Abstract—In this paper, we study the problem of an optimal oculomotor control during the execution of visual search tasks. We introduce a computational model of human eye movements, which takes into account various constraints of the human visual and oculomotor systems. In the model, the choice of the subsequent fixation location is posed as a problem of a stochastic optimal control, which relies on reinforcement learning methods. We show that if biological constraints are taken into account, the trajectories simulated under a learned policy share both basic statistical properties and a scaling behavior with human eye movements. We validated our model simulations with human psychophysical eye-tracking experiments.

Index Terms—Multifractal analysis, reinforcement learning, scaling in biology, visual search.

I. INTRODUCTION

The human oculomotor system performs hundreds of thousands of eye movements per day during the execution of different behavioral tasks. In order to find the details of a visual scene related to the tasks, humans direct foveal vision to the most informative locations via saccades—high-velocity conjugate gaze shifts. Saccades are followed by a visual fixation, during which the human oculomotor system generates fixational eye movements involuntarily. Despite the remarkable achievements in the modeling of fixational eye movements and the interpretation of their fundamental properties [1]–[3], there is no comprehensive generic model of fixation selection [4]–[6], which takes into account the underlying mechanisms of visual attention [5], [7], [8] and qualitatively describes the statistical properties of saccadic eye movements during the execution of visual tasks [9]–[12].

Previously, the problem of fixation selection was studied in the framework of control models of eye movements [9], [11], [13]. In control models, the observer gathers information about the world during each fixation, integrates information over all fixations into a belief state, and makes a choice of the next location on which to fixate.

This choice is governed by the policy of gaze allocation—a function that specifies the action of decision maker in a certain belief state. It was shown that the policy based on information maximization criteria [9] generates trajectories that share basic statistical properties with human eye movements. In this paper, we set the goal of developing a control model of fixation selection that is capable of interpreting the scaling behavior of human eye movements [10], [12], [14], [15] and provides a human level of performance to a computational agent.

In contrast to the previous research on control models, we take into account the inherent uncertainty of human oculomotor system and the duration of saccadic eye movements. It is well known that any motor action of humans is executed with random error, which increases with movement magnitude [16], [17]. Despite the oculomotor system having developed a correction mechanism for saccade errors [18], these result in inevitable temporal costs. Furthermore, the duration of saccades is empirically correlated with their magnitude as well [19]. These factors result in situations where the observer has to choose between more informative remote (and riskier) locations and those nearby (but less informative ones). We show that if these constraints are taken into consideration, the trajectories simulated under a learned policy share both basic statistical properties and scaling behavior with human eye movements, which is not achievable with the conventional infomax model [9].

On the basis of our results, we argue that we have made the following contribution.

1) The formulation of the biologically plausible model of gaze allocation in the human observer from the point of view of stochastic optimal control. The representation of the model in the form of partially observable Markov decision process (PO-MDP) and the proposal of a heuristic policy.

2) The development of robust and high performance algorithms of simulation of PO-MDP. The implementation of reinforcement learning algorithms of policy optimization and numerical estimation of the optimal policy of gaze allocation.

3) The comprehensive statistical analysis of simulated trajectories and data from our psycho-physical experiments. The policy, which is learned with the policy gradient REINFORCE algorithm, shows the highest level of statistical similarity with human eye movements. In our experiments we discovered the dependency of the mean saccade length and \( q \)-order Hurst exponent on visibility of the target, which was explained by our model.
II. MODEL OF IDEAL OBSERVER

In this section, we formulate the model of the ideal observer, which aims to localize the single target object on the stationary 2-D image. We represent the model in the form of PO-MDP, which is summarized by flow chart on Fig. 1.

A. World State

At the beginning of each episode the target object appears randomly at one of $L$ possible locations. We assume that the target is placed on background noise or surrounded by distractors, which are placed on vacant locations. The world state $S_n$ is represented as a tuple

$$S_n = (m, A_n, t_n)$$

where $m$ is a location of the target on the image and $A_n$ is gaze fixation location that changes with the number of step $n$, and $t_n$ is time passed from the start of a trial and the step $n$.

If the observer fixates the gaze on the location of target $A_n = m$

$$A_n = m$$

the visual task is considered to be accomplished. This formulation of the terminal state reflects the necessity to foveate the target in order to extract as much information about its identity and details as possible. The location of target $m$ does not change during a trial.

B. Update of Belief State

The decision making of the observer is modeled as PO-MDP with a belief state $p_n$—a discreet probability distribution function of target location given all observations received up to the step $n$. Because the observer is instructed that the target appears randomly, the initial belief state $p_0$ is a discrete uniform distribution.

On each step $n$ observer receives the observation vector $W_n = (W_{1,n}, \ldots, W_{L,n})$, whose elements represent the perceptual evidence that the target is at corresponding locations. The probability distribution function is updated using Bayesian inference [20]

$$p_n(l) = \frac{p_{n-1}(l)p(W_n|l, A_n)}{\sum_k p_{n-1}(k)p(W_n|k, A_n)}$$

where $l$ is the index of the location and $p(W|l, A)$ is an observation model. In order to take into account the uncertainty of the processing of perceptual information within the neural circuits of the observer, we follow the “noisy observation” paradigm [9]. In this paradigm the observation model $p(W|l, A)$ reflects the presence of the observer’s internal sources of inefficiency, such as physical neural noise on all stages of information processing. According to the perceptual model [11] the observation $W$ may be represented as a random variable with Gaussian distribution with mean depending on the location $m$ of the center of target on the lattice

$$p(W|l, A) = \prod_l p(W_l|A) = \prod_l N(W_l; \delta_{l,m}, \frac{1}{F(||l-A||)})$$

where $\delta_{l,i}$ is Kronecker delta, $N(x, \mu, \nu)$ is a value of Gaussian function with mean $\mu$ and variance $\nu$ for argument $x$; $||l-A||$ is Euclidean distance between the locations $l$ and the current fixation $A$, and $F$ is a fovea-peripheral operating characteristic (FPOC) [13]. FPOC is a function that represents the dependence of a signal-to-noise ratio on the eccentricity. Fig. 2 demonstrates FPOC calculated for several values of RMS contrast of the background $1/f$ noise: $\epsilon_n \in (0.1, 0.15, 0.2, 0.25)$ and a single value of RMS contrast of target $\epsilon_t = 0.2$. The calculation are based on the analytical expressions from [11]. The signal-to-noise ratio has a peak at fovea and decreases rapidly with eccentricity.

In our simulations we consider only the case of the rotationally symmetric FPOC. This assumption is not correct for human observers, and better generic model of FPOC can be found in [21]. The broken circular symmetry of FPOC inevitably results in the asymmetry of the visual search process [22]. We simplify the model of FPOC, because in this paper we focus our attention more on the temporal structure of eye movements rather than on their spatial distribution.

C. Execution of Saccades

The decision of which location to fixate next, $D_n$, is made on each step of PO-MDP according to the policy of gaze allocation $\mu$

$$D_n = \mu(p_n).$$
After making the decision, the coordinates of the next fixation location \( A_{n+1} \) are defined by the execution function

\[
A_{n+1} = \alpha(D_n) = D_n + J_n
\]

where \( J_n \) is a Gaussian-distributed random error with zero mean and standard deviation \( v \) defined in [17]

\[
v = \zeta_0 + \zeta_1 ||D_n - A_n||.
\]

The error of the saccade execution is proportional to intended saccade amplitude \( ||D_n - A_n|| \) given in degrees, the value of parameters: \( \zeta_0 = 0.87 \text{deg} \) and \( \zeta_1 = 0.084 \) (from [17]).

The next step of PO-MDP starts after the transition to the location \( A_{n+1} \). This decision making model may be extended in order to take into account the extraction of visual information between the moment of making the decision where to fixate next and completion of the saccade. This possibility is discussed in Section 4 in the supplementary material.

D. Duration of Steps

After each consequent step the time variable \( t \) of world state (1) is updated in a deterministic way

\[
t_n = t_{n-1} + \Theta(n)
\]

where \( \Theta(n) \) is a duration of step \( n \). The duration of time step \( \Theta(n) \) is considered as a total time, which is required for the relocation of the gaze from a previous location \( A(n-1) \) to the current one \( A(n) \) and the extraction of visual information from the location \( A(n) \). Therefore, we consider \( \Theta(n) \) as a sum of durations of the fixation \( \Theta_{fix}(n) \) and the saccade \( \Theta_{sac}(n) \). According to the literature, both of these time intervals are empirically correlated with a magnitude of the saccade preceding the fixation [19], [23], [24]. The duration of saccadic eye movements \( \Theta_{sac}(n) \) in range of magnitudes from 1.5° to 30° is possible to approximate as [25]

\[
\Theta_{sac}(n) = \tau_{sac}||A_n - A_{n-1}||^{0.4}
\]

where \( \tau_{sac} = 21 \text{ms} \cdot \text{deg}^{-0.4} \). Besides the magnitude of saccade, the fixation duration \( \Theta_{fix}(n) \) is influenced by various factors as a discriminability of the target [26], its complexity and the visual task of the observer [24], [27]. However, if the observer is correctly informed about the targets’ properties before the task execution and performs the visual task without any interruptions, the contribution of these factors to the fixation duration (with exception of magnitude) is constant during each trial. The eye-tracking experiments with the fixations tasks [23], [24], [28] found that the dependence of fixation duration on saccade amplitude is linear

\[
\Theta_{fix}(n) = ||A_n - A_{n-1}|| \tau_{fix} + \Theta_{0,fix}
\]

with a slope \( \tau_{fix} = 6 \text{ms/deg} \). The constant \( \Theta_{0,fix} = 250 \text{ms} \) is an intercept, averaged from values from eye-tracking data [29], [30]. Finally, the duration of step \( n \) is

\[
\Theta(n) = \Theta_{sac}(n) + \Theta_{fix}(n).
\]

The values of parameters used in simulations are consistent with our estimates from the eye-tracking experiments: \( \tau_{sac} = 20 \pm 3 \text{ms} \cdot \text{deg}^{-0.4} \), \( \tau_{fix} = 5.8 \pm 1.8 \text{ms/deg} \), \( \Theta_{0,fix} = 241 \pm 42 \text{ms} \). Within this range of the parameters’ values we did not find a substantial difference in the estimates of the learned policy of gaze allocation (18).

E. Value Function

Given the initial world state \( S_0 \), we define the cost function for policy \( \mu \) as an expectation of a random variable \( V \)

\[
V_{\mu}(S_0) = E[V|\mu, S_0].
\]

The random variable \( V \) denotes the cost and is defined by

\[
V \equiv c \sum_{n=0}^{N} \Theta(n) = cT_N
\]

where \( N \) is a total number of steps in the episode, and \( c \) is a time cost constant.

Formulation of the cost function as a temporal cost differentiates this paper from the previous works [11], [13], [31], which formulate the cost function as a number of steps of PO-MDP. We show below, that the policy \( \mu \) optimized for the cost function with the reward defined in (13) generates the sequences of actions with statistical characteristics close to the human saccadic eye movements.

III. POLICY OF GAZE ALLOCATION

A. Infomax Approach

In this section, we describe two heuristic policies related to the model of entropy limit minimization searcher [9]. We define the information gain on the step \( n+1 \) as: \( \Delta I(n+1) = H(p_n) - H(p_{n+1}) \), where \( H(\cdot) \) is Shannon entropy. The heuristic policy \( \pi_0 \) is defined as a policy which chooses such decision \( D_n \) that maximizes the expected information gain \( \Delta I(n+1) \)

\[
\pi_0(p_n) = D_n = \arg \max_D [E[\Delta I(n+1)]].
\]

The term \( E[\Delta I(n+1)] \) is calculated analytically in [9] for the case of the saccadic eye movements without uncertainty \( (A_{n+1} = D_n) \)

\[
E[\Delta I(n+1)] = \frac{1}{2} (p_n * F^2)(D_n)
\]
where sign $\ast$ denotes a convolution operator, and $F$ is FPOC represented as a radially symmetric 2-D function: $F(A) \equiv F(\|A\|)$. Expression (15) gives an approximate value of the expected information gain in the case of the stochastic saccadic placement (6).

Fig. 3 illustrates the decision-making process, which corresponds to the policy $\pi_0$. The color map (left) represents the function of the expected information gain (15). The blue cross corresponds to the location of the current fixation on the step $n$. The observer makes a decision to fixate at the location defined by the policy: $D_n = \pi_0(p_n)$. This decision results in a saccadic eye movement to location $A_{n+1} = \alpha(D_n)$ marked by the green cross. After receiving the observation at the step $n + 1$, observer updates the belief state and evaluates the information gain for the next decision. In this particular situation, the target is absent in the vicinity of $A_{n+1}$, and the observation resulted in the decline of probability $p_{n+1}$ in the area around the fixation (the green cross). This area is effectively inhibited from the subsequent fixations due to low probability $p_{n+1}$. The size of this area is defined by the values of FPOC (in this case $\varepsilon_1 = 0.2$ and $\varepsilon_n = 0.1$).

The trajectories generated with infomax greedy policy match the basic properties of human eye movements [9]. However, the policy (15) does not consider the correlation between the magnitude of saccades and the durations of steps of MDP. We show later that the policy $\pi_0$ is inferior to the policy that optimizes the expected rate of information gain $E[\Delta I(n + 1)/\Theta(n + 1)]$

$$\pi_1(p_n) = \arg \max_D E[\Delta I(n + 1)/\Theta(n + 1)].$$

(16)

Using the expression for $E[\Delta I(n + 1)]$ (15), for the deterministic saccadic placement ($A_{n+1} \equiv D_n$)

$$\pi_1(p_n) = D_n = \arg \max_D \left[ (p_n \ast F^2)(D_n)/\Theta(n + 1) \right].$$

(17)

The policy $\pi_1$ is called “infomax rate” in the text below. The performance of these two heuristic policies will be compared with a performance of the policy learned with reinforcement learning algorithms.

### B. Optimal Policy Estimation

In this section, we describe the evaluation of the policy of gaze allocation that optimizes the cost function (12) for any starting world state $S_0$. We start with the representation of the stochastic policy $\mu$ in the on [20]

$$\mu(D, p) = \frac{\exp(f(D, p))}{\sum_l \exp(f(l, p))}$$

(18)

where $f(D, p)$ is a function of expected reward gain after making the decision $D$ with the belief state $p$. In this paper, we limit the search of $f(D, p)$ to a convolution [20] of belief state $p$

$$f(D, p) = \sum_l K(D - l)p(l).$$

(19)

In Section 1, in the supplementary material, we justify this choice of the policy and evaluate the form of a kernel function $K$ that allows us to effectively solve the optimization problem with the policy gradient algorithms. Our task is the search of the kernel function $K$, which corresponds to the policy that optimizes the cost function $V_\mu$

$$K^* = \arg \min_K V_{\mu(K)}(S_0)$$

(20)

for any starting world state $\forall S_0$. The policy $\mu(K^*)$ is called the optimal policy of gaze allocation.

We approach the optimization problem (20) with an algorithm named “REINFORCE with optimal baseline” [32] according to the procedure described in Section 1 in the supplementary material. The performance of REINFORCE was compared with one of the optimization algorithms named “policy gradient parameter exploration” (PGPE) adopted...
from [33]. The algorithm of REINFORCE with an optimal baseline belongs to the class of the likelihood ratio methods, whereas PGPE is related to the finite difference methods. Despite the distinction between these two approaches, both algorithms give a close estimation of the optimal policy (Section 1 in the supplementary material). We simulated trajectories for the data analysis in Section IV using the solution provided by REINFORCE due its better performance comparing to PGPE. Fig. 4 demonstrates the decision making process under the policy learned for FPOC corresponding to conditions $e_1 = 0.2$ and $e_n = 0.1$. At the step $n$, observer fixates the location marked by a blue cross. The policy $\mu$ defines a probability density function of a decision $D$ where to fixate next (18). Observer chooses the decision $D_n$ according to the policy, which results in a saccadic eye movement to location $A_{n+1} = \alpha(D_n)$ (the green cross). As well as in the case of dynamics under the heuristic policy $\pi_0$, previously visited locations are inhibited from the subsequent fixations. Note that movements to remote locations are inhibited by the radial function. This results in co-directed short movements, which are also characteristic of human observer.

IV. BASIC PROPERTIES OF TRAJECTORIES

In this section, we discuss the statistical properties of trajectories generated with the learned policy $\mu$ and the heuristic policies $\pi_0$ and $\pi_1$. The simulations were performed on the grid with size 128 x 128 that corresponds to the visual field with size of 15 x 15 deg in the psychophysical experiment. In order to justify our computational model, we reproduced the psychophysical experiments from [9]. The detailed description of the experiments can be found in Section 3 in the supplementary material.

A. Performance

Although this computational model was not designed for an exact prediction of a response time of human observers, it demonstrates a high level of consistency in a performance of the visual task execution with human observers. The performance was measured as an average time to reach the target (the mean response time) and as a percentage of the correct fixations on target’s location on an N-alternative forced choice task (N-AFC). The unsuccessful trials from the psychophysical experiments were excluded from the consideration. We found that the number of the unsuccessful trials grows with the contrast of noise: 2.3%, 5.7%, 9.8%, and 16.4% for the corresponding numbers of the contrast $e_n = (0.1, 0.15, 0.2, 0.25)$.

Fig. 5(left) demonstrates the percentage of correct fixations on the target location for the experimental conditions: $e_l = 0.2$ and $e_n = 0.15$. Means and standard errors of the response time of the human observers is presented on Fig. 5(right) together with means of the response time for three policies estimated from $10^4$ episodes of PO-MDP. The learned policy outperforms two heuristics and the human observers both in the mean response time and the percentage of the correct fixations for all experimental conditions. Human observers significantly outperformed the infomax rate for the experimental conditions: $e_n = (0.2, 0.25)$ (Student’s $t$-test $p < 0.05$) and the infomax greedy for the conditions $e_n = (0.15, 0.2, 0.25)$ ($p < 0.05$) on the mean response time, which was previously found in [9] and [22]. In the same time the learned policy outperformed the human observers significantly for the condition $e_n = 0.25$, while for other conditions $t$-test did not reject hypothesis that distributions have equal means at 5% significance level.

B. Amplitude Distribution

Fig. 6(left) shows the length distributions of saccades of the human observers and the simulated agents performing the visual search task corresponding to the experimental conditions: $e_n = 0.2$ and $e_l = 0.2$.

The distributions for all policies and the human observers exhibit an ascent between 0 deg and maximum around 2 deg. The difference in the behavior of the distributions starts from 4 deg. In this experimental conditions the share of the saccades
Fig. 5. Performance of the human observers and the simulated agents. The learned policy outperforms two heuristics both in the mean completion time and the percentage of correct responses in N-AFC task (left) for the experimental conditions: \( e_0 = 0.2 \) and \( e_1 = 0.2 \). The dependence of mean completion time (right) for the learned policy resembles one for the human observer.

Fig. 6. Histograms of the length distribution of the saccadic events (left) for trajectories generated under the policies \( \pi_0, \pi_1, \) and \( \mu_{\text{conv}} \) and the human eye movements corresponding to the experimental conditions: \( e_0 = 0.2 \) and \( e_1 = 0.2 \). The data was binned with the resolution of 0.1 deg. The distribution function for all policies and human observer exhibits an ascent between 0 deg and maximum around 2 deg. The distribution of length corresponding to the infomax greedy \( \pi_0 \) stabilizes after 4 deg and declines only after 10 deg. It is not consistent with length distribution of human saccadic eye movements, which is concave on an interval [4.0\(^\circ\), 14.0\(^\circ\)]. The mean length of saccades decreases with \( e_0 \) (right). It is immediate consequence of the decrease of the width of FPOC with \( e_0 \), which defines the area of inhibition from the subsequent fixations.

The mean length of the saccades was estimated from \( 10^4 \) episodes of PO-MDP for all three policies and compared with the mean length of the saccades of the human observers [see Fig. 6 (right)]. According to our results, the mean length of the saccades decreases with \( e_0 \), which is consistent with our simulations. It is an immediate consequence of the decrease of values of FPOC with the increase of the RMS contrast of noise, which is illustrated on Fig. 2. The amplitude of the signal exceeds the amplitude of noise within the circle area with
C. Geometrical Persistence

In this section, we analyze the distribution of the directional angle \( \theta_d \) (this notation was introduced in [10]) of the human saccadic eye movements and the simulated trajectories. The directional angle is the angle between two consequent saccades, and, therefore, can be defined as \( \theta_d = \tan^{-1}\left(\frac{y_{n+1}/x_{n+1}}{y_n/x_n}\right) \), where \((y_n, x_n)\) are the coordinates of \(n\)th fixation. According to this definition, the movement is related to a persistent one if the directional angle is close to 0 or \(2\pi\). The angles with the values close to \(\pi\) correspond to anti-persistent movements.

The distributions of the directional angle were calculated for the trajectories generated by Markov decision process with the policies \(\pi_0\), \(\pi_1\), and \(\mu\). Fig. 7(left) demonstrates the distribution of the directional angle of the saccadic events for the human observers and the simulated trajectories for \(e_n = 0.2\) and \(e_t = 0.2\). The infomax greedy policy \(\pi_0\) generates the trajectories with stable anti-persistent movements, because the policy \(\pi_0\) chooses the next fixation location without taking the current location into consideration. Due to the inhibitory behavior of infomax, it is much less likely to choose the nearby location instead of remote and relatively unexplored ones. Only geometrical borders limit the choice of the next fixation, which results in fixations on the opposite side of the visual field (as the most remote point, see Fig. 3).

In contrast, the decision process under the learned policy \(\mu\) tends to preserve the direction of the movement. The dynamic of the system under the policy \(\mu\) is quite similar to self-avoiding random walk model described in [1]. Due to the asymptotic behavior of the kernel function \(K(x, y)\), the reward gain from the remote locations is suppressed, meanwhile, the locations, which are already visited, are also inhibited (see Fig. 4). This results in short-range self-avoiding movements, which demonstrate the persistent behavior [1], [37], and, therefore, the probability distribution of the directional angle \(\theta_d\) is biased toward values 0 or \(2\pi\). According to Fig. 7(left), the dynamics under the heuristics \(\pi_1\) is also characterized as a persistent random walk. The learned policy \(\mu\) has, in general, a stronger radial ranking of locations than \(\pi_1\), which results in a shorter range of saccades, and a repulsion, caused by inhibition, becomes more relevant. The distribution of average length of saccades depending on \(\theta_d\) is shown on Fig. 7(right).

In our experiments we discovered that the geometrical persistence depends on the visibility of target (on FPOC in the simulations). We measured the share of the saccades, which retain the direction of the previous movement: \(\cos(\theta_d) > 0\). This quantity is called “persistence coefficient.”
Fig. 8. Share of saccades, which retain the direction of the previous movement: $\cos(\theta_f) > 0$, is called the persistence coefficient. This quantity demonstrates the dependence of the persistence on the visibility of the target. As it was mentioned previously, the average saccade length is decreasing with the growth of the RMS contrast (6). Therefore, the linear term (10) in the duration of the steps becomes less relevant, and the decision making becomes more agnostic about temporal costs (closer to the information greedy $\pi_0$). The decline of the persistent coefficient is also a characteristic of the human eye movements, which was not covered in the previous research.

For statistical analysis of simulated trajectories we use a multifractal detrended fluctuation analysis (MF-DFA) [38], which is a widely used method for detection of long-range correlations in stochastic time-series. It has found successful applications in the field of bioinformatics [39], [40], nano, and geo-physics [41]. This method is based on the approximation of trends in time-series and the subtraction of detected trends (detrending) from original data on different scales. The detrending allows deducting the undesired contribution to long-range correlation, which is a result of nonstationarities of physical processes. We use the package provided by Ihlen [42] for all our estimations of the generalized Hurst exponent in this section.

In Section 2, in the supplementary material, we thoroughly explain the details of multifractal analysis. Section 2.2 in the supplementary material presents the details of MF-DFA algorithm. In Sections 2.2 and 2.3, in the supplementary material, we explain how MF-DFA is performed over simulated trajectories. The results of multifractal analysis of human eye movements are presented in Section V-A. Section V-B summarizes our findings and compares generalized Hurst exponent of simulated trajectories to the one of human eye movements for different experimental conditions.

### A. Multifractality of Human Eye-Movements

We perform MF-DFA over the difference of time series of the human gaze positions and in order to compare the estimated generalized Hurst exponent with the simulations. The differentiated time series was estimated from raw data of coordinates of the gaze fixations $A = \{(x_1, y_1), \ldots (x_N, y_N)\}$ with the resolution of 7 ms.

\[
\Delta X = \{(x_2 - x_1), \ldots, (x_N - x_{N-1})\} \quad (21)
\]

\[
\Delta Y = \{(y_2 - y_1), \ldots, (y_N - y_{N-1})\}. \quad (22)
\]

The time series $\Delta X$ and $\Delta Y$ were estimated for each trial with certain experimental conditions and concatenated over all participants. After this, we represent the differentiated time series in the following way: $\Delta X = \{F_1, S_1, \ldots, F_{m-1}, S_{m-1}, F_m\}$, where $F_i$ and $S_i$ correspond to the sequences of the movements during time interval of $i$th fixation and saccade, respectively [10]. We separate the differentiated time series on the fixational and the saccadic time series

\[
\Delta X_F = \{F_1, 0_{s_1}, \ldots, 0_{s_{m-1}}, F_m\} \quad (23)
\]

\[
\Delta X_S = \{0_{f_1}, S_1, \ldots, 0_{f_{m-1}}, S_m, 0_{f_m}\} \quad (24)
\]

where $0_n$ corresponds to zero array with the length $n$, and $f_m$ and $s_m$ are the lengths of corresponding sequences $F_m$ and $S_m$.

Fig. 9 demonstrates the scaling of the $q$-order fluctuation function $F_q(s)$ (see Section 2 in the supplementary material). This graph is a result of the application of MF-DFA over the horizontal concatenated differentiated time series $\Delta X$ of the human scan-paths for the experimental conditions: $e_i = 0.2$ and $e_n = 0.25$. The red, blue, and green lines correspond to the linear approximation of function $\log_2(F_q(s))$ for the orders $q = \{-10;\ 0;\ 10\}$. The scaling of $F_q(s)$ exhibits the crossover on a time scale of 256 ms. The crossover separates
the “lower” and “upper” regimes mentioned in [10]. According to Amor et al. the crossover is caused by the presence of two different generative mechanisms of eye movements. The lower regime is related to fixational eye movements (which is supported by the value of crossover scale \(s_{\text{cros}}\) being close to the average fixation duration), and upper regime—to saccadic ones. The crossover in the scaling of \(F_q(s)\) was observed for all experimental conditions.

In order to distinguish between two different types of multifractality [38] we calculated the generalized Hurst exponent \(H_{\text{shuf}}(q)\) for the shuffled differentiated time series. The first type of multifractality is a consequence of a broad probability density function for the values of time series. If only multifractality of the first type presents in time series: \(H(q) = H_{\text{shuf}}(q)\). The second type of multifractality is caused by the difference in correlation between large and small fluctuations, which is a scenario described in [10]. In this case \(H_{\text{shuf}}(q) = 0.5\) and \(H(q) = 0.5 + H_{\text{corr}}(q)\), where \(H_{\text{corr}}(q)\) is (negative) positive for the long-range correlation (anti-)correlation. If both types of multifractality present in time series: \(H(q) = H_{\text{shuf}}(q) + H_{\text{corr}}(q)\).

Fig. 10 demonstrates our estimates of the Hurst exponent of the shuffled time series \(H_{\text{shuf}}(q)\) (top) and the correlational Hurst exponent \(H_{\text{corr}}(q)\) (bottom) for the horizontal (left) and the vertical components (right). We estimated both exponents for the saccades (green dashed line) and FEM (purple dashed line) in the upper and the lower regimes of scales, respectively. As well as a previous graph 9, this one is a result of an application of MF-DFA over the concatenated differentiated time series of the human eye movements for the experimental conditions: \(c_1 = 0.2\) and \(c_2 = 0.25\). The behavior of \(H_{\text{shuf}}(q)\) for the full time series and the saccadic time series in the upper regime corresponds to the one mentioned in [38, eq. (27)]

\[
H(q) \sim \begin{cases} 
1/q & (q > \alpha) \\
1/\alpha & (q \leq \alpha) 
\end{cases}
\] (25)

with \(\alpha \sim 1\). Equation (25) was derived for time series of uncorrelated random values with the power law distribution

\[
P = \begin{cases} 
\alpha x^{-(\alpha + 1)} & x \geq 1 \\
0 & x < 1 
\end{cases}
\] (26)

One can see a similarity of the function (26) with the distribution of the amplitude of the saccadic events for humans (see Fig. 6). The amplitude distribution of the saccades demonstrates the power law behavior on the interval \([4,0^\circ, 14,0^\circ]\) with \(\alpha \approx 1\). The probability distribution function (26) also reflects an absence of saccades with the length lower than minimal one. Therefore, the first type of multifractality of the saccadic time series is caused by the broad probability distribution of saccade magnitude.

The difference in the long-range correlation of large and small fluctuations is reflected by \(H_{\text{corr}}(q)\) [Fig. 10 (bottom)]. Due to the properties of fluctuation function for the positive (negative) \(q\)-orders the main contribution are coming from segments containing the large (small) fluctuations [38]. The positive (negative) long-range correlation \(H_{\text{corr}}(q) > 0\) is, therefore, a characteristic of the small (large) fluctuations in the upper regime for the saccadic and the full time series. These results are consistent with the distribution of the average length of saccade to the directional angle [see Fig. 7 (right)], which also indicates the difference in the persistence of large and small saccades. Therefore, we confirm here that the small saccadic eye movements demonstrate the long-range correlations as well as fixational eye movements.

The time series of FEM demonstrates the monofractal behavior and the positive correlations with \(H \approx 0.8\) in the lower regime of scales [10]. However, the behavior of both \(H_{\text{corr}}(q)\) and \(H_{\text{shuf}}(q)\) for the full time series in the lower regime indicates the presence of multifractals of both types. At the present moment we have no explanation of the multifractality in the lower regime and leave this problem for a future work.
Fig. 10. Hurst exponent of the shuffled time series $H_{shuf}(q)$ (top) and the correlational Hurst exponent $H_{corr}(q)$ (bottom) for the horizontal (left) and the vertical components (right) of the human eye movements. As well as a previous graph (Fig. 9), this one is a result of an application of MF-DFA over the concatenated human scan-paths for the experimental conditions: $e_l = 0.2$ and $e_r = 0.25$. The behavior of $H_{shuf}(q)$ for both horizontal and vertical shifts for full scales corresponds to the one mentioned in [38, eq. (27)]. We assume that multifractality of the first type is caused by an asymptotic behavior of the amplitude distribution of saccades (see Fig. 6). The difference in the long-range correlation of large and small fluctuations is reflected by $H_{corr}(q)$ (bottom). Due to the properties of the fluctuation function for positive (negative) $q$-orders the main contribution are coming from segments containing small (large) fluctuations [38]. The positive (negative) long range correlation ($H_{corr}(q) > 0$) is, therefore, a characteristic of small (large) fluctuations in the upper and the full scales regimes for both directions. In general, these results are consistent with the distribution of the average length of saccades to the directional angle [see Fig. 7 (right)], which also indicates the difference in persistence of large and small saccades.

B. Dependence on Visibility

In this section, we present a comparison of the generalized Hurst exponent for the human eye movements in the upper regime and the simulated trajectories under the learned policy. As well as in the case of the geometrical persistence, we claim the quantitative properties of the statistical persistence depend on the visibility of the target.

We estimated the correlational Hurst exponent $H_{corr}(q)$ and the Hurst exponent of the shuffled time series $H_{shuf}(q)$ for the differentiated trajectories of the human eye movements for all levels of the RMS contrast of background noise: $e_n \in (0.1, 0.15, 0.2, 0.25)$. Fig. 11(left) shows $H_{corr}(q)$ (left) of simulated trajectories (blue) under the learned policy $\mu$ and the correlational Hurst exponent for the human eye movements (pink) averaged over two directions: $H_{corr}(q) = (H_{corr}^x(q) + H_{corr}^y(q))/2$ in the upper regime. The correlational Hurst exponents for the negative $q$-orders declines with the growth of the RMS contrast of background noise both for the human eye movements and the simulated trajectories. This indicates the weakening of the correlation between small fluctuations. For the positive $q$-orders the correlational Hurst exponent is less affected by the change of the visibility of target. The $H_{corr}(q)$ for $q = 10$ stabilized on values 0.04 and −0.12 for human eye movements and the simulated trajectories correspondingly. In general, the correlations weaken with the growth of the RMS contrast, which is consistent with the decline of the geometrical persistence (Fig. 8). The decline of the Hurst exponent with the increase of difficulty of visual search task was also observed in the previous work [12].

The Hurst exponent of the shuffled time series [Fig. 11 (right)], as well as the correlational Hurst exponent, demonstrates the decline with the growth of the RMS contrast for the negative $q$-orders both for the human eye movements and the simulated trajectories. In Section V-A, we mentioned that the behavior of $H_{shuf}(q)$ resembles the one related to time series of random values with the power law distribution (26). The average value of this time series equals $1/(\alpha - 1)$ for $\alpha > 1$. The increase of $\alpha$ results both in the decrease of the average value in time series and the decrease of the value of $H_{shuf}(q) \sim 1/\alpha$ for $q < 0$. Therefore, the average value...
in time series and the values of $H_{\text{shuf}}(q)$ for the negative $q$-orders are correlated in the assumption of the power-law distribution. Previously we found the decrease of the average saccade length with the growth of RMS of background noise (Fig. 6), which is consistent with the decrease of values of $H_{\text{shuf}}(q)$ for negative $q$-orders. We assume that this correlation is caused by the power-law asymptotic behavior of the length distribution of human eye movements (26).

VI. CONCLUSION

We have presented a computational model of the ideal observer that both qualitatively and quantitatively describes the human visual behavior during the execution of the visual search task. The basis of this model is the observer’s representation of the constraints of its own visual and oculomotor systems. We demonstrated that a consideration of the temporal costs and uncertainty of the execution of saccades results in the dramatic change of the basic statistical properties and the scaling behavior of the simulated time series.

We performed the multifractal analysis of our data and discovered the presence of two types of multifractality both in time series of the human eye movements and the model simulations. The multifractality caused by the broad amplitude distribution of the saccades (the first type of multifractality) makes a significant contribution to the multifractal behavior of time series, which was not covered in the previous work [10]. After the estimation of the correlational part of the Hurst exponent [38] we confirmed the presence of the long-range positive correlations of the small saccades in the upper regime. On the contrary, the large saccades exhibit the weak long-range anti-correlations for the model simulations and the human eye movements in the upper regime. As well as in the case of the geometrical persistence, we found that the long-range correlations between eye movements weaken with the decline of the target’s visibility, which is consistent with the previous work on this topic [12].

In this paper, we focused our attention more on the persistence of eye movements rather than on their spatial distribution. That is why we did not consider the factors that are not directly related to the tradeoff between the temporal costs and the expected information gain. We estimate the optimal policy under the assumption that the visual search process is characterized by shift-rotational symmetry [20], which was not observed in the previous work with similar experimental settings [22]. The symmetry of the visual search can be broken by angular dependency of FPOC in both cases of normal controls and patients with vision disabilities [43]. We plan to include the angular dependency to radial and smoothing functions of policy in order to consider the asymmetry of the visual field in our future works. Our next goal will be a development of decision making model that operates in continuous time and, therefore, capable to control fixation durations. In Section 4, in the supplementary material we already discussed the possibility of extension of current observation model in which the observer makes decisions based on observations from the labile stage of the current fixation and the nonlabile stage of the previous one.

To sum up, this framework provides an elegant explanation of scaling and persistent dynamic of the voluntary saccades from an optimality point of view. It clearly demonstrates that control models are able to describe human eye movements far beyond their basic statistical properties.

ACKNOWLEDGMENT

The authors would like to thank M. Hansard, I. Mareschal, and I. Patras for useful discussions and comments on this paper, and I. Miflah for help with setting of eye-tracking device.

REFERENCES

[1] R. Engbert, K. Mergenthaler, P. Sinn, and A. Pikovsky, “An integrated model of fixational eye movements and microsaccades,” Proc. Nat. Acad. Sci., vol. 108, no. 39, pp. E765–E770, 2011.
Alexander Vasilyev was born in Yekaterinburg, Russia, in 1991. He received the B.Sc. and M.Sc. degrees in applied mathematics and physics from the Moscow Institute of Physics and Technology, Dolgoprudny, Russia, in 2012 and 2014, respectively. He is currently working toward the Ph.D. degree at the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, U.K.

His current research interest includes modeling of human eye-movements during execution of visual tasks.