Estimation of Neural Inputs and Detection of Saccades and Smooth Pursuit Eye Movements by Sparse Bayesian Learning

Federico Wadehn[1], David J. Mack[2], Thilo Weber[1], Hans-Andrea Loeliger[1].

***Abstract*— Eye movements reveal a great wealth of information about the visual system and the brain. Therefore, eye movements can serve as diagnostic markers for various neurological disorders. For an objective analysis, it is crucial to have an automatic and robust procedure to extract relevant eye movement parameters. An essential step towards this goal is to detect and separate different types of eye movements such as fixations, saccades and smooth pursuit. We have developed a model-based approach to perform signal detection and separation on eye movement recordings, using source separation techniques from sparse Bayesian learning. The key idea is to model the oculomotor system with a state space model and to perform signal separation in the neural domain by estimating sparse inputs which trigger saccades. The algorithm was evaluated on synthetic data, neural recordings from rhesus monkeys and on manually annotated human eye movement recordings with different smooth pursuit paradigms. The developed approach shows a high noise-robustness, provides saccade and smooth pursuit parameters, as well as estimates of the position, velocity and acceleration profiles. In addition, by estimating the input to the oculomotor system, we obtain an estimate of the neural inputs to the oculomotor muscles.**

I. Introduction

The eyes are, metaphorically speaking, a window to the soul. Similarly, eye movements can be seen as a window to the brain [1]. In this work, we will focus on *smooth pursuit eye movements* (SPEM), fixations and saccades. During *fixations*, information is gathered from a detail (the target) in the visual environment. Fast, jerky eye movements known as *saccades*, separate phases of fixation and guide the gaze from one detail of interest to the next [2]. If the target starts moving, SPEM will try to keep it in the center of gaze. Interestingly, if the motion pattern is predictable, humans can (almost) completely overcome the latency of the oculomotor system and perfectly follow the target [3]. This suggests the creation of an internal representation of the target motion. All goal-directed eye movements described above involve brain-wide control networks, making their study valuable in diverse fields ranging from marketing to neuroscience to clinical applications [1]. Especially in the latter setting, the precise temporal profiles of eye movements are of interest as several neurodegenerative diseases affect parameters such as saccadic peak velocity [4]. The tight relationship between saccade amplitude, duration and peak velocity, known as the *main sequence* [5], has diagnostic value for various neurological disorders. Similarly, the quality of SPEM is indicative of a patient's motion processing ability.

To compute such parameters, event detection algorithms have to extract saccades, fixations and SPEM from eye movement recordings. This is not a simple task due to various artefacts, like blinks, baseline drift and sensor noise. Unwary filtering of the signal to remove noise, can severely degrade the quality of the extracted parameters [6]. The velocity profile is commonly obtained by numerical differentiation, which inevitably amplifies noise. To distinguish different types of eye movements, various event detection algorithms have been proposed [7], [8]. Since saccades are much faster than other eye movements, the simplest detection algorithms are based on thresholding of the velocity (or acceleration) profile to tell saccades and non-saccades apart –an approach known as *IVT* [7]. Due to their simplicity, IVT variants are still regularly in use [9]. Such approaches, however, have difficulties in the presence of high or varying noise levels and during SPEM.

In this paper we present a model-based approach, grounded in oculomotor physiology, for the estimation of neural inputs, detection of saccades, fixations and SPEM. A minimal working example of the code is made available[1]. The oculomotor plant is modeled via a linear state space model, whose (sparse) inputs trigger the neural control signals generating the forces of the oculomotor muscles. Relying on techniques from sparse Bayesian learning [10], [11], the algorithm can detect these triggers and separate saccades from SPEM.

II. Model of the Oculomotor System

The model consists of two parts: The oculomotor plant and the two neural input systems (see Fig. 1). The complete model and each subsystem will be described with a linear state space model (SSM) of the form:

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_{k-1} \\ y_k &= Cx_k + z_k, \end{aligned} \tag{1}$$

with $k \in \{1, ..., N\}$, where $N \gg 1$ denotes the duration of the recording, u_k the input, z_k the measurement noise and y_k the output of the system.

A. Mechanistic Oculomotor Plant Model

Bahill [12] developed a physiological oculomotor plant model for horizontal eye movements. The model describes

[1]Signal and Information Processing Laboratory, ETH Zurich, Zurich, Switzerland. Email: {loeliger, wadehn}@isi.ee.ethz.ch and thiweber@ee.ethz.ch

[2]University Hospital Zurich, Zurich, Switzerland. Current Email: david-jule.mack@iis.ee.ethz.ch

[1]https://bitbucket.org/FedericoW/embc18_saccade_spem_code

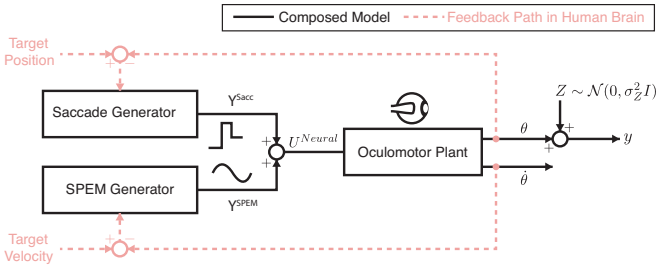

Fig. 1: Neural systems feeding into the oculomotor plant. The (noise-free) output is the angular position θ of the eye ($\widehat{=}\ x_1^{\text{Plant}}$ of oculomotor plant SSM).

the interaction of the bulbus oculi (mass) with two extra-ocular muscles (medial and lateral rectus) via a mass-spring-damper system of the form:

$$x_k^{\text{Plant}} = A_{\text{Plant}} \cdot x_{k-1}^{\text{Plant}} + B_{\text{Plant}} \cdot u_{k-1}^{\text{Neural}} \tag{2}$$

$$y_k = C_{\text{Plant}} \cdot x_k^{\text{Plant}} + z_k, \tag{3}$$

where A_{Plant} and B_{Plant} are obtained by discretizing the 6th order continuous-time SSM from [12]. The neural input u^{Neural} ($\widehat{=}$ output of neural control system) is responsible for generating the force of the agonist and antagonist muscles. The output of the oculomotor plant state space model y_k is the (noisy) eye position measured by eye tracker at time k. The first four components of the state vector $x_k \in \mathbb{R}^6$ are the horizontal eye position, the (effective) agonist and antagonist muscle length and the velocity of the eye. Since the output of the model is the eye position, the read-out matrix is $C_{\text{Plant}} = (1, 0, 0, 0, 0, 0)$. The active state tension (force [12]) of the agonist and antagonist muscles are represented by the last two components $x_k^{(5)}$ and $x_k^{(6)}$. Since only the combined effect of the agonist and antagonist muscle can be estimated, we reduced the model to a 5th order SSM by merging the agonist and antagonist dynamics. The new state $x_k^{(5)}$ represents the difference in active state tension between the agonist and antagonist muscle pair.

B. Neural Input Model

Depending on the neural input u^{Neural}, the oculomotor plant (2) can realize fixations, saccades and SPEM. In the following, we will introduce the two subsystems generating the neural inputs for saccades, fixations (with random jitter) and SPEM.

1) Saccades and Fixations Model: Saccades are thought to be generated by exciting the extra-ocular muscles with a pulse-step firing pattern [13]. A pulse-step signal can be modeled with a first-order hold state [10], which is triggered by sparse inputs u^{Sacc}:

$$x_k^{\text{Sacc}} = x_{k-1}^{\text{Sacc}} + u_{k-1}^{\text{Sacc}} \tag{4}$$

$$y_k^{\text{Sacc}} = x_k^{\text{Sacc}} \tag{5}$$

Figure 2 shows a simulated pulse-step neural input signal $u^{\text{Neural}} = y^{\text{Sacc}}$ that is integrated by the oculomotor plant. This results in the depicted saccade position θ and velocity $\dot{\theta}$ profiles (cf. Fig. 2 top and center).

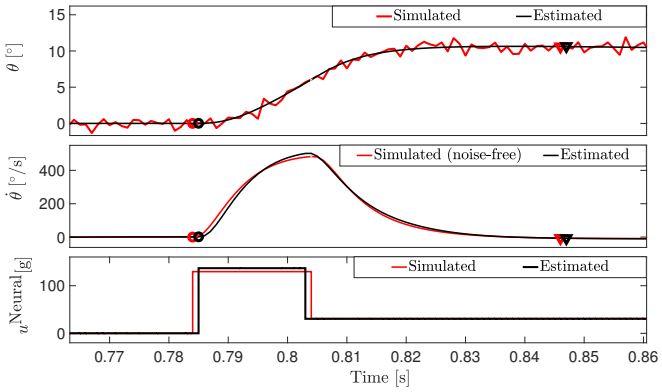

Fig. 2: Top: Simulated 10° saccade with additive Gaussian noise and estimated position profile. The dots and triangles indicate the start and end of the saccade. Center: Simulated velocity (noise-free) and estimated velocity from noisy position profile. Bottom: Simulated and estimated neural input triggering the saccade. The neural inputs to the muscles are in force units (g = grams) as in [12] to avoid a cumbersome conversion to spike rates.

2) SPEM Models: SPEM can be induced by controlled experiments in which a subject is instructed to visually follow a moving target. To achieve low latency tracking, the brain builds an internal representation of the target motion. We account for this internal representation by the following SPEM model:

$$x_k^{\text{SPEM}} = A_{\text{SPEM}} \cdot x_{k-1}^{\text{SPEM}} + B_{\text{SPEM}} \cdot u_{k-1}^{\text{SPEM}} \tag{6}$$

$$y_k^{\text{SPEM}} = C_{\text{SPEM}} \cdot x_k^{\text{SPEM}} \tag{7}$$

Different motion profiles of the target: Sinusoidal (Fig. 3), ramp (Fig. 4), parabolic, etc., can be realized with suitable system matrices. To be agnostic about the type of elicited SPEM, one can use a random walk on the velocity [10] with

$$A_{\text{SPEM}} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \quad B_{\text{SPEM}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{8}$$

driven by an i.i.d. Gaussian input u^{SPEM} and a read-out matrix $C_{\text{SPEM}} = (1, 0)$. For a sinusoidal pattern with target frequency $\omega_{\text{SPEM}} = 2\pi f_{\text{SPEM}}$ and sampling period T_s, we use the following system matrices:

$$A_{\text{SPEM}} = \begin{bmatrix} 1 & T_s & 0 \\ 0 & \cos(\omega_{\text{SPEM}} T_s) & \sin(\omega_{\text{SPEM}} T_s) \\ 0 & -\sin(\omega_{\text{SPEM}} T_s) & \cos(\omega_{\text{SPEM}} T_s) \end{bmatrix}, \; B_{\text{SPEM}} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{9}$$

and $C_{\text{SPEM}} = (1, 0, 0)$. In this case, the state vector of the SPEM system can be seen as a rotating phasor. Figure 3 shows the signal separation results on real eye tracker data in which a human subject was instructed to visually follow a sinusoidally moving target (a bright dot that moves from the left to the right side of the screen).

3) Force Input: The neural input to the oculomotor plant (2) is the sum of the output signals of the SPEM and saccade subsystems:

$$u_k^{\text{Neural}} = y_k^{\text{Sacc}} + y_k^{\text{SPEM}}. \tag{10}$$

III. Estimation of Neural Inputs

The interconnected system, consisting of the oculomotor plant and the SPEM/saccade subsystems (in total three

SSMs), can compactly be written as a single SSM of the form (1). When the inputs to the SSM are given (in terms of their means and variances), state estimates, i.e. the (adaptively filtered) position, velocity and force profiles, can readily be obtained via Kalman smoothing [14]. In this paper, however, the inputs are unknown and have to be estimated too. For non-sparse inputs, the maximum a posteriori estimate (MAP) can be obtained by Kalman smoothing or equivalently Gaussian message passing [15], [16]. For this, we model the SPEM input u^{SPEM} as i.i.d. Gaussians with unknown means, but fixed variance. The MAP estimate $\hat{u}^{\text{SPEM}}$ is the mean of the posterior distribution $p(u^{\text{SPEM}}|y_{1:N})$, which can easily be computed by Kalman smoothing [10], [15].

Saccades in contrast, are triggered by sparse inputs u^{Sacc}. To model such sparse inputs, we resort to the main idea from sparse Bayesian learning [11]: Unknown sparse quantities are modeled as zero-mean Gaussians with unknown variance:

$$u_k^{\text{Sacc}} \sim \mathcal{N}(0, \sigma^2_{u_k^{\text{Sacc}}}), \quad k \in \{1, ..., N\}. \tag{11}$$

Maximum likelihood estimation (or MAP with a suitable prior) of the unknown variances is known to be sparsifying [10], [11]. The unknown parameters (here unknown variances $\sigma^2_{u^{\text{Sacc}}}$) can be estimated via the expectation maximization (EM) algorithm [17], where the E-step of the EM-algorithm boils down to Kalman smoothing. In the M-step, the parameters of the unknown inputs u^{SPEM} and u^{Sacc} are iteratively updated using first and second moments (means and variances) computed in the E-step [10], [11].

Saccade and SPEM Parameter Extraction: The estimated position θ and velocity $\dot{\theta}$ profiles (cf. Fig. 2) can be used to detect saccade onsets and endings, e.g., using an IVT algorithm. Given the saccadic onsets and endings as well as the estimated velocity profile, the saccadic main sequence parameters can be computed. From the estimated SPEM signal y^{SPEM} (cf. Fig. 3) parameters such as the instantaneous phase and amplitude can be obtained.

IV. Results

We evaluated the presented algorithm on synthetic data as well as on real data from humans and rhesus monkeys. All procedures followed the guidelines set by the National Institutes of Health and national law and were approved by local ethics committees.

Simulated Data: Table I shows the saccade detection performance (precision and recall) of our algorithm on synthetic data. The data was generated with the oculomotor plant model presented in sec. II under varying measurement noise noise levels σ_Z. In addition, we assessed how well parameters of interest such as the saccadic peak velocity are captured.

Human Subjects: Figure 3 and 4 show the separation of saccades and SPEM for a sinusoidally moving target and for a ramp paradigm. The estimated amplitude and phase of the SPEM could serve as diagnostic markers, since in many neurodegenerative diseases the relevant neural control networks are impaired such that zero-latency tracking is not possible anymore. Furthermore, we evaluated the signal separation and saccade detection performance of our algorithm on 19 recordings of horizontal eye movements from [8]. In this experiment subjects were instructed to follow a ramp paradigm as shown in Fig. 4. Table II shows the performance of the algorithm against expert manual annotations from [8].

Estimation of Neural Signals in Monkeys: Previously published neuronal data from rhesus monkeys [19] was used to test the quality of the estimated neural inputs. The data consists of eye position recordings sampled at 1 kHz with simultaneously recorded firing rates of single neurons of the abducens nucleus. For the estimation, we used the oculomo-

Noise level	Precision	Recall	Rel. peak vel.
$\sigma_Z = 0.05°$	1	1	0.99
$\sigma_Z = 0.2°$	1	1	0.91
$\sigma_Z = 0.8°$	1	0.76	0.77
$\sigma_Z = 3.2°$	0.96	0.46	0.87

TABLE I: Precision, recall and rel. peak velocity $v^{\text{est}}_{\text{max}}/v^{\text{sim}}_{\text{max}}$ estimated on synthetic data with varying noise levels. The data consists of sinusoidal SPEM and 50 saccades whose amplitudes are logarithmically distributed between $0.5°$ and $30°$.

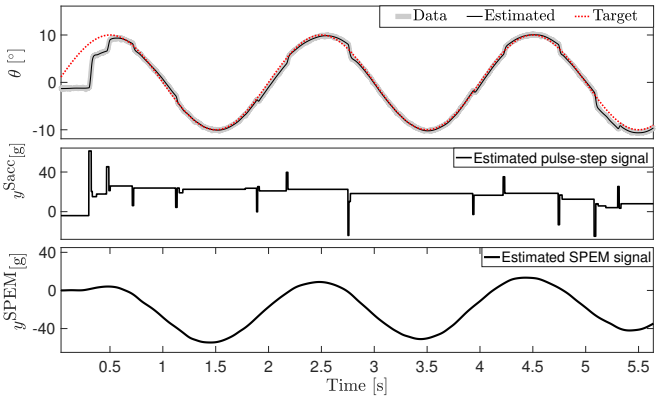

Fig. 3: Separation of SPEM and (catch-up) saccades in real eye tracking data from [18] with model (9). In the experiment a sinusoidally moving target was shown (viewing distance: 57 cm, amplitude: 10 °, frequency: 0.5 Hz). The data were recorded at 1 kHz using a limbustracker (Skalar IRIS, Skalar Medical B.V., Delft, Netherlands).

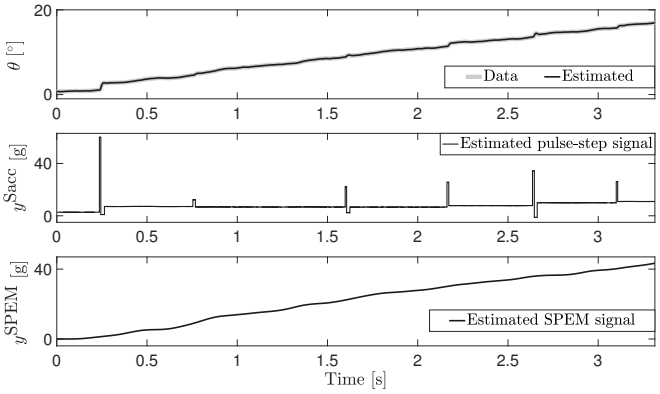

Fig. 4: Separation of SPEM and (catch-up) saccades in real eye tracking data taken from [8] with model (8).

# Recordings	Cohen's κ	Precision	Recall
19	0.76	0.94	0.94

TABLE II: Cohen's κ, precision and recall on real data with manual annotations of saccades from [8] as baseline. The total number of annotated saccades was 60, of which 58 were detected.

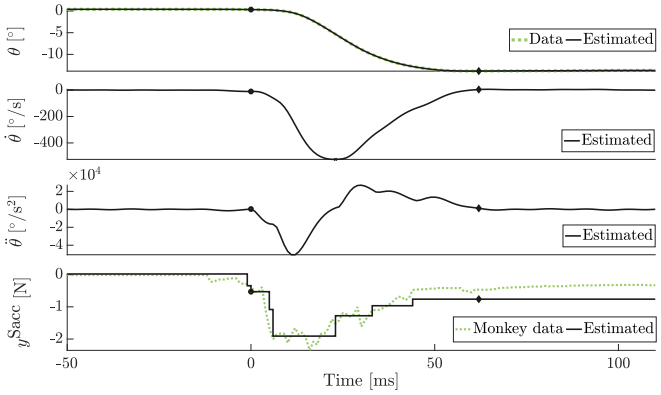

Fig. 5: Top and bottom: Low-noise recording of eye position (scleral search coil) and neural firing rate from single abducens motoneuron (tungsten microelectrodes, FHC Inc.) in a rhesus monkey during a saccade. Top to bottom: Estimated position, velocity, acceleration and neural input signal. As in [20] the neural firing rate is converted to the force unit (N = Newton).

tor plant parameters for rhesus monkeys described in [20]. From Fig. 5 we see that the shape (especially the decay) of the estimated neural input matches well the real recordings. Note that the recording is from a single motoneuron, whereas the estimated input represent the aggregated neural input of many motoneurons to the oculomotor plant. To compare the shape, we rescaled the neural recording to fit the magnitude of the estimated neural input as done in [20].

V. Discussion and Conclusion

We have presented a model-based approach that relies on sparse input estimation to separate saccades, fixations and smooth pursuit eye movements. In addition to smoothed position, velocity, acceleration and force profiles, the algorithm is capable of estimating the neural input to the oculomotor muscles. The algorithm is robust against noise and, due to the model-based approach, highly accurate in detecting saccades. The current model considers each eye separately and its model parameters are tailored to the horizontal gazing direction. A natural extension would include a system identification step, where patient-specific parameters of the oculomotor plant are learned directly from the recordings. For this, an EM algorithm akin to the one described for learning input variances, can be envisioned. Similarly, using the same framework, blinks which are currently not accounted for, could be modeled with an outlier noise source [21]. Finally, due to the modular structure of the approach, a further natural extension is to build a model that captures both eyes and both gaze directions simultaneously. A starting point could be the 2D oculomotor plant model [22]. Further interesting investigations include the evaluation of the discriminative power of the extracted saccade and SPEM parameters, in particular for detecting neurodegenerative diseases.

Acknowledgement

We thank Prof. Uwe Ilg for his comments and for providing the sinusoidal SPEM data. We also thank Prof. Hans-Peter Thier and Dr. Peter Dicke of the Hertie-Institute for Clinical Brain Research for the neural recordings.

References

[1] A. G. Shaikh and D. S. Zee, "Eye movement research in the twenty-first century- a window to the brain, mind, and more," *The Cerebellum*, 2017.
[2] A. L. Yarbus, *Eye movements and vision*. Springer, New York, 1967.
[3] A. T. Bahill and J. D. McDonald, "Smooth pursuit eye movements in response to predictable target motions," *Vision Research*, vol. 23, no. 12, 1983.
[4] D. S. Zee, L. M. Optican, J. D. Cook, D. A. Robinson, and W. K. Engel, "Slow saccades in spinocerebellar degeneration." *Archives of Neurology*, vol. 33, no. 4, 1976.
[5] A. T. Bahill, M. R. Clark, and L. Stark, "The main sequence, a tool for studying human eye movements," *Mathematical Biosciences*, vol. 24, no. 3-4, 1975.
[6] D. J. Mack, S. Belfanti, and U. Schwarz, "The effect of sampling rate and lowpass filters on saccades–a modeling approach," *Behavior Research Methods*, vol. 49, no. 6, 2017.
[7] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols." in *Proc. of the Symposium on Eye Tracking Research & Applications*, Palm Beach Gardens, Florida, USA.
[8] R. Andersson, L. Larsson, K. Holmqvist, M. Stridh, and M. Nyström, "One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms," *Behavior Research Methods*, vol. 49, no. 2, 2017.
[9] L. Larsson, M. Nyström, and M. Stridh, "Detection of saccades and postsaccadic oscillations in the presence of smooth pursuit." *IEEE Trans. Biomed. Eng.*, vol. 60, no. 9, 2013.
[10] H.-A. Loeliger, L. Bruderer, H. Malmberg, F. Wadehn, and N. Zalmai, "On sparsity by NUV-EM, Gaussian message passing, and Kalman smoothing," in *Information Theory and Applications Workshop (ITA)*, San Diego, California, USA, 2016.
[11] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *Journal of Machine Learning Research*, vol. 1, 2001.
[12] A. T. Bahill, J. R. Latimer, and B. T. Troost, "Linear homeomorphic model for human movement," *IEEE Trans. Biomed. Eng.*, no. 11, 1980.
[13] A. F. Fuchs and E. S. Luschei, "Firing patterns of abducens neurons of alert monkeys in relationship to horizontal eye movement." *Journal of Neurophysiology*, vol. 33, no. 3, 1970.
[14] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013, vol. 3.
[15] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proc. of the IEEE*, vol. 95, no. 6, 2007.
[16] F. Wadehn, L. Bruderer, D. Waltisberg, T. Keresztfalvi, and H.-A. Loeliger, "Sparse-input detection algorithm with applications in electrocardiography and ballistocardiography," in *Int. Conf. on Bio-inspired Systems and Signal Processing*, Lisbon, Portugal, 2015.
[17] Z. Ghahramani and G. E. Hinton, "Parameter estimation for linear dynamical systems," University of Totronto, Tech. Rep., 1996.
[18] U. J. Ilg. Schülerlabor Neurowissenschaften. Retrieved on Jan. 2018. [Online]. Available: www.neuroschool-tuebingen-schuelerlabor.de
[19] M. Prsa, P. W. Dicke, and P. Thier, "The absence of eye muscle fatigue indicates that the nervous system compensates for non-motor disturbances of oculomotor function," *Journal of Neuroscience*, vol. 30, no. 47, 2010.
[20] W. Zhou, X. Chen, and J. Enderle, "An updated time-optimal 3rd-order linear saccadic eye plant model," *International Journal of Neural Systems*, vol. 19, no. 05, 2009.
[21] F. Wadehn, L. Bruderer, J. Dauwels, V. Sahdeva, H. Yu, and H.-A. Loeliger, "Outlier-insensitive Kalman smoothing and marginal message passing," in *24th European Signal Processing Conference (EUSIPCO)*, Budapest, Hungary, 2016.
[22] O. Komogortsev, C. Holland, S. Jayarathna, and A. Karpov, "2D linear oculomotor plant mathematical model: Verification and biometric applications," *ACM Trans. on Applied Perception*, vol. 10, no. 4, 2013.