Eye-tracker algorithms to detect saccades during static and dynamic tasks: a structured review

Samuel Stuart[1], Aodhan Hickey[2], Rodrigo Vitorio[3], Karen Welman[4], Stacy Foo[5,6], David Keen[7] and Alan Godfrey[8]*

[1]Department of Neurology, Oregon Health & Science University, Portland, Oregon, United States of America
[2]Department of Health Intelligence, HSC Public Health Agency, Belfast, UK
[3]Department of Physical Education, São Paulo State University, São Paulo, Brazil
[4]Department of Sport Science, Movement Lab, Stellenbosch University, Stellenbosch, South Africa
[5]School of Human Sciences (Exercise and Sport Science), The University of Western Australia, Perth, Australia
[6]Physical Education and Sports Science, National Institute of Education, Nanyang Technological University, Singapore
[7]School of Life Sciences, University of Warwick, Warwick, UK
[8]Department of Computer and Information Science, Northumbria University, Newcastle upon Tyne, UK

***Correspondence to**:
Alan Godfrey, PhD
Department of Computer and Information Science
Northumbria University
Newcastle upon Tyne
NE1 8ST
Tel: 0191 227 3642
E-mail: alan.godfrey@northumbria.ac.uk

Word count: 3996

Abstract: 165

Tables: 2

Figures: 1

Abstract

Eye-tracking devices have become widely used as clinical assessment tools in a variety of applied-scientific fields to measure saccadic eye movements. With the emergence of multiple static and dynamic devices, the concurrent need for algorithm development and validation is paramount. This review assesses the prevalence of current saccade detection algorithms, their associated validation methodologies and the suitability of their application. Medline, Embase, PsychInfo, Scopus, IEEEXplore and ACM Digital Library databases were searched. Two independent reviewers and an adjudicator screened articles describing the detection of saccades from raw infrared/video-based eye-tracker data. Thirteen articles were screened and met the inclusion criteria. Overall, the majority of reviewed saccadic detection algorithms used simple velocity-based classifications with static eye-tracking systems. Studies demonstrated validity but are limited by the static nature of testing. Heterogeneity in system design, proprietary and bespoke algorithmic methods used, processing strategies, and outcome reporting is evident. This paper suggests the use of a more standardised methodology to facilitate experimental validity and improve comparison of results across studies.

Keywords: Algorithm, Detection, Eye-movements, Eye-tracker, Saccades

1 Introduction

Eye movements, specifically saccades, form the basis for visual exploration as they rapidly shift the fovea of the eye between areas of interest within the environment (Stuart et al., 2014a, Otero-Millan et al., 2008). Saccadic measurement has become prominent in many research fields, as deficient saccadic function can be used to understand cognitive (Hutton, 2008) or visual (Ibbotson and Krekelberg, 2011) processes, and help with neurological examination/diagnosis (Termsarasab et al., 2015) or to understand functional deficits in everyday tasks (e.g. walking) (Stuart et al., 2017a, Stuart et al., 2018). Saccadic eye movements are generally classified or examined through a range of spatiotemporal and kinematic outcome variables, including velocity, acceleration, number/frequency, timing and duration (Stuart et al., 2014b, Baloh et al., 1975, Boghen et al., 1974). Previously temporal measures of saccades (i.e. velocity, amplitude etc.) have been obtained with high-resolution invasive scleral search coils (Kimmel et al., 2012) or non-invasive electrooculography (EOG) (Stuart et al., 2016b) techniques, but such methods lack the spatial component of eye movement tracking (i.e. what is an individual looking at) and are influenced by a host of physiological processes (e.g. muscle activity, discomfort etc.). More recently a shift to non-invasive methods for comprehensively extracting quantitative temporal and spatial information are infrared/video-based eye-tracker devices (Stuart et al., 2014a). Although these devices do not measure eye orientation, they provide more accurate data than EOG during dynamic tasks (Stuart et al., 2016b) and reportedly have less variability in saccadic measurements than scleral coils (Smeets and Hooge, 2003), which were previously thought to be the most precise device for measuring eye movements. Video recordings from eye and scene cameras monitor eye movements and map them onto the external environment within various study protocols (Stuart et al., 2015, Holmqvist et al., 2011, Duchowski, 2007). The aforementioned outcomes are traditionally collected within static (e.g. seated or standing) tasks (Spooner et al., 1980, Forssman et al., 2017, Stuart et al., 2016a), but technological advances have recently facilitated more dynamic (e.g. navigating complex environments) test conditions (Stuart et al., 2017a, Stuart et al., 2018, Franchak and Adolph, 2010, Hayhoe and Ballard, 2005, Bardi et al., 2015, Matthis et al., 2018, Matthis and Hayhoe, 2015).

Accurate and reliable measurements of saccadic behaviours during various tasks in healthy (Munoz et al., 1998, Boghen et al., 1974, Tong et al., 2017) and atypical populations are paramount for informing clinical decisions (Stuart et al., 2017a, Stuart et al., 2018, Vidal et al., 2012). Detecting and classifying saccades from eye-tracker signals generally requires an algorithm to decipher eye movement signals and quantify outcomes. Most eye-tracker manufacturers provide ‘black-box’ signal processing and analytical programmes for detecting and measuring saccades. Importantly, this does not allow researchers to assess the specific methods involved (e.g. threshold settings and/or data reduction techniques) that are fundamental towards achieving high-quality and accurate of outcomes (Stuart et al., 2016a). However, some eye-tracker instruments provide raw Cartesian coordinate data (*x*, *y*) that can be exported via their proprietary software and can be used to process signals using customised algorithms leading to greater flexibility and sensitivity for more participant-specific outcomes (Schneider et al., 2009). Consequently, heterogeneous development in this field has led to a plethora of algorithms being created (Salvucci and Goldberg, 2000). The latter study reported that saccadic detection algorithms include velocity, dispersion and area-based methods can be used within eye-tracking protocols, but the authors provided no recommendations for algorithm use. In the process of developing robust data processing algorithms, it is often helpful to have informed recommendations. Thus, we examined previous work that developed and/or evaluated algorithms designed to derive/detect saccades from infrared/video-based eye-tracking data and aimed to provide recommendations concerning algorithm design and study methodologies.

2 Methods

2.1 Search strategy

The key terms were “eye-tracking”, “saccades” and “algorithm” that were searched using ‘wildcard’ markers (e.g. eye-track*, saccad*, algorithm*) within title, abstract and keywords (until May 2018). Key terms were matched and exploded with medical subject headings (MeSH) in each separate database where appropriate. Databases searched included Medline (from 1946), Embase (from 1974), PsychInfo (from 1806), Scopus, IEEEXplore and ACM Digital Library.

2.2 Inclusion and exclusion criteria

Studies were relevant if they incorporated terminologies that focused on development and/or evaluation of algorithms to detect saccades within eye-tracking data in the title, abstract and/or keywords. Articles were included if they reported development of algorithms to detect saccades from raw eye-tracker data. Articles were excluded if the methodologies involved (1) equipment other than infrared/video-based eye-trackers, (2) general details of commercial computer software packages (i.e. not detailing algorithm functionality), and/or (3) exclusive descriptions of fixations (including micro-saccades) and/or smooth pursuit eye movement detection/measurements. Articles written in English were considered for review, and conference abstracts, case studies, reviews, book chapters, commentaries, discussion papers and/or editorials were excluded.

2.3 Data extraction

An initial title screen for relevant articles was performed by one reviewer (SS) once the database results were combined. Subsequently, both the titles and abstracts of the selected articles were reviewed by two independent reviewers (SS, AH) and confirmed by a third reviewer (AG). A full-text review was required to establish whether the study had met the review criteria if the title and/or abstract were ambiguous. Information including the purpose of the study and algorithm design details were extracted and synthesised into a table format (SS). Data entries were then confirmed by a second reviewer (AH). Only data pertaining to the detection and measurement of saccades were extracted.

3 Results

3.1 The evidence base

The search strategy yielded 156 articles, excluding duplicates (Figure 1 - Adapted from Moher et al. (2009)). The initial screening resulted in 41 articles of interest of which 13 were identified for inclusion by the first reviewer (SS) and 14 by the second reviewer (AH), resulting in 1 disagreement. A consensus was made for the inclusion of 13 articles for review following a consultation with the third reviewer (AG). Reasons for exclusion at abstract and full-text stage are highlighted in Figure 1 and summarised here:

1. No quantitative analysis of saccade detection algorithm (Salvucci and Goldberg, 2000, Radant and Hommer, 1992),
2. Description of software packages (Cercenelli et al., 2017, Andreu-Perez et al., 2016),
3. No infrared/video-based eye-tracker (Sauter et al., 1991),
4. Non-saccadic detection/measurement only (i.e. measurement of fixations, smooth pursuits, micro-saccades etc.) (Larsson et al., 2015, Duchowski et al., 2002, Falkmer et al., 2008, Gómez-Poveda and Gaudioso, 2016, Holland et al., 2012, Holland and Komogortsev, 2013, Komogortsev and Karpov, 2013, Kübler et al., 2014, Pauler et al., 1996, Pedrotti et al., 2011, Toivanen, 2016, Zhang et al., 2009),
5. Eye-tracking device development or frequency comparison (Espinosa et al., 2015, Karn, 2000, Leube et al., 2017, Lyamin and Cherepovskaya, 2017, Price et al., 2009, Pruehsner and Enderle, 2002, Pruehsner et al., 2003, Stuart et al., 2017b) and
6. Did not involve human participants (Konig and Buffalo, 2014, Corrigan et al., 2017, Behrens and Weiss, 1992).

<Figure 1>

3.2 Study design

All included reviewed studies did not provide detailed descriptions of their study participants including, basic visual capabilities (acuity or contrast sensitivity), use of corrective eye-wear, cognitive abilities, and/or age-range. The majority of the studies involved healthy individuals yet no specific inclusion and/or exclusion criteria for their eye-tracking experiments were described, Table 1. Furthermore, one study adopted a dynamic study protocol that involved walking (Stuart et al., 2014b) and another used a functional task of driving (Tafaj et al., 2012a), while the remaining studies focused on static seated eye-tracking tasks (Table 2).

<Table 1>

3.3 Algorithm design

Algorithms reported within the reviewed articles used a range of identification techniques to derive saccades from eye movement signals. However, all of the algorithms used velocity, acceleration, dispersion or adaptive-based algorithms either independently or combined, Table 2.

The most popular algorithm design involved velocity thresholds-based for saccade detection (Andersson et al., 2017, Komogortsev et al., 2010, Komogortsev and Karpov, 2013, Larsson et al., 2013b, Liston et al., 2013, Nyström and Holmqvist, 2010, Santini et al., 2016, Stuart et al., 2014b, Zemblys et al., 2018, Diaz et al., 2013, Kumar et al., 2008). However, the specific threshold values of the algorithms varied between studies (i.e. detection velocity 20-300°/s, Table 2), with several providing no details of velocity-thresholds used (Tafaj et al., 2012a, Diaz et al., 2013) or using adaptive algorithms that change the thresholds based upon the data (Zemblys et al., 2018, Larsson et al., 2013a, Nyström and Holmqvist, 2010). One dynamic study used a higher detection velocity (250°/s) in order to account for vestibular-ocular reflexes that may affect data quality (Stuart et al., 2014b). Several studies examined many algorithms with the same eye movement data to examine different algorithm accuracies (Andersson et al., 2017, Komogortsev et al., 2010, Komogortsev and Karpov, 2013).

<Table 2>

3.4 Instruments

Static eye-tracking devices were the primary tools used within the studies (Andersson et al., 2017, Komogortsev and Karpov, 2013, Kumar et al., 2008, Larsson et al., 2013a, Liston et al., 2013, Rozado et al., 2012, Zemblys et al., 2018, Nyström and Holmqvist, 2010), while mobile or head-mounted eye-trackers were used in four studies (Santini et al., 2016, Stuart et al., 2014b, Tafaj et al., 2012a, Diaz et al., 2013).

3.5 Computer coding language

MATLAB® (MathWorks, MA, USA) was the preferred coding language used to implement eye-tracker saccade detection algorithms across the studies (Andersson et al., 2017, Larsson et al., 2013b, Liston et al., 2013, Nyström and Holmqvist, 2010, Santini et al., 2016, Stuart et al., 2014b, Zemblys et al., 2018), while C# (Microsoft Corporation, WA, USA) and Python™ (Python Software Foundation, DE, USA) were used in two other studies (Diaz et al., 2013, Tafaj et al., 2012b). Four studies did not, however, provide details of the coding language used to develop their algorithms (Komogortsev et al., 2010, Komogortsev and Karpov, 2013, Rozado et al., 2012, Kumar et al., 2008) (Table 2). There were also few studies that combined eye-tracking software pre-processing (e.g. iViewX™, SensoMotoric Instruments, Germany) with custom computer coding to evaluate their eye-tracking data (Diaz et al., 2013, Larsson et al., 2013b).

3.6 Algorithm outcomes and evaluation

The reviewed studies provided a range of algorithm-derived saccadic outcomes, including velocity, acceleration, amplitude, duration, timing, and number/frequency. However, outcome reporting varied between the studies, with some studies providing comprehensive details (Larsson et al., 2013a, Stuart et al., 2014b, Zemblys et al., 2018) and others not providing any outcomes (Diaz et al., 2013, Kumar et al., 2008, Rozado et al., 2012), Table 2.

Additionally, the evaluation of algorithm accuracy was performed in an unstandardized manner. Several studies used human 'expert' coders as a reference for their algorithms, which involved inspection of eye-tracker signals and videos to compare human saccade detection to algorithm detection and measurement. Other studies reported a range of algorithm comparisons using previously published algorithms as a validation comparison. Performance of some algorithms (e.g. Nyström & Holmqvist (NH) or Identification by Hidden Markov Model (IHMM)) was dependent upon the evaluation method, with different results for human or previous algorithm detection (Andersson et al., 2017).

Studies sparsely reported statistical analysis methods, but largely used correlation coefficients (e.g. Cohens Kappa or Interclass correlation coefficients) or percentages of matched events from direct comparisons to humans or other algorithms. While most studies demonstrated good to excellent agreement between methodologies it is unclear if the statistical analysis is accurate due to the small sample sizes involved and the lack of basic statistical reporting (e.g. normality of data, reasons for non-parametric or parametric analysis, statistical software used etc.). Furthermore, studies did not examine statistical bias, assumed that relative and absolute agreement were captured within the same test, and did not provide limits of agreement or equivalent tests.

4 Discussion

This review examined 13 studies that developed algorithms to detect saccadic eye movements within static or dynamic infrared video-based eye-tracker signals. This review has demonstrated that there is limited robust evidence on the development of a standardised approach to saccadic detection algorithm. Similarly, methodological limitations of previous studies impact the ability to understand and implement earlier algorithms with current technology. Further work is warranted to establish appropriate study and algorithm design to accurately tailor the saccadic detection methodology to static or dynamic tasks.

4.1 Study design

All studies provided sparse details of study protocols/design which limits the ability of future researchers to replicate methodological procedures and algorithms. The shortcomings in methodologies and barriers to transferability of algorithms is further impacted by the limited range of populations involved, only two studies involved atypical participants (Stuart et al., 2014b, Tafaj et al., 2012a), Table 2. The reviewed studies provided exclusive details regarding algorithm development, but lacked basic inclusion and exclusion criteria for their participants. This may limit the applicability of their algorithms across populations as understanding and in-depth analysis of the quality of the sample is lacking. Furthermore, without inclusion of alterative populations, it is unclear if the developed saccadic detection algorithms could be accurately applied to a range of participants. Indeed, one study did not disclose any information about their participants (Diaz et al., 2013) and therefore results must be considered with caution when implementing the algorithm. The sample size of the participant groups also varied with most studies involving small groups (<30 participants), which can impact the generalisability of results and limits the power of statistical analysis. It is also difficult to know whether results would be influenced by a greater number of participants or recorded eye movements, which should be considered within future algorithm assessments.

Many algorithms were developed during static (in particular seated) tasks, with only one study examining algorithm performance during a dynamic task (walking) (Stuart et al., 2014b) and another during a functional task (driving) (Tafaj et al., 2012a). With the increasing use of mobile eye-tracking devices there is a need to shift focus to evaluate saccadic detection algorithms during dynamic and/or functional tasks to ensure robust and accurate outcomes across a range of tasks. This is important as static-based eye-tracker algorithms may not perform as well during dynamic tasks (Stuart et al., 2017b).

4.2 Algorithm design

4.2.1 Standardisation

Algorithm suitability and adaptability is dependent upon many factors including (1) type of task (i.e. static or dynamic), (2) data quality (i.e. clean or noisy), (3) eye-tracking devices (i.e. sampling frequency, binocular or monocular, fixed or mobile with rigid or flexible eye-cameras), (4) required saccadic outcomes, (5) features available for saccade identification (i.e. root mean square, bivariate contour ellipse area etc.), and (6) the research question. Currently there is no standardised approach for deriving saccadic information from eye-tracker signals, although some reviewed articles had considered human 'expert' coders as the reference standard (Andersson et al., 2017, Zemblys et al., 2018). However, our findings suggest the use of velocity-based thresholds could be the foundation for an automated/objective standard as most studies adopted this approach. There was, however, little consensus or justification for the specific thresholds used for accurate saccade detection as they were either undisclosed or ranged from 20-300°/s (Table 2). Although, dynamic testing reportedly used a higher velocity-based threshold (>250°/s) to rule of small eye movements due to vestibular-ocular reflexes that may occur when moving. However, this substantial variation makes direct comparisons across studies difficult. Therefore, deciding upon an optimal velocity threshold for saccadic detection is challenging due to differences in study methodologies and instrumentation. Future research should report thresholds used and potentially evaluate the influence of differing threshold levels across different populations.

4.2.2 Velocity-based threshold methodologies

Our findings suggest that simple velocity-based threshold techniques are adequate for both static and dynamic testing conditions, which may be enhanced with the addition of other spatiotemporal and/or kinematic parameters (e.g. acceleration, dispersion (amplitude) and/or duration thresholds). However, a single filtering methodology or threshold for every eye-tracking device/signal and participant is unlikely to produce accurate results. It is important to mention that filtering or smoothing eye-tracking signals can result in an artificial reduction of saccade peak velocities and broadening of saccade durations, which can differ dependent on the eye movement signal and the filtering technique used (Diaz et al., 2013). Similarly due to between participant variability, physiological considerations of eye movement classification may be required in order to obtain a reasonable threshold range for each individual (Komogortsev et al., 2010, Salvucci and Goldberg, 2000), or at least be based upon a wide range of normative data. Although some studies suggest the use of adaptable algorithms to customise the thresholds and filters for the individual signals to manage the aforementioned issues (Nyström and Holmqvist, 2010, Zemblys et al., 2018), further work is required to establish which features are essential for robust and accurate saccade detection from various eye-tracking signals.

4.2.3 Self-validation methodologies

With the emergence of advanced data analytics and intelligent systems, the most complex algorithms used in the observed eye-tracking studies included probabilistic methods, including (1) Bayesian decision theory (i.e. pattern classification when underlying patterns are known), (2) online mixture methods (i.e. involves generating repeated realisation of random variables from probabilities of a known simple distribution) and (3) machine learning (i.e. data driven-predictions based on training sets of input observations). These techniques provide a unique opportunity for self-validation in the absence of a comparative measure. The provision of open-source data-sets from previously validated study paradigms could greatly improve the development of these new analysis methods by providing learning data, and encourages consistency in experimental set-ups between simple and complex methods. As data modelling methods become more widely used in bio-informatics studies, it is important to explore their suitability for application in advanced eye-tracking data systems. However to date, with the presence of effective simple velocity-based algorithms, studies lacked justification for further levels of algorithm complexity.

4.3 Instruments and coding language

Several studies showed the application of static-based algorithms to dynamic tasks, which was made possible with the use of mobile eye-tracking technology. However, the progression from static to dynamic tasks require some pragmatic considerations. Static eye-trackers have higher sampling frequencies and require constrained movements and it is unknown whether the algorithms developed for these devices can be robustly applied to mobile devices with lower sampling frequencies and unrestricted movements. Our findings indicated that there were limited methodological descriptions such as eye-tracker hardware features, software versions, and even the pixel-to-degree conversion factors used. We found that methodologies varied across several studies (sampling frequencies ranging from 25 to 50Hz) that used a similar device Dikablis®, Essential and Professional Glasses Ergoneers GmbH, Berlin, Germany) (Santini et al., 2016, Stuart et al., 2014b, Tafaj et al., 2012a). As minimum of 50Hz is required for accurate saccadic detection with robust outcomes (Holmqvist et al., 2011, Andersson et al., 2010), this suggest that inaccurate and technically flawed algorithms have been developed (Santini et al., 2016). Furthermore, only one study provided detailed descriptions of the pixel-to-degree conversion factor used (Stuart et al., 2014b), which is an essential input for accurate eye movement velocity calculations. This input factor is determined from the relative distance between the eye and the eye camera, and is often standardised for use across all participants. Such technique may be inappropriate especially for mobile eye-trackers with flexible eye cameras (e.g. Dikablis Professional Glasses, Ergoneers GmbH, Germany) as the eye positions relative to the eye cameras can vary significantly across participants, impacting both the eye movement signals and velocity output data (Santini et al., 2016), and ultimately influencing the performance of the algorithm developed.

All-but-two algorithms were developed with MATLAB®, Table 2. Adoption of a single language facilitates more immediate replication and direct comparison of study findings across independent groups. Alternatively, this limits development or more widespread adoption. As this field matures, new coding environments will be used. Yet care must be taken on the replication of algorithms between platforms (e.g. MATLAB® vs. Python™) as implementation differences on recursive functions can lead to accumulation errors (Ladha et al., 2016). However, exact replication of algorithms may only be possible with the provision of specific coding scripts used to derive the saccadic detection and metrics. None were provided by any of the reviewed articles. Future algorithm development would benefit from open access publication of algorithm code to allow independent validation and/or application to a variety of devices, tasks, participants and eye movement signals. Alternatively, algorithms protected by intellectual property or within iterative stages of development could be best represented using pseudocode (Komogortsev and Karpov, 2013, Salvucci and Goldberg, 2000) or mathematical notation (Kim et al., 2018), aided by the provision of flow diagrams (Stuart et al., 2014a, Stuart et al., 2017b).

4.4 Evaluation of algorithm outcomes

Accuracy of some algorithm outputs depended on the particular evaluation method used (i.e. human experts or other previously validated algorithms), with a single human expert being more similar to newly developed algorithms, i.e. knowing exact signal characteristics and nuances. However, human coding is time-consuming and may bias the data (Andersson et al., 2017) as the human coder is often involved in the algorithm development, leading to inflated agreement. Further work is required to develop optimal means of evaluating algorithm performance. This is complicated as many of the reviewed algorithms are not entirely similar. The aims of the algorithms are similar (i.e. obtain meaningful outcomes from eye movement signals), but the means by which outcomes are obtained varies, with some examining saccades exclusively while others may require fixation-related data and/or attempt to identify all eye movement events. Improvements were found within the reviewed algorithms when combinations of saccade-fixation detection filters and thresholds were used, alongside the presence of setting adjustments or addition of adaptive threshold adjustments. However, the extent of these improvements appeared relatively minor when compared with human coders, and thus the added benefits over simple saccade detection methods (i.e. identification with velocity-based threshold) is unknown.

Another accuracy limiting factor is the diverse range that exists in both occulography systems and their associated algorithms. Comparisons between research paradigms, especially during a period of rapid innovation within the field, is complex when comparisons were not made between 'like-for-like' systems. The variety that already exists in research grade devices including, static and mobile units, and their wide variety of hardware designs and housing set-ups makes direct comparison between systems difficult. This can potentially confound the processing of raw data as the physical design of hardware systems often generates noise within the resultant data itself, especially in mobile eye-tracking devices (Stuart et al., 2016a). The absence of standardised processing and analysis pipelines implied that not all algorithms were assessed robustly against reference measures or validated algorithms, with either no comparisons being made (Kumar et al., 2008) or subjective reports of adaptive algorithms (Tafaj et al., 2012a). Some studies also reported that there were notable differences found between certain algorithms (Nyström and Holmqvist, 2010) and that meaningful saccadic data can be lost depending on thresholds set (Liston et al., 2013). Therefore, there is little evidence within the reviewed studies of an established 'gold standard' system or algorithm for saccade detection.

4.5 Recommendations for future research

Based on the findings of this review, we make the following recommendations for future research into saccade detection algorithms within static and dynamic conditions:

1. Robust study design reporting is required, which should include adequate detail for the results to be replicated.
2. Comparison of algorithms within both static and mobile eye-tracking is required to allow appropriate algorithm selection based upon robust methodological findings.
3. Algorithms need to be assessed within the specific context of the study, for example assessment during dynamic rather than static tasks for mobile eye-tracker signals.
4. Determination of appropriate saccadic thresholds dependent upon sampling frequency, device, task and participant group.
5. Adopt and report suitable eye-tracker calibration methods prior to testing.
6. Open access code publication for further validation by other researchers.

5 Conclusions

In summary, we reviewed studies pertaining to saccade detection within infrared/video-based eye-trackers. No consensus was found for the optimal means of detecting saccades, but the majority of algorithms used a velocity-based threshold identification method. This is likely due to the lack of a 'gold standard' saccade detection method and reflects the difficulty of developing an algorithm without a robust comparator. Future work is required to establish a more harmonised reporting format with transparent saccade detection algorithms.

Conflicts of Interest

The authors declare no conflicts of interest.

REFERENCES

ANDERSSON, R., LARSSON, L., HOLMQVIST, K., STRIDH, M. & NYSTRÖM, M. 2017. One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods,* 49**,** 616-637.

ANDERSSON, R., NYSTRÖM, M. & HOLMQVIST, K. 2010. Sampling frequency and eye-tracking measures: how speed affects durations, latencies, and more. *Journal of Eye Movement Research,* 3.

ANDREU-PEREZ, J., SOLNAIS, C. & SRISKANDARAJAH, K. 2016. EALab (Eye Activity Lab): a MATLAB Toolbox for Variable Extraction, Multivariate Analysis and Classification of Eye-Movement Data. *Neuroinformatics,* 14**,** 51-67.

BALOH, R. W., SILLS, A. W., KUMLEY, W. E. & HONRUBIA, V. 1975. Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology,* 25**,** 1065-1065.

BARDI, L., DI GIORGIO, E., LUNGHI, M., TROJE, N. F. & SIMION, F. 2015. Walking direction triggers visuo-spatial orienting in 6-month-old infants and adults: An eye tracking study. *Cognition,* 141**,** 112-120.

BEHRENS, F. & WEISS, L. R. 1992. An algorithm separating saccadic from nonsaccadic eye movements automatically by use of the acceleration signal. *Vision Research,* 32**,** 889-893.

BOGHEN, D., TROOST, B., DAROFF, R., DELL'OSSO, L. & BIRKETT, J. 1974. Velocity characteristics of normal human saccades. *Investigative Ophthalmology & Visual Science,* 13**,** 619-623.

CERCENELLI, L., TIBERI, G., CORAZZA, I., GIANNACCARE, G., FRESINA, M. & MARCELLI, E. 2017. SacLab: A toolbox for saccade analysis to increase usability of eye tracking systems in clinical ophthalmology practice. *Computers in Biology and Medicine,* 80**,** 45-55.

CORRIGAN, B. W., GULLI, R. A., DOUCET, G. & MARTINEZ-TRUJILLO, J. C. 2017. Characterizing eye movement behaviors and kinematics of non-human primates during virtual navigation tasks. *Journal of vision,* 17**,** 15.

DIAZ, G., COOPER, J., KIT, D. & HAYHOE, M. 2013. Real-time recording and classification of eye movements in an immersive virtual environment. *Journal of Vision Vol 13(12), 2013, ArtID 5,* 13.

DUCHOWSKI, A., MEDLIN, E., COURNIA, N., MURPHY, H., GRAMOPADHYE, A., NAIR, S., VORAH, J. & MELLOY, B. 2002. 3-D eye movement analysis. *Behavior Research Methods, Instruments, and Computers,* 34**,** 573-591.

DUCHOWSKI, A. T. 2007. Eye tracking methodology. *Theory and practice,* 328.

ESPINOSA, J., ROIG, A. B., PÉREZ, J. & MAS, D. 2015. A high-resolution binocular video-oculography system: Assessment of pupillary light reflex and detection of an early incomplete blink and an upward eye movement. *BioMedical Engineering Online,* 14.

FALKMER, T., DAHLMAN, J., DUKIC, T., BJÄLLMARK, A. & LARSSON, M. 2008. Fixation identification in centroid versus start-point modes using eye-tracking data. *Perceptual and Motor Skills,* 106**,** 710-724.

FORSSMAN, L., ASHORN, P., ASHORN, U., MALETA, K., MATCHADO, A., KORTEKANGAS, E. & LEPPÄNEN, J. M. 2017. Eye-tracking-based assessment of cognitive function in low-resource settings. *Archives of disease in childhood,* 102**,** 301-302.

FRANCHAK, J. M. & ADOLPH, K. E. 2010. Visually guided navigation: Head-mounted eye-tracking of natural locomotion in children and adults. *Vision research,* 50**,** 2766-2774.

GÓMEZ-POVEDA, J. & GAUDIOSO, E. 2016. Evaluation of temporal stability of eye tracking algorithms using webcams. *Expert Systems with Applications,* 64**,** 69-83.

HAYHOE, M. & BALLARD, D. 2005. Eye movements in natural behavior. *Trends in cognitive sciences,* 9**,** 188-194.

HOLLAND, C., KOMOGORTSEV, O. & TAMIR, D. Identifying usability issues via algorithmic detection of excessive visual search. Conference on Human Factors in Computing Systems - Proceedings, 2012. 2943-2952.

HOLLAND, C. D. & KOMOGORTSEV, O. V. Complex eye movement pattern biometrics: Analyzing fixations and saccades. Proceedings - 2013 International Conference on Biometrics, ICB 2013, 2013.

HOLMQVIST, K., NYSTRÖM, M., ANDERSSON, R., DEWHURST, R., JARODZKA, H. & VAN DE WEIJER, J. 2011. *Eye tracking: A comprehensive guide to methods and measures*, OUP Oxford.

HUTTON, S. B. 2008. Cognitive control of saccadic eye movements. *Brain Cogn,* 68**,** 327-40.

IBBOTSON, M. & KREKELBERG, B. 2011. Visual perception and saccadic eye movements. *Current opinion in neurobiology,* 21**,** 553-558.

KARN, K. S. "Saccade pickers" vs. "fixation pickers": The effect of eye tracking instrumentation on research. Proceedings of the Eye Tracking Research and Applications Symposium 2000, 2000. 87-88.

KIM, S., POLLANEN, M., REYNOLDS, M. G. & BURR, W. S. Identification of Errors in Mathematical Symbolism and Notation: Implications for Software Design. International Congress on Mathematical Software, 2018. Springer, 297-304.

KIMMEL, D., MAMMO, D. & NEWSOME, W. 2012. Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers in Behavioral Neuroscience,* 6.

KOMOGORTSEV, O. V., GOBERT, D. V., JAYARATHNA, S., KOH, D. H. & GOWDA, S. M. 2010. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *IEEE Transactions on Biomedical Engineering,* 57**,** 2635-2645.

KOMOGORTSEV, O. V. & KARPOV, A. 2013. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior Research Methods,* 45**,** 203-215.

KONIG, S. D. & BUFFALO, E. A. 2014. A nonparametric method for detecting fixations and saccades using cluster analysis: removing the need for arbitrary thresholds. *Journal of neuroscience methods,* 227**,** 121-131.

KÜBLER, T. C., BUKENBERGER, D. R., UNGEWISS, J., WÖRNER, A., ROTHE, C., SCHIEFER, U., ROSENSTIEL, W. & KASNECI, E. Towards automated comparison of eye-tracking recordings in dynamic scenes. 2014 5th European Workshop on Visual Information Processing (EUVIP), 10-12 Dec. 2014 2014. 1-6.

KUMAR, M., KLINGNER, J., PURANIK, R., WINOGRAD, T. & PAEPCKE, A. Improving the accuracy of gaze input for interaction. Eye Tracking Research and Applications Symposium, ETRA 2008, 2008 Savannah, GA. 65-68.

LADHA, C., DIN, S. D., NAZARPOUR, K., HICKEY, A., MORRIS, R., CATT, M., ROCHESTER, L. & GODFREY, A. Toward a low-cost gait analysis system for clinical and free-living assessment. 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 16-20 Aug. 2016 2016. 1874-1877.

LARSSON, L., M. NYSTROM & STRIDH, M. 2013a. Detection of Saccades and Postsaccadic Oscillations in the Presence of Smooth Pursuit. *IEEE Transactions on Biomedical Engineering,* 60**,** 2484-2493.

LARSSON, L., NYSTRÖM, M., ANDERSSON, R. & STRIDH, M. 2015. Detection of fixations and smooth pursuit movements in high-speed eye-tracking data. *Biomedical Signal Processing and Control,* 18**,** 145-152.

LARSSON, L., NYSTROM, M. & STRIDH, M. 2013b. Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. *IEEE Transactions on Biomedical Engineering,* 60**,** 2484-2493.

LEUBE, A., RIFAI, K. & WAHL, S. 2017. Sampling rate influences saccade detection in mobile eye tracking of a reading task. *Journal of Eye Movement Research,* 10.

LISTON, D. B., KRUKOWSKI, A. E. & STONE, L. S. 2013. Saccade detection during smooth tracking. *Displays,* 34**,** 171-176.

LYAMIN, A. V. & CHEREPOVSKAYA, E. N. 2017. An Approach to Biometric Identification by Using Low-Frequency Eye Tracker. *IEEE Transactions on Information Forensics and Security,* 12**,** 881-891.

MATTHIS, J. & HAYHOE, M. 2015. Eye, head, and foot tracking during locomation over real-world complex terrain. *Journal of vision,* 15**,** 1322-1322.

MATTHIS, J. S., YATES, J. L. & HAYHOE, M. M. 2018. Gaze and the Control of Foot Placement When Walking in Natural Terrain. *Current Biology,* 28**,** 1224-1233. e5.

MOHER, D., LIBERATI, A., TETZLAFF, J. & ALTMAN, D. G. 2009. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine,* 151**,** 264-269.

MUNOZ, D., BROUGHTON, J., GOLDRING, J. & ARMSTRONG, I. 1998. Age-related performance of human subjects on saccadic eye movement tasks. *Experimental brain research,* 121**,** 391-400.

NYSTRÖM, M. & HOLMQVIST, K. 2010. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior research methods,* 42**,** 188-204.

OTERO-MILLAN, J., TRONCOSO, X. G., MACKNIK, S. L., SERRANO-PEDRAZA, I. & MARTINEZ-CONDE, S. 2008. Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. *Journal of vision,* 8**,** 21-21.

PAULER, D. K., ESCOBAR, M. D., SWEENEY, J. A. & GREENHOUSE, J. 1996. Mixture models for eye-tracking data: A case study. *Statistics in Medicine,* 15**,** 1365-1376.

PEDROTTI, M., LEI, S., DZAACK, J. & RÖTTING, M. 2011. A data-driven algorithm for offline pupil signal preprocessing and eyeblink detection in low-speed eye-tracking protocols. *Behavior Research Methods,* 43**,** 372-383.

PRICE, D., KAPUTA, D., SIERRA, D. A. & ENDERLE, J. Infrared-based eye-tracker system for saccades. Bioengineering, Proceedings of the Northeast Conference, 2009.

PRUEHSNER, W. R. & ENDERLE, J. D. 2002. The operating version of the eye tracker, a system to measure saccadic eye movements. *Biomedical Sciences Instrumentation,* 38**,** 113-118.

PRUEHSNER, W. R., LIEBLER, C. M., RODRIGUEZ-CAMPOS, F. & ENDERLE, J. D. 2003. The Eye Tracker System - A system to measure and record saccadic eye movements. *Biomedical Sciences Instrumentation,* 39**,** 208-213.

RADANT, A. D. & HOMMER, D. W. 1992. A quantitative analysis of saccades and smooth pursuit during visual pursuit tracking. A comparison of schizophrenics with normals and substance abusing controls. *Schizophrenia Research,* 6**,** 225-235.

ROZADO, D., SAN AGUSTIN, J., RODRIGUEZ, F. B. & VARONA, P. 2012. Gliding and saccadic gaze gesture recognition in real time. *Transactions on Interactive Intelligent Systems,* 1.

SALVUCCI, D. D. & GOLDBERG, J. H. Identifying fixations and saccades in eye-tracking protocols. *In:* SPENCER, S. N. & SPENCER, S. N., eds. Proceedings of the Eye Tracking Research and Applications Symposium 2000, 2000 Palm Beach Gardens, FL. 71-78.

SANTINI, T., FUHL, W., KÜBLER, T. & KASNECI, E. Bayesian identification of fixations, saccades, and smooth pursuits. Eye Tracking Research and Applications Symposium (ETRA), 2016. 163-170.

SAUTER, D., MARTIN, B. J., DI RENZO, N. & VOMSCHEID, C. 1991. Analysis of eye tracking movements using innovations generated by a Kalman filter. *Medical and Biological Engineering and Computing,* 29**,** 63-69.

SCHNEIDER, E., VILLGRATTNER, T., VOCKEROTH, J., BARTL, K., KOHLBECHER, S., BARDINS, S., ULBRICH, H. & BRANDT, T. 2009. Eyeseecam: An eye movement–driven head camera for the examination of natural visual exploration. *Annals of the New York Academy of Sciences,* 1164**,** 461-467.

SMEETS, J. B. & HOOGE, I. T. 2003. Nature of variability in saccades. *J Neurophysiol,* 90**,** 12-20.

SPOONER, J. W., SAKALA, S. M. & BALOH, R. W. 1980. Effect of aging on eye tracking. *Archives of Neurology,* 37**,** 575-576.

STUART, S., ALCOCK, L., GALNA, B., LORD, S. & ROCHESTER, L. 2014a. The measurement of visual sampling during real-world activity in Parkinson's disease and healthy controls: A structured literature review. *Journal of neuroscience methods,* 222**,** 175-188.

STUART, S., ALCOCK, L., GODFREY, A., LORD, S., ROCHESTER, L. & GALNA, B. 2016a. Accuracy and re-test reliability of mobile eye-tracking in Parkinson's disease and older adults. *Medical Engineering and Physics,* 38**,** 308-315.

STUART, S., GALNA, B., DELICATO, L. S., LORD, S. & ROCHESTER, L. 2017a. Direct and indirect effects of attention and visual function on gait impairment in Parkinson's disease: influence of task and turning. *European Journal of Neuroscience*.

STUART, S., GALNA, B., LORD, S. & ROCHESTER, L. 2015. A protocol to examine vision and gait in Parkinson's disease: impact of cognition and response to visual cues. *F1000Research,* 4.

STUART, S., GALNA, B., LORD, S., ROCHESTER, L. & GODFREY, A. Quantifying saccades while walking: validity of a novel velocity-based algorithm for mobile eye tracking. Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference. , 2014b Chicago, Il., USA.: IEEE, 5739-5742.

STUART, S., HICKEY, A., GALNA, B., LORD, S., ROCHESTER, L. & GODFREY, A. 2016b. iTrack: instrumented mobile electrooculography (EOG) eye-tracking in older adults and Parkinson’s disease. *Physiological measurement,* 38**,** N16.

STUART, S., HICKEY, A., GALNA, B., LORD, S., ROCHESTER, L. & GODFREY, A. 2017b. iTrack: instrumented mobile electrooculography (EOG) eye-tracking in older adults and Parkinson's disease. *Physiological measurement,* 38**,** N16-N31.

STUART, S., LORD, S., GALNA, B. & ROCHESTER, L. 2018. Saccade frequency response to visual cues during gait in Parkinson's disease: the selective role of attention. *European journal of neuroscience,* 47**,** 769-778.

TAFAJ, E., KASNECI, G., ROSENSTIEL, W. & BOGDAN, M. Bayesian online clustering of eye movement data. 7th Eye Tracking Research and Applications Symposium, ETRA 2012, 2012a Santa Barbara, CA. 285-288.

TAFAJ, E., KASNECI, G., ROSENSTIEL, W. & BOGDAN, M. Bayesian online clustering of eye movement data. Eye Tracking Research and Applications Symposium (ETRA), 2012b. 285-288.

TERMSARASAB, P., THAMMONGKOLCHAI, T., RUCKER, J. C. & FRUCHT, S. J. 2015. The diagnostic value of saccades in movement disorder patients: a practical guide and review. *Journal of Clinical Movement Disorders,* 2**,** 14.

TOIVANEN, M. 2016. An advanced Kalman filter for gaze tracking signal. *Biomedical Signal Processing and Control,* 25**,** 150-158.

TONG, M. H., ZOHAR, O. & HAYHOE, M. M. 2017. Control of gaze while walking: Task structure, reward, and uncertainty. *Journal of vision,* 17**,** 28-28.

VIDAL, M., TURNER, J., BULLING, A. & GELLERSEN, H. 2012. Wearable eye tracking for mental health monitoring. *Computer Communications,* 35**,** 1306-1311.

ZEMBLYS, R., NIEHORSTER, D. C., KOMOGORTSEV, O. & HOLMQVIST, K. 2018. Using machine learning to detect events in eye-tracking data. *Behavior Research Methods,* 50**,** 160-181.

ZHANG, Y., FU, H., LIANG, Z., ZHAO, X., CHI, Z., FENG, D. & ZHAO, X. Eye movement data modeling using a genetic algorithm. 2009 IEEE Congress on Evolutionary Computation, 18-21 May 2009 2009. 1038-1044.

Figures

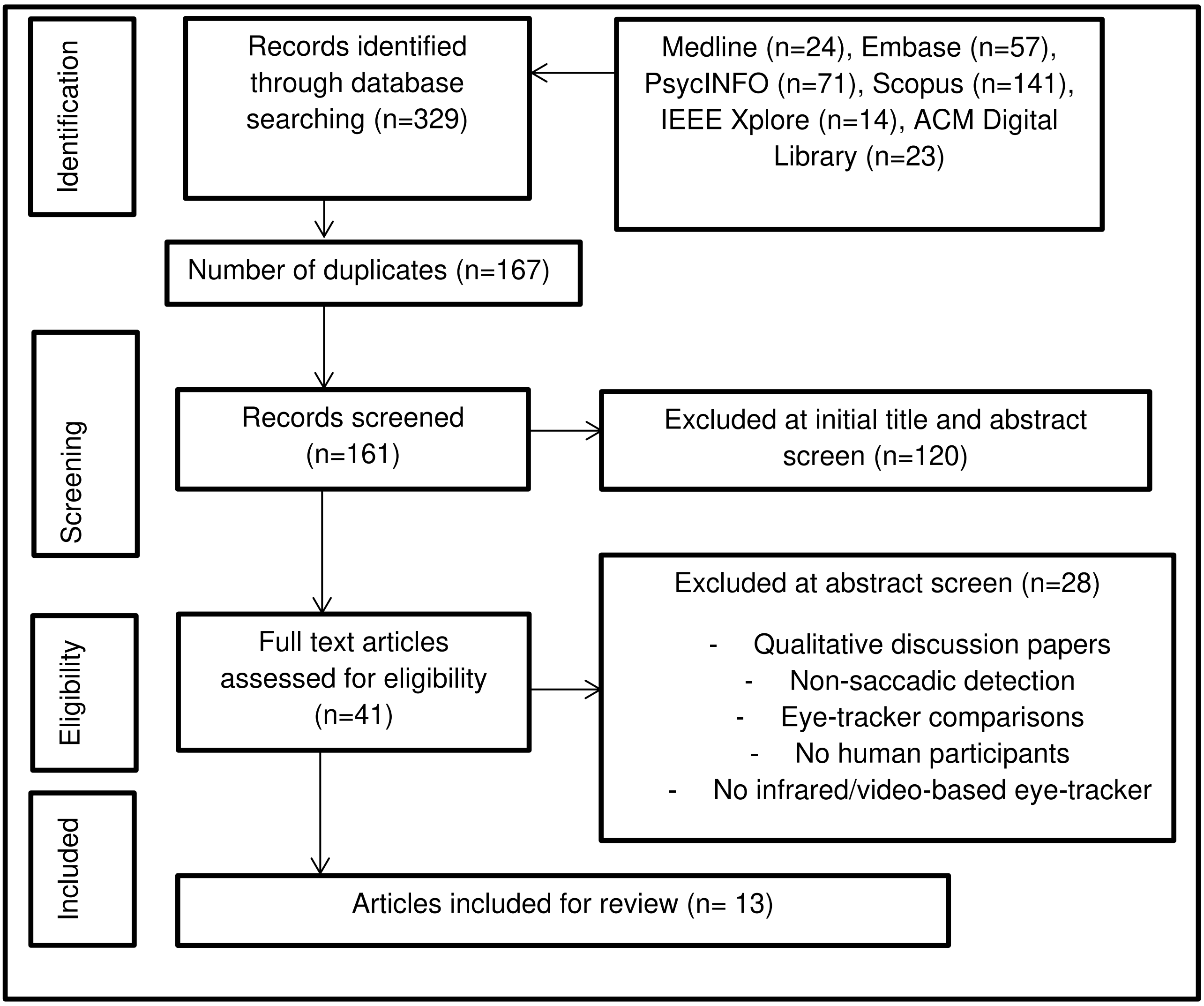

Figure 1. *PRISMA flow chart of study design. This illustrates the yield of the search strategy at each stage of the study selection process.*

TABLES

***Table 1**: Study participant characteristics, inclusion and exclusion criteria*

Reference	Participants	Inclusion criteria	Exclusion criteria
Andersson et al. (2017)	• 34 participants	NR	NR
Diaz et al. (2013)	NR	NR	NR
Komogortsev and Karpov (2013)	• 11 participants	• 18-25 years • No prior eye-tracking experience	NR
Komogortsev et al. (2010)	• 22 participants (9 males, 13 females, 21.2 ± 3.1 years)	• 18-25 years • Positional accuracy during calibration better than 1.7^0 • Less than 20% invalid data	NR
Kumar et al. (2008)	• 15 participants (11 males, 4 females, 26 years)	NR	NR
Larsson et al. (2013a)	• 33 participants (31.2 ± 9.9 years)	• Students	NR
Liston et al. (2013)	NR	NR	NR
(Nyström and Holmqvist, 2010)	• 300 participants (reading task) • 10 participants (scene perception task)	• Students • Manually observed high quality eye-tracker data	NR
Rozado et al. (2012)	• 10 participants	NR	NR
Santini et al. (2016)	• Six participants (4 males, 2 females, 31.5 ± 2.6 years). Two participants wore corrective glasses for myopia	NR	NR
Stuart et al. (2014b)	• Five participants with Parkinson's disease • Five healthy older adults control participants	• >50 years	NR
Tafaj et al. (2012a)	• Homonymous visual field deficit participants • Glaucoma participants • Control participants	NR	NR
Zemblys et al. (2018)	• Five participants. Only one participant was used for the testing dataset	NR	NR

* NR – Not reported

Table 2: *Study characteristics including algorithm methodologies*

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
Andersson et al. (2017)	Evaluated 10 eye-movement classification algorithms, and compared against data from human coders	Static	Static eye-tracker (500Hz, (iViewX™ Hi-Speed 1250, SensoMotoric Instruments GmbH, Berlin, Germany)	Various algorithms that involved: 1. Fixation Dispersion Algorithm based on Covariance (CDT): Veneri et al. (2010) fixation dispersion algorithm 2. Engbert & Mergenthalser (EM): Derived from algorithm by Engbert and Kliegl (2003) 3. Identification by Dispersion-Threshold (IDT): based on Widdel (1984) data reduction algorithm 4. Identification by Kalman Filter (IKF): Generates predicted eye velocity signal based on acceleration modelled as white noise 5. Identification by Minimal Spanning Tree (MST): Creates a 'data-tree' that generates multiple data branches 6. Identification by Velocity Threshold (IVT): Velocity computed for every eye position, then compared against a threshold and marked as a FIX orpart of a SAC 7. Identification by Hidden Markov Model (IHMM): Similar to IVT. Classes fixations or saccades but implements a Viterbi sampler (Forney, 1973) and a re-estimation algorithm (Baum et al., 1970) 8. Nyström & Holmqvist (NH): adaptive algorithm that adjusts velocity threshold based on presence of noise 9. Binocular Individual	Bilateral filter Kalman filter (*Required for IKF algorithm*)	Variable thresholds used CDT: 13-sample window EM: SAC Velocity parameter separating from noise. SAC Duration ≥6 samples. IDT: FIX Duration ≥55ms, FIX Dispersion ≥2.7°. IKF: Chi-square threshold =3.75, 5-window sample, deviation value = 1000. IMST: SAC Amplitude ≥0.6°, 200-sample window. IHMM: SAC Velocity ≥45°/s. IVT: SAC Velocity ≥45°/s. NH: FIX Duration ≥55ms. BIT: FIX Duration ≥56ms.	MATLAB® SAC: Humans, EM, IDT, IKF, IMST, IHMM, IVT, NH, LNS Outcomes: SAC Velocity SAC Duration Accuracy (Cohen's Kappa, κ, comparison to humans) EM: 0.64-0.67 IDT: 0.26-0.45 IKF: 0.46-0.58 IMST: 0.30-0.54 IHMM: 0.60-0.71 IVT: 0.63-0.76 NH: 0.60-0.68 LNS: 0.75-0.81

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
				Threshold (BIT): van der Lans et al. (2011) algorithm examines velocity of both eyes simultaneously 10. Larsson, Nyström & Stridh (LNS): Adaptive algorithm using velocity-based thresholds			
Diaz et al. (2013)	Detection of SAC when using a head-mounted virtual reality display to counteract vestibulo-ocular reflex	Static virtual reality in a controlled setting.	Mobile infra-red eye-tracker (60Hz, Arrington Research Inc., Scottsdale, AR, USA)	Velocity threshold-based classification	Median (3-units wide) and Gaussian (3-units) filters to remove signal outliers and jitter respectively	NR	Python™, Viewpoint software (version 2.9.2.6) and EYE-TRAC®6 (Applied Science Laboratories, USA)
Komogortsev and Karpov (2013)	Developed new and modified existing eye-tracking algorithms for ternary classification (i.e. FIX, SAC, and smooth pursuits).	Static in a controlled setting. Manual observations vs. proposed algorithms	Static eye-tracker (1000Hz, EyeLink 1000, SR Research, Canada)	Various algorithms that involved: 1. Modified the IVT algorithm with a second threshold to create the velocity and velocity threshold identification (IVVT) 2. Velocity and movement pattern identification (IVMP) 3. The newly proposed velocity and dispersion threshold identification (IVDT). Separation of SAC similar to IVVT and IVMP but then separates smooth pursuits from FIX by applying modified dispersion threshold within a temporal window of size Tw	Manual filtering of noise	SAC Velocity ≥70°/s. IVVT: FIX Velocity ≤26°/s, FIX Amplitude >3.5°, FIX Duration <4ms. IVMP: Magnitude of movement Tw =0.2, Temporal window =120-140ms. IVDT: Dispersion threshold TD =2°, Temporal window =110-150ms.	NR Extension of the I-VT (IVVT) does not provide meaningful classification. The IVMP and IVDT provided classification close to manual observations but the latter were only done for 3 participants (out of 10). The IVDT had smaller performance variability and dependence on

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
							thresholds vs. IVMP.
Komogortsev et al. (2010)	Evaluated five algorithms to classify oculomotor behaviour. To develop a common basis for comparing algorithms through qualitative and quantitative scores	Static in a controlled setting.	Static eye-tracker (120Hz, Tobii x120, Tobii Technology AB, Stockholm, Sweden)	Various algorithms that involved: 1. IVT 2. IHMM: 3. IKF: 4. IMST: 5. IDT	Kalman filter (Required for IKF algorithm)	Variable thresholds used IVT: SAC Velocity =5-300°/s. IHMM: SAC Velocity =5-300°/s. IMST: Dispersion =0.033-2°. IKF: Threshold values =1-60. IDT: Dispersion =0.033-2°. FIX Duration ≥100ms.	NR SAC number, amplitude and quantitative score
Kumar et al. (2008)	Real-time SAC detection and fixation smoothing	Static in a controlled setting. Red balloons (n=20) were presented on a screen. The balloon popped when looked at and moved locations.	Static eye-tracker (Tobii 1750, Tobii Technology AB, Stockholm, Sweden)	Gaze movement threshold where two gaze points separated by a Euclidean distance of more than a given SAC threshold. Similar to IVT with modifications: measure displacement relative to current estimate of fixation location and; look ahead one measurement and reject movements over SAC threshold.	Kalman filter	SAC smoothing algorithm introduces an additional 20ms latency at SAC thresholds	NR
Larsson et al. (2013)	To detect SAC and post-saccadic oscillations in the presence of smooth pursuits	Static in a controlled setting. Compared to manually annotated eye movements and another algorithm (Nyström and Holmqvist, 2010). Testing consisted of images, texts, moving dots, short video clips, and a scrolling text.	Static eye-tracker (500Hz, (iViewX™ Hi-Speed 1250, SensoMotoric Instruments, GmbH, Berlin, Germany)	Combines saccade detection in the acceleration domain with specialised on and offset criteria for saccades and post-saccadic oscillations (PSO). This is in 3 stages: pre-processing, detection and PSO.	9-point binocular calibration in iViewX™ followed by 4 validation targets. Pre-processing excluded 3 different disturbances: screen outliers, blinks (700ms threshold) and one-sample spikes. For the latter a median filter (length = 3)	Adaptable SAC Velocity and Acceleration thresholds. Sample-to-sample Velocity ≤20% of Peak Velocity for SAC start/end to be classified. Time between SAC ≥20ms. SAC Duration ≥6ms.	MATLAB® and iViewX™ SAC number, duration and peak velocity Good agreement between algorithm and manual observations (κ ≈ 0.8).

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
					was used to avoid supressing PSO. PSO described by an all-pole model.		
Liston et al. (2013)	Three stage algorithm to detect SAC during smooth pursuits	Static in a controlled environment. Presented an ideal-observer analysis to benchmark detection performance vs. manual observations.	Static video-based eye-tracker (ISCAN RK-726 PCI), ISCAN, Burlington, USA)	Three stages: 1. Median filter to process eye-velocity trace; 2. Linear detector based ideal observer approach measuring SAC likelihood as well as threshold to flag possible saccades 3. Clustering to mitigate effects of uncertainty and tracker noise	Non-linear median filter, sliding window of odd size which computes the median velocity within the window Algorithm parameters: window size, SAC amplitude threshold, min. SAC duration, min refractory period	SAC Velocity =0.20 to 60°/s, with ~30ms cross-correlation template.	MATLAB®. SAC velocity, amplitude and duration. Efficiency varied as a function of minimum duration (20-40%). Threshold too high, small SAC not be detected. Threshold too low, small number of false SAC.
Nyström and Holmqvist (2010)	Developed an adaptive algorithm to detect SAC, while overcoming glissades.	Static in a controlled setting. Compared against two other algorithms (Salvucci and Goldberg, 2000, and Smeets and Hooge, 2003)	Static eye-tracker (500Hz, (iViewX^{TM} Hi-Speed 1250, SensoMotoric Instruments, GmbH, Berlin, Germany)	Adaptable velocity threshold-based classification, with removal of glissades.	Savitzky-Golay finite impulse response smoothing filter, 2nd order with a length of ×2 min. saccade duration.	Used an iterative data-driven approach to find SAC velocity threshold based on initial value = 100-300°/s (≤1000°/s), SAC Acceleration ≤100000°/s^2, SAC Duration ≥10ms, FIX Duration ≥40ms. Window length for glissade search = SAC end + min. fixation duration	MATLAB®. SAC velocity, acceleration and duration. Notable differences between algorithms. Detection of glissades feasible and found to occur in high proportions of participants

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
Rozado et al. (2012)	Detection of gliding gaze gestures and SAC gaze gestures in real-time	Static in a controlled setting. Testing for suitable combination using Needleman-Wunsch algorithm and Hierarchical Temporal Memory networks	Head-mounted ITU gaze tracker with infra-red webcam (30Hz, Sandberg Nightcam 2, Sanberg A/S, Denmark)	Combination of saccadic gaze gesture and Needleman-Wunsch (movement pattern recognition) was found to be best	Pilot testing informed larger test. Findings suggested use of saccadic gaze gestures only and not gliding. Testing involved 3 tasks: accuracy, browsing and velocity.	Dwell time ≥ 500ms	NR Approx. 91-95% dependent upon dwell time
Santini et al. (2016)	Ternary classification of eye movements.	Static in a controlled setting. Collected 24 datasets involving both induced and natural eye movements compared with IVDT algorithm (Komogortsev and Karpov, 2013)	Mobile eye-tracker (Dikablis Professional Glasses, 60Hz, Ergoneers GmbH, Berlin, Germany). Only monocular data was collected at 30Hz.	Bayesian online mixture model to automatically track the visual attention in dynamic visual scenes. Utilises Gaussian, Gamma and Dirichlet distributions and is IVDT: velocity and dispersion based.	No calibration step was performed by using pupil position signal as input and so applied unjittering function (Stampe, 1993)	IBDT: SAC Duration ≥ 80ms. IVDT: FIX Duration ≥ 100ms.	MATLAB®. SAC velocity, amplitude and duration. Precision, specificity (95.6%, 95.4%) compared to IVDT (89.6%, 92.1%)
Stuart et al. (2014b)	To detect SAC during walking	Dynamic in a controlled laboratory setting. 5m walking compared with manual classifications.	Mobile infra-red eye-tracker (50Hz, Dikablis Essential Glasses, Ergoneers GmbH, Berlin, Germany)	Velocity and acceleration threshold-based classification. Utilises point to point change of x/y co-ordinates, conversion of pixels to degrees	Initial calibration procedure involving SAC at 5° distance while standing static	SAC Velocity: ≥ 240°/s, < 1000°/s. SAC Acceleration: > 3000°/s², <100000°/s². SAC Amplitude: ≥ 5°. SAC Duration: ≤100ms.	MATLAB®. SAC number, frequency, velocity, amplitude, direction, duration and timing. Good reliability with ICC (2,1) between PD (0.94) and HC (0.94) participants. SAC detection: PD (85%), HC (81%).

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy
Tafaj et al. (2012)	To detect SAC and FIX points in real-time (online) dynamic scenes during	Static while performing a driving task. General observations from algorithm deployment.	Mobile infra-red eye-tracker (25Hz, Dikablis Essential Glasses, Ergoneers GmbH, Berlin, Germany)	Identified by Random Forest Classifier (IRF): produces many decision trees that can use the 14 features provided to derive SAC.	3-point calibration procedure No pre-processing was, however, performed by the equipment used)	NR	C# and Infer.NET. The latter was also utilised Variational Message Passing for probalistic inference. SAC points Algorithm showed promise to adapt quickly
Zemblys et al. (2018)	Introduce a new design principle for saccade detection using machine learning for selection of appropriate thresholds. Compared against human coder and NH algorithm (Nyström and Holmqvist, 2010), as well as across different sampling frequencies.	Static in a controlled environment. Tracked a silver 0.2° dot with 2x2 pixel black centre, which jumped around a 7x7 grid, with pauses of 1s at each position.	Static eye-tracker (1000Hz) (EyeLink 1000)	IRF: produces many decision trees that can use the 14 features provided to derive SAC. **14 features:** - sampling frequency - root mean square (rms) - standard deviation (SD) - bivariate contour ellipse area (bcea) - dispersion - velocity/acceleration - median distance - mean distance - Rayleigh test - i2mc (Identification by two-means clustering) - rms/SD/bcea difference	Low-pass Butterworth filter with cut-off frequency of x0.8 Nyquist frequency of the new data rate, with a 20ms-window. Data re-sampled at 60, 120, 200, 250, 300, 500 and 1250Hz.	No user adaptable settings from the heuristics used by the IRF.	MATLAB® Saccade velocity, acceleration, amplitude and duration. Accuracy (κ): IRF: 0.91 (vs. humans), 0.58 (vs. NH algorithm)

Reference	Aims	Task and test conditions	Device	Algorithm	Pre/Post Processing	Algorithm Thresholds	Coding platform, outcomes and accuracy

FIX – Fixation, min. – minimum, NR – Not reported, SAC – Saccade, vs. – versus. PD – Parkinson's disease, HS – healthy controls