

# 視運動性眼振に基づく注意切り替えのリアルタイム推定：情報入力への基礎的検討

清水 美羽<sup>\*,\*\*\*</sup>・金 成 慧<sup>\*\*,\*\*\*,†</sup>

<sup>\*</sup>筑波大学大学院理工情報生命学術院システム情報工学研究群知能機能システム学位プログラム 〒305-8573 茨城県つくば市天王台 1-1-1

<sup>\*\*</sup>東北大学総合知インフォマティクス研究センター 〒980-8577 宮城県仙台市青葉区片平 2-1-1

<sup>\*\*\*</sup>前所属：宇都宮大学工学部基盤工学科 〒321-8585 栃木県宇都宮市陽東 7-1-2

## Real-Time Estimation of Attentional Switching Based on Optokinetic Nystagmus: A Preliminary Study Toward Information Input

Miyu SHIMIZU<sup>\*,\*\*\*</sup> and Kei KANARI<sup>\*\*,\*\*\*,†</sup>

<sup>\*</sup>Degree Program in Intelligent and Mechanical Interaction Systems, Systems and Information Engineering, Graduate School of Science and Technology, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573

<sup>\*\*</sup>Advanced Institute of So-Go-Chi (Convergence Knowledge) Informatics, Tohoku University 2-1-1 Katahira, Aoba-ku, Sendai, Miyagi 980-8577

<sup>\*\*\*</sup>Former Affiliation: Department of Fundamental Engineering, School of Engineering, Utsunomiya University, 7-1-2 Yoto, Utsunomiya, Tochigi 321-8585

This study explores a novel information-input interface utilizing optokinetic nystagmus (OKN) as an indicator of visual attention. Unlike gaze-based methods, OKN can reflect attentional shifts without requiring stable fixation. We implemented a two-step approach: real-time estimation using a simple velocity threshold and offline analysis combining seven criteria to distinguish OKN from saccades and blinks. In experiments, real-time estimation achieved correct character input in 60.3% of trials. Offline analysis improved detection accuracy, yielding mean precision of 0.85, recall of 0.80, and F1 (the harmonic mean of precision and recall) of 0.81 across participants. These findings demonstrate that OKN enables reliable detection of attentional switching and provide a foundation for developing a fast, non-invasive communication system for users, including individuals with serve motor or speech impairments.

**Key words:** optokinetic nystagmus, attention switching, information input interface, communication support, visual attention

### はじめに

重度の身体障害や発話障害をもつ人々にとって、意志を伝達する手段の確保は、社会生活における重要な課題である。世界保健機関の報告によれば、何らかの障害を抱えている人は全世界人口の約15%を占めていると推定されており<sup>1)</sup>、その中には身体障害や感覚障害を抱えた人も含まれる。そのため、非言語的で身体運動を必要としない意思伝達手法の研究開発は、多くの人々の生活支援に貢献しうる重要な研究テーマである。

代表的な技術の1つであるブレインコンピューターイン

ターフェース (brain-computer interfaces: BCI)<sup>2)</sup> は、脳の神経信号を利用して情報入力を行うため、重度の障害をもつ人でも意思表示が可能である。しかし、多くのBCIは侵襲的方法を採用しているため、外科手術に伴うリスクや、電極の洗浄といったメンテナンスが必要である。さらに脳波信号はノイズの影響を受けやすく、同じユーザーであっても短時間で活動パターンが変化するため、継続的な学習が必要となる<sup>3,4)</sup> など、実用化には課題が残されている。

その他の手法として、眼球運動を利用した視線入力<sup>5)</sup> は非侵襲的かつ比較的低コストで利用可能である。しか

<sup>†</sup> kei.kanari.a2@tohoku.ac.jp

し、視線入力は固視を安定して維持する必要があるため、瞬きや不随意的なサッカードの影響で入力精度が低下することがある。例えば、不随意的なサッカードには、固視中にも自然に発生する微小サッカード (microsaccades)<sup>6)</sup> や、注意の逸れや外因的の刺激によって意図せず生じる大きなサッカードが含まれる。このように固視に依存する点は、より汎用的な利用を目指す上での課題であり、固視に依存せず利用者がどこに注意を向けているかを検出できる新たな手法の開発が求められている。

こうした流れの中で、近年、視線移動を用いない新しいインターフェースとして、視覚的注意に伴う瞳孔反応を利用した手法が提案されている<sup>7,8)</sup>。これらの先行研究では、瞳孔が視線位置だけでなく、視覚的注意を向けた対象の明るさに応じて変化するという知見<sup>9-11)</sup> を利用している。例えば、白背景上に文字が提示されている場合、その文字に注意を向けると瞳孔が収縮し、この変化を手がかりとして入力対象の文字を特定することができる。この特性を利用することで、視線移動を伴わない意思伝達が可能である。しかし、これらの手法には依然として解決すべき課題が残されている。例えば、Mathôtら<sup>7)</sup> の研究では、注意対象の明るさに応じた瞳孔収縮を利用することで視線移動を伴わない文字入力が可能であることが示されたものの、瞳孔反応の潜時が数100ミリ秒から1秒程度と遅いため、入力速度が制限されることが指摘されている。また、Mutoら<sup>8)</sup> は輝度を周期的に変調させた刺激に対する瞳孔径の同期の変動を利用する手法を報告したが、利用できる周波数帯が低いため、入力文字候補数を増やすと確定までに要する時間が長くなるという制約があった。さらに、いずれの方法も周囲光条件や利用者ごとの反応特性に依存するため、環境や個人差に左右されず安定した入力性能を実現することは容易ではない。

そこでわれわれは、瞳孔反応よりも潜時が短く、迅速な注意検出が可能とされる視運動性眼振 (optokinetic nystagmus: OKN) に着目した。OKNは持続的な運動刺激に対して誘発される眼球運動であり、運動物体を追うように低速で動く緩徐相と、その反対方向に高速で戻る急速相が交互に繰り返される特徴をもつ。先行研究では、OKNが視覚的注意と密接に関連し、視線位置と注意位置が異なる場合でも注意を向けた運動刺激に対応して発生すること<sup>12)</sup>、また運動方向の異なる刺激が同一位置に重ねて提示された場合でも、注意を向けた刺激の方向に応じて選択的に誘発されること<sup>13)</sup> が報告されている。これらの知見に基づき、注意を切り替えたタイミングを推定し、入力文字の推定を可能にした<sup>14)</sup>。

しかし、アイトラッカーで測定したOKNはサッカードや瞬きなどと形態が類似しており、リアルタイムで高精度に識別することは容易ではない。また、OKNの検出に関する既存研究<sup>15-17)</sup> の多くはオフライン解析を前提とし、緩徐相の0交差や回帰により、広視野・一定速度刺激下で生じる持続的で整ったOKN (安定した緩徐相と急速相) を同定する手法が中心である。これらは十分な区間長と安定した速度推定を前提とする。一方、本研究が対象とするのは、ターゲットが特定方位に到達した瞬間の注意切り替えに伴って生じる短持続の一過性OKNであり、出現は断続的・非定常で、かつ局所的な空間セクター内での短時間トリガーをその場で確定する必要がある。このため、先行手法をそのまま適用すると、短窓では緩徐相推定が不安定となり (遅延・誤判定の増加)、サッカード・瞬きの混入にも脆弱となる。

以上の課題を踏まえて、本研究の目的は、OKNを用いた注意切り替え検出を情報入力インターフェースへ応用するための基盤技術として、OKN急速相に基づく単純なオンライン指標による注意切り替え推定の有効性と限界を定量化するとともに、オフライン同定により推定精度がどの程度改善可能かを検証し、今後の高精度リアルタイム化に向けた課題を整理することである。この目的を達成するため、本手法ではリアルタイム推定 (オンライン) とオフライン同定を組み合わせた2段階設計を採用した。まず第1段階 (オンライン) では、予備実験でOKN急速相速度が全参加者で100 deg/s以上となることを確認した知見に基づき、リアルタイムで取得した眼球位置の速度が100 deg/sを超えた時刻を即時に検出・記録した。これは、先行研究<sup>14)</sup> が緩徐相の0交差を事後に検出して入力時刻を定義したのに対し、急速相の速度閾値による即時検出を用いる点に特徴がある。なお、当該時刻が入力有効範囲に含まれるかの判定は記録データのオフライン解析で実施した。続く第2段階 (オフライン) では、取得データに対して時間・速度・加速度 (周期や交互パターンを含む) に基づく7指標を統合し、瞬きやサッカードによる誤判定を抑えてOKNを高精度に同定することで、オンライン推定の限界を補完しうる上限性能を評価する。さらに、本オフライン手法はラベル付きデータセットの構築にも活用でき、将来的な機械学習モデルの学習などを通じた高精度リアルタイム推定の実現に資する可能性がある。また、本研究で検討したOKN同定手法は、将来的に非侵襲的な意思伝達インターフェースにおける注意切り替え推定の精度向上に活用できる可能性がある。ただし、システムとしての高速化および信頼性の達成は本稿の範囲外であり、今後の課題とする。

## 1. 実験方法

### 1.1 参加者

予備実験には、19～22歳の大学生7名（うち女性6名、平均年齢19.8歳）が参加した。本実験には、予備実験に参加した3名を含む18～22歳の大学生8名（うち女性6名、平均年齢19.5歳）が実験に参加した。なお、本実験参加者8名のうち2名は前処理後の眼球位置信号における欠損率（NaN サンプル数 / 総サンプル数）が35%以上であったため解析から除外し、最終的に6名のデータを解析対象とした。ここで欠損は、瞬きやトラッカーロストなどにより有効値が得られなかったフレームをNaNとして符号化したものを指す。35%という閾値は、本装置・前処理系での既存研究<sup>13,14</sup>に基づく経験的基準である。すべての参加者は1.0以上の視力（矯正視力を含む）を有していた。参加者は全員、書面によるインフォームドコンセントを得たうえで実験に参加した。本研究は、宇都宮大学ヒトを対象とした研究に関する倫理審査委員会の承認を受けて実施された（登録番号：H22-0022）。

### 1.2 装置

実験は照度65 lxの部屋で行われた。刺激は液晶ディスプレイ（EIZO ColorEdge CS2420-Z, 1920×1200 pixels, 51.84×32.4 cm, 48.9×31.7°）にリフレッシュレート60 Hzで提示された。参加者とディスプレイとの視距離は57 cmで、参加者の頭は顎台によって固定された。刺激の作成や提示には、数値解析用ソフトウェアMATLAB（R2017b, MathWorks, Inc.）と心理物理学実験用ソフトウェアPsychophysics Toolbox<sup>18,19</sup>を備えたMacBook Pro（macOS Sierra, 10.12.6, Apple）を用いた。眼球位置はビデオベースのアイトラッカーiRecHS2<sup>20</sup>を用いて、サンプリングレート500 Hzで測定した。

### 1.3 刺激

実験刺激の例をFig. 1に示す。画面中央に黒い固視点（直径0.54 deg, 輝度0.26 cd/m<sup>2</sup>）、右方向に運動する緑色（RGB [0, 255, 0], 輝度153.33 cd/m<sup>2</sup>）のランダムドットパターン（直径0.54 deg, 密度3 dot/deg<sup>2</sup>, 速度16.2 deg/s）、上方向に運動する青色（RGB [0, 0, 255], 輝度9.08 cd/m<sup>2</sup>）のランダムドットパターン（直径0.54 deg, 密度3 dot/deg<sup>2</sup>, 速度16.2 deg/s）が半径5.4 degの円内にだけ提示された。固視点を中心とした半径12.15 degと6.48 degの2つの円周上にローマ字が配置された。外円には五十音の各行（あ行・か行・さ行…）に対応するローマ字の頭文字「a」「k」「s」「t」「n」「h」「m」「y」「r」「w」が36°間隔で配置され、速度30π/180 rad/sで右回転した。内円にはローマ字入力の母音に対応する「a」「i」

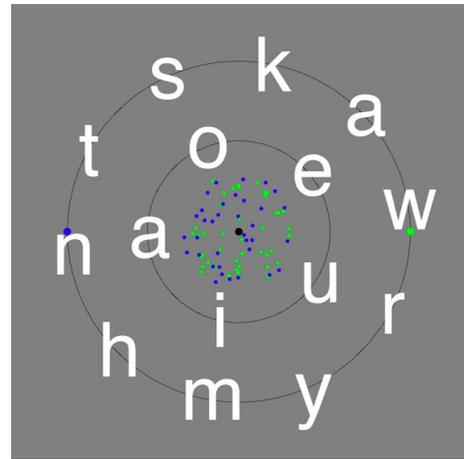


Fig. 1 Example of stimuli used in the experiment. The input characters were arranged with consonants on the outer circle and vowels on the inner circle. At the center, a fixation point was presented together with a moving dot pattern. On the outer circle, a green dot and a blue dot were displayed.

「u」「e」「o」が72°間隔で配置され、速度60π/180 rad/sで左回転した。各文字の色は白（RGB [255, 255, 255], 輝度222.33 cd/m<sup>2</sup>）だった。外円の3時方向と9時方向にはそれぞれ緑点と青点が提示された。本実験では外円文字での検証を行い、緑点に入力したい文字が到達した際に、固視点から緑色の運動ドットに注意を切り替えた。一方、青点は一般化可能な二重リング型UI（user interface）を想定し、緑点の反対側に注意切り替え点を用意するデザイン要素として配置した。将来的には、青点位置にターゲットが到達した際に上方向ドットへの注意切り替えでも入力できる拡張を想定するが、本研究の評価対象は緑点のみであり、青点に関連する上方向ドットの運動は解析対象外とした。なお、内円はUI構造の提示を目的とした表示であり、ターゲット文字としては用いず評価対象外とした。背景はグレー（RGB [128, 128, 128], 輝度49.26 cd/m<sup>2</sup>）だった。

### 1.4 手順

実験の1試行の手順をFig.2に示す。試行の最初に教示文が提示された（Fig. 2 (a)）。参加者がテンキーを押すことで、入力画面が提示された。始めに、参加者は中央の固視点を凝視した（Fig. 2 (b)）。その後、指定された入力文字が緑の点に到達したら注意を右方向に運動するドットに切り替えた（Fig. 2 (c)）。そして、指定文字が緑の点を過ぎたら再び固視点を凝視した（Fig. 2 (d)）。文字は必ず「a」が0時の位置から回転を開始した。ターゲットとなる外円のローマ字の頭文字10条件（「a」「k」「s」「t」「n」「h」「m」「y」「r」「w」）に対して、各3回繰り返した。ターゲットとなる文字はランダムに選ばれた。

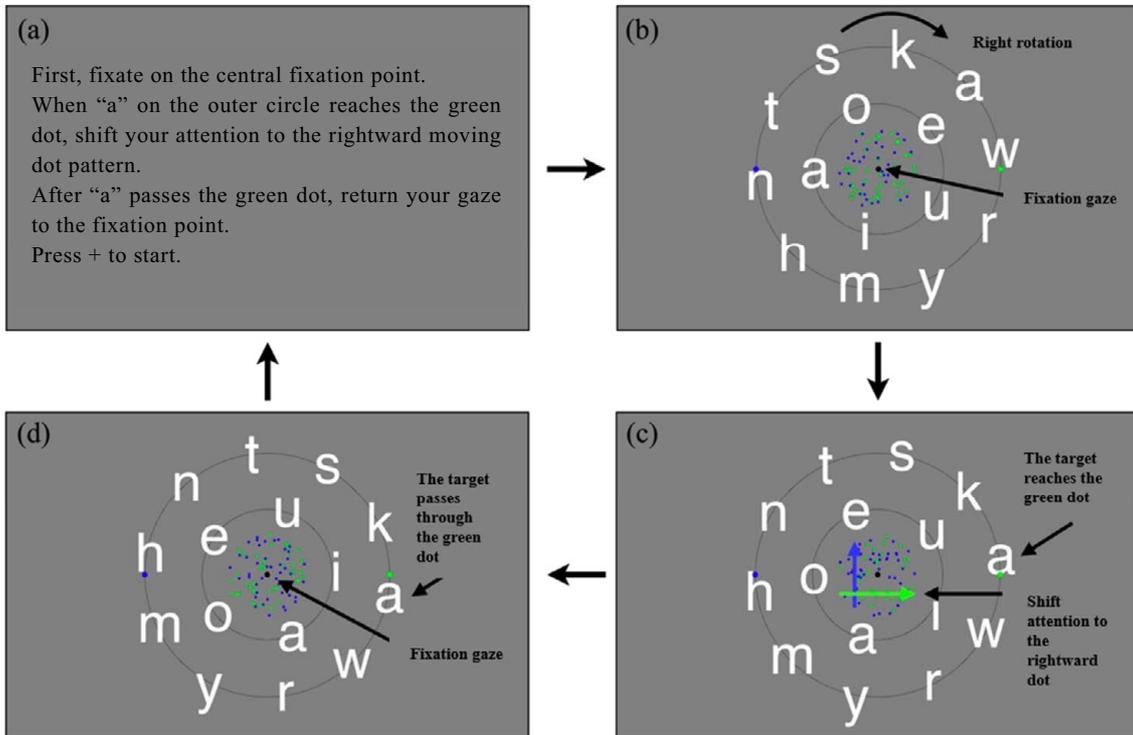


Fig. 2 Procedure of a single trial in the experiment. (a) At the beginning of each trial, an instruction text was presented. (b) Participants were required to fixate on the central fixation point. (c) When the target reached the green dot, participants shifted their attention from the fixation point to the rightward moving dot pattern. (d) After the target passed the green dot, they were instructed to return their gaze to the fixation point.

### 1.5 リアルタイム推定

先行研究<sup>14)</sup>が緩徐相の0交差を事後に検出して入力時刻を定義したのに対し、本研究では、急速相の速度閾値(100 deg/s)によりその場で注意切り替えの時刻を検出・保存することを設計の前提とした。

iRecHS2で測定された水平方向の眼球位置(deg)のデータをDAコンバーター(AO-1608L-LPE, CONTEC)を用いて電圧としてリアルタイムに出力した。出力データはデータロガー(LabJack U3-LV)を通して解析用のPCに入力し、MATLABを用いて眼球位置(deg)を微分することで速度(deg/s)に変換した。本実験と同様の刺激、手順で予備実験を行った結果、注意を向けたときに発生したOKNの急速相速度はすべての参加者で、100 deg/s以上であることが明らかになった。そこで本実験では、100 deg/s以上のOKN急速相速度が出た場合、注意を固視点から運動ドットに切り替えたと定義した。固視点と文字の中心を結ぶ線を基準とし、そこから両側±18°の範囲をその文字が入力できる有効範囲とした(Fig. 3)。OKNから推定された入力時間とターゲット文字が実際に緑の点に到達するまでの時間の差を分析することで、入力文字の推定精度を検討した。

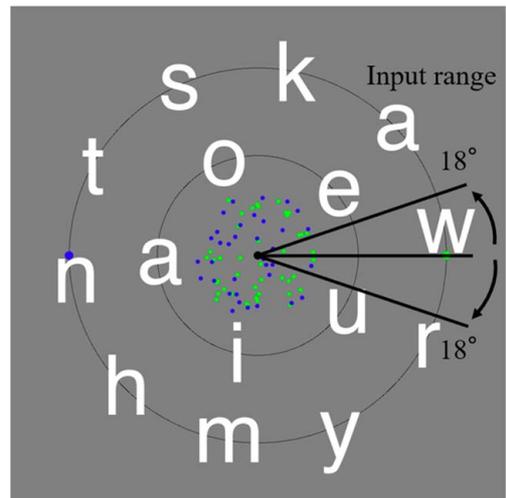


Fig. 3 Character input range. The input range of a character was defined as  $\pm 18^\circ$  on either sides of the line connecting the fixation point and the character center.

### 1.6 オフライン分析

リアルタイム入力を実現するためには、瞬きによるノイズやサカードなどの眼球運動とOKNとを切り分ける必要がある。したがって、注意の切り替えタイミングをより正確に推定するためには、1.5節(リアルタイム推定)の

ように眼球運動の速度情報のみを利用するだけでは OKN を十分に検出することができない。そこで、OKN とその他の眼球運動を区別するため、OKN の特性を検討した。リアルタイム推定の際に得られた眼球運動データをもとに、次の段落で示す 7 つの指標を用いてオフラインで OKN を検出した。まず、視線位置を目視で確認し、緩徐相と急速相が交互に出現するパターンをもとに、各試行における OKN の発生を判断した。なお、OKN の有無に対する判断は、各参加者において全試行の波形特徴（振幅など）を参照しながら、一貫した基準に基づいて行った。これを基準として、検出手法によって抽出された OKN 候補との一致を評価し、true positive (TP), false positive (FP), false negative (FN) を算出した。さらに、検出精度の定量的評価として、目視で判定された OKN との一致に基づいて、precision (適合率), recall (再現率) および F1 スコア (調和平均) を算出した。各評価指標における閾値は、誤検出による誤入力のリスクを低減することを優先し、検出結果の適合率が最大となるよう参加者ごとに最適化して設定した。

OKN の検出にあたり、先行研究<sup>15,16)</sup>に基づいて 7 つの評価指標を設定した。以下にその概要を示す。

- [1] 眼球運動の加速度：加速度の閾値を設定し、眼球位置変化の方向の切り替わり点を検出した。これは、OKN に特徴的な緩急相と急速相の切り替わり点を検出するための指標である（閾値：5000~25000 deg/s<sup>2</sup>）。
- [2] 加速度ピーク間の時間差：[1] で検出された切り替わり点間の時間間隔に着目した。OKN では、この間隔が比較的一定であるのに対し、サッカーボールなどは突発的に発生するため時間差が不規則になる（閾値：200~766.66 ms）。
- [3] 緩徐相と急速相の連続的発生：OKN では緩徐相と急速相が連続して発生することから、その連続出現を評価指標とした。多くの参加者では閾値を 3 回以上に設定したが、一部の参加者では 1 回のみ出現にとどまったため、その場合は閾値を 1 回以上に設定した。
- [4] 緩徐相と急速相の切り替わり眼球位置の比較：緩徐相から急速相への切り替わり点と急速相から緩徐相への切り替わり点における眼球位置を比較した。緩徐相と急速相が繰り返し発生する場合、本実験では刺激が右方向に運動していたため、前者は後者より右側に位置するはずであり、この整合性を評価した。
- [5] 緩徐相速度と急速相速度の比較：急速相速度が緩徐相速度より大きくなることを利用した。
- [6] 眼球位置変化の線形性：緩徐相・急速相はともに線形

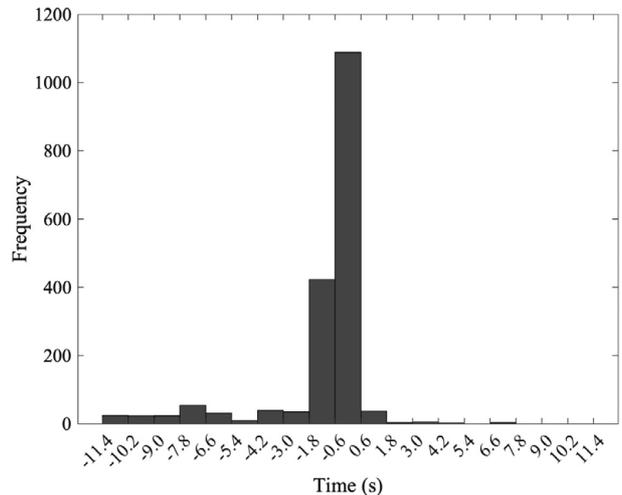


Fig. 4 Results of real-time estimation. The horizontal axis represents time, and the vertical axis indicates the number of data samples. A value of 0 s corresponds to the moment when the estimated input time coincided with the time the center of the target character reached the green dot. An observation within the interval of  $-0.6$  to  $0.6$  s was regarded as evidence that the attentional shift occurred within the effective input range of the target character.

であると仮定し、各区間に線形回帰を適用して近似直線を算出し、そこから残差を用いて線形性を評価した。

- [7] 眼球位置の逆方向変化の検出：OKN では緩徐相・急速相の途中で逆戻りが生じないことを前提に評価した。

## 2. 結 果

### 2.1 リアルタイム推定の結果

Fig. 4 に、全参加者における OKN から推定された入力時間とターゲット文字が実際に緑の点に到達した時間との差の結果を示す。横軸は時間を示し、0 s は推定された入力時間とターゲット文字の中心が緑の点に到達した時間が一致していることを示す。 $-0.6 \sim 0.6$  s の範囲に入っていれば、ターゲット文字の有効範囲内で注意の切り替えが行われていたことを示す。横軸は頻度を示す。結果として、最頻値の区間は  $-0.6 \sim 0.6$  s であり、その割合は参加者ごとに 42.9% (109/254), 62.4% (267/428), 50.9% (140/275), 89.2% (148/166), 54.9% (179/326), 68.9 (246/357) であった。全参加者では 60.3% (1089/1806) であった。次に頻度が高かった区間は  $-1.8 \sim -0.6$  s であり、その割合は 23.4% (423/1806) であった。また、OKN 検出時刻 ( $t_{\text{detect}}$ ) とターゲット到達時刻 ( $t_{\text{target}}$ ) を最近傍マッチングにより対応付け、許容窓 ( $\pm 0.6$  s) 内のペアについて検出時刻誤差 (timing error) を  $\Delta t = t_{\text{detect}} - t_{\text{target}}$  と定義し

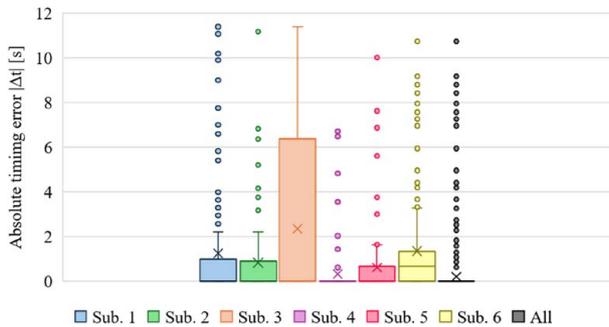


Fig. 5 Distribution of timing error  $\Delta t$ .  $\Delta t = t_{\text{detect}} - t_{\text{target}}$ . Boxes show median and IQR (25th-75th percentiles). Whiskers extend to the most extreme data points within  $1.5 \times \text{IQR}$  from the quartiles; points beyond are plotted as outliers.

た、全参加者を通した  $|\Delta t|$  の中央値は 0.00 s, interquartile range (IQR: 四分位範囲) は 0.90 s であった (Fig. 5). 参加者によってはばらつきが大きい例も認められた。さらに、全参加者の平均絶対誤差 (MAE) は 0.19, 二乗平均平方根誤差 (RMSE) は 0.99 であった。

## 2.2 オフライン分析の結果

各参加者における precision (適合率) と recall (再現率) を Table 1 に示す。precision の全参加者の平均値は 0.85 ( $SD = 0.011$ ), Recall は 0.80 ( $SD = 0.040$ ), F1 スコアは 0.81 ( $SD = 0.026$ ) であった。Fig. 6 にオフライン解析の例を示す。上段の 2 つのパネルは眼球運動の加速度を、下段の 3 つのパネルは眼球位置を示す。横軸は時間、縦軸は加速度または眼球位置を示し、眼球位置は縦軸において値が増加する方向が右方向に対応する。図中の点は、OKN として検出されたデータを示す。18.5 s 付近までの区間は文字入力試行に入る前の眼球運動であり、教示文を読む際の眼球運動やノイズが含まれていたが、OKN とは区別されていた。その後の区間でも、瞬きに起因すると考えられる眼球位置の大きく急激な変化が観察されたが、これも意図通りに除外されていた。一方、OKN として検出されたデータでは、右方向の緩徐相と左方向への急速相が確認され、OKN が適切に意図通り検出できていることがわかる。以上より、本研究で設定した 7 つの指標によって、OKN とその他のノイズや眼球運動を効果的に区別できていることが示された。

## 2.3 垂直成分 OKN のオフライン解析

本研究の評価対象外であった垂直方向のドットパターンに対応した OKN の発生有無を確認するため、水平方向の OKN 検出に使用したものと同一の時系列データで追加解析を行った。瞬き除去後の信号に平滑化を施し、時間微分により速度を算出し、さらに速度の時間微分から加速度を算出した。評価指標と閾値は水平方向と同一とし、(i) 加速度ピーク検出、(ii) 周期制約 (0.20~0.66 s)、(iii) 山-谷-山 (または谷-山-谷) の交互パターンと幾何条件 (中央点が両側より  $\geq 0.1$  deg 高い)、(iv) 方向一貫性 (偏差  $\leq 1.5708$  rad) を適用した。極性判定のみ垂直用に置き換え、上向き OKN を緩徐相が正 ( $>0$ )、急速相が負 ( $<0$ ) と定義した。次に、水平成分 OKN と中心時刻の一致 ( $\pm 0.30$  s) および区間オーバーラップ ( $\pm 0.10$  s) により、同時発生を照合した。

これらの結果を Table 2 に示す。結果より、水平方向 OKN との同時発生の系統性はいずれの参加者においても確認されず、ターゲット文字入力時に垂直方向ドットパターンが入力を阻害する兆候はみられなかった。一方で、試行全体では垂直成分 OKN の発生が参加者あたり 2~31 回 (平均値 13.3 回) 発生した。

## 3. 考察

注意切り替えのリアルタイム推定では、急速相の速度のみに基づいて OKN を検出し、そのタイミングから指定した入力文字の範囲内に正しく入っているかどうかの評価を行った。その結果、発生した OKN が 60.3% の割合で正しい文字入力内に入っていた。これより、急速相の速度のみに基づいた場合、入力文字の推定はある程度有効であるが、約 40% は誤入力してしまう可能性があることが示唆された。このことは、速度情報のみを指標とする方法では注意切り替えの推定精度に限界があり、より信頼性の高い推定のためには追加の指標を組み合わせる必要があると考えられる。

リアルタイム推定における OKN の検出精度が低かった理由として、瞬きによるノイズや単純なサッカードなどの眼球運動と OKN 急速相との区別が考慮されていなかったことが考えられる。そこで、オフラインでの OKN の検出

Table 1 OKN detection results obtained from offline analysis. Precision, recall, and F1 score are shown for each participant as well as the average across participant.

	Sub. 1	Sub. 2	Sub. 3	Sub. 4	Sub. 5	Sub. 6	Average
Precision	0.70	0.93	0.91	0.81	1.00	0.76	0.85
Recall	0.58	0.93	1.00	0.48	0.97	0.84	0.80
F1 Score	0.63	0.93	0.95	0.56	0.98	0.78	0.81

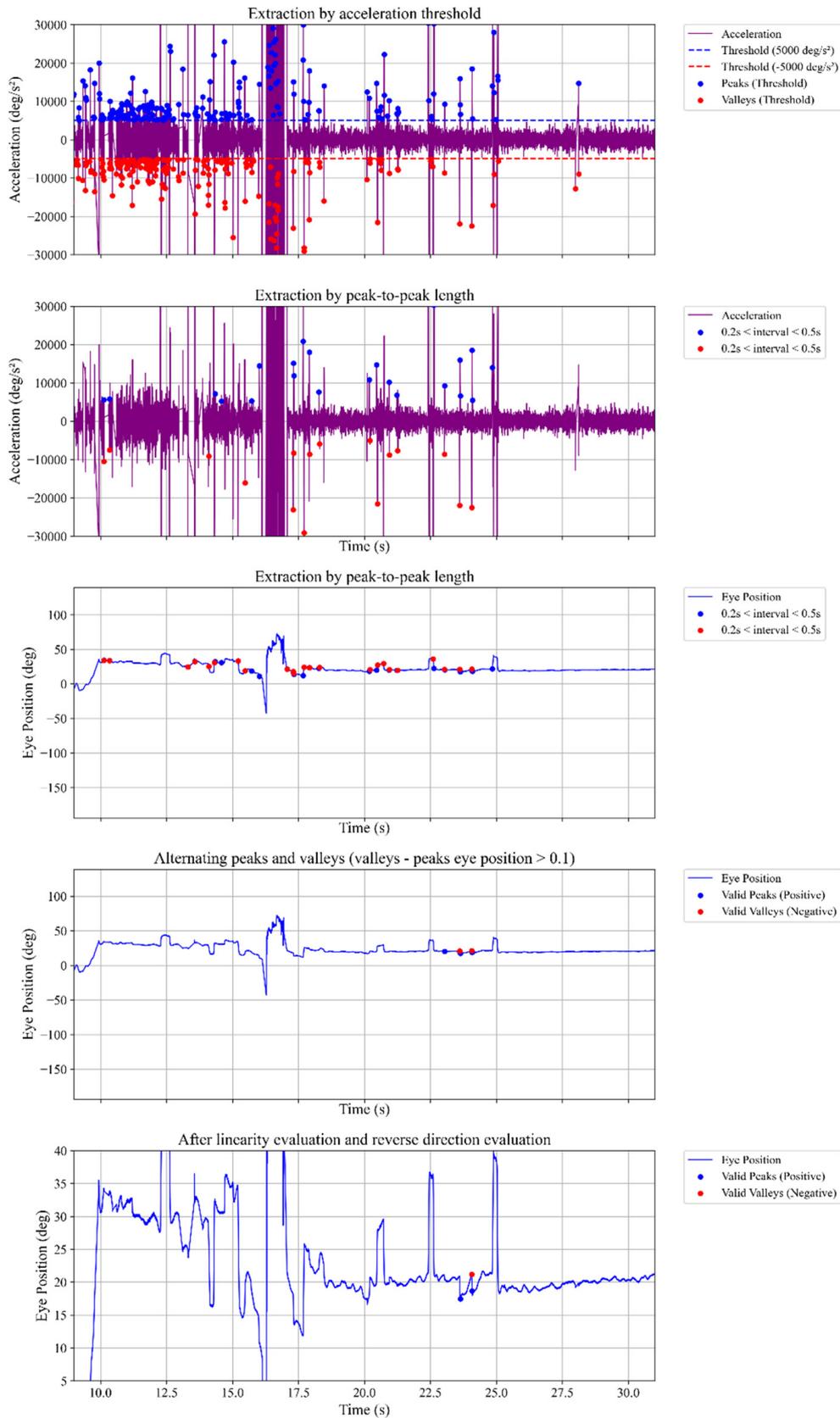


Fig. 6 Example of offline analysis. The horizontal axis represents time, and the vertical axis represents either acceleration or eye position. The first panel shows the data extracted using the acceleration threshold. The second panel presents the acceleration data further filtered according to the criterion described in [1]. The third panel illustrates the corresponding eye position. The fourth panel shows the eye position data obtained by evaluating the sequential occurrence of slow and fast phases. The fifth panel presents the data finally identified as OKN, based on the linearity of eye position and changes in eye movement direction.

Table 2 Vertical OKN detections and co-occurrence with horizontal OKN. Counts of vertical OKN per participant and co-occurrence with horizontal OKN (timing match  $\pm 0.30$  s, overlap  $\geq 0.10$  s). Extraction followed the same criteria as horizontal OKN.

	Sub. 1	Sub. 2	Sub. 3	Sub. 4	Sub. 5	Sub. 6	Average
Detection count	31	13	7	9	2	18	13.3
Match count	0	0	0	0	0	0	0.0

では、先行研究<sup>15,16)</sup>を参考に7つの指標を用いて検討した。その結果、全参加者における検出精度の平均値は precision (適合率) で 0.85 ( $SD = 0.011$ ), recall (再現率) で 0.80 ( $SD = 0.040$ ), F1 スコア (調和平均) で 0.81 ( $SD = 0.026$ ) であった。これより、リアルタイム推定で誤検出されていた眼球運動は検出されにくくなり、約 80% の精度で OKN を正しく検出できることが示された。また、同程度の割合で注意切り替えのタイミングを見逃すことなく検出できることも確認された。一方で、一部の参加者では recall が約 50% にとどまった。データを確認したところ、眼球運動が線形的でなく一部で停滞していたり、緩徐相の速度が一定でなかったりする例がみられた。リアルタイム推定の結果、参加者の OKN 応答は実際に入力文字が到達する前に生じており、これは先行研究で報告された瞳孔反応の早期出現と一致していた<sup>14)</sup>。その要因として、次の文字の到達を予期できたことが考えられる。このような予期的応答は、注意切り替え検出における時間ずれを発生させる要因となるため、今後の推定精度向上には、入力範囲の補正や文字を決定する範囲でのみ提示する工夫などが必要である。

垂直成分 OKN のオフライン解析では、水平成分の OKN のリアルタイム推定の際と同様の評価指標・閾値で OKN を検出した。その結果、試行中の垂直成分 OKN の検出数は平均 13.3 回、水平成分 OKN との同時発生数はいずれの参加者でも 0 回であった。同時発生が確認されなかったことから、水平方向におけるターゲット文字入力時に垂直方向のドットパターンが入力を阻害することはなかったと考えられる。一方で、いずれの参加者でも試行中に垂直成分 OKN が発生していたことから、意図していないタイミングで垂直方向の文字入力が行われてしまう可能性を示唆している。そのため、垂直方向のドットパターンと水平方向のドットパターンの速度に差をつけて解析の際にそれを考慮する方法など対策をする必要がある。

本研究のオフライン推定における検出精度は、precision および recall でいずれも 8 割以上であり、これは機械学習を用いた Juhola ら<sup>16)</sup>の手法と同程度の水準であるといえる。さらに、先行研究がオフライン解析を前提としていたのに対し、本研究ではリアルタイム推定とオフライン分析

からなる 2 段階アプローチを提案し、リアルタイムな文字入力に OKN 検出を応用可能であることを示した。また、Essig ら<sup>17)</sup>のリアルタイム化を目指した研究では、OKN 発生の有無を刺激提示中に判定するのみであり、発生タイミングの検出までは行っていない。これに対して本研究では、OKN の発生タイミングを検出し、それに基づいて文字入力を可能とした点、さらに運動刺激が提示されていない教示画面などの区間においても、瞬きやサッカードを誤って OKN と判定せずにおおむね正しく分類できた点に優位性がある。

瞳孔反応ベースの注意入力に関する先行研究<sup>7,8)</sup>と比較すると、本研究の OKN を即時トリガーとして用いるアプローチは、短潜時および時間分解能の観点から有効と考えられる。瞳孔反応は有用である一方で、潜時が長い(数 100 ミリ秒から 1 秒)、時間分解能が低く連続入力に不向き、照明・輝度変動の影響を受けやすいという特性が知られている。これに対し、本研究の OKN トリガーは、急速相速度の閾値により短潜時かつ高い時間精度で注意切り替えの時刻を即時に与えられ、さらに運動方向・方位情報と直結することから、高速な逐次入力や候補の迅速な絞り込みに適する。実験結果として、10 候補 ( $chance = 10\%$ ) に対しリアルタイム正答率 60.3% を示し、オフライン同定では  $F1 = 0.81$  まで高精度化できた。一方で、瞳孔入力の代表例では、4 択で 91.0% (約 20.2 s/ 選択, ITR (information transfer rate) 4.55 bits/min), 8 択で 87.6% (約 28.0 s/ 選択, ITR 4.86 bits/min)<sup>7)</sup>, 12 候補構成で平均 85.4% (条件により 87.5%, 一部条件で 62.5%)<sup>8)</sup> が報告されており、高精度だが判断窓は比較的長い傾向がある。以上より、瞳孔法と OKN 法は相補的であり、実用上は OKN (即時トリガー) と瞳孔 (確定) のハイブリッド構成が有望である。

本研究にはいくつか限界がある。第一に、OKN の検出率を向上させるためには、参加者ごとの個人差を考慮する必要がある。特に、参加者 4 および参加者 6 においては、OKN が連続的に発生しないケースが多くみられた。また、ほとんどの参加者において、データの一部では緩徐相が十分に線形でなく、明瞭な OKN が観察されない例も確認された。これらの問題に対応するためには、まず各参加者に最適化された刺激条件を見出し、検出しやすい OKN

を誘発することが求められる。第二に、本手法の2段階目のアプローチであるオフライン分析では、OKNの7つの指標を用いて80%程度の検出率を達成したが、処理はオフラインであった。実際の使用環境でリアルタイムに同等の検出率を実現するためには改善が必要である。特に、[2]加速度ピーク間の時間差や[3]緩徐相と急速相の連続的発生といった指標は、複数フレームにわたるデータを参照する必要があるため、短時間での判定が難しい。

これらの課題に対しては、機械学習の導入が有効である可能性が高い。機械学習を用いれば、速度、振幅、加速度、連続性など複数の特徴量を統合的に扱うことができ、個人差やデータのばらつきに対しても頑健な検出が期待できる。また、時系列モデルを利用すれば、短い時間窓でも緩徐相と急速相の依存関係を学習できるため、従来オフラインでしか捉えられなかったパターンをリアルタイムに検出できる。さらに、瞬きやサッカードなどOKNと形態が類似するノイズに対しても、多次元的な特徴量の組み合わせに基づき高精度に識別できると考えられる。リアルタイム推定の解析において、OKN検出時刻とターゲット文字到達時刻の差分を箱ひげ図に示すことでデータの個人差の可視化と定量化を行った。結果からも、参加者によってデータのばらつき度合いに差があった。一般化に向けては、頭部非固定や環境変動を含む条件へ段階的にデータ収集を拡張し、当面は個別最適化と転移的適応により個体差・条件差への適応を図る計画である。まず参加者層と計測条件を広げた小規模コホートで妥当性を再確認し、必要に応じて照明変動やキャリブレーションずれなどの攪乱を系統的に導入して外的妥当性を検証する。なお、本稿で用いた短窓の因果特徴は、そのままリアルタイム実装へ展開可能である。今後はこうした学習の手法とデータ拡張を併用することで、個人差やリアルタイム処理の制約を段階的に克服し、高精度かつ実用的なOKNベースの意思伝達システムの構築を目指す。

## まとめ

本研究では、視運動性眼振（OKN）を指標とする情報入力インターフェースの実現可能性を検討した。実験では、視覚的注意の切り替えにより発生するOKNをリアルタイムで検出し、入力文字の推定に応用した。その結果、急速相の速度に基づいたリアルタイム推定では、60.3%の割合で正しく文字入力を行えることが示唆された。一方で、瞬きやサッカードなどによる誤検出が精度低下の要因となった。そこで、オフライン解析では、7つの指標を組み合わせてOKNと他の眼球運動を区別したところ、全参

加者平均でprecision 0.85, recall 0.80, F1スコア 0.81を達成し、高精度な検出が可能であることを確認した。

しかし、個人差により一部の参加者では検出が安定しない例もみられ、またオフライン処理に依存している点は実用化に向けた課題である。今後は機械学習を導入することで複数の特徴量を統合的に扱い、個人差やノイズに頑健なリアルタイム検出を可能にすることが期待される。本研究は、高速かつ非侵襲的な意思伝達システムの実現に向けた基盤を提供するものであり、重度障害者を含む幅広いユーザーへのコミュニケーション支援に貢献する可能性を示した。

本研究は、JST START プロジェクト推進型SBIR フェーズ1 支援 JPMJST2352, JSPS 科研費 JP22K17928, JP25K15190 の助成を受けたものである。

## 文 献

- 1) World Health Organization and World Bank: World Rep. Disabil., **1** (2011) 1-325.
- 2) N. Birbaumer: Psychophysiology, **43** (2006) 517-532.
- 3) S. Guo, Y. Wang, X. Zhang and B. Tang: Biomed. Signal Process. Control, **100** (2025) 107122.
- 4) H. Raza, D. Rathee, S. M. Zhou, H. Cecotti and G. Prasad: Neurocomputing, **343** (2019) 154-166.
- 5) T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Fray: IEEE Trans. Syst. Man Cybern., **19** (1989) 1527-1534.
- 6) H. B. Barlow: J. Physiol., **116** (1952) 290.
- 7) S. Mathôt, J. B. Melmi, L. Van Der Linden and S. Van der Stigchel: Plos One, **11** (2016) e0148805.
- 8) Y. Muto, H. Miyoshi and H. Kaneko: Plos One, **15** (2020) e0226991.
- 9) P. Binda, M. Pereverzeva and S. O. Murray: J. Neurosci., **33** (2013) 2199-2204.
- 10) S. Mathôt, L. Van der Linden, J. Grainger and F. Vitu: Plos One, **8** (2013) e78168.
- 11) M. Naber, G. A. Alvarez and K. Nakayama: Front. Psychol., **4** (2013) 919.
- 12) K. Kanari, K. Sakamoto and H. Kaneko: Plos One, **12** (2017) e0175453.
- 13) K. Kanari and H. Kaneko: J. Opt. Soc. Am. A, **38** (2021) 149-156.
- 14) 金成慧, 猪股千聖: 視覚の科学, **44** (2023) 35-43.
- 15) M. Juhola, H. Aalto, T. Jutila and T. P. Hirvonen: Ann. Biomed. Eng., **39** (2011) 973-982.
- 16) M. Juhola, H. Aalto, H. Joutsijoki and T. P. Hirvonen: Adv. Artif. Neural Syst., **2013** (2013) 972412.
- 17) P. Essig, Y. Sauer and S. Wahl: Transl. Vis. Sci. Technol., **10** (2021) 12.
- 18) D. H. Brainard: Spat. Vis., **10** (1997) 433-436.
- 19) D. G. Pelli: Spat. Vis., **10** (1997) 437-442.
- 20) K. Matsuda, T. Nagami, Y. Sugase, A. Takemura and K. Kawano: "A widely applicable real-time mono/binocular eye tracking system using a high frame-rate digital camera," *Proc. of the International Conference on Human-Computer Interaction* (Cham, 2017) pp. 593-608.