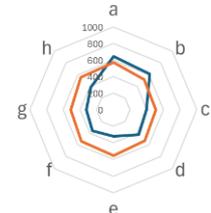


図4. 被験者Bの特性

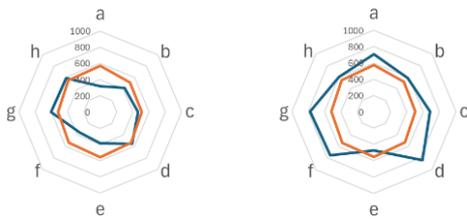


図 5. 被験者 C の特性

図 6. 被験者 D の特性

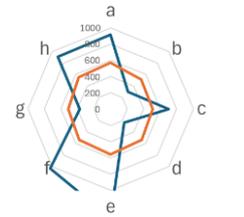


図 7. 被験者 E の特性

これより、被験者の特性はバラバラであり、5名の被験者間では共通特徴によるグループ分けができなかった。したがって、瞬きに関する個人差は大きいことが分かり、それは同じ年齢層、性別であっても変わらない事も分かった。

3.2 判別精度

モデル3種と補正有無を含めた6つのパターンにおける被験者全体平均の判別精度は、表2、各パターンにおける被験者ごとの判別精度は図8で示す。

表 2. 被験者全体の判別精度

平均	補正有	補正無し
共通モデル(1)	75.28	75.48
個人モデル(2)	77.86	77.51
複合モデル(3)	84.15	83.57

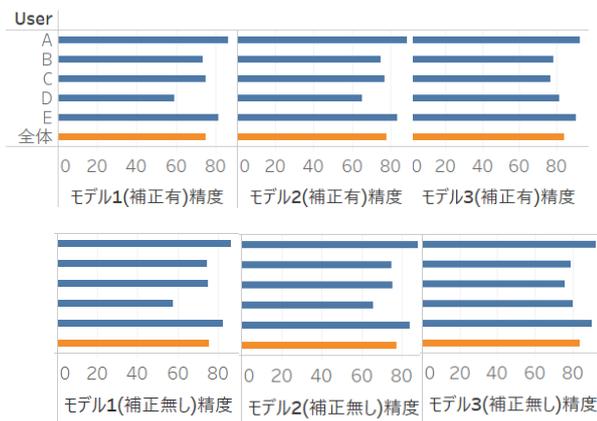


図 8. 各パターンにおける被験者ごとの判別精度

判別精度は、共通モデル→個人モデル→複合モデルの順に高いということ、補正有と補正無しでは差が無い事が分かる。

そこで、各モデルの補正有と無しについてt検定を行った所、全てのモデルにおいてp値が0.05を上回り、「補正の有無による判別精度には有意差がない」といえる。次に、補正有の方の3種のモデル間で分散分析を行なった所、p

値は $0.002 < 0.05$ となり「3種のモデル間の判別精度には有意差がある」といえる。

また、判別の内訳としては、共通モデルは自然瞬目の誤判別は少なく随意性瞬目は多い。個人モデルは自然瞬目の誤判別が少なく、随意性瞬目は共通モデルに比べ少ないがそれでも多い。複合モデルでは、自然性瞬目の誤判別が他2つと比較し僅かに多いが数回に収まる程度に少なく、随意性瞬目の誤判別はほどほどに少ない。

3.3 補正の強弱

実験で行う補正の強さを(2)式の2倍に変えてみた結果が、次の表3となる。

表 3. 補正の強弱による判別精度

平均	補正有	補正無し
共通モデル	86.17	85.00
個人モデル	84.95	84.95
複合モデル	77.18	81.88

基本的には表2と同様の結果が得られているが、複合モデルにおいて非常に精度が低下していることが分かる。自然性瞬きの誤判別が、全実験を通して最も多かった事が特徴として挙がる。

3.4 モデル評価

実験に使用したそれぞれのモデルにおいて、特徴量の重要度の順位を表4で示す。太字は上位4位に入っていることを示す。

表 4. 特徴量の重要度

特徴量	共通	モデル1_A	モデル2_A	モデル1_B	モデル2_B
ear_min	1	1	1	1	1
duration_ms	2	4	2	3	3
ear_integrated	3	2	4	6	3
ear_diff	4	3	3	3	4
closure_ratio	5	5	5	2	5
gaze_shift	6	6	6	6	6
blink_speed	7	7	7	6	7
after_saccade	8	8	8	3	8

特徴量	モデル1_C	モデル2_C	モデル1_D	モデル2_D	モデル1_E	モデル2_E
ear_min	4	1	1	1	3	1
duration_ms	1	2	3	2	1	2
ear_integrated	2	3	4	3	4	3
ear_diff	3	4	6	4	4	4
closure_ratio	5	5	2	5	2	5
gaze_shift	6	6	7	6	4	6
blink_speed	8	7	5	7	4	7
after_saccade	7	8	8	8	8	8

表4より、ear_min, duration_ms, ear_integrated, ear_diffの4つは重要度が高いといえる。

その4つに絞り、どのような傾向で瞬目の判別を行っていたのかは、まず、ear_minでは、値が大きいと自然性瞬目/値が小さいと随意性瞬目と判別していた。duration_msでは、値が小さいと自然性瞬目/値が大きいと随意性瞬目、ear_integratedでは、値が小さいと自然性瞬目/値が大きいと随意性瞬目、ear_diffでは、値が小さいと自然性瞬目/値が大きいと随意性瞬目と判別していた。

また、複合モデルの際の判別精度と各種の特徴量の相関係数を算出した所、表5のようになった。

表5. 判別精度と特徴量の相関係数

duration_ms	ear_min	ear_diff	after_saccade	gaze_shift	ear_integrated	blink_speed	closure_ratio
0.08	-0.02	-0.03	-0.25	0.41	0.20	-0.24	0.06

表5より、gaze_shiftには正の相関があり、ear_integratedとblink_speed, after_saccadeにはそれぞれ正と負の弱い相関がある事が分かる。

4. 考察

4.1 瞬目の個人差

今回実験を5名に行ったが、性別や年齢が同じ条件であったのにも関わらずその瞬目特性は非常にバラエティー豊かであった。被験者の数をさらに増やした場合、瞬目特性傾向によるグループ分けが可能になるかもしれないが、5名の対象のみでは難しかった。判別精度が平均よりも高かった被験者Aと被験者E、低かった被験者Bと被験者C、Dの2つのグループと瞬目特性を比較しても特に共通要素は見当たらず、各特徴量との精度の相関も全体的に弱い。特に被験者Aと被験者Eに関しては精度が高かったが、特徴量の傾向が真逆であった。したがって、ある要素が大きめだから精度も高くなるということが無いように瞬目特性の大きさは、判別精度への影響がきわめて軽微であると考えられる。ただし、モデルの判別基準には深く関係するため瞬目の個人差への対処が必要不可欠である点は変わらないと考える。

また、被験者Aと被験者Eにおける高精度判別は、学習データの量ではなく、学習データに含まれる自然性瞬目と随意性瞬目の割合などの学習データとしての質の影響が大きいと考えられるが、明確な要因は分からなかった。

4.2 補正

今回、瞬目の個人差への取り組みの一つとして瞬目データへの補正を行った。これは、学習モデル構築時と実験の際の瞬目データの値が大きく異なっただとしてモデルが機能するという意図だったが、結果として、判別精度の数値への寄与はなかった。しかし、影響がなかった訳ではなく、誤判別の内訳に影響を及ぼしていた。例えば、複合モデルにおいて、補正有りの場合は、随意瞬目と誤判別しやすく、補正無しの場合は、自然性瞬目と誤判別しやすかった。つまり、自然性瞬目の判別や随意性瞬目の判別と分けて捉えると、判別への寄与はあったと考えられる。

4.3 モデル

今回実験を行った所、複合モデルが最も精度が良く、次に個人モデル、共通モデルと続いた。

まず、共通モデルと個人モデルにおいては、個人モデルの方が学習データは少なかったのにも関わらず個人モデルの方が判別精度は良い。したがって、個人特化の学習データの方が、学習効率が良いと考えられる。

次に、複合モデルが最も精度が高かったことに関して、個人モデルは学習精度が高いものの学習データの多様性が乏しいことによる対応力の低さが共通モデルによって補完されたことが要因だと考える。

上記より、より精度の高い判別モデルを構築する際は、汎用モデルよりも個人に特化したモデルを組み込む方が良いと考えた。

また、視線操作のしやすさの大きな課題であった自然性瞬目の随意性瞬目との混同に対して、適切にアプローチできる可能性が示され、やりやすさへの貢献が期待できる。

5. おわりに

本研究では、瞬目判別モデルにおける個人差への対応の適用について、各モデルの比較検討を行った。結果より、個人差に対応した個人モデルの適用は、瞬目判別モデルの精度向上に寄与することが分かった。

6. 今後の課題

今後の課題としては、次の3点が挙げられる。

1点目は、学習データの増量であり、特に個人モデル等に使用する学習データを拡充させることができれば前述の対応力の低さの改善が見込める。ただ、モデルの良し悪しは、個人モデル作成の手間や多くの人に作製することを前提にした場合の最終的な学習データ量などの視点から総合的に判断する必要がある。

2点目は、実験方法の改善であり、現状の実験方法では各種の瞬きの数をコントロールできていない。今実験においては随意性瞬目の割合が10~30%と幅が広く、もっと画一的な条件での比較が求められる。例えば、随意性瞬目の値をコントロールし、自然性瞬目は必要な数を実験データからランダムに抽出するなどを考えている。

3点目は補正条件の改善であり、より適切な補正值の導出条件を考え、誤判別を引き起こすことがないように改善したいと考えている。

参考文献

- [1] 夏目達也, 内村裕也, クリスチャンアルサーテクルツ, 柴田史久, 木村朝子, “VR 空間操作コマンドとしてのアイジェスチャ UI の特性分析”, 情報処理学会論文誌, Vol.64, No.2, 312-325, 2023
- [2] 佐藤寛修, 阿部清彦, 松野省吾, 大山実, “畳み込みニューラルネットワークによる開眼閉眼状態の識別”, 電気学会電子・情報・システム部門大会講演論文集, ROMBUNNO.OS1-6, 2018.9.5
- [3] 杉山敏子, 田多英興, “成人における内因性瞬目の年齢差と性差”, 生理心理学と精神生理学, Vol.25, No.3, 255-265, 2007
- [4] Dlib, Boost, <https://dlib.net/>, 最終閲覧日 2025. 7. 18
- [5] LightGBM, Microsoft, <https://lightgbm.readthedocs.io/en/stable/>, 最終閲覧日 2025. 7. 18