

Published in final edited form as:

J Neurosci Methods. 2014 April 30; 227: 121–131. doi:10.1016/j.jneumeth.2014.01.032.

A Nonparametric Method for Detecting Fixations and Saccades Using Cluster Analysis: Removing the Need for Arbitrary Thresholds

Seth D. König^{a,b,e} and Elizabeth A. Buffalo^{b,c,d}

^aWallace H. Coulter Department of Biomedical Engineering at the Georgia Institute of Technology and Emory University, 313 Ferst Drive, Atlanta, Georgia 30332, USA

^bYerkes National Primate Research Center, 954 Gatewood Road, Atlanta, Georgia 30329, USA

^cDepartment of Neurology, Emory University School of Medicine, 1440 Clifton Road, Atlanta, GA 30322, USA

^dDepartment of Physiology and Biophysics, University of Washington, Seattle, WA 98195, USA

^eGraduate Program in Neurobiology and Behavior, University of Washington, Seattle, WA 98195, USA

Abstract

Background—Eye tracking is an important component of many human and non-human primate behavioral experiments. As behavioral paradigms have become more complex, including unconstrained viewing of natural images, eye movements measured in these paradigms have become more variable and complex as well. Accordingly, the common practice of using acceleration, dispersion, or velocity thresholds to segment viewing behavior into periods of fixations and saccades may be insufficient.

New Method—Here we propose a novel algorithm, called Cluster Fix, which uses k-means cluster analysis to take advantage of the qualitative differences between fixations and saccades. The algorithm finds natural divisions in 4 state space parameters—distance, velocity, acceleration, and angular velocity—to separate scan paths into periods of fixations and saccades. The number and size of clusters adjusts to the variability of individual scan paths.

Results—Cluster Fix can detect small saccades that were often indistinguishable from noisy fixations. Local analysis of fixations helped determine the transition times between fixations and saccades.

© 2014 Elsevier B.V. All rights reserved.

Corresponding Author: Name: Elizabeth A. Buffalo, ebuffalo@uw.edu, Phone number: 1-404-712-9431.

Present postal address: Washington NPRC, Buffalo Lab, 1959 NE Pacific Street, Box 357290, Seattle, WA 98195

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest

The authors do not report any conflicts of interest.

Comparison with Existing Methods—Because Cluster Fix detects natural divisions in the data, predefined thresholds are not needed.

Conclusions—A major advantage of Cluster Fix is the ability to precisely identify the beginning and end of saccades, which is essential for studying neural activity that is modulated by or time-locked to saccades. Our data suggest that Cluster Fix is more sensitive than threshold-based algorithms but comes at the cost of an increase in computational time.

Keywords

saccade detection; fixations; eye tracking; cluster analysis; viewing behavior

2. Introduction

Rigorous analysis of eye movements dates back to the seminal work of Alfred Yarbus (Yarbus, 1967). Today, eye tracking is used to determine the location of visual attention (Duchowski, 2002; McAlonan et al., 2008; Lee et al., 2011), measure memory (Smith et al., 2006; Smith and Squire, 2008; Hannula and Ranganath, 2009; Jutras et al., 2009; Richmond and Nelson, 2009; Hannula et al., 2010; Jutras and Buffalo, 2010; Hannula et al., 2012; Killian et al., 2012), detect cognitive impairments (Crutcher MD, 2009; Lagun et al., 2011; Zola et al., 2013), and evaluate visual search strategies (Najemnik and Geisler, 2005; Dewhurst R, 2012). The development of non-invasive infrared eye-tracking technologies has further enhanced the value and feasibility of collecting viewing behavior across a large range of experimental tasks.

Commonly, viewing behavior, represented as a scan path, is parsed into periods of fixations and saccades using a variety of algorithms. The most widely used algorithms employ velocity and/or acceleration thresholds to detect the occurrences of saccades because the velocity and acceleration of the eye are much greater during a saccade than during a fixation (Otero-Millan et al., 2008; Nystrom and Holmqvist, 2010; Kimmel et al., 2012). Threshold-based algorithms have the benefit of being intuitive, quick, and easy to implement. Other popular algorithms use density or dispersion and areas of interest (Tatler BW, 2005; Ito et al., 2011). Variants and combinations of these algorithms include mechanisms to correct for errors in eye tracking such as blinks and other temporary losses of signal (Wass et al., 2013).

Despite a significant increase in the use of eye movements in neuroscience, there have been very few advances in the algorithms used to detect fixations and saccades (Salvucci and Goldberg, 2000). We could only find a few instances of algorithms that deviated significantly from the most widely used algorithms. Unfortunately, many of these alternative algorithms still employ a velocity threshold to detect potential saccades, followed by additional techniques including principal component analysis to distinguish between smooth pursuit, saccades, and noise (Berg et al., 2009; Liston et al., 2012). One exception is (Urruty et al., 2007) which used dispersion and projection clustering into arbitrary subspace to detect fixations. To the best of our knowledge, these algorithms have not been adopted in subsequent studies.

Algorithms employing velocity and acceleration thresholds for saccade detection may be sufficient for simple tasks in which subjects make predictable saccades towards a stationary target; however, more complex oculomotor tasks such as unconstrained viewing of natural scenes or dynamic stimuli may produce more variable eye movements (Andrews and Coppola, 1999; Hayhoe and Ballard, 2005; Berg et al., 2009; Rayner, 2009). A major source of this variability arises from the variability in saccade amplitude which is strongly correlated with the peak velocity of the saccade (Otero-Millan et al., 2008; Martinez-Conde et al., 2009; Martinez-Conde et al., 2013). Velocity thresholds may not accurately parse highly variable scan paths into periods of fixations and saccades since saccade amplitudes and thus their peak velocity are not constrained in free viewing. Further, many algorithms employ arbitrary thresholds based on qualitative human observations which can vary across research laboratories and even from one experiment to the next within a laboratory. Finally, computed viewing behavior statistics including saccade rate, fixation duration, saccade duration, and saccade amplitude vary not only according to experimental variables but also by the method used to calculate them (Duchowski, 2007; Shic et al., 2008; Nystrom and Holmqvist, 2010). Therefore, there exists a need for a more accurate, sensitive, non-arbitrary, and completely automated saccade detection algorithm. Such an algorithm could constitute a “gold standard” for detecting fixations and saccades from scan paths so that viewing behavior could be accurately compared across experiments, laboratories, and algorithms.

Here we present a novel algorithm, called Cluster Fix, which applies k-means cluster analysis to parse scan paths into fixations and saccades. There are several clear qualitative differences between fixations and saccades—saccades are temporally short with a high velocity whereas fixations are longer in duration with a slower velocity. These qualitative differences translate into quantitative differences and the occupation of different regions in state space. Cluster Fix makes no assumptions about the arrangement of scan paths in state space, requires no human inputs, and includes only duration thresholds as free parameters.

3. Methods

3.1 Eye Tracking

Scan paths were obtained at 200 Hz using an infrared eye tracker (ISCAN) from rhesus macaques freely viewing 288 images of natural scenes. Eye tracking data were collected from 4 adult male macaques seated head-fixed in a dimly illuminated room 60 cm away from a 19” CRT monitor with a refresh rate of 120 Hz. Images of natural scenes were 600 by 800 pixels large and subtended 25 by 33 degrees of visual angle (dva). Experimental control software (CORTEX <http://dally.nimh.nih.gov/>) displayed images for 10 seconds each. Initial calibration of the infrared eye tracking system consisted of a 9 point calibration task. Drift was tracked throughout the experiment by presenting additional calibration trials between image viewing trials. We excluded from further analysis any eye tracking data more than 50 pixels (2 dva) outside of the image. Blinks were rarely observed in our data so we did not make any corrections other than the exclusion of data outside of the image. Standard blink correction techniques should work with Cluster Fix if scan paths are evaluated in a piece-wise manner ignoring blinks and as long as at least one fixation is

present in each evaluated portion of the scan path (Supplementary Figure 1). All experiments were carried out in accordance with the National Institutes of Health guidelines and were approved by the Emory University Institutional Animal Care and Use Committee and Emory Institutional Review Board.

3.2 Cluster Fix Algorithm

The Cluster Fix algorithm was written in MATLAB and is available as supplemental material. Table 1 contains the procedural outline detailing the major processes achieved by the algorithm. To avoid filtering artifacts, eye traces were buffered prior to filtering, filtered, and then the buffers were removed. First, horizontal and vertical eye traces from the viewing of each image were individually pre-processed using a polyphase implementation (MATLAB function RESAMPLE) to up-sample the data from 200 Hz to 1000 Hz and then filtered using a 60th order low pass filter with a cutoff frequency of 30 Hz. These pre-processing steps were used to remove noise from the scan path while retaining prominent features of saccades. These pre-processing steps followed the method used previously in our laboratory to remove noise for saccade detection with a velocity threshold. However, these pre-processing steps could be replaced by any pre-processing steps that sufficiently increase the signal-to-noise ratio.

Next, the absolute value of 4 state space parameters—distance, velocity, acceleration, and angular velocity—were calculated for every time point. Velocity and acceleration were computed as the first and second derivative of position, respectively. Distance was measured as the Euclidian distance between the position of the scan path at a time point to the position of the scan path two time points later. Angular velocity was calculated as the difference in the angle of scan path from one time point to the next. Angular velocity was subtracted from 360 degrees so that lower values were associated with fixations. For each state space parameter, any values greater than 3 standard deviations above the mean were set to 3 standard deviations above the mean, and all values were then normalized from 0 to 1.

Cluster Fix globally clustered every time point in state space into k number of clusters. We determined the appropriate number of clusters using the average silhouette width (MATLAB function SILHOUETTE). The silhouette width measures the average ratio of inter- and intra-cluster distances to determine the appropriate number of clusters. Higher ratio values indicate that points within clusters were closer to each other than points outside of their respective clusters. We chose the number of possible clusters to be from 2 to 5 clusters because in a typical scan path there is at least 1 fixation and 1 saccade, and in the most complex scan path we can divide fixations into 2 separate clusters and saccades into 3 separate clusters. Fixations can be subdivided into 2 clusters: one with low angular velocity and one with high angular velocity. Saccades can be subdivided into 3 clusters: low velocity but high acceleration, low acceleration but high velocity, and high velocity and high acceleration.

To reduce the number of computations, SILHOUETTE was used iteratively on 10% of the time points to determine the k , between 2 and 5, that produced the highest ratio or within 90% of the highest ratio. We found no difference between using only 10% of the time points or all the time points except a reduction in the number of computations. In the case where

the ratio was high for several k , the largest k was used. Once the appropriate number of clusters was identified, clusters were determined using k-means cluster analysis on all the time points (Figure 1A–B). Five replicates were performed for determining the appropriate number of clusters and for clustering of all the time points. The cluster with the lowest sum of the mean velocity and acceleration was classified as a cluster consisting of fixation time points. Because fixations were often divided into 2 clusters, one with high angular velocity and one with low angular velocity angular velocity, additional fixation clusters were determined by finding clusters whose mean velocity and acceleration were within 3 standard deviations of the mean of the first fixation cluster. All other clusters were classified as saccade clusters (Figure 1C–D). Fixation periods shorter than 25 ms in duration were also reclassified as saccades.

To increase the sensitivity of the algorithm to smaller amplitude saccades, the algorithm reevaluated each fixation locally using the same method applied in global clustering (Figure 2). The concept of local re-clustering is to analyze data at the appropriate scale (i.e. in between 2 “large” saccades detected by global clustering) to remove the over shadowing effects of the larger variability in the whole or global data. In local re-clustering, time points 50 ms (approximately the average saccade duration) prior to and following a detected fixation were re-clustered with the detected fixation. SILHOUETTE was used iteratively on 20% of the time points to determine the k , between 1 and 5. The median of the best k was chosen for the final number of clusters. The additional possibility of only finding 1 optimal cluster was added in case the evaluated portion of the scan path only contained a single fixation and no saccades. For each cluster, the median velocity and median acceleration were identified. Then, the cluster with the lowest sum of these two values was considered to consist of fixation time points. Because the number of time points in each cluster was relatively small, measures of the mean and standard deviation of each cluster were more sensitive to outliers. Therefore, additional fixation clusters were determined by finding clusters whose median velocity and acceleration overlapped with the first fixation cluster in velocity and acceleration state space.

Time points that fell within saccade clusters identified using local re-clustering were classified as saccade time points in the global clusters. Any fixations shorter than 5 ms in duration were also temporarily classified as saccades. This duration criterion should not be considered a free parameter but simply accounted for incorrectly detected fixations that occurred at velocity or acceleration peaks and troughs where acceleration or velocity was near 0, respectively. Next, Cluster Fix flagged saccades less than 10 ms in duration as classification errors and reclassified them as fixations. Then, Cluster Fix flagged fixations that were less than 25 ms in duration as classification errors and reclassified them as saccades. These duration criteria operate independently, but infrequently these criteria were used together when a very short “fixation” (i.e. less than 25 ms) was adjacent to another very short “fixation” to create a fixation with a duration greater than 25 ms. Finally, fixation and saccade time periods were down-sampled to the acquisition frequency of 200 Hz.

4. Results

Local re-clustering was performed after global clustering because the variability in each state space parameter within a single fixation or saccade was much smaller than the variability across all fixations or saccades. As shown in Figure 3, previously detected fixations were reevaluated via local re-clustering to ensure that the algorithm was more sensitive to smaller saccades and to increase the specificity for determining the transition time between fixations and saccades. For this representative scan path, global clustering and even liberal thresholds could not distinguish three smaller saccades from the noisy fixation.

Surprisingly, local re-clustering revealed some overlap between saccades and fixations in global state space (Figure 1E–F). The overlap in state spaces is due to the detection of smaller saccades whose velocity and acceleration profiles were similar to that of a noisy fixation. A reduction in the size of the fixation cluster was also observed due to an increase in specificity of the transition time between fixations and saccades.

Velocity and/or acceleration thresholds appear to often miss smaller saccades that are similar to noisy fixations. While it is problematic to compare methods for detecting fixations because every method will produce different results and no gold standard currently exists, it is possible to demonstrate the utility of this novel method by showing that arbitrary thresholds in velocity and acceleration state space cannot achieve the same sensitivity as Cluster Fix. On the same set of scan paths, we implemented a basic velocity and acceleration threshold algorithm which detected saccades above various velocity thresholds with the additional constraint that each saccade had to contain a maximum acceleration above a specific threshold. The same pre-processing of scan paths for Cluster Fix was performed prior to the implementation of each threshold algorithm. Additionally, the threshold algorithms included a local re-classification in which fixation time points juxtaposed to saccades with an acceleration greater than the acceleration threshold became reclassified as saccade time points. Finally the threshold algorithm applied a saccade and a fixation duration threshold of 10 ms and 25 ms, respectively. As seen in Table 2 the values selected for the thresholds included both arbitrary thresholds and thresholds dependent on the variability of the scan path (i.e. mean + std). Importantly, the choice of thresholds dramatically altered the computed behavioral statistics. Compared to Cluster Fix, the threshold algorithms appeared to omit smaller amplitude saccades as indicated by larger average saccade arc length and increased distances between fixations. Because threshold algorithms omitted potential small amplitude saccades, the number of detected fixations and saccades decreased.

Visual inspection of velocity profiles and the arrangement of scan paths in state space also revealed that velocity and acceleration thresholds often omitted potential smaller saccades and classified them as fixations (Figure 4). The omission of potential smaller saccades occurred either when the velocity during a saccade did not reach the threshold or when the saccade's velocity did not exceed the threshold for a sufficient duration. Further, thresholds appeared to be insufficient for precisely identifying the onset and offset of saccades. This is particularly evident when visualized in 2D state space (Figure 4C). There exist an infinite number of combinations of thresholds, but no combination of these types of thresholds could

identify saccades and fixations that overlap in state space. From the 2D state space plot, we observed that an optimal velocity and acceleration threshold would be a nonlinear function of both velocity and acceleration instead of singular threshold values.

A key advantage of Cluster Fix was that the algorithm impartially identified saccade start and end times based on the 4 state space parameters. Similar to the threshold algorithm that we implemented, additional computation could be added to threshold-based algorithms, such as determining when the saccade reaches a minimum velocity or acceleration. However, this may require additional arbitrary thresholds.

Visual inspection of the scan paths supported the same conclusions as above in which it appeared that Cluster Fix correctly identified saccades of various amplitudes not identified by the threshold-based algorithms (Figure 5). Various thresholds omitted different probable saccades (Figure 5: purple circles), and several of the probable smaller saccades were consistently omitted across different thresholds (Figure 5: orange circles); however, none of the algorithms appeared to misclassify fixations. Visual inspection across multiple scan paths from the same monkey and scan paths from different monkeys indicated that while some thresholds may appear to be sufficient for one scan path, that same threshold may not be sufficient for another scan path or subject (Supplementary Figure 2).

In addition to the velocity and acceleration threshold-based algorithms, we implemented two common dispersion-based algorithms: a simple dispersion threshold algorithm and a minimum spanning tree algorithm (MATLAB code from (Komogortsev et al., 2010; Komogortsev and Karpov, 2013)). In the dispersion threshold algorithm, we used a window of 25 ms comparable to the minimum expected fixation duration. In the minimum spanning tree algorithm we used a 100 ms window comparable to the longest expected saccade duration. Unlike the dispersion threshold algorithm, the minimum spanning tree algorithm required fixation (25 ms) and saccade duration (10 ms) thresholds to accurately analyze scan paths. Similar to the velocity and acceleration threshold algorithms, the threshold values drastically altered computed behavioral statistics (Table 2) and the accuracy of these algorithms (Supplementary Figure 3). In general, dispersion-based algorithms had difficulty detecting the onset and offset of saccades. As indicated by the behavioral statistics in Table 2, the dispersion-based algorithms also detected a larger number of fixations and saccades. Visual inspection of the scan paths revealed that the lower dispersion threshold values caused the algorithms to detect noisy portions of fixations as saccades inappropriately dividing fixations into multiple “fixations” with shorter durations.

The k-means clustering algorithm employed by Cluster Fix is an optimization algorithm attempting to find a global minimum in the distances between points in a cluster and the clusters' centroids. The k-means algorithm initiates randomly and then iteratively computes clusters. Replications of this iterative process are used to increase the chance of finding the best clusters. Because this process is random and an optimal solution is not guaranteed to be found, the results produced by Cluster Fix may have some inconsistencies across multiple applications to the same scan path. We have developed a bootstrapping method to determine the consistency of the results produced by Cluster Fix (Supplementary Figure 4). We found that Cluster Fix consistently detects saccades with amplitudes greater than 1.5 dva.

However, saccades with smaller amplitudes were less consistently detected, and the consistency of their detection appeared to be related to the noisiness of the surrounding scan path. Smaller saccades may be detected more consistently with a higher acquisition frequency. Further, approximately 92% of all time points were always classified as a fixation or a saccade across multiple applications, and 97% of time points were consistently classified as a fixation or a saccade across at least 90% of applications.

Lastly, we calculated average parameter values across multiple images for individual monkeys. Although the average parameter values changed substantially for saccades, these values changed less for fixations (Figure 6), which could be useful for real-time fixation detection. The average parameter values for fixations and saccades occupied completely separate regions of state space. In fact, a support vector machine classified which regions of state space belonged to fixations versus saccades with greater than 99% accuracy for each monkey individually ($n = 4$) trained on as little as 5% of the data.

5. Discussion

Here, we have described a novel algorithm using cluster analysis to detect periods of fixations and saccades to improve the analysis of highly variable scan paths. Both global and local cluster analysis were necessary to detect small saccades that were often indistinguishable from noisy fixations or hidden by the large variability in saccade amplitudes. Global clustering offers an initial pass on classifying scan paths into fixations and saccades while local re-clustering refines these results, allowing for the detection of smaller saccades and a determination of the transition times between fixations and saccades. Cluster Fix removed the need for determining thresholds and removed the need for any assumptions regarding the arrangement of parameters in state space. The algorithm assumed that individual clusters are of Gaussian distributions since k-means was used, but we did not make any assumptions about the arrangement or size of these clusters in state space. Moreover, clusters changed with the amount of noise and variability found in the scan path allowing the algorithm to adapt to individual scan paths. One caveat is that Cluster Fix still requires fixation and saccade duration thresholds as free parameters. In theory, additional classifiers could distinguish between fixations and saccades without the need for duration thresholds.

We are confident that Cluster Fix can correctly identify smaller saccades because these smaller saccades were identified by finding natural divisions in scan paths, and these smaller saccades occupied different regions of state space than fixations. Visual inspection of the scan paths supports this conclusion as well. Identification of smaller saccades was more difficult when the smaller saccades were surrounded by highly variable fixations in comparison to smaller saccades surrounded by less noisy fixations. Further, identification of these smaller saccades often served to break up longer “fixations,” detected by global clustering, into more appropriate length fixations.

A common practice in the analysis of human eye movement data is to operationally define saccades that are smaller than 1 dva in amplitude as microsaccades (Engbert and Kliegl, 2003; Otero-Millan et al., 2008; Martinez-Conde et al., 2013). Here we did not distinguish

between microsaccades and saccades albeit approximately 4% of the observed saccades had amplitudes of less than 1 dva (not equivalent to saccade arc length in Table 2). We observed that a large portion of these saccades were detected during “complex fixations” in which the eye covered a substantially larger area of the image for longer durations than most fixations. Often these “complex fixations” were observed when monkeys looked at complex objects such as faces and may have included rapid eye movements not clearly distinguishable at a sampling rate of 200 Hz. Therefore, we do not know if Cluster Fix is sufficient for detecting microsaccades or other fixational eye movements often observed at higher sampling frequencies (Martinez-Conde et al., 2009).

Future work could improve Cluster Fix by adding parameters or adaptations to detect eye movements that are neither fixations nor saccades such as smooth pursuit. The largest challenge in detecting smooth pursuit would be determining the best way to consolidate clusters. With Cluster Fix each cluster represents tangible aspects of the scan path. Typically, in our binary classification, we find a saccade cluster with high velocity and low acceleration, and this cluster often represents periods of consecutive time points in the scan path. With the right selection of clusters and duration thresholds we see no reason why one could not identify periods of smooth pursuit with Cluster Fix.

For a given subject, parameter values across multiple images were consistent for fixations, and fixation clusters did not overlap in state space with saccade clusters. This consistency suggests the plausibility of analyzing viewing behavior across multiple images using a single scan path as a template. During a practice trial, fixation and saccade clusters could quickly be determined as well as which areas of state space belong to each cluster using a classification algorithm such as a support vector machine. For subsequent experimental trials, for each time point, the position in state space could be classified as a fixation or saccade. With the addition of a duration threshold, these classifications could be used to determine when fixations and saccades occur in near real-time. Real-time detection of eye movements is essential for gaze-contingent experiments in which trial parameters depend on the subject’s viewing behavior. Other applications of Cluster Fix could include using the algorithm to determine proper velocity and acceleration thresholds for novel or complex tasks instead of trying to determine these thresholds empirically.

Determining the number of parameters to use is an important aspect of parameter selection. We found that the removal of distance from Cluster Fix drastically decreased the specificity and sensitivity of the algorithm for small saccades and made the algorithm more susceptible to noise. Likewise, the removal of angular velocity from the algorithm produced similar results despite mostly distinguishing between periods of fixation with high angular velocity and periods of fixation with low angular velocity. We postulate that the addition of these parameters may help classify time points properly when the other parameters cannot. Alternatively, these additional parameters may help determine the appropriate number and size of clusters in state space. There is no constraint on the number of parameters used in Cluster Fix though a minimum of 2 parameters, namely velocity and acceleration, would be highly recommended.

When choosing state space parameters, we took what we qualitatively observed as differences between fixations and saccades and turned these into quantifiable parameters. Distance was selected because, during a fixation, data points are compact and close to each other while data points are more dispersed during a saccade. Density-dispersion algorithms explicitly take advantage of this parameter. Velocity and acceleration parameters were chosen because the eye moves faster during a saccade than during a fixation. The angular velocity parameter accounts for the smooth linear-like movements of saccades while fixations appear irregular (at 200 Hz) with position fluctuating around an attended location.

A great deal of effort has been expended on determining or developing methods for determining the proper thresholds for threshold-based algorithms. Even if one threshold is found to be optimal for a scan path for one subject in one experiment, that threshold is not guaranteed to be appropriate for another subject, experiment, or even the same subject on another trial. Our data support the idea that thresholds are suboptimal for detecting fixations and saccades. Velocity and acceleration thresholds are relatively fast and easy to implement but suffer from detection errors. Dispersion-based algorithms can be implemented relatively quickly in real-time and appear to detect the presence of fixations with relatively high accuracy. However, dispersion-based algorithms fail to accurately measure the transition times between fixations and saccades which are extremely important in understanding potential neural correlates of eye movements.

A primary goal of this work was to create an algorithm that would be easy to implement across laboratories for a variety of behavioral tasks. For this reason, the algorithm was devised in MATLAB using commonly available functions. However, this algorithm may be enhanced through the use of better clustering algorithms as well as alternative ways for determining the appropriate number of clusters such as statistical measures (e.g. explanation of variance), visual inspection, the separation of clusters, and stability of clusters with resampling (Ben-Hur et al., 2002; Dudoit and Fridlyand, 2002; Pham et al., 2005). Cluster Fix requires approximately 240 seconds to analyze 720 seconds of eye data sampled at 200 Hz (MATLAB 2012b, Intel Core Xeon Processor 2.80 GHz with 16 GB RAM). This is significantly slower than an algorithm using velocity and acceleration thresholds which requires only a few seconds to analyze the same amount of data. The two dispersion-based algorithms analyzed scan paths substantially slower than Cluster Fix, but optimization of these algorithms could substantially increase the speed of analysis. In Cluster Fix nearly 80% of the computation time is devoted to determining the number of clusters of which nearly 90% is devoted to local re-clustering. Cluster Fix sacrifices time for a significant increase in sensitivity and specificity. Cluster Fix may run faster by parallelizing local re-clustering or compiling Cluster Fix in a different programming language optimized for cluster analysis.

One limitation of the k-means clustering algorithm is that the clusters are determined by the distance between the points in the clusters and the clusters' centroids. K-means initiates with random points and iteratively computes clusters until the distance between points in the clusters to the centroid of the clusters converges. The convergence most often occurs at a local minimum. To increase the probability that a better minimum is found, k-means can be replicated several times. We used 5 replicates for determining the appropriate number of

clusters and for performing k-means to cluster all the time points. Occasionally, only poor local minima are found or convergence does not occur in a reasonable number of computations thus producing less than optimal clustering. This produces some variability in the results and affects the detection of smaller saccades since these are less distinguishable from noise. For this reason, each time Cluster Fix is applied to the same scan path the results may vary slightly. If one wanted to be absolutely certain that they detected all of the smaller amplitude saccades, Cluster Fix results could be averaged over several applications. Additionally, more replications could be performed to increase the consistency of the results but at a cost of computational time. Other clustering algorithms may solve some of the limitations of k-means clustering, but previous attempts to use hierarchical clustering were unsuccessful.

6. Conclusion

A primary advantage of Cluster Fix is the ability to properly detect the onset and offset of saccades. Several neural mechanisms are aligned in time to saccades including saccadic suppression in the LGN (Ross et al., 2001), modulation in firing rates of neurons outside of the early visual areas (Krekelberg et al., 2003; Crapse and Sommer, 2012), and phase reset of thetaoscillations (Rajkai et al., 2008; Jutras and Buffalo, 2010; Hoffman et al., 2013; Jutras and Buffalo, 2013; Jutras et al., 2013). More accurate scan path analysis may help improve the detection of these neural phenomena as well as reduce variability in data possibly attributable to improper saccade detection. Finally, the general concepts behind Cluster Fix could be extended to the tracking of animal positions, complex dynamical systems, and the motion of body parts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors would like to thank Michael Jutras and Nathan Killian for the basic MATLAB code used to analyze eye tracking data; Esther Tonea and William Li for the creation of image sets; and Megan Jutras for helping with collecting, organizing, and analyzing the behavioral data. We would also like to thank Niklas Wilming for his helpful comments on the manuscript. Funding for this work was provided by the National Institute of Mental Health Grants MH080007 (to E.A.B.) and MH093807 (to E.A.B.); National Center for Research Resources Grant P51RR165 (currently the Office of Research Infrastructure Programs/OD P51OD11132); and the National Institute of Health 5T90DA032466-02 and 5T90DA032436-03 (S.D.K).

References

- Andrews TJ, Coppola DM. Idiosyncratic characteristics of saccadic eye movements when viewing different visual environments. *Vision Research*. 1999; 39:2947–2953. [PubMed: 10492820]
- Ben-Hur A, Elisseeff A, Guyon I. A stability based method for discovering structure in clustered data. *Pac Symp Biocomput*. 2002:6–17. [PubMed: 11928511]
- Berg DJ, Boehnke SE, Marino RA, Munoz DP, Itti L. Free viewing of dynamic stimuli by humans and monkeys. *J Vis*. 2009; 9:19, 11–15. [PubMed: 19757897]
- Crapse TB, Sommer MA. Frontal eye field neurons assess visual stability across saccades. *J Neurosci*. 2012; 32:2835–2845. [PubMed: 22357866]

- Crutcher MDC-HR, Manzanares CM, Lah JJ, Levey AI, Zola SM. Eye tracking during a visual paired comparison task as a predictor of early dementia. *Am J Alzheimers Dis Other Demen.* 2009; 24:258–266. [PubMed: 19246573]
- Dewhurst RNM, Jarodzka H, Foulsham T, Johansson R, Holmqvist K. It Depends on How You Look at It: Scanpath Comparison in Multiple Dimensions with MultiMatch, a Vector-based Approach. *Behavior research methods.* 2012; 44:1079–1100. [PubMed: 22648695]
- Duchowski AT. A breadth-first survey of eye-tracking applications. *Behav Res Meth Ins C.* 2002; 34:455–470.
- Duchowski, AT. *Eye Tracking Methodology: Theory and Practice.* 2nd Edition. Springer; 2007.
- Dudoit S, Fridlyand J. A prediction-based resampling method for estimating the number of clusters in a dataset. *Genome Biol.* 2002; 3
- Engbert R, Kliegl R. Microsaccades uncover the orientation of covert attention. *Vision Res.* 2003; 43:1035–1045. [PubMed: 12676246]
- Hannula DE, Ranganath C. The Eyes Have It: Hippocampal Activity Predicts Expression of Memory in Eye Movements. *Neuron.* 2009; 63:592–599. [PubMed: 19755103]
- Hannula DE, Baym CL, Warren DE, Cohen NJ. The Eyes Know: Eye Movements as a Veridical Index of Memory. *Psychol Sci.* 2012; 23:278–287. [PubMed: 22327015]
- Hannula DE, Althoff RR, Warren DE, Riggs L, Cohen NJ, Ryan JD. Worth a glance: using eye movements to investigate the cognitive neuroscience of memory. *Front Hum Neurosci.* 2010; 4
- Hayhoe M, Ballard D. Eye movements in natural behavior. *Trends Cogn Sci.* 2005; 9:188–194. [PubMed: 15808501]
- Hoffman KL, Dragan MC, Leonard TK, Micheli C, Montefusco-Siegmund R, Valiante TA. Saccades during visual exploration align hippocampal 3–8 Hz rhythms in human and non-human primates. *Frontiers in Systems Neuroscience.* 2013; 7
- Ito J, Maldonado P, Singer W, Grun S. Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. *Cereb Cortex.* 2011; 21:2482–2497. [PubMed: 21459839]
- Jutras MJ, Buffalo EA. Recognition memory signals in the macaque hippocampus. *Proc Natl Acad Sci U S A.* 2010; 107:401–406. [PubMed: 20018683]
- Jutras MJ, Buffalo EA. Oscillatory correlates of memory in non-human primates. *Neuroimage.* 2013
- Jutras MJ, Fries P, Buffalo EA. Gamma-band synchronization in the macaque hippocampus and memory formation. *J Neurosci.* 2009; 29:12521–12531. [PubMed: 19812327]
- Jutras MJ, Fries P, Buffalo EA. Oscillatory activity in the monkey hippocampus during visual exploration and memory formation. *Proc Natl Acad Sci.* 2013; 110:13144–13149. [PubMed: 23878251]
- Killian NJ, Jutras MJ, Buffalo EA. A map of visual space in the primate entorhinal cortex. *Nature.* 2012; 491:761–764. [PubMed: 23103863]
- Kimmel DL, Mammo D, Newsome WT. Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Front Behav Neurosci.* 2012; 6:49. [PubMed: 22912608]
- Komogortsev OV, Karpov A. Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior research methods.* 2013; 45:203–215. [PubMed: 22806708]
- Komogortsev OV, Gobert DV, Jayarathna S, Koh DH, Gowda SM. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *Ieee T Bio-Med Eng.* 2010; 57:2635–2645.
- Krekelberg B, Kubischik M, Hoffmann KP, Bremmer F. Neural correlates of visual localization and perisaccadic mislocalization. *Neuron.* 2003; 37:537–545. [PubMed: 12575959]
- Lagun D, Manzanares C, Zola SM, Buffalo EA, Agichtein E. Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. *J Neurosci Meth.* 2011; 201:196–203.
- Lee B, Pesaran B, Andersen RA. Area MSTd neurons encode visual stimuli in eye coordinates during fixation and pursuit. *J Neurophysiol.* 2011; 105:60–68. [PubMed: 20980545]
- Liston DB, Krukowski AE, Stone LS. Saccade detection during smooth tracking. *Displays.* 2012

- Martinez-Conde S, Otero-Millan J, Macknik SL. The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nat Rev Neurosci*. 2013; 14:83–96. [PubMed: 23329159]
- Martinez-Conde S, Macknik SL, Troncoso XG, Hubel DH. Microsaccades: a neurophysiological analysis. *Trends Neurosci*. 2009; 32:463–475. [PubMed: 19716186]
- McAlonan K, Cavanaugh J, Wurtz RH. Guarding the gateway to cortex with attention in visual thalamus. *Nature*. 2008; 456:391–394. [PubMed: 18849967]
- Najemnik J, Geisler WS. Optimal eye movement strategies in visual search. *Nature*. 2005; 434:387–391. [PubMed: 15772663]
- Nystrom M, Holmqvist K. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behav Res Methods*. 2010; 42:188–204. [PubMed: 20160299]
- Otero-Millan J, Troncoso XG, Macknik SL, Serrano-Pedraza I, Martinez-Conde S. Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. *J Vis*. 2008; 8:21. 21-18. [PubMed: 19146322]
- Pham DT, Dimov SS, Nguyen CD. Selection of K in K-means clustering. *P I Mech Eng C-J Mec*. 2005; 219:103–119.
- Rajkai C, Lakatos P, Chen CM, Pincze Z, Karmos G, Schroeder CE. Transient cortical excitation at the onset of visual fixation. *Cerebral Cortex*. 2008; 18:200–209. [PubMed: 17494059]
- Rayner K. Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol*. 2009; 62:1457–1506.
- Richmond J, Nelson CA. Relational memory during infancy: evidence from eye tracking. *Dev Sci*. 2009; 12:549–556. [PubMed: 19635082]
- Ross J, Morrone MC, Goldberg ME, Burr DC. Changes in visual perception at the time of saccades. *Trends in Neurosciences*. 2001; 24:113–121. [PubMed: 11164942]
- Salvucci, DD.; Goldberg, JH. Proceedings of the 2000 symposium on Eye tracking research & applications. Palm Beach Gardens, Florida, USA: ACM; 2000. Identifying fixations and saccades in eye-tracking protocols; p. 71-78.
- Shic, F.; Scassellati, B.; Chawarska, K. Proceedings of the 2008 symposium on Eye tracking research & applications. Savannah, Georgia: ACM; 2008. The incomplete fixation measure; p. 111-114.
- Smith CN, Squire LR. Experience-dependent eye movements reflect hippocampus-dependent (aware) memory. *J Neurosci*. 2008; 28:12825–12833. [PubMed: 19036976]
- Smith CN, Hopkins RO, Squire LR. Experience-dependent eye movements, awareness, and hippocampus-dependent memory. *J Neurosci*. 2006; 26:11304–11312. [PubMed: 17079658]
- Tatler BWBR, Gilchrist ID. Visual correlates of fixation selection: effects of scale and time. *Vision Research*. 2005; 45:643–659. [PubMed: 15621181]
- Urruty T, Lew S, Ithadaddene N, Simovici DA. Detecting eye fixations by projection clustering. *ACM Trans Multimedia Comput Commun Appl*. 2007; 3:1–20.
- Wass SV, Smith TJ, Johnson MH. Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behav Res Methods*. 2013; 45:229–250. [PubMed: 22956360]
- Yarbus, A. Eye movements and vision. Vol. 1967. New York: 1967.
- Zola SM, Manzanares CM, Clopton P, Lah JJ, Levey AI. A behavioral task predicts conversion to mild cognitive impairment and Alzheimer's disease. *Am J Alzheimers Dis Other Demen*. 2013; 28:179–184. [PubMed: 23271330]

Highlights

- Standard saccade detection algorithms are insufficient for complex viewing tasks
- Cluster Fix detects saccades with k-means cluster analysis in scan path state space
- Global and local criteria increase the accuracy of saccade detection
- A gold standard is required to compare viewing behavior across experiments

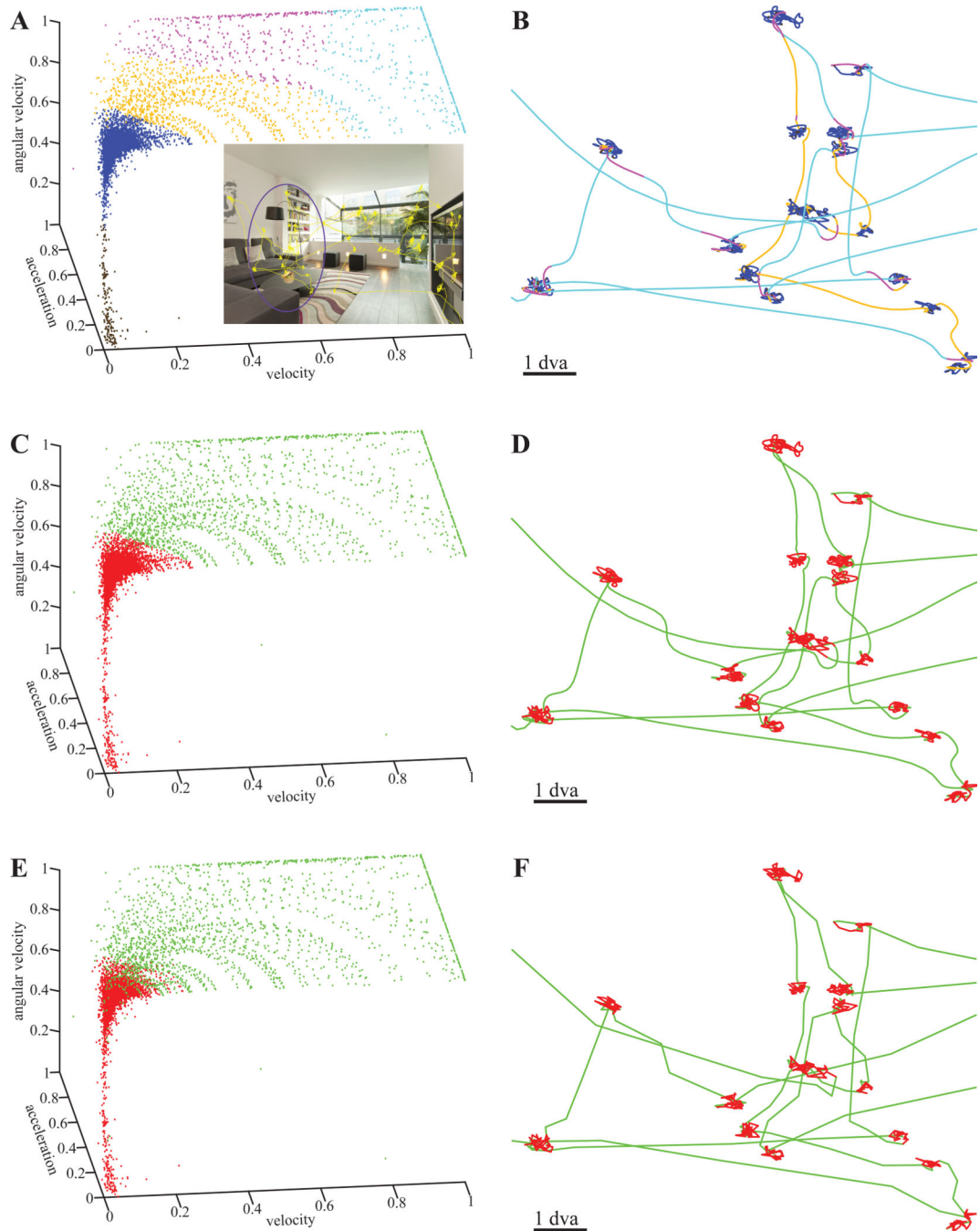


Figure 1. Global clustering in scan path state space

A) Global clustering identified the appropriate number of clusters in 4 state space parameters—distance (not shown), velocity, acceleration, and angular velocity—normalized from 0 to 1. Each dot represents a single time point (1 ms) from a representative scan path (yellow lines overlaying image in inset) with each color representing a different cluster. Blue and brown clusters represent time points with low acceleration and velocity. These two clusters represent the two states of fixation (high and low angular velocity). Pink, gold, and turquoise clusters represent clusters with higher velocity and acceleration. These three

clusters represent the three states of saccades. **B)** Representative section of the up-sampled scan path globally clustered in A (indicated by purple circle in inset image). Each cluster represents a different portion of the scan path across multiple fixations or saccades. **C)** Fixation clusters (red) were determined as the clusters with the lowest velocity and acceleration. All other clusters were classified as saccades (green). **D)** The same portion of the scan path now parsed into fixations and saccades. **E)** Following local re-clustering many of the fixation time points were reclassified as the beginning and end of saccades. Additionally, reclassification reduced the size of the fixation cluster and resulted in some fixation and saccade time points overlapping in state space. **F)** Representative section of the final raw scan path parsed into fixations and saccades.

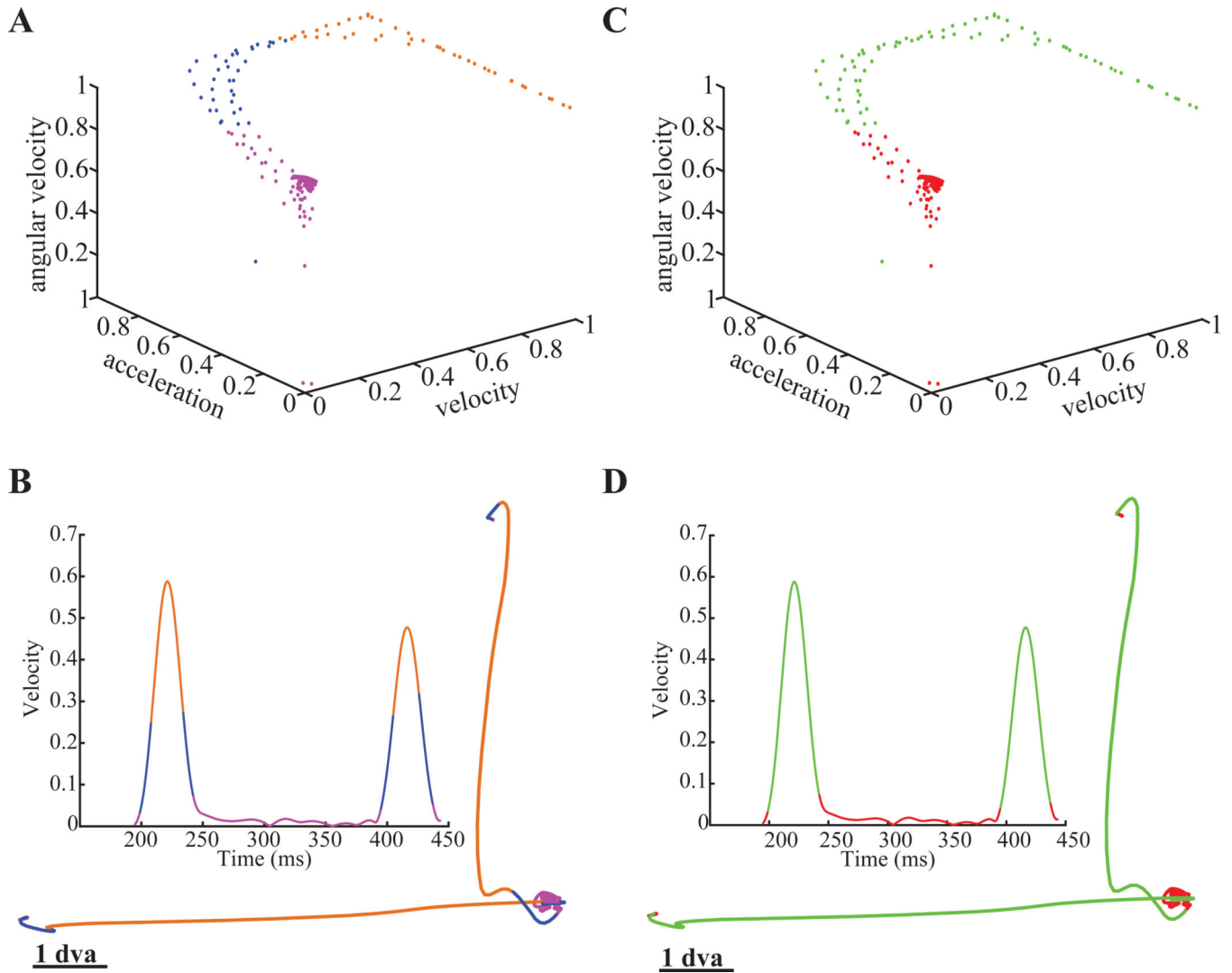


Figure 2. Local re-clustering of a detected fixation in state space

A) Initially, local reclustering identified the appropriate number of clusters between a fixation and surrounding portions of the scan path typically including 2 saccades. **B)** The velocity profile from the same portion of scan path. (**Inset**) Each cluster represented a different portion of the up-sampled scan path. **C–D** show the clusters, scan path, and velocity profile now classified as fixations (red) or saccades (green). Cluster Fix classified the clusters with the lowest velocity and acceleration as fixations. All other clusters were classified as saccades.

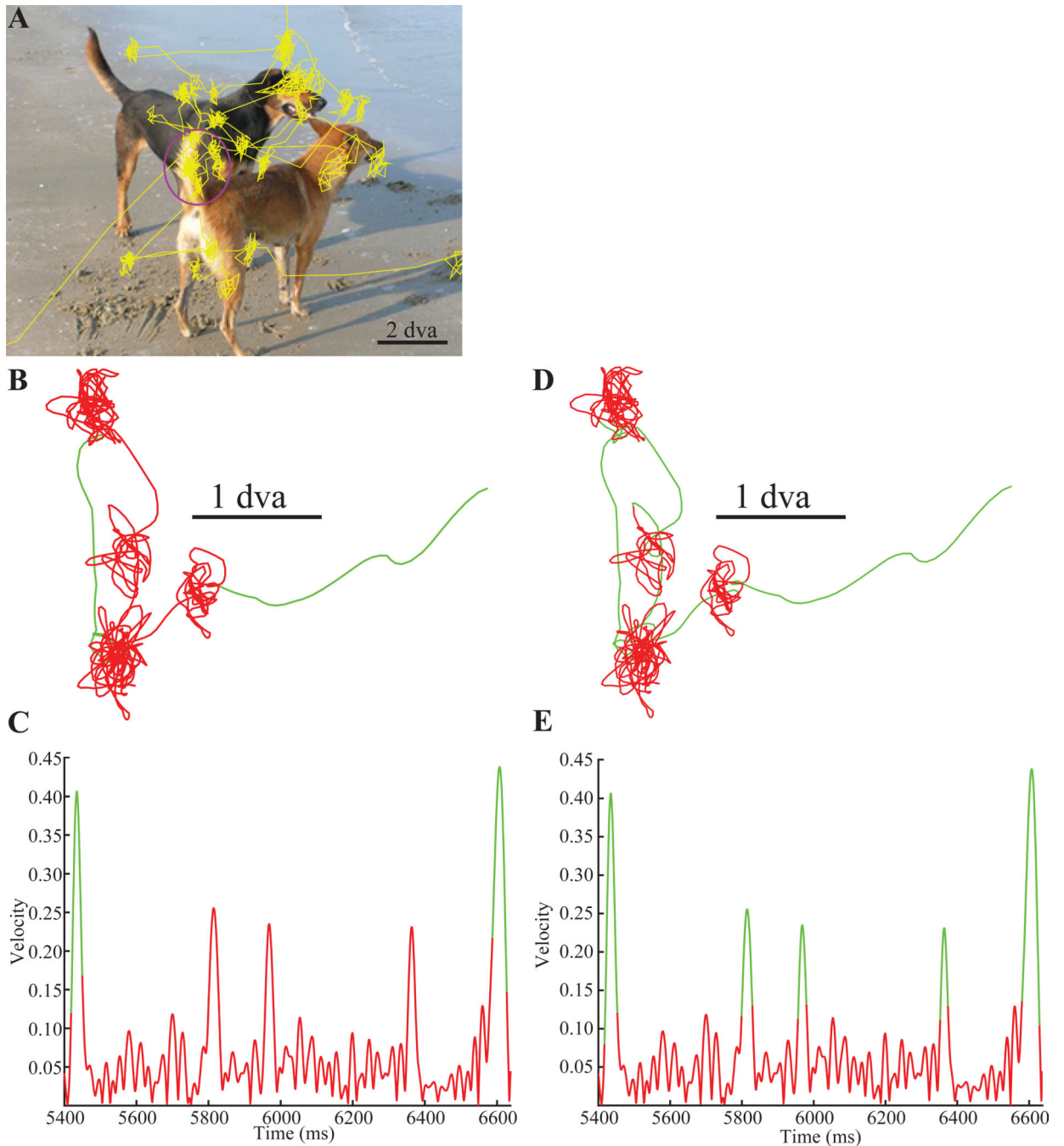


Figure 3. Local re-clustering can detect smaller saccades hidden by large global variability

A) A representative scan path (yellow) overlaying the viewed image. **B)** In a portion of the scan path (purple circle in A), global clustering only identified a single fixation (red) surrounded by 2 larger saccades (green). **C)** The velocity profile during this portion of the scan path revealed that the fixation detected by global clustering was highly variable with several potential saccades having velocities just above this variance. **D–E)** Local re-clustering of the fixation identified 3 additional saccades, which separated the single fixation detected by global clustering into 4 distinct fixations. Local re-clustering also reclassified

some of the fixation time points near the two larger saccades—initially detected by global clustering—as saccades.

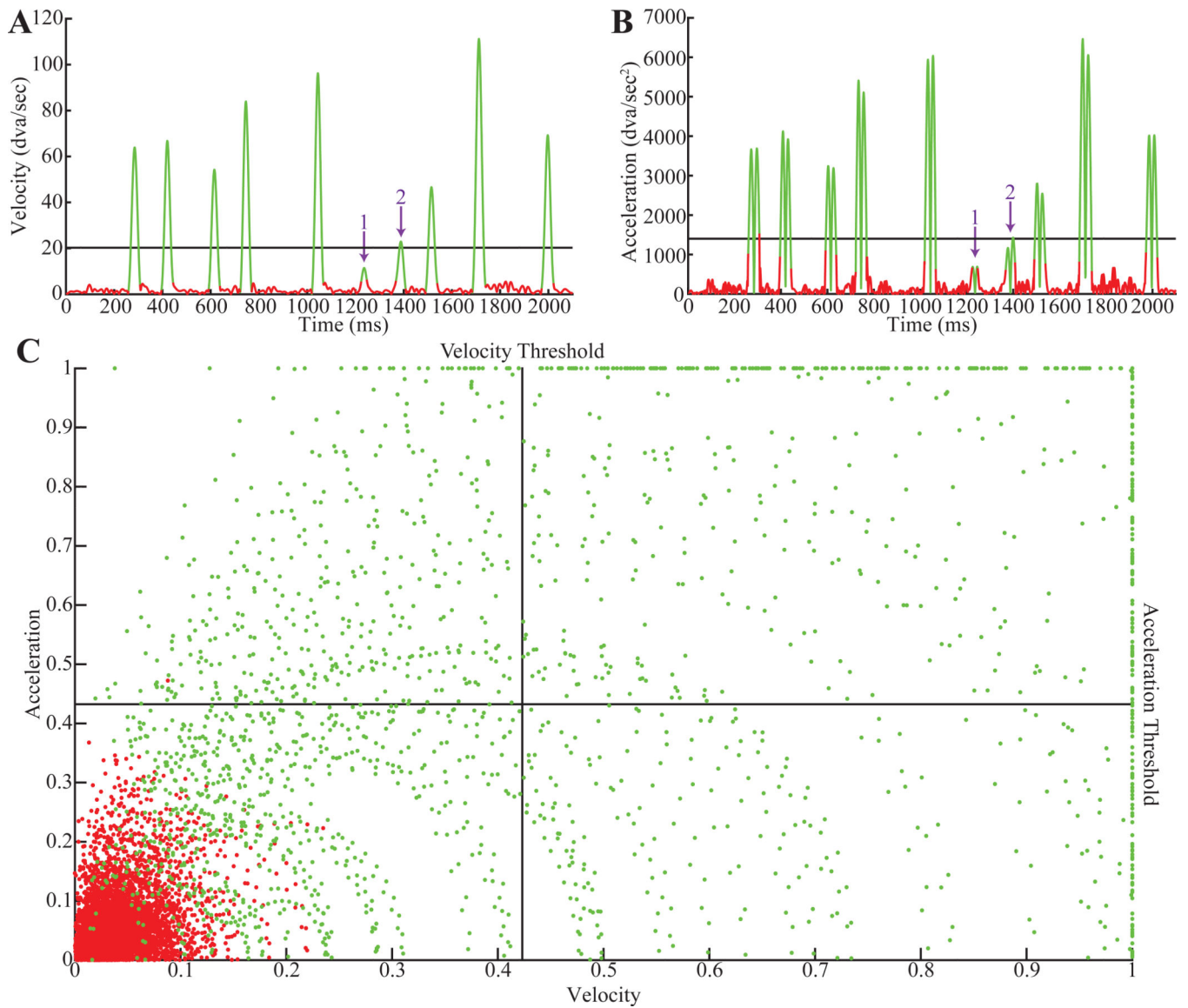


Figure 4. Thresholds vs. Cluster Fix

Fixations (red) and saccades (green) determined through Cluster Fix for an example segment of a scan path. Velocity (A) and acceleration (B) thresholds (black lines, mean + std) appear to miss smaller amplitude saccades either due to velocity or acceleration values not exceeding the threshold (purple ↓ #1) or values did not exceed the threshold for a sufficient duration (purple ↓ #2). C) A 2D state space plot indicates that single-value thresholds may misclassify fixation and saccade time points which may be critical to properly identify the onset and offset of saccades. Note that because velocity and accelerations values were normalized to the mean plus 3 times the standard deviation many saccade time points are located on the edge of the plot. Also, note that up-sampling and low pass filtering creates an appearance of arc-like patterns in state space.

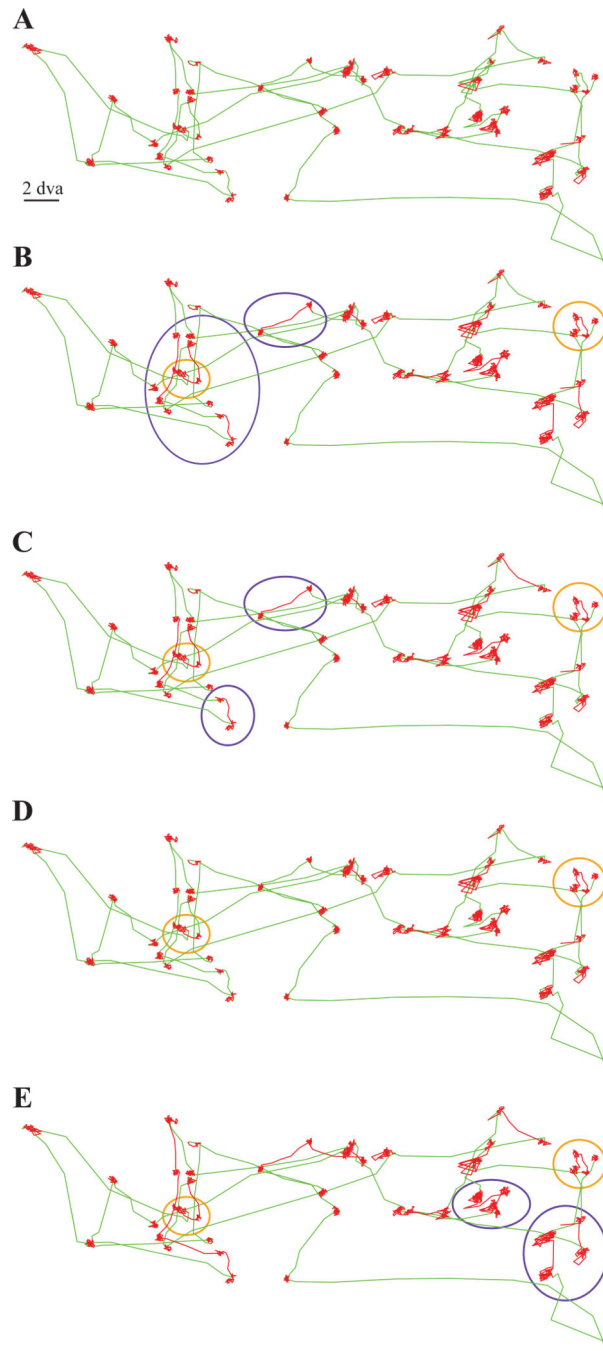


Figure 5. Visual comparison of detected fixations and saccades by algorithm
 Compared to Cluster Fix (A), velocity and acceleration threshold algorithms (B–E) appeared unable to detect smaller saccades (red: fixations, green: saccades). The velocity and acceleration algorithms are as follows: B) mean + standard deviation, C) $30^\circ/\text{s}$ & $8000^\circ/\text{s}^2$, D) $75^\circ/\text{s}$ & $5000^\circ/\text{s}^2$, and E) $100^\circ/\text{s}$ & $8500^\circ/\text{s}^2$. Purple circles highlight some of the key discrepancies between Cluster Fix and the threshold-based algorithms, and orange circles indicate potential saccades commonly missed across multiple threshold-based algorithms.

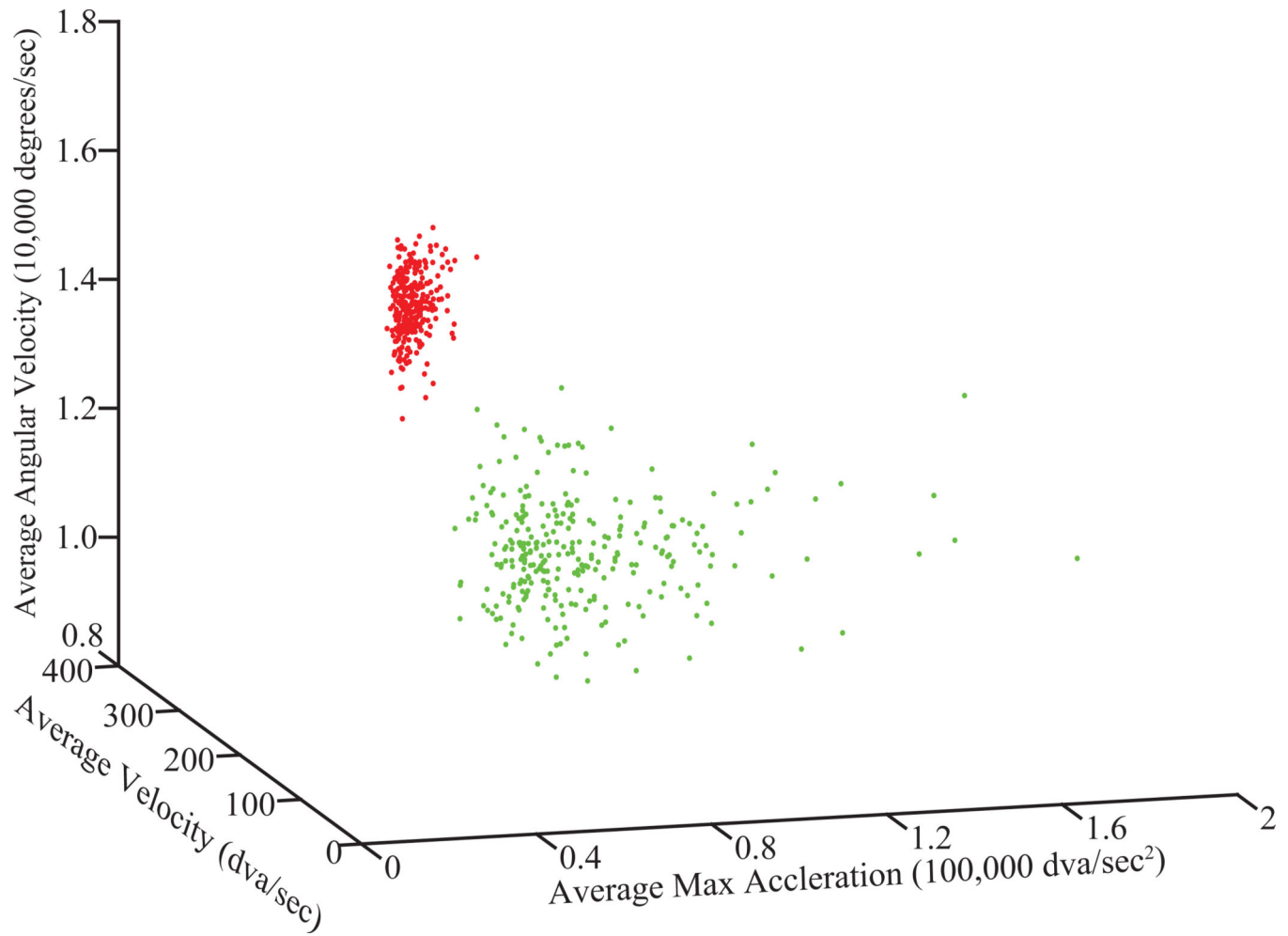


Figure 6. Single subject cluster means were consistent across scan paths

The average velocity, peak acceleration, and angular velocity during fixations (red) across multiple scan paths from a monkey viewing 288 different images remained consistent, but values for these parameters during saccades (green) were more variable. Each dot represents the average value for all fixations or saccades from one scan path.

Table 1

Cluster Fix Procedural Outline

1	Pre-process and Filter
	a. Up-sample horizontal and vertical eye traces from 200 Hz to 1000 Hz
	b. Low pass filter with a cutoff frequency of 30 Hz
2	Calculate distance, velocity, acceleration, and angular velocity for every time point
3	Move Outliers and Normalize
	a. Move outliers greater than the mean + 3*std to the mean + 3*std
	b. Individually normalize the 4 state space parameters to be from 0 to 1
4	Global Clustering
	a. Determine the number of clusters
	b. Cluster using k-means
	c. Determine fixation clusters and saccades clusters
	d. Reclassify fixations with durations less than 25 ms as saccades
5	Local Re-clustering
	a. Compare detected fixations to adjacent portions of the scan path
	b. Determine the number of clusters
	c. Cluster using k-means
	d. Determine fixation clusters and saccade clusters
6	Reclassify global fixation time points that were locally determined to be saccades
7	Consolidate using duration thresholds
	a. Classify fixations with durations less than 5 ms as saccades
	b. Reclassify saccades with durations less than 10 ms as fixations
	c. Reclassify fixations with durations less than 25 ms as saccades
8	Post-processing
	a. Down-sample to acquisition frequency of 200 Hz

Table 2

Computed Behavioral Statistics by Algorithm and Threshold

Algorithm	Threshold	Fixation Duration (ms)	Saccade Duration (ms)	Saccade Arc Length (dva)	Distance Between Fixations (dva)	Fixations per Image	Instantaneous Saccade Rate (Hz)
Cluster Fix	N/A	177.0 ± 97.0	52.4 ± 12.8	7.5 ± 4.6	6.0 ± 4.6	35.8 ± 7.6	4.6 ± 1.8
	mean + std (global*)	296.9 ± 185.7	48.5 ± 7.0	9.1 ± 4.6	8.0 ± 4.5	24.3 ± 5.4	3.4 ± 1.4
Velocity and Acceleration	30 °/s	8,000 °/s ²	280.8 ± 179.3	54.5 ± 9.4	9.7 ± 4.7	24.5 ± 6.3	3.5 ± 1.5
	75 °/s	5,000 °/s ²	247.3 ± 144.5	50.7 ± 7.9	8.5 ± 4.5	28.6 ± 6.3	3.8 ± 1.4
	100 °/s	8,500 °/s ²	315.1 ± 201.8	46.8 ± 7.5	9.7 ± 4.7	22.3 ± 6.0	3.3 ± 1.4
Dispersion	0.5 dva	137.2 ± 80.4	37.6 ± 18.3	5.9 ± 4.6	5.0 ± 4.6	44.5 ± 9.3	6.1 ± 3.8
	0.75 dva	183.5 ± 94.2	32.6 ± 13.0	6.1 ± 4.5	5.8 ± 4.5	37.9 ± 7.5	4.9 ± 2.5
	1 dva	215.7 ± 107.3	29.2 ± 10.7	6.3 ± 4.5	6.4 ± 4.5	33.9 ± 6.6	4.3 ± 1.7
	1.25 dva	242.3 ± 121.2	26.6 ± 9.2	6.4 ± 4.6	6.8 ± 4.6	31.1 ± 6.5	4.1 ± 1.5
	mean + std (local [†])	112.18 ± 48.2	27.1 ± 6.6	3.3 ± 3.1	3.5 ± 4.4	65.3 ± 8.2	7.6 ± 2.5
Minimum Spanning Tree	0.5 dva	122.2 ± 70.9	52.1 ± 23.5	6.3 ± 4.9	4.7 ± 4.6	47.4 ± 8.8	6.2 ± 3.2
	0.75 dva	154.9 ± 82.7	45.0 ± 15.2	6.5 ± 4.6	5.4 ± 4.6	41.1 ± 7.8	5.2 ± 2.4
	1 dva	183.3 ± 93.0	42.3 ± 12.2	6.8 ± 4.4	6.1 ± 4.5	35.8 ± 7.0	4.5 ± 1.7
	1.25 dva	207.6 ± 106.3	39.8 ± 10.0	6.9 ± 4.4	6.3 ± 4.5	34.5 ± 6.6	4.4 ± 1.6

Values are mean ± std

* Global refers to calculating threshold based on the entire scan path.

[†] Local refers to calculating the threshold for the window being analyzed. Window was 100 ms long. A larger window size (e.g. 200 ms) will produce results similar to a threshold of 0.75 dva.