Unsupervised clustering method to detect microsaccades

Jorge Otero-Millan
Department of Neurobiology, Barrow Neurological Institute, Phoenix, AZ, USA
Department of Signal Theory and Communications, University of Vigo, Vigo, Spain

Jose L. Alba Castro
Department of Signal Theory and Communications, University of Vigo, Vigo, Spain

Stephen L. Macknik
Department of Neurosurgery, Barrow Neurological Institute, Phoenix, AZ, USA
Department of Neurobiology, Barrow Neurological Institute, Phoenix, AZ, USA

Susana Martinez-Conde
Department of Neurobiology, Barrow Neurological Institute, Phoenix, AZ, USA

Microsaccades, small involuntary eye movements that occur once or twice per second during attempted visual fixation, are relevant to perception, cognition, and oculomotor control and present distinctive characteristics in visual and oculomotor pathologies. Thus, the development of robust and accurate microsaccade-detection techniques is important for basic and clinical neuroscience research. Due to the diminutive size of microsaccades, however, automatic and reliable detection can be difficult. Current challenges in microsaccade detection include reliance on set, arbitrary thresholds and lack of objective validation. Here we describe a novel microsaccade-detecting method, based on unsupervised clustering techniques, that does not require an arbitrary threshold and provides a detection reliability index. We validated the new clustering method using real and simulated eye-movement data. The clustering method reduced detection errors by 62% for binocular data and 78% for monocular data, when compared to standard contemporary microsaccade-detection techniques. Further, the clustering method's reliability index was correlated with the microsaccade-detection error rate, suggesting that the reliability index may be used to determine the comparative precision of eye-tracking devices.

Introduction

Saccades are rapid eye movements that change the line of sight between successive points of fixation during visual scanning of a scene. Their range of behaviors encompasses both voluntary and involuntary shifts of fixation (Leigh & Zee, 2006; McCamy, Macknik, & Martinez-Conde, 2014). Small involuntary saccades, called microsaccades, interrupt attempted fixation of a visual target once or twice a second (Martinez-Conde, Macknik, & Hubel, 2004; Martinez-Conde, Otero-Millan, & Macknik, 2013; Rolfs, 2009; see Figure 1A, B).

Recent years have seen a surge in microsaccade research (Martinez-Conde et al., 2004; Martinez-Conde et al., 2013; Rolfs, 2009). Microsaccades have been linked to the perceptual restoration of faded images (Costela, McCamy, Macknik, Otero-Millan, & Martinez-Conde, 2013; Martinez-Conde, Macknik, Troncoso, & Dyar, 2006; McCamy et al., 2012; Troncoso, Macknik, & Martinez-Conde, 2008), the visual scanning of small targets (Otero-Millan, Macknik, Langston, & Martinez-Conde, 2013), and the correction of fixation position (Otero-Millan, Macknik, & Martinez-Conde, 2012; Otero-Millan et al., 2011), among other functions (Martinez-Conde et al., 2013). Microsaccades can reflect attentional shifts (Engbert & Kliegl, 2003; Hafed & Clark, 2002), modulate the activity of visual neurons (Martinez-Conde, Macknik, & Hubel, 2000; 2002), and show distinctive characteristics in patients affected with ophthalmic and neurological diseases (Chen et al., 2010; Kapoula et al., in press; Martinez-Conde, 2006; Otero-Millan, Schneider, Leigh, Mack-

Citation: Otero-Millan, J., Alba Castro, J. L., Macknik, S. L., & Martinez-Conde, S. (2014). Unsupervised clustering method to detect microsaccades. *Journal of Vision, 14*(2):18, 1–17, http://www.journalofvision.org/content/14/2/18, doi:10.1167/14.2.18.

doi: 10.1167/14.2.18 Received May 2, 2013; published February 25, 2014 ISSN 1534-7362

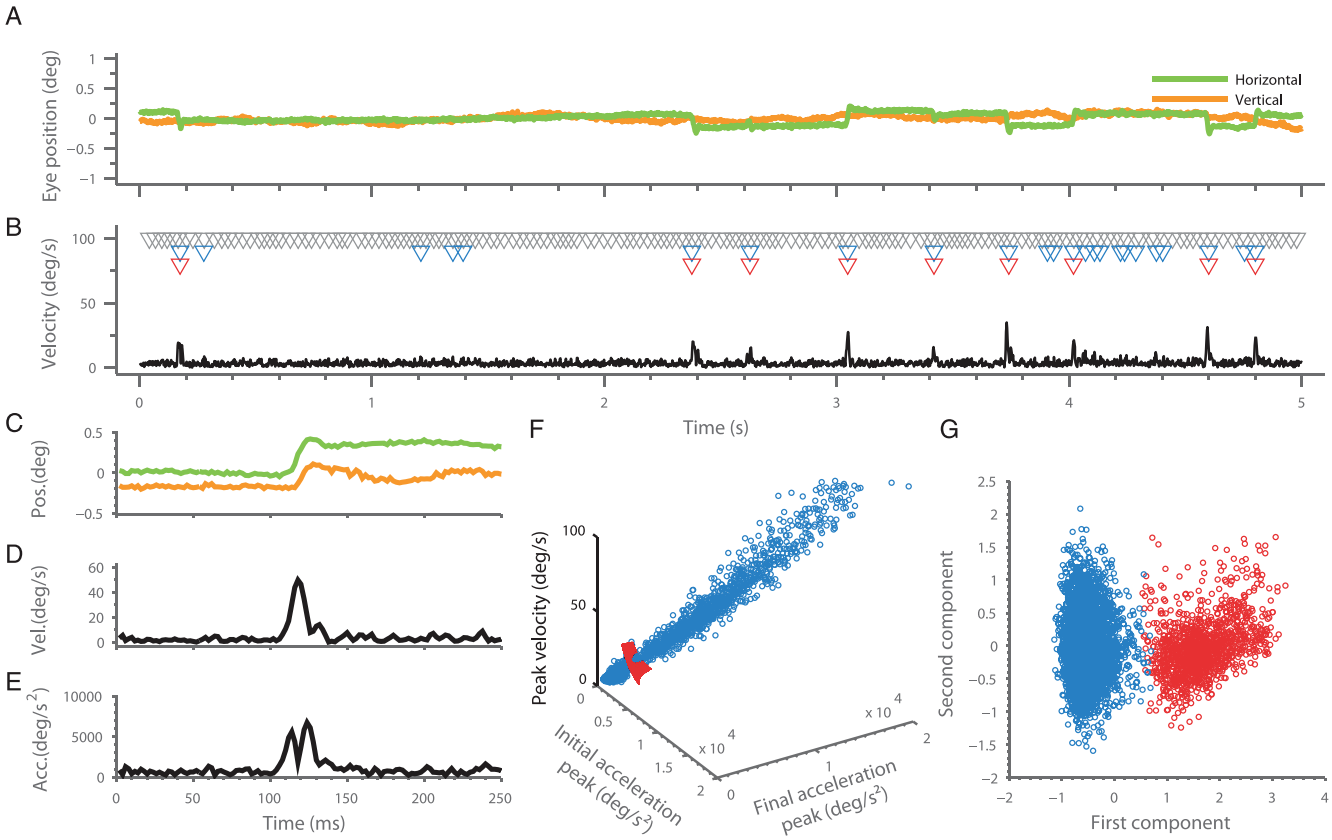

Figure 1. New clustering method for microsaccade detection. (A) 5 s of eye-position recordings. (B) Eye velocity. Gray triangles indicate velocity peaks, blue triangles indicate microsaccade candidates, and red triangles indicate microsaccades identified by the clustering method. We note that we chose the trace example in (A) specifically to illustrate borderline cases, where the velocity peaks accompanying microsaccades are not necessarily clear. (C–E) Eye position, velocity magnitude, and acceleration magnitude during a microsaccade. (F) Scatter plot showing peak velocity, initial acceleration peak, and final acceleration peak for all microsaccade candidates in one recording. The red surface represents the boundary that separates microsaccades from noisy events. (G) Scatter plot showing two uncorrelated components of the features used in the clustering after normalization. Red dots indicate microsaccades and blue dots indicate noisy events.

nik, & Martinez-Conde, 2013; Otero-Millan et al., 2011; Serra, Liao, Martinez-Conde, Optican, & Leigh, 2008). Thus, improved microsaccade detection has great potential value for the visual and oculomotor research community.

Microsaccade magnitudes may range from 3 min of arc (Zuber, Stark, & Cook, 1965) to 1° (Engbert, 2006; Otero-Millan, Troncoso, Macknik, Serrano-Pedraza, & Martinez-Conde, 2008), and microsaccade velocities range from 3°/s to 100°/s. Typical microsaccade durations are around 20 ms. These and other microsaccadic parameters can vary greatly across subjects, however.

All saccades, including microsaccades, may be distinguished from other eye movements (e.g., slow intersaccadic drift) based on their higher velocity. Thus, automatic saccade-detection methods have relied typically on simple velocity thresholds (e.g., 20°/s; Bahill, Brockenbrough, & Troost, 1981). See Komogortsev, Jayarathna, Koh, and Gowda (2010) and Salvucci and Goldberg (2000) for reviews on saccade-detection methods.

Variable levels of noise in the recording system (e.g., in the signal itself, video, or voltage) and physiological artifacts (including head movements, or changes in pupil size in the case of video trackers) can hinder automatic saccade detection. Modern video-based eye trackers are more susceptible to physiological and nonphysiological sources of noise than classical methods such as the scleral search-coil technique (Robinson, 1963). Due to their small magnitudes and relatively slow speeds, microsaccades are among the most challenging saccades to identify reliably.

Engbert and Kliegl (2003) pioneered new saccade-detection methods that are robust to variable recording conditions and sensitive enough to detect microsaccades. Their automatic microsaccade-detecting method has been the workhorse for the entire field for a decade, contributing critically to the proliferation of human microsaccade studies and rapid replication of results

(Martinez-Conde et al., 2013). A critical feature of this method is a velocity threshold that adapts to the level of noise in the data (see Methods and Engbert & Kliegl, 2003, for a detailed description). A subsequent refinement introduced the option of considering only binocular microsaccades (i.e., microsaccades occurring simultaneously in both eyes; Engbert, 2006) to reduce potentially false positives (a prevalent problem in velocity-threshold-based detection methods; Nyström & Holmqvist, 2010).

More recently, Bettenbühl et al. (2009) developed a new method that combines wavelet and principal-components analyses, which characterizes the waveforms of microsaccades in a set of fundamental components. Not broadly used as of this writing, this method is restricted to horizontal eye movements and requires blink-free recordings. Alternative microsaccade-detection methods have used a combination of velocity and direction-change thresholds (Martinez-Conde et al., 2000) or velocity and acceleration thresholds with subsequent examination by an expert to reduce detection errors (Hafed, Goffart, & Krauzlis, 2009).

Here we set out to develop a novel method for the detection of small saccades in noisy eye-movement records that—unlike previous methods—did not require the setting of a particular threshold. We developed a new unsupervised method to detect saccadic eye movements, including the smallest microsaccades produced during attempted fixation, based on clustering techniques. This method has three main advantages over present microsaccade-detecting methods: First, it provides an index of the reliability of the detection outcome (related to the recording's signal-to-noise ratio). Second, it does not require the setting of an arbitrary threshold. Rather, it characterizes both microsaccades and the level of noise in the data to automatically find a boundary between them. Third, it does not rely strongly on the binocularity of microsaccades.

We validated the performance of this clustering method against the method developed by Engbert and Kliegl (2003)—which we will refer to as the E&K method for brevity—using eye-movement recordings from a commercial video tracker, and we found an improved and more robust performance.

Methods

A new method for microsaccade detection

Microsaccades are the fastest eye movements during attempted fixation, a feature that facilitates their detection. Microsaccade rates range typically between one and four per second (Martinez-Conde et al., 2004; Martinez-Conde, Macknik, Troncoso, & Hubel, 2009; Otero-Millan, Macknik, et al., 2013; Otero-Millan et al., 2008), a feature not used by current detection methods. The new detection method presented here relies on both microsaccade velocities and rates.

In any classification problem it is important to characterize not only the events of interest (e.g., microsaccades) but also those events that can be identified erroneously as events of interest (e.g., noise arising from the measuring system or from physiological sources). The proposed method automatically selects a set of candidate events that contains all the true microsaccades plus an undetermined, but upper bounded, number of nonmicrosaccade events, and then uses an unsupervised clustering technique to find the boundary between the two event populations. An implementation of the new method is available for download at: http://smc.neuralcorrelate.com/software/microsaccade-detection/.

Event detection

The proposed method identifies peaks of high velocity in the eye-movement recordings as potential microsaccade candidates. Microsaccade rates are typically below four per second (Martinez-Conde et al., 2004). Thus, we select only the highest velocity peaks necessary to obtain a rate of microsaccade candidates of five per second, so as to ensure the inclusion of both true microsaccades and nonmicrosaccadic eye movements.

We estimate the horizontal and vertical instantaneous eye velocity from the eye position following Equation 1 as in Engbert and Kliegl (2003):

$$v_i = \frac{F_s}{6}(x_{i+2} + x_{i+1} - x_{i-1} - x_{i-2}), \quad (1)$$

where x_i is the eye position (horizontal or vertical) at time i, v_i is the instantaneous eye velocity (horizontal or vertical) at time i, and F_S is the sampling rate. This operation is equivalent to smoothing the eye position with a triangular window (normalized Barlett window) of six samples and then differentiating to obtain the velocity. To maintain proper alignment between the velocity and position signals, the output must correspond with the center of the window. The length of the window may be adapted depending on the data-collection sampling rate to maintain a constant bandwidth—for example, by using a six-sample window (0, 1/6, 2/6, 2/6, 1/6, 0) for data recorded at 500 Hz and a 12-sample window (0, 1/12, 2/12, 3/12, 4/12, 5/12, 5/12, 4/12, 3/12, 2/6, 1/6, 0) for data recorded at 1000 Hz. The bandwidth of this triangular smoothing filter is approximately 100 Hz, which matches the bandwidth of fixational saccades (Findlay, 1971).

Then, we calculate the velocity magnitude from the horizontal and vertical components:

$$v = \sqrt{v_x^2 + v_y^2}. \quad (2)$$

If binocular recordings are available, we average the velocity magnitudes of the left and the right eye. We calculate the acceleration magnitude from the horizontal and vertical velocity components in the same way we calculate the velocity magnitude from horizontal and vertical eye-position components (Equations 1 and 2). From here on, we refer to velocity and acceleration magnitudes as simply velocity and acceleration.

To select the microsaccade candidates, we find the velocity local maxima that are separated by at least 30 ms (to avoid dynamic overshoots; Abadi, Scallan, & Clement, 2000; Otero-Millan et al., 2011; Otero-Millan et al., 2008) and then select the highest values to obtain an average rate of five candidates per second throughout each trial (Figure 1B). We define the beginning and the end of the peak (i.e., microsaccade candidate) as the last sample before the peak, and the first sample after the peak, below a 3°/s velocity threshold. This threshold corresponds to the slowest microsaccades reported in studies using measuring systems with a high signal-to-noise-ratio (Zuber et al., 1965).

Clustering

Clustering refers to the problem of classifying a set of observations into different groups, so that elements within a group share more similarities among themselves than with the elements in other groups. Clustering is an unsupervised classification problem, because the true groups that observations belong to are unknown, thereby precluding supervised training. Unsupervised classification methods have been used in spike sorting, genetics, and certain eye-movement analyses (Jorde & Wooding, 2004; Lewicki, 1998; Vidal, Bulling, & Gellersen, 2012).

One standard algorithm for cluster analysis is *k*-means. This consists of an iterative algorithm that assigns observations to *k* groups or clusters to minimize the within-cluster variability, that is, the sum of distances from each observation to the center of its own cluster. Each observation is characterized by a vector of features, and the separation between observations is typically measured with euclidean distance. Here we used the implementation of the algorithm from the function kmeans of the Statistical Toolbox within the MATLAB framework (MathWorks, Inc., Natick, MA).

For each microsaccade candidate, we consider the following features: peak velocity, initial acceleration peak, and final acceleration peak (Figure 1C through E). The initial acceleration peak is the highest acceleration value prior to the velocity peak, and the final acceleration peak is the highest acceleration value after the velocity peak, within the limits of the microsaccade. Velocity and acceleration are commonly used parameters to detect microsaccades (Engbert & Kliegl, 2003; Hafed et al., 2009).

The distributions of all these parameters are very skewed towards small values, so to facilitate the clustering, we use their logarithms and normalize them by subtracting the mean and dividing by the standard deviation (operation known as z-score). Let x_i be the vector of features for the microsaccade candidate i,

$$x_i = \Big(zscore\Big(\log(vel_i)\Big), zscore\Big(\log(acc\_start_i)\Big), zscore\Big(\log(acc\_stop_i)\Big)\Big), \quad (3)$$

and X be the matrix formed by all the candidates:

$$X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}. \quad (4)$$

Because all the listed characteristics are highly correlated, we obtain the principal components before applying the clustering. If C is the covariance matrix of X, M the mean vector of X, V the matrix formed by the eigenvectors of C, and D the diagonal matrix with the inverse square root of the eigenvalues of C, then the new matrix X^* corresponds with the uncorrelated components of X:

$$X^* = (X - M) * (V * D). \quad (5)$$

Next, we use only the first p columns of X^* (which account for the most variance) and apply the *k*-means algorithm to this new set of observations (Figure 1F, G). We select $p = 1$, 2, or 3 depending on how many columns have an eigenvalue larger than 5% of the maximum eigenvalue (to ensure that the selected uncorrelated components of X explain most of the variance). To select the initial condition for the algorithm (i.e., the starting center of each cluster), we divide the data into groups of equal number of candidates by sorting them by peak velocity, and we calculate the mean features within each group. To select the value of K, the number of clusters, we test multiple values (2, 3, and 4) and select the one with the smallest average silhouette (Rousseeuw, 1987; see next subsection for details). The algorithm is able to detect more than two potential clusters (i.e., microsaccades and noise) because the sources of noise are diverse; thus noisy events may separate into multiple clusters.

Then we select the cluster with the largest average magnitude as the true microsaccade cluster.

Quality control

Internal or external methods may estimate the goodness of the separation between clusters. Internal methods employ only the information already used for the clustering, that is, the candidates' features. External methods use additional information such as true labels of the observations (i.e., the true group that they belong to). Because we do not have the true labels of our data, here we chose an internal method.

We defined a detection reliability index based on the mean silhouette (Rousseeuw, 1987). The silhouette of each observation is defined as

$$s_i = \frac{b_i - a_i}{\max(a_i, b_i)}, \quad (6)$$

where a_i is the average distance between observation x_i and all the elements in its own cluster. If there are only two clusters, b_i is the average distance between observation x_i and the elements of the other cluster. If there are more than two clusters, b_i is the shortest of all average distances between observation x_i and the elements of each cluster, excluding its own.

The silhouette value for each observation ranges from -1 to $+1$ and is a measurement of the similarity between that observation and other observations in its own cluster, compared to its similarity to observations in other clusters. The average silhouette across observations provides a measurement of the distance between clusters.

In our method, this metric serves to estimate the separation between microsaccades and noise, so that one may discard data with high levels of noise that, if used, would produce unreliable results.

Human-data collection

The eye-movement database in this study included recordings from 20 adult subjects (12 men, eight women) with normal or corrected-to normal vision. Each subject participated in one or two sessions of approximately 50 min each, amounting to a total of 24 recordings. When two sessions from the same subject were used, different eye-tracking systems or settings were used in each session. All subjects fixated a small target on a computer monitor at a distance of 57 cm for 30- or 45-s trials. The data were collected as part of previously reported studies where subjects maintained fixation on a centrally presented target, with or without simultaneously performing a perceptual task (Martinez-Conde et al., 2006; McCamy et al., 2012; Otero-Millan et al., 2012; Troncoso, Macknik, & Martinez-Conde, 2008; Troncoso, Macknik, Otero-Millan, & Martinez-Conde, 2008). Experiments were performed under the guidelines of the Barrow Neurological Institute's Institutional Review Board (protocol 04BN039), and written informed consent was obtained from each participant. Fifteen of the subjects were naïve to the purposes of the experiment and were paid $15 per session.

Eye position was recorded noninvasively in both eyes with a fast video-based eye-movement monitor (EyeLink II or EyeLink 1000; SR Research, Ottawa, Ontario, Canada) at 500 samples per second (instrument noise of 0.01° RMS, per manufacturer specifications). We identified blink periods as the portions of the data where the pupil information was missing. We added 200 ms before and after each period to further include the initial and final parts of the blink, where the pupil is partially occluded. We moreover removed portions of the data corresponding to very fast decreases and increases in pupil area (20 units per sample, approximately 0.5 mm^2) plus the 200 ms before and after. Such periods are probably due to partial blinks, where the pupil is never fully occluded (thus failing to be identified as a blink by the eye-tracker software; Troncoso, Macknik, Otero-Millan, et al., 2008). We note that this method to remove partial blinks is specific to EyeLink systems, and that other eye-tracking systems may require different methods to remove partial blinks.

Other microsaccade-detection methods

To validate the performance of the present method, we also detected microsaccades with a widely used and accepted objective and automatic method of microsaccade detection (Engbert & Kliegl, 2003). The E&K method requires the setting of a sensitivity factor λ, which, multiplied by an estimation of the level of noise in the data, determines the final value of the velocity threshold. Here we used $\lambda = 6$ when performing analyses with a single λ value. When comparing multiple values of λ we used $\lambda = (2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 10, 12, 15, 20)$. To reduce the amount of potential noise (Engbert, 2006), we analyzed only binocular microsaccades (that is, microsaccades with a minimum overlap of one data sample in both eyes; Laubrock, Engbert, & Kliegl, 2005; Engbert, 2006; Rolfs, Laubrock, & Kliegl, 2006; Troncoso et al., 2008a). We also imposed a minimum intersaccadic interval of 20 ms so that dynamic overshoots observed right after microsaccades were not categorized as extra microsaccades (Møller, Laursen, Tygesen, & Sjølie, 2002; Otero-Millan et al., 2008; Troncoso, Macknik, & Martinez-Conde, 2008b).

We also tested an additional recent method that selects the value of λ automatically for the E&K algorithm (Engbert & Mergenthaler, 2006; Mergenthaler & Engbert, 2010). To do this, we calculated the instantaneous eye velocity of every data sample within each trial as

$$v_i = F_s(x_{i-1} - x_i), F_s = sampling\,frequency. \quad (7)$$

Then we randomly shuffled the velocity samples in time, to remove all correlations between contiguous data samples and between left and right eye, thus obtaining the surrogate velocities. To construct surrogate eye-position data from the surrogate velocities, we computed the integral of the surrogate velocity. We note that the distribution of the velocity samples in the surrogate data is exactly the same as in the original data. Finally, we applied the same detection algorithm to both surrogate and original data, varying the values of λ. For most λ values, there were more microsaccades detected in the original data than in the surrogate data. We selected the λ that resulted in the largest difference between the number of microsaccades detected in the original and surrogate data.

Manual labeling

To measure the performance of the clustering method, we created a new data set with the microsaccades detected by an expert operator (JO-M). The expert—albeit involved in the algorithm's development—had no knowledge of the exact details of the final method at the time of the labeling, which occurred during the early stages of the research. The expert inspected a visual graphic interface that presented 10 s of eye-position traces (horizontal and vertical) at a time, therefore precluding the direct use of the microsaccade peak-velocity and acceleration features that the method came to rely on for detection. The vertical axis and the size of the window were kept constant. The expert pointed to the location of each microsaccade with a mouse. The beginning and end of each microsaccade were then detected automatically around the point selected. First we found the nearest velocity peak to the point selected by the expert. Then we defined the beginning and the end of the peak as the last sample before the peak, and the first sample after the peak, below a 3°/s velocity threshold.

Simulated-data generation

We simulated eye-movement recordings that included microsaccades and noise mimicking those of each actual recording. First, we extracted a template of the average shape of a microsaccade by averaging the velocity waveform of 100 different microsaccades normalized to a peak velocity of 1°/s. Next, we randomly assigned onset times following an ex-Gaussian distribution (mean = 150 ms, σ = 50 ms, τ = 300 ms) and random peak velocities following a log-normal distribution (mean = 40°/s, σ = 13°/s) to create a sequence of microsaccades. In some of the simulations, microsaccade magnitudes (and therefore peak velocities) were constant for all microsaccades in the sequence.

Then we added noise to this signal. To generate the noise, we used an autoregressive process of order 10. To estimate the coefficients of the process, we modeled the data from microsaccade-free portions of each recording using the Yule–Walker method (function aryule, MathWorks, Inc). To produce a more realistic (i.e., not constant) level of noise in the simulated recordings, we moreover applied a multiplicative low-frequency component of white noise filtered at 1 Hz (Figure 2).

ROC curves

We used a receiver-operating-characteristic (ROC) analysis (Green & Swets, 1966) to evaluate the performance of the different microsaccade-detection methods. This analysis makes no assumptions about the underlying distributions. To obtain the ROC curve, we plotted the probability of true positives as a function of the probability of false positives for all possible criterion levels. The area under the ROC curve provides a measure of the discriminability of two signals and is directly related to the overlap of the two distributions of the property that is being compared (Green & Swets, 1966).

For the E&K method, the criterion level corresponds to the parameter λ. The clustering method does not have a criterion level, because it yields the optimal operating point by design. To obtain different points in the ROC curve for the clustering method, we parametrically modified the distance between each point and the center of the microsaccade cluster, thereby reducing or increasing the number of candidates detected as microsaccades. In this scenario, the total number of true negatives is arbitrary, because any point in time may be counted as a nonmicrosaccade event. Thus, we held the total number of true negatives constant and equal to five per second.

Results

We developed a novel method to detect the small involuntary saccades produced during attempted gaze

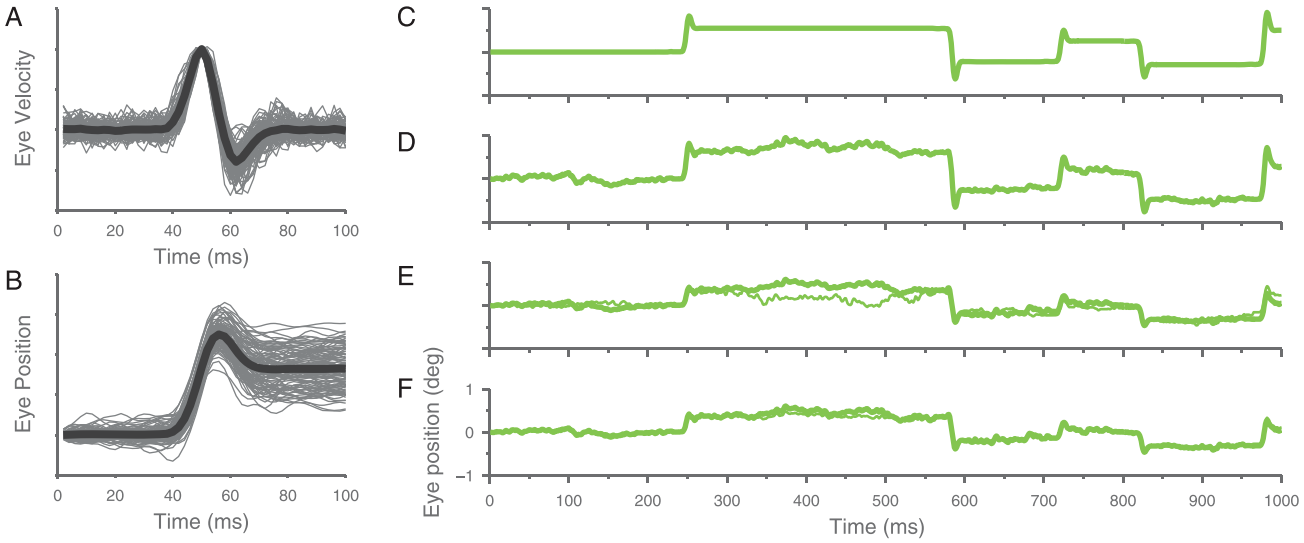

Figure 2. Simulated data. (A, B) Template extraction. Individual microsaccades used to extract the template (gray) and template (black). (C) Signal including only microsaccades, no noise. (D) Data with autoregressive noise. (E) Binocular data with independent autoregressive noise and additional multiplicative low-frequency noise components. (F) Binocular data as in (E), but with correlated (70%) noise components.

fixation: microsaccades. Unlike previous saccade-detection methods (Bahill et al., 1981; Engbert & Kliegl, 2003; Martinez-Conde et al., 2000), the present method does not require the setting of arbitrary thresholds. Rather, it characterizes automatically both microsaccadic and nonmicrosaccadic events (including noise) and uses clustering techniques to find the ideal boundary between them. Our new clustering method also provides an index of the reliability of the result, in relationship to the signal-to-noise ratio in the raw data.

To validate our method, we compared its performance to the most popular contemporary method of microsaccade detection (Engbert & Kliegl, 2003), applying a three-pronged approach: First, we analyzed qualitatively the distributions of microsaccade properties (i.e., magnitudes, peak velocities) obtained with the E&K method and the clustering method on real eye-movement data. Second, we compared the performance of both detection methods against an expert's manual detection (JO-M), using the same eye-movement data set. Finally, we evaluated the performances of both methods on an artificially generated data set.

Qualitative validation

Many saccade-detection methods classify the fast portions of eye-movement recordings as saccades by applying a set velocity threshold to the data. The E&K method offers the critical advantage of a threshold that adapts to the level of noise in the data, but it requires the user to indicate a factor (λ, typically 6 but sometimes 4 or 5; see Engbert & Mergenthaler, 2006; McCamy et al., 2012; Mergenthaler & Engbert, 2010) that is multiplied by the standard deviation of the eye-movement velocity to obtain the final velocity threshold (see Methods for details).

To determine the effects of different λ values in the E&K method, and to compare them with the results from the clustering method, we used a qualitative approach based on the shape of the resultant distributions of saccadic parameters. Nyström and Holmqvist (Nyström & Holmqvist, 2010) used this approach previously to evaluate the quality of automatic saccade detection.

Very low (permissive) λ values produced many false positives and bimodal distributions of microsaccade magnitudes and peak velocities, where the first mode corresponds to false positives (noise or slow eye movements) and the second mode to true microsaccades (Figure 3). In an ideal method, all detected microsaccades would be included in the second mode and none in the first mode.

Figure 3A and B shows the average distributions of microsaccade amplitudes and peak velocities for all recordings combined. Low λ values (<4) produce many microsaccades in the first mode (i.e., false positives). High λ values (>7) decrease the size of the second mode, thereby reducing the number of true microsaccades detected (i.e., more false negatives). The standard value of $\lambda = 6$ works well, but some false positives remain in the first mode (especially clear in Figure 3B). The clustering method results in a single (i.e., second) mode, suggesting an improved performance when compared to set λ values (Figure 3A, B).

Figure 3C through E illustrates how individual recordings may produce varied results. In Figure 3C,

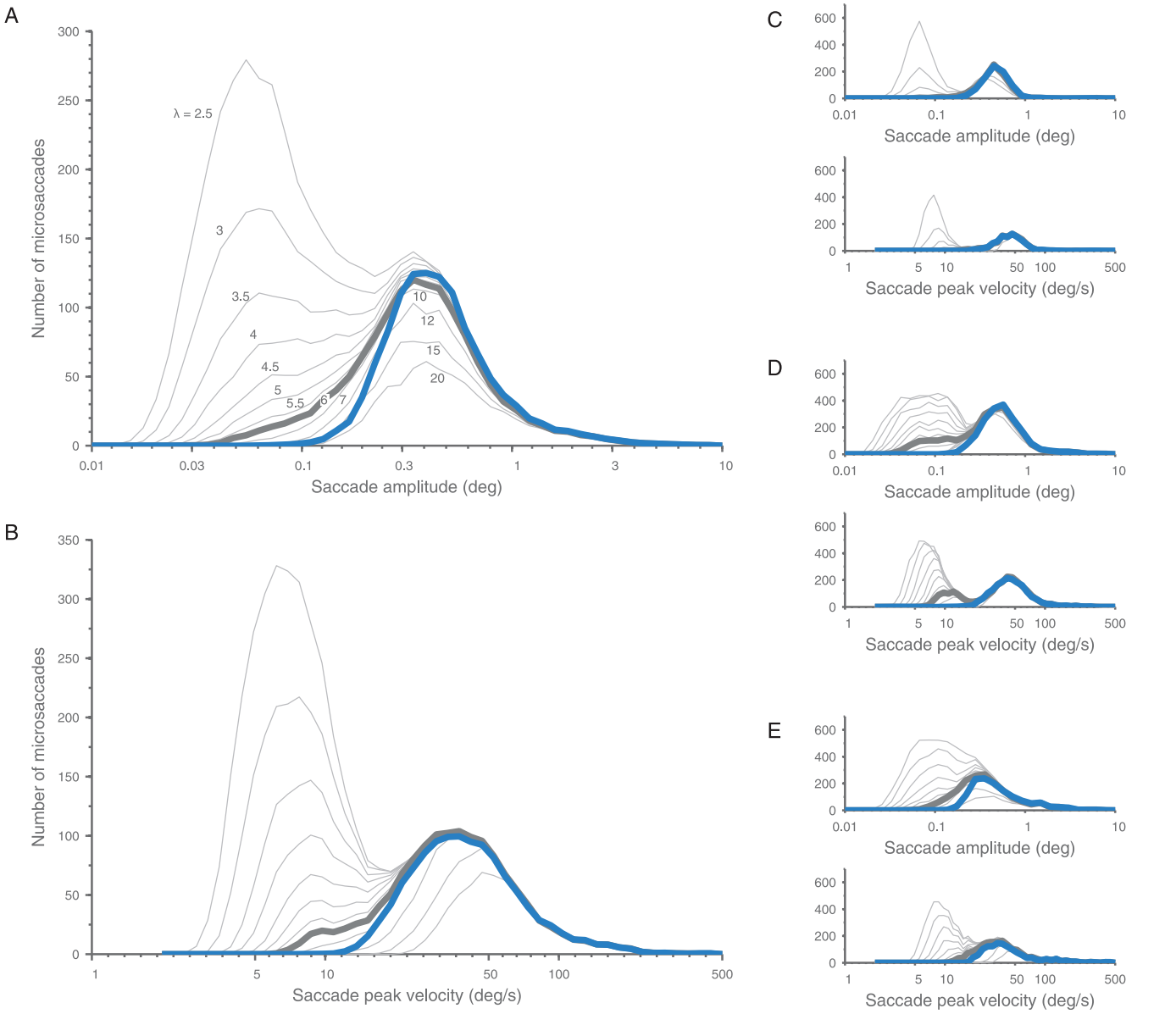

Figure 3. Bimodal distributions of microsaccade properties. (A) Average microsaccade amplitude distribution across all recordings ($n = 24$). (B) Average microsaccade peak-velocity distribution across all recordings. (C–E) Microsaccade amplitude and peak-velocity distributions for three individual recordings, with different qualitative results. (A–E) Thin gray lines correspond to different E&K λ values (from 2.5 to 20, left to right). The thick gray line corresponds to $\lambda = 6$. The blue line indicates the distribution obtained with the clustering method.

where the bimodality is very obvious, both the E&K and the clustering methods work very well. In Figure 3D the bimodality is also clear, but only the clustering method performs well. In Figure 3E, the bimodality is not clear and it is hard to know which method, if any, produces an optimal performance.

Quantitative validation

To quantify the performance of different microsaccade-detection methods, one would ideally use a data set where the occurrences of true microsaccades are labeled correctly. Such a data set is not possible to obtain, unfortunately, given that microsaccades are involuntary and that a flawless method of recording and detection is not yet available. Thus, we opted for expert manual labeling of the microsaccades in our data set (see Methods) and used these labels to calculate the number of errors (false positives and false negatives) from each detection method. Other microsaccade studies have used expert validation to correct the results of automated detection (Hafed et al., 2009).

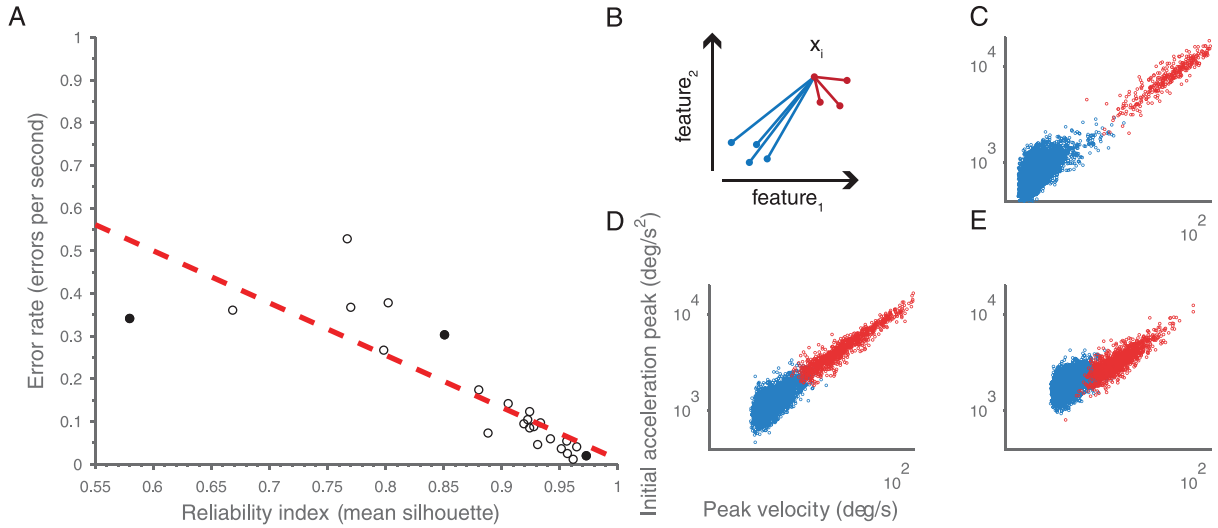

Figure 4. Reliability index. (A) Correlation between detection reliability index (mean silhouette) and error rate. Each circle represents one recording ($n = 24$). (B) Schematic of the calculation of the silhouette in a sample. (C–E) Examples of high, medium, and low detection reliability (i.e., high, medium, and low values of the mean silhouette). Clusters are maximally separated in (C), indicating high detection reliability, and show substantial overlap in (E), indicating low detection reliability. Examples correspond to recordings highlighted in black in (A).

To measure detection performance, we defined the error rate as the number of errors (false positives and false negatives) per unit of time. A false positive was a microsaccade detected automatically that did not overlap in time with any microsaccade detected manually. A false negative was a microsaccade detected manually that did not overlap in time with any microsaccade detected automatically.

The clustering method not only identifies microsaccades automatically, but it also provides a detection reliability index (mean silhouette), related to the signal-to-noise ratio in the recordings (see Methods). Figure 4 shows that the error rate is inversely correlated to the mean silhouette of each recording ($R = -0.9$, $p = 0.000000003$), thus suggesting that our metric is a good predictor of the reliability of the detection. A recording with a high mean silhouette is more likely to result in few detection errors.

Next, we calculated the error rate across recordings for the E&K ($\lambda = 6$) and clustering methods. Figure 5A shows the results for all recordings, sorted by mean silhouette (higher silhouette first). The clustering method shows an improved performance in most cases (19 out of 24 recordings). The median error rate in the E&K method is 0.25 errors per second; the median error rate in the clustering method is 0.1 errors per second (we used median error rates to reduce the influence of outliers; Figure 5B). The cluster method's improvement in overall detection error rate results from a lesser number of false positives than in the E&K method (Figure 5C). False negatives were slightly more prevalent with the cluster method than with the E&K method, however (Figure 5D).

The performance of the E&K method with $\lambda = 6$ varied across recordings (see also Figure 3), suggesting the possibility that other λ values might have been preferable in some instances. To address this issue, Engbert and Mergenthaler (Engbert & Mergenthaler, 2006; Mergenthaler & Engbert, 2010) developed an additional method to select the best λ value for any given recording. Their method, based on surrogate data, selects the value of λ that maximizes an estimation of difference between the numbers of true and false positives (see Methods). Figure 6A shows the error rates obtained when selecting the best λ for each recording: Some individual error rates remain quite large, but the median error rate improves slightly. We also calculated the error rates when selecting for each recording the λ value that produced the smallest error rate a posteriori. This resulted in a much lower median error rate, but still slightly higher than with the clustering method. Thus, the clustering method selected the optimum boundary between true microsaccades and noisy or other nonmicrosaccadic events. (See Figure 6B for ROC curves for the different methods.)

Some experimental conditions preclude binocular recordings, due to the task, the setup, or the limitations of the eye tracker. Thus, it is important to evaluate the performance of microsaccade-detection methods in monocular recordings as well. Here we applied the same data set as in the previous analyses, but using the recordings from one eye only. Figure 7 shows the performance of the E&K ($\lambda = 6$) and clustering methods for monocular recordings. The clustering method was more robust and outper-

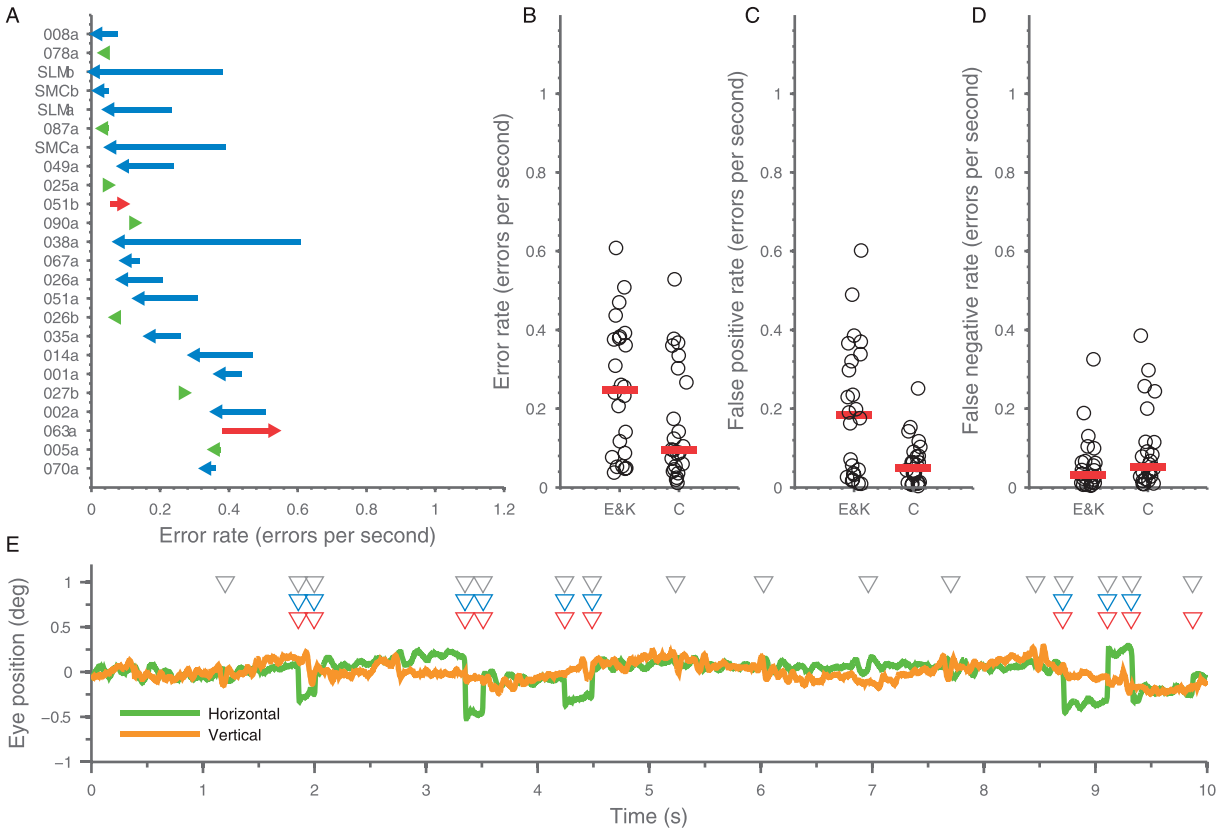

Figure 5. Error rates in E&K and clustering methods. (A) Arrows indicate the difference in error rate between E&K and clustering method in each recording. Blue: Clustering method results in a smaller error rate. Green: Clustering method results in an equal error rate. Red: Clustering method results in a larger error rate. (B) Median error rates (red lines) for each method. C) Median false-positive rates (red lines) for each method. (D) Median false-negative rates (red lines) for each method (E&K, $\lambda = 6$; $C =$ clustering method). (E) Example of microsaccade detection with the E&K and clustering methods. Gray triangles indicate microsaccades detected by E&K with $\lambda = 6$, blue triangles indicate microsaccades detected with the clustering method, and red triangles indicate microsaccades labeled manually.

formed the E&K method in most recordings (17 of 24), reducing the median error rate from 1.1 to 0.25 errors per second.

Validation using simulated data

Next, we assessed the performance of each method using simulated data. We generated artificial eye-position traces, including microsaccades (replicating a template extracted for each recording; see Methods) and different forms of noise (see Methods for details).

First, we generated monocular data and added simple autoregressive noise. Both methods behaved well in this condition, for moderate levels of noise (Figure 8A). Next, we added a multiplicative term of low-frequency noise, which made the overall noise level variable through time, thus resulting in a more realistic scenario (since in real situations the level of noise may vary with subject movement or changes in pupil size). This condition resulted in performance degradation for the E&K method (Figure 8B). Then we generated binocular data and added independent noise signals to each eye (in addition to the multiplicative term; Figure 8C). This improved dramatically the performance of the E&K method, which used the binocular information in a very efficient way. Finally, we added a 70% correlation between the noise signals in each eye, rather than using completely independent noise in each eye. This is also a more realistic scenario, because intersaccadic drift or noise and artifacts are usually correlated between the two eyes (Figure 8E). In this situation, the clustering method behaved better than the E&K method. Comparison of false-positive and false-negative rates for varying microsaccade velocities also showed an advantage for the clustering method over the E&K method (Figure 8F).

These combined results suggest that the E&K method works very well in circumstances in which there is simple noise uncorrelated between the two eyes, but

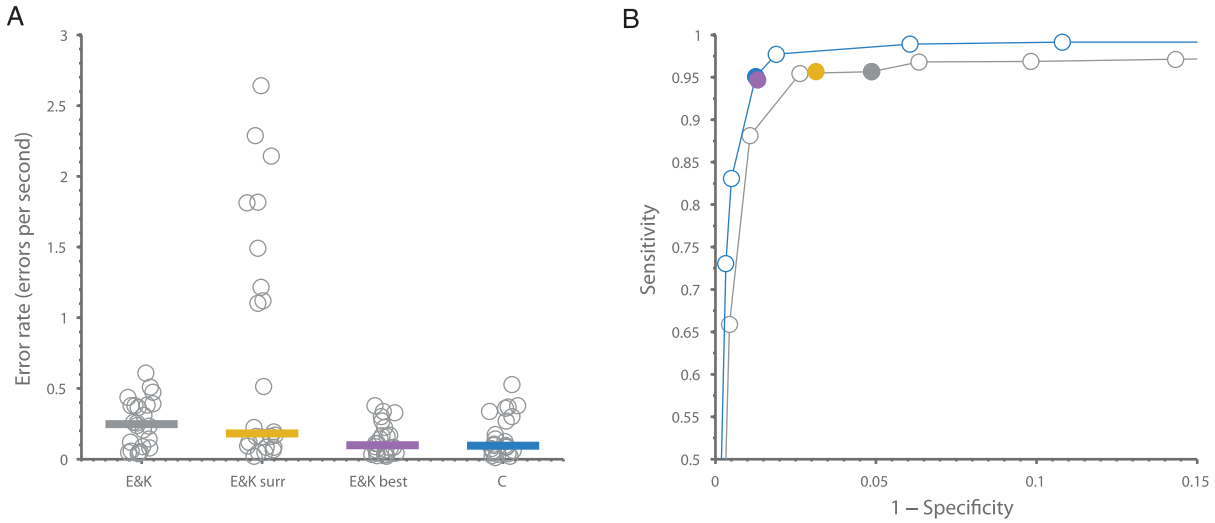

Figure 6. Error rates in E&K and clustering methods. (A) Median error rates (thick lines) for each method (E&K, $\lambda = 6$; E&K surr, λ indicated by the surrogate method for each recording; E&K best, λ chosen a posteriori to minimize error rates; C = clustering method). (B) ROC curves for the E&K (gray) and clustering (blue) methods. The solid blue circle indicates the actual result of the clustering method, without manipulating the distances between each point and the center of the microsaccade cluster; see Methods for details; the solid gray circle indicates E&K, $\lambda = 6$; other colors as in (A).

that the clustering method performs better when the noise level varies throughout the recording and is correlated between the two eyes (a more realistic scenario).

We note that one should not assume that microsaccades are the only signals that correlate between the two eyes (Figure 8E). Reliance on this assumption may partly explain the suboptimal performance of the surrogate method (Figure 6).

Detection reliability as a function of microsaccade magnitude

The performance of any microsaccade-detection method will necessarily degrade as microsaccade magnitude decreases. Thus, it is important to determine detection reliability for microsaccades of different magnitudes and to estimate the proportion of micro-

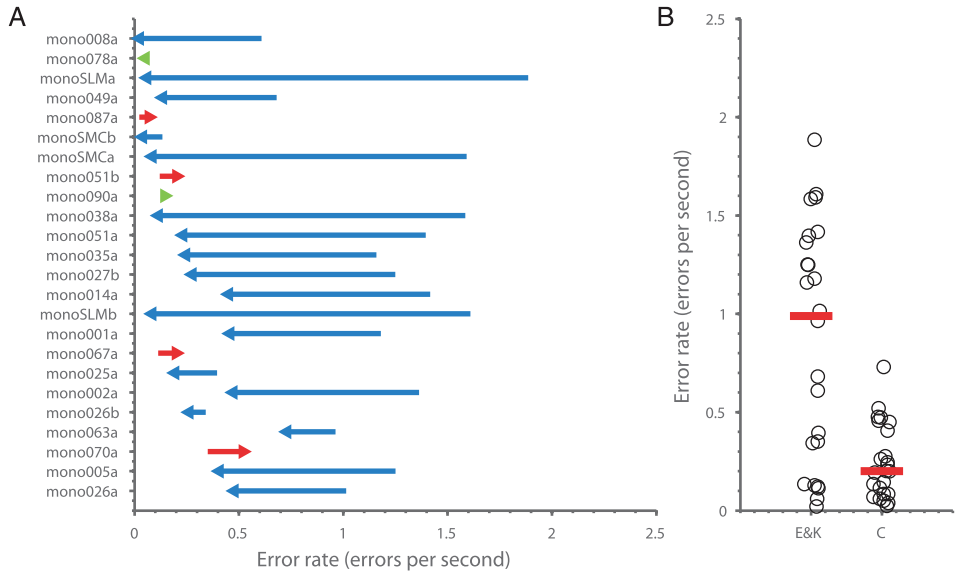

Figure 7. Error rates in the E&K and clustering methods, using monocular data. (A) Arrows indicate the difference in error rate between E&K and clustering method in each recording. Blue: Clustering method results in a smaller error rate. Green: Clustering method results in an equal error rate. Red: Clustering method results in a larger error rate. (B) Median error rates (red lines) for each method (E&K, $\lambda = 6$; C = clustering method).

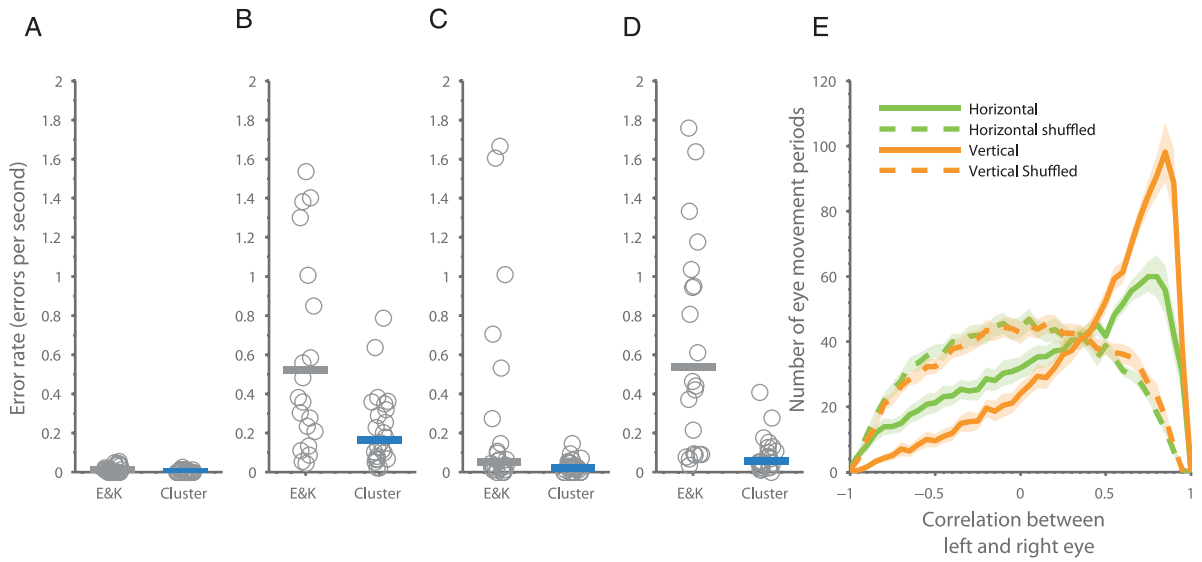

Figure 8. Performance of E&K and clustering methods with simulated data (E&K, $\lambda = 6$). (A) Monocular data with autoregressive noise. (B) Monocular data with autoregressive noise and additional multiplicative low-frequency noise. (C) Binocular data with independent noise components generated as in (B). (D) Binocular data as in (C), but with correlated (70%) noise components. (A–D) Horizontal lines represent median error rates across recordings. (E) Distributions of correlation coefficients between the left and right eye positions during microsaccade-free periods. Solid lines: The correlation coefficients distributions are skewed towards 1 (especially in the case of the vertical component), indicating that eye-position data are correlated between the two eyes, even in the absence of microsaccades. Dashed lines: Distributions expected by chance. We calculated the chance distributions by correlating the left-eye positions during microsaccade-free periods with the right-eye positions of different (randomly selected) microsaccade-free periods.

saccades that may not be detected reliably due to their small magnitude.

We found that both the E&K method and the clustering method detected reliably microsaccades with magnitudes over 0.2°. Using manual labeling as the gold standard, both methods detected microsaccades of approximately 0.2° more than 75% of the time, and more than 75% of the detected microsaccades were true microsaccades (Figure 9A, B). Consistent with the results in Figure 5, the E&K method was more sensitive, at the cost of more false positives (that is, it detected more small microsaccades but also more false alarms), whereas the clustering method was more precise, at the cost of more false negatives (that is, more of the detected microsaccades were true microsaccades, but it missed more small microsaccades; Figure 9A, B).

Most of the microsaccades in the recordings (>90%) were larger than 0.2°, with 0.2° already being in the tail end of the distribution (Figure 9C), which suggests that even if the manual detection missed some of the smaller microsaccades, their number could not have been very large.

To estimate the number of microsaccades missed by the manual detection, we generated new simulations as described in the Methods section, but now each simulation contained microsaccades of a fixed magnitude. Then we performed manual detection of microsaccades on this simulated data set. Detection performance dropped below 75% for microsaccades less than 0.1° (Figure 9D).

The simulations with fixed microsaccade magnitudes also allowed us to compare the performances of the E&K and clustering methods as a function of microsaccade magnitude. We generated families of ROC curves, one for each microsaccade magnitude (Figure 9E, F). The area under the ROC curve gives an indication of each method's performance. The clustering method was equivalent to, or outperformed, the E&K method for all microsaccade magnitudes (Figure 9G).

Discussion

We present a new method to detect microsaccades based on clustering techniques. Our analyses showed that this method performs better than current standard methods of microsaccade detection (Engbert & Kliegl, 2003), with a 62% improvement in the case of binocular data and a 77% improvement with monocular data (Figures 5, 7). In this section, we discuss the additional advantages that the clustering method incorporates. We moreover review the possible ways to validate eye-movement classification methods, including the clustering method. Finally, we propose that the detection

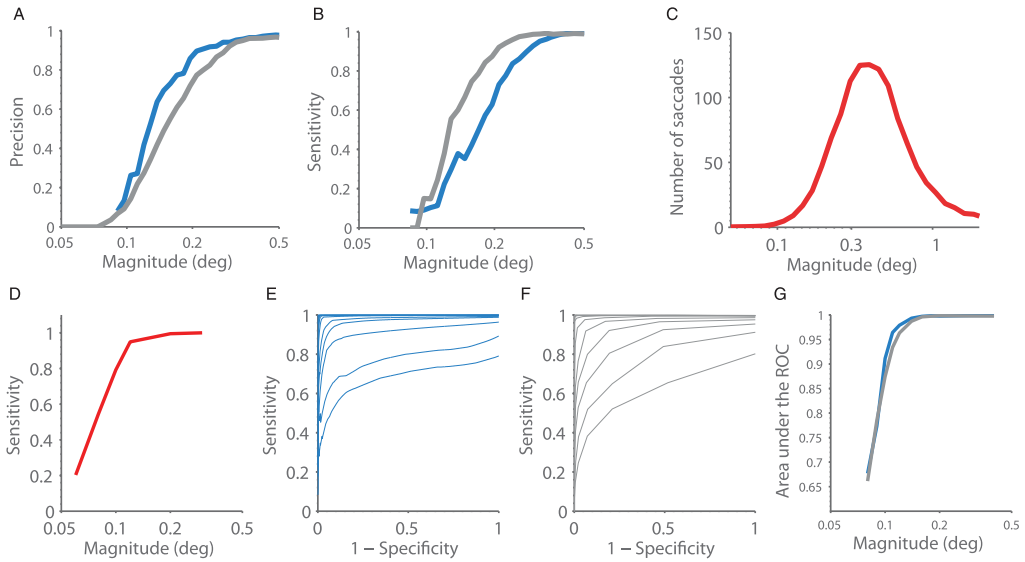

Figure 9. Performance of E&K (gray) and clustering (blue) methods for different microsaccade magnitudes. (A) Precision of both methods as a function of microsaccade magnitude (probability that a detected microsaccade is a true microsaccade, as identified by manual detection). Dashed lines indicate the 0.75 probability threshold. (B) Sensitivity of both methods as a function of microsaccade magnitude (probability of detection of a true microsaccade, as identified by manual detection). Dashed lines indicate the 0.75 probability threshold. (C) Average microsaccade-magnitude distribution obtained with manual detection. (D) Sensitivity of manual detection on simulated data, as a function of microsaccade magnitude (probability of manual detection of a simulated microsaccade). (E, F) Family of ROC curves obtained with the clustering (E) and E&K (F) methods in simulations of fixed-magnitude microsaccades (from 0.05° to 0.4°). (G) Area under the ROC curve as a function of microsaccade magnitude for the E&K and clustering method (ROC curves from E and F).

reliability index provided by the clustering method can have significant value as a means to determine the comparative precision of eye-tracking devices.

Advantages of the clustering method

Most current automatic saccade- or microsaccade-detection methods require the setting of a sensitivity parameter, such as a velocity threshold. Several of these methods—with the E&K algorithm as a particularly valuable example—have been used to excellent effect in the last decade. Yet one cannot be certain that the setting is optimal for any given data set, due to the lack of objective validation. The clustering method we have developed does not require such a parameter because it finds, automatically and objectively, a boundary between true microsaccades and nonmicrosaccadic events, including intersaccadic drift, noise, and artifacts such as head movements and changes in pupil size. We note that, even though signals related to head movements and changes in pupil size are generally slower than those from saccades, the smallest microsaccades (i.e., <0.2°) are also quite slow (i.e., <20°/s). Further, depending on the specific hardware and eye-tracking algorithm used, head movements and pupil-size changes may produce apparent fast changes (i.e., the corneal reflection coming in and out of the pupil can introduce a quick change in the position of the center of mass of the pupil).

The clustering method takes advantage of the fact that microsaccade rates usually range from 0.5 to four per second to select a set of microsaccade candidates (at a rate of five per second) that contain all true microsaccades and a comparable number of non-microsaccadic events. This assumption is sound, given that most subjects (even those with neurological pathologies) produce average saccade rates that are no higher than four per second (Abadi & Gowen, 2004; Otero-Millan, Schneider, et al., 2013; Otero-Millan et al., 2011) and no lower than 0.5 per second (Martinez-Conde et al., 2004; Martinez-Conde et al., 2009). Indeed, most healthy subjects exhibit saccade rates between one and two per second, although some subjects can suppress microsaccades down to only 0.5 per second (Rolfs, 2009). Even in free-viewing tasks in which subjects are allowed to make all types of saccades, saccadic rates usually do not exceed 4 per second (Otero-Millan, Macknik, et al., 2013; Otero-Millan et al., 2008). Thus, the clustering method can be applied to detect saccades of any size in any type of perceptual or oculomotor task.

Further, the clustering method does not rely on the binocularity of microsaccades. When binocular data is

available, it averages the eye position across the eyes to reduce noise, but this feature is not critical for the method's performance (Figure 7). This results in a double advantage: First, the method works well even if only monocular data are available (Figure 7). Second, when binocular data are available, the method does not discard potential monocular microsaccades, or microsaccades with a very small component in one of the eyes (Van Horn & Cullen, 2012), as most previous methods do.

An additional unique feature of the clustering method is that it provides an index of the signal-to-noise ratio in the data, which may be used to quantify the performance of various eye-tracking systems or to warn the experimenter about high levels of noise and thus the potential need to discard data.

General principles to validate methods of microsaccade detection

Performance validation is a recurring problem for microsaccade- and saccade-detection methods. Several options have been used in the past:

Comparison with labeled data. This is the most straightforward approach to evaluating detection methods, as it compares the detected events against the true values, that is, the true microsaccade occurrences. In the case of microsaccade detection, it requires an expert to label every single microsaccade in the data, a strategy that poses potential problems related to the level of expertise of the labeler and his or her subjectivity, and it is time consuming when long recordings must be labeled. Despite these difficulties, and given the intrinsic difficulties of microsaccade detection, expert manual labeling is sometimes used to reduce the numbers of false positives and false negatives (Hafed et al., 2009).

Comparison with stimuli. In the situation in which subjects are instructed to follow a jumping target, the saccade-detection results can be checked against the target jumps (Komogortsev et al., 2010; Salvucci & Goldberg, 2000). This method poses potential problems, however, including variability in the subjects' reaction times and occasional corrective saccades. More importantly, it poses two important challenges for microsaccade detection: (a) Visual fixation of a stationary target includes no target jumps by definition, thus precluding their use to validate the results of automatic microsaccade detection. (b) In nonfixation conditions, such as during the guided viewing of small target jumps, one could use the target displacements to validate the detection of microsaccades following the target's jumps but would miss any additional microsaccades produced in between jumps.

Distribution metrics or qualitative exploration. Different descriptive statistics or graphical representations can be used to identify detection problems. For instance, one may identify values outside physiologically plausible ranges, such as saccade durations that are too short or too long or velocities that are too high. Visual inspection can moreover point to potential problems such as bimodalities in the distributions of saccade parameters (Nyström & Holmqvist, 2010).

Simulated data. Simulated data are advantageous in that they can be labeled objectively, and one can moreover control the level and properties of the noise. The main drawback of this option is its limited realism.

Each of these approaches has advantages and limitations, and so we opted to use all of them (except for the comparison with stimuli, which is not applicable to microsaccades) to validate our method.

A new index of eye-tracking precision?

Two main parameters are important in any eye-tracking system: precision and accuracy. Precision defines the magnitude of the minimum change in eye position that the system can detect. Accuracy refers to the system's ability to indicate the correct eye position without a bias. In the case of saccade detection, precision is the critical parameter. Low accuracy will affect the characterization but not the detection of saccades.

Many factors may influence the quality of a recording (Holmqvist, Nyström, & Mulvey, 2012), such as subject characteristics, operator expertise, amount of head or body movement, position of the camera or source of illumination, and critically, the specific eye-tracking system in use.

Because one cannot distinguish changes in eye position that correspond to true eye movements from noise or artifacts, the standard assessment of an eye-tracking system's precision relies on the use of (stationary) artificial eyes, where all the eye movements recorded are necessarily due to noise or artifacts.

Our novel clustering method can solve this important problem by providing an objective index of an eye-tracking system's precision. This reliability index, based on mean silhouette, is correlated with the error rate of different recordings (Figure 4). Reliability indices obtained with different eye trackers for the same subject, in the same experimental and task conditions, will reflect the precision of each eye-tracking system. The eye-tracking system with the highest reliability index will be the most precise and will produce the fewest detection errors.

Keywords: fixation, saccades, microsaccade detection, eye-movement classification

Acknowledgments

This study was supported by the Barrow Neurological Foundation (to SLM and SMC) and the National Science Foundation (Award 0726113 to SLM and Awards 0852636 and 1153786 to SMC). We thank Behrooz Kousari for technical assistance.

Commercial relationships: none.
Corresponding author: Susana Martinez-Conde.
Email: smart@neuralcorrelate.com.
Address: Department of Neurobiology, Barrow Neurological Institute, Phoenix, AZ.

References

Abadi, R. V., & Gowen, E. (2004). Characteristics of saccadic intrusions. *Vision Research, 44*(23), 2675–2690, doi:10.1016/j.visres.2004.05.009.

Abadi, R. V., Scallan, C. J., & Clement, R. A. (2000). The characteristics of dynamic overshoots in square-wave jerks, and in congenital and manifest latent nystagmus. *Vision Research, 40*(20), 2813–2829.

Bahill, A., Brockenbrough, A., & Troost, B. (1981). Variability and development of a normative data base for saccadic eye movements. *Investigative Ophthalmology and Visual Science, 21*(1), 116–125, http://www.iovs.org/content/21/1/116. [PubMed] [Article]

Bettenbühl, M., Paladini, C., Mergenthaler, K., Kliegl, R., Engbert, R., & Holschneider, M. (2009). Microsaccade characterization using the continuous wavelet transform and principal component analysis. *Journal of Eye Movement Research*, *3*(5), 1–14.

Chen, A. L., Riley, D. E., King, S. A., Joshi, A. C., Serra, A., Liao, K., … Leigh, R. J. (2010). The disturbance of gaze in progressive supranuclear palsy: Implications for pathogenesis. *Frontiers in Neurology, 147,* 1–19, doi:10.3389/fneur.2010.00147.

Costela, F. M., McCamy, M. B., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2013). Microsaccades restore the visibility of minute foveal targets. *PeerJ, 1,* e119, doi:10.7717/peerj.119.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research, 154,* 177–192.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035–1045, doi:10.1016/S0042-6989(03)00084-1.

Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences, USA, 103*(18), 7192–7197, doi:10.1073/pnas.0509557103.

Findlay, J. M. (1971). Frequency analysis of human involuntary eye movement. *Biological Cybernetics, 8*(6), 207–214, doi:10.1007/BF00288749.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Huntington, NY: Wiley.

Hafed, Z. M., & Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research, 42*(22), 2533–2545, doi:10.1016/S0042-6989(02)00263-8.

Hafed, Z. M., Goffart, L., & Krauzlis, R. J. (2009, February 13). A neural mechanism for microsaccade generation in the primate superior colliculus. *Science, 323*(5916), 940–943, doi:10.1126/science.1166112.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. *Proceedings of the symposium on eye tracking research and applications, 12,* 45–52, doi:10.1145/2168556.2168563.

Jorde, L. B., & Wooding, S. P. (2004). Genetic variation, classification and "race." *Nature Genetics, 36,* S28–S33, doi:10.1038/ng1435.

Kapoula, Z., Yang, Q., Otero-Millan, J., Xiao, S., Macknik, S. L., Lang, A., … Martinez-Conde, S. (in press). Distinctive features of microsaccades in Alzheimer's disease and in mild cognitive impairment. *Age*.

Komogortsev, O. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2010). Qualitative and quantitative scoring and evaluation of the eye movement classification algorithms. *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, 10,* 65–68. New York: ACM. doi:10.1145/1743666.1743682.

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research*, *45*(6), 721–730. doi:10.1016/j.visres.2004.09.029.

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements*. New York: Oxford University Press.

Lewicki, M. S. (1998). A review of methods for spike sorting: The detection and classification of neural action potentials. *Network: Computation in Neural Systems, 9*(4), 53–78.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. In S. Martinez-Conde, S. L. Macknik, L. M. Martinez, J.-M. Alonso, & P. U. Tse (Eds.), *Progress in Brain Research* (Vol. 154, Part A, pp. 151–176). Elsevier. Available at http://dx.doi.org/10.1016/S0079-6123(06)54008-7.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Nature Neuroscience*, *3*(3), 251–258.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2002). The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proceedings of the National Academy of Sciences*, *99*(21), 13920–13925.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, *5*(3), 229–240.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron, 49*(2), 297–305, doi:10.1016/j.neuron.2005.11.033.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: A neurophysiological analysis. *Trends in Neurosciences, 32*(9), 463–475, doi:10.1016/j.tins.2009.05.006.

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience, 14*(2), 83–96, doi:10.1038/nrn3405.

McCamy, M. B., Macknik, S. L., & Martinez-Conde, S. (2014). Natural eye movements and vision. In P. Editor (Ed.), *The new visual neurosciences* (pp. 849–863). Cambridge, MA: MIT Press.

McCamy, M. B., Otero-Millan, J., Macknik, S. L., Yang, Y., Troncoso, X. G., Baer, S. M., ... Martinez-Conde, S. (2012). Microsaccadic efficacy and contribution to foveal and peripheral vision. *Journal of Neuroscience, 32*(27), 9194–9204, doi:10.1523/JNEUROSCI.0515-12.2012.

Mergenthaler, K., & Engbert, R. (2010). Microsaccades are different from saccades in scene perception. *Experimental Brain Research, 203*(4), 753–757, doi:10.1007/s00221-010-2272-9.

Møller, F., Laursen, M., Tygesen, J., & Sjølie, A. (2002). Binocular quantificaiton and characterization of microsaccades. *Graefe's Archive for Clinical and Experimental Ophthalmology, 240*(9), 765–770, doi:10.1007/s00417-002-0519-2.

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods, 42*(1), 188–204, doi:10.3758/BRM.42.1.188.

Otero-Millan, J., Macknik, S. L., Langston, R. E., & Martinez-Conde, S. (2013). An oculomotor continuum from exploration to fixation. *Proceedings of the National Academy of Sciences, USA, 110*(15), 6175–6180, doi:10.1073/pnas.1222715110.

Otero-Millan, J., Macknik, S. L., & Martinez-Conde, S. (2012). Microsaccades and blinks trigger illusory rotation in the "rotating snakes" illusion. *Journal of Neuroscience, 32*(17), 6043–6051, doi:10.1523/JNEUROSCI.5823-11.2012.

Otero-Millan, J., Schneider, R., Leigh, R. J., Macknik, S. L., & Martinez-Conde, S. (2013). Saccades during attempted fixation in Parkinsonian disorders and recessive ataxia: From microsaccades to square-wave jerks. *PLoS ONE, 8*(3), e58535, doi:10.1371/journal.pone.0058535.

Otero-Millan, J., Serra, A., Leigh, R. J., Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2011). Distinctive features of saccadic intrusions and microsaccades in progressive supranuclear palsy. *Journal of Neuroscience, 31*(12), 4379–4387, doi:10.1523/JNEUROSCI.2600-10.2011.

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*(14):21, 1–18, http://www.journalofvision.org/content/8/14/21, doi:10.1167/8.14.21. [PubMed] [Article]

Robinson, D. A. (1963). A method of measuring eye movements using a scleral coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics, 10,* 137–145.

Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research, 49*(20), 2415–2441, doi:10.1016/j.visres.2009.08.010.

Rolfs, M., Laubrock, J., & Kliegl, R. (2006). Shortening and prolongation of saccade latencies following microsaccades. *Experimental Brain Research, 169*(3), 369–376, doi:10.1007/s00221-005-0148-1.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics, 20,* 53–65, doi:10.1016/0377-0427(87)90125-7.

Salvucci, D., & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. In P. Editor (Ed.), *Proceedings of the 2000 symposium on eye tracking research & applications* (pp. 71–78).

Palm Beach Gardens, FL: ACM. doi:10.1145/355017.355028.

Serra, A., Liao, K., Martinez-Conde, S., Optican, L. M., & Leigh, R. J. (2008). Suppression of saccadic intrusions in hereditary ataxia by memantine. *Neurology, 70*(10), 810–812, doi:10.1212/01.wnl.0000286952.01476.eb.

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2008a). Microsaccades counteract perceptual filling-in. *Journal of Vision, 8*(14):15, 1–9, http://www.journalofvision.org/content/8/14/15, doi:10.1167/8.14.15. [PubMed] [Article]

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008b). Microsaccades drive illusory motion in the Enigma illusion. *Proceedings of the National Academy of Sciences, USA, 105*(41), 16033–16038, doi:10.1073/pnas.0709389105.

Van Horn, M. R., & Cullen, K. E. (2012). Coding of microsaccades in three-dimensional space by premotor saccadic neurons. *Journal of Neuroscience, 32*(6), 1974–1980, doi:10.1523/JNEUROSCI.5054-11.2012.

Vidal, M., Bulling, A., & Gellersen, H. (2012). Detection of smooth pursuits using eye movement shape features. In P. Editor (Ed.), *Proceedings of the symposium on eye tracking research and applications* (pp. 177–180). New York: ACM. doi:10.1145/2168556.2168586.

Zuber, B. L., Stark, L., & Cook, G. (1965, December 10). Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science, 150*(3702), 1459–1460, doi:10.1126/science.150.3702.1459.