



# LUND UNIVERSITY

## Detection of oculomotor events using random forest

Zemblys, Raimondas; Niehorster, Diederick C; Holmqvist, Kenneth

2016

### *Document Version:*

Publisher's PDF, also known as Version of record

[Link to publication](#)

### *Citation for published version (APA):*

Zemblys, R., Niehorster, D. C., & Holmqvist, K. (2016). *Detection of oculomotor events using random forest*. Abstract from The Scandinavian Workshop on Applied Eye Tracking 2016, Turku, Finland.

### *Total number of authors:*

3

### **General rights**

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

## **Detection of oculomotor events using random forest**

Raimondas Zemblys<sup>1</sup>, Diederick C. Niehorster<sup>2</sup>, & Kenneth Holmqvist<sup>2</sup>

<sup>1</sup> Siauliai University, <sup>2</sup> Lund University

Event detection is a challenging stage in eye movement data analysis. The major drawback of current event detection methods is that parameters have to be adjusted based on eye movement data quality. However, a fully automated classification of raw gaze samples as fixations, saccades or other oculomotor events can be achieved using a machine learning approach. Any already manually or algorithmically detected events can be used to train a model to produce similar classification of other data without the need to set parameters. In this study we explore the application of random forest machine learning technique for the detection of fixations, saccades and post-saccadic oscillations (PSOs). Specifically, we train separate classification models using manually (MAN) and automatically (NH, Nystrom and Holmqvist, 2010) annotated data utilizing features used in classical deterministic algorithms: velocity, acceleration, dispersion, directional data, etc. The NH model achieved 94% (Cohen's Kappa,  $K=81.2$ ) classification accuracy on test set of automatically annotated data and 90.6% ( $K=62.2$ ) on manually annotated data. The MAN model achieved 93.2% ( $K=71.8$ ) and 95.3% ( $K=87.2$ ) accuracy, respectively. As a baseline, the agreement between the NH and MAN training data was 92.1% ( $K=66.5$ ). Our findings furthermore reveal that PSOs are the most challenging events to classify, because even expert annotators have a hard time tagging onsets and offsets of PSOs. For example, the NH model misclassifies 1/3 of the PSO samples as fixation and almost 1/5th as saccade samples, compared to the NH algorithm output itself. Variable importance analysis further indicates that for the NH model, velocity, acceleration and directional information are the most important features to achieve classification, while the MAN model utilizes standard deviation and RMS instead of directional information.

We conclude that machine learning techniques are a very good substitute for current state of the art event detection algorithms and manual coding, resulting in very similar classification. We will further extend the classifier to achieve good classification in noisy data and data with a low sampling rate, and to detect other eye-movement events (such as smooth pursuit and nystagmus) by including other training sets.