

A nonlinear generalization of the Savitzky-Golay filter and the quantitative analysis of saccades

Weiwei Dai

Department of Electrical and Computer Engineering,
Tandon School of Engineering, New York University,
Brooklyn, NY, USA



Ivan Selesnick

Department of Electrical and Computer Engineering,
Tandon School of Engineering, New York University,
Brooklyn, NY, USA



John-Ross Rizzo

Department of Neurology, School of Medicine,
New York University, New York, NY, USA



Janet Rucker

Department of Neurology, School of Medicine,
New York University, New York, NY, USA



Todd Hudson

Department of Neurology, School of Medicine,
New York University, New York, NY, USA



The Savitzky-Golay (SG) filter is widely used to smooth and differentiate time series, especially biomedical data. However, time series that exhibit abrupt departures from their typical trends, such as sharp waves or steps, which are of physiological interest, tend to be oversmoothed by the SG filter. Hence, the SG filter tends to systematically underestimate physiological parameters in certain situations. This article proposes a generalization of the SG filter to more accurately track abrupt deviations in time series, leading to more accurate parameter estimates (e.g., peak velocity of saccadic eye movements). The proposed filtering methodology models a time series as the sum of two component time series: a low-frequency time series for which the conventional SG filter is well suited, and a second time series that exhibits instantaneous deviations (e.g., sharp waves, steps, or more generally, discontinuities in a higher order derivative). The generalized SG filter is then applied to the quantitative analysis of saccadic eye movements. It is demonstrated that (a) the conventional SG filter underestimates the peak velocity of saccades, especially those of small amplitude, and (b) the generalized SG filter estimates peak saccadic velocity more accurately than the conventional filter.

Introduction

The smoothing and differentiation of time series is important in numerous fields. The Savitzky-Golay (SG) filter (Savitzky & Golay, 1964) is widely used for this purpose, especially for biomedical data. For example, the SG filter has been advocated for electroencephalography and electrocardiography (Azami, Mohammadi, & Bozorgtabar, 2012), elastography (Luo, Bai, He, & Ying, 2004), near infrared spectroscopy (Schneider & Kovar, 2003), functional magnetic resonance imaging (Geissler et al., 2007), speech enhancement (Shajeesh, Kumar, Pravena, & Soman, 2012), and eye movement analysis (Nyström & Holmqvist, 2010; Quaia, Joiner, FitzGibbon, Optican, & Smith, 2010; Shaikh, Xu-Wilson, Grill, & Zee, 2010).

Despite its widespread use, the SG filter obeys an unavoidable trade-off between noise suppression and signal distortion, as does any linear filter. Correspondingly, the SG filter tends to oversmooth time series that exhibit abrupt deviations such as sharp waves or steps, yet these are often of physiological interest (e.g., a saccade is an abrupt change in eye position). Consequently, the SG filter can lead to physiological parameters being systematically underestimated (e.g., the peak angular velocity of the eye during a saccade). Accurate estimation of such

Citation: Dai, W., Selesnick, I., Rizzo, J.-R., Rucker, J., & Hudson, T. (2017). A nonlinear generalization of the Savitzky-Golay filter and the quantitative analysis of saccades. *Journal of Vision*, 17(9):10, 1–15, doi:10.1167/17.9.10.

doi: 10.1167/17.9.10

Received March 6, 2017; published August 16, 2017

ISSN 1534-7362 Copyright 2017 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.



Downloaded From: <http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/936403/> on 08/20/2017

parameters is of interest, since they are used in clinical research and practice.

Saccades are rapid eye movements redirecting sight from one object of interest to another (Leigh & Zee, 2015). These movements are brief, rarely lasting more than 80 ms (Leigh & Zee, 2015). Advances in the study of anatomy and pathology of the brain and eye movements have reinforced the substantial utility of saccades in clinical practice. Abnormalities of saccades offer important clues in the diagnosis of a number of neurological disorders (Ramat, Leigh, Zee, & Optican, 2007). Quantitative saccade parameters include latency, amplitude, duration, peak velocity, and peak acceleration, etc. Among these measures, the peak velocity of saccades is of particular interest. Slow saccadic eye movements may suggest a brainstem lesion in excitatory saccade burst neurons (Horn & Büttner-Ennever, 1998; Horn, Büttner-Ennever, Suzuki, & Henn, 1995; Ramat, Leigh, Zee, Shaikh, & Optican, 2008). Examples of conditions that typically cause saccadic slowing include progressive supranuclear palsy (Chen et al., 2010; Leigh & Riley, 2000) or spinocerebellar ataxia type 2 (Garbutt, Harwood, Kumar, Han, & Leigh, 2003; Wadia & Swami, 1971). Additional factors, such as saccade adaption and visual salience, may also affect saccade peak velocity (Ethier, Zee, & Shadmehr, 2008; Schütz, Braun, & Gegenfurtner, 2011); even subtle reductions in the peak velocity of otherwise normal saccades have been observed in states of mental fatigue (Di Stasi et al., 2012). Thus, accurate determination of saccade peak velocity is critical to the understanding of normal and pathological saccade behavior.

We propose a generalization of the SG filter that can track abrupt deviations in time series more accurately than the conventional SG filter. The generalized SG filter leads to more accurate parameter estimates when it is desirable to measure abrupt changes, such as the peak velocity of saccadic eye movements. The proposed filtering methodology models a time series as the sum of two component time series: a low-frequency time series for which the conventional SG filter is well suited, and a second time series that exhibits instantaneous deviations (e.g., sharp waves, steps, or more generally, discontinuities in a higher-order derivative). We formulate the generalized SG filter via sparse signal modeling. The output of the filter is expressed in terms of the solution to a sparse-regularized linear inverse problem, and is calculated via convex optimization. The generalized SG filter is thus nonlinear and can overcome the limitations of the conventional (linear) SG filter.

We model an abrupt deviation in a time series as a discontinuity either in the underlying signal, in the derivative of the underlying signal, or in a higher order derivative of the underlying signal. In turn, the next-

higher order derivative of the time series exhibits impulses (isolated values of high amplitude). Such a time series therefore has a *sparse* component (a time series is said to be *sparse* if it is mostly zero in value, except for isolated values of high amplitude). Accordingly, in the proposed generalization of the SG filter, we adopt tools from sparse signal processing (Elad, 2010). An early instance of filtering based on a sparse signal model is *total variation denoising* (Rudin, Osher, & Fatemi, 1992), which models a step signal as having a sparse first-order derivative.

The proposed generalization of the SG filter is closely related to the *sparsity-assisted signal smoothing* (SASS) method (Selesnick, 2015, 2017; Selesnick, Graber, Pfeil, & Barbour, 2014). The SASS method can be considered a generalization of the digital Butterworth filter (a particular *recursive* filter), while the filter proposed in this work constitutes a generalization of the SG filter (a particular *nonrecursive* filter). Additionally, SASS was developed only for smoothing, while the filter proposed here is also developed for differentiation, which is important for the estimation of velocity time series and associated parameters (e.g., peak velocity).

In this article, we demonstrate that the proposed generalization of the SG filter can estimate peak saccadic velocity from eye position time series more accurately than the conventional SG filter. The evaluation of the proposed filter uses (a) simulated data based on a recently developed parametric model for saccadic eye movements (Dai, Selesnick, Rizzo, Rucker, & Hudson, 2016), (b) physiologic data of saccadic eye movements that have been objectively recorded and manually identified by two experts, Nyström and Andersson (2017), and (c) data recorded in our lab (Rizzo et al., 2016).

Methods

The SG smoothing filter

The SG smoothing filter is a particular low-pass filter defined by two parameters that we shall denote K and M (Orfanidis, 1995; Savitzky & Golay, 1964; Schafer, 2011b). The SG filter can be defined and implemented as a weighted moving average, i.e., a finite impulse response (FIR) filter. We denote by $x(n)$ the time series to be estimated, and we suppose the observed time series is given by $y(n) = x(n) + w(n)$ where $w(n)$ represents additive white Gaussian noise. Applied to time series $y(n)$, the SG filter produces an output time series $\hat{x}(n)$ given by

$$\hat{x}(n) = \sum_{k=-M}^M h(k)y(n-k) \quad (1)$$

where M is a parameter of the filter and $h(n)$ is the impulse response of the filter (defined for $|n| \leq M$).

The SG filter is defined via a least-squares polynomial approximation problem. Specifically, the SG filter defines $\hat{x}(0)$ (i.e., the output at time $n = 0$) to be the value of the coefficient of a polynomial of order K that best fits the time series data $y(n)$ over the interval $|n| \leq M$. Given the time series $y(n)$ and the parameters K and M , define the polynomial

$$p(n) = \sum_{k=0}^K c_k n^k \quad (2)$$

as the polynomial minimizing the square error E ,

$$E = \sum_{n=-M}^M (y(n) - p(n))^2. \quad (3)$$

The output value $\hat{x}(0)$ of the SG filter at time $n = 0$ is defined to be $p(0)$ —that is, the value of the best-fitting polynomial at $n = 0$. Hence, for the SG filter, we have $\hat{x}(0) = p(0) = c_0$ where c_0 is a coefficient of the best-fitting polynomial. The output values $\hat{x}(n)$ at time instants other than $n = 0$ are similarly defined. This process wholly determines the impulse response $h(n)$ of the filter (Orfanidis, 1995). We note that the impulse response of the SG filter is symmetric—that is, $h(-n) = h(n)$. Given M , the parameter K must be chosen so that $0 \leq K \leq 2M$.

The transfer function of a filter with impulse response $h(n)$ is defined as

$$H(z) = \sum_n h(n) z^{-n} \quad (4)$$

and the frequency response is given by $H(e^{j\omega})$. Since the impulse response of the SG filter is symmetric, its frequency response is real-valued, and it is a *zero-phase* filter. Figure 1a shows the frequency response of the SG smoothing filter for several parameter values. The cut-off frequency of the SG filter depends on the parameters K and M , as described by Schafer (2011a). We highlight a property of the SG smoothing filter we use below.

Property 1

Let $H(z)$ be the transfer function of a SG smoothing filter with polynomial order parameter K . Then $H(z)$ can be expressed as

$$H(z) = 1 - (1 - z^{-1})^{K+1} Q(z) \quad (5)$$

where $Q(z)$ is the transfer function of another FIR filter. This property implies that the frequency response

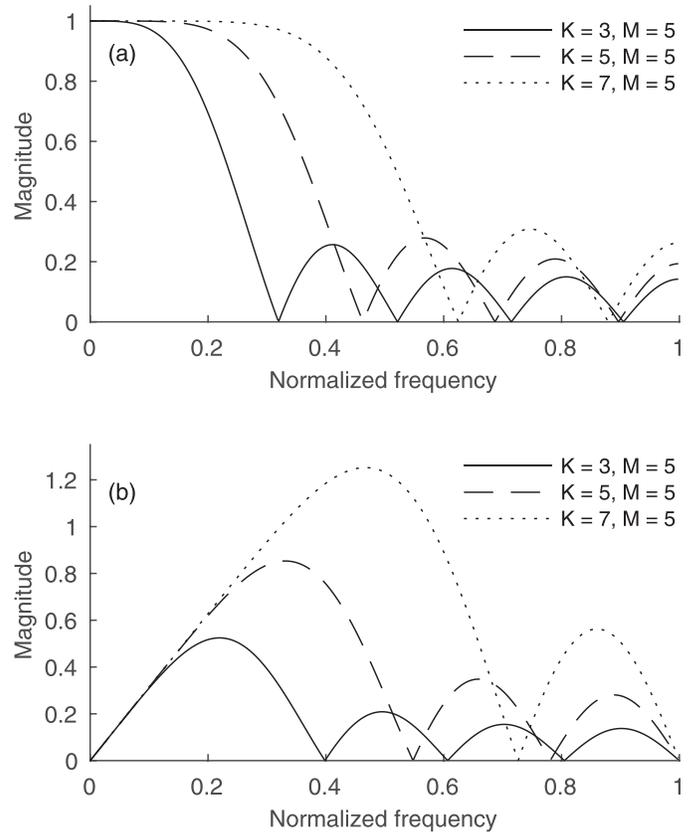


Figure 1. The frequency response of Savitzky-Golay (a) smoothing and (b) differentiation filters for several parameter values.

of the SG smoothing filter is flat (of order K) at $\omega = 0$ (Orfanidis, 1995). The flatness of the frequency response can be seen in Figure 1a.

The SG differentiation filter

SG filters are used not only for smoothing but also for differentiation of time series. Like the SG smoothing filter, the SG differentiation filter is a particular zero-phase low-pass differentiation filter defined by two parameters, K and M , and it can be implemented as an FIR filter. The SG differentiation filter is defined via the differentiation of the polynomial that best fits the data, determined already for the SG smoothing filter. Differentiation of the polynomial p in Equation 2 yields

$$p'(n) = \sum_{k=1}^K \bar{c}_k n^{k-1} \quad (6)$$

where $\bar{c}_k = kc_k$. Like the SG smoothing filter, the impulse response $\bar{h}(n)$ of the SG differentiation filter is determined by parameters K and M . Figure 1b shows the frequency response of the SG differentiation filter

estimate \hat{x}_1 ,

$$\hat{x}_2 = H(y - \hat{x}_1). \quad (23)$$

In turn, we express an estimate of the unknown time series x as

$$\hat{x} = \hat{x}_1 + \hat{x}_2 \quad (24)$$

$$= \hat{x}_1 + H(y - \hat{x}_1) \quad (25)$$

$$= (I - H)\hat{x}_1 + Hy \quad (26)$$

$$= G\hat{x}_1 + Hy \quad (27)$$

$$= RD_S\hat{x}_1 + Hy \quad (28)$$

where G is the high-pass filter (Equation 10), which admits the factorization (Equation 17). By assumption, $D_S\hat{x}_1$ is sparse. Hence, we define $\hat{u} = D_S\hat{x}_1$ and write

$$\hat{x} = R\hat{u} + Hy. \quad (29)$$

Since we model \hat{u} as sparse, we estimate \hat{u} via standard sparse-regularized least squares (Elad, 2010). That is, \hat{u} is calculated as the solution to the optimization problem

$$\hat{u} = \arg \min_u \left\{ \frac{1}{2} \|y - (Ru + Hy)\|_2^2 + \lambda \|u\|_1 \right\} \quad (30)$$

where

$$\|x\|_2^2 := \sum_n x(n)^2 \quad (31)$$

$$\|x\|_1 := \sum_n |x(n)| \quad (32)$$

and λ is a positive parameter. The parameter λ should be chosen according to the noise level. Several algorithms are available to solve Equation 30, e.g., the forward-backward splitting algorithm (Combettes & Pesquet, 2011). Once \hat{u} is obtained, the estimated time series \hat{x} is then given by Equation 29.

A generalization of the differentiation filter

In this section, we use the framework developed in the previous section to extend the SG differentiation filter.

We write the derivative of the time series x as x' and we express it as

$$x' = \bar{F}x \quad (33)$$

where \bar{F} is a zero-phase full-band differentiation filter (i.e., its impulse response is symmetric and it accurately differentiates high-frequency sinusoids).

The ideal full-band differentiation filter cannot be realized by an FIR filter, but it can be well approximated by one (Al-Alaoui, 2007; Kumar, Roy, & Shah, 1992; Pei & Wang, 2001; Rabiner & Schafer, 1974; Selesnick, 2002). In fact, the SG differentiation filter with parameters $K = 2M$ is an approximate full-band differentiation filter. Therefore, we let \bar{F} be a zero-phase full-band differentiation FIR filter. A problem in practice with full-band differentiators is that they severely amplify noise. Hence, by a differentiation filter, we mean by default a *low-pass* differentiation filter. The SG filter with parameters $K < 2M$ is a low-pass differentiation filter.

Let $\bar{h}(n)$ be the impulse response of the SG differentiation filter with parameters K and M . If the input of the filter is the time series y , then the output of the filter is given by the weighted moving average,

$$\hat{x}'(n) = \sum_{k=-M}^M \bar{h}(k)y(n-k). \quad (34)$$

As in Equation 8, we express the output via matrix-vector multiplication,

$$\hat{x}' = \bar{H}y \quad (35)$$

where \bar{H} is a matrix determined by the SG differentiation filter.

In the previous section, we subtracted the zero-phase low-pass filter H from the identity, I , to define a zero-phase high-pass filter G (i.e., $G = I - H$). Here, we subtract the zero-phase low-pass differentiation filter \bar{H} from a zero-phase full-band differentiation filter \bar{F} , to define

$$\bar{G} = \bar{F} - \bar{H}. \quad (36)$$

The filter \bar{G} is a zero-phase high-pass differentiation filter. Here, we let \bar{F} be a SG differentiation filter with parameters $K = 2M$. Then the transfer function of \bar{F} can be written as

$$\bar{F}(z) = (1 - z^{-1}) + (1 - z^{-1})^{2M+1}V(z) \quad (37)$$

where $V(z)$ is the transfer function of another FIR filter. Using Equations 7 and 37, we write the transfer function of filter \bar{G} as

$$\bar{G}(z) = \bar{F}(z) - \bar{H}(z) \quad (38)$$

$$= (1 - z^{-1}) + (1 - z^{-1})^{2M+1}V(z) - (1 - z^{-1}) - (1 - z^{-1})^{K+1}U(z) \quad (39)$$

$$= (1 - z^{-1})^{2M+1}V(z) - (1 - z^{-1})^{K+1}U(z). \quad (40)$$

If $K \leq 2M$, then

$$\bar{G}(z) = (1 - z^{-1})^{K+1} W(z) \quad (41)$$

where $W(z) = (1 - z^{-1})^{2M-K} V(z) - U(z)$.

If $S \leq K + 1$, then using Equations 13 and 41, we may factor $D_S(z)$ out of $\bar{G}(z)$,

$$\bar{G}(z) = (1 - z^{-1})^{K+1} W(z) \quad (42)$$

$$= (1 - z^{-1})^{K+1-S} W(z) D_S(z) \quad (43)$$

$$= \bar{R}(z) D_S(z) \quad (44)$$

where $\bar{R}(z) = (1 - z^{-1})^{K+1-S} W(z)$.

The two-component model

We assume the time series x to be differentiated can be modeled as $x = x_1 + x_2$ as in the previous section. Given the noisy time series $y = x + w = x_1 + x_2 + w$, we seek an estimate \hat{x}' of the unknown differentiated time series x' . Since we assumed x_2 is a signal for which conventional SG filters are well suited, its derivative can be well estimated by the SG differentiation filter \bar{H} —that is, $x'_2 \approx \bar{H}(x_2 + w)$. Consequently,

$$x'_2 \approx \bar{H}(x_2 + w) \quad (45)$$

$$= \bar{H}(y - x_1). \quad (46)$$

Hence, we propose to estimate x'_2 in terms of an estimate \hat{x}_1 ,

$$\hat{x}'_2 = \bar{H}(y - \hat{x}_1). \quad (47)$$

In turn, an estimate of the unknown time series x' can be expressed as

$$\hat{x}' = \hat{x}'_1 + \hat{x}'_2 \quad (48)$$

$$= \bar{F}\hat{x}_1 + \bar{H}(y - \hat{x}_1) \quad (49)$$

$$= (\bar{F} - \bar{H})\hat{x}_1 + \bar{H}y \quad (50)$$

$$= \bar{G}\hat{x}_1 + \bar{H}y \quad (51)$$

$$= \bar{R}D_S\hat{x}_1 + \bar{H}y \quad (52)$$

$$= \bar{R}\hat{u} + \bar{H}y \quad (53)$$

where \bar{G} is the high-pass filter (Equation 36, which admits the factorization in Equation 44) and $\hat{u} = D_S\hat{x}_1$. Once \hat{u} is obtained as the solution to optimization problem (Equation 30), the estimated time series \hat{x}' is given by Equation 53. The framework developed here generalizes SG filters for higher order differentiation as well.

The generalized SG filter: Summary

We summarize the procedure for the proposed generalized SG filter. Let y denote the time series data to be filtered.

1. Set the parameters, K and M of the conventional SG filter. They must be positive integers with $K < 2M$.
2. Set the additional parameters S and λ . Parameter S must be a positive integer with $S \leq K + 1$, which determines the type of abrupt deviation in the time series. Parameter λ must be a positive real number, which is nominally set proportional to the noise level.
3. Define linear filters:
 - a. Define H to be the conventional SG smoothing filter.
 - b. Define $R(z) = [1 - H(z)] / (1 - z^{-1})^S$.
 - c. Define \bar{H} to be the conventional SG differentiation filter.
 - d. Define $\bar{R}(z) = [\bar{F}(z) - \bar{H}(z)] / (1 - z^{-1})^S$.
4. Solve problem (Equation 30) to obtain \hat{u} .
5. Set $\hat{x} = R\hat{u} + Hy$ (smoothing).
6. Set $\hat{x}' = \bar{R}\hat{u} + \bar{H}y$ (differentiation).

The filters H , \bar{H} and \bar{F} can be determined using any realization of the conventional SG filter. For the entire calculation, Matlab software is available from the authors.

Results

In this section, we apply the proposed filtering method for the problem of peak saccadic velocity estimation, and we evaluate its accuracy. First, we simulate saccadic eye movements (horizontal angular displacement) using a recently developed parametric model (Dai et al., 2016). This saccade model has three parameter: η , c , and A . The formula for the saccade model is

$$s(t) = cf(\eta t/c) - cf(\eta t/c - A/c) \quad (54)$$

where f is defined as

$$f(t) = \begin{cases} t + 0.25e^{-2t}, & t \geq 0 \\ 0.25e^{2t}, & t \leq 0. \end{cases} \quad (55)$$

It has been shown that this saccade model can well approximate physiologic saccades (Dai et al., 2016). In this model, the saccade $s(t)$ has an amplitude of A and a peak velocity of

$$V_p = \eta(1 - e^{-A/c}). \quad (56)$$

We note that the relation between peak velocity V_p and

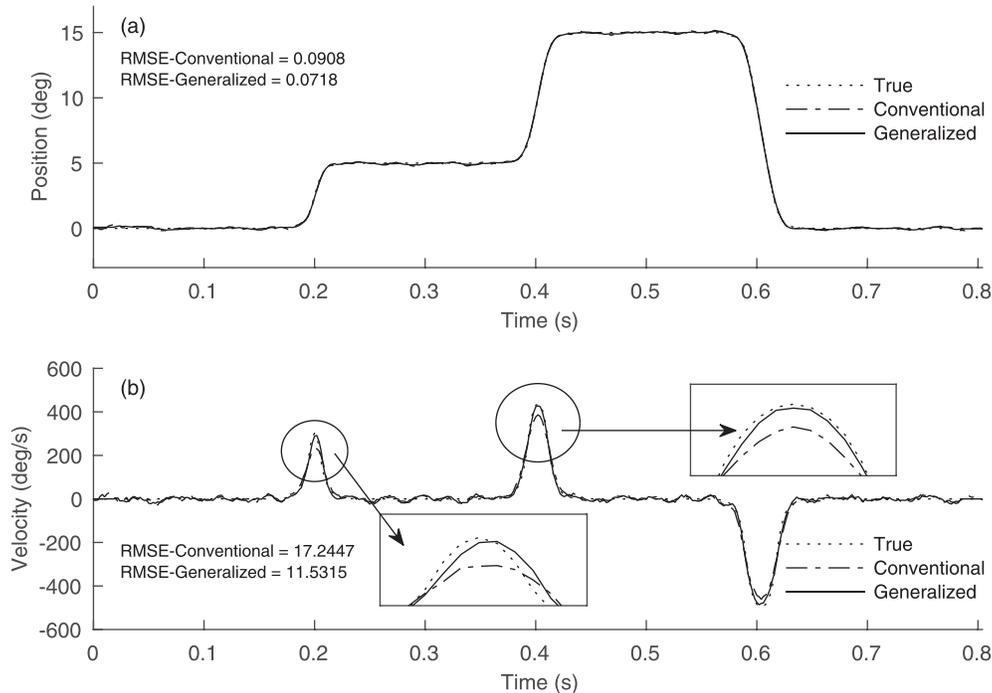


Figure 2. Smoothing (a) and differentiation (b) of simulated saccades of 5°, 10°, and 15° amplitude. Unlike the conventional SG filter, the generalized SG filter does not underestimate the saccade peak velocity.

amplitude A in Equation 56 is the “main sequence” formula proposed by Baloh, Sills, Kumley, and Honrubia (1975).

The use of simulated saccades is useful in this work because true saccade velocities are unknown for human subjects. However, by simulating saccades with known velocity, the error of a velocity estimation method can be quantitatively evaluated. The simulation consists of horizontal saccades made to the right and then back to the central angular position with amplitude 5°, 10°, and 15°, as shown in Figure 2. We simulate data at a sampling rate of 500 samples/s to match the recording condition in our lab. (We used the EyeLink 1000 Plus eye tracker.) While the eight most common sampling rates for eye trackers are 30, 50, 60, 120, 240, 250, 500, and 1000 Hz (Mack, Belfanti, & Schwarz, 2017), eye trackers with low sampling rates are not suitable for the study of saccade dynamics, such as saccade peak velocity. A sampling rate of 333 Hz is considered sufficient for studying position, velocity, and acceleration of human saccadic eye movements (Bahill, Kallman, & Lieberman, 1982; Inchingolo & Spanio, 1985). To this simulated time series, we add a zero-mean white Gaussian random process $\mathcal{N}(0, 0.2)$ to model noise. We use both the conventional and proposed generalized SG filters for smoothing and differentiation. For the conventional SG filter, we use parameters $K = 2$, $M = 5$, as recommended by Nyström and Holmqvist (2010) for

eye movement analysis. For the generalized SG filter, we use parameters $K = 3$, $M = 10$, $S = 4$.

The generalized SG differentiation filter estimates the peak saccadic velocity more accurately than the conventional SG differentiation filter (Figure 2). The conventional SG filter underestimates the peak velocity, especially for small-amplitude saccades. We also evaluate the filters in terms of the root mean square error (RMSE) between the noiseless and filtered time series (Figure 2). The RMSE values further indicate that the generalized SG filter performs smoothing and differentiation more accurately than the conventional SG filter.

To evaluate the generalized SG filter across a range of noise levels, we simulate eye movement data as in Figure 2, but vary the noise standard deviation, σ . We set λ in Equation 30 proportional to σ , because greater λ results in stronger smoothing. Across a range of noise levels, the generalized SG filter estimates position, velocity, acceleration, and saccade peak velocity more accurately (i.e., with lower RMSE) than the conventional SG filter (Figure 3).

To further investigate the accuracy of peak saccadic velocity estimation using the conventional and generalized SG differentiation filters, we simulate 500 saccades of 5° and 15° at a sampling rate of 500 samples/s and add zero-mean white Gaussian random process $\mathcal{N}(0, 0.2)$ to model noise. The distribution of estimated saccade peak velocity is shown as histograms in Figure 4 for each case. The histograms reveal

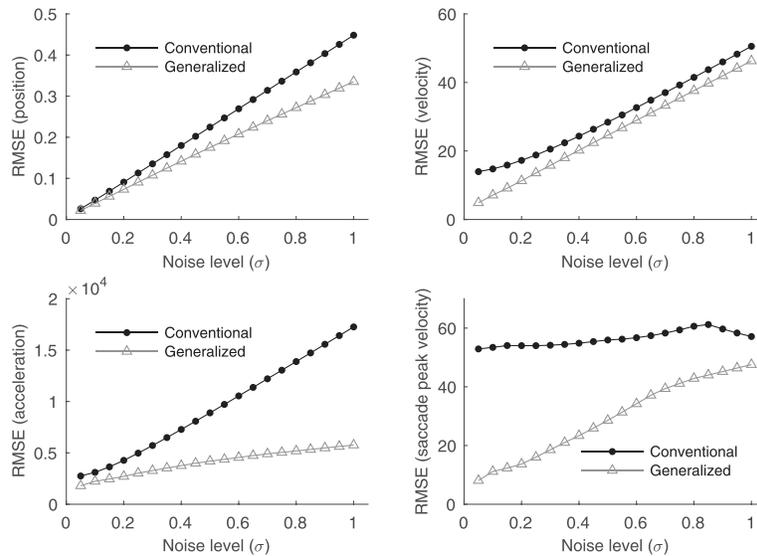


Figure 3. Evaluation of conventional and generalized SG filter across noise level (σ). The generalized SG filter has better RMSE than the conventional SG filter.

that the conventional SG filter is biased; it systematically underestimates the saccade peak velocity. The histograms show that the peak saccadic velocity, as estimated by the generalized SG filter, is generally closer to the true value (indicated by the dashed line).

Systematic underestimation of peak saccadic velocity leads to distortion of the main sequence relationship, an important diagnostic tool for clinicians (Bahill, Clark, & Stark, 1975). To investigate the relative

impact of the filter, we simulate eye movement data comprising 50 saccades of various amplitudes and with random intersaccadic intervals and add a zero-mean white Gaussian random process $\mathcal{N}(0, 0.1)$ to model noise. Figure 5a shows the true main sequence curve and its 95% prediction interval (the true peak velocities and amplitudes are known, since the saccades are simulated). The main sequence curve follows an exponential equation (Equation 56) where V_p is the

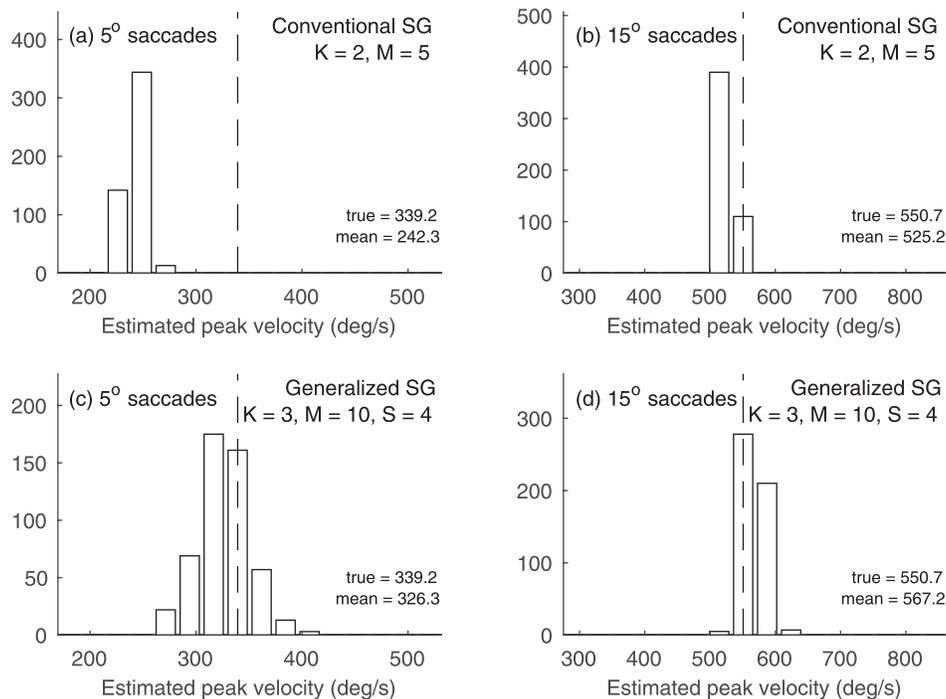


Figure 4. Histogram of peak saccadic velocity of 5° and 15° saccades. (a–b) Estimated using conventional SG filter. (c–d) Estimated using generalized SG filter. Dashed lines indicate the true peak velocities.

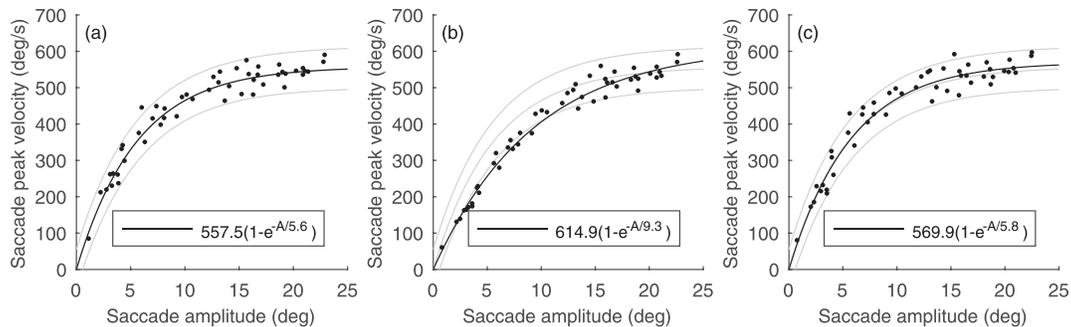


Figure 5. Main sequence relationship between peak velocity and amplitude of simulated saccades. (a) True values. (b) Calculated using conventional SG filter. (c) Calculated using generalized SG filter. The generalized SG filter yields a more accurate main sequence than the conventional SG filter.

saccade peak velocity, A is the saccade amplitude, and η and c are parameters to be estimated (Baloh et al., 1975). Figure 5b and c show the main sequence curves calculated using the conventional SG filter and generalized SG filter, respectively. It can be observed that the conventional SG filter results in a distorted main sequence, and the best-fitting main sequence curve differs substantially from the true main sequence curve. This is because the peak velocities of small-amplitude saccades are underestimated. The generalized SG filter results in a main sequence curve that more accurately estimates the true one.

To examine the generalized SG filter for different individuals, we simulate eye movements as in Figure 5a, but vary the main sequence parameter η or c . We

vary the main sequence parameters because different individuals have different main sequence relations. The shape of the simulated saccade (Equation 54) depends on the parameters η and c . Therefore, we vary either parameter η or c to represent saccades made by 20 individuals. For each individual, we simulate eye movement data comprising 50 saccades of various amplitudes and with random intersaccadic intervals and add zero-mean white Gaussian noise $\mathcal{N}(0, 0.1)$. The main sequence parameters calculated using the conventional SG filter and generalized SG filter are listed in Table 1. It can be observed that the generalized SG filter estimates the main sequence parameters more accurately than the conventional SG filter.

File ID	Parameter η			Parameter c		
	True	Conventional (%)	Generalized (%)	True	Conventional (%)	Generalized (%)
1	510.0	550.0 (8)	506.4 (-1)	6.0	8.7 (45)	5.4 (-9)
2	530.0	593.0 (12)	542.2 (2)	6.0	9.5 (58)	5.9 (-2)
3	550.0	591.2 (7)	558.1 (1)	6.0	8.8 (47)	5.9 (-2)
4	570.0	631.8 (11)	581.8 (2)	6.0	9.7 (62)	6.0 (1)
5	590.0	656.8 (11)	599.9 (2)	6.0	10.0 (67)	5.9 (-1)
6	610.0	688.7 (13)	625.7 (3)	6.0	10.5 (76)	6.3 (4)
7	630.0	694.7 (10)	641.1 (2)	6.0	10.2 (70)	6.0 (0)
8	650.0	708.4 (9)	664.9 (2)	6.0	10.0 (67)	6.1 (1)
9	670.0	735.9 (10)	684.1 (2)	6.0	10.4 (74)	6.2 (4)
10	690.0	760.2 (10)	704.5 (2)	6.0	10.6 (77)	6.1 (2)
11	600.0	662.6 (10)	623.7 (4)	4.2	8.6 (105)	4.7 (11)
12	600.0	660.5 (10)	618.2 (3)	4.6	8.7 (89)	4.8 (5)
13	600.0	657.3 (10)	614.9 (2)	5.0	9.0 (79)	5.3 (5)
14	600.0	655.2 (9)	609.5 (2)	5.4	9.2 (71)	5.4 (0)
15	600.0	663.2 (11)	611.3 (2)	5.8	9.9 (70)	5.9 (2)
16	600.0	675.0 (12)	612.0 (2)	6.2	10.5 (69)	6.2 (1)
17	600.0	651.7 (9)	610.7 (2)	6.6	10.0 (52)	6.6 (0)
18	600.0	659.7 (10)	612.4 (2)	7.0	10.7 (53)	7.0 (1)
19	600.0	653.1 (9)	605.3 (1)	7.4	10.7 (45)	7.2 (-3)
20	600.0	664.6 (11)	607.2 (1)	7.8	11.5 (48)	7.6 (-2)

Table 1. Main sequence parameters, η and c , of simulated saccades estimated using the conventional and generalized SG filters.

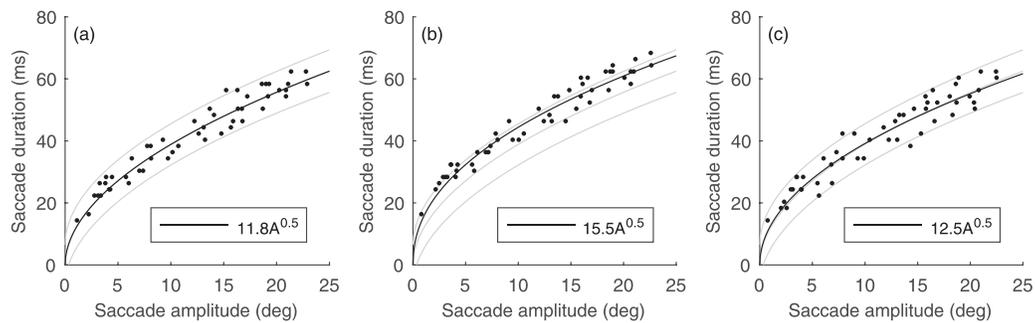


Figure 6. Main sequence relationship between duration and amplitude of simulated saccades. (a) True values. (b) Calculated using conventional SG filter. (c) Calculated using generalized SG filter. The generalized SG filter estimates the true relationship more accurately than the conventional SG filter.

The oversmoothing of eye movement data also leads to prolonged estimates of saccadic duration. Although this phenomenon is not as prominent as the estimation of peak saccadic velocity, it affects the empirical relationship between saccade duration and amplitude. Here, we define the onset and end of a saccade using a velocity threshold of $30^\circ/\text{s}$. As shown in Figure 6b, the main sequence relationship calculated using the conventional SG filter is different from the reference one, shown in Figure 6a. In comparison, the generalized SG filter yields a more accurate estimate of the true main sequence relationship as shown in Figure 6c.

We also compare the conventional and generalized SG filters as applied to real eye movement data. Although no “ground truth” for the real saccade peak

velocity or amplitude is available, it is nevertheless informative to observe that the two filters yield different results, demonstrating the impact of the utilized filter. Figure 7 shows part of an eye movement time series recorded at a sampling rate of 500 samples/s while the participant views an image. This time series is part of a publicly available dataset for which eye movements have been manually annotated (Larsson, Nyström, & Stridh, 2013). Many automatic saccade detection algorithms have been developed and evaluated (Andersson, Larson, Holmqvist, Stridh, & Nyström, 2017; Behrens, MacKeben, & Schröder-Preikschat, 2010; Dai et al., 2016; Engbert & Kliegl, 2003; Komogortsev, Gobert, Jayarathna, Koh, & Gowda, 2010; König & Buffalo, 2014; Nyström & Holmqvist, 2010; Otero-Millan, Castro, Macknik, &

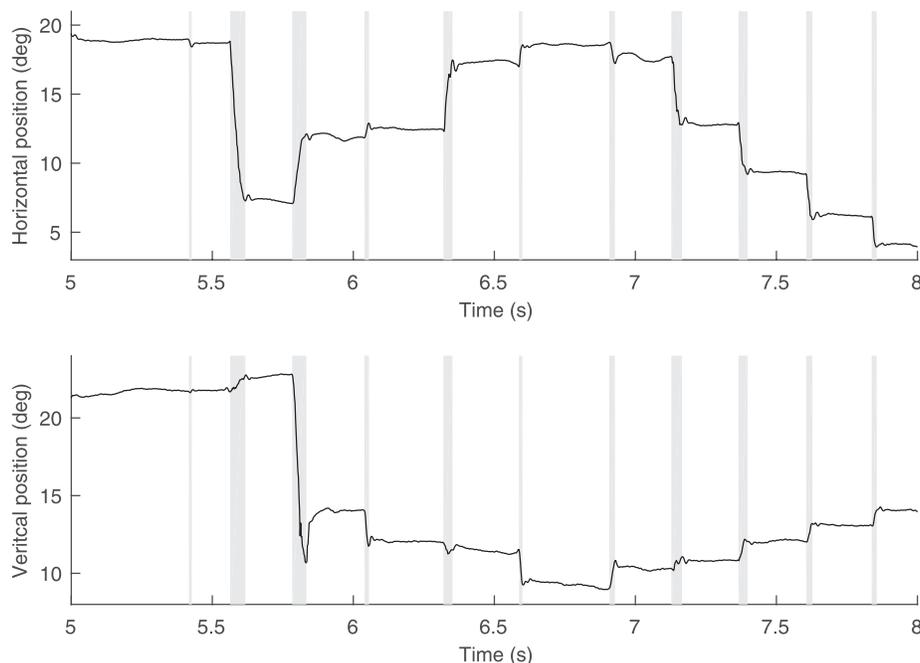


Figure 7. Eye movement of one participant viewing an image. Saccades identified by experts are indicated by shaded gray areas. The annotated data is publicly available online.

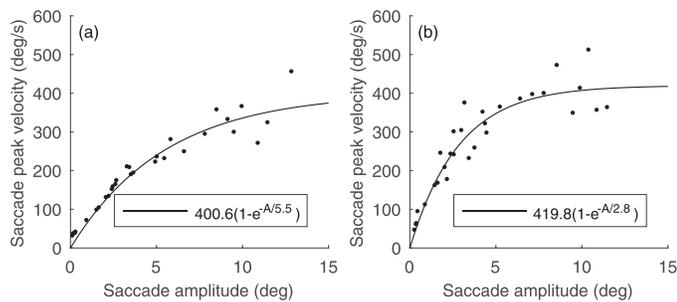


Figure 8. Main sequence relationship between peak velocity and amplitude of real saccades. (a) Calculated using conventional SG filter. (b) Calculated using generalized SG filter.

Martinez-Conde, 2014). Since our focus is on methods to accurately estimate saccade parameters rather than methods to reliably detect saccades, we use the manual annotation instead of an automatic saccade detection algorithm to identify saccades. The horizontal and vertical components of eye movement time series are processed separately. The velocity and amplitude of each eye movement is then calculated by taking the square root of the sum of squares of the horizontal and vertical components (Nyström & Holmqvist, 2010). Using the conventional and generalized SG filters to calculate saccade peak velocity and amplitude, we obtain the values illustrated in Figure 8. The best-fitting exponential curves are calculated based on the respective estimated saccade parameters. It can be observed that the two filters yield distinctly different results. In particular, the main sequence curve obtained using the generalized SG filter is substantially steeper than that obtained using the conventional SG filter. This is consistent with the results using simulated saccades: The conventional SG filter tends to underestimate the peak velocity of small-amplitude saccades.

We further investigate the conventional and generalized SG filter as applied to real eye movement data of 20 individuals. The data was recorded using the EyeLink 1000 Plus eye tracker and sampled at a rate of 500 samples/s while the participants read numbers on a test card from left to right and top to bottom (Rizzo et al., 2016). Saccades are manually annotated by author WD for this study. The main sequence parameters calculated using the conventional SG filter and generalized SG filter are listed in Table 2. The parameter c estimated by the conventional SG filter is larger than that estimated by the generalized SG filter (8.10 ± 2.07 vs. 5.76 ± 2.17 , $p = 0.001$, two-sample t test). The parameter η estimated using either the conventional or generalized SG filter is similar. The same observation is found when the filter is applied to different data: The conventional SG filter tends to underestimate the peak velocity of small-amplitude saccades. Therefore, we conclude that the main

File ID	Parameter η		Parameter c	
	Conventional	Generalized	Conventional	Generalized
1	486.2	530.5	7.8	5.8
2	495.5	563.2	5.8	3.7
3	515.2	552.7	9.7	7.8
4	573.2	568.5	9.2	7.1
5	579.2	625.7	7.9	5.6
6	478.2	453.3	8.6	6.0
7	530.0	522.6	8.7	5.7
8	515.8	501.0	8.2	5.8
9	453.6	463.3	6.0	4.0
10	500.9	473.7	9.5	6.7
11	617.2	604.1	12.9	10.1
12	427.3	369.7	6.9	3.8
13	546.5	540.7	7.3	3.8
14	467.2	498.4	5.8	3.8
15	494.7	440.1	8.0	4.2
16	412.1	425.8	5.8	4.1
17	429.2	547.2	8.8	9.8
18	715.1	697.2	12.7	9.7
19	457.9	444.9	6.5	4.2
20	477.9	504.5	5.8	3.5

Table 2. Main sequence parameters, η and c , of real saccades estimated using the conventional and generalized SG filters.

sequence estimated using the generalized SG filter is more accurate than the one estimated using the conventional SG filter.

Discussion

The SG filter is a particular (linear) low-pass filter widely used for smoothing and differentiation of time series. Other commonly used filters for eye movement data is the moving average filter and the Butterworth filter (Mack et al., 2017). They are used for data smoothing and usually followed by a two-point central difference filter for data differentiation. All three filters are linear filters and hence obey an unavoidable trade-off between noise suppression and signal distortion. It should also be recognized that linear low-pass filters tend to oversmooth abrupt deviations from a smooth trend. Therefore, quantitative measures of signal extremes (e.g., the peak value of a signal derivative) might be underestimated when calculated using the SG filter, or any other linear filter. The proposed nonlinear generalization of the SG filter aims to mitigate this issue, by modeling a signal component having discontinuities in its order- S derivative of some prescribed order S .

Historically, the relationship between saccade peak velocity and amplitude has been a useful metric in

clinical practice for the detection and assessment of various disease states; however, saccades of different amplitudes have different power spectra (Harris, Wallman, & Scudder, 1990). Hence, the use of a conventional low-pass filter with a fixed cut-off frequency may be suboptimal for the analysis of eye movement data comprising saccades of various sizes. One comprehensive study used a saccade model to evaluate the three commonly used linear filters (moving average, Savitzky-Golay, and Butterworth) as applied to eye movement data (Mack et al., 2017). Their results suggest various filters should be used for saccades of different sizes. However, eye movement data usually contains saccades of various sizes. For example, we make small intraline saccades and large interline saccades during reading. Recent studies showed that chronic concussion patients and multiple sclerosis patients made more saccades to complete a rapid number naming test, and these extra saccades appear to be small (Hainline et al., 2017; Rizzo et al., 2016). As shown in the Results section, the conventional SG filter tends to underestimate the peak velocity of saccades, especially small-amplitude saccades. The proposed nonlinear generalization of the SG filter may more accurately estimate saccade peak velocity and other saccade parameters.

In the Results section, we used white Gaussian random processes to model noise. Studies of artificial eyes for the measurement of noise in eye trackers showed that the noise is white (it has the same energy at all frequencies) when artificial eyes are used (Coey, Wallot, Richardson, & Van Orden, 2012; Wang, Mulvey, Pelz, & Holmqvist, 2017). Both studies point out that human data produce pink noise (the energy decreases as the frequency increases). This may be because the human eye is never still and always makes micro movements, which can be similar in amplitude to the noise. We conducted numerical experiments adding pink noise instead of white noise in the simulations. The generalized SG filter was found to more accurately estimate peak velocity in pink noise as well as white noise.

Conclusions

In the quantitative analysis of saccades, the limitations of linear low-pass filters should be taken into account, given their potential to negatively impact objective metrics. This article demonstrates through simulation using a recently developed parametric saccade model that the SG filter, a widely used linear low-pass filter, may systematically underestimate the peak velocity of saccades. This article proposes a nonlinear generalization of the SG filter that can more

accurately track deviations from the smooth background of an underlying signal. The generalized SG filter was shown to estimate saccade peak velocity more accurately than the conventional SG filter, which is commonly used for this purpose. In turn, the generalized SG filter leads to a more accurate estimation of the main sequence (i.e., the relationship between the saccade peak velocity and amplitude). Given the clinical significance of quantitative saccade analysis in disease diagnosis, the generalized SG filters may be useful in the study of eye movement data, particularly saccades.

Keywords: Savitzky-Golay filter, smoothing, differentiation, sparsity, saccade peak velocity, quantitative saccade analysis

Acknowledgments

This work was supported in part by the National Science Foundation under Grant No. 1525398.

Commercial relationships: none.

Corresponding author: Weiwei Dai.

Email: wd471@nyu.edu.

Address: Department of Electrical and Computer Engineering, Tandon School of Engineering, New York University, Brooklyn, NY, USA.

References

- Al-Alaoui, M. A. (2007). Linear phase low-pass IIR digital differentiators. *IEEE Transactions on Signal Processing*, 55(2), 697–706.
- Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., & Nyström, M. (2017). One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods*, 49(2), 616–637.
- Azami, H., Mohammadi, K., & Bozorgtabar, B. (2012). An improved signal segmentation using moving average and Savitzky-Golay filter. *Journal of Signal and Information Processing*, 3(1), 39–44.
- Bahill, A. T., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences*, 24(3), 191–204.
- Bahill, A. T., Kallman, J. S., & Lieberman, J. E. (1982). Frequency limitations of the two-point central difference differentiation algorithm. *Biological Cybernetics*, 45(1), 1–4.

- Baloh, R. W., Sills, A. W., Kumley, W. E., & Honrubia, V. (1975). Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology*, 25(11), 1065–1070.
- Behrens, F., MacKeben, M., & Schröder-Preikschat, W. (2010). An improved algorithm for automatic detection of saccades in eye movement data and for calculating saccade parameters. *Behavior Research Methods*, 42(3), 701–708.
- 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*, 1, 147.
- Coey, C. A., Wallot, S., Richardson, M. J., & Van Orden, G. (2012). On the structure of measurement noise in eye-tracking. *Journal of Eye Movement Research*, 5(4), 1–10.
- Combettes, P. L., & Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, & H. Holikowicz (Eds.), *Fixed-point algorithms for inverse problems in science and engineering* (pp. 185–212). New York: Springer-Verlag.
- Dai, W., Selesnick, I., Rizzo, J.-R., Rucker, J. C., & Hudson, T. E. (2016, December). A parameteric model for saccadic eye movement. In *IEEE Signal Processing in Medicine and Biology Symposium* (pp. 1–6), doi:10.1109/SPMB.2016.7846860. New York: IEEE.
- Di Stasi, L. L., Renner, R., Catena, A., Cañas, J. J., Velichkovsky, B. M., & Pannasch, S. (2012). Towards a driver fatigue test based on the saccadic main sequence: A partial validation by subjective report data. *Transportation Research Part C: Emerging Technologies*, 21(1), 122–133.
- Elad, M. (2010). *Sparse and redundant representations: From theory to applications in signal and image processing*. New York: Springer.
- Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9), 1035–1045.
- Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Changes in control of saccades during gain adaptation. *Journal of Neuroscience*, 28(51), 13929–13937.
- Garbutt, S., Harwood, M. R., Kumar, A. N., Han, Y. H., & Leigh, R. (2003). Evaluating small eye movements in patients with saccadic palsies. *Annals of the New York Academy of Sciences*, 1004(1), 337–346.
- Geissler, A., Gartus, A., Foki, T., Tahamtan, A. R., Beisteiner, R., & Barth, M. (2007). Contrast-to-noise ratio (CNR) as a quality parameter in fMRI. *Journal of Magnetic Resonance Imaging*, 25(6), 1263–1270.
- Hainline, C., Rizzo, J.-R., Hudson, T. E., Dai, W., Birkemeier, J., Raynowska, J., . . . Rucker, J. C. (2017). Capturing saccades in multiple sclerosis with a digitized test of rapid number naming. *Journal of Neurology*, 264(5), 989–998.
- Harris, C. M., Wallman, J., & Scudder, C. A. (1990). Fourier analysis of saccades in monkeys and humans. *Journal of Neurophysiology*, 63(4), 877–886.
- Horn, A. K. E., & Büttner-Ennever, J. A. (1998). Premotor neurons for vertical eye movements in the rostral mesencephalon of monkey and human: Histologic identification by parvalbumin immunostaining. *Journal of Comparative Neurology*, 392(4), 413–427.
- Horn, A. K. E., Büttner-Ennever, J. A., Suzuki, Y., & Henn, V. (1995). Histological identification of premotor neurons for horizontal saccades in monkey and man by parvalbumin immunostaining. *Journal of Comparative Neurology*, 359(2), 350–363.
- Inchingolo, P., & Spanio, M. (1985). On the identification and analysis of saccadic eye movements—A quantitative study of the processing procedures. *IEEE Transactions on Biomedical Engineering*, 9, 683–695.
- Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D. H., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering*, 57(11), 2635–2645.
- König, S. D., & Buffalo, E. A. (2014). A nonparametric method for detecting fixations and saccades using cluster analysis: Removing the need for arbitrary thresholds. *Journal of Neuroscience Methods*, 227, 121–131.
- Kumar, B., Roy, S. C. D., & Shah, H. (1992). On the design of FIR digital differentiators which are maximally linear at the frequency $\pi/p, p \in \{\text{positive integers}\}$. *IEEE Transactions on Signal Processing*, 40(9), 2334–2338.
- Larsson, L., Nyström, M., & Stridh, M. (2013). Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit. *IEEE Transactions on Biomedical Engineering*, 60(9), 2484–2493.
- Leigh, R. J., & Riley, D. E. (2000, March). Eye

- movements in parkinsonism: It's saccadic speed that counts. *Neurology*, 54(5), 1020–1021.
- Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements*. Cambridge, UK: Oxford University Press.
- Luo, J.-W., Bai, J., He, P., & Ying, K. (2004). Axial strain calculation using a low-pass digital differentiator in ultrasound elastography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 51(9), 1119–1127.
- Luo, J.-W., Ying, K., He, P., & Bai, J. (2005). Properties of Savitzky-Golay digital differentiators. *Digital Signal Processing*, 15(2), 122–136.
- Mack, D. J., Belfanti, S., & Schwarz, U. (2017). The effect of sampling rate and lowpass filters on saccades—A modeling approach. *Behavior Research Methods*, E-pub ahead of print, doi:10.3758/s13428-016-0848-4.
- Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*, 42(1), 188–204.
- Orfanidis, S. J. (1995). *Introduction to signal processing*. Upper Saddle River, NJ: Prentice-Hall.
- Otero-Millan, J., Castro, J. L. A., Macknik, S. L., & Martinez-Conde, S. (2014). Unsupervised clustering method to detect microsaccades. *Journal of Vision*, 14(2):18, 1–17, doi:10.1167/14.2.18. [PubMed] [Article]
- Pei, S.-C., & Wang, P.-H. (2001). Closed-form design of maximally flat FIR Hilbert transformers, differentiators, and fractional delayers by power series expansion. *IEEE Transactions on Circuits and Systems I*, 48(4), 389–398.
- Quaia, C., Joiner, W. M., FitzGibbon, E. J., Optican, L. M., & Smith, M. A. (2010). Eye movement sequence generation in humans: Motor or goal updating? *Journal of Vision*, 10(14):28, 1–31, doi:10.1167/10.14.28. [PubMed] [Article]
- Rabiner, L. R., & Schafer, R. W. (1974). On the behavior of minimax relative error FIR digital differentiators. *Bell Labs Technical Journal*, 53(2), 333–361.
- Ramat, S., Leigh, R. J., Zee, D. S., & Optican, L. M. (2007). What clinical disorders tell us about the neural control of saccadic eye movements. *Brain*, 130(1), 10–35.
- Ramat, S., Leigh, R. J., Zee, D. S., Shaikh, A. G., & Optican, L. M. (2008). Applying saccade models to account for oscillations. *Progress in Brain Research*, 171, 123–130.
- Rizzo, J.-R., Hudson, T. E., Dai, W., Birkemeier, J., Pasculli, R. M., Selesnick, I., . . . Rucker, J. C. (2016). Rapid number naming in chronic concussion: Eye movements in the King-Devick test. *Annals of Clinical and Translational Neurology*, 3(10), 801–811.
- Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1–4), 259–268.
- Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8), 1627–1639.
- Schafer, R. W. (2011a). On the frequency-domain properties of Savitzky-Golay filters. In *Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop* (pp. 54–59). New York: IEEE.
- Schafer, R. W. (2011b). What is a Savitzky-Golay filter? *IEEE Signal Processing Magazine*, 28(4), 111–117.
- Schneider, R. C., & Kovar, K.-A. (2003). Analysis of ecstasy tablets: Comparison of reflectance and transmittance near infrared spectroscopy. *Forensic Science International*, 134(2), 187–195.
- Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision*, 11(5):9, 1–30, doi:10.1167/11.5.9. [PubMed] [Article]
- Selesnick, I. (2002). Maximally flat low-pass digital differentiators. *IEEE Transactions on Circuits and Systems II*, 49(3), 219–223.
- Selesnick, I. (2015). Sparsity-assisted signal smoothing. In *Excursions in harmonic analysis* (Vol. 4, pp. 149–176). Basel, Switzerland: Birkhäuser (Springer).
- Selesnick, I. (2017, March). Sparsity-assisted signal smoothing (revisited). In *IEEE International Conference on Acoustics, Speech, Signal Processing* (pp. 4546–4550), doi:10.1109/ICASSP.2017.7953017. New York: IEEE.
- Selesnick, I., Graber, H. L., Pfeil, D. S., & Barbour, R. L. (2014). Simultaneous low-pass filtering and total variation denoising. *IEEE Transactions on Signal Processing*, 62(5), 1109–1124.
- Shaikh, A. G., Xu-Wilson, M., Grill, S., & Zee, D. S. (2010). ‘Staircase’ square-wave jerks in early Parkinson’s disease. *British Journal of Ophthalmology*, 95(5), 705–709, doi:10.1136/bjo.2010.179630.
- Shajeesh, K. U., Kumar, S., Pravena, D., & Soman, K. P. (2012). Speech enhancement based on Savitzky-Golay smoothing filter. *International Journal of Computer Applications*, 57(21), 39–44.

Wadia, N. H., & Swami, R. K. (1971). A new form of heredo-familial spinocerebellar degeneration with slow eye movements (nine families). *Brain*, *94*(2), 359–374.

Wang, D., Mulvey, F. B., Pelz, J. B., & Holmqvist, K. (2017). A study of artificial eyes for the measurement of precision in eye-trackers. *Behavior Research Methods*, *49*(3), 947–959.