2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Non-parametric analysis of eye-tracking data by anomaly detection

Daniel Jansson, Olov Rosén, and Alexander Medvedev
Department of Information Technology, Uppsala University
Uppsala, Sweden
E-mail: daniel.jansson@it.uu.se, olov.rosen@it.uu.se, alexander.medvedev@it.uu.se

***Abstract*— A non-parametric approach for distinguishing between individuals by means of recorded eye movements is suggested. The method is based on the principles of stochastic anomaly detection and relies on measured data for probability distribution estimation and evaluation. For visual stimuli that excite the essential nonlinear dynamics of the human oculomotor system, mean gaze trajectories and characterizations of their uncertainty are approximated per individual using eye-tracking data. With this information, eye-tracking profiles are established, against which independently acquired data sets are statistically tested to evaluate the probability that they belong to said profiles. Both Gaussian function fitting and kernel density estimation (KDE) techniques are used for distribution estimation and novel means for testing observations against general distributions are suggested. It is shown that the presented method yields promising results in terms of individual classification based on eye movements. Further, using the KDE method for trajectory distribution estimation provides better selectivity compared to normal distribution fitting.**

I. Introduction

Studies of human eye movements have been greatly simplified by the introduction of modern video-based eye-tracking techniques. It has indeed been the topic of numerous recent studies and several attempts have been made to use eye-tracking data to quantify the functionality of the human oculomotor system.

There are different types of eye movement (the two most commonly mentioned being *saccades* and *smooth pursuit*) [2], all of which are governed by complex neuromuscular systems. Research has shown that various medical conditions impair the oculomotor system to different degrees. For example, Huntington's Chorea [1], Schizophrenia [10] and Parkinson's Disease [3] compromise the smooth pursuit system (SPS), motivating the search for accurate quantification methods, which could then be used as diagnosing or even staging tools.

If measures can be found that effectively distinguish between healthy individuals using eye movements, they are also likely to detect the alternations in the oculomotor system resulting from a clinical condition.

An attempt at such a measure, mentioned in previous research [6], [7], is the *smooth pursuit gain* (SPG). The SPG is defined by the ratio of the angular velocity of the eye to that of a moving target during smooth pursuit and has been claimed to vary between individuals of different age. It is also said to be lower in subjects with Parkinson's Disease [6]. Although this is a promising result, the SPG is not an exhaustive measure of the oculomotor system's function as it is descriptive of only one point in the SPS frequency characteristics. In [8], it was suggested to model the SPS as a dynamical system and use eye-tracking data to identify its unknown parameters, thus introducing the dynamic SPG (DSPG) measure.

DSPG yields full dynamic models per individual, but may be unreliable if the chosen model structure is indescriptive of the underlying system. Problems may also arise when the studied data set contains saccades or other types of eye movement not governed by the smooth pursuit mechanism and hence not properly described by the selected model. It is therefore of interest to find a non-parametric approach to use as a supplement to the model-based methods. Non-parametric methods rely entirely on data and no modeling assumptions about the studied system are made.

In this paper, such a model-free approach for distinguishing between healthy individuals using recorded eye movements is derived. It is based on the principles of anomaly detection and uses statistical methods to find deviating data sets. A visual stimulus consisting of a white circle moving along a specially designed trajectory on a black background on a computer monitor is presented to a test subject multiple times. The gathered eye-tracking data are then used to obtain approximate statistical properties of the resulting gaze trajectories in the form of a mean trajectory and distribution estimates at each time instant. The mean trajectory and the distribution estimates give an individual eye-tracking profile against which an independently acquired data set can be statistically tested in order to evaluate the probability of the latter to belong to the profile. In order to carry out such tests for arbitrarily distributed data, a way of efficiently testing observations against general distributions is proposed.

The paper is composed as follows: In Sec. II an overview of the method of generating eye-tracking profiles from data is presented. Sec. III describes how normal and KDE distributions can be approximated from data. The definition of the outlier region and means to find it for normal and general distributions is given in Sec. IV. The experimental setup and conducted experiments are described in Sec. V and the corresponding results are given in Sec. VI. Finally, conclusions are stated and the methods and results are discussed in Sec. VII.

II. METHOD

A. Visual stimulus

The visual stimuli used herein were generated using the method in [9]. The stimulus generation method presented therein gives input sequences with the rich spectral and amplitude excitation needed to accurately identify the nonlinear Wiener-type models used to portray the SPS. The stimuli presented to the test subject are smooth random movements of a white circle in a 25 cm $\times$ 25 black background window on a computer monitor.

B. Method overview

Assume that N_s data sets of eye movements are recorded from a test subject tracking the same trajectory of the visual stimulus multiple times on different occasions. Due to the complex nature of the oculomotor system, the response to a visual stimulus will not be the same for repeated exposures. Hence, the N_s data sets will not be equal. For each of the N_t time instances at which the gaze direction is sampled, there will be N_s data points, one from each set of recorded eye movements. Since horizontal and vertical gaze direction coordinates are logged separately, the data points will have two components.

The data points at time instance k will be seemingly random with some expected value, and can thus be seen as N_s observations of a two-dimensional stochastic variable, $X(k)$. Note that there will be one stochastic variable per time instance. If the stochastic variables $X(k)$ are assumed to be independent, the distribution of $X(k)$ for each k can be estimated from data. The assumption that $X(k)$ are independent is restrictive, but is shown to produce sufficient results for the purposes of this study. The distribution of $X(k)$ will depend on the trajectory of the visual stimulus, but also on the individual tracking it. If the probability density function (PDF) of $X(k)$ for $k = 1, 2, \ldots, N_t$ is known for an individual, it is possible to determine whether a given data set is likely to come from the same individual or not. For each of the N_t time instances, a hypothesis test with the null hypothesis that the data are indeed observations of $X(k)$, can be carried out. If the number of time instances, in which the deviation of the data from the considered distribution is high, the data set is deemed to be the gaze trajectory of a different individual. The approach generalizes in a straightforward manner to the case of a group of test subjects sharing a property, such as e.g. healthy persons or persons of a certain age.

In practice, the distributions of $X(k)$, $k = 1, 2, \ldots, N_t$, are not known, but can be estimated from data. The simplest way is to use a histogram. However, since the data are two-dimensional, a large number of data points is needed to achieve sufficiently small bin widths for reliable statistical testing. To acquire a large number of data points, a test subject would have to track the same visual stimulus a large number of times, which would be time-consuming and tedious. Therefore, a way of acquiring smooth two-dimensional PDF estimates from relatively small sample data sets must be found.

III. DISTRIBUTION ESTIMATION

The normal distribution (Gaussian distribution) is commonly used to approximate the statistical properties of the underlying stochastic variable of a given data set. However, normal distribution fitting may not always give satisfactory results, depending on the studied quantity.

Testing for normality can be done in different ways. A histogram of the data can be inspected to resemble a normal distribution PDF. A Q-Q plot may also indicate that the data is not normally distributed [13]. Here, the Q-Q plot and the Lilliefors test are used. The Lilliefors test is an adaptation of the Kolmogorov-Smirnov test for normality [5].

Should the data appear to not be normally distributed, an alternative method to distribution estimation is *kernel density estimation* (KDE). In KDE nothing is presumed about the underlying distribution to be estimated. It is thus a non-parametric way of estimating an unknown PDF, in contrast to the parametric method of normal distribution fitting.

The details of the two approaches for PDF estimation are given below.

A. Normal distribution estimation

Assume that the sample $(x_1,\ x_2,\ \ldots,\ x_{N_s})$, $x_i \in \mathbf{R}^2$ is drawn from some distribution. Normal distribution fitting is a parametric way of estimating the unknown PDF.

The general expression for a two-dimensional normal distribution function is

$$\phi(x) = \frac{1}{2\pi\sqrt{|\Sigma|}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}. \tag{1}$$

where $\mu \in \mathbf{R}^2$ is the mean and $\Sigma \in \mathbf{R}^{2\times 2}$ is the covariance matrix of the distribution. The matrix determinant is denoted $|\cdot|$. Fitting (1) to the data is a matter of finding estimates of μ and Σ. The distribution mean, μ, can be estimated by the sample mean

$$\hat{\mu} = \frac{1}{N_s} \sum_{i=1}^{N_s} x_i, \tag{2}$$

and Σ by the sample covariance matrix

$$\hat{\Sigma} = \frac{1}{N_s - 1} \sum_{i=1}^{N_s} (x_i - \hat{\mu})(x_i - \hat{\mu})^T. \tag{3}$$

Once $\hat{\mu}$ and $\hat{\Sigma}$ are computed, they can be inserted into (1) to give the resulting PDF estimate.

B. Kernel density estimation

KDE is a non-parametric method used to obtain a smooth estimate of the PDF of an unknown distribution. Assume that the observations $(x_1,\ x_2,\ \ldots,\ x_{N_s})$, $x_i \in \mathbf{R}^2$ are drawn from some unknown distribution with probability density $f : \mathbf{R}^2 \rightarrow \mathbf{R}$. The kernel density estimator of f is given by

$$\hat{f}(x) = \frac{1}{N_s h} \sum_{i=1}^{N_s} K\left(\frac{x - x_i}{h}\right), \tag{4}$$

where K is the chosen kernel function and h is a user parameter called the *bandwidth*. Here, the kernel was chosen to be the two-dimensional standard normal density function

$$\phi(x) = \frac{1}{2\pi} e^{-\frac{1}{2}x^T x}. \quad (5)$$

Using this kernel, the kernel density estimator can be rewritten as

$$\hat{f}(x) = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{1}{2\pi h} e^{-\frac{1}{2h^2}(x-x_i)^T(x-x_i)} = \frac{1}{N_s} \sum_{i=1}^{N_s} \psi(x - x_i), \quad (6)$$

where $\psi(x)$ is a two-dimensional normal density function with standard deviation h and mean 0. The estimate of the probability density f is thus the mean of N_s two-dimensional Gaussian functions with standard deviations h (note that the kernel is circular) and means given by the observations x_i.

The choice of the bandwidth h has great impact on the resulting kernel estimator. In the one-dimensional case, if the underlying distribution being estimated is Gaussian, an optimal choice of h can be derived [11]. This is not possible in the general case, and an appropriate value must be found experimentally.

IV. Evaluating the outlier region

Assume that an observation, $x \in \mathbf{R}^2$, is made and that it must be determined whether it is likely to be an observation of a given random variable X, or not. A hypothesis test with the null hypothesis:

- H_0: x is an observation of X

must be carried out. One way to do this is to define an *outlier region*, S, of the random variable, being the set of all possible observations deemed unlikely to come from the considered distribution, i.e all x for which H_0 is rejected. The probability that an observation of X lies in S should be low. Define α such that

$$P(X \in S) = \int_S f(x)dx = \alpha. \quad (7)$$

Consequently, α is the probability with which an observation of the considered random variable is deemed (incorrectly) to be from some other distribution. The choice of α will influence the size of the outlier region S. Evaluation of the outlier region is performed depending on the considered distribution. Methods for testing observations against normal distributions are well-known and widely used.

A. Normal distribution

For the case when the two-dimensional random variable X has a normal distribution with mean μ and covariance Σ, the outlier region can be derived as follows: Let $Y = \Sigma^{-\frac{1}{2}}(X - \mu)$, so that $Y \in N(0, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix. Form the variable

$$Z = (X - \mu)^T \Sigma^{-1} (X - \mu) = Y^T Y = Y_1^2 + Y_2^2, \quad (8)$$

where Y_i denotes the i:th component of Y. Z is the square of the Mahalanobis distance between X and μ and is a sum of the squares of two standard normal random variables and will therefore have a χ_2^2 distribution [4]. The outlier region of X can thus be defined as

$$S_1 = \left\{x : (x - \mu)^T \Sigma^{-1} (x - \mu) \geq \chi_2^2(1 - \alpha)\right\} \quad (9)$$

where $\chi_2^2(p)$ is the quantile function for probability p of the χ_2^2 distribution. Eq. (9) shows that the outlier region, S_1, is the exterior of an ellipse centered around the distribution mean. Thus, for a given observation x, H_0 is rejected if $x \in S_1$, which is easily established by inserting x into (9).

B. General distribution

The kernel density estimates of the PDF are not Gaussian functions and the outlier region must be derived differently than in the normal distribution case. The PDF may have several peaks and consequently the Mahalanobis distance from the distribution mean is no longer a plausible test quantity. The outlier region may be non-convex or even disjoint.

The outlier region of a random variable X with probability density f is given by

$$S_2 = \left\{x : \int_{S_2} f dx = \alpha, f(x) \leq f(x_c), \forall x_c \in S_2^c\right\}, \quad (10)$$

where S_2^c is the complement set of S_2. If f is positive and continuous, S_2 will be uniquely defined by (10). This can be realized by considering the integral

$$J = \int_Q f(x)dx \quad (11)$$

where

$$Q = \{x : f(x) \leq \gamma\}, \quad (12)$$

and γ is some positive constant. If f is positive and continuous, increasing γ in (12) will continuously increase the value of J in (11). For some value of γ, denoted by γ_T, J will be equal to α and thus Q will be equal to S_2. Since there is only one value of γ which yields $J = \alpha$, S_2 is unique.

Determining whether a given observation x is part of S_2 is not straightforward from (10). Finding S_2 analytically is in general not possible. The following method is suggested for numerically calculating the outlier region:

- Evaluate f for the finite set of uniformly spaced grid points $\{x_i\}_{i=1}^M$ to obtain $\{f_i\}_{i=1}^M$.
- Let $\{f_{(i)}\}_{i=1}^M$ be $\{f_i\}_{i=1}^M$ sorted in ascending order.
- Find K such that $\sum_{i=1}^K f_{(i)} \leq \frac{\alpha}{A} < \sum_{i=0}^{K+1} f_{(i)}$, where A is the area of a grid element.
- An approximation, $\hat{S}_2$, of S_2 is then given by

$$\hat{S}_2 = \{x_i : f(x_i) \leq f_{(K)} = \gamma_T\}. \quad (13)$$

In simpler terms: Keep summing the smallest elements of $\{f_i\}_{i=1}^M$ while the sum does not exceed $\frac{\alpha}{A}$. Denote the largest term in the sum by γ_T. Let S_2 be the set of all x_i corresponding to the summed f_i. For a given observation x, H_0 is rejected if $x \in \hat{S}_2$, which is easily checked by inserting x into (13).

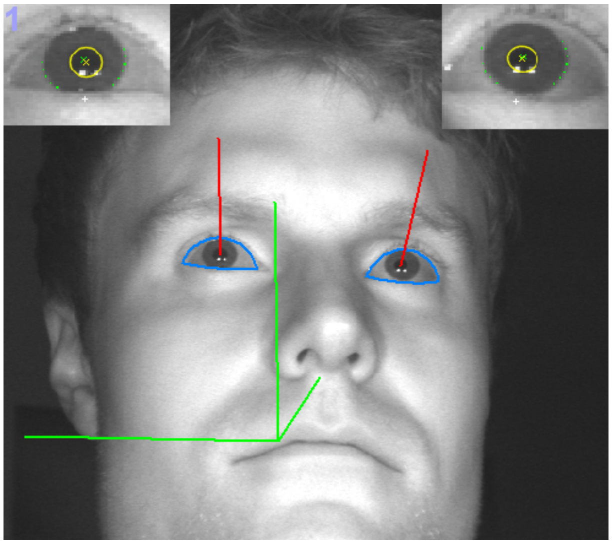

Fig. 1. Screen shot from the eye-tracking software.

V. Experiment

Gaze direction data of test subjects attempting to track the moving circle on a computer monitor were recorded using a video-based eye tracker from Smart Eye AB, Sweden. The eye tracker output is the distance in centimeters (horizontal and vertical components separately) between the monitor center and the point where the gaze direction line intersects the monitor. Eye-tracking data were sampled at a sampling frequency of $f_s = 60$ Hz. Fig. 1 shows a screen shot from the eye-tracking software, illustrating how the eye tracker evaluates the gaze direction.

The conducted experiment involved four test subjects:

- P1: Man, 26 years old
- P2: Man, 27 years old
- P3: Man, 54 years old
- P4: Man, 54 years old

P1 was first asked to track the $T = 26\ s$ long visual stimulus 25 times spread over a one week period to avoid memorization of the stimulus trajectory. This generated 25 data sets, each with $N_t = Tf_s = 26 \cdot 60 = 1560$ time samples. In the same manner, 25 data sets were gathered from P3. Further, P2 and P4 were asked to track the same stimulus 5 times each. A total of 60 data sets were acquired.

VI. Results

Distributions at each time instance were estimated from the 20 first data sets of both P1 and P3. The rest of the recorded data sets were used to compare with the estimated distributions, i.e to test whether they are observations from the distributions in question or not. The 20 first data sets of P1 and P3 will be referred to as the *estimation sets*. The remaining 5 sets of P1 and P3, and the 5 data sets from P2 and P4, will be referred to as the *testing sets*. The goal is to investigate the ability of the presented method to determine whether the testing sets come from P1, P3 or neither.

A. Testing for normality

The distribution of the data varies between different time instances and a test for normality must be done in each instance. For illustrative purposes, the estimation sets of P1 were examined at time instance 1200. The data at several other time instances had similar characteristics. Fig. 2 shows a Q-Q plot and a histogram of the horizontal data (x-coordinate) for time instance 1200. Judging from these two plots, the data may not be normally distributed. Carrying out a Lilliefors test indeed showed that the data were not from a normal distribution at the 95% significance level.

Performing the Lilliefors test for all of the 1560 time instances in the estimations sets of P1 and P3 showed that the data were not likely to be normally distributed in 39.7% of the cases for P1 and 44.2% for P3.

B. Distribution estimation

Using the estimation sets of P1, two-dimensional PDFs for each of the 1560 time instances were estimated using the two methods presented in Sec. III: Normal distribution fitting and kernel density estimation. The PDF in each time step was thus estimated from 20 observations of the underlying stochastic variable. The same was done for the estimation sets of P3.

For the KDE method, five different choices of the bandwidth h were used. The five choices of the bandwidth were $h = \{0.1\hat{\sigma},\ 0.25\hat{\sigma},\ 0.5\hat{\sigma},\ \hat{\sigma},\ 2\hat{\sigma}\}$, where $\hat{\sigma}$ is either the sample standard deviation of the horizontal data or the vertical data, depending on which of the two is larger. The bandwidth thus varies between the time instances. As an example, the estimated normal distribution and the corresponding KDE with $h = 0.5\hat{\sigma}$, for P1 in time instance 1200 are shown in Fig. 3.

C. Outlier detection

The testing sets were compared to the distributions estimated in Sec. VI-B. The number of outliers in each testing set was computed by inserting the observation at every time instance into the expressions for the outlier regions S_1 and S_2, with significance level $\alpha = 0.01$. For every distribution estimate, the mean number of outliers in each data set was calculated. The mean number of outliers in the testing sets pertaining to test subject Pk is denoted by m_{Pk}.

Fig. 4 shows a heat map of the estimated KDE distribution for P1 at time instance 1200 along with the outlier region and an observation from a testing set of P2 lying outside the region.

The results of comparison of the data sets to the distribution of P1 are presented in Tab. I. The results of comparison of the data sets to the distribution of P3 are presented in Tab. II. The numbers are given in percent of the total number of time steps. The entries in Tab. I and Tab. II thus give the amount of time during which the gaze direction of the test subjects deviated significantly from the mean trajectory of the estimation sets from P1 and P3 respectively.

It is evident from Tab. I that the values of the number of outliers in the testing sets of P1 are significantly lower than

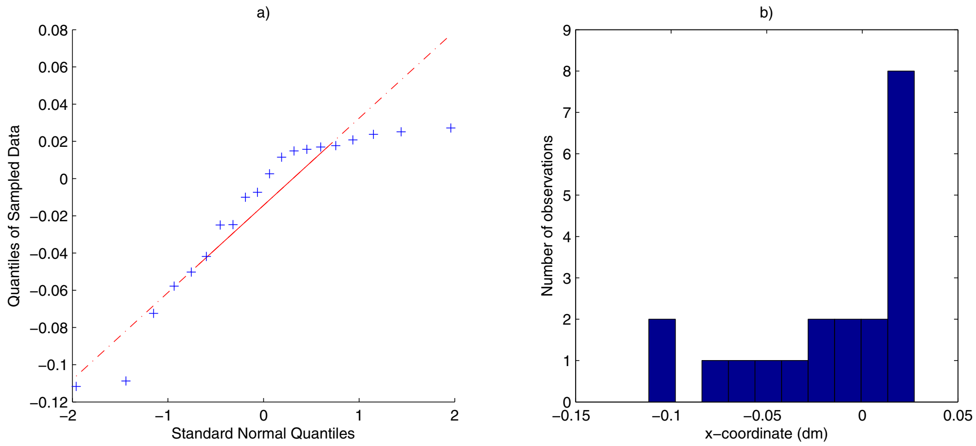

Fig. 2. a) The Q-Q plot and b) the histogram for the x-coordinate of the estimation sets from P1 at time instance 1200. The red line in the Q-Q plot indicates a normal distribution.

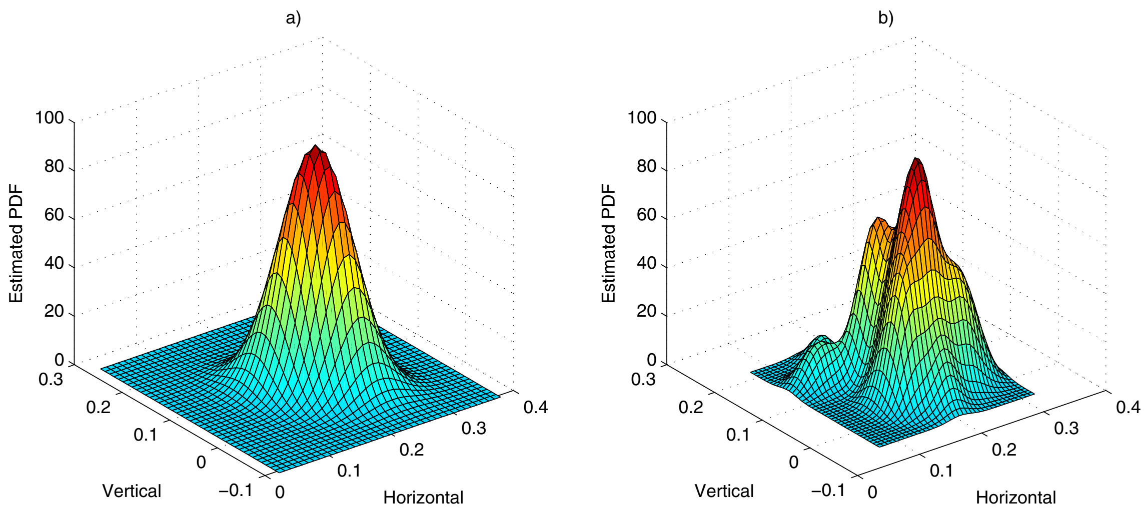

Fig. 3. a) The estimated normal distribution and b) the KDE with $h = 0.5\hat{\sigma}$ at time instance 1200 (using the estimation data of P1.

Dist. type	h	m_{P1} (%)	m_{P2} (%)	m_{P3} (%)	m_{P4} (%)
KDE	$0.1\hat{\sigma}$	36	46	55	66
	$0.25\hat{\sigma}$	4.9	21	32	48
	$0.5\hat{\sigma}$	0.6	15	27	37
	$\hat{\sigma}$	0.1	0.9	5.6	15
	$2\hat{\sigma}$	0	0.5	0.7	1.8
Normal	-	0.5	0.6	5.1	11

TABLE I

The means, m_{P1}, m_{P2}, m_{P3}, m_{P4}, of the number of outliers in the testing sets when comparing to the distribution estimates of P1. The means are given as percent of the total number of time steps.

Dist. type	h	m_{P1} (%)	m_{P2} (%)	m_{P3} (%)	m_{P4} (%)
KDE	$0.1\hat{\sigma}$	54	48	27	65
	$0.25\hat{\sigma}$	30	15	4.4	51
	$0.5\hat{\sigma}$	15	10	1.1	35
	$\hat{\sigma}$	2.7	1.4	0.9	7.4
	$2\hat{\sigma}$	0.4	0.4	0	2.5
Normal	-	8.0	5.1	2.1	15

TABLE II

The means, m_{P1}, m_{P2}, m_{P3}, m_{P4}, of the number of outliers in the testing sets when comparing to the distribution estimates of P3. The means are given as percent of the total number of time steps.

in the other test subjects. In Tab. II, the same can be said about the number of outliers in the testing sets of P3.

VII. Conclusions and Discussion

This paper is aimed at providing tools for distinguishing between individuals on the basis of their recorded eye movements. The suggested method relies on anomaly detection

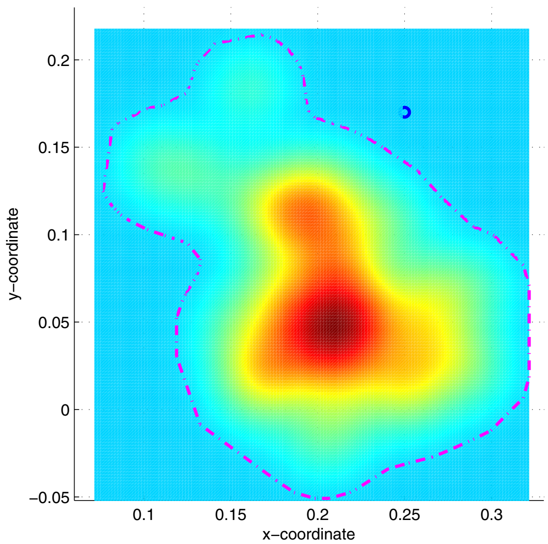

Fig. 4. The KDE estimate of the PDF at time instance 1200 from the estimation data of P1. Red indicates high values. The approximated outlier region is the exterior of the dashed line. The circle shows the gaze direction of P2 at this time instance.

and is non-parametric. The results of Sec. VI highlight several properties of the method.

Tab. I shows that P2, P3 and P4 are best distinguished from P1 when using the KDE method with $h = 0.5\hat{\sigma}$. For this choice of the bandwidth h, there is a most significant difference between the number of outliers in the testing sets of P1 compared to the sets of the other three test subjects. The number of outliers in the testing sets of P1 is expected to be low since the comparing distributions were estimated from the gaze direction data of this test subject. In Tab. II, it can be seen that the same holds true for the testing sets of P3 in comparison to the other examined individuals.

Tab. I shows that, for $h = 0.5\hat{\sigma}$, P3 and P4 differ more from P1, in terms of number of outliers, than P2 does. Furthermore, the results in Tab. II indicate that P2 and P1 differ less from P3 than P4 does. These two results suggest that the eye-tracking profiles of P1 and P2 are similar, but that the profiles of P3 and P4 differ from P1 and P2 and from each other. A possible explanation for this is the age of the test subjects. P1 and P2 are close in age and young relative to P3 and P4. The function of the oculomotor system impairs with age and may do so to different degrees.

Judging from Fig. 2 and the Lilliefors test, the data are not always well-modeled by a normal distribution. Hence, using Gaussian functions to approximate the PDFs at each time instance should impair the performance of the suggested method. By studying Tab. I and Tab. II it is apparent that this is indeed the case. The number of outliers in the testing sets does not vary significantly between individuals when normal distribution fitting is used. If the testing sets of P1 are deemed to come from P1, so should the testing sets of P2.

Tab. I and Tab. II also show that the choice of bandwidth for the kernel function in the KDE method has a significant effect on the results. When the bandwidth is low, the difference in the number of outliers between different test subjects is not large, making it difficult to distinguish between them. A low bandwidth in the kernel function will result in a PDF estimate with several narrow spikes, each centered at an observation. A high bandwidth will increase the smoothing and the estimated PDF will appear almost Gaussian.

The effect of the visual stimulus on the results of the presented method is not examined in this paper. The visual stimulus used was chosen because it had proven to give good results in parameter estimation of nonlinear dynamic models of the SPS [9]. There may exist visual stimuli more suitable for non-parametric anomaly detection.

In conclusion, with the method presented in this paper, individuals were successfully distinguished from each other on the basis of their recorded eye movements. The results of this work are promising, but since the number of participating test subjects was small, the results are only indicative of the potential of the presented method.

Acknowledgment

This study is in part financed by Advanced Grant 247035 from European Research Council entitled "*Systems and Signals Tools for Estimation and Analysis of Mathematical Models in Endocrinology and Neurology*".

References

[1] G. Avanzini, F. Girotti, T. Carazeni, R. Spreafico. Oculomotor disorders in Huntington's chorea *Journal of Neurology, Neurosurgery and Psychiatry*, 42:581–589, 1979.

[2] R. Dodge Five types of eye movements in the horizontal meridian plane of the field of regard *American Journal of Physiology*, 8:307–329, 1903.

[3] J. M. Gibson, R. Pimlott, C. Kennard. Ocular motor and manual tracking in Parkinson's disease and the effect of treatment *Journal of Neurology*, 50:853–860, 1987.

[4] M. Hazewinkel Chi-squared distribution *Encyclopedia of Mathematics, Springer*, 2001.

[5] H. Lilliefors On the Kolmogorov-Smirnov test for normality with mean and variance unknown *Journal of American Statistical Association*, 62: 399-402, 1967

[6] S. Marino, E. Sessam, G. Di Lorenzo, P. Lanzafame, G. Scullica, A. Bramanti, F. La Rosa, G. Iannizzotto, P. Bramanti, P. Di Bella Quantitative Analysis of Pursuit Ocular Movements in Parkinson's Disease by Using a Video-Based eye-tracking System. *European Neurology*, 58:193–197, 2007.

[7] D. Jansson, A. Medvedev, H. W. Axelson Mathematical modeling and grey-box identification of the human smooth pursuit mechanism *IEEE Multi-conference on Systems and Control*, Yokohama, Japan, 2010.

[8] D. Jansson, A. Medvedev Dynamic Smooth Pursuit Gain Estimation from eye-tracking Data *IEEE Conference on Decision and Control*, Orlando, Florida, 2011.

[9] D. Jansson, A. Medvedev Visual Stimulus Design in Parameter Estimation of the Human Smooth Pursuit System from Eye-Tracking Data *IEEE American Control Conference*, Washington D.C., 2013.

[10] A. B. Sereno, P. S. Holzman Antisaccades and Smooth Pursuit Eye Movements in Schizophrenia *Biological Psychiatry*, 37:394–401, 1995.

[11] B. W. Silverman Density Estimation for Statistics and Data Analysis *London: Chapman & Hall/CRC*, ISBN 0-412-24620-1, 1998

[12] T. Wigren MATLAB Software for Recursive Identification of Wiener Systems *Systems and Control, Department of Information Technology, Uppsala University*, 2007.

[13] M.B. Wilk, R. Gnanadeskian Probability plotting methods for the analysis of data *Biometrika*, 55:1–17, 1968.